



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

**Understanding SAF-A's role in chromatin
organisation throughout the cell cycle and in a rare
neurodevelopmental disorder**

Elena Lazarova



THE UNIVERSITY
of EDINBURGH

Thesis submitted for the degree of *Doctor of Philosophy*

The University of Edinburgh

June 2024

Declaration

The work presented in this thesis is the original work of the author. Contribution by others has been clearly stated. This thesis has been written entirely by the author and has not been submitted in whole or in part for any other degree or professional qualification.

Elena Lazarova

6th June 2024

Acknowledgements

I would like to thank Nick Gilbert for providing guidance and support throughout my PhD and teaching me old-school techniques. This thesis uses many techniques and materials developed by Ryu-Suke Nozawa and I am grateful to him for teaching me at the beginning of my PhD journey. Thank you to the past and present Gilbert lab members (Catherine, Waad, Elaine, Jamilla, Mattia, Jon, Jim, Tasnim, Olivia, Patrick, James, Rafal, Charles, Adam, Issy, and Lora) for making a supportive environment.

I am also grateful for my thesis committee (Javier Caceres, Sara Buonomo, David FitzPatrick) providing valuable input to my PhD journey. Thank you to the AIR team (Matt Pearson, Laura Murphy, Ann Wheeler, and James Iremonger) for continuous support for my live cell microscopy. Many thanks to Kumiko Samejima for providing RPE1 CDK1as cells, as well as Steven Wilson providing HNRNPU-AID. I am grateful to Lizzie Freyer from the FACS facility for always helping with my experiments and providing emotional support. Thank you to Cath Heath for sharing insights about HNRNPU/SAF-A and giving me access to her embargoed thesis. Grateful to Andrew Wood for CRISPR advice, to Shelagh Boyle for FISH advice, and to Atlanta Cook for mutational advice and PyMOL help.

To the fellow PhD friends I made along the way – Waad, Laura, Alan, and Arianna – you made this a fun journey and I'm glad we all supported each other during our respective thesis writing challenges. I am grateful to everyone that encouraged me along the way – the Hop Trop dancers, Rock & Metal society members, and people at the Salisbury Place. Very grateful to my invaluable counsellor, Alison, and to my friends Jana, Isabelle, Adèle, Makaela, and Evelyn for letting me vent and cheering me on. Thank you to my fantastic flatmate, Ariana, for the emotional support and understanding during a difficult writing process. And to Adam, who only joined my life less than two years ago yet has been one of my biggest supporters and the safest place in Edinburgh for me.

Finally, I am eternally grateful to past Elena who worked so hard and did so much to the point of burnout and mental health difficulties. You were more than enough, and I feel so proud of you when I was writing this thesis and sorting through the huge amount of work you did. I am also grateful for mamo Maria who supported me in her own way and believed I could do this every step of the way and always showed me cute pics and videos of my cat Marie.

Lay Summary

All living organisms contain DNA in their cells for information storage and transfer. How DNA is packaged within cells is, therefore, an important research question. SAF-A is a protein that helps DNA unpackage to allow access to the stored information for cells to function properly. This research project focused on understanding SAF-A and its role in organising our DNA when cells divide. I investigated SAF-A's behaviour when cells finish dividing, a process called the mitotic exit, which is not well understood. By studying SAF-A, I aimed to uncover its importance for the development of our brain and how problems with SAF-A function can lead to neurodevelopmental disorders. My findings showed that SAF-A is important for organising the DNA after cell division by interacting with specific molecules called RNAs. Both RNA and SAF-A make mesh-like structures that ensure the DNA is properly arranged. When SAF-A was absent or not functioning correctly, the DNA organisation was disrupted, which could contribute to the cellular problems observed in some individuals with neurodevelopmental disorders.

To study SAF-A, I used advanced imaging techniques to observe SAF-A's behaviour in living cells during different stages of cell division. I found that SAF-A is recruited to the DNA during a specific stage called cytokinesis, which is when the cell finishes dividing. Interestingly, SAF-A's recruitment was dependent on the cell's transcriptional activity, which is when genes are "read" to produce RNAs. This suggests that the DNA needs to be actively transcribed during cell division to recruit SAF-A and maintain proper DNA organisation. Moreover, I explored the effects of SAF-A mutations found in patients with neurodevelopmental disorders. By introducing these mutations into human cells, I could observe how they affected the organisation of DNA. Each SAF-A mutation had a different impact on the DNA structure, further highlighting the importance of SAF-A for normal brain development. Understanding SAF-A's role in DNA organisation during the mitotic exit and its connection to neurodevelopmental disorders brings us closer to finding potential treatments and therapeutic targets for these conditions.

Abstract

Large-scale chromatin organisation during interphase and mitosis has been extensively studied; however, the mitotic exit remains poorly understood. This thesis aims to elucidate the involvement of the essential protein Scaffold Attachment Factor A (SAF-A; also called HNRNPU) in chromatin organisation during the mitotic exit and its impact on cellular phenotypes, particularly in the context of HNRNPU-related neurodevelopmental disorder. SAF-A regulates interphase chromatin structures in a transcription-dependent manner. The protein can decompact gene-active regions by interacting with chromatin-associated RNAs and binding ATP to form protein oligomers. SAF-A's eviction from mitotic chromosomes suggests it might have a precise recruitment to gene-active regions for establishing and maintaining the RNA mesh after mitosis. Additionally, perturbations in SAF-A function may result in faulty chromatin architecture at gene-active regions, contributing to cellular phenotypes observed in patients with HNRNPU-related neurodevelopmental disorder.

To investigate whether transcriptional events during telophase or in early G₁ recruit SAF-A, I used a live cell imaging system and co-localisation analysis. Imaging cells going through mitosis showed that SAF-A is recruited to chromatin during cytokinesis and inhibiting transcription at that time delayed its recruitment to early G₁. This allowed the examination of SAF-A's association with RNAs during mitosis and mitotic exit using synchronized cells. In comparison with monomeric SAF-A, more oligomeric SAF-A was associated with new and old RNA during mitosis and early G₁, showing that the protein regulates large-scale chromatin organisation as soon as it is recruited during mitosis. Next, I investigated the interaction between SAF-A and RNA and whether SAF-A depletion affects the RNA mesh in interphase. Depleting SAF-A with RNAi treatment resulted in a reduced RNA mesh density and connectivity. Finally, the effect of SAF-A mutations, from patients with the HNRNPU-related neurodevelopmental disorder, on chromatin organisation was explored. Human cell lines expressing mutated SAF-A proteins were generated, and their impact on chromatin decompaction was quantified using FISH with each mutant having a different phenotype. This thesis sheds light on the behavior of SAF-A at the mitotic exit and provides insights into its role in large-scale chromatin organisation. Furthermore, it investigates the cellular phenotypes associated with SAF-A patient mutations, contributing to our understanding of the underlying mechanisms and potential therapeutic targets for HNRNPU-related neurodevelopmental disorder.

Table of Contents

DECLARATION.....	II
ACKNOWLEDGEMENTS	III
LAY SUMMARY	IV
ABSTRACT.....	V
TABLE OF CONTENTS	VI
LIST OF FIGURES.....	X
LIST OF TABLES.....	X
ABBREVIATIONS	XII
1. CHAPTER 1: INTRODUCTION	17
1.1. Chromatin organisation and the cell cycle.....	17
1.1.1. Chromatin structural components	21
1.1.1.1. Histones.....	21
1.1.1.2. Chromatin remodellers.....	24
1.1.1.3. Chromatin-associated RNAs and the RNA mesh.....	26
1.1.1.4. RNA binding proteins	31
1.1.1.5. Other proteins involved in chromatin organisation	33
1.1.2. Chromatin organisation during interphase	38
1.1.3. Chromatin organisation during mitosis	43
1.1.4. Chromatin organisation at the mitotic exit and in early G ₁	46
1.2. SAF-A.....	49
1.2.1. SAF-A and chromatin organisation during interphase.....	50
1.2.2. SAF-A and chromatin organisation during mitosis	54
1.2.3. Other SAF-A functions	55
1.2.4. SAF-A mutations in a rare neurodevelopmental disorder	56
1.3. Hypothesis and aims.....	61
2. CHAPTER 2: MATERIALS AND METHODS	63
2.1. Reagents, solutions, and buffers.....	63
2.2. General methods.....	69
2.2.1. RNA purification and quantification.....	69
2.2.2. DNA purification and quantification.....	70
2.2.3. Bacterial transformation and culture.....	71
2.2.4. PCR amplification	71
2.2.5. Agarose gel electrophoresis.....	73
2.2.6. DNA extraction from agarose gels.....	73
2.2.7. Restriction enzyme digestion	74
2.2.8. Gibson Assembly	74
2.2.9. Ligation of DNA fragments	74
2.2.10. Sanger sequencing	74
2.2.11. RT-qPCR.....	74
2.3. Cell culture	75

2.3.1. Cell lines	75
2.3.2. Cell passage	76
2.3.3. Freezing and thawing cells.....	76
2.3.4. Experiment preparation.....	77
2.3.5. DNA transfection	77
2.3.6. RNA interference.....	79
2.3.7. Growth assay.....	80
2.3.8. CRISPR Cas9 stable cell line generation	80
2.3.9. Stable cell line generation using the FlpIn system	82
2.4. <i>Protein analysis</i>	82
2.4.1. Protein lysate preparation	82
2.4.2. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)	82
2.4.3. Western blotting	83
2.5. <i>Microscopy techniques</i>	84
2.5.1. Protein immunofluorescence.....	84
2.5.2. Click chemistry.....	85
2.5.3. dSTORM microscopy	86
2.5.4. Live cell imaging	87
2.6. <i>SAF-A oligomer crosslinking</i>	88
2.7. <i>SAF-A immunoprecipitation</i>	88
2.8. <i>Cell synchronisation</i>	89
2.8.1. Synchronising cells at the G ₁ /S boundary and S phase.....	89
2.8.2. Synchronising cells at the G ₂ /M boundary and M phase.....	90
2.8.3. Cell cycle assessment with propidium iodide staining	91
2.9. <i>Fluorescence in-situ hybridisation (FISH)</i>	91
2.9.1. Fosmid probe details.....	91
2.9.2. Purification of fosmid DNA	92
2.9.3. Nick translation of FISH probes.....	93
2.9.4. Preparation of interphase cells on slides	93
2.9.5. Hybridisation of FISH probes	93
2.9.6. Imaging and analysis of FISH probes.....	94
2.10. <i>Nucleoside incorporation and quantification</i>	95
2.10.1. Radiolabelling with [5- ³ H]-uridine	95
2.10.1.1. Northern blot.....	95
2.10.1.2. Liquid scintillation counting	95
2.10.2. Metabolic labelling with 5-ethynyl uridine	96
2.10.3. Metabolic labelling with 4-thiouridine.....	96
2.10.3.1. Incorporation and quantification.....	96
2.10.3.2. Lysolecithin permeabilisation	96
2.10.3.3. Thiol modification.....	96
2.10.3.4. One step RT-PCR	97
2.10.3.5. Library preparation and sequencing	99
2.10.3.6. SLAM-seq analysis.....	100
2.10.3.7. Nuclear fractionation.....	100
2.11. <i>SAF-A mutation analysis</i>	101
2.11.1. Protein conservation analysis.....	101
2.11.2. 1D-to-3D method.....	101

2.11.3. Mutational density	101
3. CHAPTER 3: UNDERSTANDING SAF-A'S ROLE IN CHROMATIN DECOMPACTION AT MITOTIC EXIT.....	102
3.1. <i>Introduction</i>	102
3.2. <i>Live cell imaging of cells during mitosis using SNAP-tag</i>	105
3.2.1. Optimising the cellular and imaging conditions.....	105
3.2.2. Co-localisation of SAF-A and DNA from mitosis to early-G ₁	111
3.2.3. Co-localisation of DNA and HNRNPC1 from mitosis to early-G ₁	117
3.2.4. Co-localisation of DNA and HNRNPA1 from mitosis to early-G ₁	121
3.3. <i>SAF-A's interaction with RNA during the cell cycle</i>	125
3.3.1. SAF-A oligomerisation.....	125
3.3.2. SAF-A oligomer immunoprecipitation	127
3.3.3. SAF-A and RNA immunoprecipitation.....	128
3.3.4. Cell synchronisation to obtain a pure mitotic cell population.....	135
3.3.5. SAF-A and total RNA immunoprecipitation during mitosis and early G ₁	144
3.3.6. Interaction with newly transcribed RNAs during mitosis and early G ₁	146
3.4. <i>Discussion</i>	146
4. UNDERSTANDING THE NATURE OF SAF-A'S INTERACTION WITH RNA	154
4.1. <i>Introduction</i>	154
4.2. <i>Methods for SAF-A depletion</i>	156
4.2.1. Generating a mini-Auxin-Inducible Degron (mAID) tagged SAF-A cell line	156
4.2.2. Generating a degradation tag (dTAG) tagged SAF-A cell line.....	163
4.2.3. RNAi depletion of SAF-A at an earlier time point	165
4.3. <i>SAF-A and the RNA mesh</i>	167
4.3.1. Super-resolution microscopy of nuclear proteins.....	167
4.3.2. Super-resolution microscopy of SAF-A with transcription inhibition	171
4.3.3. Visualising nascent RNA with 5-EU	173
4.3.4. Super-resolution microscopy of SAF-A and RNA	176
4.3.5. Super-resolution microscopy of RNA with SAF-A depletion.....	181
4.4. <i>Amplicon SLAM-seq</i>	183
4.4.1. Method optimisation	183
4.4.2. Pilot experiment	188
4.4.3. Comparison between different thiol modification methods.....	195
4.4.4. Nascent RNA dynamics	199
4.5. <i>Discussion</i>	206
5. CHAPTER 5: SAF-A PATIENT MUTATIONS AND THEIR EFFECT ON ITS ROLE IN CHROMATIN ORGANISATION	215
5.1. <i>Introduction</i>	215
5.2. <i>Selection of patient mutations</i>	217
5.2.1. Protein conservation	217
5.2.2. Mutations in the general population.....	219
5.2.3. Analysis of clinical variants in the literature	221
5.2.4. Criteria for patient mutation selection for further experiments.....	229
5.3. <i>Effect of SAF-A mutations in human cell lines</i>	231
5.3.1. Making cell lines expressing SAF-A mutants.....	231
5.3.2. Cell growth of patient mutation cells.....	240

5.3.3. Changes in chromatin decompaction due to patient mutations.....	246
5.3.4. Effect on SAF-A oligomerisation.....	249
5.4. <i>Discussion</i>	252
6. CHAPTER 6: DISCUSSION.....	258
6.1. <i>Summary</i>	258
6.2. <i>SAF-A is recruited to chromatin during mitotic exit in a transcription-dependent manner</i>	259
6.3. <i>Oligomeric SAF-A and monomeric SAF-A differ in RNA association during the cell cycle</i>	260
6.4. <i>SAF-A depletion affects the nascent interconnected RNA mesh</i>	261
6.5. <i>SAF-A mutations from patients with the HNRNPU-related neurodevelopmental disorder</i>	262
6.6. <i>Concluding remarks</i>	263
REFERENCES	264
APPENDIX A	311
APPENDIX B	318
APPENDIX C	320
APPENDIX D	325

List of Figures

Figure 1.1. Overview of chromatin organisation and accessibility.....	20
Figure 1.2. The dynamic nuclear mesh.....	27
Figure 1.3. Cohesin complex structure and dynamics.....	35
Figure 1.4. Chromatin fibre models and the large-scale genomic organisation....	40
Figure 1.5. Cell division and mitotic condensation.....	44
Figure 1.6. SAF-A and orthologs protein diagram.....	50
Figure 1.7. SAF-A interphase organisation.....	53
Figure 1.8. SAF-A and RNA localisation during mitosis.....	55
Figure 1.9. SAF-A mutations and potential mechanism.....	58
Figure 3.1. Re-establishment of gene expression during late mitosis and early G ₁	103
Figure 3.2. Optimisation of live cell imaging conditions of FlpIn 293 T-REx cells expressing SNAP-SAF-A.....	107
Figure 3.3. Synchronisation aids for live cell imaging of FlpIn 293 T-REx cells....	110
Figure 3.4. Live cell imaging of SNAP-SAF-A and co-localisation with DNA in FlpIn 293 T-REx cells.....	113
Figure 3.5. Live cell imaging of SNAP-HNRNPC1 and co-localisation with DNA in FlpIn 293 T-REx cells.....	119
Figure 3.6. Live cell imaging of SNAP-HNRNPA1 and co-localisation with DNA in FlpIn 293 T-REx cells.....	123
Figure 3.7. SAF-A endogenous protein oligomerisation.....	126
Figure 3.8. SAF-A protein oligomer immunoprecipitation.....	127
Figure 3.9. SAF-A RNA immunoprecipitation.....	129
Figure 3.10. [5- ³ H]U-RNA incorporation.....	131
Figure 3.11. <i>In vivo</i> coupling of SAF-A with [5- ³ H]U-RNA.....	134
Figure 3.12. Cell synchronisation of different cell types.....	137
Figure 3.13. Testing of mitotic degron (MD) as a synchronisation aid.....	140
Figure 3.14. Cell synchronisation of the HCT 116 cell line with a double thymidine block.....	142
Figure 3.15. SAF-A coupled with [5- ³ H]U-RNA during mitosis and early G ₁	145
Figure 4.1. Depletion of transiently transfected FLAG-mAID-SAF-A in HCT 116 cells.....	157
Figure 4.2. Tagging endogenous SAF-A with mAID in HCT 116 Tet-OsTIR1 cells.....	158
Figure 4.3. Cell growth of mAID-SAF-A and HCT 116 Tet-OsTIR1 cells.....	160
Figure 4.4. Testing of HCT116 TIR1 cells expressing SAF-A-AID.....	162
Figure 4.5. Depletion of transiently transfected SAF-A tagged with FKBP12 ^{F36V} in HCT 116 cells.....	164
Figure 4.6. SAF-A depletion using RNAi in RPE-1 and HCT 116 cells.....	166
Figure 4.7. Super-resolution microscopy of nuclear proteins using dSTORM.....	169
Figure 4.8. Effect of transcription inhibition on the SAF-A structures.....	172

Figure 4.9. Visualising RNA with click chemistry.....	174
Figure 4.10. Two-colour super-resolution microscopy of RNA and SAF-A using dSTORM.....	178
Figure 4.11. Effect of SAF-A depletion on the RNA mesh.....	182
Figure 4.12. 4sU incorporation in FlpIn 293 T-REx and HCT 116 cells.....	184
Figure 4.13. Lysolecithin trial.....	186
Figure 4.14. Nascent RNA labelling with 4sU and amplicon sequencing.....	189
Figure 4.15. Comparison of thiol modification methods.....	196
Figure 4.16. Amplicon sequencing of RNA in the cytoplasm and nucleus.....	201
Figure 5.1. <i>In silico</i> approach.....	216
Figure 5.2. Protein sequence conservation of full-length SAF-A.....	218
Figure 5.3. Population-wide mutations in SAF-A.....	219
Figure 5.4. Pathogenic mutations in SAF-A overlayed on protein sequence conservation.....	224
Figure 5.5. Pathogenic mutations in SAF-A.....	227
Figure 5.6. Pathogenic mutation in Walker A motif of SAF-A.....	230
Figure 5.7. Expression of SAF-A patient mutants in HCT 116 cells.....	232
Figure 5.8. FlpIn 293 T-REx cell line generation with selected SAF-A patient mutations.....	234
Figure 5.9. SAF-A depletion using RNAi in FlpIn 293 T-REx cells.....	236
Figure 5.10. SAF-A depletion using different concentrations of RNAiMAX reagent in FlpIn 293 T-REx cells.....	238
Figure 5.11. HNRNPUL1 depletion using RNAi.....	239
Figure 5.12. Cell growth of FlpIn 293 T-REx cell lines depleted of SAF-A with RNAi.....	242
Figure 5.13. Patient mutation effect on cell growth of HCT 116 cells depleted of SAF-A with RNAi.....	245
Figure 5.14. Patient mutation rescue of chromatin compaction after HCT116 cells are depleted of SAF-A using RNAi.....	248
Figure 5.15. Effect of patient mutations on SAF-A protein oligomerisation.....	251
Figure 1.A. Vector maps of plasmids used for FlpIn 293 T-REx cell line generation.....	311
Figure 2.A. Vector maps of plasmids used in DNA transfections.....	313
Figure 3.A. Vector maps of plasmids used in CRISPR Cas9 cell line generation.....	316
Figure 4.A. Vector maps of plasmids used for amplification or digestion to make other plasmids.....	317

List of Tables

Table 1.1. SAF-A mutations by type and protein region.....	60
Table 2.1. Reagents.....	63
Table 2.2. List of antibodies.....	66
Table 2.3. Kits.....	67
Table 2.4. Solutions and buffers.....	68
Table 2.5. Primers for plasmid PCR amplification.....	71
Table 2.6. Primers for genomic DNA PCR amplification.....	72
Table 2.7. Primers for inverse PCR amplification.....	72
Table 2.8. Primers for genotyping PCR amplification.....	73
Table 2.9. Primers for qPCR.....	75
Table 2.10. Plasmids for mammalian transfection.....	78
Table 2.11. Target sites and gRNA sequences to target CRISPR Cas9 to the ATG codon of SAF-A.....	81
Table 2.12. Plasmids for cell line generation using CRISPR Cas9.....	81
Table 2.13. Western blotting antibodies.....	84
Table 2.14. Protein immunofluorescence antibodies.....	85
Table 2.15. dSTORM antibodies.....	86
Table 2.16. SNAP-tag fluorescent substrates.....	88
Table 2.17. Synchronisation methods.....	90
Table 2.18. Fosmid probes for FISH.....	92
Table 2.19. Primers for pilot RT-PCR.....	97
Table 2.20. Barcoded primers for RT-PCR.....	98
Table 2.21. Indexed primers for PCR amplification.....	99
Table 3.1. Average length of mitotic phases for live cell imaging conditions.....	112
Table 3.2. Propidium iodide and H3S10ph signal after a 2xThy block release.....	141
Table 4.1. Sequencing reads of amplicons for each 4sU treatment condition.....	193
Table 5.1. Population-wide synonymous mutation ratios per domain.....	220
Table 5.2. Population-wide missense mutation ratios per domain.....	220
Table 5.3. SAF-A pathogenic mutations in clinical patients.....	221
Table 5.4. All pathogenic mutation ratios per domain.....	228
Table 5.5. Only substitution and deletion pathogenic mutation ratios per domain	228
Table A. Gibson Assembly and ligation DNA constructs.....	318
Table B. HNRNPU-related neurodevelopmental disorder patients and phenotype information from literature.	325

Abbreviations

1NM-PP1	4-amino-1-tert-butyl-3-(1'-naphthylmethyl) pyrazolo[3,4-d]pyrimidine
2xThy	Double Thymidine Block
3C	Chromosome Conformation Capture
3D	Three-Dimensional
4sU	4-Thiouridine
5-EU	5-Ethynyl Uridine
5-Ph-IAA	5-Phenyl-Indole-3-acetic acid
[5-3H]-U	Tritiated Uridine
Å	Angstrom
AAA	ATPases Associated with diverse cellular Activities
actD	Actinomycin D
AD	Alzheimer's Disease
AF	AlexaFluor
AID	Auxin-Inducible Degron
ALS	Amyotrophic Lateral Sclerosis
ATP	Adenosine Triphosphate
AUX/IAA	Auxin/Indole-3-acetic acid
bp	Base Pairs
Blast	Blasticidin S
BM(PEG) ₂	1,8-bismaleimido-diethyleneglycol
BSA	Bovine Serum Albumin
caRNA	Chromatin-associated RNA
CDK1	Cyclin-dependent Kinase 1
ChIP	Chromatin Immunoprecipitation
Chr	Chromosome
ChromEMT	ChromEM Tomography
circRNA	Circular RNA
CLIP	Cross-linking Immunoprecipitation
cpm	Counts per Minute
Cryo-EM	Cryogenic Electron Microscopy
CTCF	CCCTC-Binding Factor
DAPI	4,6-diamino-phenylindole
DDD	Deciphering Developmental Diseases
ddH ₂ O	Double Distilled Water
DDR	DNA Damage Response
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
Dox	Doxycycline
DRB	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DSS	Disuccinimidyl Suberate
dSTORM	Direct Stochastic Optical Reconstruction Microscopy
dTAG	Degradation Tag
DTT	Dithiothreitol
ECL	Enhanced Chemiluminescence
EDTA	Ethylenedinitrilotetraacetic Acid

EdU	5-Ethynyl-2'-deoxyuridine
EGFP	Enhanced Green Fluorescent Protein
EM	Electron Microscopy
EMSA	Electrophoretic Mobility Shift Assay
EMT	Electron Microscope Tomography
EtBr	Ethidium Bromide
EtOH	Ethanol
FACS	Fluorescently Activated Cell Sorting
FAIRE-MS	Formaldehyde-Assisted Isolation of Regulatory Elements-coupled Mass Spectrometry
FCS	Foetal Calf Serum
FISH	Fluorescent <i>in situ</i> Hybridisation
FRT	Flp Recombination Target
γH2A.X	H2A.X phosphorylated at serine 139
GMBS	N-γ-maleimidobutyl-oxysuccinimide ester
gnomAD	Genome Aggregation Database
gRNA	Guide RNA
HDR	Homology-Directed Repair
HELB	DNA Helicase B
HF	High Fidelity
Hi-C	High-throughput Chromosome Conformation Capture
hnRNP	Heterogeneous Nuclear Ribonucleoprotein
HPLC	High Performance Liquid Chromatography
hr	Hour
HR	Homologous Recombination
HRP	Horseradish Peroxidase
hTERT	Human Telomerase Reverse Transcriptase
Hyg	Hygromycin B
IAA	2-Iodoacetamide
IDR	Intrinsically Disordered Region
IF	Immunofluorescence
IP	Immunoprecipitation
kb	Kilo Base
KD	Knockdown
kDa	Kilo Dalton
KH	HNRNPK-Homology
LAD	Lamina-Associated Domain
LB broth	Luria-Bertani Broth
LINE	Long Interspersed Nuclear Element
LL	Lysolecithin
LLPS	Liquid-Liquid Phase Separation
lncRNA	Long Non-coding RNA
LOF	Loss of Function
LTR	Long Terminal Repeat
mAID	Mini Auxin-Inducible Degron
MD	Mitotic Degron
miRNA	Micro RNA
mRNA	Messenger RNA
MSA	Multiple Sequence Alignment
MW	Molecular Weight

ncRNA	Non-coding RNA
NGS	Next Generation Sequencing
NHEJ	Non-homologous End Joining
NMD	Nonsense Mediated Decay
OD	Optical Density
Oma	Orthologous Matrix
ORF	Open Reading Frame
OsTIR1	<i>Oryza Sativa</i> Transport Inhibitor Response 1
PAR-CLIP	Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
ph	Phosphorylation
PI	Propidium Iodide
pLDDT	Per-residue Confidence Score
pLI	Probability of Loss of Function Intolerance
PMSF	Phenylmethyl Sulfonyl Fluoride
Pol II	RNA Polymerase II
Poly(A)	Poly(Adenine)
pre-mRNA	Precursor mRNA
PSF	Point-spread function
PTM	Post-Translational Modification
PVDF	Polyvinylidene Difluoride
qPCR	Quantitative Polymerase Chain Reaction
RAP-seq	RNA Antisense Purification Sequencing
RBD	RNA Binding Domain
RBP	RNA Binding Protein
RIP	RNA-Immunoprecipitation
RNA	Ribonucleic Acid
RNAi	RNA-interference
ROI	Region of Interest
RPE1 cells	Retinal Pigment Epithelial cells
rpm	Revolutions per Minute
rRNA	Ribosomal RNA
RRM	RNA Recognition Motif
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
RT-qPCR	Reverse Transcriptase Quantitative Polymerase Chain Reaction
SAF-A	Scaffold Attachment Factor A
SCN	Suprachiasmatic Nucleus
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
SINE	Short Interspersed Elements
siRNA	Short Interfering RNA
SLAM-seq	Thiol(SH)-Linked Alkylation for the Metabolic Sequencing of RNA
SNP	Single Nucleotide Polymorphism
snRNA	Small Nuclear RNA
snoRNA	Small Nucleolar RNA
SPRY	SPLa and the Ryanodine Receptor
SSC	Saline Sodium Citrate

STED	Stimulated Emission Depletion
STORM	Stochastic Optical Reconstruction Microscopy
TAD	Topologically Associated Domain
TAE	Tris-Acetate-EDTA
TBE	Tris/Borate/EDTA
TE	Tris/EDTA
tRNA	Transfer RNA
trp	Triptolide
UMI	Unique Molecular Identifier
UV	Ultraviolet
Van complex	Ribonucleoside Vanadyl Complex
WA	Walker A
WB	Walker B
WT	Wild Type
Xi	Inactive X Chromosome
XIST	X-inactive specific transcript

1. Chapter 1: Introduction

1.1. Chromatin organisation and the cell cycle

The organisation of chromatin during interphase follows a hierarchical order, starting with DNA wrapped around histone octamers, which, along with linker histones, form nucleosomes, the fundamental units of chromatin structure (Figure 1.1A). In its unfolded state (for example, in low salt conditions), the chromatin fibre is often described as “beads-on-a-string” with individual nucleosomes having a 10 nm diameter. Through electrostatic interactions, DNA and histone modifications (see Section 1.1.1.1.), chromatin remodelling proteins (see Section 1.1.1.2.), non-coding RNA interactions and RNA binding proteins (see Sections 1.1.1.3. and 1.1.1.4.), and the functions of other important proteins (see Section 1.1.1.5.), chromatin further assembles into higher-order structures (Clark and Kimura, 1990; Bannister and Kouzarides, 2011; Cech and Steitz, 2014; Meller *et al.*, 2015; Clapier *et al.*, 2017; Edwards *et al.*, 2017). SAF-A (Scaffold Attachment Factor-A; see Section 1.2.), an RNA binding protein is hypothesised to interact with non-coding RNAs and could form a dynamic protein-RNA mesh which regulates the interphase chromatin structure (Nozawa *et al.*, 2017; see Section 1.2.1.). *In vitro*, the higher-order chromatin structure resembles a 30-nm diameter regularly packaged fibre, while *in vivo* it has a disordered nature with fibres between 5 and 24 nm in diameter (Langmore and Paulson, 1983; Eltsov *et al.*, 2008; Maeshima *et al.*, 2010; Maeshima *et al.*, 2019). These fibres are arranged into topologically associated domains (sub-TADs and TADs; see Section 1.1.2.) which associate with each other based on gene activity which finally leads to the largest unit of chromatin organisation – the chromosome (Cremer and Cremer, 2010; see Section 1.1.3.).

Proliferating cells have four distinct and sequential cell cycle phases (Gap 1/G₁, Synthesis/S, Gap 2/G₂, Mitosis/M; Figure 1.1B) which require cell cycle checkpoints to be fulfilled before passing through them and ensure that the DNA is accurately replicated before division. The duration of these phases can vary between different cell types like stem cells having a shorter G₁ phase duration compared to somatic cells (Becker *et al.*, 2006). Environmental factors can also influence cell cycle

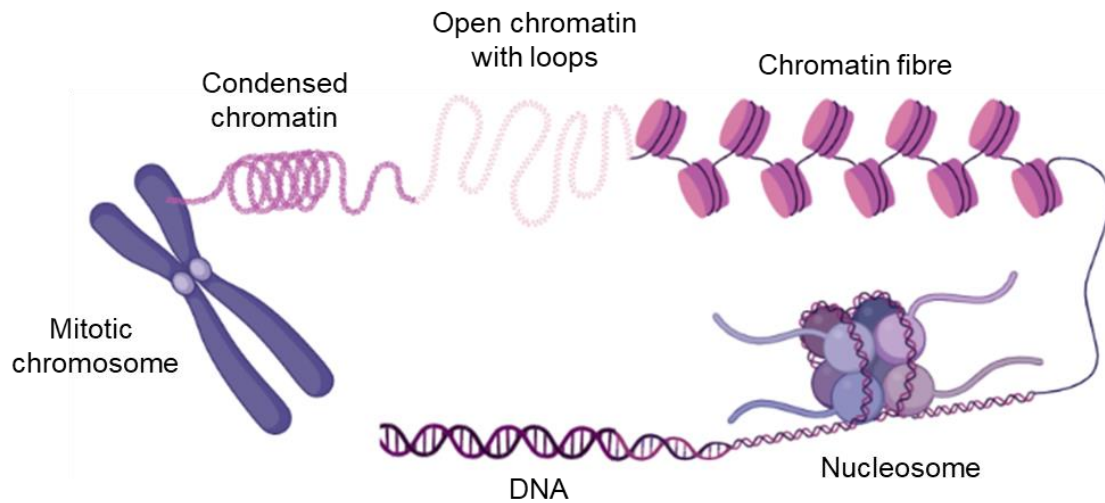
duration like DNA damage prolonging G₁ and G₂ phases (Chao *et al.*, 2017). Cell cycle duration can even vary across single cells derived from the same clone while being grown in the same conditions (Dawson *et al.*, 1965, Chao *et al.*, 2019, Stallaert *et al.*, 2022). The cell cycle checkpoints describe various requirements that need to be fulfilled before the cells can progress to the next phase. These requirements include appropriate cell size at the end of G₁ and G₂, repair of DNA damage before S-phase and mitosis, stabilisation of replication during S-phase, completion of replication, and correct mitotic spindle attachment during mitosis (Barnum and O'Connell, 2014). If cells do not meet these requirements, they are arrested temporarily at the cell cycle checkpoint to allow more time to meet the criteria. Checkpoint control is dependent on complexes made up of a catalytic subunit, the cyclin-dependent kinase (CDK), and a regulatory subunit, the cyclin. The cyclin subunit has specific expression and proteolysis patterns to ensure the complexes are active during the correct cell cycle phase. For example, the G₂/M checkpoint requires the increased activity of CDK1 and cyclin B complex which is dependent on cyclin B levels peaking in G₂ phase and phosphorylation of CDK1 (Gallant and Nigg, 1992). Additional regulation of the CDK-cyclin complexes is dependent on phosphorylation and dephosphorylation of CDKs, inhibition or activation by small molecules from the INK4 and Cip/Kip families, and transcriptional regulation by transcription factor complexes RB-E2F, DREAM, and MMB-FOXM1 (Sherr and Roberts, 1999; Fischer *et al.*, 2016).

The accessibility, structure, and arrangement of chromatin change throughout the cell cycle (Figure 1.1B). Gradually, chromatin becomes more accessible to DNaseI through G₁ phase, peaking during the early S-phase followed by a sharp drop during G₂ and mitosis (Pederson, 1972). The organisation of chromatin varies throughout G₁ phase as chromatin is highly mobile as the chromosomes settle in their respective territories within the nucleus (see Section 1.1.2.). Experiments studying the 3D conformation of chromatin showed that the regions that initiate replication first are closely located despite being set megabases apart on chromosomes (Moindrot *et al.*, 2012). During the G₁ phase, histone marks like acetylation and methylation are deposited on chromatin to reflect the cell's transcriptional programme and to specify the origins of replication where DNA replication begins in S-phase (Raynaud *et al.*, 2014). When the replication fork is established in S-phase, the chromatin fibre is

loosened ahead of it by nucleosome displacement. The maximum accessibility in S-phase coincides with the gradual relaxation of chromatin structure to facilitate replication with DNA partially exposed as old histones are displaced and re-deposited or new ones are incorporated (Cusick *et al.*, 1983; Alabert and Groth, 2012; see Section 1.1.1.1.). Transcription during S-phase can impair DNA replication due to formation of RNA-DNA hybrids, stalled RNA polymerase or topological stress on the chromatin (Deshpande and Newlon, 1996; Huertas and Aguilera, 2003; Tuduri *et al.*, 2009; Saponaro *et al.*, 2014). Therefore, to temporally separate transcription and replication, cells could delay transcription start site replication until G₂/M while continuing to transcribe active genes until they are replicated (Wang *et al.*, 2021). During the G₂ phase, chromatin is prepared for mitosis while it maintains the overall organisation of chromatin established in the late G₁ phase (Lu *et al.*, 2010). Finally, during mitosis, the chromatin is the least accessible as it is condensed into discrete and highly organised chromosomes (see Section 1.1.3.). During mitosis, there was a low level of transcription detected by pulse-labelling of nascent RNA transcripts with a surge of transcriptional activity appearing during telophase before the re-establishment of gene expression throughout G₁ phase (Hsiung *et al.*, 2016; Palozola *et al.*, 2017; see Section 1.1.4.).

SAF-A's role in large-scale chromatin organisation has primarily been investigated during interphase (G₁, S, G₂). Its role in the context of mitosis has only recently been researched while its function at mitotic exit has not been described before (Sharp *et al.*, 2020; see Section 1.2.2.). How mutations in SAF-A, which lead to a rare neurodevelopmental disorder, affect the hypothesised dynamic nuclear mesh it forms is also an area of speculation (Dugger *et al.*, 2023; see Section 1.2.4.). The multiple unknowns about this protein and its hypothesised chromatin organisation function, together with other researched functions (see Section 1.2.3.), make it an appealing target to study to further understand its role in the cell cycle.

A



B

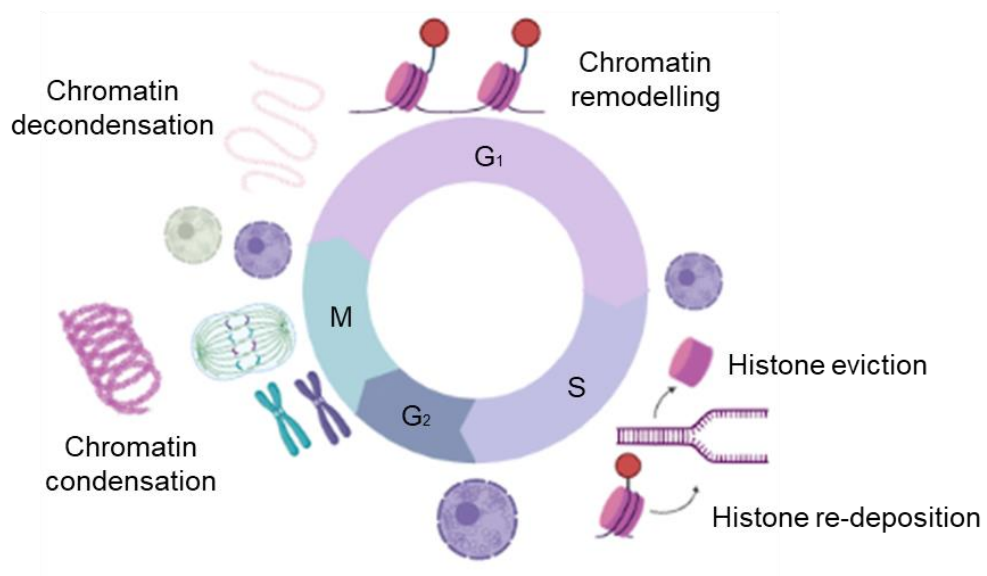


Figure 1.1. Overview of chromatin organisation and accessibility. A) The hierarchy of chromatin organisation is depicted, which can be read starting from the highest compaction (mitotic chromosome) or the lowest compaction (DNA) level. A double strand of DNA is wrapped around eight core histone proteins to make up the nucleosome (Chapter 1.1.1.1.). Nucleosomes interact further to create a chromatin fibre that can be looped by cohesin to facilitate transcriptional processes (Chapters 1.1.1.5. and 1.1.2.). Chromatin is further compacted by condensin to form mitotic chromosomes during prophase (Chapters 1.1.1.5. and 1.1.3.). B) A schematic representation of the cell cycle with distinct phases of interphase (G_1 , S, G_2), and mitosis (M). Chromatin is condensed during mitosis to allow successful segregation of DNA between the daughter cells. During G_1 phase, the transcriptional programming from before mitosis is re-enacted via chromatin remodellers and histone marks are recreated. During S phase, chromatin is loosened to allow replication as histones are evicted and re-deposited, and histone post-translational modifications are recreated. Figure created in BioRender.

1.1.1. Chromatin structural components

1.1.1.1. Histones

The nucleosome is the first level of chromatin organisation where 145 to 147 base pairs (in mammals) of negatively charged DNA are wound around a positively charged core histone scaffold which allows for regulation of key cellular processes like transcription, replication, and DNA repair (Luger *et al.*, 1997). The core histone scaffold is an octamer and it includes two H2A-H2B dimers and an H3-H4 tetramer but can also incorporate histone variants such as H2A.X, H3.3, and CENP-A which play important roles in specific cellular processes (Biterge and Schneider, 2014). The globular histone-fold domains of the core histones facilitate nucleosome assembly by hydrophobic associations, while with their positively charged lysine and arginine residues they bind the negatively charged DNA through electrostatic interactions (Felsenfeld, 1978; Sandman and Reeve, 2000). Additionally, there are core histone N- and C-terminal tails extending out from their globular domains which often have post-translational modifications (PTMs) that affect chromatin structure and function (Kouzarides, 2007). Adjacent nucleosomes are linked together by the linker histone H1 (or its variant H5), which protects the linker DNA and stabilises its interaction with the nucleosome (Hill, 2001). Similar to the core histones, H1 contains a globular domain, C-terminal and N-terminal tails, and can undergo PTMs. The globular domain binds to the DNA, while the tails contribute to stabilising the binding of the linker histones to the core histones and the linker DNA and facilitate packaging into a higher-order chromatin structure (Allan *et al.*, 1980; Allan *et al.*, 1986).

While the canonical histones are incorporated into nucleosomes during DNA replication and are expressed during S-phase, the incorporation of histone variants is not dependent on replication and instead they can be expressed throughout the cell cycle (Albig and Doenecke, 1997). The incorporation of core histone variants is regulated by specific histone chaperones like CAF-1, RCAF, and NAP-1/2 (Krude and Keller, 2001; Biterge and Schneider, 2014). Histone chaperones are acceptors of newly synthesised histones or those ejected by the replication fork during S-phase and deposit them onto the newly replicated DNA. For example, the acetylated H3-H4 dimer is placed by the chaperone CAF-1 which is recruited by the replication clamp PCNA to sites of active replication (Shibahara and Stilman, 1999). By interacting with

chromatin modifiers and having structural differences from canonical histones, the variant histones can affect nucleosomal stability and modulate the chromatin conformation. Histone H2A has the most variants which include H2A.X, H2A.Z, H2A.Bbd, and macroH2A. Out of them, macroH2A is the largest histone variant, close to three times the size of canonical H2A, and it possesses an affinity for NAD metabolites (Till and Ladurner, 2009). Nucleosomes containing macroH2A closely resemble canonical nucleosomes, except for a small amino acid difference in the L1 loop which increases nucleosome stability and further compacts chromatin (Chakravarthy and Luger, 2006; Muthurajan *et al.*, 2011). This effect links macroH2A to transcriptional repression, as it is also enriched on the transcriptionally inactive X chromosome (Xi) and the inactive alleles of imprinted genes (Costanzi and Pehrson, 1998; Choo *et al.*, 2006). In contrast, H2A.X is linked to destabilising nucleosomal structure and its phosphorylation on serine 139 recruits DNA damage repair machinery (Rogakou *et al.*, 1998; Li *et al.*, 2010). Histone H3 has two variants, H3.3 and CENP-A, which also have different roles and effects on nucleosomal stability. H3.3 differs from the canonical H3.1 protein by five amino acids and replaces it at transcriptionally active genes, resulting in an open chromatin conformation (Ahmad and Henikoff, 2002; Chen *et al.*, 2013). CENP-A is specifically recruited to centromeres, compacting the nucleosomal arrays at those sites and facilitating the recruitment of kinetochore components (Régnier *et al.*, 2005; Panchenko *et al.*, 2011). Linker histone H1 also has 6 variants which are expressed in a replication-dependent manner (H1.1 to H1.5 and H1t expressed in testis), as well as 5 variants in humans which are expressed throughout the cell cycle and incorporated into the chromatin independently of replication (H1.0 and H1.X expressed in somatic cells, H1T2 and H1LS1 expressed in testis, and H1oo expressed in ovaries; Izzo *et al.*, 2008). The germline-specific variants are associated with chromatin condensation at different stages of spermatogenesis and fertilization (Yan *et al.*, 2003; Martianov *et al.*, 2005; Mizusawa *et al.*, 2010). H1.0 might have a role in regulating pluripotency genes during differentiation, while H1.X is enriched in more compact chromatin domains (Happel *et al.*, 2005; Terme *et al.*, 2011).

PTMs on the N- and C-terminal tails of histones also modulate chromatin structure directly by altering the charge of protein residues, thereby affecting DNA interactions, or indirectly, by recruiting enzyme complexes which recognize the modifications.

There are eight different classes of modifications which can be performed on over sixty different residues (Kouzarides, 2007). The complex crosstalk between modifications on the same or different histone tails has been likened to a histone code. The most studied modifications are phosphorylation, acetylation, and methylation, with the last adding extra complexity due to residues being able to be mono-, di-, or trimethylated. Phosphorylation can occur on serine, threonine, and tyrosine residues of histone tails, which can directly affect chromatin structure due to its negative charge or recruit reader or effector proteins with phosphor-binding domains (Rossetto *et al.*, 2012). As mentioned earlier, the histone variant H2A.X can be phosphorylated at serine 139 (γ H2A.X) as part of the DNA-damage response pathways by the kinases ATM, ATR, and DNA-PK (Shroff *et al.*, 2004). In response to a DNA break, the mark spreads over several megabases of DNA, creating a platform for retaining DNA damage signalling and repair proteins like MDC1 (Rogakou *et al.*, 1999; Stucki *et al.*, 2005). It is also thought that γ H2A.X can recruit chromatin remodelling complexes like INO80 and SRCAP, which would make DNA more accessible to aid the repair process (Karagiannis and El-Osta, 2007). After the DNA break has been repaired, the mark is removed via histone exchange facilitated by INO80 and by dephosphorylation by the phosphatases PP2A, PP4, PP6, and Wip1 (van Attikum *et al.*, 2007; Rossetto *et al.*, 2012). H2A.X is also phosphorylated on tyrosine 142 which is also important for the DNA damage response. This phosphorylation is done by the WICH complex and dephosphorylated by the EYA1/3 phosphatase (Xiao *et al.*, 2009). In contrast to γ H2A.X, the levels of H2A.XY142ph decrease during the DNA damage response, which is important for maintaining the former's DNA repair foci (Cook *et al.*, 2009). If mammalian cells are deficient in H2A.X, they become highly sensitive to ionizing radiation with increased levels of genomic instability (Bassing *et al.*, 2002).

As histones are highly conserved proteins critical for DNA packaging, mutations in them would lead to potentially severe disease or might not be tolerated. Searching the DECIPHER database of genetic variants and associated phenotypes, not every gene for canonical histone proteins is associated with a disease (Firth *et al.*, 2009). For instance, 17 genes are coding for H2A but out of them only 2 (H2AC6 and H2AC21) have been implicated in patients with a disease. Similarly, among the 18 genes encoding H2B, only 1 (H2BC21) has been linked to disease, for H3 out of 15

genes – only 2 (H3C12 and H3-4), and for H4 out of 15 genes – only 3 (H4C2, H4C3, H4C11). Mutations in histones have also been linked to multiple tumour types and can occur in both the globular domains and at their tails, affecting PTMs (Nacev *et al.*, 2019). For example, mutations at the nucleosome anchoring point are frequently observed, particularly at E56 and E59 in H2A and at E113 in H2B, with E56K/Q found in uterine, ovarian, lung, renal, head and neck, pancreatic and rectal cancer (Zhao *et al.*, 2016). Out of the histone variants, DECIPHER shows a high loss-of-function (LOF) intolerance for the genes responsible for macroH2A with a probability of LOF intolerance (pLI) scores of 0.74 (MACROH2A1) and 0.98 (MACROH2A2). A pLI score closer to 1 indicated a higher intolerance to mutations which might lead to haploinsufficiency (Karczewski *et al.*, 2020). The gene H2AZ1, which encodes the H2A.Z histone variant, has a pLI of 0.88. The high pLI scores underscore the importance of macroH2A and H2A.Z in transcriptional repression. Mutations in the coding genes of the histone variant H3.3 have been associated with Bryant-Li-Bhoj neurodevelopmental syndrome 1 and 2 (OMIM 619720 and 619721) which led to altered PTMs and differences in gene regulation (Bryant *et al.*, 2020). Linker histones and their variants also have variable associations with diseases. Truncation mutations in the gene coding for H1.4 have been implicated in Rahman syndrome (OMIM 617537) and autism with the C-terminal tail truncation possibly causing a defect in epigenetic regulation (Tatton-Brown *et al.*, 2017; Duffney *et al.*, 2018). Many of these links to patient mutations have occurred only in recent years, suggesting that more canonical core and linker histones, as well as their variant genes, may be implicated in human diseases in the future.

1.1.1.2. *Chromatin remodellers*

Chromatin interacts with a wide range of proteins that have a multitude of roles. Out of these, chromatin remodellers rank high on the scale of importance as they can change nucleosome positioning or composition in an active process utilising ATP hydrolysis. These changes adjust the accessibility of the DNA which is vital for cellular processes like transcription, DNA repair, and cell cycle progression. Some remodellers have been briefly mentioned so far and they belong to the SWI/SNF protein family which can be divided further into four subfamilies – CHD, INO80, ISWI, and SWI/SNF (Gorbalenya and Koonin, 1993). Out of them, the INO80 subfamily is

the most evolutionarily conserved due to a high degree of homology in the ATPase subunit (Morrison and Shen, 2009). They are transcriptional regulators with the two most well-known complexes being INO80 (consisting of INO80, RUVBL1, RUVBL2, β -Actin, actin-related proteins, and non-conserved subunits) and SRCAP (consisting of SRCAP, RUVBL1 and RUVBL2, GAS41, β -Actin, actin-related proteins, and non-conserved subunits). The presence of Ruv-B-like helicases (RUVBL1, RUVBL2) is what distinguishes the INO80 subfamily from other chromatin remodelling families. The RUVBL proteins were originally discovered as DNA repair factors which hints at the complexes' role in the DNA damage response pathways (Tsaneva *et al.*, 1992). The SRCAP complex replaces the canonical H2A-H2B dimers with an H2A.Z histone variant at promoters while INO80 catalyses the opposite reaction (Mizuguchi *et al.*, 2004; Wong *et al.*, 2007; Papamichos-Chronakis *et al.*, 2011). INO80 also removes the histone variant H2A.X after a DNA strand break is repaired, although both complexes are recruited to γ H2A.X to facilitate chromatin access for other DNA repair proteins (van Attikum *et al.*, 2007). In cells depleted of INO80 there are DNA repair deficiencies, specifically affecting homologous recombination (HR) which links INO80 to 5'-3' resection of double-strand break ends (Gospodinov *et al.*, 2011). SRCAP has similarly been linked to DNA-end resection (Dong *et al.*, 2014). The cellular localisation of INO80 changes throughout the cell cycle, with it being mainly in the cytoplasm during G₁ then localising to chromatin at replication forks during S-phase (Hur *et al.*, 2010). Absence of the complex leads to slower progression during the S-phase, replication fork collapse, and H2A.Z misincorporation.

The SWI/SNF subfamily of complexes is the prototype chromatin remodellers with the three main complexes in mammals being BAF, PBAF, and GBAF/ncBAF. They can move nucleosomes by sliding as well as eject them to create nucleosome-depleted regions for transcriptional regulation during cell differentiation and lineage specification (Boeger *et al.*, 2004; Dechassa *et al.*, 2010). The first two complexes contain the ATPase subunit BRG1/BRM, shared core subunits like BAF47, BAF57, BAF60a/b/c, BAF155, and BAF170, and other complex-specific subunits. GBAF differs from them by the exclusion of some canonical subunits and the inclusion of BRD9 and GLTSCR1 or GLTSCR1L (Alpsoy and Dykhuizen, 2018). The complexes are highly enriched at enhancers and play a role in modulating enhancer accessibility for transcription factor binding (Alver *et al.*, 2017; Wang *et al.*, 2017). In

different cell types, the BAF complex can exchange subunits to generate cell-type specific complexes. For example, at cell cycle exit and differentiation of neural progenitor cells switch from containing BAF53a, BAF45a/d, and SS18 to the neuron specific BAF53b, BAF45 b/c, and CREST with switch perturbation leading to a cell cycle progression block (Lessard *et al.*, 2007; Braun *et al.*, 2021). PBAF complexes differ in chromatin-binding modules' isoforms (PHF10) depending on whether cells are actively proliferating or differentiating. In hematopoietic stem cells, PHF10 is required for population maintenance, while in proliferating myelocytes the isoform PHF10-P is used, other complexes include PHF10-S in myeloid progenitors and their progeny, with PHF10-Ss being a component of the PBAF complex in neutrophils (Krasteva *et al.*, 2017; Viryasova *et al.*, 2019). The above examples show the heterogeneity of SWI/SNF complexes in different cells and their importance. They have also been linked to the DNA damage repair pathways with BAF and PBAF implicated in HR and non-homologous end joining (NHEJ; Brownlee *et al.*, 2015). Specifically, the common ATPase subunit, BRG1, is recruited to DNA breaks and assists in repair mechanisms (Qi *et al.*, 2015). It also has been shown to reduce nucleosome density at DNA damage sites and to induce the phosphorylation of histone variant H2A.X to further promote repair (Park *et al.*, 2006; Chen *et al.*, 2019). Mutations in SWI/SNF subunits have been implicated in nearly a quarter of cancers, presumably owing to the importance of the complexes in DNA repair and transcriptional regulation (Mittal and Roberts, 2020). Core subunit mutations are also frequently associated with neurodevelopmental disorders with mutations in BAF47 and BRG1/BRM causing variations of Coffin-Siris syndrome (OMIM 614608 and 614609) while mutations in BAF250a/b, a BAF-specific subunit, have been deemed to be the most frequent cause of *de novo* intellectual disability (Tsurusaki *et al.*, 2012; Kosho *et al.*, 2014; Fitzgerald *et al.*, 2015). As the mutations are usually heterozygous and lead to LOF, they implicate a dosage-sensitive role for the subunits in the developing human brain.

1.1.1.3. *Chromatin-associated RNAs and the RNA mesh*

Approximately 95% of synthesised RNA remains in the nucleus instead of being exported to the cytoplasm for protein translation (Nozawa and Gilbert, 2019). The remaining RNA is found to be associated with chromatin, hence termed chromatin-

associated RNA (caRNA), and together with RNA-binding proteins forms a dynamic nuclear mesh (Figure 1.2). RNA itself is dynamic and rapidly turned over, with different types having multiple roles in the cell. Apart from rRNAs, mRNAs, and tRNAs, nuclear RNA types include the non-coding RNAs (ncRNAs) such as small nuclear RNAs (snRNA) that are components of the spliceosome, small nucleolar RNAs (snoRNA) responsible for ribosome biogenesis, long ncRNAs (lncRNA) like XIST, HOTAIR, and NEAT1 which have diverse functions, circular RNAs (circRNA) implicated in cell differentiation and homeostasis, microRNAs (miRNA) which can regulate or repress mRNA translation, long terminal repeats (LTR) representing endogenous retroviruses, long interspersed nuclear element-1 (LINE-1) and short interspersed elements (SINE) which are non-LTR groups of transposable elements (Palazzo and Lee, 2015).

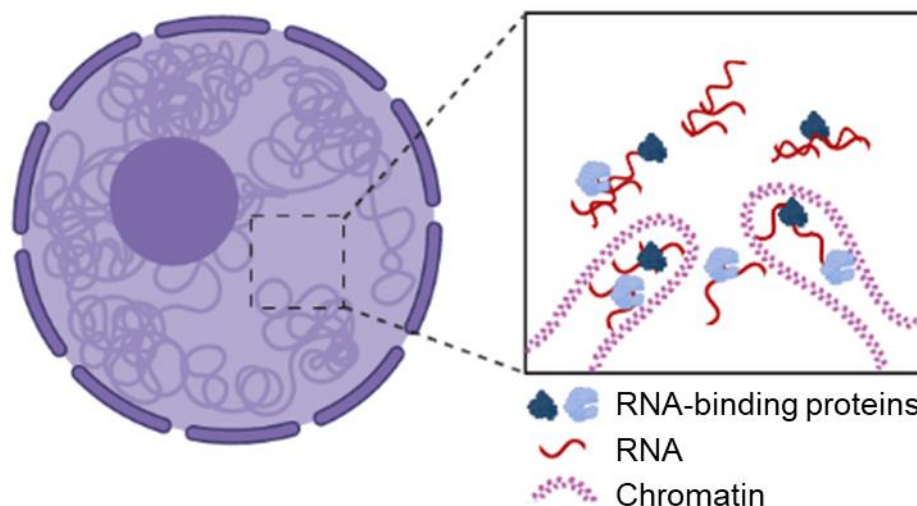


Figure 1.2. The dynamic nuclear mesh. In the nuclei of mammalian cells, RNA-binding proteins have been recently shown to form a dynamic nuclear mesh with ncRNA that affects interphase chromatin structure. This mesh so far has been shown to maintain an open chromatin structure to possibly regulate transcription. Figure created in BioRender.

Some types of nuclear RNAs have specific functions, while a substantial portion (spliced introns, intergenic RNA) is often considered to be a by-product of transcription as they have been found around transcription factories composed of clustered and transiently immobilised RNA polymerase II (RNA pol II) complexes

based at discrete genomic sites (Rieder *et al.*, 2012; Caudron-Herger *et al.*, 2015). Most human genes contain introns with a large part of these sequences being over 50 kb in length. Transcribing such large introns takes approximately 15 min for RNA pol II, making it a very time-consuming and energetically costly process for the cell (Shepard *et al.*, 2009; Jonkers *et al.*, 2014). These long by-products would rapidly accumulate if they weren't degraded by the exosome and the exonuclease XRN2 (Miki and Großhans, 2013). Before they are degraded, non-exported nuclear RNAs are hypothesised to assist in chromatin organisation. For certain ncRNAs the effect is clear – XIST (X-inactive specific transcript) is found to localise at the inactive female X chromosome (Xi) territory and forms the Barr body, which is crucial for silencing the Xi (Clemson *et al.*, 1996). Although XIST is expressed from the soon-to-be inactive X chromosome, it needs to migrate across the Xi which is challenging in the high viscosity of the nucleus. A study used RNA antisense purification sequencing (RAP-seq) to show that transiently expressed XIST migrates along the surface of the X chromosome by utilising the local 3D architecture (Engreitz *et al.*, 2013). Once at distal Xi sites, XIST needs to be secured to it which is a role attributed to SAF-A in mice (see Section 1.2.3.) but might be performed by multiple RNA binding proteins (RBPs) in human cells, including the SAF-A paralog HNRNPUL1 (Hasegawa *et al.*, 2010; Chu *et al.*, 2015; Kolpa *et al.*, 2016; Sakaguchi *et al.*, 2016). Albeit a study using the super-resolution microscopy method 3D-SIM found that XIST does not co-localise with the chromatin it is thought to repress, but it does co-localise with SAF-A, which has been implicated in lncRNA localisation (Smeets *et al.*, 2014). Another example of an ncRNA with a role is the abundant and long-lived RNA C₀T-1 comprised of highly repeated sequences, such as LINE-1. C₀T-1 RNAs associate with euchromatin foci in interphase via tethering facilitated by SAF-A and their loss leads to aberrant chromatin compaction and distribution (Hall *et al.*, 2014; see Section 1.2.3.). These repeat sequences do not appear to migrate but remain near their site of synthesis and stabilize in the vicinity of chromatin. C₀T-1 RNAs are also absent from the Barr body, which might mean that certain ncRNAs are associated with chromatin compaction while others are linked to chromatin decompaction (Hall and Lawrence, 2016). RNA molecules can also be a part of distinct cellular condensates like the nucleolus, stress granules, nuclear speckles, and paraspeckles (Roden and Gladfelter, 2021). These membraneless cellular bodies are thought to be formed by liquid-liquid phase separation (LLPS), the mixing of two chemically similar

yet molecularly distinct liquid phases (Li *et al.*, 2012; Bergeron-Sandoval *et al.*, 2016; Hnisz *et al.*, 2017). For example, the cooperative interactions between multivalent molecules of RNA and RNA-binding proteins were shown to form the phase-separated P granules in *Caenorhabditis elegans* germ cells (Brangwynne *et al.*, 2009). Paraspeckles are also formed via LLPS induced by the lncRNA NEAT1 by recruiting NONO dimers through its repetitive middle domain (Clemson *et al.*, 2009; Yamazaki *et al.*, 2018). Using RNA structure mapping methods, NEAT1 was found to have ample local and distal RNA-RNA interactions important for paraspeckle formation (Lin *et al.*, 2018). Phase-separated condensates were also shown to have liquid-like properties like fusion and fission and allowed protein exchange to occur in seconds instead of in minutes or longer (Weidtkamp-Peters *et al.*, 2008; Brangwynne *et al.*, 2009; Brangwynne *et al.*, 2011). However, experiments sometimes extrapolate the phase separation behaviour of molecules *in vitro* to functional LLPS *in vivo*. For example, stably folded proteins in specific buffers and non-physiological conditions have been shown to phase separate as a liquid, specifically before crystallising into a solid phase (Dumetz *et al.*, 2008; Zhang *et al.*, 2012). Therefore, the ability of proteins to phase separate *in vitro* with or without RNA does not necessarily translate to the formation of functional cellular condensates in healthy cells.

As RNA is highly negatively charged due to its phosphate backbone, it has been implicated in a type of LLPS between oppositely charged polyelectrolytes called complex coacervation (Overbeek and Voorn, 1957; Lin *et al.*, 2019). Due to this, RNA might also have a role in the genome-wide distribution of ncRNAs in the compartmentalisation of the nucleus by promoting the phase separation of proteins with intrinsically disordered regions (IDRs) or RNA binding domains (RBDs). Such proteins can bind RNA, with low levels of nucleic acids promoting phase separation while high levels inhibit it (Maharana *et al.*, 2018; Henninger *et al.*, 2021). After transcription initiation, low RNA levels might promote condensate formation via electrostatic interactions (Henninger *et al.*, 2021). Then as the condensates promote bursts of transcription, the higher levels of RNA could promote condensate dissolution. This forms equilibrium-based positive and negative feedback loops to regulate phase separation. Interestingly, RNA has also been shown to antagonise recruitment of proteins to chromatin. In the case of PRC2, a repressive chromatin modifier, RNA competes with it for binding to nucleosomes and blocking transcription

or degrading RNA allows PRC2 to bind to active genes (Beltran *et al.*, 2016; Hosogane *et al.*, 2016). Furthermore, it has been shown that nascent transcripts with non-coding sequences can create a scaffold structure that prevents DNA compaction (Creamer *et al.*, 2021). This non-chromatin scaffold can be maintained when chromatin, soluble proteins, and most nuclear RNA are depleted, reminiscent of the hypothesised dynamic RNA mesh. While RNA-involved phase separation is an attractive explanation for the compartmentalisation and stabilisation of transcriptional processes, no direct link between RNA and the transcription factories has been established. Furthermore, the role of nuclear RNAs has been mostly investigated in higher eukaryotes so it is not known how conserved their recently realised role is in regulating chromatin organisation in other organisms.

There are numerous diseases linked to the disruption of RNA processing either via changes to the RNA structure or their interactions with RBPs. One review mentions at least 40 cis- and trans-acting mutations that lead to disease due to defects in pre-miRNA splicing, whilst ncRNAs have also been found to be implicated in diseases (Cooper *et al.*, 2009). Deregulation of lncRNAs has been linked to cancer albeit it is not known whether this is a driver or consequence of oncogenesis (Li *et al.*, 2013). HOTAIR by itself might be able to promote metastasis of cancer by affecting the chromatin structure and MALAT1 has been linked to the pathology of vascular diseases and cardiac dysfunction as it is downregulated by known disease risk factors (Gupta *et al.*, 2010; Yan *et al.*, 2019). Recently many bioinformatic methods for associating human lncRNAs with diseases have been developed but further biological studies are required to investigate any of these promising results (Chen *et al.*, 2015; Lan *et al.*, 2017). CircRNAs, have also been found to be abundant in the neurons, and are co-expressed with Alzheimer's disease (AD) genes (Dube *et al.*, 2019). Their lack of free hydroxyl ends could make them resistant to exonucleases, thus accumulating with age in brain cells. They could then act as sponges for miRNA, sequestering them to regulate their function, but to complicate the picture, it was shown circRNAs can be translated *in vivo*, showing that we still lack much knowledge ncRNAs (Hansen *et al.*, 2013; Legnini *et al.*, 2017; Pamudurti *et al.*, 2017).

1.1.1.4. RNA binding proteins

During transcription in transcription factories, nascent RNA undergoes processing near RNA-binding proteins (RBPs), which are abundant on chromatin. This allows the correct processing (splicing, polyadenylation, capping) to quickly generate mature RNA that is ready for nuclear export and subsequent translation (Jackson, 2005; Singh and Padgett, 2009). RBPs are also involved in RNA localisation, export, storage, and decay. To interact with RNA, RBPs need to have an RBD like RNA recognition motifs (RRM), HNRNPK-homology (KH) domains, or RG/RGG motifs (Ray *et al.*, 2013). The RRM is found in more than half of known RBDs and consists of a β -sheet with four strands, which interacts with RNA, and two α -helices in the order $\beta\alpha\beta\beta\alpha\beta$ (Cléry *et al.*, 2008; Cook *et al.*, 2011). KH domains are the second most common RBD and consist of 65-70 amino acids and are composed of three β -strands and three α -helices in the order $\beta\alpha\alpha\beta\beta\alpha$ with RNA binding to a conserved glycine–lysine–X–glycine motif between the first and second α -helices (Musco *et al.*, 1996; Valverde *et al.*, 2008). The RG/RGG motif is an RNA binding segment rich in arginine-glycine-glycine and/or arginine-glycine repeats which are positively charged and can form hydrogen bonds and cation– π interactions with RNA molecules (Kiledjian and Dreyfuss, 1992; Corley and Gready, 2008). Another role for RBPs which has recently gained traction is their involvement in phase separation with the RNA mesh to promote RNA polymerase II binding and transcription (Figure 1.2). For instance, it has been proposed that the paraspeckle protein PSPC1 utilises RNA-dependent condensate formation to facilitate the phosphorylation of RNA pol II and stabilize its interaction with chromatin. In support of this, the degradation of PSPC1 has been found to disrupt RNA pol II binding and transcription (Shao *et al.*, 2022). Given that only approximately 1% of RNA pol II binding events lead to transcriptional elongation, this stabilisation of the enzyme is required for successful gene expression. This mechanism could also attract further RBPs to regulate transcription as many of them preferentially localise to transcriptionally active regions of chromatin (Xiao *et al.*, 2019). An interesting mechanism emerges, wherein high concentrations of nascent transcripts can initially inhibit transcription but subsequently recruit RBPs that then promote transcription through phase separation. RNA decay could then play an important role to regulate this mechanism, with XRN2 and the exosome inhibiting phase separation. Although not all RBP-facilitated phase separation might be involved in gene regulation. For example, the RBP FUS, which is involved in

transcription, splicing, mRNA transport, and translation, has an RG/RGG motif that allows it to phase separate in the nucleus to create droplets required for DNA repair initiation (Murthy *et al.*, 2019; Levone *et al.*, 2021).

Initially, it was proposed that the interphase chromatin is maintained by a rigid nuclear matrix (Berezney and Coffey, 1974). However, further research has led to the emergence of the concept of a dynamic matrix structure (Ma *et al.*, 1999; Hancock, 2000). This dynamic structure is hypothesised to be formed by heterogeneous nuclear ribonucleoproteins (hnRNPs) through their interactions with nuclear RNAs and chromatin (Romig *et al.*, 1992; Michieletto and Gilbert, 2019; Figure 1.2.). These hnRNPs represent a well-studied family of RBPs which are known to have a role in the post-transcriptional control of RNA (Dreyfuss *et al.*, 1993). They contain KH and RG/RGG motifs to facilitate RNA binding, together with auxiliary protein domains. A recent study using mass spectrometry coupled with formaldehyde-assisted isolation of regulatory elements (FAIRE-MS) found that HNRNPA2/B1 and HNRNPD were highly enriched on mitotic chromatin where they might be retained at key positions for transcription initiation in the daughter cells (Ye *et al.*, 2023). Other enriched factors had IDRs which links them to LLPS via self-association or interaction with RNAs and chromatin organisation during interphase. An example is HNRNPA1, which forms stress granules in the cytoplasm via its IDR and might also form pathological fibrils in cells (Molliex *et al.*, 2015; Gui *et al.*, 2019). SAF-A (HNRNPU) is another hnRNP family member and was shown to be involved in large-scale interphase chromatin organisation by potentially forming an RNA-protein oligomer mesh thereby facilitating the decompaction of transcriptionally active chromatin (Nozawa *et al.*, 2017; see Section 1.2.1.). However, unlike HNRNPA2/B1 and HNRNPD, SAF-A appears to be excluded from mitotic chromosomes, suggesting that it may perform its role after nuclear assembly (Sharp *et al.*, 2020).

Loss-of-function mutations in RBPs have been found to cause many diseases potentially by perturbing a tissue-specific expression of critical RNA targets. A trinucleotide repeat mutation in the 5'-untranslated region of the FMR1 gene leads to the loss of its protein product involved in mRNA translations at neuronal synapses (Wang *et al.*, 2012). This loss causes fragile X syndrome, the most common inherited intellectual disability disorder. Another example is mutations in PRPF3, PRPF8, PRPF31, and RP9, which cause forms of retinitis pigmentosa (Mordes *et al.*, 2006).

They are components of the spliceosome, which when mutated, reduce the level of spliceosome function in photoreceptor neurons and reduce the amounts of important highly expressed mRNAs like rhodopsin, thus impeding neuronal function (Mordes *et al.*, 2007). Gain-of-function mutations in RBPs can also form RNA, protein or RNA-protein aggregates which are often implicated in neurological diseases. An example is amyotrophic lateral sclerosis (ALS) which can be caused by mutations in various RBPs like hnRNPs (HNRNPA1, HNRNPA2/B1), FUS, or TDP43, and are implicated in phase separation processes (Lagier-Tourenne *et al.*, 2010; Kim *et al.*, 2013). HNRNPA1 and HNRNPA2/B1 were also found to be mutated in their IDRs in ALS, which increases their tendency to self-associate into stress granules and sequester RNA (Kim *et al.*, 2013). In summary, these proteins sequester RNA to negatively affect RNA metabolism, which could be a common disease mechanism for other RBPs with IDRs leading to degenerative diseases.

1.1.1.5. *Other proteins involved in chromatin organisation*

Cohesin

To facilitate large-scale chromatin organisation, the nucleosome fibre is folded through DNA looping and DNA-DNA interactions. Important players in this process are the SMC (Structural Maintenance of Chromosomes) complexes, which encircle DNA via an ATP-dependent mechanism (Jeppsson *et al.*, 2014). Cohesin is one class of SMC complex which holds the two sister chromatids together after DNA replication, a vital mechanism for chromosome segregation during cell division (Peters and Nishiyama, 2012). It also plays an important role in loop formation during interphase, DNA repair, and transcriptional regulation. In cells that undergo mitosis, the complex is made up of two subunits, SMC1 α and SMC3, a kleisin-type subunit RAD21, and a non-kleisin subunit STAG1 or STAG2, which together maintain a ring-like configuration (Figure 1.3A). Cells that undergo meiosis have additional subunits, the cohesin subunit SMC1 β , the kleisin subunits RAD21, RAD21L, and REC8, and the non-kleisin subunit STAG3 (Rankin, 2015). Loading the cohesin complex onto chromosomes requires a cohesin loader complex composed of NIPBL2 and MAU2, which performs complex crosstalk with chromatin remodellers BAF/PBAF and

nucleosomes at promoter sequences to access the DNA (Chao *et al.*, 2017; Muñoz *et al.*, 2022; Figure 1.3B). During early G₁, cohesin is loaded onto chromatin by the cohesin loader complex in an ATP-dependent process (Zhang *et al.*, 2019; Figure 1.3B). For cohesin to be then unloaded from the chromatin, it requires a cohesin regulatory complex consisting of WAPL and PDS5 (Losada *et al.*, 2005, Tedeschi *et al.*, 2013; Figure 1.3B). During S-phase, the structure of the cohesin complex encircles chromatin fibres in a topological interaction after acetylation by the acetyltransferases ESCO1 and ESCO2, which are vital for sister chromatid cohesion (Zhang *et al.*, 2008; Whelan *et al.*, 2011; Murayama & Uhlmann, 2014). Cohesin stably holds the two sister chromatids until it is released from them by phosphorylation by the kinases CDK1 and Aurora-B in prophase and cleavage of RAD21 by separase in anaphase (Hauf *et al.*, 2001; Hauf *et al.*, 2005). The PDS5-WAPL complex also assists in cohesin release during mitosis (Tedeschi *et al.*, 2013). After deacetylation by HDAC8, the cohesin complexes can be re-used in G₁ of the daughter cells (Deardorff *et al.*, 2012). Cohesin binding was also found to coincide with CTCF binding sites, a sequence-specific binding protein also involved in large-scale chromatin structure (Parelho *et al.*, 2008; see Section 1.1.2. and Figure 1.4D). Cohesin is often observed to be enriched at the base of putative chromatin loops and TAD boundaries (Lieberman-Aiden, *et al.*, 2009; Dixon *et al.*, 2012). This could be due to the ability of monomeric and dimeric cohesin to entrap distal DNA on the same strand to physically bring it together (Skibbens, 2016). Deletion of WAPL has been shown to result in chromatin loop extensions while degradation of the RAD21 cohesin subunit eliminated nearly all loops which highlights the importance of cohesin in chromatin loop formation (Haarhuis *et al.*, 2017; Rao *et al.*, 2017). Mutations in the cohesin complex or its regulatory complexes lead to diseases called cohesinopathies. An example of a cohesinopathy is Cornelia de Lange syndrome (OMIM 122470) which can be caused by mutations in the cohesin loader NIPBL2 and by mutations in HDAC8 that inhibit cohesin deacetylation (Krantz *et al.*, 2004; Liu *et al.*, 2009; Deardorff *et al.*, 2012). Mutations in the cohesin complex have also been implicated in chromatid cohesion defects leading to aneuploidy in cancers (Solomon *et al.*, 2011).

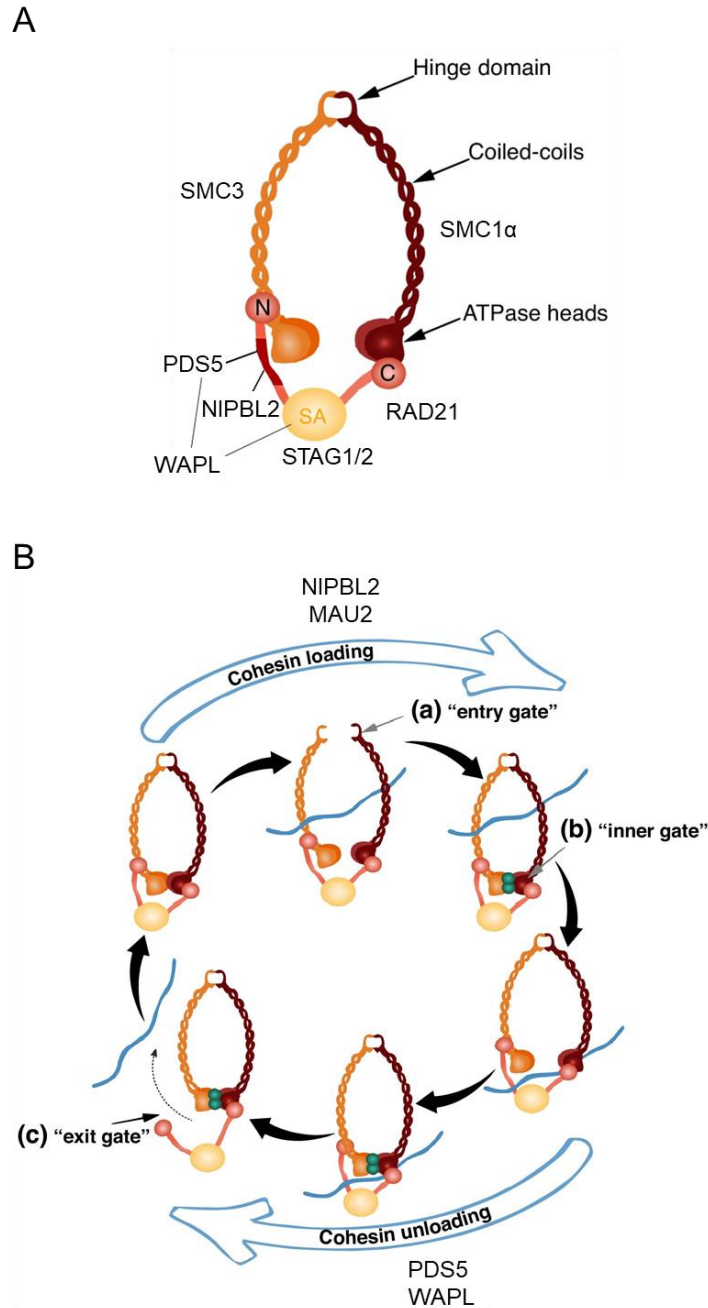


Figure 1.3. Cohesin complex structure and dynamics. A) Schematic representation of the structure of the cohesin complex in human mitotic cells. The SMC proteins, SMC1 α and SMC3, interact through their hinges and are bridged by the kleisin subunit RAD21 at their ATPase heads to form a ring-like structure. The non-kleisin (SA) subunits, STAG1 and STAG2, interact with RAD21. The interaction sites for NIPBL2, PDS5, and WAPL are shown with black lines. B) Schematic representation of cohesin dynamics in G₁ phase. The chromatin fibre (blue) is entrapped by an "entry gate" formed by the SMC proteins' hinge domains (a). The two SMC ATPase heads are engaged by ATP (green) with assistance from NIPBL2 and MAU2. This forms an "inner gate" that is opened by ATP hydrolysis (b). The chromatin fibre then passes through an "exit gate" located between the coiled-coil region of SMC3 and the N-terminus of RAD21 (c). Unloading of cohesin from chromatin is mediated by PDS5 and WAPL. Images adapted from Morales and Losada, 2018.

Condensin

Condensin is another SMC complex with a similar structure to cohesin. Mammalian cells have two condensin complexes, I and II, which share SMC2 and SMC4 as main subunits. Condensin I has the kleisin-type subunit CAP-H and non-kleisin subunits CAP-D2 and CAP-G, while condensin II has respectively CAP-H2, and CAP-D3 and CAP-G2 (Hirano and Mitchison, 1994; Hirano *et al.*, 1997; Ono *et al.*, 2003; Yeong *et al.*, 2003). The SMC subunits interact with the kleisin subunit to form a closed ring that might loop DNA in a similar way to cohesin complexes during interphase (Onn *et al.*, 2007; Jeppsson *et al.*, 2014; Krepel *et al.*, 2018; Ryu *et al.*, 2020; Lee *et al.*, 2020; Lee *et al.*, 2022). In addition, the ATPase activity of the SMC subunits is thought to dissociate the condensin complex and perturbation of this process leads to the loss of condensin complex function (Lammens *et al.*, 2004; Hudson *et al.*, 2008). Thus, similarly to cohesin, condensin can topologically enclose DNA and has DNA translocase activity. Condensin has been shown to reconstitute mitotic chromatids from *Xenopus laevis* together with core histones, their chaperones, and topoisomerase II *in vitro* (Shintomi *et al.*, 2015; Sakai *et al.*, 2018). The role of condensin is to compact the chromosomes during mitosis (Gibcus *et al.*, 2018). Specifically, the mitotic role of condensin I is to compact the chromosomes laterally, while condensin II compacts the chromosomes longitudinally with the ratio of the two complexes being key to the final shape (Shintomi and Hirano, 2011). Condensin II can create large loops of ~400 kb via rapid loop extrusion while condensin I creates smaller loops of ~80 kb within the larger loops (Elbatsh *et al.*, 2019). There is an order to the packaging as well - condensin I binds to mitotic chromosomes after nuclear envelope breakdown in prometaphase, while condensin II is found on chromatin from interphase and prophase (Gerlich *et al.*, 2006; Green *et al.*, 2012). At the end of mitosis, condensin I is left in the cytoplasm while condensin II needs to be continuously unloaded from the chromatin by MCPH1 (Houlard *et al.*, 2021). Mutations in MCPH1 have been shown to cause microcephaly, a neurodevelopmental disease with a characteristically reduced head circumference, and patient cells show premature chromosome condensation during the G₂ phase due to condensin II binding to chromatin prematurely (Neitzel *et al.*, 2002). Microcephaly can also be caused by mutations in condensin subunits due to excessive DNA/chromatin bridges during chromosome separation (Martin *et al.*,

2016; Pang *et al.*, 2022). Furthermore, mutations in condensin subunits have been linked to glioma pathogenesis and progression by promoting glioma cell proliferation and invasion and impairing the response of immune cells (Pang *et al.*, 2022).

Topoisomerases

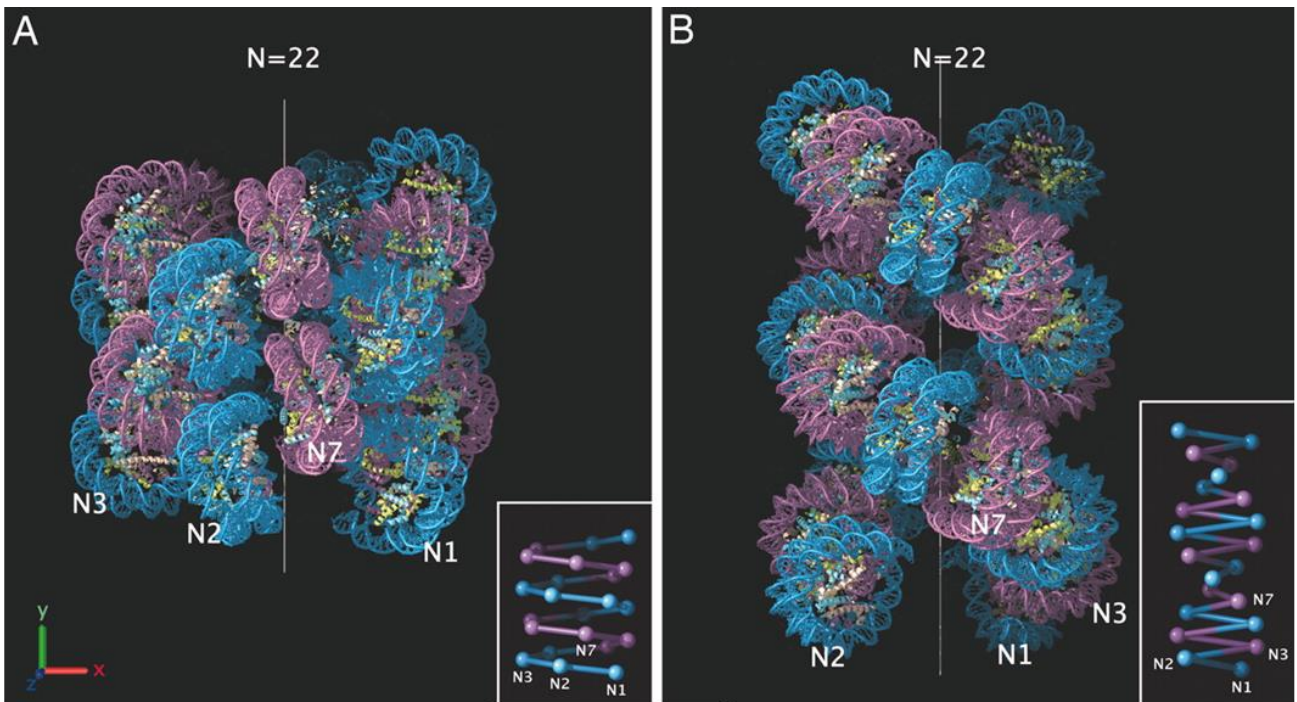
During many cellular processes involving DNA, the DNA often becomes supercoiled, either by under-twisting (negative supercoiling) or over-twisting (positive supercoiling). Topoisomerases are enzymes that can release topological strain, they can also disentangle DNA knots, precatenanes, and catenanes generated during replication, transcription, and DNA repair. Many different topoisomerases exist but they are broadly classified into two groups – type I and type II – depending on whether they cut one or two DNA strands. Further classification of subtypes depending on structure and function yields five family groups across all organisms – type IA, IB, IC, IIA, and IIB. In humans, the type IA enzymes TOP3A and TOP3B resolve DNA entanglements that occur during replication and repair by passing a single strand of DNA through a cleaved DNA strand (Lopez *et al.*, 2005). TOP3B can also handle folded RNA cleavage in the cytoplasm when the RNA molecules are knotted or catenated (Saha *et al.*, 2020). TOP3A is the only topoisomerase that can resolve hemicatenanes during replication and HR, and within mitochondrial DNA (Lee *et al.*, 2019; Bizard and Hickson, 2020). The human type IB topoisomerase, TOP1 can nick and rotate a DNA strand to relax both kinds of supercoiling in segments without nucleosomes and it has been found within mitotic chromosomal loops (Koster *et al.*, 2005; Salceda *et al.*, 2006; Pommier *et al.*, 2022). There is also a mitochondria-specific type IB topoisomerase, TOP1MT, which is responsible for the maintenance of the hundreds of copies of ~16,000 bp circular mitochondrial DNA. Similarly to type IA, type IIA topoisomerases (TOP2A and TOP2B in humans) passage two DNA strands through a cleaved DNA duplex which allows them to resolve both types of supercoiling, DNA precatenanes, and chromosomal intertwining (Liu *et al.*, 1980; Keszthelyi *et al.*, 2016). TOP2A specifically is essential for DNA replication progression and disentangling sister chromatids in daughter cells (Lee and Berger, 2019). As condensin complexes generate chromatin loops to compact the chromosomes, TOP2A relaxes the positive supercoiling and its absence arrests cells

at the G₂-M boundary (McClendon *et al.*, 2005; Lee and Berger, 2019). Another use for TOP2A is in facultative heterochromatin formation and resolution in tandem with the BAF chromatin remodelling complex (Miller *et al.*, 2017). Similarly, TOP2B is important for mitosis, heterochromatin regulation, and DNA repair (Austin *et al.*, 2018). All topoisomerases, except TOP1MT so far, are connected to various disorders and cancers, showing their importance in healthy chromatin organisation (Pommier *et al.*, 2022).

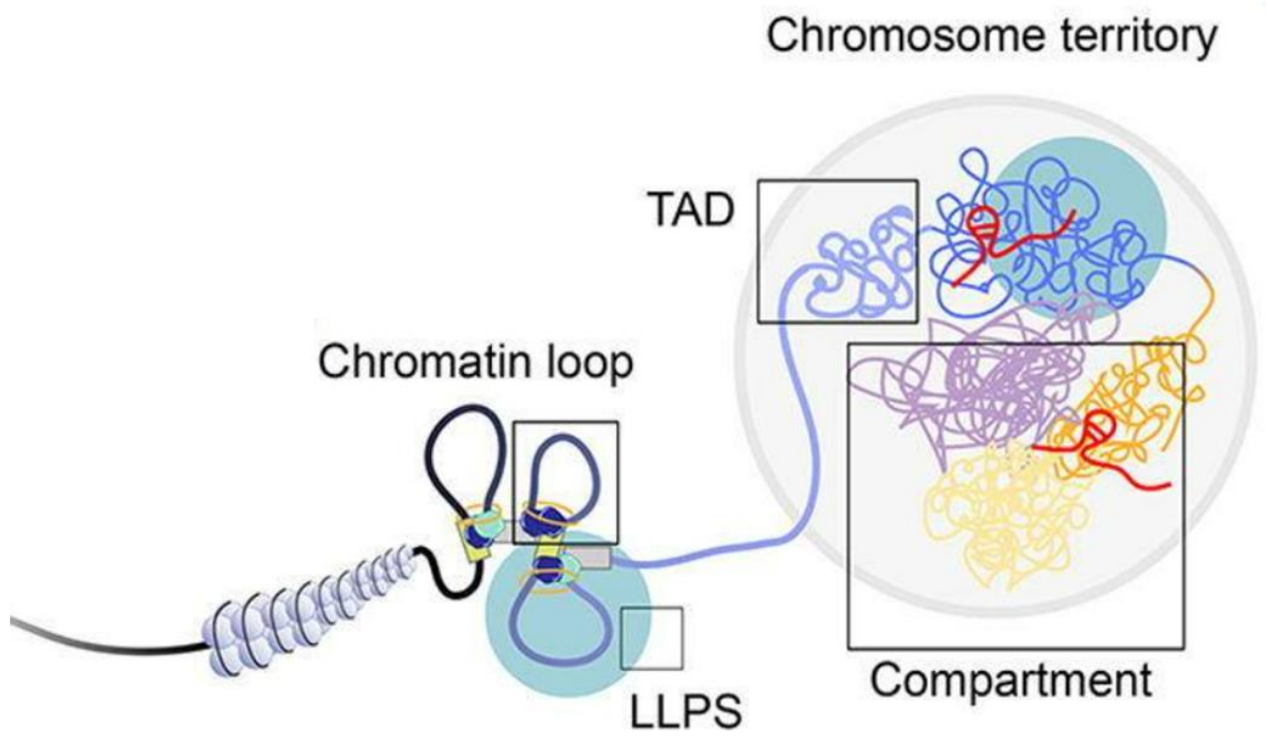
1.1.2. Chromatin organisation during interphase

The basic level of packaging is the winding of DNA around histones to form nucleosomes which can be manipulated by chromatin remodellers to either promote or inhibit the binding of RNA pol II and transcription. Beyond nucleosomes, the next level of packaging involves the assembly of arrays of nucleosomes along with linker histones through electrostatic interactions, resulting in the formation of a 30-nm chromatin fibre. However, the precise mechanisms underlying the assembly of the 30-nm chromatin fibre remain incompletely understood, and two classical models have been proposed (Figure 1.4A and B). The first model is the solenoid or one-start model discovered using electron microscopy (EM) and isolated chromatin from chicken erythrocytes in the presence of magnesium and linker histone H1 (Finch and Klug, 1976). According to this model, consecutive nucleosomes interact with each neighbour and a bend in the linker DNA leads to the presence of 6 nucleosomes per helical turn of the fibre. Cryo-EM and single-molecule force spectroscopy experiments have provided further support for the solenoid model (Robinson *et al.*, 2006; Kruithof *et al.*, 2009). The second model is the zig-zag or two-start model where nucleosomes interact with their second neighbour (Worcel *et al.*, 1981; Woodcock *et al.*, 1984). In this model, the linker DNA does not bend but instead weaves between nucleosome stacks. Reconstituted chromatin and X-ray crystallography studies have observed the zig-zag model (Dorigo *et al.*, 2003; Schalch *et al.*, 2005). Some evidence suggests that both models might co-exist within the same chromatin fibre and their presence may depend on whether the region is transcriptionally active or inactive (Wu *et al.*, 2007; Grigoryev *et al.*, 2009; Risca *et al.*, 2016; Gilbert, 2019). However, *in vivo* experiments using cryo-EM imaging of

unfixed chromosomes initially found 11-nm fibres, while electron microscopy tomography (EMT) together with DNA labelling showed disrupted fibres from 5 to 24 nm without disturbing the chromatin structure (Dubochet *et al.*, 1988; Ou *et al.*, 2017). These experiments do not support the existence of either model of the 30-nm chromatin fibre, which instead could be an artefact of the harsh conditions required to isolate and reconstitute chromatin. Other studies propose that the 30-nm chromatin fibre may exist in specific regions of the nucleus, such as heterochromatin regions enriched in satellite sequences with regularly positioned nucleosomes or centromeric chromatin (Gilbert and Allan, 2001; Kruithof *et al.*, 2009; Naughton *et al.*, 2010).



C



D

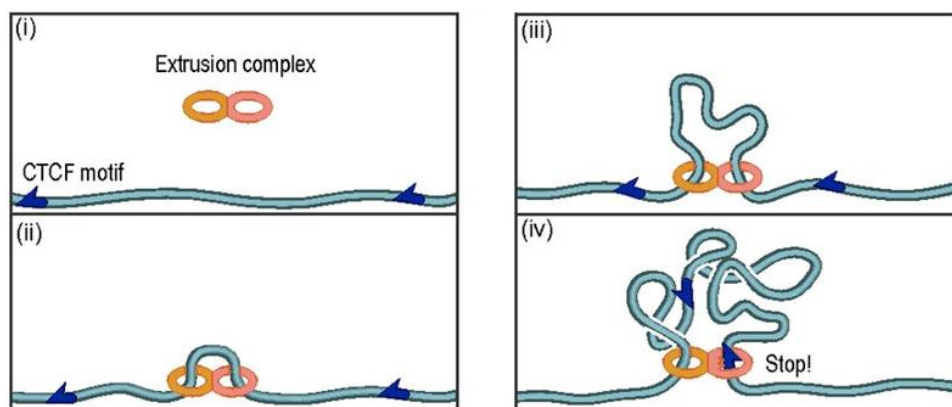


Figure 1.4. Chromatin fibre models and the large-scale genomic organisation. A) and B) show the two popular models for 30 nm chromatin fibres – solenoid (A) and zig-zag (B) with key nucleosomes highlighted (N1, N2, N3, and N7) to showcase the differences in each arrangement. From Robinson *et al.*, 2006. C) A schematic representation of large-scale genomic organisation starting from DNA and the nucleosomal fibre (left) forming chromatin loops, topologically associated domains (TADs), chromosome territories and compartments. Liquid-liquid phase separation (LLPS) is shown as a green bubble at the scale of chromatin loops and territories. From Ding and Jhang, 2023. D) Loop extrusion model showing the cohesin complex loading at a random locus and forming a tiny loop in the chromatin fibre (i and ii). The loop grows as the two subunits of the cohesin complex move in opposite directions along the DNA fibre (iii). When a subunit encounters a CTCF binding motif facing towards it, loop extrusion is halted (iv). From Sanborn *et al.*, 2015.

A higher level of chromatin organisation involves chromatin domains called TADs, which are separated by short DNA motifs known as insulators and influenced by different regulatory proteins and differential histone PTMs (Figure 1.4C). TADs are characterised by an enrichment of interactions within them and are hypothesised to be formed by loop extrusion. Techniques like chromatin conformation capture (3C) and its genome-wide variant Hi-C, which resolve proximity-ligated DNA, have enabled interactions within the genome to be visualised (Dekker *et al.*, 2002; Lieberman-Aiden *et al.*, 2009). Hi-C studies have identified TADs that are bordered by the regulatory protein CTCF (Dixon *et al.*, 2012). Initially, CTCF was found to bind to insulators and hypothesised to mark the boundaries between active and inactive chromatin domains (Yusufzai *et al.*, 2004; Cuddapah *et al.*, 2009). CTCF is also involved in chromatin loop formation by blocking the cohesin loop extrusion process and limiting loop size when the 3' ends of CTCF binding motives face towards the loop (Figure 1.4D; Sanborn *et al.*, 2015; Nora *et al.*, 2017). TADs also correlate well with active and repressive marks on the chromatin they encompass (Sexton *et al.*, 2012). An important characteristic of TADs is their ability to bring distal genetic elements, which are not linearly close, into spatial proximity, adding an additional layer of gene regulation (Bickmore and van Steensel, 2013).

The higher-order organisation of chromatin beyond the 30-nm fibre is not fully understood but it has been found that its density and mobility vary from G₁ to G₂ phase. In the early G₁ phase, after the separation of daughter cells, chromatin exhibits high mobility until the formation of chromosome territories (Belmont and Bruce, 1994; Walter *et al.*, 2003). Chromosome territories represent the distinct volumes each chromosome occupies within the interphase nucleus and were initially thought to be physically separated by an interchromatin domain (Figure 1.4C; Cremer

and Cremer, 2001). However, it has been discovered that chromosomal DNA can intermingle, and interchromosomal interactions are important for active genes located between chromosomal territories, such as the MHC II locus (Branco and Pombo, 2006). Upon transcriptional activation, the MHC II locus is not confined solely to the territory of chromosome 6 but can also be found within chromosomes 1, 2, and 9. Generally, gene-rich chromosomes tend to be closer to the interior of the nucleus, while gene-poor chromosomes and heterochromatic regions are closer to the nuclear envelope (Boyle *et al.*, 2001). The positioning of chromosomal territories varies among different tissues but shows greater similarity among closely related cell types (Parada *et al.*, 2004). The chromosomal territories themselves have an internal order where the actively transcribed regions appear to be closer to the surface of a territory, while the inactive segments are sequestered internally (Boyle *et al.*, 2011). However, it is important to note that this order is not absolute, as regions containing ubiquitously expressed and tissue-specific genes are not always found at the periphery of a chromosome territory and their location appears to be conserved across species (Mahy *et al.*, 2002). Further differences arise whether a region is highly repetitive and inactive or actively transcribed. For example, repetitive regions of satellite DNA might have a regular helical conformation and be more compact than regular chromatin (Gilbert and Allan, 2001). Meanwhile, gene-rich euchromatin has a more disrupted structure specifically at transcription start sites, probably to facilitate easier RNA pol II binding (Gilbert *et al.*, 2004; Naughton *et al.*, 2010). This means that regions with high gene density are associated with high transcriptional activity and have more decompacted chromatin relative to gene-poor regions, which have more compact chromatin domains. Interestingly, core centromeric regions have a disrupted chromatin structure when actively transcribed (Naughton *et al.*, 2022). In summary, chromatin organisation during interphase is heterogeneous and depends on factors such as local gene density, ongoing cellular processes like transcription or replication, and the chromatin environment of specific regions.

1.1.3. Chromatin organisation during mitosis

Mitosis is a well-studied phase of the cell cycle where DNA is compacted and readily visualised by light microscopy, revealing individual chromosomes (Paweletz, 2001; Figure 1.5A). Each chromosome is packaged by topoisomerases and condensins with sister chromatids held together by cohesins to ensure the successful division of DNA between the daughter cells. The compaction of mitotic chromosomes compared to interphase chromatin varies from 2-fold to 50-fold depending on the method used for measurement, such as fluorescent H2B occupancy signal or chromosomal loci distances (Belmont, 2006; Vagnarelli, 2012). However, imaging using electron microscopy tomographic analysis suggests that chromatin is only compacted by 40% during mitosis (Ou *et al.*, 2017). Additionally, chromatin compaction might not remain constant, for example live cell imaging of fluorescent H2B from prophase to metaphase showed that the chromatin became less dense between mid-prophase to late prophase and then became more dense in metaphase (Liang *et al.*, 2015). Chromosome conformation capture experiments have also been used to probe the mitotic chromatin structure - contact maps from cells in mitosis showed an absence of TADs (Naumova *et al.*, 2013). Instead, there was a linear decrease in contact probability as genomic distances increased, up to 10 Mb whilst at longer distances, there was a reduction in contact frequencies. This observation has also been observed in cells undergoing a synchronised entry to mitosis showed a sudden loss of nuclear compartments in early prophase and subsequent loss of TADs (Gibcus *et al.*, 2018).

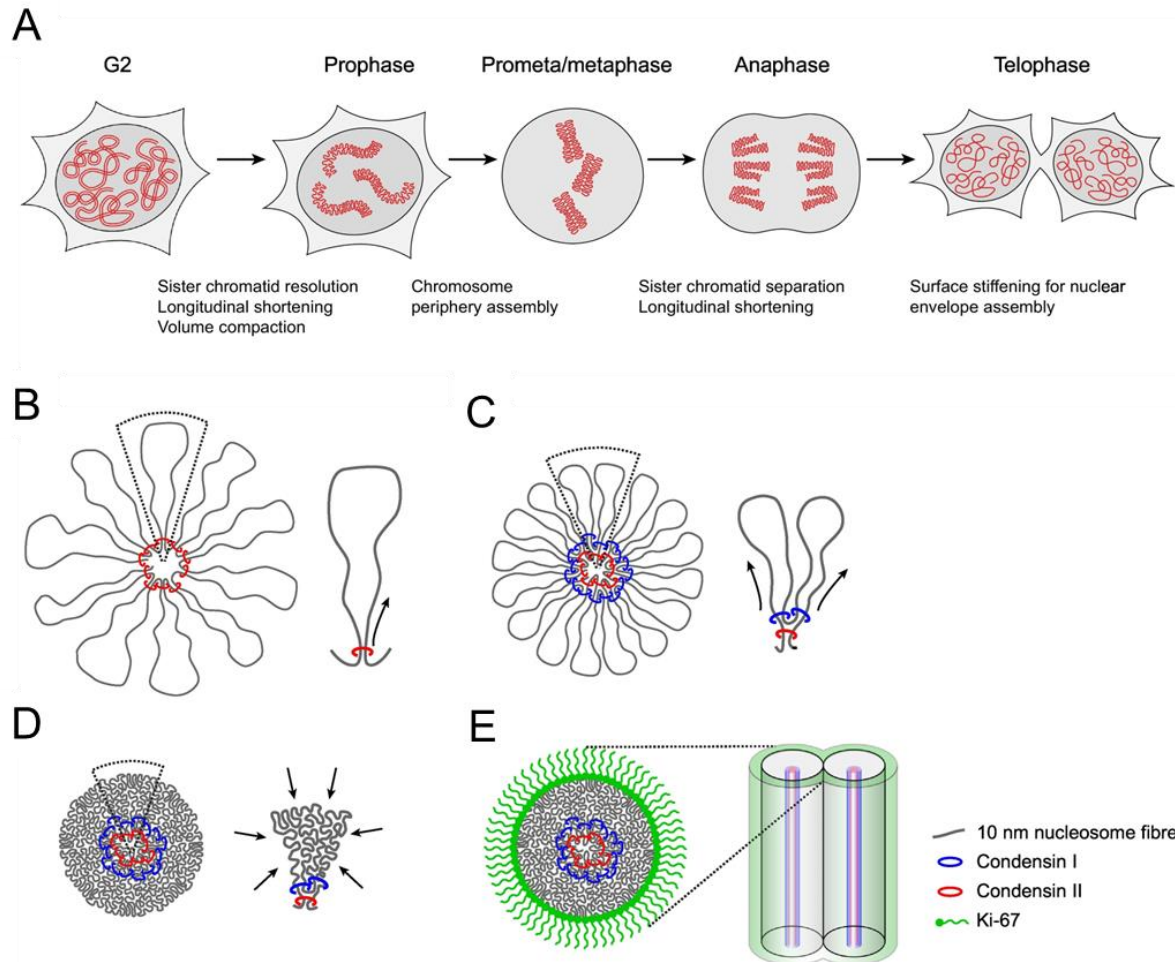


Figure 1.5. Cell division and mitotic condensation. A) Illustration of DNA (in red) structure during cell division. B) to E) are a schematic cross section of a chromatid during prophase. B) Condensin II (in red) extrudes DNA loops. C) Condensin I (in blue) binds after condensin II and it extrudes smaller loops within the original ones. D) Nucleosomal interactions further compact chromatin. E) Ki-67 (in green) creates brush-like structures on the compacted chromatin's surface during late prophase. From Batty and Gerlich, 2019.

While condensation is a fact, how mitotic chromosomes are organised is still unknown and various models have been proposed. Based on electron microscopy imaging of isolated chromosomes, there is a proteinaceous chromosome scaffold around which mitotic chromatin is looped, known as the radial loop model (Paulson and Laemmli, 1977). To image the metaphase chromosomes, they were depleted of histones and spread on a treated surface. This remaining core maintained the general shape of a chromosome from which they observed that the DNA extended as loops. When treated with proteases or urea, the structure was disrupted, hinting that non-histone proteins were involved in chromosome structure. Further research discovered that some of the non-histone proteins were topoisomerase II α , condensin complexes, and cohesin complexes (Earnshaw *et al.*, 1985; Hirano and Mitchison,

1994; Saitoh *et al.*, 1994). More recent electron microscopy studies show that up to 40% of the chromosomal volume is not chromatin but a layer of proteins and RNA (Paulson *et al.*, 2021). This is maintained by Ki-67 forming brushlike structures on the chromatin surface starting from late prophase through to anaphase, coinciding with when we can observe the rod-like chromosome shapes (Cuylen *et al.*, 2016; Figure 1.5E). Albeit Ki-67's role in shaping chromosomes might not be essential, as deletion mutants still allow chromosome formation albeit with swelling (Booth *et al.*, 2014; Takagi *et al.*, 2018). Depletion of Ki-67 leads to the loss of the perichromosomal compartment, and its impact becomes evident after cell division, where altered interactions between chromosomes result in reduced nucleoli numbers (Booth *et al.*, 2014).

Although condensin complexes are important for chromosome folding (Section 1.1.1.5., Figure 1.5B and C), condensed chromosomes can still form in cells depleted of condensins, but they have reduced structural integrity (Vagnarelli *et al.*, 2006). This observation led to the discovery of chromatin condensation factors with unclear functions such as Repo-Man and topoisomerase IIA. Their inhibition leads to slowed chromatin decondensation with no prominent structural defects (Gorbsky, 1994; Vagnarelli *et al.*, 2006). Therefore, chromatin condensation independent of condensin complexes and condensation factors is possibly driven by PTMs on histone tails causing attractive forces to compact prophase chromatin (Wilkins *et al.*, 2014). Overall, during mitosis chromatin is condensed by attractive forces and molecular factors, causing chromosomes to reach their highest compaction level in anaphase (Mora-Bermúdez *et al.*, 2007; Figure 1.5D). A specific histone PTM associated with condensation is the phosphorylation of serine 10 on the tail of H3 (H3S10ph) by Aurora-B kinase (Wei *et al.*, 1999). The process starts before mitosis and peaks during metaphase with dephosphorylation starting in anaphase, preceding the visible decondensation of chromosomes. Mutating the site of phosphorylation or blocking kinase activity perturbed chromosome dynamics during mitosis in organisms like *Tetrahymena thermophila* and *Xenopus laevis*, and led to improper segregation possibly due to defects in chromatin condensation (Wei *et al.*, 1999; de la Barre *et al.*, 2000). However, in other organisms such as *Saccharomyces cerevisiae* and *Caenorhabditis elegans* there was no direct link found between ablation of H3S10ph and chromosome condensation (Hsu *et al.*, 2000). In addition, in

humans, no link has been discovered between loss of H3S10ph and mitotic condensation when reconstructing nucleosomal array *in vitro* (Equilibrina *et al.*, 2015).

As mitosis progresses, it becomes apparent that the structure of chromosomes requires active maintenance rather than being a static entity. Depletion of proteins required for chromosome condensation, such as core condensin subunits SMC2 and SMC4, condensin I kleisin subunit CAP-H, and topoisomerase IIA, results in the loss of the recognizable chromosome structure (Hirano and Mitchison, 1994; Piskadlo *et al.*, 2017; Samejima *et al.*, 2018; Nielsen *et al.*, 2020). These findings indicate that mitotic chromosomes are dynamic structures that rely on active protein processes to maintain their overall organisation. However, the precise mechanisms by which the cell ceases maintenance of chromosome shape in late anaphase and telophase, leading to decondensation, and how interphase chromatin is subsequently established, remain unclear.

1.1.4. Chromatin organisation at the mitotic exit and in early G₁

Upon mitotic entry, metazoan chromatin is subjected to significant condensation, forming mitotic chromosomes that vary from 1.4- to 50-fold greater compaction relative to interphase chromosomes, depending on the method of visualisation (Belmont, 2006; Ou *et al.*, 2017). Concomitantly, transcription factors, chromatin remodelling complexes and transcriptional complexes, that were associated with interphase chromatin, are evicted. These changes to the chromatin structure, coupled with nucleus disassembly, have been historically thought to interrupt nuclear processes upon entry to mitosis, but more recent work has shown how transcription is restarted before the end of mitosis, during telophase, in bursts exceeding interphase transcriptional levels (Prasanth *et al.*, 2003; Hsiung *et al.*, 2016). While these structural changes of chromatin at the mitosis-G₁ transition are cytologically well described, the underlying molecular mechanisms and significance of transcriptional reactivation patterns to cell programming and fate are less so. What has been recognised so far is that chromosomes visibly decondense between late anaphase and telophase. However, the re-establishment of what we know as the

interphase chromatin organisation is not achieved until later in G₁ when the chromosomal territories have stabilised and TADs are restored (Dileep *et al.*, 2015).

Mitotic chromosome decondensation is an active process that is not just a reversal of condensation events but is still relatively unstudied compared to condensation. Decondensation involves the inactivation of mitotic kinases, the reversal of phosphorylation events by phosphatases, and potentially the participation of chromatin decondensation factors. The precise role of PP2A protein phosphatase in mitotic exit remains unclear, although its depletion does lead to a delay in this process (Schmitz *et al.*, 2010). Similarly, the depletion of mitotic kinases such as Aurora-B by ATPase p97 has uncertain implications for mitotic exit (Ramadan, *et al.*, 2007). Experiments conducted in *Xenopus* cell-free extracts highlighted potential chromatin decondensation factors, including RuvB-like ATPases, DRG GTPases, and their DFRP regulatory factors (Magalska *et al.*, 2014; Schellhaus, 2017). Depletion of individual factors such as RuvBL1, RuvBL2, DRG1, DRG2, DFRP1, and DFRP2 resulted in further condensation of cell-free mitotic chromosomes, while their reintroduction led to chromosome decondensation (Magalska *et al.*, 2014; Schellhaus, 2017). Moreover, introducing mutant forms of ATPases or GTPases, or removing ATP and GTP nucleosides, inhibited chromatin decondensation, providing evidence for the active nature of this process. In higher organisms, the well-conserved ATPases RuvBL1 and RuvBL2 are part of chromatin remodelling complexes such as Ino80 and SWR-C which might explain their chromatin decondensation properties in *Xenopus* (Nano and Houry, 2013). RuvBL1 and RuvBL2 overexpression also has a strong link to tumour growth and poor prognosis in many cancer types including liver, colorectal, and gastric cancers (Assimon *et al.*, 2019). Out of the GTPases and their regulatory factors, only DRG2 has been studied in higher organisms. DRG2 knockdown has been shown to cause a cellular arrest at the G₂-to-mitosis boundary due to impaired microtubule dynamics (Jang *et al.*, 2016; Dang *et al.*, 2018). However, the involvement of these factors in chromatin decondensation in organisms other than *Xenopus* remains to be demonstrated, suggesting the presence of more complex decondensation processes in these organisms.

While most transcription factors, transcriptional complexes and architectural proteins are evicted from mitotic chromosomes during prophase, a small number of molecules are retained on the chromatin (Kadauke and Blobel, 2013; Festuccia *et al.*, 2017). Subsequently, most evicted proteins re-bind to the chromatin during late anaphase or telophase while the chromosomes are still morphologically condensed, showing our lack of understanding of the mechanism of the chromatin decondensation (Hsiung *et al.*, 2016). For example, the architectural protein CTCF, which is important for interphase higher-order chromatin structure, is evicted from most sites during mitosis and resumes full binding by telophase (Sekiya *et al.*, 2017; Zhang *et al.*, 2019). Similarly, RNA pol II re-enters the newly formed nucleus by late telophase, an event linked to bursts of transcriptional activation, differing from transcription levels during interphase, at half of the active genes (Hsiung *et al.*, 2016). But another important architectural protein, cohesin, is fully evicted from mitotic chromosomes and is rebound at a slower rate (Zhang *et al.*, 2019). As cohesin and CTCF are important for chromatin loop extrusion and establishment of TADs, the slower re-binding of cohesin fits with cells re-establishing the domains about 2 hours after mitosis (Dileep *et al.*, 2015). Other events during the mitosis-to-G₁ transition are tethering of heterochromatin to the nuclear lamina during telophase by PRR14, lamina-associated domain re-establishment, and long-range chromosome interactions (Poleshko *et al.*, 2013; Kind *et al.*, 2013; Dileep *et al.*, 2015). Overall, the kinetics of the re-establishment of some nuclear processes and architecture have been investigated, but their direct link to chromatin decondensation at the end of mitosis has not been elucidated. Furthermore, how chromatin decondensation sets up the transcriptional program of daughter cells has not been mechanistically explained.

1.2. SAF-A

SAF-A was first identified in 1984 and is the largest heterogenous ribonuclear protein at approximately 90 kDa (Dreyfuss *et al.*, 1984; Kiledjian and Dreyfuss, 1992). The nuclear protein is an abundant component of the nuclear matrix and has been previously linked to pre-mRNA splicing and XIST-mediated transcriptional silencing as well as other nuclear processes (see Section 1.2.3.; Xiao *et al.*, 2012; McHugh *et al.*, 2015). It has several specialised domains responsible for its function in chromatin organisation (Figure 1.6). These domains include: a putative DNA binding domain (SAP domain) at the N-terminus, which facilitates low-affinity transient binding to DNA sequences like scaffold attachment regions (Fackelmayer *et al.*, 1994; Göhring, 1997); a SPRY domain with an unknown function and potential involvement in RNA binding (Miles, 2022); a catalytic AAA+ domain that may be involved in ATP binding and hydrolysis, potentially linked to oligomerisation (Nozawa *et al.*, 2017); a C-terminal RNA binding segment rich in arginine-glycine(-glycine) repeats (RG/RGG motif), enabling interaction with various cellular RNAs (Chong *et al.*, 2018); and intrinsically disordered regions (IDRs) between the SAP and SPRY domains, which likely play a crucial role in facilitating protein interactions (van der Lee *et al.*, 2014). Further research suggests that the AAA+ domain might instead be a P-loop NTPase domain, and might not be able to drive oligomerisation, at least *in vitro* (Miles, 2022). It is important to comment that the protein has two orthologs – HNRNPUL1 and HNRNPUL2 which have the same identifiable protein domains described above and are well conserved apart from the RNA binding segment (Miles, 2022). Yet they have not been shown to have a clear role in chromatin organisation like SAF-A. HNRNPUL1 has been suggested to have a role in transcription regulation and DNA repair whilst HNRNPUL2 might be involved in the DNA damage response (Kzhyshkowska *et al.*, 2003; Hong *et al.*, 2013; Allen, 2015).

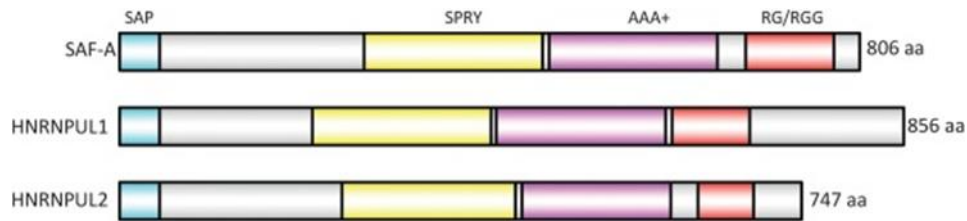


Figure 1.6. SAF-A and orthologs protein diagram. Diagram showing the domain similarities between SAF-A and its two orthologs, HNRNPUL1 and HNRNPUL2. They all have SAP (blue), SPRY (yellow), AAA+ (purple), and RG/RGG (red) motifs. From Nozawa and Gilbert, 2019.

1.2.1. SAF-A and chromatin organisation during interphase

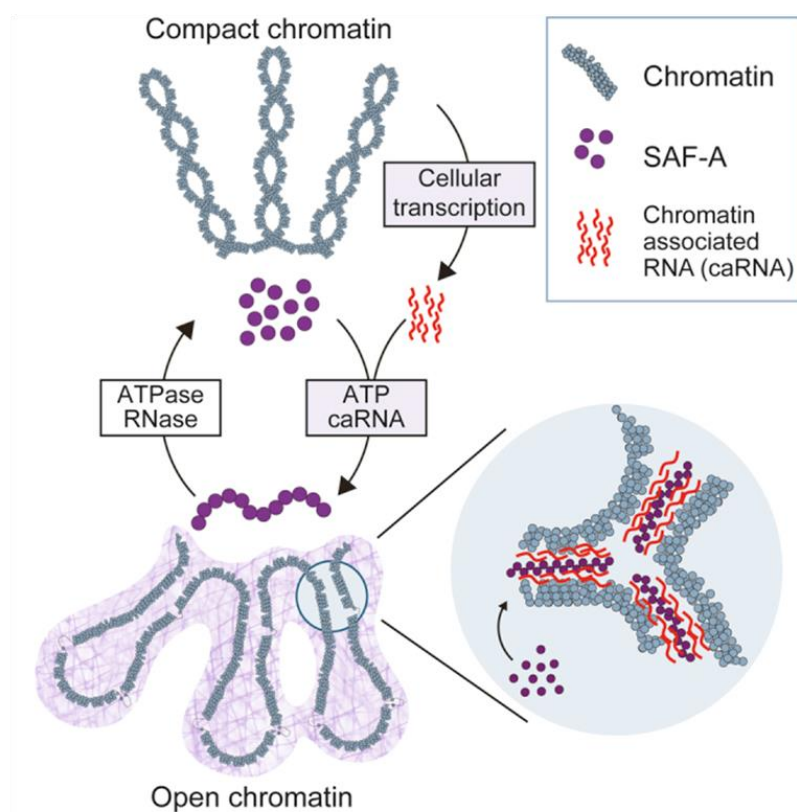
Recently the Gilbert lab demonstrated that SAF-A could act as a scaffold for interphase chromatin decompaction (Figure 1.7A). Using fluorescence *in-situ* hybridisation (FISH) as a probe for large scale chromatin structure, siRNA-mediated depletion of SAF-A in human RPE1 cells resulted in gene-rich/active/euchromatin regions becoming compacted whilst there was a minimal effect on gene-poor/inactive/heterochromatin regions (Nozawa *et al.*, 2017; Figure 1.7B). A study in mouse hepatocytes (AML12 cells) has similarly shown that global depletion of SAF-A with siRNA re-defines TAD boundaries, increases lamina-associated domains (LAD), and reduces chromatin loop formation (Fan *et al.*, 2018). As mentioned previously (Section 1.1.2.), transcription has been shown to decompact large-scale chromatin structures whilst transcription inhibition causes compaction of gene-rich/active regions. These findings provide for a potential link between transcription and chromatin decompaction, and enabled us to speculate that caRNAs might be required for SAF-A function. The identity of these caRNAs might not be as important as their abundance as previous CLIP-seq experiments show SAF-A binds to exons and introns at levels consistent with the respective amounts of these RNA species (Xiao *et al.*, 2012; Huelga *et al.*, 2012; Yugami *et al.*, 2020; Van Nostrand *et al.*, 2020). These studies also suggest that binding is non-specific as there is no well-defined consensus binding sequence or motif. Another key player in chromatin decompaction was the ability of SAF-A to oligomerise through its AAA+ domain, shown via crosslinking a FLAG-tagged version of the protein followed by fractionation on a SDS-PAGE gel. It was subsequently shown that SAF-A oligomerisation was

disturbed by treatment with α -amanitin (transcription inhibitor), RNase A/T1 (RNA hydrolase/degrader), and apyrase (ATP-diphosphohydrolase). Therefore, oligomer formation was dependent on transcription, RNA availability, and ATP availability. Mutations in the AAA+ domain at key residues (Walker A motif, Walker B motif) fixed SAF-A in a monomeric or oligomeric form by inhibiting its ability to bind ATP or hydrolyse ATP respectively (Nozawa *et al.*, 2017). The Walker A motif typically consists of GXXXXGK[T/S] amino acids as is required for ATP binding while the Walker B motif consists of hhhhD[D/E] amino acids and is required for ATP hydrolysis (Miller and Enemark, 2016). Mutations in these motifs proved to be useful tools for studying the function of SAF-A with further FISH experiments where SAF-A was depleted from cells using siRNA knockdown and a rescue assay was performed by transfecting with either wild-type or mutant versions of the protein. While re-introduction of wild-type SAF-A rescued the compact phenotype of gene-rich regions, SAF-A-WA (Walker A mutant; K491A) was unable to decompact the active chromatin regions while SAF-A-WB (Walker B mutant; D561A) held them in a constitutively open position, even after transcription inhibition (Nozawa *et al.*, 2017). Furthermore, deletion of either the RNA or DNA binding domains of SAF-A also leads to a reduction in the ability of the protein to decompact chromatin.

