

Novel proteomics approaches to study
S-acylation in biomolecular condensate
regulation and disease

Amina Nigmatulina

A thesis submitted for the degree of

Doctor of Philosophy

Imperial College London

Department of Chemistry

June 2025

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND). Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law

Statement of originality

I, Amina Nigmatulina, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Biomolecular condensates are membraneless organelles formed through liquid-liquid phase separation (LLPS), driven by multivalent interactions. These dynamic compartments play key roles in a wide range of cellular processes, and their dysregulation is linked to numerous diseases. Whilst many protein post-translational modifications (PTMs) are known to regulate condensate assembly and disassembly, the role of lipid PTMs, particularly S-acylation, remains largely unexplored. S-acylation involves the covalent attachment of long-chain fatty acids (C16-C22) onto cysteines, mediated by zDHHC enzymes and reversed by acyl protein thioesterases (APTs). S-acylation regulates myriad cellular functions, and its dysregulation is implicated in many diseases. Given its hydrophobic and reversible nature, it was hypothesised that S-acylation could influence condensate dynamics via hydrophobic interactions. However, studying S-acylation is challenging due to its inherent biochemical properties, enzyme redundancy, and limited analytical tools.

Here, I initially developed a novel double enrichment proteomic method to systematically study S-acylation dynamics within stress granules (SGs), a model condensate system. This method combined APEX2 proximity labelling and S-acyl exchange (APEX-AX) techniques to investigate S-acylation changes during SG assembly upon stress. However, initial experiments revealed the incompatibility of the APEX2 approach with S-acylation profiling due to lability of the thioester bond. Separately, addressing the broader scarcity of analytical tools, I developed the **S**-acylation **P**rofilng by **P**hosphonate Tagging **P**roteomics (**SAPPHIRE**) method, which enables high-throughput and site-specific analysis of S-acylation on a proteome-wide scale. SAPPHIRE demonstrated improved sensitivity and specificity compared to existing S-acylproteomic techniques, enabling the identification of >4,000 putative S-acylation sites across multiple human cell lines. Furthermore, SAPPHIRE facilitated pioneering analyses of S-acylation patterns in human patient tissues and provided the first site-specific S-acylation profile in *T. gondii* parasites. Collectively, this thesis addresses key challenges in studying S-acylation and establishes SAPPHIRE as a powerful platform for comprehensive analysis of S-acylation across diverse biological systems and disease contexts.

Acknowledgements

Reflecting on the past four years, I am most grateful to have met so many incredible people, who have supported me and enriched my life in ways I never imagined.

Firstly, I would like to thank Ed, for all the amazing opportunities you have entrusted me with, for your encouragement, insightful and brilliant guidance, and for nurturing confidence and independence in me as a scientist. But mostly thank you for your kindness and understanding, with visa, political or any other challenges, you have made all that is possible (and more!) for a supervisor to help lessen the stresses of the situation. I feel incredibly lucky to have been your student and a part of your group for all these years. I also want to thank Dan, for being my MRes supervisor and introducing me to the group. Thank you for the fun, late-night Cohort 1 shifts, which showed me just how enjoyable experimental science could be.

A huge thank you to Manuel, Sarah, Jack and Anne, for generously reviewing chapters of this thesis. Your feedback and encouraging words really helped me get this across the finish line.

I have been incredibly lucky to work every day alongside my friends – a privilege so rare in adulthood. To Matt and Thomas, thank you for the unexpected friendship that quickly became the most fun and supportive space filled with chaotic conversations, three-jug nights, crosswords, scientific discussions and hype talks, you really are not like other guys. To Shradda and Lame, my day one MRes ladies, thank you for your unwavering friendship and support, I genuinely don't think I could've gotten through it all without you. Thank you for every rant, coffee break, dinner and brunch. To Sarah, thank you for your boundless kindness, for inviting me to lunch one day and never stopping since, and all the fun times, from simple pub days to crazy sculpture visuals and James Blake. To Jack, thank you for all the proteomics guidance, the much-needed London tour during this write up, and tennis matches. An honourable mention goes to you and Matt for the iconic Dresden trip, which will forever be one of the highlights of my PhD (3BM forever). To Archie, thank you for all the fun lunches and lab conversations, some of which helped me through difficult moments. To Zhangping and Fangyuan, the best office neighbours one could hope for, thank you for your support, calmness, kindness and laughs, my PhD journey is largely so amazing because of our bay. To Anne, Aidan, Xinyue, Amrita, Tom J, Tom B, Anya, Dasha, Edward, Erika, Masha, James, Janine, Jana, Adam, Vanessa, Daisy, the Crick crew, and everyone else in the group, thank you for creating such a fun, friendly and warm atmosphere, I have learned so much from all of you.

To my friends outside the lab: to Adiya, thank you for our lifelong friendship that I've leaned on time and again. Thank you for making my time in Almaty, Brighton and London so memorable, from F1 Sundays, Peggy Gou days, and the insane Italy trip. To Angelina, Diana, Bishka, thank you for always being there.

To Robbie, I am most grateful to have met you during my PhD, you have been my peace and light in this journey. Thank you for your incredible heart, kindness and patience (especially during the write-up period!). Thank you for making life feel magical, for all the unforgettable trips we've had (and your incredible planning), and for being the best trial audience for all my posters and presentations, you know more about S-acylation than anyone else in fintech. Here's to celebrating another milestone together, and many more to come.

To my incredible family, to Dasha, you've been my rock for as long as I can remember, and I am so grateful we got to live together in London these few years. Thank you for always being there, for all the "horses" outings, Victus soul craziness, and every rant about anything and everything. To Markhabat and Lesha, thank you for your continuous support, for all the amazing times in Portugal and for making us a padel family. To Dasha, Platon and Aileen, thank you for being the most wonderful nieces and nephew, all the fun we have when we are together washes away any stress and makes life brighter. To Тетя Надя, спасибо тебе за твою любовь и поддержку, которую я всегда ощущаю, даже через страны.

Lastly and most importantly, thank you to my parents, who have supported every dream, inspired every step, and stood by me through it all. Спасибо вам родители за вашу безграничную любовь, поддержку и веру в меня. Спасибо за то, что всегда были рядом, вдохновляли и подталкивали меня не сдаваться и продолжать свой путь в науке. Все мои достижения и успехи я посвящаю вам. Ваша непоколебимая вера в меня — моя самая большая мотивация и источник вдохновения.

Statement of Collaboration

The results of this thesis have been supported by collaborations with members of the Tate Group at Imperial College London. The key contributions are listed below:

- Dr. Aidan Pidd performed all tissue extractions for experiments presented in **Chapter 4.2**
- Dr. Janessa Grech performed *Toxoplasma gondii* maintenance and lysis for experiment presented in **Chapter 4.3**
- Dr. Anne Schuhmacher carried out THP1 cell culture, treatments and lysis for the pyroptosis experiment in **Chapter 4.4**
- Daria Berezina carried out peptides loading onto Evotips (Evosep) for all proteomics experiments
- Dr. Jianan Lu or Dr. Jack Houghton ran samples on the LC-MS/MS
- Matthew White optimised the Oasis-HLB 96-well plate desalting protocol
- U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1 and U2OS-APEX2-NES cell lines used in **Chapter 2** were designed, generated and gifted by the Eran Hornstein lab at Weizmann Institute of Science

Publications and Presentations

Publications:

- Nigmatulina A, White M.E.H., Schuhmacher A., Houghton J.W., Tate E.W. **SAPPHIRE: A novel proteomics platform for high-throughput identification of protein S-acylation using phosphonate handles.** *In preparation*

Oral presentations:

- Nigmatulina A. *et al.* **SAPPHIRE: A novel proteomics platform for high-throughput identification of protein S-acylation sites using phosphonate handles.** *Presented at the 14th International Activity Based Protein Profiling Meeting (London, UK) – April 2025*
- Nigmatulina A. *et al.* **Probing the role of S-acylation in biomolecular condensate regulation.** *Presented at Imperial College London Postgraduate Symposium (London, UK) – June 2024*

Poster presentations:

- Nigmatulina A. *et al.* **SAPPHIRE: A novel proteomics platform for high-throughput identification of protein S-acylation sites using phosphonate handles.** *Presented at the 23rd Human Proteome Organization (HUPO) World Congress (Dresden, DE) – October 2024*
- Nigmatulina A. *et al.* **Probing the role of S-acylation in biomolecular condensate regulation.** *Presented at EMBL: Cellular mechanisms driven by phase separation (Heidelberg, DE) – May 2024*
- Nigmatulina A. *et al.* **Probing the role of S-acylation in biomolecular condensate regulation.** *Presented at Imperial College London Postgraduate Symposium (London, UK) – July 2023*

Table of Contents

Copyright Declaration	2
Statement of originality.....	2
Abstract	3
Acknowledgements	4
Statement of Collaboration	6
Publications and Presentations	7
Table of Contents	8
List of Figures	13
List of Tables	17
Abbreviations	18
Chapter 1. Introduction	23
1.1 Biomolecular Condensates	24
1.1.1 Liquid-Liquid Phase Separation.....	25
1.2 Dynamic regulation of biomolecular condensates	26
1.2.1 Post translational modifications	26
1.2.2 Membrane surfaces.....	28
1.3 Condensate dysregulation as a therapeutic target	29
1.4 Protein lipidation	30
1.5 S-acylation	32
1.5.1 zDHHC and APT enzymes as regulators of dynamic S-acylation	33
1.5.2 S-acylation in health and disease.....	34
1.6 Tools to study S-acylation	36
1.6.1 Mass spectrometry-based tools.....	36
1.6.2 Small molecules.....	38
1.6.3 Challenges in studying S-acylation.....	40
1.7 Hypothesis	40

1.8	Project aims	42
Chapter 2.		
<i>Investigating the role of S-acylation in stress granules by chemical proteomics</i>		45
2.1	Introduction	46
2.2	Bioinformatic analysis of SG proteome and S-acylation	48
2.2.1	Curation of the SG protein composition list	48
2.2.2	Comparison against SwissPalm	51
2.3	Conceptualisation of a double enrichment proteomics approach	52
2.4	Validation of APEX2 activity and SG localisation	55
2.4.1	Cell line and activity validation	55
2.4.2	Confocal Microscopy	57
2.5	APEX-AX for S-acylated SG proteome identification	59
2.5.1	Pilot experiment	59
2.5.2	APEX-AX: Phosphonate as a mass tag	61
2.6	Discussion	65
2.6.1	Double or nothing: Challenges in double enrichment proteomics	65
2.6.2	Challenges of profiling S-acylation	67
2.7	Chapter summary	68
<i>Chapter 3. SAPPHIRE: a novel proteomics platform for sensitive and high-throughput analysis of S-acylation</i>		70
3.1	Introduction	71
3.2	SAPPHIRE Proof-of-Concept	72
3.2.1	SAPPHIRE Conceptualisation	72
3.2.2	Pilot Experiment	74
3.3	Optimisation of the SAPPHIRE method	77
3.3.1	IMAC Enrichment	77
3.3.2	Free cysteine alkylation	79
3.3.3	Increasing the throughput and ease of the workflow	81
3.4	Profiling S-acylation sites in HEK293T cells with SAPPHIRE	82

3.4.1	Results.....	83
3.4.2	Functional and structural analysis of S-acylated sites identified by SAPPHIRE.....	85
3.4.3	S-acylation consensus sequence and cysteine accessibility analysis	87
3.5	Western blot validation of the identified sites	91
3.6	Benchmarking SAPPHIRE against existing technologies	92
3.6.1	Site-specific ABE vs SAPPHIRE.....	92
3.6.2	Nanographite fluoride-based solid-phase extraction (nGF-SPE) vs SAPPHIRE	96
3.7	Discussion	99
3.7.1	Strengths.....	99
3.7.2	Limitations	100
3.7.3	The complexity of hydroxylamine-sensitive S-acyl intermediates and non- enzymatic modifications	102
3.8	Chapter summary and future work	104
Chapter 4. Application of SAPPHIRE for S-acylated proteome profiling across disease models and organisms		107
4.1	Introduction	108
4.2	Tumour tissue analysis	108
4.2.1	Breast cancer	109
4.2.2	Colorectal cancer.....	111
4.2.3	Conclusions.....	115
4.3	<i>Toxoplasma gondii</i>	115
4.3.1	Site-specific S-acylated proteome profiling	116
4.3.2	Conclusions.....	118
4.4	Pyroptosis	118
4.4.1	GSDMD Cys191 S-acylation.....	121
4.4.2	NLRP3 S-acylation.....	125
4.4.3	MCC950 blocks GSDMD palmitoylation	126
4.4.4	Conclusions.....	127

4.5	Stress Response.....	128
4.5.1	Limitations	130
4.5.2	Conclusions.....	132
4.6	Discussion	133
4.6.1	Site discrepancy and enzyme redundancy.....	133
4.6.2	Site summary	135
4.7	Chapter summary.....	136
Chapter 5. Conclusions and Future Perspectives.....		138
Chapter 6. Materials and Methods.....		142
6.1	General Procedures.....	143
6.1.1	Cell culture	143
6.1.2	Acetonitrile (MeCN)-based protein precipitation	143
6.1.3	MeCN-based protein precipitation: 96-well plate format	144
6.1.4	SDS-PAGE	144
6.1.5	Western blot.....	145
6.1.6	Patient tissue preparation.....	145
6.2	Cell Treatments	146
6.2.1	Pyroptosis	146
6.2.2	<i>Toxoplasma gondii</i>	146
6.2.3	Stress Response.....	146
6.3	SAPPHIRE Proteomics	147
6.3.1	Buffers	147
6.3.2	Cell Lysis.....	147
6.3.3	Free Cysteine alkylation and acyl-phosphonate exchange	147
6.3.4	Dephosphorylation	148
6.3.5	96-well plate-based Oasis-HLB desalting.....	148
6.3.6	Automated Zirconium (IV) Immobilised Metal Affinity Chromatography	149
6.4	SAPPHIRE version 1 (SAPPHIREv1) and optimisation	149
6.5	APEX2 proximity labelling.....	150
6.5.1	APEX2 in-cell biotinylation and cell lysis.....	150

6.5.2	Western Blotting.....	151
6.5.3	Confocal Microscopy.....	151
6.5.4	APEX-AX Proteomics.....	152
6.6	Cysteine-centric and intact S-acylation ID.....	154
6.6.1	Acyl-PEG Exchange (APE).....	154
6.6.2	Site-specific Acyl-Biotin Exchange (ssABE).....	155
6.6.3	nanographite fluoride-based solid-phase extraction (nGF-SPE)	156
6.7	General proteomics methods	157
6.7.1	Whole Proteome sample preparation.....	157
6.7.2	Evosep-timsTOF LC-MS/MS: DDA enrichment proteomics.....	157
6.7.3	Evosep-timsTOF LC-MS/MS: DIA whole proteome	158
6.7.4	nanoElute2-timsTOF LC-MS/MS: DDA nGF-SPE.....	158
6.8	Proteomics data analysis.....	159
6.8.1	Fragpipe	159
6.8.2	DIA-NN: Whole proteome data.....	160
6.8.3	Statistical analysis and visualisation	160
6.9	Bioinformatic analysis	161
6.9.1	Cysteine PTM, functional and pPSE annotation	161
6.9.2	Identification of S-acylation Motifs	162
6.9.3	Clinical variants annotations	162
6.9.4	Functional enrichment analysis	162
References.....		165
Chapter 7. Appendix.....		190

List of Figures

Chapter 1

Figure 1.1. Biomolecular condensates in eukaryotic cells	24
Figure 1.2. Liquid–liquid phase separation	26
Figure 1.3. PTMs and membranes as dynamic regulators of biomolecular condensates.....	27
Figure 1.4. Overview of constitutive lipid post-translational modifications	31
Figure 1.5. Overview of protein S-acylation	32
Figure 1.6. Structural feature of zDHHHC20	34
Figure 1.7. Mass spectrometry-based tools to probe S-acylation.....	37
Figure 1.8. Small molecule tools to study S-acylation.	39
Figure 1.9. S-acylation as a putative regulator of biomolecular condensates	41

Chapter 2

Figure 2.1. Stress granule formation and architecture	46
Figure 2.2. Comparison of stress granule composition datasets from five independent studies.	50
Figure 2.3. Comparative analysis of stress granule proteins with the SwissPalm S-acylation database	51
Figure 2.4. Schematic of APEX-AX: a double enrichment proteomics platform to identify stress-dependent changes in S-acylation of stress granule (SG) proteins	54
Figure 2.5. APEX2 construct design and cell line validation	55
Figure 2.6. APEX2 activity requires both biotin-phenol and H ₂ O ₂	56
Figure 2.7. APEX2-G3BP1 localises to stress granules in response to stress	59
Figure 2.8. Modified APEX-AX workflow: phosphonate as a mass tag.	62
Figure 2.9. Assessment of labelling and enrichment efficiency	63

Figure 2.10. APEX-AX identifies stress-dependent components but is unsuitable for profiling S-acylation.	64
---	----

Chapter 3

Figure 3.1. Schematic overview of the SAPPHIREv1 proteomics workflow.....	75
Figure 3.2. Pilot experiment results for SAPPHIREv1	76
Figure 3.3. High rate of false-positive identifications in the SAPPHIREv1 dataset.....	77
Figure 3.4. Optimisation and automation of Zr-IMAC enrichment	78
Figure 3.5. Optimisation of free cysteine alkylation conditions	80
Figure 3.6. Schematic of the optimised SAPPHIRE S-acylproteomic workflow.....	83
Figure 3.7. Site-specific S-acylated proteome profiling in HEK293T cells using the optimised SAPPHIRE protocol.	84
Figure 3.8. Site-specific S-acylated proteome profiling of HEK293T cells using SAPPHIRE...	84
Figure 3.9. Functional enrichment analysis of putative S-acylated proteins identified by SAPPHIRE in HEK293T cells.....	86
Figure 3.10. Post-translational modification (PTM) and functional annotation of putative S-acylation sites in HEK293T cells identified using SAPPHIRE	87
Figure 3.11. Surface exposure analysis of putative S-acylated cysteine sites identified in HEK293T cells using SAPPHIRE	90
Figure 3.12. Western blot analysis of putative S-acylation sites using the acyl-mPEG exchange (APE) assay in HEK293T cells.....	92
Figure 3.13. Schematic of the site-specific acyl-biotin exchange (ssABE) workflow.	93
Figure 3.14. Site-specific S-acylated proteome profiling in HEK293T cells using site-specific acyl-biotin exchange (ssABE)	94
Figure 3.15. Comparison of putative S-acylated sites and proteins identified by SAPPHIRE, ssABE, and the SwissPalm database	95

Figure 3.16. Limited S-acylated site overlap is observed between SAPPHIRE, ssABE, literature and SwissPalm datasets.	96
Figure 3.17. Comparison of S-acylated sites identified by nGF-SPE and SAPPHIRE methods.	97
Figure 3.18. Overview of hydroxylamine-sensitive cysteine events that could be captured by SAPPHIRE.....	103

Chapter 4

Figure 4.1. Site-specific S-acylated proteome profiling of breast cancer patient tissue with SAPPHIRE.....	109
Figure 4.2. Identification of breast cancer tumour-enriched S-acylation sites	110
Figure 4.3. SAPPHIRE identification of colorectal cancer tumour-enriched S-acylation sites	112
Figure 4.4. Comparison of colorectal cancer patients' tumour-enriched vs healthy-enriched S-acylated sites	113
Figure 4.5. KEGG pathway enrichment analysis of colorectal cancer tumour-enriched S-acylated sites.	114
Figure 4.6. Principal component analysis of colorectal cancer patient samples profiling with SAPPHIRE.....	115
Figure 4.7. Site-specific S-acylated proteome profiling of <i>Toxoplasma gondii</i> with SAPPHIRE	117
Figure 4.8. Comparative analysis of <i>T. gondii</i> S-acylation sites identified by SAPPHIRE	117
Figure 4.9. GSDMD Cys191 S-acylation is the key checkpoint in pyroptosis	120
Figure 4.10. Schematic of the experimental conditions for profiling global changes in S-acylation in response to pyroptosis	121
Figure 4.11. Site-specific S-acylated proteome profiling of THP1 cells during pyroptosis induction with SAPPHIRE	122
Figure 4.12. GSDMD Cys191 is S-acylated in response to pyroptosis.....	123

Figure 4.13. GSDMD Cys191 S-acylation during pyroptosis occurs independently of GSDMD expression.....	125
Figure 4.14. MCC950 blocks S-acylation of GSDMD Cys191.....	127
Figure 4.15. Identification of arsenite stress-dependent changes in global S-acylation	129
Figure 4.16. Site-specific S-acylated proteome profiling of HEK293T cells in response to arsenite stress using SAPPHIRE	131
Figure 4.17. Principal component analysis (PCA) of HEK293T SAPPHIRE profiling in response to stress.	132