It is not entirely clear what attracts SAF-A to gene-active regions and how this differs from its reported role in securing XIST to inactive chromatin. It has been shown that SAF-A can interact with chromatin, other hnRNP proteins, RNA pol II, RNA, and telomeres (Göhring and Fackelmayer, 1997; Vizlin-Hodzic *et al.*, 2011; Nozawa *et al.*, 2017; Izumi and Funa, 2019; Sharp *et al.*, 2020; Miles *et al.*, 2022). However, in many of these cases, the interaction does not appear to be of high affinity nor stable, which might mean SAF-A is merely in the vicinity of these molecules and proteins. SAF-A has an IDR, a domain which is frequently implicated in phase separation but no evidence for this ability has been found yet for SAF-A. If SAF-A is involved in sequestering nuclear processes by creating condensates and/or an RNA mesh, that would naturally put it in the vicinity of RNA, DNA, and proteins. Super-resolution microscopy using the dSTORM method does suggest that SAF-A forms an interconnected mesh of clusters which differs from that of other RBPs like SC-35 and HNRNPC, but further experiments are needed to show if the mesh is organised or created by ncRNAs (Marenda *et al.*, 2021). Or it could simply be a stabiliser of

transcriptionally active chromatin, recruited to these sites by initial bursts of transcription. Its differing roles in inactive and active chromatin regions could also be regulated by post-translational mechanisms. SAF-A is equally inherited by the daughter cells after mitosis and it has a diverse set of post-translational modifications, including phosphorylation, acetylation, and ubiquitylation (Samejima *et al.*, 2022). However, post-translational modifications have not been studied in detail to understand their role in SAF-A regulation. Still, recent research into SAF-A's role in chromatin organisation during interphase suggests that it is RNA, specifically repeat-rich ncRNA, which might attract SAF-A to actively transcribed regions (Creamer *et al.*, 2021). This would put SAF-A-mediated chromatin decompaction downstream of transcription, which is confirmed when harshly inhibiting transcription (Nozawa *et al.*, 2017; Skalska *et al.*, 2021). However, less severe and more precise experiments are required to confirm this, especially as RNA's role in genome organisation and phase separation is still being investigated.

A



B

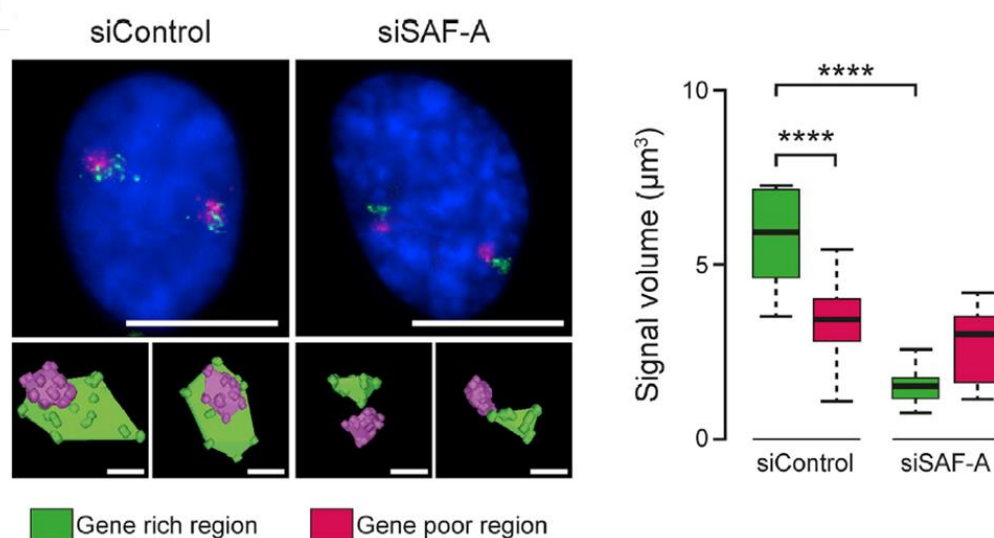


Figure 1.7. SAF-A interphase organisation. A) Schematic representation of the hypothesised method behind SAF-A opening up chromatin. SAF-A is present as monomer (purple) and in the presence of ATP and chromatic-associated RNAs (caRNA, in red) it forms oligomers and keeps chromatin in an open conformation. The zoomed-in area shows oligomeric SAF-A and caRNA forming a mesh-like structure to facilitate an open chromatin state. B) Fluorescence *in-situ* hybridisation (FISH) experiment with labelled oligo probes covering either a gene-rich (green) or gene-poor (red) region in the presence of short interfering RNAs. Left are confocal images of representative nuclei with scale bars of 10 μm . Below is the 3D reconstruction of regions with scale bars of 1 μm . Right are box plots quantifying the volume occupied by the regions with over 20 nuclei counted per condition. From Nozawa *et al.*, 2017.

1.2.2. SAF-A and chromatin organisation during mitosis

Fixed cell imaging during mitosis has revealed that SAF-A is excluded from mitotic chromosomes compared to its diffuse nuclear staining pattern in interphase (Sharp *et al.*, 2020; Figure 1.8). This exclusion is triggered by the phosphorylation of two serine residues (Ser 14 and Ser 26) in the SAP domain by Aurora-B kinase during prophase. Aurora-B is important for early mitosis events such as cohesin removal and condensin I loading for chromosome formation (Losada *et al.*, 2002; Lipp *et al.*, 2007). Preventing phosphorylation at Ser 14 and Ser 26 via point mutations maintains SAF-A binding to the chromatin which then leads to segregation defects during anaphase. Phosphorylation of serine 59 by Polo-like kinase 1 is also only present during mitosis and its disruption leads to the accumulation of unaligned chromosomes during anaphase but its function is otherwise unknown (Douglas *et al.*, 2015). SAF-A might be co-localised with nuclear RNA during mitosis, which are also excluded from mitotic chromosomes (Sharp *et al.*, 2020). Therefore, Sharp and colleagues hypothesise that it is the phosphorylation of SAF-A and subsequent eviction from chromatin that sequesters the RNA mesh away from mitotic chromatin. As RNAs can disrupt chromatin structure, and antagonise or promote the binding of proteins, it is important to separate them from chromosomes. If SAF-A-RNA complexes are not excluded from chromosomes that could disrupt the chromosome formation process and thus cause segregation defects. However, further experiments are needed to show whether it is SAF-A that clears nuclear RNA from chromosomes. Following mitosis, SAF-A and some nuclear RNAs are proposed to be relocated back to chromatin, the timing of which has not yet been investigated (Blower *et al.*, 2023). It remains unknown whether SAF-A re-localisation to chromatin actively promotes the decompaction of gene loci to facilitate transcription or if transcriptional activity triggers SAF-A re-localisation. Recent studies demonstrating bursts of transcriptional activity during mitosis raise questions regarding the significance of this spike in activity. Genes marked by H3 lysine 27 acetylation are preferentially primed for RNA pol II transcription re-initiation and are predominantly located in active regions of the genome (Hsiung *et al.*, 2016). This transient transcriptional surge may contribute to the formation of a nuclear RNA mesh involving SAF-A, although further investigation is required to establish its precise role.

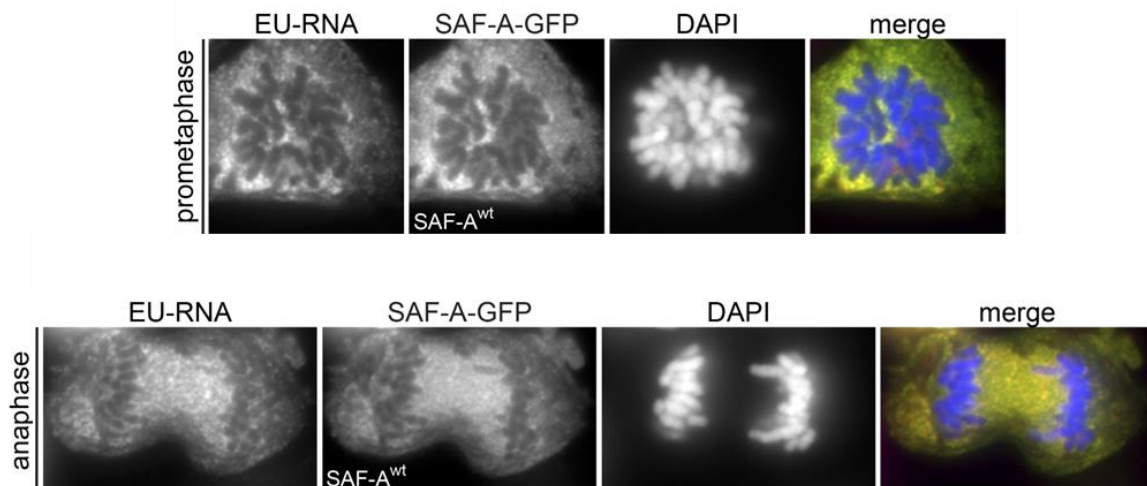


Figure 1.8. SAF-A and RNA localisation during mitosis. Microscopy images of cell nuclei during prometaphase and anaphase of cells transiently expressing SAF-A^{wt}-GFP and immunostained for GFP (green). 5-ethynyl uridine (EU-RNA) labelling detects nuclear RNA and DAPI labelling detects DNA. From Sharp *et al.*, 2020.

1.2.3. Other SAF-A functions

Apart from SAF-A's role in chromatin organisation, it has also been shown to be a key component of the nuclear scaffold and has been implicated in transcription elongation, RNA splicing, and lncRNA localisation. When SAF-A was discovered, it was shown to bind to scaffold-attachment region DNA in chromatin, hence its name, Scaffold Attachment Factor-A (Romig *et al.*, 1992; Fackelmayer *et al.*, 1994; Göhring & Fackelmayer, 1997). The binding of SAF-A to these DNA regions helps to maintain the nuclear architecture. When SAF-A was cleaved by Caspase-3 during apoptosis, the link between DNA and proteins was broken which collapsed the cell nucleus (Göhring *et al.*, 1997; Kipp *et al.*, 2000). SAF-A can also interact with different proteins that regulate transcription elongation as part of protein complexes. Together with β -actin and PCAF (a histone acetyltransferase), SAF-A associated with the phosphorylated C-terminal domain (serine 2 or 2 and 5) of RNA pol II and was required for active transcription elongation (Kukalev *et al.*, 2005; Obrdlik *et al.*, 2008). In contrast, SAF-A binding on its own to RNA pol II instead inhibited the C-terminal domain phosphorylation by TFIIH and Cdk9 thus repressing transcription elongation (Kim and Nikodem, 1999). In terms of RNA splicing, SAF-A can bind to the introns of

pre-mRNAs, is necessary for exon inclusion and is implicated in exon definition events (Huelga *et al.*, 2012; Xiao *et al.*, 2012; Ye *et al.*, 2015). SAF-A was also involved in the maturation of the 17S U2 small nuclear ribonucleoprotein giving it an additional indirect role in splicing (Xiao *et al.*, 2012). Finally, as mentioned before, SAF-A is important for the localisation of lncRNAs like XIST (see Section 1.1.1.3.), C₀T-1, FIRRE, and PANDA. C₀T-1 RNA can localise to euchromatin by tethering to it via SAF-A while a dominant negative mutant version of SAF-A (Cysteine 280) disrupted the nuclear scaffold and dispersed the RNAs (Fackelmayer *et al.*, 1994; Hall *et al.*, 2014). Similarly, SAF-A depletion caused XIST to diffuse through the nucleoplasm (Hasegawa *et al.*, 2010). The lncRNA FIRRE can interact with SAF-A through its RNA-repeating domain and is thus retained in the nucleus (Hacisuleyman *et al.*, 2014). Knockdown of SAF-A resulted in the loss of proximity between FIRRE and its chromatin binding sites (Hacisuleyman *et al.*, 2014). Other RNAs that contain an RNA-repeating domain were also found to be retained in the nucleus by SAF-A (Hacisuleyman *et al.*, 2016). Finally, PANDA requires SAF-A to localise it to senescence genes in both proliferating and senescent cells (Puvvula *et al.*, 2014). PANDA can then interact with the polycomb repressive complexes PRC1 and PRC2 to inhibit senescence or with the transcriptional factor NF- κ B to promote it (Puvvula *et al.*, 2014). Most of the mentioned SAF-A functions rely on its dual interaction with RNAs and chromatin though it is not known how SAF-A potentially switches between the different functions. SAF-A's function in transcriptional elongation was dependent on different protein binding partners while other functions might require changes in its PTMs which have not been extensively studied so far.

1.2.4. SAF-A mutations in a rare neurodevelopmental disorder

In the DECIPHER database, SAF-A has the maximum pLI score of 1, meaning that it is highly intolerant of LOF mutations and insufficient dosage of the protein in cells would leave to haploinsufficiency. This has been corroborated by mouse studies where a hypomorphic mutation of SAF-A was shown to cause early embryonic lethality with apparent abnormalities in mouse embryos after 6.5 days of development (Roshon and Ruley, 2005). In humans, SAF-A initially stood out as a candidate gene

for a neurodevelopmental disorder with large-scale deletions in the 1q44 chromosome region due to its known involvement in embryonic brain development (Caliebe *et al.*, 2010). A couple of years later, microdeletions within the 1q44 region covering SAF-A were linked to the development of a brain disorder with symptoms including early-onset epileptic seizures, developmental delay, and intellectual disability (Ballif *et al.*, 2012; Nagamani *et al.*, 2012; Allen *et al.*, 2013; Gupta *et al.*, 2014). Only in the past decade has exome analysis of patients with similar symptoms or patient-parent trios of families with autism spectrum disorder found *de novo* SAF-A mutations (Bramswig *et al.*, 2017; Leduc *et al.*, 2017; Lim *et al.*, 2017; Yates *et al.*, 2017). Over 100 specific mutations in patients have been reported where it is clear they are the principal cause of the disorder (Figure 1.9A). The phenotypes fit well with the expression pattern of SAF-A – it is found in the developing and adult brain, in the cerebellum, heart, kidney, and liver. Currently, limited research has been conducted on the mechanisms underlying these protein mutations and the HNRNPU-related neurodevelopmental disorder (OMIM 602869), with information primarily limited to the severe patient phenotype and some variability among individuals.

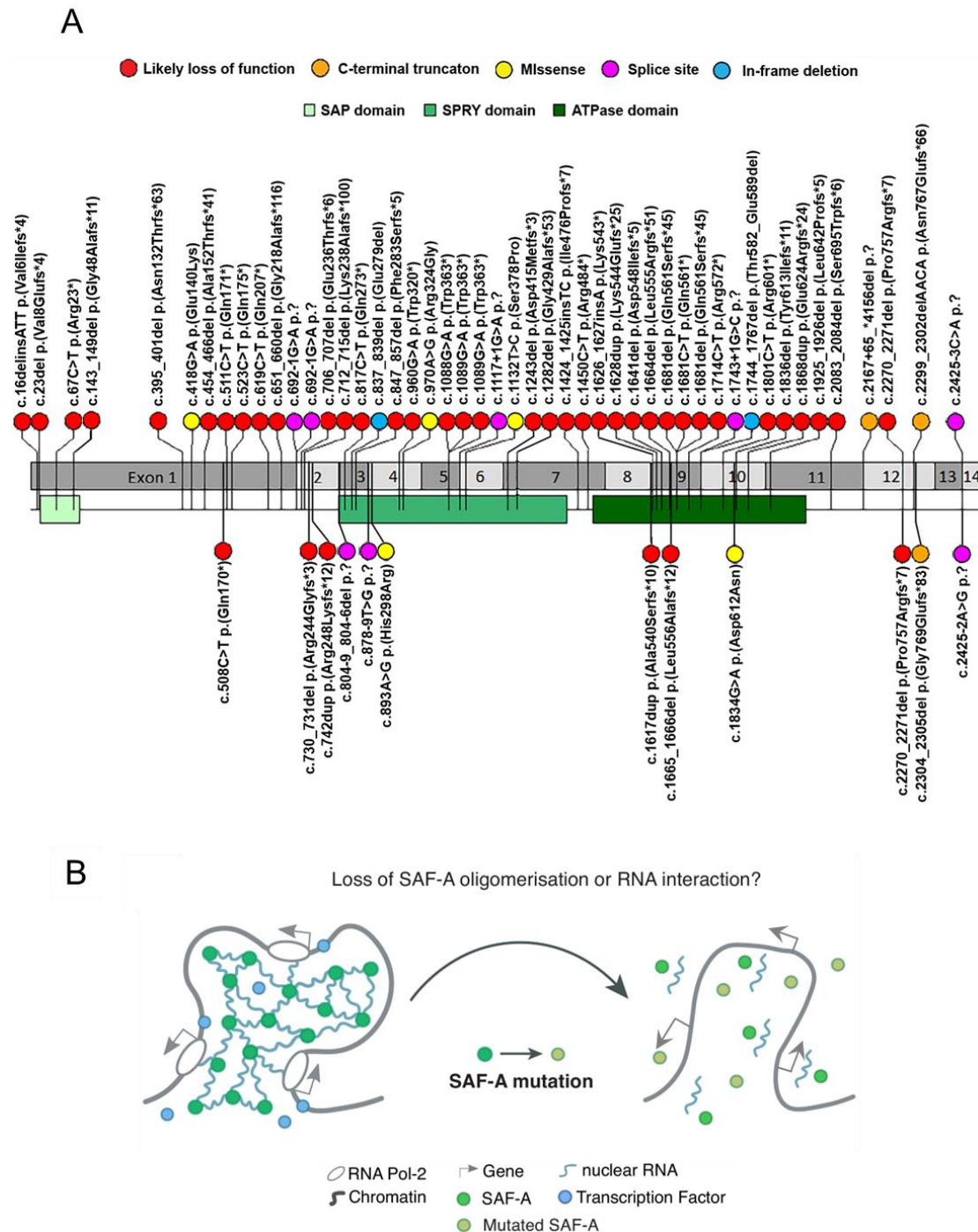


Figure 1.9. SAF-A mutations and potential mechanism. A) Plot showing most SAF-A patient mutations discovered to date. From Taylor *et al.*, 2022. B) Schematic of potential mechanism behind SAF-A missense mutations - affecting nuclear mesh formation and decompaction of active chromatin regions. Adapted from Marenda *et al.*, 2022.

In a mouse study, a heterozygous frameshift LOF mutation from the deletion of SAF-A exons 3 to 6 resulted in 8 out of 11 animals exhibiting a reversed activity pattern, with significantly reduced activity at night and increased activity during the day with the remaining animals having completely arrhythmic behaviour in constant darkness conditions (Lai *et al.*, 2018). This mutation is hypothesised to lead to haploinsufficiency with visibly less protein observed in the suprachiasmatic nucleus (SCN). This links to previous research that found SAF-A expressed in a circadian manner in mouse SCN and that it has a role in regulating the oscillatory transcription of *Bmal1* in cells (Tamaru *et al.*, 2003; Onishi *et al.*, 2008). In the 2018 study, when SAF-A was depleted, there were significant differences found in the expression levels of the clock genes, *Avp* and *Vip*, which regulate the circadian rhythm. As SAF-A might be involved in maintaining an RNA mesh around active gene regions, including those of Clock genes, a perturbation of its protein levels could affect the chromatin decompaction. The less open chromatin could then reduce the amount of bound *Bmal1*/Clock to the *Avp* and *Vip* promoters. A more recent study in mice introduced a heterozygous frameshift mutation (113 bp deletion in exon 1) in SAF-A leading to a truncated version of the protein that was not expressed, possibly due to the nonsense-mediated decay (NMD) pathway (Dugger *et al.*, 2023). The truncated SAF-A mutant was not detected by Western blot yet there was no significant reduction in overall SAF-A protein and mRNA levels. This hints that this essential protein might have a compensatory mechanism to boost mRNA levels and protein production to near-normal levels. Even with this compensation, the mice still had under- and overexpressed genes, specifically in the pyramidal cells in the subiculum of the hippocampus. The differentially regulated genes were linked to neurodevelopmental phenotypes such as developmental delay, epilepsy, and autism. However, the mice did not show any brain abnormalities like reduced size or aberrant corpus callosum morphology similar to Lai *et al.*, (2008) showing no changes to SCN morphology. At the same time, the mutant mice in Dugger *et al.*, (2023) were smaller than wild-type, had developmental delay and had an increased predisposition to seizures but none of that was observed in the Lai *et al.*, (2008) study. These results hint at how the neurodevelopmental disorder associated with SAF-A mutations develops in patients where insufficient SAF-A might affect chromatin organisation and influence key developmental events.

Most *de novo* mutations observed in patients are frameshifts and protein truncations which probably lead to haploinsufficiency in the developing brain by activating the NMD pathway (Table 1.1). However, we do not know if single-codon deletions and missense mutations would result in the decay of the altered protein (Figure 1.9B). In neurodevelopmental diseases, it is common to have dominant negative mutations, where the protein retains its correct folding, but mutated key residues hinder its functional ability (Backwell and Marsh, 2022). This could be the case with mutations in the SPRY-ATPase domain, where there have been fewer missense mutations found than expected. The ATPase function could be perturbed in a correctly folded SAF-A protein, which would not trigger the cell's compensatory mechanism for maintaining SAF-A abundance, as seen with mutations resulting in protein truncation. Consequently, this would affect SAF-A's overall ability to decompact transcriptionally active regions at the appropriate time or keep them constitutively open, similar to the effects of WA and WB motif mutations. While most SAF-A mutations in patients so far are *de novo*, there has been a rare case of inheriting a SAF-A mutation, possibly in an autosomal dominant manner (Wang *et al.*, 2020). Other hnRNP proteins have also been implicated in neurodevelopmental disorders, as mentioned in Section 1.1.1.4. An interesting case was found in ALS where SAF-A and HNRNPQ came up as top candidates for disease pathology but were not mutated (Bakkar *et al.*, 2018). Their mRNA and protein distributions were altered as well as their protein levels. This was then linked to SAF-A and HNRNPQ being hypothetical interacting partners of proteins often mutated in ALS like FUS, TDP-43, and Ataxin2 but as of yet, no known mechanism of how SAF-A mutations or perturbation contribute to neurodevelopmental disorders has been discovered.

Protein region	Nº of total mutations	Nº of frameshifts	Nº of truncations	Nº of substitutions	Nº of deletions	Nº of splicing variants
SAP domain	5	4	0	0	1	0
IDR 1	32	19	7	4	0	2
SPRY domain	22	4	7	5	3	3
Connection	3	1	1	1	0	0
AAA+ domain	32	19	7	2	2	2
IDR 2	1	0	0	0	1	0
RNA domain	19	10	3	2	1	3
Introns	0	0	0	0	0	0
Total	114	57	25	14	8	10

Table 1.1. SAF-A mutations by type and protein region.

1.3. Hypothesis and aims

Large-scale chromatin organisation has been investigated for decades in interphase and mitosis and even mitotic entry, but the mitotic exit has remained a poorly studied event. It is known that daughter cells re-establish the chromatin architecture to facilitate their cellular programming, yet the protein actors and processes involved have not been elucidated. SAF-A is an essential protein with strong evidence that it has a significant role in large-scale chromatin organisation in interphase. Its dissociation from mitotic chromosomes led me to hypothesise that it needs to be recruited in a precise manner to active gene regions to establish and/or maintain the RNA mesh for high levels of transcription. SAF-A's recruitment to active gene regions could be initiated by high levels of transcriptional bursting during telophase or by the upregulation of genes in G₁. However, SAF-A's links to the re-establishment of transcription have not been investigated. In addition, the mechanism that drives the phenotype in the HNRNPU-related neurodevelopmental disorder could also be linked to SAF-A's role in large-scale chromatin organisation. Perturbation of the mechanism could lead to faulty chromatin architecture at gene active regions. Therefore, the overall aim of this thesis was to probe the behaviour of SAF-A at the mitotic exit and investigate the cellular phenotypes of key patient mutations. To achieve this, I established three aims:

- 1) To describe SAF-A's role in chromatin decompaction at mitotic exit:

In Chapter 3, I use live cell imaging to analyse the distribution of SAF-A during mitosis and mitotic exit and then quantified the results with colocalisation analysis. This enabled me to check whether SAF-A is associated with RNAs during those times using synchronised cells.

- 2) To understand SAF-A's interaction with RNA:

In Chapter 4, I investigated how SAF-A depletion affects the RNA mesh by using super-resolution microscopy and use a metabolic labelling method to investigate the kinetics of nascent RNA.

3) To investigate the effect of SAF-A patient mutations on chromatin organisation:

In Chapter 5, I analysed the available literature on patient mutations and highlighted key mutations to investigate further. I then made human cell lines expressing mutated proteins and quantified their effect on chromatin decompaction using FISH.

2. Chapter 2: Materials and methods

2.1. Reagents, solutions, and buffers

Reagent	Product code
1 Kb Plus DNA Ladder	10787018, Invitrogen
1,8-bismaleimido-diethyleneglycol (BM(PEG) ₂)	22336, Thermo Scientific
2,2,2-Trifluoroethylamine (TFEA)	269042, Sigma-Aldrich
2-mercaptoethanol	M3148, Sigma-Aldrich
2-propanol (Isopropanol)	109827, Sigma-Aldrich
3-Indoleacetic acid	I3750, Sigma-Aldrich
30% Acrylamide	SBL-20-2100-05, Severn Biotech
37% Formaldehyde	F8775, Sigma-Aldrich
3X FLAG® Peptide	F4799, Sigma-Aldrich
4-thiouridine (4sU)	T4509, Sigma-Aldrich
5-ehynyl uridine (5-EU)	CLK-N002-10, Jena Bioscience
Aarl	ER1581, Thermo Scientific
Acetic acid (glacial)	695092, Sigma-Aldrich
Acid-phenol:Chloroform	AM9720, Invitrogen
Actinomycin D	A4262, Sigma-Aldrich
Advanced DMEM/F12	12634010, Gibco
Afel	R0652S, New England Biolabs
Agarose	A9539, Sigma-Aldrich
Alel-v2	R0685S, New England Biolabs
Amersham™ Hybond™-XL	RPN2020S, Cytiva
Ammonium Persulfate (APS)	A3678, Sigma-Aldrich
AMPure XP Beads	A63882, Beckman Coulter
Aphidicolin	178273, Calbiochem
Atto 488 Azide	72709, Sigma Aldrich
BamHI-HF	R3136S, New England Biolabs
BbsI	FD1014, Promega
BglII	R0144S, New England Biolabs
BisBenzimide H 33342 Trihydrochloride (Hoechst 33342)	B2261, Sigma-Aldrich
Blasticidin S Hydrochloride	B001, TOKU-E
Boric acid/borate	B0394, Sigma-Aldrich
Bovine Serum Albumin (BSA)	B2518, Sigma-Aldrich
BstXI	R0113S, New England Biolabs
Catalase from bovine liver	C30-100MG, Sigma-Aldrich
Chloramphenicol	C0378, Sigma-Aldrich
Chloroform	288306, Sigma-Aldrich
Countess™ Cell Counting Chamber Slides	C10228, Invitrogen
D-(+)-glucose	G6152, Sigma-Aldrich
DAPI (4',6-diamidino-2-phenylindole)	D8417, Sigma-Aldrich
DH5α Competent Cells for Subcloning	EC0111, Thermo Scientific
Dimethyl sulfoxide (DMSO)	900645, Sigma-Aldrich

Disuccinimidyl suberate (DSS)	21655, Thermo Scientific
DL-Dithiothreitol (DTT)	D9779, Sigma-Aldrich
DNA polymerase I, Large (Klenow) Fragment	M0210S, New England Biolabs
DNase I	04716728001, Roche
dNTP Set	10297018, Invitrogen
Doxycycline hyclate	D9891, Sigma-Aldrich
DRAQ5™	424101, BioLegend
dTAG-13	6605, Tocris
Dulbecco's Modified Eagle Medium (DMEM), phenol-red	41965039, Gibco
Dulbecco's Modified Eagle Medium, no phenol-red	31053028, Gibco
dUTP (ChromaTide AF594-5-dUTP)	C11400, Invitrogen
dUTP (SEEBRIGHT Green 500)	ENZ-NUC117-0025, Enzo Lifesciences
Dynabeads™ M-280 Sheep Anti-Mouse IgG	11202D, Invitrogen
EcoRV-HF	R3195S, New England Biolabs
Epredia™ SuperFrost Plus™ Adhesion slides	10149870, Thermo Scientific
Ethanol	51976, Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	E9884, Sigma-Aldrich
Ethidium Bromide (EtBr)	E1510, Sigma-Aldrich
Formamide	F9037, Sigma-Aldrich
FseI	R0588S, New England Biolabs
G50 Sephadex Spin Columns	G50DNARO, Roche
Gibson Assembly® Master Mix	E2611S, New England Biolabs
Glucose Oxidase from Aspergillus niger	G7141, Sigma-Aldrich
Glycogen	G1508, Sigma-Aldrich
HindIII-HF	R3104S, New England Biolabs
Human Cot-1 DNA	15279011, Invitrogen
Hygromycin B	10843555001, Roche
Hyperfilm™ ECL™ film (18 × 24 cm)	28906837, Cytiva
Immobilon®-P PVDF Membrane	IPVH00010, Millipore
Iodoacetamide	I1149, Sigma-Aldrich
KpnI-HF	R3142S, New England Biolabs
L-alpha-Lysophosphatidylcholine from egg yolk (Lysolecithin, LL)	L4129, Sigma-Aldrich
L-ascorbic acid	95209, Sigma-Aldrich
LightCycler® 480 SYBR Green I Master	04707516001, Roche
Lipofectamine™ RNAiMAX Transfection Reagent	13778150, Invitrogen
Lysozyme	L6876, Sigma-Aldrich
Magnesium Chloride	M9272, Sigma-Aldrich
Methanol	179337, Sigma-Aldrich
μ-Dish, 35 mm high glass bottom	81158, Ibidi
MOPS	69947, Sigma-Aldrich
Neomycin	N6910, Sigma-Aldrich
Nocodazole	487928, Sigma-Aldrich
Nonidet™ P-40 (NP-40) substitute	74385, Sigma-Aldrich
NotI-HF	R3189S, New England Biolabs

Nunc™ Lab-Tek™ II Chambered Coverglass	155360, Thermo Scientific
NuPAGE™ 3 to 8%, Tris-Acetate, 1.0–1.5 mm, Mini Protein Gels	EA03755BOX, Invitrogen
NuPAGE™ 4 to 12%, Bis-Tris, 1.0–1.5 mm, Mini Protein Gels	NP0322BOX, Invitrogen
NuPAGE™ LDS Sample Buffer (4X)	NP0007, Invitrogen
NuPAGE™ MOPS SDS Running Buffer (20X)	NP0001, Invitrogen
NuPAGE Tris-Acetate SDS Running Buffer (20X)	LA0041, Invitrogen
Opti-MEM™ I Reduced Serum Medium	31985070, Gibco
Pang Supersolution	43100024
Paraformaldehyde	158127, Sigma-Aldrich
pEGFP-C1	6084-1, Clontech
pEZ-M03-hnRNPC1-eGFP	EX-H345-M03, tebu-bio
Phase Lock Gel Heavy Tubes	0032.005.152, Eppendorf
Phenol:Chloroform:Isoamyl alcohol	15593031, Invitrogen
Phenylmethylsulfonyl fluoride (PMSF)	78830, Sigma-Aldrich
Picolyl-Azide-Sulfo-Cy5	CLK-1177-1, Jena Bioscience
Pierce™ ECL Western Blotting Substrate	32106, Thermo Scientific
pmCherry-C1	632524, Clontech
Pmel	R0560S, New England Biolabs
Poly-L-lysine solution	P4832, Sigma-Aldrich
Ponceau S	P7170, Sigma-Aldrich
Potassium chloride	P9541, Sigma-Aldrich
PrestoBlue™ Cell Viability Assay	A13262, Invitrogen
Propidium iodide (PI)	P4170, Sigma-Aldrich
Proteinase K	P81025, New England Biolabs
PshAI	R0593S, New England Biolabs
PureLink™ RNase A	12091039, Invitrogen
Puromycin Dihydrochloride	A1113803, Gibco
Gel Loading Dye, Purple (6X)	B7024S, New England Biolabs
QIAshredder Column	79656, Qiagen
RNase Cocktail™ RNase A / T1	AM2286, Invitrogen
RNasin® Plus Ribonuclease Inhibitor	N2611, Promega
RO-3306	15149, Cayman Chemical Company
Roche LightCycler® 480 SYBR Green I Master	04887352001, Roche
Salmon Sperm	AM9680, Ambion
Sbfl-HF	R3642S, New England Biolabs
SeeBlue Plus2 Pre-stained Protein Standard	LC5925, Sigma-Aldrich
SiR-DNA	SC007, Spirochrome
SNAP-Cell® Oregon Green®	S9104S, New England Biolabs
SNAP-Cell® TMR-Star	S9105S, New England Biolabs
Sodium Acetate	567422, EMD Millipore
Sodium Bicarbonate	S5761, Sigma-Aldrich
Sodium Chloride	S9888, Sigma-Aldrich
Sodium periodate	419610050, Acros Organics
Sodium Dodecyl Sulphate (SDS)	BP2436200, Fisher Scientific

Spel	R0133S, New England Biolabs
Stealth RNAi™ siRNA Negative Control, Med GC	12935300, Invitrogen
Stealth RNA target sequence: hnRNPUL1	HSS145966, Invitrogen
Stealth RNA target sequence: SAF-A	HSS104917, Invitrogen
Stealth RNA target sequence: SAF-A 2	HSS104918, Invitrogen
Storage Phosphor Screen BAS-IP TR 2040 E	28-9564-81, Cytiva
TapeStation D1000 High Sensitivity Reagents	5067-5585, Agilent
TapeStation D1000 High Sensitivity ScreenTape	5067-5584, Agilent
TapeStation RNA High Sensitivity Ladder	5067-5581, Agilent
TapeStation RNA High Sensitivity Sample Buffer	5067-5580, Agilent
TapeStation RNA High Sensitivity ScreenTape	5067-5579, Agilent
TEMED (N,N,N',N'-Tetramethylethylenediamine)	T7024, Sigma-Aldrich
Thymidine	T9250, Sigma-Aldrich
Triptolide	T3652, Sigma-Aldrich
Tris	678937, Sigma-Aldrich
Tris base	TRIS-RO, Sigma-Aldrich
Tris-glycine	T9784, Sigma-Aldrich
Triton™ X-100	T8787, Sigma-Aldrich
TRIzol Reagent	T9424, Sigma-Aldrich
Trypan Blue Solution, 0.4%	15250061, Gibco
Tween-20	P1379, Sigma-Aldrich
Trypsin-EDTA solution (10X)	T4174, Sigma-Aldrich
UltraPure™ Agarose	16500500, Invitrogen
Uridine [5-3H]	ART0154, American Radiolabeled Chemical
VECTASHIELD® Antifade Mounting Medium	H-1000, Vector Laboratories
Whatman 3MM Chr Chromatography Paper	3030-392, Cytiva
Xbal	RO145S, New England Biolabs
XcmI	R0533S, New England Biolabs

Table 2.1. Reagents.

Antibody	Product code	Use
AffiniPure Fab Fragment Donkey Anti-Mouse IgG (H+L)	715-007-003, Jackson Immuno Research	dSTORM
AffiniPure F(ab') ₂ Fragment Donkey Anti-Rabbit IgG (H+L)	711-006-152, Jackson Immuno Research	dSTORM
Anti-Histone H3 (phospho S10) antibody	ab47297, Abcam	Western blotting
Anti-hnRNP C1/C2 Antibody (4F4)	sc-32308, Santa Cruz Biotechnology	dSTORM
Anti-hnRNP U, clone 3G6	05-1516, Sigma-Aldrich	Western blotting
Anti-hnRNP-U Polyclonal	a300-689a, Bethyl Laboratories	Immunofluorescence

Anti-HNRNPUL1 antibody	HPA046290, Atlas antibodies	Western blotting
Anti-SC35	ab11826, Abcam	dSTORM
GAPDH (14C10) Rabbit mAb	2118, Cell Signalling	Western blotting
Goat anti-Mouse IgG (H+L), Alexa Fluor 488	A11029, Invitrogen	Immunofluorescence
Goat anti-Rabbit IgG (H+L), Alexa Fluor 647	A21244, Invitrogen	Immunofluorescence
Goat Anti-rabbit IgG-HRP	A0545, Sigma-Aldrich	Western blotting
Monoclonal ANTI-FLAG® M2	F1804, Sigma-Aldrich	Western blotting, immunoprecipitation, immunofluorescence
Phospho-Histone H3 (Ser10) (D2C8) XP® Rabbit mAb (Alexa Fluor® 488 Conjugate)	3465S, Cell Signaling Technology	Flow cytometry cell cycle assessment
Sheep anti-mouse IgG-HRP	NXA931, GE Healthcare	Western blotting

Table 2.2. List of antibodies. Information about how these antibodies were used can be found in Section 2.4.3., 2.5., and 2.8.3. Working dilutions can be found in Table 2.13, 2.14, and 2.15.

Kit	Product code
E.Z.N.A. Plasmid Mini Kit I	D6942-01, OMEGA bio-tek
Lipofectamine™ 3000	L3000015, Invitrogen
NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®	E7103, New England Biolabs
One Step PrimeScript™ III RT-PCR Kit	R600A, Takara Bio
RNeasy kit	74004, Qiagen
SuperScript™ III Reverse Transcriptase	18080093, Invitrogen
Q5 Hot Start High-Fidelity DNA Polymerase	M0493S, New England Biolabs
Qiaprep Spin Miniprep Kit	27104, Qiagen
QIAquick Gel Extraction Kit	28704, Qiagen
QIAquick PCR Purification Kit	28104, Qiagen
Qubit dsDNA BR Assay Kit	Q32853, Thermo Scientific
Qubit dsDNA HS Assay Kit	Q32854, Thermo Scientific
Qubit RNA BR Assay Kit	Q10211, Thermo Scientific
Qubit RNA HS Assay Kit	Q32855, Thermo Scientific
T4 DNA Ligase	M0202S, New England Biolabs
TURBO DNase	AM2239, Invitrogen
ZymoPURE II Plasmid Midiprep Kit	D4200, Zymo Research

Table 2.3. Kits.

Solution	Components
15% Main Polyacrylamide Gel	8 ml total volume: 22.5% ddH ₂ O, 50% of 30% Acrylamide, 25% ml 1.5 M Tris (pH 8.8), 1% of 10% SDS, 1% of 10% APS, 0.1% TEMED
4% Stacking Polyacrylamide Gel	5 ml total volume: 60% ddH ₂ O, 13.4% of 30% Acrylamide, 25% 0.5 M Tris (pH 6.8), 1% of 10% SDS, 1% of 10% APS, 0.1% TEMED
Formaldehyde Gel Buffer (10X)	200 mM MOPS, 50 mM NaOAc, 10 mM EDTA, pH to 7.0
Formaldehyde Gel Running Buffer	1X formaldehyde gel buffer, 0.74% formaldehyde, to 1 L with H ₂ O
Freezing Media	48% Ham's F-10 Nutrient Mix, 40% Tryptose Phosphate Broth, 12% DMSO. Prepared by IGC Technical Services.
Hypotonic Buffer	20 mM Tris-HCl (pH 7.4), 10 mM KCl, 2 mM MgCl ₂ , 1 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF
Isotonic Buffer	20 mM Tris-HCl (pH 7.4), 150 mM KCl, 2 mM MgCl ₂ , 1 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF
Luria-Bertani (LB) Broth	Per 1 L: 10 g tryptone, 5 g yeast extract, 10 g NaCl, 1 g glucose, pH to 7 with NaOH. Prepared by IGC Technical Services.
Oligomerisation Assay Buffer	125 mM NaCl, 50 mM Tris (pH 7.5) or 50 mM PIPES pH 7, 5 mM MgCl ₂ , 0.05% Triton X-100, 1 mM PMSF, 12.5 µM DTT, 1% (v/v) RNasin ribonuclease inhibitor
Oxygen Scavenger Imaging Buffer	8.9 ml total volume: 5.3 ml of 200 mM Tris and 50 mM NaCl, 2 ml of 40% glucose, 200 µl GLOX (160 µl of 200 mM Tris and 50 mM NaCl, 40 µl catalase, 18 mg glucose oxidase), 1.32 ml of 1 M 2-mercaptoethanol
Phosphate Buffer Saline (PBS)	160 mM NaCl, 3 mM KCl, 8 mM Na ₂ HPO ₄ , 1 mM KH ₂ PO ₄ from tablets. Per 1 Litre: 10 tablets of Dulbecco's Unipath (Oxoid). Prepared by IGC Technical Services.
RNA Gel Loading Buffer (5X)	0.16% (v/v) saturated bromophenol blue, 4 mM EDTA (pH 8.0), 2.66% formaldehyde, 20% (w/v) glycerol, 32.4% formamide, 40% (v/v) 10X formaldehyde gel buffer
Sodium Phosphate Buffer	200 ml total volume of 0.5 M solution: 93.2 ml of 1 M Na ₂ HPO ₄ and 6.8 ml of 1 M NaH ₂ PO ₄ , H ₂ O to 200 ml
SSC (20X)	3 M Sodium Chloride, 300 mM Sodium Citrate. Prepared by IGC Technical Services.
TAE Buffer (10X)	400 mM Tris Base, 200 mM Glacial Acetic Acid, 100 mM EDTA
TBE Buffer (10X)	890 mM Tris base, 890 mM Borate, 20 mM EDTA
TBS-T Buffer	50 mM Tris (pH 7), 150 mM NaCl, 0.05%

	(v/v) Tween-20
TE Buffer	10 mM Tris-HCl (pH 7.6), 1 mM EDTA. Prepared by IGC Technical Services.
Transfer Buffer for Western Blots	24 mM g Tris Base, 14.4 g glycine (192 mM), 20% methanol (v/v). Dissolved in water before methanol addition.
Wash Buffer (Immunoprecipitation)	125 mM NaCl, 50 mM Tris (pH 7.5) or 50 mM PIPES pH 7, 1% (v/v) RNasin Plus ribonuclease inhibitor

Table 2.4. Solutions and buffers.

All solutions and buffers were prepared in deionised water unless otherwise stated.

2.2. General methods

2.2.1. RNA purification and quantification

RNA for RT-qPCR, Northern blotting, and 4sU quantification was extracted from cells and purified using an RNeasy kit with an on-column DNase digestion step, as per the manufacturer's instructions. RNA for other experiments was extracted using TRIzol reagent and chloroform. Cells were lysed directly on plates by adding TRIzol, scraping the plates with a pipette tip, transferring the liquid to a tube, and leaving to stand for 5 min at RT. 20% (v/v) chloroform was added to the samples before shaking to mix and leaving to stand for 3 min at RT. Samples were centrifuged at 12,000 x *g* for 15 min at 4°C. The top aqueous layer was transferred to a new tube (0.1 mM DTT was added if RNA contained 4sU) and 50% (v/v) isopropanol was added before mixing and leaving to stand at RT for 10 min. Samples were centrifuged at 12,000 x *g* for 20 min at 4°C and the supernatant was removed. 1 ml of 75% ethanol was added (0.1 mM DTT was added if RNA contained 4sU) before vortexing and centrifuging at 7,500 x *g* for 5 min at 4°C. Further centrifugation was performed if the pellet floated. The supernatant was removed and the pellet was left to air dry. RNA was dissolved in RNase-free ddH₂O (1 mM DTT was added if RNA contained 4sU) by heating the pellet at 60°C for 10 min. Samples were treated with TURBO DNase according to the manufacturer's instructions, followed by another round of 50% (v/v) isopropanol precipitation. Samples were kept at -70°C between further experiments.

Input and elute samples from SAF-A immunoprecipitation experiments were treated with TURBO DNase I at 37°C for 30 min to digest DNA then with proteinase K for 1 hr at 55°C to digest proteins. RNA was then extracted using an equal volume of phenol:chloroform (with isoamyl alcohol, 25:24:1) or acid-phenol:chloroform (with IAA, 125:24:1) and vortexing to mix. Samples were centrifuged at 12,000 x *g* for 15

min at 4°C. The top aqueous layer was transferred to a new tube for isopropanol precipitation. Samples were analysed on an agarose gel or TapeStation RNA High Sensitivity ScreenTape.

RNA was quantified using a Nanodrop 8000 by measuring sample optical density (OD) at 260 nm or by using a Qubit Flex Fluorometer RNA Broad Range or High Sensitivity assay kit, as per the manufacturer's instructions.

2.2.2. DNA purification and quantification

Genomic DNA from mammalian cell cultures grown in a 6 or 10-cm dish was extracted using a lysis buffer (150 mM NaCl, 1% SDS, 10 mM EDTA) supplemented with 2.5% (v/v) RNase A/T1 cocktail and incubated for 30 min at 37°C. The lysed cells were then incubated with 100 µg/ml proteinase K for 1 hr at 55°C. Samples were put through a syringe twice before adding an equal volume of phenol:chloroform (with isoamyl alcohol, 25:24:1) and vortexing to mix. Samples were centrifuged at maximum speed for 10 min at RT. The top aqueous layer was transferred to a new tube for a repeat addition of an equal volume of phenol:chloroform, then centrifuged as before. The same was repeated with an equal volume of chloroform. To the final supernatant, 2.5x 100% ethanol was added with 10% sodium acetate and was mixed by vortexing. Samples were placed on dry ice for 20 min and then centrifuged at maximum speed for 10 min at 4°C. The supernatant was removed and 500 µl of 70% ethanol was used to wash the pellet. Samples were centrifuged at maximum speed for 1 min at 4°C and the supernatant was removed. The pellet was left to air dry at RT and dissolved in TE buffer (Table 2.4.) then stored at -20°C.

Cells grown in a 96-well plate were lysed with 50 µl lysis buffer (100 mM Tris-HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl) supplemented with 100 µg/ml proteinase K. The plates were left at 37°C overnight before removing the lysis buffer into a 0.5 ml Eppendorf tube. 50 µl of ice-cold isopropanol was added to each tube and they were left on a shaker for 10 min to mix. The tubes were centrifuged at maximum speed for 10 min and the supernatant was removed. The pellet was washed with 70% ethanol and samples were centrifuged at maximum speed for 1 min. The supernatant was removed, and pellets were left to air dry at RT. The pellets were dissolved in 50 µl TE buffer (Table 2.4.) and stored at 4°C before genotyping PCR amplification (see Section 2.2.4.).

For small-scale isolation of plasmid DNA from 5 ml overnight cultures, a Qiaprep Spin Miniprep kit or E.Z.N.A. plasmid mini kit I was used following the manufacturer's instructions. For medium-scale isolation of plasmid DNA from 50 ml overnight cultures, the ZymoPURE plasmid midiprep kit was used following the manufacturer's

instructions. Plasmid DNA was analysed using restriction enzyme digest (see Section 2.2.7.) and Sanger sequencing (see Section 2.2.10.).

DNA was quantified using a Nanodrop 8000 by measuring sample OD at 260 nm or by using a Qubit Flex Fluorometer dsDNA Broad Range or High Sensitivity assay kit, as per the manufacturer's instructions.

2.2.3. Bacterial transformation and culture

DH5 α competent cells were used for bacterial transformation of plasmids of interest or ligated constructs according to the manufacturer's instructions. Resulting cell cultures were spread over LB-agar plates supplemented with ampicillin, chloramphenicol, or kanamycin. Liquid cultures (5 or 50 ml) were prepared by inoculating LB-Broth (supplemented with antibiotics; Table 2.4.) with a bacterial colony picked from an LB-agar plate with a 20 μ l pipette tip. Cultures were incubated in a shaking incubator at 220 rpm at 37°C overnight. LB-agar plates were prepared by IGC Technical Services.

2.2.4. PCR amplification

PCR amplification from selected plasmids (Table 2.5.) was performed using Q5 Hot Start HF polymerase with 0.1 ng plasmid DNA in a 50 μ l reaction according to the manufacturer's instructions. To amplify samples the following cycling conditions were used in a Bio-Rad Thermal Cycler PTC-240: 98°C for 30 s; then 35 cycles of 98°C for 10 s, recommended T_m for 30 s, 72°C for 30 s; final elongation at 72°C for 2 min. The melting temperature (T_m) for each primer set was calculated using the NEB T_m calculator. PCR reactions were gel purified (see Section 2.2.6.).

Primer	Primer sequence
hnRNPA1 forward primer (with NotI site)	5'-TAAATGCGGCCGCGCATGTC TAAGTCAGAGTCTCCTAAAG-3'
hnRNPA1 reverse primer (with BamHI site)	5'-AGCCAGGATCCTTAAAAT CTTCTGCCACTGCCATAG-3'
hnRNPC1 forward primer (with NotI site)	5'-TAAATGCGGCCGCGCGGCC AGCAACGTTACCAACAAG-3'
hnRNPC1 reverse primer (with BamHI site)	5'-ATTAGGATCCTTAAG AGTCATCCTCGCCATTGG-3'
XRN2 forward primer (with FseI site)	5'-GCACCATGGCCGGCCAAATGGGA-3'

XRN2 reverse primer (with NotI site)	5'-CGCGGCCGCAAGGCTAT TCCAATTGTATCTTCCTG-3'
EGFP forward primer (with FseI site)	5'-CATGGCCGGCCTCGCCA CCATGGTGAGCAAGGGC-3'
EGFP reverse primer (with NotI site)	5'-CGCGGCCGCAAGGCTCT TGACAGCTCGTCCATGC-3'

Table 2.5. Primers for plasmid PCR amplification.

PCR amplification from purified genomic DNA (Table 2.6.) was performed using Q5 Hot Start HF polymerase with 50 ng DNA in a 50 µl reaction according to the manufacturer's instructions. To amplify samples the following cycling conditions were used in a Bio-Rad Thermal Cycler PTC-240: 98°C for 3 min; then 35 cycles of 98°C for 30 s, 72°C for 60 s; final elongation at 72°C for 4 min. PCR reactions were gel purified (see Section 2.2.6.).

Primer	Primer sequence
Forward primer (with HindIII site)	5'-TCGATAAGCTTGCCAAAC GAGTATAGCCACCAAAGACCC-3'
Reverse primer (with XbaI site)	5'-AATTAATCTAGAGGCT GTTGCTGCTGCGTCGCCGG-3'

Table 2.6. Primers for genomic DNA PCR amplification.

Inverse PCR amplification from plasmid DNA (Table 2.7.) was performed using Q5 Hot Start HF polymerase with 0.1 ng plasmid DNA in a 50 µl reaction according to the manufacturer's instructions. To amplify samples the following cycling conditions were used in a Bio-Rad Thermal Cycler PTC-240: 98°C for 3 min; then 35 cycles of 98°C for 30 s, 64°C for 45 s, 72°C for 160 s; final elongation at 72°C for 5 min. PCR reactions were gel purified (see Section 2.2.6.).

Primer	Primer sequence
Forward primer (with SpeI site)	5'-AATTAACTAGTAGTT CCTCGCCTGTTAATGT-3'
Reverse primer (with EcoRV site)	5'-TCGATGATATCTGTG AGTGCCCCGATTCACC-3'

Table 2.7. Primers for inverse PCR amplification.

Genotyping PCR amplification from genomic DNA (Table 2.8.) was performed using Q5 Hot Start HF polymerase with 10 µl genomic DNA from FLAG-mAID-SAF-A cell clones in a 50 µl reaction according to the manufacturer's instructions. To amplify samples the following cycling conditions were used in a Bio-Rad Thermal Cycler

PTC-240: 98°C for 3 min; then 35 cycles of 98°C for 30 s, 71°C for 45 s, 72°C for 60 s; final elongation at 72°C for 5 min. PCR reactions were PCR purified using the QIAquick PCR purification kit as per the manufacturer's instructions. Samples were eluted in 20 µl of DNase-free water and stored at -20°C.

Primer	Primer sequence
Internal forward primer	5'-ACCATGAAAAAGCCTGAACTCACCG-3'
Reverse primer	5'-GGCTGTTGCTGCTGCGTCGCCGG-3'

Table 2.8. Primers for genotyping PCR amplification.

2.2.5. Agarose gel electrophoresis

For RNA samples, gels with the required percentage of agarose were prepared in 1X formaldehyde gel buffer (Table 2.4.) and cooled before adding formaldehyde to 0.65% and left to polymerise. Gels were equilibrated into 1X formaldehyde running buffer (Table 2.4.) at RT before use. Samples were diluted in RNA gel loading buffer (Table 2.4.) and 0.5 µg/ml ethidium bromide (EtBr) then heated to 65°C for 5 min. They were chilled before being loaded into wells and run at 100 V in 1X formaldehyde running buffer for 2 hr at RT. A pump was used to circulate the buffer. Gels were imaged using a Fuji FLA-5100 laser scanner.

For DNA samples, gels with the required percentage of agarose were prepared using 1% agarose in 1X TBE buffer (for analysis; Table 2.4.) or 1X TAE buffer (for gel extraction; Table 2.4.) supplemented with 0.5 µg/ml EtBr. Gels were equilibrated in 1X of the respective buffer at RT before use. Samples were diluted in 1X purple gel loading dye and run together with a 1 kb Plus DNA ladder for 1 to 1.5 hr at 100 V at RT. Gels were imaged using a camera and Syngene UV lightbox.

2.2.6. DNA extraction from agarose gels

DNA fragments were separated by gel electrophoresis (see Section 2.2.5.) and imaged using UV light. Gel slices containing the expected DNA bands were excised with a scalpel and stored in a 2 ml Eppendorf tube on ice. DNA was extracted from the gel slices using the QIAquick gel extraction kit according to the manufacturer's instructions.

2.2.7. Restriction enzyme digestion

Plasmids and DNA constructs were digested with restriction enzymes according to NEBcloner recommendations for single or double enzyme reactions. In the case of AarI, the reaction was set up according to the manufacturer's instructions. Reactions were 20 to 50 µl, depending on DNA concentration. Restriction digestion products were analysed and/or purified by agarose gel electrophoresis.

2.2.8. Gibson Assembly

Gibson Assembly was performed using the Gibson Assembly master mix as per the manufacturer's instructions.

2.2.9. Ligation of DNA fragments

DNA ligation was performed with the T4 DNA Ligase kit as per the manufacturer's instructions. The insert and vector were ligated in a 3:1 molar ratio.

2.2.10. Sanger sequencing

Sanger DNA sequencing was performed by the IGC Technical Services using the appropriate primers. DNA sequencing data was analysed using SnapGene.

2.2.11. RT-qPCR

To reverse transcribe RNA into cDNA, the SuperScript III Reverse Transcriptase kit was used as per the manufacturer's instructions. qPCR for each sample was performed in a 10 µl volume in triplicate on a 96-well plate. Each well contained a cDNA sample, SYBR Green I Master Mix, and 200 nM primers (Table 2.9.). To amplify samples the following cycling conditions were used in a LightCycler 480: 95°C for 2 min; then 45 cycles of 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s.

Primer	Primer sequence
ACTG1 forward	5'-GCCTCCAGGTTTCTCATTT-3'
ACTG1 reverse	5'-GGGTGTTGAAGGTCTCAAA-3'
EEF1A1 forward	5'-CGTCATTGGACACGTAGATT-3'
EEF1A1 reverse	5'-ACCTAATGATTCTGCTGGTAAA-3'

Table 2.9. Primers for qPCR.

2.3. Cell culture

2.3.1. Cell lines

hTERT RPE-1 – Human telomerase reverse transcriptase (hTERT) immortalised human retinal pigment epithelial cells. A near-diploid cell line of female origin with one derivative X-chromosome with additional chromosomal material at the end of the q-arm. The other chromosomes are karyotypically normal.

hTERT RPE-1 CDK1_{as} - hTERT RPE-1 cell line with an F80G mutation in the cyclin-dependent kinase 1 (CDK1) gene to confer bulky ATP analogue sensitivity (as; Saldivar *et al.*, 2018). Addition of the ATP analogue 4-amino-1-tert-butyl-3-(1'-naphthylmethyl) pyrazolo[3,4-d]pyrimidine (1NM-PP1) inactivates CDK1 in a reversible manner (Hochegger *et al.*, 2007). Cells and 1NM-PP1 were provided by Kumiko Samejima (post-doctoral researcher in the Earnshaw lab).

HCT 116 – Human colorectal carcinoma cell line of male origin with a mutation in codon 13 of the *RAS* proto-oncogene. A near-diploid cell line with karyotypic abnormalities such as monosomic chromosomes 10 and 18.

HCT 116 Tet-OsTIR1 – HCT 116 cell line that contains the Tet-inducible (Tet) Transport Inhibitor Response 1 (TIR1) protein from *Oryza Sativa* (Os; Natsume *et al.*, 2016; Yesbolatova *et al.*, 2019). A HCT 116 Tet-OsTIR1 cell line used in this thesis has a mini auxin-inducible degron (mAID) tag integrated at the N-terminus of the SAF-A gene. The addition of auxin or synthetic indole-3-acetic acid (IAA) targets SAF-A-AID for degradation by the proteasome.

HCT 116 TIR1 - HCT 116 cell line that constitutively expressed OsTIR1 protein (Heath, 2018). Cells provided by Stuart Wilson.

HNRNPU-AID (SAF-A-AID) – HCT 116 cell line that constitutively expressed OsTIR1 protein. In addition, the SAF-A gene has a C-terminal integration of the auxin-inducible degron (AID) tag (Heath, 2018). The addition of auxin or IAA targets SAF-A-AID for degradation by the proteasome. Cells provided by Stuart Wilson.

HEK-293 – Human embryonic kidney cell line with epithelial morphology. The cell line is hypotriploid with karyotypic abnormalities such as four copies of chromosomes 17 and 22.

FlpIn 293 T-REx – HEK-293 cell line used to make Tet-inducible stable cell lines. Contains one Flp Recombination Target (FRT) site to allow integration of a gene construct of choice. The addition of the antibiotic doxycycline (Dox) expresses the integrated gene product. FlpIn 293 T-REx cell lines used in this thesis include expression of SNAP-SAF-A, SNAP-HNRNPC1, SNAP-HNRNPA1, FLAG-SAF-A, FLAG-SAF-A P487S, FLAG-SAF-A R714S, FLAG-SAF-A Y805*, FLAG-SNAP-SAF-A P487S, FLAG-SNAP-SAF-A R714S, FLAG-SNAP-SAF-A Y805*. SNAP-SAF-A and FLAG-SAF-A expressing FlpIn 293 T-REx cells were made by Ryu-Suke Nozawa (former post-doctoral researcher in the Gilbert lab).

2.3.2. Cell passage

RPE-1 and RPE-1 CDK1_{as} cells were cultured in Advanced DMEM/F12 media containing phenol-red, supplemented with 10% foetal calf serum, 1% L-glutamine, 0.3% (w/v) sodium bicarbonate, and 1% pen/strep. HCT 116 cells and cell lines, FlpIn 293 T-REx cells and cell lines, and HEK-293 cells were cultured in Dulbecco's Modified Eagle Medium containing phenol-red, supplemented with 10% foetal calf serum and 1% pen/strep. FlpIn 293 T-REx cell lines were maintained in the presence of 100 µg/ml hygromycin and 15 µg/ml blasticidin antibiotics. HCT 116 TIR1 cells were maintained in the presence of 10 µg/ml blasticidin, while SAF-A-AID cells required the addition of 150 µg/ml hygromycin, 800 µg/ml neomycin, and 1 µg/ml puromycin. Cells were maintained in T75 flasks at 37°C with 5% CO₂ in a cell culture incubator. To passage, cells were washed with PBS, detached with trypsin and re-seeded in a new flask at a 1:8 to 1:10 dilution. Cells were typically passaged every 3 days.

2.3.3. Freezing and thawing cells

For freezing, an 80% confluent flask of cell culture was trypsinised, washed with PBS, and centrifuged at 300 x *g* for 4 min to form a cell pellet. The pellet was resuspended in 2 ml of freezing media (Table 2.4.). Subsequently, the cell suspension was divided into two cryotubes, with each containing 1 ml of suspension, and frozen at -80°C for storage for less than 3 months. For long-term storage, the cells were transferred to liquid nitrogen.

For thawing, frozen cell lines were thawed rapidly by placing the cryotubes at 37°C. The thawed cells were then resuspended in 15 ml of growth medium and transferred to T75 flasks. The flasks containing the resuspended cells were incubated at 37°C with 5% CO₂ in a cell culture incubator overnight. Subsequently, the medium was removed, and the cell culture was washed with PBS. Fresh media was added to the flask to ensure the removal of any residual DMSO. Recovered cells were kept in an incubator for at least one week and passaged as required before using them for experiments.

2.3.4. Experiment preparation

Experiments were carried out by plating cells in their relevant media for 24 hr prior to other steps. For HEK-293 and FlpIn 293 T-REx cells and cell lines, plates or slides were first coated with 0.01% poly-L-lysine solution (MW 150-300 kDa) by leaving the solution for 10 min, aspirating it, washing with PBS thrice and leaving the plates/slides to dry. Poly-L-lysine increases the electrostatic interaction between negatively charged ions of the cell membrane and the positively charged sites on the plastic plate/glass slide surface thus promoting cell adhesion. Experiments involving dSTORM microscopy and live cell imaging used phenol-red free DMEM (supplemented with 2 mM L-glutamine, 5% double charcoal-stripped FCS, 1% pen/strep) to remove phenol-red autofluorescence at 440 nm. IAA was added to cells expressing mAID-SAF-A or SAF-A-AID at a final concentration of 500 µM for 2 to 24 hr. Cells expressing FKBP12^{F36V}-tagged proteins were incubated with 100 nM to 2 µM dTAG-13 for 24 hr.

2.3.5. DNA transfection

Cells were plated the day before use to give a 60% confluency (for imaging experiments) or 80% confluency (for other experiments) at the time of transfection. Lipofectamine 3000 diluted in Opti-MEM reduced serum media was used to transfect DNA (Table 2.10.) according to the manufacturer's instructions.

Plasmid	Origin
pCMV-FlpIn-3xFLAG-mAID-SAF-A-P2A-OsTIR1 (877)	From pCMV-FlpIn-mAID-3xFLAG-SAF-A-P2A-OsTIR1 (made by Ryu-Suke Nozawa; digested with PshAI and AarI) and 3xFLAG-mAID fragment assembled using Gibson Assembly
pCMV-dTAG-3xFLAG-SAF-A (880)	From pCMV-FlpIn-mAID-3xFLAG-SAF-A-P2A-OsTIR1 (made by Ryu-Suke Nozawa; digested with AfeI and PmeI to remove P2A-OsTIR1 then with SbfI-HF and KpnI-HF) and dTAG N-term sequence assembled using Gibson Assembly
pCMV-SAF-A-dTAG-3xFLAG (882)	From pCMV-FlpIn-3xFLAG-mAID-SAF-A-P2A-OsTIR1 (digested with SbfI-HF and KpnI-HF to remove 3xFLAG-mAID then with AfeI and PmeI to remove P2A-OsTIR1) and dTAG C-term sequence assembled using Gibson Assembly
pCMV-EGFP-HELB (886)	From pCMV-EGFP-C1 (digested with HindIII-HF and KpnI-HF) and HELB sequence (digested with HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-mCherry-HELB (887)	From pCMV-mCherry-C1 (digested with HindIII-HF and KpnI-HF) and HELB sequence (digested with HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-FlpIn-3xFLAG-SNAP-SAF-A (758)	Made by Ryu-Suke Nozawa (former post-doctoral researcher in the Gilbert lab)
pCMV-FlpIn-3xFLAG-SAF-A (722)	Made by Ryu-Suke Nozawa
pCMV-FlpIn-3xFLAG-SAF-A-1-666aa(deltaRGG) (768)	Made by Ryu-Suke Nozawa
pCMV-FlpIn-3xFLAG-SAF-A-K491A(WA) (732)	Made by Ryu-Suke Nozawa
pCMV-FlpIn-3xFLAG-SNAP-hnRNPA1 (891)	From hnRNPA1 sequence PCR amplified from the pET-A1 plasmid (from Caceres lab) and pCMV-FlpIn-3xFLAG-SNAP-SAF-A (digested with NotI-HF and BamHI-HF) ligated with T4 ligase
pCMV-FlpIn-3xFLAG-SNAP-hnRNPC1 (893)	From hnRNPC1 sequence PCR amplified from the pEZ-M03-hnRNPC1-eGFP plasmid and pCMV-FlpIn-3xFLAG-SNAP-SAF-A (digested with NotI-HF and BamHI-HF) ligated with T4 ligase
pCMV-XRN2-dTAG-3xFLAG (894)	From pCMV-SAF-A-dTAG-3xFLAG (digested with FseI and NotI-HF) and XRN2 sequence PCR amplified from pCMV-FlpIn-XRN2-mAID (made by Ryu-Suke Nozawa) ligated with T4 ligase
pCMV-EGFP-dTAG-3xFLAG (895)	From pCMV-SAF-A-dTAG-3xFLAG (digested with FseI and NotI-HF) and EGFP sequence PCR amplified from pCMV-EGFP-C1 ligated with T4 ligase
pCMV-EGFP-hMD (899)	From pCMV-EGFP-C1 (digested with HindIII-HF and KpnI-HF) and hMD sequence (digested with HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-EGFP-mMD (900)	From pCMV-EGFP-C1 (digested with HindIII-HF and KpnI-HF) and mMD sequence (digested with

	HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-mCherry-hMD (901)	From pCMV-mCherry-C1 (digested with HindIII-HF and KpnI-HF) and hMD sequence (digested with HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-mCherry-mMD (902)	From pCMV-mCherry-C1 (digested with HindIII-HF and KpnI-HF) and mMD sequence (digested with HindIII-HF and KpnI-HF) ligated with T4 ligase
pCMV-FlpIn-3xFLAG-SNAP-SAF-A P487S (930)	From pCMV-FlpIn-3xFLAG-SNAP-SAF-A (made by Ryu-Suke Nozawa; digested with AclI-v2 and BstXI) and SAF-A 423-556 aa fragment with c.1459CtoT assembled using Gibson Assembly
pCMV-FlpIn-3xFLAG-SNAP-SAF-A R714S (931)	From pCMV-FlpIn-3xFLAG-SNAP-SAF-A (made by Ryu-Suke Nozawa; digested with XcmI and Sbf-HF) and SAF-A 677-806 aa fragment with c.2199GtoT assembled using Gibson Assembly
pCMV-FlpIn-3xFLAG-SNAP-SAF-A Y805* (932)	From pCMV-FlpIn-3xFLAG-SNAP-SAF-A (made by Ryu-Suke Nozawa; digested with XcmI and Sbf-HF) and SAF-A 677-806 aa fragment with c.2414_2415delinsGA assembled using Gibson Assembly
pCMV-FlpIn-3xFLAG-SAF-A P487S (933)	From pCMV-FlpIn-3xFLAG-SAF-A (made by Ryu-Suke Nozawa; digested with AclI-v2 and BstXI) and SAF-A 423-556 aa fragment with c.1459CtoT assembled using Gibson Assembly
pCMV-FlpIn-3xFLAG-SAF-A R714S (934)	From pCMV-FlpIn-3xFLAG-SAF-A (made by Ryu-Suke Nozawa; digested with XcmI and Sbf-HF) and SAF-A 677-806 aa fragment with c.2199GtoT assembled using Gibson Assembly
pCMV-FlpIn-3xFLAG-SAF-A Y805* (935)	From pCMV-FlpIn-3xFLAG-SAF-A (made by Ryu-Suke Nozawa; digested with XcmI and Sbf-HF) and SAF-A 677-806 aa fragment with c.2414_2415delinsGA assembled using Gibson Assembly

Table 2.10. Plasmids for mammalian transfection. Vector maps can be found in Appendix A and DNA constructs used in Gibson Assembly can be found in Appendix B.

2.3.6. RNA interference

To deplete proteins from cells using RNA interference (RNAi), siRNA oligos were transfected using the RNAiMAX transfection reagent. 10 nM of siRNA were added to 600 µl of Opti-MEM reduced serum media and mixed by inversion. 16 µl or 36 µl of RNAiMAX were added to 600 µl of Opti-MEM reduced serum media and mixed by inversion. The two mixtures were mixed and incubated for 20 min at RT. Of the final mixture, 300 µl was added to 6-well plates or 600 µl to 4-well plates and cells were incubated for 24 to 72 hr, depending on the experiment. Protein depletion was

visualised using protein immunofluorescence (see Section 2.5.1.) or Western blotting (see Section 2.4.3.).

2.3.7. Growth assay

The proliferation of mammalian cell populations was measured using the PrestoBlue™ Cell Viability Reagent, a cell-permeable resazurin-based solution. Metabolically active cells reduce the resazurin (blue, nonfluorescent) to resorufin (red, highly fluorescent). Resorufin has an excitation peak of 569 and an emission peak of 586. Cells to be assayed were grown in black clear-bottomed plates to reduce background and autofluorescence. The cell viability reagent was used according to the manufacturer's instructions. Fluorescence measurements were taken between the 544 and 590 nm wavelengths in a BMG LABTECH FLUOstar® Omega microplate reader. Each condition had 4 to 12 technical replicates which were plotted as box plots using R. Statistical significance between different conditions was tested using the Wilcoxon test. * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

2.3.8. CRISPR Cas9 stable cell line generation

To generate HCT116 Tet-OsTIR1 cells expressing 3xFLAG-mAID-SAF-A, homology-directed repair (HDR) with a donor plasmid was used. To do this, the CRISPR Cas9 protein was directed to the start codon of endogenous SAF-A by guide RNAs (gRNA) targeting within 2 bp of the ATG site (Table 2.11.). To generate the donor plasmid (Table 2.12.), a KIF4 knock-in construct was digested with the restriction enzymes SpeI and BglII with the larger fragment gel purified then assembled using Gibson Assembly to remove the mAID-HaloTag and replace with 3xFLAG-mAID (Appendixes A and B). The resulting plasmid was then digested with EcoRV and SpeI to gel purify the HygR-P2A-3xFLAG-mAID fragment for later insertion into the donor plasmid. The KIF4 knock-in construct was digested with HindIII-HF and XbaI with the smaller fragment gel purified and ligated with PCR amplified SAF-A 5' UTR and exon 1 (Appendix A; gel purified then digested with HindIII-HF and XbaI then PCR purified) then transformed into E. coli for plasmid DNA generation. The construct containing SAF-A 5' UTR and exon 1 was used for inverse PCR amplification to mutate the protospacer adjacent motif and digested with EcoRV and SpeI before ligation with the HygR-P2A-3xFLAG-mAID fragment to create the SAF-A knock-in construct. The donor plasmid was linearised with XbaI and HindIII before transfection. To make the gRNA constructs (Table 2.12.), the phosphorylated oligos (Table 2.11.) were annealed by mixing 100 μ M of oligo 1 and 2 and annealing at 95°C for 5 min then a

step-wise reduction of 0.5-1°C per second until 25°C. The PX459_V2 plasmid was digested with BbsI according to the manufacturer's instructions and gel purified. The digested PX459_V2 plasmid was ligated with the annealed gRNA constructs with the addition of BbsI in the ligation mix.

gRNA	Target Site	Oligo 1	Oligo 2
gRNA 1	ATTAACAGGC GAGGAACTCA TGG	5'-/5Phos/ CACC G ATTAACAGG CGAGGAACTCA-3'	3'-C TAATTGTC CGCTCCTTGAGT CAAA /5Phos/-5'
gRNA 2	CAGGCGAGGA ACTCATGGTG AGG	5'-/5Phos/ CACC G CAGGCGAG GAACTCATGGTG-3'	3'-C GTCCGCT CCTTGAGTACCAC CAAA /5Phos/-5'
gRNA 3	AGGCGAGGAA CTCATGGTGA GGG	5'-/5Phos/ CACC G AGGCGAGG AACTCATGGTGA-3'	3'-C TCCGCTC CTTGAGTACCACT CAAA /5Phos/-5'

Table 2.11. Target sites and gRNA sequences to target CRISPR Cas9 to the ATG codon of SAF-A. Overhands for cloning are bolded.

Plasmid	Origin
PX-459-V2	From Sproul lab
PX-459-V2-gRNA1 (869)	From PX_459_V2 digested with BbsI and ligated with gRNA 1
PX-459-V2-gRNA2 (870)	From PX_459_V2 digested with BbsI and ligated with gRNA 2
PX-459-V2-gRNA3 (871)	From PX_459_V2 digested with BbsI and ligated with gRNA 3
SAF-A-knock-in-HygR-3xFLAG-mAID (872)	From KIF4 knock-in mAID-Halo-tag construct (provided by Kumiko Samejima; Appendix A) digested as described above

Table 2.12. Plasmids for cell line generation using CRISPR Cas9.

HCT116 Tet-OsTIR1 cells were transfected with 500 ng linearised donor plasmid and 150 ng of each gRNA plasmid together with Lipofectamine 3000 (see Section 2.3.5.). A GFP-containing donor plasmid was used as a positive control. After a day, cells were trypsinised and diluted 1:100, 1:500, and 1:1000 times into 10 cm dishes. More 10 cm dishes were plated with the least dense looking plate dilution. After a day, 100 µg/ml of the selection antibiotic hygromycin was added to cell plates. Cell media supplemented with antibiotics was changed every 3-4 days until colony formation. Visible colonies (~1 mm) were transferred to flat-bottomed 96-well plates using a 20 µl pipette tip filled with 10 µl of trypsin. To transfer the colonies, they were scraped with the pipette tip and were trypsinised. 100 µl of media supplemented with 100 µg/ml hygromycin was added to each well and the trypsinised colonies were resuspended with light pipetting. After confluency, wells were split into a 96-well plate

(for genomic DNA extraction, see Section 2.2.2.) and a 24-well plate. Once the 24-well plate was confluent, the cells were split into a 12-well plate. This was repeated for a 6-well plate and finally a T25 flask from where cells were frozen and kept at -70°C until PCR amplification confirmed the successful integration of the HygR-3xFLAG-mAID-P2A construct (see Section 2.2.4.). Clones with successful integration were thawed and grown in 6-well plates in the presence of 500 µM IAA and 0.2 µg/ml Dox to confirm FLAG-mAID-SAF-A expression by Western blotting (see Section 2.4.3.).

2.3.9. Stable cell line generation using the FlpIn system

FlpIn 293 T-REx cell lines expressing a protein of interest were generated with pOG44 and FlpIn plasmids (Table 2.10. – 891, 893, 930 to 935) following the manufacturer's instructions.

2.4. Protein analysis

2.4.1. Protein lysate preparation

Mammalian cells were typically grown in 6-well plates until reaching 80% confluency. Whole-cell nuclear extracts from cells were prepared using 1x NuPage LDS sample buffer with 12.5 mM DTT. A volume of 1 ml of LDS buffer was used for every 1×10^7 cells. The cell lysates were then scraped into microfuge tubes using a cut-off P1000 tip and incubated at 70°C for 10 min. Subsequently, the samples were sonicated using the Soniprep 150 probe sonicator at ~3.5 µm for 30 s and stored at -20°C until further analysis.

2.4.2. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were separated on NuPAGE 4-12% (typically) Bis-Tris, homemade 15% polyacrylamide (for detection of histone proteins) or NuPAGE 3-8% Tris-Acetate (for detection of SAF-A oligomers) mini gels. Protein extracts were thawed, spun at maximum speed for 30 s to remove debris, and then mixed with 1x LDS sample buffer before loading onto the gel. 3 to 10 µl of SeeBlue Plus2 Pre-Stained Protein

Standard were also loaded onto the gel to estimate protein size. NuPAGE Novex mini gel system tanks were filled with 1x NuPAGE MOPS SDS Running buffer for Bis-Tris and polyacrylamide gels and electrophoresis was performed at 150 V for 60 min. For Tris-Acetate gels, 1x NuPAGE Tris-Acetate SDS Running Buffer was used, and electrophoresis was performed at 150 V for 180 min.

15% polyacrylamide gels (Table 2.4.) were made by firstly mixing everything for the main gel in order with TEMED added last and the main gel mix was transferred to a gel frame. 70% ethanol was added on top of the gel to smoothen the top. Then the stacking gel (Table 2.4.) was prepared by mixing everything in order, the same as the main gel. Desired gel combs were added, and bubbles were cleared. The gel was left to harden at room temperature before removing the gel combs and using the gel immediately for electrophoresis.

2.4.3. Western blotting

Whole-cell nuclear extracts were separated by size using SDS-PAGE, following the procedure described in section 2.4.2. After electrophoresis, the proteins on the gels were transferred onto Immobilon-P polyvinylidene fluoride (PVDF) membranes through an electrophoretic transfer at 100 V for 90 min at 4°C, using a transfer buffer (Table 2.4.). Prior to the transfer, the PVDF membrane was briefly washed in methanol and rinsed in the transfer buffer. Following the transfer, the PVDF membrane was washed three times for 5 min each in 1x TBS-T buffer (Table 2.4.) and then blocked in 5% (w/v) Marvel (original dried semi-skimmed milk powder) or 5% BSA (for anti-phosphate residue antibodies) in 1x TBS-T for 10 min at RT. The membrane was then incubated with primary antibodies (Table 2.13.) diluted in the blocking solution for 2 hr at RT or overnight at 4°C. After incubation, the membrane was washed three times for 5 min each in TBS-T to rinse excess antibodies and then probed with secondary antibodies (Table 2.13.) conjugated to horseradish peroxidase (HRP) in the blocking solution for 2 hr at RT or overnight at 4°C. The HRP bound to the membrane was detected using the Pierce™ enhanced chemiluminescence (ECL) Western Blotting Substrate according to the manufacturer's instructions. The membrane was either exposed to photographic film and developed using a Konika SRX-101A developer or using an Invitrogen iBright™ CL1500 Imaging System scanner without photographic film. Band intensity was quantified using the ImageJ software's Label Peaks function and plotted using Microsoft Excel as raw values or normalised using GAPDH.

Antibody	Concentration	Secondary	Concentration
Anti-FLAG	1:10,000	Anti-mouse HRP	1:5,000
Anti-GAPDH	1:5,000	Anti-rabbit HRP	1:5,000
Anti-H3S10ph	1:5,000	Anti-rabbit HRP	1:5,000
Anti-hnRNP U, clone 3G6	1:10,000	Anti-mouse HRP	1:5,000
Anti-HNRNPUL1	1:2,000	Anti-rabbit HRP	1:5,000
Anti-SAF-A S59ph	1:3,000	Anti-rabbit HRP	1:5,000

Table 2.13. Western blotting antibodies.

2.5. Microscopy techniques

2.5.1. Protein immunofluorescence

Human cells were typically grown on 22x22 mm coverslips in 6-well plates containing 2 ml media until reaching 60% confluency. Before fixation, cells expressing SNAP-tagged proteins were treated with a SNAP-Cell fluorescent substrate (Table 2.16.) for 30 min. Cells were washed twice with PBS and fixed with 4% PFA for 10 min at RT. They were washed again with PBS and permeabilised in 0.2% Triton X-100. Cells were washed twice with PBS and either stained in 50 µg/ml DAPI immediately before mounting (if no antibodies were used) or blocked (if antibodies were used). Blocking was done with 5% BSA in PBS for 10 min at RT. Primary antibodies were added onto the coverslips at the desired concentration (Table 2.14.) in PBS and incubated for 4 hr at RT in a dark, humidified chamber. Primary antibody was removed by washing coverslips twice in PBS at RT. Secondary antibody was added onto the coverslips at the desired concentration (Table 2.14.) in PBS and incubated for 2 hr at RT in the dark. Secondary antibody was removed by washing coverslips twice in PBS at RT. 50 µg/ml DAPI was used to detect DNA and nuclei by staining cells for 10 min at RT. Coverslips were drained and mounted in Vectashield antifade mounting medium on top of Eprelia™ SuperFrost Plus™ Adhesion slides and sealed with clear nail varnish. Fluorescence images were acquired either on a Zeiss Axiolmager M2 fluorescence microscope or a Zeiss Axioplan II Epifluorescence microscope using a 63x or 100x objective. For fluorescent images acquired on a Zeiss Axiolmager M2 fluorescence microscope, the following were used: Plan-Apochromat objectives (Carl Zeiss, Cambridge, UK), Photometrics Prime BSI CMOS camera (Photometrics Ltd, Tuscon AZ), a Zeiss Colibri 7 LED light source, together with Zeiss filter sets 90 HE, 92 HE, 96 HE, 38 HE and 43 HE (Carl Zeiss UK, Cambridge, UK). Image capture was performed in Zeiss Zen 3.5 software. For epifluorescent images acquired on a Zeiss Axioplan II fluorescence microscope, the following were used: Photometrics Coolsnap HQ2 CCD camera (Teledyne Photometrics, Tuscon, AZ), Plan-Neofluar/apochromat objective lenses (Carl Zeiss, Cambridge, UK), a Mercury Halide

fluorescent light source (Exfo Excite 120, Excelitas Technologies) and Chroma #83000 triple band pass filter set (Chroma Technology Corp., Rockingham, VT) with the single excitation and emission filters installed in motorised filter wheels (Prior Scientific Instruments, Cambridge, UK). Image capture was performed using Micromanager (Version 1.4).

Antibody	Concentration	Secondary	Concentration
Anti-FLAG	1:1000	Donkey AF488 anti-mouse	1:500
Polyclonal anti-hnRNP U	1:200	Donkey AF647 anti-rabbit	1:500

Table 2.14. Protein immunofluorescence antibodies.

2.5.2. Click chemistry

Human cells were typically grown on 22x22 mm coverslips in 6-well plates containing 2 ml media until reaching 60% confluency. Before fixation, cells were incubated with the uridine analogue 5-ethynyl uridine (5-EU, see Section 2.10.2.). Cells were washed twice with PBS and fixed with 4% PFA for 10 min at RT. They were washed again with PBS and permeabilised in 0.2% Triton X-100. Cells were washed twice with PBS and either stained in 50 µg/ml DAPI immediately before mounting (if no antibodies were used) or blocked (if antibodies were used). Blocking was done with 5% BSA in PBS for 10 min at RT. Coverslips were incubated while covered with click reaction mix (50 µl per coverslip) in a dark humidified chamber for 30 min at RT. The click reaction mix was prepared in 250 µl by mixing 222 µl PBS, 2.5 µl 200 mM CuSO₄, 1 µl fluorescently labelled azide and 25 µl 200 mg/ml L-ascorbic acid and allows the addition of the azide group onto the alkyne group of 5-EU. Coverslips were washed in PBS twice to remove unused azide groups. 50 µg/ml DAPI was used to detect DNA and nuclei by staining cells for 10 min at RT. Coverslips were drained and mounted in Vectashield antifade mounting medium on top of Eprepia™ SuperFrost Plus™ Adhesion slides and sealed with clear nail varnish. Fluorescence images were acquired either on a Zeiss Axiomager M2 fluorescence microscope or a Zeiss Axioplan II Epifluorescence microscope using a 63x or 100x objective, as in sub-section 2.5.1.

2.5.3. dSTORM microscopy

RPE-1 cells were typically grown in 8-well Lab-Tek II Chambered Coverglass (No. 1.5 borosilicate glass bottom) until reaching 80% confluency. For transcription inhibition during interphase, 1 μ M triptolide was added 4 hr before fixing. Cells were washed twice with PBS and fixed with 4% PFA for 10 min at RT. They were washed thrice with PBS and permeabilised in 0.2% Triton X-100. Cells were washed thrice with PBS and blocked with 1% BSA in PBS for 10 min at RT. Primary antibodies were added onto the cells at the desired concentration (Table 2.15.) in PBS and incubated for 2 hr at RT in a dark, humidified chamber. Primary antibody was removed by washing the cells thrice in PBS at RT. Secondary antibody was added onto the cells at the desired concentration (Table 2.15.) in PBS and incubated for 1 hr at RT in the dark. The secondary antibodies were made by F(ab')₂ fragment donkey anti-rabbit or anti-mouse IgG conjugated to a fluorophore (CF568 or CF647) at a ratio of 1:1. Secondary antibody was removed by washing cells thrice in PBS at RT. For samples incubated with 5-EU (see Section 2.10.2.), click chemistry was performed (see Section 2.5.2.) by incubating wells with 50 μ l of click reaction mix containing Cy5 picolyl azide. Oxygen scavenger imaging buffer (Table 2.4.) was used to fill the wells supplemented with 50 μ g/ml DAPI to stain nuclei. A cover glass and nail varnish were used to seal the 8-well plate to prevent oxygen inflow. Images were taken on the same day after sealing the 8-well plate.

Antibody	Concentration	Secondary	Concentration
Polyclonal anti-hnRNP U	1:100	Donkey CF568 anti-rabbit	1:100
Polyclonal anti-hnRNP U	1:100	Donkey CF647 anti-rabbit	1:100
Anti-hnRNP C1/C2	1:1000	Donkey CF647 anti-mouse	1:100
Anti-SC-35	1:500	Donkey CF647 anti-mouse	1:100

Table 2.15. dSTORM antibodies.

dSTORM images were acquired on a Nikon Eclipse Ti2 inverted microscope (Nikon UK Ltd, Kingston Upon Thames, UK) with laser total internal reflection fluorescence system (TIRF) using a SR HP Apo TIRF 100x 1.49NA objective lens and an additional 1.5x optical zoom. Laser light is provided via a Nikon LU-NV laser bed with 405, 488, 561, 640 nm laser lines. Widefield illumination is provided by a CoolLED PE-300 LED light source (CoolLED Ltd, Andover, UK). Images were acquired with a Hamamatsu Orca-Fusion CMOS camera (325-6, Sunayama-cho, Naka-ku, Hamamatsu City, Shizuoka Pref., 430-8587, Japan) using Nikon NIS-Elements

Advanced Research software. Image capture was performed as a stack of 20,000 frames at 19-ms exposure time using the NIS-Elements software. Post-processing analysis and SuperStructure analysis were performed as in Marena *et al.*, 2021. Each condition had 6 to 15 independent nuclei which were plotted as box plots using R or Python. Statistical significance between different conditions was tested using the Wilcoxon test. * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

2.5.4. Live cell imaging

Human cells were typically grown in 35-mm glass bottom dishes containing 2 ml media until reaching 60% confluency. A day before imaging, cells were transfected with plasmids containing the desired protein of interest (see Section 2.3.5.) and treated with 1 $\mu\text{g/ml}$ to express the transfected protein or FRT site integrated protein. If cells underwent synchronisation prior to imaging, 10 μM RO-3306 was added 24 hours before imaging. For transcription inhibition during mitosis, 1 μM triptolide was added 1 hr before imaging. L-ascorbic acid (100 or 500 μM) was added as a protective agent against photodamage an hour before imaging. To visualise DNA, cells were stained with either 2 mg/ml Hoechst 33342 for 30 min, 1 or 2 μM SiR-DNA for 60 min, or 5 μM DRAQ5 for 30 min.

Before imaging, cells expressing SNAP-tagged proteins were treated with a SNAP-Cell fluorescent substrate (Table 2.16.) for 30 min. Cells were washed thrice with PBS and pre-warmed cell media was re-added to the dishes. The pre-warmed cell media contained 1 μM SiR-DNA, 100 μM L-ascorbic acid and/or 1 μM triptolide if they were used before washing. Images of live cells undergoing mitosis were acquired every minute over 1 hr on a Nikon Eclipse Ti-E inverted microscope (Nikon Instruments UK Ltd, Kingston Upon Thames, UK) with a Photometrics Prime BSI camera sCMOS (Teledyne Photometrics, 3440 E. Britannia Drive, Suite 100, Tucson, AZ 85706; HDR conversion gain, 16-bit, 1x1 binning, variable exposure time per experiment) using a 63x or 100x oil objective and a Semrock BrightLine LED-DA/FI/TR/Cy5-4X4M-B (Sedat) laser filter set (IDEX Health & Science, LLC, Center of Excellence, 1180 John Street, Rochester, New York, 14586, United States; variable laser power per experiment). Widefield illumination was provided by a CoolLED PE-4000 LED light source (CoolLED Ltd, Andover, UK). The microscope had an Okolab bold line stage top incubation chamber (Okolab S.R.L, Ottaviano, NA, Italy), that was pre-warmed, together with objectives and filled with 5% CO_2 before imaging. Image capture was performed in NIS-Elements software (Nikon UK Ltd, Kingston-on-Thames, UK).

Substrate	Concentration (IF)	Concentration (Live cell)
OregonGreen	1:2000	1:500
SiR647	1:2000	1:500
TMR-Star	1:2000	1:500

Table 2.16. SNAP-tag fluorescent substrates.

Images were analysed in ImageJ. Pearson correlation coefficient was calculated using the Coloc2 plugin with a macro written by Laura Murphy (Appendix C). Statistical significance between mitotic stages was tested using the Wilcoxon signed-rank test for paired samples. * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

2.6. SAF-A oligomer crosslinking

Cells used for SAF-A immunoprecipitation and RNA purification were PBS washed and crosslinked using 800 mJ of UV radiation at 254 nm to crosslink RNA and proteins. Cells were then detached with trypsin and washed twice in PBS before counting the cell number using cell counting chamber slides in the Countess Automated Cell Counter according to the manufacturer's instructions. Cells were divided into tubes containing 1×10^6 cells and permeabilised with 100 μ l oligomerisation assay buffer (Table 2.4.) including Tris (for BM(PEG)₂) or PIPES (for DSS) for 10 min at 37°C. Crosslinkers (0.1 mM, 0.3 mM, and 1 mM for Western blotting, 1 mM only for IP) were added for 5 min at RT. Cells were then processed for Western blotting (see Section 2.4.3.) or SAF-A immunoprecipitation (see Section 2.7.).

2.7. SAF-A immunoprecipitation

Magnetic beads coated in anti-FLAG M2 antibody were prepared by washing aliquots of beads thrice in PBS, then incubating with oligomerisation assay buffer (Table 2.4.) and antibody overnight at 4°C. Cells were crosslinked as in section 2.8. Proteins were extracted using the Soniprep 150 probe sonicator at $\sim 2 \mu$ m for 10 cycles for 30 s on/30 s off and centrifuged at maximum speed for 10 min. 10% of the supernatant was collected as an input sample for further processing for Western blotting or RNA isolation. Anti-FLAG M2 antibody-coated magnetic beads were washed twice with oligomerisation assay buffer (or wash buffer for RNA isolation; Table 2.4.) before adding the protein sample and incubating overnight at 4°C. The beads were washed twice again with oligomerisation assay buffer (or wash buffer for RNA isolation) to

produce elute samples before processing for Western blotting or RNA isolation. Samples for RNA isolation were eluted with wash buffer containing 100 µg/ml 3xFLAG peptide by incubating the samples 5x3 min with rotation before removing the supernatant at a magnetic rack.

2.8. Cell synchronisation

2.8.1. Synchronising cells at the G₁/S boundary and S phase

HCT 116 and RPE-1 CDK1_{as} cells were synchronised at the G₁/S boundary by a double thymidine block (2xThy; Table 2.17.). A high concentration of thymidine in the cell media inhibits DNA synthesis by blocking the reduction of cytidine diphosphate to deoxycytidine diphosphate (Bostock *et al.*, 1971). Cells were plated in 6 or 10-cm dishes at a density of 35,000 (RPE-1 CDK1_{as}) or 100,000 cells/ml (HCT 116) and cultured overnight. Media containing 2 mM or 2.5 mM thymidine was then added to the cells for 18 to 20 hr to block the cell cycle during the S phase. Cells were washed twice in PBS and released in normal growth media for 8 hr to allow cells to progress through the S phase. 2 mM or 2.5 mM thymidine was added to the cell media for a further 14 to 16 hr to block the cell cycle and retain cells at the G₁/S boundary. After the final block, cells were washed in PBS twice and released in pre-warmed media. Flow cytometry to assess the cell cycle progress of synchronised cells was performed on samples cell populations collected every 1 or 2 hr up to 13 hr following release (see Section 2.8.3.).

RPE-1 cells were synchronised during the early S phase by the addition of aphidicolin, a DNA polymerase α inhibitor that blocks replication (Pedrali-Noy *et al.*, 1980; Table 2.17.). Cells were plated in 6 cm dishes at a density of 50,000 cells/ml and cultured overnight. Media containing 15 µM aphidicolin was added to the cells for 24 hr to block the cell cycle during the S phase. Cells were washed twice in PBS and released in pre-warmed media. Flow cytometry to assess the cell cycle progress of synchronised cells was performed on samples cell populations collected every 2 hr up to 18 hr following release (see Section 2.8.3.).

Cell line	Primary block	Secondary block	Tertiary block	Arrest phase
HCT 116	2 mM thymidine for 18 hr and 8 hr release	2 mM thymidine for 14 hr	N/A	End of G ₁
HCT 116	2 mM thymidine for 24 hr and 3 hr release	10 µM RO-3306 for 12 hr	N/A	End of G ₂
HCT 116	2 mM thymidine for 24 hr and 3 hr release	25 ng/ml nocodazole for 12 hr	N/A	Pro-metaphase
RPE-1	15 µM aphidicolin for 24 hr	N/A	N/A	Early S phase
RPE-1 CDK1 _{as}	2.5 mM thymidine for 20 hr and 8 hr release	2.5 mM thymidine for 16 hr	N/A	End of G ₁
RPE-1 CDK1 _{as}	2.5 mM thymidine for 20 hr and 8 hr release	2.5 mM thymidine for 16 hr and 7 hr release	5 µM 1NM-PP1 for 5 hr	End of G ₂

Table 2.17. Synchronisation methods.

2.8.2. Synchronising cells at the G₂/M boundary and M phase

HCT 116 cells were synchronised after a single thymidine block and RO-3306 addition at the G₂/M boundary or a thymidine block and nocodazole addition during pro-metaphase (Table 2.17.). RO-3306 selectively inhibits the activity of CDK1 and entry to mitosis while nocodazole inhibits microtubule formation to arrest cells at pro-metaphase before chromosome line up on the metaphase plate (Brabander *et al.*, 1986; Vassilev, 2006). Cells were plated in 6 cm dishes at a density of 100,000 cells/ml and cultured overnight before adding media containing 2 mM thymidine for 24 hr to block the cell cycle during the S phase. Cells were washed twice in PBS and released in normal growth media for 3 hr before the addition of 10 µM RO-3306 or 25 ng/ml nocodazole for 12 hr to retain cells at the G₂/M boundary or pro-metaphase. After the final block, cells were washed in PBS thrice and released in pre-warmed media. Flow cytometry to assess the cell cycle progress of synchronised cells was performed on samples cell populations collected every 2 hr up to 6 hr following release (see Section 2.8.3.).

RPE-1 CDK1_{as} cells were synchronised at the G₂/M boundary after a 2xThy block and 1NM-PP1 addition (Table 2.17.). Cells were plated and treated as in section 2.8.1 with the addition of a 7 hr release after the second thymidine block. They were then incubated with 5 µM 1NM-PP1, a bulky ATP analogue that inhibits analogue sensitive CDK1, to retain cells at the G₂/M boundary. After the final block, cells were washed in PBS thrice and released in pre-warmed media. Flow cytometry to assess the cell cycle progress of synchronised cells was performed on samples cell populations collected every 30 min up to 1.5 hr following release (see Section 2.8.3.).

2.8.3. Cell cycle assessment with propidium iodide staining

Staining with propidium iodide (PI) was used to assess the cell cycle distribution of cell populations after synchronisation (see Section 2.8.1. and 2.8.2.). PI is a DNA intercalating fluorescent dye that has a fluorescence emission maximum of 617 nm when bound to nucleic acids. Cells were trypsinised and centrifuged at 300 x *g* for 4 min. They were washed in PBS and centrifuged again before resuspension in 70% ethanol and PBS mix to fix and permeabilise the cells at a concentration of 1×10^6 cells/ml. The 70% ethanol and PBS mix was pipetted slowly while the cell pellet was gently vortexed to avoid rupturing the cells and ensure a uniform fixation. Fixed cells were stored at 4°C for up to a month.

Before cell cycle analysis with PI, the cells were centrifuged at 300 x *g* for 4 min and washed twice in PBS. The cell pellets were resuspended at a concentration of 1×10^7 cells/ml in PBS containing 5 µg/ml PI and 1 µg/ml RNase A for 30 min at RT in a dark chamber. Before cell cycle analysis with PI and H3 Ser10 phosphorylation (H3S10ph) antibodies, the cells were centrifuged at 300 x *g* for 4 min and washed twice in PBS. Cells were blocked in 1% BSA for 10 min with rotation. After a PBS wash, cells were incubated with an AF 488 conjugated antibody for H3S10ph at a 1:100 concentration in PBS for 45 min with rotation in a dark chamber. Cells were washed with PBS and pelleted then resuspended at a concentration of 1×10^7 cells/ml in PBS containing 1% BSA, 5 µg/ml PI and 1 µg/ml RNase A for 30 min at RT in a dark chamber. Cell cycle assessment was performed on an LSR Fortessa analyser by measuring the signals in the 695/740 channel for PI and the 488/525 channel for H3S10ph/AF 488 with assistance from Elizabeth Freyer (Flow Cytometry Facility Manager).

2.9. Fluorescence in-situ hybridisation (FISH)

2.9.1. Fosmid probe details

Fosmid probes (Table 2.18.) were purchased from BacPac resources and the probe position was previously verified on metaphase chromosomes. Gene density was previously calculated from Ensembl 87.

Position/ name	Clone name	Sequence name	Position (hg38)	Properties
11p15.5	CI 11p15-49	n/a	chr11:736369	gene rich
11p15.5	H19-IGF2	n/a	chr11:2070000	gene rich
11p15.1	WI2-2224K21	G248P87004F11	chr11:17,488,939- 17,528,052	gene rich
11p15.1	WI2-3168N23	G248P8017G12	chr11:19,034,283- 19,072,361	gene rich
11p14.1	WI2-906E12	G248P8190C6	chr11:29,684,660- 29,731,106	gene poor
11p14.1	WI2-1673H3	G248P83109D2	chr11:31,195,720- 31,235,826	gene poor

Table 2.18. Fosmid probes for FISH.

2.9.2. Purification of fosmid DNA

To prepare fosmid DNA, a bacterial stab from a glycerol stock was inoculated onto an LB-agar plate to isolate single colonies. Using a 20 µl pipette tip, a single bacterial colony was picked and inoculated into 5 ml LB-broth supplemented with 0.1% (v/v) chloramphenicol and incubated overnight at 37°C with shaking (220 rpm). The bacterial culture was centrifuged (spin out rotor) at 5000 x g for 10 min at 4°C and the supernatant was removed. The pellet was resuspended in 200 µl resuspension buffer (50 mM D(+)glucose, 25 mM Tris pH 8, 10 mM EDTA) with 60 µl 5 mg/ml lysozyme and 5 µl RNase A/T1 cocktail then incubated at RT for 5 min. 400 µl of lysis buffer (200 mM NaOH, 1% SDS) was added and samples were mixed and left on ice for 5 min until samples were clear. 300 µl of acetate buffer (3 M potassium acetate, 11.5% glacial acetic acid) was added and samples were mixed then vortexed to break up the flocculent and left on ice for 5 min. Samples were centrifuged at maximum speed for 5 min at 4°C and the supernatant was transferred to a new tube. An equal volume of phenol:chloroform (with isoamyl alcohol, 25:24:1) was added, and samples were mixed and centrifuged at maximum speed for 4 min. The aqueous layer was transferred to a new tube and an equal volume of chloroform was added. Samples were mixed and centrifuged at maximum speed for 4 min at 4°C. The aqueous layer was transferred to a new tube and an equal volume of isopropanol was added. Samples were placed on dry ice for 30 min and then were centrifuged at maximum speed for 12 min at 4°C. The supernatant was removed and the pellet was washed in 500 µl of 70% ethanol then centrifuged at maximum speed for 5 min at 4°C. The supernatant was removed and the pellet was left to air dry. The pellet was resuspended in 30 µl TE buffer (Table 2.4.).

2.9.3. Nick translation of FISH probes

Fosmid probes were nick translated by mixing 6 µl of purified fosmid DNA (see above) with 2.5 µl each of 0.5 mM dATP, 0.5 mM dCTP, and 0.5 mM dGTP, and 2.5 µl of labelled dUTP (AF 594-5-dUTP or Green 500 dUTP), 1 µl of 1:10 DNase I, 1 µl of DNA polymerase I in 1X nick translation salts (50 mM Tris pH 7.5, 10 mM MgSO₄, 0.1 mM DTT, 50 µg/ml BSA) to a total volume of 20 µl. The reaction mix was incubated at 16°C for 90 min and quenched by the addition of 2 µl 20% SDS and 3 µl of 500 mM EDTA. Volume was brought up to 90 µl with TE buffer (Table 2.4.) and passed through a G50 Sephadex spin column according to the manufacturer's instructions to remove unincorporated nucleotides. Probes were protected from light and stored at 4°C before hybridisation.

2.9.4. Preparation of interphase cells on slides

Cells for FISH experiments were grown on slides in 4-well dishes. After other experimental treatments, cells were washed thrice in PBS and fixed in 4% PFA for 10 min at RT. They were washed thrice with PBS again and permeabilised in 0.2% Triton X-100 (made in PBS) for 10 min at RT. Then they were washed thrice with PBS again before leaving in PBS overnight at 4°C, air drying and transfer to -70°C.

2.9.5. Hybridisation of FISH probes

Slides were taken out of -70°C and briefly thawed before washing in PBS once and then in 2X SSC (Table 2.4.) once. They were incubated with 100 µg/ml RNase A in 2X SSC for 60 min at 37°C. Slides were washed with 2X SSC and dehydrated through an ethanol series and air-dried. They were warmed in a 70°C oven for 5 min before denaturing in 70% formamide in 2X SSC (pH 7.5) for 30 min (HCT116 and FlpIn 293 T-REx cells) or 55 min (RPE-1 cells) at 80°C. They were then dehydrated through an ethanol series and warmed at 37°C on a hot plate for probe addition. 80 to 100 ng of each probe pair was mixed with 10 µg C₀t1 DNA and 5 µg sonicated salmon sperm DNA in 2X the volume of 100% ethanol. The mix was centrifuged in a vacuum centrifuge for 45 min with heat to remove the liquid and resuspended in 12 µl of hybridisation mix (6 µl deionised formamide, 1.2 µl 20X SSC, 0.1 µl Tween-20, 1.6 µl dextran sulphate, and H₂O) for 30 min at RT. Probes were denatured for 5 min at 80°C, re-annealed for 15 min at 37°C and snap cooled on ice. The probes were added to coverslips and put on the warmed slides, sealed with Pang supersolution, then incubated in a 37°C overnight in the dark. Slides were washed for 5 min in 4X

SSC with 0.1% Tween-20 and 0.5 µg/ml DAPI at 37°C. Excess liquid was drained and slides were mounted with coverslips using Vectashield antifade mounting medium and sealed with clear nail varnish.

2.9.6. Imaging and analysis of FISH probes

Fluorescence images from Chapter 4 were acquired using a Photometrics Coolsnap HQ2 CCD camera (Photometrics Ltd, Tuscon, AZ) and a Zeiss AxioImager A1 fluorescence microscope with a Plan Apochromat 100x 1.4NA objective, a Nikon Intensilight Mercury based light source (Nikon UK Ltd, Kingston-on-Thames, UK) and a Chroma #89014ET (3 colour) single excitation and emission filter (Chroma Technology Corp., Rockingham, VT) with the excitation and emission filters installed in Sutter motorised filter wheels (Sutter Instrument, Novato, CA). A piezoelectrically driven objective mount (PIFOC model P-721, Physik Instruments GmbH & Co, Karlsruhe) was used to control movement in the z dimension. Hardware control, image capture and analysis were performed using Nikon NIS-Elements software (Nikon UK Ltd, Kingston-on-Thames, UK). Images were deconvolved using a calculated point-spread function (PSF) in Volocity version 6.2.1. (Quorum Technologies, Puslinch, Canada) and fast deconvolution. Images were analysed in Imaris version 9.6.0. using the Spots feature to identify fluorescent probe centroids. The distance between probe pairs was measured using the Measurement tool. Statistical significance between mitotic stages was tested using the Wilcoxon signed-rank test for paired samples. * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

Fluorescence images from Chapter 5 were acquired using a Teledyne Photometrics Prime BSI sCMOS camera (Teledyne Photometrics Ltd, Tuscon, AZ) and a Zeiss AxioImager Z2 fluorescence microscope with a Plan Apochromat 100x 1.46NA, Plan Apochromat 100x 1.4NA objective lens. Fluorescence illumination was provided by an Excelitas Xylis LED light source (Excelitas Technologies Corp. 200 West Street, 4th Floor East, Waltham, MA 02451 United States). Fluorescence filters were single-band cubes (Carl Zeiss UK, Cambridge, UK). The excitation and emission filters are installed in Sutter motorised filter wheels (Sutter Instrument, Novato, CA). Image capture and deconvolution were performed using Zeiss Zen 3.5 acquisition software. Images were deconvolved using a calculated PSF and automatic deconvolution in Zeiss Zen 3.5 Acquisition software (Carl Zeiss UK, Cambridge, UK). They were then analysed in Imaris version 9.6.0. using the Spots feature to identify fluorescent probe centroids. The distance between probe pairs was measured using the Measurement tool. Statistical significance between mitotic stages was tested using the Wilcoxon signed-rank test for paired samples. * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

2.10. Nucleoside incorporation and quantification

2.10.1. Radiolabelling with [5-³H]-uridine

Cells were incubated with 20 µCi/ml [5-³H]U for 15 min or overnight before washing off radiolabel with three PBS washes and incubating cells at 37°C in pre-warmed media for a specific time depending on the experiment (0 to 3 hr). Samples with actD treatment had 50 ng/ml actD added to inhibit rRNA production 1 hr prior to radiolabel incubation. For Northern blotting, RNA was collected using an RNeasy kit while after SAF-A immunoprecipitation RNA was collected using acid-phenol:chloroform extraction (see Section 2.2.1.).

2.10.1.1. Northern blot

Total RNA extracts were separated by size using agarose gel electrophoresis, following the procedure described in section 2.2.5. After electrophoresis, the RNA on the gels were transferred onto positively charged Hybond-XL nylon membranes through upward capillary transfer. Prior to the transfer, the nylon membrane and gel were briefly washed in 10X SSC. For upward capillary transfer, a support was put in a vessel and topped with a wick of blotting paper (Whatman 3MM paper). The gel was placed on top of the paper, followed by the nylon membrane and a stack of paper towels. A weight of ~500 g was placed on top of the stack. Transfer buffer (10X SSC) was added to the vessel so that the wick is touching the liquid but not the gel. The transfer was left overnight. Following the transfer, the RNA was crosslinked to the membrane by 70 mJ of UV radiation at 254 nm. The membrane was rinsed in 1X SSC and left to dry before incubation in a storage phosphor screen for ³H detection overnight. Screens were imaged using a Fuji FLA-5100 laser scanner.

2.10.1.2. Liquid scintillation counting

To measure the specific activity of RNA radioactively labelled with ³H, a low energy β emitter, 5 µl of purified RNA samples was diluted in 4 ml of liquid scintillator. Negative control samples were made of 5 µl of H₂O diluted in 4 ml of liquid scintillator. Samples were counted twice on a Hidex 300 SL running Laura for PET software for 10 min to produce counts per minute (cpm) values. The specific activity of samples was calculated by dividing the average cpm values by the concentration of RNA using Microsoft Excel. Samples used in Figure 3.10 were counted by Jon Stocks (current PhD student in the Gilbert lab).

2.10.2. Metabolic labelling with 5-ethynyl uridine

RPE-1 cells were incubated with 1 mM 5-EU for 15 min before washing off with three PBS washes and incubating cells at 37°C in pre-warmed media for a specific time depending on the experiment (0 to 6 hr). Samples with actD treatment had 50 ng/ml actD added to inhibit rRNA production 1 hr prior to nucleoside incubation.

2.10.3. Metabolic labelling with 4-thiouridine

2.10.3.1. Incorporation and quantification

Cells were incubated with 100 µM or 1 mM 4-thiouridine (4sU) for 15 min before washing off with two PBS washes and incubating cells at 37°C in pre-warmed media for a specific time depending on the experiment (0 to 2 hr). Samples with actD treatment had 50 ng/ml actD added to inhibit rRNA production 1 hr prior to nucleoside incubation. 50 µl of purified RNA (see Section 2.2.1.) incubated with 4sU was separated by reverse phase HPLC on a Dionex Ultimate 3000 HPLC system by Bernie Ramsahoye (Clinical Haematologist and Molecular Biologist at the IGC). Nucleotides in RNA were detected at their peak absorbances – guanosine monophosphate at 254 nm and 4-thiouridine monophosphate at 323 nm. Nucleotide quantities were measured by the area under each peak at the respective peak UV absorbances.

2.10.3.2. Lysolecithin permeabilisation

Samples with lysolecithin (LL) treatment had the respective concentration of LL (diluted in chloroform or ethanol) added to briefly permeabilise cells for 2 min at 4°C or 37°C. LL treatment included 1 mM CaCl₂ to promote cell attachment. Cells were washed twice with PBS after LL treatment. PBS mixed with 0.1% Trypan Blue was added for 1 min to stain permeabilised cells. Cells were washed twice with PBS before adding pre-warmed media and imaging on a microscope immediately and after 20 min incubation at 37°C. Cells were imaged with a Bio-Rad ZOE™ Fluorescent Cell Imager microscope.

2.10.3.3. Thiol modification

Carboxyamidomethylation was performed on 10 to 20 µg of RNA from 4sU metabolic labelling experiments. RNA samples were incubated with 50% DMSO, 10 mM iodoacetamide, and 50 mM sodium phosphate buffer (pH 8; Table 2.4.) for 15 min at

50°C. The reaction was quenched by adding 20 mM DTT. Samples were then ethanol precipitated.

Oxidation and amino lysis were performed on 10 µg of RNA from 4sU metabolic labelling experiments. RNA samples were mixed with 1 mM EDTA (pH 8), 600 mM 2,2,2-trifluoroethylamine (TFEA), and 100 mM sodium acetate buffer (pH 5.2). Sodium periodate was added to the reactions at a final concentration of 10 mM and incubated for 60 min at 45°C with rotation. The reaction was quenched by adding 300 mM KCl and 300 mM sodium acetate (pH 5.2) and leaving samples at 4°C for 10 min. Samples were then ethanol precipitated.

2.10.3.4. One step RT-PCR

RT-PCR reactions were performed using the one step PrimeScript™ III RT-PCR kit according to the manufacturer's instructions. RT-PCR for each sample was performed in a 25 µl volume containing an RNA sample, PrimeScript™ III RT-PCR mix, and 200 nM primers (Tables 2.19. and 2.20.). To amplify samples the following cycling conditions were used in a Bio-Rad Thermal Cycler PTC-240 thermocycler: 52°C for 5 min, 95°C for 10 s; then 11 cycles of 95°C for 5 s, 64°C for 31 s. cDNA was separated by gel electrophoresis (see Section 2.2.5.) and imaged using UV light. Samples were PCR purified using the QIAquick PCR purification kit as per the manufacturer's instructions before DNA concentration was measured.

Primer	Primer sequence
ACTG1 exon-exon forward	5'-CGTGTTTCCTTCCATCGT-3'
ACTG1 exon-exon reverse	5'-GCCAGATCTTCTCCATGTC-3'
ACTG1 exon-intron forward	5'-GCCTCCAGGTTTCTCATTT-3'
ACTG1 exon-intron reverse	5'-GGGTGTTGAAGGTCTCAA-3'
ACTG1 intron-intron forward	5'-CCGGCTTGATTCTGACA-3'
ACTG1 intron-intron reverse	5'-CTCACACGCCACAACAT-3'
EEF1A1 exon-exon forward	5'-CGTCATTGGACACGTAGATT-3'
EEF1A1 exon-exon reverse	5'-TTCACGCTCAGCTTTCAG-3'
EEF1A1 exon-intron forward	5'-CGTCATTGGACACGTAGATT-3'
EEF1A1 exon-intron reverse	5'-ACCTAATGATTCTGCTGGTAAA-3'
EEF1A1 intron-intron forward	5'-CCTCGTGCTTGAGTTGAG-3'
EEF1A1 intron-intron reverse	5'-GATCTTGCCCCGCATTTA-3'
GAPDH exon-exon forward	5'-ACTCTGGTAAAGTGGATATTGT-3'
GAPDH exon-exon reverse	5'-GGAAGATGGTGATGGGATTT-3'
GAPDH exon-intron forward	5'-TCCACCCATGGCAAATTC-3'
GAPDH exon-intron reverse	5'-GGGACACAAGGTTACCATATAC-3'
GAPDH intron-intron forward	5'-AAGAGCTAGGAAGGACAGG-3'
GAPDH intron-intron reverse	5'-CTCAAAGGGCAGGAGTAAAG-3'
LDHA exon-exon forward	5'-TCAGCAAGAGGGAGAAAGCC-3'
LDHA exon-exon reverse	5'-CCAAGCCACGTAGGTCAAGA-3'

Table 2.19. Primers for pilot RT-PCR.

Primer	Primer sequence
ACTG1 exon-exon forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN NAGCCGTGTTTCCTTCCATCGT-3'
ACTG1 exon-exon reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN GTGCCAGATCTTCTCCATGTC -3
ACTG1 exon-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN NAAGCCTCCAGGTTTCTCATTT-3'
ACTG1 exon-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN CGGGGTGTTGAAGGTCTCAAA-3'
ACTG1 intron-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TGCCCGGCTTGATTTCTGACA-3'
ACTG1 intron-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN CATGCTCACACGCCACAACAT 3'
EEF1A1 exon-exon forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TCGTCATTGGACACGTAGATT-3'
EEF1A1 exon-exon reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN ACGTTACGCTCAGCTTTTCAG 3'
EEF1A1 exon-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TCGTCATTGGACACGTAGATT-3'
EEF1A1 exon-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN ACCTAATGATTCTGCTGGTAAA-3'
EEF1A1 intron-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TCGCCTCGTGCTTGAGTTGAG-3'
EEF1A1 intron-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN GCAGATCTTGCCCCGCATTTA-3'
GAPDH exon-exon forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN ACTCTGGTAAAGTGGATATTGT-3'
GAPDH exon-exon reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN TGGAAGATGGTGATGGGATTT-3'
GAPDH exon-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TGATTCCACCCATGGCAAATTC-3'
GAPDH exon-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN GGGACACAAGGTTACCATATAC-3'
GAPDH intron-intron forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN NAAAAGAGCTAGGAAGGACAGG-3'
GAPDH intron-intron reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN ACTCAAAGGGCAGGAGTAAAG-3'
LDHA exon-exon forward	5'-ACACTCTTTCCCTACACGACGCTCTTCCG ATCT NNNN TCAGCAAGAGGGAGAAAGCC-3'
LDHA exon-exon reverse	5'-AGACGTGTGCTCTTCCGATCT NNNN CCAAGCCACGTAGGTCAAGA-3'

Table 2.20. Barcoded primers for RT-PCR. Unique molecular identifiers in bold.

2.10.3.5. Library preparation and sequencing

Libraries were prepared for sequencing using 200 ng of cDNA and NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® according to the manufacturer's instructions. qPCR to estimate the amplification cycles required for each sample was performed in a 10 µl volume in duplicate on a 96-well plate. Each well contained a cDNA sample, 1X SYBR Green I, 1X NEB Next Ultra II Q5 PCR master mix, 1.25 µM P7 consensus index reverse primer for each condition, and 1.25 µM P5 consensus forward primer (Table 2.21.). To amplify samples the following cycling conditions were used in a LightCycler 480: 75°C for 5 min, 98°C for 30 s; then 20 cycles of 98°C for 10 s, 65°C for 75 s; finally, 65°C for 5 min. The cycle number at which fluorescence had reached 25% of its maximum was used to estimate the number of amplification cycles required. Samples were then amplified by the estimated number of amplification cycles in a 50 µl volume containing 1X NEB Next Ultra II Q5 PCR master mix, 1.25 µM P7 consensus index reverse primer for each condition, and 1.25 µM P5 consensus forward primer. Amplified DNA was PCR purified using the QIAquick PCR purification kit as per the manufacturer's instructions. Library size selection was performed using AMPure XP beads as per the manufacturer's instructions. Libraries were separated on a Tapestation D1000 High Sensitivity ScreenTape. Sequencing was performed on the NextSeq 550 platform using the NextSeq 500/550 Mid-Output v2.5 flow cell by the Genetics Core at the Edinburgh Clinical Research Facility.