Appendix

Figure 7.1. Clinical variant annotation of putative S-acylation sites in HEK293T cells identified using SAPPHIRE.....	192
Figure 7.2. S-acylation conserved sequence motif analysis. Using the putative S-acylated Cysteine sites identified in HEK293T cells by SAPPHIRE.....	194
Figure 7.3. Whole proteome analysis of colorectal cancer patient tissues	195
Figure 7.4 Global protein expression changes between healthy and tumour tissues of colorectal cancer patients.	196
Figure 7.5 SAPPHIRE identification of colorectal cancer tumour-enriched S-acylation sites.	197
Figure 7.6. Site-specific S-acylated proteome profiling of colorectal cancer patient tissue with SAPPHIRE.....	199
Figure 7.7. Whole proteome analysis of <i>Toxoplasma gondii</i>).....	200
Figure 7.8. Whole proteome analysis of THP1 cells in response to pyroptosis	201

List of Tables

Chapter 2

Table 2.1. Summary of proximity labelling proteomics technologies	48
Table 2.2. Summary of stress granule composition studies	49
Table 2.3. Summary of double enrichment proteomics approaches	65

Chapter 3

Table 3.1. Summary of LC-MS/MS-based S-acylproteomic approaches.....	73
Table 3.2. Summary of zDHHC enzymes' S-acylated sites identified in HEK293T cells with SAPPHIRE.....	85
Table 3.3. S-acylation consensus motif analysis	88
Table 3.4. Comparison of the SAPPHIRE and ssABE experimental features.....	96
Table 3.5. S-palmitoylation sites identified by replication of the nGF-SPE workflow	98

Chapter 5

Table 6.1. List of antibodies used in this study.	163
Table 6.2. SAPPHIRE experiments experimental conditions.....	164

Appendix

Table 7.1. Stress-dependent G3BP1 interactors identified by APEX-AX	191
Table 7.2. Putative S-acylated Cysteines identified by SAPPHIRE in HEK293T cells associated with pathogenic clinical variants	193
Table 7.3. Peptide quantification for CRC samples	197
Table 7.4. Top 10 colorectal cancer tumour-enriched S-acylated sites ranked by mean log ₂ FoldChange	198

Abbreviations

ABE	Acyl-biotin exchange
Acyl-RAC	Acyl resin-assisted capture
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
APE	Acyl-mPEG exchange
APEX-AX	APEX2 proximity labelling coupled to acyl-exchange proteomics
APT	Acyl protein thioesterase
Arg	Arginine
BG	Background
BP	Biotin Phenol
BSA	Bovine serum albumin
CAA	Chloroacetamide
CAM	Carbamidomethylation
CANX	Calnexin
CoA	Coenzyme A
CRC	Colorectal cancer
C_{sat}	Threshold concentration
CuAAC	Copper-catalysed alkyne-azide cycloaddition
Cys	Cysteine
CysPAT	Cysteine-specific phosphonate adaptable tag
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DMSO	Dimethyl sulfoxide
DTDP	2,2'-dithiodipyridine
EDTA	Ethylenediaminetetraacetic acid
eIF	Eukaryotic translation initiation factor
ER	Endoplasmic reticulum
EtOH	Ethanol
FA	Formic acid
FBS	Foetal bovine serum
FC	Fold change
FG	Foreground
FRAP	Fluorescence recovery after photobleaching

FTD	Frontotemporal dementia
FUS	Fused in sarcoma
G3BP1	Ras GTPase-activating protein-binding protein 1
GA	Glycolic acid
GO	Gene ontology
GPI	Glycosylphosphatidylinositol
GPX4	Glutathione peroxidase 4
GSDMD	Gasdermin D
HA	Hydroxylamine
hnRNPA	Heterogeneous nuclear ribonucleoprotein A
IAA	Iodoacetamide
IMAC	Immobilised metal affinity chromatography
IP	Immunoprecipitation
ISR	Integrated stress response
KD	Knockdown
kDa	Kilodalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
KF	KingFisher
KO	Knockout
LC/MS-MS	Liquid chromatography tandem mass spectrometry
LFQ	Label-free quantification
LLPS	Liquid-liquid phase separation
LPS	Lipopolysaccharide
Lys	Lysine
MeCN	Acetonitrile
MeOH	Methanol
MLCC	Metabolic labelling click chemistry
MMTS	Methylmethanethiosulfonate
MS	Mass spectrometry
NEM	N-ethyl maleimide
NEMhyd	Hydrated NEM
NES	Nuclear export signal
nGF-SPE	nanographite fluoride-based solid-phase extraction
NLRP3	NOD-like receptor protein 3
NMT	<i>N</i> -myristoyl transferase

NMTi	NMT inhibitor
OSM	Open search mode
PARylation	Poly-ADP-ribosylation
PASEF	Parallel accumulation serial fragmentation
PAT	Phosphonate adaptable tag
PCA	Principal component analysis
PCR	Polymerase chain reaction
pCys	Phosphonated Cysteine
PEG	Polyethylene glycol
PL	Proximity labelling
PM	Plasma membrane
PMA	Phorbol-12-myristate-13-acetate
pPSE	prediction-aware part-sphere exposure
PSM	Peptide-spectrum match
PTM	Posttranslational modification
RBP	RNA-binding protein
ROS	Reactive oxygen species
RT	Room temperature
SAPPHIRE	S-acylation Profiling by Phosphonate Tagging proteomics
SAPPHIREv1	SAPPHIRE version 1
SDS	Sodium dodecyl sulfate
Ser	Serine
SG	Stress granule
SILAC	Stable isotope labelling with amino acids in cell culture
SPD	Samples per day
TCEP	tris(2-carboxyethyl)phosphine
TDP-43	TAR DNA-binding protein 43
TEA	Triethanolamine
TFA	Trifluoroacetic acid
Thr	Threonine
TIA1	T-cell intracellular antigen-1
timsTOF	Trapped ion mobility spectrometry -Time of flight
TMT	Tandem mass tag
Tyr	Tyrosine
WT	Wildtype

zDHHC

Zinc-finger Asp-His-His-Cys motif

Chapter 1.

Introduction

1.1 Biomolecular Condensates

Cellular function depends on the spatiotemporal coordination of complex biochemical processes within the densely packed cellular space. To achieve this, the cellular milieu is organised into distinct compartments enabling concentration and sequestration of specific proteins and nucleic acids. Whilst the importance of classical phospholipid bilayer membrane-bound organelles (MBOs), such as the nucleus, mitochondria and endoplasmic reticulum (ER), have long been recognised, it is becoming increasingly appreciated that many essential cellular processes occur in compartments lacking a delimiting membrane (**Fig. 1.1**). These membraneless organelles (MLOs), historically referred to as cellular bodies, droplets, speckles, granules, puncta, or assemblies, are now collectively referred to as biomolecular condensates¹.

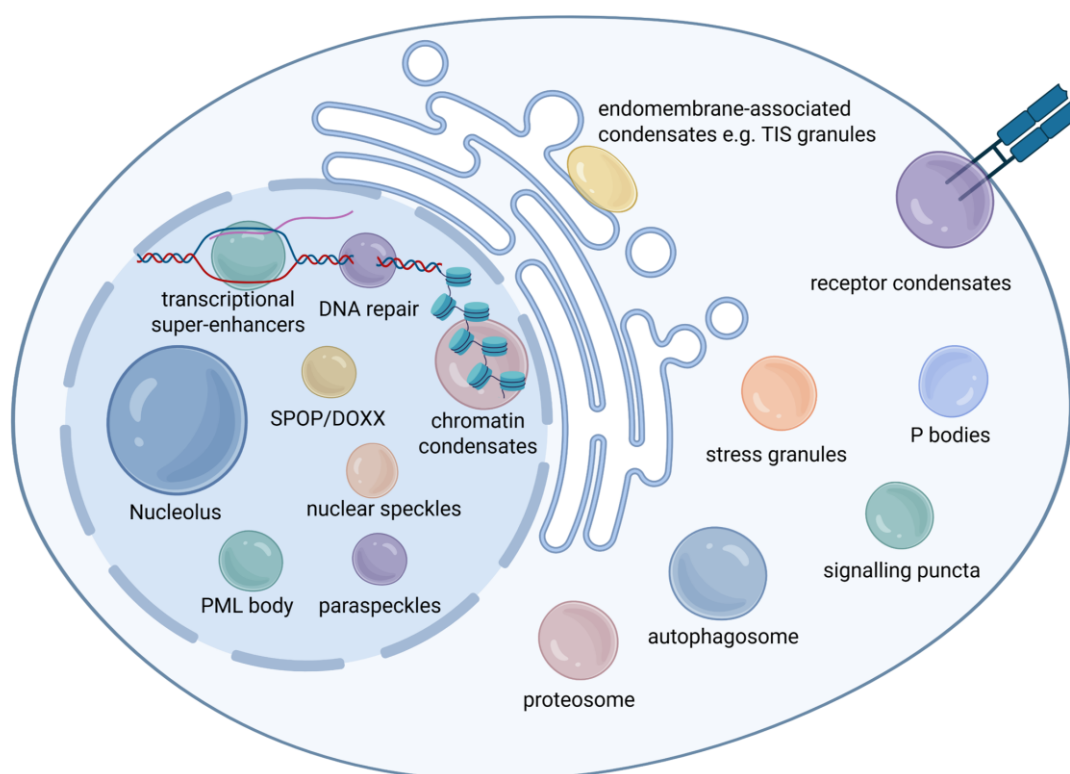


Figure 1.1. Biomolecular condensates in eukaryotic cells. Schematic of the membraneless organelles formed through liquid-liquid phase separation (biomolecular condensates) within the cytoplasm and nucleoplasm of the cell. Figure created with Biorender.com.

Even though the first observation of a condensate, the nucleolus, dates back to the 1830s, the field of condensate research has only recently become an area of substantial interest^{2,3}.

Biomolecular condensates range from nanometres to micrometres in size and are prevalent throughout both the cytoplasm and nucleoplasm including P granules, stress granules (SGs), transcriptional superenhancers, and DNA damage repair sites among others^{1,4}. Formed through liquid-liquid phase separation (LLPS), condensates exhibit liquid-like features and may form transiently or persist indefinitely depending on physiological or pathological context⁵.

On a molecular scale, condensates function to concentrate or exclude specific factors, either increasing reaction kinetics and specificity, or reducing kinetics and inhibition respectively⁶. At the cellular level, condensates are involved in myriad functions, including transcriptional regulation, cytosolic and nuclear stress responses, autophagy, trafficking of cellular components, RNA regulation and processing, protein quality control, as well as the formation of pre- and postsynaptic condensates in neuronal synaptic sites^{1,4,6,7}.

1.1.1 Liquid-Liquid Phase Separation

Biomolecular LLPS occurs in a concentration-dependent manner, driven by weak intra- and intermolecular interactions, such as hydrophobic, electrostatic, cation- π , and π - π , between multivalent biomolecules (**Fig. 1.2a**)^{8–11}. Multivalency is the key determinant of LLPS, and is a feature of proteins with repetitive motifs, or intrinsically disordered regions (IDRs), as well as RNA and DNA. At low concentrations, even multivalent molecules will exist in solution. However, their increasing concentration leads to a growing non-stoichiometric macromolecular assembly, which decreases in solubility whilst simultaneously promoting additional interactions (**Fig. 1.2b**)¹. Phase transition then occurs in a switch-like manner at the critical threshold concentration (C_{sat}), above which the thermodynamic forces favour phase separation to minimise the free energy¹².

The formation and composition of biomolecular condensates can be conceptualised through a general scaffold-and-client framework. Scaffold components are essential initiators of condensate assembly, whilst the client molecules are subsequently recruited to the condensates through interactions with the scaffolds. Clients are typically dispensable yet typically constitute the bulk of the condensate. A given condensate, can contain hundreds of distinct molecules which can exchange between the condensate and the surrounding environment and determine its material properties¹³.

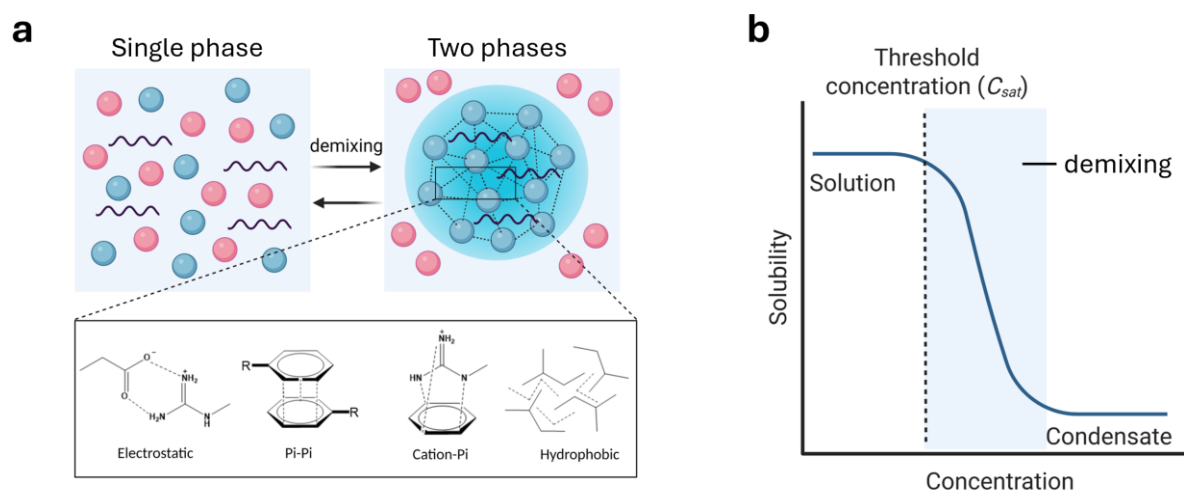


Figure 1.2. Liquid-liquid phase separation. **a)** Above a critical threshold concentration (C_{sat}), weak interactions among multivalent biomolecules can shift the free energy landscape to favour demixing into a condensed phase, driving condensate formation. **b)** At low concentrations, individual molecules remain in solution. As concentration increases, the solubility of the growing network of interacting molecules decreases until demixing occurs at C_{sat} . Each additional interactor provides further binding sites, attracting more molecules. This positive feedback loop underlies the switch-like behaviour of phase separation (blue curve) once a critical concentration is reached. Figure created with BioRender.com.

1.2 Dynamic regulation of biomolecular condensates

In addition to multivalency and concentration, several key parameters influence the thermodynamics of LLPS, including temperature, salt and pH. Whilst control over some thermodynamic factors lies outside of cellular control, cells evolved diverse mechanisms to modulate these parameters, enabling spatiotemporal regulation of condensate assembly, dissolution, composition and biophysical properties, such as viscosity and surface tension¹⁴. These mechanisms of dynamic regulation include control of protein expression, post-transcriptional and post-translational modifications (PTMs), membrane surfaces, active regulation and RNA species^{13,14}. For the purposes of this thesis, only regulation by PTMs and membranes will be discussed here.

1.2.1 Post translational modifications

PTMs can directly tune the valency and strength of interactions of both scaffold and client proteins (**Fig. 1.3a**). Common PTMs involved in condensate regulation include

phosphorylation, acetylation, methylation, SUMOylation, ubiquitination, poly-ADP-ribosylation (PARylation) and glycosylation^{14–16}.

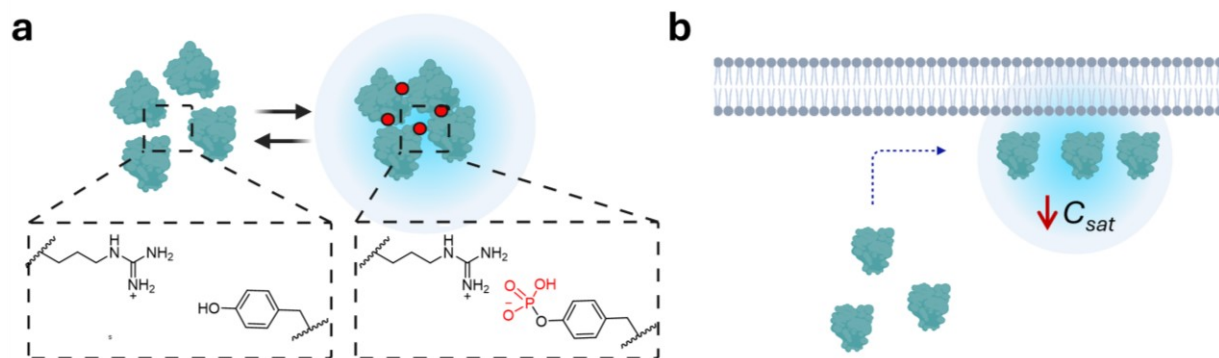


Figure 1.3. PTMs and membranes as dynamic regulators of biomolecular condensates. a) PTMs, such as phosphorylation (in red) can both promote and disrupt condensate formation by modulating multivalent interactions. **b)** membrane surface may regulate condensate formation by decreasing the C_{sat} . Figure created with Biorender.com.

Phosphorylation enables context-dependent condensate regulation, due to its dynamic and reversible kinetics. Phosphorylation adds a negative charge onto Serine/Threonine or Tyrosine residues (Ser/Thr/Tyr), to either promote or disrupt electrostatic interactions with other molecules e.g. negatively charged RNA (**Fig. 1.3a**). For instance, phosphorylation inhibits LLPS of fused in sarcoma (FUS), cytoplasmic polyadenylation element-binding protein (CPEB4), TAR DNA-binding protein 43 (TDP-43) and heterogeneous nuclear ribonucleoprotein A2 (hnRNPA2), but promotes LLPS of TIA1 and Tau^{17,18}. Moreover, phosphorylation can create specific binding sites, as exemplified by Grb2 proteins interacting only with the phosphorylated T cells receptor (linker for activation of T cells (LAT) via SH2 domains¹⁹.

In contrast, Arginine (Arg) methylation is less dynamic and primarily neutralises the positive charge, promoting condensate dissolution²⁰. Arg methylation has been shown to disrupt cation- π interactions in FUS and hnRNPA2²¹. Similarly, Lysine (Lys) acetylation neutralises positively charged lysine residues, thereby also disrupting cation- π interactions²². For example, Lys acetylation on N-terminal tails of histones reverses heterochromatin condensates to promote active transcription^{4,23}.

Condensates can be simultaneously regulated by multiple PTMs. SGs, for example, are known

to be regulated by eight different PTMs to date²⁴. The main SG scaffold, G3BP1, is dynamically acetylated and deacetylated by HDAC6 and CPB/p300 at Lys376, methylated at Arg435/447/460 by PRMT1 and 5 and demethylated by JMJD6, and phosphorylated at Ser149/232²⁵. This complex and concurrent regulation of a single scaffold protein within a single condensate, highlights the complexity of condensate regulation.

1.2.2 Membrane surfaces

Even though biomolecular condensates are primarily defined as membraneless compartments, interactions with phospholipid bilayer membrane surfaces have emerged as important regulators of condensate assembly^{14,26,27}. Membrane-associated condensates can form at the plasma membrane (PM) and other phospholipid bilayer membranes within the cell (endomembranes).

Membrane-associated condensates follow the same thermodynamic principles, but their restriction of molecular diffusion to a two-dimensional plane, lowers the C_{sat} (**Fig. 1.3b**). Biochemical reconstitution experiments that this reduction in dimensionality can reduce the C_{sat} by up to 30 fold²⁸. Increasing number of signalling pathways are reported to involve membrane-associated condensates. The Nephhrin–Nck–N-WASP signalling system was the first identified example, which forms phase separated clusters at the PM, required for regulating actin polymerisation in kidney podocyte cells²⁹. Another well-known example involves LLPS of pLAT, Grb2 and Sos1, which increases Sos1 dwell time, crucial for downstream RAS signalling³⁰. Additionally, the Wnt/ β -catenin signalling pathway is mediated through membrane-associated Dishevelled condensates^{26,30,31}.

Apart from the PM, endomembrane-associated condensates serve significant regulatory function. For instance, RNA-binding protein TIS11B undergoes LLPS to form TIS Granules that are intertwined with the ER, forming a TIGER domain which facilitates specific protein-protein interactions³². Furthermore, condensates may “hitchhike” on moving MBOs for long-distance transport. For example, annexin A11 (ANXA11) tethers SGs to the lysosomes, that themselves rely on microtubules for movement^{27,33}. Increased dwelling time may also act as a negative regulator, as exemplified with stimulator of interferon genes (STING)³⁴. Membrane-associated condensates can also act as regulators by modulating membrane properties such as curvature, and function in membrane repair. For example, SG form at membrane damage sites, acting as “plugs” to stabilise and enable membrane repair³⁵.

Despite significant advancements in understanding membrane-mediated regulation of condensates, many questions remain unanswered, particularly regarding the mechanisms governing protein recruitment to specific endomembranes.

1.3 Condensate dysregulation as a therapeutic target

Given the vast functional scope of condensates, it is unsurprising that their dysregulation and scaffold mutations are linked to neurodegeneration, cancer, cardiomyopathy, and viral infections⁵.

In cancer, aberrant formation, mislocalisation and composition of condensates have been observed, most notably for nuclear condensates³⁶. Particularly, cancer cells acquire transcriptional super-enhancer condensates, such as EWS-FLI in Ewing's sarcoma, Yes-associated protein (YAP)/PDZ-binding motif (TAZ) (YAP/TAZ), and MYC, to drive oncogene expression^{37–39}. For example, disruption of YAP condensates increases sensitivity to anti-PD1 immunotherapy in mouse lung adenocarcinoma model⁴⁰. Mutations within the SPOP protein, a tumour suppressor and a key scaffold for the SPOP/DAXX nuclear condensates, results in the accumulation of the DAXX pro-survival protein implicated in breast, endometrial, gastric and prostate cancers, whilst mutations within the promyelocytic leukaemia protein (PML) bodies nuclear are associated with tumorigenesis⁴¹. Beyond nuclear condensates, mutations within SG proteins and aberrant SG kinetics are associated with multiple cancers and are hypothesised to confer stress resistance to promote survival^{42,43}.

Protein aggregation is a hallmark of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), Parkinson's disease (PD), and Alzheimer's disease (AD). Increasing evidence suggests that biomolecular condensates act as crucibles for pathological aggregates^{43,44}. Notably, mutations or dysregulated PTMs of SG-associated RNA-binding proteins (RBPs) such as α -synuclein, FUS, Tau, TDP-43, and hnRNPA1, may drive the transitions from dynamic liquid-like states to irreversible aggregates^{45–48}. For example, loss of FUS Arg methylation promotes FUS aggregation and is often observed in FTD, whilst aggregation of TDP-43 is induced by acetylation, with acetylated TDP-43 lesions reported in ALS patients' spinal cords^{21,49}. While the precise cascade linking aggregation to neurodegeneration remains unclear, it is hypothesised that these aggregates are either directly toxic or contribute to disease through the loss of protein function^{45,47}.

Given this growing evidence, condensate-modifying therapeutics (c-mods) represent a new frontier in drug discovery⁵. C-mods have the potential to prevent condensate formation, reverse aberrant condensates or modify their material properties^{5,50}. Moreover, existing therapeutic small molecules have been shown to partition into super-enhancer condensates. For instance, cisplatin, a commonly used antineoplastic intercalating agent, partitions into transcriptional condensates, resulting in higher selectivity for super-enhancer DNA, potentially explaining its efficacy in cancer over normal cells⁵¹. Understanding the “chemical grammar” that dictates small molecule condensate partitioning would enable guided drug design with higher efficacy⁵⁰.

1.4 Protein lipidation

Protein lipidation is a diverse class of PTMs involving the covalent attachment of lipid or lipid-like groups, and can be broadly classified into two classes based on substrate localisation⁵². The first category encompasses lipid modifications involved in outer membrane or secretory targeting, such as glycosylphosphatidylinositol (GPI) anchor and *O*-cholesterylation. The second, and a more extensively studied category, refers to intracellular lipidation, particularly the addition of fatty acyl or polyisoprenyl moieties. Intracellular lipidation can be further categorised into constitutive, glycine *N*-myristoylation and Cysteine (Cys) *S*-prenylation (farnesylation and geranylgeranylation), and dynamic, Cys *S*-acylation (**Fig. 1.4, Fig 1.5**). These PTMs occur on thousands of proteins, and primarily modulate their affinity for cellular membranes, but also function, stability and interactome.

N-myristoylation is the irreversible transfer of a saturated C14:0 fatty acid (myristate) from myristoyl-coenzyme A (CoA) onto the free amine of an N-terminal Glycine (**Fig. 1.4a**)⁵³. In humans, *N*-myristoylation occurs on over 200 substrates, and is catalysed by *N*-myristoyl transferase 1 and 2 (NMT1 and NMT2) via a stable amide bond, with NMT1 accounting for the majority of cellular *N*-myristoylation⁵⁴. This PTM occurs primarily co-translationally, following removal of initiator methionine by methionine aminopeptidases (MetAPs) early in translation, but can also occur post-translationally. Dysregulated *N*-myristoylation is implicated in several cancers, infectious diseases, and metabolic disorders, with some NMT inhibitors (NMTi) currently undergoing clinical trials as anticancer agents⁵⁴.

S-prenylation involves the attachment of farnesyl (C15) or geranylgeranyl (C20) isoprenoids

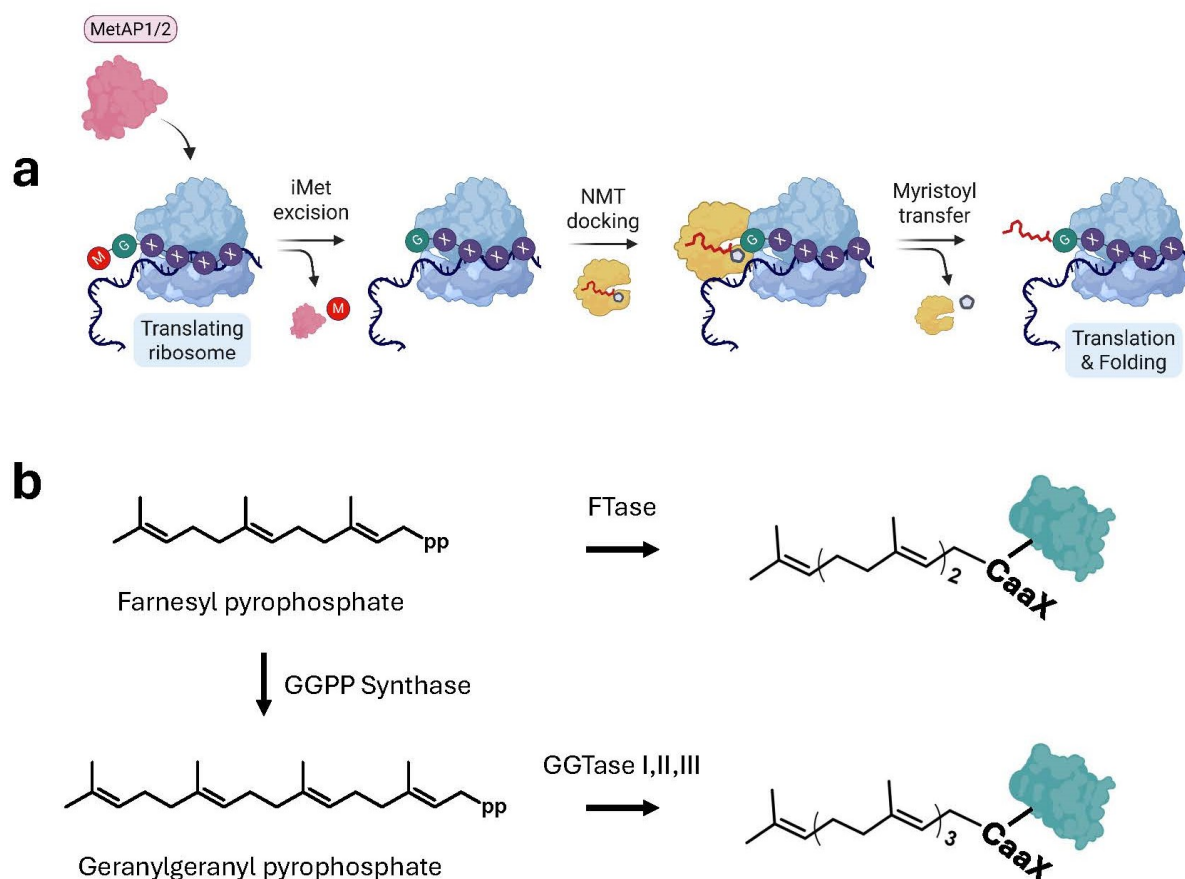


Figure 1.4. Overview of constitutive lipid post-translational modifications. **a)** Mechanism of glycine *N*-myristoylation. *N*-myristoylation occurs at the ribosome, with the initiator methionine (iMet) first removed by MetAP-mediated cleavage. NMT1 or 2 then transfers a myristoyl group from the myristoyl-CoA onto the newly revealed N-terminal glycine. **b)** Mechanism of cysteine *S*-prenylation. *S*-prenylation occurs at the C-terminal “CAAX” motifs, catalysed by FTase or GGTase I–III. Figure created with Biorender.com.

onto Cys via a thioether bond, mediated by the respective farnesyltransferase (FTase) and geranylgeranyl transferases I, II and III (GGTases I–III) (**Fig. 1.4b**)⁵². *S*-prenylation typically occurs at a C-terminal motif known as the CaaX box (a = aliphatic amino acid, X = any amino acid), whereby the X typically determines the type of prenylation. Most prenylated proteins belong to the small GTPase family, including members of the RAS, Rab, and Rho subfamilies. Prevalence of RAS activating mutations in cancer and their reliance on prenylation for PM localisation and oncogenic signalling, poses prenyltransferases and prenyl biosynthesis as attractive therapeutic targets⁵⁵. Moreover, dysregulated *S*-prenylation is linked to several neurological and infectious diseases, such as PD, AD and hepatitis C, among others^{52,56,57}. However, selective FT inhibitor (FTI) - tipifarnib, failed in clinical trials for head and neck

squamous cell carcinoma due to adverse side effects, but is currently under investigation as anticancer combination therapy^{54,58,59}. Notably, FTI Lonafarnib is the only FDA-approved drug for inhibiting a lipid transferase to date, used for the treatment of Hutchinson-Gilford progeria syndrome (HGPS) and processing-deficient progeroid laminopathies, demonstrating the potential for safe and efficacious targeting of protein lipidation^{54,60}.

1.5 S-acylation

The most unique of the lipid PTMs, is the dynamic S-acylation, which involves the attachment of long chain fatty acids onto Cys via a labile thioester bond (**Fig. 1.5**)⁶¹. The S-acyl chains may be both saturated and unsaturated and range from 14 to 22 carbons in length, though the most prevalent lipid is the C16:0 palmitate⁶². Due to this fatty acid selectivity bias, S-acylation is often referred to as “palmitoylation”, which does not reflect its heterogeneity. S-acylation is therefore the preferred term for this form of lipidation and will be used in this thesis.

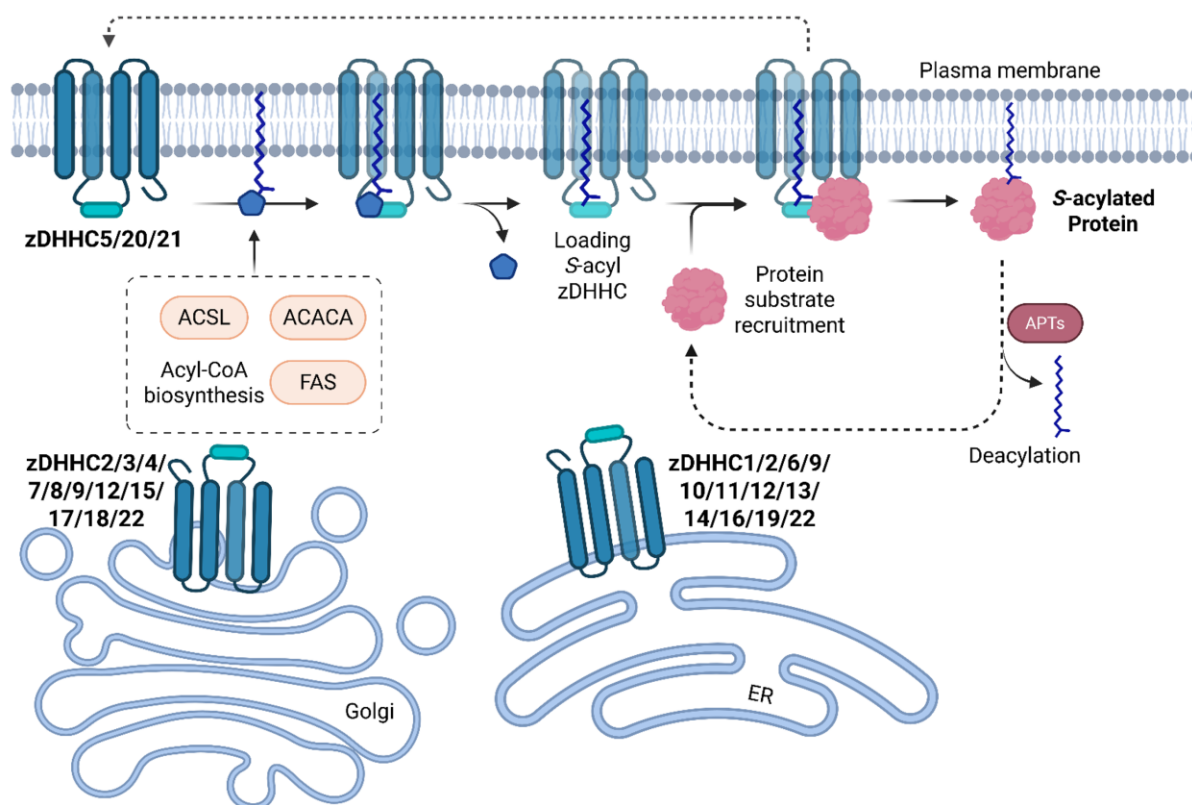


Figure 1.5. Overview of protein S-acylation. Fatty acyl-CoAs (derived from lipid biosynthesis) are first loaded onto the zDHHC enzyme, through auto-acylation of their catalytic cysteine. The acyl chain is then transferred onto the cysteine of the protein substrate, followed by regeneration of apo-zDHHC. Acylprotein thioesterases (APTs) hydrolyse the thioester modification, enabling reversible and dynamic cycle. Most zDHHCs reside on the Golgi apparatus, Endoplasmic Reticulum (ER), and plasma membranes. Figure created with BioRender.com

S-acylation stands out for three main reasons:

1. Reversibility – it is the only reversible lipid PTM
2. Lack of consensus sequence - it is currently believed that S-acylation can occur on any accessible Cys, independent of any proximal residues
3. Diverse substrate scope - S-acylation occurs on thousands of proteins (estimated 20% of the human proteome) with diverse functional and biochemical properties⁶³

Like other lipid PTMs, S-acylation confers hydrophobicity to proteins thereby facilitating their trafficking and interaction with the PM and endomembranes but can also affect the stability, protein-protein interactions, conformational changes and lysosomal degradation of its substrates⁶¹.

1.5.1 zDHHC and APT enzymes as regulators of dynamic S-acylation

S-acylation is catalysed by a family of zinc-finger and Asparagine-Histidine-Histidine-Cysteine (DHHC) motif containing protein acyltransferases (zDHHCs). The human genome encodes for 23 zDHHC isoforms, each possessing at least four transmembrane domains and a cytoplasmic DHHC motif. Most ZDHHCs reside at the Golgi, ER, and to a lesser extent at the PM (**Fig. 1.5**)⁶⁴. zDHHCs have been demonstrated to function via a two-step ping-pong mechanism. First, the active site Cys of the DHHC motif reacts with acyl-coenzyme A (acyl-CoA) forming an enzyme-acyl intermediate (auto-acylation) (**Fig. 1.5**)⁶⁵. Crystal structure of human zDHHC20 revealed that the lipid resides within the hydrophobic cavity formed by the transmembrane domains, which is proposed to govern the acyl chain specificity for each enzyme (**Fig. 1.6**)⁶⁶. The acyl chain is then transferred from the DHHC Cys to a Cys of the target protein, regenerating the enzyme. However, it has been previously demonstrated that some zDHHCs (e.g. zDHHC17) do not get auto-acylated, but instead may rely on direct substrate binding for acyl chain transfer, although the precise alternative mechanisms remain unknown⁶⁷. Interestingly, zDHHCs themselves can be regulated by other zDHHCs alongside other PTMs. Furthermore, specific isoforms such as ZDHHC5, 7, 14, and 18, require interaction with accessory proteins like GOLGA7 (also known as GCP16) for their stability^{68–70}.

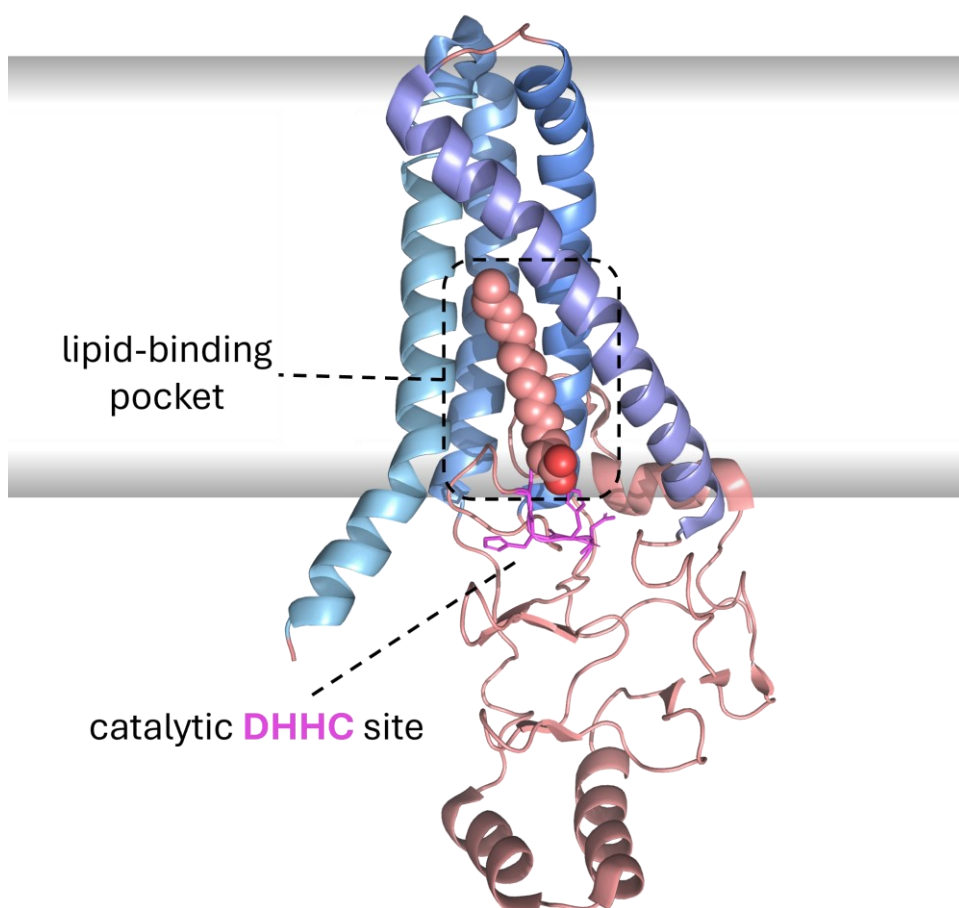


Figure 1.6. Structural features of zDHHC20. Crystal structure of the zDHHC20 enzyme irreversibly inhibited by the lipid-mimic 2-bromopalmitate (in light red). The four transmembrane helices form a lipid binding cavity. The catalytic Asp-His-His-Cys motif (magenta) is located at the interface of the cytosol and the lipid bilayer. The figure was generated in PyMOL using PDB ID: 6BML.

Deacylation is mediated by acyl protein thioesterases (APTs) of the serine hydrolase superfamily including APT1 and 2, as well as α/β -hydrolase domain-containing protein 10 (ABHD10), ABHD17A, ABHD17B and ABHD17C. Similarly to zDHHCs, APTs have distinct substrate and chain length specificities. Although primarily cytosolic, APTs have been shown to undergo cycles of S-acylation/deacylation, often required for its deacylation catalytic function⁷¹.

1.5.2 S-acylation in health and disease

Given its vast substrate scope, it is unsurprising that S-acylation is crucial in myriad cellular

processes, and its dysregulation is implicated in various cancers⁵⁴. The dynamic acylation/deacylation cycle regulates key signalling pathways involved in tumorigenesis, such as the RAS superfamily (HRAS, NRAS and KRAS4A), receptor tyrosine kinases (RTK), and EGFR, some of which require cooperation with other lipid modifications⁷²⁻⁷⁵. For instance, PM translocation of HRAS relies on *S*-acylation of two Cys (181 and 184) and *S*-prenylation of one Cys (186), whilst non-receptor tyrosine kinase YES requires prior *N*-myristoylation^{73,76}. Beyond signalling, *S*-acylation functions in protein homeostasis. For example, programmed death ligand 1 (PD-L1) *S*-acylation by zDHHC3 and zDHHC9 confers protection from ubiquitination and lysosomal degradation, and disruption of this event (through Cys mutation or zDHHC knockdown) led to increased tumorigenesis^{77,78}. Conversely, NLRP3 *S*-acylation by zDHHC12 induces its degradation through chaperone-mediated autophagy⁷⁹. Furthermore, *S*-acylation plays important roles in programmed cell death, with *S*-acylation of Gasdermin D (GSDMD), BAX and GPX4, implicated in pyroptosis, apoptosis and ferroptosis respectively. *S*-acylation is also involved in the regulation of nuclear proteins, such as signal transducer and activator of transcription 3 (STAT3), which requires *S*-acylation by zDHHC3 and 7 and deacylation by APT2 for nuclear localisation and induction of pro-inflammatory and pro-tumorigenic gene expression⁸⁰. Conversely, p53 is *S*-acylated by zDHHC1 to elicit its tumour suppressor function⁸¹. Interestingly, non-enzymatic auto *S*-acylation of the transcriptional enhanced associate domain (TEAD) family has also been reported and presents a promising therapeutic avenue⁸¹.

In addition to the dysregulation of specific pathways, altered expression of the *S*-acylation machinery is frequently observed in tumours, with examples including downregulation of zDHHC22 in breast cancer and zDHHC2 overexpression in tyrosine kinase inhibitor resistant clear cell renal cell carcinoma^{82,83}. Moreover, aberrant *S*-acylation has been demonstrated to confer resistance to cancer therapy, with zDHHC2-mediated *S*-acylation of mitochondrial acylglycerol kinase (AGK) resulting sunitinib resistance in renal cell carcinoma, whereas zDHHC16-mediated *S*-acylation of proprotein convertase subtilisin/kexin type 9 (PCSK9) linked to sorafenib resistance in liver cancer⁵⁴. Beyond cancer, *S*-acylation is linked with inflammatory disorders, as well as diseases of the central nervous system, particularly with Huntington's disease, PD, and AD^{84,85}. Overall, this accumulating evidence underscores the therapeutic potential of targeting the *S*-acylation machinery for numerous diseases.