Primer	Primer sequence
P5 consensus forward	5'-AATGATACGGCGACCACCGAGATCTACACT CTTTCCCTACACGACGCTCTTCCGATCT-3'
P7 consensus index 1 reverse	5'-CAAGCAGAAGACGGCATACGAGAT TTCCGTATAG GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 2 reverse	5'-CAAGCAGAAGACGGCATACGAGAT GAGGCTTAGA GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 3 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CTGTCCGTTT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 4 reverse	5'-CAAGCAGAAGACGGCATACGAGAT GTCTTGTAAG GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 5 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CCACCTTGGT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 6 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CTACAAGAGT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 7 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CCTTGTAAGT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 8 reverse	5'-CAAGCAGAAGACGGCATACGAGAT GTCTCAGTAG GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 9 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CTACAGTGAC GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 10 reverse	5'-CAAGCAGAAGACGGCATACGAGAT TGTACGCCTA GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 11	5'-CAAGCAGAAGACGGCATACGAGAT CTAGAGCAGT

reverse	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 12 reverse	5'-CAAGCAGAAGACGGCATACGAGAT TTCTGGTAGG GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 13 reverse	5'-CAAGCAGAAGACGGCATACGAGAT TTAATCCGGT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 14 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CTACTGTTGT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 15 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CCAGTGAGTC GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
P7 consensus index 16 reverse	5'-CAAGCAGAAGACGGCATACGAGAT CATGTAACAC GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'

Table 2.21. Indexed primers for PCR amplification. Index sequence in bold.

2.10.3.6. SLAM-seq analysis

Basecall data produced by the NextSeq 550 is automatically uploaded to BaseSpace, a cloud-based data management and analysis service provided by Illumina where it is converted into FASTQ files to allow analysis. The resulting sequences were aligned to the reference index of amplified regions using bowtie v2.3.5.1. The nucleotides at each base position were calculated using the function pileup2acgt from the Sequenza-utils package for Python or using the readr v.1.4.0 package for R. The respective number or percentage of base conversions were calculated for each condition and plotted as box plots or line graphs using R.

2.10.3.7. Nuclear fractionation

To isolate cytoplasmic and nuclear fractions from samples, cells were washed with PBS and trypsinised after 4sU incubation. Cells were centrifuged at 300 x g for 3 min to remove media and washed with PBS. They were centrifuged again to remove PBS and were resuspended in ice-cold hypotonic buffer (Table 2.4.) supplemented with 1% (v/v) RNase inhibitor by pipetting them up and down slowly. An equal volume of ice-cold isotonic buffer (Table 2.4.) supplemented with 0.1% NP-40 was added to the tubes and the samples were mixed by inversion. The tubes were incubated at 4°C for 3 min then they were centrifuged at 720 x g at 4°C for 3 min. The supernatant was collected as a cytoplasmic fraction. The resulting cell pellet was broken by tapping and was collected as a nuclear fraction.

2.11. SAF-A mutation analysis

2.11.1. Protein conservation analysis

A multiple sequence alignment (MSA) of the Oma group 1163611 of SAF-A orthologs in metazoan species was generated in CLUSTAL format using MAFFT FFT-NS-i v7.487. The v4 AlphaFold model of SAF-A was downloaded from UniProt and was used together with generated MSA as input to the ConSurf server to generate a protein surface conservation PyMOL script (Altenhoff *et al.*, 2020; Yariv *et al.*, 2023). The surface conservation script was run in PyMOL over the AlphaFold model of SAF-A.

2.11.2. 1D-to-3D method

To map mutational variants onto a SAF-A's 3D structure as spheres of increasing sizes and colour intensity to match mutational frequency, the 1D-to-3D program was used following instructions provided by the Cook lab (Deák and Cook, 2022).

2.11.3. Mutational density

To calculate the average mutational density in a protein domain (V_d) compared to the whole protein (V_p), the V_d/V_p calculator program was used following instructions provided by the Cook lab (Deák and Cook, 2022).

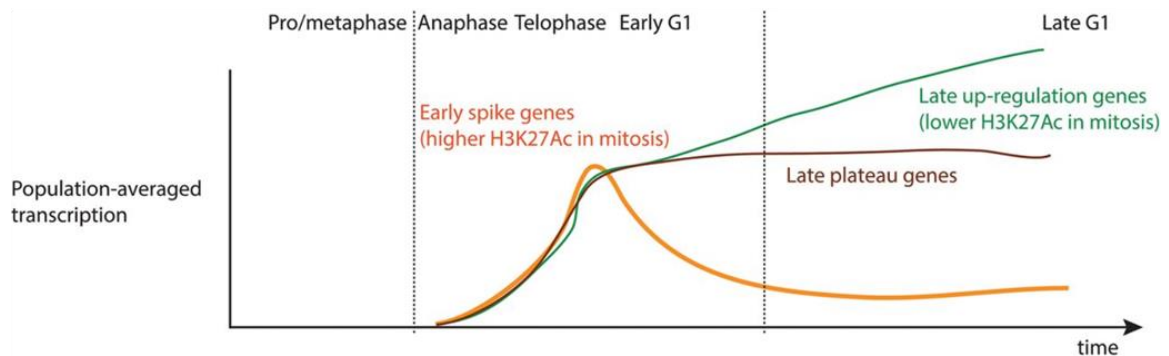
3. Chapter 3: Understanding SAF-A's role in chromatin decompaction at mitotic exit

3.1. Introduction

Recent research has put forward a role for SAF-A in organising large-scale chromatin structure during interphase by forming a mesh that is dependent on transcription and RNA availability (Nozawa *et al.*, 2017; Fan *et al.*, 2018). Further evidence shows that ncRNA is responsible for recruiting SAF-A to transcriptionally active chromatin (Creamer *et al.*, 2021; Skalska *et al.*, 2021). However, these experiments were performed using harsh treatments including high concentrations of transcription inhibitors over a long time to deplete the cells of RNAs. There is though a stage during the cell cycle when transcription is briefly stopped and resumes after a short period – mitosis. During this time SAF-A is excluded from mitotic chromatin, based on observations from imaging in fixed cells (unpublished lab data, Sharp *et al.*, 2020). As most SAF-A research has been undertaken during interphase, it is not known when it might start to form a mesh that can subsequently decompact active chromatin regions. Imaging cells as they transition through mitosis could therefore reveal how SAF-A is recruited to transcriptionally active regions. Current knowledge suggests that transcription restarts during telophase, wherein RNA pol II can bind to transcription start sites simultaneously with other transcription factors (Prasanth *et al.*, 2003). Additionally, RNA pol II interacts with proteins retained during mitosis such as the TATA-binding protein (TBP) resulting in a surge of transcriptional activity in specific genes like *Chchd3*, *Zeb2*, and *Runx1* in mice (Hsiung *et al.*, 2016; Figure 3.1A). Apart from this transcriptional burst, there are genes that reach a plateau in activity during early G₁, like *Asb1*, and genes that exhibit increasing levels of transcription throughout the G₁ phase, such as *Mavs*. This trend could be attributed to the re-establishment of chromatin loops facilitated by cohesin rebinding (Zhang *et al.*, 2019). It is then possible that transcription at this point during and after mitosis could be responsible for recruiting SAF-A to maintain a decompact chromatin state. In this Chapter, I set out to establish when SAF-A is recruited to chromatin and whether it occurs concomitantly with transcriptional re-activation in the daughter nuclei. To achieve this, I used a FlpIn 293 T-REx cell line expressing the SNAP-

tagged SAF-A gene inserted at the Flp Recombination Target (FRT) site under the control of a doxycycline (Dox) dependent promoter. After activation by drug treatment, single cells were imaged during mitosis.

A



B

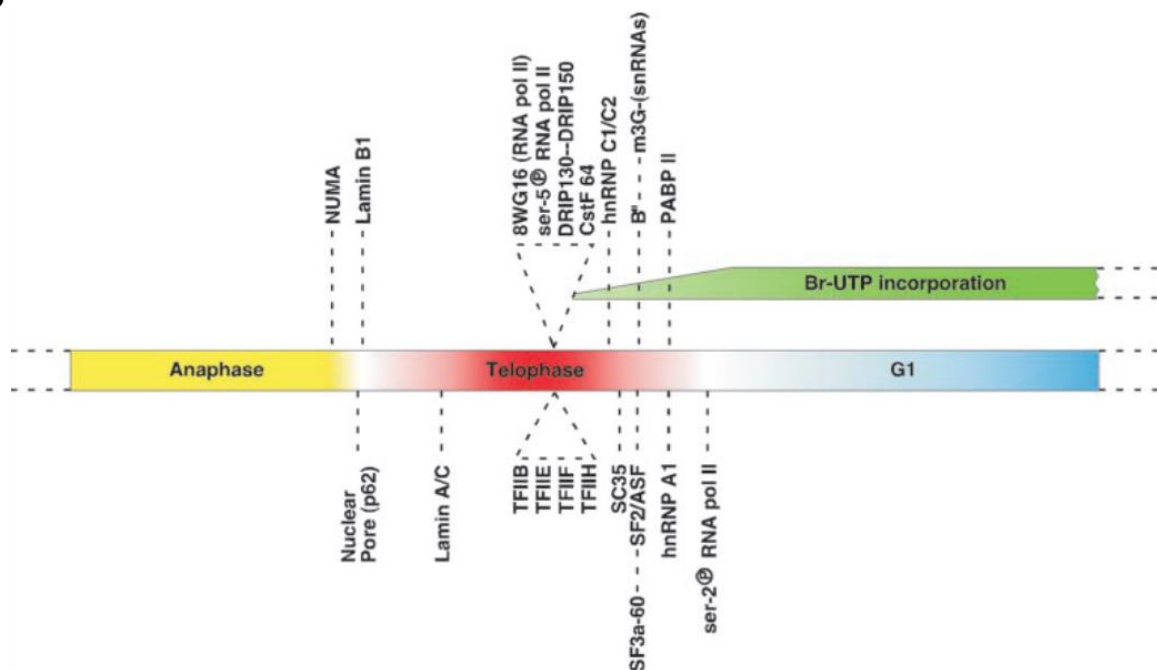


Figure 3.1. Re-establishment of gene expression during late mitosis and early G₁. A) A model of transcriptional re-activation from mitosis throughout the G₁ phase created from summarising the genome-wide transcriptional data from an averaged cell population. The figure is taken from Hsiung *et al.*, 2016. B) A schematic representing the order of re-incorporation of gene expression machinery into dividing cell chromatin and the daughter nuclei. The figure is taken from Prasanth *et al.*, 2003.

In addition to SAF-A, other proteins in the hnRNP family have different patterns of recruitment into the daughter nuclei and from this behaviour can be divided into early and late groups (Figure 3.1B). The early group of factors are found on chromatin together with RNA pol II and other transcription factors before nuclear envelope formation – an example is HNRNPC1/C2 (Pinol-Roma and Dreyfuss, 1991). In contrast, the late group is transported in after nuclear envelope formation and is dependent on transcription. This group includes HNRNPA1, HNRNPA2, HNRNPB1, HNRNPB2, HNRNPE, HNRNPH, and HNRNPL. When transcription is inhibited with a high dose of actinomycin D (actD) or 5,6-dichloro-1- β -D-ribofuranosylbenzimidazole (DRB), the HNRNPA1 of the late group is sequestered in the cytoplasm. Live cell imaging has confirmed that HNRNPC1 and HNRNPA1 are indeed early and late respectively in terms of chromatin recruitment (Prasanth *et al.*, 2003). Surprisingly though, recent research using FAIRE-MS showed that HNRNPA2 and HNRNPB1 are already enriched on mitotic chromatin, an observation at odds with previous findings (Ye *et al.*, 2023). Despite this, I used SNAP-tagged versions of HNRNPC1 and HNRNPA1 as controls to compare SAF-A's recruitment to chromatin. I also used brief treatment with triptolide, a transcriptional inhibitor, to perturb nascent RNA transcription to investigate whether transcription is required for SAF-A recruitment. While actD and DRB inhibit RNA elongation and change the phosphorylation of RNA pol II, triptolide inhibits transcription initiation by binding to the XRP subunit of RNA pol II, inhibiting its ATPase activity and preventing the formation of a transcription bubble (Bensaude, 2011). Therefore, triptolide inhibits transcription at an earlier step which coupled with fast action (typically less than an hour) and low concentration makes it an ideal compound for targeted transcription inhibition.

In this Chapter, I also investigated whether SAF-A tends to bind RNA as a monomer or oligomer during the cell cycle. To achieve this, I used the sulfhydryl-to-sulfhydryl crosslinker 1,8-bismaleimido-diethyleneglycol (BM(PEG)₂) to differentiate between oligomeric and monomeric SAF-A coupled with protein immunoprecipitation and RNA purification. After establishing a timeline for SAF-A recruitment to transcriptionally active chromatin using live cell imaging, I used a synchronised population of cells to investigate whether SAF-A is bound to RNA during mitosis. Sharp and colleagues hinted that this is the case with a proximity ligation assay in single cells (Sharp *et al.*, 2020). In contrast, I wanted to take a more global approach and analyse whether in

HCT 116 cells that synchronously go through mitosis if SAF-A can be immunoprecipitated with RNA labelled with [5-³H]-uridine ([5-³H]U). The incorporation of [5-³H]U was compared between HCT 116 and FlpIn 293 T-REx cells using scintillation counting, performed by Jon Stocks (current PhD student) on samples collected by me. The length of treatment of cells with [5-³H]U can also then be used to determine whether SAF-A interacts with nascent or total RNA during and after mitosis. For these experiments, HCT 116 cells were selected as they can be reliably synchronised with a double thymidine block, have excellent transfection efficiency, and are suitable for proteomic and RNA studies. Furthermore, treating cells with the sulfhydryl-to-sulfhydryl crosslinker 1,8-bismaleimido-diethyleneglycol (BM(PEG)₂) can also be used to discriminate between oligomeric or monomeric SAF-A interacting with RNAs (Nozawa *et al.*, 2017).

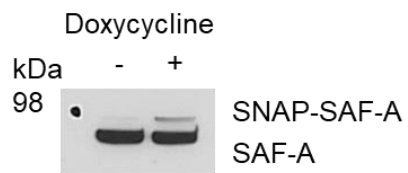
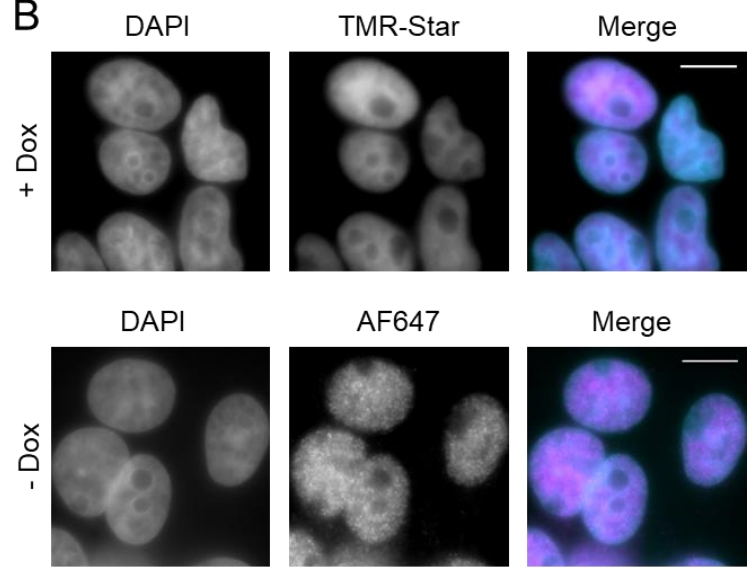
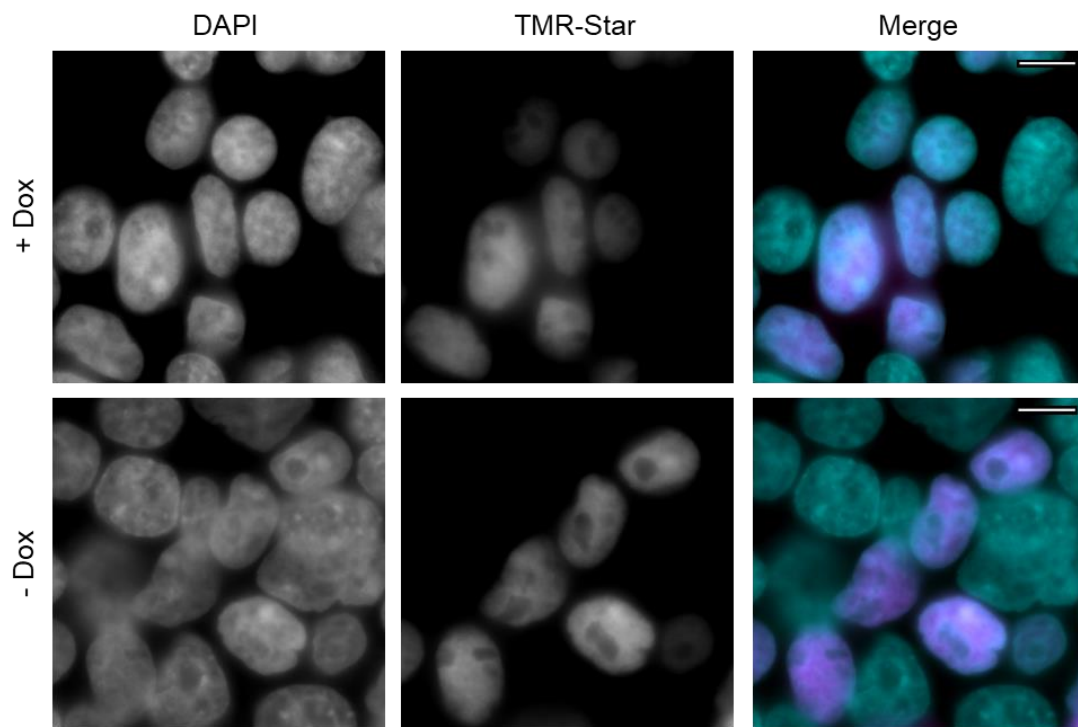
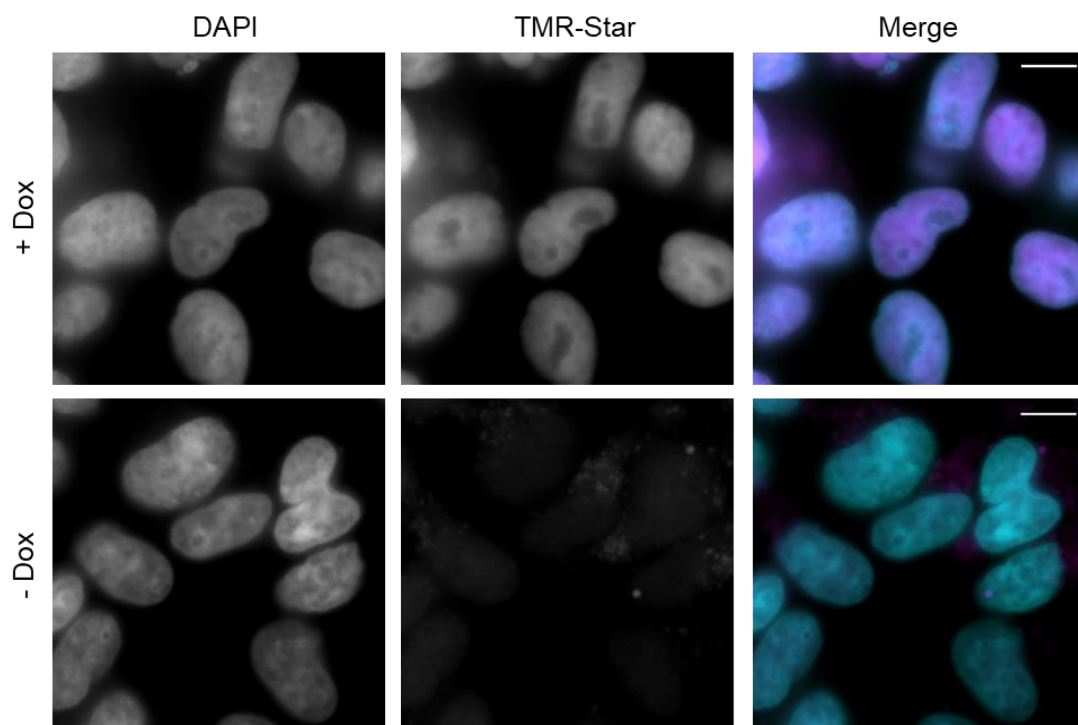
3.2. Live cell imaging of cells during mitosis using SNAP-tag

3.2.1. Optimising the cellular and imaging conditions

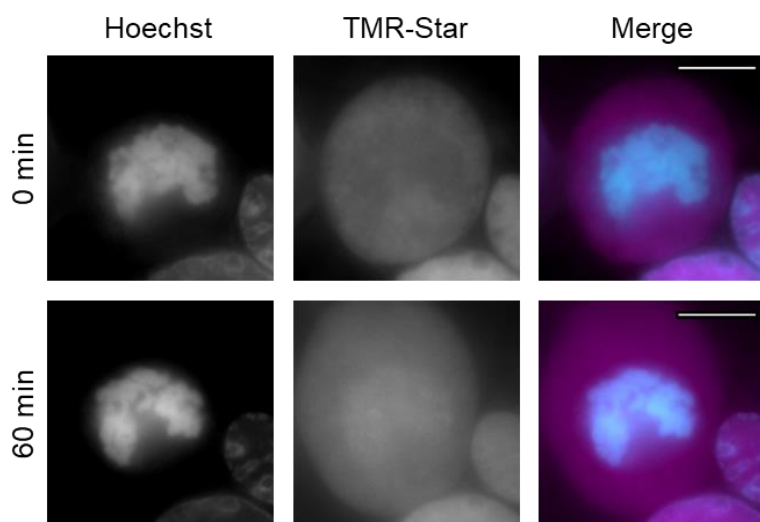
To image SAF-A during mitosis and at the mitotic exit I utilised a FlpIn 293 T-REx cell line available in the Gilbert lab with a live cell imaging construct where a SNAP-tag is attached to the N-terminus of SAF-A. The SNAP-tag, derived from the human O⁶-alkylguanine-DNA-alkyltransferase protein, with the addition of a membrane permeable chemically inert fluorescent probe, allows for the real-time imaging of a protein of interest when the cells are treated with the probe for 30 min before imaging (Mollwitz *et al.*, 2012). In Figure 3.2A, probing the lysate of cells treated with 1 µg/ml Dox for 24 hr with an anti-SAF-A antibody revealed the SNAP-SAF-A construct with an expected molecular weight of 109 kDa. A faint band was also present in the absence of Dox, but the predominant protein was wild-type SAF-A (WT; estimated to migrate at approximately 90 kDa). Immunofluorescence confirmed that the distribution of SNAP-SAF-A (visualised with a TMR-Star red probe) is comparable to that of the WT SAF-A (visualised with an anti-SAF-A primary antibody and an Alexa Fluor 647/AF 647 far-red secondary antibody) with no obvious effect of the slight overexpression of the protein (Figure 3.2B). When incubating FlpIn 293 T-REx cells

for 3 days with Dox, it became evident that the SNAP-SAF-A construct did not exhibit ubiquitous expression. Interestingly, even the untreated cells showed a positive signal for the SNAP-tagged protein (Figure 3.2C). Treatment with selection antibiotics hygromycin and blasticidin was then used to purify the cell line before further live cell imaging tests. This resulted in the ubiquitous expression of SNAP-SAF-A in cells treated with Dox for 3 days and no expression in untreated cells (Figure 3.2D).

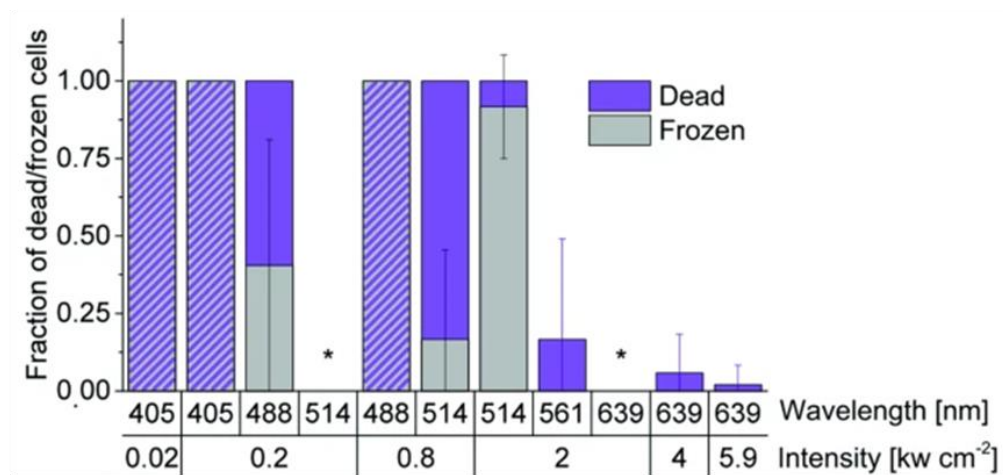
Initial live cell imaging revealed that cells did not progress through mitosis when imaged for 60 minutes (Figure 3.2E; imaging conditions - 365 nm laser power at 10% for 100 ms, 580 nm laser power at 50% for 200 ms). A likely issue was the use of a 365 nm laser required for exciting the Hoechst 33342 DNA stain typically used in live cell imaging. This laser is reported to cause short laser wavelength phototoxicity even at low intensities and for short pulses which would “freeze” and kill cells (Wäldchen *et al.*, 2015; Figure 3.2F). Imaging a “frozen” cell in mitosis for over 60 min showed that it did not progress through mitosis, even with reduced laser intensity (0.02 to 0.2 kW/cm) and exposure time. The 365 nm laser also bleached the red SNAP ligand TMR-Star which was an obstacle to visualising SNAP-SAF-A (Figure 3.2G). To circumvent this, I opted to use a far-red DNA stain instead, since the 635 nm laser required for excitation was not found to be phototoxic even at a high laser intensity of 6 kW/cm (Figure 3.2F). I tested the commonly used DRAQ5 and SiR-DNA far-red DNA stains (Figure 3.2H; imaging conditions - DRAQ5 635 nm laser power at 20% for 100 ms and for SiR-DNA 635 nm laser power at 50% for 40 ms). Cells treated with DRAQ5 did not progress through mitosis which might be due to its cytotoxicity even at low concentrations (Lukinavičius, *et al.*, 2015). In contrast, cells treated with SiR-DNA allowed me to successfully image cells going through mitosis. However, many cells continued to be stuck and not progress through mitosis, therefore I repeated the experiment with the addition of L-ascorbic acid an hour prior to imaging as a protective agent against photodamage (Figure 3.2I; imaging conditions - 635 nm laser power at 73% for 100 ms; Wäldchen *et al.*, 2015; Harada *et al.*, 2023). L-ascorbic acid is thought to suppress mitotic phototoxicity by neutralising superoxide radicals and by directly interacting with DNA to reduce photo-induced double-stranded DNA breaks (Harada *et al.*, 2023). Both tested concentrations (100 μ M and 500 μ M) of L-ascorbic acid permitted successful mitotic progression so for subsequent experiments samples were supplemented with 100 μ M L-ascorbic acid.

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

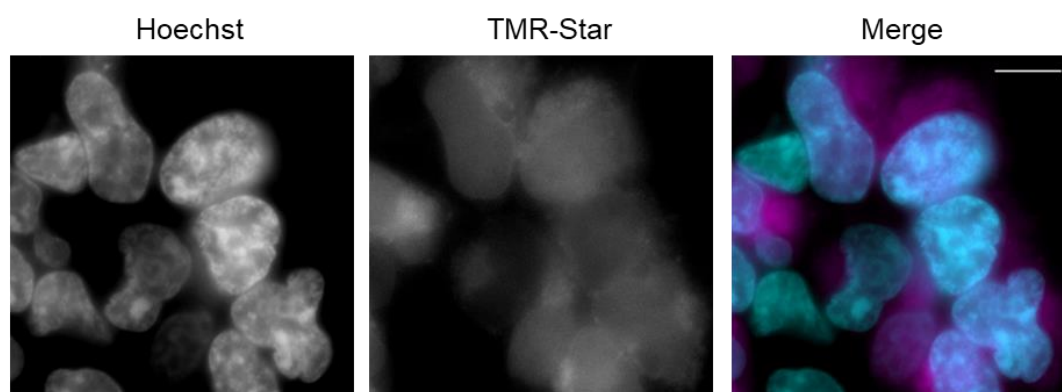
E



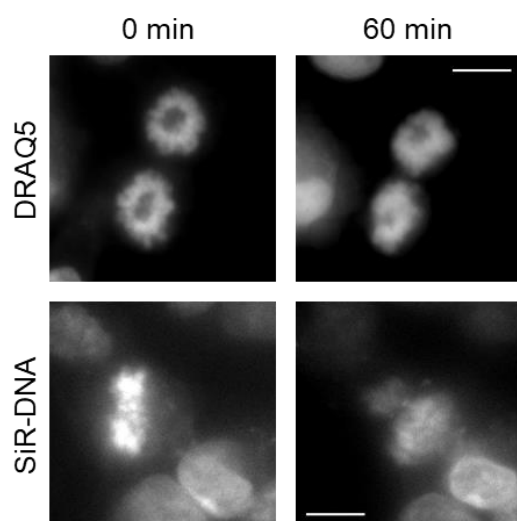
F



G



H



I

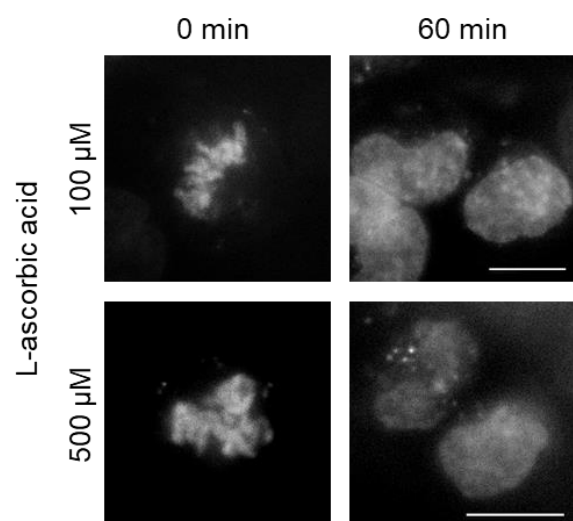
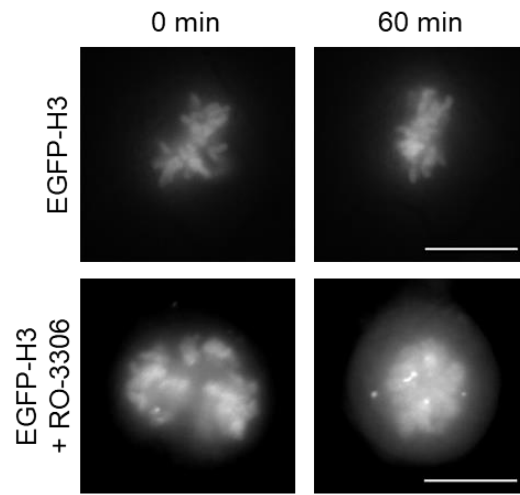


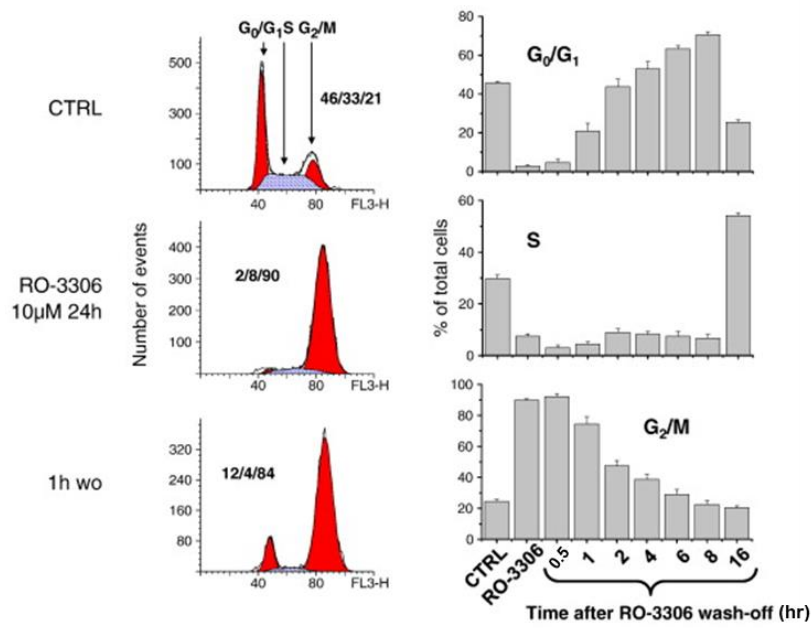
Figure 3.2. Optimisation of live cell imaging conditions of FlpIn 293 T-REx cells expressing SNAP-SAF-A. A) Western blot of cells cultured in the absence or presence of doxycycline (Dox) for 24 hr and probed with an antibody that recognised SAF-A protein. B) Fixed cell immunofluorescent staining of DNA and SAF-A in the absence or presence of Dox for 24 hr. SNAP-SAF-A in cells treated with Dox was visualised using the red SNAP ligand probe, TMR-Star. Endogenous SAF-A in cells without Dox treatment was visualised using a rabbit anti-SAF-A antibody and a mouse anti-rabbit Alexa Fluor 647 (AF647) secondary antibody. DNA was visualised by DAPI staining. C) and D) Fixed cell immunofluorescent staining of DNA and SAF-A in the absence or presence of Dox for 72 hr. SNAP-SAF-A was visualised using the red SNAP ligand probe, TMR-Star while DNA was visualised by DAPI staining. D) Images were taken after a 2-week treatment with the antibiotics hygromycin B and blasticidin S. E) Live cell imaging of a cell “frozen” in mitosis with representative images taken at the start of imaging and 60 min later. DNA was visualised using Hoechst 33342 and SNAP-SAF-A was visualised with TMR-Star. F) Observed cell death due to phototoxicity from laser wavelength and intensity. Results based on continuous wave (cw) irradiation for 240 s. Figure taken from Wäldchen *et al.*, 2015. G) Live cell imaging of a cell population showing bleaching of TMR-Star signal after 365 nm irradiation. DNA was visualised using Hoechst 33342 and SNAP-SAF-A was visualised with TMR-Star. H) Live cell imaging of cells in mitosis for 60 min with images taken every minute. Showing images from the start and end of imaging. DNA was visualised in cells with the far-red probes DRAQ5 and SiR-DNA. I) Live cell imaging of cells during mitosis for 60 min with images taken every minute. Images shown are from the start and end of imaging. Cells were incubated with 100 μ M or 500 μ M L-ascorbic acid for 1 hr before imaging and imaged in its presence. Scale bar is 10 μ m.

To image a larger number of cells while minimizing light exposure, cell synchronization at the G₂/M boundary was attempted using RO-3306 for a 24-hr period and washing off the CDK1 inhibitor before imaging (Vassilev, 2006). In Figure 3.3A (imaging conditions - 470 nm laser power 50% for 30 ms), cells were transfected with EGFP-H3 to visualise chromatin to circumvent the potential 365 nm laser cytotoxicity. However, with and without RO-3306 treatment, cells remained arrested during mitosis, similar to Figure 3.2E. Considering that imaging with a 470 nm laser to excite EGFP likely caused mitotic arrest due to phototoxicity (Figure 3.2F) and that that RO-3306 treatment might delay mitotic entry (Figure 3.3B), the strategy of cell synchronization to increase the number of imaged cells was discontinued. An alternative strategy was to monitor the cell cycle stage using a molecular marker. One approach was to use the C-terminus of DNA-helicase B (HELB) coupled to a fluorescent tag which localises to the nucleus or cytoplasm depending on the cell stage (Spencer *et al.*, 2013; Figure 3.3C). Constructs were made with HELB fused downstream of EGFP or mCherry and were used to identify cells in different stages of the cell cycle (see Table 2.10.). However, as other experiments showed that FlpIn 293 T-REx cells were sensitive to phototoxicity, the range of lasers used in the live cell experiment was limited to two, 580 nm for red stains like TMR-Star, a SNAP ligand, and 635 for far-red stains like SiR-DNA. Although, HELB is a useful tool for identifying cells at specific stages in the cell cycle in many other experiments, it was difficult to use it for my studies as there were not sufficient additional laser lines available.

A



B



C

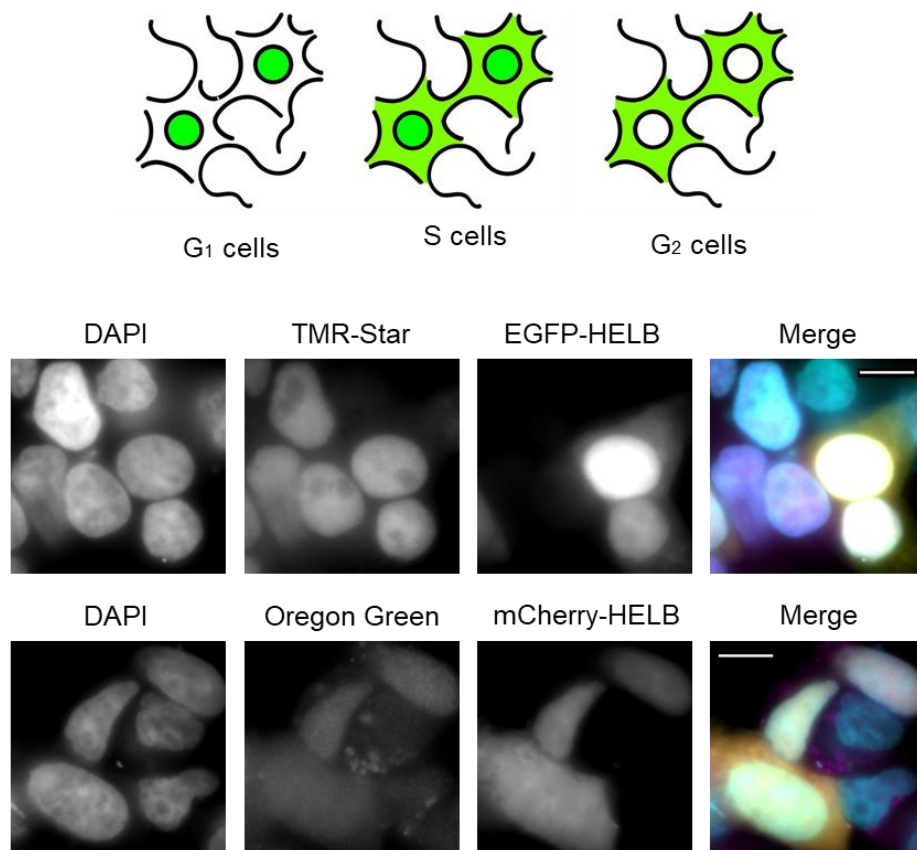


Figure 3.3. Synchronisation aids for live cell imaging of FlpIn 293 T-REx cells. A) Live cell imaging of cells in mitosis after transfection with EGFP-H3 with or without 24-hr 10 μ M RO-3306 treatment to arrest cells at the G₂/M boundary. RO-3306 treated cells were washed with PBS prior to imaging. Cells were imaged for 60 min and images were taken every minute. B) Flow cytometry of HEK-293 cells' DNA content to determine the population cell stage of control (CTRL) and RO-3306 treated cells. On the left is the representative of cell cycle distribution in different conditions (two means 1 hr after RO-3306 washout) with numbers indicating the percentage of cells in different cell stages (G₁, S, and G₂/M). On the right are histograms showing the distribution of cells in different cell cycle stages in different conditions and at different times after release from arrest with RO-3306. Figure taken and adapted from Borowiec *et al.*, 2014. C) Top is a schematic representation of cells expressing the C-terminus DNA-helicase B (HELB) coupled to a fluorescent protein and its localisation in the cells. The bottom images are fixed immunofluorescent cells expressing SNAP-SAF-A (TMR-Star and Oregon Green) or fluorescent-tagged HELB (EGFP-HELB and mCherry-HELB). DNA is visualised with DAPI. Scale bar is 10 μ m.

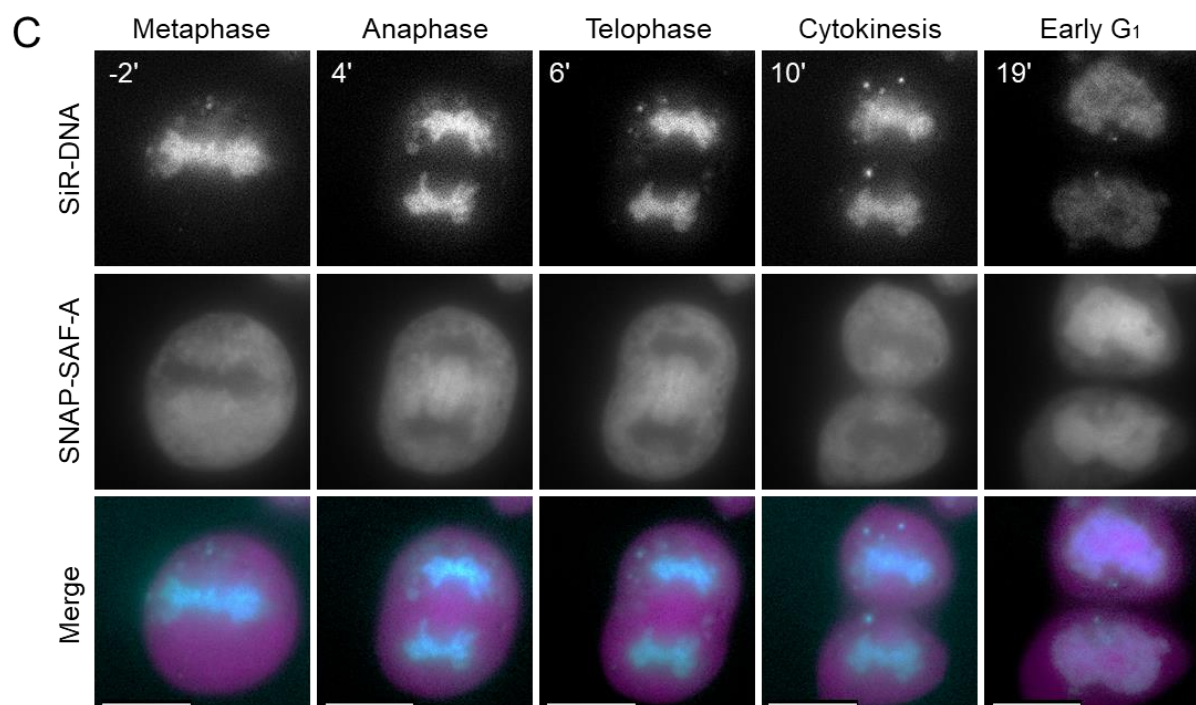
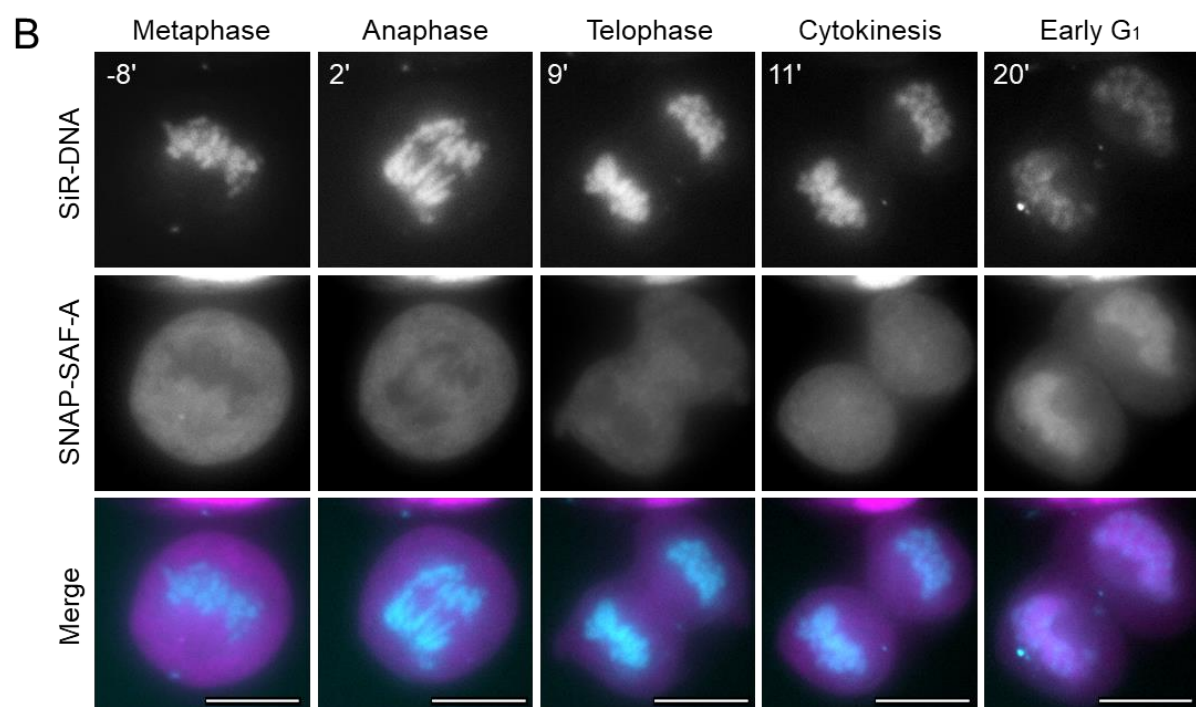
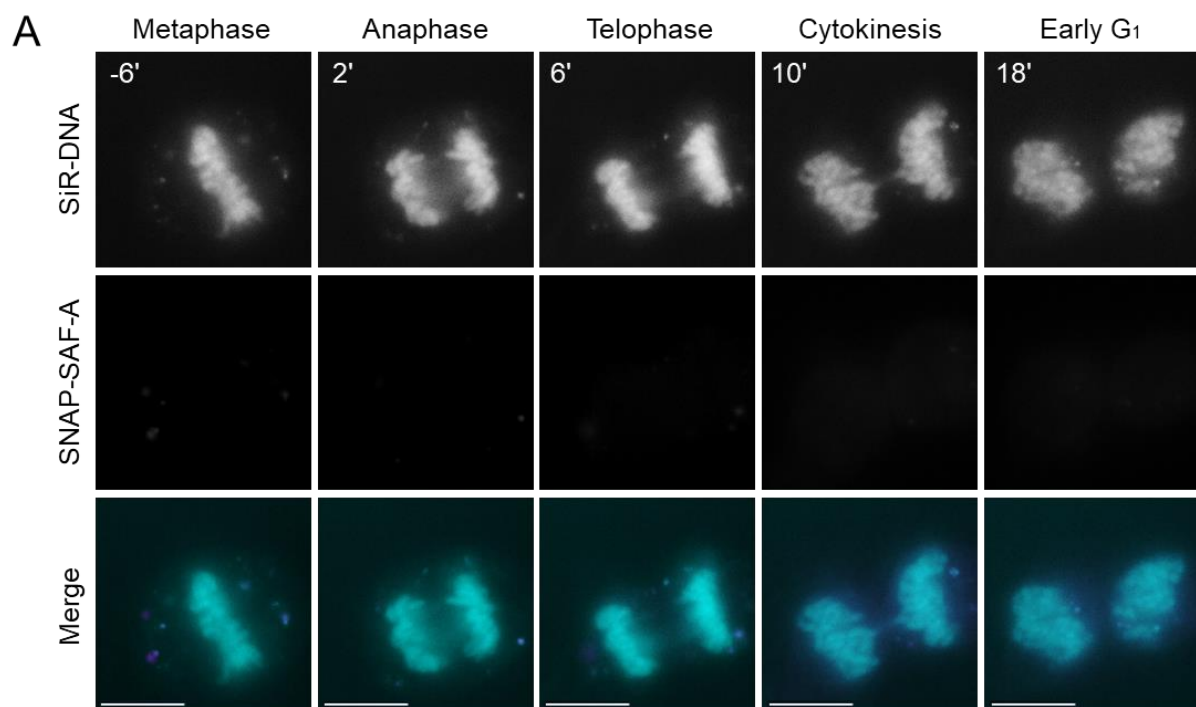
3.2.2. Co-localisation of SAF-A and DNA from mitosis to early-G₁

After optimising the conditions for live cell imaging, I imaged FlpIn 293 T-REx cells expressing SNAP-SAF-A transitioning through mitosis (Figure 3.4). Cells underwent the expected cytologically observable phases of mitosis, from prophase to cytokinesis. Due to challenges with cell selection, not all cells were observed from the start of prophase. Therefore, all timelapse live cell images shown in this chapter (Figures 3.4, 3.5, and 3.6 with imaging conditions were 580 nm laser power at 14% for 30 ms and 635 nm laser power at 73% for 100 ms) include only metaphase, anaphase, telophase, and cytokinesis, hence referred to as the mitotic phases. Also, the start of anaphase was used as the beginning of the mitotic timeline in these images because all cells were imaged entering this phase, unlike metaphase. All imaged live cells in this chapter had comparable lengths of anaphase, telophase, and cytokinesis when analysed using the Wilcoxon signed-rank test (Table 3.1). Figure 3.4A confirms that cells not treated with Dox do not express SNAP-SAF-A. This lack of expression is consistent with the continuous treatment of passaging cells with hygromycin and blasticidin, as mentioned previously. A 24-hr treatment with 1 μ g/ml Dox was sufficient to image SNAP-SAF-A's behaviour during mitosis in live cells (Figure 3.4B). During metaphase, anaphase, and telophase, SNAP-SAF-A is excluded from mitotic chromatin as previously reported (Sharp *et al.*, 2020). Co-localisation between the DNA and protein signal can be observed starting from cytokinesis and continuing as the daughter cells exit division in the early G₁ phase.

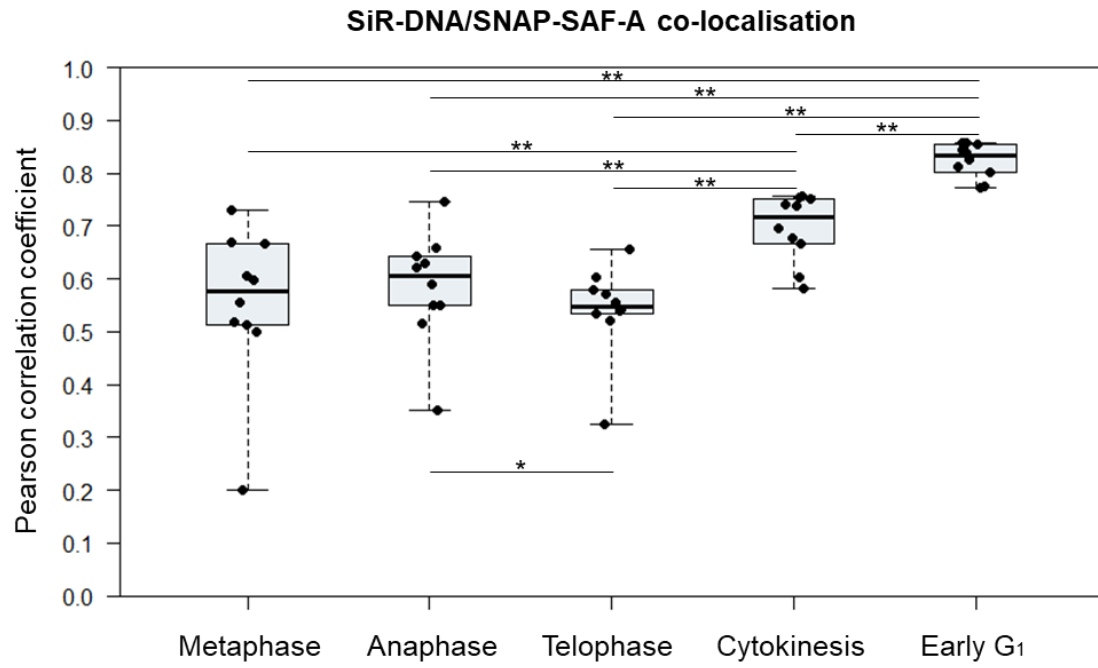
This visible overlap between SAF-A and DNA signal might be due to the protein weakly binding to chromatin or due to nuclear import following nuclear envelope reassembly but these imaging experiments are unable to differentiate between the two. During the early G₁ phase, the SNAP-SAF-A signal is co-localised with the DNA signal. The timing of when SNAP-SAF-A re-establishes the characteristic nuclear diffuse pattern, as observed in Figure 3.2B, remains unknown. Imaging was not conducted for more than 30 minutes following mitotic exit, preventing observation of this process. The co-localisation between the DNA and protein signal during the mitotic phases and early G₁ of 10 cells were quantified using the Pearson correlation coefficient, generated by the Coloc2 plugin from ImageJ and presented as a box plot in Figure 3.4D. Statistical analysis (Wilcoxon signed-rank test) of the results showed statistically significant differences in Pearson correlation coefficient values between early G₁ and the mitotic phases ($p = 0.002$), cytokinesis and the rest of the mitotic phases ($p = 0.002$ vs metaphase and telophase, $p = 0.006$ vs anaphase), and between anaphase and telophase ($p = 0.01$). The lowest co-localisation during DNA and SNAP-SAF-A is during telophase, from where it increases to higher than the rest of mitosis levels in cytokinesis and early G₁. This coincides with a visible overlap seen in cells starting from cytokinesis.

Condition	Anaphase	Telophase	Cytokinesis
No Dox (Figure 3.4A)	5.5'	4'	1.6'
SAF-A (Figure 3.4B)	5.4'	3.9'	1.7'
SAF-A + triptolide (Figure 3.4C)	5.5'	4.1'	1.7'
HNRNPC1 (Figure 3.5C)	6'	4'	1.6'
HNRNPC1 + triptolide (Figure 3.5D)	5.4'	4'	1.6'
HNRNPA1 (Figure 3.6C)	6.6'	4.2'	1.8'
HNRNPA1 + triptolide (Figure 3.6D)	7'	4.2'	1.8'

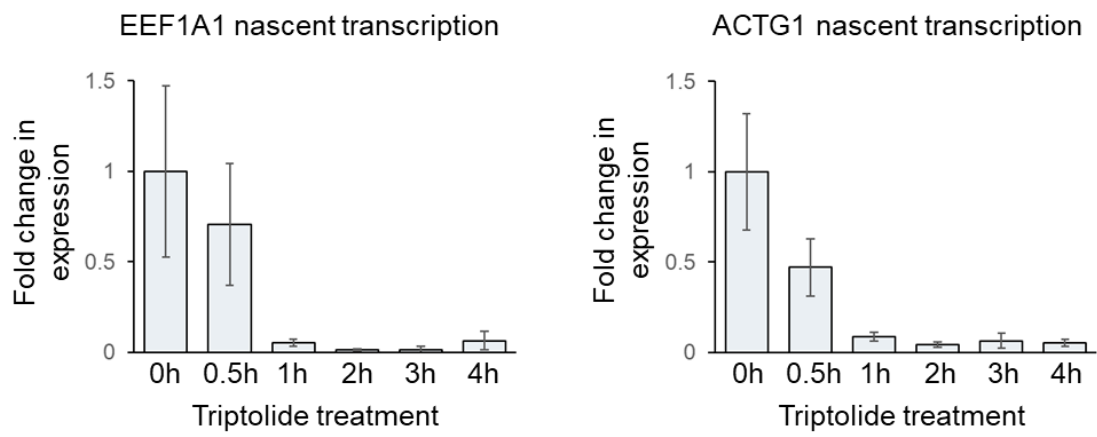
Table 3.1. Average length of mitotic phases for live cell imaging conditions. Each average is a measure of 6 cells (Figure 3.4A), 10 cells (Figures 3.4B and 3.4C), or 5 cells (Figures 3.5C, 3.5D, 3.6C, and 3.6D). Statistical analysis with the Wilcoxon signed-rank test revealed no statistical differences between the different conditions for each mitotic phase.



D



E



F

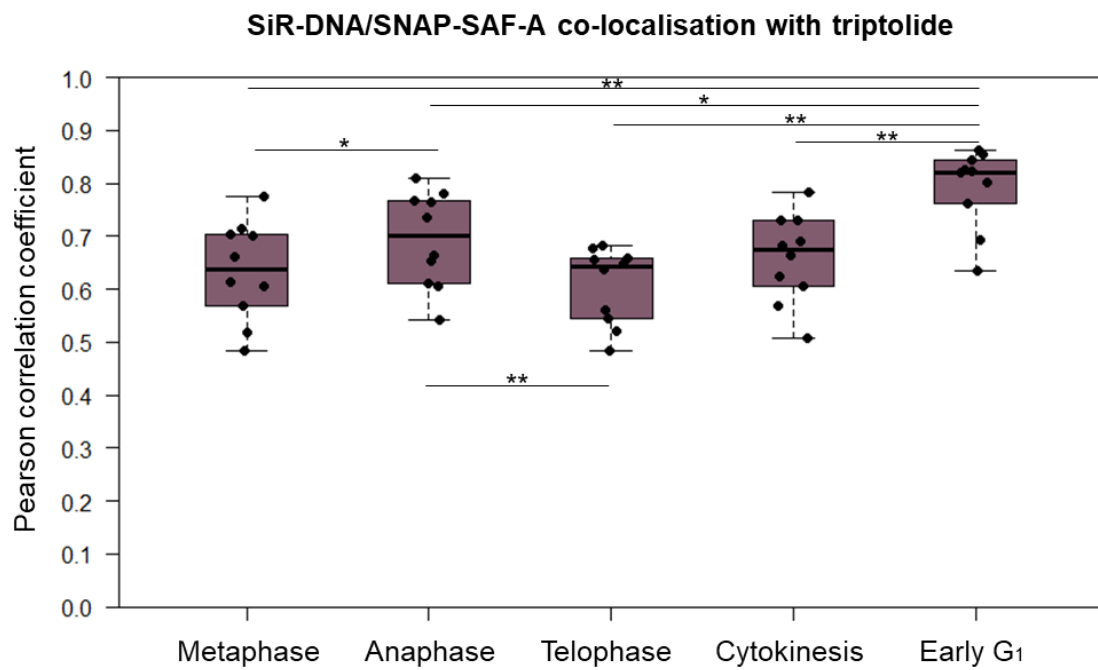


Figure 3.4. Live cell imaging of SNAP-SAF-A and co-localisation with DNA in FlpIn 293 T-REx cells. In figures A to C the chosen images are representative of different mitotic stages and early G₁ is shown as 10 min after cell division (except in the case of figure B where it is 8 min). The timer in the top left of images corresponds to the time before/after start of anaphase. A) Representative cell imaged without doxycycline (Dox) and treated with SiR-DNA to visualise DNA and TMR-Star to check off-target expression. B) Representative cell imaged after 24-hr treatment with Dox. TMR-Star was used to visualise SNAP-SAF-A while DNA was visualised with SiR-DNA. C) Representative cell imaged after 24-hr treatment with Dox and 1 hr treatment with triptolide. The cell underwent mitosis in the presence of triptolide. D) Box plot of Pearson correlation coefficient values generated from the Coloc2 ImageJ plugin. Each stage on the x-axis is the average of 10 single cells' values with differences between them analysed using the Wilcoxon signed-rank test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the respective stages (* = $p \leq 0.05$, ** = $p \leq 0.01$). E) RT-qPCR using exon-intron primers for EEF1A1 (left) and ACTG1 (right). Each sample was normalised using 18S expression as a loading control and then normalised to the 0h control time point using the $\Delta\Delta Ct$ method. The y-axis shows the fold change in expression calculated with $2^{-\Delta\Delta Ct}$. The x-axis shows the triptolide treatment time course. Data show mean and s.e. of 3 technical replicates. F) As D) but for cells treated with triptolide for 1 hr. Scale bar is 10 μm .

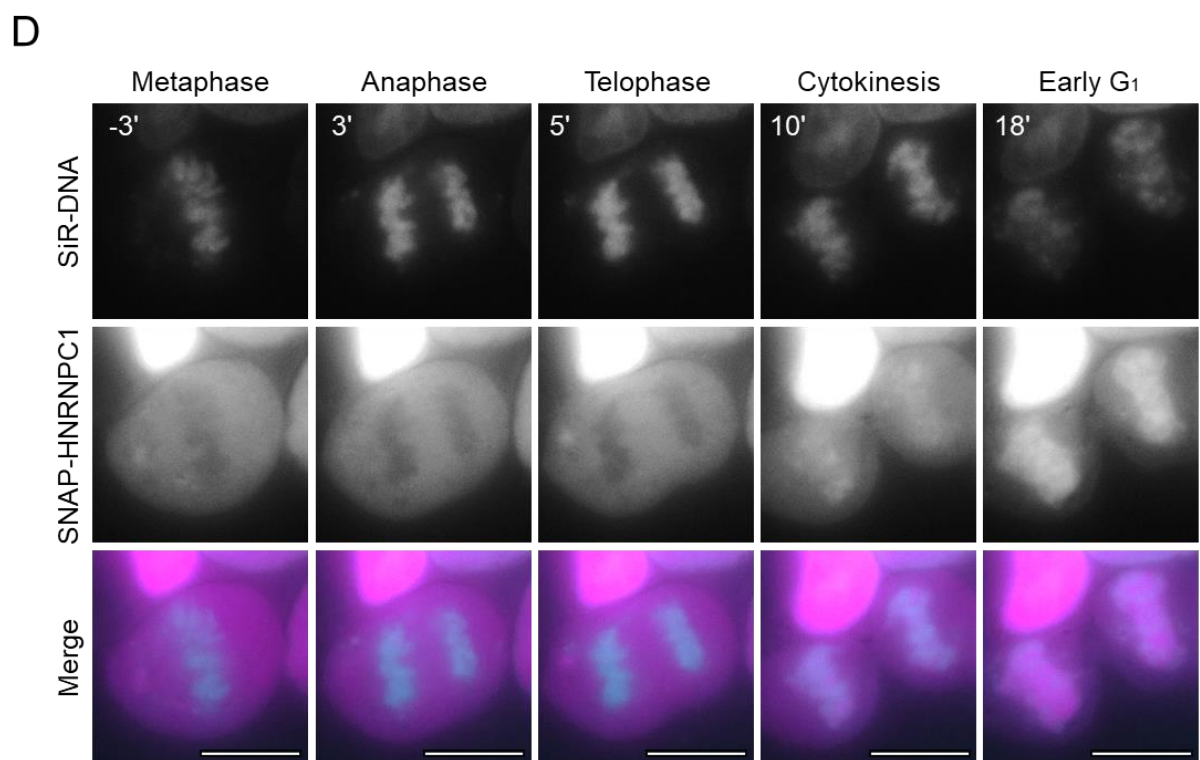
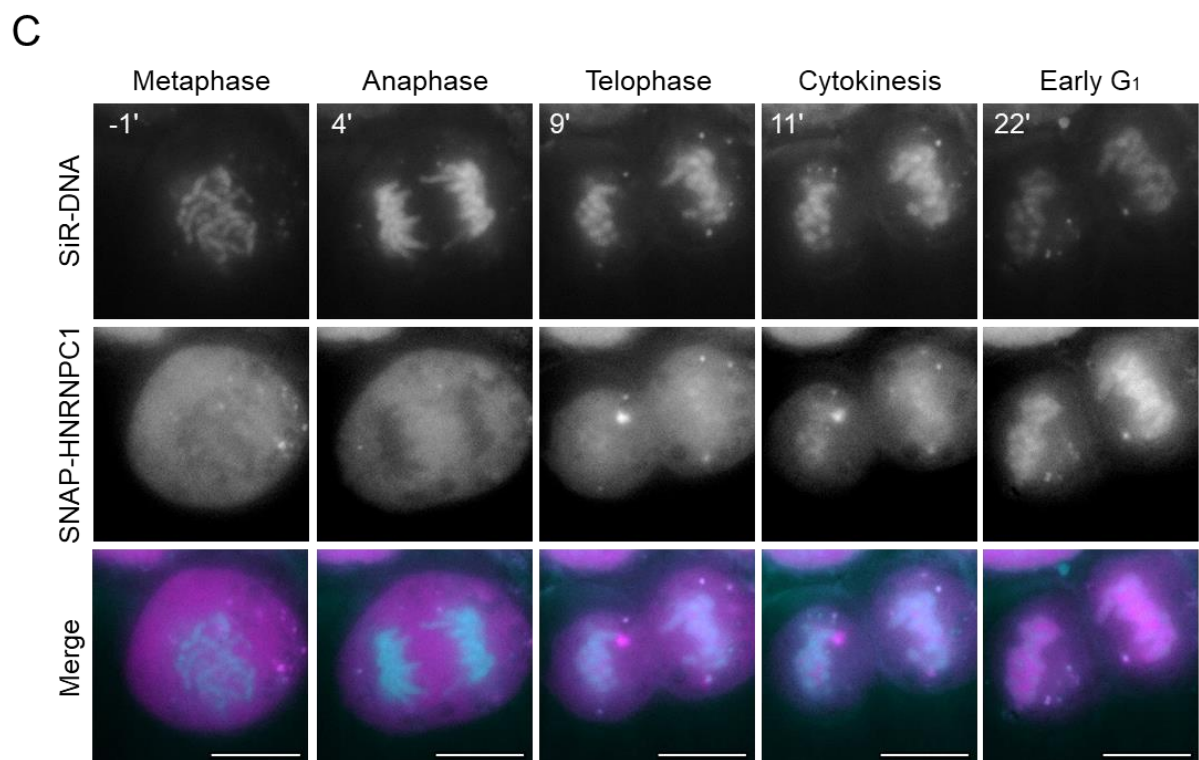
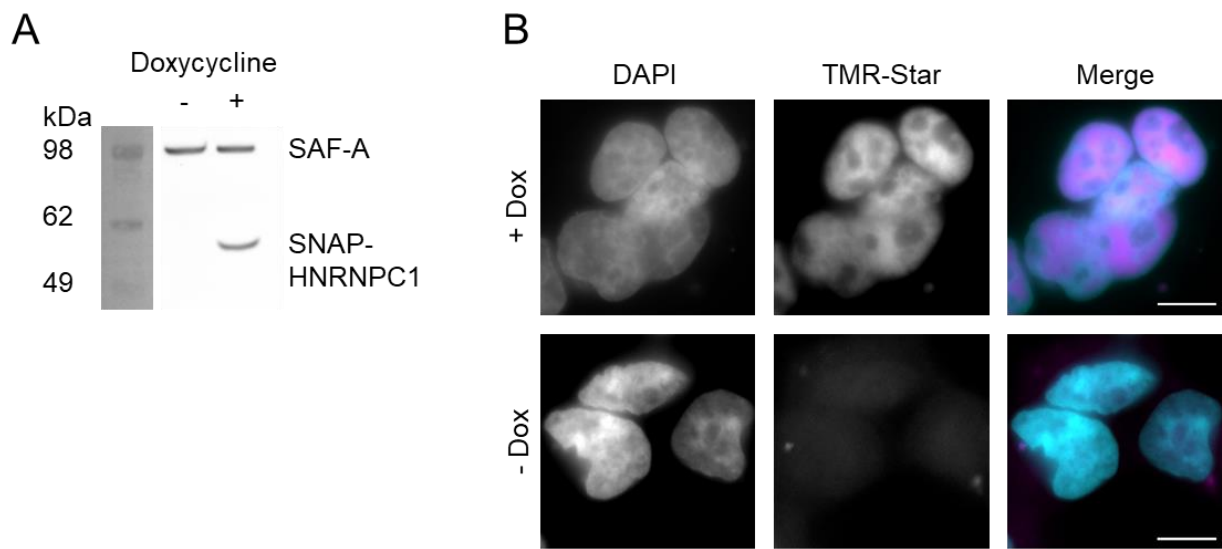
After these initial studies, I wanted to investigate whether perturbation of nascent transcription during mitosis affected SAF-A behaviour. To determine the optimal conditions for transcription inhibition, cells were treated with triptolide for different times and monitored by RT-qPCR of pre-spliced RNA from the highly expressed genes *EEF1A1* and *ACTG1* (Figure 3.4E). Cells were collected at time 0, after 0.5 h, 1 h, 2 h, 3 h or 4 h of 1 μ M triptolide treatment and RNA was extracted. After reverse transcription using random primers, the cellular levels of *EEF1A1* and *ACTG1* nascent transcripts were estimated using qPCR with primers spanning exon-intron boundaries. After 0.5 h triptolide treatment, there is a 0.25- and 0.5-fold change in *EEF1A1* and *ACTG1* nascent transcript levels respectively. While after 1 to 4 h of treatment, there is a 0.9-fold change in transcript levels, confirming the reported rapid effect of triptolide. Based on these results, 1 hr triptolide treatment prior to imaging was performed to visualise SNAP-SAF-A in the absence of transcription during mitosis (Figure 3.4C). It was observed that cells treated with triptolide exhibited exclusion of SNAP-SAF-A from metaphase through cytokinesis, in contrast to cells under normal conditions. In early G₁ SNAP-SAF-A was seen to be enriched in the DNA signal. Quantification of SAF-A and DNA co-localisation in multiple cells and subsequent statistical analysis showed similar to normal conditions statistically significant differences in Pearson correlation coefficient values between early G₁ and the mitotic phases in 10 imaged cells (Figure 3.4F; Wilcoxon signed-rank test, $p = 0.01$ vs metaphase, $p = 0.04$ vs anaphase, $p = 0.004$ vs telophase, and $p = 0.002$ vs cytokinesis). However, there were no significant differences in Pearson correlation coefficient values between cytokinesis and the other mitotic phases. Telophase was similarly observed to have the lowest co-localisation scores with statistically significant differences in Pearson correlation coefficient values between anaphase and telophase ($p = 0.002$). There were also statistically significant differences in Pearson correlation coefficient values found between metaphase and anaphase ($p = 0.05$). Overall, triptolide treatment delayed the mitotic binding of SNAP-SAF-A during cytokinesis but did not significantly alter the enrichment of co-localisation with DNA during early G₁.

3.2.3. Co-localisation of DNA and HNRNPC1 from mitosis to early-G₁

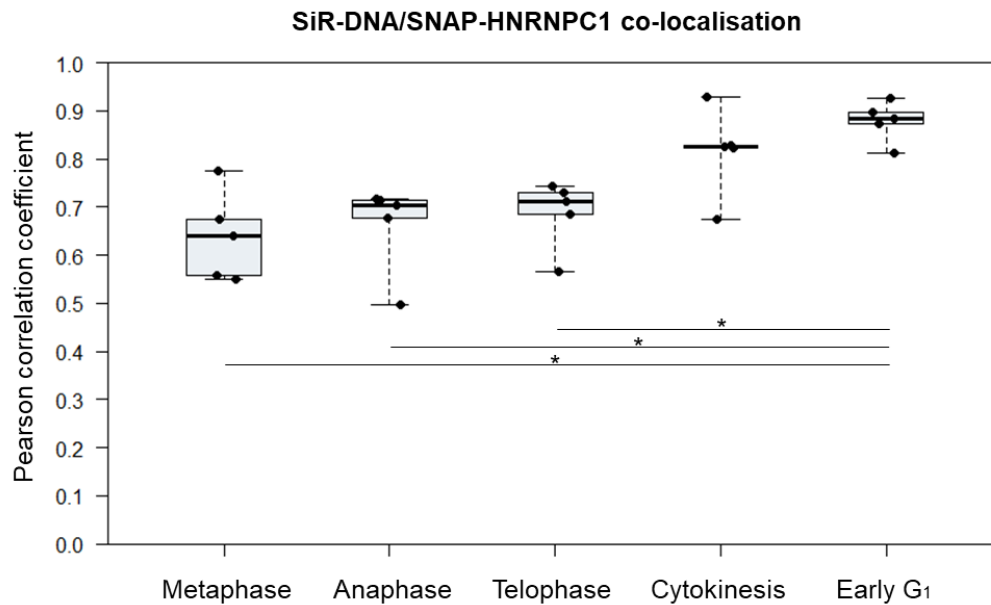
To validate the functionality of my live cell imaging system and ensure that the observed results with SNAP-SAF-A were not due to artifacts, I aimed to replicate the study using another protein whose behavior during mitosis has been documented. Previously HNRNPC1 has been investigated in fixed cells and with live cell imaging (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). To perform live cell imaging to examine HNRNPC1 during mitosis, FlpIn 293 T-REx cells were generated with a SNAP-HNRNPC1 gene integrated at the FRT site. FlpIn 293 T-REx cells were transfected with a Flp recombinase expression plasmid pOG44 and an expression plasmid containing a FRT site and SNAP-HNRNPC1 according to the system manual. After treatment with antibiotics hygromycin and blasticidin, cells grown in the presence or absence of 1 µg/ml Dox were lysed and probed with an anti-FLAG antibody and an anti-SAF-A antibody (Figure 3.5A). Only the cells treated with Dox expressed the expected 59 kDa SNAP-HNRNPC1 as visualised by Western blotting (Figure 3.5A). Both treated and untreated cells expressed comparable amounts of endogenous SAF-A as a control whilst only cells treated with Dox expressed SNAP-HNRNPC1. Cells were also probed with the red SNAP ligand TMR-Star to check SNAP-HNRNPC1 distribution in the nucleus using immunofluorescence (Figure 3.5B). The observed distribution of the protein was comparable to images of HNRNPC1's diffuse distribution in the literature (Siomi and Dreyfuss, 1995; Lee *et al.*, 2004).

After confirming the correct expression and distribution of SNAP-HNRNPC1 in the generated cell line, I proceeded to investigate its behavior during mitosis using live cell imaging. Cells were treated with Dox for 24 hr and imaged going through mitosis for an hour after treatment with TMR-Star and SiR-DNA to visualise SNAP-HNRNPC1 and DNA respectively. During metaphase and anaphase, SNAP-HNRNPC1 was excluded from chromosomes (Figure 3.5C). The protein and DNA signal began to overlap during telophase and was continuously co-localised during cytokinesis and early G₁. This is consistent with previous evidence showing HNRNPC1 becomes localised with mitotic chromatin during telophase at a similar time as RNA Pol II (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). In Figure 3.5E, the Pearson correlation coefficient values from five cells are presented as a box

plot. The statistical analysis revealed statistically significant differences in Pearson correlation coefficient values between early G₁ and metaphase, anaphase, and telophase (Wilcoxon signed-rank test, $p = 0.03$). The median value of co-localisation was lowest in metaphase, and there was an increase in co-localisation between the protein and DNA throughout the remaining phases of mitosis.



E



F

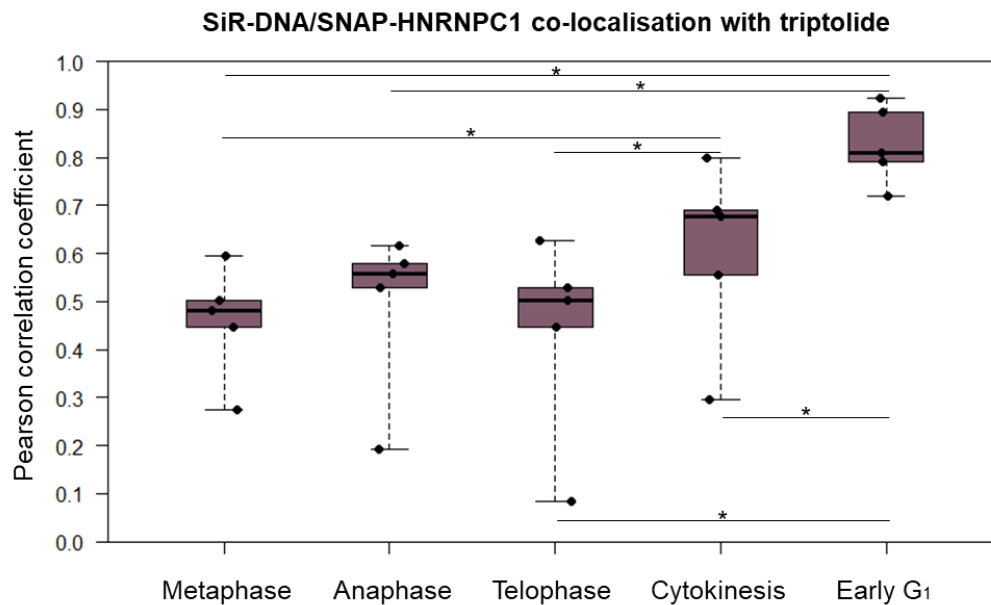


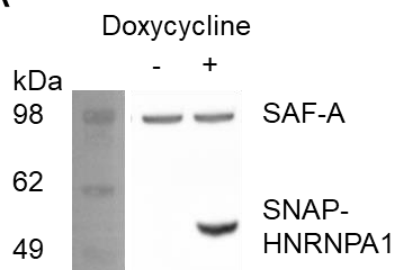
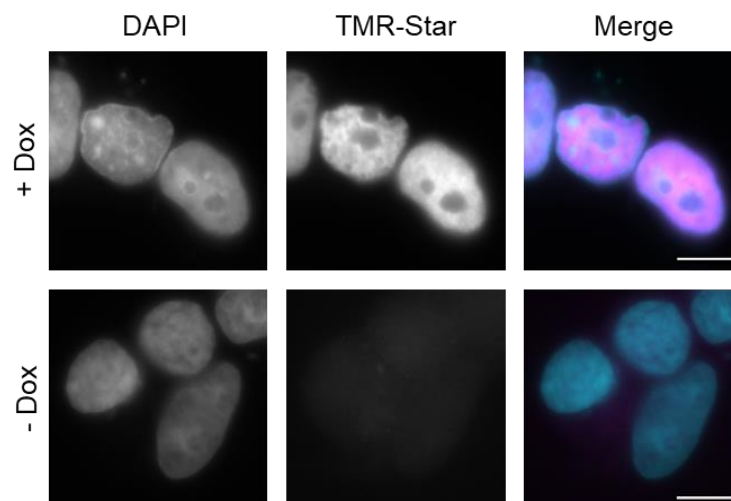
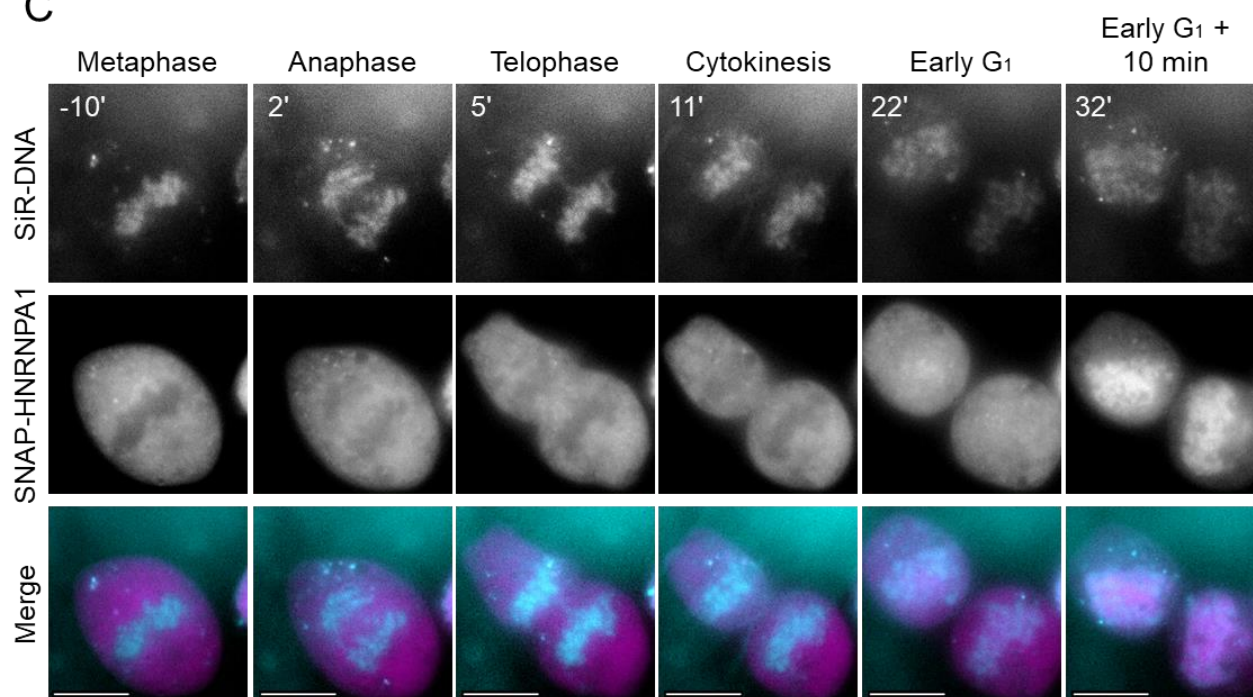
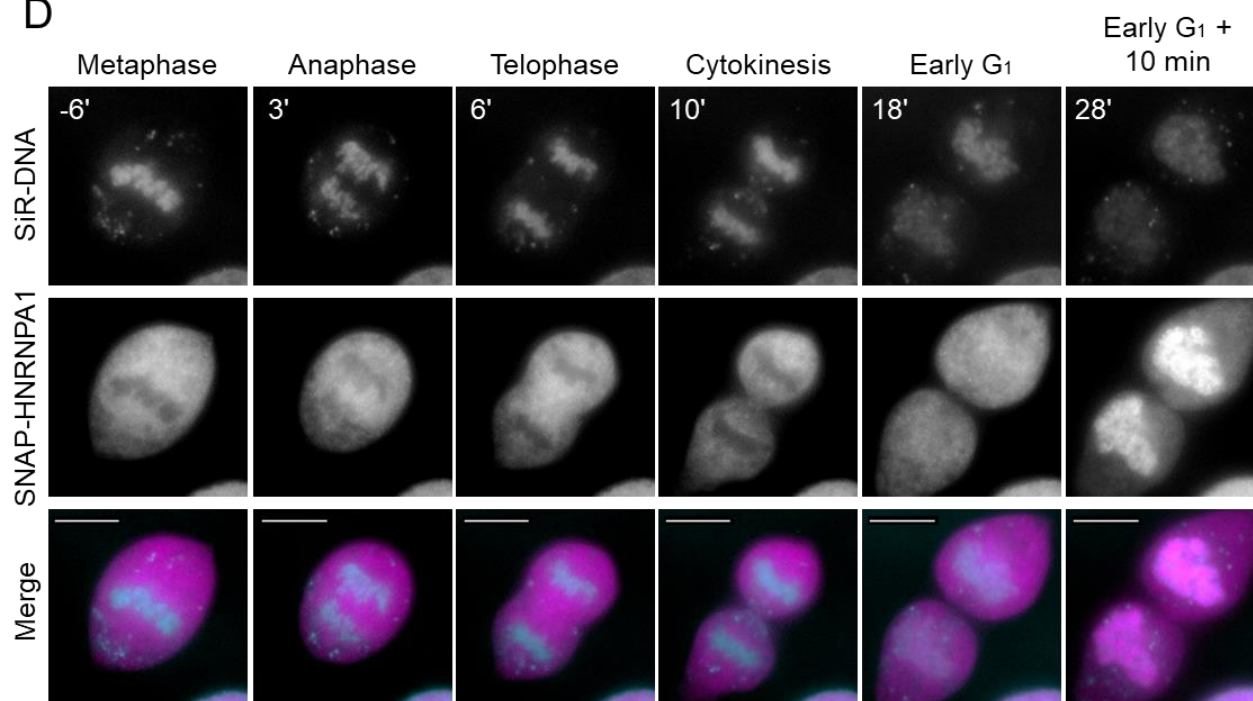
Figure 3.5. Live cell imaging of SNAP-HNRNPC1 and co-localisation with DNA in FlpIn 293 T-REx cells. In figures C and D the chosen images are representative of different mitotic stages and early G₁ is shown as 10 min after cell division. The timer in the top left of images corresponds to the time before/after start of anaphase. A) Western blot of cells cultured in the absence or presence of doxycycline (Dox) for 24 hr and probed for FLAG-tagged proteins (SNAP-HNRNPC1) and SAF-A as a loading control. B) Fixed cell immunofluorescent staining of DNA and SNAP-HNRNPC1 in the absence or presence of Dox for 24 hr. SNAP-HNRNPC1 was visualised using the red SNAP ligand probe, TMR-Star. DNA was visualised by DAPI staining. C) Representative cell imaged after 24-hr treatment with Dox. TMR-Star was used to visualise SNAP-HNRNPC1 while DNA was visualised with SiR-DNA. D) Representative cell imaged after 24-hr treatment with Dox and 1 hr treatment with triptolide. The cell underwent mitosis in the presence of triptolide. E) Box plot of Pearson correlation coefficient values generated from the Coloc2 ImageJ plugin. Each stage on the x-axis is the average of 5 single cells' values with differences between them analysed using the Wilcoxon signed-rank test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the respective stages (* = $p \leq 0.05$). F) As E) with values from cells treated with triptolide for 1 hr. Scale bar is 10 μm .

To investigate the effect of nascent transcription, I inhibited mitotic transcription using 1 μ M of triptolide for 1 hr before imaging. Cells were then imaged going through mitosis for a further hour. After transcription inhibition, colocalisation between SNAP-HNRNPC1 and SiR-DNA was delayed until cytokinesis unlike the co-localisation observed during telophase in control cells (Figures 3.5C and 3.5D). The Pearson correlation coefficient values for co-localisation between SNAP-HNRNPC1 and DNA from 5 cells treated with triptolide were plotted and statistically significant differences in Pearson correlation coefficient values between the mitotic phases and early G₁ were tested with the Wilcoxon signed-rank test (Figure 3.5F). Similarly to cells in the absence of triptolide, there were statistically significant differences in Pearson correlation coefficient values between early G₁ and the mitotic phases, including cytokinesis ($p = 0.03$). These findings support the notion that inhibiting mitotic transcription leads to co-localisation between SNAP-HNRNPC1 and DNA at a later mitotic phase. In addition, there were statistically significant differences in Pearson correlation coefficient values observed between cytokinesis and telophase ($p = 0.03$), and cytokinesis and metaphase ($p = 0.03$). The lowest median value for co-localisation was found during metaphase, similar to cells without transcription inhibition.

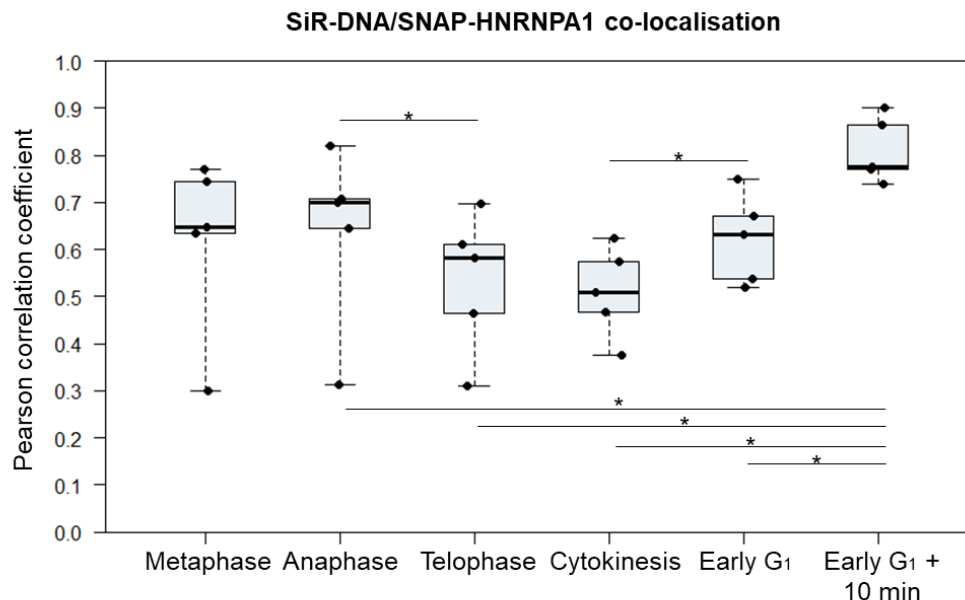
3.2.4. Co-localisation of DNA and HNRNPA1 from mitosis to early-G₁

Another test for the live cell imaging system was with HNRNPA1 which had a different mitotic behaviour and response to mitotic transcription inhibition when compared to HNRNPC1 (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). Like with HNRNPC1, a SNAP-HNRNPA1 construct was integrated at the FRT site of FlpIn T-REx cells. The expression of SNAP-HNRNPA1 was confirmed by probing cell lysates with an anti-FLAG and an anti-SAF-A antibody (Figure 3.6A). Only the cells treated with 1 μ g/ml Dox expressed the expected 53 kDa SNAP-HNRNPA1 protein while both lysates had similar amounts of endogenous SAF-A. SNAP-HNRNPA1 cells were incubated with the red SNAP ligand TMR-Star (Figure 3.6B) and the observed distribution was comparable to the diffuse distribution reported in the literature (Siomi and Dreyfuss, 1995; Choi, 2017).

After confirming the correct expression of SNAP-HNRNPA1 and its distribution in the nucleus, I imaged single cells going through mitosis. To perform this, cells were treated with Dox treatment for 24 hr and incubated with TMR-Star and SiR-DNA before imaging to visualise SNAP-HNRNPA1 and DNA. SNAP-HNRNPA1 was observed to be dissociated from the DNA in all mitotic phases but then there was visible overlap between the protein and DNA after cell division in early G₁ which became enriched over time (Figure 3.6C). This behaviour is similar to what has been previously described in the literature where the protein re-enters the nucleus after nuclear assembly at the end of cell division (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). After imaging 5 separate cells going through mitosis, the Pearson correlation coefficient values for co-localisation between SNAP-HNRNPA1 and DNA were plotted and statistical differences were calculated using the Wilcoxon signed-rank test (Figure 3.6E). There were statistically significant differences in Pearson correlation coefficient values observed between the mitotic phases of anaphase and telophase ($p = 0.03$), between cytokinesis and early G₁ ($p = 0.03$), and between early G₁ after 10 min and the other plotted phases except for metaphase ($p = 0.03$). The lowest observed median value for co-localisation was in cytokinesis where the signal increased significantly during the two observed G₁ time points.

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

E



F

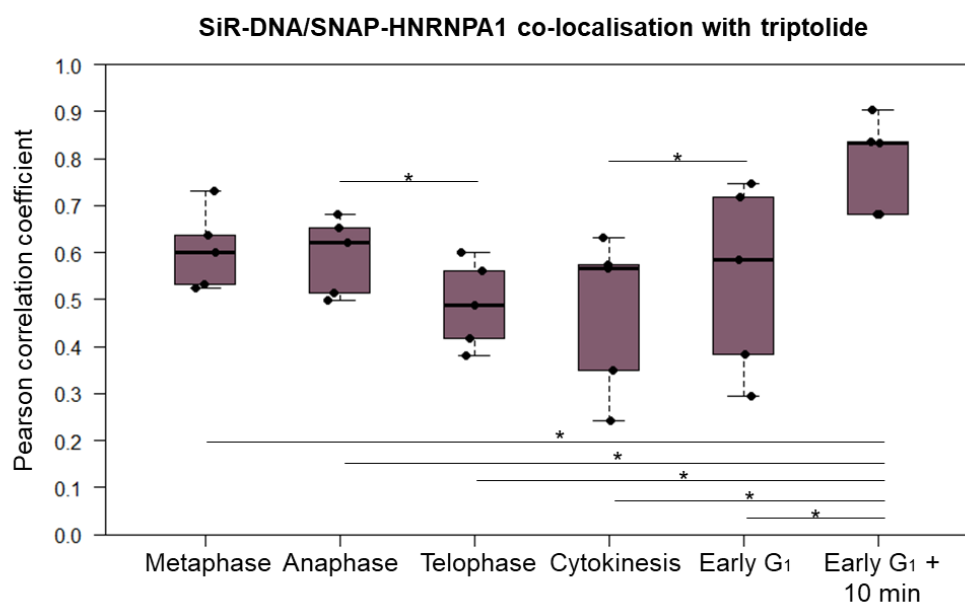


Figure 3.6. Live cell imaging of SNAP-HNRNPA1 and co-localisation with DNA in FlpIn 293 T-REx cells. In figures C and D the chosen images are representative of different mitotic stages and early G₁ is shown as 10 min after cell division. The early G₁ stage + 10 min signifies 20 min after cell division. The timer in the top left of images corresponds to the time before/after start of anaphase. A) Western blot of cells cultured in the absence or presence of doxycycline (Dox) for 24 hr and probed for FLAG-tagged proteins (SNAP-HNRNPA1) and SAF-A as a loading control. B) Fixed cell immunofluorescent staining of DNA and SNAP-HNRNPA1 in the absence or presence of Dox for 24 hr. SNAP-HNRNPA1 was visualised using the red SNAP ligand probe, TMR-Star. DNA was visualised by DAPI staining. C) Representative cell imaged after 24-hr treatment with Dox. TMR-Star was used to visualise SNAP-HNRNPA1 while DNA was visualised with SiR-DNA. D) Representative cell imaged after 24-hr treatment with Dox and 1 hr treatment with triptolide. The cell underwent mitosis in the presence of triptolide. E) Box plot of Pearson correlation coefficient values generated from the Coloc2 ImageJ plugin. Each stage on the x-axis is the average of 5 single cells' values with differences between them analysed using the Wilcoxon signed-rank test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the respective stages (* = $p \leq 0.05$). F) As E) with values from cells treated with triptolide for 1 hr. Scale bar is 10 μ m.

To investigate the effect of nascent transcription, I inhibited mitotic transcription using 1 μ M of triptolide for 1 hr before imaging. Cells were then imaged going through mitosis for a further hour. After transcription inhibition, SNAP-HNRNPA1 was observed to be excluded from mitotic chromatin during cell division and co-localised with DNA in early G₁, similarly to cells without transcription inhibition (Figure 3.6D). The overlap between the protein and DNA was more pronounced 10 minutes later into G₁. Plotting the Pearson correlation coefficient values for co-localisation between DNA and SNAP-HNRNPA1 from the observed cells revealed similar significant differences in Pearson correlation coefficient values from the Wilcoxon signed-rank test as above (Figure 3.6F). There were statistically significant differences in Pearson correlation coefficient values observed between early G₁ after 10 min and the mitotic phases, including metaphase this time ($p = 0.03$). Anaphase and telophase again had statistically significant differences in Pearson correlation coefficient values ($p = 0.03$) as well as cytokinesis and early G₁ ($p = 0.03$).

3.3. SAF-A's interaction with RNA during the cell cycle

3.3.1. SAF-A oligomerisation

To investigate how SAF-A monomer and oligomer bind differently to RNA, protein-protein crosslinkers were utilised in an oligomerisation assay (Nozawa *et al.*, 2017). The sulfhydryl-to-sulfhydryl crosslinker BM(PEG)₂ crosslinks cysteine residues with a spacer of 14.7Å while the primary amine crosslinker disuccinimidyl suberate (DSS) crosslinks lysine residues with a spacer of 11.4Å. FlpIn 293 T-REx cells were treated with trypsin and lightly permeabilised, before addition of crosslinker at concentrations ranging from 0 to 1 mM. The cells were lysed and the proteins were separated on a tris-acetate protein gel and visualised with anti-SAF-A antibodies by Western blotting (Figure 3.7). As the differently crosslinked samples were not fractionated on the same gel, the intensity of oligomer bands between experiments would not be a good comparison. Therefore, the amount and intensity of oligomer bands were compared to monomer bands for each crosslinker. Using this comparison, BM(PEG)₂ crosslinking (Figure 3.7A) yielded more oligomers at increasing concentrations

compared to DSS (Figure 3.7B). Due to this, BM(PEG)₂ was used as a crosslinker in further experiments to increase the SAF-A oligomer isolated from protein samples.

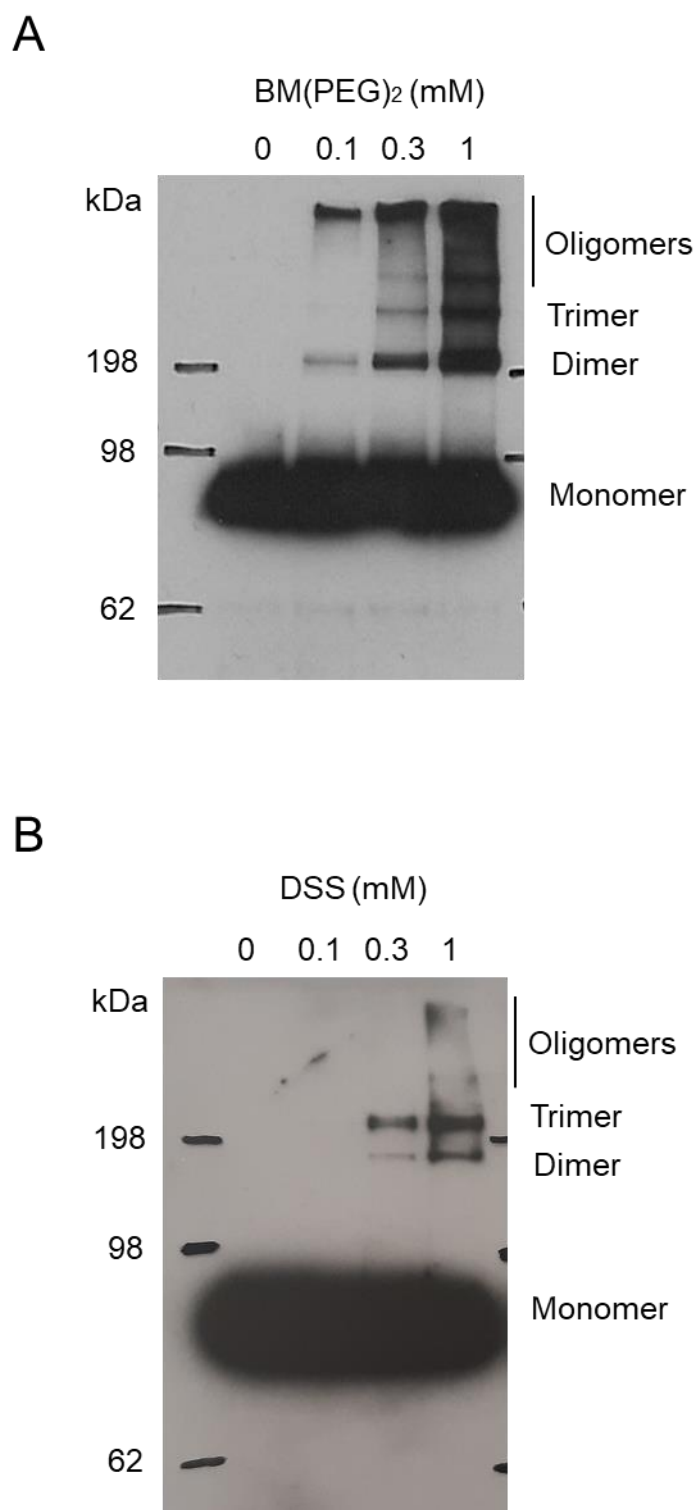


Figure 3.7. SAF-A endogenous protein oligomerisation. Western blots of FlpIn 293 T-REx cells in the absence of Dox which were crosslinked with increasing concentrations of BM(PEG)₂ (A) or DSS (B). The lysed samples were fractionated by SDS-PAGE on a 3-8% tris-acetate gel and probed with anti-SAF-A antibodies.

3.3.2. SAF-A oligomer immunoprecipitation

After successfully visualising SAF-A oligomers using the BM(PEG)₂ crosslinker, the next step was to isolate oligomeric proteins for further experiments. FlpIn 293 T-REx cells expressing Dox-inducible FLAG-tagged SAF-A were treated with trypsin and lightly permeabilised. They were then either treated with DMSO or 1 mM BM(PEG)₂ for 10 minutes to crosslink SAF-A oligomers. After sonication to disrupt the nuclear mesh, input samples were taken and lysed for subsequent visualisation. FLAG-tagged SAF-A was immunoprecipitated with magnetic beads coated in anti-FLAG antibodies and then directly lysed before preparing them for gel loading. The lysates were probed with anti-SAF-A antibodies to visualise both endogenous and FLAG-tagged SAF-A (Figure 3.8). SAF-A trimer was visible with crosslinker treatment but not with DMSO treatment but higher MW oligomers were not clearly visible on blots even after running the gel for extended periods of time.

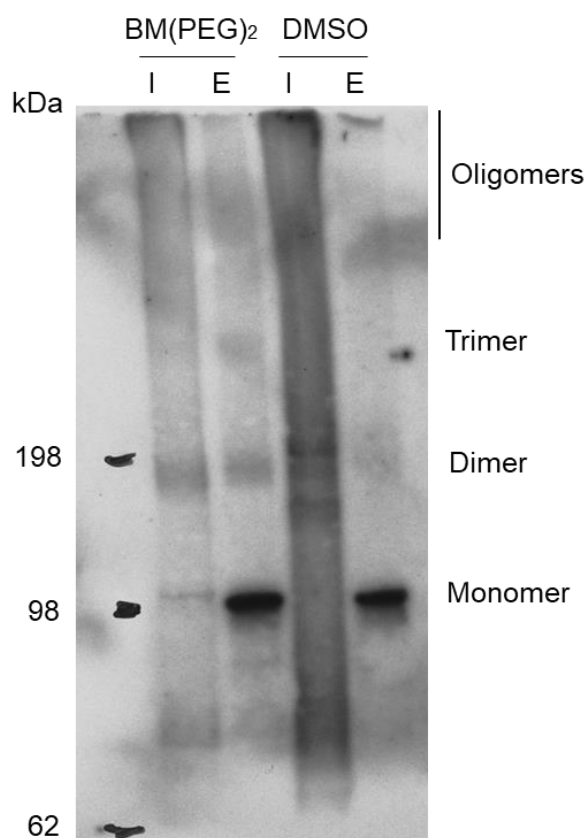
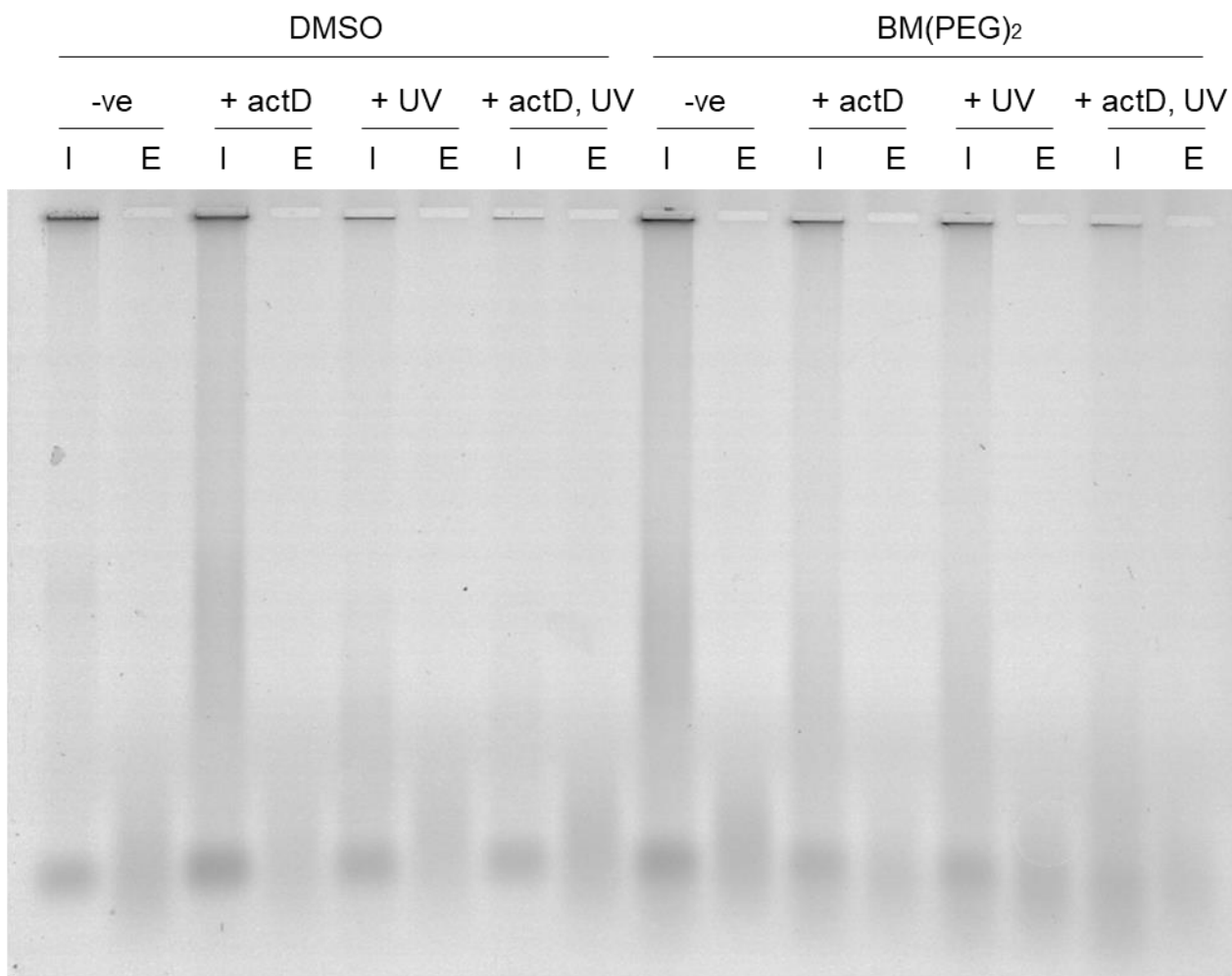


Figure 3.8. SAF-A protein oligomer immunoprecipitation. Western blots of FlpIn 293 T-REx cells expressing FLAG-SAF-A in the presence of Dox which were crosslinked with BM(PEG)₂ or treated with DMSO. Input samples (I) were collected after sonication and centrifugation, before addition to FLAG-tagged magnetic beads. Elute samples (E) were eluted from FLAG-tagged magnetic beads. The lysed samples were fractionated by SDS-PAGE on a 3-8% tris-acetate gel and probed with anti-SAF-A antibodies.

3.3.3. SAF-A and RNA immunoprecipitation

After successfully visualising and immunoprecipitating SAF-A oligomers, I wanted to purify RNA bound to SAF-A to compare the amount of bound RNA between monomeric and oligomeric SAF-A. I purified RNA from FlpIn 293 T-REx cell samples in the presence of a low-dose of actD (used to inhibit rRNA production), and/or UV crosslinking to see if a 0Å crosslinker increased the amount of purified RNA. The samples were treated with DMSO or 1 mM BM(PEG)₂ to produce SAF-A oligomer-enriched samples. FLAG-SAF-A was expressed in the cells (after Dox treatment) and immunoprecipitated using magnetic beads. RNA was extracted from the input (pre-beads) and elute (post-beads) using phenol:chloroform after proteinase K treatment. My first attempt showed genomic DNA contamination on a formaldehyde agarose gel (Figure 3.9A) as high MW smears in the gel wells. The experiment was repeated with an acid phenol RNA isolation to better separate the RNA from the DNA. This yielded lower levels of RNA, which were not visible on a formaldehyde agarose gel, so the samples were visualised using a TapeStation (Figure 3.9B). Whilst these methods did not readily allow the amounts of immunoprecipitated RNA to be compared, they did show that both monomeric and oligomeric SAF-A can be immunoprecipitated with RNA and subsequently isolated for other experiments. This technique was used in tandem with [5-³H]U radioactive RNA labelling (as explained below) to investigate if monomeric or oligomeric SAF-A is preferentially bound to nascent or total RNA during different cell cycle phases.

A



B

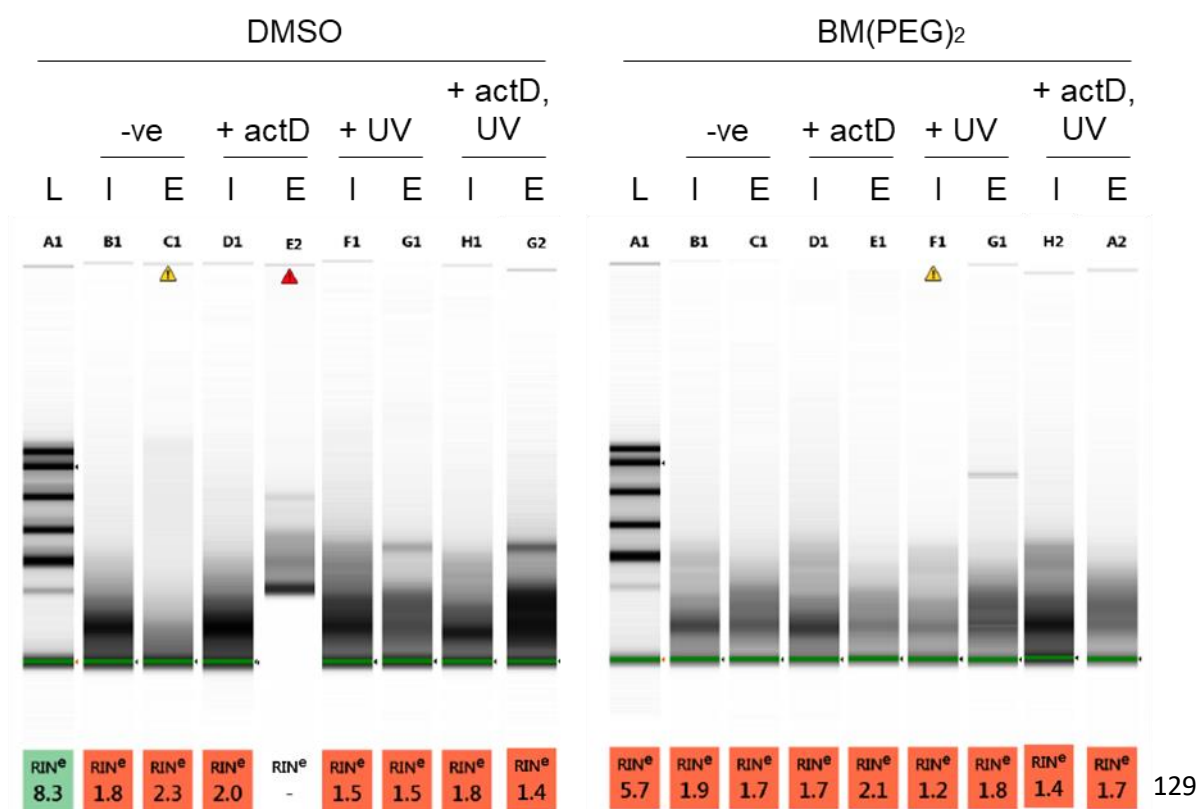
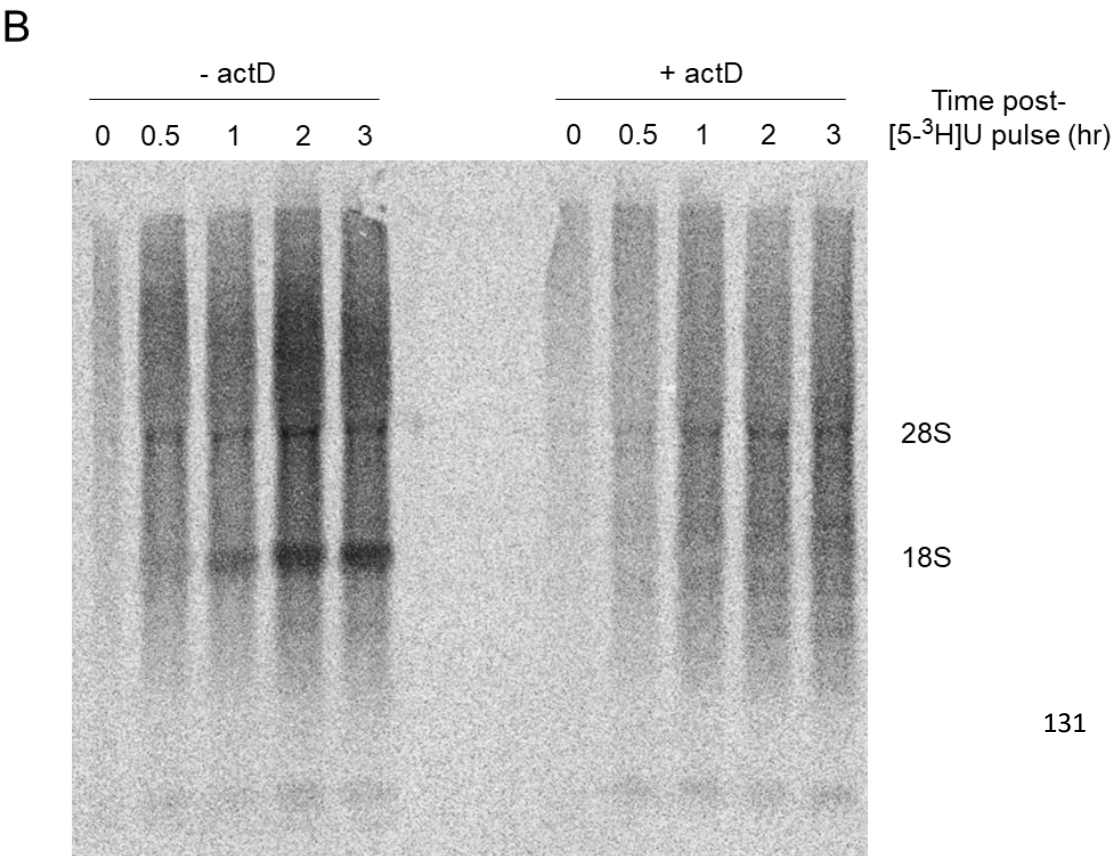
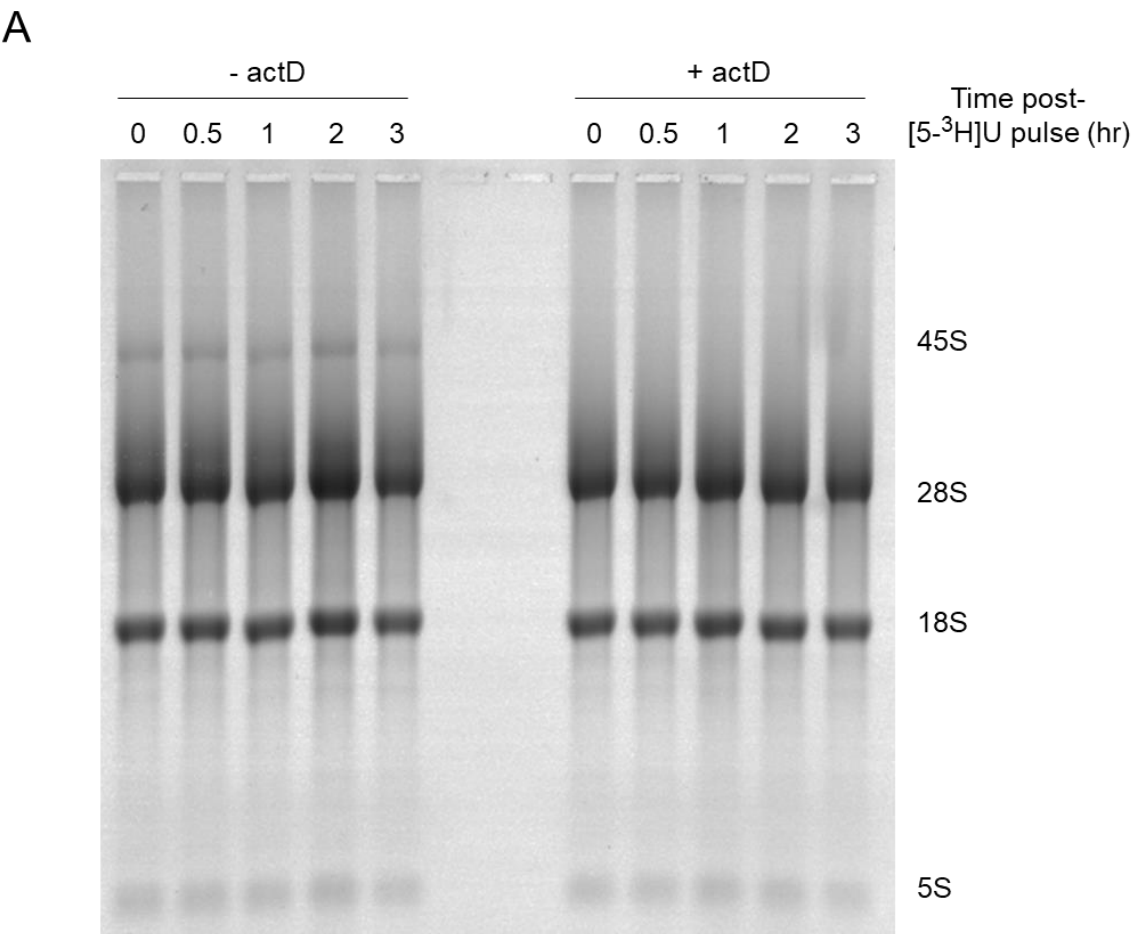


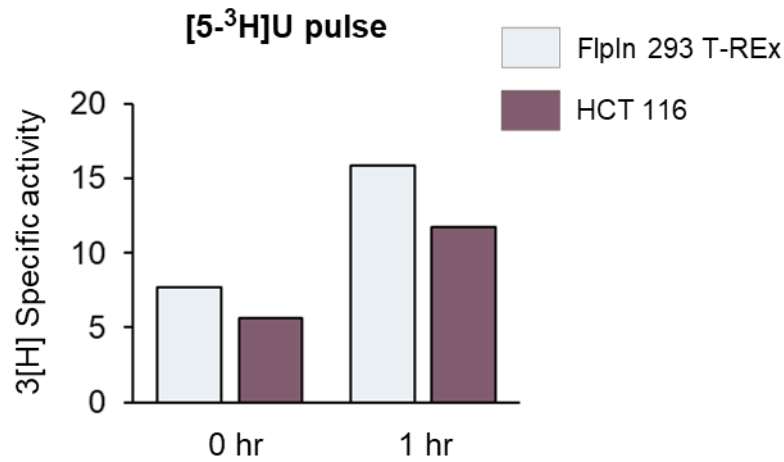
Figure 3.9. SAF-A RNA immunoprecipitation. Agarose gel (A) or TapeStation analysis (B) of purified RNA from FlpIn 293 T-REx cells expressing Dox-inducible FLAG-SAF-A which were treated with either actD or UV, both or none (-ve). Samples were then treated with DMSO or crosslinked with BM(PEG)₂. Input samples (I) correspond to purified RNA before FLAG-SAF-A immunoprecipitation. Elute samples (E) were purified RNA extracted from immunoprecipitated FLAG-SAF-A using anti-FLAG magnetic beads.