1.6 Tools to study S-acylation

1.6.1 Mass spectrometry-based tools

Historically, S-acylation was studied using radiolabelling with [^3H]-palmitate, a time-consuming and potentially hazardous method^{86,87}. As such, emergence of MS-based proteomics has been pivotal in driving S-acylation (and other lipid PTMs) research. Proteome-scale analysis of S-acylation, or any PTM in general, requires enrichment from complex biological samples, such as cell or tissue lysates. However, S-acylation is non-antigenic, precluding its enrichment with immunoprecipitation (IP) approaches⁶¹. Moreover, direct analysis of intact S-acylated peptides is precluded by the inherent hydrophobicity and lability of this PTM^{87,88}. These challenges prompted the development of elaborate enrichment strategies to enable S-acylproteomic profiling.

Currently, two main approaches exist to profile S-acylation on a proteome-wide scale (S-acylproteomics) – lipid-centric and cysteine-centric (**Fig. 1.7a,b**). Cysteine-centric approaches exploit the selectivity of NH_2OH (hydroxylamine/HA) for thioester bond cleavage and are applicable to any tissue or cell lysates. These lysates are first treated with an alkylating agent, typically N-ethylmaleimide (NEM) or chloroacetamide (CAA), to block all free Cys (**Fig. 1.7a**). Subsequently, the S-acyl chains are cleaved with neutral HA, and the newly revealed Cys are then conjugated to a Cys-reactive affinity tag or resin, followed by the relevant enrichment, digestion and LC-MS/MS analysis^{87,89}. The most common cysteine-centric methods include acyl-biotin exchange (ABE) (e.g., biotin-HPDP) and acyl-resin assisted capture (acyl-RAC) (thiopropyl Sepharose resin), however the cysteine-centric toolbox has grown in recent years. Cysteine-centric methods are going to be discussed in more detail in **Chapter 3.1**.

Lipid-centric approaches on the other hand, are based on the metabolic labelling with a fatty acid analogue followed by click chemistry (MLCC) strategy (**Fig. 1.7b**). Typically, an azide- or alkyne-functionalised lipid analogue (usually palmitate) is added to cell culture medium and metabolically incorporated into native S-acylation sites by endogenous S-acylation machinery⁵³. After cell lysis, labelled proteins are conjugated to a biotin analogue by copper-catalysed alkyne-azide cycloaddition (CuAAC) click chemistry. The *in vitro* biotinylated proteins are then enriched with streptavidin beads, digested into peptides by trypsin and

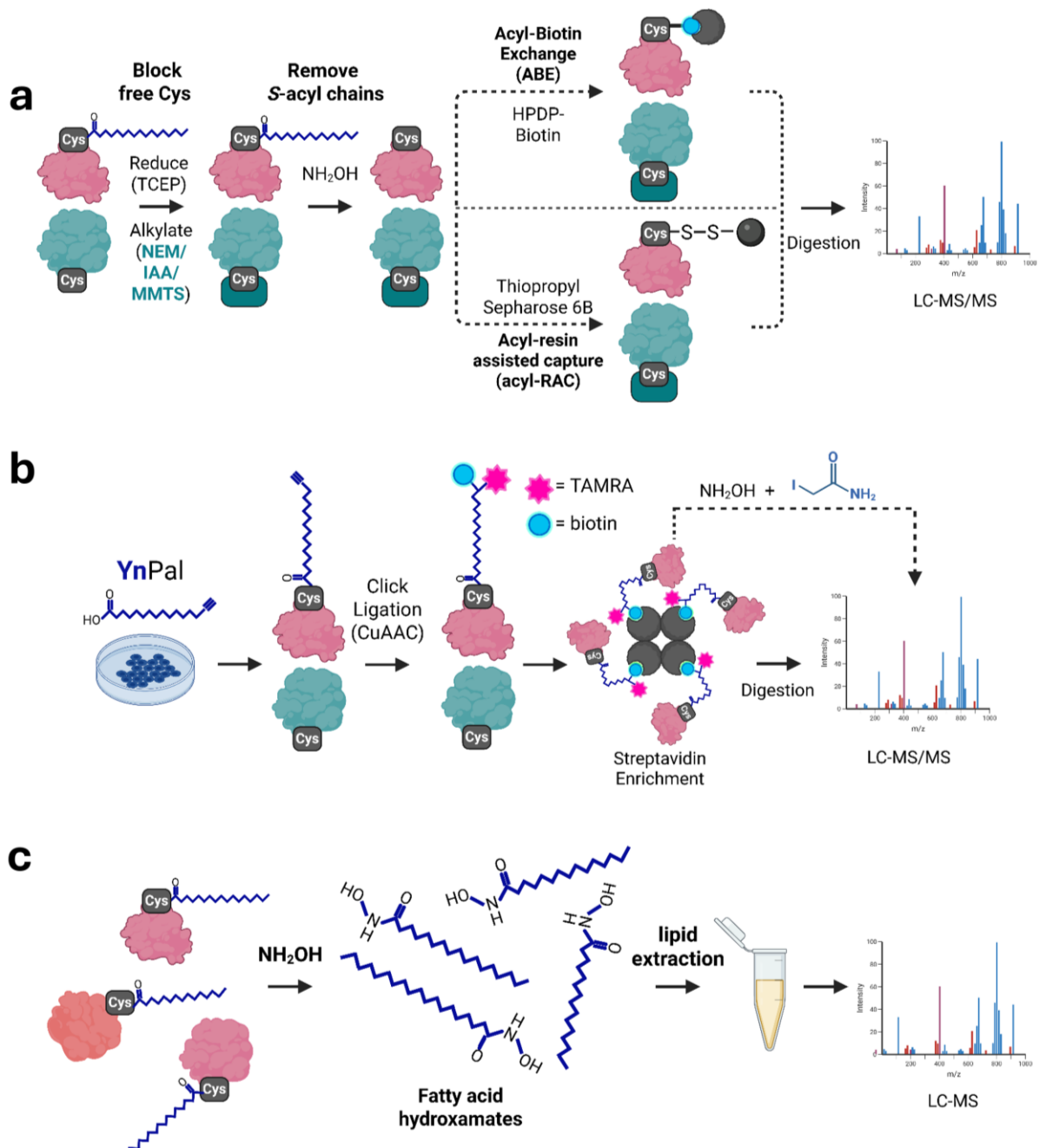


Figure 1.7. Mass spectrometry-based tools to probe S-acylation. **a)** Cysteine-centric methods: Free thiols on cysteines are initially blocked using alkylating agents such as iodoacetamide (IAA), N-ethylmaleimide (NEM), or methyl methanethiosulfonate (MMTS). The acyl chains are then selectively cleaved using neutral hydroxylamine (NH_2OH), exposing previously modified cysteines, which are subsequently labelled with a cysteine-reactive biotin tag (as in acyl-biotin exchange, ABE) or captured using a thiol-reactive Sepharose resin (as in acyl-RAC). Labelled proteins are enriched, digested into peptides, and analysed by LC-MS/MS. **b)** Lipid-centric approaches: These methods involve metabolic incorporation of alkynyl-fatty acid analogues (YnPal). Following cell lysis, click chemistry is used to attach biotin to S-acylated proteins, allowing enrichment with streptavidin beads. Site specificity can be achieved through thioester hydrolysis to release proteins from the beads, followed by mass tag labelling of the liberated thiols. **c)** Profiling fatty acyl chain heterogeneity: Treatment of S-acylated proteins with NH_2OH releases the lipid moieties as fatty acid hydroxamates (FAHs), which can be detected and quantified by LC-MS. Figure created with BioRender.com.

analysed on the LC-MS/MS. Although MLCC has been instrumental in driving S-acylation research, it is limited to cell culture models and favours substrates high turnover rates^{87,90}. Moreover, these functionalised probes may also incorporate into *N/O*-fatty acylated proteins and GPI-anchors, resulting in false positive identifications⁹⁰. To circumvent this, some MLCC approaches introduce an additional HA cleavage off the bead for specificity^{91–93}. Although most optimised methods lack site-specific resolution, a site-specific MLCC method was introduced, whereby S-acylated proteins are released from the streptavidin beads with HA, and subsequently labelled with iodoacetamide (IAA)⁹⁰. A recent addition from our group and others, is the chemical genetic “bump-hole” approach, which is the only system enabling enzyme-specific assignment by utilising a fatty acid analogue with a chemical “bump” with specificity for a specific mutant zDHHC^{94,95}. The optimised “bump-hole” strategy is available for zDHHC20, with initial proof of concept for zDHHC3/7/11/15. Both cysteine- and lipid-centric approaches require extensive workflows, especially for site identification, and can be prone to false positive identifications (see **Chapter 3.1**). Notably, both acyl-exchange and MLCC approaches are also widely utilised for gel-based validation experiments.

In addition to these lipid- and cysteine-centric approaches, a recent method using nanographite fluoride-based solid-phase extraction (nGF-SPE), enabled direct enrichment and analysis of intact S-acylation, identifying over 700 sites in HeLa cells (discussed in **Chapter 3.6.2**)⁹⁶. However, despite the seminal implications of this approach, it has not been replicated by any other groups.

Beyond identification of the substrate profiles, separate MS-based methods are utilised for profiling the diversity of acyl chains in a given context, as this information is lost in S-acylproteomics approaches. These methods similarly rely on HA treatment but subsequently focus on the LCMS-based identification of the resulting fatty acid hydroxamates (**Fig. 1.7c**)⁶². These methods have since been applied to human cells, tissues and *C. elegans*, revealing enzyme-specific fatty acid preferences (e.g. C18:0 zDHHC17 and shorter chains for zDHHC3)^{62,97}.

1.6.2 Small molecules

One of the biggest limiting factors in S-acylation research, is the absence of selective inhibitors or probes for the zDHHC enzymes. One inhibitor that continues to be used to probe S-acylation is 2-bromopalmitate (2-BP), which although engages with some zDHHCs, suffers

from broad reactivity with membrane-bound and lipid metabolism proteins, resulting in hundreds of off-targets targets including APTs (**Fig. 1.8**)^{98–100}. Its use arises from the lack of better alternatives but is not recommended due to its off-target profile¹⁰¹. Another inhibitor, Cyano-myrcylamide (CMA), demonstrated slightly improved selectivity over 2-BP and lack of APT inhibition, but still reported more than 200 off-targets¹⁰⁰. It is now generally recognised that fatty acid mimetics and their derivatives lead to numerous off-targets¹⁰⁰. More recently, the discovery of a zDHHC8-specific small molecule inhibitor PF-670462, zDHHC3-specific inhibition by antibiotic avibactam and zDHHC20-specific inhibition by small molecule SD-066-4 have been claimed in three independent reports^{102–104}. However, all of these studies lack the necessary structure-activity-relationship and selectivity data to promote their use as selective inhibitors for any of the aforementioned zDHHCs.

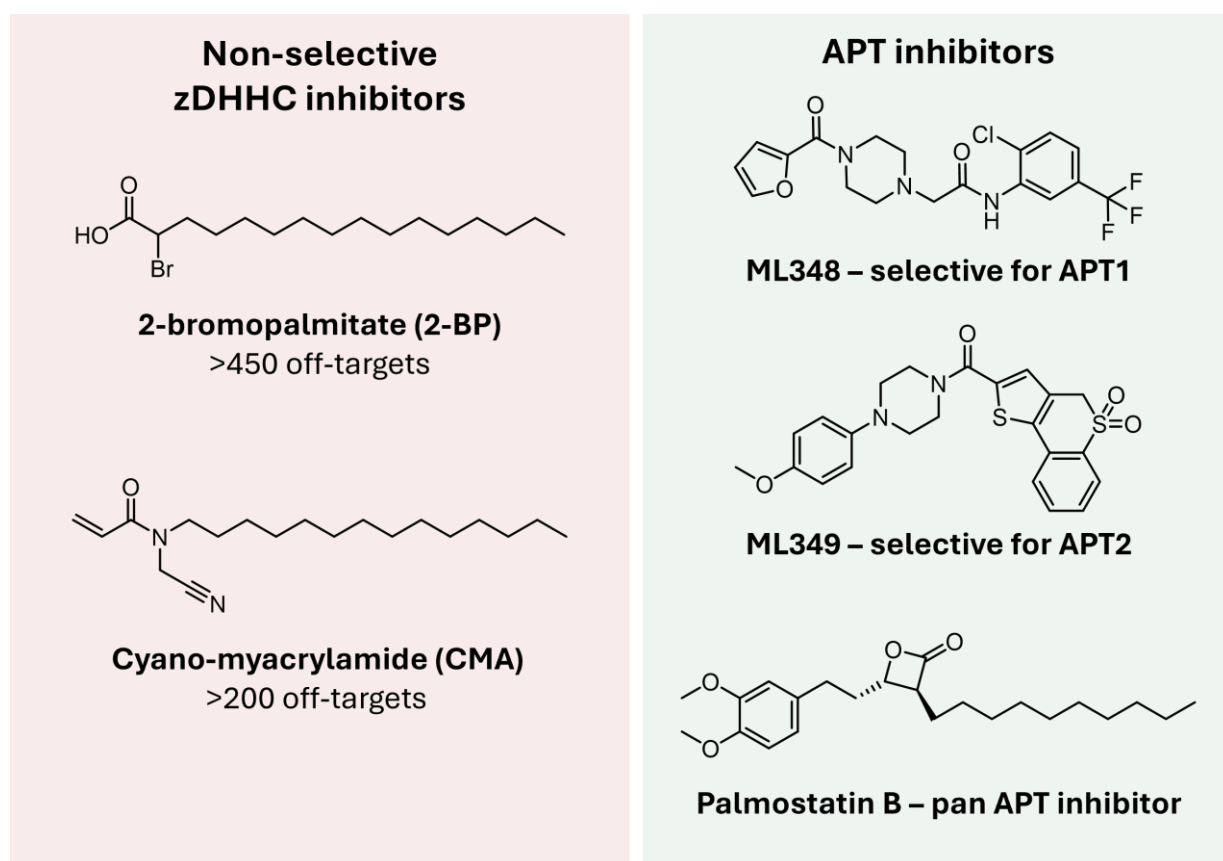


Figure 1.8. Small molecule tools to study S-acylation.

In contrast, several inhibitors are available for APTs, including a broad APT inhibitor Palmostatin B, and selective inhibitors ML348 and ML349 targeting APT1 and 2, respectively (**Fig. 1.8**)^{105,106}.

1.6.3 Challenges in studying S-acylation

S-acylation was first reported in 1979 using radiolabelling and autoradiography⁸⁶. Although much progress has been made into our understanding of this PTM, the study of S-acylation continues to face three main challenges – paucity of analytical and chemical tools, enzyme functional redundancy, and diversity of substrates.

As described above, identification of S-acylated proteins on the global scale relies on indirect proteomics methods, which typically lack site specificity and are prone to false positives. Subsequently, in the absence of selective zDHHC inhibitors, validation of putative S-acylated proteins and their biological significance, requires knowing either the modification Cys site or the cognate zDHHC to enable either the respective Cys mutagenesis or KD of the event-specific zDHHC. Both strategies are labour and time intensive, with the precise assignment of the cognate zDHHC alone requiring individual KD or overexpression of each zDHHC. These challenges are further amplified by a degree of functional redundancy exhibited by the 23-member zDHHC family. As such, compensatory mechanisms in response to KD, may lead to incorrect or variable assignments, as is the case with GSDMD, NLRP3, and GPX4 (further discussed in **Chapter 4.6.1**).

Lastly, the diverse protein substrate scope and lack of a clear consensus sequence, further limit curation of a high-confidence S-acylated proteome database. The only currently available S-acylation database, SwissPalm, lacks stringent curation criteria and likely contains many false positives⁶³.

1.7 Hypothesis

The underlying principle of condensate formation relies on weak multivalent interactions, including hydrophobic, electrostatic, π - π , and cation- π interactions (**Fig. 1.2a**). The importance of hydrophobic interactions particularly, is widely recognised and demonstrated through simulations and experiments with 1,6-hexanediol, a widely used aliphatic alcohol known to disrupt hydrophobic interactions¹¹. As discussed in this Chapter, the roles of many PTMs and membrane interactions have been extensively investigated, although the precise mechanisms for protein recruitment to membranes remain unclear. However, the involvement of protein lipidation remained largely unexplored.

Lipid PTMs significantly enhance protein hydrophobicity and affinity for membranes, two properties potentially critical for condensate formation. Given that dynamic and transient nature of condensates, and a similarly dynamic and reversible nature of S-acylation, it was hypothesised to play key roles in condensate formation (**Fig. 1.9**):

Hypothesis 1 posits that S-acylation enhances protein hydrophobicity, thereby increasing interaction valency and promoting condensate formation via hydrophobic interactions.

Hypothesis 2 suggests S-acylation may nucleate condensate formation through membrane recruitment

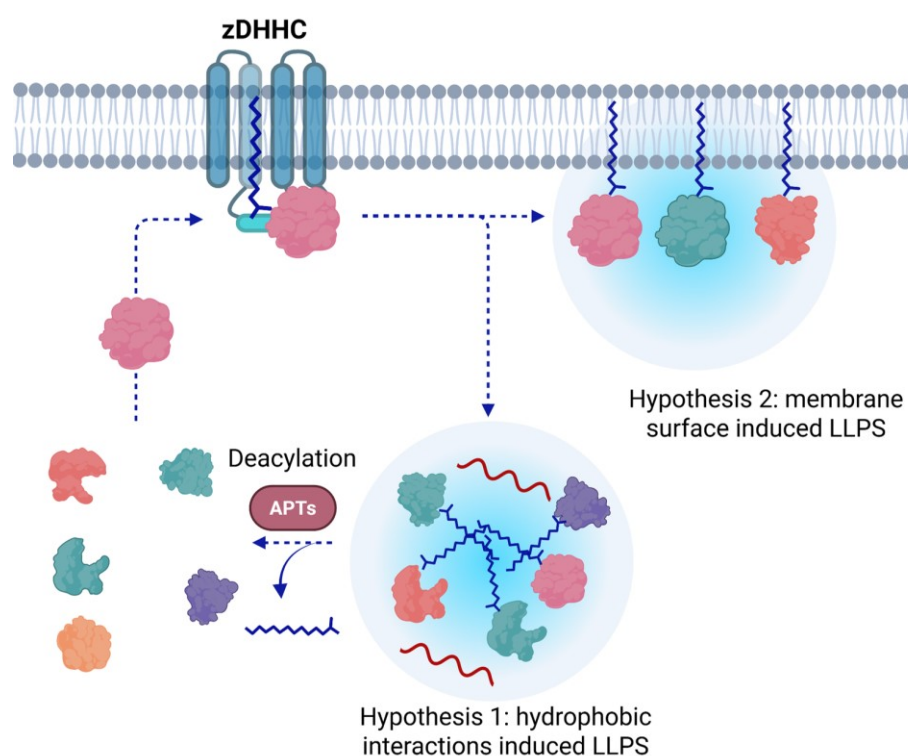


Figure 1.9. S-acylation as a putative regulator of biomolecular condensates. S-acylation is proposed to promote condensate assembly through two primary mechanisms. The first hypothesis suggests that the addition of a hydrophobic fatty acid onto a protein increases its multivalency for hydrophobic interactions, thereby enhancing its association with condensates. The second hypothesis builds on S-acylation's primary role in membrane localisation of its substrates, proposing that S-acylation facilitates the formation of membrane-associated condensates by membrane targeting. Figure created with BioRender.com.

At the conception of this PhD project and hypothesis generation, no prior studies have been published regarding the role of S-acylation in biomolecular condensates. Since then, zDHHC19-mediated S-acylation of p62 condensates (Cys289 and Cys290) has been

demonstrated to function in their autophagic membrane localisation, whereas deacylation by APT2 suppressed p62-mediated autophagy¹⁰⁷. Although S-acylation did not directly contribute to the LLPS behaviour of p62, it was critical for their localisation and function. In addition to this, during the preparation of this thesis, another study was published which directly corroborates **hypothesis 1**, demonstrating that S-acylation of NLRP3 (Cys130 and 261) by zDHHC7, directly promoted its LLPS in the cells by lowering the C_{sat} , independent of membrane recruitment, and was necessary for NLRP3 activation¹⁰⁸.

1.8 Project aims

Given the general lack of research into the role of S-acylation in biomolecular condensates, this project aimed to investigate **hypothesis 1** in a model condensate system. SGs were selected as model condensates due to their wide characterisation in the field, ease of chemical inducibility, as well as association with multiple disease states (discussed in **Chapter 2.1**). Leveraging the expertise of our research group in MS-based proteomics, this project initially aimed to develop a novel double-enrichment proteomics platform to identify stress-dependent changes in the S-acylation profiles of SG proteins. Upon successful identification and validation of the stress-dependent sites, microscopy-based workflows, such as assessment of SG dynamics by fluorescence recovery after photobleaching (FRAP) and temporal profiling with live-cell imaging, would be performed to assess their functional contribution for SG assembly and dynamics.