To analyse the amount of bound RNA to oligomeric and monomeric SAF-A, I first tested [5-³H]U incorporation into RNA. HCT 116 cells were treated with radioactively labelled uridine for 15 min and after 0 to 3 hr cells were lysed and total RNA was purified. Some cell samples were also treated with low-dose actD before [5-³H]U incubation to stop radioactive uridine incorporation into rRNA. Total RNA was separated on a formaldehyde agarose gel (Figure 3.10A) which showed different MW smears of RNA with the strongest signal coming from structural rRNAs including 28S, 18S, and 5S. Samples treated with actD had a clear absence of the 45S nucleolar precursors of 28S and 18S rRNA. The RNA from the gel was transferred onto a membrane and exposed on a phosphor screen to capture the ionising radiation of the samples and imaged on a FUJI FLA-5100 phosphoimager (Figure 3.10B). Most of the radioactive signal accumulated in 28S and 18S structural rRNA over time. In the presence of actD there was no 18S signal and a very weak 28S signal. As most nascent RNA produced is rRNA, low-dose actD treatment enabled me to instead focus on other RNA species in the cells and is used in subsequent immunoprecipitation, microscopy, and sequencing experiments. The amount of [5-³H]U incorporation in HCT 116 and FlpIn 293 T-REx cells was detected using a scintillation counter by Jon Stocks. Samples were produced by either incubating the cells with [5-³H]U for 15 min and isolating RNA immediately or 1 hr later (nascent RNA) or by incubating the cells overnight and isolating the RNA in the morning (total RNA). Each sample was read twice on a scintillation counter to give the average counts per minute (cpm) which was then normalised using the respective RNA concentration measured in the samples. The resulting specific activity of the samples' ionising radiation is presented as a bar plot (Figures 3.10C and 3.10D). For both nascent and total RNA, FlpIn 293 T-REx cells had more signal than HCT 116 cells. The difference was smaller for nascent RNA (Figure 3.10C) with FlpIn 293 T-REx cells having 1.4-fold more specific activity at 0 hr and 1 hr than HCT 116 cells. For total RNA (Figure 3.10D) FlpIn 293 T-REx cells had 5.6-fold more specific activity than HCT 116 cells. The total RNA-specific activity for HCT 116 cells was only 2-fold

more than the 1 hr signal which shows either a limit to incorporation or fast cell growth leading to confluency and reduced uptake of nucleotides.



C



D

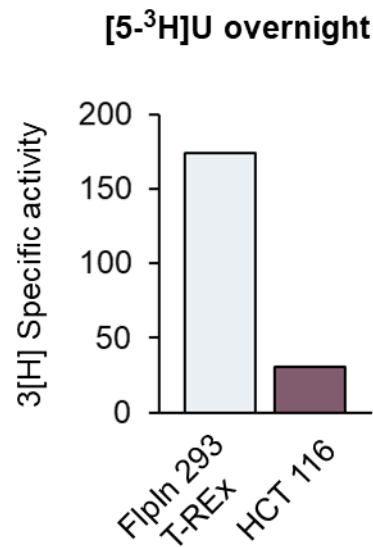


Figure 3.10. [5-³H]U-RNA incorporation. A) Formaldehyde gel of total purified RNA from HCT 116 cells incubated with [5-³H]U for 15 min, washed with PBS and harvested at 0, 0.5, 1, 2, or 3 hr. Cells were also treated with low-dose actD for 1 hr prior to pulse. B) Northern blot of agarose gel in A). C) Bar plot showing the [5-³H]U specific activity in total RNA collected from FlpIn 293 T-REx and HCT 116 cells. Samples were taken at 0 and 1 hr after 15 min [5-³H]U pulse. Each bar plot represents a sample size of N = 1. D) Same as C) but samples were taken after overnight incubation with [5-³H]U.

After testing [5-³H]U incorporation into RNA in HCT 116 cells I wanted to investigate SAF-A:RNA association when monomeric versus oligomeric during interphase. To achieve this, [5-³H]U labelling was incorporated into SAF-A monomer and oligomer immunoprecipitation (with BM(PEG)₂ crosslinking) and RNA isolation. HCT 116 cells were transfected with a construct containing FLAG-SAF-A or FLAG-SAF-A-ΔRGG (only amino acids 1 to 666, without RNA C-term domain) 24 hr before harvesting them. I hypothesised that the FLAG-SAF-A-ΔRGG construct would have less or no RNA binding overall than WT SAF-A due to absence of the RGG motif for RNA binding. Cells were treated with [5-³H]U overnight to label all RNA (total RNA) or treated with actD before a 15 min pulse of [5-³H]U (nascent RNA). Protein and RNA were crosslinked together using UV and cells were trypsinised and lightly permeabilised before treatment with DMSO or BM(PEG)₂ for oligomer crosslinking. Samples were sonicated to disrupt the nuclear mesh and increase protein pulldown efficiency. Immunoprecipitation of FLAG-tagged proteins was performed with magnetic beads coated in anti-FLAG antibodies and the protein was eluted with 3xFLAG peptide. RNAs were then purified from the input and elute samples. 5 µl of each sample was diluted in solvent and analysed on a scintillation counter as described previously. The ratio of the specific activity for the input and elute samples for each condition were presented as bar plots (Figure 3.11). This approach was used to take into account and normalise the specific activity of the input samples. Nascent RNA binding (Figure 3.11A) was surprisingly similar between the two proteins, with FLAG-SAF-A-ΔRGG having 1.3-fold more signal when monomeric and both having highly similar signal levels when more oligomeric protein is immunoprecipitated. For both constructs there was more RNA pulled down with monomeric SAF-A than with oligomeric (5.1-fold more for FLAG-SAF-A and 6.2-fold more for FLAG-SAF-A-ΔRGG) hinting that asynchronous interphase cells might have reduced transcriptional activity that drives SAF-A oligomer formation. The total RNA signal (Figure 3.11B) was more pronounced for monomeric FLAG-SAF-A by 1.5-fold compared to oligomeric and clearer for oligomeric FLAG-SAF-A-ΔRGG by 4.8-fold compared to monomeric. More RNA was immunoprecipitated with oligomeric SAF-A (2.4-fold more) in the presence of FLAG-SAF-A-ΔRGG compared with FLAG-SAF-A which was unexpected as the absence of the RGG motif should lead to reduced levels of protein oligomerisation (Nozawa *et al.*, 2017). Meanwhile, more RNA was isolated with monomeric SAF-A (3-fold more) in the presence of FLAG-SAF-A

compared with FLAG-SAF-A- Δ RGG. The results showed a higher signal for FLAG-SAF-A- Δ RGG compared to FLAG-SAF-A in some conditions, possibly due to being crosslinked with endogenous WT protein.

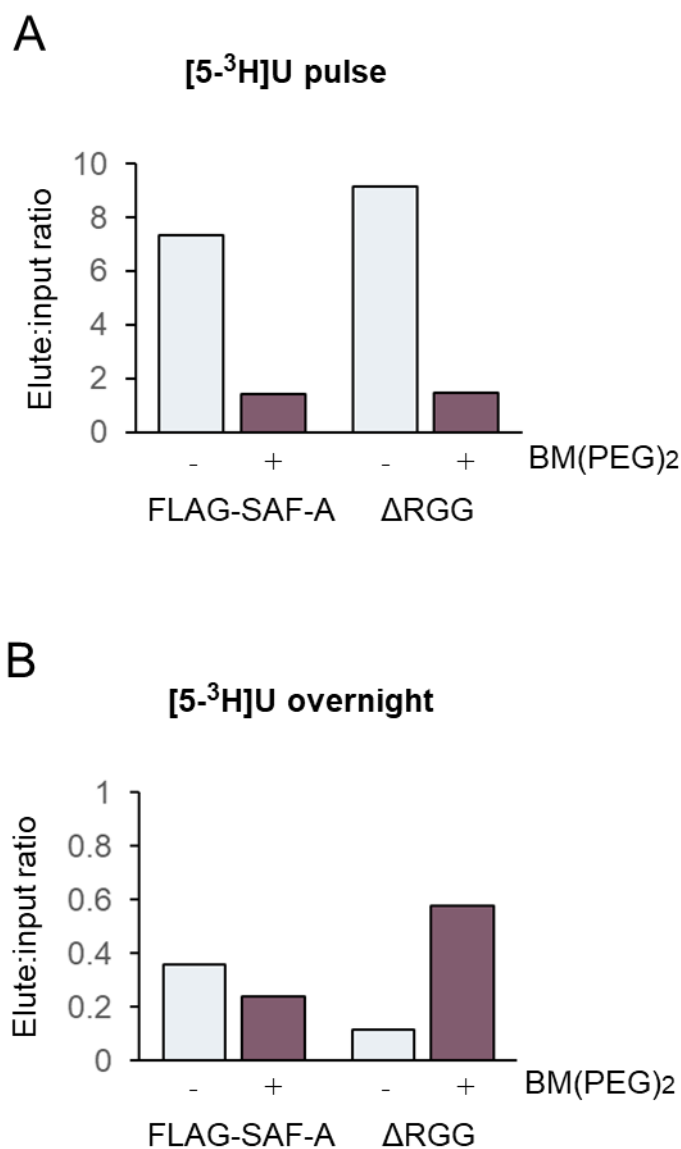


Figure 3.11. *In vivo* coupling of SAF-A with [5-³H]U-RNA. Bar plots show the ratio of elute [5-³H]U specific activity to input [5-³H]U specific activity. Input samples were purified RNA before immunoprecipitating FLAG-SAF-A. Elute samples are purified RNA from immunoprecipitated FLAG-SAF-A using magnetic beads. The specific activity was calculated from the measured ionising radiation divided by the RNA amount in each sample. DMSO-treated samples are in ice blue, BM(PEG)₂ crosslinked samples are in violet. Each bar plot represents a sample size of N = 1. A) Ratio of [5-³H]U specific activity after a 15 min pulse treatment with [5-³H]U of cell populations. B) Ratio of [5-³H]U specific activity after overnight treatment with [5-³H]U of cell populations.

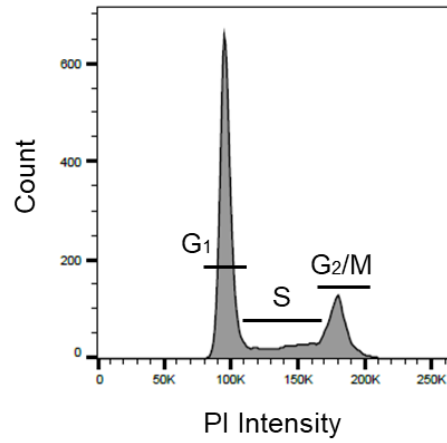
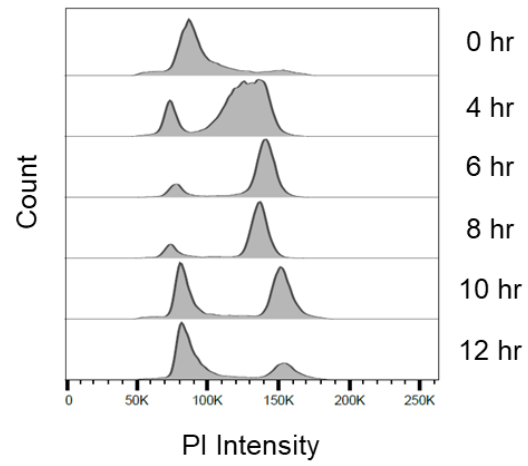
3.3.4. Cell synchronisation to obtain a pure mitotic cell population

After observing that SNAP-SAF-A begins to co-localise with DNA during cytokinesis at the end of cell division, I wanted to investigate whether SAF-A is interacting with nascent and/or total RNA during and after mitosis. As SAF-A is excluded from mitotic chromatin during most of mitosis and forms a cloud around the chromosomes, it could still be interacting with RNA. During the reformation of the nucleus in early G₁, SAF-A co-localises with the DNA signal, raising questions about its increased interaction with nascent RNA or its general RNA-binding activity, considering its recruitment to transcriptionally active chromatin. To check the status of RNA associated with SAF-A during these cell cycle phases, it would be important to have a high proportion of cells progressing through and exiting mitosis synchronously. To make a synchronised population, cells are often incubated with blocking agents that arrest cells at specific cell cycle stages such as nocodazole in pro-metaphase. I tested multiple previously reported cell synchronisation strategies with different blocking agents in three cell lines, to find out which method has the highest proportion of cells entering and exiting mitosis simultaneously (Table 2.17).

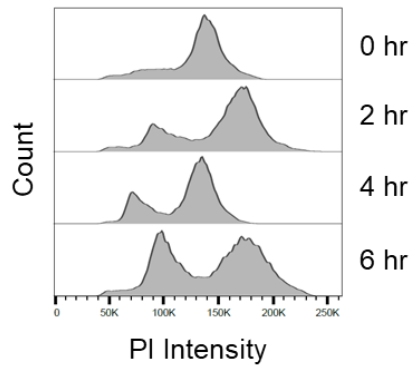
As HCT 116 cells have a shorter cell cycle compared to non-cancer cell lines, I tested arresting them at the end of the G₁ phase using a 2 mM double thymidine block (2xThy) and releasing them to follow their cell cycle progression. A 2xThy block inhibits DNA synthesis and two rounds of treatment are required to ensure cells are arrested at the end of the G₁ phase and not throughout the S phase (Bostock *et al.*, 1971). To check the cell cycle stage of a cell population, cells are collected at certain time points, usually every 1 to 2 hr after release from arrest and fixed in ethanol and PBS. Then they are treated with 1 mg/ml propidium iodide (PI) to visualise the DNA content in the cells and analysed via flow cytometry to determine the ratio of cells in each cell phase. After release from the 2xThy block, it takes around 8-10 hr for cells to enter mitosis, which they then exit completely after another 2-4 hr (Figure 3.12A). The flow cytometry traces of synchronised cells shown in Figures 3.12 and 3.14 are not perfectly aligned as different samples have incorporated different amounts of PI before analysis which moves the values to the left or right compared to the first sample. However, it does not detract from subsequent analysis as the peaks are still clearly visible. Other blocking reagents used with HCT 116 cells were nocodazole, an

inhibitor of microtubule formation, and RO-3306, a selective inhibitor of CDK1 (Ho *et al.*, 2001; Vassilev, 2006). Attempts to synchronise HCT 116 cells closer to mitosis after a single thymidine block by using 10 μ M RO-3306 to arrest cells at the end of G₂ or 25 ng/ml nocodazole for pro-metaphase arrest, resulted in a poor release where cells exited mitosis over 2 to 8 hr (RO-3306) or 2-4 hr (nocodazole) (Figure 3.12B and 3.12C). I also tested the synchronisation of RPE-1 cells with 15 μ M aphidicolin, which has been previously used in the lab to induce replication stress and arrests cells in the early S phase (Boteva, 2017, Boteva *et al.*, 2020). After release in the early S phase, the cells exited from mitosis 16 to 18 hr later but were not synchronised (Figure 3.12D).

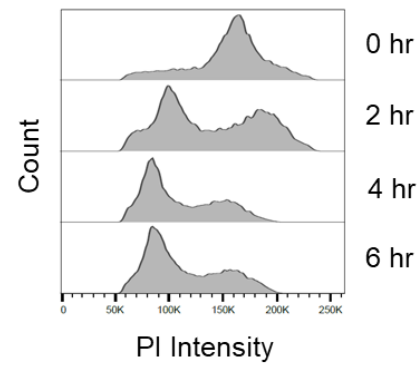
A

Time post
2xThy block

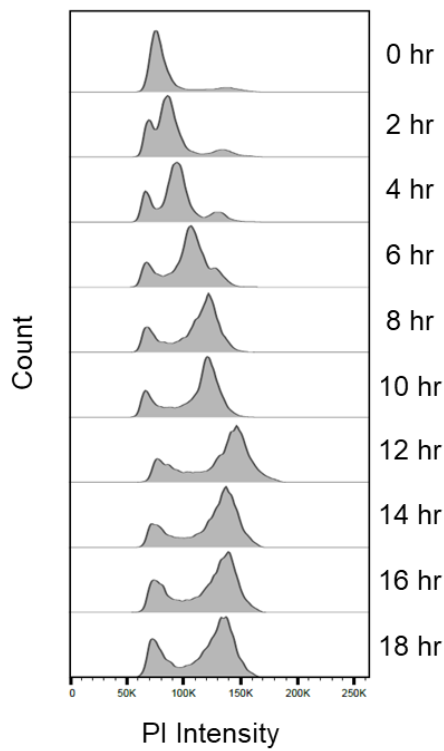
B

Time post
RO-3306 block

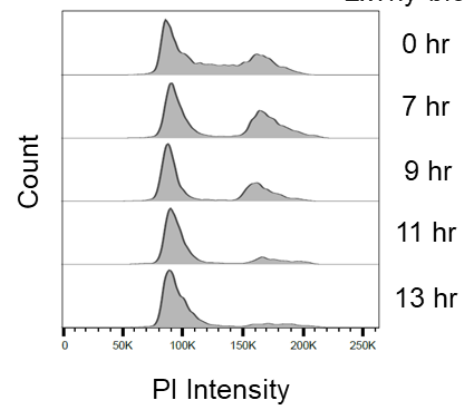
C

Time post
nocodazole block

D

Time post
aphidicolin block

E

Time post
2xThy block

F

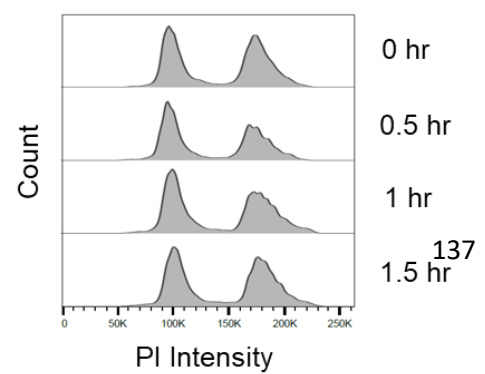
Time post
1NM-PP1 block

Figure 3.12. Cell synchronisation of different cell types. Methods used are described in Table 2.17. A) Cell cycle profile (left) of asynchronous HCT 116 cells showing manual gating for flow cytometry of synchronised cell populations. PI histograms of HCT 116 cells treated with a 2 mM thymidine block to synchronise at the G₁/S boundary and tracked with flow cytometry after release (right). Cell cycle progression was tracked at the shown intervals. B) PI histograms of HCT 116 cells treated with a 2 mM thymidine block followed by a 10 µM RO-3306 block to synchronise at the G₂/M boundary. C) PI histograms of HCT 116 cells treated with a 2 mM thymidine block followed by a 25 ng/ml nocodazole to synchronise during pro-metaphase. D) PI histograms of RPE-1 cells treated with 15 µM aphidicolin to synchronise during the early S phase. E) PI histograms of RPE-1 CDK1_{as} cells treated with 2.5 mM thymidine block to synchronise at the G₁/S boundary. F) PI histograms of RPE-1 CDK1_{as} cells treated with 5 µM 1NM-PP1 following a 2.5 mM thymidine block to synchronise cells at the G₂/M boundary.

Recently, another method to synchronise RPE-1 cells was established using an F80G mutation in the CDK1 gene to make it sensitive to the ATP analogue 1NM-PP1, also known as RPE-1 CDK1 analogue sensitive cells (RPE-1 CDK1_{as}; Hochegger, *et al.*, 2007; Salvidar *et al.*, 2018). These cells arrest at the end of G₂, similarly to RO-3306 and are reported to immediately enter mitosis upon washout of 1NM-PP1. Normally, 1NM-PP1 incubation is limited to 5 hr as the centrioles were reported to increase in number during longer incubation times. This was hypothesised to be too short a time for sufficient arrest for total population synchronisation, therefore I coupled a 5 μ M 1NM-PP1 treatment with a 2xThy block in RPE-1 CDK1_{as} cells. When the cells were arrested with a 2.5 mM 2xThy block, there was a large proportion of cells detected in the S and G₂/M phase signalling the block was not optimal (Figure 3.12E). A similar experiment with a 5 mM 2xThy block to increase the specificity of the block resulted in similar poor synchronicity (data not shown). Releasing RPE-1 CDK1_{as} cells for 7 hr after 2xThy block and treating them with 5 μ M 1NM-PP1 to enrich them at the end of G₂ resulted in poor synchronisation at G₂/M phase and slow progress through mitosis (Figure 3.12F). Therefore, RPE-1 CDK1_{as} were unsuitable for my cell synchronisation goals.

In addition to cell synchronisation methods, I explored the potential use of a mitotic degron (MD) tag to facilitate the FACS sorting of mitotic cells after synchronisation. The MD tag spans residues 13 to 91 of the mouse cyclin B1 gene (Zhang *et al.*, 2019). It is reported to be expressed in mitotic cells, with pro-metaphase cells having the strongest signal and a weaker signal being observed in anaphase and telophase. After creating constructs where the reported mouse MD (mMD) and the same region of the human cyclin B1 gene MD (hMD) were put downstream of fluorescent tags, EGFP and mCherry. HCT 116 cells were transiently transfected with the constructs and imaged to check the expression and localisation of the tagged MDs (Figure 3.13). The results showed that the mCherry-tagged MDs were predominantly localised in the cytoplasm rather than the nucleus and were excluded from mitotic chromosomes. Conversely, the EGFP-tagged MDs were expressed in both the cytoplasm and nucleus of interphase cells, contrary to the expected mitotic cell-specific expression pattern. As a result, I decided not to pursue further experimentation with these MD constructs due to their suboptimal performance in HCT 116 cells.

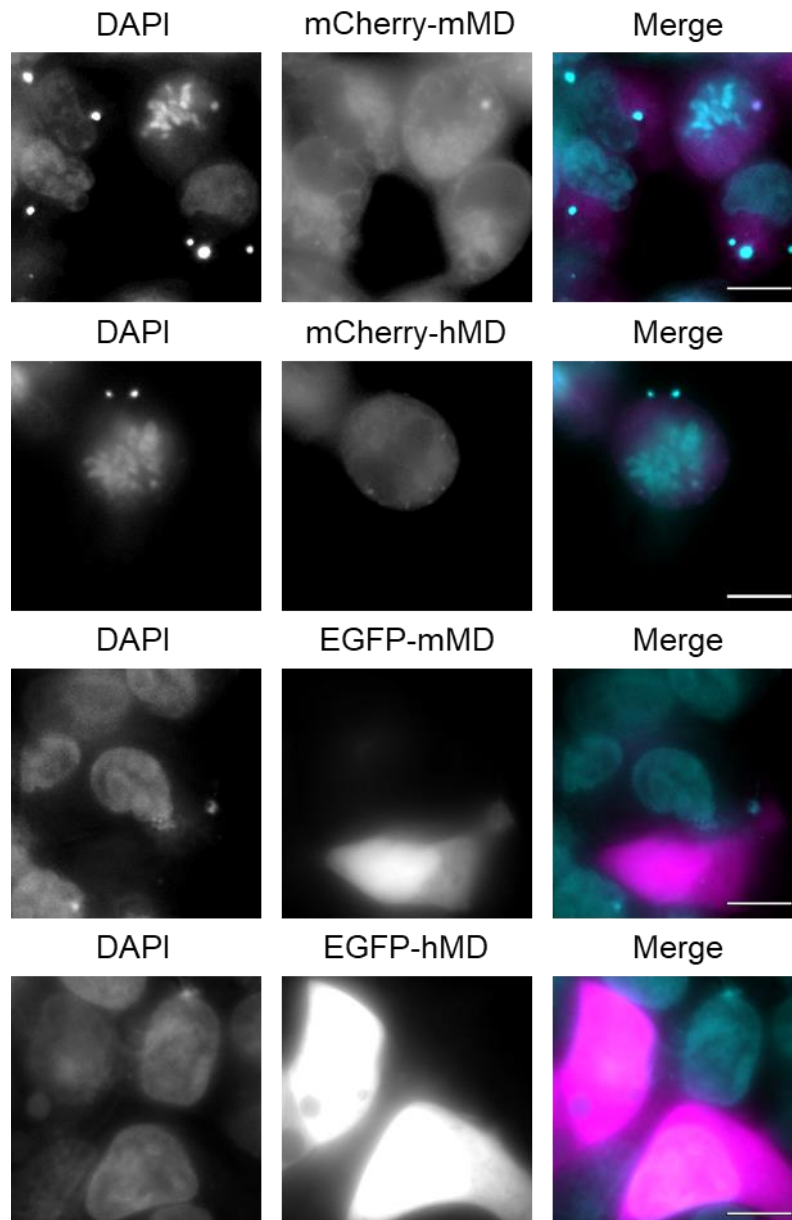


Figure 3.13. Testing of mitotic degron (MD) as a synchronisation aid. Fixed cell immunofluorescent staining of DNA and MD fluorescently tagged constructs. Cells were transfected with the mouse (mMD) or human (hMD) residues 13 to 91 of the cyclin B1 gene expressed downstream of a fluorescent tag (EGFP or mCherry). The images were taken the day after transfection. DNA was visualised with DAPI. Scale bar is 10 μ m.

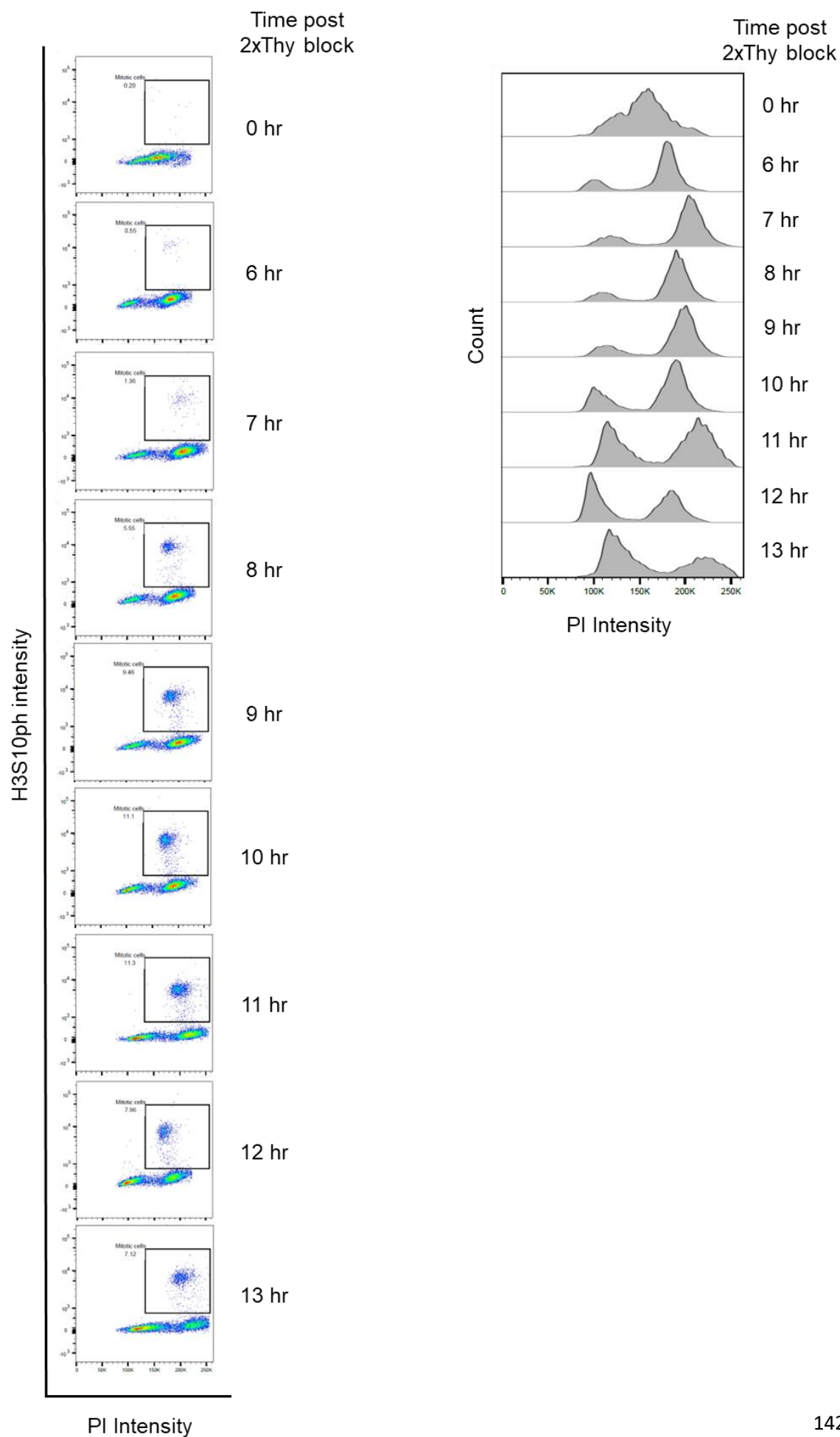
After these experiments with the MD tags were unsuccessful, I continued to use a 2xThy block in HCT 116 cells for cell synchronisation. I further analysed the time points at which HCT 116 cells enter and exit mitosis with FACS (Figure 3.14A). In addition to PI, cells were treated with an anti-H3S10ph antibody conjugated to Alexa

Fluor 488 fluorophore to mark cells in mitosis. Table 3.2 shows that cells were entering mitosis about 8 hr (5.6% cells with H3S10ph signal) after release from a 2xThy block with the highest concentration of mitotic cells seen at 10 and 11 hr after release (11.1% and 11.3% cells with H3S10ph signal respectively). By the 12-hr time point, more cells had exited mitosis than entered (8% of cells with H3S10ph signal). The mitotic cell signal accounted for only a fraction of the cells that had undergone DNA duplication and were in the G₂/M phase: 75.9% at 8 hr, 63.2% at 10 hr, and 45% at 12 hr. Consequently, this cell synchronisation protocol and the 0-hr, 10-hr, and 12-hr time points were selected for subsequent experiments focused on investigating the behaviour of SAF-A during mitosis and early G₁ (Figure 3.14B). Using this method, I was able to examine SAF-A protein levels and phosphorylation status of serine 59 (S59) from the end of G₁ through mitosis to the next G₁ phase over 12 hr. Western blotting with an anti-SAF-A antibody showed no change in the abundance of the protein over the cell cycle (Figure 3.14C). S59 on SAF-A was found to be phosphorylated during mitosis after probing with an anti-SAF-AS59ph antibody, consistent with another study (Douglas *et al.*, 2015). Notably, phosphorylation of SAF-A at S59 was also observed when cells were arrested in the G₁ phase and after 12 hr. The presence of mitotic cells in the latter sample was confirmed by probing for H3S10ph levels.

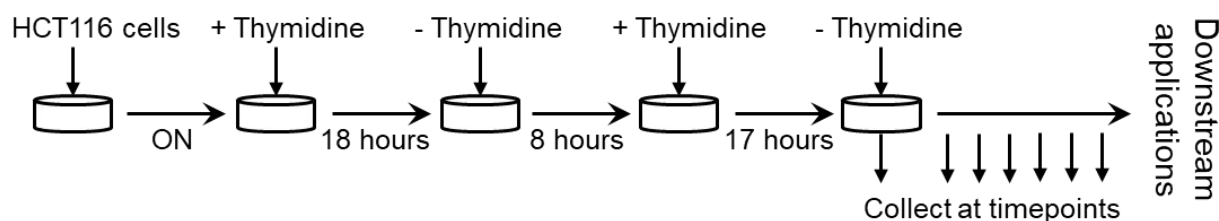
Time post 2xThy block release	% of cells in G2/M (PI signal)	% of cells in mitosis (H3S10ph AF488 signal)
0 hr	36.7%	0.2%
6 hr	70.5%	0.6%
7 hr	69.5%	1.4%
8 hr	75.9%	5.6%
9 hr	72.3%	9.5%
10 hr	63.2%	11.1%
11 hr	53.1%	11.3%
12 hr	45.3%	8%
13 hr	31.7%	7.1%

Table 3.2. Propidium iodide and H3S10ph signal after a 2xThy block release.

A



B



C

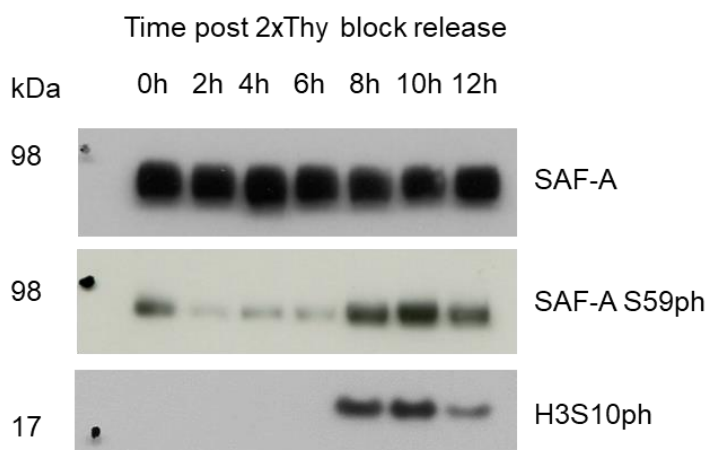
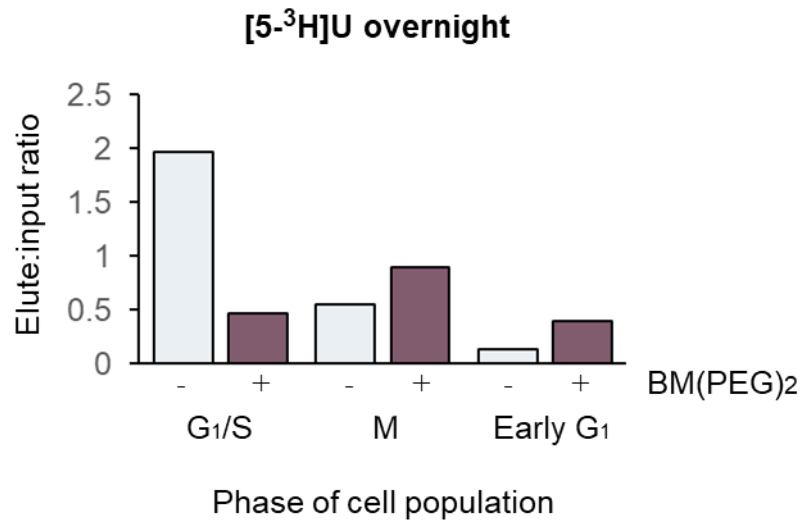


Figure 3.14. Cell synchronisation of the HCT 116 cell line with a double thymidine block. A) Flow cytometry analysis of cell populations following release from a double 2 mM thymidine block (2xThy) to synchronise at the G₁/S boundary. Samples were incubated with anti-H3S10ph antibody to identify dividing cells. The left graphs show H3S10ph intensity versus PI intensity (DNA content) enabling the identification of mitotic cells at different time points. The right graphs are PI histograms of the cell populations at different time points. B) Diagram of synchronisation protocol used in thesis experiments. HCT 116 were plated and grown overnight (ON) then treated with 2 mM thymidine for 18 hr to arrest cells during the S phase and at the G₁/S boundary. Cells were released for 8 hr to clear the S phase and then treated with the same dose of thymidine for 17 hr to arrest cells at the G₁/S boundary. After washing out the second thymidine dose, cell populations were collected at different timepoints depending on the experimental design and used for downstream experiments. C) Western blot of cells released from a 2xThy block and collected at the time points shown after release. Cell lysates were probed for SAF-A, SAF-A phosphorylated on Serine 59, and H3S10ph.

3.3.5. SAF-A and total RNA immunoprecipitation during mitosis and early G₁

After testing different cell synchronisation methods, I assessed whether SAF-A is associated with RNA at different cell cycle stages. I used the 2xThy block method to synchronise HCT 116 cells (Figure 3.14B) and transfected them with a construct containing FLAG-SAF-A 24 hr before collecting them. They were also treated with [5-³H]U overnight to radioactively mark RNA for subsequent analysis. Cells were collected at 0-hr, 10-hr, and 12-hr post-release from the block to obtain cell populations enriched at the G₁/S phase boundary, mitosis and early G₁ respectively. Before collection, the cells were UV crosslinked at 254 nm to fix the protein and RNA. They were then treated with trypsin and lightly permeabilised to increase the uptake of subsequent crosslinker treatment. Cells at each time point were treated with DMSO or 1 mM BM(PEG)₂ for 10 minutes to crosslink proteins at slightly further distance apart. Input samples were taken after sonication to disrupt the nuclear interactions and increase the efficiency of protein pulldown. Immunoprecipitation of FLAG-SAF-A was performed with magnetic beads coated in anti-FLAG antibodies, the protein was eluted with 3xFLAG peptide then RNAs were purified from the input and elute samples. After measuring the RNA concentration, 5 µl of each sample was diluted in solvent and then analysed on a scintillation counter to measure the incorporated radioactivity. The average cpm values were normalised using the RNA concentration of each sample to obtain the specific activity of the input and elute RNA (Figure 3.15A). In Figure 3.15A the highest ratio was seen for the DMSO-treated cells at the G₁/S phase boundary while the BM(PEG)₂-treated cells had a 4-fold smaller ratio. This shows that there is less RNA pulled down with oligomeric SAF-A at this time compared to monomeric. During mitosis and the early G₁ phase there is a 1.6-fold and 3-fold larger ratio of BM(PEG)₂-treated cells than with DMSO. That suggests there is more RNA pulled down with oligomeric SAF-A than with monomeric.

A



B

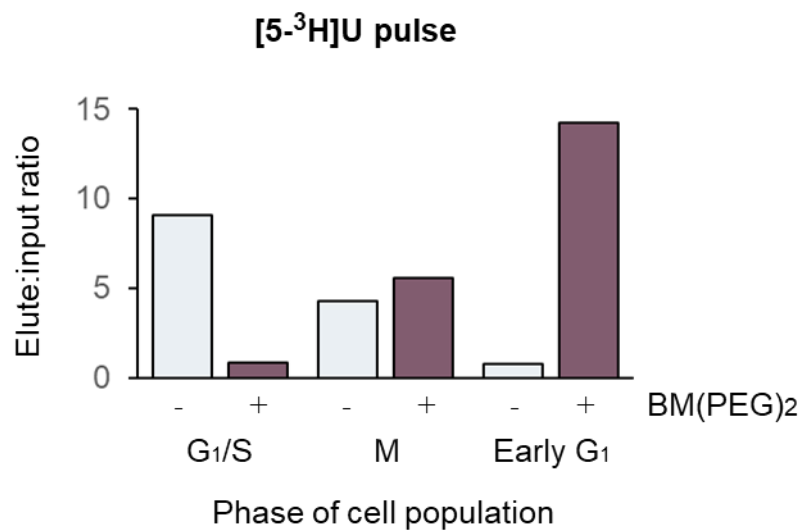


Figure 3.15. SAF-A coupled with [5-³H]U-RNA during mitosis and early G₁. Bar plots show the ratio of elute [5-³H]U specific activity to input [5-³H]U specific activity. Input samples were purified RNA before immunoprecipitating FLAG-SAF-A. Elute samples are purified RNA from immunoprecipitated FLAG-SAF-A using magnetic beads. The specific activity was calculated from the measured ionising radiation divided by the RNA amount in each sample. DMSO-treated samples are in ice blue, BM(PEG)₂ crosslinked samples are in violet. Cell populations were synchronised to the G₁/S phase, M phase, and early G₁ using a 2xThy block as described in Figure 3.9B. Each bar plot represents a sample size of N = 1. A) Ratio of [5-³H]U specific activity after overnight treatment with [5-³H]U of cell populations. B) Ratio of [5-³H]U specific activity after a 15 min pulse treatment with [5-³H]U of cell populations.

3.3.6. Interaction with newly transcribed RNAs during mitosis and early G₁

To see whether SAF-A interacts with nascent RNA a similar experiment as the previous paragraph was conducted except that cells were treated with [5-³H]U 15 minutes before UV crosslinking to mark only newly transcribed RNA. Prior to collection, cells were also treated for an hour with a low dose of actD to inhibit rRNA synthesis. The rest of the method is as described above. Figure 3.15B shows the specific activity ratios between input and elute for each condition presented as a bar plot. DMSO-treated cells collected at the G₁/S phase boundary had 11-fold more signal than BM(PEG)₂-treated cells. Cells during mitosis had 1.3-fold more signal in the BM(PEG)₂-treated cells than the DMSO-treated ones. Similarly, cells in the early G₁ phase had 19-fold more signal in the former treatment. Therefore, the overall trend seen in Figure 3.10A where at the G₁/S phase boundary there is more RNA associated with monomeric SAF-A and then more RNA is associated with oligomeric SAF-A later in the cell cycle is reproduced in this experiment (Figure 3.15B).

3.4. Discussion

In this chapter, I aimed to investigate the timing of SAF-A in decompacting chromatin after cell division to further understand its role in large-scale chromatin organisation. I was able to image live cells progressing through mitosis after initial troubleshooting by using only red and far-red lasers coupled with L-ascorbic acid (Figure 3.2). Ultimately, the addition of L-ascorbic acid made it possible to image cells during mitosis likely due to its ability to suppress the effect of phototoxicity in cells (Harada *et al.*, 2023). As I did not do a photodamage assay (imaging of γH2A.X foci, measuring reactive oxygen species levels), I do not know if the imaged cells might have damage present, such as DNA double-stranded breaks. To achieve the aim of this chapter, I initially examined the co-localisation of SNAP-SAF-A with the DNA signal during mitosis using live cell imaging microscopy. The results revealed that the co-localisation begins as the daughter cells separate during cytokinesis (Figures 3.4B and 3.4D), which is consistent with findings from fixed cell imaging indicating the

exclusion of SAF-A from chromatin during mitosis (Sharp *et al.*, 2020, unpublished data). Furthermore, I investigated the effect of transcriptional inhibition during mitosis on the timing and degree of SAF-A and DNA co-localisation during the early G₁ phase (Figures 3.4C and 3.4F). The findings indicated that while the start of co-localisation was delayed, the degree of co-localisation remained largely unchanged. This suggests that the recruitment of SAF-A to decondensed chromatin is partly dependent on transcriptional activity, but ultimately occurs after the reassembly of the nuclear envelope. Whether additional expression of SAF-A or the addition of triptolide affected the progression through mitosis could not be determined because none of the cells were imaged from the start of prophase. However, all cells were imaged for the full length of anaphase, telophase, and cytokinesis which showed that increased expression of SAF-A or a 1 hr treatment with triptolide did not significantly affect the length of these mitotic phases (Table 3.1). From the live cell imaging (Figure 3.4B), there is a cloud of SNAP-SAF-A around the newly formed nucleus, and it is unknown whether it is reimported or whether new SAF-A will be made to compensate for the loss of protein. In comparison, recent research showed that SAF-A was completely localised in the nucleus after nuclear envelope reassembly but was cytoplasmic in early G₁ while the nuclear envelope was still reforming in DLD-1 cells (Blower *et al.*, 2023). It is important to note that this co-localisation between chromatin and SAF-A can be due to SAF-A being reimported to the daughter nucleus after nuclear envelope assembly but does not necessarily equal chromatin binding. As SAF-A seems to be actively excluded from mitotic chromosomes by phosphorylation at Ser 14 and 16 by Aurora-B kinase, removal of the marks could also be responsible for recruiting the protein into the nucleus independently from transcription by facilitating chromatin binding (Sharp *et al.*, 2020). It is currently unknown which phosphatase is responsible for the reversal of PTMs on SAF-A and the timing of events.

To validate the reliability of the SNAP-SAF-A results, I conducted live cell imaging of SNAP-HNRNPC1 and HNRNPA1 in the presence and absence of a transcriptional inhibitor. Additional expression of HNRNPC1 or HNRNPA1, with or without triptolide, did not significantly affect the length of anaphase, telophase, or cytokinesis (Table 3.1). There were potentially increased mitotic defects in some cells expressing SNAP-HNRNPA1, with or without triptolide, that prolonged their anaphase. However, more cells need to be imaged from those conditions, including control cells, to

investigate whether this anaphase prolongment is statistically significant. HNRNPC1 and HNRNPA1 are appropriate controls to use as they belong to the same protein family as SAF-A and they show different behaviour during and after cell division. The proteins also have key structural differences – SAF-A is the only HNRNP family member with a SPRY domain that might be important for RNA binding (Miles, 2022). SAF-A also has an RG/RGG repeat motif for RNA binding while HNRNPC1 has an RRM, and HNRNPA1 has 2 RRMs and an RG/RGG repeat motifs (The UniProt Consortium, 2023). My imaging confirmed the co-localisation with DNA for them to begin in telophase and early G₁ respectively which was as previously reported (Figures 3.5C, 3.5E, 3.6C, and 3.6E; Prasanth *et al.*, 2003). To investigate the effects of transcription inhibition on protein and DNA co-localisation I treated the cells with triptolide for a short time prior to imaging. Results were comparable with previous findings for HNRNPC1 but not for HNRNPA1 (Figures 3.5D, 3.5F, 3.6D, and 3.6F; Pinol-Roma and Dreyfuss, 1991). While HNRNPC1's delay of co-localisation with DNA after transcription inhibition has not been previously reported, there is no other change in behaviour, and it is highly co-localised with DNA in early G₁. HNRNPA1 was reported to not re-enter the nucleus after transcription inhibition which was not consistent with my results where the protein still co-localises with DNA after mitosis. However, these previous experiments were using fixed cell imaging of cells that had just exited mitosis and were in early G₁ which leads to possible differences in results when compared to live cell imaging. Firstly, cell fixation takes time to set in which makes it possible to miss short-lasting or transient protein-DNA or protein-protein interactions that are visible in live cells (Schmiedeberg *et al.*, 2009). Secondly, the precise timing of cell exit from mitosis is unknown, which may lead to missed co-localisation events between proteins and DNA that occur at later time points. For example, in my experiments, the SNAP-HNRNPC1 signal is highly co-localised with DNA 10 minutes after cell division finished compared to 20 minutes for SNAP-HNRNPA1. Finally, in my transcription inhibition experiments, I opted to use triptolide, a potent inhibitor of transcription initiation known for its rapid activity and efficacy at lower concentrations. This choice deviates from the commonly employed transcription elongation inhibitors, such as actD and DRB, which require higher concentrations and exhibit slower effects on transcription elongation. It is worth noting that different transcription inhibitors can exert diverse effects on protein behaviour, potentially leading to variations in experimental outcomes. Thus, it would

be intriguing to investigate the behaviour of SNAP-SAF-A, SNAP-HNRNPC1, and SNAP-HNRNPA1 in the presence of actD or DRB. By exploring the effects of these different inhibitors, we can gain a more comprehensive understanding of how transcription inhibition influences the behaviour of these proteins and their co-localisation with DNA.

In this chapter, I aimed to investigate the connection between SAF-A and RNA organisation by examining how SAF-A binds RNA. To achieve this, I sought to differentiate between RNAs bound to oligomeric SAF-A, which has been reported to decompact transcriptionally active chromatin, and RNAs bound to monomeric SAF-A (Nozawa *et al.*, 2017). To enable the extraction of oligomeric SAF-A under harsh conditions, I used the protein-protein crosslinker BM(PEG)₂, which covalently links SAF-A proteins, thereby stabilising the oligomers (Figures 3.7A and 3.8). Although DSS was also tested as a crosslinker, it did not yield a comparable level of oligomeric SAF-A signal as detected by Western blotting when compared to BM(PEG)₂ (Figure 4.7B). This difference in oligomer isolation could be attributed to BM(PEG)₂ crosslinking proteins at a larger distance than DSS (14.7Å compared to 11.4Å), indicating that SAF-A oligomers are not tightly bound protein complexes. Alternatively, the differential effectiveness of the crosslinkers could be due to their amino acid group specificity, with BM(PEG)₂ crosslinking Cys-Cys pairs and DSS crosslinking Lys-Lys pairs. Previous research on oligomer stabilisation has demonstrated that the optimal detection of oligomeric SAF-A signal is achieved with BM(PEG)₂ and N-γ-maleimidobutyryl-oxysuccinimide ester (GMBS), the latter crosslinking cysteines and lysines (Nozawa *et al.*, 2017). Therefore, although SAF-A has fewer cysteines than lysines (13 versus 70), crosslinking cysteines to other cysteines or lysines improves the stabilisation of protein oligomers.

To distinguish between nascent and total RNA, I combined SAF-A crosslinking, immunoprecipitation, and RNA purification with [5-³H]U radiolabelling (Figure 3.11). Initially, my plan was to use FlpIn 293 T-REx cells expressing FLAG-SAF-A or FLAG-SAF-A-ΔRGG to investigate the association of nascent and total RNA with monomeric or oligomeric SAF-A. However, due to timing constraints and cell line issues, I performed the experiment using HCT 116 cells and transfected FLAG-tagged SAF-A constructs. Fortunately, HCT 116 cells displayed similar levels of [5-³H]U incorporation to FlpIn 293 T-REx cells, making them a suitable alternative

(Figure 3.10). Surprisingly, after purifying the RNA associated with the FLAG-tagged SAF-A constructs, I observed a comparable signal between FLAG-SAF-A and FLAG-SAF-A- Δ RGG when cells were pulsed with [5-³H]U (Figure 3.11). I had expected FLAG-SAF-A- Δ RGG to exhibit less or no RNA association compared to WT SAF-A due to the absence of the RGG motif necessary for RNA binding. When FLAG-SAF-A- Δ RGG was previously immunoprecipitated after a 15 min pulse labelling of RNA with 5-EU, it did not elute together with RNA unlike WT SAF-A (Nozawa *et al.*, 2017). One possible explanation is that FLAG-SAF-A- Δ RGG interacts with endogenous SAF-A and thus remains in proximity to RNA. Another possibility is that SAF-A interacts with RNA through alternative domains, such as the SPRY domain (Miles, 2022). Both FLAG-tagged constructs displayed a higher radiolabelled RNA signal for monomeric SAF-A compared to oligomeric SAF-A (BM(PEG)₂-crosslinked cells), and the signal levels were similar to those observed in cells arrested at the G₁/S phase boundary (Figure 3.15B). These results suggest that the RNA formed during a 15 min pulse of [5-³H]U may be too early to interact with oligomeric SAF-A or induce its oligomerisation. Monomeric SAF-A likely interacts with RNA before forming oligomers, a process that may require more time than the 15 min pulse. Super-resolution microscopy revealed prominent co-localisation of SAF-A and RNA 30 min after 5-ethynyl uridine (5-EU) incubation (Figures 4.10D and 4.10E), indicating that more oligomeric SAF-A may interact with nascent RNA at that time. In contrast, 5-EU pulse labelling experiments showed that immunoprecipitated SAF-A displayed the lowest amount of bound RNA immediately after labelling and the highest amount 3 hr after labelling (Nozawa *et al.*, 2017). Therefore, to investigate the timing of SAF-A oligomer formation, it would be beneficial to repeat this experiment with different time points after initial radiolabel incubation. For future experiments, it is advisable to use FlpIn 293 T-REx cells as they incorporate [5-³H]U more efficiently compared to HCT 116 cells (Figure 3.10C) and do not require protein transfection that might lead to protein overexpression. Additionally, a better negative control for this experiment could involve using a construct with the Walker A and Walker B mutations which exclusively produce monomeric or oligomeric SAF-A due to reduced ATP binding or hydrolysis, respectively (Nozawa *et al.*, 2017). Finally, when HCT 116 cells were radiolabelled overnight to label total RNA, similar signal intensities were observed between monomeric and oligomeric SAF-A (FLAG-SAF-A), or there was more signal with oligomeric SAF-A (FLAG-SAF-A- Δ RGG) (Figure 3.11B). The eluted samples

exhibited less signal compared to the input, which may indicate unsuccessful elution or immunoprecipitation. Therefore, to draw conclusive remarks about these results, this experiment needs to be repeated with a different cell line and negative controls.

After I showed that SAF-A co-localises with DNA at the mitotic exit, I synchronised HCT 116 cells to enrich them in mitosis and early G₁ and check whether SAF-A is associated with RNA during those cell cycle phases. It is difficult to achieve a high degree of synchronisation but I observed that close to 20% of cells with doubled DNA also had an H3S10ph signal corresponding to mitosis at 10 hr post 2xThy release (Table 3.2). The real number of cells in mitosis might be higher as the multiple wash steps required for fixing cells before FACS analysis might wash out the loosely attached mitotic cells. Therefore, I used the 10 hr post-release time point for looking at mitotic cells. The percentage of cells in the G₂/M phase was reduced by a third and mitotic cells by a fourth at 12 hr compared to 10 hr post 2xThy release and used that time point for looking at cells in the early G₁ phase (Table 3.2). The 13-hr time point had an increased reduction of G₂/M phase cells (half) and mitotic cells (a third) compared to the 10 hr time point. However, due to the long nature of these experiments and the required downstream processing, I selected the 12-hr time point instead. When using these time points coupled with crosslinking, [5-³H]U incorporation into RNA, SAF-A pulldown and RNA purification, I then looked at the association between SAF-A and RNA during the cell cycle (Figure 3.15). The cells arrested at the G₁/S phase boundary had the highest RNA association for monomeric SAF-A compared to oligomeric SAF-A (BM(PEG)₂-crosslinked cells) likely due to the block preventing entry into the S phase. In asynchronous cells, there is a wave of transcription that peaks at the G₁/S phase boundary and is repressed during the S phase (Bertoli *et al.*, 2013). Blocking cell cycle progression at this stage might lead to aberrant transcription expression or repression and it does not reflect the actual transcriptional pattern in either phase (Cooper, 1997). However, the result for cells at the G₁/S phase boundary was similar to that of asynchronous cells (which are generally in G₁ phase as seen in Figure 3.12A) in interphase where there was a higher RNA association for monomeric SAF-A instead of oligomeric SAF-A showing that this might be a cell-type specific pattern (Figure 3.11). Cells enriched for mitotic signal had similar amounts of both monomeric and oligomeric SAF-A-RNA association, albeit it is not known whether this is due to contaminant cells in either G₁

or G₂ phase. Therefore, it is difficult to conclude with certainty whether SAF-A is associated with RNA using my experimental method, however SAF-A association with RNA during mitosis is highly likely as a recent experiment with BrU-labeled RNA and a proximity ligation assay has shown (Sharp *et al.*, 2020). Finally, in early G₁ there is a higher signal from oligomeric SAF-A compared to monomeric which corresponds to the increase in transcription at that cell cycle stage. These results are currently preliminary and should be repeated if a better cell synchronisation technique is developed or with experiments that use less cell material, and include a better negative control where transcription is inhibited and a positive control of asynchronous cells.

Cell synchronisation was also useful for checking the abundance of SAF-A during the cell cycle (Figure 3.14C). The protein abundance remains the same in HCT 116 cells as PTMs (S59 phosphorylation) change during the cell cycle. However, in chicken DT40 cells that are synchronised in prometaphase, released, then fixed, lysed, and run on a mass spectrometer, the SAF-A signal was halved after mitosis showing that it is equally inherited by the daughter cells (Samejima *et al.*, 2022). These differences could be due to cell type but SAF-A is an abundant ubiquitously expressed and essential protein which hints that its behaviour should be similar in most cells. Furthermore, the chicken and human SAF-A protein sequences are highly identical (81%) when checked using BLASTp (NCBI Resource Coordinators, 2018). As my cell synchronisation is not complete, there is probably a mixed cell population in the probed lysates that makes it difficult to see many clear differences. Ideally, I would have liked to be able to sort cells in mitosis to achieve an enriched mitotic population, however, attempts using MD as a synchronisation aid were unsuccessful (Figure 3.13). Another method I could have utilised was mitotic shake-off where dishes with synchronised cells are shaken to detach cells in mitosis and isolated (Fox *et al.*, 1987). However, while this technique is useful for microscopy or protocols where a low number of cells is enough, it was not suitable for my proteomic and transcriptomic experiments.

Overall, in this chapter, I showed using live cell imaging that SNAP-SAF-A co-localises with chromatin during cytokinesis in a manner partly dependent on transcription. However, it is unclear whether this represents true chromatin or RNA binding or just the nuclear import of SAF-A after nuclear envelope formation. I also

examined SAF-A's association with RNA at the mitotic exit and early G₁ phase using protein-protein crosslinkers and radiolabelling techniques. SAF-A was associated with nascent and total RNA in both monomeric and oligomeric forms, with higher association in the monomeric form at the G₁/S phase boundary. During mitosis and the early G₁ phase, there was more oligomeric SAF-A associated with nascent and total RNA. This shows that SAF-A binds to RNA during and after mitosis to facilitate large-scale chromatin organisation. The interactions between SAF-A and nascent RNA are further explored in Chapter 4.

4. Understanding the nature of SAF-A's interaction with RNA

4.1. Introduction

The interaction between SAF-A and RNA plays a crucial role in organizing RNA within cells. This ranges from SAF-A anchoring XIST to the Xi in mice to being recruited to transcriptionally active chromatin regions by ncRNA (Hasegawa *et al.*, 2010; Nozawa *et al.*, 2017; Fan *et al.*, 2018; Creamer *et al.*, 2021; Skalska *et al.*, 2021). Previous CLIP-seq experiments have shown that SAF-A does not exhibit a specific preference for RNA species and binds to the most abundant type, usually introns (Xiao *et al.*, 2012; Huelga *et al.*, 2012; Yugami *et al.*, 2020; Van Nostrand *et al.*, 2020). SAF-A also has no consensus sequence for binding to RNA through its RGG motif which explains its ability to non-specifically bind RNA, albeit weakly (Thandapani *et al.*, 2013). However, the mechanism by which SAF-A regulates chromatin organisation in an RNA-dependent manner remains unclear. To investigate the effect of SAF-A depletion on RNA organisation, I set out to establish a rapid method that could be used instead of RNAi for depleting SAF-A from cells. Previously, 24-hr RNAi treatment to deplete SAF-A was found to have no off-target effects but after 48 to 72 hr of RNAi there were transcriptional changes as measured by RNAseq (Nozawa *et al.*, 2017). For experiments such as examining the effect of depleting SAF-A on chromatin structure, 24 hr is a sufficient period for RNAi treatment. However, more complex experiments like expressing a mutant variant of SAF-A and synchronising cells often require longer RNAi treatment times. To improve upon RNAi, I utilised a degron tag that targets the protein for degradation and often works more rapidly than the targeted depletion of mRNA. The lab previously made a mini-Auxin-Inducible Degron (mAID) tagged SAF-A construct that I further developed and introduced into the HCT 116 Tet-OsTIR1 cell line using chromosome engineering with the CRISPR-Cas9 system (Natsume *et al.*, 2016).

RNA and RNA binding proteins (RBPs) are essential components of the dynamic nuclear mesh and play a vital role in chromatin organisation (see Chapters 1.1.1.3. and 1.1.1.4.). Previous studies have indicated that depleting SAF-A does not appear to affect the quantity of RNA in the nucleus (Nozawa *et al.*, 2017). I sought to

investigate if SAF-A depletion impacts the RNA mesh by altering caRNA distribution or its organisation. To explore this, I visualised nascent caRNA using 5-ethynyl uridine (5-EU) in cells transfected with siRNAs targeting SAF-A mRNA for degradation. Furthermore, I utilised Direct Stochastic Optical Reconstruction Microscopy (dSTORM), a super-resolution visualisation technique, to examine the distribution of SAF-A and RNA within cells. dSTORM enables the observation of nuclear proteins beyond the diffraction limit and using a newly developed method for analysing the data we were able to show SAF-A can form interconnected protein clusters (Marenda *et al.*, 2021). This collaborative study involved cell preparation for dSTORM and microscopy performed by me, with SuperStructure method development conducted by Mattia Marenda and Davide Michieletto, and assistance from Sebastian van der Linde and Nick Gilber. We then employed super-resolution imaging, coupled with RNAi to knock down SAF-A expression or triptolide to inhibit transcription, to investigate the relationship between SAF-A availability, transcription, and nascent RNA organisation. The subsequent dSTORM images were the result of a joint effort between me, Mattia Marenda, and Davide Michieletto with cluster analysis done by Mattia Marenda using SuperStructure.

Another method I developed to investigate the kinetics of nascent RNA in cells involved thiol(SH)-linked alkylation for the metabolic sequencing of RNA (SLAM-seq) using short 4-thiouridine (4sU) incubations, in contrast to other studies which used several hours to a day long incubation (Herzog *et al.*, 2017). 4sU incorporation into RNA was assayed by purifying RNA from cells, degrading to single nucleotides and then analysing the samples by HPLC columns (performed by Bernie Ramsahoye). An improvement to the published method for analysing RNA kinetics involved deep sequencing of exons and introns from highly expressed genes, instead of low-depth whole genome sequencing. This allowed us to calculate the lifetime of nascent intronic RNA using a Poisson mixture model. This collaborative study involved me performing all laboratory experiments to generate the final sequencing library, with bioinformatic analysis undertaken by Graeme Grimes and Nick Gilbert with statistical modelling conducted by Hannes Becher.

4.2. Methods for SAF-A depletion

4.2.1. Generating a mini-Auxin-Inducible Degron (mAID) tagged SAF-A cell line

To achieve faster protein degradation for time-sensitive experiments, a popular approach is to attach a degron tag onto a protein of interest. The lab started trialling the mAID tag construct which I further developed into a functional transient transfection construct. This tag is a 68 aa fragment of the full-length auxin-inducible degron (AID) tag, which allows for rapid protein degradation in the presence of auxin (or synthetic indole-3-acetic acid, IAA) and OsTIR1 protein (Figure 4.1A; Natsume *et al.*, 2016). First, I transfected HCT 116 cells with a construct with the mAID tag connected to the N-terminus of SAF-A and the OsTIR1 gene separated by a P2A site. I treated the cells with 500 μ M IAA for 24 hr and probed the cell lysates with anti-FLAG and anti-GAPDH antibodies to check the availability of the proteins on a Western blot (Figure 4.1B). Bands were visible of the predicted ~100 kDa size for FLAG-mAID-SAF-A and they were responsive to IAA treatment. When the Western blot signal of FLAG-mAID-SAF-A was normalised to the GAPDH signal and plotted as a bar graph, it was visible that the protein was 80% degraded over a 10-hr period which was over twice as fast as by standard RNAi (Figure 4.6C). The later time points of 12, 18, and 24 hr had higher relative intensity compared to the low point of 10 hr, possibly due to FLAG-mAID-SAF-A overexpression overworking the proteasomal degradation system or cell growth with no access to IAA.

After confirming that the FLAG-mAID-SAF-A construct was functional, I wanted to generate a cell system where SAF-A could be depleted to see the state of the chromatin at different points of the cell cycle, as well as to test mutated versions of the protein. To do so, I stably transfected the HCT 116 Tet-OsTIR1 cell line with a homology donor plasmid to tag the N-terminus of the endogenous SAF-A gene with 3xFLAG and mAID and to insert the hygromycin B (HygB) antibiotic resistance gene separated by a P2A site (following the protocol in Yesbolatova *et al.*, 2019; Figure 4.2A). The transfection also contained 3 gRNAs targeting the ATG start codon of SAF-A in HCT 116 to direct the Cas9 enzyme to cut the DNA sequence in that area.

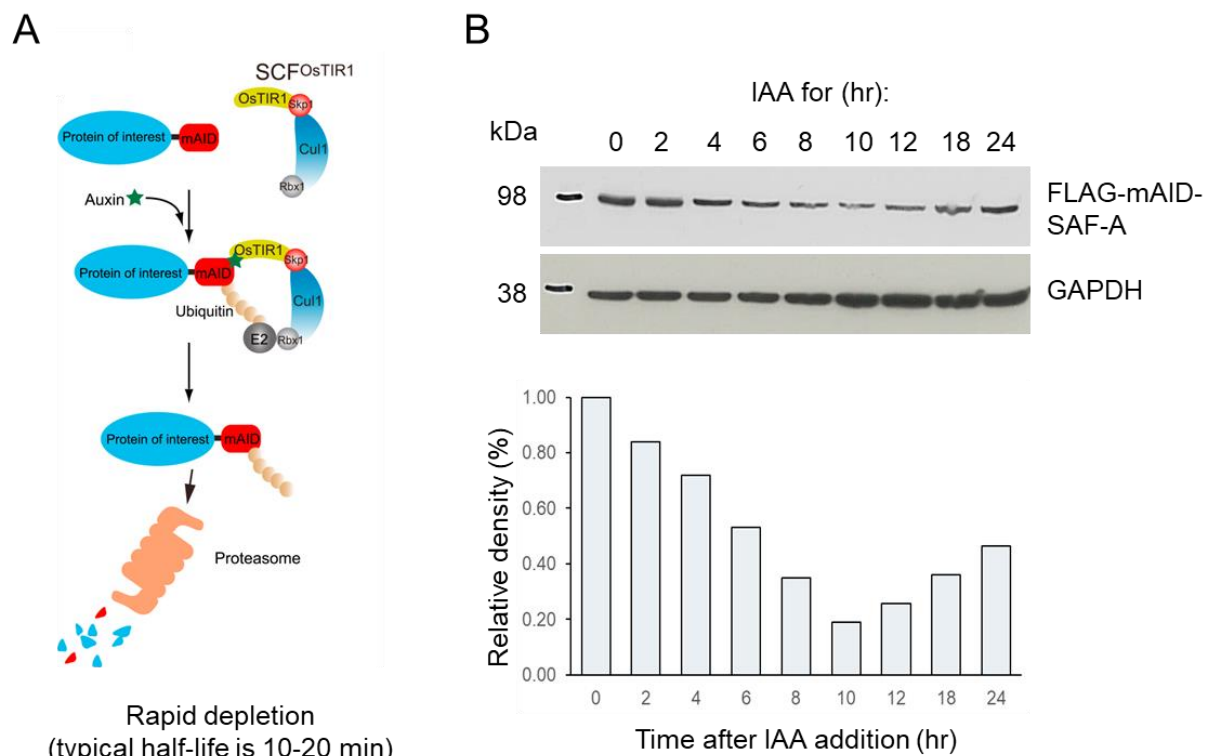
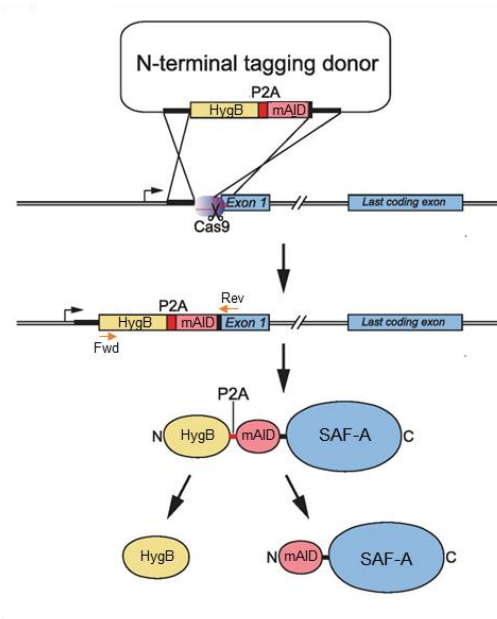


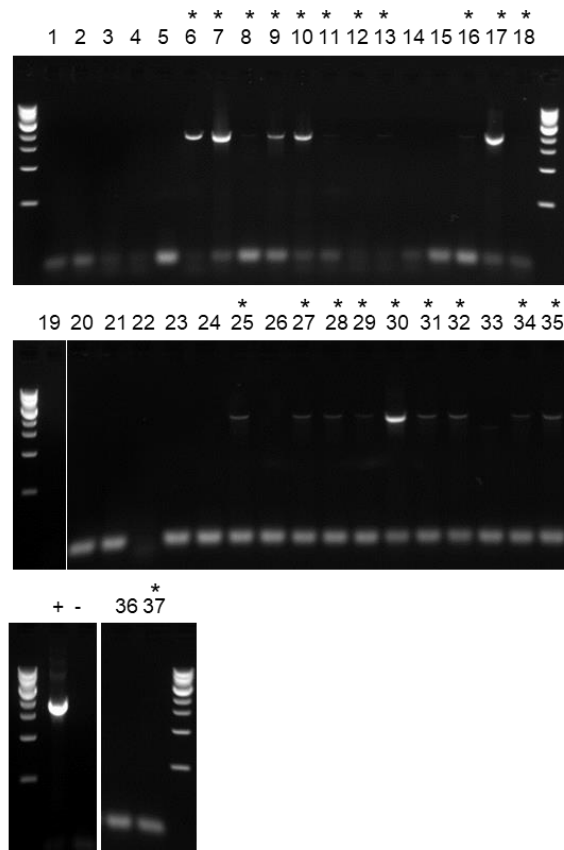
Figure 4.1. Depletion of transiently transfected FLAG-mAID-SAF-A in HCT 116 cells. A) Schematic illustration of the function of the mAID degron system. Taken from Natsume *et al.*, 2016. B) Western blot (top) of HCT 116 cell lysates transfected with FLAG-mAID-SAF-A and treated with 500 μ M IAA for the above times. Lysates were probed with anti-FLAG and anti-GAPDH antibodies. Bar plot (bottom) of the intensity of FLAG-mAID-SAF-A bands normalised to GAPDH intensity to show the relative density of Western blot bands.

In that way I could select cells that successfully underwent homologous recombination and contained the tagged protein without expressing a protein fused to a selection marker. After culturing cells in antibiotics for two weeks, I picked 37 colonies, out of which 21 had a band of the expected 2500 bp size (Figure 4.2B). These clones were sequenced using Sanger sequencing to confirm the in-frame insertion of HygB, P2A, 3xFLAG and mAID. The clones were then defrosted, 6 of them did not recover, but the remaining 15 were grown in the presence of IAA for 24 hr to check for successful degradation as well as tag insertion (Figure 4.2C). 12 clones showed a positive signal for FLAG and were further checked whether the tag was found on one or both alleles (Figure 4.2D) with clone 10 as a negative control. The heterozygous clones (7, 9, 12, 13, 18, 29, 31, and 37) had a stronger endogenous SAF-A signal and a very faint band of FLAG-mAID-SAF-A above it that was difficult to see due to the small kDa difference (arrows in Figure 4.2D). The homozygous clones (clones 8, 25, and 35) also had a weaker signal. Clone 8 had a higher band size than expected and was not responsive to IAA treatment. Clone 11 had no visible signal in Figure 4.2D but due to its FLAG signal in Figure 4.2C it was hypothesised to be a homozygous clone.

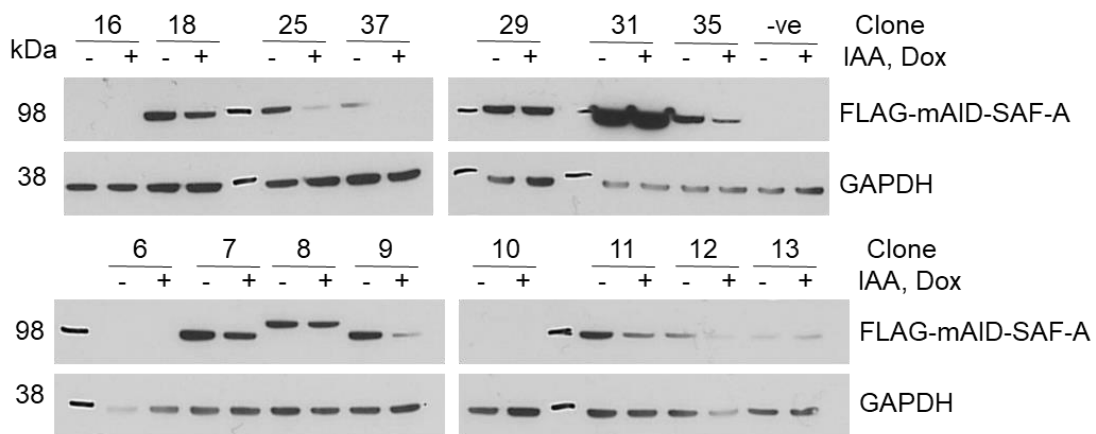
A



B



C



D

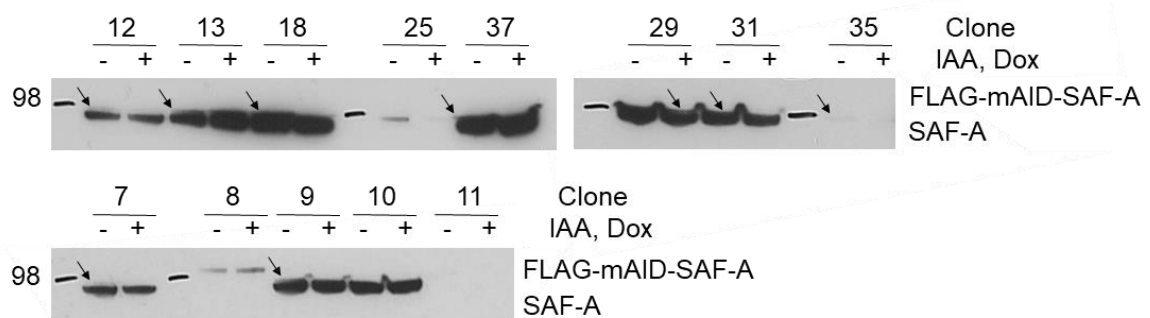


Figure 4.2. Tagging endogenous SAF-A with mAID in HCT 116 Tet-OsTIR1 cells. A) Schematic illustration of the method used to tag the N-terminus of endogenous SAF-A. Orange arrows show forward (Fwd) and reverse (Rev) primers used in B). Adapted from Yesbolatova *et al.*, 2019. B) Agarose gels of standard PCR for detection of mAID tag insertion in endogenous SAF-A in clones grown under Hygromycin B antibiotic selection. The expected band size of successful insertion (marked with *) is 2500 bp. The ladder used is NEB 1kb Plus. C) Western blots with inserts with and without culturing in 500 μ M IAA and doxycycline for 24 hours. Probed with anti-FLAG antibodies. GAPDH was used as a loading control. D) Same as C) except that the blot was probed with only anti-FLAG and anti-SAF-A antibodies. Black arrows show FLAG-mAID-SAF-A signal.

As SAF-A is an essential protein for cells, prolonged treatment with IAA and subsequent degradation of the protein was hypothesised to affect cell viability. To test the requirement of a homozygous clone (clone 25) for SAF-A, it was cultured in the presence of Dox and IAA or DMSO over 3 days. I incubated cells at 24-hr intervals with the PrestoBlue reaction reagent to measure population proliferation. Plotting the measurements of fluorescence between 544 and 590 nm to detect the availability of resazurin on a box plot revealed no differences between DMSO- or IAA-treated cells at 24 and 48 hr when using a Wilcoxon test (Figure 4.3A). However, there were statistically significant differences in cell growth found at 72 hr of treatment ($p = 0.03$). In comparison, when the same clone was treated with siRNA to KD SAF-A for 72 hr, statistically significant differences in cell proliferation were observed at 48 hr ($p = 0.029$; Figure 4.3B). At 72 hr, control cells continued to proliferate while KD cells were at a similar level of fluorescence compared to 24 hr earlier. There was also a statistically significant difference in cell growth between the control and treatment samples at 72 hr ($p = 7.6 \times 10^{-6}$). When the parental HCT 116 Tet-OsTIR1 cell line was treated with siRNA to KD SAF-A as above, there were statistical differences observed at 48- and 72- hr ($p = 0.031$, $p = 0.0009$; Figure 4.3C). From this data it appeared that the stable cell line expressing mAID-SAF-A was not very responsive to IAA treatment and homozygote degradation but was responsive to siRNA KD of endogenous SAF-A. These results were not supportive of the stable cell line being used for further experiments at that time. In addition, after several passages, the cells lost the mAID tag on one copy of the SAF-A gene and thus became heterozygous showing that the tag is not well tolerated on the N-terminus of the gene (data not shown).

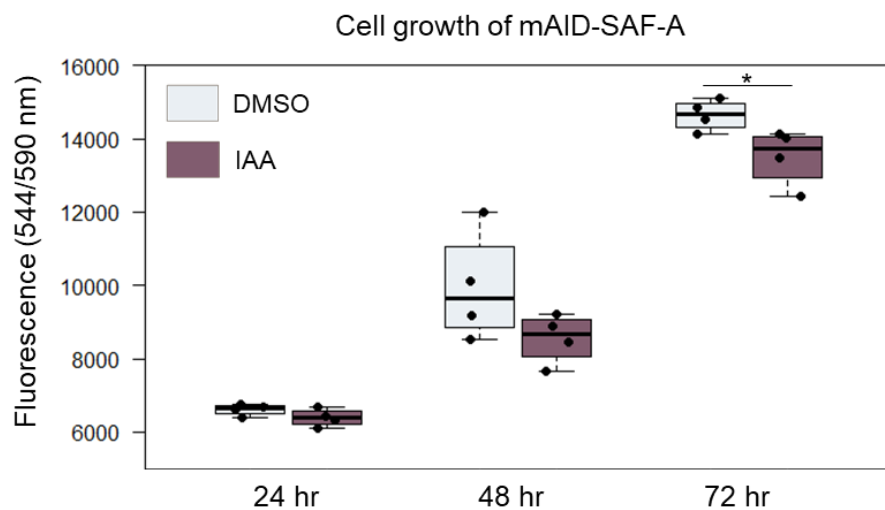
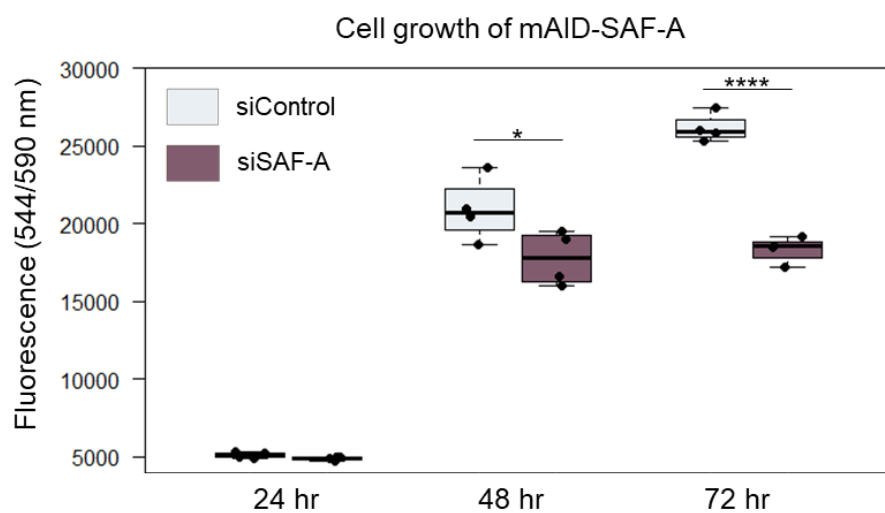
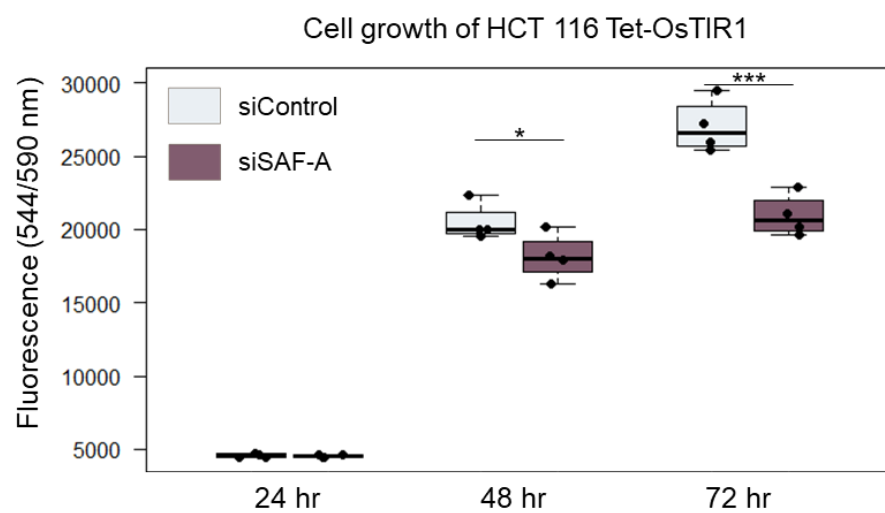
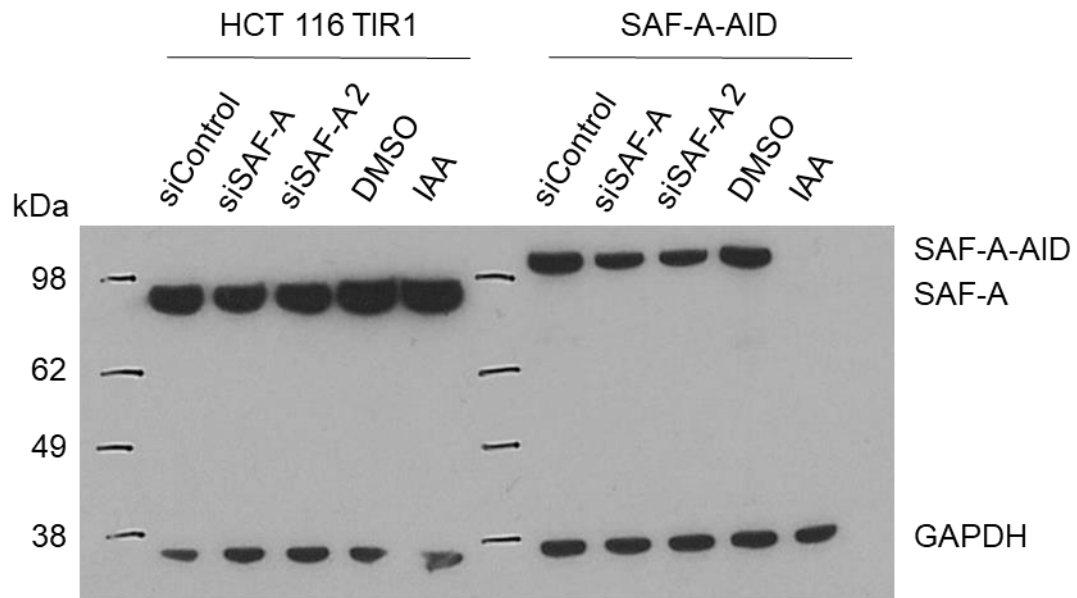
A**B****C**

Figure 4.3. Cell growth of mAID-SAF-A and HCT 116 Tet-OsTIR1 cells. Values on the y-axis were taken with a plate reader measuring fluorescence at 544/590 nm wavelength following a 10-minute treatment with the PrestoBlue reagent. Each box plot consists of 4 technical replicates and shows the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the different conditions at each time point (Wilcoxon test, * = $p \leq 0.05$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$). A) Cell growth of stably generated mAID-SAF-A cells grown in the presence of doxycycline and DMSO or 500 μ M IAA for the respective times. B) Cell growth of mAID-SAF-A cells treated with RNAi for the respective times. C) As B) but with HCT 116 Tet-OsTIR1 cells.

I also had the opportunity to use a SAF-A-AID degron cell line made in the Wilson lab which was homozygous for SAF-A tagged on the C-terminus with the full-length AID tag (Heath, 2018). This cell line was derived from a parental HCT 116 cell line constitutively expressing the OsTIR1 gene which lead to slightly reduced expression levels of SAF-A-AID compared to endogenous SAF-A (Figure 4.4A). The SAF-A AID cell line was responsive to 500 μ M IAA and after 24 hr of treatment there was no protein visible on the Western blot when probed with an anti-SAF-A antibody. Furthermore, both the parental and degron cell lines were resistant to 24-hr RNAi with siRNA targeting SAF-A. However, after two weeks of passaging in the presence of the required selection antibiotics (hygromycin, neomycin/G418, and puromycin), probing the cell lysate of SAF-A-AID cells with anti-SAF-A antibody revealed that the cell line was no longer homozygous for SAF-A-AID (Figure 4.4B).

A



B

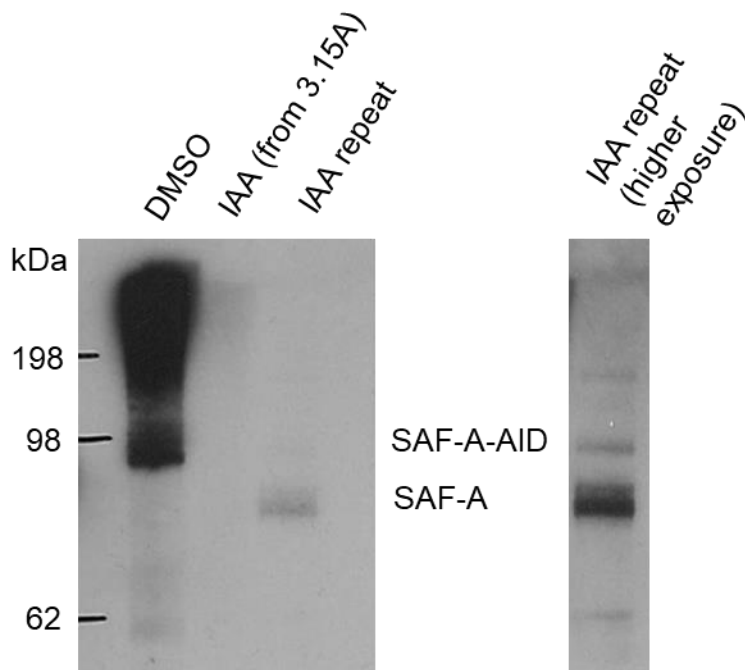


Figure 4.4. Testing of HCT116 TIR1 cells expressing SAF-A-AID. A) Western blot of paternal HCT 116 cells and cells expressing SAF-A-AID treated with RNAi (siControl, siSAF-A, siSAF-A 2), DMSO or 500 μ M IAA for 24 hours. Cell lysates were probed with anti-SAF-A and anti-GAPDH antibodies. B) Western blots of SAF-A-AID cells treated with DMSO and IAA two weeks after defrosting. Sample IAA (from 3.15A) is the same as A) sample SAF-A-AID – IAA. Cell lysates were probed with anti-SAF-A antibody.

4.2.2. Generating a degradation tag (dTAG) tagged SAF-A cell line

After my experiments indicated that mAID and AID were unsuitable for making a degron-tagged version of SAF-A, I researched alternative inducible degradation methods. A new approach is called dTAG, where the FKBP12 mutant protein (F36V) sequence is attached to the protein of interest and a specific targeting molecule is used to mark it for degradation by the native proteasome (Nabet *et al.*, 2018). The advantage of the dTAG system over the AID system is that no other genetic changes are required beyond tagging the protein, while for the AID system the cells need to express the OsTIR1 gene for the degradation to work (Figure 4.5A). To test the dTAG system, I created SAF-A constructs tagged with FKBP12^{F36V} under tetR control to test their degradation when treated with dTAG-13. Initial data did not show any degradation when treated with Dox and at the recommended 100 nM dTAG-13 concentration after cell lysates were probed with an anti-FLAG antibody (Figure 4.5B). They also had additional protein bands on the protein blots indicating construct instability. I further increased the concentration of dTAG-13 to 500, 1000 and 2000 nM to increase the protein degradation and saw no change in its abundance when probing cell lysates with an anti-FLAG antibody (Figure 4.5C). As I did not have a working positive control, I created two C-terminally tagged constructs for XRN2 and EGFP to test if the lack of degradation is due to SAF-A or the tag. When using a range of concentrations for dTAG-13, I was able to degrade half of XRN2 at the lowest concentration and none of EGFP at any concentration according to Western blot results (Figure 4.5D, data not shown).

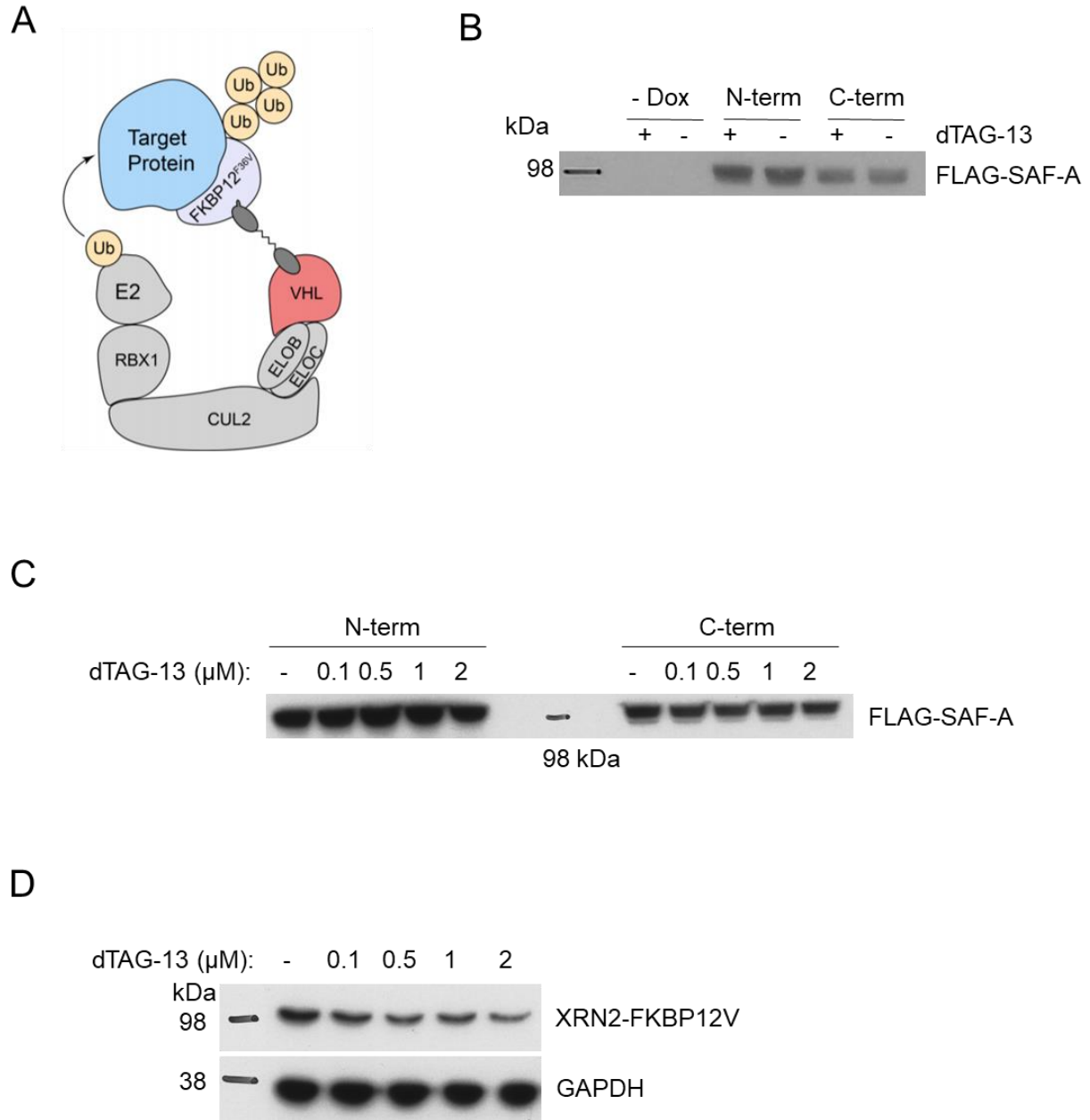
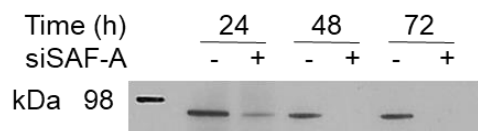


Figure 4.5. Depletion of transiently transfected SAF-A tagged with FKBP12^{F36V} in HCT 116 cells. A) Schematic illustration of the dTAG degron system. Taken from Nabet *et al.*, 2018. B) Western blot of cell lysates transfected with FKBP12^{F36V}-FLAG-SAF-A (N-term) or SAF-A-FKBP12^{F36V}-FLAG (C-term) and treated with 100 nM dTAG-13 and doxycycline (Dox) for 24 hours. Lysates were probed with anti-FLAG antibody. C) as B) but treated with different concentrations of dTAG-13. D) Western blot of cell lysates transfected with XRN2-FKBP12^{F36V}-FLAG and treated with different concentrations of dTAG-13 and Dox for 24 hours. Cell lysates were probed with anti-FLAG and anti-GAPDH antibodies.

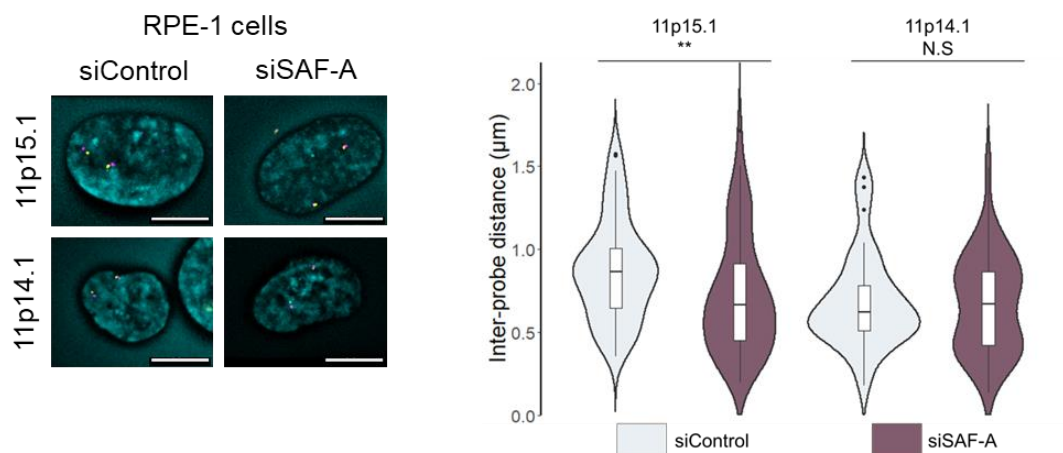
4.2.3. RNAi depletion of SAF-A at an earlier time point

Previous experiments in the lab to elucidate the effect of SAF-A depletion on chromatin structure utilised RNAi over 48 to 72 hr in RPE-1 or FlpIn 293 T-REx cells (Nozawa *et al.*, 2017). In this experiment, I aimed to examine whether compaction of gene-active regions could be observed within a shorter timeframe of 24 hr, during which SAF-A may not be fully depleted. I knocked down SAF-A in RPE-1 and HCT 116 cells using RNAi and checked the levels of the endogenous protein by probing cell lysates with an anti-SAF-A antibody (Figure 4.6A and 4.6C). In RPE-1 cells, after 24 hr of knockdown (KD) there was still protein visible albeit at a greatly reduced amount. At 48- and 72-hr KD there was no protein visible compared to control samples. Similarly, for HCT 116 most of the protein was knocked down at 24 hr when compared to the control and there was none visible at 48- and 72-hr KD. To assess an effect on large-scale chromatin structure, I performed fluorescence-in-situ hybridisation (FISH) with probes for a gene-rich (chromosome 11 p15.1) and a gene-poor (chromosome 11 p14.1) region to see whether a 24 hr SAF-A KD affected interphase chromatin decompaction (Figure 4.6B and 4.6D). 3D images of cells were taken with z-stack intervals of 20 nm and images were deconvolved in Volocity using the Fast Restoration option (see Section 2.9.6.). Using Imaris, each probe in an image was detected and modelled using an estimation of the point-spread function (PSF) of light in 3D. The probe centroids were calculated, and high-quality probe pairs were manually selected for measurement (Figure 4.6B and 4.6D). For RPE-1 cells there was a statistically significant difference in FISH probe distance between siControl and siSAF-A samples in the gene-rich region (Wilcoxon test, $p = 0.006$) and no differences in the gene-poor region, similar to Nozawa *et al* (2017). In HCT 116 there was statistically significantly more compaction in the gene-rich region (Wilcoxon test, $p = 8.1 \times 10^{-7}$) and no differences in the gene-poor region. Overall, my results indicate that treatment of cells with only 24 hr RNAi could be utilised in HCT 116 cells in other experiments (see Chapters 5.3.3.) where endogenous SAF-A is knocked down and a different version of the protein is expressed.

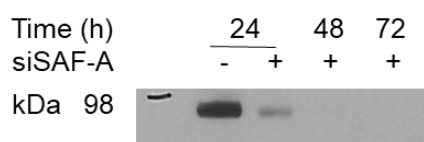
A



B



C



D

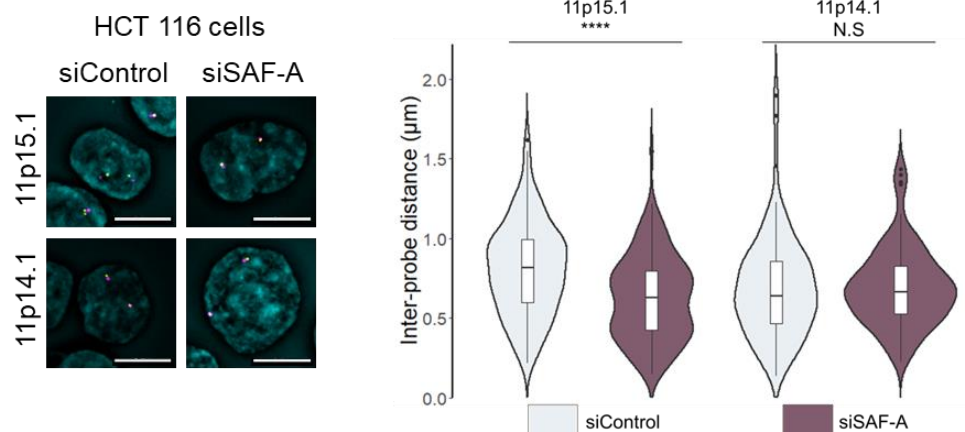


Figure 4.6. SAF-A depletion using RNAi in RPE-1 and HCT 116 cells. A) Western blot of RPE-1 cell lysates treated with siControl (-) or siSAF-A (+) for 24, 48, or 72 hours. B) Representative images (left) of FISH with probes for gene-rich (11p15.1) and gene-poor (11p14.1) regions of the genome in RPE-1 cells treated with siControl or siSAF-A. Distribution of inter-probe distances between probe pairs. Each violin plot (read from left to right) represents a probe pair sample size of $N_1 = 50$, $N_2 = 61$, $N_3 = 33$, $N_4 = 37$. Wilcoxon test, ** = $p \leq 0.01$. C) As A for HCT 116 cells. D) As B for HCT 116 cells. Each violin plot (read from left to right) represents a probe pair sample size of $N_1 = 132$, $N_2 = 127$, $N_3 = 89$, $N_4 = 87$. Wilcoxon test, **** = $p \leq 0.0001$. Scale bar is 8 μm .

4.3. SAF-A and the RNA mesh

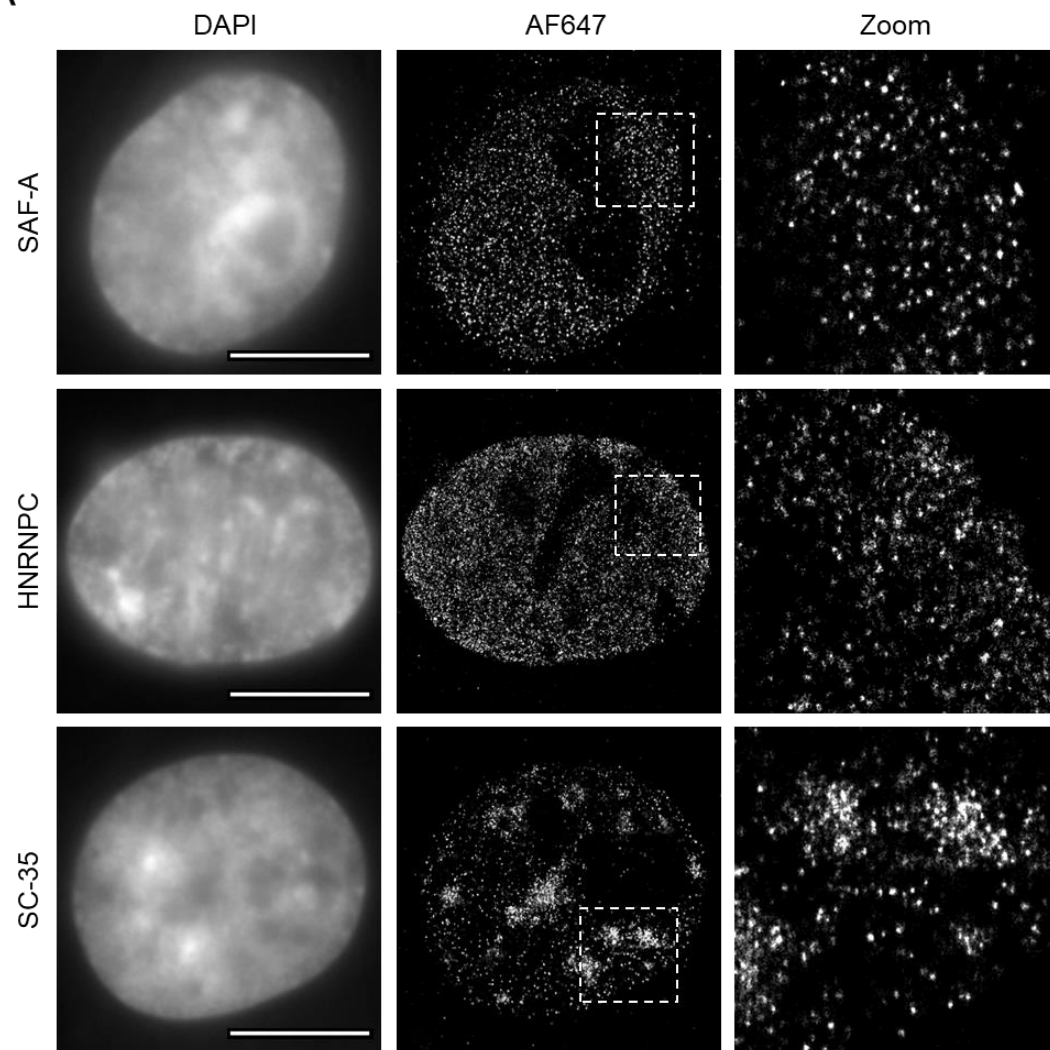
4.3.1. Super-resolution microscopy of nuclear proteins

Single-molecule localisation microscopy is a super-resolution microscopy technique that surpasses the 200 nm diffraction limit of most light microscopy techniques. By calculating a list of 3D positions of localisation events and reconstructing an image from these positions, imaging techniques such as dSTORM can achieve a resolution of approximately 20 nm (Henriques *et al.*, 2011; van de Linde *et al.*, 2011). To use dSTORM, the experimental steps are similar to those of conventional fixed cell immunofluorescence for studying a protein of interest. However, in dSTORM, the secondary antibody is conjugated to a single fluorophore that through its imaging properties “blinks”, releasing photons. This blinking behaviour is captured in a series of frames so that each blink generally corresponds to a single protein molecule. If the blinking frequency is appropriately balanced – not too high or too low - it allows for the precise localisation of numerous protein molecules within a cell. To reconstruct an image, clustering algorithms are necessary, which are generally based on principles from a range of fields, including statistics, graph theory, computational geometry, and machine learning (Khater *et al.*, 2020).

Previous experiments in the lab with super-resolution STED (stimulated emission depletion) microscopy were performed to observe the organisation patterns of wild-type SAF-A and SAF-A Walker A (WA) and Walker B (WB) mutants (Nozawa *et al.*, 2017). SAF-A WB mutants were found to exhibit a granular staining pattern due to an increased size of SAF-A oligomers compared to SAF-A WT and WA. To further understanding of SAF-A's nuclear organisation, I used dSTORM to image SAF-A's nuclear organisation in high resolution and two other nuclear proteins as a comparison, HNRNPC and SC-35. These proteins are all involved in RNA processing with the former two proteins being of the same family and having an apparently similar diffuse distribution. The latter protein, SC-35, is known to form distinct nuclear speckles. Using super-resolution microscopy, I collected experimental data and could explore the protein distribution of these three proteins (Figure 4.7A). Another advantage of this type of microscopy over standard light microscopy methods is the

ability to perform calculations on the 3D positions to determine protein clusters, cluster density and their connectivity. Generally, these density-based clustering algorithms rely on subjectively defined parameters like the popular DBSCAN algorithm does (Ester *et al.*, 1996). A method, SuperStructure, was developed using the experimental data I collected for SAF-A, HNRNPC and SC-35 to allow the detection and quantification of connected clusters without user-defined parameters and to define the density of molecules within these clusters (Marenda *et al.*, 2021). Using SuperStructure allowed us to forego subjectively defined parameters for cluster size and detect clusters and super-clusters of nuclear proteins by multiple runs of the DBSCAN algorithm to detect the number of clusters within a radius distance from a fluorescent point that increases with each run (Figures 4.7B). The program then outputs a curve which initially follows a Poisson function (with Gaussian decay of the number of connections at a low radius with many clusters) that represents connectivity within clusters. If clusters are connected, the curve then follows a different Poisson function with exponential decay as the radius increases, meaning the connected clusters merge quickly to reduce the total number of clusters. Therefore, the higher the decay length is, the more connected protein clusters are. The decay length is then normalised by the cluster density contribution to allow for comparison between datasets with different cluster densities. The data showed that SAF-A forms an interconnected mesh of dense clusters while HNRNPC forms a less connected mesh with less dense clusters (Figure 4.7B, Wilcoxon test, $p = 0.002$). Meanwhile, SC-35 has two distinct distribution patterns, one of the highly connected dense clusters (intra) corresponding to the nuclear speckle structures that have a similar density to SAF-A clusters but a higher connectiveness (Wilcoxon test, $p = 0.002$), and a second state (inter) that has less connectivity and density than the HNRNPC mesh (Wilcoxon test, $p = 0.002$). Overall, while SAF-A, HNRNPC and SC-35 speckles form visually similar clusters, SuperStructure analysis showed that they have different cluster connectiveness and density.

A



B

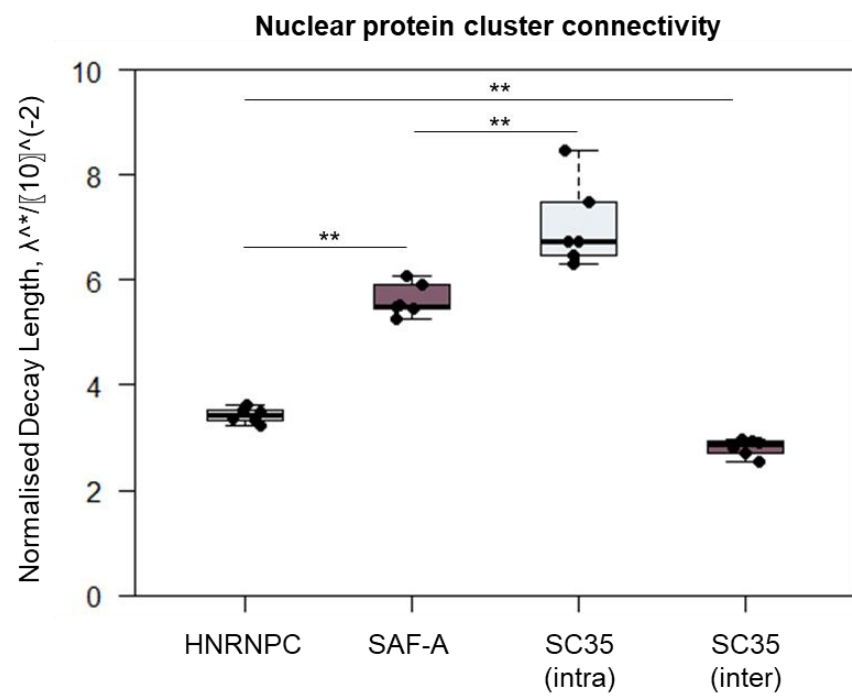


Figure 4.7. Super-resolution microscopy of nuclear proteins using dSTORM. A) Reconstructed super-resolution dSTORM microscopy images of fixed cell immunofluorescent staining of DNA and a nuclear protein. DNA was visualised by DAPI staining. Nuclear proteins were visualised with a rabbit antibody (SAF-A) or mouse antibody (HNRNPC, SC-35) and a donkey anti-rabbit or anti-mouse CF647 secondary antibody. Dashed white line squares represent zoomed in regions. B) Box plot of normalised decay length (λ^*) for protein clusters found in reconstructed dSTORM images. Each box on the x-axis is the average of 6 single cells' decay lengths (λ) divided by detected cluster density (ρ_{cl}) with differences between them analysed using the Wilcoxon test (Marenda *et al.*, 2021). Decay length is obtained from SuperStructure-generated curves from the entire cell nucleus except in the case of HNRNPC and SC-35 where local regions were analysed instead (an average of 5 values per cell). Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the respective proteins (** = $p \leq 0.01$). Scale bar is 10 μm .

4.3.2. Super-resolution microscopy of SAF-A with transcription inhibition

Previous research showed that transcription regulates SAF-A oligomerisation in the presence of caRNAs (Nozawa *et al.*, 2017). To investigate if the observed connected SAF-A clusters (Figure 4.7) consisted of oligomeric SAF-A structures, I inhibited transcription over a long period of time to deplete cells of most nuclear RNA. RPE-1 cells were treated with triptolide for 4 hr to inhibit nascent RNA transcription and degrade much of the RNA that would be associated with SAF-A (Figure 4.8A). The 4 hr treatment with triptolide was based on experiments performed in the lab by Mattia Marenda and Davide Michieletto where they observed changes to SAF-A domain fragments (e.g. SPRY-AAA, SPRY-AAA-RGG) organisation after 4 hr transcription inhibition with α -amanitin and triptolide (data not shown). While there were no obvious visual differences between the conditions, SuperStructure analysis revealed differences in cluster size and connectivity (Figures 4.8B and C). In Figure 4.8C, the cluster connectivity is shown by decay rate instead of decay length for better visual representation, with a higher value reflecting higher connectivity. SAF-A cluster size was reduced by a statistically significant amount in cells with inhibited transcription compared to control cells (Figure 4.8B; Wilcoxon test, $p = 7.2 \times 10^{-9}$). Connectivity between SAF-A clusters was also reduced in cells treated with triptolide compared to the control (Figure 4.8C; Wilcoxon test, $p = 0.0004$). Therefore, the high connectivity of the SAF-A mesh and the protein cluster sizes are transcriptionally dependent. This loss of connectivity and reduction in cluster size could be due to the loss of RNA bound to SAF-A which reduces SAF-A oligomers to monomers which shows it is likely that SAF-A clusters contain oligomeric SAF-A structures.

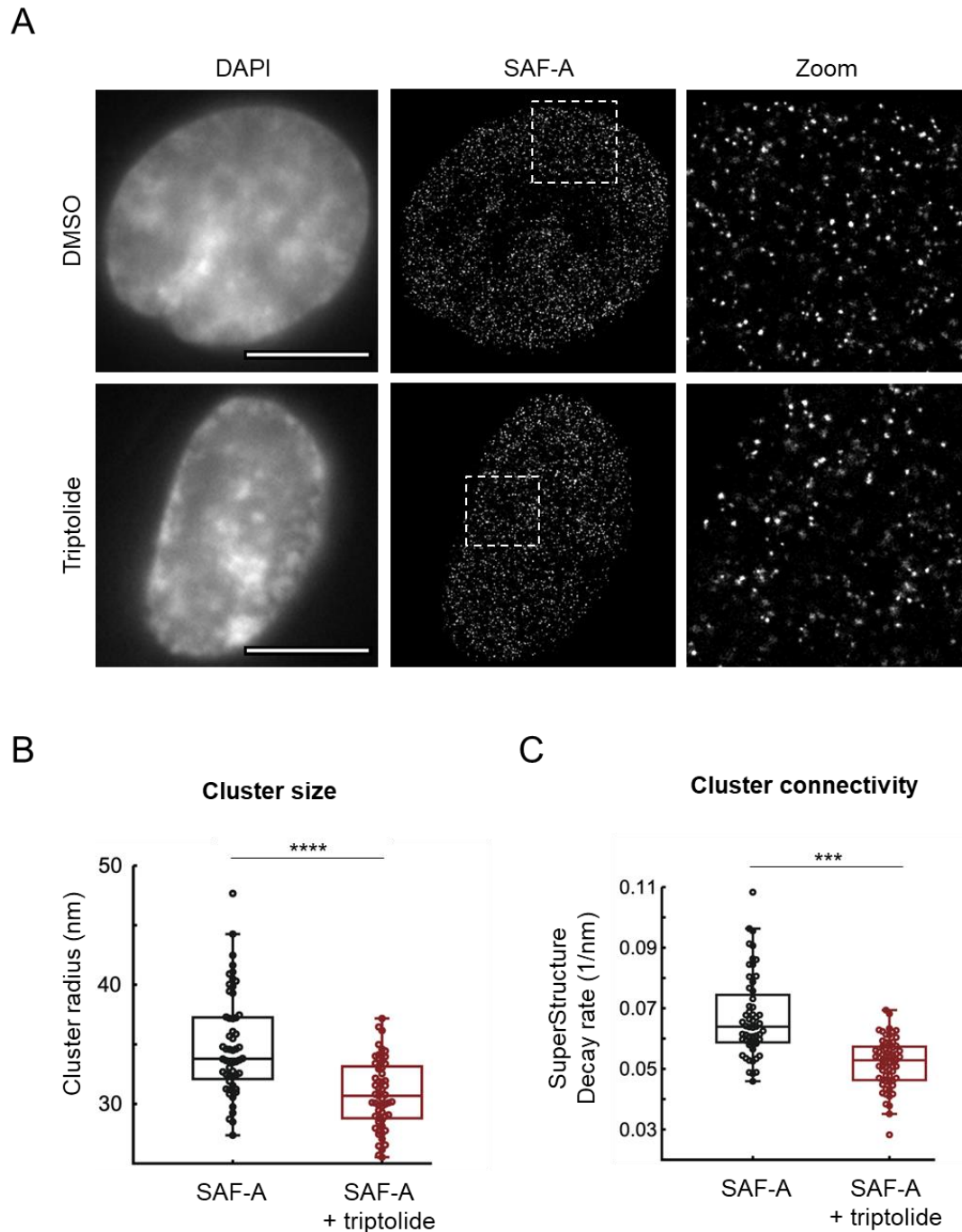
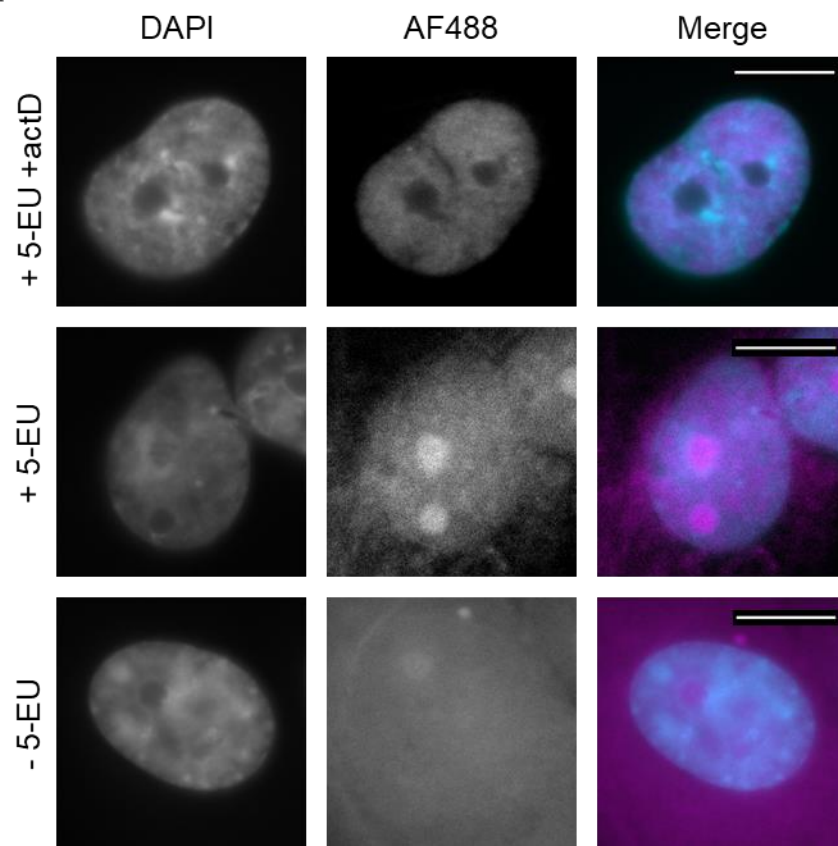


Figure 4.8. Effect of transcription inhibition on the SAF-A structures. A) Reconstructed super-resolution dSTORM microscopy images of fixed cell immunofluorescent staining of DNA and SAF-A in cells treated with DMSO or triptolide for 4 hr. DNA was visualised by DAPI staining. SAF-A was visualised with a rabbit antibody and a donkey anti-rabbit CF647 secondary antibody. Dashed white line squares represent zoomed in regions. B) Box plot of cluster radius of local regions of the SAF-A mesh calculated from reconstructed dSTORM images with SuperStructure analysis. The local regions were 3 to 5 per cell, for a total of 14 untreated cells and a total of 15 triptolide-treated cells. Each box plot on the x-axis is the average of the local regions with differences between the two conditions analysed using the Wilcoxon test. Boxes show the median, lower and upper quartiles, and max and min values with each cell region value shown as a coloured circle. C) Same as B) but for decay rate (τ) calculated by SuperStructure analysis ($\tau = 1/\lambda$). Box plots were generated in collaboration with Mattia Marenda. Statistically significant differences are marked with stars and bars between the respective conditions (** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$). Scale bar is 10 μm .

4.3.3. Visualising nascent RNA with 5-EU

My previous studies and other experiments from the group supported the idea that SAF-A is able to form an RNA-protein mesh in the nucleus. To test this idea I wanted to examine how SAF-A depletion affected RNA mesh organisation. While SAF-A depletion does not change the overall quantity of RNA in the nucleus it could affect the distribution or organisation of RNA including caRNA. To do this experiment I used fixed cell immunofluorescence where nascent RNA was visualised using 5-EU labelling within a short time window of 15 min in the presence of a low dose of actD to inhibit rRNA production. To visualise the RNA, the RPE-1 cells were fixed and treated with a click reaction to attach a fluorescently labelled azide probe (Figure 4.9). If rRNA production is not inhibited, the 5-EU signal accumulates in the nucleolus because of new rRNA synthesis (Figure 4.9A) which makes it difficult to visualise the RNA mesh. After low-dose actD treatment and a 5-EU pulse, the lifetime of nascent RNA can be visualised, as the label does not appear to be recycled in cells (Figure 4.9B). The signal increased over time with the strongest RNA signal observed at 0.5 and 1 hr after the 5-EU pulse. The signal then reduced and had mostly disappeared at 3 hr post-treatment. These results are consistent with previous research (Nozawa *et al.*, 2017).

A



B

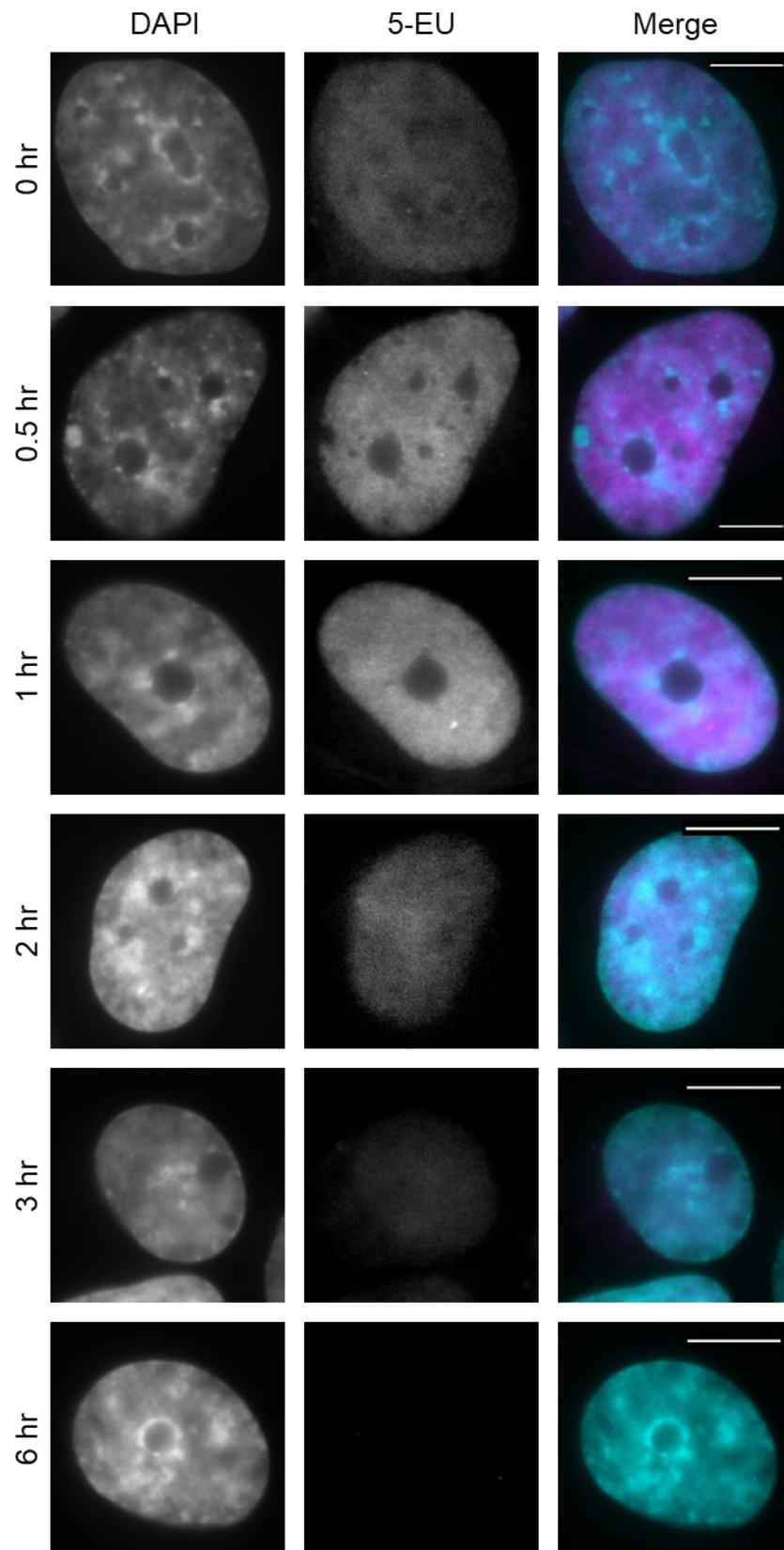
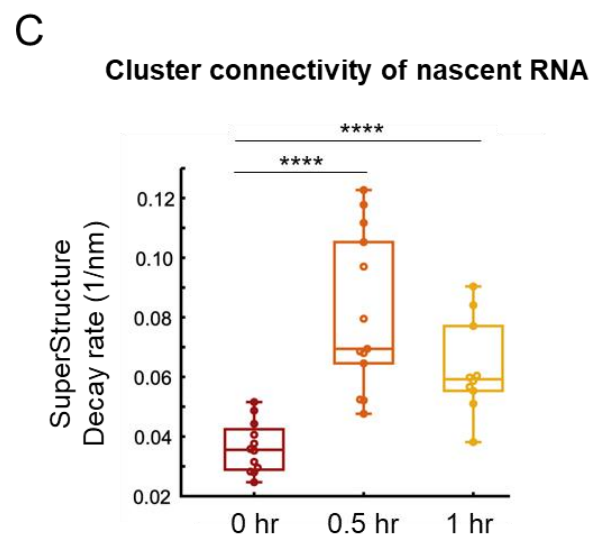
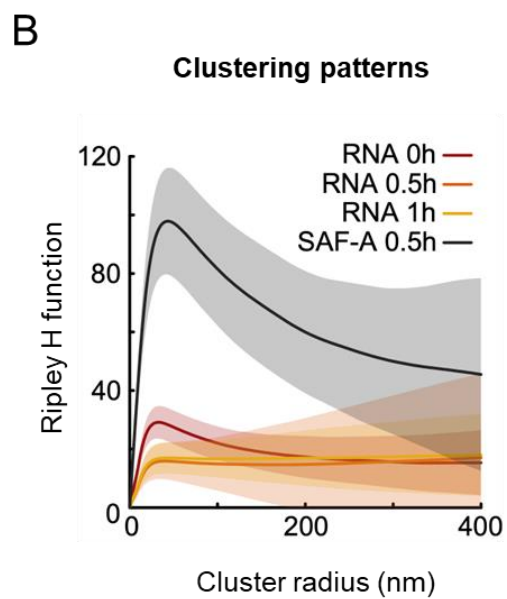
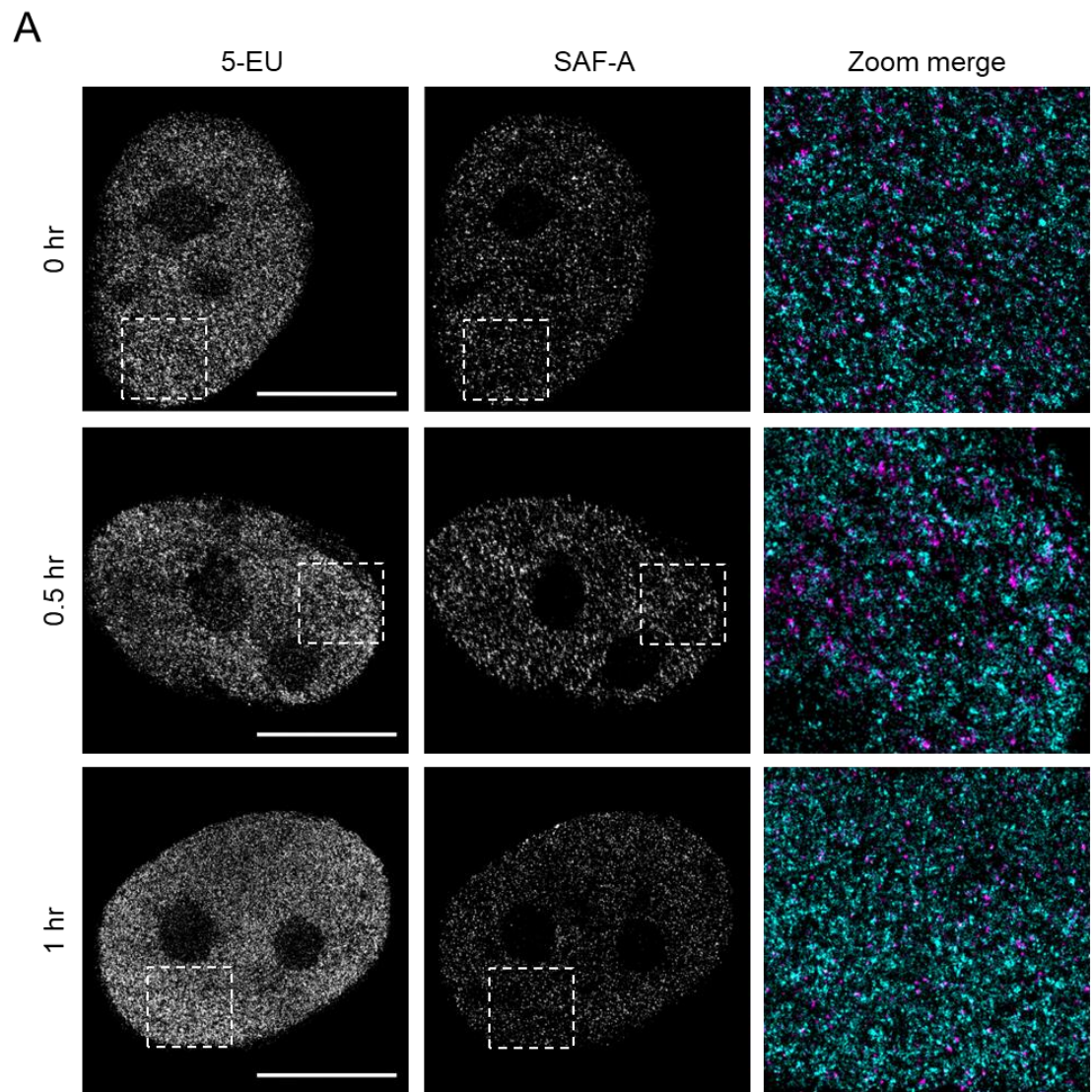


Figure 4.9. Visualising RNA with click chemistry. Immunofluorescence staining of fixed RPE-1 cells. DNA was visualised by DAPI staining. RNA was visualised with 5-EU incorporation followed by conjugation to AF488 azide using click chemistry. A) Cells were pre-treated with actD and/or pulse labelled for 15 min with 5-EU. Cells were fixed 1 hr after pulse labelling. B) Cells were treated with actD then pulse labelled for 15 min with 5-EU before fixation at specific time points. Scale bar is 10 μm .

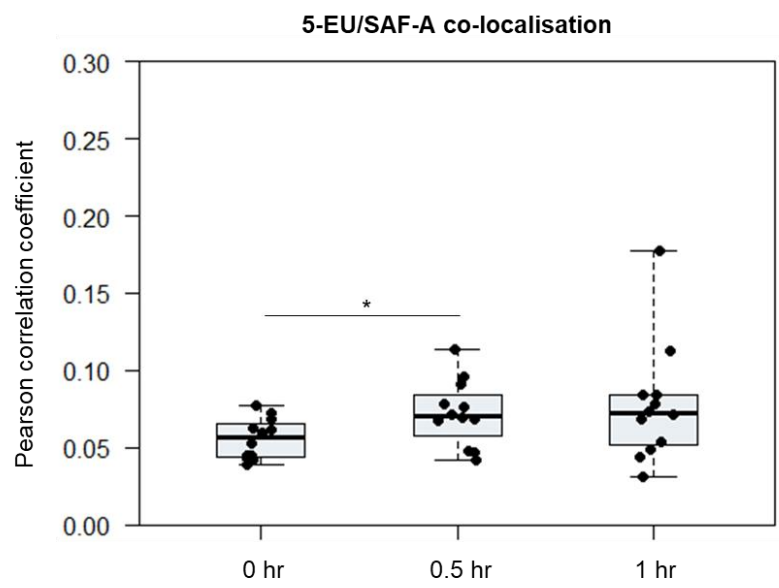
4.3.4. Super-resolution microscopy of SAF-A and RNA

Super-resolution microscopy can also be used to visualise nuclear RNA distribution by labelling with 5-EU and conjugating the nucleoside with click chemistry to a blinking fluorescent azide. Changes in nascent RNA organisation (Figure 4.9B) were followed after a brief pulse of 5-EU in the presence of low-dose actD to prevent enrichment of the signal in nucleolar rRNA. A similar experiment was performed in RPE-1 cells pulsed with 5-EU for 15 min, washed and fixed at 0, 0.5 or 1 hr after treatment (Figure 4.10A). Initially, nascent RNA forms smaller clusters compared to SAF-A's when looking at their cluster radiuses (Figure 4.10B). Over time though the RNA clusters became more diffuse, whilst simultaneously the connectivity between RNA clusters increased with time, with clusters at 0.5 and 1 hr after 5-EU incubation having statistically significant differences in SuperStructure decay rate compared to 0 hr (Figure 4.10C; Wilcoxon test, $p = 1.8 \times 10^{-6}$ between 0 and 0.5 hr, $p = 6.4 \times 10^{-5}$ between 0 and 1 hr). In the same experiment, SAF-A was simultaneously labelled to allow the co-localisation between the SAF-A clusters and the nascent RNA over time to be determined (Figure 4.10A). By eye, it is difficult to examine the co-localisation of SAF-A and RNA unlike with imaging from other light microscopy techniques used for SAF-A and DNA co-localisation (Chapter 3.2). To quantify the co-localisation between the RNA and SAF-A signal at different time points of 12 cells each was quantified using Pearson correlation coefficient, generated by the Coloc2 plugin from ImageJ (Figure 4.10D). The closer the values are to 1, the higher the co-localisation. The average values for the three conditions were below 0.1 which shows the difficulty of analysing co-localisation at such a high resolution using this plugin. RNA and SAF-A at 0.5 hr post-5-EU treatment had the highest co-localisation compared to 0 hr based on a Wilcoxon test ($p = 0.03$). An alternative way to visualise co-localisation for super-resolution images is to co-localise clusters identified with the DBSCAN

algorithm (Figure 4.10E). This method found that the maximum co-localisation was at 0.5 hr post-5-EU treatment, the same as the results from Coloc2 co-localisation. Cluster co-localisation found that 1 hr post-5-EU treatment RNA and SAF-A had the least co-localisation which differed from the Coloc2 co-localisation where the 1 hr sample was between 0 and 0.5 hr in terms of co-localisation. The Coloc2 plugin focuses on the signal overlap from two channels which was more challenging to measure with the single points produced by single-molecule localisation microscopy. Therefore, the DBSCAN co-localisation observations provide a more accurate assessment in this case, as they compare cluster signal overlap. As the DBSCAN algorithm also has a history of use in super-resolution microscopy to determine clusters and co-localisation, I have more confidence in the results produced by it than with the Coloc2 plugin for this data.



D



E

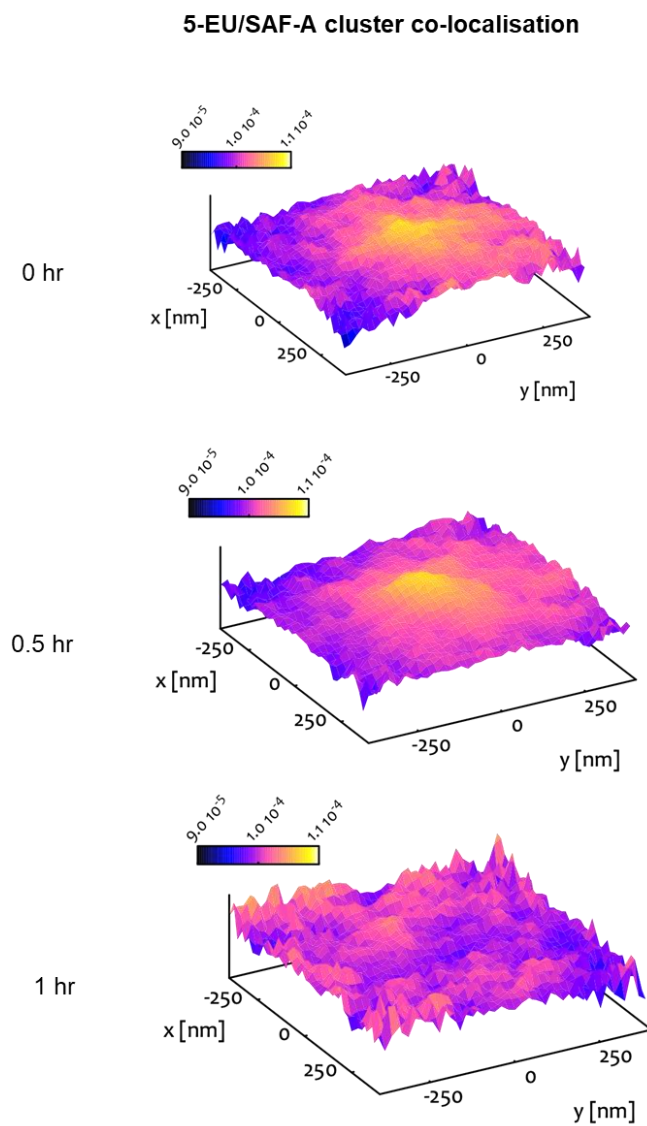
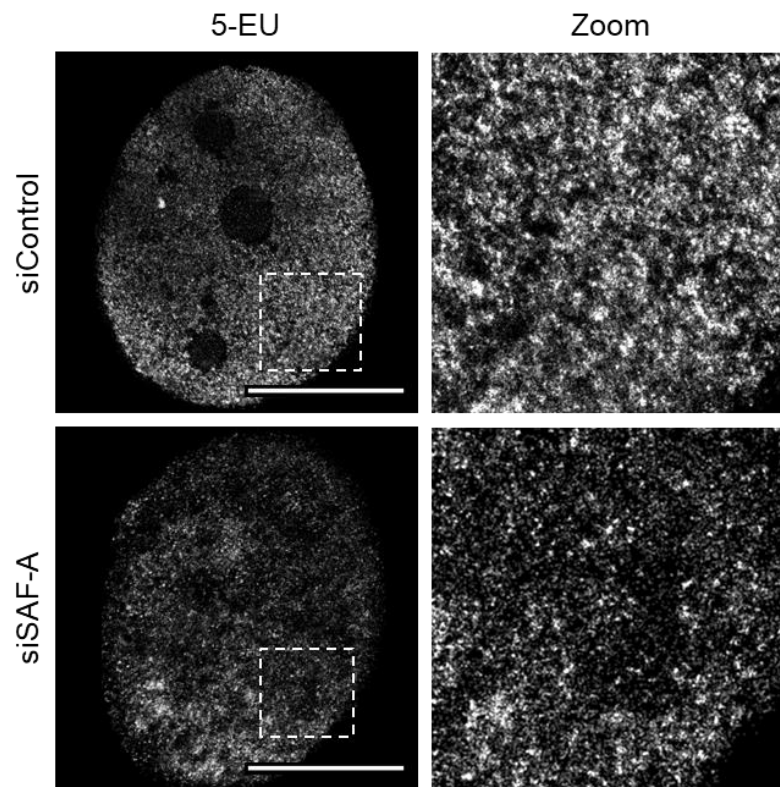


Figure 4.10. Two-colour super-resolution microscopy of RNA and SAF-A using dSTORM. A) Reconstructed super-resolution dSTORM microscopy images of fixed cell immunofluorescent staining of RNA and SAF-A in cells incubated with 5-EU for 15 min and fixed immediately or after 30 or 60 min. RNA was visualised with 5-EU integration and conjugation with Cy5 picolyl azide using click chemistry. SAF-A was visualised with a rabbit antibody and a donkey anti-rabbit CF568 secondary antibody. Dashed white line squares represent zoomed in regions. RNA is visualised in cyan and SAF-A is displayed in magenta in the merged images. B) Line plot of clustering patterns of nascent RNA at different timepoints and SAF-A. Ripley's H function values show the cumulative deviation of points from randomness across all distances. Cluster radius was analysed with SuperStructure. Values are calculated from 12 cells for 0 hr timepoint and 14 cells for the other conditions. C) Box plot of decay rate (τ) for nuclear RNA clusters found in reconstructed dSTORM images with a SuperStructure analysis ($\tau = 1/\lambda$). Each box plot on the x-axis is the average of 12 (0 hr) or 14 (0.5 hr, 1 hr) with differences between the three conditions analysed using the Wilcoxon test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a coloured circle. D) Box plot of Pearson correlation coefficient values generated from the Coloc2 ImageJ plugin. Each box on the x-axis is the average of 12 single cells' values with differences between them analysed using the Wilcoxon test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. E) 3D graph of co-localisation between SAF-A and RNA clusters analysed using the DBSCAN algorithm. The brighter the colour, the more co-localisation there is while darker colours show less co-localisation. Statistically significant differences are marked with stars and bars between the respective conditions (* = $p \leq 0.05$, **** = $p \leq 0.0001$). Scale bar is 10 μm . The plots in B), C) and E) were generated in collaboration with Mattia Marenda.

4.3.5. Super-resolution microscopy of RNA with SAF-A depletion

I wanted to determine whether the depletion of SAF-A affected the organisation of the nascent RNA mesh in the nucleus using super-resolution microscopy. An experiment in RPE-1 cells, where cells are treated with siControl or siSAF-A for 24 hr and then the RNA was visualised using 5-EU, allowed me to compare the signal between the two conditions in individual cells at a high resolution (Figure 4.11A). Visually, cells depleted of SAF-A had a weaker clustered RNA signal compared to the strong RNA signal visible 1 hr after 5-EU incorporation. Analysis with SuperStructure showed that cells depleted of SAF-A had reduced cluster density (data not shown) and cluster connectivity between clusters compared to control cells (Figure 4.11B; Wilcoxon test, $p = 0.0001$). Therefore, depletion of SAF-A appears to disrupt the RNA mesh and causes it to disperse throughout the nucleus, weakening the signal of RNA clusters. Previously it was shown that SAF-A depletion did not affect nascent RNA levels visualised with 5-EU or 4sU (Nozawa *et al.*, 2017). Therefore, the observed differences between siControl and siSAF-A treated cells in my experiment should not be due to overall changes in transcription but due to SAF-A depletion.

A



B

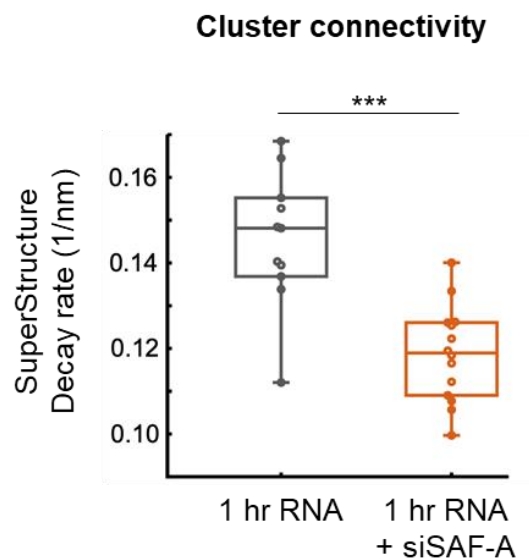


Figure 4.11. Effect of SAF-A depletion on the RNA mesh. A) Reconstructed super-resolution dSTORM microscopy images of fixed cell immunofluorescent staining of RNA in cells treated with siControl or siSAF-A for 24 hr. RNA was visualised after a 15 min 5-EU integration, fixation after 1 hr, and conjugation with Cy5 picolyl azide using click chemistry. B) Box plot of decay rate (τ) for nuclear RNA clusters found in reconstructed dSTORM images with a SuperStructure analysis ($\tau = 1/\lambda$). Each box plot on the x-axis is the average of 11 (1 hr RNA) or 15 (1 hr RNA + siSAF-A) with differences between the three conditions analysed using the Wilcoxon test. Boxes show the median, lower and upper quartiles, and max and min values with each cell value shown as a coloured circle. Box plot was generated in collaboration with Mattia Marenda. Statistically significant differences are marked with stars and bars between the respective conditions (***) = $p \leq 0.001$). Scale bar is 10 μm .

4.4. Amplicon SLAM-seq

4.4.1. Method optimisation

Previous CLIP-seq experiments that sequence the total RNA obtained from crosslinked SAF-A show that SAF-A prefers binding the most abundant RNA species (Xiao *et al.*, 2012; Huelga *et al.*, 2012; Yugami *et al.*, 2020; Van Nostrand *et al.*, 2020). However, those experiments do not differentiate between old and newly synthesised RNA whilst it has been shown that nascent RNA is important for the formation of scaffolds that prevent DNA compaction and attract SAF-A to actively transcribed regions (Nozawa *et al.*, 2017; Creamer *et al.*, 2021; Skalska *et al.*, 2021). During the course of my research, I have used uridine analogues (e.g. 5-EU, [5-³H]U) to investigate the interaction of SAF-A with nascent RNA. Following from this, I wanted to identify the nascent RNA molecules bound to SAF-A, which could be of potential molecular importance for driving SAF-A's recruitment to actively transcribed regions. One technique that allows for this identification is SLAM-seq, which involves the chemical modification of 4sU nucleosides integrated into RNA, enabling the subsequent differentiation between nascent and old transcripts during sequencing (Herzog *et al.*, 2017). Other comparable techniques, such as TimeLapse-seq and TUC-seq, have been used to estimate mRNA degradation rates and have demonstrated a similar performance to SLAM-seq (Boileau *et al.*, 2021). The original published SLAM-seq experiments used long labelling times (at least 2 hr) that were too long for labelling nascent RNA. Consequently, I explored various strategies to facilitate 4sU uptake over short periods of time to focus on nascent RNA.

I initially focused on improving 4sU uptake and incorporation by increasing nucleoside incorporation 10-fold to 1 mM and using HCT 116 cells instead of RPE-1 cells which are reported to overexpress the equilibrative nucleoside transporter 1 (ENT1) to increase nucleoside import (Snezhkina *et al.*, 2016). Cells were grown at 60% confluency and used a fresh aliquots of 4sU were used with protection from light to prevent degradation. After pulsing cells with 4sU for 15 min, I collected RNA at 0 and 1 hr intervals and analysed 4sU incorporation by HPLC (Figure 4.12). A similar experiment was repeated in FlpIn 293 T-REx cells and compared to HCT 116 values

(Figure 4.12). Although HCT 116 cells had a 1.2-fold higher signal than FlpIn 293 T-REx cells, this difference was sufficiently small that it was acceptable to use FlpIn 293 T-REx cells because they express an inducible FLAG-tagged SAF-A necessary for protein immunoprecipitation.

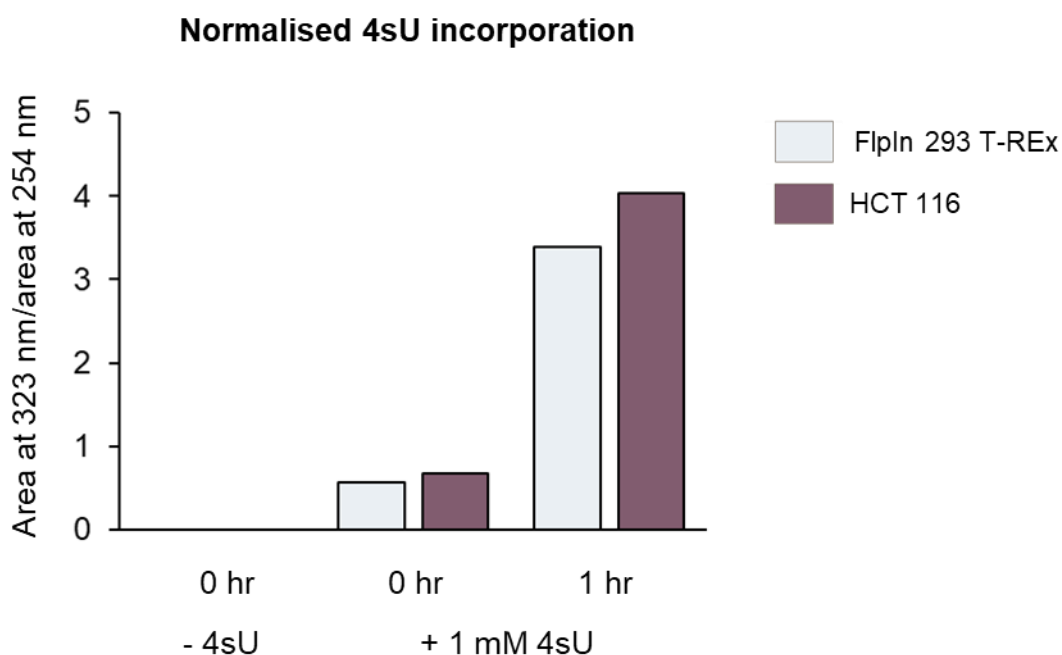
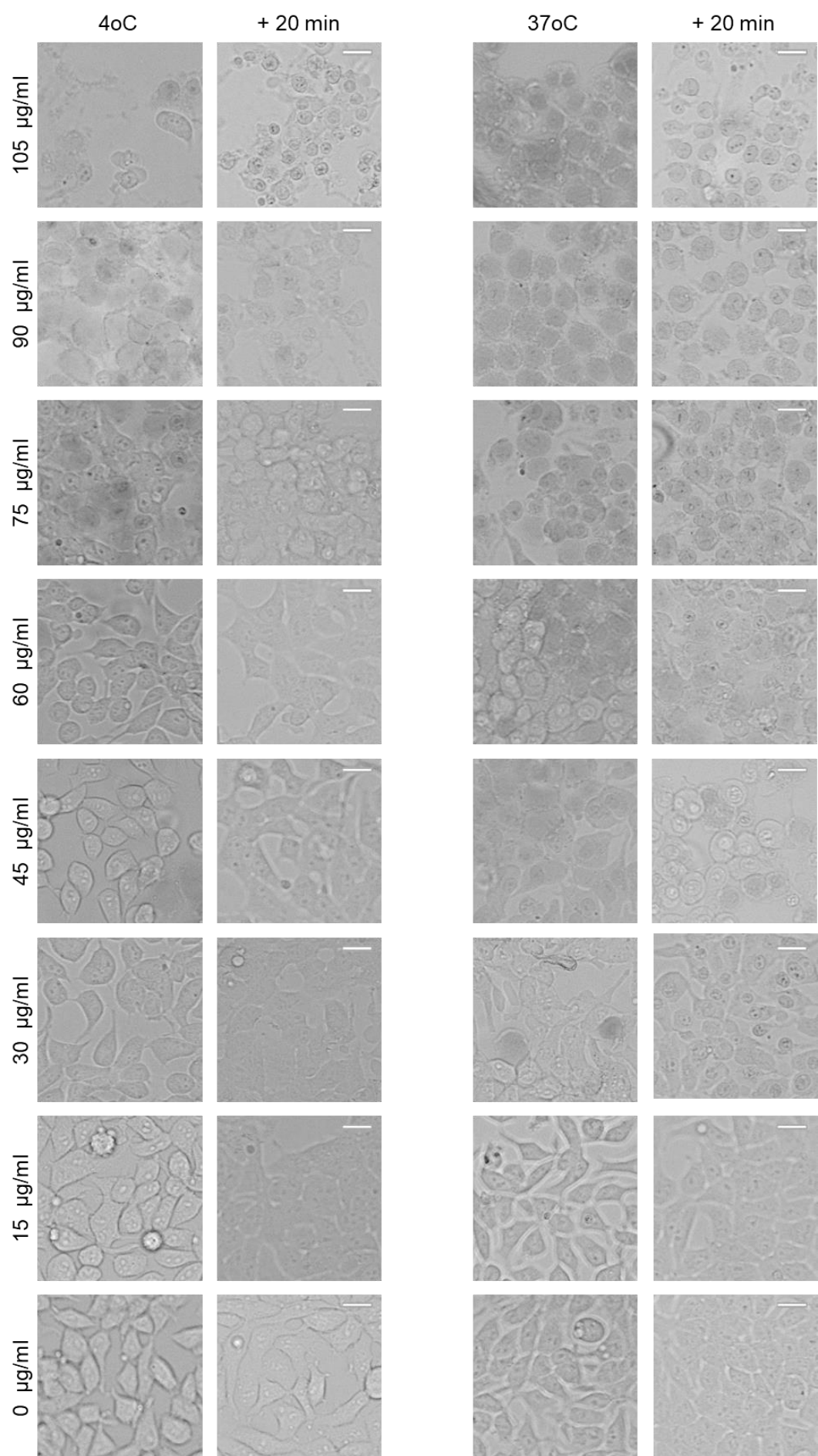


Figure 4.12. 4sU incorporation in FlpIn 293 T-REx and HCT 116 cells. Bar plot of HPLC values for 4sU/G area. FlpIn 293 T-REx and HCT 116 cells were incubated with 1 mM 4sU for 15 min then total RNA was collected after wash (0 hr) or 1 hr later. Each bar plot represents a sample size of N = 1.

To further improve 4sU uptake, I tested permeabilising cells with lysolecithin, a fatty acid that has previously been shown to increase the uptake of radiolabelled thymidine 5 to 10-fold after a short incubation (Miller *et al.*, 1979). Lysolecithin concentrations ranging from 0 µg/ml to 105 µg/ml were tested on FlpIn 293 T-REx cells. The incubations were performed at 37°C and 4°C with 1 mM CaCl₂ to ensure cells remain attached. After a 2 min incubation with lysolecithin, 0.1% Trypan Blue was added for 1 min, washed off and the cells were imaged under a brightfield microscope (Figure 4.13A). Usually, only dead cells uptake Trypan Blue, but

lysolecithin permeabilisation facilitated its uptake. The cells were then incubated for 20 min and imaged again to see if they successfully expelled the dye, a sign that the permeabilisation is reversible and would not affect cell survival. At 4°C there was uptake of Trypan Blue only at 75 to 105 µg/ml lysolecithin concentration, while for 37°C uptake is visible in cells down to 30 µg/ml concentration. At 37°C, most cells took up Trypan Blue when treated with 60 to 105 µg/ml lysolecithin. Trypan Blue was expelled at all concentrations in the two temperature conditions as the dye is not visible in cells 20 min later. Based on these results, I decided the optimal lysolecithin concentration for this cell type was between 60 and 105 µg/ml, with a 1 min incubation at 37°C instead of a cold shock. Lysolecithin was dissolved in chloroform for the imaging test and changed the secondary structure of the poly-L-lysine used to help FlpIn 293 T-REx cells attach thus forming bubbles that detached the cells and made imaging more challenging (Shibata *et al.*, 1991). Further experiments instead used lysolecithin dissolved in ethanol without any adverse effects on the poly-L-lysine. Lysolecithin concentrations were then tested whether they affected cell proliferation with a PrestoBlue assay. Cells were treated with 60 to 105 µg/ml of lysolecithin for 2 min at 37°C then incubated for 24 hr before addition of the PrestoBlue reaction reagent. Fluorescence was measured at 544/590 nm which revealed no difference in cell growth 24 hr after lysolecithin treatment, except for at the highest concentration of 105 µg/ml (Figure 4.13B, Wilcoxon test, $p = 0.039$).

A



B

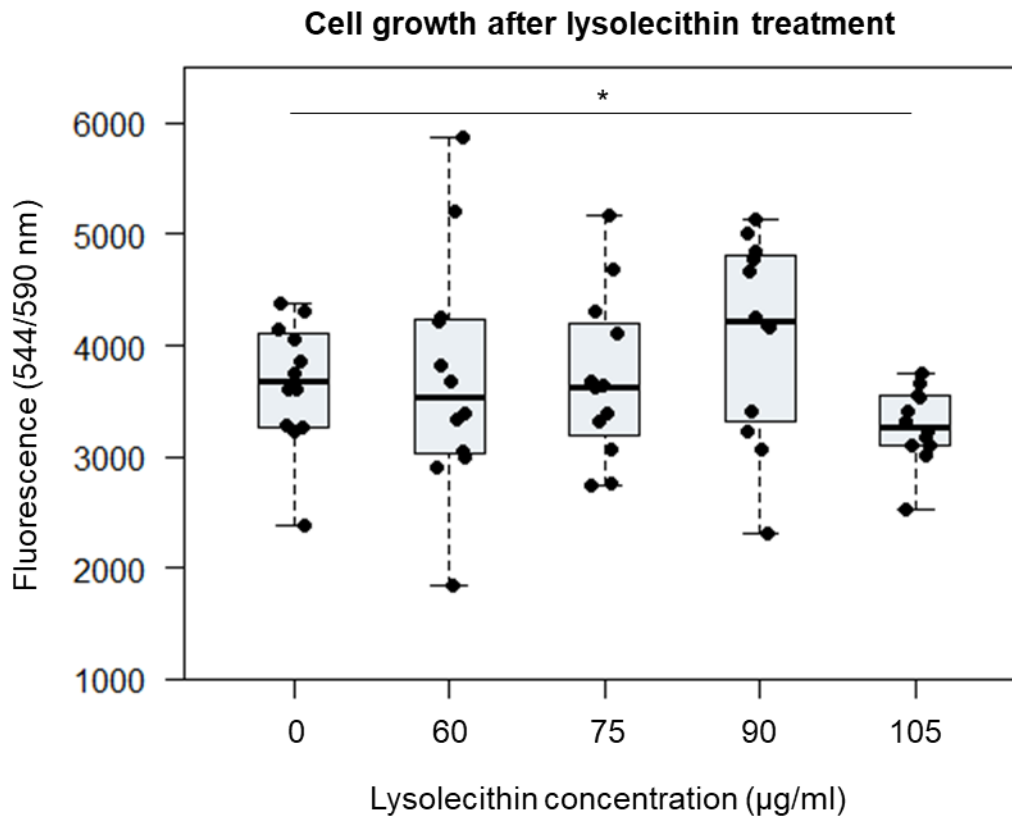
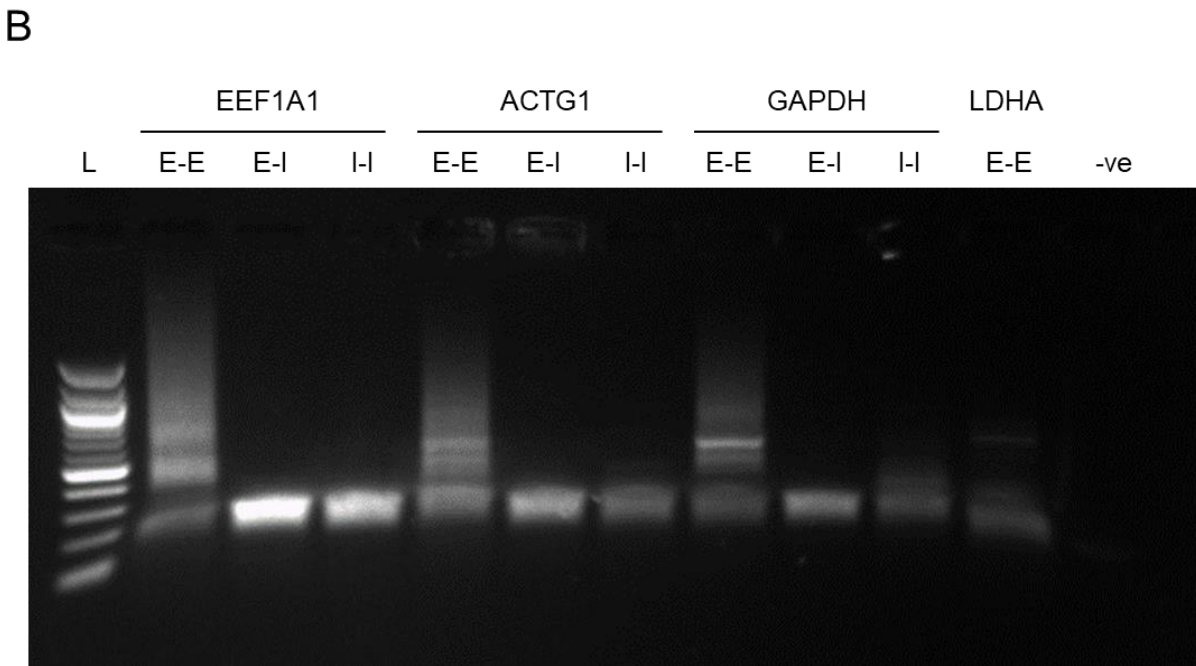
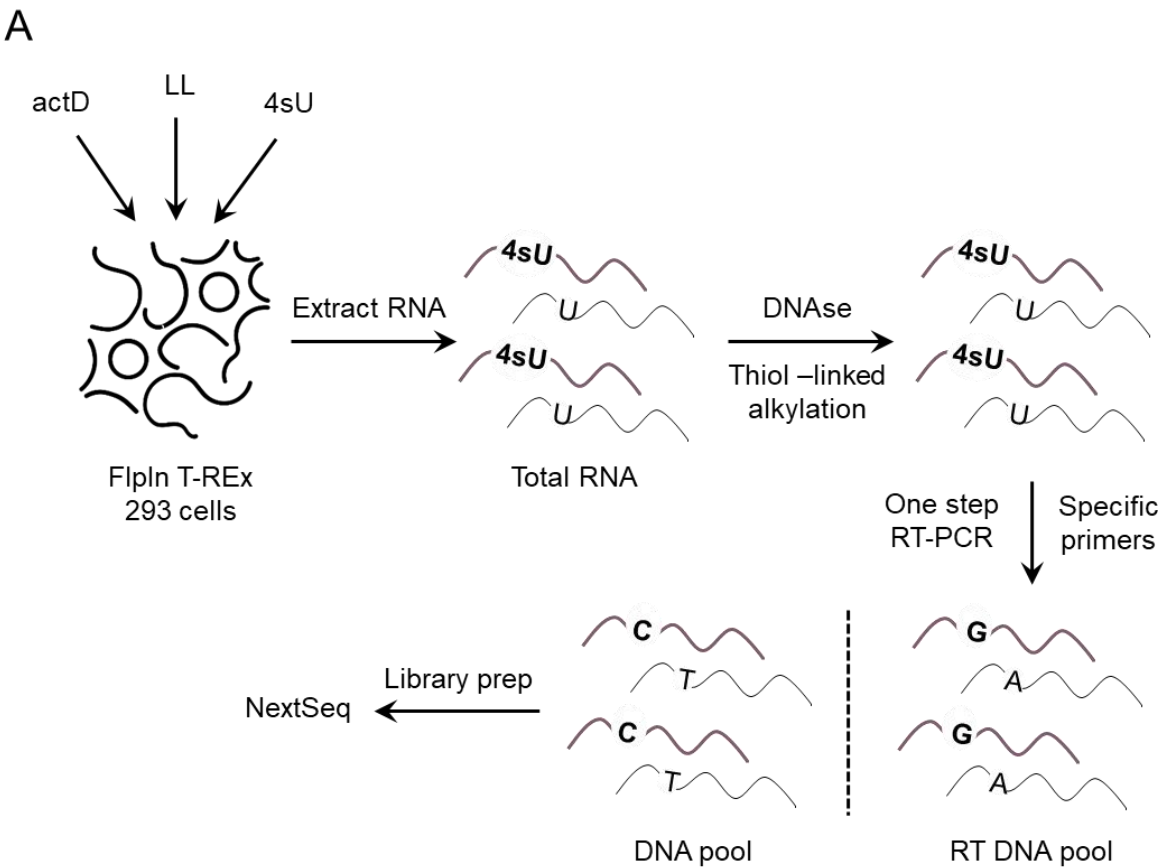


Figure 4.13. Lysolecithin trial. A) Brightfield images of FlpIn 293 T-REx cells treated with different concentrations of lysolecithin (0 to 105 µg/ml) at 4°C or 37°C then briefly treated with 0.1% Trypan Blue. Imaged after treatment with Trypan Blue and 20 min later. Scale bar is 20 µm B) Cell proliferation of FlpIn 293 T-REx cells 24 hr after treatment with different concentrations of lysolecithin shown as a box plot of fluorescence values. Values on the y-axis were taken with a plate reader measuring fluorescence at 544/590 nm wavelength following a 10-minute treatment with the PrestoBlue reagent. Each box plot consists of 12 technical replicates and shows the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the different conditions at each time point (Wilcoxon test, * = $p \leq 0.05$).

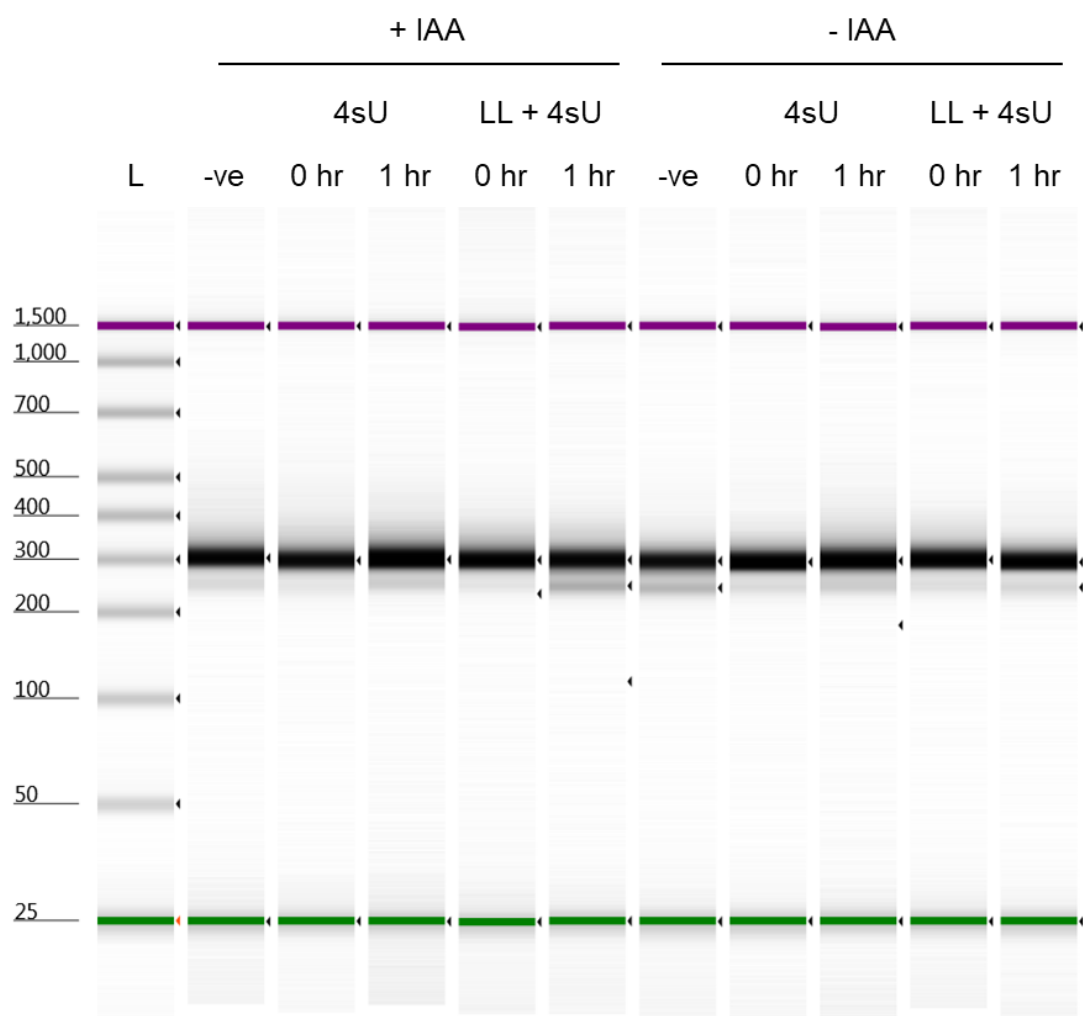
4.4.2. Pilot experiment

After optimising 4sU uptake, I conducted a pilot SLAM-seq experiment. FlpIn 293 T-REx cells were pre-treated with actD (to suppress rRNA production) and then incubated for 15 min with 4sU (Figure 4.14A). No 4sU addition was used as a negative control. Some cells were briefly permeabilised with 75 µg/ml lysolecithin before 4sU incubation. RNA was collected immediately after a 4sU pulse or after a 1 hr incubation and purified using a Qiagen column with an on-column DNaseI digestion. 10 µg of RNA from each condition was modified by thiol-linked alkylation in the presence of 2-iodoacetamide (IAA), with the remainder of the RNA treated with the reaction mix in the absence of IAA as a thiol alkylation negative control. Thiol-linked alkylation by IAA works by covalently attaching a carboxyamidomethyl group to 4sU which increases the bulk of the nucleotide to increase the likelihood of the reverse transcriptases to read the nucleotide as a C instead of a U. This results in a T-to-C conversion rate of over 94% when 4sU labelled transcripts are thiol modified and undergo RT-PCR (Herzog *et al.*, 2017). The chemically modified RNA was used in a one-step RT-PCR reaction with primers amplifying highly expressed genes. Genes were selected from gene expression data in BioGPS, a gene annotation and expression profile portal (Wu *et al.*, 2016). The three highest expressed genes for HEK-293 T-REx cells were eukaryotic translation elongation factor 1 alpha 1 (EEF1A1; 45000 to 65000 transcripts), γ -actin (ACTG1; 24000 to 26000 transcripts), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH; 13000 to 18000 transcripts). For each gene, three primer pairs were designed to amplify 150 to 176 bp regions in an exon, intron or spanning an exon-intron junction. Together with a positive control primer pair used in the lab for lactate dehydrogenase A (LDHA; 5th most highly expressed gene from HEK-293 T-REx dataset with 9000 to 15000 transcripts), I used the primers targeting highly expressed genes to amplify amplicons from control RNA in a test one-step RT-PCR reaction (Figure 4.14B). All primer pairs produced the expected amplicon products (150 to 176 bp) with higher bands visible (~500 bp) for exon-exon primers due to the amplification of pre-mRNA products that have included introns. The RT-PCR reaction was repeated with ten primer pairs for each condition. The amplicons for each condition were pooled together to prepare

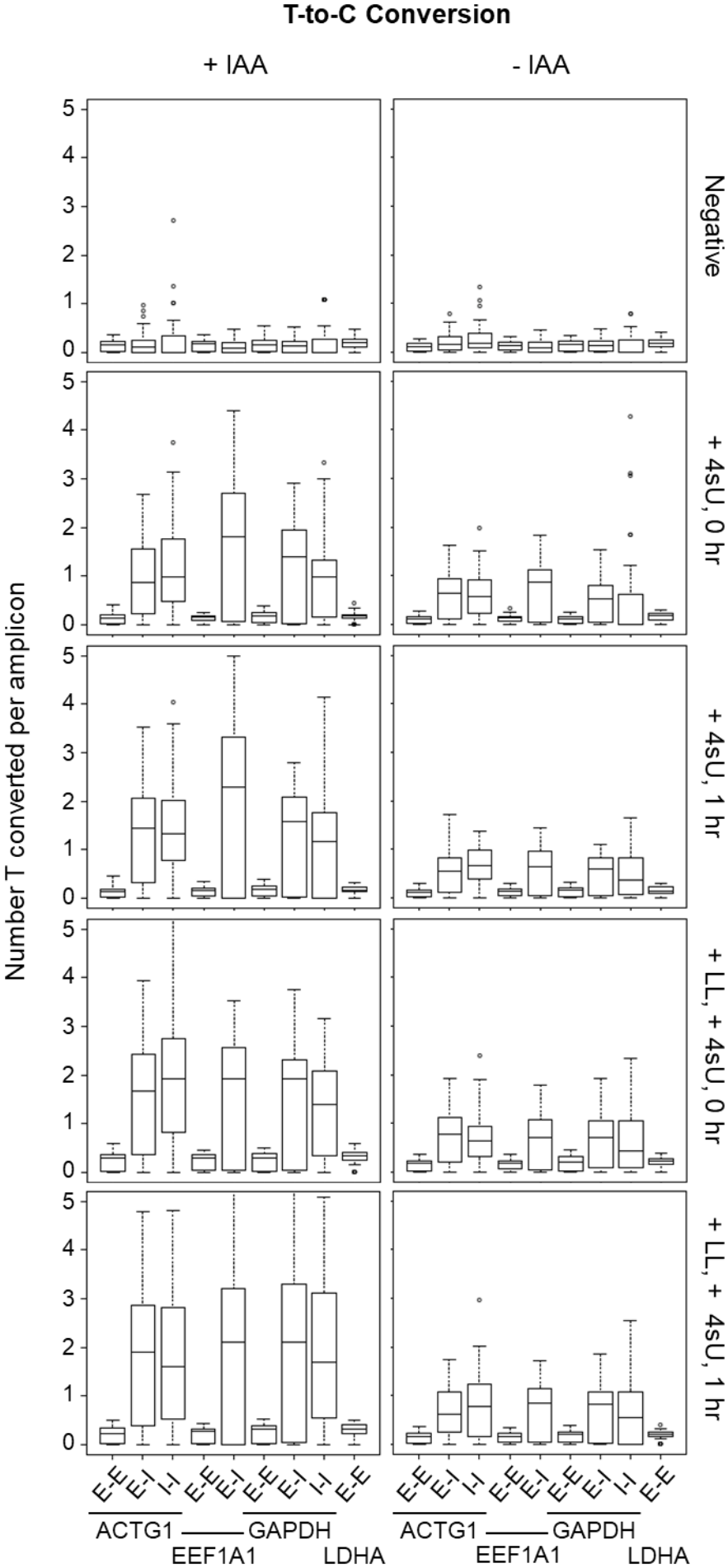
DNA libraries for sequencing which yielded a final library product at around 300 bp (Figure 4.14C). The final libraries were pooled together and sent for sequencing.



C



D



E

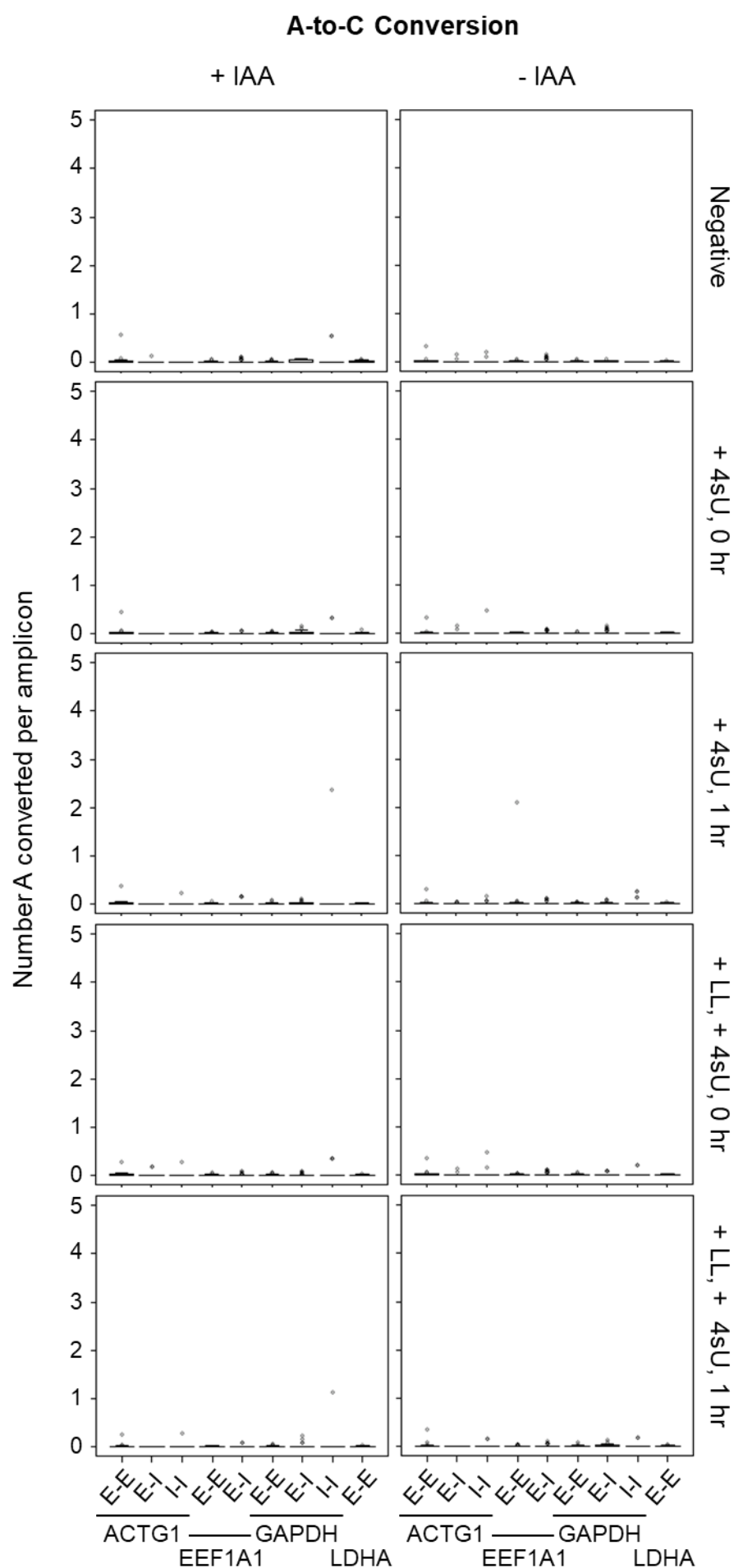


Figure 4.14. Nascent RNA labelling with 4sU and amplicon sequencing. A) Schematic representation of method workflow. B) Agarose gel of RT-PCR using specific primers for exon-exon (E-E), exon-intron (E-I) or intron-intron (I-I) amplicons of highly expressed genes. LDHA E-E as a positive control. The ladder used is NEB 1 kb Plus. The regions of interest expected size varies from 150 to 176 bp. C) TapeStation analysis of pooled amplicons for each cell condition after barcoding with a library prep kit. Final library size varies from 290 to 310 bp. D) Bar plots for amplicons under each cell condition that have a T-to-C conversion. E) Same as D) but for A-to-C conversion. Bar plots were generated in collaboration with Nick Gilbert.

Sample	IAA	EEF1A1			ACTG1			GAPDH			LDHA	Total
		E-E	E-I	I-I	E-E	E-I	I-I	E-E	E-I	I-I	E-E	
Negative	+	1699978	3884	0	314015	1651	600	454165	5455	372	214382	2694502
4sU 0 hr	+	1985059	3042	0	444772	910	1024	544115	5923	610	370539	3355994
4sU 1 hr	+	1080985	1246	0	455214	966	896	473314	11621	346	225600	2250188
LL + 4sU 0 hr	+	1410286	5319	0	393779	1076	728	313322	5105	580	257827	2388022
LL + 4sU 1 hr	+	923028	2207	0	369856	1052	750	272311	2889	358	195871	1768322
Negative	-	1110237	3998	0	394541	3819	2101	355255	11123	787	239327	2121188
4sU 0 hr	-	946500	4542	0	534046	2462	1734	443746	4419	334	249311	2187094
4sU 1 hr	-	865927	4393	0	378764	4929	3664	381367	13796	1610	176472	1830922
LL + 4sU 0 hr	-	1214893	4294	0	640532	3046	1264	447158	2558	963	393446	2708154
LL + 4sU 1 hr	-	982651	4217	0	467996	1960	1296	343319	7640	1136	345491	2155706

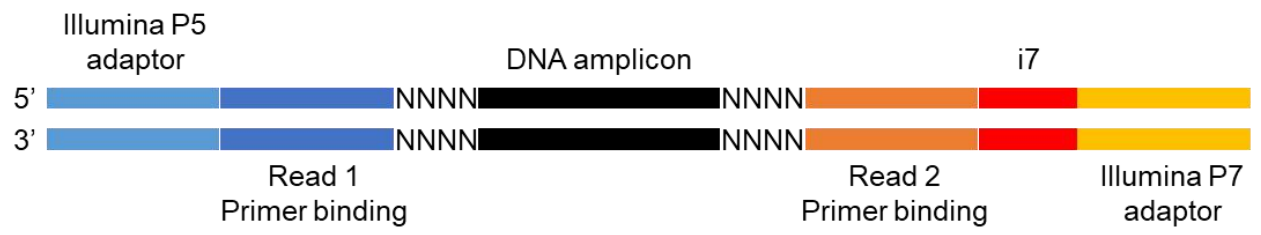
Table 4.1. Sequencing reads of amplicons for each 4sU treatment condition. Exon-exon (E-E), exon-intron (E-I) or intron-intron (I-I) amplicons. LL is lysolecithin, 4sU is 4-thiouridine, the time is collection after 4sU incubation, IAA is 2-iodoacetamide.

The sequencing reads were successfully mapped with bowtie2 (>98% per sample). They revealed an exon bias with exonic reads being on average 57- to 568-fold more than intronic reads (Table 4.1). There were 1.5-fold to 10-fold more exon-intron amplicons read than intron-intron. This revealed a sequencing bias, possibly due to more available gene exons than introns for amplification. Also, while there was a product visible via agarose gel electrophoresis for the intron-intron amplicon of EEF1A1, there were no sequencing reads registered in any condition (Figure 4.14B, Table 4.1). The number of T-to-C conversions per amplicon for each condition was calculated using the function `pileup2acgt` from the `Sequenza-utils` package for Python (Figure 4.14D). The negative controls showed a low base rate of T-to-C conversions in the absence of 4sU due to reverse transcriptase base reading error. The RNA samples which underwent thiol modification had a higher number of T-to-C conversions than the unmodified samples – on average 3- to 5-fold more. This was expected as in the absence of alkylation, transcripts labelled with 4sU had 10-11% of T-to-C conversions in previous research (Herzog *et al.*, 2017). This is why it is important to sequence the RNA in the absence of alkylation to establish a base rate of T-to-C conversions for 4sU. RNA collected 1 hr after 4sU incubation had more T-to-C conversions than RNA collected immediately after the incubation – on average 1.5- to 2-fold more. Additionally, newly made intronic RNA had the highest number of T-to-C conversions compared to exonic at both time points. These results indicate that RNAs specific only to the three highly expressed genes might have different half-lives between their exons and introns that are further investigated in Section 4.4.4. Cells that were treated with lysolecithin before 4sU incubation had on average 1.5- to 2-fold more T-to-C conversions per amplicon when comparing the same time points. This suggested that a brief permeabilisation with lysolecithin does increase 4sU uptake in FlpIn 293 T-REx cells. When looking at A-to-C conversions as a comparison there were no differences between the conditions (Figure 4.14E). Overall, the pilot experiment showed successful 4sU incorporation that can be used as a readout for newly synthesised RNA.

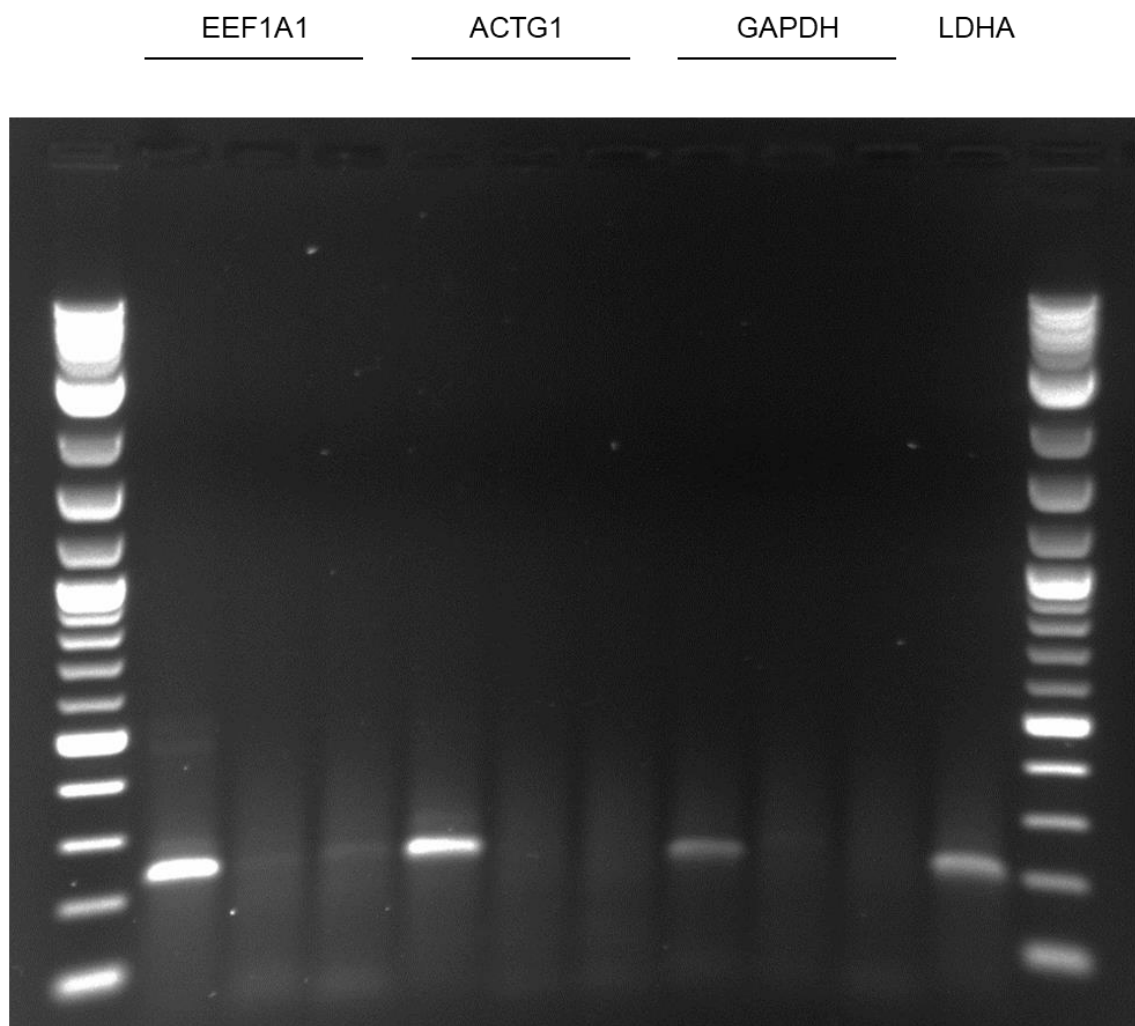
4.4.3. Comparison between different thiol modification methods

The difference in the number of T-to-C conversions between samples treated with IAA and untreated samples was 3- to 5-fold more in the pilot experiment which was less than the 9-fold difference reported in the literature (Herzog *et al.*, 2017). I, therefore, wanted to test a different chemical conversion method and compare it to thiol alkylation with IAA. Another method is to use TFEA, an amino nucleophile, with NaIO₄, an oxidant, which convert 4sU to N4-trifluoroethylcytosine recognized by reverse transcriptase as a cytosine (Schofield *et al.*, 2018; Qiu *et al.*, 2020). I treated FlpIn 293 T-REx cells with actD for 1 hr, then a brief lysolecithin permeabilisation before 4sU incubation for 15 min. RNA was collected from these cells with TRIzol-chloroform extraction immediately after 4sU treatment or after a 2-hr incubation. The RNA was purified, treated with DNaseI to remove contaminant DNA molecules, then 10 µg of each condition and a negative (no 4sU) control underwent either thiol-linked alkylation with IAA treatment or aminolysis and oxidation with TFEA and NaIO₄ treatment. New primer pairs targeting the three most highly expressed genes and LDHA in FlpIn 293 T-REx cells were developed to include unique molecular identifiers (UMIs) – four random bases before and after each amplicon (Figure 4.15A). The primers also included consensus sequences for binding to the second set of primers required for sequencing which also have a barcode (i7). The barcodes were used to identify each condition when the final DNA libraries were pooled together. These new primer sets were used for one-step RT-PCR of the chemically modified and unmodified RNA samples. Unlike in the pilot experiment, the input RNA for each primer pair was balanced by using 500 ng for intronic primers and 50 ng for exonic primers. All primer pairs produced the expected amplicon products (208 to 234 bp) when separated on an agarose gel (Figure 4.15B). The intronic amplicons had a weaker signal than exonic amplicons. The amplicons for each condition were pooled together to prepare DNA libraries for sequencing which yielded a final library product from 290 to 310 bp and after final pooling of all samples, the libraries were sequenced.

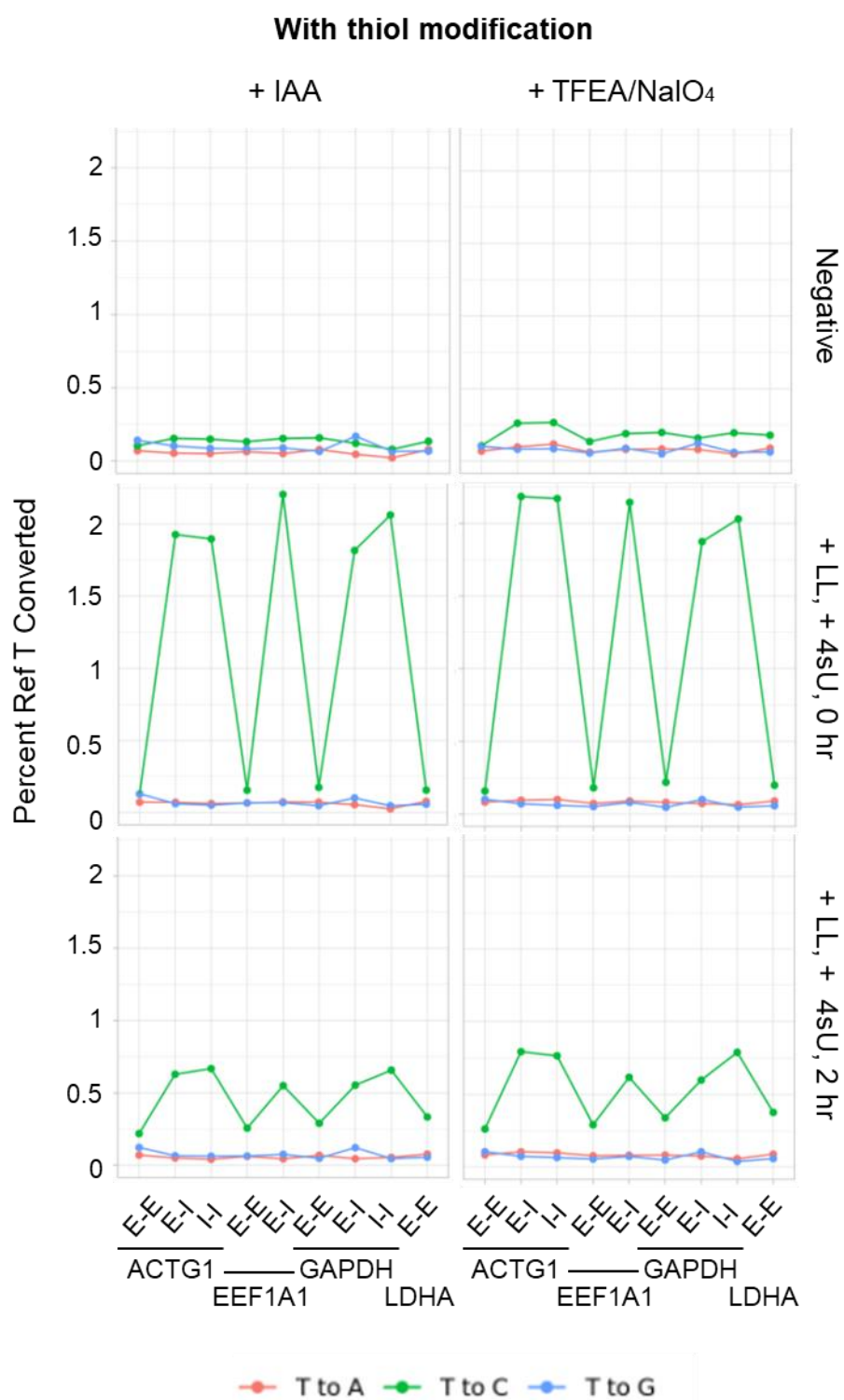
A



B



C



D

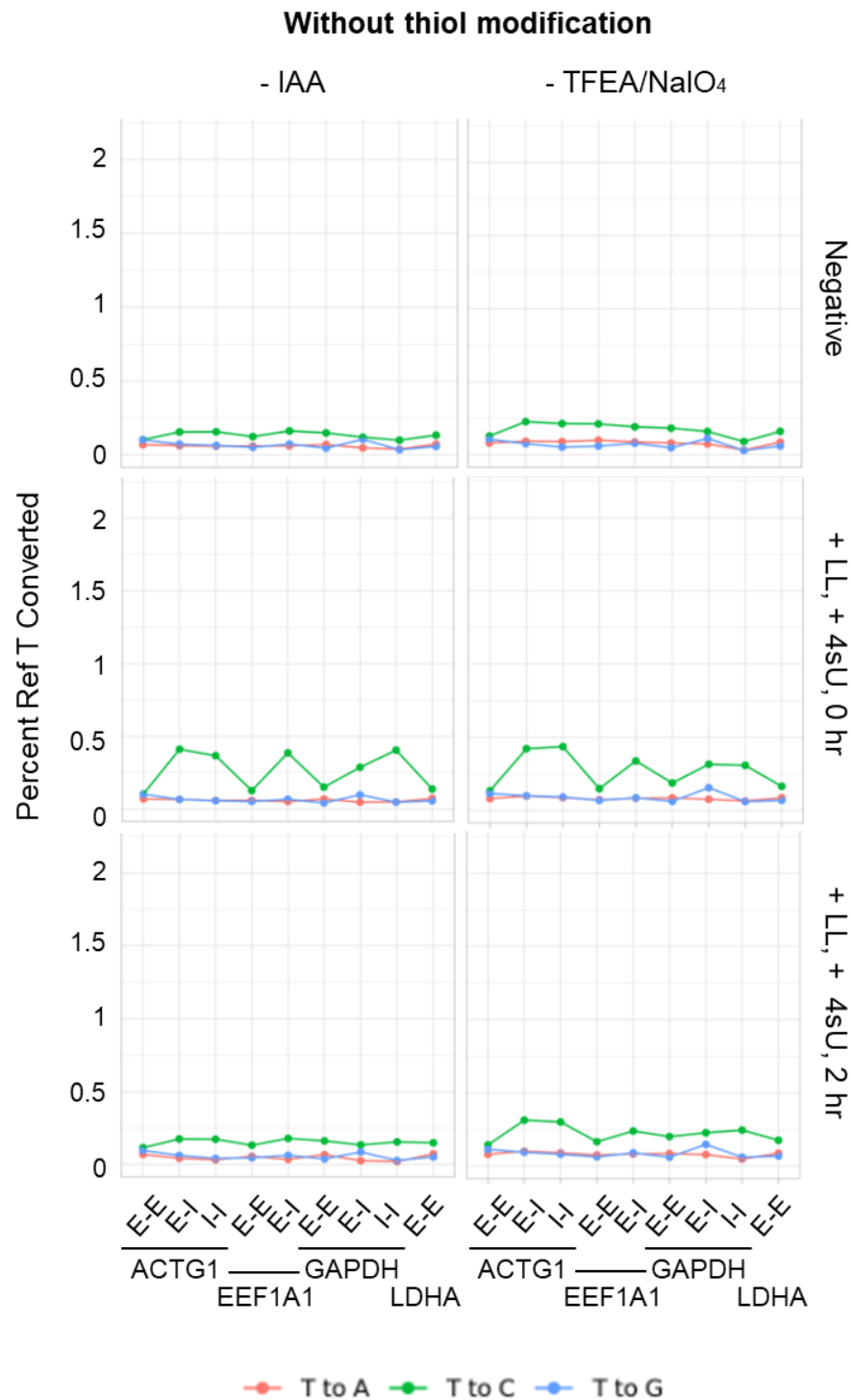


Figure 4.15. Comparison of thiol modification methods. A) Schematic representation of library products made with custom primers. B) Agarose gel of RT-PCR using specific primers for exon-exon (E-E), exon-intron (E-I) or intron-intron (I-I) amplicons of highly expressed genes. LDHA E-E as a positive control. The ladder used is NEB 1 kb Plus. The regions of interest expected size varies from 208 to 234 bp. C) Line plots for amplicons with average percentage of T converted to C, A, or G after thiol modification with different methods under different cell conditions. D) Same as C) but without the active reagents for thiol modification added. Line plots were generated in collaboration with Graeme Grimes.

The sequencing results again showed a read number imbalance between exons and introns and again there were no reads for EEF1A1 intron-intron amplicons even when there was a weak band visible in the agarose gel (data not shown; Figure 4.15B). The percentage of T-to-C conversions per amplicon for each condition was calculated using the readr package in R (Figures 4.15C and 4.15D). There were no significant differences in percentage of T-to-C conversions between thiol alkylation with IAA or thiol modification with TFEA/NaIO₄ (Figure 4.15C). As the latter reagents are highly volatile and toxic, and offer no advantages over IAA, further experiments were done using thiol alkylation by IAA. Consistent with the pilot experiment, there is a base rate of T-to-C conversion in the absence of 4sU and without thiol modification treatment (Figure 4.15D). None of the conditions had an increase in T-to-A or T-to-G conversions which confirms that 4sU incorporation affects only T-to-C conversions. Also consistent with the pilot experiment, the intronic RNA showed the highest percentage of T-to-C conversions compared to the exonic RNA in all conditions due to the read number imbalance leading to more exonic reads than intronic. When RNA was treated with thiol modification reagents, the highest percentage of T-to-C conversions are seen at 0 hr after 4sU incubation compared to 2 hr later (Figure 4.15C). The 2 hr time point coupled with 0 hr and 1 hr could allow us to follow nascent RNA dynamics, therefore I used these time points in the next section.

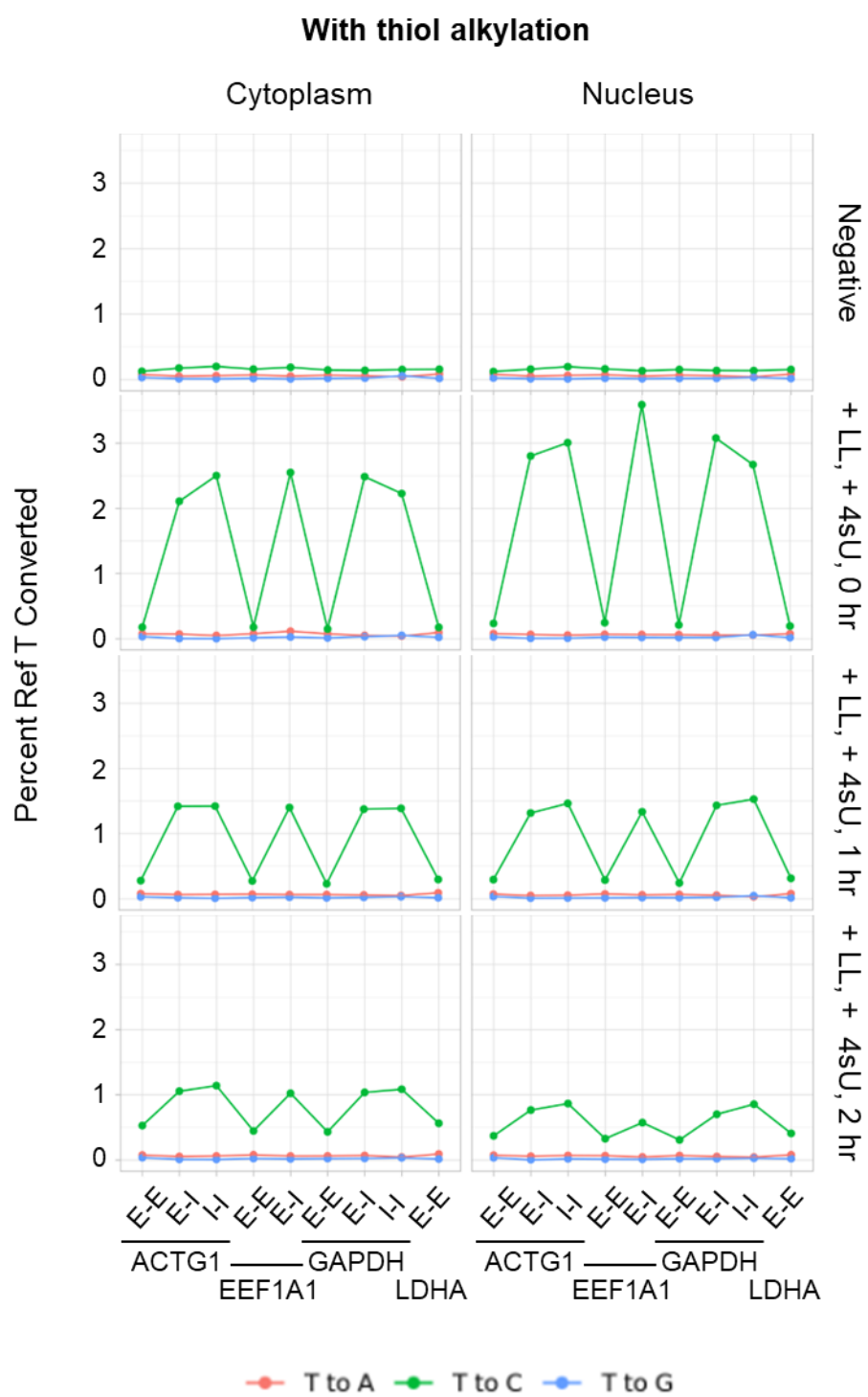
4.4.4. Nascent RNA dynamics

My previous experiments showed an interesting difference in T-to-C conversions in introns and exons over time where they reduced over time in introns and pre-mRNAs (exon-intron) but increased for exons (mRNA). Therefore, I supposed that using SLAM-seq on amplicons of highly expressed genes would allow us to understand the RNA dynamics of those genes which I hypothesised would differ in the nucleus and in the cytoplasm. This is important as the introns of these highly expressed genes might contribute to ncRNA nuclear organisation involved in large-scale chromatin organisation and following their dynamics could help us understand how ncRNA structures are formed (Creamer *et al.*, 2021). To achieve this, I performed an experiment focusing on nascent RNA at 0, 1 and 2 hr after 4sU incubation and

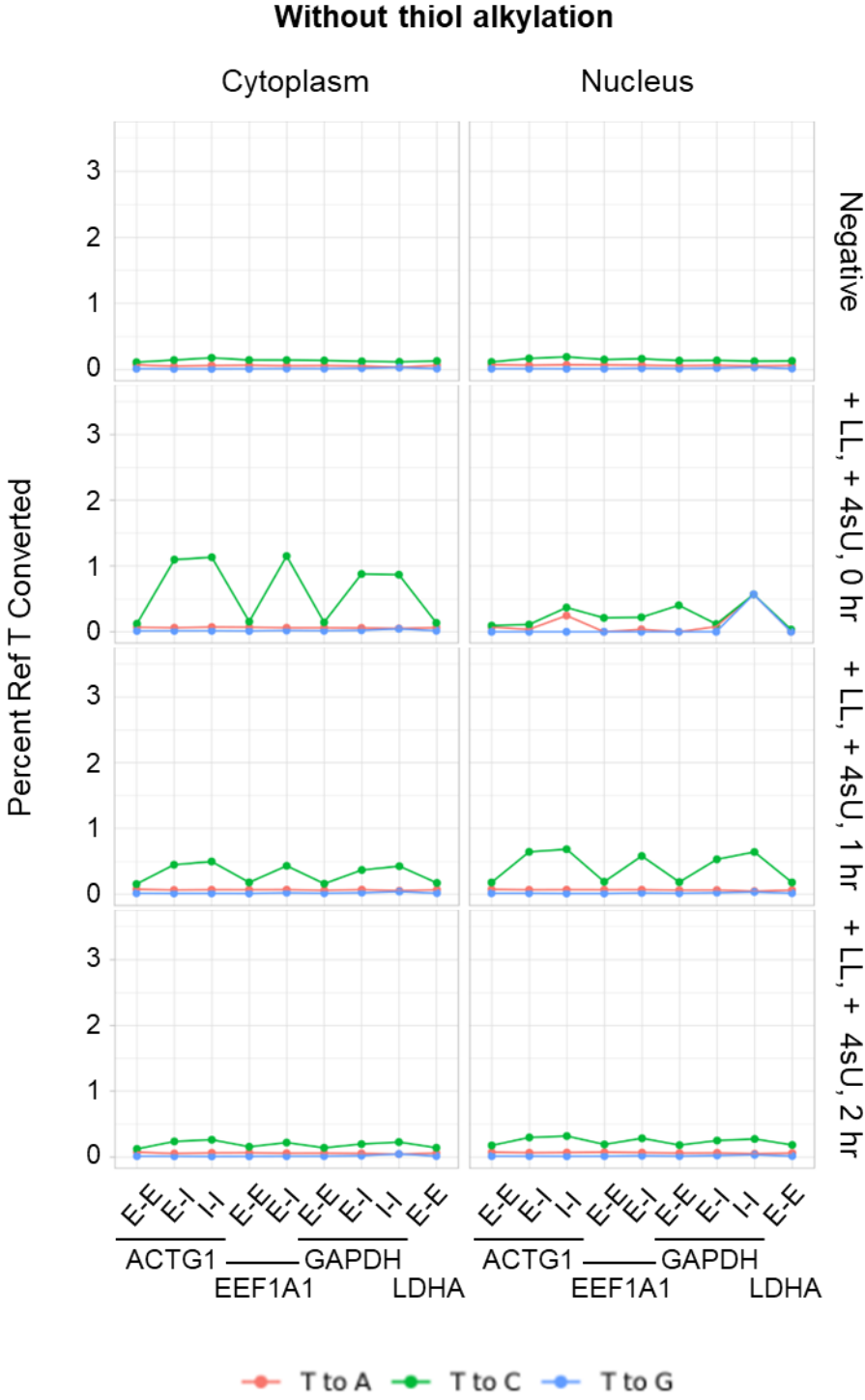
separated nuclei from the cytoplasm to differentiate nuclear from cytoplasmic RNA. First, FlpIn 293 T-REx cells were treated with actD for 1 hr, permeabilised with lysolecithin and then incubated with 4sU for 15 min. At 0, 1 and 2 hr cells were trypsinised and then washed in a mild hypotonic buffer with detergent to slightly swell the cells and lyse the cell membrane whilst keeping the nuclear membrane intact. After incubation on ice, nuclei were centrifuged and the supernatant was collected as a cytoplasmic fraction, while the cell pellet was resuspended to be used as a nuclear fraction. RNA was collected from these fractions with TRIzol-chloroform extraction and then treated with DNaseI to remove contaminating DNA molecules. 20 µg of RNA from each condition and a negative sample (no 4sU) was treated with IAA to induce thiol-linked alkylation of the 4sU bases. Primers encoding UMIs were used for one-step RT-PCR of the chemically modified and unmodified RNA samples. In this experiment, the input RNA for each primer pair was balanced by using 1000 ng for intronic primers compared to 5 ng for exon-exon EEF1A1, 15 ng for exon-exon ACTG1 and GAPDH, and 24 ng for LDHA. The amplicons for each condition were pooled together to prepare DNA libraries for sequencing which yielded a final library product from 290 to 310 bp. The final libraries were pooled together and sequenced.

Analysis of the reads after sequencing showed a balanced number of reads for each amplicon and the percentage of T-to-C conversions per amplicon for each condition was calculated using the readr package in R (Figures 4.16A and 4.16B). For all timepoints the T-to-C conversion percentages were similar between cytoplasmic and nuclear RNA except in the case of 0 hr nuclear RNAs where there was a 1.2- to 1.4-fold increase in the percentage of T-to-C conversions compared to cytoplasmic RNAs. It was surprising that cytoplasmic and nuclear samples had highly similar T-to-C conversion percentages as I expected that the percentages should be reduced for introns and exon-intron amplicons in the cytoplasm due to splicing occurring in the nucleus before mRNA export to the cytoplasm. One possible explanation is that the nuclear fraction is contaminated with residual mRNA potentially attached to the nuclear membrane via the endoplasmic reticulum.

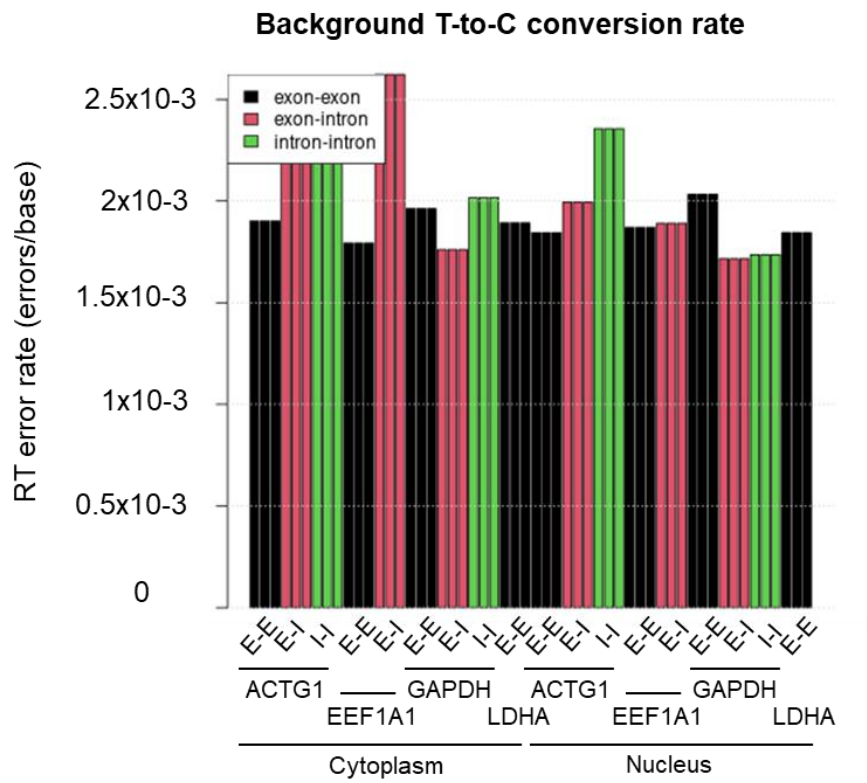
A



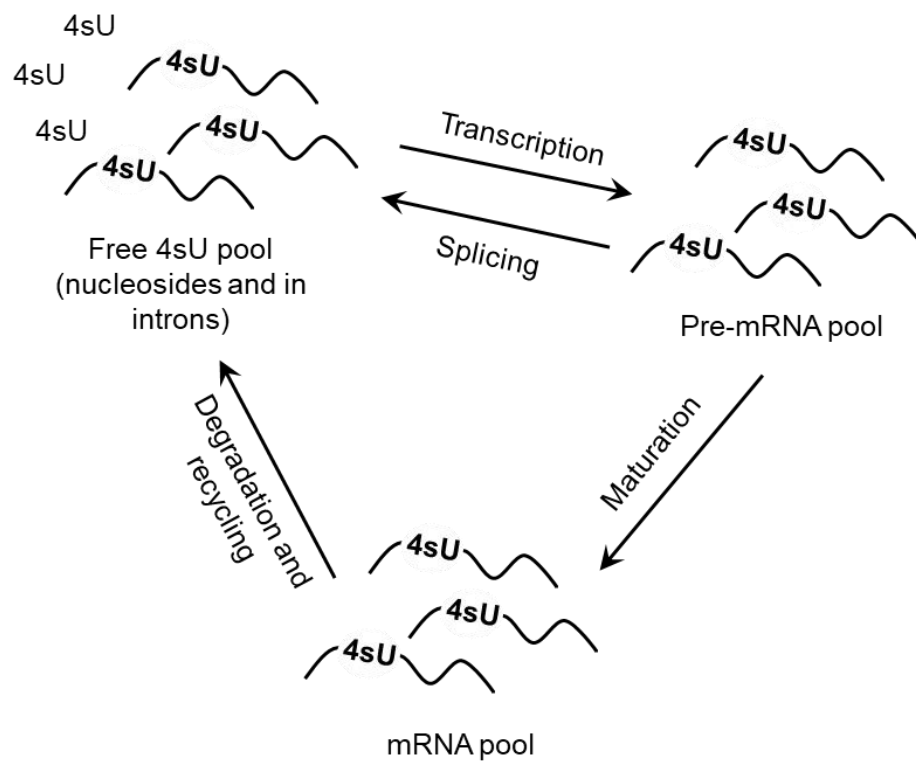
B



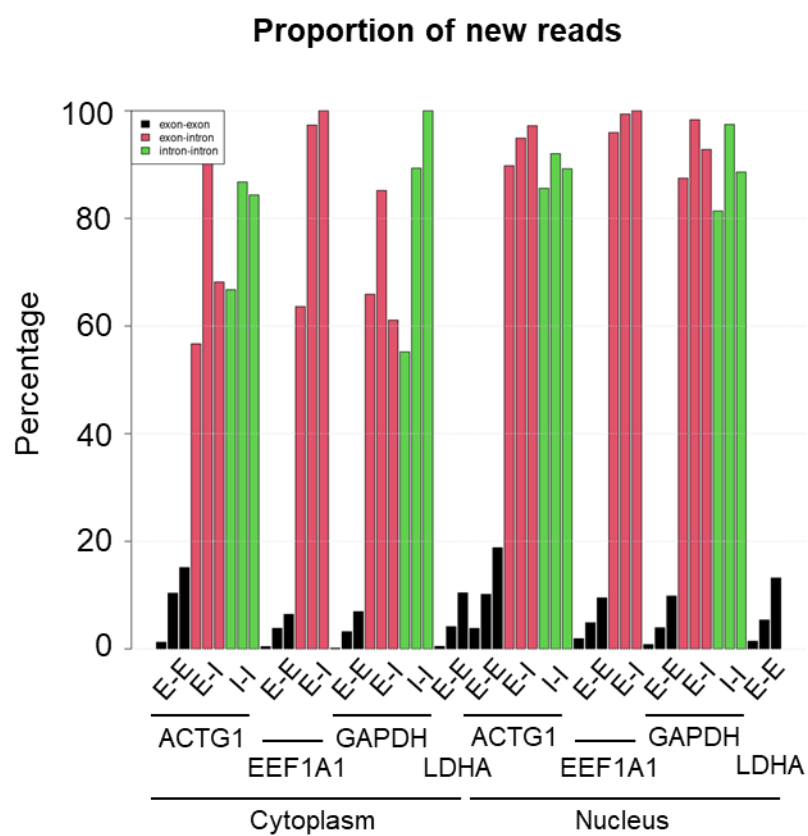
C



D



E



F

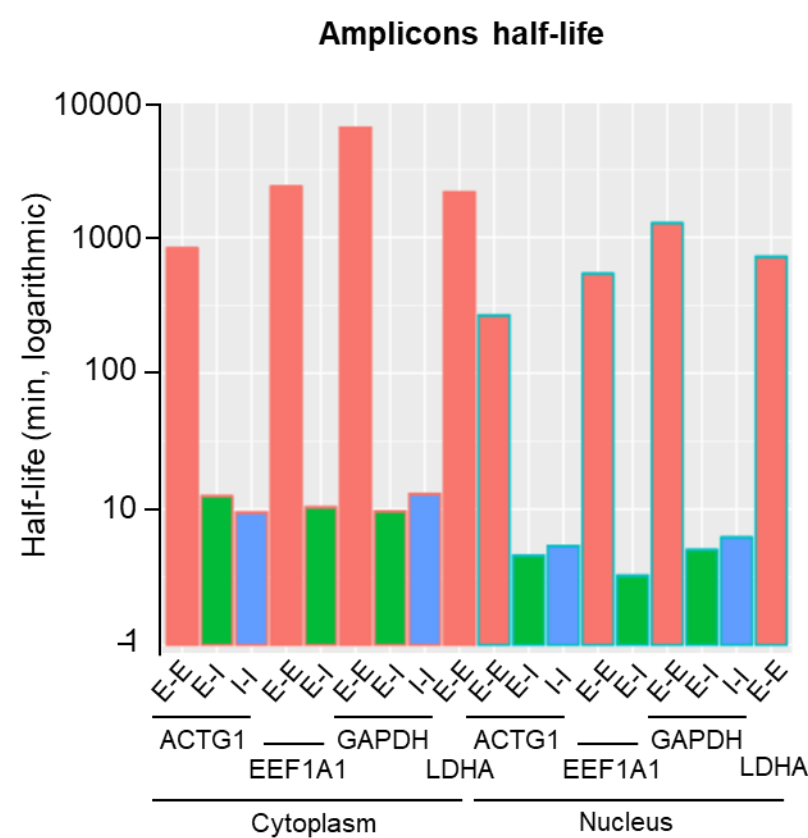


Figure 4.16. Amplicon sequencing of RNA in the cytoplasm and nucleus. A) Line plots for amplicons with average percentage of T converted to C, A, or G after cell fractionation and thiol alkylation under different cell conditions. B) Same as A) but without the active reagent for thiol alkylation added. C) Bar plot of background T-to-C conversion rate calculated from samples without 4sU incorporation and thiol alkylation. No timepoints, therefore the three bars per condition are the same. D) Schematic representation of 4sU and RNA kinetics in cells. E) Bar plot of the proportion of new RNA molecules calculated by subtracting the probability of background T-to-C conversions from the probability of 4sU incorporation after fitting a Poisson mixture distribution. For each amplicon, the first bar is the 0 hr post-4sU treatment value, then 1 hr and 2 hr. F) Bar plot of amplicon half-life calculated by using the amount of old and new amplicons at 0 hr in the presence of 4sU and thiol alkylation. Line plots were generated in collaboration with Graeme Grimes and bar plots were generated in collaboration with Hannes Becher.

Analysis of the sequencing data allowed me to make several conclusions. Firstly, as expected, there was an increase in T-to-C conversion when cells were incubated with 4sU, with a bigger increase when the RNA also underwent thiol alkylation (Figure 4.16A). Secondly, there were no increases in T-to-A or T-to-G conversions except in the case of the nuclear sample collected 0 hr after 4sU treatment that did not undergo thiol alkylation. That sample had an uncharacteristically low overall read number - 663 total reads compared to >2,400,000 for other conditions – which might explain the differences in base conversion. Thirdly, like in the previous two experiments, there is a low-level T-to-C conversion in the absence of 4sU and without thiol modification treatment (Figure 4.16B). This result was used to obtain a background estimate of T-to-C conversions due to reverse transcriptase error by fitting a Poisson distribution (Figure 4.16C). This background error rate was estimated to be 2×10^{-3} errors/base on average. Finally, similar to the pilot and thiol modification comparison experiment, the intronic RNAs in this experiment showed the highest percentage of T-to-C conversions compared to the exonic RNA even when the read imbalance between RNA origin was addressed. This was potentially due to most exon reads consisting primarily of mature mRNA which would have the least T-to-C conversions due to being present in the cells from before 4sU incubation (old reads) while intronic reads consisted mostly of new RNAs labelled with 4sU (new reads; Figure 4.16D). At later timepoints in the experiment (1 and 2 hr), the percentage of T-to-C conversions for exonic reads slowly increased which showed that 4sU is recycled and incorporated into new mRNAs as previous mRNAs and free introns are degraded and recycled. This process can be modelled using a Poisson distribution and used to calculate the percentage of new RNA reads. Additionally,

there was a low-level T-to-C conversion due to reverse transcriptase error rate which was used to produce a background error estimate with a Poisson distribution. Therefore, the number of T-to-C conversions in new reads was a mixture of two Poisson distributions due to nascent RNA reads forming two populations – one from 4sU incorporation and from reverse transcriptase error. To determine the proportion of new reads and old reads for intronic and exonic RNAs, a Poisson mixture model was used (Figure 4.16E). Between 50% to 100% of cytoplasmic intron reads and 80% to 100% of nuclear intron reads were attributed to new molecules while exon reads varied from 0% to 20%. The proportion of new exon and cytoplasmic intron molecules generally increased over time with the highest values seen at 2 hr post-4sU incubation, except in the case of ACTG1 exon-intron and GAPDH exon-intron reads. In comparison, the new nuclear intron molecules remained at similar levels over time. The differences in the proportion of new reads between cytoplasmic and nuclear samples was further investigated by estimating the half-life of amplicons using the remaining proportion of old reads at 0 hr after a 15 min 4sU incubation (Figure 4.16F). Exon reads had the highest half-life ranging from 829 to 6419 min in the cytoplasm and from 267 to 1268 min in the nucleus. Intron reads had much shorter half-lives than exons, ranging from 9 to 13 min in the cytoplasm and from 3 to 6 min in the nucleus. The differences between amplicon half-life in the cytoplasm and nucleus was statistically significant only for introns showing that while the T-to-C conversion percentages were similar in both compartments, there were differences in RNA kinetics (Wilcoxon test, $p = 0.008$).

4.5. Discussion

In this chapter, I aimed to investigate the connection between SAF-A and RNA organisation by examining whether SAF-A depletion affects the RNA mesh. To achieve this, I aimed to integrate a mAID degron tag with SAF-A to achieve faster degradation levels with less off-target effects than RNAi to check the state of chromatin organisation and to use mutant versions of SAF-A. While mAID-SAF-A was responsive to IAA treatment when transiently transfected, it was not responsive in a stably generated cell line from HCT 116 Tet-OsTIR1 cells (Figures 4.1A and

4.3A). Furthermore, the tagged alleles showed greatly reduced SAF-A expression levels which might make the homozygous cells less viable than the heterozygous cells (Figure 4.2C and 4.2D). Therefore, if the population of cells wasn't fully monoclonal, the heterozygous cells could outcompete the homozygous cells which can show on a Western blot as loss of the tag from one SAF-A allele (Figure 4.4B). In the literature, another mAID-tagged SAF-A cell line seemed to suffer from instability and slightly reduced expression levels (Sharp *et al.*, 2020). It is unclear whether the tag itself leads to protein instability or whether it is the presence of auxin in cell media that constitutively reduces the tagged protein levels e.g., leaky degron (Natsume *et al.*, 2016). The latter might be the case in the constitutively expressed OsTIR1 cells from the Stuart Wilson lab (Figure 4.4A; Heath, 2018). It shouldn't be the case for the cell lines I generated as the OsTIR1 gene is under tetracycline repression and should only be expressed when doxycycline is added. However, recent research showed that the high dose of IAA required for degradation was toxic in some cell types and a mutated version of the OsTIR1 gene (F74G) was developed that used a much lower concentration of a 5-phenyl-indole-3-acetic acid (5-Ph-IAA) instead (Yesbolatova *et al.*, 2020). Therefore, a stably generated cell line using cells expressing the mutated OsTIR1 gene might circumvent the leaky degron, basal degradation, and IAA toxicity of the current method. The dTAG system also circumvents these issues as it uses the native proteasome and a low dose of a synthetic molecule to achieve degradation but in my hands SAF-A was not degraded when fused to the FKBP12^{F36V} tag (Figures 4.5B and 4.5C). The tag did work on XRN2 though which might be of further use in the Gilbert lab. Due to the issues with degron tags, I instead utilised RNAi in other experiments to check the RNA organisation when SAF-A is depleted (Figure 4.11) and the chromatin structure in the absence of SAF-A and the presence of mutants (Chapter 5.3.3.). According to previous RNAseq experiments, 24-hr siSAF-A treatment has no off-target effects (Nozawa *et al.*, 2017). When I tested this length of siSAF-A KD, it produced a statistically significant difference in the gene-rich 11p15.1 region in both RPE-1 and HCT 116 cells (Figure 4.6) which was comparable to the 48-hr KD in previous literature which indicates that for future experiments SAF-A can be knocked down using a 24 hr RNAi treatment (Nozawa *et al.*, 2017).

In this chapter, the dSTORM super-resolution microscopy technique was employed to visualise the distribution of SAF-A with higher resolution compared to conventional

light microscopy methods. RPE-1 cells were imaged using antibodies for SAF-A, HNRNPC or SC-35, and data analysis with the SuperStructure method unveiled the presence of clusters formed by these proteins, exhibiting varying sizes and interconnectivity (Marenda *et al.*, 2021; Figure 4.7). Notably, SAF-A formed denser and more interconnected clusters compared to HNRNPC, another HNRNP family member, despite both proteins exhibiting diffuse patterns of nuclear distribution when imaged using traditional light microscopy (Figures 3.2B and 3.5B). The differences in size and connectivity between SAF-A and HNRNPC clusters might be attributed to their different abilities to oligomerise. Previous studies have shown that SAF-A can form high molecular weight oligomers, that are important for transcriptionally active regions' decompaction, in the presence of caRNAs and ATP while HNRNPC forms tetramers that play a role in RNA maturation (Barnett *et al.*, 1989; Cieniková *et al.*, 2015; Nozawa *et al.*, 2017). SC-35, an important part of the spliceosome and RNA maturation, has not been reported undergo oligomerisation. However, SC-35 has been shown to localise to nuclear "speckle" domains enriched in other splicing components and RNA. These speckles likely correspond to the highly connected dense clusters observed with dSTORM microscopy while the less connected clusters of SC-35 might be free protein that is not associated with other splicing components (Moen *et al.*, 1995; Figure 4.7B). Furthermore, we investigated the effect of transcription inhibition on SAF-A's nuclear distribution. Previous studies have demonstrated that the removal of most nascent RNA by transcription inhibition with α -amanitin or flavopiridol released the insoluble SAF-A into soluble fractions of SAF-A and treatment with α -amanitin reduced SAF-A immunofluorescence levels (Nozawa *et al.*, 2017). In contrast, treatment with triptolide did not visibly decrease SAF-A levels but resulted in reduced clusters size and connectivity (Figure 4.8). Since SAF-A requires interaction with RNA to form oligomers, the decrease in cluster size and connectivity could be attributed to the loss of oligomeric SAF-A. This notion is supported by FRAP and FCS experiments conducted by the Gilbert lab, which demonstrated that transcription inhibition increased SAF-A's diffusivity (data not shown). Another explanation for the reduction in SAF-A cluster size and connectivity could be that RNA scaffolds determine SAF-A's localisation in the nucleus and loss of RNA releases the protein from chromatin as seen when cells are treated with RNases for 1 min or 10 min (Nozawa *et al.*, 2017; Creamer *et al.*, 2021). The impact of transcription inhibition on HNRNPC is currently unknown and would be an interesting

comparison to SAF-A. Additionally, transcription inhibition has been shown to not affect the characteristic SC-35 speckles observed with traditional light microscopy possibly due to the persistence of mRNA transcripts within speckles even after prolonged transcriptional arrest (Shopland *et al.*, 2002; Nozawa *et al.*, 2017).

Before visualising nascent RNA distribution with super-resolution dSTORM microscopy, fixed RPE-1 cells were incubated with 5-EU for 15 min and imaged at different time points after incubation to track the lifecycle of nascent RNA (Figure 4.9B). The strongest nascent RNA signal was observed 30 to 60 min after 5-EU incubation, consistent with previous experiments involving 5-EU pulsing (Nozawa *et al.*, 2017). To track nascent RNA using dSTORM, RPE-1 cells were incubated with 5-EU for 15 min and immediately imaged or imaged after 0.5 or 1 hr (Figure 4.10A). Analysis of the data using the SuperStructure algorithm revealed that immediately after transcription (0 hr), nascent RNA exhibited a different clustering pattern and a larger cluster radius compared to RNA 30 to 60 min later (0.5 and 1 hr; Figure 4.10B). Additionally, nascent RNA clusters showed higher connectivity at the later time points compared to 0 hr (Figure 4.10C). The increase in RNA cluster radius over time suggests that nascent RNA continues to aggregate after transcription, forming larger and more connected clusters possibly through electrostatic interactions and liquid-liquid phase separation (LLPS; Henninger *et al.*, 2021). Thus, the differences in clustering patterns at 0 hr compared to later timepoints could correspond to initial condensate formation during transcription initiation. It is worth noting that if the experiment was conducted 2 to 3 hr after 5-EU incubation, it might have revealed a reduction in cluster radius and connectivity, and potentially a different clustering pattern or lack of as RNA condensates dissolve after reaching a high level of RNA concentration. However, this hypothesis, does not consider RNA degradation by XRN2 or the nuclear exosome, which could contribute to the reduction in nascent RNA clusters (Tomecki *et al.*, 2010; Turowski and Tollervey, 2015). Interestingly, SAF-A predominantly co-localises with nascent RNA at the 0.5 hr time point, followed by the 0 hr and 1 hr time points when analysed using the DBSCAN clustering algorithm (Figure 4.10E). Therefore, although nascent RNA exhibited the strongest signal at 0.5 and 1 hr, along with similar clustering patterns, size, and connectivity, it may only interact with SAF-A at the earlier time point. This finding contradicts previous research where immunoprecipitated SAF-A showed increasing levels of 5-EU

labelled RNA that plateaued 2-3 hr after pulse labelling (Nozawa *et al.*, 2017). Therefore, the co-localisation of SAF-A with nascent RNA at 0.5 and 1 hr after 5-EU pulse labelling might necessarily not correlate with RNA interaction. It would be interesting to repeat the dSTORM experiment with 2 and 3 hr timepoints after 5-EU pulse labelling to investigate for co-localisation between SAF-A and RNA at those timepoints. Previous studies have indicated that SAF-A depletion does not appear to affect the quantity of RNA in the nucleus, while transcription inhibition affects the organisation of SAF-A (Nozawa *et al.*, 2017). To further investigate the effect of SAF-A depletion on the nascent RNA mesh, dSTORM microscopy was used, focusing on RNA 1 hr after 5-EU pulse labelling which exhibited a strong signal and an increased cluster radius and connectivity compared to 0 hr. Cells partially depleted of SAF-A after 24 hr of RNAi treatment showed reduced RNA cluster connectivity compared to control cells, suggesting SAF-A's involvement in nascent RNA organisation (Figure 4.11B). Interestingly, SAF-A had the lowest co-localisation with nascent RNA at 1 hr after 5-EU treatment in my research but had a high level of 5-EU signal in bound RNA in previous research (Nozawa *et al.*, 2017). It is possible that the loss of SAF-A reduced RNA phase separation and its connectivity, but this hypothesis requires further investigation by depleting SAF-A from cells and visualising nascent RNA at additional timepoints after 5-EU pulse labelling.

In this Chapter, I aimed to investigate the binding preference of SAF-A to nascent RNA through sequencing methods. To accomplish this, I used 4sU labelling of nascent RNA in combination with SLAM-seq thiol alkylation. This approach introduces T-to-C conversions in nascent RNA, allowing differentiation between newly synthesised or “new” RNA (after 4sU labelling) and pre-existing or “old” RNA (before 4sU labelling). I was not able to perform experiments with SAF-A immunoprecipitation due to time constraints but was able to execute several experiments focusing on T-to-C conversions in amplicons of highly expressed genes. There were several key differences in my experiments compared to the original SLAM-seq protocol (Herzog *et al.*, 2017), Firstly, I used a 15 min labelling time with a high dose of 4sU, focusing on a small number of recently transcribed amplicons rather than a longer incubation period spanning several hours to a day. This was coupled with a low-dose actD treatment to inhibit nascent rRNA synthesis and thus stop 4sU incorporation into rRNA. Concentrations of 4sU above 50 μ M have been

shown to inhibit processing of rRNA and ribosome production leading to a nucleolar stress response (Burger *et al.*, 2013). Given that rRNA synthesis represents a significant portion of total RNA metabolism, accounting for up to 70% of nascent transcripts, the actD treatment was only administered for 1 hr prior to 4sU labelling (Rabani *et al.*, 2011). Therefore, RNA samples collected immediately after 15 min of 4sU labelling should not contain labelled rRNAs, however it is unknown whether samples collected 1 or 2 hr after may have 4sU-labelled rRNAs due to RNA nucleotide recycling and its potential impact on rRNA processing. Moreover, the impact of the high dose of 4sU used in these experiments on cell growth of FlpIn 293 T-REx cells was not tested. Secondly, I optimised 4sU incorporation in FlpIn 293 T-REx cells by adding a lysolecithin permeabilisation step to enhance incorporation efficiency (Figures 4.13 and 4.14D). The average increase of T-to-C conversions per amplicon in lysolecithin-treated cells was 1.5- to 2-fold more compared to untreated cells. This increase was smaller than expected compared to the previously reported 5- to 10-fold increase in radiolabelled thymidine incorporation into DNA (Miller *et al.*, 1979). This disparity could be attributed to cell type specific differences, as the increase was tested in BHK cells while I used a HEK-293 derived cell line. Ideally, I would have tested for the differences in 4sU incorporated between lysolecithin-treated and untreated cells, but my dot blotting experiments and HPLC column analysis were unsuccessful (data not shown). A third deviation from the original SLAM-seq protocol was sequencing of amplicons from highly expressed genes (EEF1A1, ACTG1, GAPDH, and LDHA) to increase sequencing depth (Figure 4.14). If this experiment were to be repeated with SAF-A immunoprecipitation to investigate nascent RNA kinetics of RNA bound to the protein, whole-genome sequencing would be used instead of specific amplicon primers. Although I prepared whole-genome samples of total RNA that were sequenced, the data has not been analysed due to time constraints, and therefore, the quality and signal-to-noise ratio of the data remain unknown.

The prepared libraries were sequenced on the NextSeq platform, generating paired-end reads that were successfully mapped to the target amplicons. The presence of 4sU and thiol alkylation resulted in an increased number of T-to-C conversions in the amplicons, confirming the successful incorporation and modification of 4sU (Figures 4.14D, 4.15, and 4.16). No A-to-C, T-to-A or T-to-G conversions were detected,

indicating that the experiments did not introduce other base changes. While only a limited number of nucleotide substitutions were shown for simplicity, it is ideal to test all possible nucleotide substitutions, as done in other 4sU metabolic labelling and thiol modification studies such as SLAM-seq, TimeLapse-seq, and TUC-seq (Herzog *et al.*, 2017; Riml *et al.*, 2017; Schofield *et al.*, 2018; Qiu *et al.*, 2020). 4sU incorporation and sequencing were tested using two different thiol modification methods – thiol alkylation with IAA with aminolysis and oxidation with TFEA/NaIO₄ (Figure 4.15). A previous study had demonstrated that IAA thiol modification led to lower library complexity compared to TFEA/NaIO₄ treatment of pooled beads corresponding to single cells (Qiu *et al.*, 2020). However, in their experiment, IAA thiol modification was performed for 1 hr at 45°C, whereas in the original SLAM-seq protocol and my experiments, a 15 min treatment at 50°C was used (Herzog *et al.*, 2017). In my study, both methods yielded a similar percentage of T-to-C conversions. Furthermore, the thiol modification methods experiment successfully introduced UMIs into the sequencing library, enabling accurate identification of the final product (Figures 4.15C and 4.15D). Analysis of UMIs in the sequencing reads did not reveal any over-amplification of specific reads (data not shown).

To investigate the kinetics of nascent RNA, I used metabolic labelling by incubating cells with 4sU and sequenced RNA at 0, 1, and 2 hr after incubation (Figure 4.16). My focus was on nuclear RNAs, so I fractionated FlpIn 293 T-REx cells to obtain cytoplasmic and nuclear fractions. It is worth noting that most other studies utilizing 4sU metabolic labelling focus on total RNA and do not separate nuclear and cytoplasmic fractions. Surprisingly, sequencing of the amplicons revealed similar percentages of T-to-C conversions between the cytoplasmic and nuclear fractions (Figures 4.16A and 4.16B). However, when the proportion of new reads and amplicon half-lives were calculated using a Poisson mixture model, differences were observed between cytoplasmic and nuclear RNAs (Figures 4.16E and 4.16F). Firstly, I observed that intron amplicons had a higher proportion of new RNA molecules compared to exon amplicons, which was expected since mRNAs contribute more to old reads due to their longer half-lives (Figure 4.16E). The proportion of new RNA molecules in exon amplicons increased over time as transcription continued, while the proportion in intron amplicons remained relatively constant in nuclear samples or showed an increase for cytoplasmic samples. During its lifetime, this nascent intronic

RNA may contribute to the RNA mesh and potentially interact with the SAF-A network, as has been reported for ncRNAs (Creamer *et al.*, 2021). Secondly, the half-lives of amplicons in the cytoplasmic and nuclear fraction were calculated and there was a statistically significant difference in intron half-lives between the two fractions (Figure 4.16F). The examined highly expressed genes are known ubiquitously expressed housekeeping genes with long mRNA half-lives, which is consistent with the observed half-lives for exon-exon amplicons in the cytoplasm (Yang *et al.*, 2003; Badreldin *et al.*, 2021). Although the half-lives for exon-exon amplicons in the nucleus were noticeably shorter, the difference was not statistically significant. The shorter half-lives are likely due to mRNA export to the cytoplasm. Considering that most genes are transcribed at an average rate of approximately 3500 bp per min, EEF1A1, ACTG1, and GAPDH would take 1 to 2 min to be transcribed into pre-mRNAs, while LDHA would take over 4 min (Fuchs *et al.*, 2014). Meanwhile, their pre-mRNAs and introns persist in the cytoplasm and nucleus for times longer than their hypothetical transcription times, suggesting that the observed half-lives are not attributed to aborted transcripts or failed mRNAs. During this period, the pre-mRNAs and introns might contribute to RNA organisation. However, it was surprising that pre-mRNAs and introns had longer half-lives in the cytoplasm than the nucleus. The nuclear intron half-lives observed for ACTG1 and GAPDH were close to the typically observed 5 min for most introns (Baboo and Cook, 2014). It is possible they had twice as long half-lives in the cytoplasm than in the nucleus due to the absence of the 5'-3' exoribonuclease XRN2 in the cytoplasm, leaving their degradation solely to the exosome machinery (Miki and Großhans, 2013). Finally, in future experiments, it would be important to consider the background T-to-C conversion rate in the absence of 4sU and thiol alkylation. In this study, the background error rate was estimated by fitting a Poisson distribution (Figure 4.16C). The PrimeScript III RTase used in the 4sU incorporation and sequencing experiments exhibited a background error rate of 2×10^{-3} errors/base for T-to-C conversions, indicating a relatively low reverse transcriptase fidelity for sequencing experiments. Alternative RTases such as MMLV or AMV RTases, which have error rates of 3.3 to 5.9×10^{-4} errors/base, would be more suitable (Yasukawa *et al.*, 2017). Therefore, it would be interesting to repeat this experiment using a reverse transcriptase with higher fidelity and employing SAF-A immunoprecipitation to examine the identity and dynamics of nascent RNAs bound to SAF-A.