However, initial efforts at investigating this hypothesis were significantly hindered by limitations in existing methods for S-acyl site identification on a proteome-wide scale. To address this gap, the latter half of this PhD solely focused on developing a novel S-acylproteomic approach utilising phosphonate handles for site-specific profiling of S-acylation. The main themes of this thesis are therefore structured as follows:

- **Chapter 2:** I investigate the role of S-acylation in SGs, by conceptualising and attempting a double-enrichment proteomics workflow, integrating APEX2 proximity labelling and S-acylproteomics approaches for the *in situ* identification of S-acylated sites within assembled SGs.

- **Chapter 3:** I conceptualise, optimise and apply a novel proteomics platform - **S-A**cylation **P**rofilng by **P**Hosphonate taggIng p**R**ot**E**omics (**SAPPHIRE**) – to enable streamlined identification of global S-acylation with site-specific resolution.
- **Chapter 4:** I apply SAPPHIRE for comprehensive profiling of the S-acylated proteome profiling in diverse tissues and cells, associated with disease-relevant contexts, including investigation of global changes in S-acylation in response to stress.

Chapter 2.

Investigating the role of S-acylation in stress granules by chemical proteomics

2.1 Introduction

SGs are cytoplasmic biomolecular condensates that rapidly assemble in response to stress and dissolve upon return to normal growth conditions (**Fig. 2.1**). Their formation is typically initiated by global translational arrest, mediated by phosphorylation of eukaryotic translation initiation factor 2 alpha (eIF2 α), leading to widespread polysome disassembly¹⁰⁹. This event releases a large pool of unbound mRNA into the cytosol, enabling interactions with specific scaffold proteins to form initial SG structures known as “cores” or nanoscopic “seeds”. Ras GTPase-activating protein-binding protein 1 (G3BP1) serves as the central scaffold for SGs, functioning as a molecular switch to trigger LLPS in response to increased intracellular RNA concentrations, although in some cases SGs may occur independently of G3BP1¹¹⁰.

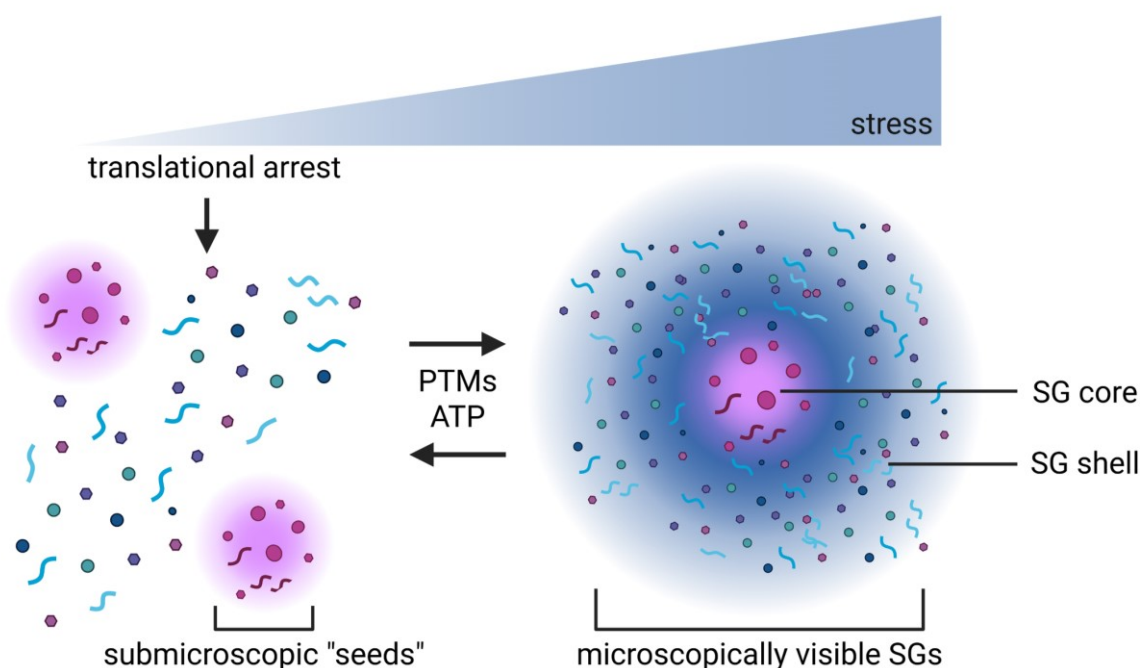


Figure 2.1. Stress granule formation and architecture. Upon cellular stress, translational arrest releases a pool of unbound mRNA into the cytoplasm, where it is incorporated into stress granules (SGs). A pre-existing protein interaction network serves as a seed for SG assembly. The formation and dynamics of SGs are modulated by post-translational modifications (PTMs) and active cellular processes. Figure made with Biorender.com.

To date, no universally accepted function for SGs exists, though many have been proposed, including sequestration and protection of RNAs and proteins from degradation, conservation of cellular energy, translational quality control, stabilisation and repair of damaged

endolysosomal membranes, and suppression of cell death pathways through sequestration of proapoptotic and pyroptotic factors^{35,111–118}. Nevertheless, the precise biological significance and function of SGs remain elusive.

The thermodynamics of SG assembly and disassembly, as well as their molecular composition, are tightly regulated and vary depending on the stress type^{119–121}. Their dysregulation results in aberrant SG kinetics, implicated in many disease states, highlighting their potential as a therapeutic strategy^{47,122,123}. Under chronic stress conditions, SGs often lose their liquid like-properties, acting as nucleation sites for insoluble aggregates, a hallmark of neurodegenerative disorders⁴⁴. Mutations of key SG proteins, including TDP-43, FUS and TIA1 are notably linked with ALS and FTD^{45,124–128}. Furthermore, overexpression of SG components, including G3BP1/2, has been reported in several tumours, including pancreatic cancer, sarcoma, hepatocellular carcinoma and malignant gliomas^{42,122,129–132}. It is hypothesised that SGs confer stress adaptation as well as chemo and radiotherapy resistance, promoting survival within the stressful tumour microenvironment (nutrient deprivation, hypoxia, increased reactive oxygen species)^{133–136}.

Historically, the study of SG composition, and subcellular compartments in general, has predominantly relied on targeted approaches such as fluorescent imaging and immunoprecipitation (IP), which require prior knowledge of candidate proteins¹³⁷. More recently, proximity labelling (PL) proteomics emerged as a powerful and unbiased tool to map endogenous interactions with spatiotemporal resolution. PL typically involves a genetic fusion of a protein “bait”, to a labelling enzyme that converts small molecule substrates into reactive species that covalently label proximal biomolecules¹³⁸. The most utilised enzymes include engineered peroxidases (APEX2, split-APEX) that create phenoxy radicals upon stimulation with H₂O₂, and biotin ligases (BioID, TurboID, miniTurbo, split-TurboID, LOV-Turbo) that generate biotin-AMP esters^{139–142}. More recently, photocatalytic PL technologies have emerged, most notably “microenvironment mapping” (μMap), which leverages iridium (Ir) photocatalysts to convert diazirines into highly reactive carbenes through Dexter energy transfer upon blue light irradiation^{143,144}. Each technology features distinct labelling times, radii and amino acid specificity, as summarised in **Table 2.1**. However, regardless of the enzyme/catalyst employed, these approaches typically rely on biotin-based substrates - biotin-phenol (BP) for APEX2, biotin for BioID and derivatives, and biotin diazirine for

photocatalytic approaches. Biotinylated proteins are subsequently enriched using ‘avidin-based beads, digested into peptides and analysed by LC-MS/MS.

Table 2.1. Summary of proximity labelling proteomics technologies

Name	Type	Size (kDa)	Labelling time	Labelling sites	Labelling radius	Advantages	Limitations
APEX2 ¹³⁹	Peroxidase	28	1 min	Tyr, Trp, Cys, His	~20 nm	High temporal resolution: Applicability to both protein and RNA	Toxicity of H ₂ O ₂ → limited application <i>in vivo</i>
BioID ¹⁴¹	Biotin Ligase	35	6-24 h	Lys	<10 nm	<i>In vivo</i> applicability	Low catalytic activity → poor temporal resolution
TurboID ¹⁴²	Biotin Ligase	35	10 min	Lys	<10 nm	Highest activity biotin ligase, <i>In vivo</i> applicability	High biotin affinity
miniTurbo ¹⁴²	Biotin Ligase	28	10 min	Lys	<10 nm	Smallest bioID	Lower catalytic activity and stability compared to TurboID
μMap ¹⁴⁴	Iridium catalyst	N/A	5-15 min	Broad spectrum, inserts into C-H bonds	<4 nm	No enzyme required	Known ligand required

As outlined in **chapter 1.7**, our hypothesis posits that S-acylation may contribute to the multivalent interactions driving LLPS and therefore condensate formation (**Fig. 1.9**). In this Chapter, SGs will serve as a model condensate system to probe the previously unexplored role of S-acylation in condensate dynamics. Specifically, this Chapter will outline efforts to develop a double enrichment proteomics platform, termed APEX-AX, combining APEX2 PL and cysteine-centric S-acyl exchange (AX) methods, to systematically map S-acylation states within the SG proteome.

2.2 Bioinformatic analysis of SG proteome and S-acylation

2.2.1 Curation of the SG protein composition list

Prior to undertaking any proteomics approaches, a bioinformatic analysis was first conducted to explore the link between S-acylation and SG proteins using existing datasets. This analysis was prioritised due to the lack of any previous studies linking S-acylation with condensate

regulation, including SGs. Although numerous studies have reported SG compositions, the only available database for mammalian SGs (msgp.pt) has not been updated since 2018¹⁴⁵. Thus, a comprehensive and updated SG proteome dataset was curated using data from six independent MS-based proteomic studies. Given that SG composition can vary depending on the stress, cell type, and identification technique, the selected studies feature diverse stressors, cell models and methodology, including four PL and two biochemical isolation approaches^{120,146–150}. These studies are summarised in **Table 2.2**, yielding a curated list of 1110 SG proteins.

Table 2.2. Summary of stress granule composition studies

Article	Method	Stress	Cell Model	Protein IDs
Jain <i>et al.</i> 2016 ¹⁴⁹	Biochemical isolation	500 μ M NaAsO ₂	U2OS	317
Youn <i>et al.</i> 2018 ¹⁴⁶	Bio ID proximity labelling	No Stress	Stable HEK293T Flp-In T-Rex (139 bait cell lines mapping to 119 unique genes)	96
Markmiller <i>et al.</i> 2018 ¹²⁰	APEX2 proximity labelling	500 μ M NaAsO ₂ 500 nM thapsigargin	HEK293 and NPCs - CRISPR Knock-in of bait (APEX2-G3BP1/NES)	221
Marmor-Kollet <i>et al.</i> 2020 ¹⁵¹	Multi-Bait APEX2 proximity labelling	400 μ M NaAsO ₂	U2OS cells - CRISPR KO of endogenous baits, stable expression of APEX2-tagged baits (G3BP1/FMR1/FXR1/NES)	240
Hu <i>et al.</i> 2023 ¹⁵⁰	Biochemical isolation	Heat shock – 43 °C	U2OS	442
Pan <i>et al.</i> 2024 ¹⁴⁸	μ Map proximity labelling	500 μ M NaAsO ₂	Stable HEK293T Flp-In T-Rex (HaloTag-G3BP1/NES)	371
Total number of reported SG proteins = 1150				

Interestingly, all PL-based studies reported the presence of a pre-existing interactions network even in the absence of stress. For instance, one study utilising a single APEX2-G3BP1 approach (single-bait APEX2) showed minimal interactome changes upon stress¹²⁰. Another study utilising BioID with hundreds of baits (multi-bait BioID) study revealed stable G3BP1- and EIF4 α 1-bait interactomes in response to various stressors (NaAsO₂, eIF4 α inhibitors

pateamine A and hippuristanol), whereas other baits like TIAL1 (TIAR) exhibited higher stress dependence¹⁴⁶. In contrast, the P body interactomes, interrogated in the same study, were distinct pre-and post-stress. Similar observations were reported by another APEX2 study utilising three baits (multi-bait APEX2), whereby G3BP1-, but not FMR1-/FXR1-baits, exhibited minimal stress-dependent interactome changes¹⁴⁷. Collectively, these observations suggested that SGs exist in submicroscopic structures (“seeds”) or dynamically cycle in interactions even under normal conditions, rapidly coalescing into visible granules upon stress¹⁴⁷. To circumvent this, both the single-bait APEX2 and multi-bait APEX2 approaches, as well as the μ Map approach, utilised a complementary construct containing the relative PL fused to nuclear export signal (NES), which facilitated comparison of stressed conditions to cytoplasmic labelling. The BioID approach on the other hand, profiled the SG “cores” in the absence of stress.

Comparative analysis of the SG proteomes reported by the six aforementioned studies revealed a limited overlap of the reported SG proteins, with only 85 proteins (7.7%) found in four or more datasets and 15 found in all six studies (**Fig. 2.2**). These discrepancies highlight

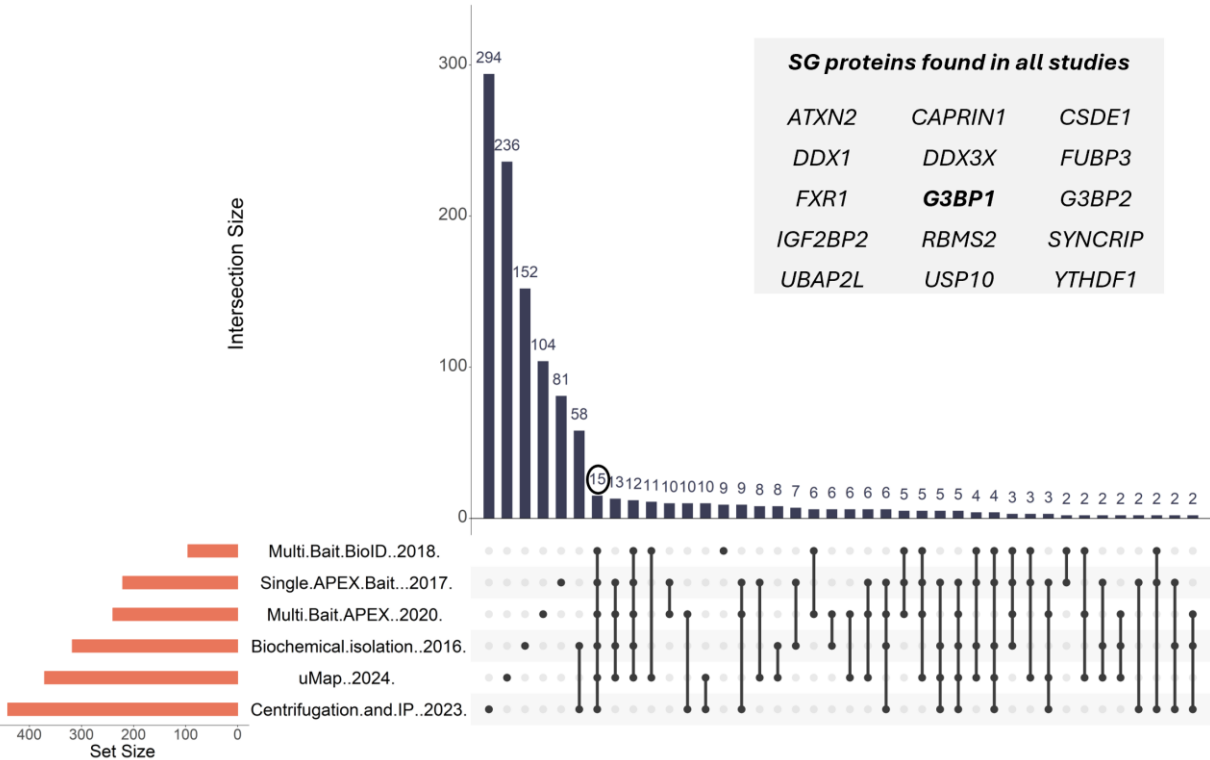


Figure 2.2. Comparison of stress granule composition datasets from five independent studies. UpSet plot showing the overlap of proteins identified as stress granule (SG) components across five different studies. Only 15 proteins were consistently identified in all five datasets.

the expected variability in SG composition depending on the stressor, cell type and model (knock-in vs stable expression), and experimental methodology. For example, SG cores typically measure at ~200-400 nm in diameter, whilst the labelling radius of some PL technologies is less than ~4 nm (**Table 2.1**)¹²⁴. Furthermore, the biochemical isolation and BioID approaches focused on SG cores, whereas APEX2 and μ Map studies profiled both core and stress-dependent “shell” assemblies. This compositional variability underscores the complexity of the dynamic kinetics and architecture of SGs.

For all experiments throughout this thesis, any proteins referred to as SG proteins are defined as identified in this curated database.

2.2.2 Comparison against SwissPalm

The SG proteome list was then cross-referenced against SwissPalm (Release 4), an open-access database manually curated with published S-acyl proteome datasets⁶³. Release 4 contained data from 38 S-acyl proteomic studies across 20 species (19 studies for humans), accumulating more than 18000 proteins and >6000 sites. Comparative analysis revealed that most SG proteins (70%) were found in at least one S-acyl proteomic study (**Fig. 2.3a**). When

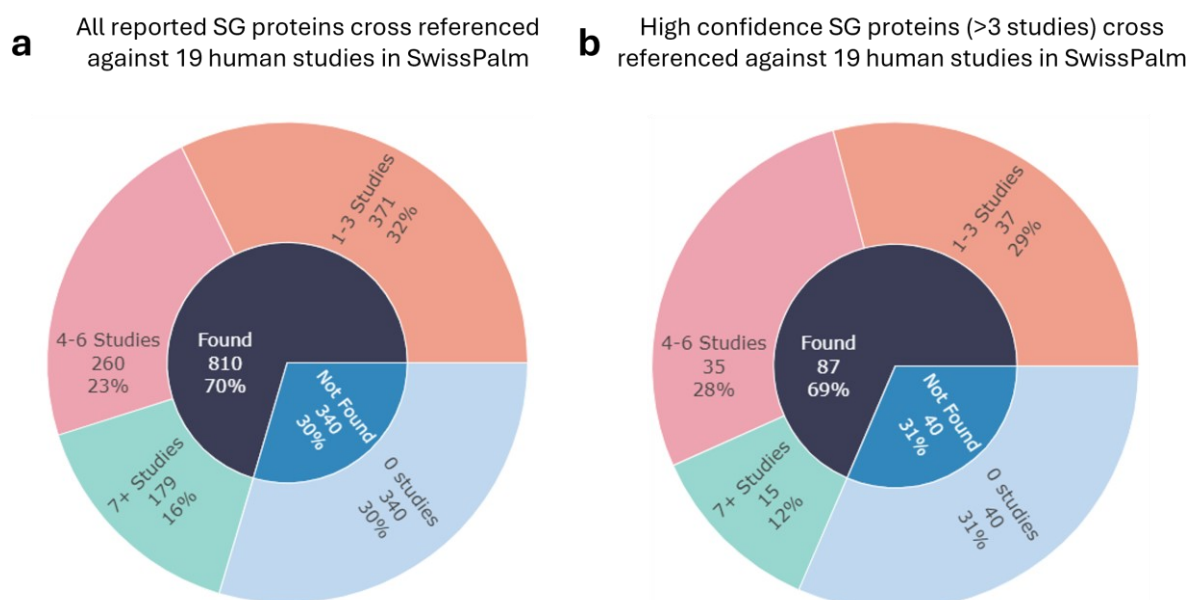


Figure 2.3. Comparative analysis of stress granule proteins with the SwissPalm S-acylation database. **a)** The 1150 SG proteins curated from five independent studies SwissPalm (Release 4) was utilised for comparison. **b)** High confidence SG proteins found in at least three of the SG composition studies cross-referenced against SwissPalm.

narrowing down to high-confidence SG proteins, which are defined as identified in more than three studies, the same trend was observed with 69% found in at least one study (**Fig. 2.3b**). Moreover, 179 proteins were found in more than seven S-acyl proteomic studies, bringing higher confidence to their true S-acylation status (**Fig. 2.3**).

According to our hypothesis, S-acylation of SG proteins would be induced upon stress (**Fig. 1.9**). Given that none of the S-acyl proteomic studies deposited onto SwissPalm investigated stress-dependent changes, relevant S-acylation events may not be represented. Furthermore, it is important to note that the SwissPalm database contains many false positive proteins due to the current limitations associated with S-acyl proteomic approaches, and lack of stringent curation criteria. Nonetheless, SwissPalm remains the only S-acylation substrate database to date.

For the remainder of this thesis, a protein or site stated as “previously found/reported on SwissPalm” refers to proteins or sites found in at least one S-acylproteomic or targeted S-acylation study, according to Release 4 of the SwissPalm database.

2.3 Conceptualisation of a double enrichment proteomics approach

With the bioinformatic analysis in support of our hypothesis, we then pursued the development of a novel double enrichment proteomics platform to identify stress-dependent S-acylation events within assembled SGs. It was hypothesised that combining PL, to specifically label SG components, with an S-acylproteomic technique for selective tagging of S-acylated Cys would achieve this objective (**Fig. 2.4**). Crucially, site-specific S-acylation identification was a key requirement for the combined workflow, as downstream hit validation would involve individual Cys mutagenesis, given the lack of selective inhibitors for the zDHHC enzymes.