Overall, in this chapter, I used super-resolution imaging to show SAF-A's nuclear distribution and the formation of connected protein clusters. These clusters were denser and more interconnected compared to HNRNPC's, another HNRNP family member, which suggests that they're possibly formed by protein oligomerisation. Inhibiting transcription altered SAF-A's nuclear distribution during interphase and reduced cluster size and connectivity which is linked to SAF-A's dependency on RNA interaction for oligomer formation. Depletion of SAF-A using RNAi resulted in reduced density and connectivity of RNA clusters which highlighted the importance of SAF-A in nascent RNA organisation. More precise perturbation of SAF-A function was achieved in Chapter 5 by using clinical mutations and their impact was quantified in Chapter 5.

5. Chapter 5: SAF-A patient mutations and their effect on its role in chromatin organisation

5.1. Introduction

A decade ago, the first *de novo* SAF-A mutations that cause the HNRNPU-related neurodevelopmental disorder (OMIM 602869) were discovered. Patients with microdeletions in the 1q44 chromosomal region causing loss of the SAF-A gene or with smaller mutations within the gene itself had a common phenotype of early-onset epileptic seizures, developmental delay, and intellectual disability (Ballif *et al.*, 2012; Nagamani *et al.*, 2012; Allen *et al.*, 2013; Gupta *et al.*, 2014; Bramswig *et al.*, 2017; Leduc *et al.*, 2017; Lim *et al.*, 2017; Yates *et al.*, 2017). Some patients also had the less common symptoms of hypotonia, heart defects and kidney issues. Since then, more mutations have been discovered but the molecular basis responsible for the disease is unknown. Most of the research has also focused on loss of function (LOF) mutants that are hypothesised to lead to haploinsufficiency as many of these mutations either truncate the protein or cause a frameshift (Table 1.1). They are though important for understanding the disorder and revealed that there is insufficient protein compensation when one copy of SAF-A is mutated which subsequently leads to transcriptional dysregulation in mice and human cortical organoids (Lai *et al.*, 2018; Dugger *et al.*, 2020; Ressler *et al.*, 2022). Whether this transcriptional dysregulation is due to SAF-A's role in chromatin organisation, RNA metabolism or RNA mesh homeostasis is unknown (Marenda *et al.*, 2022). In addition, none of the LOF variants studied matched those that patients had which might cause different symptoms than those observed in mice. Therefore, in this Chapter, I curated mutations from patients with the HNRNPU-related neurodevelopmental disorder and selected those that might affect SAF-A's role in chromatin organisation. I therefore focused on pathogenic missense variants and single codon deletions in conserved protein residues instead of protein truncations and frameshifts that might otherwise lead to loss of the protein due to instability. To do this, I followed the 1D-to-3D method (Figure 5.1) to visualise synonymous and missense variants in the general population that are available in the Genome Aggregation Database (gnomAD; Deák and Cook, 2022). I then identified SAF-A protein domains that were depleted of variants using a

ratio between the average density of variants in the whole protein (Vp) and the average density of missense variants in a protein domain (Vd) using a program calculator. I repeated these same analyses for a curated list of pathogenic mutations available in clinical literature. Coupled with conservation data, literature on SAF-A domain function, and the above analyses I selected three SAF-A mutants to investigate further with cell-based assays. These mutants are P487S/P506S, R714S/R733S, Y805*/Y824* (SAF-A isoform/SAF-A full length).

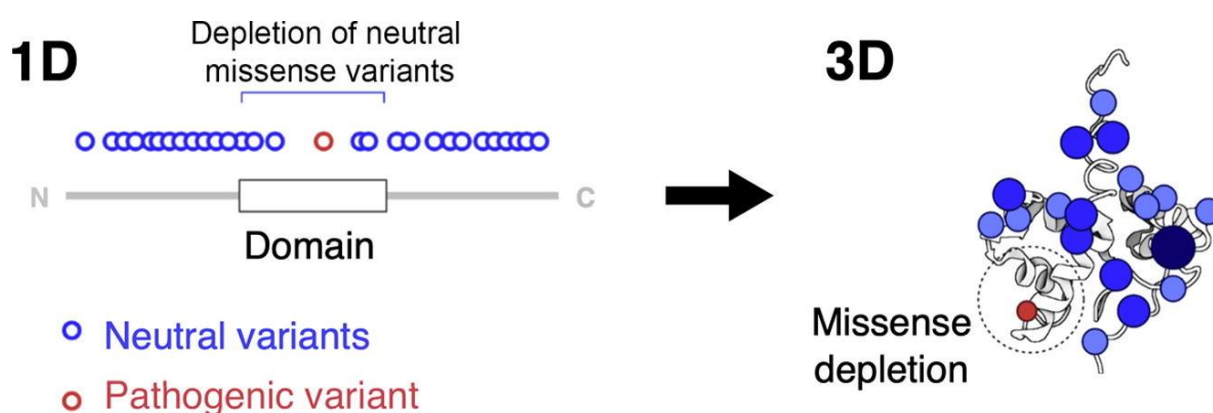


Figure 5.1. *In silico* approach. A schematic showing the 1D-to-3D approach for mapping population-wide missense variants onto a protein structure to identify depleted regions and gain insights into protein function. The figure is taken from Deák and Cook, 2022.

Expression of clinical mutations in human cells is a good alternative available to investigate the molecular effect of mutants when patient-derived material can be difficult to obtain. This is the case for patients with HNRNPU-related neurodevelopmental disorder as the nature of the disorder makes consent and ethical approval difficult to obtain. Therefore, I established FlpIn 293 T-REx cell lines expressing key patient mutations, similar to the cell lines used in Chapter 3.2 for live cell imaging of SAF-A. Plasmids expressing either the FLAG-tagged or FLAG-SNAP-tagged mutants were inserted at the FRT site under Dox control in FlpIn 293 T-REx cells. Proliferation of the cells depleted of endogenous SAF-A and expressing the mutant version of the protein was investigated using the PrestoBlue assay (see Methods 2.3.7.). Whilst the effect of the mutants on chromatin organisation was investigated using FISH probes for the gene-rich region 11p15.5 (see Methods 2.9.) and by conducting an oligomer formation assay (see Methods 2.6.).

5.2. Selection of patient mutations

5.2.1. Protein conservation

A multiple sequence (MSA) alignment between full-length SAF-A and its orthologs was generated using the Oma group 1163611 of metazoan species for the protein (Altenhoff *et al.*, 2020). The MSA was used in tandem with the AlphaFold model for SAF-A (AF-Q00839-F1) as input to the ConSurf server to produce a surface conservation PyMOL script that was then mapped onto the SAF-A protein structure (Yariv *et al.*, 2023; Figure 5.2 and 5.4). Highlighted SAF-A domains (Figure 5.2; SAP - residues 8 to 42; SPRY – residues 267 to 464; AAA+ - residues 488 to 672) and regions of interest have the highest conservation scores which suggests they are the most functionally conserved. Meanwhile, most of the intrinsically disordered regions (IDR 1 - residues 43 to 266; IDR 2 - residues 673 to 825) which have no secondary protein structure are the least conserved except for the RGG motif (residues 714 to 739) that is important for RNA binding. Surprisingly, the C-terminal tail tip (residues 823 to 825) of SAF-A is shown to be nested between the SPRY (close to residues 278 to 281 and 298 to 304), and AAA+ (close to residues 571 to 573) domains in the AlphaFold model, and may have a functional role. The mentioned C-terminal tail tip, SPRY and AAA+ residues appear to be very well conserved in SAF-A orthologs.

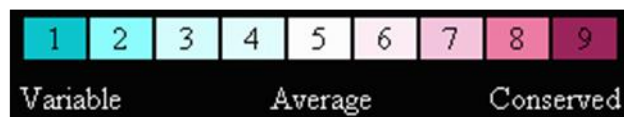
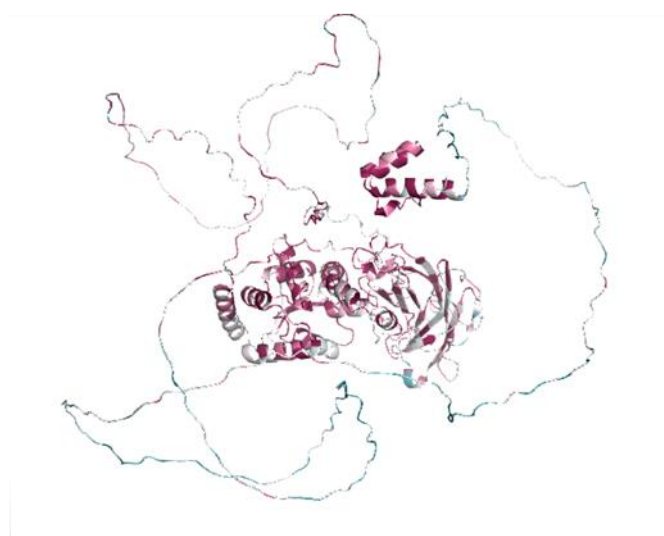
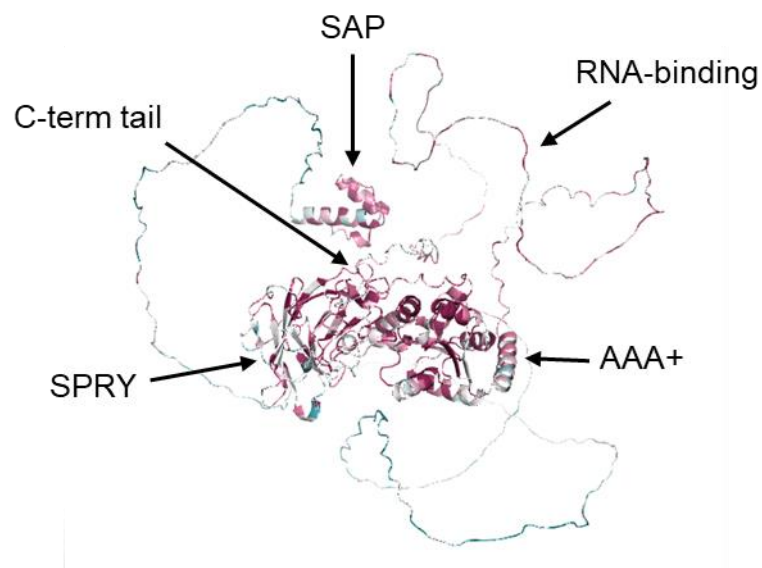


Figure 5.2. Protein sequence conservation of full-length SAF-A. Predicted structure of SAF-A (AlphaFold, AF-Q00839-F1) annotated with sequence conservation in metazoa. The key for conservation scores is at the bottom of the figure. Domains and regions are shown with arrows.

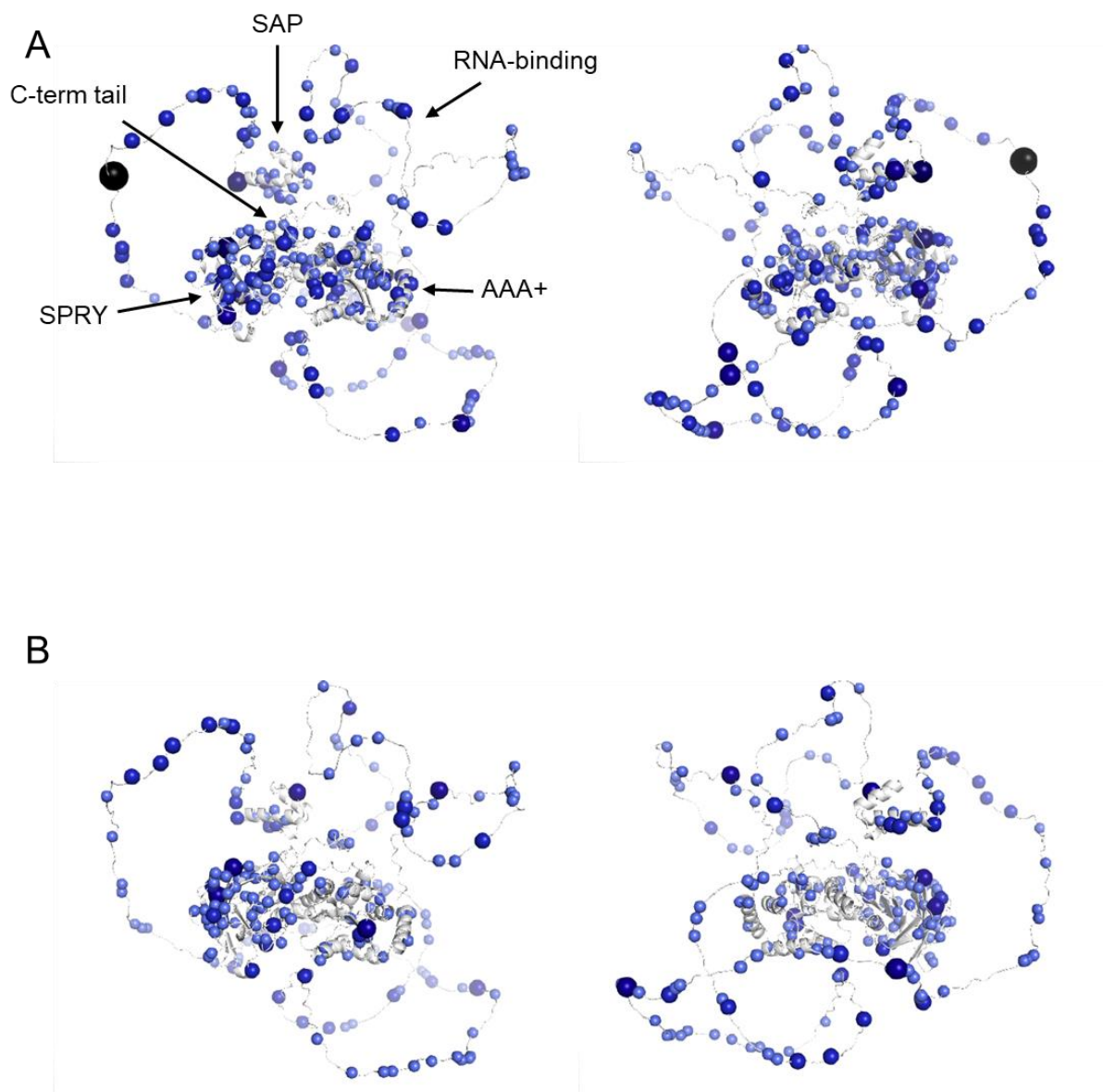


Figure 5.3. Population-wide mutations in SAF-A. Predicted structure of SAF-A (AlphaFold, AF-Q00839-F1) annotated with synonymous (A) and missense (B) variants. The size and darker colour of the annotations represent the frequency of the variants.

5.2.2. Mutations in the general population

Next, I visualised the SAF-A population-wide synonymous (Figure 5.3A) and missense (Figure 5.3B) variants using data from gnomAD overlaid on the AlphaFold model using the 1D-to-3D method (Deák and Cook, 2022). By eye, the AAA+ domain appears to be slightly depleted of synonymous variants and more depleted of missense variants. I therefore calculated the Vd/Vp ratio (average density

of variants in a protein domain/average density of variants in the whole protein) for domains and regions of interest in full-length SAF-A to reveal differences in variant depletion (Table 5.1 and 5.2). The ratios confirmed the previous observation that the AAA+ domain is partly depleted of synonymous variants (V_d/V_p ratio = 0.79) and much more depleted of missense variants (V_d/V_p ratio = 0.42). In addition, the IDR 2 region was partly depleted of synonymous variants (0.82) but not depleted of missense variants (0.96), while the SAP domain and actin-binding region were missense depleted (0.44 and 0.48 respectively). The regions with lower V_d/V_p ratios could be mutation-intolerant and mutations in them are more likely to lead to pathogenesis.

Fragment	Residue number	Vd/Vp ratio
SAF-A	1 to 825	$V_p = 0.24$
SAP domain	8 to 42	1.2
IDR 1	43 to 266	1.16
SPRY domain	267 to 464	1.02
Connection	465 to 487	1.46
AAA+ domain	488 to 672	0.79
Actin-binding	611 to 626	1.05
IDR 2	673 to 825	0.82
RGG box	714 to 739	1.13

Table 5.1. Population-wide synonymous mutation ratios per domain.

Fragment	Residue number	Vd/Vp ratio
SAF-A	1 to 825	$V_p = 0.26$
SAP domain	8 to 42	0.44
IDR 1	43 to 266	1.33
SPRY domain	267 to 464	1.15
Connection	465 to 487	1.68
AAA+ domain	488 to 672	0.42
Actin-binding	611 to 626	0.48
IDR 2	673 to 825	0.96
RGG box	714 to 739	1.49

Table 5.2. Population-wide missense mutation ratios per domain.

5.2.3. Analysis of clinical variants in the literature

To investigate pathogenic mutations in missense-depleted regions of the protein, I curated the available clinical data on SAF-A mutations in patients (Table 5.3). From Table 1.1, the majority (>80%) of mutations (frameshifts, truncations, splicing variants) are hypothesised to lead to a truncated protein and haploinsufficiency. The remaining mutations (missense substitutions and single codon deletions) might lead either to haploinsufficiency or dominant negative effects due to reduced protein function. Therefore, I compared the severity of phenotypes between the two groups of mutations. In terms of phenotype, most clinical data is linked to cases of autism or intellectual disability and not all patients have phenotypical information beyond that. Due to the phenotypes being incomplete and different clinical classifications on the severity of the phenotype, there was no obvious difference in phenotypes between the two groups of mutations.

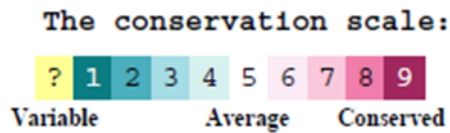
DNA mutation	Amino acid change	Protein region	Reference
Chr1:245,025,709-245,133,797dup	chr 1 duplication including first 3 exons of HNRNPU	SAP domain + Acidic/disordered + SPRY domain	Bramswig <i>et al.</i> , 2017
c.16delinsATT	p.Val6Ilefs*4 + inherited (x2)	SAP domain	Depienne <i>et al.</i> , 2017; Chérot <i>et al.</i> , 2018
c.23del1	p.Val8Glufs*4	SAP domain	Yates <i>et al.</i> , 2017
c.29_30del	p.Lys10del	SAP domain	Gillentine <i>et al.</i> , 2021
c.67C>T	p.Arg23*	SAP domain	Durkin <i>et al.</i> , 2020
c.76del	p.Ser26LeufsTer35	SAP domain	Shimada <i>et al.</i> , 2018
c.143_149del7	p.Gly48Afs*11	IDR 1	Song <i>et al.</i> , 2021
c.149_156del	p.Arg50fs*	IDR 1	Gillentine <i>et al.</i> , 2021
c.198del	p.Ala67fs*	IDR 1	Gillentine <i>et al.</i> , 2021
c.253dupG	p.Glu85Glyfs*	IDR 1	Epilepsy Genetics Initiative, 2019
c.324_328del	Glu109Argfs*34	IDR 1	Gillentine <i>et al.</i> , 2021
c.325G>C	Glu109Gln	IDR 1	Gillentine <i>et al.</i> , 2021
c.334dup	p.Ala112Glyfs*33	IDR 1	Wang <i>et al.</i> , 2020
c.395_401del7	p.Asn132Thrfs*63	IDR 1	Durkin <i>et al.</i> , 2020
c.401_402delAC	Asp134Glyfs*10	IDR 1	Fernández-Marmiesse <i>et al.</i> , 2019
c.418G>A	p.Glu140Lys	IDR 1	King <i>et al.</i> , 2014/Yates <i>et al.</i> , 2017
c.454_466del	p.Ala152Thrfs*41	IDR 1	Durkin <i>et al.</i> , 2020
c.469G>C	p.Gly157Arg	IDR 1	Krumm <i>et al.</i> , 2015
c.481C>T	p.Gln161*	IDR 1	Gillentine <i>et al.</i> , 2021
c.508C>T	p.Gln170* x2	IDR 1	Feliciano <i>et al.</i> , 2019; Taylor <i>et al.</i> , 2022
c.511C>T	p.Gln171*	IDR 1	Hamdan <i>et al.</i> , 2014/ Bramswig <i>et al.</i> , 2017
c.523C>T	p.Gln175*	IDR 1	Bramswig <i>et al.</i> , 2017
c.547_560del	p.Ala183fs*	IDR 1	Gillentine <i>et al.</i> , 2021

c.596dup	p.Pro200fs*	IDR 1	Gillentine <i>et al.</i> , 2021
c.619C>T	p.Gln207*	IDR 1	Durkin <i>et al.</i> , 2020
c.622C>T	p.Gln208T	IDR 1	Salinas <i>et al.</i> , 2021
c.651_660del	p.Gly218Alafs*118	IDR 1	Leduc <i>et al.</i> , 2017
c.692-1G>A	p.? splice site variant	IDR 1	Depienne <i>et al.</i> , 2017; Durkin <i>et al.</i> , 2020
c.706_707del	p.Glu236Trpfs*6 x3	IDR 1	Farwell <i>et al.</i> , 2015; Durkin <i>et al.</i> , 2020 x2
c.712_715del	p.Lys238Alafs*100	IDR 1	Durkin <i>et al.</i> , 2020
c.730_731delAG	p.Arg244Glyfs*3	IDR 1	Taylor <i>et al.</i> , 2022
c.742dup	p.Arg248Lysfs*12	IDR 1	Taylor <i>et al.</i> , 2022
c.804-9_804-6del	p.?	IDR 1	Taylor <i>et al.</i> , 2022
c.817C>T	p.Gln273* x2	SPRY domain	Bramswig <i>et al.</i> , 2017; Demos <i>et al.</i> , 2019
c.837_839delAGA	p.Glu279del	SPRY domain	Durkin <i>et al.</i> , 2020
c.847_857del	p.Phe283Serfs*5	SPRY domain	Durkin <i>et al.</i> , 2020
c.872_877delIATACTT	p.Thr292_Y293del	SPRY domain	Gillentine <i>et al.</i> , 2021
c.878A>G	p.Tyr293Cys	SPRY domain	Hinokuma <i>et al.</i> , 2020
c.878-9T>G	p.?	SPRY domain	Taylor <i>et al.</i> , 2022
c.893A>G	p.His298Arg	SPRY domain	Taylor <i>et al.</i> , 2022
c.960G>A	p.Trp320*	SPRY domain	Yates <i>et al.</i> , 2017
c.970A>G	p.Arg324Gly	SPRY domain	Bramswig <i>et al.</i> , 2017
c.971_973delGAG	p.Arg324_Ala325delinsThr	SPRY domain	Johannesen <i>et al.</i> , 2020
c.1050_1051del	p.Thr351Lysfs*4	SPRY domain	Wang <i>et al.</i> , 2020
c.1088G>A	p.Trp363*	SPRY domain	Durkin <i>et al.</i> , 2020
c.1089G>A	p.Trp363* x2	SPRY domain	Leduc <i>et al.</i> , 2017; Durkin <i>et al.</i> , 2020
c.1117+1G>A	Splice donor variant	SPRY domain	Yates <i>et al.</i> , 2017
c.1132T>C	p.Ser378Pro	SPRY domain	Bramswig <i>et al.</i> , 2017
c.1142A>G	p.Tyr381Cys	SPRY domain	Satterstrom <i>et al.</i> , 2020
c.1171_1172TG[1]	p.Cys391_Gln392delins*	SPRY domain	Gillentine <i>et al.</i> , 2021
c.1243del	p.Asp415Metfs*3	SPRY domain	Taylor <i>et al.</i> , 2022
c.1282delC	p.Gly429Alafs*53	SPRY domain	Song <i>et al.</i> , 2021
c.1368A>C	p.Gln456Asn	SPRY domain	Gillentine <i>et al.</i> , 2021
c.1424_1425insTC	p.Ile476Profs*7	Connection	Yates <i>et al.</i> , 2017
c.1435G>A	p.Val479Ile	Connection	Gillentine <i>et al.</i> , 2021
c.1450C>T	p.Arg484*	Connection	Durkin <i>et al.</i> , 2020
c.1516C>T	p.Pro506Ser	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.1615-1G>A	Canonical splice site	AAA+ domain	Need <i>et al.</i> , 2012/Zhu <i>et al.</i> , 2015
c.1617dup	p.Ala540Serfs*4	AAA+ domain	Taylor <i>et al.</i> , 2022
c.1626_1627insA	p.Lys543Leufs*11	AAA+ domain	Yates <i>et al.</i> , 2017
c.1626_1627insTT	p.Lys543Leufs*11	AAA+ domain	DDD Study <i>et al.</i> , 2017
c.1641delA	p.Asp548Ilefs*5	AAA+ domain	Durkin <i>et al.</i> , 2020
c.1664del	p.Leu555Argfs*51	AAA+ domain	Yates <i>et al.</i> , 2017
c.1665_1666delGT	p.Leu556Alafs*12	AAA+ domain	Taylor <i>et al.</i> , 2022
c.1681delC	p.Gln561Serfs*44	AAA+ domain	DDD Study <i>et al.</i> , 2017
c.1681del1	p.Gln561Serfs*45 x2	AAA+ domain	Depienne <i>et al.</i> , 2017; Durkin <i>et al.</i> , 2020
c.1681C>T	p.Gln561* x2	AAA+ domain	Depienne <i>et al.</i> , 2017; Wang <i>et al.</i> , 2020
c.1714C>T	p.Arg572* x2	AAA+ domain	Leduc <i>et al.</i> , 2017; Wang <i>et al.</i> , 2020
c.1743+1G>C	p.? splice site variant	AAA+ domain	Durkin <i>et al.</i> , 2020
c.1744_1767del	p.Thr582_Gln589del	AAA+ domain	Allen <i>et al.</i> , 2013; Bramswig <i>et al.</i> , 2017

c.1755_1756insCCTCT	p.Ala586Profs*22	AAA+ domain	DDD Study <i>et al.</i> , 2017
c.1801C>T	p.Arg601*	AAA+ domain	Durkin <i>et al.</i> , 2020
c.1812dup	p.Arg605*	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.1813_1814dup	p.Val606*	AAA+ domain	Wang <i>et al.</i> , 2020
c.1834G>A	p.Asp612Asn	AAA+ domain	Taylor <i>et al.</i> , 2022
c.1836del	p.Tyr613Ilefs*11	AAA+ domain	Durkin <i>et al.</i> , 2020
c.1846_1847insCA	p.Lys616Thrfs*3	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.1852C>T	p.Gln618*	AAA+ domain	Wang <i>et al.</i> , 2020
c.1865_1867del	p.Glu622del	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.1868dup	p.Glu624Argfs*24	AAA+ domain	de Kovel <i>et al.</i> , 2016, Depienne <i>et al.</i> , 2017
c.1920delT	p.Phe640fs*	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.1925_1926del	p.Leu642Profs*5	AAA+ domain	Durkin <i>et al.</i> , 2020
c.2014delG	p.Glu672Lysfs*	AAA+ domain	Gillentine <i>et al.</i> , 2021
c.2083_2084del	p.Ser695Trpfs*6	IDR 2	Durkin <i>et al.</i> , 2020
c.2085dup	p.Gly696fs*	IDR 2	Gillentine <i>et al.</i> , 2021
c.2167+35_*4156del	p.? deletion of last three exons	IDR 2 + RGG box	Durkin <i>et al.</i> , 2020
c.2199G>T	p.Arg733S	RGG box	Gillentine <i>et al.</i> , 2021
c.2197_2199delAGG	p.Arg733del	RGG box	Oates <i>et al.</i> , 2018
c.2270_2271delCT	p.Pro757Argfs*7 x2	IDR 2	Leduc <i>et al.</i> , 2017; Taylor <i>et al.</i> , 2022
c.2277dup	p.Pro760Serfs*5	IDR 2	Spagnoli <i>et al.</i> , 2021
2299_3202AACA	p.Asn767fs* x2	IDR 2	Gillentine <i>et al.</i> , 2021
c.2299_2302del	p.Asn767Glufs*66	IDR 2	Monroe <i>et al.</i> , 2016/Depienne <i>et al.</i> , 2017
c.2302_2303delAG	p.Gly769fs*	IDR 2	Gillentine <i>et al.</i> , 2021
c.2304_2305del	p.Gly769fs*, p.Gly769Glufs*83	IDR 2	Gillentine <i>et al.</i> , 2021; Taylor <i>et al.</i> , 2022
c.2319_2320del	p.Glu774fs*	IDR 2	Gillentine <i>et al.</i> , 2021
c.2365C>T	p.Arg789* x2	IDR 2	Veith <i>et al.</i> , 2014; Wang <i>et al.</i> , 2020
c.2408A>G	p.Asn803Ser	IDR 2	Gillentine <i>et al.</i> , 2021
c.2425-2A>G	n/a x2	IDR 2	Gillentine <i>et al.</i> , 2021; Taylor <i>et al.</i> , 2022
c.2425-3C>A	n/a	IDR 2	Depienne <i>et al.</i> , 2017
c.2471T>C, c.2472T>A	p.Tyr824*	IDR 2	Carvill <i>et al.</i> , 2013

Table 5.3. SAF-A pathogenic mutations in clinical patients. Full table with phenotypes is available in Appendix D.

224



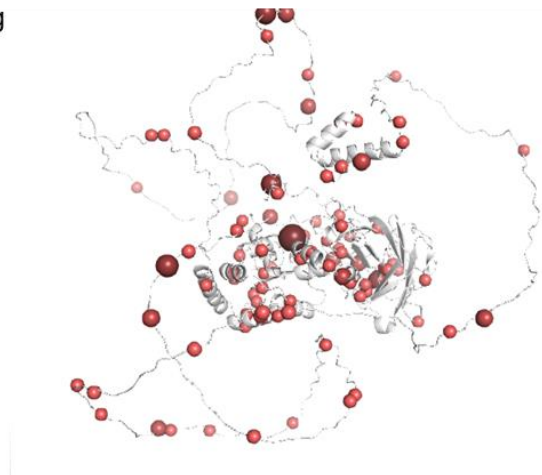
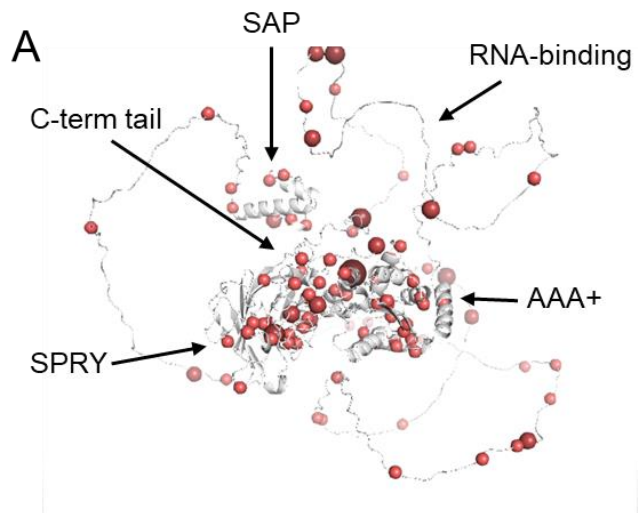
- e** - An exposed residue according to the NACSES algorithm.
- b** - A buried residue according to the NACSES algorithm.
- f** - A predicted functional residue (highly conserved and exposed).
- s** - A predicted structural residue (highly conserved and buried).
- X** - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.
- M** - substitution mutation
- /** - deletion mutation
- *** - truncation mutation
- ~** - splicing mutation
- M*** - frameshift mutation
- (xN)** - number of mutations

Figure 5.4. Pathogenic mutations in SAF-A overlayed on protein sequence conservation. Multiple sequence alignment (MSA) used for CONSURF figure generation (Figure 5.2) overlayed with clinical patient pathogenic mutations gathered from the literature. The key for conservation scores is in the legend together with the MSA annotation key.

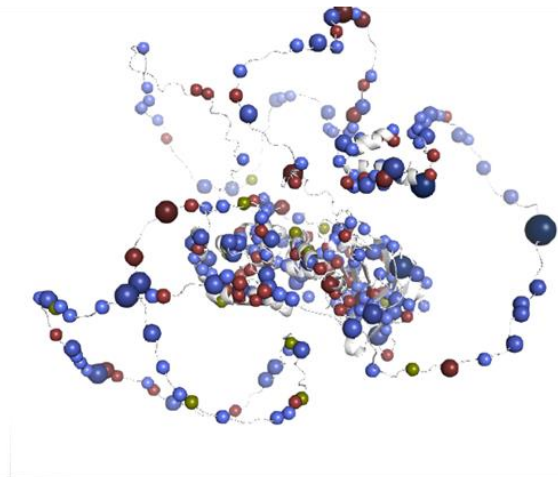
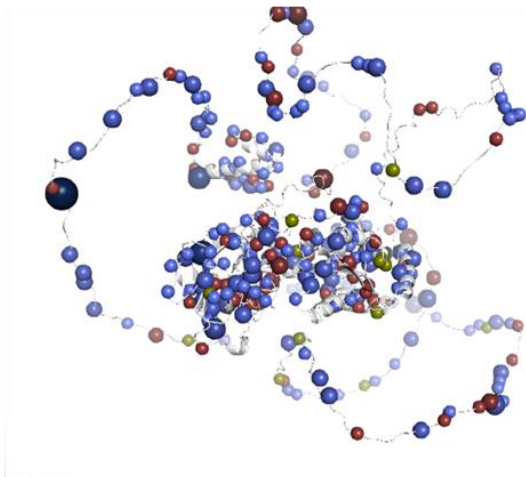
Next, I highlighted the MSA of SAF-A orthologs (Figure 5.4) with the curated pathogenic mutations to check whether they overlap with conserved residues. Generally, there were more mutations in the well-conserved regions of the protein with missense and deletion variants being found almost exclusively in those regions. Subsequently I visualised the pathogenic mutation on the full-length SAF-A 3D protein structure (Figure 5.5A) using the variant locations from Table 5.3. By eye, it is difficult to see whether there is an enrichment for pathogenic mutations in any domain but the RGG motif region that is responsible for RNA binding appears to be depleted of mutations. When comparing pathogenic mutations with synonymous

variants (Figure 5.5B) or missense variants (Figure 5.5C) there is a minimal degree of overlap observed. Specifically, only 15 pathogenic mutations were observed on the same amino acid as synonymous variants and 16 pathogenic mutations similarly overlapped with missense variants. Only 2 mutations were at the same location in the protein among all three variant types.

Calculating the Vd/Vp ratio for pathogenic mutations (Table 5.4) confirms the previous observation that the RGG motif is depleted of mutations ($Vd/Vp = 0.36$). The SPRY domain and IDR 2 region are also slightly depleted of pathogenic mutations (0.86 and 0.74). The connection region (residues 465 to 487) between the SPRY and AAA+ domains was partly enriched in mutations (1.24) together with the AAA+ domain (1.23). Specifically, the actin-binding region of SAF-A within the AAA+ domain was highly enriched in mutations (3.56). Comparing the Vd/Vp ratios of missense variants (Table 5.2) to those of pathogenic mutations highlighted that the missense-depleted regions of the AAA+ domain and actin-binding region were more enriched in pathogenic mutations in general. When calculating the Vd/Vp ratio for only missense and single codon deletion pathogenic mutations (Table 5.5) there are some key differences. The actin-binding region is still highly enriched for mutations (5.16) but the AAA+ domain in general is slightly depleted (0.89). Both IDRs are depleted of mutations (IDR 1 - residues 43 to 266, $Vd/Vp = 0.74$; IDR 2 $Vd/Vp = 0.54$) while the rest of the domains and connecting regions were more enriched in mutations (SAP domain $Vd/Vp = 1.18$; SPRY domain 1.67; connection 1.79).



B



C

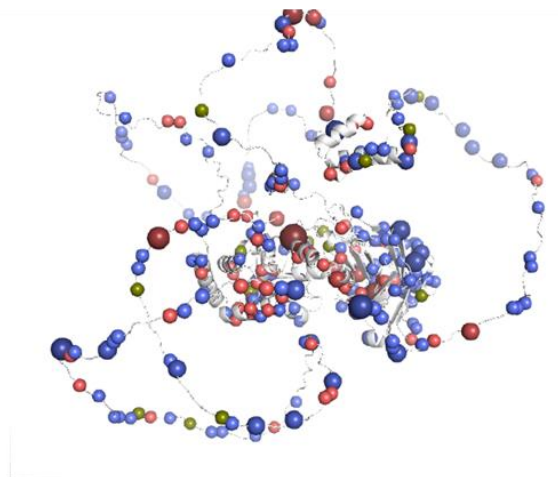
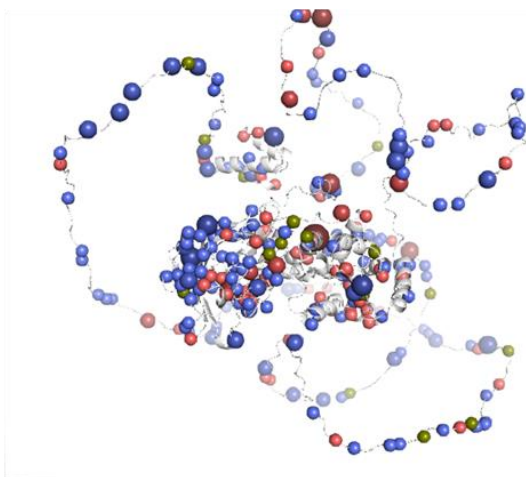


Figure 5.5. Pathogenic mutations in SAF-A. Predicted structure of SAF-A (AlphaFold, AF-Q00839-F1) annotated with pathogenic (A) variants. Overlap between pathogenic (in red) and synonymous (in blue, B) or missense (in blue, C). Where both types of variants occur, they are coloured green. The size and darker colour of the annotations represent the frequency of the variants.

Fragment	Residue number	Vd/Vp ratio
SAF-A	1 to 825	Vp = 0.11
SAP domain	8 to 42	1.08
IDR 1	43 to 266	1.06
SPRY domain	267 to 464	0.86
Connection	465 to 487	1.24
AAA+ domain	488 to 672	1.23
Actin-binding	611 to 626	3.56
IDR 2	673 to 825	0.74
RGG box	714 to 739	0.36

Table 5.4. All pathogenic mutation ratios per domain.

Fragment	Residue number	Vd/Vp ratio
SAF-A	1 to 825	Vp = 0.02
SAP domain	8 to 42	1.18
IDR 1	43 to 266	0.74
SPRY domain	267 to 464	1.67
Connection	465 to 487	1.79
AAA+ domain	488 to 672	0.89
Actin-binding	611 to 626	5.16
IDR 2	673 to 825	0.54
RGG box	714 to 739	1.59

Table 5.5. Only substitution and deletion pathogenic mutation ratios per domain.

5.2.4. Criteria for patient mutation selection for further experiments

To select which mutations would be investigated further, I formulated criteria. The mutations needed to be on conserved residues, ideally, a substitution or deletion variant, not available in the general population, and have patient phenotype data available. Firstly, I disregarded the sites of pathogenic mutations that overlapped with the missense data from gnomAD – V6Ifs*, R50Hfs*, A67fs*, G157R, E170*, E175*, G218Afs*, splice site variant at c.692, E279del, T292_Y293del, H298R, G429Afs*, R572*, A586Pfs*, R605*, and R789*. Then I used the previously generated Vd/Vp ratios (Tables 5.1, 5.2, 5.4, 5.5) to determine missense-depleted regions which were enriched in pathogenic mutations. Out of them, the actin-binding region was the most enriched with a substitution (D612N) and deletion mutation (E622del). The patient with the D612N mutation had an additional genetic defect – loss of the clone RP11-38023 at Xp11.23 therefore the mutation was disregarded as a candidate (Table 5.3; Taylor *et al.*, 2022) and the E622del mutation had no information about the patient phenotypes so was not prioritised as a candidate (Table 5.3; Gillentine *et al.*, 2021). A substitution mutation in the AAA+ domain (P506S) within the Walker A motif was highlighted as a candidate due to its potential importance for ATP hydrolysis. It was also highly conserved and there was patient phenotype data (Figure 5.6, Table 5.3; Gillentine *et al.*, 2021). The SAP domain was slightly enriched for mutations and had a single codon deletion, K10del, with available patient phenotype data which was a potential candidate (Table 5.3; Gillentine *et al.*, 2021). The SPRY domain had five missense (Y293C, H298R, R324G, S378P, and Y381C) mutations and three deletions (E279del, T292_Y293del, R324_A325delinsT) mutations which made the most enriched domain in them. Out of them, three sites were available in the general population as missense variants so they were removed as candidates (E278del, T292_Y293del, H298R). The patient with the R324G mutations had an additional genetic defect at KDM6B therefore was not considered further (Bramswig *et al.*, 2017). Similarly, the patient with the S378P mutation had a maternal duplication at 8p23.1 and a maternal deletion at 13q31.3 so was not considered further (Bramswig *et al.*, 2017). The connection region between the SPRY and AAA+ domains was enriched in pathogenic variants and had a single substitution, V479I. The RGG motif was also enriched in substitution and deletion mutants. It had one substitution,

R733S, and two deletions, c.2167+35_*4156del and R733del. The former deletion resulted in the deletion of the last three exons so it might lead to haploinsufficiency, therefore it was not used for further analysis. The latter deletion and substitution were at the same site, part of an RGG repeat in that region, therefore it was prioritised for further analysis. Finally, a truncation mutation, Y824*, was considered as a candidate because it would result in deleting the final two amino acids of SAF-A, two tyrosines, which were highly conserved and nested close to the SPRY and AAA+ domains in the AlphaFold model (Figure 5.2 and 5.4). In summary, after following my criteria, I had nine candidates to choose from. I decided to stick to substitution mutations in general but also continue with the Y824* truncation mutant as it might be important for protein function and/or stability. Out of the five remaining missense mutations I picked R733S as it was disrupting an RGG repeat and might affect the RNA binding capabilities of SAF-A, and P506S as it was in the Walker A motif and might affect ATP hydrolysis like the WA mutation (Nozawa *et al.*, 2017). In the experiments described in this chapter, the mutations will be referred to as P487S, R714S, and Y805* because the constructs used derived from the 806 amino acid isoform of SAF-A which has a deletion of amino acids 212 to 230 compared to another protein isoform that is 825 amino acids in length, and which is commonly used for numbering of mutations. Previous work in the Gilbert lab has indicated that the 806 aa isoform of SAF-A is likely the most abundant form in cells, and by Western blot only one isoform of the protein is apparent (Nozawa *et al.*, 2017).

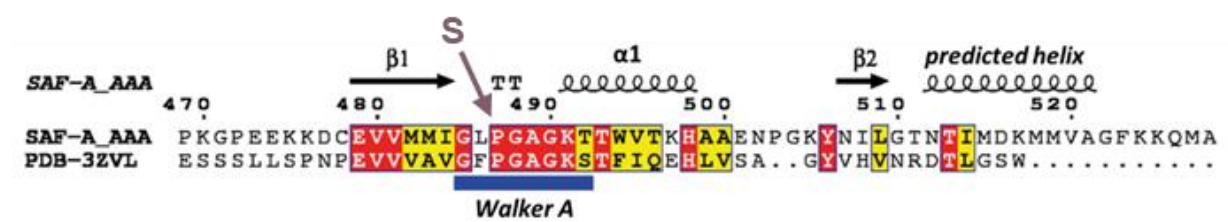


Figure 5.6. Pathogenic mutation in Walker A motif of SAF-A. Conserved amino acids (red) and similar amino acids (yellow) between SAF-A AAA+ domain and mammalian PNK (PDB-3ZVL). Walker A motif substitution is highlighted in purple. Figure from Nozawa *et al.*, 2017.

5.3. Effect of SAF-A mutations in human cell lines

5.3.1. Making cell lines expressing SAF-A mutants

The three mutations for further analysis were introduced into FlpIn transformation plasmids containing FLAG- or FLAG-SNAP-SAF-A by using restriction digests and Gibson Assembly. Their presence was confirmed with sequencing and immunofluorescence of transiently transfected HCT 116 cells (Figure 5.7). The FLAG-tagged mutant plasmids were compared to transiently expressed FLAG-SAF-A after overnight incubation, fixation, permeabilization, block, and incubation with an anti-FLAG primary antibody and an anti-mouse Alexa Fluor 647 far-red secondary antibody (Figure 5.7A). The wild-type SAF-A and mutant variants' localisation in the nucleus was highly similar with different fluorescence intensities due to the nature of transient transfection. The FLAG-SNAP-tagged mutant plasmids were compared to transiently expressed FLAG-SNAP-SAF-A treated similarly as above except that the cells were incubated with the red SNAP probe, TMR-STAR (Figure 5.7B). Their localisation was comparable to wild-type SAF-A and again, not all cells expressed the same amount of transfected protein.

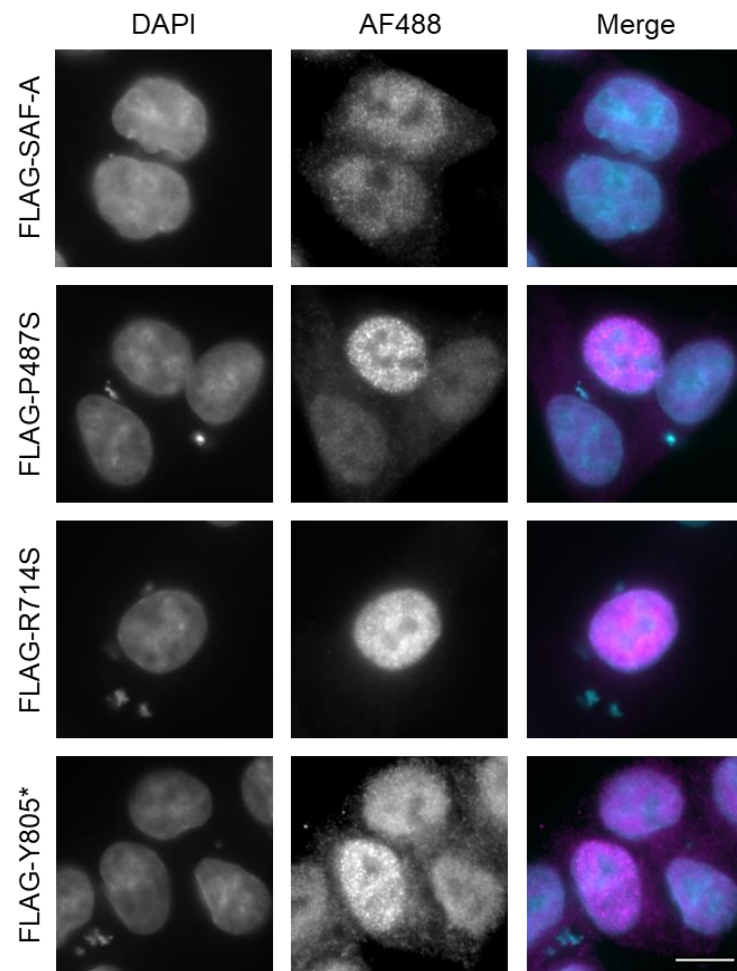
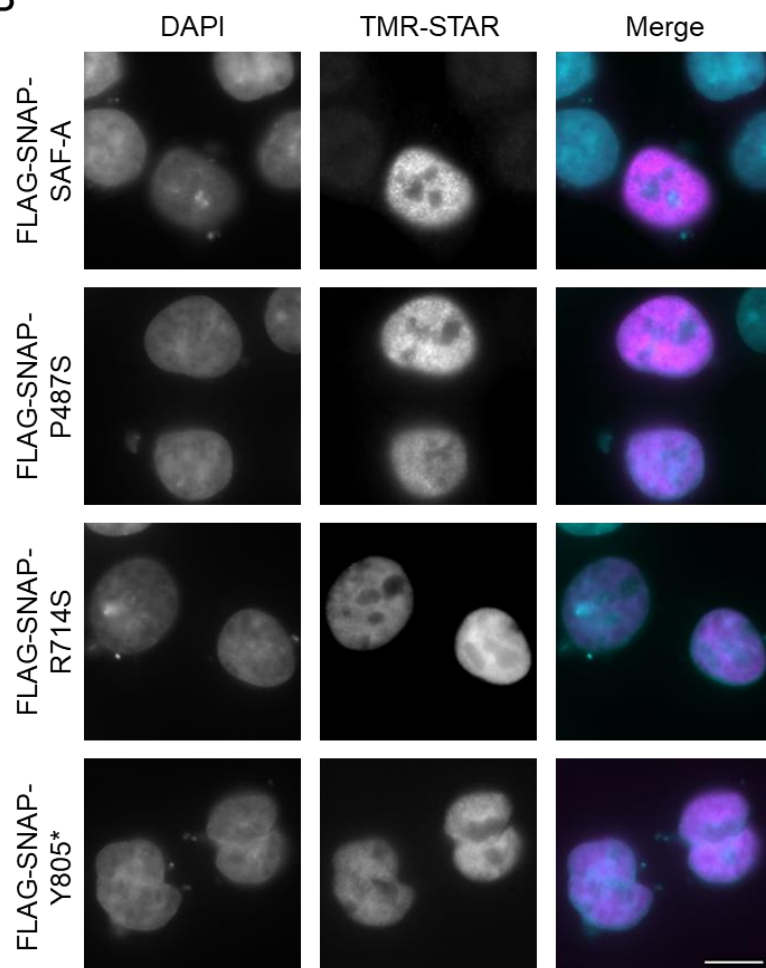
A**B**

Figure 5.7. Expression of SAF-A patient mutants in HCT 116 cells. DNA was visualised by DAPI staining. A) FLAG-tagged SAF-A and patient mutation variants were visualised using a mouse anti-FLAG antibody and a rabbit anti-mouse Alexa Fluor 488 (AF488) secondary antibody. B) FLAG-SNAP-tagged SAF-A and patient mutation variants were visualised using the red SNAP ligand probe, TMR-STAR. The white scale bars represent 10 μ m.

After confirming the plasmids' sequence and expression, they were integrated into FlpIn 293 T-REx cell lines at the FRT site using the Flp recombinase expression plasmid, pOG44. After treatment with hygromycin and blasticidin antibiotics, cells grown in the presence or absence of 1 μ g/ml Dox were lysed and probed with an anti-SAF-A antibody and anti-GAPDH antibodies (Figure 5.8). In the absence of Dox, there was no higher band visible than wild-type SAF-A (estimated at 90 kDa) for cells transformed with patient mutants, similar to the parental cell line (Figure 5.8A). In the presence of Dox, there was a slightly higher band visible for FLAG-tagged mutants (estimated at 93 kDa; Figure 5.8B) and a faint higher band for FLAG-SNAP-tagged mutants (estimated at 112 kDa; Figure 5.8C). This confirmed the successful integration of FlpIn plasmids at the FRT site and their expression in the presence of Dox.

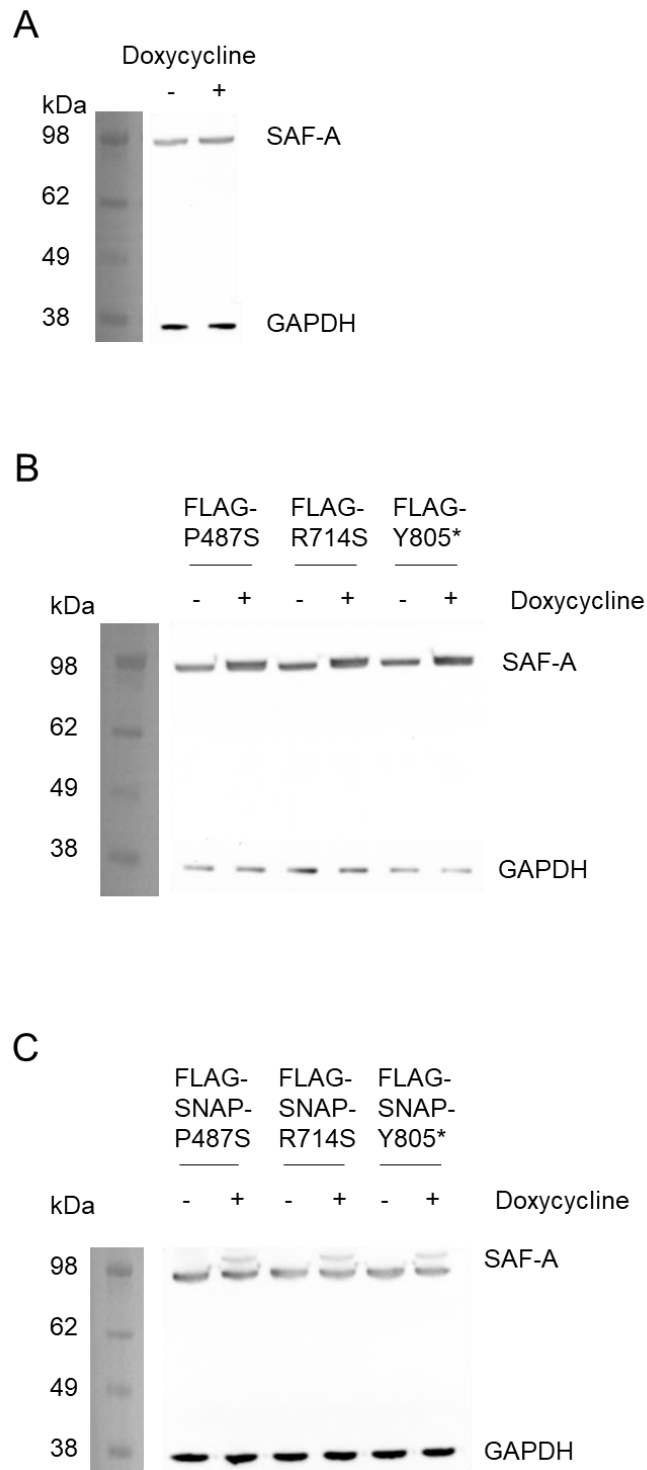


Figure 5.8. FlpIn 293 T-REx cell line generation with selected SAF-A patient mutations. Western blots were probed with anti-SAF-A antibody and GAPDH as a loading control. Cells were cultured in the absence or presence of doxycycline for 24 hours. A) Parental FlpIn 293 T-REx cells. B) FlpIn 293 T-REx cells with inserted FLAG-tag SAF-A mutants at the FRT site. C) FlpIn 293 T-REx cells with inserted FLAG-tag SNAP-SAF-A mutants at the FRT site.

To examine the effect of the SAF-A mutants in a homozygous state, endogenous SAF-A was knocked down using RNAi. Parental FlpIn 293 T-REx cells were incubated with siControl as a positive control and two siRNAs targeting SAF-A (siSAF-A and siSAF-A 2) over 1 to 3 days to determine the timeline of depletion (Figure 5.9). Cells treated with different siRNAs and incubated for different time points were lysed and probed for SAF-A and GAPDH. The 72-hr treatment with all 3 siRNAs but mostly with siControl left a slightly lower than expected band for SAF-A due to unknown reasons. When the band intensity was quantified and normalised to the GAPDH signal as a loading control, the 72 hr siControl band had a higher intensity than the two other time points (1.34 compared to 1 for 24 hr and 0.88 for 48 hr). For the 72 hr siSAF-A and siSAF-A 2, they also had the highest intensity compared to the other two time points (1.02 for 72 hr siSAF-A, compared to 0.6 and 0.57 for 24 and 48 hr; 0.76 for 72 hr siSAF-A 2, compared to 0.55 and 0.41 for 24 and 48 hr). The 24 and 48 hr treatments resulted in a 40% to 60% reduction in protein, which was less than expected compared to other cell lines and did not achieve a full KD of SAF-A (HCT 116 and RPE-1 in Chapter 3.4.1). The siSAF-A 2 siRNA achieved a slightly better KD but it cannot be used further as the SAF-A mutants are also sensitive to this siRNA so would also be knocked down, unlike with siSAF-A which they have been mutated to be resistant to.

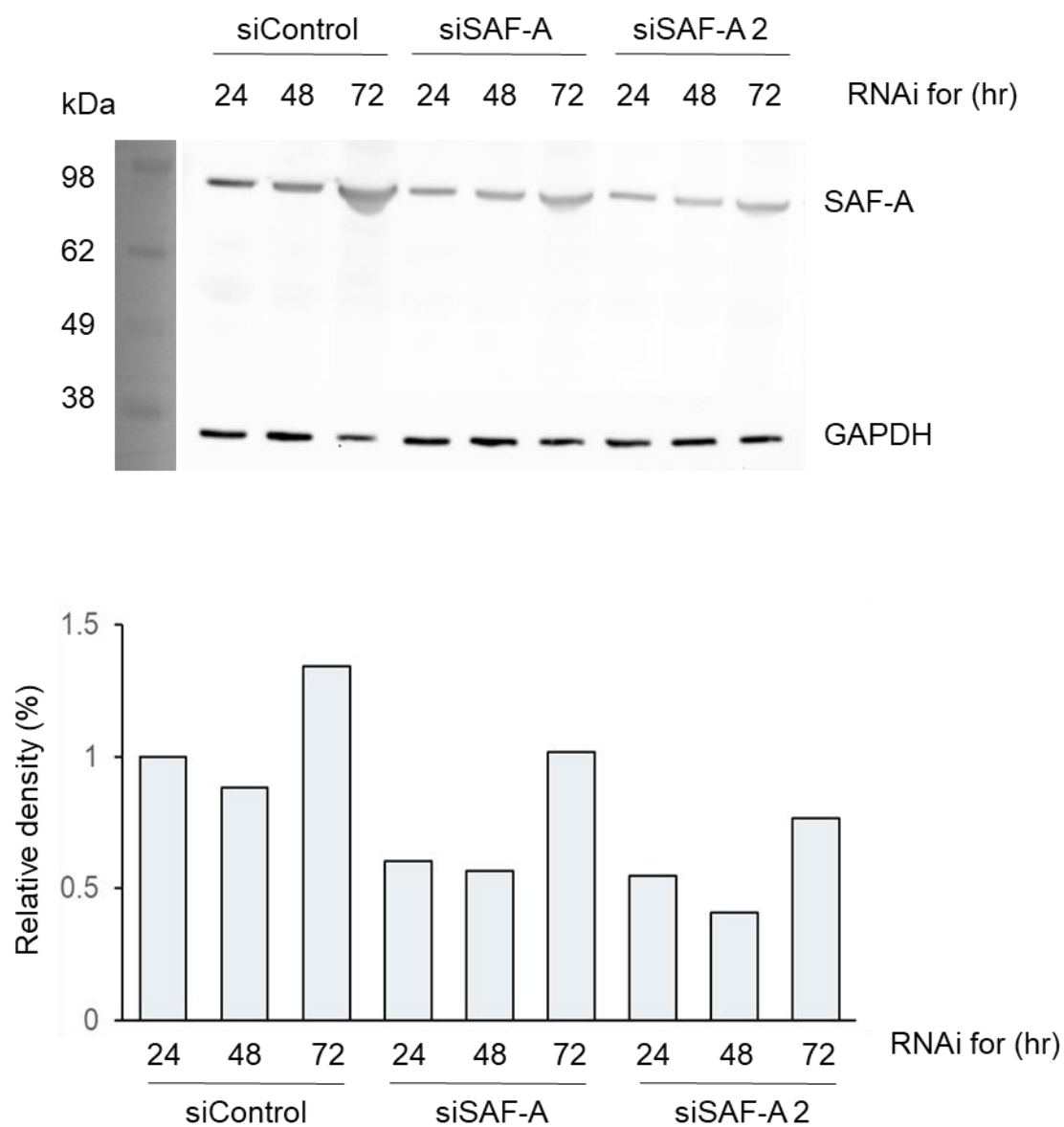


Figure 5.9. SAF-A depletion using RNAi in FlpIn 293 T-REx cells. Western blot (top) of FlpIn 293 T-REx cells treated with siControl, siSAF-A or siSAF-A 2 for 24, 48 or 72 hours. Probed with anti-SAF-A antibody and GAPDH as a loading control. Bar plot (bottom) of the intensity of SAF-A bands normalised to GAPDH intensity to show the relative density of Western blot bands for each condition.

To improve the knockdown of endogenous SAF-A, the concentration of transfection reagent, RNAiMAX was increased from 16 μ l to 36 μ l according to manufacturer's recommendations. Previous experiments had successfully used 16 μ l of RNAiMAX as it was shown to be sufficient in literature (Nozawa *et al.*, 2017). The same experiment as above was repeated with the original amount of transfection reagent (Figure 5.10A) and the new increased amount (Figure 5.10B). The observation in Figure 5.9 where the 72 hr time points had a much higher band intensity than the other time points was not seen for the repeat transfection (0.91 compared to 1 for 24 hr and 0.9 for 48 hr) and was partially true for the 72 hr siControl cells when treated with 36 μ l RNAiMAX (1.13 compared to 1 for 24 hr and 0.95 for 48 hr). The repeat transfection had a 40% to 60% reduction with siSAF-A and a 60% to 80% reduction with siSAF-A 2. Meanwhile, the increased amount of transfection reagent had a smaller effect with a 10% to 40% reduction with siSAF-A and a 25% to 70% reduction with siSAF-A 2. Therefore, the old amount of 16 μ l RNAiMAX continued to be used in latter experiments. As the KD with siSAF-A in FlpIn 293 T-REx was not as pronounced as with other cell lines (HCT 116 and RPE-1 in Chapter 3.4.1) it would not have been ideal to study the effect of mutants in isolation. However, results might be closer to the patient phenotypes where they still have one wild-type copy of the protein which should account for half of the overall protein and one mutant copy account for the rest. Therefore, a siSAF-A KD with a 40% to 60% reduction, as is the case here, can be used to investigate heterozygous mutation patient phenotypes.

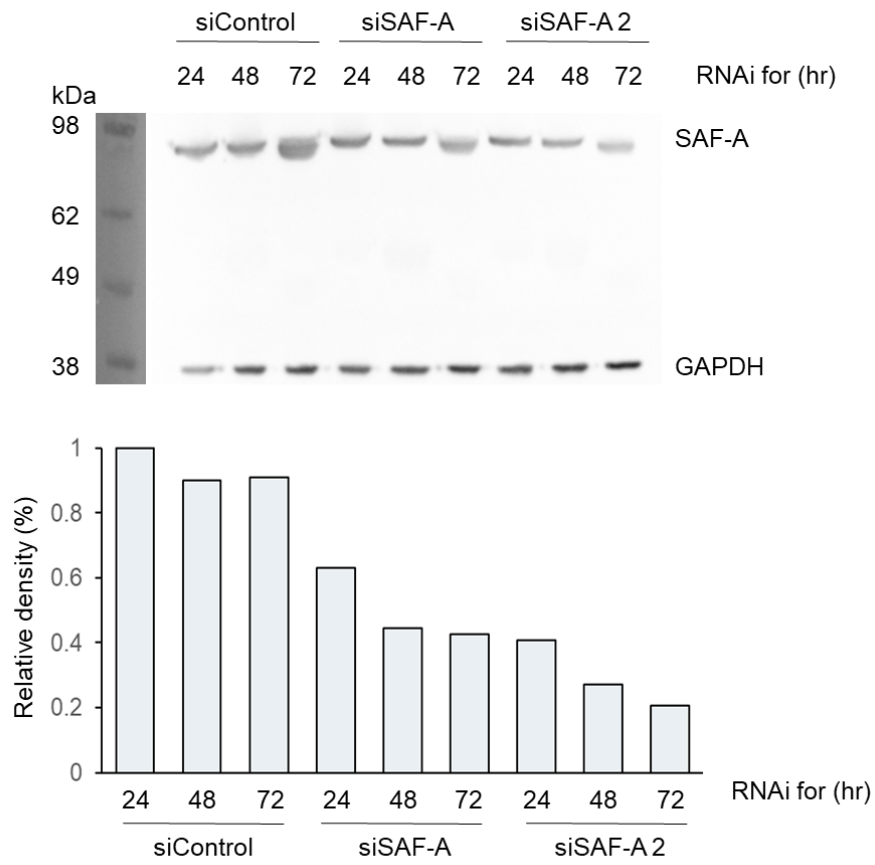
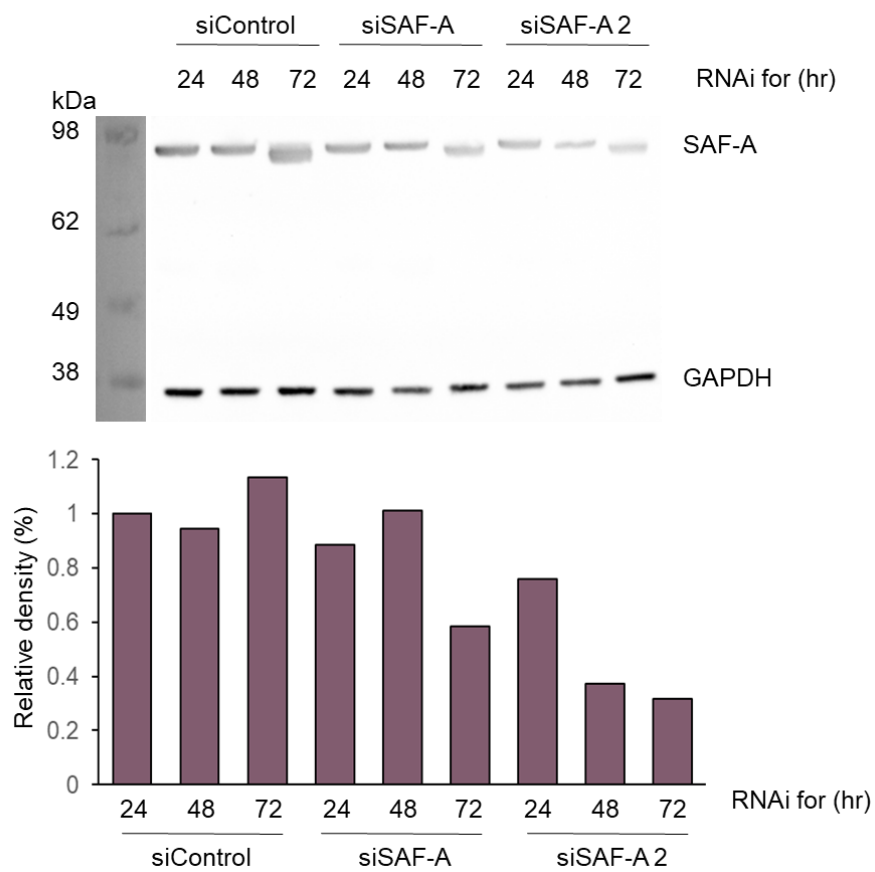
A**B**

Figure 5.10. SAF-A depletion using different concentrations of RNAiMAX reagent in FlpIn 293 T-REx cells. A) Western blot (top) of FlpIn 293 T-REx cells treated with siControl, siSAF-A or siSAF-A 2 for 24, 48 or 72 hours. Probed with anti-SAF-A antibody and GAPDH as a loading control. Bar plot (bottom) of the intensity of SAF-A bands normalised to GAPDH intensity to show the relative density of Western blot bands for each condition. B) same as A) but with 36 μ l RNAiMAX.

Finally, to check whether FlpIn 293 T-REx cells were generally more resistant to RNAi or just to SAF-A KD, they were transfected with siRNA targeting HNRNPUL1, a SAF-A paralog. To compare the effectiveness of the KD, HCT 116 cells were also used. The two cell lines were lysed after 48 hr treatment with either siControl, siSAF-A or siHNRNPUL1 and probed with anti-HNRNPUL1 and anti-GAPDH antibodies (Figure 5.11). The Western blot showed that the siRNA treatment for HNRNPUL1 resulted in a successful KD for both cell lines, indicating there is something unique about SAF-A in FlpIn 293 T-REx cells.

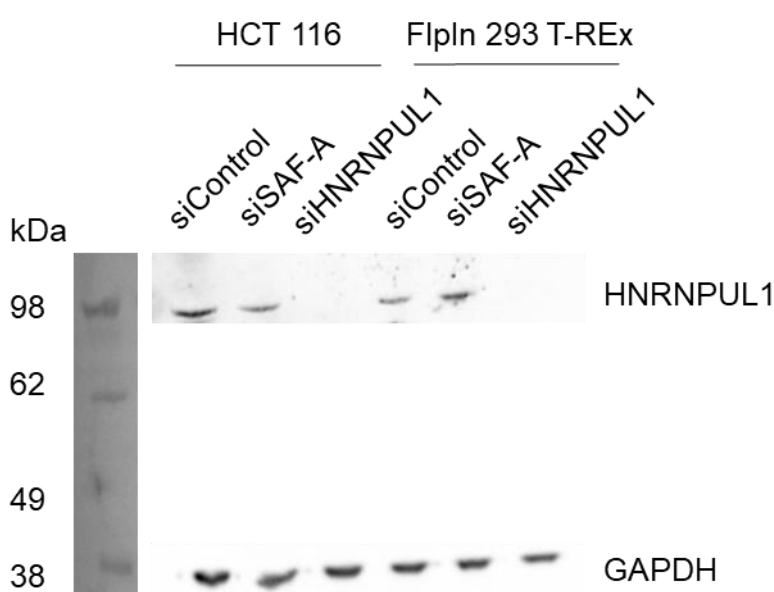


Figure 5.11. HNRNPUL1 depletion using RNAi. Western blot of HCT 116 and FlpIn 293 T-REx cells treated with siControl, siSAF-A or siHNRNPUL1 for 48 hours. Probed with anti-HNRNPUL1 antibody and GAPDH as a loading control.

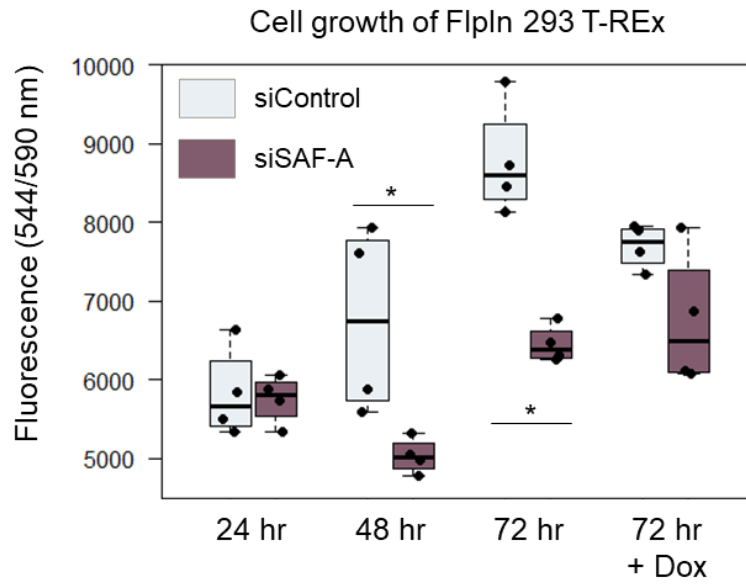
5.3.2. Cell growth of patient mutation cells

As SAF-A depletion has been shown to affect cell proliferation (Chapter 4.2.1), measuring cell growth after RNAi KD and with or without mutant variant expression can show whether the mutants had an effect on the cells. In a simple rescue experiment, parental FlpIn 293 T-REx cells, cells expressing FLAG-SAF-A, and cells expressing the three FLAG-tagged mutant variants (P487S, R714S, Y805*) were transfected with siControl or siSAF-A for 24, 48 or 72 hr. At the specific time points, they were incubated with PrestoBlue to measure the cell population growth. Each cell line was also separately incubated with Dox for 24 hr after 48 hr of RNAi treatment and then incubated with PrestoBlue. I hypothesised that the parental cell line would be affected by the siSAF-A KD at the later time points as previous Western blots showed a drop in SAF-A abundance after 24 hr of KD (Figure 5.9 and 5.10). The parental cell line showed no difference in cell growth at 24 hr but there were statistically significant differences found at 48 and 72 hr of treatment when using a Wilcoxon test ($p = 0.029$) which corresponded to the continuing reduction in SAF-A levels (Figure 5.12A). However, when the parental cells were treated with Dox for 24 hr, there were no statistically significant differences. As they do not express any additional protein, it is possible the Dox treatment affected the viability of the control cells but it is unusual.

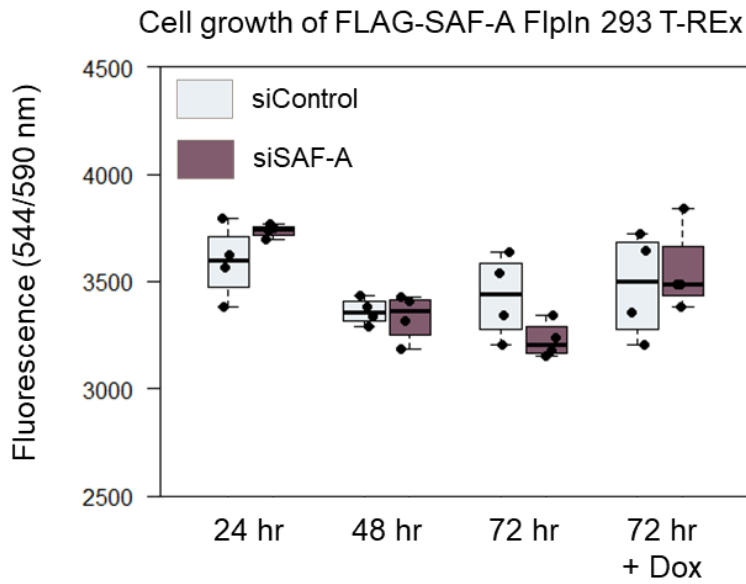
Next, I hypothesised that if the patient mutations negatively affected SAF-A function, there should be a statistically significant difference in cell proliferation between siControl and siSAF-A samples at the 72 hr RNAi time point when Dox is added 24 hr earlier. First, I tested FLAG-SAF-A expressing cells which should rescue the SAF-A depletion phenotype seen at 48 and 72 hr post-siSAF-A treatment in the parental cell line as the expressed protein is the wild-type variant. However, there were no statistically significant differences for the cells at those time points with or without Dox treatment (Figure 5.12B). It is possible that the RNAi treatment was unsuccessful for the FLAG-SAF-A expressing cell line unlike with the parental cell line (Figure 5.12A). Continuing with the patient mutants, the cells carrying the P487S SAF-A mutant exhibited a similar pattern to the FLAG-SAF-A cells, showing no difference between siControl and siSAF-A at 48 and 72 hr (Figure 5.12C). However, a statistically significant difference in cell growth was observed at 24 hr between siControl and

siSAF-A (Wilcoxon test, $p = 0.029$), unlike in the parental cell line. In contrast, the cells carrying the R714S SAF-A mutant did not show any statistically significant differences at any time point, except during Dox treatment, where the siSAF-A value was significantly higher than the siControl (Figure 5.12D, Wilcoxon test, $p = 0.03$). The cells carrying the Y805* mutant exhibited statistically significant differences only at 72 hr of RNAi treatment (Figure 5.13E, Wilcoxon test, $p = 0.029$). Interestingly, the Y805* mutant displayed a potential rescue in the knockdown phenotype when treated with Dox, as there was no statistical difference between the conditions and the average siSAF-A value increased compared to the 72-hr treatment without Dox. However, since the other cell lines did not show any effect of RNAi knockdown compared to the parental cell line, these results remain inconclusive. Overall, the cell lines with additional SAF-A expression showed a stunted growth rate compared to the control, as both conditions for each time point exhibited similar fluorescence values, while the fluorescence values of the parental cell lines continued to increase, at least for the siControl values.

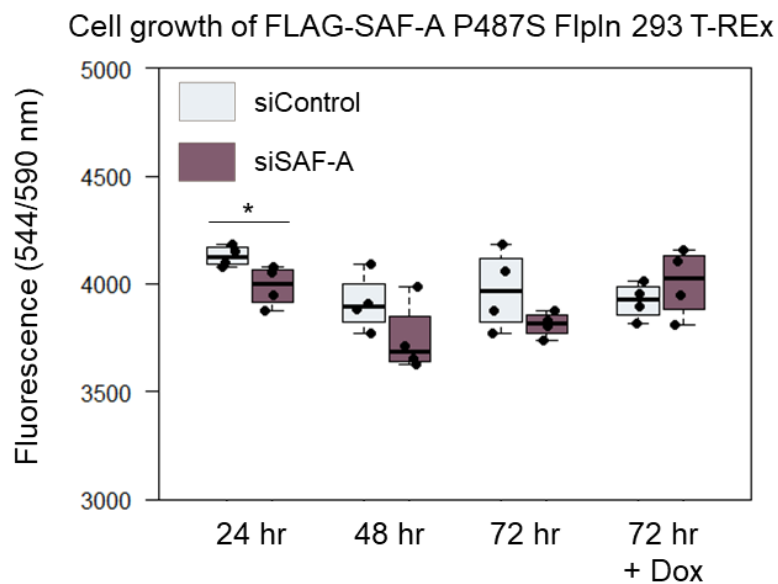
A



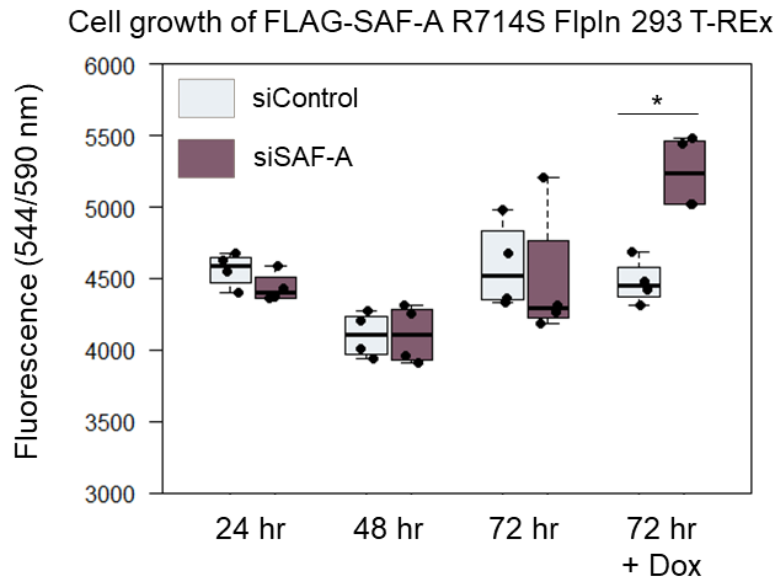
B



C



D



E

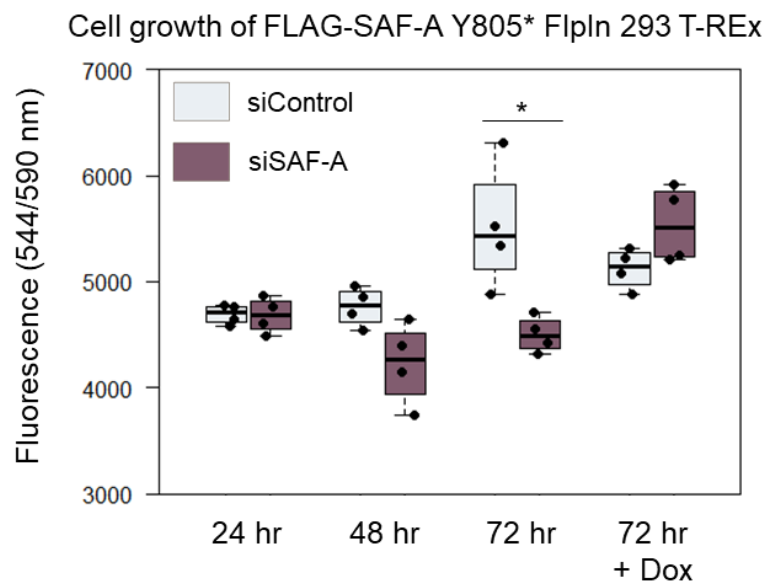


Figure 5.12. Cell growth of FlpIn 293 T-REx cell lines depleted of SAF-A with RNAi. Values on the y-axis were taken with a plate reader measuring fluorescence at 544/590 nm wavelength following a 10-minute treatment with the PrestoBlue reagent. Each box plot consists of 4 technical replicates and shows the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the different conditions at each time point (Wilcoxon test, * = $p \leq 0.05$). Legends show the RNAi treatment for each cell line. A) Cell growth of parental FlpIn 293 T-REx. B) Cell growth of FlpIn 293 T-REx cells expressing FLAG-SAF-A in the presence of doxycycline. C to E) Same as B) but expressing the P487S mutant, R714S, and Y805* mutants respectively.

As measuring cell growth in the FlpIn 293 T-REx cell lines was inconclusive, a similar experiment was designed using the HCT 116 cell line and transient plasmid transfections instead. Endogenous SAF-A in the HCT 116 can be successfully knocked down using siSAF-A as previously shown in Chapter 3.4.1. Therefore, I transfected HCT 116 cells with siControl or siSAF-A for 24, 48 or 72 hr and incubated the cells with PrestoBlue to measure their proliferation (Figure 5.13A). There were no significant differences between the two conditions at 24 and 48 hr but there was a statistically significant difference at 72 hr (Wilcoxon test, $p = 3.6 \times 10^{-5}$). This is consistent with SAF-A being depleted when probed with anti-SAF-A antibody after RNAi for 72 hr (Figure 3.11C). The fluorescent values of siSAF-A cells at 72 hr post-RNAi were close to those at 48 hr, which suggests that the cell population had stopped growing at the latter time point. I then transfected HCT 116 cells 48 hr after RNAi treatment with either siControl or siSAF-A with plasmids expressing either FLAG-tagged SAF-A, WA mutant SAF-A, P487S mutant SAF-A, R714S mutant SAF-A or Y805* mutant SAF-A and measured their fluorescent values after PrestoBlue incubation 24 hr post-transfection (Figure 5.13B). All the transfectants showed a statistically significant difference in cell growth between the siControl and siSAF-A conditions (Wilcoxon test, $p = 0.03$) which shows that there was no full rescue of the KD phenotype on cell proliferation. It is possible that 24 hr were not sufficient for the cells to fully recover from 48 hr of siSAF-A KD as even the plasmid expressing endogenous SAF-A (FLAG-SAF-A) did not fully rescue the phenotype.

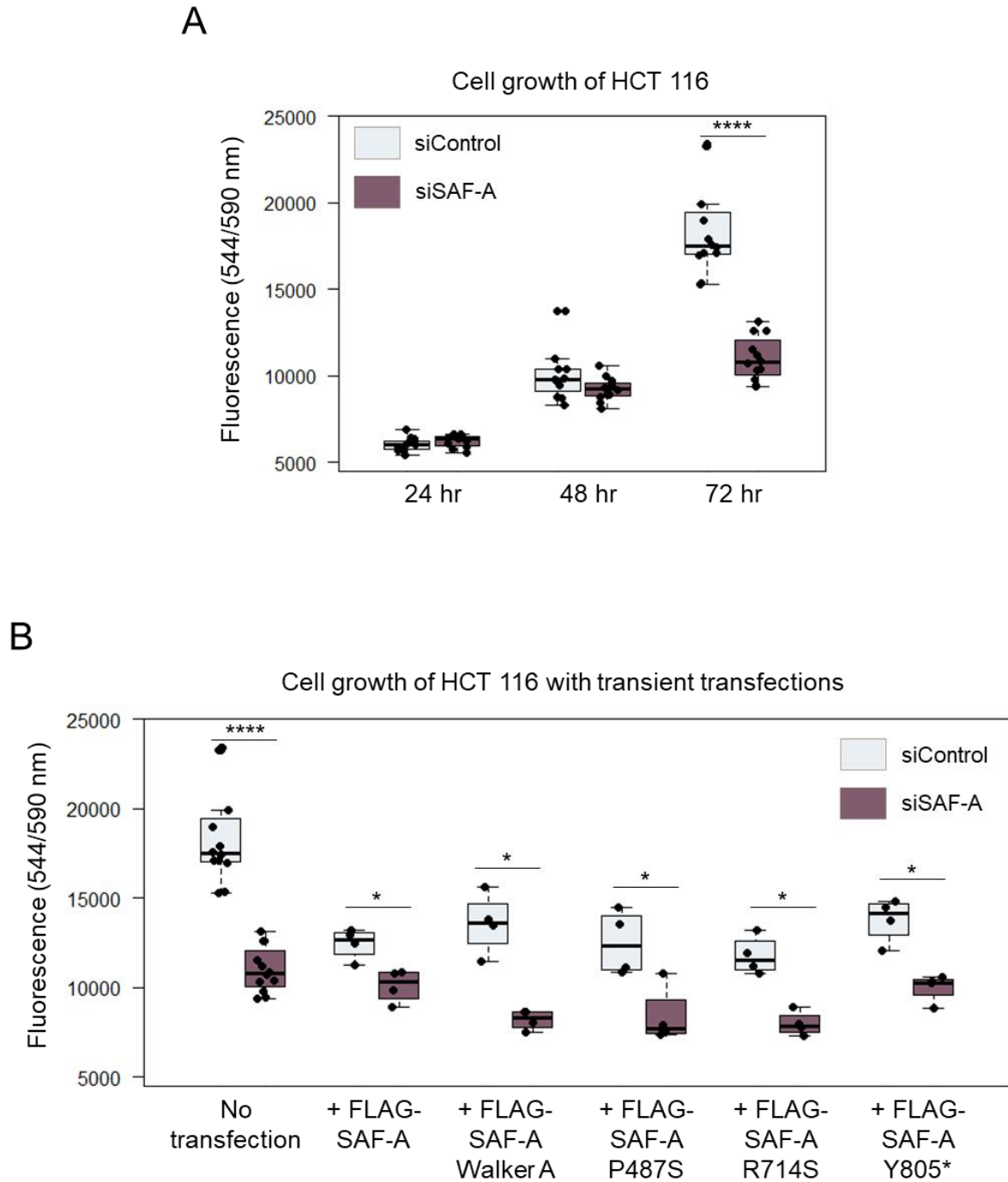


Figure 5.13. Patient mutation effect on cell growth of HCT 116 cells depleted of SAF-A with RNAi. Values on the y-axis were taken with a plate reader measuring fluorescence at 544/590 nm wavelength following a 10-minute treatment with the PrestoBlue reagent. Each box plot consists of 12 (untransfected) or 4 (transfected) technical replicates and shows the median, lower and upper quartiles, and max and min values with each cell value shown as a black dot. Statistically significant differences are marked with stars and bars between the different conditions at each time point (Wilcoxon test, * = $p \leq 0.05$, **** = $p \leq 0.0001$). Legends show the RNAi treatment for each transfection. A) Cell growth of HCT 116 cells treated depleted of SAF-A with RNAi. B) Cell growth of HCT 116 cells transfected with different FLAG-tagged SAF-A plasmids at 48 hours post-RNAi treatment.

5.3.3. Changes in chromatin decompaction due to patient mutations

As differences in cell proliferation showed inconclusive results, I decided to investigate whether the mutants could have a dominant negative effect on large-scale chromatin structure. To achieve this, I performed FISH with probes across a gene-rich region on chromosome 11 (p15.5) to see whether the mutants would rescue a SAF-A KD phenotype on interphase chromatin decompaction. Previous experiments in Chapter 3.4.1. showed that after a 24 hr SAF-A KD in HCT 116 cells, siSAF-A samples were significantly more compacted than siControl when probing in a gene-rich region (Figure 3.11D). Therefore, HCT 116 cells were treated with siSAF-A for 24 hr and then transfected with plasmids expressing FLAG-tagged SAF-A, and SAF-A mutants (WA, P487S, R714S, Y805*) for 24 hr post-transfection. Comparing the chromatin decompaction in cells treated with siControl or siSAF-A for 48 hr showed that there was a statistically significant difference in this gene-rich region (Figure 5.14; Wilcoxon test, replicate 1 $p = 8.3 \times 10^{-12}$, replicate 2 $p = 4 \times 10^{-7}$) like in region 11p15.1 in Chapter 3.4.1. Expressing FLAG-SAF-A rescued the compaction phenotype as there was no statistically significant difference with siControl but there was with siSAF-A (Wilcoxon test, replicate 1 $p = 2.3 \times 10^{-5}$, replicate 2 $p = 1.1 \times 10^{-8}$). The opposite happened when expressing the WA mutant SAF-A where there were statistically significant differences with siControl but none with siSAF-A (Wilcoxon test, replicate 1 $p = 3.5 \times 10^{-10}$, replicate 2 $p = 1.8 \times 10^{-5}$). The results of the patient mutants fell in between FLAG-SAF-A and WA results. Firstly, the P487S mutant rescued the chromatin compaction phenotype as it was statistically different to siSAF-A but not to siControl (Wilcoxon test, replicate 1 $p = 3.1 \times 10^{-7}$, replicate 2 $p = 3.5 \times 10^{-7}$). Next, the R714S mutant partly rescued the chromatin compaction phenotype as it was statistically different to both siControl and siSAF-A (Wilcoxon test, replicate 1 vs siControl $p = 0.0003$, replicate 2 vs siControl $p = 0.01$, replicate 1 vs siSAF-A $p = 0.001$, replicate 2 vs siSAF-A $p = 0.002$). Finally, the Y805* mutant failed to rescue the phenotype as it was statistically different only to siControl (Wilcoxon test, replicate 1 $p = 1.7 \times 10^{-8}$, replicate 2 $p = 1.2 \times 10^{-10}$). Compared to the rescued when siSAF-A cells are transfected with FLAG-SAF-A, transfection with Walker A (Wilcoxon test, replicate 1 $p = 0.0002$, replicate 2 $p = 8.9 \times 10^{-7}$), R714S (Wilcoxon test, replicate 1 $p = 0.002$, replicate 2 $p = 0.0007$), and Y805* (Wilcoxon test, replicate 1 $p = 0.005$,

replicate 2 $p = 3.5 \times 10^{-15}$) were all statistically different. Overall, as the cell growth results previously were inconclusive on whether there was a patient phenotype, screening for chromatin compaction rescue in cells showed that each mutant had different effects. P487S led to chromatin decompaction re-establishment like with wild-type FLAG-SAF-A, R714S led to partial decompaction, and Y805* did not decompact chromatin similar to the WA mutant.

Figure 5.14. Patient mutation rescue of chromatin compaction after HCT116 cells are depleted of SAF-A using RNAi. Representative images (top) of FISH with probes for the gene-rich 11p15.5 region of the genome in HCT116 cells treated with siControl or siSAF-A and transfected with FLAG-tagged SAF-A plasmids at 24 hours post-RNAi treatment. The white scale bars represent 10 μ m. Distribution (bottom) of inter-probe distances between probe pairs. Each violin plot (read from left to right, values from 2 biological replicates) represents a probe pair sample size of $N_1 = 186/104$, $N_2 = 128/100$, $N_3 = 100/189$, $N_4 = 163/121$, $N_5 = 80/183$, $N_6 = 168/253$, $N_7 = 265/292$. Wilcoxon test, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$. Replicates in different colours as shown, differences between transfection and siControl are shown with black asterisks, between transfection and siSAF-A are shown with red asterisks and between FLAG-SAF-A transfection and the other transfectants with blue asterisks.

5.3.4. Effect on SAF-A oligomerisation

Another experiment to investigate the patient mutant phenotype is the oligomerisation assay. In this assay SAF-A is immunoprecipitated in the presence or absence of the crosslinker BM(PEG)₂ and loaded on a protein gel to visualise whether it forms oligomers (Nozawa *et al.*, 2017). It was previously reported endogenous SAF-A forms oligomers in the presence of caRNA and ATP which can be seen when crosslinked and certain mutations disrupt this behaviour. Specifically, the WA mutant led to exclusively monomeric SAF-A and the WB mutant was primarily oligomeric. Therefore, comparing the amount of monomeric and oligomeric endogenous SAF-A when patient mutants are expressed in FlpIn 293 T-REx cells can show whether the variants affect the oligomerisation ability of the protein. P487S was hypothesised to affect SAF-A similarly to the WA mutant as it was also in the Walker A motif. The R714S mutant might also have an effect as it could alter the RNA binding capacity of SAF-A and thus reduce the amount of oligomeric protein. The Y805* mutant could also affect oligomerisation if the C-tail end of SAF-A was important for stabilisation of the AAA+ domain due to its vicinity to the missing amino acids.

FlpIn 293 T-REx cells expressing FLAG-tagged SAF-A and patient mutants were treated with trypsin and lightly permeabilised before crosslinking. They were then either treated with DMSO or 1 mM BM(PEG)₂ for 10 minutes to crosslink SAF-A oligomers. After sonication to disrupt the nuclear mesh, input samples were taken and lysed for subsequent visualisation. Then FLAG-tagged SAF-A was immunoprecipitated with magnetic beads coated in anti-FLAG antibodies and then

directly lysed before preparing it for gel loading. The lysates were probed with anti-SAF-A antibodies to visualise both endogenous and FLAG-tagged SAF-A (Figure 5.15). The BM(PEG)₂ crosslinking was successful (Figure 5.15A) as there was nearly no monomeric SAF-A visible (estimated 90 kDa for endogenous, 93 kDa for FLAG-SAF-A) in elute samples compared to DMSO treatment (Figure 5.15B). The immunoprecipitation was also successful when comparing the input versus elute samples. In both conditions, the input samples had multiple unknown bands starting from under SAF-A until the end of the gel (between 90 and 45 kDa) which were absent in the eluted fraction. There was one band at around 45 kDa in the eluted DMSO-treated samples that is of unknown identity. FlpIn 293 T-REx cells with an inserted FLAG-SAF-A in the absence of Dox had no higher band in the monomer range when treated with DMSO due to lack of expression. Similarly, the Y805* mutant had no higher band which showed there was a problem with its expression in this experiment. There are only a dimer band visible for the Y805* mutant when crosslinked with BM(PEG)₂ which is likely endogenous SAF-A. The wild-type, P487S and R714S mutants all had higher MW bands visible for dimeric, trimeric, and potentially oligomeric SAF-A which shows that these patient mutations might not interfere with the ability of the protein to oligomerise. However, there was a protein signal visible in the gel well at the top and the high MW oligomers have not been separated well which is where the main differences in oligomerisation can be seen in previous research (Nozawa *et al.*, 2017). Attempts to conduct gel electrophoresis for longer times either at room temperature or at 4°C resulted in less clear results that were difficult to interpret (data not shown).

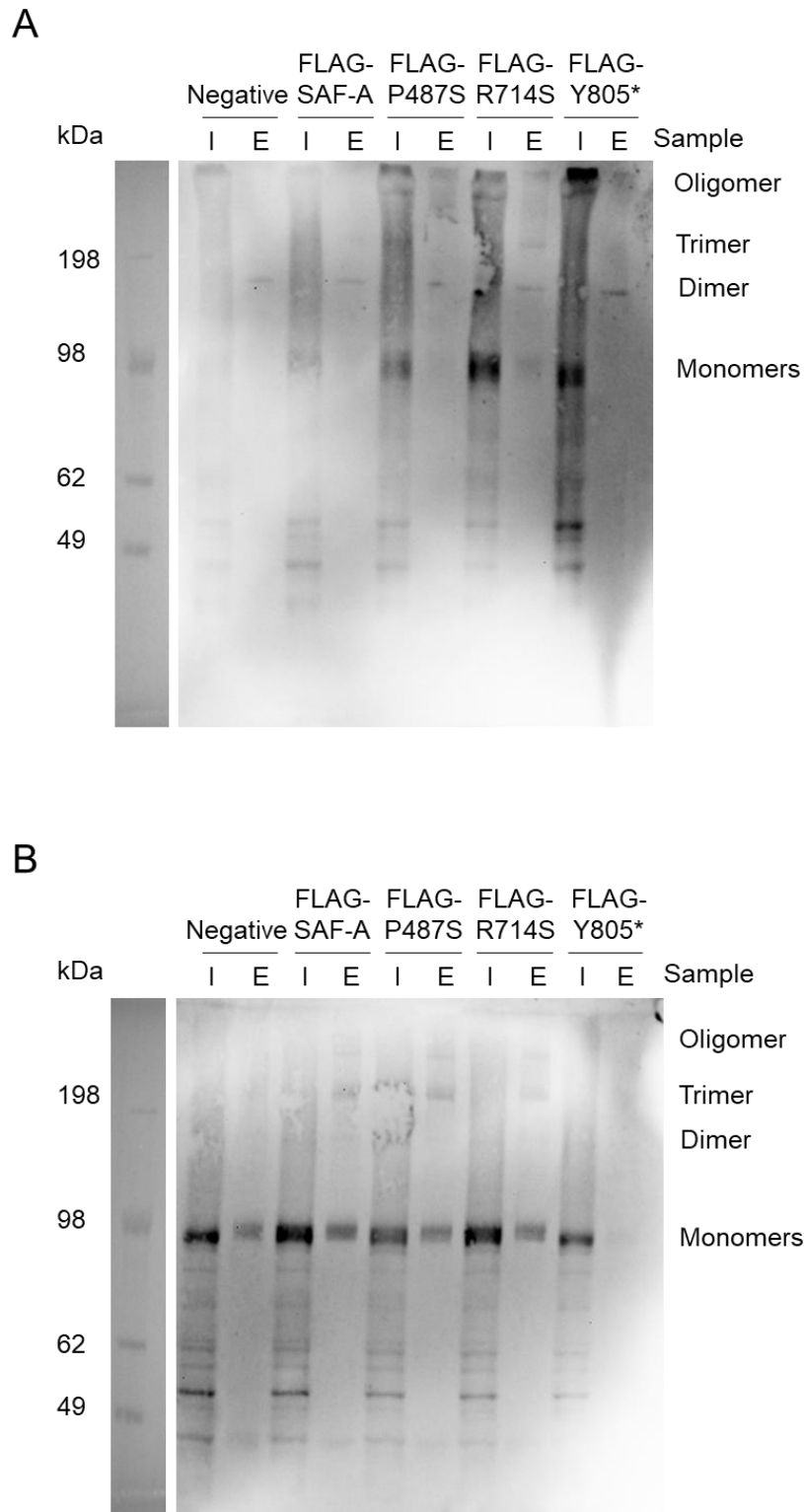


Figure 5.15. Effect of patient mutations on SAF-A protein oligomerisation. Western blots of FlpIn 293 T-REx cells expressing FLAG-SAF-A or patient mutants in the presence of Dox which were crosslinked with BM(PEG)₂ (A) or treated with DMSO (B). The negative sample is the parental cell line without Dox treatment. Input samples (I) were collected after sonication and centrifugation, before addition to FLAG-tagged magnetic beads. Elute samples (E) were eluted from FLAG-tagged magnetic beads. The lysed samples were fractionated by SDS-PAGE on a 3-8% tris-acetate gel and probed with anti-SAF-A antibodies.

5.4. Discussion

In this chapter, I aimed to investigate the effect of clinical mutations found in SAF-A on its role in chromatin organisation to further understand the cause and phenotype of the HNRNPU-related neurodevelopmental disorder. I first selected which patient mutants to investigate in cells combining protein conservation data, variants in the general population, clinical data on pathogenic variants and literature on SAF-A protein domains. Due to time limitations, I could investigate only three mutants, so I formulated selection criteria to potentially choose the mutants that are most likely to have an effect different from LOF leading to haploinsufficiency. The three patient mutations that were selected for further experiments were P487S/P506S, R714S/R733S, and Y805*/Y824* (SAF-A isoform/SAF-A full length). There were limitations to my criteria that could have led me to overlook other interesting mutants. For example, I disregarded sites of overlap between pathogenic mutations and population-wide missense variants even if the mutation was not the same. While that specific missense variant in the general population might not have led to the development of the disorder, the pathogenic mutant clearly has. In addition, some of my reasoning was based on the predicted model of the protein by AlphaFold, specifically the importance of the C-tail of the protein. According to AlphaFold, the positioning of the C-tail end with amino acids 824 and 825 had a low confidence score (a per-residue confidence score (pLDDT) between 70 and 50) and the IDRs of SAF-A had a very low confidence score (pLDDT below 50) in general. In comparison, the SAP, SPRY and AAA+ domains had very high (pLDDT over 90) or confident model scores (pLDDT between 90 and 70). While AlphaFold has high protein prediction for folded proteins it performs less well with IDRs (Strodel, 2021). The true structure of the protein might not have the C-tail end of the protein linking back to the SPRY and AAA+ domain but there has been no successful crystallisation of the full-length protein to investigate this. Finally, I did not consider the structural effect of the candidate mutations and whether they might destabilise the folding of the protein. In general, non-conservative missense mutations are more likely to lead to destabilisation by affecting the secondary protein structure. Out of the candidate missense mutations (Y293C, Y381C, V479I, P506S, R733S) only P506S and R733S can be classified as non-conservative due to the change of a nonpolar side chain to a

polar one for the former and a positively charged one to a polar one for the latter. However, conservative mutations should still not be disregarded as they could cause dominant-negative or gain-of-function phenotypes (Backwell and Marsh, 2022). Further investigation of variant function can be performed using FoldX for stability calculations of pathogenic mutations and different variant effect predictors (Gerasimavicius *et al.*, 2022).

To perform further experiments with the chosen patient mutations, they were integrated into FlpIn 293 T-REx cells under Dox control. Ideally, coupling the expression of the mutants with KD of endogenous SAF-A with RNAi would have allowed me to isolate the effect of the variants. However, the cell line was resistant to full SAF-A KD and could achieve at best 60% KD when treated with siSAF-A for over 48 hr (Figure 5.10A). When comparing to test experiments I had done early on in HEK-293 cells, SAF-A was also not fully knocked down after 72 hr of RNAi treatment (data not shown). Other labs have also confirmed having limited success with the cell line and SAF-A KD. While this amount of KD would not have isolated the effect of the patient mutants, it would be closer to the biological reality where potentially half of the protein is wild-type and half is mutated. When the cell proliferation of the cell lines was measured using a PrestoBlue assay (Figure 5.12), there did not appear to be a mutant phenotype. Although, in general, the RNAi KD of SAF-A did not result in similar results in the reduction of cell growth across all the cell lines. A better alternative would be to use a cell line that is more susceptible to SAF-A KD and has consistent results. Due to time limitations, I could only transiently transfect HCT 116 cells which I had previously used to show the chromatin compaction phenotype due to SAF-A KD (see Chapter 4.2.3.). These results also proved to be inconclusive (Figure 5.13) as there was not a sufficient rescue with FLAG-SAF-A. Expressing the protein variants for longer than 24 hr after 48 hr of SAF-A KD would allow sufficient time to differentiate between potential phenotypes. It is possible that certain cell types are more susceptible to reduced SAF-A levels than others. Therefore, another cell line that could be utilised instead is RPE-1 FlpIn T-REx where there is an FRT site for inserting a gene of interest. In previous experiments I have been able to KD SAF-A in RPE-1 cells which had an effect on chromatin decompaction (see Chapter 4.2.3.), therefore it would be interesting to try these assays in these cells. An alternative cell line could be one closer to human neurons to accurately identify the potential

molecular mechanisms of pathogenic SAF-A variants. An example is SH-SY5Y cells which can be terminally differentiated into mature neurons and are often used for studying Parkinson's disease (Kovalevich and Langford, 2013; Xicoy *et al.*, 2017). These cells would be a good model as it is possible that the HNRNPU-related neurodevelopmental disorder develops when neural cells propagate or differentiate with altered transcriptional states. Evidence for that can be found in a recent study of a heterozygous frameshift mutant in human cortical organoids (Ressler *et al.*, 2022). There was a higher burden of dysregulated genes found in precursor cell populations than in more mature neurons. These results show that certain cell types are more susceptible to reduced SAF-A levels at least in 6.5-week-old organoids. Therefore, investigating the effect of SAF-A mutants on these key cellular processes could elucidate the cause of aberrant neuronal development.

As comparing changes in cell growth when expressing mutants in cells depleted of SAF-A showed inconsistent results, other experiments were used to investigate molecular phenotypes. One of these was FISH in a gene-rich region in cells depleted of SAF-A and expressing mutants. In Chapter 3.4.1, I had shown that 24 hr SAF-A KD, while not completely depleting the cells of the protein, compacted the gene-rich region 11p15.1 and had no effect on the gene-poor region 11p14.1 in both HCT 116 and RPE-1 cells. In this chapter, I depleted HCT 116 cells of SAF-A due to their more profound chromatin compaction phenotype and higher transfection efficiency. After a 24-hr KD and a subsequent 24-hr expression of plasmids with FLAG-tagged versions of SAF-A, I measured the distances between two loci in the gene-rich 11p15.5 region (Figure 5.14). My positive control (FLAG-SAF-A) and negative control (FLAG-SAF-A WA) confirmed that the experiment was successful as each of them respectively rescued the chromatin compaction or had no effect on it. The FISH experiment was able to highlight the different phenotypes of the patient mutants. I hypothesised that the P487S mutant would have a similar effect to the WA mutant as it is within the WA motif. However, the P487S mutant rescued the chromatin compaction phenotype when SAF-A is partly depleted, unlike the WA mutant. It is possible that the P487S mutant instead has an effect closer to the WB mutant, which causes chromatin to be constitutively open in the absence of transcription (Nozawa *et al.*, 2017). As the chromatin decompaction role of SAF-A is thought to be downstream of transcription, the wild-type protein cannot decompact gene-rich chromatin with transcription

perturbed, in the absence of RNA or ATP. However, the WB mutant decompacts chromatin even in the absence of transcription. To investigate whether P487S or other mutants have this effect, the experiment can be repeated with pre-treatment of the slides, after mutant expression, with a transcription inhibitor like α -amanitin or triptolide and compare chromatin structure to the WB phenotype. The R714S mutant partially rescued the chromatin compaction phenotype as it was statistically different to both the siSAF-A and siControl measurements. The mutation is within the first RGG repeat in the RGG motif of SAF-A which is an important motif for RNA-binding due to the strong positive charge and presence of aromatic residues (Kiledjian and Dreyfuss, 1992). I hypothesised that this mutation could affect the RNA-binding capacity of the protein but possibly not significantly as it is only one out of five RGG repeats altered in the RGG motif and one out of 13 RG and RGG repeats in IDR 2 in total. As the results show a partial rescue from chromatin compaction, the mutant effect is stronger than hypothesised. However, in patients, this might not be the molecular mechanism responsible for the neurodevelopmental disorder as only half of the proteins will have reduced RNA-binding capabilities. Finally, the Y805* mutant was hypothesised to destabilise the SAF-A protein due to the removal of the C-tail end of the protein nested between the SPRY and AAA+ domains in the predicted AlphaFold protein model. Indeed, this mutant failed to rescue the chromatin compaction phenotype and had the most severe compaction result, comparable to SAF-A protein depletion and the WA mutant. However, it is unknown if this protein will be expressed in patient cells or will be subject to NMD due to destabilisation thus leading to haploinsufficiency. Overall, a repeat of this experiment in the FlpIn 293 cells with partial KD could show whether half of the protein being mutant is enough to affect chromatin decompaction in a dominant negative manner for the R714S and Y805* mutants. To further explore the effects of the P487S mutant, it would be valuable to conduct a similar experiment in the presence of a transcription inhibitor like triptolide. This would help determine whether the mutant maintains chromatin in a constitutively decompacted state similar to the WB mutant. Overall, it might also have been better for this experiment to have been undertaken in a cell line with integrated mutants either using CRISPR or the FlpIn system as plasmid transfection led to many imaging artefacts such as clumping of signals around the cell nucleus and multiple signals within the nucleus. While this was circumvented by disregarding cells with

such aberrant signals in the experiment shown here, it would be easier, faster, and more reproducible in another cell system.

Another experiment to investigate the potential molecular mechanisms behind patient mutations was to do an oligomerisation assay (Figure 5.15). This enables the visualisation of monomeric and oligomeric SAF-A and whether mutants disrupt its ability to form oligomers or to de-oligomerise back into monomers (Nozawa *et al.*, 2017). The results showed oligomer formation for all mutants except Y805* where there did not appear to be FLAG-tagged protein expression for an unknown reason. While it could be that the mutation led to the destabilisation of the protein, it can be visualised on the Western blot following Dox addition (Figure 5.8B and C). Repeats of the oligomerisation assay also showed FLAG-tagged Y805* mutant SAF-A so in this case there might have been a technical error (data not shown). For the other mutants, there were no apparent differences in the monomer or oligomer abundance with or without a crosslinker when compared to wild-type FLAG-SAF-A. A better way to represent the results would have been to also compare them to the WA mutant where there is a reduction in oligomers as a negative control. The Gilbert lab already has a FlpIn 293 T-REx cell line expressing the WA mutant which I did not include at the time. In addition to the WA mutant, the protein lysates could be separated with the immunoprecipitation samples on one gel to compare the protein abundance more directly between conditions. Another method to test oligomerisation and RNA binding involves using a purified segment of the SAF-A protein containing the SPRY and AAA+ domains, mixed with an RNA oligomer and ATP, to explore protein-RNA interactions through an electrophoretic mobility shift assay (EMSA). An EMSA experiment performed in the Gilbert lab with wild-type SAF-A SPRY-AAA fragments or containing WA or WB mutations showed how WA led to reduced RNA binding and WB led to increased and irreversible RNA binding compared to wild-type SAF-A (data not shown). Additional investigations into the molecular mechanisms of patient mutants could involve visualizing their interactions with RNA by using dSTORM microscopy to examine the mutated SAF-A variants and assess the connectedness of protein clusters, thereby determining the mutants' impact on SAF-A cluster density (Marenda *et al.*, 2021). This can also be coupled with RNA mesh visualisation to show if the mutants change the nuclear RNA distribution like SAF-A KD does in Chapter 4.3. Another way to assess alteration in the RNA binding capacity of mutant

proteins is to use radioactive RNA analogues and protein immunoprecipitation as undertaken in Chapter 3.3. Finally, SAF-A mutants have been shown to affect the mitotic progression and might lead to DNA inheritance defects (Douglas *et al.*, 2015; Sharp *et al.*, 2020). Passage through mitosis is also a critical time for cell fate transitions and transcriptional reprogramming (Soufi and Dalton, 2016). To investigate the potential influence of patient mutants on mitotic progression, live cell imaging can be used to observe cells expressing SNAP-tagged mutants as described in Chapter 3.2. For this purpose, FlpIn 293 T-REx cell lines expressing SNAP-tagged mutants have been generated, enabling the visualisation of mutant SAF-A behaviour during mitosis (Figure 5.8C). However, imaging of SAF-A patient mutants during mitosis has not yet been carried out.

Overall, in this chapter, I investigated the impact of the SAF-A mutations P487S/P506S, R714S/R733S, and Y805*/Y824* (SAF-A isoform/SAF-A full length) on chromatin organisation. When SAF-A is depleted from cells using RNAi, the P487S mutant could rescue chromatin compaction, the R714S mutant showed partial rescue, and the Y805* mutant failed to rescue, indicating a potentially severe effect on chromatin compaction similar to complete SAF-A depletion. The Y805* mutant might destabilize the protein, and its expression was undetectable in some assays. Other mutants did not show significant differences in oligomer formation compared to the wild-type SAF-A. Mutants like P487S might maintain chromatin in a decompacted state without transcription, while R714S could affect RNA-binding capacity due to its location in the RNA-binding RGG motif. The Y805* mutation likely impacts protein stability and function due to its position in the predicted AlphaFold model.

6. Chapter 6: Discussion

6.1. Summary

This thesis aimed to investigate the role of SAF-A in chromatin organisation during mitotic exit and early G₁ and its implications for cellular phenotypes, specifically in the context of the HNRNPU-related neurodevelopmental disorder. While both SAF-A and large-scale chromatin organisation have been studied during interphase and mitosis, the processes underlying mitotic exit remains poorly understood. I hypothesised that SAF-A would be recruited to chromatin during telophase or at the mitotic exit when transcription restarts to facilitate the decompaction of gene-active regions.

In Chapter 3, I optimised live cell imaging of SNAP-SAF-A, SNAP-HNRNPC1, and SNAP-HNRNPA1 to observe their behaviour during mitosis under control conditions and with transcription inhibition. Co-localisation analysis and imaging revealed that SAF-A is recruited to chromatin during cytokinesis. Transcription inhibition delayed SAF-A recruitment to early G₁. To examine SAF-A's association with RNA during the cell cycle, I optimized cell synchronization and performed SAF-A oligomer and RNA immunoprecipitation. The results showed that increased amounts of oligomeric SAF-A were associated with both new and old RNA compared to monomeric SAF-A during mitosis and early G₁, indicating that the protein regulates large-scale chromatin organisation immediately upon recruitment during mitosis.

In Chapter 4, I investigated the interaction between SAF-A and RNA and whether SAF-A depletion affects the RNA mesh in interphase. Several methods to deplete SAF-A from cells were tested, with RNAi proving to be the most reliable. By using super-resolution microscopy, SAF-A was visualised as a dense interconnected network of clusters, which were reduced and less connected in the presence of transcription inhibition. Nascent RNA formed a network of smaller clusters that exhibited different clustering patterns at different time points. Depleting SAF-A with RNAi treatment reduced the density and connectivity of RNA clusters. Additionally, a method based on SLAM-seq was developed to investigate nascent RNA, focusing on amplicons from highly expressed genes.

In Chapter 5, SAF-A mutations found in patients with the HNRNPU-related neurodevelopmental disorder were introduced into human cell lines. The pathogenic mutations were analysed, and those meeting specific criteria were selected for further investigation. The mutants did not significantly affect cell viability or SAF-A protein oligomerization; however, they displayed varying degrees of impact on chromatin decompaction at a gene-rich region.

Overall, this thesis provides insights into the role of SAF-A in chromatin organisation during the mitotic exit and early G₁. The findings contribute to our understanding of SAF-A's behaviour and its implications for HNRNPU-related neurodevelopmental disorder.

6.2. SAF-A is recruited to chromatin during mitotic exit in a transcription-dependent manner

Live cell imaging and co-localisation analysis revealed that SAF-A exhibits high co-localisation with chromatin during cytokinesis unlike in the other mitotic phases. This differed from other HNRNP family proteins, HNRNPC1 and HNRNPA1, which were highly co-localised with chromatin during telophase and early G₁ respectively. Previous studies have classified HNRNP family proteins into early and late recruitment groups based on their timing of chromatin association (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). My live cell imaging experiments confirmed these previous observations that HNRNPC1 is recruited early, possibly along with RNA pol II, as the nuclear envelope reforms during telophase, while HNRNPA1 is recruited later, after nuclear envelope formation in early G₁ (Hetzer, 2010). Therefore, by comparing SAF-A and chromatin association to that of HNRNPC1's and HNRNPA1's, I can conclude that the recruitment of SAF-A initiates before the completion of nuclear envelope formation.

Previous studies have shown that the recruitment of late group HNRNP proteins is affected by transcription inhibition, leading to their sequestration in the cytoplasm (Pinol-Roma and Dreyfuss, 1991; Prasanth *et al.*, 2003). Surprisingly, my live cell imaging and co-localization analysis showed the opposite effect. In the presence of

triptolide, a transcription initiation inhibitor, the early recruitment of HNRNPC1 was delayed until cytokinesis, while the recruitment of HNRNPA1 was not affected. However, it is important to note that triptolide treatment might influence RNA pol II binding since it inhibits transcription initiation, thereby affecting the recruitment of HNRNPC1. Repeating the live cell imaging experiments with a transcription elongation inhibitor, such as DRB, could yield different results. The recruitment of SAF-A to chromatin was also affected by triptolide treatment, delaying its initiation until early G₁. Nevertheless, the delayed recruitment did not impact the increased enrichment of SAF-A and HNRNPC1 in the nucleus during early G₁. This suggests that the proteins are primarily recruited by other methods once the nuclear envelope is established, with partial recruitment linked to transcription or RNA pol II binding. It is possible that a small fraction of SAF-A and HNRNPC1 are required for chromatin organisation and RNA processing before the full re-establishment of the nuclear envelope. Additionally, the initial recruitment and accumulation of SAF-A during early G₁ precedes the establishment of cohesin loops, which occurs gradually throughout the G₁ phase (Zhang *et al.*, 2019). Future live cell imaging studies could explore the behaviour of SAF-A during early G₁ as it re-establishes its characteristic diffuse pattern and compare it with cohesin loop establishment timing to gain further insights into large-scale chromatin organization during this phase.

6.3. Oligomeric SAF-A and monomeric SAF-A differ in RNA association during the cell cycle

Synchronising cells using a 2xThy block enabled investigation into the binding of monomeric and oligomeric SAF-A to RNA during mitosis and early G₁. Immunoprecipitation of SAF-A in the presence of the crosslinker BM(PEG)₂ resulted in enrichment of oligomeric SAF-A, consistent with previous findings in the lab (Nozawa *et al.*, 2017). Combining protein immunoprecipitation with RNA radiolabelling using [5-³H]U allowed for comparison of the amount of labelled RNA bound to SAF-A monomers (without crosslinking) and oligomers. Both total RNA and nascent RNA were preferentially associated with oligomeric SAF-A during early G₁ and monomeric SAF-A during cell cycle arrest at the G₁/S phase boundary. This

observation may be attributed to the transcriptional activity of cells at those time points; early G₁ cells might exhibit active transcription, unlike arrested cells, which could drive the formation of SAF-A oligomers to decompact chromatin. To confirm the abundance of SAF-A oligomers or monomers during early G₁ and at the G₁/S phase boundary, a Western blotting experiment could be conducted. In addition, RNA association results at the G₁/S phase boundary were similar to asynchronous cell results when pulsed with [5-³H]U, hinting that it might be the case that oligomeric SAF-A does not interact with nascent RNA 15 min after its transcription. In fact, experiments using dSTORM to examine SAF-A and RNA co-localization demonstrated the highest co-localization 30 minutes after 5-EU incubation. Before repeating the experiment with cell synchronization, it would be interesting to compare RNA bound to SAF-A oligomers and monomers at different time points following [5-³H]U incubation. The results obtained during mitosis showed similar levels of total and nascent RNA signals between monomeric and oligomeric SAF-A. However, given that approximately 20% of cells may potentially be in mitosis at that time, these results may not be conclusive for the mitotic phase. Ideally, a cell sorting experiment could generate a highly enriched population of mitotic cells, providing a better understanding of SAF-A and RNA association during mitosis.

6.4. SAF-A depletion affects the nascent interconnected RNA mesh

The dSTORM super-resolution microscopy technique, combined with data analysis using the SuperStructure method, revealed that SAF-A forms dense and highly connected protein clusters during interphase (Marenda *et al.*, 2021). However, transcription inhibition using triptolide disrupted these SAF-A clusters, leading to a reduction in their size and connectivity. These clusters may represent hotspots of oligomeric SAF-A at transcriptionally active chromatin regions. Currently, we lack a method to distinguish between oligomeric and monomeric SAF-A when imaging cells. One possible approach could be to express FLAG-tagged versions of WA and WB SAF-A mutants, which could be visualized using dSTORM. By comparing the cluster density and connectivity of these mutants to wild-type SAF-A, we could gain insights into the role of oligomeric SAF-A in chromatin organisation.