APEX2 PL was selected as the labelling enzyme, and G3BP1 as the bait protein due to its established effectiveness of the APEX2-G3BP1 strategy in profiling the SG proteome (**Table 2.2**), and high temporal resolution, which would enable future interrogation of SG dynamics^{120,147}. As such, a U2OS cell line stably expressing APEX2 fused to the N-terminus of G3BP1, separated by a V5 linker, was selected (**Fig. 2.5a**). Notably, this construct is downstream of a tetracycline/doxycycline-inducible promoter, conferring the ability to titrate protein expression levels. This cell line additionally lacks endogenous G3BP1, and its

functionally redundant homologue G3BP2, achieved through CRISPR-Cas9-mediated genetic knock out (KO) (U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP, hereinafter referred to as APEX2-G3BP1). Together with the doxycycline inducibility, these features enable precise control over G3BP1-APEX2 concentration in the cell, preventing spontaneous SG formation due to overexpressed G3BP1¹¹⁰. Additionally, due to the pre-stress sub-microscopic assemblies typically observed in SG PL proteomic studies, an additional U2OS cell line stably expressing APEX2 fused to NES (hereinafter referred to as APEX2-NES), under a tetracycline-inducible promoter was selected (**Fig. 2.5a**). Both cell lines were designed, generated and kindly gifted to us by the Eran Hornstein lab at Weizmann Institute of Science, and were previously used to temporally map SG compositional changes, identifying SUMOylation as a regulator of SG disassembly¹⁴⁷.

As discussed in **Chapter 1.6.1**, direct mass spectrometric detection of S-acylation is hindered by its inherent lability and hydrophobicity. Therefore, integration of either an MLCC or a cysteine-centric AX step into the workflow was necessary (**Fig. 1.7**). MLCC approaches were deemed unsuitable for coupling to PL for two reasons. Firstly, all existing MLCC approaches utilise and are optimised for biotin-based pulldowns, clashing with the prerequisite biotinylation and streptavidin enrichment steps of the APEX2 workflow^{53,90,152,153}. Secondly, MLCC approaches lack the temporal resolution for capturing the hypothesised rapid changes in S-acylation, as the full assembly of NaAsO₂-induced SGs occurs in less than 30 mins¹⁵⁴. Consequently, an AX step following in-cell APEX2 labelling and cell lysis, was hypothesised to offer greater compatibility with the proposed workflow. However, the most commonly used AX-based methods, ABE and acyl-RAC, were similarly unsuitable due to either reliance on biotin or high protein input requirements (discussed in **Chapter 3.1**), as well as a lack of commercially available 6B-Sepharose resin (**Table 3.1**).

Due to these limitations, alternative AX tags for the alkylation of liberated thiols following HA-mediated acyl cleavage were explored. Although the cysteine-centric toolbox has expanded in recent years (**Table 3.1**), few site-specific tagging approaches (alternative to ABE and acyl-RAC) existed at the time of this workflow's conceptualisation. Concurrently to this PhD, a separate project within our research group was employing a Cys reactive phosphonate adaptable tag (CysPAT) for reactive Cys profiling. This CysPAT probe, comprising an iodoacetamide group for thiol reactivity and a phosphonic acid group enabling enrichment of Cys-containing peptides by immobilised metal affinity chromatography (IMAC), was hypothesised to be compatible with the proposed approach (**Fig. 2.4**)¹⁵⁵. The final design of the

workflow therefore involved APEX2-mediated PL, acyl-phosphonate exchange (AX) reaction, streptavidin enrichment, subsequent on-bead digestion and IMAC-based peptide enrichment (**Fig. 2.4**). This approach, termed APEX-AX, would therefore enrich for and identify S-acylated SG proteins with site-specific resolution. Nonetheless, potential caveats of this workflow include possible steric hindrance impeding accessibility of the HA/CysPAT labelling to certain Cys due to proximal biotinylated Tyrosines. Moreover, the effects of the H_2O_2 treatment on the Cys modification landscape, particularly stability of the thioester bond, remain unknown.

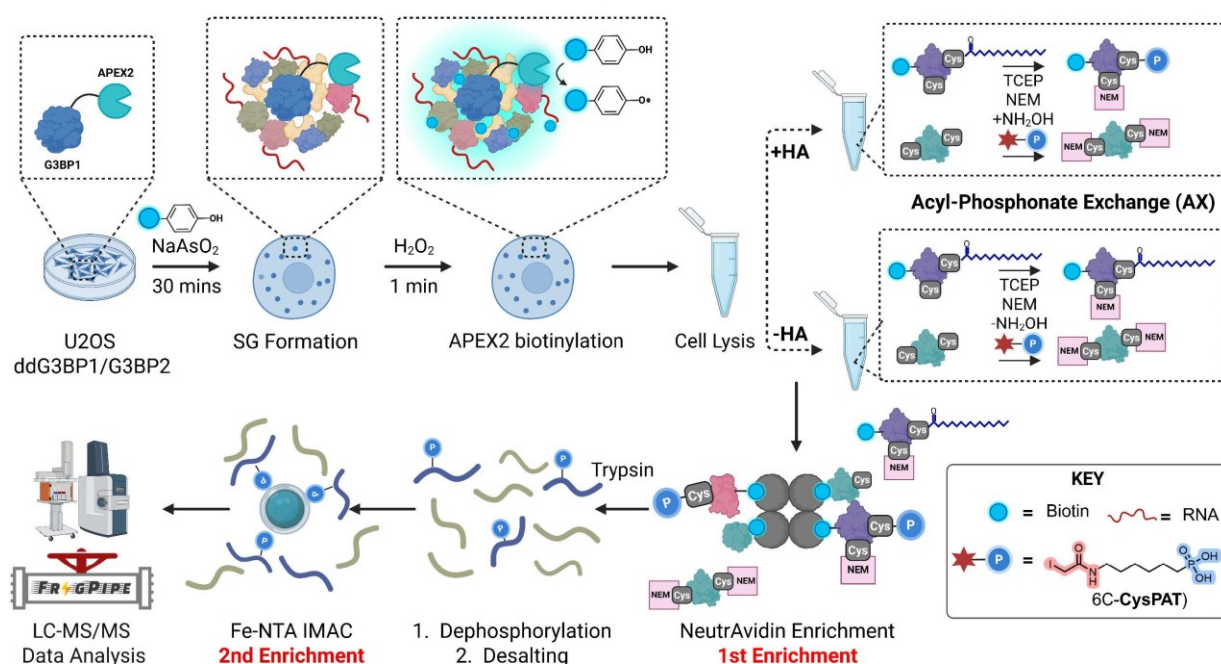


Figure 2.4. Schematic of APEX-AX: a double enrichment proteomics platform to identify stress-dependent changes in S-acylation of stress granule (SG) proteins. U2OS cells lacking endogenous G3BP1/2 are induced with doxycycline (50 ng/mL) to express an APEX2-G3BP1 fusion protein. The following day, cells are treated with 500 μM sodium arsenite (NaAsO_2) for 30 min to induce SG formation, concurrently with biotin-phenol (cyan). After 30 minutes, APEX2-mediated biotinylation is initiated by adding 1 mM H_2O_2 for 1 min, followed by quenching and cell lysis. Lysates are then subjected to the acyl-phosphonate exchange (AX) reaction, which involves blocking free cysteines with N-ethylmaleimide (NEM). Samples are split into two aliquots: one treated with hydroxylamine (NH_2OH) and one untreated (negative control). Simultaneously, 100 μM CysPAT is added. If a cysteine is S-acylated, it becomes labelled with a phosphonate group. Biotinylated proteins are then enriched with NeutrAvidin beads, and digested on-bead. Resulting peptides are treated with a phosphatase to remove endogenous phosphorylation, desalted, and subjected to Fe-NTA immobilised metal affinity chromatography (IMAC) to enrich phosphonate-labelled (i.e. S-acylated) peptides. Peptides are then identified by LC-MS/MS. This workflow ensures identification of SG-localised, S-acylated proteins with site-specific resolution. Figure made with Biorender.com.

The conceptualisation of the APEX-AX workflow has further emphasised the pressing need for robust, site-specific S-acylproteomic technologies. This lack of alternative tagging approaches led to the selection of CysPAT, a probe developed and solely utilised for Cys profiling, and unexplored for identification of S-acylation^{155,156}. Its advantageous site-specificity and scalability presented an opportunity for its employment in S-acylproteomics, regardless of APEX-AX. Thus, the focus of this PhD project later shifted towards developing a novel cysteine-centric proteomics platform, termed SAPPHIRE, for site-specific S-acylated proteome profiling through leveraging the CysPAT probe. The conceptual development, optimisation and application of this strategy are comprehensively described in **Chapter 3** and **Chapter 4**.

2.4 Validation of APEX2 activity and SG localisation

2.4.1 Cell line and activity validation

Upon receiving the engineered U2OS cell lines, the correct sequences of the genomically integrated APEX2-G3BP1 and APEX2-NES constructs were validated by genomic extraction, polymerase chain reaction (PCR) and subsequent Sanger sequencing (data not shown). The APEX2-G3BP1 cell line was additionally evaluated for KO of endogenous G3BP1 and stable expression of APEX2-G3BP1 by Western blot analysis (**Fig. 2.5b**). Construct expression was

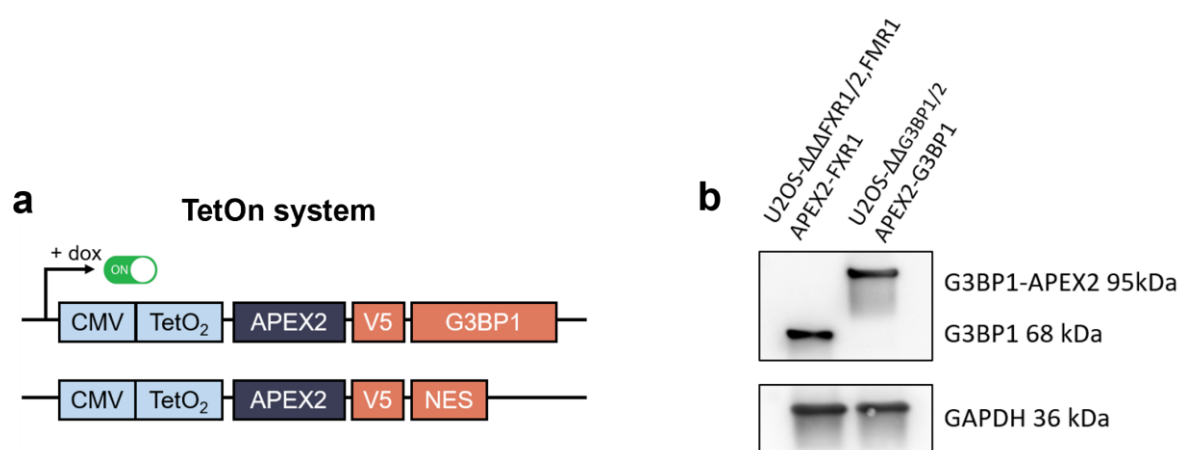


Figure 2.5. APEX2 construct design and cell line validation. **a)** Schematic of the two APEX2 constructs used in this study: APEX2-G3BP1 and APEX2-NES, both under a tetracycline/doxycycline-inducible promoter for controlled expression. **b)** Western blot validation of endogenous G3BP1 knock-out in U2OS- $\Delta\Delta$ G3BP1/2 cells expressing APEX2-G3BP1. Only the 95 kDa fusion construct is detected in this cell line.

induced for 24 hours with a low concentration of doxycycline (50 ng/mL), previously reported to equate endogenous expression of G3BP1¹⁴⁷. A U2OS cell line lacking endogenous FXR1/FXR2/FMR1 and expressing APEX2-FMR1 was used as the G3BP1 wild type (WT) control¹⁴⁷. As expected, only a single 95 kDa band, corresponding to the molecular weight of APEX2-G3BP1, was detected, and its relative expression closely matched that of endogenous G3BP1. Given the critical role of protein concentration on LLPS, preventing overexpression artifacts is of critical importance. Thus, this doxycycline concentration (50 ng/ mL) was used in all subsequent experiments for both constructs.

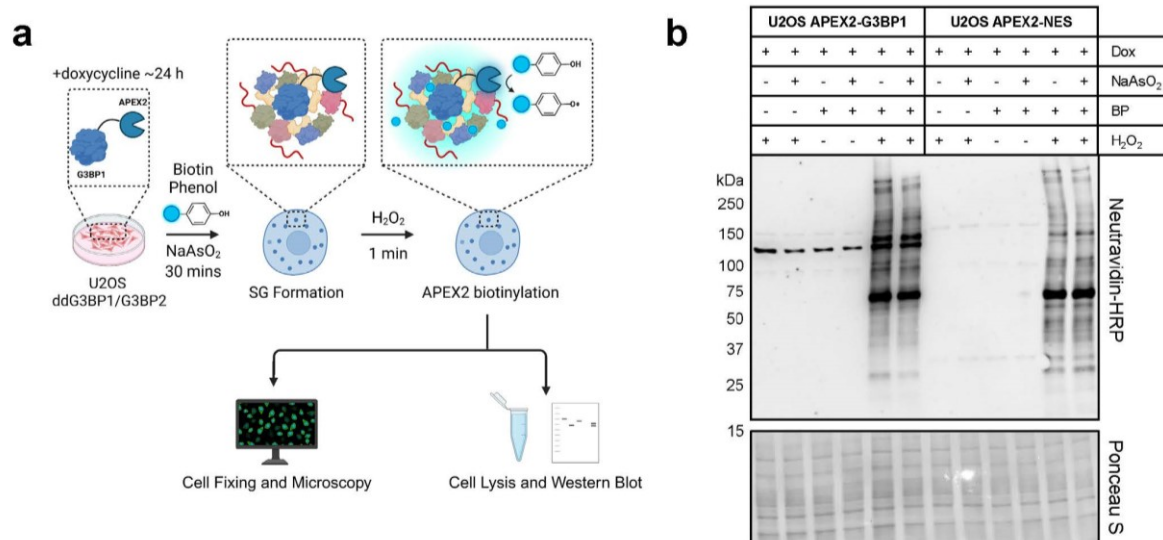


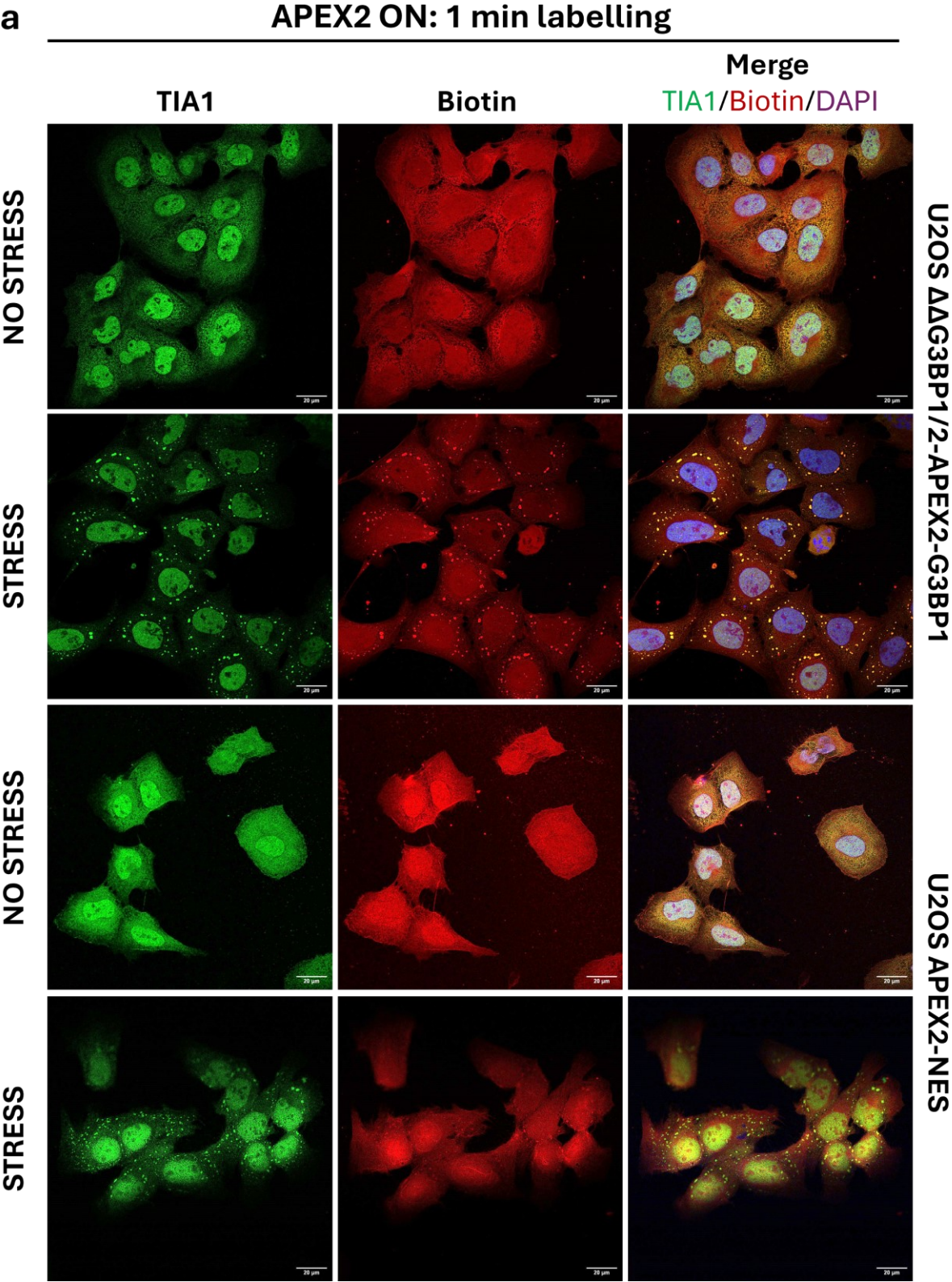
Figure 2.6. APEX2 activity requires both biotin-phenol and H₂O₂. **a)** Schematic of the experimental workflow used to validate APEX2 construct activity and subcellular localisation in U2OS cells expressing either APEX2-NES or APEX2-G3BP1. Two experimental workflows were performed: In both, cells were first induced with doxycycline (50 ng/mL) to express the respective fusion protein. The following day, cells were treated with 500 μ M sodium arsenite (NaAsO₂) for 30 minutes to induce stress granule (SG) formation, concurrently with biotin-phenol (cyan) due to similar required incubation times. After 30 minutes, APEX2-mediated biotinylation was initiated by adding 1 mM H₂O₂ for 1 minute, followed by quenching. For activity validation, cells were lysed and analysed by western blot. For localisation analysis, cells were fixed, permeabilised, immunostained, and imaged using confocal microscopy. Figure made with Biorender.com. **b)** Western blot using NeutrAvidin-HRP showing that APEX2-dependent biotinylation occurs only in the presence of both biotin-phenol and H₂O₂.

Next, to validate the biotinylation activity of the APEX2 constructs, a Western blot analysis using Neutravidin-HRP conjugate was performed (**Fig. 2.6a**). Both cell lines were treated with doxycycline for 24 hours, followed by treatment with 500 μ M NaAsO₂ and 500 μ M BP for 30

mins. APEX2 labelling was then induced by treatment with H_2O_2 for one minute before quenching and cell lysis. Control conditions, omitting either BP, H_2O_2 or NaAsO_2 , were also included. As expected, robust APEX2 biotinylation activity was observed only in the presence of both BP and H_2O_2 , with only bands corresponding to endogenously biotinylated proteins, at 130 kDa and 75 kDa, detected when either reagent was omitted (**Fig. 2.6b**)¹³⁹. Furthermore, the overall biotinylation patterns remained uncompromised between NaAsO_2 stressed and non-stressed conditions in both cell lines.

2.4.2 Confocal Microscopy

The correct localisation and activity of the APEX2 constructs were then assessed using immunofluorescence microscopy (**Fig. 2.6a**). Both cell lines were treated as described above, and following quenching, cells were fixed, permeabilised and immunostained for T-cell intracellular antigen-1 (TIA1), a well-characterised SG marker (green channel), biotinylation using a Neutravidin-TexasRed conjugate, and nuclei using 4',6-diamidino-2-phenylindole (DAPI) (blue channel)¹⁴⁷. Pleasingly, SG formation and SG-localised biotinylation were observed in APEX2-G3BP1 cells only under stress, whereas no SGs and a diffused biotinylation signal occurred under normal conditions (**Fig. 2.7a**). In contrast, APEX2-NES cells displayed a diffused biotinylation pattern under both stress and non-stress conditions, consistent with the cytoplasmic distribution of the construct. Importantly, no biotinylation signal was detected when BP was omitted (APEX2 off), whilst SG formation still occurred (**Fig. 2.7b**). Collectively, these results validated the correct activity and localisation of the APEX2 constructs in both cell lines.



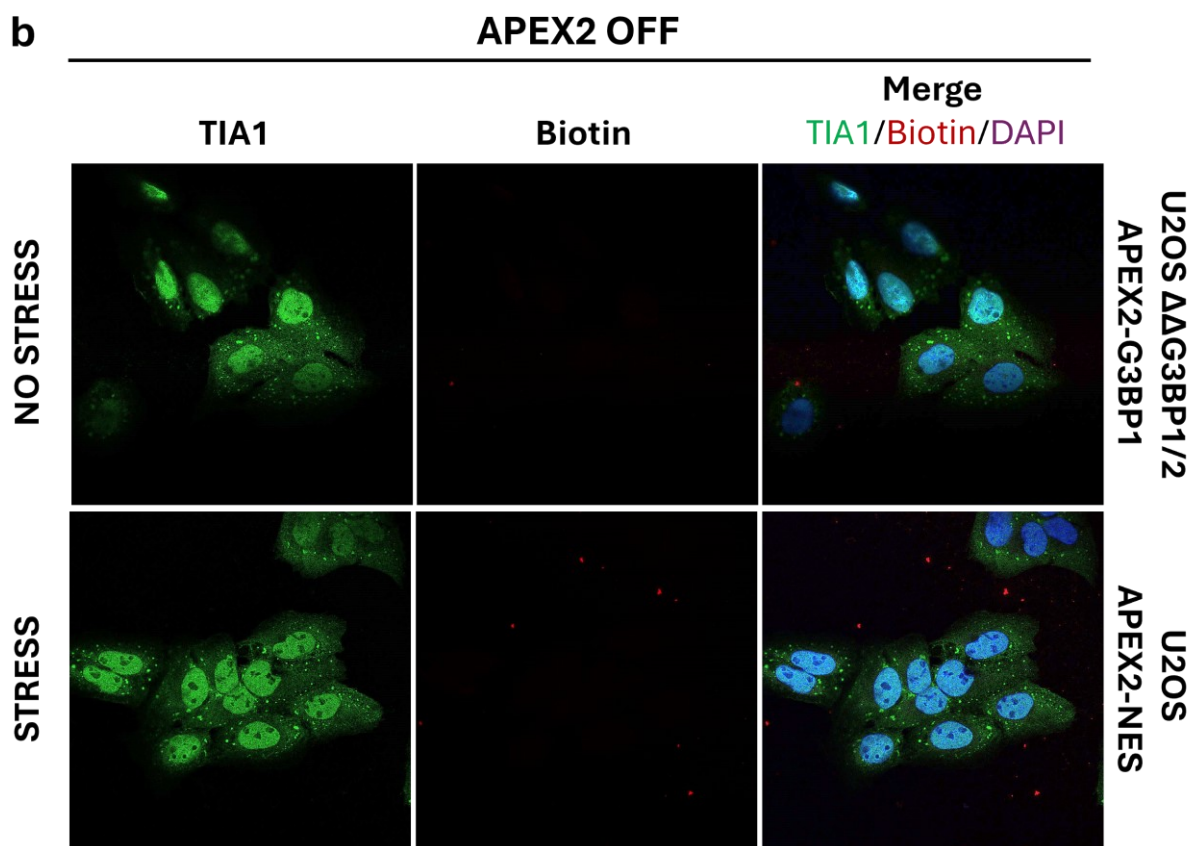


Figure 2.7. APEX2-G3BP1 localises to stress granules in response to stress Confocal microscopy showing **a)** APEX2-G3BP1 and APEX2-NES activity in U2OS cells, with or without sodium arsenite stress (NaAsO₂ 500 μM, 30 min). Immunofluorescence depiction of TIA1, Neutravidin-Texas-red staining of biotinylated proteins at the proximity of the APEX bait, and merged signal demonstrating the precise localization of the APEX activity in stress granules (SGs). Lens, ×63; scale bar, 10 μm. **b)** Lack of APEX2-G3BP1 and APEX2-NES biotinylating activity in U2OS cells when biotin phenol was omitted.

2.5 APEX-AX for S-acylated SG proteome identification

2.5.1 Pilot experiment

With the validated constructs in hand, a pilot APEX-AX experiment was conducted. As described throughout this chapter, the primary objective of the APEX-AX approach was the identification of stress-dependent changes in S-acylation patterns of SG proteins, rather than comprehensive profiling of the composition of SGs. Moreover, given the general paucity of

double-enrichment proteomics methods, extensive optimisation of the workflow was anticipated. Therefore, only the APEX2-G3BP1 cell line was employed for the initial experiment to reduce experimental complexity, thereby minimising human errors and costs, particularly given the lack of established high-throughput proteomics protocols in our lab at the time. Control conditions such as APEX2-NES and an 'APEX2-off' condition, achieved by omitting either BP or H₂O₂, to account for non-specific NeutrAvidin bead binding, will be critical for future studies. However, their inclusion was not prioritised at this early stage until a fully optimised workflow has been established.

The pilot experiment aimed to explore the feasibility of the APEX-AX concept without extensive optimisation of either the PL or AX parts of the workflow. As such, general conditions recommended for the APEX2 PL and acyl-exchange workflows were selected with comprehensive description of this experiment's workflow is detailed in **Methods 6.5.4.2.1**. Briefly, APEX2-G3BP1 cells were induced with doxycycline, either NaAsO₂ *stressed* or left untreated (no stress condition) for 30 mins, and APEX2 biotinylation was induced for one minute (**Fig. 2.4**). Cells were then lysed, split into four technical replicates and subjected to the acyl-phosphonate exchange reaction. AX conditions were not optimised prior to this experiment and were mostly derived from the previously reported acyl-RAC and acyl-mPEG exchange (APE) protocols, with some modifications (see **Methods 6.2.2**)^{90,157}. Notably, HA cleavage and CysPAT labelling reactions were conducted concurrently, using a reduced concentration of CysPAT (100 µM), as used for reactive Cys profiling, compared to 1 mM mPEG-maleimide concentration used in APE. The AX reaction involved sequential incubation with TCEP for disulfide reduction and NEM for free Cys alkylation, with an ACN-based protein precipitation after each step. Each sample was then split into two aliquots, a no HA control – where only the CysPAT probe was added, and a HA sample where both HA and CysPAT were added to facilitate phosphonate labelling. Subsequently, all samples were subjected to another ACN-based precipitation and 500 µg protein per sample was enriched using Neutravidin agarose beads. followed by on-bead digestion and peptide dephosphorylation with Quick CIP phosphatase. All samples were subsequently desalted using Pierce™ peptide desalting spin columns (Thermo) and subjected to IMAC enrichment using the Fe-NTA phosphopeptide enrichment columns (Thermo). Notably, pre-desalting peptide quantification was below the detection limit of the fluorometric peptide quantification kit (ThermoFisher). The eluted peptides were then analysed using the Evosep One coupled to a Bruker trapped ion

mobility spectrometry time-of-flight mass spectrometer (timsTOF HT) LC-MS/MS set up. Raw data search was then performed using the MSFragger search engine with NEM- and phosphonate-modified Cys (pCys) set as variable modifications (**Fig. 2.4**).

After filtering for common contaminants and three valid values in at least one condition, only 16 peptides across all samples were identified, none of which had the phosphonate modification. It was hypothesised that the extensive workflow described above, led to negligible peptide amounts even prior to elution from the IMAC columns. This initial failure further underscores the challenges of APEX-AX design and optimisation requirements. Firstly, no intermediate checks were performed to assess the APEX2 labelling and Neutravidin enrichment efficiency. Whilst the labelling conditions have been validated throughout this chapter, the Neutravidin enrichment conditions were derived from a previous proteomics experiment utilising this cell line (data not shown). Nevertheless, without direct analysis, such as a Neutravidin-HRP Western blot of lysates pre- and post-pulldown, it remains unknown whether the PL portion of the workflow requires further optimisation. Secondly, the CysPAT AX reaction was not optimised, which could have led to inefficient phosphonate labelling, accounting for the lack of protein/peptide MS identifications. Thirdly, systematic optimisations to determine sufficient protein and peptide input requirements for the Neutravidin and IMAC enrichments respectively, would be necessary.

As mentioned previously, the CysPAT probe has never been applied for S-acylproteomics. Thus, the establishment of CysPAT AX alone was anticipated to involve an extensive round of optimisation, given the general challenges associated with cysteine-centric approaches (discussed in **Chapter 3.1**). Collectively, this motivated the development of SAPPHIRE, detailed comprehensively in **Chapter 3**.

2.5.2 APEX-AX: Phosphonate as a mass tag

Due to the failure of the double-enrichment strategy, the use of CysPAT as a mass tag, rather than a secondary enrichment handle, was hypothesised to circumvent the observed peptide input challenges whilst still conferring SG protein S-acyl site identification. Following the establishment and optimisation of SAPPHIRE, the revised workflow was applied to APEX2-G3BP1 cells in four biological replicates (**Fig. 2.8**). After APEX2 labelling and lysis, phosphonate AX reaction was performed as per the optimised SAPPHIRE protocol, involving double NEM alkylation and “one-pot” HA-CysPAT labelling steps. Subsequently, 600 µg of

protein per sample was enriched using Neutravidin beads, tryptic digested and analysed by LC-MS/MS.

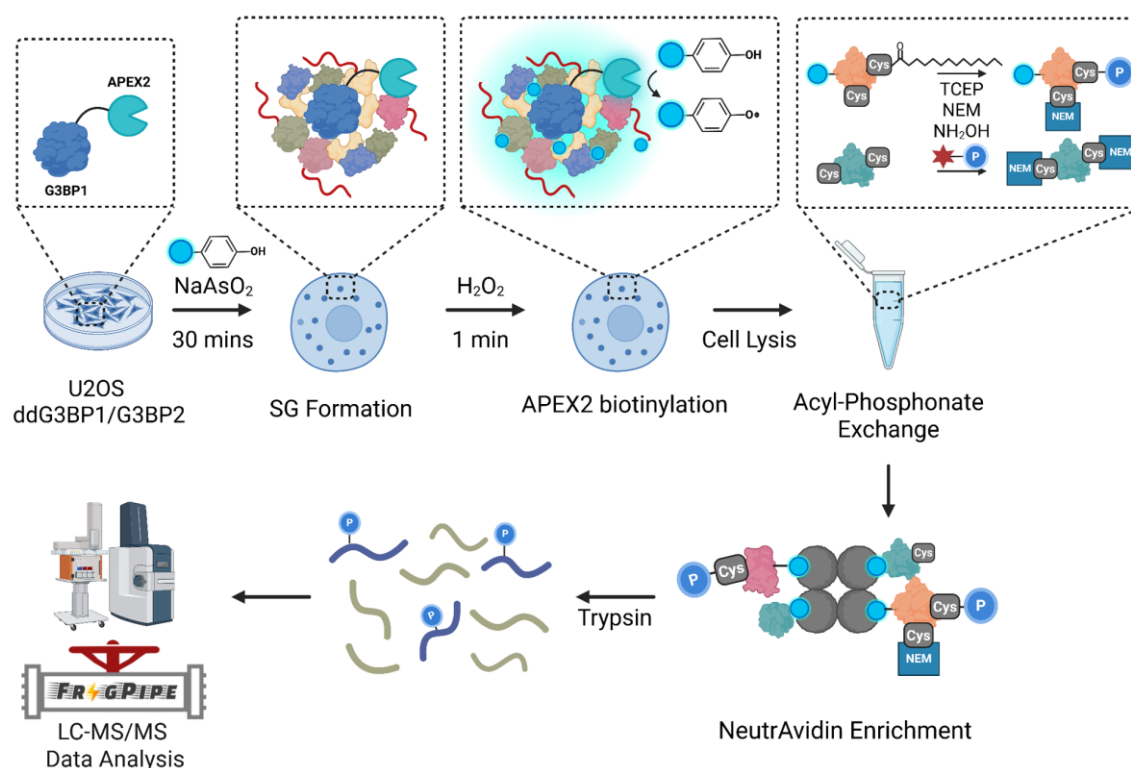


Figure 2.8. Modified APEX-AX workflow: phosphonate as a mass tag. U2OS cells lacking endogenous G3BP1/2 are induced with doxycycline (50 ng/mL) to express an APEX2-G3BP1 fusion protein. The following day, cells are treated with 500 μ M sodium arsenite (NaAsO₂) for 30 min to induce SG formation, concurrently with biotin-phenol (cyan). After 30 minutes, APEX2-mediated biotinylation is initiated by adding 1 mM H₂O₂ for 1 min, followed by quenching and cell lysis. Lysates are then subjected to the acyl-phosphonate exchange (AX) reaction, which involves blocking free cysteines with N-ethylmaleimide (NEM). Samples are split into two aliquots: one treated with hydroxylamine (NH₂OH) and one untreated (negative control). Simultaneously, 100 μ M CysPAT is added. If a cysteine is S-acylated, it becomes labelled with a phosphonate group. Biotinylated proteins are then enriched with NeutrAvidin beads and digested on-bead. The resulting peptides are then identified by LC-MS/MS. A phosphonated Cys modifications is then searched for using the Fragpipe software. Figure made with Biorender.com.

APEX2 labelling and pulldown efficiency were evaluated using a Neutravidin-HRP blot, demonstrating efficient enrichment of biotinylated proteins, predominantly retained in the beads fraction and minor presence in the supernatant (**Fig. 2.9a**). Furthermore, principal component analysis (PCA) pleasingly showed a clear separation of the stress and no stress

clusters, indicating good reproducibility across biological replicates and differential proteomic profiles between tested conditions (**Fig. 2.9b**).

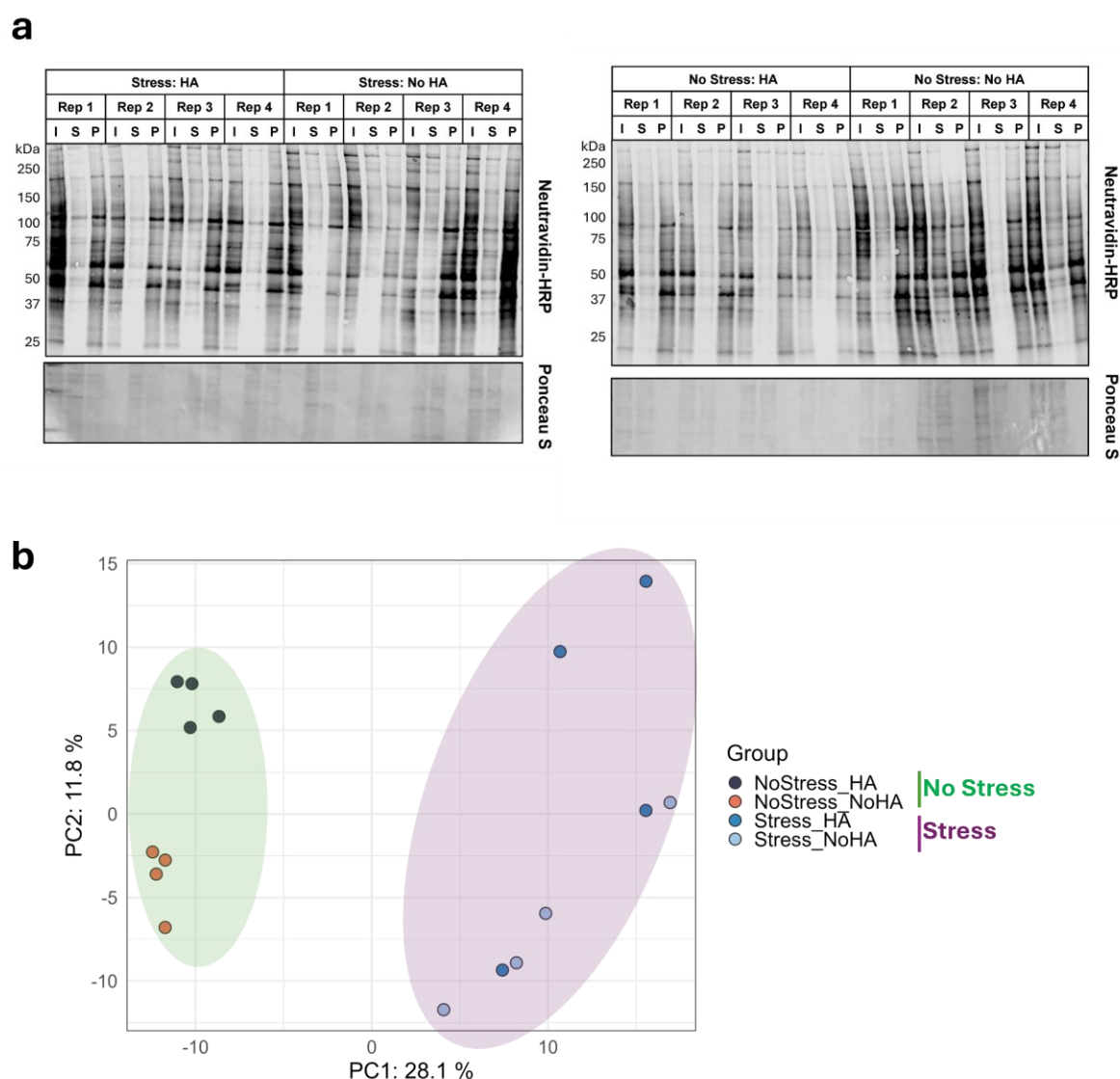
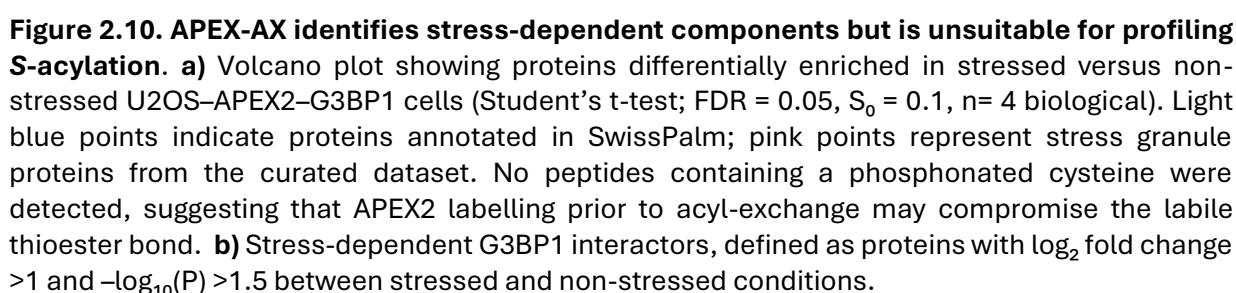


Figure 2.9. Assessment of labelling and enrichment efficiency. a) Western blot using NeutrAvidin–HRP to assess biotinylation in input (I), supernatant (S), and pulldown (P) fractions from stressed and non-stressed samples used for modified APEX-AX proteomics. Efficient enrichment is indicated by the absence of biotinylated proteins in the supernatant. **b)** Principal component analysis (PCA) showing clear separation between stressed and non-stressed samples, indicating robust biological distinction.

In total, 1221 proteins were identified across all samples, with 39 significantly enriched in the stress compared to no stress condition (**Fig. 2.10a,b**). Of these stress-dependent interactors, 33 had been previously identified as a SG protein, and 31 were found within the SwissPalm



64

bioinformatic analysis, as well as the identification of over 1900 putative Cys belonging to 597 SG proteins in HEK293T cells (please refer to results presented in **Chapter 3.4.1**). The complete absence of phosphonate-modified Cys here suggests a potential destabilisation of the labile thioester bond by H_2O_2 , required for APEX2 activity, precluding downstream phosphonate tagging. However, investigation of this hypothesis remained beyond the timeline of this PhD's project.

2.6 Discussion

2.6.1 Double or nothing: Challenges in double enrichment proteomics

Double tagging or double enrichment approaches remain relatively uncommon in the field of proteomics and PTM analysis. At the beginning of this PhD and generation of this hypothesis, only one double enrichment proteomics technology has been reported¹⁵⁸. Since then, three additional methods have emerged, resulting in just four approaches that combine PL with a second tag/enrichment reported to date. These include subcellular-specific uncaging-assisted biotinylation and mapping of phosphoproteome (SubMAPP), local cysteine capture (Cys-LoC) and local cysteine oxidation (Cys-LOx), Surface Cysteine enrichment (Cys-Surf) and Transit ID^{158–161}. These methods are summarised in **Table 2.3**.

Table 2.3. Summary of double enrichment proteomics approaches

Name	Year of publication	1 st enrichment	2 nd enrichment	Protein input per sample
APEX-AX proposal	2021	Protein level: Biotin-Neutravidin	Peptide level: Phosphonate-IMAC	500-600 µg
SubMAPP	2021	Protein level: Biotin-Streptavidin	Peptide level: Phosphate-IMAC	10 mg
CysLOx/CysLoC	2023	Protein level: Biotin-Streptavidin	Peptide level: Biotin-Neutravidin	500 µg
TransitID	2023	Protein-level: Anti-fluorescein IP	Protein-level: Biotin-Streptavidin	2 mg
Cys-Surf	2023	Protein level: Biotin-Streptavidin	Peptide level: Biotin-Neutravidin	750 µg

SubMAPP technology was developed for mapping phosphorylation within subcellular compartments, and most closely resembles the APEX-AX enrichment strategy, as it involves photoactivatable TurboID-mediated labelling, and subsequent combination of biotin-streptavidin and phosphopeptide Titanium dioxide (TiO₂) IMAC enrichments¹⁵⁸. In contrast, TransitID involves TurboID biotinylation and APEX2 labelling employing an optimised alkyne phenol probe, click-chemistry to azide-fluorescein, and subsequent sequential anti-fluorescein IP and biotin/streptavidin enrichments¹⁶¹. TransitID was similarly used for profiling protein trafficking between the nucleolus and SGs and vice versa, utilising G3BP1-APEX2 and G3BP1-LOV-Turbo constructs. Conversely, the Cys-LoC/Cys-LOx and Cys-Surf approaches demonstrated a double biotin enrichment strategy^{159,160}.