Nascent RNA imaged using dSTORM also formed an interconnected network similar to SAF-A but with a smaller cluster radius. Newly transcribed RNAs at 0 hours after 5-EU incubation exhibited the largest cluster radius compared to RNAs observed 30 to 60 minutes later. However, these early RNAs displayed the lowest connectivity between clusters. Over time, RNA clusters became more connected, leading to the characteristic strong signal of nascent RNA at later time points. Hence, the clustering patterns of nascent RNA differ depending on their age. When examining the co-localization of SAF-A signal with nascent RNAs, the highest degree of co-localization was observed at 30 minutes after 5-EU incubation, followed by 0 hours and 1 hour. The functional significance of the increased co-localization between SAF-A and nascent RNA remains unknown, as it is unclear whether this interaction triggers SAF-A oligomer formation for chromatin decompaction.

6.5. SAF-A mutations from patients with the HNRNPU-related neurodevelopmental disorder

Through the analysis of SAF-A missense and synonymous variants in the general population, as well as pathogenic mutations using the 1D-to-3D method and Vd/Vp calculations, I identified potential mutations that could lead to a non-haploinsufficiency cell phenotype in patients with HNRNPU-related neurodevelopmental disorder. FlpIn 293 T-REx cell lines were treated with siSAF-A and expressed the P487S, R714S, and Y805* SAF-A patient mutants. Additionally, HCT 116 cells were transfected with plasmids for patient mutant expression. These experiments revealed no changes in cell viability compared to wild-type SAF-A. The findings from the FlpIn 293 T-REx cell lines could be in line with patient mutant cells, where one copy of the SAF-A protein is mutated, due to partial depletion using RNAi. This suggests that patient cells may not exhibit altered viability with a heterozygous mutant. However, in HCT 116 cells, it appeared that there was insufficient time for the production of adequate levels of SAF-A protein, as even transfection with FLAG-SAF-A did not rescue the cell viability phenotype. Moreover, the mutants did not show differences in SAF-A oligomer formation compared to wild-type SAF-A.

Nevertheless, notable differences were observed when endogenous SAF-A was depleted using RNAi in HCT 116 cells, and the distance between loci in the gene-rich region 11p15.5 was measured. Y805* mutants failed to rescue the chromatin compaction phenotype, similar to WA mutants described in previous research (Nozawa *et al.*, 2017). On the other hand, R714S mutants exhibited partial rescue of the phenotype, while P487S mutants rescued the phenotype fully similarly to wild-type SAF-A. These results suggest that there may be other mutations in patients that interfere with SAF-A's role in large-scale chromatin organization. It would be valuable to repeat the experiments in Chapter 5 using neuronal cells, such as SH-SY5Y cells, to further investigate the molecular mechanisms underlying the HNRNPU-related neurodevelopmental disorder.

6.6. Concluding remarks

This thesis proposes a role for SAF-A in regulation of large-scale chromatin organisation as the daughter cells exit mitosis, thereby establishing a link between protein behaviour during mitosis and interphase. The findings presented in this thesis also suggest that certain SAF-A mutations found in patients might disrupt the protein's ability to facilitate the decompaction of gene-rich regions without altering its overall abundance.

References

- Ahmad, Kami, and Steven Henikoff. 2002. 'The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly'. *Molecular Cell* 9 (6): 1191–1200. [https://doi.org/10.1016/S1097-2765\(02\)00542-7](https://doi.org/10.1016/S1097-2765(02)00542-7).
- Alabert, Constance, and Anja Groth. 2012. 'Chromatin Replication and Epigenome Maintenance'. *Nature Reviews Molecular Cell Biology* 13 (3): 153–67. <https://doi.org/10.1038/nrm3288>.
- Albig, Werner, and D. Doenecke. 1997. 'The Human Histone Gene Cluster at the D6S105 Locus'. *Human Genetics* 101 (3): 284–94. <https://doi.org/10.1007/s004390050630>.
- Allan, J., P. G. Hartman, C. Crane-Robinson, and F. X. Aviles. 1980. 'The Structure of Histone H1 and Its Location in Chromatin'. *Nature* 288 (5792): 675–79. <https://doi.org/10.1038/288675a0>.
- Allan, James, Tanya Mitchell, Nerina Harborne, Lothar Bohm, and Colyn Crane-Robinson. 1986. 'Roles of H1 Domains in Determining Higher Order Chromatin Structure and H1 Location'. *Journal of Molecular Biology* 187 (4): 591–601. [https://doi.org/10.1016/0022-2836\(86\)90337-2](https://doi.org/10.1016/0022-2836(86)90337-2).
- Allen, Andrew S., Samuel F. Berkovic, Patrick Cossette, Norman Delanty, Dennis Dlugos, Evan E. Eichler, Michael P. Epstein, et al. 2013. 'De Novo Mutations in Epileptic Encephalopathies'. *Nature* 501 (7466): 217–21. <https://doi.org/10.1038/nature12439>.
- Allen, James. 2015. 'A Role for hnRNPUL2 in Linking the Management of DNA Topology with the DNA Damage Response'. Doctoral Thesis.
- Alpsoy, Aktan, and Emily C. Dykhuizen. 2018. 'Glioma Tumor Suppressor Candidate Region Gene 1 (GLTSCR1) and Its Paralog GLTSCR1-like Form SWI/SNF Chromatin Remodeling Subcomplexes'. *Journal of Biological Chemistry* 293 (11): 3892–3903. <https://doi.org/10.1074/jbc.RA117.001065>.
- Altenhoff, Adrian M, Clément-Marie Train, Kimberly J Gilbert, Ishita Mediratta, Tarcisio Mendes de Farias, David Moi, Yannis Nevers, et al. 2021. 'OMA Orthology in 2021: Website Overhaul, Conserved Isoforms, Ancestral Gene Order and More'. *Nucleic Acids Research* 49 (D1): D373–79. <https://doi.org/10.1093/nar/gkaa1007>.
- Alver, Burak H., Kimberly H. Kim, Ping Lu, Xiaofeng Wang, Haley E. Manchester, Weishan Wang, Jeffrey R. Haswell, Peter J. Park, and Charles W. M. Roberts. 2017. 'The SWI/SNF Chromatin Remodelling Complex Is Required for Maintenance of Lineage Specific Enhancers'. *Nature Communications* 8 (1): 14648. <https://doi.org/10.1038/ncomms14648>.
- Assimon, Victoria A., Yangzhong Tang, Jesse D. Vargas, Grace J. Lee, Zhi Yong Wu, Kenny Lou, Bing Yao, et al. 2019. 'CB-6644 Is a Selective Inhibitor of the RUVBL1/2

- Complex with Anticancer Activity'. *ACS Chemical Biology* 14 (2): 236–44.
<https://doi.org/10.1021/acscchembio.8b00904>.
- Attikum, Haico van, Olivier Fritsch, and Susan M Gasser. 2007. 'Distinct Roles for SWR1 and INO80 Chromatin Remodeling Complexes at Chromosomal Double-Strand Breaks'. *The EMBO Journal* 26 (18): 4113–25.
<https://doi.org/10.1038/sj.emboj.7601835>.
- Austin, Caroline A., Ka C. Lee, Rebecca L. Swan, Mushtaq M. Khazeem, Catriona M. Manville, Peter Cridland, Achim Treumann, Andrew Porter, Nick J. Morris, and Ian G. Cowell. 2018. 'TOP2B: The First Thirty Years'. *International Journal of Molecular Sciences* 19 (9): 2765. <https://doi.org/10.3390/ijms19092765>.
- Baboo, Sabyasachi, and Peter R Cook. 2014. "Dark Matter" Worlds of Unstable RNA and Protein'. *Nucleus* 5 (4): 281–86. <https://doi.org/10.4161/nucl.29577>.
- Backwell, Lisa, and Joseph A. Marsh. 2022. 'Diverse Molecular Mechanisms Underlying Pathogenic Protein Mutations: Beyond the Loss-of-Function Paradigm'. *Annual Review of Genomics and Human Genetics* 23 (1): 475–98.
<https://doi.org/10.1146/annurev-genom-111221-103208>.
- Badreldin, Amr A., Leila Bagheri, Bangke Zhang, A. Noelle Larson, and Andre J. van Wijnen. 2021. 'Relative mRNA and Protein Stability of Epigenetic Regulators in Musculoskeletal Cell Culture Models'. *Gene* 766 (January):145032.
<https://doi.org/10.1016/j.gene.2020.145032>.
- Bakkar, Nadine, Tina Kovalik, Ileana Lorenzini, Scott Spangler, Alix Lacoste, Kyle Sponaule, Philip Ferrante, Elenee Argentinis, Rita Sattler, and Robert Bowser. 2018. 'Artificial Intelligence in Neurodegenerative Disease Research: Use of IBM Watson to Identify Additional RNA-Binding Proteins Altered in Amyotrophic Lateral Sclerosis'. *Acta Neuropathologica* 135 (2): 227–47. <https://doi.org/10.1007/s00401-017-1785-8>.
- Ballif, Blake C., Jill A. Rosenfeld, Ryan Traylor, Aaron Theisen, Patricia I. Bader, Roger L. Ladda, Susan L. Sell, et al. 2012. 'High-Resolution Array CGH Defines Critical Regions and Candidate Genes for Microcephaly, Abnormalities of the Corpus Callosum, and Seizure Phenotypes in Patients with Microdeletions of 1q43q44'. *Human Genetics* 131 (1): 145–56. <https://doi.org/10.1007/s00439-011-1073-y>.
- Bannister, Andrew J., and Tony Kouzarides. 2011. 'Regulation of Chromatin by Histone Modifications'. *Cell Research* 21 (3): 381–95. <https://doi.org/10.1038/cr.2011.22>.
- Barnett, Stanley F., Daniel L. Friedman, and Wallace M. LeStourgeon. 1989. 'The C Proteins of HeLa 40S Nuclear Ribonucleoprotein Particles Exist as Anisotropic Tetramers of (C1)₃ C2'. *Molecular and Cellular Biology* 9 (2): 492–98.
<https://doi.org/10.1128/mcb.9.2.492-498.1989>.

- Barnum, Kevin J., and Matthew J. O'Connell. 2014. 'Cell Cycle Regulation by Checkpoints'. In *Cell Cycle Control: Mechanisms and Protocols*, edited by Eishi Noguchi and Mariana C. Gadaleta, 29–40. Methods in Molecular Biology. New York, NY: Springer. https://doi.org/10.1007/978-1-4939-0888-2_2.
- Barre, Anne-Elisabeth de la, Véronique Gerson, Stéphanie Gout, Martina Creaven, C. Dave Allis, and Stefan Dimitrov. 2000. 'Core Histone N-termini Play an Essential Role in Mitotic Chromosome Condensation'. *The EMBO Journal* 19 (3): 379–91. <https://doi.org/10.1093/emboj/19.3.379>.
- Bassing, Craig H., Katrin F. Chua, JoAnn Sekiguchi, Heikyung Suh, Scott R. Whitlow, James C. Fleming, Brianna C. Monroe, et al. 2002. 'Increased Ionizing Radiation Sensitivity and Genomic Instability in the Absence of Histone H2AX'. *Proceedings of the National Academy of Sciences* 99 (12): 8173–78. <https://doi.org/10.1073/pnas.122228699>.
- Batty, Paul, and Daniel W. Gerlich. 2019. 'Mitotic Chromosome Mechanics: How Cells Segregate Their Genome'. *Trends in Cell Biology* 29 (9): 717–26. <https://doi.org/10.1016/j.tcb.2019.05.007>.
- Becker, Klaus A., Prachi N. Ghule, Jaclyn A. Therrien, Jane B. Lian, Janet L. Stein, Andre J. van Wijnen, and Gary S. Stein. 2006. 'Self-Renewal of Human Embryonic Stem Cells Is Supported by a Shortened G1 Cell Cycle Phase'. *Journal of Cellular Physiology* 209 (3): 883–93. <https://doi.org/10.1002/jcp.20776>.
- Belmont, A S, and K Bruce. 1994. 'Visualization of G1 Chromosomes: A Folded, Twisted, Supercoiled Chromonema Model of Interphase Chromatid Structure.' *Journal of Cell Biology* 127 (2): 287–302. <https://doi.org/10.1083/jcb.127.2.287>.
- Belmont, Andrew S., Michael B. Braunfeld, John W. Sedat, and David A. Agard. 1989. 'Large-Scale Chromatin Structural Domains within Mitotic and Interphase Chromosomes in Vivo and in Vitro'. *Chromosoma* 98 (2): 129–43. <https://doi.org/10.1007/BF00291049>.
- Beltran, Manuel, Christopher M. Yates, Lenka Skalska, Marcus Dawson, Filipa P. Reis, Keijo Viiri, Cynthia L. Fisher, et al. 2016. 'The Interaction of PRC2 with RNA or Chromatin Is Mutually Antagonistic'. *Genome Research* 26 (7): 896–907. <https://doi.org/10.1101/gr.197632.115>.
- Bensaude, Olivier. 2011. 'Inhibiting Eukaryotic Transcription'. *Transcription* 2 (3): 103–8. <https://doi.org/10.4161/trns.2.3.16172>.
- Berezney, Ronald, and Donald S. Coffey. 1974. 'Identification of a Nuclear Protein Matrix'. *Biochemical and Biophysical Research Communications* 60 (4): 1410–17. [https://doi.org/10.1016/0006-291X\(74\)90355-6](https://doi.org/10.1016/0006-291X(74)90355-6).

- Bertoli, Cosetta, Jan M. Skotheim, and Robertus A. M. de Bruin. 2013. 'Control of Cell Cycle Transcription during G1 and S Phases'. *Nature Reviews Molecular Cell Biology* 14 (8): 518–28. <https://doi.org/10.1038/nrm3629>.
- Bickmore, Wendy A., and Bas van Steensel. 2013. 'Genome Architecture: Domain Organization of Interphase Chromosomes'. *Cell* 152 (6): 1270–84. <https://doi.org/10.1016/j.cell.2013.02.001>.
- Biterge, Burcu, and Robert Schneider. 2014. 'Histone Variants: Key Players of Chromatin'. *Cell and Tissue Research* 356 (3): 457–66. <https://doi.org/10.1007/s00441-014-1862-4>.
- Bizard, Anna H., and Ian D. Hickson. 2020. 'The Many Lives of Type IA Topoisomerases'. *Journal of Biological Chemistry* 295 (20): 7138–53. <https://doi.org/10.1074/jbc.REV120.008286>.
- Blower, Michael D., Wei Wang, and Judith A. Sharp. 2023. 'Differential Nuclear Import Regulates Nuclear RNA Inheritance Following Mitosis'. *Molecular Biology of the Cell* 34 (4): ar32. <https://doi.org/10.1091/mbc.E23-01-0004>.
- Boeger, Hinrich, Joachim Griesenbeck, J. Seth Strattan, and Roger D Kornberg. 2004. 'Removal of Promoter Nucleosomes by Disassembly Rather Than Sliding In Vivo'. *Molecular Cell* 14 (5): 667–73. <https://doi.org/10.1016/j.molcel.2004.05.013>.
- Boileau, Etienne, Janine Altmüller, Isabel S Naarmann-de Vries, and Christoph Dieterich. 2021. 'A Comparison of Metabolic Labeling and Statistical Methods to Infer Genome-Wide Dynamics of RNA Turnover'. *Briefings in Bioinformatics* 22 (6): bbab219. <https://doi.org/10.1093/bib/bbab219>.
- Booth, Daniel G., Masatoshi Takagi, Luis Sanchez-Pulido, Elizabeth Petfalski, Giulia Vargiu, Kumiko Samejima, Naoko Imamoto, et al. 2014. 'Ki-67 Is a PP1-Interacting Protein That Organises the Mitotic Chromosome Periphery'. *eLife* 3 (May): e01641. <https://doi.org/10.7554/eLife.01641>.
- Borowiec, Anne-Sophie, Gabriel Bidaux, Rachida Tacine, Pauline Dubar, Natascha Pigat, Philippe Delcourt, Olivier Mignen, and Thierry Capiod. 2014. 'Are Orai1 and Orai3 Channels More Important than Calcium Influx for Cell Proliferation?' *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1843 (2): 464–72. <https://doi.org/10.1016/j.bbamcr.2013.11.023>.
- Bostock, C. J., D. M. Prescott, and J. B. Kirkpatrick. 1971. 'An Evaluation of the Double Thymidine Block for Synchronizing Mammalian Cells at the G1-S Border'. *Experimental Cell Research* 68 (1): 163–68. [https://doi.org/10.1016/0014-4827\(71\)90599-4](https://doi.org/10.1016/0014-4827(71)90599-4).

- Boteva, Lora. 2017. 'Investigating Transcription, Replication and Chromatin Structure in Determining Common Fragile Site Instability', July.
<https://era.ed.ac.uk/handle/1842/28803>.
- Boteva, Lora, Ryu-Suke Nozawa, Catherine Naughton, Kumiko Samejima, William C. Earnshaw, and Nick Gilbert. 2020. 'Common Fragile Sites Are Characterized by Faulty Condensin Loading after Replication Stress'. *Cell Reports* 32 (12): 108177.
<https://doi.org/10.1016/j.celrep.2020.108177>.
- Boyle, Shelagh, Susan Gilchrist, Joanna M. Bridger, Nicola L. Mahy, Juliet A. Ellis, and Wendy A. Bickmore. 2001. 'The Spatial Organization of Human Chromosomes within the Nuclei of Normal and Emerin-Mutant Cells'. *Human Molecular Genetics* 10 (3): 211–20. <https://doi.org/10.1093/hmg/10.3.211>.
- Boyle, Shelagh, Matthew J. Rodesch, Heather A. Halvensleben, Jeffrey A. Jeddloh, and Wendy A. Bickmore. 2011. 'Fluorescence in Situ Hybridization with High-Complexity Repeat-Free Oligonucleotide Probes Generated by Massively Parallel Synthesis'. *Chromosome Research* 19 (7): 901–9. <https://doi.org/10.1007/s10577-011-9245-0>.
- Brabander, M. De, G. Geuens, R. Nuydens, R. Willebrords, F. Aerts, J. De Mey, and J. R. McIntosh. 1986. 'Microtubule Dynamics during the Cell Cycle: The Effects of Taxol and Nocodazole on the Microtubule System of Pt K2 Cells at Different Stages of the Mitotic Cycle'. In *International Review of Cytology*, edited by G. H. Bourne, J. F. Danielli, and K. W. Jeon, 101:215–74. International Review of Cytology. Academic Press. [https://doi.org/10.1016/S0074-7696\(08\)60250-8](https://doi.org/10.1016/S0074-7696(08)60250-8).
- Bramswig, Nuria C., Hermann Josef Lüdecke, Fadi F. Hamdan, Janine Altmüller, Filippo Beleggia, Nursel H. Elcioglu, Catharine Freyer, et al. 2017. 'Heterozygous HNRNPU Variants Cause Early Onset Epilepsy and Severe Intellectual Disability'. *Human Genetics* 136 (7): 821–34. <https://doi.org/10.1007/s00439-017-1795-6>.
- Branco, Miguel R, and Ana Pombo. 2006. 'Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations'. *PLoS Biology* 4 (5): e138.
<https://doi.org/10.1371/journal.pbio.0040138>.
- Brangwynne, Clifford P., Christian R. Eckmann, David S. Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A. Hyman. 2009. 'Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation'. *Science* 324 (5935): 1729–32.
<https://doi.org/10.1126/science.1172046>.
- Brangwynne, Clifford P., Timothy J. Mitchison, and Anthony A. Hyman. 2011. 'Active Liquid-like Behavior of Nucleoli Determines Their Size and Shape in *Xenopus Laevis* Oocytes'. *Proceedings of the National Academy of Sciences* 108 (11): 4334–39.
<https://doi.org/10.1073/pnas.1017150108>.

- Braun, Simon M. G., Ralitsa Petrova, Jiong Tang, Andrey Krokhotin, Erik L. Miller, Yitai Tang, Georgia Panagiotakos, and Gerald R. Crabtree. 2021. 'BAF Subunit Switching Regulates Chromatin Accessibility to Control Cell Cycle Exit in the Developing Mammalian Cortex'. *Genes & Development*, February. <https://doi.org/10.1101/gad.342345.120>.
- Brownlee, Peter M., Cornelia Meisenberg, and Jessica A. Downs. 2015. 'The SWI/SNF Chromatin Remodelling Complex: Its Role in Maintaining Genome Stability and Preventing Tumourigenesis'. *DNA Repair*, Cutting-edge Perspectives in Genomic Maintenance II, 32 (August):127–33. <https://doi.org/10.1016/j.dnarep.2015.04.023>.
- Bryant, Laura, Dong Li, Samuel G. Cox, Dylan Marchione, Evan F. Joiner, Khadija Wilson, Kevin Janssen, et al. 2020. 'Histone H3.3 beyond Cancer: Germline Mutations in Histone 3 Family 3A and 3B Cause a Previously Unidentified Neurodegenerative Disorder in 46 Patients'. *Science Advances* 6 (49): eabc9207. <https://doi.org/10.1126/sciadv.abc9207>.
- Burger, Kaspar, Bastian Mühl, Markus Kellner, Michaela Rohrmoser, Anita Gruber-Eber, Lukas Windhager, Caroline C Friedel, Lars Dölken, and Dirk Eick. 2013. '4-Thiouridine Inhibits rRNA Synthesis and Causes a Nucleolar Stress Response'. *RNA Biology* 10 (10): 1623–30. <https://doi.org/10.4161/rna.26214>.
- Caliebe, Almuth, Hester Y. Kroes, Jasper J. van der Smagt, José I. Martin-Subero, Holger Tönnies, Ruben van 't Slot, Rutger A. J. Nievelstein, et al. 2010. 'Four Patients with Speech Delay, Seizures and Variable Corpus Callosum Thickness Sharing a 0.440 Mb Deletion in Region 1q44 Containing the HNRPU Gene'. *European Journal of Medical Genetics* 53 (4): 179–85. <https://doi.org/10.1016/j.ejmg.2010.04.001>.
- Carvill, Gemma L., Sinéad B. Heavin, Simone C. Yendle, Jacinta M. McMahon, Brian J. O'Roak, Joseph Cook, Adiba Khan, et al. 2013. 'Targeted Resequencing in Epileptic Encephalopathies Identifies de Novo Mutations in CHD2 and SYNGAP1'. *Nature Genetics* 45 (7): 825–30. <https://doi.org/10.1038/ng.2646>.
- Caudron-Herger, Maiwen, Peter R. Cook, Karsten Rippe, and Argyris Papantonis. 2015. 'Dissecting the Nascent Human Transcriptome by Analysing the RNA Content of Transcription Factories'. *Nucleic Acids Research* 43 (14): e95. <https://doi.org/10.1093/nar/gkv390>.
- Cech, Thomas R., and Joan A. Steitz. 2014. 'The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones'. *Cell* 157 (1): 77–94. <https://doi.org/10.1016/j.cell.2014.03.008>.
- Chakravarthy, Srinivas, and Karolin Luger. 2006. 'The Histone Variant Macro-H2A Preferentially Forms "Hybrid Nucleosomes" *♦'. *Journal of Biological Chemistry* 281 (35): 25522–31. <https://doi.org/10.1074/jbc.M602258200>.

- Chao, Hui Xiao, Randy I Fakhreddin, Hristo K Shimerov, Katarzyna M Kedziora, Rashmi J Kumar, Joanna Perez, Juanita C Limas, et al. 2019. 'Evidence That the Human Cell Cycle Is a Series of Uncoupled, Memoryless Phases'. *Molecular Systems Biology* 15 (3): e8604. <https://doi.org/10.15252/msb.20188604>.
- Chao, Hui Xiao, Cere E. Poovey, Ashley A. Privette, Gavin D. Grant, Hui Yan Chao, Jeanette G. Cook, and Jeremy E. Purvis. 2017. 'Orchestration of DNA Damage Checkpoint Dynamics across the Human Cell Cycle'. *Cell Systems* 5 (5): 445-459.e5. <https://doi.org/10.1016/j.cels.2017.09.015>.
- Chao, William C. H., Yasuto Murayama, Sofía Muñoz, Andrew W. Jones, Benjamin O. Wade, Andrew G. Purkiss, Xiao-Wen Hu, et al. 2017. 'Structure of the Cohesin Loader Scc2'. *Nature Communications* 8 (1): 13952. <https://doi.org/10.1038/ncomms13952>.
- Chen, Ping, Jicheng Zhao, and Guohong Li. 2013. 'Histone Variants in Development and Diseases'. *Journal of Genetics and Genomics, Epigenetics: Development, Diseases and Memories*, 40 (7): 355–65. <https://doi.org/10.1016/j.jgg.2013.05.001>.
- Chen, Xing. 2015. 'KATZLDA: KATZ Measure for the lncRNA-Disease Association Prediction'. *Scientific Reports* 5 (1): 16840. <https://doi.org/10.1038/srep16840>.
- Chen, Yu, Haiping Zhang, Zhu Xu, Huanyin Tang, Anke Geng, Bailian Cai, Tao Su, et al. 2019. 'A PARP1–BRG1–SIRT1 Axis Promotes HR Repair by Reducing Nucleosome Density at DNA Damage Sites'. *Nucleic Acids Research* 47 (16): 8563–80. <https://doi.org/10.1093/nar/gkz592>.
- Chérot, E., B. Keren, C. Dubourg, W. Carré, M. Fradin, A. Lavillaureix, A. Afenjar, et al. 2018. 'Using Medical Exome Sequencing to Identify the Causes of Neurodevelopmental Disorders: Experience of 2 Clinical Units and 216 Patients'. *Clinical Genetics* 93 (3): 567–76. <https://doi.org/10.1111/cge.13102>.
- Choi, Mieyoung. 2017. 'Analysis of the Role of RGG box of human hnRNP A1 protein'. *Journal of the Korea Academia-Industrial cooperation Society* 18 (12): 575–80. <https://doi.org/10.5762/KAIS.2017.18.12.575>.
- Chong, P. Andrew, Robert M. Vernon, and Julie D. Forman-Kay. 2018. 'RGG/RG Motif Regions in RNA Binding and Phase Separation'. *Journal of Molecular Biology, Phase Separation in Biology and Disease*, 430 (23): 4650–65. <https://doi.org/10.1016/j.jmb.2018.06.014>.
- Choo, Jung Ha, Jeong Do Kim, Jae Hoon Chung, Lisa Stubbs, and Joomyeong Kim. 2006. 'Allele-Specific Deposition of macroH2A1 in Imprinting Control Regions'. *Human Molecular Genetics* 15 (5): 717–24. <https://doi.org/10.1093/hmg/ddi485>.
- Chu, Ci, Qiangfeng Cliff Zhang, Simão Teixeira da Rocha, Ryan A. Flynn, Maheetha Bharadwaj, J. Mauro Calabrese, Terry Magnuson, Edith Heard, and Howard Y.

- Chang. 2015. 'Systematic Discovery of Xist RNA Binding Proteins'. *Cell* 161 (2): 404–16. <https://doi.org/10.1016/j.cell.2015.03.025>.
- Cieniková, Zuzana, Sandrine Jayne, Fred Franz Damberger, Frédéric Hai-Trieu Allain, and Christophe Maris. 2015. 'Evidence for Cooperative Tandem Binding of hnRNP C RRM s in mRNA Processing'. *RNA* 21 (11): 1931–42. <https://doi.org/10.1261/rna.052373.115>.
- Clapier, Cedric R., Janet Iwasa, Bradley R. Cairns, and Craig L. Peterson. 2017. 'Mechanisms of Action and Regulation of ATP-Dependent Chromatin-Remodelling Complexes'. *Nature Reviews Molecular Cell Biology* 18 (7): 407–22. <https://doi.org/10.1038/nrm.2017.26>.
- Clark, David J., and Takeshi Kimura. 1990. 'Electrostatic Mechanism of Chromatin Folding'. *Journal of Molecular Biology* 211 (4): 883–96. [https://doi.org/10.1016/0022-2836\(90\)90081-V](https://doi.org/10.1016/0022-2836(90)90081-V).
- Clemson, C M, J A McNeil, H F Willard, and J B Lawrence. 1996. 'XIST RNA Paints the Inactive X Chromosome at Interphase: Evidence for a Novel RNA Involved in Nuclear/Chromosome Structure.' *Journal of Cell Biology* 132 (3): 259–75. <https://doi.org/10.1083/jcb.132.3.259>.
- Clemson, Christine M., John N. Hutchinson, Sergio A. Sara, Alexander W. Ensminger, Archa H. Fox, Andrew Chess, and Jeanne B. Lawrence. 2009. 'An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles'. *Molecular Cell* 33 (6): 717–26. <https://doi.org/10.1016/j.molcel.2009.01.026>.
- Cook, Kate B., Hilal Kazan, Khalid Zuberi, Quaid Morris, and Timothy R. Hughes. 2011. 'RBPDB: A Database of RNA-Binding Specificities'. *Nucleic Acids Research* 39 (suppl_1): D301–8. <https://doi.org/10.1093/nar/gkq1069>.
- Cook, Peter J., Bong Gun Ju, Francesca Telese, Xiangting Wang, Christopher K. Glass, and Michael G. Rosenfeld. 2009. 'Tyrosine Dephosphorylation of H2AX Modulates Apoptosis and Survival Decisions'. *Nature* 458 (7238): 591–96. <https://doi.org/10.1038/nature07849>.
- Cooper, Stephen. 1997. 'G1 and S Phase Gene Expression Cannot Be Analyzed in Mammalian Cells Synchronized by Inhibition'. *Microbial & Comparative Genomics* 2 (4): 269–73. <https://doi.org/10.1089/omi.1.1997.2.269>.
- Cooper, Thomas A., Lili Wan, and Gideon Dreyfuss. 2009. 'RNA and Disease'. *Cell* 136 (4): 777–93. <https://doi.org/10.1016/j.cell.2009.02.011>.
- Corley, Susan M., and Jill E. Gready. 2008. 'Identification of the RGG Box Motif in Shadoo: RNA-Binding and Signaling Roles?' *Bioinformatics and Biology Insights* 2 (January): BBI.S1075. <https://doi.org/10.4137/BBI.S1075>.

- Costanzi, Carl, and John R. Pehrson. 1998. 'Histone macroH2A1 Is Concentrated in the Inactive X Chromosome of Female Mammals'. *Nature* 393 (6685): 599–601. <https://doi.org/10.1038/31275>.
- Creamer, Kevin Michael, Heather Jill Kolpa, and Jeanne Bentley Lawrence. 2021. 'Nascent RNA Scaffolds Contribute to Chromosome Territory Architecture and Counter Chromatin Compaction'. *Molecular Cell* 81 (17): 3509–3525.e5. <https://doi.org/10.1016/j.molcel.2021.07.004>.
- Cremer, T., and C. Cremer. 2001. 'Chromosome Territories, Nuclear Architecture and Gene Regulation in Mammalian Cells'. *Nature Reviews Genetics* 2 (4): 292–301. <https://doi.org/10.1038/35066075>.
- Cremer, Thomas, and Marion Cremer. 2010. 'Chromosome Territories'. *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889. <https://doi.org/10.1101/cshperspect.a003889>.
- Cuddapah, Suresh, Raja Jothi, Dustin E. Schones, Tae-Young Roh, Kairong Cui, and Keji Zhao. 2009. 'Global Analysis of the Insulator Binding Protein CTCF in Chromatin Barrier Regions Reveals Demarcation of Active and Repressive Domains'. *Genome Research* 19 (1): 24–32. <https://doi.org/10.1101/gr.082800.108>.
- Cusick, M. E., K. S. Lee, M. L. DePamphilis, and P. M. Wassarman. 1983. 'Structure of Chromatin at Deoxyribonucleic Acid Replication Forks: Nuclease Hypersensitivity Results from Both Prenucleosomal Deoxyribonucleic Acid and an Immature Chromatin Structure'. *Biochemistry* 22 (16): 3873–84. <https://doi.org/10.1021/bi00285a024>.
- Cuylen, Sara, Claudia Blaukopf, Antonio Z. Politi, Thomas Müller-Reichert, Beate Neumann, Ina Poser, Jan Ellenberg, Anthony A. Hyman, and Daniel W. Gerlich. 2016. 'Ki-67 Acts as a Biological Surfactant to Disperse Mitotic Chromosomes'. *Nature* 535 (7611): 308–12. <https://doi.org/10.1038/nature18610>.
- Dang, Thao, Soo Hwa Jang, Sung Hoon Back, Jeong Woo Park, and In Seob Han. 2018. 'DRG2 Deficiency Causes Impaired Microtubule Dynamics in Hela Cells'. *Molecules and Cells* 41 (12): 1045–51. <https://doi.org/10.14348/molcells.2018.0129>.
- Dawson, K. B., H. Madoc-Jones, and E. O. Field. 1965. 'Variations in the Generation Times of a Strain of Rat Sarcoma Cells in Culture'. *Experimental Cell Research* 38 (1): 75–84. [https://doi.org/10.1016/0014-4827\(65\)90429-5](https://doi.org/10.1016/0014-4827(65)90429-5).
- Deák, Gauri, and Atlanta G. Cook. 2022. 'Missense Variants Reveal Functional Insights Into the Human ARID Family of Gene Regulators'. *Journal of Molecular Biology* 434 (9): 167529. <https://doi.org/10.1016/j.jmb.2022.167529>.
- Dechassa, Mekonnen Lemma, Abdellah Sabri, Santhi Pondugula, Stefan R. Kassabov, Nilanjana Chatterjee, Michael P. Kladde, and Blaine Bartholomew. 2010. 'SWI/SNF

Has Intrinsic Nucleosome Disassembly Activity That Is Dependent on Adjacent Nucleosomes'. *Molecular Cell* 38 (4): 590–602.
<https://doi.org/10.1016/j.molcel.2010.02.040>.

Dekker, Job, Karsten Rippe, Martijn Dekker, and Nancy Kleckner. 2002. 'Capturing Chromosome Conformation'. *Science* 295 (5558): 1306–11.
<https://doi.org/10.1126/science.1067799>.

Demos, Michelle, Ilaria Guella, Conrado DeGuzman, Marna B. McKenzie, Sarah E. Buerki, Daniel M. Evans, Eric B. Toyota, et al. 2019. 'Diagnostic Yield and Treatment Impact of Targeted Exome Sequencing in Early-Onset Epilepsy'. *Frontiers in Neurology* 10. <https://www.frontiersin.org/articles/10.3389/fneur.2019.00434>.

Depienne, Christel, Caroline Nava, Boris Keren, Solveig Heide, Agnès Rastetter, Sandrine Passemard, Sandra Chantot-Bastaraud, et al. 2017. 'Genetic and Phenotypic Dissection of 1q43q44 Microdeletion Syndrome and Neurodevelopmental Phenotypes Associated with Mutations in ZBTB18 and HNRNPU'. *Human Genetics* 136 (4): 463–79. <https://doi.org/10.1007/s00439-017-1772-0>.

Deshpande, Atul M., and Carol S. Newlon. 1996. 'DNA Replication Fork Pause Sites Dependent on Transcription'. *Science* 272 (5264): 1030–33.
<https://doi.org/10.1126/science.272.5264.1030>.

Dileep, Vishnu, Ferhat Ay, Jiao Sima, Daniel L. Vera, William S. Noble, and David M. Gilbert. 2015. 'Topologically Associating Domains and Their Long-Range Contacts Are Established during Early G1 Coincident with the Establishment of the Replication-Timing Program'. *Genome Research* 25 (8): 1104–13.
<https://doi.org/10.1101/gr.183699.114>.

Ding, Tianyi, and He Zhang. 2022. 'Novel Biological Insights Revealed from the Investigation of Multiscale Genome Architecture'. *Computational and Structural Biotechnology Journal* 21 (December):312–25.
<https://doi.org/10.1016/j.csbj.2022.12.009>.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. 'Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions'. *Nature* 485 (7398): 376–80.
<https://doi.org/10.1038/nature11082>.

Dong, Shunli, Jinhua Han, Hongxia Chen, Ting Liu, Michael S. Y. Huen, Yeran Yang, Caixia Guo, and Jun Huang. 2014. 'The Human SRCAP Chromatin Remodeling Complex Promotes DNA-End Resection'. *Current Biology* 24 (18): 2097–2110.
<https://doi.org/10.1016/j.cub.2014.07.081>.

Dorigo, Benedetta, Thomas Schalch, Kerstin Bystricky, and Timothy J. Richmond. 2003. 'Chromatin Fiber Folding: Requirement for the Histone H4 N-Terminal Tail'. *Journal of Molecular Biology* 327 (1): 85–96. [https://doi.org/10.1016/S0022-2836\(03\)00025-1](https://doi.org/10.1016/S0022-2836(03)00025-1).

- Douglas, Pauline, Ruiqiong Ye, Nicholas Morrice, Sébastien Britton, Laura Trinkle-Mulcahy, and Susan P. Lees-Miller. 2015. 'Phosphorylation of SAF-A/hnRNP-U Serine 59 by Polo-Like Kinase 1 Is Required for Mitosis'. *Molecular and Cellular Biology* 35 (15): 2699–2713. <https://doi.org/10.1128/mcb.01312-14>.
- Dreyfuss, G, Y D Choi, and S A Adam. 1984. 'Characterization of Heterogeneous Nuclear RNA-Protein Complexes in Vivo with Monoclonal Antibodies'. *Molecular and Cellular Biology* 4 (6): 1104–14. <https://doi.org/10.1128/mcb.4.6.1104-1114.1984>.
- Dreyfuss, Gideon, Michael J. Matunis, Serafin Pinol-Roma, and Christopher G. Burd. 1993. 'hnRNP PROTEINS AND THE BIOGENESIS OF mRNA'. *Annual Review of Biochemistry* 62 (1): 289–321. <https://doi.org/10.1146/annurev.bi.62.070193.001445>.
- Dube, Umber, Jorge L. Del-Aguila, Zeran Li, John P. Budde, Shan Jiang, Simon Hsu, Laura Ibanez, et al. 2019. 'An Atlas of Cortical Circular RNA Expression in Alzheimer Disease Brains Demonstrates Clinical and Pathological Associations'. *Nature Neuroscience* 22 (11): 1903–12. <https://doi.org/10.1038/s41593-019-0501-5>.
- Dubochet, Jacques, Marc Adrian, Jiin-Ju Chang, Jean-Claude Homo, Jean Lepault, Alasdair W. McDowall, and Patrick Schultz. 1988. 'Cryo-Electron Microscopy of Vitrified Specimens'. *Quarterly Reviews of Biophysics* 21 (2): 129–228. <https://doi.org/10.1017/S0033583500004297>.
- Duffney, Lara J., Purnima Valdez, Martine W. Tremblay, Xinyu Cao, Sarah Montgomery, Allyn McConkie-Rosell, and Yong-hui Jiang. 2018. 'Epigenetics and Autism Spectrum Disorder: A Report of an Autism Case with Mutation in H1 Linker Histone HIST1H1E and Literature Review'. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 177 (4): 426–33. <https://doi.org/10.1002/ajmg.b.32631>.
- Dugger, Sarah A., Ryan S. Dhindsa, Gabriela De Almeida Sampaio, Andrew K. Ressler, Elizabeth E. Rafikian, Sabrina Petri, Verity A. Letts, et al. 2023. 'Neurodevelopmental Deficits and Cell-Type-Specific Transcriptomic Perturbations in a Mouse Model of HNRNPU Haploinsufficiency'. *PLOS Genetics* 19 (10): e1010952. <https://doi.org/10.1371/journal.pgen.1010952>.
- Durkin, Anna, Shadi Albaba, Andrew E. Fry, Jenny E Morton, Andrew Douglas, Ana Beleza, Denise Williams, et al. 2020. 'Clinical Findings of 21 Previously Unreported Probands with HNRNPU-Related Syndrome and Comprehensive Literature Review'. *American Journal of Medical Genetics, Part A* 182 (7): 1637–54. <https://doi.org/10.1002/ajmg.a.61599>.
- Earnshaw, W C, B Halligan, C A Cooke, M M Heck, and L F Liu. 1985. 'Topoisomerase II Is a Structural Component of Mitotic Chromosome Scaffolds.' *Journal of Cell Biology* 100 (5): 1706–15. <https://doi.org/10.1083/jcb.100.5.1706>.

- Edwards, John R., Olya Yarychivska, Mathieu Boulard, and Timothy H. Bestor. 2017. 'DNA Methylation and DNA Methyltransferases'. *Epigenetics & Chromatin* 10 (1): 23. <https://doi.org/10.1186/s13072-017-0130-8>.
- Elbatsh, Ahmed M. O., Eugene Kim, Jorine M. Eeftens, Jonne A. Raaijmakers, Robin H. van der Weide, Alberto García-Nieto, Sol Bravo, et al. 2019. 'Distinct Roles for Condensin's Two ATPase Sites in Chromosome Condensation'. *Molecular Cell* 76 (5): 724-737.e5. <https://doi.org/10.1016/j.molcel.2019.09.020>.
- Eltsov, Mikhail, Kirsty M. MacLellan, Kazuhiro Maeshima, Achilleas S. Frangakis, and Jacques Dubochet. 2008. 'Analysis of Cryo-Electron Microscopy Images Does Not Support the Existence of 30-Nm Chromatin Fibers in Mitotic Chromosomes in Situ'. *Proceedings of the National Academy of Sciences* 105 (50): 19732–37. <https://doi.org/10.1073/pnas.0810057105>.
- Engreitz, Jesse M., Amy Pandya-Jones, Patrick McDonel, Alexander Shishkin, Klara Sirokman, Christine Surka, Sabah Kadri, et al. 2013. 'The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome'. *Science* 341 (6147): 1237973. <https://doi.org/10.1126/science.1237973>.
- Equilibrina, Ilma, Elena Krayukhina, Kota Inoue, Yuya Ishikawa, Akihiro Kawamoto, Takayuki Kato, Isao Suetake, et al. 2015. 'The Role of Phosphorylation of Histone H3 at Serine 10 in Chromatin Condensation in Vitro'. *Chromosome Science* 18 (1–2): 9–14. <https://doi.org/10.11352/scr.18.9>.
- Ester, Martin, Hans-Peter Kriegel, Jorg Sander, and Xiaowei Xu. 1996. 'A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.' *Kdd* 96 (34): 226–31.
- FACKELMAYER, Frank O., Kirsten DAHM, Andrea RENZ, Uwe RAMSPERGER, and Arndt RICHTER. 1994. 'Nucleic-acid-binding Properties of hnRNP-U/SAF-A, a Nuclear-matrix Protein Which Binds DNA and RNA in Vivo and in Vitro'. *European Journal of Biochemistry* 221 (2): 749–57. <https://doi.org/10.1111/j.1432-1033.1994.tb18788.x>.
- Fan, Hui, Pin Lv, Xiangru Huo, Jicheng Wu, Qianfeng Wang, Lu Cheng, Yun Liu, et al. 2018. 'The Nuclear Matrix Protein HNRNPU Maintains 3D Genome Architecture Globally in Mouse Hepatocytes'. *Genome Research* 28 (2): 192–202. <https://doi.org/10.1101/gr.224576.117>.
- Farwell, Kelly D., Layla Shahmirzadi, Dima El-Khechen, Zöe Powis, Elizabeth C. Chao, Brigitte Tippin Davis, Ruth M. Baxter, et al. 2015. 'Enhanced Utility of Family-Centered Diagnostic Exome Sequencing with Inheritance Model-Based Analysis: Results from 500 Unselected Families with Undiagnosed Genetic Conditions'. *Genetics in Medicine* 17 (7): 578–86. <https://doi.org/10.1038/gim.2014.154>.

- Feliciano, Pamela, Xueya Zhou, Irina Astrovskaya, Tychele N. Turner, Tianyun Wang, Leo Brueggeman, Rebecca Barnard, et al. 2019. 'Exome Sequencing of 457 Autism Families Recruited Online Provides Evidence for Autism Risk Genes'. *Npj Genomic Medicine* 4 (1): 1–14. <https://doi.org/10.1038/s41525-019-0093-8>.
- Felsenfeld, Gary. 1978. 'Chromatin'. *Nature* 271 (5641): 115–22. <https://doi.org/10.1038/271115a0>.
- Fernández-Marmiesse, Ana, Iria Roca, Felicitas Díaz-Flores, Verónica Cantarín, M^a Socorro Pérez-Poyato, Ana Fontalba, Francisco Laranjeira, et al. 2019. 'Rare Variants in 48 Genes Account for 42% of Cases of Epilepsy With or Without Neurodevelopmental Delay in 246 Pediatric Patients'. *Frontiers in Neuroscience* 13. <https://www.frontiersin.org/articles/10.3389/fnins.2019.01135>.
- Festuccia, Nicola, Inma Gonzalez, Nick Owens, and Pablo Navarro. 2017. 'Mitotic Bookmarking in Development and Stem Cells'. *Development (Cambridge)* 144 (20): 3633–45. <https://doi.org/10.1242/dev.146522>.
- Finch, J T, and A Klug. 1976. 'Solenoidal Model for Superstructure in Chromatin.' *Proceedings of the National Academy of Sciences* 73 (6): 1897–1901. <https://doi.org/10.1073/pnas.73.6.1897>.
- Firth, Helen V., Shola M. Richards, A. Paul Bevan, Stephen Clayton, Manuel Corpas, Diana Rajan, Steven Van Vooren, Yves Moreau, Roger M. Pettett, and Nigel P. Carter. 2009. 'DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources'. *The American Journal of Human Genetics* 84 (4): 524–33. <https://doi.org/10.1016/j.ajhg.2009.03.010>.
- Fischer, Martin, Patrick Grossmann, Megha Padi, and James A. DeCaprio. 2016. 'Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F Target Gene Analyses Identifies Cell Cycle Gene Regulatory Networks'. *Nucleic Acids Research* 44 (13): 6070–86. <https://doi.org/10.1093/nar/gkw523>.
- Fitzgerald, T. W., S. S. Gerety, W. D. Jones, M. van Kogelenberg, D. A. King, J. McRae, K. I. Morley, et al. 2015. 'Large-Scale Discovery of Novel Genetic Causes of Developmental Disorders'. *Nature* 519 (7542): 223–28. <https://doi.org/10.1038/nature14135>.
- Fox, Michael H., Robert A. Read, and Joel S. Bedford. 1987. 'Comparison of Synchronized Chinese Hamster Ovary Cells Obtained by Mitotic Shake-off, Hydroxyurea, Aphidicolin, or Methotrexate'. *Cytometry* 8 (3): 315–20. <https://doi.org/10.1002/cyto.990080312>.
- Fuchs, Gilad, Yoav Voichek, Sima Benjamin, Shlomit Gilad, Ido Amit, and Moshe Oren. 2014. '4sUDRB-Seq: Measuring Genomewide Transcriptional Elongation Rates and Initiation Frequencies within Cells'. *Genome Biology* 15 (5): R69. <https://doi.org/10.1186/gb-2014-15-5-r69>.

- Gallant, P., and E. A. Nigg. 1992. 'Cyclin B2 Undergoes Cell Cycle-Dependent Nuclear Translocation and, When Expressed as a Non-Destructible Mutant, Causes Mitotic Arrest in HeLa Cells'. *The Journal of Cell Biology* 117 (1): 213–24.
<https://doi.org/10.1083/jcb.117.1.213>.
- Gerasimavicius, Lukas, Benjamin J. Livesey, and Joseph A. Marsh. 2022. 'Loss-of-Function, Gain-of-Function and Dominant-Negative Mutations Have Profoundly Different Effects on Protein Structure: Implications for Variant Effect Prediction'. bioRxiv. <https://doi.org/10.1101/2021.10.23.465554>.
- Gerlich, Daniel, Toru Hirota, Birgit Koch, Jan-Michael Peters, and Jan Ellenberg. 2006. 'Condensin I Stabilizes Chromosomes Mechanically through a Dynamic Interaction in Live Cells'. *Current Biology* 16 (4): 333–44.
<https://doi.org/10.1016/j.cub.2005.12.040>.
- Gibcus, Johan H., Kumiko Samejima, Anton Goloborodko, Itaru Samejima, Natalia Naumova, Johannes Nuebler, Masato T. Kanemaki, et al. 2018. 'A Pathway for Mitotic Chromosome Formation'. *Science* 359 (6376): eaao6135.
<https://doi.org/10.1126/science.aao6135>.
- Gilbert, Nick. 2019. 'Biophysical Regulation of Local Chromatin Structure'. *Current Opinion in Genetics & Development*, Genome Architecture and Expression, 55 (April):66–75. <https://doi.org/10.1016/j.gde.2019.06.001>.
- Gilbert, Nick, and James Allan. 2001. 'Distinctive Higher-Order Chromatin Structure at Mammalian Centromeres'. *Proceedings of the National Academy of Sciences* 98 (21): 11949–54. <https://doi.org/10.1073/pnas.211322798>.
- Gilbert, Nick, Shelagh Boyle, Heike Fiegler, Kathryn Woodfine, Nigel P. Carter, and Wendy A. Bickmore. 2004. 'Chromatin Architecture of the Human Genome: Gene-Rich Domains Are Enriched in Open Chromatin Fibers'. *Cell* 118 (5): 555–66.
<https://doi.org/10.1016/j.cell.2004.08.011>.
- Gillentine, Madelyn A., Tianyun Wang, Kendra Hoekzema, Jill Rosenfeld, Pengfei Liu, Hui Guo, Chang N. Kim, et al. 2021. 'Rare Deleterious Mutations of HNRNP Genes Result in Shared Neurodevelopmental Disorders'. *Genome Medicine* 13 (1): 63.
<https://doi.org/10.1186/s13073-021-00870-6>.
- Göhring, Frank, and Frank O. Fackelmayer. 1997. 'The Scaffold/Matrix Attachment Region Binding Protein hnRNP-U (SAF-A) Is Directly Bound to Chromosomal DNA in Vivo: A Chemical Cross-Linking Study'. *Biochemistry* 36 (27): 8276–83.
<https://doi.org/10.1021/bi970480f>.
- Göhring, Frank, Birgit L. Schwab, Pierluigi Nicotera, Marcel Leist, and Frank O. Fackelmayer. 1997. 'The Novel SAR-Binding Domain of Scaffold Attachment Factor A (SAF-A) Is a Target in Apoptotic Nuclear Breakdown'. *The EMBO Journal* 16 (24): 7361–71. <https://doi.org/10.1093/emboj/16.24.7361>.

- Gorbalenya, Alexander E., and Eugene V. Koonin. 1993. 'Helicases: Amino Acid Sequence Comparisons and Structure-Function Relationships'. *Current Opinion in Structural Biology* 3 (3): 419–29. [https://doi.org/10.1016/S0959-440X\(05\)80116-2](https://doi.org/10.1016/S0959-440X(05)80116-2).
- Gorbsky, G J. 1994. 'Cell Cycle Progression and Chromosome Segregation in Mammalian Cells Cultured in the Presence of the Topoisomerase II Inhibitors ICRF-187 [(+)-1,2-Bis(3,5-Dioxopiperazinyl-1-Yl)Propane; ADR-529] and ICRF-159 (Razoxane)'. *Cancer Research* 54 (4): 1042–48.
- Gospodinov, Anastas, Thomas Vaissiere, Dragomir B. Krastev, Gaëlle Legube, Boyka Anachkova, and Zdenko Herceg. 2011. 'Mammalian Ino80 Mediates Double-Strand Break Repair through Its Role in DNA End Strand Resection'. *Molecular and Cellular Biology* 31 (23): 4735–45. <https://doi.org/10.1128/MCB.06182-11>.
- Green, Lydia C., Paul Kalitsis, Tsz M. Chang, Miri Cipetic, Ji Hun Kim, Owen Marshall, Lynne Turnbull, et al. 2012. 'Contrasting Roles of Condensin I and Condensin II in Mitotic Chromosome Formation'. *Journal of Cell Science* 125 (6): 1591–1604. <https://doi.org/10.1242/jcs.097790>.
- Grigoryev, Sergei A., Gaurav Arya, Sarah Correll, Christopher L. Woodcock, and Tamar Schlick. 2009. 'Evidence for Heteromorphic Chromatin Fibers from Analysis of Nucleosome Interactions'. *Proceedings of the National Academy of Sciences* 106 (32): 13317–22. <https://doi.org/10.1073/pnas.0903280106>.
- Gui, Xinrui, Feng Luo, Yichen Li, Heng Zhou, Zhenheng Qin, Zhenying Liu, Jing Gu, et al. 2019. 'Structural Basis for Reversible Amyloids of hnRNPA1 Elucidates Their Role in Stress Granule Assembly'. *Nature Communications* 10 (1): 2006. <https://doi.org/10.1038/s41467-019-09902-7>.
- Gupta, Rajnish A., Nilay Shah, Kevin C. Wang, Jeewon Kim, Hugo M. Horlings, David J. Wong, Miao-Chih Tsai, et al. 2010. 'Long Non-Coding RNA HOTAIR Reprograms Chromatin State to Promote Cancer Metastasis'. *Nature* 464 (7291): 1071–76. <https://doi.org/10.1038/nature08975>.
- Gupta, Rekha, Meenal Agarwal, Vijay R. Boqqula, Rajendra V. Phadke, and Shubha R. Phadke. 2014. 'Hemiconvulsion-Hemiplegia-Epilepsy Syndrome with 1q44 Microdeletion: Causal or Chance Association'. *American Journal of Medical Genetics, Part A* 164 (1): 186–89. <https://doi.org/10.1002/ajmg.a.36198>.
- Haarhuis, Judith H. I., Robin H. van der Weide, Vincent A. Blomen, J. Omar Yáñez-Cuna, Mario Amendola, Marjon S. van Ruiten, Peter H. L. Krijger, et al. 2017. 'The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension'. *Cell* 169 (4): 693–707.e14. <https://doi.org/10.1016/j.cell.2017.04.013>.
- Hacisuleyman, Ezgi, Loyal A. Goff, Cole Trapnell, Adam Williams, Jorge Henao-Mejia, Lei Sun, Patrick McClanahan, et al. 2014. 'Topological Organization of Multichromosomal

- Regions by the Long Intergenic Noncoding RNA Firre'. *Nature Structural & Molecular Biology* 21 (2): 198–206. <https://doi.org/10.1038/nsmb.2764>.
- Hacisuleyman, Ezgi, Chinmay J. Shukla, Catherine L. Weiner, and John L. Rinn. 2016. 'Function and Evolution of Local Repeats in the Firre Locus'. *Nature Communications* 7 (1): 11021. <https://doi.org/10.1038/ncomms11021>.
- Hall, Lisa L., Dawn M. Carone, Alvin V. Gomez, Heather J. Kolpa, Meg Byron, Nitish Mehta, Frank O. Fackelmayer, and Jeanne B. Lawrence. 2014. 'Stable COT-1 Repeat RNA Is Abundant and Is Associated with Euchromatic Interphase Chromosomes'. *Cell* 156 (5): 907–19. <https://doi.org/10.1016/j.cell.2014.01.042>.
- Hall, Lisa L, and Jeanne B Lawrence. 2016. 'RNA as a Fundamental Component of Interphase Chromosomes: Could Repeats Prove Key?' *Current Opinion in Genetics & Development*, Genome architecture and expression, 37 (April):137–47. <https://doi.org/10.1016/j.gde.2016.04.005>.
- Hamdan, Fadi F., Myriam Srour, Jose Mario Capo-Chichi, Hussein Daoud, Christina Nassif, Lysanne Patry, Christine Massicotte, et al. 2014. 'De Novo Mutations in Moderate or Severe Intellectual Disability'. Edited by Gregory M. Cooper. *PLoS Genetics* 10 (10): e1004772. <https://doi.org/10.1371/journal.pgen.1004772>.
- Hancock, Ronald. 2000. 'A New Look at the Nuclear Matrix'. *Chromosoma* 109 (4): 219–25. <https://doi.org/10.1007/s004120000077>.
- Hansen, Thomas B., Trine I. Jensen, Bettina H. Clausen, Jesper B. Bramsen, Bente Finsen, Christian K. Damgaard, and Jørgen Kjems. 2013. 'Natural RNA Circles Function as Efficient microRNA Sponges'. *Nature* 495 (7441): 384–88. <https://doi.org/10.1038/nature11993>.
- Happel, Nicole, Ekkehard Schulze, and Detlef Doenecke. 2005. 'Characterisation of Human Histone H1x' 386 (6): 541–51. <https://doi.org/10.1515/BC.2005.064>.
- Harada, Tomoki, Shoji Hata, Rioka Takagi, Takuma Komori, Masamitsu Fukuyama, Takumi Chinen, and Daiju Kitagawa. 2023. 'An Antioxidant Screen Identifies Ascorbic Acid for Prevention of Light-Induced Mitotic Prolongation in Live Cell Imaging'. *Communications Biology* 6 (1): 1–12. <https://doi.org/10.1038/s42003-023-05479-6>.
- Hasegawa, Yuko, Neil Brockdorff, Shinji Kawano, Kimiko Tsutui, Ken Tsutui, and Shinichi Nakagawa. 2010. 'The Matrix Protein hnRNP U Is Required for Chromosomal Localization of Xist RNA'. *Developmental Cell* 19 (3): 469–76. <https://doi.org/10.1016/j.devcel.2010.08.006>.
- Hauf, Silke, Elisabeth Roitinger, Birgit Koch, Christina M. Dittrich, Karl Mechtler, and Jan-Michael Peters. 2005. 'Dissociation of Cohesin from Chromosome Arms and Loss of Arm Cohesion during Early Mitosis Depends on Phosphorylation of SA2'. *PLOS Biology* 3 (3): e69. <https://doi.org/10.1371/journal.pbio.0030069>.

- Hauf, Silke, Irene C. Waizenegger, and Jan-Michael Peters. 2001. 'Cohesin Cleavage by Separase Required for Anaphase and Cytokinesis in Human Cells'. *Science* 293 (5533): 1320–23. <https://doi.org/10.1126/science.1061376>.
- Heath, Catherine. 2018. 'The Localisation of Non-Coding RNA in the Human Cell'. Phd, University of Sheffield. <https://etheses.whiterose.ac.uk/22498/>.
- Henninger, Jonathan E., Ozgur Oksuz, Krishna Shrinivas, Ido Sagi, Gary LeRoy, Ming M. Zheng, J. Owen Andrews, et al. 2021. 'RNA-Mediated Feedback Control of Transcriptional Condensates'. *Cell* 184 (1): 207–225.e24. <https://doi.org/10.1016/j.cell.2020.11.030>.
- Henriques, Ricardo, Caron Griffiths, E. Hesper Rego, and Musa M. Mhlanga. 2011. 'PALM and STORM: Unlocking Live-Cell Super-Resolution'. *Biopolymers* 95 (5): 322–31. <https://doi.org/10.1002/bip.21586>.
- Herzog, Veronika A., Brian Reichholf, Tobias Neumann, Philipp Rescheneder, Pooja Bhat, Thomas R. Burkard, Wiebke Wlotzka, Arndt Von Haeseler, Johannes Zuber, and Stefan L. Ameres. 2017. 'Thiol-Linked Alkylation of RNA to Assess Expression Dynamics'. *Nature Methods* 14 (12): 1198–1204. <https://doi.org/10.1038/nmeth.4435>.
- Hesselberth, Jay R. 2013. 'Lives That Introns Lead after Splicing'. *WIREs RNA* 4 (6): 677–91. <https://doi.org/10.1002/wrna.1187>.
- Hetzer, Martin W. 2010. 'The Nuclear Envelope'. *Cold Spring Harbor Perspectives in Biology* 2 (3): a000539. <https://doi.org/10.1101/cshperspect.a000539>.
- Hill, David A. 2001. 'Influence of Linker Histone H1 on Chromatin Remodeling'. *Biochemistry and Cell Biology* 79 (3): 317–24. <https://doi.org/10.1139/o01-098>.
- Hinokuma, Nodoka, Mitsuko Nakashima, Hideyuki Asai, Kazuyuki Nakamura, Shinjiro Akaboshi, Masataka Fukuoka, Masami Togawa, et al. 2020. 'Clinical and Genetic Characteristics of Patients with Doose Syndrome'. *Epilepsia Open* 5 (3): 442–50. <https://doi.org/10.1002/epi4.12417>.
- Hirano, Tatsuya, Ryuji Kobayashi, and Michiko Hirano. 1997. 'Condensins, Chromosome Condensation Protein Complexes Containing XCAP-C, XCAP-E and a Xenopus Homolog of the Drosophila Barren Protein'. *Cell* 89 (4): 511–21. [https://doi.org/10.1016/S0092-8674\(00\)80233-0](https://doi.org/10.1016/S0092-8674(00)80233-0).
- Hirano, Tatsuya, and Timothy J. Mitchison. 1994. 'A Heterodimeric Coiled-Coil Protein Required for Mitotic Chromosome Condensation in Vitro'. *Cell* 79 (3): 449–58. [https://doi.org/10.1016/0092-8674\(94\)90254-2](https://doi.org/10.1016/0092-8674(94)90254-2).
- Hnisz, Denes, Krishna Shrinivas, Richard A. Young, Arup K. Chakraborty, and Phillip A. Sharp. 2017. 'A Phase Separation Model for Transcriptional Control'. *Cell* 169 (1): 13–23. <https://doi.org/10.1016/j.cell.2017.02.007>.

- Ho, Yuan-Soon, Jiing-Shium Duh, Jjiang-Huei Jeng, Ying-Jan Wang, Yu-Chih Liang, Chien-Huang Lin, Chia-Jen Tseng, Cheng-Fei Yu, Rong-Jane Chen, and Jen-Kun Lin. 2001. 'Griseofulvin Potentiates Antitumorigenesis Effects of Nocodazole through Induction of Apoptosis and G2/M Cell Cycle Arrest in Human Colorectal Cancer Cells'. *International Journal of Cancer* 91 (3): 393–401. [https://doi.org/10.1002/1097-0215\(200002\)9999:9999::AID-IJC1070>3.0.CO;2-#](https://doi.org/10.1002/1097-0215(200002)9999:9999::AID-IJC1070>3.0.CO;2-#).
- Hochegger, Helfrid, Donniphath Dejsuphong, Eiichiro Sonoda, Alihossein Saberi, Eeson Rajendra, Jane Kirk, Tim Hunt, and Shunichi Takeda. 2007. 'An Essential Role for Cdk1 in S Phase Control Is Revealed via Chemical Genetics in Vertebrate Cells'. *Journal of Cell Biology* 178 (2): 257–68. <https://doi.org/10.1083/jcb.200702034>.
- Hong, Zehui, Jie Jiang, Jiaolong Ma, Shikui Dai, Ting Xu, Hui Li, and Akira Yasui. 2013. 'The Role of hnRPU1 Involved in DNA Damage Response Is Related to PARP1'. *PLOS ONE* 8 (4): e60208. <https://doi.org/10.1371/journal.pone.0060208>.
- Hosogane, Masaki, Ryo Funayama, Matsuyuki Shirota, and Keiko Nakayama. 2016. 'Lack of Transcription Triggers H3K27me3 Accumulation in the Gene Body'. *Cell Reports* 16 (3): 696–706. <https://doi.org/10.1016/j.celrep.2016.06.034>.
- Houlard, Martin, Erin E Cutts, Muhammad S Shamim, Jonathan Godwin, David Weisz, Aviva Presser Aiden, Erez Lieberman Aiden, Lothar Schermelleh, Alessandro Vannini, and Kim Nasmyth. 2021. 'MCPH1 Inhibits Condensin II during Interphase by Regulating Its SMC2-Kleisin Interface'. Edited by Adèle L Marston, Kevin Struhl, Adèle L Marston, and Andrew J Wood. *eLife* 10 (December): e73348. <https://doi.org/10.7554/eLife.73348>.
- Hsiung, C. C.S., Caroline R. Bartman, Peng Huang, Paul Ginart, Aaron J. Stonestrom, Cheryl A. Keller, Carolyn Face, et al. 2016. 'A Hyperactive Transcriptional State Marks Genome Reactivation at the Mitosis-G1 Transition'. *Genes and Development* 30 (12): 1423–39. <https://doi.org/10.1101/gad.280859.116>.
- Hsu, Jer-Yuan, Zu-Wen Sun, Xiumin Li, Melanie Reuben, Kelly Tatchell, Douglas K. Bishop, Jeremy M. Grushcow, et al. 2000. 'Mitotic Phosphorylation of Histone H3 Is Governed by Ipl1/Aurora Kinase and Glc7/PP1 Phosphatase in Budding Yeast and Nematodes'. *Cell* 102 (3): 279–91. [https://doi.org/10.1016/S0092-8674\(00\)00034-9](https://doi.org/10.1016/S0092-8674(00)00034-9).
- Hudson, Damien F., Shinya Ohta, Tina Freisinger, Fiona MacIsaac, Lau Sennels, Flavia Alves, Fan Lai, Alastair Kerr, Juri Rappsilber, and William C. Earnshaw. 2008. 'Molecular and Genetic Analysis of Condensin Function in Vertebrate Cells'. *Molecular Biology of the Cell* 19 (7): 3070–79. <https://doi.org/10.1091/mbc.e08-01-0057>.
- Huelga, Stephanie C., Anthony Q. Vu, Justin D. Arnold, Tiffany D. Liang, Patrick P. Liu, Bernice Y. Yan, John Paul Donohue, et al. 2012. 'Integrative Genome-Wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins'. *Cell Reports* 1 (2): 167–78. <https://doi.org/10.1016/j.celrep.2012.02.001>.

- Huertas, Pablo, and Andrés Aguilera. 2003. 'Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination'. *Molecular Cell* 12 (3): 711–21. <https://doi.org/10.1016/j.molcel.2003.08.010>.
- Hur, Shin-Kyoung, Eun-Jung Park, Ji-Eun Han, Yoon-Ah Kim, Jong-Doo Kim, Dongmin Kang, and Jongbum Kwon. 2010. 'Roles of Human INO80 Chromatin Remodeling Enzyme in DNA Replication and Chromosome Segregation Suppress Genome Instability'. *Cellular and Molecular Life Sciences* 67 (13): 2283–96. <https://doi.org/10.1007/s00018-010-0337-3>.
- Initiative, Epilepsy Genetics. 2019. 'The Epilepsy Genetics Initiative: Systematic Reanalysis of Diagnostic Exomes Increases Yield'. *Epilepsia* 60 (5): 797–806. <https://doi.org/10.1111/epi.14698>.
- Izumi, Hiroto, and Keiko Funa. 2019. 'Telomere Function and the G-Quadruplex Formation Are Regulated by hnRNP U'. *Cells* 8 (5): 390. <https://doi.org/10.3390/cells8050390>.
- Izzo, Annalisa, Kinga Kamieniarz, and Robert Schneider. 2008. 'The Histone H1 Family: Specific Members, Specific Functions?' 389 (4): 333–43. <https://doi.org/10.1515/BC.2008.037>.
- Jackson, Dean A. 2005. 'The Amazing Complexity of Transcription Factories'. *Briefings in Functional Genomics* 4 (2): 143–57. <https://doi.org/10.1093/bfpg/4.2.143>.
- Jang, Soo Hwa, Ah Ram Kim, Neung Hwa Park, Jeong Woo Park, and In Seob Han. 2016. 'DRG2 Regulates G2/M Progression via the Cyclin B1-Cdk1 Complex'. *Molecules and Cells* 39 (9): 699–704. <https://doi.org/10.14348/molcells.2016.0149>.
- Jeppsson, Kristian, Takaharu Kanno, Katsuhiko Shirahige, and Camilla Sjögren. 2014. 'The Maintenance of Chromosome Structure: Positioning and Functioning of SMC Complexes'. *Nature Reviews Molecular Cell Biology* 15 (9): 601–14. <https://doi.org/10.1038/nrm3857>.
- Johannesen, Katrine M., Natalya Nikanorova, Dragan Marjanovic, Agnieszka Pavbro, Line H.G. Larsen, Guido Rubboli, and Rikke S. Møller. 2020. 'Utility of Genetic Testing for Therapeutic Decision-Making in Adults with Epilepsy'. *Epilepsia* 61 (6): 1234–39. <https://doi.org/10.1111/epi.16533>.
- Jonkers, Iris, Hojoong Kwak, and John T Lis. 2014. 'Genome-Wide Dynamics of Pol II Elongation and Its Interplay with Promoter Proximal Pausing, Chromatin, and Exons'. Edited by Kevin Struhl. *eLife* 3 (April): e02407. <https://doi.org/10.7554/eLife.02407>.
- Kadauke, Stephan, and Gerd A. Blobel. 2013. 'Mitotic Bookmarking by Transcription Factors'. *Epigenetics and Chromatin* 6 (1). <https://doi.org/10.1186/1756-8935-6-6>.

- Karagiannis, T. C., and A. El-Osta. 2007. 'Chromatin Modifications and DNA Double-Strand Breaks: The Current State of Play'. *Leukemia* 21 (2): 195–200. <https://doi.org/10.1038/sj.leu.2404478>.
- Karczewski, Konrad J., Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, et al. 2020. 'The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans'. *Nature* 581 (7809): 434–43. <https://doi.org/10.1038/s41586-020-2308-7>.
- Kenny, Martin, and Ingmar Schoen. 2021. 'Violin SuperPlots: Visualizing Replicate Heterogeneity in Large Data Sets'. *Molecular Biology of the Cell* 32 (15): 1333–34. <https://doi.org/10.1091/mbc.E21-03-0130>.
- Keszhelyi, Andrea, Nicola E. Minchell, and Jonathan Baxter. 2016. 'The Causes and Consequences of Topological Stress during DNA Replication'. *Genes* 7 (12): 134. <https://doi.org/10.3390/genes7120134>.
- Khater, Ismail M., Ivan Robert Nabi, and Ghassan Hamarneh. 2020. 'A Review of Super-Resolution Single-Molecule Localization Microscopy Cluster Analysis and Quantification Methods'. *Patterns* 1 (3): 100038. <https://doi.org/10.1016/j.patter.2020.100038>.
- Kiledjian, M., and G. Dreyfuss. 1992. 'Primary Structure and Binding Activity of the hnRNP U Protein: Binding RNA through RGG Box.' *The EMBO Journal* 11 (7): 2655–64. <https://doi.org/10.1002/j.1460-2075.1992.tb05331.x>.
- Kim, Hong Joo, Nam Chul Kim, Yong-Dong Wang, Emily A. Scarborough, Jennifer Moore, Zamia Diaz, Kyle S. MacLea, et al. 2013. 'Mutations in Prion-like Domains in hnRNPA2B1 and hnRNPA1 Cause Multisystem Proteinopathy and ALS'. *Nature* 495 (7442): 467–73. <https://doi.org/10.1038/nature11922>.
- Kim, Myung K., and Vera M. Nikodem. 1999. 'hnRNP U Inhibits Carboxy-Terminal Domain Phosphorylation by TFIIH and Represses RNA Polymerase II Elongation'. *Molecular and Cellular Biology* 19 (10): 6833–44. <https://doi.org/10.1128/MCB.19.10.6833>.
- Kind, Jop, Ludo Pagie, Havva Ortabozkoyun, Shelagh Boyle, Sandra S. De Vries, Hans Janssen, Mario Amendola, Leisha D. Nolen, Wendy A. Bickmore, and Bas Van Steensel. 2013. 'Single-Cell Dynamics of Genome-Nuclear Lamina Interactions'. *Cell* 153 (1): 178–92. <https://doi.org/10.1016/j.cell.2013.02.028>.
- King, Daniel A., Tomas W. Fitzgerald, Ray Miller, Natalie Canham, Jill Clayton-Smith, Diana Johnson, Sahar Mansour, Fiona Stewart, Pradeep Vasudevan, and Matthew E. Hurles. 2014. 'A Novel Method for Detecting Uniparental Disomy from Trio Genotypes Identifies a Significant Excess in Children with Developmental Disorders'. *Genome Research* 24 (4): 673–87. <https://doi.org/10.1101/gr.160465.113>.

- Kipp, Michael, Birgit L. Schwab, Michael Przybylski, Pierluigi Nicotera, and Frank O. Fackelmayer. 2000. 'Apoptotic Cleavage of Scaffold Attachment Factor A (SAF-A) by Caspase-3 Occurs at a Noncanonical Cleavage Site *'. *Journal of Biological Chemistry* 275 (7): 5031–36. <https://doi.org/10.1074/jbc.275.7.5031>.
- Kolpa, Heather J., Frank O. Fackelmayer, and Jeanne B. Lawrence. 2016. 'SAF-A Requirement in Anchoring XIST RNA to Chromatin Varies in Transformed and Primary Cells'. *Developmental Cell* 39 (1): 9–10. <https://doi.org/10.1016/j.devcel.2016.09.021>.
- Kosho, Tomoki, Nobuhiko Okamoto, and Coffin-Siris Syndrome International Collaborators. 2014. 'Genotype-Phenotype Correlation of Coffin-Siris Syndrome Caused by Mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A'. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 166 (3): 262–75. <https://doi.org/10.1002/ajmg.c.31407>.
- Koster, Daniel A., Vincent Croquette, Cees Dekker, Stewart Shuman, and Nynke H. Dekker. 2005. 'Friction and Torque Govern the Relaxation of DNA Supercoils by Eukaryotic Topoisomerase IB'. *Nature* 434 (7033): 671–74. <https://doi.org/10.1038/nature03395>.
- Kouzarides, Tony. 2007. 'Chromatin Modifications and Their Function'. *Cell* 128 (4): 693–705. <https://doi.org/10.1016/j.cell.2007.02.005>.
- Kovalevich, Jane, and Dianne Langford. 2013. 'Considerations for the Use of SH-SY5Y Neuroblastoma Cells in Neurobiology'. In *Neuronal Cell Culture: Methods and Protocols*, edited by Shohreh Amini and Martyn K. White, 9–21. Methods in Molecular Biology. Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-62703-640-5_2.
- Kovel, Carolien G.F. de, Eva H. Brilstra, Marjan J.A. Van Kempen, Ruben Van't Slot, Isaac J. Nijman, Zaid Afawi, Peter De Jonghe, et al. 2016. 'Targeted Sequencing of 351 Candidate Genes for Epileptic Encephalopathy in a Large Cohort of Patients'. *Molecular Genetics and Genomic Medicine* 4 (5): 568–80. <https://doi.org/10.1002/mgg3.235>.
- Krantz, Ian D., Jennifer McCallum, Cheryl DeScipio, Maninder Kaur, Lynette A. Gillis, Dinah Yaeger, Lori Jukofsky, et al. 2004. 'Cornelia de Lange Syndrome Is Caused by Mutations in NIPBL, the Human Homolog of Drosophila Melanogaster Nipped-B'. *Nature Genetics* 36 (6): 631–35. <https://doi.org/10.1038/ng1364>.
- Krasteva, Veneta, Gerald R. Crabtree, and Julie A. Lessard. 2017. 'The BAF45a/PHF10 Subunit of SWI/SNF-like Chromatin Remodeling Complexes Is Essential for Hematopoietic Stem Cell Maintenance'. *Experimental Hematology* 48 (April): 58–71.e15. <https://doi.org/10.1016/j.exphem.2016.11.008>.
- Krepel, Dana, Ryan R. Cheng, Michele Di Pierro, and José N. Onuchic. 2018. 'Deciphering the Structure of the Condensin Protein Complex'. *Proceedings of the*

National Academy of Sciences 115 (47): 11911–16.

<https://doi.org/10.1073/pnas.1812770115>.

Krude, T., and C. Keller. 2001. 'Chromatin Assembly during S Phase: Contributions from Histone Deposition, DNA Replication and the Cell Division Cycle'. *Cellular and Molecular Life Sciences CMLS* 58 (5): 665–72. <https://doi.org/10.1007/PL00000890>.

Kruithof, Maarten, Fan-Tso Chien, Andrew Routh, Colin Logie, Daniela Rhodes, and John van Noort. 2009. 'Single-Molecule Force Spectroscopy Reveals a Highly Compliant Helical Folding for the 30-Nm Chromatin Fiber'. *Nature Structural & Molecular Biology* 16 (5): 534–40. <https://doi.org/10.1038/nsmb.1590>.

Krumm, Niklas, Tychele N Turner, Carl Baker, Laura Vives, Kiana Mohajeri, Kali Witherspoon, Archana Raja, et al. 2015. 'Excess of Rare, Inherited Truncating Mutations in Autism'. *Nature Genetics* 47 (6): 582–88. <https://doi.org/10.1038/ng.3303>.

Kukalev, Alexander, Ylva Nord, Carina Palmberg, Tomas Bergman, and Piergiorgio Percipalle. 2005. 'Actin and hnRNP U Cooperate for Productive Transcription by RNA Polymerase II'. *Nature Structural & Molecular Biology* 12 (3): 238–44. <https://doi.org/10.1038/nsmb904>.

KZHYSHKOWSKA, Julia, Andre RUSCH, Hans WOLF, and Thomas DOBNER. 2003. 'Regulation of Transcription by the Heterogeneous Nuclear Ribonucleoprotein E1B-AP5 Is Mediated by Complex Formation with the Novel Bromodomain-Containing Protein BRD7'. *Biochemical Journal* 371 (2): 385–93. <https://doi.org/10.1042/bj20021281>.

Lagier-Tourenne, Clotilde, Magdalini Polymenidou, and Don W. Cleveland. 2010. 'TDP-43 and FUS/TLS: Emerging Roles in RNA Processing and Neurodegeneration'. *Human Molecular Genetics* 19 (R1): R46–64. <https://doi.org/10.1093/hmg/ddq137>.

Lai, Beibei, Jianghuan Zou, Zhaoyu Lin, Zhipeng Qu, Anying Song, Ying Xu, and Xiang Gao. 2018. 'Haploinsufficiency of hnRNP U Changes Activity Pattern and Metabolic Rhythms'. *American Journal of Pathology* 188 (1): 173–83. <https://doi.org/10.1016/j.ajpath.2017.09.017>.

Lammens, Alfred, Alexandra Schele, and Karl-Peter Hopfner. 2004. 'Structural Biochemistry of ATP-Driven Dimerization and DNA-Stimulated Activation of SMC ATPases'. *Current Biology* 14 (19): 1778–82. <https://doi.org/10.1016/j.cub.2004.09.044>.

Lan, Wei, Min Li, Kaijie Zhao, Jin Liu, Fang-Xiang Wu, Yi Pan, and Jianxin Wang. 2017. 'LDAP: A Web Server for lncRNA-Disease Association Prediction'. *Bioinformatics* 33 (3): 458–60. <https://doi.org/10.1093/bioinformatics/btw639>.

- Langmore, JP, and Paulson JR. 1983. 'Low Angle X-Ray Diffraction Studies of Chromatin Structure in Vivo and in Isolated Nuclei and Metaphase Chromosomes'. *Journal of Cell Biology* 96 (4): 1120–31. <https://doi.org/10.1083/jcb.96.4.1120>.
- Leduc, Magalie S., Hsiao Tuan Chao, Chunjing Qu, Magdalena Walkiewicz, Rui Xiao, Pilar Magoulas, Shujuan Pan, et al. 2017. 'Clinical and Molecular Characterization of de Novo Loss of Function Variants in HNRNPU'. *American Journal of Medical Genetics, Part A* 173 (10): 2680–89. <https://doi.org/10.1002/ajmg.a.38388>.
- Lee, Byung-Gil, Fabian Merkel, Matteo Allegretti, Markus Hassler, Christopher Cawood, Léa Lecomte, Francis J. O'Reilly, et al. 2020. 'Cryo-EM Structures of Holo Condensin Reveal a Subunit Flip-Flop Mechanism'. *Nature Structural & Molecular Biology* 27 (8): 743–51. <https://doi.org/10.1038/s41594-020-0457-x>.
- Lee, Byung-Gil, James Rhodes, and Jan Löwe. 2022. 'Clamping of DNA Shuts the Condensin Neck Gate'. *Proceedings of the National Academy of Sciences* 119 (14): e2120006119. <https://doi.org/10.1073/pnas.2120006119>.
- Lee, Chong M., Guanshi Wang, Alexandros Pertsinidis, and Kenneth J. Mariani. 2019. 'Topoisomerase III Acts at the Replication Fork To Remove Precatenanes'. *Journal of Bacteriology* 201 (7): e00563-18. <https://doi.org/10.1128/JB.00563-18>.
- Lee, Hsiao-Hui, Chung-Liang Chien, Hsin-Kai Liao, Yu-Ju Chen, and Zee-Fen Chang. 2004. 'Nuclear Efflux of Heterogeneous Nuclear Ribonucleoprotein C1/C2 in Apoptotic Cells: A Novel Nuclear Export Dependent on Rho-Associated Kinase Activation'. *Journal of Cell Science* 117 (23): 5579–89. <https://doi.org/10.1242/jcs.01482>.
- Lee, Joyce H., and James M. Berger. 2019. 'Cell Cycle-Dependent Control and Roles of DNA Topoisomerase II'. *Genes* 10 (11): 859. <https://doi.org/10.3390/genes10110859>.
- Lee, Robin van der, Marija Buljan, Benjamin Lang, Robert J. Weatheritt, Gary W. Daughdrill, A. Keith Dunker, Monika Fuxreiter, et al. 2014. 'Classification of Intrinsically Disordered Regions and Proteins'. *Chemical Reviews* 114 (13): 6589–6631. <https://doi.org/10.1021/cr400525m>.
- Legnini, Ivano, Gaia Di Timoteo, Francesca Rossi, Mariangela Morlando, Francesca Briganti, Olga Sthandier, Alessandro Fatica, et al. 2017. 'Circ-ZNF609 Is a Circular RNA That Can Be Translated and Functions in Myogenesis'. *Molecular Cell* 66 (1): 22-37.e9. <https://doi.org/10.1016/j.molcel.2017.02.017>.
- Lessard, Julie, Jiang I. Wu, Jeffrey A. Ranish, Mimi Wan, Monte M. Winslow, Brett T. Staahl, Hai Wu, Ruedi Aebersold, Isabella A. Graef, and Gerald R. Crabtree. 2007. 'An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development'. *Neuron* 55 (2): 201–15. <https://doi.org/10.1016/j.neuron.2007.06.019>.

- Levone, Brunno R., Silvia C. Lenzken, Marco Antonaci, Andreas Maiser, Alexander Rapp, Francesca Conte, Stefan Reber, et al. 2021. 'FUS-Dependent Liquid–Liquid Phase Separation Is Important for DNA Repair Initiation'. *Journal of Cell Biology* 220 (5): e202008030. <https://doi.org/10.1083/jcb.202008030>.
- Li, Andra, Yaping Yu, Sheng-Chun Lee, Toyotaka Ishibashi, Susan P. Lees-Miller, and Juan Ausió. 2010. 'Phosphorylation of Histone H2A.X by DNA-Dependent Protein Kinase Is Not Affected by Core Histone Acetylation, but It Alters Nucleosome Stability and Histone H1 Binding *'. *Journal of Biological Chemistry* 285 (23): 17778–88. <https://doi.org/10.1074/jbc.M110.116426>.
- Li, Xiaolei, Zhiqiang Wu, Xiaobing Fu, and Weidong Han. 2013. 'Long Noncoding RNAs: Insights from Biological Features and Functions to Diseases'. *Medicinal Research Reviews* 33 (3): 517–53. <https://doi.org/10.1002/med.21254>.
- Liang, Zhangyi, Denise Zickler, Mara Prentiss, Frederick S. Chang, Guillaume Witz, Kazuhiro Maeshima, and Nancy Kleckner. 2015. 'Chromosomes Progress to Metaphase in Multiple Discrete Steps via Global Compaction/Expansion Cycles'. *Cell* 161 (5): 1124–37. <https://doi.org/10.1016/j.cell.2015.04.030>.
- Lieberman-Aiden, Erez, Nynke L. van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, et al. 2009. 'Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome'. *Science* 326 (5950): 289–93. <https://doi.org/10.1126/science.1181369>.
- Lim, Elaine T., Mohammed Uddin, Silvia De Rubeis, Yingleong Chan, Anne S. Kamumbu, Xiaochang Zhang, Alissa M. D’Gama, et al. 2017. 'Rates, Distribution and Implications of Postzygotic Mosaic Mutations in Autism Spectrum Disorder'. *Nature Neuroscience* 20 (9): 1217–24. <https://doi.org/10.1038/nn.4598>.
- Lin, Yang-Hsiang, Meng-Han Wu, Chau-Ting Yeh, and Kwang-Huei Lin. 2018. 'Long Non-Coding RNAs as Mediators of Tumor Microenvironment and Liver Cancer Cell Communication'. *International Journal of Molecular Sciences* 19 (12): 3742. <https://doi.org/10.3390/ijms19123742>.
- Lin, Yanxian, James McCarty, Jennifer N Rauch, Kris T Delaney, Kenneth S Kosik, Glenn H Fredrickson, Joan-Emma Shea, and Songi Han. 2019. 'Narrow Equilibrium Window for Complex Coacervation of Tau and RNA under Cellular Conditions'. Edited by Yibing Shan, Naama Barkai, and Carlos Castaneda. *eLife* 8 (April): e42571. <https://doi.org/10.7554/eLife.42571>.
- Linde, Sebastian van de, Anna Löschberger, Teresa Klein, Meike Heidebreder, Steve Wolter, Mike Heilemann, and Markus Sauer. 2011. 'Direct Stochastic Optical Reconstruction Microscopy with Standard Fluorescent Probes'. *Nature Protocols* 6 (7): 991–1009. <https://doi.org/10.1038/nprot.2011.336>.

- Lipp, Jesse J., Toru Hirota, Ina Poser, and Jan-Michael Peters. 2007. 'Aurora B Controls the Association of Condensin I but Not Condensin II with Mitotic Chromosomes'. *Journal of Cell Science* 120 (7): 1245–55. <https://doi.org/10.1242/jcs.03425>.
- Liu, Jinglan, Zhe Zhang, Masashige Bando, Takehiko Itoh, Matthew A. Deardorff, Dinah Clark, Maninder Kaur, et al. 2009. 'Transcriptional Dysregulation in NIPBL and Cohesin Mutant Human Cells'. *PLOS Biology* 7 (5): e1000119. <https://doi.org/10.1371/journal.pbio.1000119>.
- Liu, Leroy F., Chung-Cheng Liu, and Bruce M. Alberts. 1980. 'Type II DNA Topoisomerases: Enzymes That Can Unknot a Topologically Knotted DNA Molecule via a Reversible Double-Strand Break'. *Cell* 19 (3): 697–707. [https://doi.org/10.1016/S0092-8674\(80\)80046-8](https://doi.org/10.1016/S0092-8674(80)80046-8).
- Lopez, Christopher R., Shirley Yang, Richard W. Deibler, Starlight A. Ray, Jeanine M. Pennington, Russell J. DiGate, P. J. Hastings, Susan M. Rosenberg, and E. Lynn Zechiedrich. 2005. 'A Role for Topoisomerase III in a Recombination Pathway Alternative to RuvABC'. *Molecular Microbiology* 58 (1): 80–101. <https://doi.org/10.1111/j.1365-2958.2005.04812.x>.
- Losada, Ana, Michiko Hirano, and Tatsuya Hirano. 2002. 'Cohesin Release Is Required for Sister Chromatid Resolution, but Not for Condensin-Mediated Compaction, at the Onset of Mitosis'. *Genes & Development* 16 (23): 3004–16. <https://doi.org/10.1101/gad.249202>.
- Losada, Ana, Tomoki Yokochi, and Tatsuya Hirano. 2005. 'Functional Contribution of Pds5 to Cohesin-Mediated Cohesion in Human Cells and Xenopus Egg Extracts'. *Journal of Cell Science* 118 (10): 2133–41. <https://doi.org/10.1242/jcs.02355>.
- Lu, Junjie, Feng Li, Christopher S. Murphy, Michael W. Davidson, and David M. Gilbert. 2010. 'G2 Phase Chromatin Lacks Determinants of Replication Timing'. *Journal of Cell Biology* 189 (6): 967–80. <https://doi.org/10.1083/jcb.201002002>.
- Luger, Karolin, Armin W. Mäder, Robin K. Richmond, David F. Sargent, and Timothy J. Richmond. 1997. 'Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution'. *Nature* 389 (6648): 251–60. <https://doi.org/10.1038/38444>.
- Lukinavičius, Gražvydas, Claudia Blaukopf, Elias Pershagen, Alberto Schena, Luc Reymond, Emmanuel Derivery, Marcos Gonzalez-Gaitan, et al. 2015. 'SiR–Hoechst Is a Far-Red DNA Stain for Live-Cell Nanoscopy'. *Nature Communications* 6 (1): 8497. <https://doi.org/10.1038/ncomms9497>.
- Ma, Hong, Alan J. Siegel, and Ronald Berezney. 1999. 'Association of Chromosome Territories with the Nuclear Matrix: Disruption of Human Chromosome Territories Correlates with the Release of a Subset of Nuclear Matrix Proteins'. *Journal of Cell Biology* 146 (3): 531–42. <https://doi.org/10.1083/jcb.146.3.531>.

- Maeshima, Kazuhiro, Saera Hihara, and Mikhail Eltsov. 2010. 'Chromatin Structure: Does the 30-Nm Fibre Exist in Vivo?' *Current Opinion in Cell Biology*, Nucleus and gene expression, 22 (3): 291–97. <https://doi.org/10.1016/j.ceb.2010.03.001>.
- Maeshima, Kazuhiro, Satoru Ide, and Michael Babokhov. 2019. 'Dynamic Chromatin Organization without the 30-Nm Fiber'. *Current Opinion in Cell Biology*, Cell Nucleus, 58 (June):95–104. <https://doi.org/10.1016/j.ceb.2019.02.003>.
- Magalska, Adriana, AnnaKatharina Schellhaus, Daniel Moreno-Andrés, Fabio Zanini, Allana Schooley, Ruchika Sachdev, Heinz Schwarz, Johannes Madlung, and Wolfram Antonin. 2014. 'RuvB-like Atpases Function in Chromatin Decondensation at the End of Mitosis'. *Developmental Cell* 31 (3): 305–18. <https://doi.org/10.1016/j.devcel.2014.09.001>.
- Maharana, Shovamayee, Jie Wang, Dimitrios K. Papadopoulos, Doris Richter, Andrey Pozniakovsky, Ina Poser, Marc Bickle, et al. 2018. 'RNA Buffers the Phase Separation Behavior of Prion-like RNA Binding Proteins'. *Science* 360 (6391): 918–21. <https://doi.org/10.1126/science.aar7366>.
- Mahy, Nicola L., Paul E. Perry, Susan Gilchrist, Richard A. Baldock, and Wendy A. Bickmore. 2002. 'Spatial Organization of Active and Inactive Genes and Noncoding DNA within Chromosome Territories'. *The Journal of Cell Biology* 157 (4): 579–89. <https://doi.org/10.1083/jcb.200111071>.
- Marenda, Mattia, Elena Lazarova, and Nick Gilbert. 2022. 'The Role of SAF-A/hnRNP U in Regulating Chromatin Structure'. *Current Opinion in Genetics & Development* 72 (February):38–44. <https://doi.org/10.1016/j.gde.2021.10.008>.
- Marenda, Mattia, Elena Lazarova, Sebastian van de Linde, Nick Gilbert, and Davide Michieletto. 2021. 'Parameter-Free Molecular Super-Structures Quantification in Single-Molecule Localization Microscopy'. *Journal of Cell Biology* 220 (5): e202010003. <https://doi.org/10.1083/jcb.202010003>.
- Martianov, Igor, Stefano Brancorsini, Raffaella Catena, Anne Gansmuller, Noora Kotaja, Martti Parvinen, Paolo Sassone-Corsi, and Irwin Davidson. 2005. 'Polar Nuclear Localization of H1T2, a Histone H1 Variant, Required for Spermatid Elongation and DNA Condensation during Spermiogenesis'. *Proceedings of the National Academy of Sciences* 102 (8): 2808–13. <https://doi.org/10.1073/pnas.0406060102>.
- Martin, Carol-Anne, Jennie E. Murray, Paula Carroll, Andrea Leitch, Karen J. Mackenzie, Mihail Halachev, Ahmed E. Fetit, et al. 2016. 'Mutations in Genes Encoding Condensin Complex Proteins Cause Microcephaly through Decatenation Failure at Mitosis'. *Genes & Development* 30 (19): 2158–72. <https://doi.org/10.1101/gad.286351.116>.
- McClendon, A. Kathleen, A. Chapin Rodriguez, and Neil Osheroff. 2005. 'Human Topoisomerase II α Rapidly Relaxes Positively Supercoiled DNA: IMPLICATIONS

FOR ENZYME ACTION AHEAD OF REPLICATION FORKS *'. *Journal of Biological Chemistry* 280 (47): 39337–45. <https://doi.org/10.1074/jbc.M503320200>.

McHugh, Colleen A., Chun Kan Chen, Amy Chow, Christine F. Surka, Christina Tran, Patrick McDonel, Amy Pandya-Jones, et al. 2015. 'The Xist lncRNA Interacts Directly with SHARP to Silence Transcription through HDAC3'. *Nature* 521 (7551): 232–36. <https://doi.org/10.1038/nature14443>.

McRae, Jeremy F., Stephen Clayton, Tomas W. Fitzgerald, Joanna Kaplanis, Elena Prigmore, Diana Rajan, Alejandro Sifrim, et al. 2017. 'Prevalence and Architecture of de Novo Mutations in Developmental Disorders'. *Nature* 542 (7642): 433–38. <https://doi.org/10.1038/nature21062>.

Meller, Victoria H., Sonal S. Joshi, and Nikita Deshpande. 2015. 'Modulation of Chromatin by Noncoding RNA'. *Annual Review of Genetics* 49 (1): 673–95. <https://doi.org/10.1146/annurev-genet-112414-055205>.

Michieletto, Davide, and Nick Gilbert. 2019. 'Role of Nuclear RNA in Regulating Chromatin Structure and Transcription'. *Current Opinion in Cell Biology, Cell Nucleus*, 58 (June):120–25. <https://doi.org/10.1016/j.ceb.2019.03.007>.

Miki, Takashi S., and Helge Großhans. 2013. 'The Multifunctional RNase XRN2'. *Biochemical Society Transactions* 41 (4): 825–30. <https://doi.org/10.1042/BST20130001>.

Miles, Jamilla. 2022. 'Investigating the Structure and Function of SAF-A', May. <https://doi.org/10.7488/era/2236>.

Miller, Erik L., Diana C. Hargreaves, Cigall Kadoch, Chiung-Ying Chang, Joseph P. Calarco, Courtney Hodges, Jason D. Buenrostro, et al. 2017. 'TOP2 Synergizes with BAF Chromatin Remodeling for Both Resolution and Formation of Facultative Heterochromatin'. *Nature Structural & Molecular Biology* 24 (4): 344–52. <https://doi.org/10.1038/nsmb.3384>.

Miller, Justin M., and Eric J. Enemark. 2016. 'Fundamental Characteristics of AAA+ Protein Family Structure and Function'. *Archaea* 2016 (September):e9294307. <https://doi.org/10.1155/2016/9294307>.

Miller, Michael R., John J. Castellot, and Arthur B. Pardee. 1979. 'A General Method for Permeabilizing Monolayer and Suspension Cultured Animal Cells'. *Experimental Cell Research* 120 (2): 421–25. [https://doi.org/10.1016/0014-4827\(79\)90404-X](https://doi.org/10.1016/0014-4827(79)90404-X).

Mittal, Priya, and Charles W. M. Roberts. 2020. 'The SWI/SNF Complex in Cancer — Biology, Biomarkers and Therapy'. *Nature Reviews Clinical Oncology* 17 (7): 435–48. <https://doi.org/10.1038/s41571-020-0357-3>.

Mizuguchi, Gaku, Xuetong Shen, Joe Landry, Wei-Hua Wu, Subhojit Sen, and Carl Wu. 2004. 'ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1

Chromatin Remodeling Complex'. *Science* 303 (5656): 343–48.

<https://doi.org/10.1126/science.1090701>.

Mizusawa, Yuri, Naoaki Kuji, Yudai Tanaka, Mamoru Tanaka, Eiji Ikeda, Setsuko Komatsu, Shingo Kato, and Yasunori Yoshimura. 2010. 'Expression of Human Oocyte-Specific Linker Histone Protein and Its Incorporation into Sperm Chromatin during Fertilization'. *Fertility and Sterility* 93 (4): 1134–41.

<https://doi.org/10.1016/j.fertnstert.2008.11.028>.

Moen, PT, Jr, KP Smith, and JB Lawrence. 1995. 'Compartmentalization of Specific Pre-mRNA Metabolism: An Emerging View'. *Human Molecular Genetics* 4 (suppl_1): 1779–89. https://doi.org/10.1093/hmg/4.suppl_1.1779.

Moindrot, Benoit, Benjamin Audit, Petra Klous, Antoine Baker, Claude Thermes, Wouter de Laat, Philippe Bouvet, Fabien Mongelard, and Alain Arneodo. 2012. '3D Chromatin Conformation Correlates with Replication Timing and Is Conserved in Resting Cells'. *Nucleic Acids Research* 40 (19): 9470–81.

<https://doi.org/10.1093/nar/gks736>.

Molliex, Amandine, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P. Kanagaraj, Hong Joo Kim, Tanja Mittag, and J. Paul Taylor. 2015. 'Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization'. *Cell* 163 (1): 123–33. <https://doi.org/10.1016/j.cell.2015.09.015>.

Mollwitz, Birgit, Elizabeth Brunk, Simone Schmitt, Florence Pojer, Michael Bannwarth, Marc Schiltz, Ursula Rothlisberger, and Kai Johnsson. 2012. 'Directed Evolution of the Suicide Protein O6-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability'. *Biochemistry* 51 (5): 986–94. <https://doi.org/10.1021/bi2016537>.

Monroe, Glen R., Gerardus W. Frederix, Sanne M.C. Savelberg, Tamar I. De Vries, Karen J. Duran, Jasper J. Van Der Smagt, Paulien A. Terhal, et al. 2016. 'Effectiveness of Whole-Exome Sequencing and Costs of the Traditional Diagnostic Trajectory in Children with Intellectual Disability'. *Genetics in Medicine* 18 (9): 949–56.

<https://doi.org/10.1038/gim.2015.200>.

Mora-Bermúdez, Felipe, Daniel Gerlich, and Jan Ellenberg. 2007. 'Maximal Chromosome Compaction Occurs by Axial Shortening in Anaphase and Depends on Aurora Kinase'. *Nature Cell Biology* 9 (7): 822–31. <https://doi.org/10.1038/ncb1606>.

Morales, Carmen, and Ana Losada. 2018. 'Establishing and Dissolving Cohesion during the Vertebrate Cell Cycle'. *Current Opinion in Cell Biology*, Cell Nucleus, 52 (June):51–57. <https://doi.org/10.1016/j.ceb.2018.01.010>.

Mordes, Daniel, Xiaoyan Luo, Amar Kar, David Kuo, Lili Xu, Kazuo Fushimi, Guowu Yu, Paul Sternberg, and Jane Y. Wu. 2006. 'Pre-mRNA Splicing and Retinitis Pigmentosa'. *Molecular Vision* 12 (October):1259–71.

- Mordes, Daniel, Liya Yuan, Lili Xu, Mariko Kawada, Robert S. Molday, and Jane Y. Wu. 2007. 'Identification of Photoreceptor Genes Affected by PRPF31 Mutations Associated with Autosomal Dominant Retinitis Pigmentosa'. *Neurobiology of Disease* 26 (2): 291–300. <https://doi.org/10.1016/j.nbd.2006.08.026>.
- Morrison, Ashby J., and Xuetong Shen. 2009. 'Chromatin Remodelling beyond Transcription: The INO80 and SWR1 Complexes'. *Nature Reviews Molecular Cell Biology* 10 (6): 373–84. <https://doi.org/10.1038/nrm2693>.
- Muñoz, Sofía, Andrew Jones, Céline Bouchoux, Tegan Gilmore, Harshil Patel, and Frank Uhlmann. 2022. 'Functional Crosstalk between the Cohesin Loader and Chromatin Remodelers'. *Nature Communications* 13 (1): 7698. <https://doi.org/10.1038/s41467-022-35444-6>.
- Murayama, Yasuto, and Frank Uhlmann. 2014. 'Biochemical Reconstitution of Topological DNA Binding by the Cohesin Ring'. *Nature* 505 (7483): 367–71. <https://doi.org/10.1038/nature12867>.
- Murthy, Anastasia C., Gregory L. Dignon, Yelena Kan, Gül H. Zerze, Sapun H. Parekh, Jeetain Mittal, and Nicolas L. Fawzi. 2019. 'Molecular Interactions Underlying Liquid–liquid Phase Separation of the FUS Low-Complexity Domain'. *Nature Structural & Molecular Biology* 26 (7): 637–48. <https://doi.org/10.1038/s41594-019-0250-x>.
- Musco, Giovanna, Gunter Stier, Catherine Joseph, Maria Antonietta Castiglione Morelli, Michael Nilges, Toby J Gibson, and Annalisa Pastore. 1996. 'Three-Dimensional Structure and Stability of the KH Domain: Molecular Insights into the Fragile X Syndrome'. *Cell* 85 (2): 237–45. [https://doi.org/10.1016/S0092-8674\(00\)81100-9](https://doi.org/10.1016/S0092-8674(00)81100-9).
- Muthurajan, Uma M., Steven J. McBryant, Xu Lu, Jeffrey C. Hansen, and Karolin Luger. 2011. 'The Linker Region of MacroH2A Promotes Self-Association of Nucleosomal Arrays *'. *Journal of Biological Chemistry* 286 (27): 23852–64. <https://doi.org/10.1074/jbc.M111.244871>.
- Nabet, Behnam, Justin M. Roberts, Dennis L. Buckley, Joshiawa Paulk, Shiva Dastjerdi, Annan Yang, Alan L. Leggett, et al. 2018. 'The dTAG System for Immediate and Target-Specific Protein Degradation'. *Nature Chemical Biology* 14 (5): 431–41. <https://doi.org/10.1038/s41589-018-0021-8>.
- Nacev, Benjamin A., Lijuan Feng, John D. Bagert, Agata E. Lemiesz, JianJiong Gao, Alexey A. Soshnev, Ritika Kundra, Nikolaus Schultz, Tom W. Muir, and C. David Allis. 2019. 'The Expanding Landscape of "Oncohistone" Mutations in Human Cancers'. *Nature* 567 (7749): 473–78. <https://doi.org/10.1038/s41586-019-1038-1>.
- Nagamani, Sandesh C.Sreenath, Ayelet Erez, Carolyn Bay, Anjana Pettigrew, Seema R. Lalani, Kristin Herman, Brett H. Graham, et al. 2012. 'Delineation of a Deletion Region Critical for Corpus Callosal Abnormalities in Chromosome 1q43-Q44'.

European Journal of Human Genetics 20 (2): 176–79.
<https://doi.org/10.1038/ejhg.2011.171>.

Nano, Nardin, and Walid A. Houry. 2013. 'Chaperone-like Activity of the AAA{thorn} Proteins RVb1 and RVb2 in the Assembly of Various Complexes'. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368 (1617): 1–12.
<https://doi.org/10.1098/rstb.2011.0399>.

Natsume, Toyooki, Tomomi Kiyomitsu, Yumiko Saga, and Masato T. Kanemaki. 2016. 'Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors'. *Cell Reports* 15 (1): 210–18.
<https://doi.org/10.1016/j.celrep.2016.03.001>.

Naughton, Catherine, Covadonga Huidobro, Claudia R. Catacchio, Adam Buckle, Graeme R. Grimes, Ryu-Suke Nozawa, Stefania Purgato, Mariano Rocchi, and Nick Gilbert. 2022. 'Human Centromere Repositioning Activates Transcription and Opens Chromatin Fibre Structure'. *Nature Communications* 13 (1): 5609.
<https://doi.org/10.1038/s41467-022-33426-2>.

Naughton, Catherine, Duncan Sproul, Charlotte Hamilton, and Nick Gilbert. 2010. 'Analysis of Active and Inactive X Chromosome Architecture Reveals the Independent Organization of 30 Nm and Large-Scale Chromatin Structures'. *Molecular Cell* 40 (3): 397–409. <https://doi.org/10.1016/j.molcel.2010.10.013>.

Naumova, Natalia, Maxim Imakaev, Geoffrey Fudenberg, Ye Zhan, Bryan R. Lajoie, Leonid A. Mirny, and Job Dekker. 2013. 'Organization of the Mitotic Chromosome'. *Science* 342 (6161): 948–53. <https://doi.org/10.1126/science.1236083>.

NCBI Resource Coordinators. 2018. 'Database Resources of the National Center for Biotechnology Information'. *Nucleic Acids Research* 46 (D1): D8–13.
<https://doi.org/10.1093/nar/gkx1095>.

Need, Anna C., Vandana Shashi, Yuki Hitomi, Kelly Schoch, Kevin V. Shianna, Marie T. McDonald, Miriam H. Meisler, and David B. Goldstein. 2012. 'Clinical Application of Exome Sequencing in Undiagnosed Genetic Conditions'. *Journal of Medical Genetics* 49 (6): 353–61. <https://doi.org/10.1136/jmedgenet-2012-100819>.

Neitzel, Heidemarie, Luitgard M. Neumann, Detlev Schindler, Andreas Wirges, Holger Tönnies, Marc Trimborn, Alice Krebsova, Reyk Richter, and Karl Sperling. 2002. 'Premature Chromosome Condensation in Humans Associated with Microcephaly and Mental Retardation: A Novel Autosomal Recessive Condition'. *The American Journal of Human Genetics* 70 (4): 1015–22. <https://doi.org/10.1086/339518>.

Nielsen, Christian F., Tao Zhang, Marin Barisic, Paul Kalitsis, and Damien F. Hudson. 2020. 'Topoisomerase II α Is Essential for Maintenance of Mitotic Chromosome Structure'. *Proceedings of the National Academy of Sciences* 117 (22): 12131–42.
<https://doi.org/10.1073/pnas.2001760117>.

- Nora, Elphège P., Anton Goloborodko, Anne-Laure Valton, Johan H. Gibcus, Alec Uebersohn, Nezar Abdennur, Job Dekker, Leonid A. Mirny, and Benoit G. Bruneau. 2017. 'Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization'. *Cell* 169 (5): 930-944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>.
- Nozawa, Ryu Suke, Lora Boteva, Dinesh C. Soares, Catherine Naughton, Alison R. Dun, Adam Buckle, Bernard Ramsahoye, et al. 2017. 'SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin-Associated RNAs'. *Cell* 169 (7): 1214-1227.e18. <https://doi.org/10.1016/j.cell.2017.05.029>.
- Nozawa, Ryu-Suke, and Nick Gilbert. 2019. 'RNA: Nuclear Glue for Folding the Genome'. *Trends in Cell Biology* 29 (3): 201–11. <https://doi.org/10.1016/j.tcb.2018.12.003>.
- Oates, Stephanie, Shan Tang, Richard Rosch, Rosalie Lear, Elaine F. Hughes, Ruth E. Williams, Line H. G. Larsen, et al. 2018. 'Incorporating Epilepsy Genetics into Clinical Practice: A 360° evaluation'. *Npj Genomic Medicine* 3 (1): 1–11. <https://doi.org/10.1038/s41525-018-0052-9>.
- Obordlik, Ales, Alexander Kukalev, Emilie Louvet, Ann-Kristin östlund Farrants, Luca Caputo, and Piergiorgio Percipalle. 2008. 'The Histone Acetyltransferase PCAF Associates with Actin and hnRNP U for RNA Polymerase II Transcription'. *Molecular and Cellular Biology* 28 (20): 6342–57. <https://doi.org/10.1128/MCB.00766-08>.
- Onishi, Yoshiaki, Syuji Hanai, Tomoya Ohno, Yasuhiro Hara, and Norio Ishida. 2008. 'Rhythmic SAF-A Binding Underlies Circadian Transcription of the Bmal1 Gene'. *Molecular and Cellular Biology* 28 (10): 3477–88. <https://doi.org/10.1128/MCB.02227-07>.
- Onn, Itay, Nobuki Aono, Michiko Hirano, and Tatsuya Hirano. 2007. 'Reconstitution and Subunit Geometry of Human Condensin Complexes'. *The EMBO Journal* 26 (4): 1024–34. <https://doi.org/10.1038/sj.emboj.7601562>.
- Ono, Takao, Ana Losada, Michiko Hirano, Michael P. Myers, Andrew F. Neuwald, and Tatsuya Hirano. 2003. 'Differential Contributions of Condensin I and Condensin II to Mitotic Chromosome Architecture in Vertebrate Cells'. *Cell* 115 (1): 109–21. [https://doi.org/10.1016/S0092-8674\(03\)00724-4](https://doi.org/10.1016/S0092-8674(03)00724-4).
- Ou, Horng D., Sébastien Phan, Thomas J. Deerinck, Andrea Thor, Mark H. Ellisman, and Clodagh C. O'Shea. 2017. 'ChromEMT: Visualizing 3D Chromatin Structure and Compaction in Interphase and Mitotic Cells'. *Science* 357 (6349): eaag0025. <https://doi.org/10.1126/science.aag0025>.
- Overbeek, J. T. G., and M. J. Voorn. 1957. 'Phase Separation in Polyelectrolyte Solutions. Theory of Complex Coacervation'. *Journal of Cellular and Comparative Physiology* 49 (S1): 7–26. <https://doi.org/10.1002/jcp.1030490404>.

- Palazzo, Alexander F., and Eliza S. Lee. 2015. 'Non-Coding RNA: What Is Functional and What Is Junk?' *Frontiers in Genetics* 6.
<https://www.frontiersin.org/articles/10.3389/fgene.2015.00002>.
- Palozola, Katherine C., Greg Donahue, Hong Liu, Gregory R. Grant, Justin S. Becker, Allison Cote, Hongtao Yu, Arjun Raj, and Kenneth S. Zaret. 2017. 'Mitotic Transcription and Waves of Gene Reactivation during Mitotic Exit'. *Science* 358 (6359): 119–22. <https://doi.org/10.1126/science.aal4671>.
- Pamudurti, Nagarjuna Reddy, Osnat Bartok, Marvin Jens, Reut Ashwal-Fluss, Christin Stottmeister, Larissa Ruhe, Mor Hanan, et al. 2017. 'Translation of CircRNAs'. *Molecular Cell* 66 (1): 9-21.e7. <https://doi.org/10.1016/j.molcel.2017.02.021>.
- Panchenko, Tanya, Troy C. Sorensen, Christopher L. Woodcock, Zhong-yuan Kan, Stacey Wood, Michael G. Resch, Karolin Luger, S. Walter Englander, Jeffrey C. Hansen, and Ben E. Black. 2011. 'Replacement of Histone H3 with CENP-A Directs Global Nucleosome Array Condensation and Loosening of Nucleosome Superhelical Termini'. *Proceedings of the National Academy of Sciences* 108 (40): 16588–93.
<https://doi.org/10.1073/pnas.1113621108>.
- Pang, Du, Shengping Yu, and Xuejun Yang. 2022. 'A Mini-Review of the Role of Condensin in Human Nervous System Diseases'. *Frontiers in Molecular Neuroscience* 15 (August):889796. <https://doi.org/10.3389/fnmol.2022.889796>.
- Papamichos-Chronakis, Manolis, Shinya Watanabe, Oliver J. Rando, and Craig L. Peterson. 2011. 'Global Regulation of H2A.Z Localization by the INO80 Chromatin-Remodeling Enzyme Is Essential for Genome Integrity'. *Cell* 144 (2): 200–213.
<https://doi.org/10.1016/j.cell.2010.12.021>.
- Parada, Luis A., Philip G. McQueen, and Tom Misteli. 2004. 'Tissue-Specific Spatial Organization of Genomes'. *Genome Biology* 5 (7): R44. <https://doi.org/10.1186/gb-2004-5-7-r44>.
- Parelho, Vania, Suzana Hadjur, Mikhail Spivakov, Marion Leleu, Stephan Sauer, Heather C. Gregson, Adam Jarmuz, et al. 2008. 'Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms'. *Cell* 132 (3): 422–33.
<https://doi.org/10.1016/j.cell.2008.01.011>.
- Park, Ji-Hye, Eun-Jung Park, Han-Sae Lee, So Jung Kim, Shin-Kyoung Hur, Anthony N Imbalzano, and Jongbum Kwon. 2006. 'Mammalian SWI/SNF Complexes Facilitate DNA Double-Strand Break Repair by Promoting γ -H2AX Induction'. *The EMBO Journal* 25 (17): 3986–97. <https://doi.org/10.1038/sj.emboj.7601291>.
- Paulson, James R., Damien F. Hudson, Fernanda Cisneros-Soberanis, and William C. Earnshaw. 2021. 'Mitotic Chromosomes'. *Seminars in Cell & Developmental Biology, Mitosis*, 117 (September):7–29. <https://doi.org/10.1016/j.semcd.2021.03.014>.

- Paulson, James R., and U. K. Laemmli. 1977. 'The Structure of Histone-Depleted Metaphase Chromosomes'. *Cell* 12 (3): 817–28. [https://doi.org/10.1016/0092-8674\(77\)90280-X](https://doi.org/10.1016/0092-8674(77)90280-X).
- Paweletz, Neidhard. 2001. 'Walther Flemming: Pioneer of Mitosis Research'. *Nature Reviews Molecular Cell Biology* 2 (1): 72–75. <https://doi.org/10.1038/35048077>.
- Pederson, Thoru. 1972. 'Chromatin Structure and the Cell Cycle'. *Proceedings of the National Academy of Sciences* 69 (8): 2224–28. <https://doi.org/10.1073/pnas.69.8.2224>.
- Pedrali-Noy, G., S. Spadari, A. Miller-Faurès, A.O.A. Miller, J. Kruppa, and G. Koch. 1980. 'Synchronization of HeLa Cell Cultures by Inhibition of DNA Polymerase α with Aphidicolin'. *Nucleic Acids Research* 8 (2): 377–87. <https://doi.org/10.1093/nar/8.2.377>.
- Peters, Jan-Michael, and Tomoko Nishiyama. 2012. 'Sister Chromatid Cohesion'. *Cold Spring Harbor Perspectives in Biology* 4 (11): a011130. <https://doi.org/10.1101/cshperspect.a011130>.
- Piñol-Roma, Serafín, and Gideon Dreyfuss. 1991. 'Transcription-Dependent and Transcription-Independent Nuclear Transport of hnRNP Proteins'. *Science* 253 (5017): 312–14. <https://doi.org/10.1126/science.1857966>.
- Piskadlo, Ewa, Alexandra Tavares, and Raquel A Oliveira. n.d. 'Metaphase Chromosome Structure Is Dynamically Maintained by Condensin I-Directed DNA (de)Catenation'. *eLife* 6:e26120. <https://doi.org/10.7554/eLife.26120>.
- Poleshko, Andrey, Katelyn M. Mansfield, Caroline C. Burlingame, Mark D. Andrade, Neil R. Shah, and Richard A. Katz. 2013. 'The Human Protein PRR14 Tethers Heterochromatin to the Nuclear Lamina during Interphase and Mitotic Exit'. *Cell Reports* 5 (2): 292–301. <https://doi.org/10.1016/j.celrep.2013.09.024>.
- Pommier, Yves, André Nussenzweig, Shunichi Takeda, and Caroline Austin. 2022. 'Human Topoisomerases and Their Roles in Genome Stability and Organization'. *Nature Reviews Molecular Cell Biology* 23 (6): 407–27. <https://doi.org/10.1038/s41580-022-00452-3>.
- Prasanth, Kannanganattu V., Paula A. Sacco-Bubulya, Supriya G. Prasanth, and David L. Spector. 2003. 'Sequential Entry of Components of Gene Expression Machinery into Daughter Nuclei'. *Molecular Biology of the Cell* 14 (3): 1043–57. <https://doi.org/10.1091/mbc.E02-10-0669>.
- Puvvula, Pavan Kumar, Rohini Devi Desetty, Pascal Pineau, Agnès Marchio, Anne Moon, Anne Dejean, and Oliver Bischof. 2014. 'Long Noncoding RNA PANDA and Scaffold-Attachment-Factor SAFA Control Senescence Entry and Exit'. *Nature Communications* 5 (1): 5323. <https://doi.org/10.1038/ncomms6323>.

- Qi, Wenjing, Ruoxi Wang, Hongyu Chen, Xiaolin Wang, Ting Xiao, Istvan Boldogh, Xueqing Ba, Liping Han, and Xianlu Zeng. 2015. 'BRG1 Promotes the Repair of DNA Double-Strand Breaks by Facilitating the Replacement of RPA with RAD51'. *Journal of Cell Science* 128 (2): 317–30. <https://doi.org/10.1242/jcs.159103>.
- Qiu, Qi, Peng Hu, Xiaojie Qiu, Kiya W. Govek, Pablo G. Cámara, and Hao Wu. 2020. 'Massively Parallel and Time-Resolved RNA Sequencing in Single Cells with scNT-Seq'. *Nature Methods* 17 (10): 991–1001. <https://doi.org/10.1038/s41592-020-0935-4>.
- Rabani, Michal, Joshua Z. Levin, Lin Fan, Xian Adiconis, Raktima Raychowdhury, Manuel Garber, Andreas Gnirke, et al. 2011. 'Metabolic Labeling of RNA Uncovers Principles of RNA Production and Degradation Dynamics in Mammalian Cells'. *Nature Biotechnology* 29 (5): 436–42. <https://doi.org/10.1038/nbt.1861>.
- Ramadan, Kristijan, Roland Bruderer, Fabio M. Spiga, Oliver Popp, Tina Baur, Monica Gotta, and Hemmo H. Meyer. 2007. 'Cdc48/P97 Promotes Reformation of the Nucleus by Extracting the Kinase Aurora B from Chromatin'. *Nature* 450 (7173): 1258–62. <https://doi.org/10.1038/nature06388>.
- Rankin, Susannah. 2015. 'Complex Elaboration: Making Sense of Meiotic Cohesin Dynamics'. *The FEBS Journal* 282 (13): 2426–43. <https://doi.org/10.1111/febs.13301>.
- Rao, Suhas S. P., Su-Chen Huang, Brian Glenn St Hilaire, Jesse M. Engreitz, Elizabeth M. Perez, Kyong-Rim Kieffer-Kwon, Adrian L. Sanborn, et al. 2017. 'Cohesin Loss Eliminates All Loop Domains'. *Cell* 171 (2): 305–320.e24. <https://doi.org/10.1016/j.cell.2017.09.026>.
- Ray, Debashish, Hilal Kazan, Kate B. Cook, Matthew T. Weirauch, Hamed S. Najafabadi, Xiao Li, Serge Gueroussov, et al. 2013. 'A Compendium of RNA-Binding Motifs for Decoding Gene Regulation'. *Nature* 499 (7457): 172–77. <https://doi.org/10.1038/nature12311>.
- Raynaud, Cécile, Allison C. Mallory, David Latrasse, Teddy Jégu, Quentin Bruggeman, Marianne Delarue, Catherine Bergounioux, and Moussa Benhamed. 2014. 'Chromatin Meets the Cell Cycle'. *Journal of Experimental Botany* 65 (10): 2677–89. <https://doi.org/10.1093/jxb/ert433>.
- Régnier, Vinciane, Paola Vagnarelli, Tatsuo Fukagawa, Tatiana Zerjal, Elizabeth Burns, Didier Trouche, William Earnshaw, and William Brown. 2005. 'CENP-A Is Required for Accurate Chromosome Segregation and Sustained Kinetochore Association of BubR1'. *Molecular and Cellular Biology* 25 (10): 3967–81. <https://doi.org/10.1128/MCB.25.10.3967-3981.2005>.
- Ressler, Andrew K., Gabriela L. A. Sampaio, Sarah A. Dugger, Tamar Sapir, Daniel Krizay, Michael J. Boland, Orly Reiner, and David B. Goldstein. 2023. 'Evidence of Shared Transcriptomic Dysregulation of HNRNPU-Related Disorder between Human

- Organoids and Embryonic Mice'. *iScience* 26 (1).
<https://doi.org/10.1016/j.isci.2022.105797>.
- Rieder, Dietmar, Zlatko Trajanoski, and James McNally. 2012. 'Transcription Factories'. *Frontiers in Genetics* 3 (October). <https://doi.org/10.3389/fgene.2012.00221>.
- Riml, Christian, Thomas Amort, Dietmar Rieder, Catherina Gasser, Alexandra Lusser, and Ronald Micura. 2017. 'Osmium-Mediated Transformation of 4-Thiouridine to Cytidine as Key To Study RNA Dynamics by Sequencing'. *Angewandte Chemie International Edition* 56 (43): 13479–83. <https://doi.org/10.1002/anie.201707465>.
- Risca, Viviana I., Sarah K. Denny, Aaron F. Straight, and William J. Greenleaf. 2017. 'Variable Chromatin Structure Revealed by in Situ Spatially Correlated DNA Cleavage Mapping'. *Nature* 541 (7636): 237–41. <https://doi.org/10.1038/nature20781>.
- Robinson, Philip J. J., Louise Fairall, Van A. T. Huynh, and Daniela Rhodes. 2006. 'EM Measurements Define the Dimensions of the "30-Nm" Chromatin Fiber: Evidence for a Compact, Interdigitated Structure'. *Proceedings of the National Academy of Sciences* 103 (17): 6506–11. <https://doi.org/10.1073/pnas.0601212103>.
- Roden, Christine, and Amy S. Gladfelter. 2021. 'RNA Contributions to the Form and Function of Biomolecular Condensates'. *Nature Reviews Molecular Cell Biology* 22 (3): 183–95. <https://doi.org/10.1038/s41580-020-0264-6>.
- Rogakou, Emmy P., Chye Boon, Christophe Redon, and William M. Bonner. 1999. 'Megabase Chromatin Domains Involved in DNA Double-Strand Breaks in Vivo'. *Journal of Cell Biology* 146 (5): 905–16. <https://doi.org/10.1083/jcb.146.5.905>.
- Rogakou, Emmy P., Duane R. Pilch, Ann H. Orr, Vessela S. Ivanova, and William M. Bonner. 1998. 'DNA Double-Stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139 *'. *Journal of Biological Chemistry* 273 (10): 5858–68. <https://doi.org/10.1074/jbc.273.10.5858>.
- Romig, H., F.o. Fackelmayer, A. Renz, U. Ramsperger, and A. Richter. 1992. 'Characterization of SAF-A, a Novel Nuclear DNA Binding Protein from HeLa Cells with High Affinity for Nuclear Matrix/Scaffold Attachment DNA Elements.' *The EMBO Journal* 11 (9): 3431–40. <https://doi.org/10.1002/j.1460-2075.1992.tb05422.x>.
- Roshon, Michael J., and H. Earl Ruley. 2005. 'Hypomorphic Mutation in hnRNP U Results in Post-Implantation Lethality'. *Transgenic Research* 14 (2): 179–92. <https://doi.org/10.1007/s11248-004-8147-8>.
- Rossetto, Dorine, Nikita Avvakumov, and Jacques Côté. 2012. 'Histone Phosphorylation'. *Epigenetics* 7 (10): 1098–1108. <https://doi.org/10.4161/epi.21975>.
- Ryu, Je-Kyung, Allard J. Katan, Eli O. van der Sluis, Thomas Wisse, Ralph de Groot, Christian H. Haering, and Cees Dekker. 2020. 'The Condensin Holocomplex Cycles

Dynamically between Open and Collapsed States'. *Nature Structural & Molecular Biology* 27 (12): 1134–41. <https://doi.org/10.1038/s41594-020-0508-3>.

Saha, Sourav, Yilun Sun, Shar-yin Naomi Huang, Simone Andrea Baechler, Lorinc Sandor Pongor, Keli Agama, Ukhyun Jo, Hongliang Zhang, Yuk-Ching Tse-Dinh, and Yves Pommier. 2020. 'DNA and RNA Cleavage Complexes and Repair Pathway for TOP3B RNA- and DNA-Protein Crosslinks'. *Cell Reports* 33 (13): 108569. <https://doi.org/10.1016/j.celrep.2020.108569>.

Saitoh, N, I G Goldberg, E R Wood, and W C Earnshaw. 1994. 'ScII: An Abundant Chromosome Scaffold Protein Is a Member of a Family of Putative ATPases with an Unusual Predicted Tertiary Structure.' *Journal of Cell Biology* 127 (2): 303–18. <https://doi.org/10.1083/jcb.127.2.303>.

Sakaguchi, Takehisa, Yuko Hasegawa, Neil Brockdorff, Ken Tsutsui, Kimiko M. Tsutsui, Takashi Sado, and Shinichi Nakagawa. 2016. 'Control of Chromosomal Localization of Xist by hnRNP U Family Molecules'. *Developmental Cell* 39 (1): 11–12. <https://doi.org/10.1016/j.devcel.2016.09.022>.

Sakai, Yuji, Atsushi Mochizuki, Kazuhisa Kinoshita, Tatsuya Hirano, and Masashi Tachikawa. 2018. 'Modeling the Functions of Condensin in Chromosome Shaping and Segregation'. *PLOS Computational Biology* 14 (6): e1006152. <https://doi.org/10.1371/journal.pcbi.1006152>.

Salceda, Javier, Xavier Fernández, and Joaquim Roca. 2006. 'Topoisomerase II, Not Topoisomerase I, Is the Proficient Relaxase of Nucleosomal DNA'. *The EMBO Journal* 25 (11): 2575–83. <https://doi.org/10.1038/sj.emboj.7601142>.

Saldivar, Joshua C., Stephan Hamperl, Michael J. Bocek, Mingyu Chung, Thomas E. Bass, Fernanda Cisneros-Soberanis, Kumiko Samejima, et al. 2018. 'An Intrinsic S/G2 Checkpoint Enforced by ATR'. *Science* 361 (6404): 806–10. <https://doi.org/10.1126/science.aap9346>.

Salinas, Valeria, Nerina Martínez, Josefina Pérez Maturo, Sergio A. Rodríguez-Quiroga, Lucia Zavala, Nancy Medina, Hernán Amartino, et al. 2021. 'Clinical next Generation Sequencing in Developmental and Epileptic Encephalopathies: Diagnostic Relevance of Data Re-Analysis and Variants Re-Interpretation'. *European Journal of Medical Genetics* 64 (12): 104363. <https://doi.org/10.1016/j.ejmg.2021.104363>.

Samejima, Itaru, Christos Spanos, Kumiko Samejima, Juri Rappsilber, Georg Kustatscher, and William C. Earnshaw. 2022. 'Mapping the Invisible Chromatin Transactions of Prophase Chromosome Remodeling'. *Molecular Cell* 82 (3): 696-708.e4. <https://doi.org/10.1016/j.molcel.2021.12.039>.

Samejima, Kumiko, Daniel G. Booth, Hiromi Ogawa, James R. Paulson, Linfeng Xie, Cara A. Watson, Melpomeni Platani, Masato T. Kanemaki, and William C. Earnshaw. 2018. 'Functional Analysis after Rapid Degradation of Condensins and 3D-EM Reveals

Chromatin Volume Is Uncoupled from Chromosome Architecture in Mitosis'. *Journal of Cell Science* 131 (4): jcs210187. <https://doi.org/10.1242/jcs.210187>.

Sanborn, Adrian L., Suhas S. P. Rao, Su-Chen Huang, Neva C. Durand, Miriam H. Huntley, Andrew I. Jewett, Ivan D. Bochkov, et al. 2015. 'Chromatin Extrusion Explains Key Features of Loop and Domain Formation in Wild-Type and Engineered Genomes'. *Proceedings of the National Academy of Sciences* 112 (47): E6456–65. <https://doi.org/10.1073/pnas.1518552112>.

Sandman, Kathleen, and John Reeve. 2000. 'Structure and Functional Relationships of Archaeal and Eukaryal Histones and Nucleosomes'. *Archives of Microbiology* 173 (3): 165–69. <https://doi.org/10.1007/s002039900122>.

Saponaro, Marco, Theodoros Kantidakis, Richard Mitter, Gavin P. Kelly, Mark Heron, Hannah Williams, Johannes Söding, Aengus Stewart, and Jesper Q. Svejstrup. 2014. 'RECQL5 Controls Transcript Elongation and Suppresses Genome Instability Associated with Transcription Stress'. *Cell* 157 (5): 1037–49. <https://doi.org/10.1016/j.cell.2014.03.048>.

Satterstrom, F. Kyle, Jack A. Kosmicki, Jiebiao Wang, Michael S. Breen, Silvia De Rubeis, Joon-Yong An, Minshi Peng, et al. 2020. 'Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism'. *Cell* 180 (3): 568-584.e23. <https://doi.org/10.1016/j.cell.2019.12.036>.

Schalch, Thomas, Sylwia Duda, David F. Sargent, and Timothy J. Richmond. 2005. 'X-Ray Structure of a Tetranucleosome and Its Implications for the Chromatin Fibre'. *Nature* 436 (7047): 138–41. <https://doi.org/10.1038/nature03686>.

Schellhaus, Anna Katharina. 2017. 'The GTPase DRG1 Controls Microtubule Dynamics and Is Involved in Chromatin Decondensation'. <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77250?show=full>.

Schmiedeberg, Lars, Pete Skene, Aimée Deaton, and Adrian Bird. 2009. 'A Temporal Threshold for Formaldehyde Crosslinking and Fixation'. *PLOS ONE* 4 (2): e4636. <https://doi.org/10.1371/journal.pone.0004636>.

Schmitz, Michael H.A., Michael Held, Veerle Janssens, James R.A. Hutchins, Otto Hudecz, Elitsa Ivanova, Jozef Goris, et al. 2010. 'Live-Cell Imaging RNAi Screen Identifies PP2A-B55 α and Importin- β 21 as Key Mitotic Exit Regulators in Human Cells'. *Nature Cell Biology* 12 (9): 886–93. <https://doi.org/10.1038/ncb2092>.

Schofield, Jeremy A., Erin E. Duffy, Lea Kiefer, Meaghan C. Sullivan, and Matthew D. Simon. 2018. 'TimeLapse-Seq: Adding a Temporal Dimension to RNA Sequencing through Nucleoside Recoding'. *Nature Methods* 15 (3): 221–25. <https://doi.org/10.1038/nmeth.4582>.

- Sekiya, Takeshi, Kensaku Murano, Kohsuke Kato, Atsushi Kawaguchi, and Kyosuke Nagata. 2017. 'Mitotic Phosphorylation of CCCTC-Binding Factor (CTCF) Reduces Its DNA Binding Activity'. *FEBS Open Bio* 7 (3): 397–404. <https://doi.org/10.1002/2211-5463.12189>.
- Sexton, Tom, Eitan Yaffe, Ephraim Kenigsberg, Frédéric Bantignies, Benjamin Leblanc, Michael Hoichman, Hugues Parrinello, Amos Tanay, and Giacomo Cavalli. 2012. 'Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome'. *Cell* 148 (3): 458–72. <https://doi.org/10.1016/j.cell.2012.01.010>.
- Shao, Wen, Xianju Bi, Yixuan Pan, Boyang Gao, Jun Wu, Yafei Yin, Zhimin Liu, et al. 2022. 'Phase Separation of RNA-Binding Protein Promotes Polymerase Binding and Transcription'. *Nature Chemical Biology* 18 (1): 70–80. <https://doi.org/10.1038/s41589-021-00904-5>.
- Sharp, Judith A., Carlos Perea-Resa, Wei Wang, and Michael D. Blower. 2020. 'Cell Division Requires RNA Eviction from Condensing Chromosomes'. *Journal of Cell Biology* 219 (11). <https://doi.org/10.1083/JCB.201910148>.
- Shepard, Samuel, Mark McCreary, and Alexei Fedorov. 2009. 'The Peculiarities of Large Intron Splicing in Animals'. *PLOS ONE* 4 (11): e7853. <https://doi.org/10.1371/journal.pone.0007853>.
- Sherr, C. J., and J. M. Roberts. 1999. 'CDK Inhibitors: Positive and Negative Regulators of G1-Phase Progression'. *Genes & Development* 13 (12): 1501–12. <https://doi.org/10.1101/gad.13.12.1501>.
- Shibahara, Kei-ichi, and Bruce Stillman. 1999. 'Replication-Dependent Marking of DNA by PCNA Facilitates CAF-1-Coupled Inheritance of Chromatin'. *Cell* 96 (4): 575–85. [https://doi.org/10.1016/S0092-8674\(00\)80661-3](https://doi.org/10.1016/S0092-8674(00)80661-3).
- Shibata, Akira, Kiyo Morita, Takuya Yamashita, Hiroshi Kamaya, and Issaku Ueda. 1991. 'Anesthetic-Protein Interaction: Effects of Volatile Anesthetics on the Secondary Structure of Poly(L-Lysine)'. *Journal of Pharmaceutical Sciences* 80 (11): 1037–41. <https://doi.org/10.1002/jps.2600801108>.
- Shimada, Shino, Hirokazu Oguni, Yui Otani, Aiko Nishikawa, Susumu Ito, Kaoru Eto, Tomoyuki Nakazawa, et al. 2018. 'An Episode of Acute Encephalopathy with Biphasic Seizures and Late Reduced Diffusion Followed by Hemiplegia and Intractable Epilepsy Observed in a Patient with a Novel Frameshift Mutation in HNRNPU'. *Brain and Development* 40 (9): 813–18. <https://doi.org/10.1016/j.braindev.2018.05.010>.
- Shintomi, Keishi, and Tatsuya Hirano. 2011. 'The Relative Ratio of Condensin I to II Determines Chromosome Shapes'. *Genes & Development* 25 (14): 1464–69. <https://doi.org/10.1101/gad.2060311>.

- Shintomi, Keishi, Tatsuro S. Takahashi, and Tatsuya Hirano. 2015. 'Reconstitution of Mitotic Chromatids with a Minimum Set of Purified Factors'. *Nature Cell Biology* 17 (8): 1014–23. <https://doi.org/10.1038/ncb3187>.
- Shopland, Lindsay S., Carol V. Johnson, and Jeanne B. Lawrence. 2002. 'Evidence That All SC-35 Domains Contain mRNAs and That Transcripts Can Be Structurally Constrained within These Domains'. *Journal of Structural Biology* 140 (1): 131–39. [https://doi.org/10.1016/S1047-8477\(02\)00507-5](https://doi.org/10.1016/S1047-8477(02)00507-5).
- Shroff, Robert, Ayelet Arbel-Eden, Duane Pilch, Grzegorz Ira, William M. Bonner, John H. Petrini, James E. Haber, and Michael Lichten. 2004. 'Distribution and Dynamics of Chromatin Modification Induced by a Defined DNA Double-Strand Break'. *Current Biology* 14 (19): 1703–11. <https://doi.org/10.1016/j.cub.2004.09.047>.
- Singh, Jarnail, and Richard A. Padgett. 2009. 'Rates of in Situ Transcription and Splicing in Large Human Genes'. *Nature Structural & Molecular Biology* 16 (11): 1128–33. <https://doi.org/10.1038/nsmb.1666>.
- Siomi, H, and G Dreyfuss. 1995. 'A Nuclear Localization Domain in the hnRNP A1 Protein.' *Journal of Cell Biology* 129 (3): 551–60. <https://doi.org/10.1083/jcb.129.3.551>.
- Skalska, Lenka, Victoria Begley, Manuel Beltran, Saulius Lukauskas, Garima Khandelwal, Peter Faull, Amandeep Bhamra, et al. 2021. 'Nascent RNA Antagonizes the Interaction of a Set of Regulatory Proteins with Chromatin'. *Molecular Cell* 81 (14): 2944-2959.e10. <https://doi.org/10.1016/j.molcel.2021.05.026>.
- Skibbens, Robert V. 2016. 'Of Rings and Rods: Regulating Cohesin Entrapment of DNA to Generate Intra- and Intermolecular Tethers'. *PLOS Genetics* 12 (10): e1006337. <https://doi.org/10.1371/journal.pgen.1006337>.
- Smeets, Daniel, Yolanda Markaki, Volker J. Schmid, Felix Kraus, Anna Tattermusch, Andrea Cerase, Michael Sterr, et al. 2014. 'Three-Dimensional Super-Resolution Microscopy of the Inactive X Chromosome Territory Reveals a Collapse of Its Active Nuclear Compartment Harboring Distinct Xist RNA Foci'. *Epigenetics & Chromatin* 7 (1): 8. <https://doi.org/10.1186/1756-8935-7-8>.
- Snezhkina, Anastasiya Vladimirovna, George Sergeevich Krasnov, Andrew Rostislavovich Zaretsky, Alex Zhavoronkov, Kirill Mikhailovich Nyushko, Alexey Alexandrovich Moskalev, Irina Yurievna Karpova, et al. 2016. 'Differential Expression of Alternatively Spliced Transcripts Related to Energy Metabolism in Colorectal Cancer'. *BMC Genomics* 17 (Suppl 14): 1011. <https://doi.org/10.1186/s12864-016-3351-5>.
- Solomon, David A., Taeyeon Kim, Laura A. Diaz-Martinez, Joshlean Fair, Abdel G. Elkahloun, Brent T. Harris, Jeffrey A. Toretsky, et al. 2011. 'Mutational Inactivation of STAG2 Causes Aneuploidy in Human Cancer'. *Science* 333 (6045): 1039–43. <https://doi.org/10.1126/science.1203619>.

- Song, Zhenfeng, Ying Zhang, Chengqing Yang, Zhi Yi, Fei Li, Jiao Xue, Xiaofan Yang, and Baomin Li. 2021. 'De Novo Frameshift Variants of HNRNPU in Patients with Early Infantile Epileptic Encephalopathy: Two Case Reports and Literature Review'. *International Journal of Developmental Neuroscience* 81 (7): 663–68. <https://doi.org/10.1002/jdn.10115>.
- Soufi, Abdenour, and Stephen Dalton. 2016. 'Cycling through Developmental Decisions: How Cell Cycle Dynamics Control Pluripotency, Differentiation and Reprogramming'. *Development* 143 (23): 4301–11. <https://doi.org/10.1242/dev.142075>.
- Spagnoli, Carlotta, Susanna Rizzi, Grazia Gabriella Salerno, Daniele Frattini, Juha Koskenvuo, and Carlo Fusco. 2021. 'Pharmacological Treatment of Severe Breathing Abnormalities in a Case of HNRNPU Epileptic Encephalopathy'. *Molecular Syndromology* 12 (2): 101–5. <https://doi.org/10.1159/000512566>.
- Spencer, Sabrina L., Steven D. Cappell, Feng-Chiao Tsai, K. Wesley Overton, Clifford L. Wang, and Tobias Meyer. 2013. 'The Proliferation-Quiescence Decision Is Controlled by a Bifurcation in CDK2 Activity at Mitotic Exit'. *Cell* 155 (2): 369–83. <https://doi.org/10.1016/j.cell.2013.08.062>.
- Stallaert, Wayne, Katarzyna M. Kedziora, Colin D. Taylor, Tarek M. Zikry, Jolene S. Ranek, Holly K. Sobon, Sovanny R. Taylor, Catherine L. Young, Jeanette G. Cook, and Jeremy E. Purvis. 2022. 'The Structure of the Human Cell Cycle'. *Cell Systems* 13 (3): 230-240.e3. <https://doi.org/10.1016/j.cels.2021.10.007>.
- Strodel, Birgit. 2021. 'Energy Landscapes of Protein Aggregation and Conformation Switching in Intrinsically Disordered Proteins'. *Journal of Molecular Biology*, From Protein Sequence to Structure at Warp Speed: How AlphaFold Impacts Biology, 433 (20): 167182. <https://doi.org/10.1016/j.jmb.2021.167182>.
- Stucki, Manuel, Julie A. Clapperton, Duaa Mohammad, Michael B. Yaffe, Stephen J. Smerdon, and Stephen P. Jackson. 2005. 'MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks'. *Cell* 123 (7): 1213–26. <https://doi.org/10.1016/j.cell.2005.09.038>.
- Takagi, Masatoshi, Takao Ono, Toyoaki Natsume, Chiyomi Sakamoto, Mitsuyoshi Nakao, Noriko Saitoh, Masato T. Kanemaki, Tatsuya Hirano, and Naoko Imamoto. 2018. 'Ki-67 and Condensins Support the Integrity of Mitotic Chromosomes through Distinct Mechanisms'. *Journal of Cell Science* 131 (6): jcs212092. <https://doi.org/10.1242/jcs.212092>.
- Tamaru, Teruya, Yasushi Isojima, Katsuya Nagai, and Ken Takamatsu. 2003. 'Circadian Expression of hnRNP U, a Nuclear Multi-Potent Regulatory Protein, in the Murine Suprachiasmatic Nucleus'. *Neuroscience Letters* 341 (2): 111–14. [https://doi.org/10.1016/S0304-3940\(03\)00181-2](https://doi.org/10.1016/S0304-3940(03)00181-2).

- Tatton-Brown, Katrina, Chey Loveday, Shawn Yost, Matthew Clarke, Emma Ramsay, Anna Zachariou, Anna Elliott, et al. 2017. 'Mutations in Epigenetic Regulation Genes Are a Major Cause of Overgrowth with Intellectual Disability'. *The American Journal of Human Genetics* 100 (5): 725–36. <https://doi.org/10.1016/j.ajhg.2017.03.010>.
- Taylor, James, Michael Spiller, Kara Ranguin, Antonio Vitobello, Christophe Philippe, Ange-Line Bruel, Gerarda Cappuccio, et al. 2022. 'Expanding the Phenotype of HNRNPU-Related Neurodevelopmental Disorder with Emphasis on Seizure Phenotype and Review of Literature'. *American Journal of Medical Genetics Part A* 188 (5): 1497–1514. <https://doi.org/10.1002/ajmg.a.62677>.
- Tedeschi, Antonio, Gordana Wutz, Sébastien Huet, Markus Jaritz, Annelie Wuensche, Erika Schirghuber, Iain Finley Davidson, et al. 2013. 'Wapl Is an Essential Regulator of Chromatin Structure and Chromosome Segregation'. *Nature* 501 (7468): 564–68. <https://doi.org/10.1038/nature12471>.
- Terme, Jean-Michel, Borja Sesé, Lluís Millán-Ariño, Regina Mayor, Juan Carlos Izpisua Belmonte, María José Barrero, and Albert Jordan. 2011. 'Histone H1 Variants Are Differentially Expressed and Incorporated into Chromatin during Differentiation and Reprogramming to Pluripotency *'. *Journal of Biological Chemistry* 286 (41): 35347–57. <https://doi.org/10.1074/jbc.M111.281923>.
- Thandapani, Palaniraja, Timothy R. O'Connor, Timothy L. Bailey, and Stéphane Richard. 2013. 'Defining the RGG/RG Motif'. *Molecular Cell* 50 (5): 613–23. <https://doi.org/10.1016/j.molcel.2013.05.021>.
- The UniProt Consortium. 2023. 'UniProt: The Universal Protein Knowledgebase in 2023'. *Nucleic Acids Research* 51 (D1): D523–31. <https://doi.org/10.1093/nar/gkac1052>.
- Till, Susanne, and Andreas G. Ladurner. 2009. 'Sensing NAD Metabolites through Macro Domains'. *Frontiers in Bioscience-Landmark* 14 (9): 3246–58. <https://doi.org/10.2741/3448>.
- Tomecki, Rafal, Maiken S Kristiansen, Søren Lykke-Andersen, Aleksander Chlebowski, Katja M Larsen, Roman J Szczesny, Karolina Drazkowska, et al. 2010. 'The Human Core Exosome Interacts with Differentially Localized Processive RNases: hDIS3 and hDIS3L'. *The EMBO Journal* 29 (14): 2342–57. <https://doi.org/10.1038/emboj.2010.121>.
- Tsaneva, Irina R., Berndt Müller, and Stephen C. West. 1992. 'ATP-Dependent Branch Migration of Holliday Junctions Promoted by the RuvA and RuvB Proteins of E. Coli'. *Cell* 69 (7): 1171–80. [https://doi.org/10.1016/0092-8674\(92\)90638-S](https://doi.org/10.1016/0092-8674(92)90638-S).
- Tsurusaki, Yoshinori, Nobuhiko Okamoto, Hirofumi Ohashi, Tomoki Kosho, Yoko Imai, Yumiko Hibi-Ko, Tadashi Kaname, et al. 2012. 'Mutations Affecting Components of the SWI/SNF Complex Cause Coffin-Siris Syndrome'. *Nature Genetics* 44 (4): 376–78. <https://doi.org/10.1038/ng.2219>.

- Tuduri, Sandie, Laure Crabbé, Chiara Conti, Hélène Tourrière, Heidi Holtgreve-Grez, Anna Jauch, Véronique Pantescio, et al. 2009. 'Topoisomerase I Suppresses Genomic Instability by Preventing Interference between Replication and Transcription'. *Nature Cell Biology* 11 (11): 1315–24. <https://doi.org/10.1038/ncb1984>.
- Turowski, Tomasz W., and David Tollervey. 2015. 'Cotranscriptional Events in Eukaryotic Ribosome Synthesis'. *WIREs RNA* 6 (1): 129–39. <https://doi.org/10.1002/wrna.1263>.
- Vagnarelli, Paola. 2012. 'Mitotic Chromosome Condensation in Vertebrates'. *Experimental Cell Research* 318 (12): 1435–41. <https://doi.org/10.1016/j.yexcr.2012.03.017>.
- Vagnarelli, Paola, Damien F. Hudson, Susana A. Ribeiro, Laura Trinkle-Mulcahy, Jennifer M. Spence, Fan Lai, Christine J. Farr, Angus I. Lamond, and William C. Earnshaw. 2006. 'Condensin and Repo-Man–PP1 Co-Operate in the Regulation of Chromosome Architecture during Mitosis'. *Nature Cell Biology* 8 (10): 1133–42. <https://doi.org/10.1038/ncb1475>.
- Valverde, Roberto, Laura Edwards, and Lynne Regan. 2008. 'Structure and Function of KH Domains'. *The FEBS Journal* 275 (11): 2712–26. <https://doi.org/10.1111/j.1742-4658.2008.06411.x>.
- Van Nostrand, Eric L., Peter Freese, Gabriel A. Pratt, Xiaofeng Wang, Xintao Wei, Rui Xiao, Steven M. Blue, et al. 2020. 'A Large-Scale Binding and Functional Map of Human RNA-Binding Proteins'. *Nature* 583 (7818): 711–19. <https://doi.org/10.1038/s41586-020-2077-3>.
- Vassilev, Lyubomir T. 2006. 'Cell Cycle Synchronization at the G2/M Phase Border by Reversible Inhibition of CDK1'. *Cell Cycle* 5 (22): 2555–56. <https://doi.org/10.4161/cc.5.22.3463>.
- Viryasova, Galina M., Victor V. Tatarskiy, Andrey A. Sheynov, Eugene V. Tatarskiy, Galina F. Sud'ina, Sofia G. Georgieva, and Nataliya V. Soshnikova. 2019. 'PBAF Lacking PHD Domains Maintains Transcription in Human Neutrophils'. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1866 (12): 118525. <https://doi.org/10.1016/j.bbamcr.2019.118525>.
- Vizlin-Hodzic, Dzeneta, Rikard Runnberg, Jessica Ryme, Stina Simonsson, and Tomas Simonsson. 2011. 'SAF-A Forms a Complex with BRG1 and Both Components Are Required for RNA Polymerase II Mediated Transcription'. *PLOS ONE* 6 (12): e28049. <https://doi.org/10.1371/journal.pone.0028049>.
- Wäldchen, Sina, Julian Lehmann, Teresa Klein, Sebastian van de Linde, and Markus Sauer. 2015. 'Light-Induced Cell Damage in Live-Cell Super-Resolution Microscopy'. *Scientific Reports* 5 (1): 15348. <https://doi.org/10.1038/srep15348>.
- Walter, Joachim, Lothar Schermelleh, Marion Cremer, Satoshi Tashiro, and Thomas Cremer. 2003. 'Chromosome Order in HeLa Cells Changes during Mitosis and Early

G1, but Is Stably Maintained during Subsequent Interphase Stages'. *Journal of Cell Biology* 160 (5): 685–97. <https://doi.org/10.1083/jcb.200211103>.

Wang, Jianming, Patricia Rojas, Jingwen Mao, Martina Mustè Sadurni, Olivia Garnier, Songshu Xiao, Martin R. Higgs, Paloma Garcia, and Marco Saponaro. 2021. 'Persistence of RNA Transcription during DNA Replication Delays Duplication of Transcription Start Sites until G2/M'. *Cell Reports* 34 (7). <https://doi.org/10.1016/j.celrep.2021.108759>.

Wang, Tao, Steven M Bray, and Stephen T Warren. 2012. 'New Perspectives on the Biology of Fragile X Syndrome'. *Current Opinion in Genetics & Development*, Molecular and genetic bases of disease, 22 (3): 256–63. <https://doi.org/10.1016/j.gde.2012.02.002>.

Wang, Tianyun, Kendra Hoekzema, Davide Vecchio, Huidan Wu, Arvis Sulovari, Bradley P. Coe, Madelyn A. Gillentine, et al. 2020. 'Large-Scale Targeted Sequencing Identifies Risk Genes for Neurodevelopmental Disorders'. *Nature Communications* 11 (1). <https://doi.org/10.1038/s41467-020-18723-y>.

Wang, Xiaofeng, Ryan S. Lee, Burak H. Alver, Jeffrey R. Haswell, Su Wang, Jakub Mieczkowski, Yotam Drier, et al. 2017. 'SMARCB1-Mediated SWI/SNF Complex Function Is Essential for Enhancer Regulation'. *Nature Genetics* 49 (2): 289–95. <https://doi.org/10.1038/ng.3746>.

Wei, Yi, Lanlan Yu, Josephine Bowen, Martin A Gorovsky, and C. David Allis. 1999a. 'Phosphorylation of Histone H3 Is Required for Proper Chromosome Condensation and Segregation'. *Cell* 97 (1): 99–109. [https://doi.org/10.1016/S0092-8674\(00\)80718-7](https://doi.org/10.1016/S0092-8674(00)80718-7).

———. 1999b. 'Phosphorylation of Histone H3 Is Required for Proper Chromosome Condensation and Segregation'. *Cell* 97 (1): 99–109. [https://doi.org/10.1016/S0092-8674\(00\)80718-7](https://doi.org/10.1016/S0092-8674(00)80718-7).

Weidtkamp-Peters, Stefanie, Thorsten Lenser, Dmitri Negorev, Norman Gerstner, Thomas G. Hofmann, Georg Schwanitz, Christian Hoischen, Gerd Maul, Peter Dittrich, and Peter Hemmerich. 2008. 'Dynamics of Component Exchange at PML Nuclear Bodies'. *Journal of Cell Science* 121 (16): 2731–43. <https://doi.org/10.1242/jcs.031922>.

Whelan, Gabriela, Emanuel Kreidl, Gordana Wutz, Alexander Egner, Jan-Michael Peters, and Gregor Eichele. 2012. 'Cohesin Acetyltransferase Esco2 Is a Cell Viability Factor and Is Required for Cohesion in Pericentric Heterochromatin'. *The EMBO Journal* 31 (1): 71–82. <https://doi.org/10.1038/emboj.2011.381>.

Wilkins, Bryan J., Nils A. Rall, Yogesh Ostwal, Tom Kruitwagen, Kyoko Hiragami-Hamada, Marco Winkler, Yves Barral, Wolfgang Fischle, and Heinz Neumann. 2014. 'A

Cascade of Histone Modifications Induces Chromatin Condensation in Mitosis'. *Science* 343 (6166): 77–80. <https://doi.org/10.1126/science.1244508>.

Wong, Madeline M., Linda K. Cox, and John C. Chrivia. 2007. 'The Chromatin Remodeling Protein, SRCAP, Is Critical for Deposition of the Histone Variant H2A.Z at Promoters *'. *Journal of Biological Chemistry* 282 (36): 26132–39. <https://doi.org/10.1074/jbc.M703418200>.

Woodcock, C L, L L Frado, and J B Rattner. 1984. 'The Higher-Order Structure of Chromatin: Evidence for a Helical Ribbon Arrangement.' *Journal of Cell Biology* 99 (1): 42–52. <https://doi.org/10.1083/jcb.99.1.42>.

Worcel, A, S Strogatz, and D Riley. 1981. 'Structure of Chromatin and the Linking Number of DNA.' *Proceedings of the National Academy of Sciences* 78 (3): 1461–65. <https://doi.org/10.1073/pnas.78.3.1461>.

Wu, Chenyi, Andrew Bassett, and Andrew Travers. 2007. 'A Variable Topology for the 30-Nm Chromatin Fibre'. *EMBO Reports* 8 (12): 1129–34. <https://doi.org/10.1038/sj.embor.7401115>.

Wu, Chunlei, Xuefeng Jin, Ginger Tsueng, Cyrus Afrasiabi, and Andrew I. Su. 2016. 'BioGPS: Building Your Own Mash-up of Gene Annotations and Expression Profiles'. *Nucleic Acids Research* 44 (D1): D313–16. <https://doi.org/10.1093/nar/gkv1104>.

Xiao, Andrew, Haitao Li, David Shechter, Sung Hee Ahn, Laura A. Fabrizio, Hediye Erdjument-Bromage, Satoko Ishibe-Murakami, et al. 2009. 'WSTF Regulates the H2A.X DNA Damage Response via a Novel Tyrosine Kinase Activity'. *Nature* 457 (7225): 57–62. <https://doi.org/10.1038/nature07668>.

Xiao, Rui, Jia-Yu Chen, Zhengyu Liang, Daji Luo, Geng Chen, Zhi John Lu, Yang Chen, et al. 2019. 'Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription'. *Cell* 178 (1): 107-121.e18. <https://doi.org/10.1016/j.cell.2019.06.001>.

Xiao, Rui, Peng Tang, Bo Yang, Jie Huang, Yu Zhou, Changwei Shao, Hairi Li, Hui Sun, Yi Zhang, and Xiang Dong Fu. 2012. 'Nuclear Matrix Factor hnRNP U/SAF-A Exerts a Global Control of Alternative Splicing by Regulating U2 snRNP Maturation'. *Molecular Cell* 45 (5): 656–68. <https://doi.org/10.1016/j.molcel.2012.01.009>.

Xicoy, Helena, Bé Wieringa, and Gerard J.M. Martens. 2017. 'The SH-SY5Y Cell Line in Parkinson's Disease Research: A Systematic Review'. *Molecular Neurodegeneration* 12 (1): 10. <https://doi.org/10.1186/s13024-017-0149-0>.

Yamazaki, Tomohiro, Sylvie Souquere, Takeshi Chujo, Simon Kobelke, Yee Seng Chong, Archa H. Fox, Charles S. Bond, Shinichi Nakagawa, Gerard Pierron, and Tetsuro Hirose. 2018. 'Functional Domains of NEAT1 Architectural lncRNA Induce

- Paraspeckle Assembly through Phase Separation'. *Molecular Cell* 70 (6): 1038-1053.e7. <https://doi.org/10.1016/j.molcel.2018.05.019>.
- Yan, Wei, Lang Ma, Kathleen H. Burns, and Martin M. Matzuk. 2003. 'HILS1 Is a Spermatid-Specific Linker Histone H1-like Protein Implicated in Chromatin Remodeling during Mammalian Spermiogenesis'. *Proceedings of the National Academy of Sciences* 100 (18): 10546–51. <https://doi.org/10.1073/pnas.1837812100>.
- Yan, Youyou, Dandan Song, Xianjing Song, and Chunli Song. 2020. 'The Role of lncRNA MALAT1 in Cardiovascular Disease'. *IUBMB Life* 72 (3): 334–42. <https://doi.org/10.1002/iub.2210>.
- Yang, Edward, Erik van Nimwegen, Mihaela Zavolan, Nikolaus Rajewsky, Mark Schroeder, Marcelo Magasco, and James E. Darnell. 2003. 'Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes'. *Genome Research* 13 (8): 1863–72. <https://doi.org/10.1101/gr.1272403>.
- Yariv, Barak, Elon Yariv, Amit Kessel, Gal Masrati, Adi Ben Chorin, Eric Martz, Itay Mayrose, Tal Pupko, and Nir Ben-Tal. 2023. 'Using Evolutionary Data to Make Sense of Macromolecules with a "Face-Lifted" ConSurf'. *Protein Science* 32 (3): e4582. <https://doi.org/10.1002/pro.4582>.
- Yasukawa, Kiyoshi, Kei Iida, Hiroyuki Okano, Ryota Hidese, Misato Baba, Itaru Yanagihara, Kenji Kojima, Teisuke Takita, and Shinsuke Fujiwara. 2017. 'Next-Generation Sequencing-Based Analysis of Reverse Transcriptase Fidelity'. *Biochemical and Biophysical Research Communications* 492 (2): 147–53. <https://doi.org/10.1016/j.bbrc.2017.07.169>.
- Yates, T. Michael, Pradeep C. Vasudevan, Kate E. Chandler, Deirdre E Donnelly, Zornitza Stark, Simon Sadedin, Josh Willoughby, and Meena Balasubramanian. 2017. 'De Novo Mutations in HNRNPU Result in a Neurodevelopmental Syndrome'. *American Journal of Medical Genetics, Part A* 173 (11): 3003–12. <https://doi.org/10.1002/ajmg.a.38492>.
- Ye, Bingyu, Wenlong Shen, Yanchang Li, Dong Wang, Yan Zhang, Ping Li, Man Yin, et al. 2023. 'FAIRE-MS Reveals Mitotic Retention of Transcriptional Regulators on a Proteome-Wide Scale'. *The FASEB Journal* 37 (2): e22724. <https://doi.org/10.1096/fj.202201038RRR>.
- Ye, Junqiang, Nadine Beetz, Sean O'Keeffe, Juan Carlos Tapia, Lindsey Macpherson, Weisheng V. Chen, Rhonda Bassel-Duby, Eric N. Olson, and Tom Maniatis. 2015. 'hnRNP U Protein Is Required for Normal Pre-mRNA Splicing and Postnatal Heart Development and Function'. *Proceedings of the National Academy of Sciences* 112 (23): E3020–29. <https://doi.org/10.1073/pnas.1508461112>.
- Yeong, Foong May, Hans Hombauer, Kerstin S. Wendt, Toru Hirota, Ingrid Mudrak, Karl Mechtler, Thomas Loregger, et al. 2003. 'Identification of a Subunit of a Novel Kleisin-

β /SMC Complex as a Potential Substrate of Protein Phosphatase 2A'. *Current Biology* 13 (23): 2058–64. <https://doi.org/10.1016/j.cub.2003.10.032>.

Yesbolatova, Aisha, Toyoaki Natsume, Ken-ichiro Hayashi, and Masato T. Kanemaki. 2019. 'Generation of Conditional Auxin-Inducible Degron (AID) Cells and Tight Control of Degron-Fused Proteins Using the Degradation Inhibitor Auxinole'. *Methods, New Methods for Extracting Function from the Mammalian Genome*, 164–165 (July):73–80. <https://doi.org/10.1016/j.ymeth.2019.04.010>.

Yesbolatova, Aisha, Yuichiro Saito, Naomi Kitamoto, Hatsune Makino-Itou, Rieko Ajima, Risako Nakano, Hirofumi Nakaoka, et al. 2020. 'The Auxin-Inducible Degron 2 Technology Provides Sharp Degradation Control in Yeast, Mammalian Cells, and Mice'. *Nature Communications* 11 (1): 5701. <https://doi.org/10.1038/s41467-020-19532-z>.

Yugami, Masato, Hideyuki Okano, Atsushi Nakanishi, and Masato Yano. 2020. 'Analysis of the Nucleocytoplasmic Shuttling RNA-Binding Protein HNRNPU Using Optimized HiTS-CLIP Method'. Edited by Thomas Preiss. *PLoS ONE* 15 (4): e0231450. <https://doi.org/10.1371/journal.pone.0231450>.

Yusufzai, Timur M., Hideaki Tagami, Yoshihiro Nakatani, and Gary Felsenfeld. 2004. 'CTCF Tethers an Insulator to Subnuclear Sites, Suggesting Shared Insulator Mechanisms across Species'. *Molecular Cell* 13 (2): 291–98. [https://doi.org/10.1016/S1097-2765\(04\)00029-2](https://doi.org/10.1016/S1097-2765(04)00029-2).

Zhang, Fajun, Roland Roth, Marcell Wolf, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Stzucki, and Frank Schreiber. 2012. 'Charge-Controlled Metastable Liquid–Liquid Phase Separation in Protein Solutions as a Universal Pathway towards Crystallization'. *Soft Matter* 8 (5): 1313–16. <https://doi.org/10.1039/C2SM07008A>.

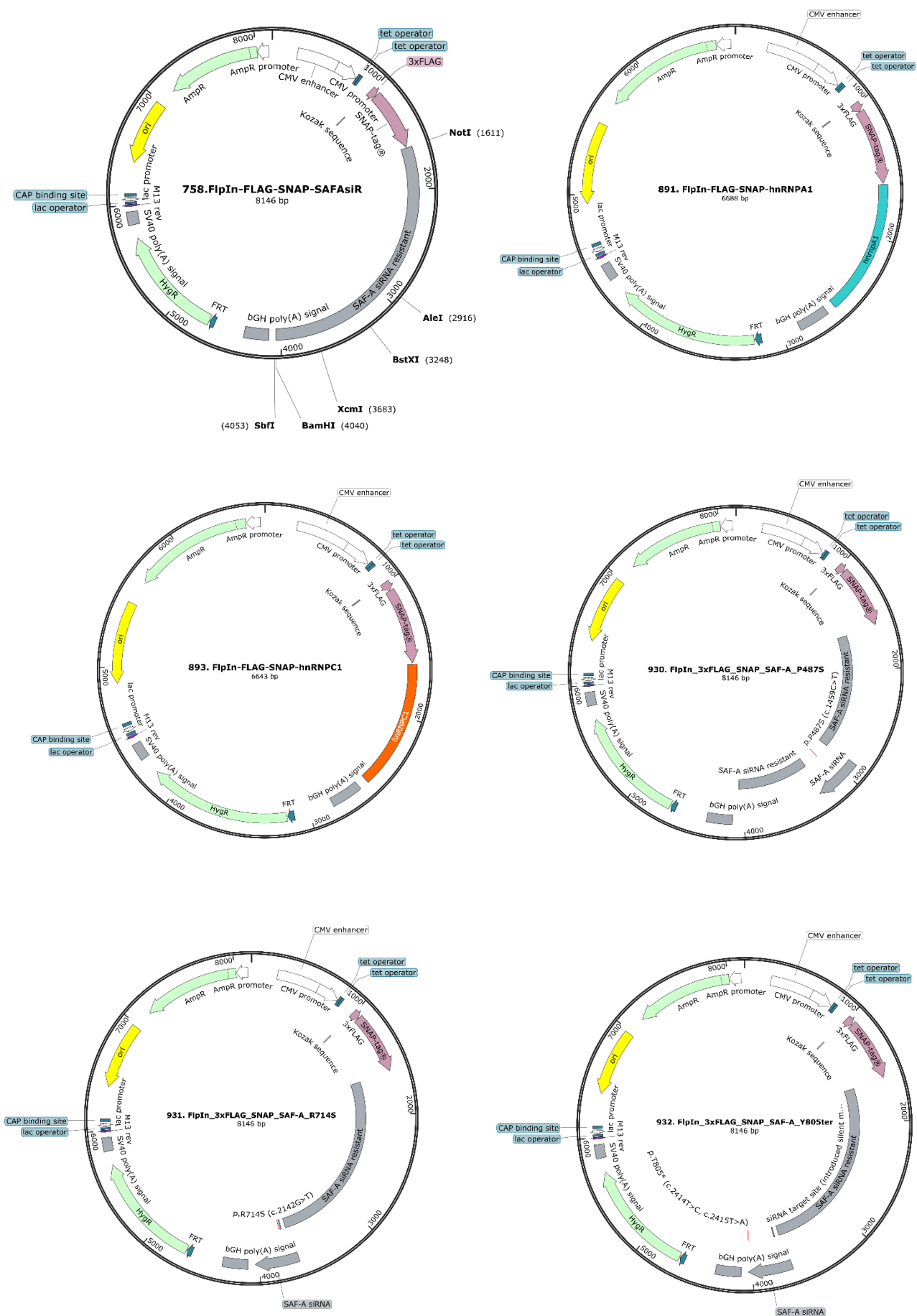
Zhang, Haoyue, Daniel J. Emerson, Thomas G. Gilgenast, Katelyn R. Titus, Yemin Lan, Peng Huang, Di Zhang, et al. 2019. 'Chromatin Structure Dynamics during the Mitosis-to-G1 Phase Transition'. *Nature* 576 (7785): 158–62. <https://doi.org/10.1038/s41586-019-1778-y>.

Zhang, Jinglan, Xiaomin Shi, Yehua Li, Beom-Jun Kim, Junling Jia, Zhiwei Huang, Tao Yang, et al. 2008. 'Acetylation of Smc3 by Eco1 Is Required for S Phase Sister Chromatid Cohesion in Both Human and Yeast'. *Molecular Cell* 31 (1): 143–51. <https://doi.org/10.1016/j.molcel.2008.06.006>.

Zhao, Siming, Stefania Bellone, Salvatore Lopez, Durga Thakral, Carlton Schwab, Diana P. English, Jonathan Black, et al. 2016. 'Mutational Landscape of Uterine and Ovarian Carcinosarcomas Implicates Histone Genes in Epithelial–Mesenchymal Transition'. *Proceedings of the National Academy of Sciences* 113 (43): 12238–43. <https://doi.org/10.1073/pnas.1614120113>.

Zhu, Xiaolin, Slavé Petrovski, Pingxing Xie, Elizabeth K. Ruzzo, Yi Fan Lu, K. Melodi McSweeney, Bruria Ben-Zeev, et al. 2015. 'Whole-Exome Sequencing in Undiagnosed Genetic Diseases: Interpreting 119 Trios'. *Genetics in Medicine* 17 (10): 774–81. <https://doi.org/10.1038/gim.2014.191>.

Appendix A



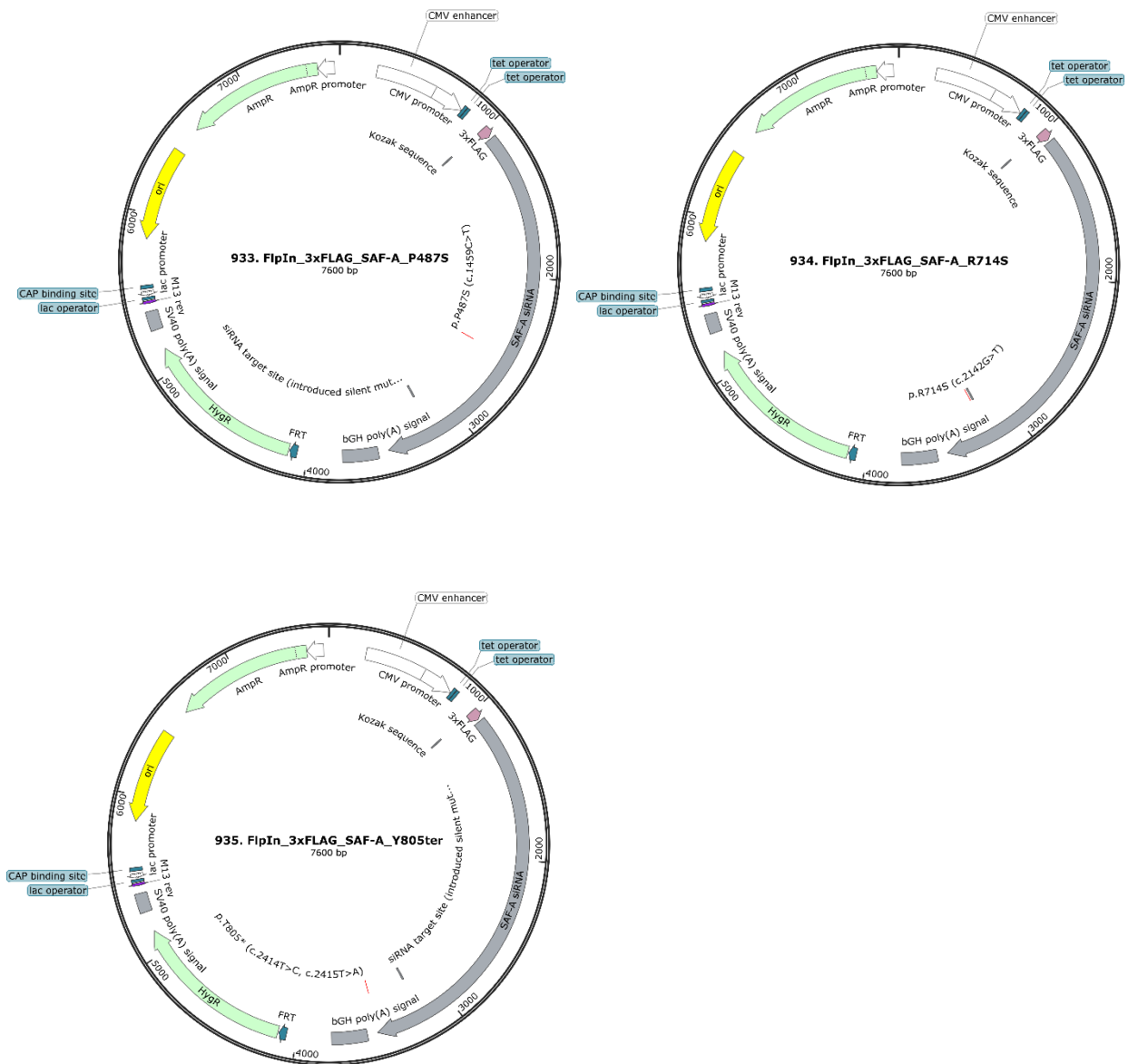
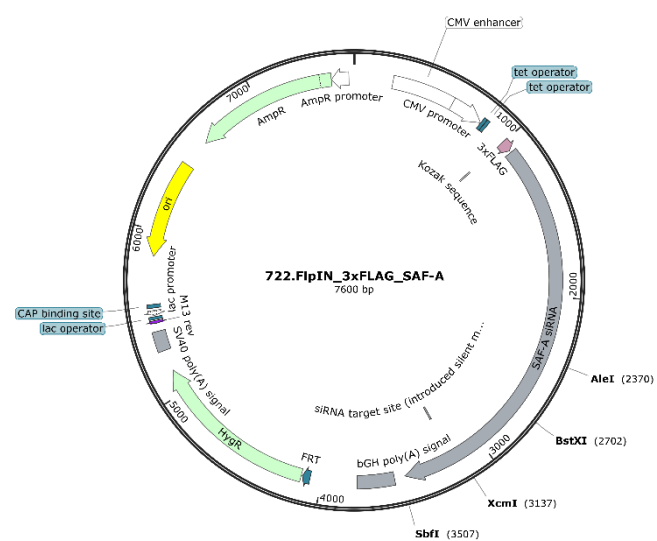
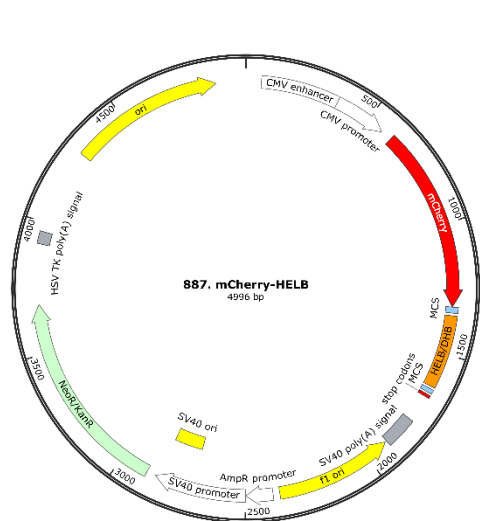
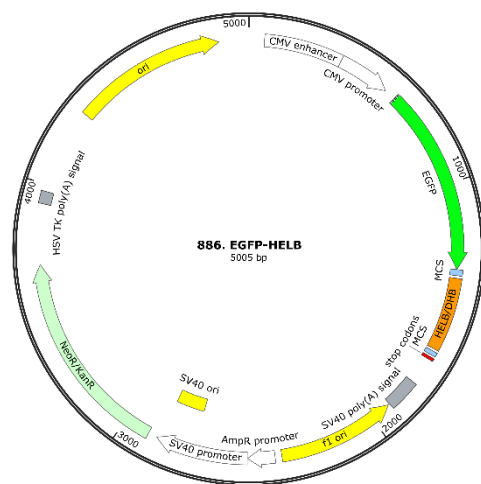
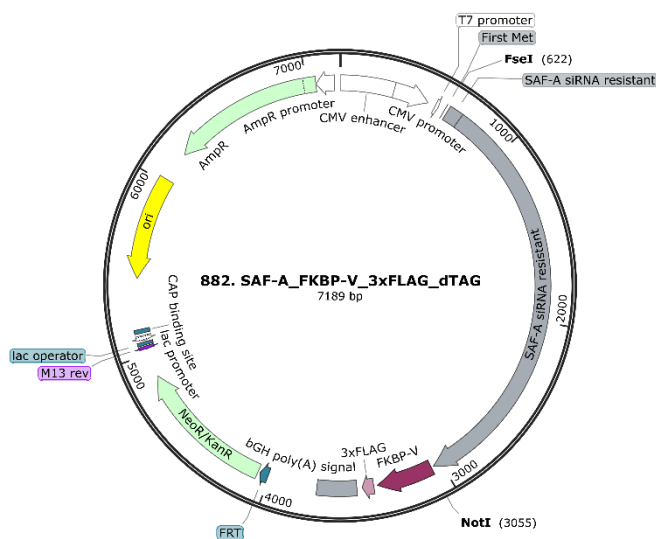
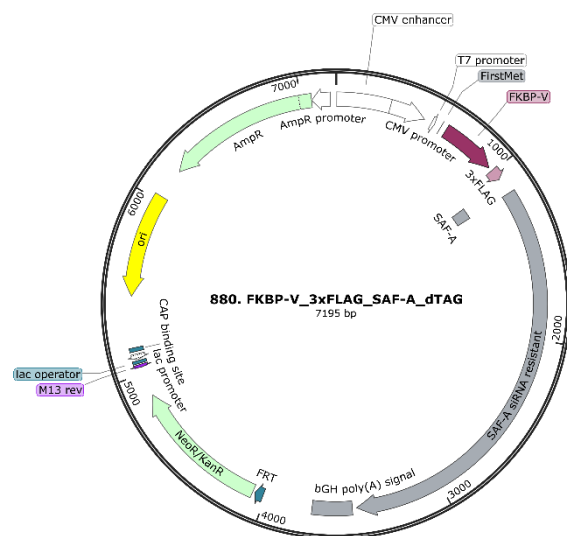
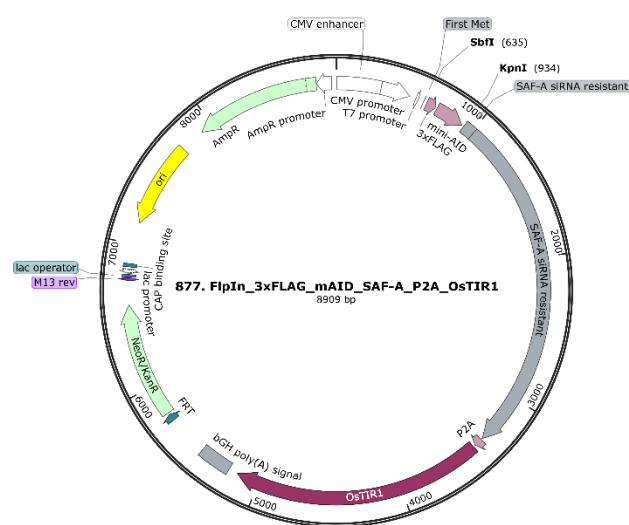
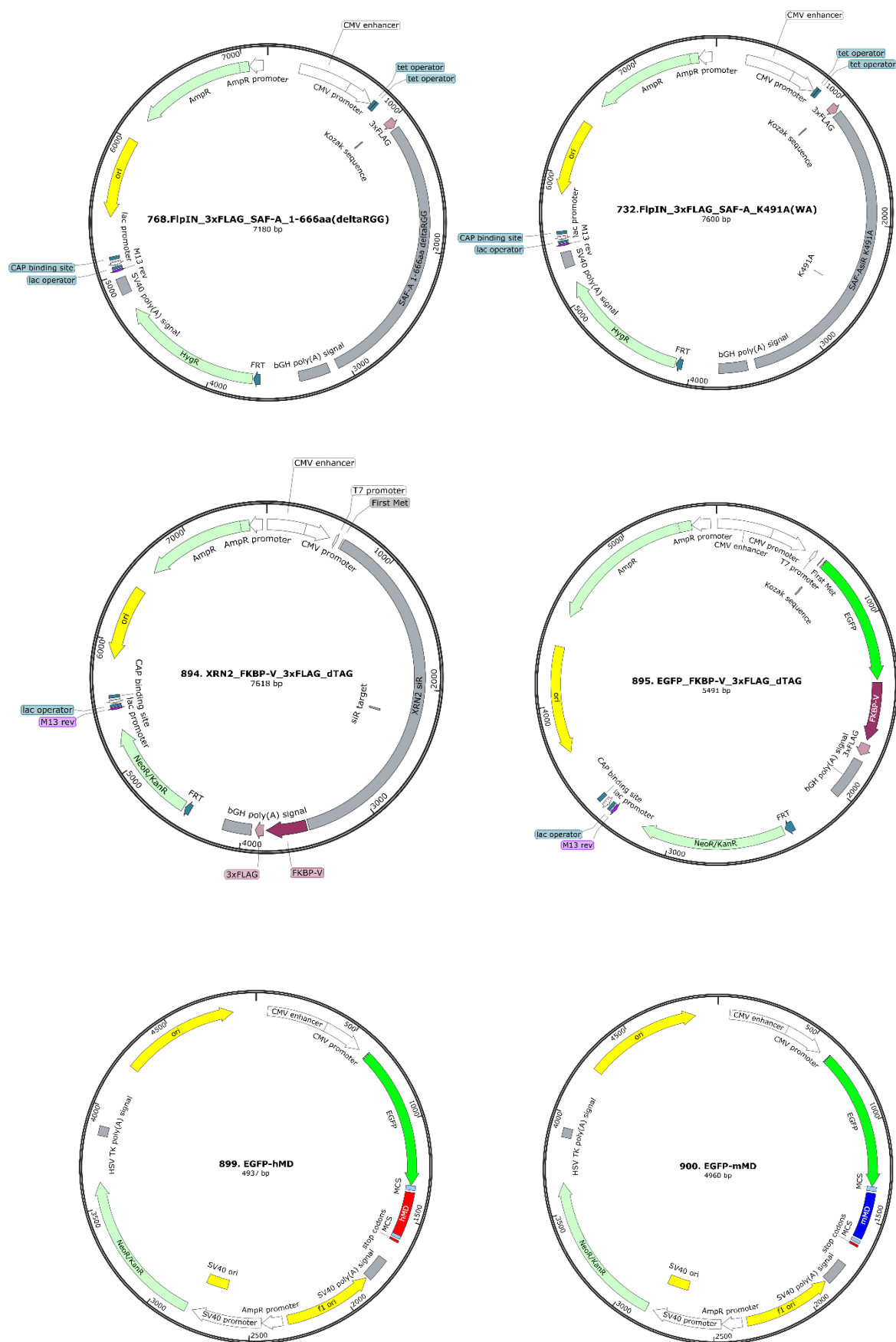


Figure 1.A. Vector maps of plasmids used for FlpIn 293 T-REx cell line generation. Enzymes used in this thesis are displayed.





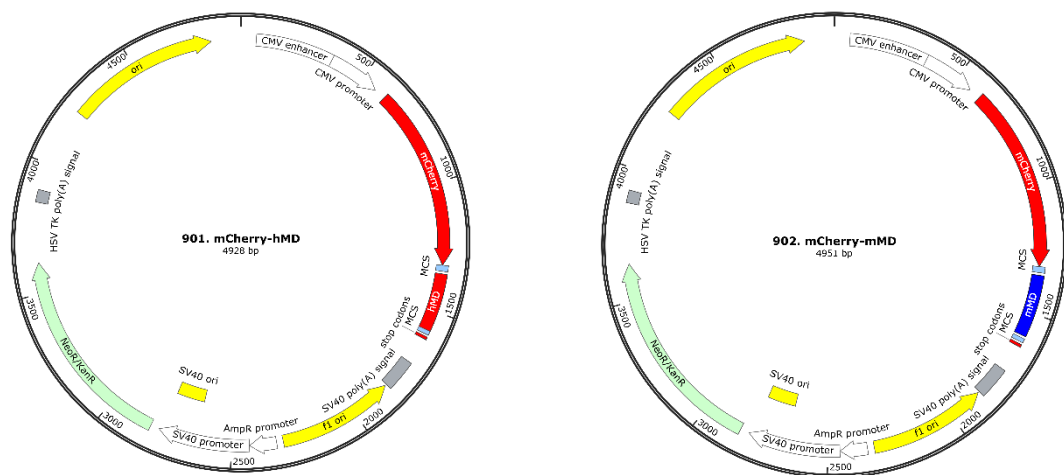


Figure 2.A. Vector maps of plasmids used in DNA transfections. Enzymes used in this thesis are displayed.



Figure 3.A. Vector maps of plasmids used in CRISPR Cas9 cell line generation.
Enzymes used in this thesis are displayed.

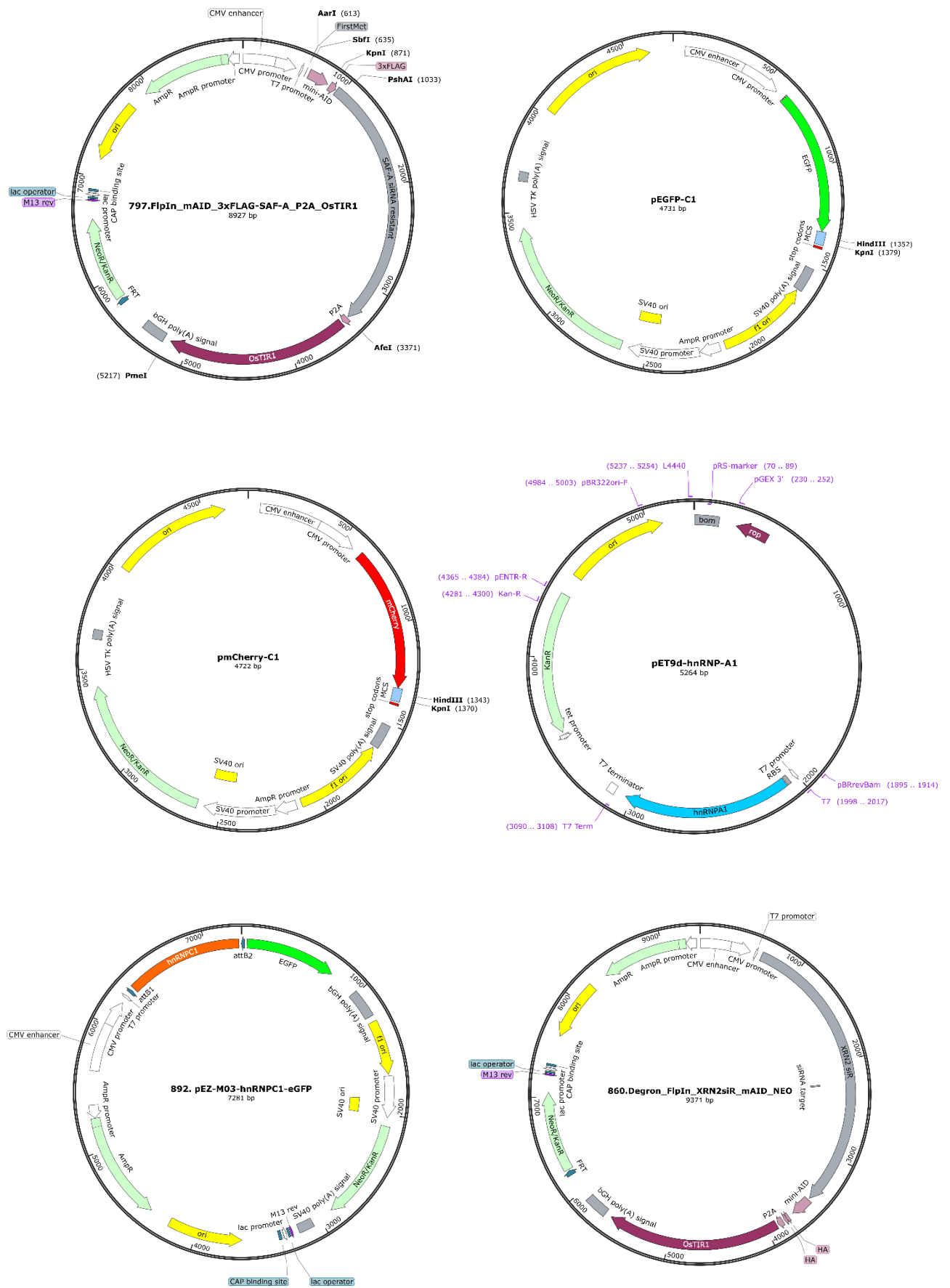


Figure 4.A. Vector maps of plasmids used for amplification or digestion to make other plasmids. Enzymes used in this thesis are displayed.

Appendix B

Name	Sequence
3xFLAG-mAID	TGGAGGAGAACCCTGGACCTAGATCTGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTCAGGAAGCGGAAAGGAGAAGAGTGCTTGTCTAAAGATCCAGCCAAACCTCCGGCCAAGGCACAAGTTGTGGGATGGCCA CCGGTGAGATCATACCGGAAGAACGTGATGGTTTCCTGCCAAA AATCAAGCGGTGGCCCGGAGGCGGCGGCGTTCGTGAAGGTAT CAATGGACGGAGCACCGTACTTGAGGAAAATCGATTTGAGGAT GTATAAAGGCGCTGGTGCAGGCGCCACTAGTGTGCGGGAGGA GGAGAAGGG
dTAG N-term	CTGCGCACCATGGCCGGCCAATTGCCTGCAGGCCTCGGAGGA GGAGTGCAGGTGGAAACCATCTCCCCAGGAGACGGGCGCACCC TTCCCAAGCGCGGCCAGACCTGCGTGGTGCCTACACCGGG ATGCTTGAAGATGGAAAGAAAGTTGATTCCTCCCGGGACAGAA ACAAGCCCTTTAAGTTTATGCTAGGCAAGCAGGAGGTGATCCG AGGCTGGGAAGAAGGGGTTGCCCAGATGAGTGTGGGTCAGAG AGCCAACTGACTATATCTCCAGATTATGCCTATGGTGCCACTG GGCACCCAGGCATCATCCCACCACATGCCACTCTCGTCTTCGA TGTGGAGCTTCTAAAACTGGAAAGCTACGATGGTGGAGGTACC GACTACAAAGACCATGACGGTGATT
dTAG C-term	AGTCAGCATTATCACCAAGGATATTATAGCCTTGCGGCCGCGGG AGTGCAGGTGGAAACCATCTCCCCAGGAGACGGGCGCACCTT CCCCAGCGCGGCCAGACCTGCGTGGTGCCTACACCGGGAT GCTTGAAGATGGAAAGAAAGTTGATTCCTCCCGGGACAGAAAC AAGCCCTTTAAGTTTATGCTAGGCAAGCAGGAGGTGATCCGAG GCTGGGAAGAAGGGGTTGCCCAGATGAGTGTGGGTCAGAGAG CCAAAGTACTGACTATATCTCCAGATTATGCCTATGGTGCCACTGGG CACCCAGGCATCATCCCACCACATGCCACTCTCGTCTTCGATG TGGAGCTTCTAAAACTGGAAAGCTACGATGGTGGAGGTACCGA CTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTA CAAGGATGACGATGACAAGCCTGCAGGCAAACCCGCTGATCAG CCTCGACTGTGCCTT
HELB	TCGATAAGCTTACACGATGACAAATGATGTCACCTGGAGCGAG GCCTCTTCGCCTGATGAGAGGACACTCACCTTTGCTGAAAGAT GGCAATTATCTTCACCTGATGGAGTAGATACAGATGATGATTTAC CAAAATCGCGAGCATCCAAAAGAACCTGTGGTGTGAATGATGAT GAAAGTCCAAGCAAAATTTTTATGGTGGGAGAATCTCCACAAGT GTCTTCCAGACTTCAGAATTTGAGACTGAATAATTTAATTCCCAG GCAACTTTTCAAGCCCACCGATAATCAAGAACTGGATGGTACC AATTAA
hMD	TCGATAAGCTTACACGGCTGAAAATAAGGCGAAGATCAACATGG CAGGCGCAAAGCGCGTTCTACGGCCCCTGCTGCAACCTCCA AGCCCGGACTGAGGCCAAGAACAGCTCTTGGGGACATTGGTA ACAAAGTCAGTGAACAACTGCAGGCCAAAATGCCTATGAAGAA GGAAGCAAAACCTTCAGCTACTGGAAAAGTCATTGATAAAAAAC TACCAAAACCTCTTGAAAAGGTACCTATGCTGGTGCCAGGATGG TACCAATTAA

mMD	TCGATAAGCTTACACGGCAGAAAATAAGGCCAAGGTCAGTATGG CAGGCGCCAAGCGTGTGCCTGTGACAGTTACTGCTGCTTCCAA GCCCCGGGCTGAGACCGAGAAGTCTCTTGGAGACATTGGTAAT AAAGTCAGCGAAGAGCTACAGGCAAGAGTGCCTCTGAAAAGG GAAGCAAAAACGCTAGGTACTGGAAAAGGTACTGTTAAAGCCC TACCAAAACCTGTAGAGAAGGTGCCTGTGTGTGAACCAGGATG GTACCAATTAA
SAF-A 423- 556 aa with c.1459CtoT	GCCACTGTTCCCGCATGTTCTCTGCCACAAGTGTGCAGTTGAA TTTAATTTTGGTCAGAAGGAAAAGCCATATTTTCCAATACCTGAA GAGTATACTTTTCATCCAGAACGTCCCCTTAGAGGATCGAGTTAG AGGACCAAAGGGGGCCTGAAGAGAAGAAAAGATTGTGAAGTTGT GATGATGATTGGCTTGTGAGGAGCTGGAAAACTACCTGGGTTA CTAAACATGCAGCAGAAAATCCAGGGAAATATAACATTCTTGGC ACAAATACTATTATGGATAAGATGATGGTGGCAGGTTTTAAGAAG CAATGGCAGATACTGGAAAAGTGAACACACTGTTGCAGAGAG CCCCCAGTGTCTTGGGAAATTTATTGAGATTGCTGCCCCGAAA GAAGCG
SAF-A 677- 806 aa with c.2199GtoT	GCAAGAACCAGTTTAACAGAGGTGGTGGCCATAGAGGACGTG GAGGATTCAATATGCGTGGTGGAAATTTTCAGAGGAGGAGCCCC AGGAAACCGAGGCGGATATAATAGGAGTGGCAACATGCCACAG AGAGGTGGTGGCGGTGGAGGAAGTGGTGGAAATCGGCTATCCA TACCCTCGTGCCCCCTGTTTTTCCTGGCCGTGGTAGTTACTCAA CAGAGGGAACTACAACAGAGGTGGAATGCCCAACAGAGGGAA CTACAACCAGAACTTCAGAGGACGAGGAAACAATCGTGGCTAC AAAAATCAATCTCAGGGCTACAACCAGTGGCAGCAGGGTCAAT TCTGGGGTCAGAAGCCATGGAGTCAGCATTATCACCAAGGATAT TATTGAGGATCCGGCCTGCAGGGATAACTCGAGTCTAGAGGGC CCGTTTAAACCCGC
SAF-A 677- 806 aa with c.2414_241 5delinsGA	GCAAGAACCAGTTTAACAGAGGTGGTGGCCATAGAGGACGTG GAGGATTCAATATGCGTGGTGGAAATTTTCAGAGGAGGAGCCCC AGGAAACCGAGGCGGATATAATAGGAGGGGCAACATGCCACAG AGAGGTGGTGGCGGTGGAGGAAGTGGTGGAAATCGGCTATCCA TACCCTCGTGCCCCCTGTTTTTCCTGGCCGTGGTAGTTACTCAA CAGAGGGAACTACAACAGAGGTGGAATGCCCAACAGAGGGAA CTACAACCAGAACTTCAGAGGACGAGGAAACAATCGTGGCTAC AAAAATCAATCTCAGGGCTACAACCAGTGGCAGCAGGGTCAAT TCTGGGGTCAGAAGCCATGGAGTCAGCATTATCACCAAGGATG ATATTGAGGATCCGGCCTGCAGGGATAACTCGAGTCTAGAGGG CCGTTTAAACCCGC

Table A. Gibson Assembly and ligation DNA constructs.

Appendix C

/*

* Written for Elena Lavarova

* Allows user to draw area on which coloc will be measured through time

* Background subtraction also performed (high background is not good for most coloc measurements)

* User can choose to analyse all frames or select specific ones

* Previous version of this macro used JaCoP but BIOP version seems much faster

* This means BIOP update site needs to be enabled by going to Help > Update > Manage Update Sites

*

*

Written by Laura Murphy

*

IGC Advanced Imaging

Resource

*

Nov 2022

*/

```
biop = getBoolean("Is the BIOP update site enabled on this version of FIJI", "Yes", "I'll check and add it if not");
```

```
if (biop==0) {
```

```
    exit("This macro needs the BIOP update site");
```

```
}
```

```
// Get input and output folders - should switch to script parameters (although should probably switch to a better programming language really)
```

```
inputFolder = getDirectory("Select the folder with your image folders");
```

```
outputFolder = getDirectory("Select the folder where you want to save the output");
```

```
// Create arrays of files in subfolder, uses function at bottom of script
```

```
dirList = newArray();
```

```
dirList = getFileTree(inputFolder, dirList);
```

```
Table.create("All images Pearsons Coefficient.csv");
```

```
// Start loop and open image with Bio-Formats
```

```
for (img = 0; img < dirList.length; img++) {
```

```
    path = dirList[img];
```

```
    run("Bio-Formats Macro Extensions");
```

```
    run("Bio-Formats Importer", "open=&path autoscale color_mode=Composite  
rois_import=[ROI manager] view=Hyperstack stack_order=XYCZT series_1");
```

```
    ori = getTitle();
```

```
    title = File.nameWithoutExtension();
```

```
    getDimensions(width, height, channels, slices, frames);
```

```
    run("Z Project...", "projection=[Max Intensity]");
```

```
    proj = getTitle();
```

```
    setTool("polygon");
```

```
    waitForUser("This projection shows the maximum projection of your time lapse  
stack\n\nDraw an ROI on the projection that covers the full area the cells will be in over time \n\nThis  
area will be used for colocalisation measurement");
```

```

// Let user know if they didn't draw a selection, full FOV will be analysed
if (selectionType() == -1){
    waitForUser("No ROI drawn, full field of view will be analysed");
}

selectWindow(ori);

// High background causing issues with coloc, using sliding paraboloid of 400 pixel as covers
largest of cells seen in test set

paraboloid = 400;

run("Subtract Background...", "rolling=" + paraboloid + " sliding");

run("Restore Selection");

close(proj);

// Make dialog

choice = getBoolean("Would you like to select specific frames to analyse or would you like to
analyse the full video?", "Pick frames", "Full stack");

if (choice==1) {

    waitForUser("Look through to note which frames you want to select");

    Dialog.create("Slices you want");

    Dialog.addString("Frames to be extracted", "14,20," + frames);

    Dialog.show();

    str = Dialog.getString();

    run("Make Substack...", "channels=1-2 frames=" + str);

    close(ori);

```

```

    }

    run("Restore Selection");

    run("Crop");

    run("BIOP JACoP", "channel_a=1 channel_b=2 threshold_for_channel_a=Otsu
threshold_for_channel_b=Otsu manual_threshold_a=0 manual_threshold_b=0 crop_rois
get_pearsons get_manders get_overlap get_li_ica get_fluorogram costes_block_size=5
costes_number_of_shuffling=100");

    run("Images to Stack", "use");

    saveAs("Tiff", outputFolder + title + "_JaCop_Results.tiff");

    run("Close");

    close(ori);

    selectWindow("Results");

    pearson = Table.getColumn("Pearson's Coefficient");

    saveAs("Results", outputFolder + title + "Coloc_Results.csv");

    selectWindow("All images Pearsons Coefficient.csv");

    Table.setColumn(title, pearson);

    saveAs("Results", outputFolder + "All images Pearsons Coefficient.csv");

    run("Close All");

    run("Clear Results");

}

// Let user know the macro is complete

Dialog.create("Progress");

```



```
Dialog.addMessage("Macro Complete!");
```

```
Dialog.show;
```

```
// File tree function
```

```
function getFileTree(dir , fileTree){
```

```
    list = getFileList(dir);
```

```
    for(f = 0; f < list.length; f++){
```

```
        if (matches(list[f], "(?i).*\\.\\.(tif|tiff|nd2|ndpi|ims|oib)$"))
```

```
            fileTree = Array.concat(fileTree, dir + list[f]);
```

```
            if(File.isDirectory(dir + File.separator + list[f]))
```

```
                fileTree = getFileTree(dir + list[f],fileTree);
```

```
        }
```

```
    return fileTree;
```

```
}
```

```
function num2array(str,delim){
```

```
    arr = split(str,delim);
```

```
    for(i=0; i<arr.length;i++) {
```

```
        arr[i] = parseInt(arr[i]);
```

```
    }
```

```
    return arr;
```

```
}
```

Appendix D

DNA mutation	Amino acid change	Protein region	Age	Gender	Neonatal problems	ID	Severity of ID	Language delay	Motor delay	Any psychiatric diagnosis?	Other psychopathology/behaviour abnormality	Hypotonia	Epilepsy	Cardiac/renal abnormality	Other abnormalities	Reference
Chr1: 245,025,709-245,133,797dup	chr 1 duplication including first 3 exons of HNRNPU	SAP domain + Acidic/disordered + SPRY domain	0.1	M	seizures, hypotonia	n/a	n/a	n/a	n/a	n/a	n/a	yes	seizures	patent ductus arteriosus	microcephaly, short stature, hypertelorism, short nose, bulbous nose tip, long philtrum, thin upper vermillion, microretrognathia, big ears, short neck, cystic lesion between posterior fossa and 4th ventricle, Dandy Walker malformation, renal pelvic ectasia, joint contractures, clonus on extremities, hypospadias, scrotal raffe, 10 ribs, closed anterior fontanelle, cryptorchidism, pes equinovarus, duplication covers first 3 exons of HNRNPU and EFCAB2	Bramwig et al., 2017
c.16delinsATT	p.Val6Ilefs*4 + inherited (paternal mosaicism)	SAP domain	6.5 8	F F	no n/a	yes yes	severe mild? Syndromic	yes n/a	yes n/a	no n/a	no n/a	yes n/a	yes yes	n/a n/a	status epilepticus, enlarged lateral ventricles, complete corpus callosum, microcephaly microcephaly, facial dysmorphism	Depienne et al., 2017; Chérot et al., 2018

c.23del1	p.Val8 Glufs*4	SAP domain	6.5	F	no	yes	severe	yes	yes	no	conspicuously happy disposition, drooling	no	1 seizure	no	facial dysmorphism, syndactyly, hypermobile, severe post natal growth retardation, lip, musculoskeletal, eye, ear, strabismus, microcephaly, mild hearing impairment, hand flapping	Yates et al., 2017
c.29_30del	p.Lys10 del	SAP domain	4	M	no	yes	mild?	yes	yes	autistic features	developmental regression	yes	no	no	ataxia, musculature, vision loss, ptosis, hepatogemaly	Gillentine et al., 2021
c.67C>T	p.Arg23*	SAP domain	2.6	F	gastronomy tube feeding	no	no	yes	yes	no	recurrent hand flapping	yes	no	no	nystagmus, strabismus, hypothyroidism, central hypoventilation, musculature (weakness)	Durkin et al., 2020
c.76del	p.Ser26 LeufsTer35	SAP domain	5	M	no	yes	n/a	yes	yes	ASD	recurrent hand flapping	no	yes	no	mild hypoplasia of corpus callosum, status epilepticus	Shimada et al., 2018
c.143_149del7	p.Gly48 Afs*11	IDR 1	1.7	F	no	no	no	yes	yes	no	no	no	yes	patent ductus arteriosus	craniofacial dysmorphisms (high forehead, frontal bossing, thick lips, low ear position)	Song et al., 2021
c.149_156del	p.Arg50 Hisfs*	IDR 1	11	M	no	yes	n/a	yes	no	ASD	yes	no	seizures	tetralogy of Fallot	cataract, strabismus	Gillentine et al., 2021
c.198del	p.Ala67 fs*	IDR 1	n/a	n/a	no	no	no	no	yes	no	no	yes	seizures	no	short stature, musculature	Gillentine et al., 2021
c.253dupG	p.Glu85 Glyfs*	IDR 1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	yes	n/a	n/a	Epilepsy Genetics Initiative, 2019

c.324 _328 del	Glu109 Argfs*3 4	IDR 1	2.5	F	no	yes	n/a	yes	yes	no	no	no	no	seizures	no	decreased body weight, short stature, hyperopia, strabismus, craniofacial dysmorphisms (bulbous nose tip, triangular face, short nose, thin upper lip)	Gillentine et al., 2021
c.325 G>C	Glu109 Gln	IDR 1	n/a	F	no	yes	n/a	yes	no	no	happy demeanor	no	no	no	no	mother has mild ID	Gillentine et al., 2021
c.334 dup	p.Ala112 Glyfs*33	IDR 1	n/a	M	no	yes	n/a	no	no	ASD	yes	no	seizures	no	no	bilateral preaxial polydactyly feet, macrotia, craniofacial dysmorphisms (ear pit, low set ears)	Wang et al., 2020
c.395 _401 del7	p.Asn132 Thrfs*63	IDR 1	10.95	M	feeding	yes	n/a	yes	yes	no	recurrent hand flapping	no	seizures	no	no	craniofacial dysmorphisms (hands, eye, strabismus, cupped ear), growth delay, decreased body weight, proportionate short stature, skeletal morphology	Durkin et al., 2020
c.401 _402 delA C	Asp134 Glyfs*10	IDR 1	15	M	n/a	yes	n/a	n/a	n/a	n/a	n/a	n/a	yes	no	no	no	Fernández-Marmiesse et al., 2019
c.418 G>A	p.Glu140 Lys	IDR 1	14	M	congenital heart disease, feeding difficulties	yes	moderate	yes	yes	Tourette syndrome	angers easily	no	no	transposition of the great vessels, tricuspid atresia	no	pointed chin, craniofacial dysmorphisms (hand, eye, strabismus, posterior helix pit, microcephaly, protruding tongue), growth delay, short stature	King et al., 2014/ Yates et al., 2017
c.454 _466 del	p.Ala152 Thrfs*41	IDR 1	8.5	F	Necrotizing enterocolitis	yes	severe	yes	yes	no	autistic features, very stubborn	no	seizures	no	no	craniofacial dysmorphisms (broad forehead, small chin, bilateral glue ear, exotropia), lordosis, constipation	Durkin et al., 2020

c.469 G>C	p.Gly15 7Arg	IDR 1	n/a	M	n/a	n/a	n/a	n/a	n/a	PDD, ASD (Gillentine)	withdrawn (Gillentine)	n / a	n/a	no	macrocephaly (Gillentine)	Krum m et al., 2015
c.481 C>T	p.Gln16 1*	IDR 1	n/a	n / a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	seizur es	n/a	n/a	Gillent ine et al., 2021
c.508 C>T	p.Gln17 0*	IDR 1	29	F	no	yes	n/a	yes	no	ASD, OCD	poor sleep	n / a	seizur es	no	no	Felicia no et al., 2019;
			34.8	F	hypot onia, feedin g difficu lties	yes	moderate	yes	yes	autism, OCD, anxiety	no	y e s	yes	no	strabismus, craniofacial dysmorphisms (thin upper vermillion, slight frontal bossing, prominent eyebrows)	Taylor et al., 2022
c.511 C>T	p.Gln17 1*	IDR 1	4	M	no	yes	severe	yes	yes	no	autistic features	y e s	epilep sy	aortic dilation, unilater al multicystic kidney	craniofacial dysmorphisms (ear, teeth, eversion of upper lip), bilateral single palmar creases, growth delay, abnormal brain imaging	Hamd an et al., 2014/ Bram wig et al., 2017
c.523 C>T	p.Gln17 5*	IDR 1	3.5	F	no	yes	severe	yes	yes	no	no	n o	seizur es	no	craniofacial dysmorphisms (ear, thumbs, microcephaly, teeth), short stature, developmental regression	Bram wig et al., 2017
c.547 _560 del	p.Ala18 3fs*	IDR 1	6	F	no	yes	n/a	yes	yes	ASD	stereotypic al hand wringing	y e s	seizur es	no	short stature, musculature, hemivertebrae, scoliosis, musculoskeletal physiology, elbow hyperextensibility, craniofacial dysmorphisms	Gillent ine et al., 2021
c.596 dup	p.Pro20 0fs*	IDR 1	n/a	n / a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	n/a	n/a	n/a	Gillent ine et

				a								a				al., 2021
c.619 C>T	p.Gln207*	IDR 1	0.7	F	hypotonia	n/a	n/a	n/a	n/a	n/a	n/a	yes	seizures	no	anteriorly placed anus, absent patellar reflexes	Durkin et al., 2020
c.622 C>T	p.Gln208T	IDR 1	19	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	yes	n/a	n/a	Salinas et al., 2021
In 806 isoform of HNRNPU, deletion from 212 to 230 aa or bp 636 to 690 (231 = 212 (-19) aa or 693 = 636 bp (-57)) [square brackets have isoform location]																
c.651_660 del [deleted]	p.Gly218Alafs*118 [deleted]	IDR 1	1.8	M	n/a	yes	n/a	yes	yes	n/a	hand-flapping	yes	seizures	n/a	macrocephaly, craniofacial dysmorphisms, developmental regression, malrotated left hippocampus	Leduc et al., 2017
c.692 - 1G>A	p.? splice site variant	IDR 1	8	F	n/a	yes	moderate-severe	yes	yes	n/a	n/a	n/a	yes	n/a	n/a	Depienne et al., 2017;
			4.1	F	n/a	n/a	n/a	n/a	yes	no	no	no	seizures	patent ductus arteriosus, atrial septal defect	prominent median palatal raphe, intrauterine growth restriction, abnormal MRI (some delayed myelination, enlargement of the frontal horns of the lateral ventricles)	Durkin et al., 2020
c.706_707 del [c.649-650del]	p.Glu236Trpfs*6 x3 [p.Glu217Trpfs*6]	IDR 1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	yes	n/a	n/a	Farwell et al., 2015;
			15.9	F	n/a	n/a	n/a	yes	yes	no	no	no	seizures	no	craniofacial dysmorphisms (branchycephaly, long eyelashes, long philtrum, low anterior hairline, up slanted palmar fissures), microcephaly, clinodactyly, decreased body weight, short stature	Durkin et al., 2020;

			23	F	n/a	n/a	n/a	yes	yes	no	abnormal temper tantrums	no	seizures	atrial septal defect (closed)	craniofacial dysmorphisms (up slanting palpebral fissures, long eyelashes, thin upper lip), broad thumbs, scoliosis, flat feet, cold feet, digestive	Durkin et al., 2020 #2
c.712_715del [c.65_65_65del]	p.Lys238Alafs*100 [p.Lys219Alafs*100]	IDR 1	5	M	n/a	n/a	n/a	yes	yes	ASD	no	yes	seizures	no	craniofacial dysmorphisms (frontal bossing, prominent metopic suture, flattened nasal bridge, short nose), bilateral postaxial polydactyly, undescended testis, decreased body weight, hypothyroidism	Durkin et al., 2020
c.730_731delAG [c.673_674delAG]	p.Arg244Glyfs*3 [p.Arg225Glyfs*3]	IDR 1	23	F	hypotonia, feeding difficulties	yes	moderate	yes	yes	no	no	yes	yes	no	synophris, long columella, arched palate, hallux vulgus	Taylor et al., 2022
c.742dup [c.685dup]	p.Arg248Lysfs*12 [p.Arg229Lysfs*12]	IDR 1	2.1	F	hypotonia	yes	moderate	yes	yes	no	no	yes	yes	no	intrauterine growth restriction, few foetal movements, thin upper vermillion, EEG abnormality	Taylor et al., 2022
c.804-9_804-6del [c.747-9_747-6del]	p.?	IDR 1	36.7	M	hypotonia, feeding difficulties	yes	mild	yes	yes	autistic features, OCD	aggressive behaviours, anxiety	yes	no (but febrile seizures)	no	thinning outer eyebrows, depressed nasal bridge, slightly high palate, swallowing problems, constipation, chronic ear infections, incisor with two roots	Taylor et al., 2022

c.817 C>T [c.76 0C>T]	p.Gln27 3* x2 [p.Gln2 54*]	SPRY domain	32	F	n/a	yes	severe	yes	no	no	no	yes	no (but fever- related at young er age)	pentha logy of Fallot	microcephaly, short stature, facial dysmorphisms (prominent metopic ridge, deep set eyes, hypertelorism, depressed nasal bridge, short nose, ateverted nares, bubulous nasal tip, long philtrum, eversion of upper lip, prominent jaw), teeth, short neck and webbing, feeding difficulties, short hands and feet, broad thumbs, growth delay	Brams wig et al., 2017;
			20	F	n/a	yes	n/a	n/a	n/a	ASD	no	n/a	yes	n/a	n/a	Demo s et al., 2019
c.837 _839 delA GA [c.78 0_78 2delA GA]	p.Glu27 9del [p.Glu2 60del]	SPRY domain	13	M	n/a	yes	moderate (IQ 52)	yes	yes	no	no	n/a	yes	n/a	craniofacial dysmorphisms (dolichocephaly, prominent forehead, down slant, slightly low hanging columella, thin upper lip), long fingers, long toes	Durkin et al., 2020
c.847 _857 del [c.79 0_80 0del]	p.Phe2 83Serfs *5 [p.Phe2 64Serfs *5]	SPRY domain	18	F	no	yes	n/a	yes	yes	n/a	drooling, bruxism	n/a	seizur es	no	short stature, stereotypical hand wringing, broad based gait, minor frontal atrophy, short fourth and fifth metacarpal and tapered finger, rhizomatic shortening, strabismus, dysmorphic facial features (wide nasal bridge), episodic hyperventilation (interspersed with apnoea associated with cyanosis)	Durkin et al., 2020

c.872_877delATTACTT [c.815_820delATTACTT]	p.Thr292_Tyr293del [p.Thr273_Tyr274del]	SPRY domain	n/a	n/a	n/a	n/a	n/a	n/a	n/a	ASD	n/a	n/a	n/a	n/a	n/a	Gillentine et al., 2021
c.878A>G [c.821A>G]	p.Tyr293Cys [p.Tyr274Cys]	SPRY domain	6	F	no	no	no	no	no	no	no	no	yes	no	suspected developmental regression	Hinokuma et al., 2020
c.878-9T>G [c.821-2T>G]	p.?	SPRY domain	4	M	feeding, jaundice	yes	moderate	yes	yes	no	no	yes	yes	moderate atrial septal defect secundum	borderline ventriculomegaly, bilateral convergent squint, long columella, four limb post-axial polydactyly, inguinal hernia, mild pulm, stenosis	Taylor et al., 2022
c.893A>G [c.836A>G]	p.His298Arg [p.His279Arg]	SPRY domain	8	F	hypotonia, feeding difficulties	yes	severe	yes	yes	ASD	no	yes	yes	no	additional genetic defect (m.15266A>G), mesial sclerosis of hippocampus, enlarged temporal lobes, strabismus, small nose, narrow mouth, round small face, frontal bossing	Taylor et al., 2022
c.960G>A [c.903G>A]	p.Trp320* [p.Trp301*]	SPRY domain	11	F	no	yes	severe	yes	yes	n/a	n/a	yes	yes	no	thick hair, prognathism, broad thumbs and great toes, truncal obesity, prominent eyebrows, elongated palpebral fissures, prominent nasal bridge, overhanging columella, thin upper lip, very sociable, hand flapping, wide-based unsteady gait, hypermobility	Yates et al., 2017

c.970 A>G [c.91 3A>G]	p.Arg324Gly [p.Arg305Gly]	SPRY domain	14.5	M	no	yes	severe	yes	no	n/a	n/a	y e s	yes	atrial septal defect, agenesi s of the right kidney, hypopla stic aorta abdom	additional genetic defect (KDM6B), growth delay, short stature, facial dysmorphisms (prominent metopic ridge, deep set eyes, hypertelorism, depressed nasal bridge, short nose, anteverted nares, bulbous nasal tip, long philtrum, thin upper vermillion, eversion of upper lip, large incisors, short neck, low set ears), corpus callosum abnormalities, cerebral atrophy, multiple lesions, atrophy of the cerebellar vermis, feeding difficulties, hearing deficit, scoliosis, delayed bone age, onychodystrophia	Brams wig et al., 2017
c.971_973 delG AG [c.914_916delG AG]	p.Arg324_Ala325delinsThr [p.Arg305_Ala306delinsThr]	SPRY domain	43	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n / a	yes	n/a	n/a	Johan nesen et al., 2020
c.1050_1051del [c.993_994del]	p.Thr351Lysfs*4 [p.Thr332Lysfs*4]	SPRY domain	n/a	F	no	yes	n/a	no	yes	ASD	no	y e s	no	no	relative microcephaly	Wang et al., 2020

c.108 8G>A [c.10 31G> A]	p.Trp36 3* [p.Trp3 44*]	SPRY domain	11.8	F	no	yes	profound	yes	yes	no	stereotypy, bruxism	n o	seizur es	no	regression, craniofacial dysmorphisms (progressive microcephaly, synophrys, widely spaces teeth), functional respiratory abnormality, regressed movement, short stature, simplified gyral patterns, reduced white matter volume	Durkin et al.,202 0
c.108 9G>A [c.10 32G> A]	p.Trp36 3* x2 [p.Trp3 44*]	SPRY domain	7	M	no	yes	n/a	yes	no	ASD	no	n o	seizur es	no	craniofacial dysmorphisms (short forehead, straight eyebrows, bulbous nose, left ear dysplastic, wide frontal incisors, developmental regression, decreased T2 shortening in the anterior temporal lobes	Leduc et al., 2017;
			n/a	F	n/a	yes	severe	yes	yes	no	bruxism	n / a	seizur es	no	craniofacial dysmorphisms (dysplastic earlobes, short palpebral fissures, broad nasal bridge, hypoplastic alae nasi, tented upper lip, prominent jaw, microcephaly), small hands and feet, tapering fingers, bilateral single palmar creases, short fifth fingers, spasticity	Durkin et al., 2020;
c.111 7+1G >A [c.10 60+1 G>A]	Splice donor variant	SPRY domain	12	F	n/a	yes	moderate	yes	yes	no	autism traits, hand flapping, special education needs	y e s	seizur es	no	craniofacial dysmorphisms (low anterior hairline, synophrys, strabismus, short upturned nose), spinal lordosis, prominent eyebrows,	Yates et al., 2017

c.113 2T>C [c.10 75T> C]	p.Ser37 8Pro [p.Ser3 59Pro]	SPRY domain	23	M	n/a	yes	severe	yes	no	no	no	yes	yes	no	additional variants (maternal duplication 8p23.1, maternal deletion 13q31.3), craniofacial dysmorphisms (deep set eyes, hypertelorism, depressed nasal bridge, short nose, anteverted nares, bulbous nasal tip, eversion of upper lip, thick incisors, accessory incisor in palatum composed of several elements, diastemas between teeth, slightly over-folded helices), wide ventricles, retarded frontal myelinisation, glandula pinealis cyst, secondary osteoporosis, obstipation, rotatory nystagmus, cortical visual impairment, brain imaging abnormalities (retarded frontal myelinisation, glandular pinealis cyst, ventriculomegaly)	Brams wig et al., 2017
c.114 2A>G [c.10 85A> G]	p.Tyr38 1Cys [p.Tyr3 62Cys]	SPRY domain	n/a	M	n/a	n/a	n/a	n/a	n/a	ASD	n/a	n/a	n/a	n/a	n/a	Satter strom et al., 2020
c.117 1_11 72TG [1] [c.111 4_111 5TG[1]]	p.Cys3 91_Gln 392deli ns* [p.Cys3 72_Gln 373deli ns*]	SPRY domain	n/a	n/a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n/a	seizur es	n/a	n/a	Gillent ine et al., 2021

c.124 3del [c.118 6del]	p.Asp4 15Metf s*3 [p.Asp3 96Metf s*3]	SPRY domain	3.2	F	hypot onia	yes	mild	yes	yes	no	no	y e s	yes	atrial septal defect	additional genetic defect (duplication (16p13.11p12.3)), centriculomegaly - asymmetry (left>right)	Taylor et al., 2022
c.128 2delC [c.12 25del C]	p.Gly42 9Alafs* 53 [p.Gly4 10Alafs *53]	SPRY domain	8	F	no	yes	moderate	yes	no	no	autistic aggressive behaviours	n o	seizur es	atrial septal defect	craniofacial dysmorphisms (high forehead, thin upper lip, thick lower lip, low ear position)	Song et al., 2021
c.136 8A>C [c.131 1>C]	p.Glu45 6Asn [p.Glu4 37Asn]	SPRY domain	n/a	n / a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	seizur es	n/a	n/a	Gillent ine et al., 2021
c.142 4_14 25ins TC [c.13 67_1 368in sTC]	p.Ile47 6Profs* 7 [p.Ile45 7Profs* 7]	Connec tion	21	F	n/a	yes	severe	yes	yes	ASD	hand flapping, aggression with violent episodes, stereotypic al hand wringing	n o	seizur es	no	hypertelorism, brachydactyly, tapering fingers, prominent eyebrows, elongated palpebral fissures, prominent nasal bridge, overhanging columella, thin upper lip, short stature, obstipation	Yates et al., 2017
c.143 5G>A [c.13 78G> A]	p.Val47 9Ile [p.Val4 60Ile]	Connec tion	3	M	no	no	n/a	no	no	no	no	n o	no	no	refractory ulcerative colitis	Gillent ine et al., 2021
c.145 0C>T [c.13 93C> T]	p.Arg48 4* [p.Arg4 65*]	Connec tion	16.2	M	no	yes	n/a	yes	yes	n/a	n/a	n / a	seizur es	no	short stature, hearing impairment, hirsutism, broad thumb, cryptorchidism	Durkin et al., 2020
Walker A box, from 503 to 511 aa [484 to 492 aa]																
c.151 6C>T [c.14 59>T]	p.Pro50 6Ser [p.Pro4 87Ser]	AAA+ domain	n/a	n / a	no	yes	n/a	no	no	ASD	no	n o	seizur es	no	growth delay	Gillent ine et al., 2021
Walker A box, from 503 to 511 aa [484 to 492 aa]																

c.1615-1G>A [c.1558-1G>A]	canonical splice site	AAA+ domain	18	M	no	yes	n/a	no	no	no	no	no	yes	no	hypoplastic 13th rib, vertebral segmentation anomalies, sagittal cleft of the vertebra, delayed skeletal maturation, bifid great toe, panhypopituitarism, hypertension, additional genetic defect in SMAD1	Need et al., 2012/ Zhu et al., 2015
c.1617dup [c.1560dup]	p.Ala540Serfs*4 [p.Ala521Serfs*4]	AAA+ domain	26.5	F	no	yes	severe	yes	yes	no	aggressive behaviours, anxiety, depression	no	yes	no	prominent epicanthus, prominent nose, cleft palate, low set ear, low-set hairline, scoliosis, hallux valgus, short hand, big phalanx, alopecia, eczema	Taylor et al., 2022
c.1626_1627insA [c.1569_1570insA]	p.Lys543Leufs*11 [p.Lys524Leufs*11]	AAA+ domain	15	F	no	yes	moderate	yes	yes	no	hand flapping, special education needs	n/a	seizures	no	craniofacial dysmorphisms (synophrys, deepset eyes, long eyelashes), puffy dry skin over hands and feet, wide based gait, thin upper lip, short philtrum	Yates et al., 2017
c.1626_1627insTT [c.1569_1570insTT]	p.Lys543Leufs*11 [p.Lys524Leufs*11]	AAA+ domain	n/a	F	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	additional genetic defects: SLC41A2 splice region variant, TANC1 missense variant	DDD Study et al., 2017
c.1641delA [c.1584delA]	p.Asp548Ilefs*5 [p.Asp529Ilefs*5]	AAA+ domain	12	F	feeding	yes	n/a	yes	yes	no	abnormal temper tantrums, stereotypical hand wringing	yes	seizures	no	small stature, mild synophrys, short nose, short fingers, short toes, growth delay, 5th finger clinodactyly, alternating divergent strabismus	Durkin et al., 2020

c.1664del [c.1607del]	p.Leu555Argfs*51 [p.Leu536Argfs*51]	AAA+ domain	8.5	F	no	yes	moderate	yes	yes	no	no	n / a	seizures	no	thin upper lip, decreased body weight, brain imaging abnormality (non-progressive T2 and FLAIR hyperintense lesions involving the deep white matter bilaterally, with sparing of the occipital lobes and basal ganglia), microcephaly	Yates et al., 2017
c.1665_1666delGT [c.1608_1609delGT]	p.Leu556Alafs*12 [p.Leu537Alafs*12]	AAA+ domain	8.8	M	hypotonia, feeding difficulties, haematemesis, jaundice	yes	n/a	yes	yes	n/a	n/a	yes	yes	no	craniofacial dysmorphisms (synophrys, depressed nasal bridge, diastasis of upper incisors, high forehead), clinodactyly of fifth finger bilaterally, preaxial polydactyly of right foot, syndactyly of second-third toe of right foot, single palmar crease bilaterally, short stature	Taylor et al., 2022
c.1681delC [c.1624delC]	p.Gln561Serfs*44 [p.Gln542Serfs*44]	AAA+ domain	n/a	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n / a	n/a	n/a	additional genetic defects: PLEKHA7 missense variant, KIF16B synonymous variant	DDD Study et al., 2017
c.1681del [c.1642del]	p.Gln561Serfs*45x2 [p.Gln542Serfs*45]	AAA+ domain	17	M	n/a	yes	severe	yes	n/a	n/a	n/a	n / a	yes	no	non-convulsive status epilepticus, thin corpus callosum	Depienne et al., 2017

			13.9	M	feeding difficulties	yes	n/a	yes	yes	n/a	n/a	n / a	seizures	no	developmental regression, bilateral valgus deformity, kyphosis, joint laxity, prominent fingertip pads, broad thumb, triangular face, wide mouth, highly arched eyebrows, epicanthus, long palpebral fissure, widely spaced teeth	Durkin et al., 2020
(mosaic) c.168 1C>T [c.16 42C>T]	p.Gln561* x2 [p.Gln542*]	AAA+ domain	6.3	M	n/a	yes	moderate	yes	yes	n/a	n/a	y e s	yes	n/a	n/a	Depienne et al., 2017; Gillentine et al., 2021
			n/a	n / a	no	yes	n/a	no	no	no	no	no	no	no	no	no
Walker B box, from 571 to 579 aa [552 to 560 aa]																
c.171 4C>T [c.16 57C>T]	p.Arg572* x2 [p.Arg553*]	AAA+ domain	7	F	n/a	yes	n/a	yes	n/a	no	hand flapping	y e s	seizures	no	short stature, strabismus, narrow brows, broad nose, developmental regression, absence of left olfactory sulcus and left olfactory bulb, suspected mild subtype of holoprosencephaly	Leduc et al., 2017;
			n/a	F	no	yes	n/a	no	no	no	no	no	no	seizures	no	agenesis septum pellucidum, dysmorphic facial features
Walker B box, from 571 to 579 aa [552 to 560 aa]																

c.1743+1G>C [c.1686+1G>C]	p.? splice site variant	AAA+ domain	10	F	feeding difficulties	yes	severe	yes	yes	no	can be affectionate but stubborn, bruxism, recurrent hand flapping	no	seizures	no	lost ability to speak but now has 15 words, craniofacial dysmorphism (narrow palpebral fissures, mild posterior rotation of ears, high palate, maligned upper incisors), slender hands and feet, feeding difficulties, breath-holding/apnoeic episodes, ataxic broad-based gait, growth delay, short stature	Durkin et al., 2020
c.1744_1767del [c.1687_1710del]	p.Thr582_Gln589del [p.Thr563_Gln570del]	AAA+ domain	17	F	no	yes	severe	yes	no	ASD	aggressive outbursts	yes	yes	no	asymmetrical smile, unsteady gait, few small foci of periventricular nodular heterotopia, short stature, subependymal gray matter heterotopia, chiari I malformation, short neck, deeply set eye, hypertelorism, low-set ears, bulbous nose tip, odontogenic keratocyst, teeth abnormalities	Allen et al., 2013; Brams et al., 2017
c.1755_1756ins CCTCT [c.1698_1699ins CCTCT]	p.Ala586Profs*22 [p.Ala567Profs*22]	AAA+ domain	18	M	n/a	yes	n/a	yes	n/a	n/a	n/a	n/a	seizures	no	other genetic defects: WDR45B missense variant, myoclonic spasms, phonic tics, abnormal face shape	DDD Study et al., 2017
c.1801C>T [c.1744C>T]	p.Arg601* [p.Arg582*]	AAA+ domain	16.7	M	no	yes	cognitive impairment	yes	n/a	autistic	n/a	n/a	seizures	no	craniofacial dysmorphisms (abnormal facial shape, prominent heavy wide arched eyebrows, triangular face, flat cheekbones), short stature	Durkin et al., 2020

c.181 2dup [c.17 55dup]	p.Arg60 5* [p.Arg5 86*]	AAA+ domain	n/a	n / a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	seizur es	n/a	n/a	Gillentine et al., 2021
c.181 3_18 14dup/c.1 815insAC [c.17 56_1 757dup/c. 1758insAC]	p.Val60 6* [p.Val5 87*]	AAA+ domain	n/a	F	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	n/a	n/a	n/a	Wang et al., 2020
c.183 4G>A [c.17 77G> A]	p.Asp6 12Asn [p.Asp5 93Asn]	AAA+ domain	18	M	feeding difficulties	yes	severe	yes	yes	likely autism	behavioral issues	yes	yes	no	additional genetic defect: Loss of clone RP11-38O23 at Xp11.23, hypertonic in extremities, hypertelorism, bilateral epicanthal folds, broad nasal root, high and narrow palate, dysplastic pinnae, posteriorly rotated low set ears, bitemporal narrowing, prominent forehead, 2nd-3rd toe syndactyly, prominent and large central incisors, spastic diplegia	Taylor et al., 2022

c.183 6del [c.17 79del]	p.Tyr61 3Ilefs*1 1 [p.Tyr5 94Ilefs* 11]	AAA+ domain	6	F	no	yes	moderate to severe	yes	yes	no	sensory issues, self-harm, aggressive towards family members, destructive to toys, no sense of danger	y e s	seizur es	no	hypermetropia, bilateral postaxial hand polydactyly, right foot polydactyly, marked ligamentous laxity, short stature, symmetrical white matter signal abnormalities throughout the brain, however, this predominantly affected the periventricular regions of both the parietal lobes and frontal lobes. There were also stable simple cysts adjacent to the temporal horn of the right lateral ventricle in the region of the hippocampus	Durkin et al., 2020
c.184 6_18 47ins CA [c.17 89_1 790in sCA]	p.Lys61 6Thrfs* 3 [p.Lys5 97Thrfs *3]	AAA+ domain	2	F	no	yes	n/a	no	no	no	no	n o	seizur es	atrial septal defect, ventricu lar septal defect	congenital sensorineural hearing impairment	Gillent ine et al., 2021
c.185 2C>T [c.17 95C> T]	p.Gln61 8* [p.Gln5 99*]	AAA+ domain	n/a	M	n/a	yes	n/a	n/a	n/a	ASD	n/a	n / a	n/a	no	n/a	Wang et al., 2020
c.186 5_18 67del [c.18 08_1 810d el]	p.Glu62 2del [p.Glu6 03del]	AAA+ domain	n/a	n / a	n/a	n/a	n/a	n/a	n/a	ASD	n/a	n / a	n/n	n/a	n/a	Gillent ine et al., 2021

c.186 8dup [c.181 1dup]	p.Glu62 4Argfs* 24 [p.Glu6 05Argfs *24]	AAA+ domain	5.5	F	n/a	yes	deficient	yes	yes	n/a	n/a	y e s	yes	no	hyperlaxity, deep-set eyes with epicanthal folds and bags below the eyes, narrow palate, grey sclerae, short 2nd digit on both hands, palmar crease on the left hand	de Kovel et al., 2016/ Depie nne et al., 2017
c.192 0delT [c.18 63del T]	p.Phe6 40fs* [p.Phe6 21fs*]	AAA+ domain	6	M	n/a	yes	n/a	yes	no	no	no	y e s	no	no	thrombocytopenia	Gillent ine et al., 2021
c.192 5_19 26del [c.18 68_1 869d el]	p.Leu6 42Profs *5 [p.Leu6 23Profs *5]	AAA+ domain	11.3	M	no	yes	n/a	no	yes	n/a	n/a	n / a	seizur es	n/a	regression, microcephaly, sparse and thin eyebrows, short palm, short foot, abnormal cranial MRI	Durkin et al., 2020
c.201 4delG [c.19 57del G]	p.Glu67 2Lysfs* [p.Glu6 53Lysfs *]	AAA+ domain	4	M	n/a	yes	n/a	no	no	ASD	no	y e s	seizur es	atrial septal defect	dysphagia, bifid uvula, left branchial cleft remnant, bilateral hydronephrosis, cryptorchidism	Gillent ine et al., 2021
c.208 3_20 84del [c.20 26_2 027d el]	p.Ser69 5Trpfs* 6 [p.Ser6 76Trpfs *6]	IDR 2	12.6	F	feedin g difficu lties	yes	n/a	yes	yes	n/a	n/a	n / a	seizur es	n/a	short stature, delayed eruption of permanent teeth, high palate, median cleft lip and palate, widely spaced teeth, childhood-onset truncal obesity, short foot	Durkin et al., 2020
c.208 5dup [c.20 28du p]	p.Gly69 6fs* [p.Gly6 77fs*]	IDR 2	n/a	n / a	n/a	yes	n/a	no	no	ASD	no	n o	seizur es	no	brain imaging abnormality	Gillent ine et al., 2021
RGG motif, from 714 to 739 aa 695 to 720 aa]																

c.216 7+35 *415 6del [c.211 0+35 *415 6del]	p.? deletion of last three exons	IDR 2 + RGG box	5.7	M	no	yes	moderate	yes	yes	no	some aggressive outbursts, no sense of danger	y e s	seizur es	trivial tricuspi d regurgit ation	craniofacial dysmorphisms (left-sided plagiocephaly, torticollis, low-set ears, prominent forehead, epicanthic folds (L>R), prominent nasal tip, sparse hair, mild hypermetropia), exotropia, unsteady gait, frequent falls	Durkin et al., 2020
c.219 9G>T [c.211 42G> T]	p.Arg73 3Ser [p.Arg7 14Ser]	RGG box	n/a	n / a	n/a	yes	n/a	no	no	ASD	n/a	n / a	seizur es	no	growth delay, eye abnormality	Gillentine et al., 2021
c.219 7_21 99del AGG [c.211 40_2 142d elAG G]	p.Arg73 3del [p.Arg7 14del]	RGG box	n/a	n / a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n / a	seizur es	n/a	n/a	Oates et al., 2018
RGG motif, from 714 to 739 aa 695 to 720 aa]																
c.227 0_22 71del CT [c.221 13_2 214d elCT]	p.Pro75 7Argfs* 7 x2 [p.Pro7 38Argfs *7]	IDR 2	7	F	n/a	yes	n/a	yes	yes	n/a	repetitive compulsive behaviour, self- stimulation	y e s	seizur es	no	strabismus, short forehead, prominent columella, overfolded helices, constipation, mild hypertrichosis, hypermobility, unilateral single palmar crease, puffy hands and feet, brachydactyly	Leduc et al., 2017;
			11	F	hypot onia, feedin g difficu lties	yes	moderate to severe	yes	yes	autism	no	y e s	yes	no	aspecific dysmorphism, tapering fingers, cutis marmorata	Taylor et al., 2022

c.227 7dup [c.22 20dup]	p.Pro76 0Serfs* 5 [p.Pro7 41Serfs *5]	IDR 2	10	M	hypot onia	yes	n/a	yes	yes	autistic feature s	hand clapping	y e s	yes	n/a	hyperventilation, apnea, hyperpnea, syncope, cyanosis, macrocephaly, hypertelorism, depressed nasal bridge, broad nasal tip, full lips, grade 1 atrioventricular block, severe oxygen desaturation	Spagn oli et al., 2021
2299 _230 _2del [2242 _224 5AAC A]	p.Asn7 67fs* [p.Asn7 48fs*]	IDR 2	n/a	n / a	no	yes	n/a	no	no	no	no	n o	no	no	short stature, macrocephaly	Gillent ine et al., 2021
c.229 9_23 02del c.229 9_23 02AA CA [c.22 42_2 245]	p.Asn7 67Glufs *66 (mosaic) [p.Asn7 48Glufs *66]	IDR 2	16.8	F	n/a	yes	moderate to severe	yes	yes	n/a	disturbed sleeping pattern	y e s	no	no	spastic diplegia, dilated ventricles (presumed aqueduct stenosis), small splenium of corpus callosum, short stature, macrocephaly	Monro e et al., 2016, Depie nne et al., 2017
c.230 2_23 03del AG [c.22 45_2 246d elAG]	p.Gly76 9fs* [p.Gly7 50Glufs *]	IDR 2	n/a	n / a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n / a	seizur es	n/a	n/a	Gillent ile et al., 2021
c.230 4_23 05del [c.22 47_2]	p.Gly76 9fs* [p.Gly7 50Glufs *]	IDR 2	5	M	n/a	yes	n/a	no	no	no	no	n o	seizur es	no	short digits, broad thumb, astigmatism, retractile testes	Gillent ile et al., 2021;

248del]	p.Gly769Glufs*83 [p.Gly750Glufs*83]		9.8	F	no	yes	moderate	yes	yes	no	hyperphagia	no	yes		ventriculomegaly, cyst on pineal gland, thin upper vermillion, short stature, 2nd-3rd toe syndactyly, cyclic neutropenia	Taylor et al., 2022
c.2319_2320del [c.2262_2263del]	pGlu774fs* [pGlu755fs*]	IDR 2	n/a	n/a	n/a	yes	n/a	no	no	no	no	yes	seizures	no	no	Gillentine et al., 2021
c.2365C>T [c.2308C>T]	p.Arg789* [p.Arg770*]	IDR 2	2.3 n/a	F M	n/a n/a	n/a yes	n/a n/a	n/a yes	n/a no	n/a ASD	n/a no	n/a no	yes seizures	n/a no	n/a regression	Veith et al., 2014; Wang et al., 2020
c.2408A>G [c.2351>G]	p.Asn803Ser [p.Asn784Ser]	IDR 2	n/a	n/a	n/a	yes	n/a	n/a	n/a	n/a	n/a	n/a	seizures	n/a	n/a	Gillentine et al., 2021
c.2425-2A>G [c.2368-2A>G]	n/a x2	IDR 2	n/a 5.5	n/a M	n/a hypotonia, feeding difficulties, reflux	yes yes	n/a moderate	n/a yes	n/a yes	n/a autism	n/a no	n/a yes	seizures no (had febrile seizures)	n/a	n/a medio-posterior atrophy, occipital atrophy, global cerebral atrophy, strabismus, hypermetropia, thin upper vermillion, large ASD, hyperlaxity	Gillentine et al., 2021 Taylor et al., 2022
c.2425-3C>A [c.2368-3C>A]	n/a	IDR 2	22	M	n/a	yes	Level of cognitive functioning - 2.5 years, level of social emotiona	yes	yes	autism	n/a	n/a	yes	no	n/a	Depienne et al., 2017

							l functionin g - 9 months									
c.247 1T>C c.247 2T>A [c.24 14T> C, c.241 5T>A]	p.Tyr82 4* [p.Tyr8 05]	IDR 2	33	M	n/a	yes	severe	n/a	n/a	n/a	n/a	n / a	seizur es	no	regression	Carvill et al., 2013

Table B. HNRNPU-related neurodevelopmental disorder patients and phenotype information from literature.