Although successful, all double enrichment approaches have faced significant challenges during their development. For example, the TransitID authors have specifically highlighted the establishment of an orthogonal label compatible with an enrichment technique as the biggest challenge, selecting a fluorescein-based tag after failures with azide-agarose and azide-FLAG enrichments¹⁶¹. Additionally, double enrichment methods (apart from Cys-LoC/Cys-LOx) require significantly higher protein inputs, with SubMAPP utilising 20-fold higher protein input than the input utilised in the pilot APEX-AX experiment (**Table 2.3**). Given the numerous protein precipitation steps required for the APEX-AX reaction prior to Neutravidin enrichment, obtaining such high protein inputs would present a practical challenge.

The attempts demonstrated herein, as well as other failed attempts in the lab (data not shown), have led to the conclusion that the use of APEX2 labelling may be fundamentally incompatible for profiling S-acylation. Whereas the pilot APEX-AX experiment likely failed due to insufficient optimisation of the extensively complex approach, the mass-tag approach with the optimised acyl-phosphonate exchange reaction, theoretically should have identified a number of S-acylated Cys. To circumvent these limitations, alternative PL strategies were considered, particularly miniTurbo and LOV-Turbo approaches. Notably, implementation of TurboID for the Cys-LoC/Cys-LOx approaches, poses significant challenges with promiscuous background labelling due to the presence of endogenous biotin, as well as newly translated TurboID¹⁵⁹. As such, establishment of a Turbo-based strategy was likewise expected to be time-consuming.

Cumulatively, these existing technologies underscore the inherent technical complexities and significant optimisation associated with double enrichment proteomics.

2.6.2 Challenges of profiling S-acylation

Probing the role of any PTM in SGs, and condensates in general, presents inherent difficulties. This is further exacerbated for S-acylation, due to the lack of specific chemical (inhibitors and probes) and analytical (antibodies, intact S-acyl ID) tools to study it. These limitations are especially apparent when compared to studies of other PTMs in SGs, which typically rely on the use of specific inhibitors or antibodies. For example, FDA-approved Olaparib and PAR antibody were utilised to study PARylation in SGs, 2D08 inhibitor for SUMOylation, SBI-0206965 for phosphorylation, TAK-243 for ubiquitylation, and MLN4924 for NEDDylation^{147,162–164}. It is important to note that the successful identification of S-acylation's role in NLRP3 LLPS, relied on a more targeted, hypothesis-driven approach, rather than broad proteomic studies¹⁰⁸. In that study, a CRISPR-Cas9-based screen identified the requirement of zDHHC7 for NLRP3 activation, leading to subsequent validation using microscopy approaches.

Even if APEX-AX or other PL-based approaches were to successfully identify a list of candidate stress-dependent S-acylated Cys within SG proteins, their downstream validation would face significant challenges due to this paucity of tools. In the absence of specific inhibitors, validation would primarily rely on site-directed mutagenesis of individual Cys, followed by microscopy-based approaches, such as FRAP and live-cell imaging, to assess their functional relevance in SG kinetics. Given the critical role that concentration plays in LLPS, PCR-based site-directed mutagenesis approaches and subsequent transient transfection, may lead to spontaneous SG formation by increasing the C_{sat} . As such, individual Cys site mutagenesis through laborious CRISPR knock-in approaches would likely be necessary. Moreover, it is possible that multiple S-acylation events contribute to SG assembly. Taken together with the anticipated optimisation of these approaches, the deconvolution of these hits and SG effects would be extremely laborious and time-consuming.

Considering these challenges alongside the timeline of this PhD, the focus of this PhD project strategically pivoted towards further optimisation and application of the SAPPHIRE approach for robust and efficient profiling of protein S-acylation.

2.7 Chapter summary

This chapter presents a proteomics-based strategy for the initial investigation of the role of S-acylation in biomolecular condensate regulation, using inducible SGs as a model condensate system. An initial bioinformatic analysis highlighted a significant overlap between known S-acylated proteins and the SG proteome, leading to the design of a novel double-enrichment proteomics approach termed APEX-AX. This approach combined APEX2-G3BP1 proximity labelling, for the identification of the SG proteome, and acyl-phosphonate exchange reactions to enable the identification of stress-dependent changes in the S-acylation states of SG proteins. While validation experiments successfully demonstrated the efficacy of APEX2-G3BP1-mediated biotinylation and localisation, initial attempts at the full APEX-AX were unsuccessful due to the complexity and insufficient optimisation of the workflow. Further efforts using a revised APEX-AX workflow utilising double tagging (biotin and phosphonate) but a single enrichment (Neutravidin), also failed in identifying any putative S-acylated Cys. However, this revised approach successfully identified 33 known SG proteins and 16 stress-dependent interactors, validating the robustness of the APEX2-G3BP1 proximity labelling component.

The inability to detect phosphonate-modified peptides likely resulted from the incompatibility between the H₂O₂-dependent APEX2 method and the detection of labile S-acylation. Given these challenges, further development of APEX-AX or alternative PL-based approaches was deemed too time-consuming. Given that the scarcity of effective non-biotin-based enrichment strategies and high-input requirements for existing S-acylproteomic techniques became evident during the development of APEX-AX, research focus shifted towards optimising and refining the acyl-phosphonate exchange technique to establish a comprehensive S-acylproteomic method. Consequently, Chapters 3 and 4 focus on the development, optimisation, and application of the **S**-acylation **P**rofilng by **P**hosphonate Tagging **P**roteomics (**SAPPHIRE**) method.

Chapter 3.

SAPPHIRE: a novel proteomics platform for
sensitive and high-throughput analysis of
S-acylation

3.1 Introduction

Protein S-acylation is a ubiquitous PTM that involves the covalent attachment of long-chain fatty acids (mainly C16:0 palmitate) onto Cys residues via a labile thioester bond. Since its initial discovery in 1979 by radiolabelling and autoradiography approaches, decades of technological progress have shed light on the extensive abundance (ca. 10-20% of the human proteome) and functional importance of this PTM in human biology^{86,87,165}. Unlike other lipid PTMs, S-acylation is reversible, functioning as a dynamic regulator of protein localisation, trafficking, and signalling, cell migration, division, polarity, metabolism and death, innate immunity, and tissue function^{54,76,101,165-170}. Aberrant S-acylation is linked to the progression of many cancers, resistance to cancer treatment, and neurodegenerative disorders, positioning it as a promising therapeutic target, particularly for the otherwise undruggable oncoproteins^{54,83,171-177}.

Despite its biological significance, the study of S-acylation in health and disease is still largely limited due to a lack of scalable tools for its proteome-scale analysis and site characterisation. Bottom-up proteomics using LC-MS/MS is a powerful and widely adopted strategy for global analysis of PTMs, yet investigating S-acylation presents four major analytical challenges. First, the inherent lability and reversibility of the thioester bond results in high tendency for the loss of the acyl moiety during sample preparation, MS/MS fragmentation or via *S*- to *N*- intramolecular acyl transfer to proximal lysine residues^{88,178-180}. Second, the substantially increased hydrophobicity of S-acylated peptides compared to their non-modified counterparts creates solubility issues and prevents their simultaneous identification in a single LC-MS/MS run⁸⁸. Third, the physicochemical properties and non-antigenicity of S-acylation, prevent direct enrichment of intact lipidated peptides from complex protein digests. Fourth, the dynamic nature of S-acylation leads to variable site stoichiometry, ranging from 1% to 100% occupancy, requiring highly sensitive methods to enable identification of low occupancy sites^{87,181}.

Collectively, these analytical hurdles have driven the development of many cysteine-centric and lipid-centric technologies that enabled discovery of thousands of S-acylation substrates and sites⁶³. Nevertheless, most existing methods are unable to characterise the precise modified site, require high protein inputs, remain labour-intensive and time consuming, suffer

from high background, and lack the scalability or automation potential required for large scale experiments (**Table 3.1**). Furthermore, despite broad application of cysteine-centric techniques across various species and tissues, human tumour samples remained unexplored, likely due to the high protein input requirements and limited patient sample availability^{62,182–186}. Characterising distinct S-acylation patterns in human tumours would build on existing evidence implicating aberrant S-acylation in cancer, further driving our understanding of its precise role and promising drug targets^{187–190}.

To address these limitations, the development of a novel, site-specific S-acylproteomic workflow applicable to diverse biological samples was pursued. Given that lipid-centric approaches require metabolic incorporation, and are therefore unsuitable for tissue-level analysis, a cysteine-centric strategy was selected, with the following aims in mind:

- 1) To improve enrichment specificity
- 2) To reduce required protein input, enabling human tumour sample analysis
- 3) To increase the scalability of the workflow to a 96-well plate
- 4) To ensure cost efficiency

Over the course of this chapter, the conceptualisation, optimisation and initial applications of this method will be discussed.

3.2 SAPPHIRE Proof-of-Concept

3.2.1 SAPPHIRE Conceptualisation

As discussed in Chapter 1, cysteine-centric S-acylproteomics typically involve reducing any disulfide bridges with TCEP, blocking of free and disulfide Cys by alkylation with *N*-ethyl maleimide (NEM), selective cleavage of thioester bonds using hydroxylamine (HA) at neutral pH, and subsequent labelling of the newly revealed Cys with a thiol-reactive biotin affinity tag (ABE and ABE-related approaches) or resin (acyl-RAC)⁸⁷. Both ABE and acyl-RAC require large protein inputs, laborious protein precipitation steps, and involve harsh washing and elution conditions, such as extreme detergents and temperatures, which may hydrolyse the labile thioester bond (**Table 3.1**).

Table 3.1. Summary of LC-MS/MS-based S-acylproteomic approaches. The indicated protein input is derived from the original publication (cited) but may differ in other publications utilising the same method. Moreover, injection volumes are not indicated in most studies, complicating direct comparison of the required protein input between the studies.

Name	Year of publication	Labelling/ Enrichment tag	Free Cysteine blockage	Site resolution?	Protein Input required	Quantification
Cysteine-centric						
ABE ¹⁹¹	2006	Biotin-HPDP	NEM	No	?	LFQ
PalnPISC ¹⁹²	2010	Biotin-HPDP	TCEP-NEM	Yes	1 mg	LFQ
Acyl-RAC ⁸⁹	2011	Thiopropyl Sepharose	MMTS	Yes	10-20 mg	iTRAQ
Dithiodipyridine-Functionalised Magnetic Nanoparticles ¹⁹³	2018	Fe ₃ O ₄ /SiO ₂ -SSPy microspheres	TCEP-NEM	No	?	LFQ
ssABE ¹⁸²	2017	Biotin-HPDP	TCEP-IAA	Yes	2 mg	LFQ/SILAC
ssABE (Sardana <i>et al.</i> ¹⁷⁵)	2023	Biotin-HPDP	TCEP and twice NEM/CAA for site ID	Yes	1.1 mg	LFQ
Low-Background ABE (lbABE) ¹⁸¹	2019	Biotin-HPDP	TCEP-NEM-DTDP	No	1 mg	LFQ
iodoTMT ¹⁸⁶	2024	iodoTMT0	TCEP-IAA	Yes	0.5 mg	TMT
Acyl-Trap ¹⁹⁴	2024	Pyridyl-disulfide quartz	NEM	Yes	>20 µg	TMT
Intact Peptide Methods						
nanographite fluoride-based solid-phase extraction (nGF-SPE) ⁹⁶	2023	nGF	TCEP-NEM	Yes	7.5 mg	LFQ
Lipid-centric						
MLCC ¹⁵³	2009	Biotin	IAA	No	1-1.5 mg	SILAC
ssMLCC ⁹⁰	2018	Biotin	NEM/IAA for site ID	Yes	2 mg	LFQ
Bump hole MLCC ⁹⁴	2024	Biotin	NEM/CAA for siteID	Generally no, but can be performed	2 mg	LFQ

Biotin enrichment is employed not only in ABE, but also in all metabolic labelling click chemistry (MLCC) approaches, making it the most common affinity-handle in

S-acylproteomics^{87,178}. In addition to the suboptimal elution conditions, biotin-streptavidin enrichments suffer from high background, raising two major concerns. Firstly, high-abundance non-S-acylated proteins, including endogenously biotinylated proteins, may obscure the signal from low-abundance S-acylated proteins, thereby limiting their identification by LC-MS/MS. Secondly, co-enrichment of non-S-acylated and S-acylated forms of the same protein can distort the quantitative ratios between biological conditions - a phenomenon known as “ratio compression”¹⁸¹. Another limitation is that biotin labelling alone does not provide site resolution and requires further workflow modifications to achieve it^{175,182,195}.

To achieve the aims set out for this method, alternative enrichment strategies were sought out. Recently, Phospho-based Adaptable Tags (PATs) emerged as alternative enrichment handles with superior sensitivity and specificity to the classical biotin-streptavidin approaches¹⁹⁶. Of particular interest was the Cysteine-specific Phosphonate Adaptable Tag (CysPAT), consisting of a sulfhydryl-reactive iodoacetamide and a phosphonic acid group, which enables enrichment of Cys-containing peptides by immobilised metal affinity chromatography (IMAC) with site specific resolution and under mild elution conditions¹⁵⁶. A modified version of CysPAT, 6C-CysPAT, featuring a 6C linker compared to the original 2C linker, demonstrated deep cysteinome profiling and was made commercially available¹⁵⁵. Given these advantages, it was hypothesised that labelling of Cys with the commercially available CysPAT following HA cleavage, may significantly improve the recovery and site-specific analysis of S-acylated peptides. Consequently, the adaptation of CysPAT for S-acylation Profilng by PHosphonate taggIng pRotEomics (**SAPPHERE**) was investigated.

3.2.2 Pilot Experiment

To assess the feasibility of the SAPPHERE concept, a simplistic proteomics workflow was devised, whereby the key steps, such as lysis, blocking of free Cys and aminolysis, were based on the most common conditions in the literature (**Table 3.1, Fig. 3.1**). Therefore, the first iteration of SAPPHERE (SAPPHEREv1) involved generating a HEK293T cell lysate in a 4% SDS-based buffer (see **Methods 6.4**), which was split into four technical replicates for the subsequent reduction of disulfide bridges and alkylation of any free Cys with TCEP and NEM respectively, followed by MeCN-based protein precipitation to remove any traces of NEM.

Each technical replicate was then split into two aliquots: a No HA control (-HA) – in which only 100 μ M 6C-CysPAT was added and an HA treatment condition (+HA) – in which both HA (final concentration 0.75M at neutral pH) and 100 μ M 6C-CysPAT were added to facilitate acyl-phosphonate exchange.

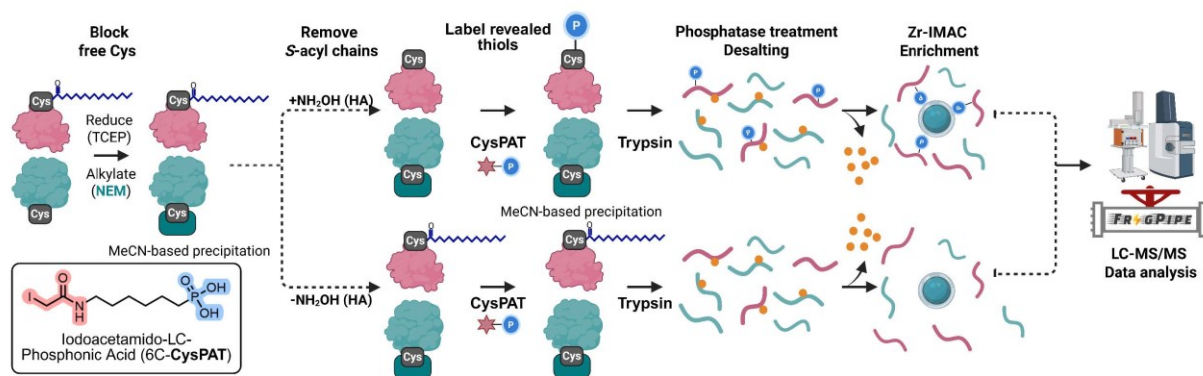


Figure 3.1. Schematic overview of the SAPPHIREv1 proteomics workflow. Protein-containing lysates are initially treated with TCEP and N-ethylmaleimide (NEM) to reduce disulfide bridges and cap any free cysteines. Excess NEM is then removed by protein precipitation and each sample is then split into two aliquots. One aliquot is treated with hydroxylamine (NH_2OH) to cleave S-acyl chains, and the newly exposed free thiols are subsequently labelled with 100 μ M CysPAT. The second aliquots is the negative control, whereby the NH_2OH is omitted, precluding labelling of S-acylated Cys with CysPAT. All samples are then subjected to a second protein precipitation and treated with trypsin. The resulting peptides are dephosphorylated using a phosphatase and desalted prior to Zr-IMAC enrichment. Figure created with Biorender.com.

Following labelling, all samples were subjected to MeCN-based precipitation, tryptic digested and desalted using the Pierce™ peptide desalting spin columns (ThermoFisher). To ensure enrichment of the phosphonated peptides over the endogenous phosphopeptides, samples were treated with QuickCIP alkaline phosphatase, which selectively removes native phosphate groups while the phosphonate label remains intact due to the enzyme-resistant P–C bond¹⁹⁶. Following dephosphorylation, peptides were desalted, quantified, and 500 μ g per sample was manually enriched using Zr-IMAC beads. Using data-dependent acquisition (DDA), enriched peptides were analysed on an Evosep One coupled to a Bruker trapped ion mobility spectrometry time-of-flight mass spectrometer (timsTOF HT) LC-MS/MS set up. The identification and quantification of peptides was performed using MSFragger search engine, with NEM-, hydrolysed NEM(NEMhyd)-, and phosphonate-modified Cys (pCys), as well as

Serine/Threonine/Tyrosine (Ser/Thr/Tyr) phosphorylation set as variable modifications (**Fig. 3.1**)¹⁹⁷.

In the pilot SAPPHIREv1 experiment, 24,846 peptides were identified. Non-modified peptides constituted the majority of the identified peptides, followed by sequences containing at least one pCys (12%), NEM- and hydrolysed NEM-modified Cys (NEM and NEMhyd) and a phosphorylated Ser/Thr/Tyr (**Fig. 3.2a,b**). Such high prevalence of non-modified background indicated low selectivity of the Zr-IMAC enrichment. Nevertheless, 420 pCys peptides exhibited significant enrichment in the +HA condition ($-\log P > 1.5$ and $\log_2 FC > 1$) corresponding to 373 unique Cys sites, of which 55 had been previously annotated as S-acylated in the SwissPalm database (**Fig. 3.3**). However, 791 (734 unique) pCys peptides were preferentially enriched ($-\log P > 1.5$ and $\log_2 FC < -1$) in the -HA condition, suggesting non-specific Cys labelling by CysPAT even in the absence of HA, likely due to incomplete blocking of free Cys. Nevertheless, a higher ratio of pCys and lower ratio of non-modified peptides observed in the +HA compared to the -HA condition, indicated that phosphonate labelling was at least partially HA-dependent (**Fig 3.2a, Fig. 3.3**).

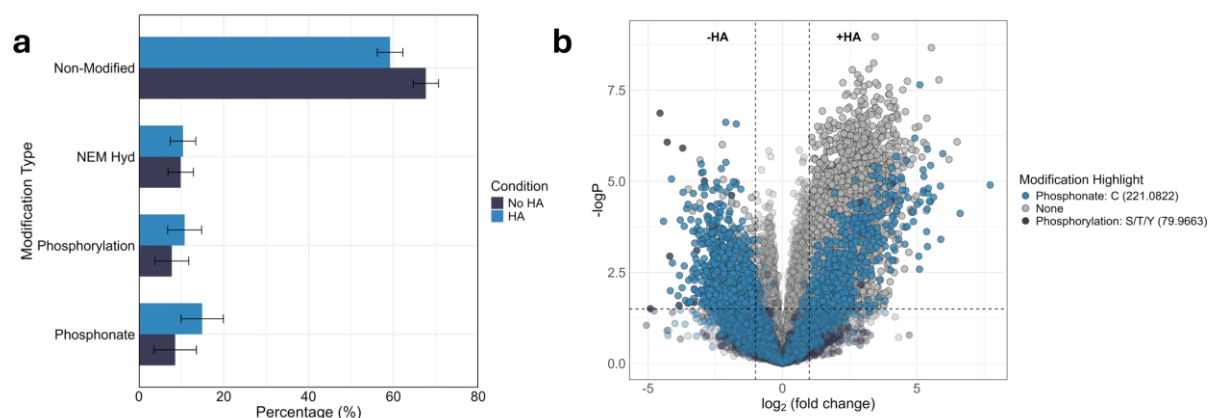


Figure 3.2. Pilot experiment results for SAPPHIREv1. a) Percentage breakdown of the top four modifications on peptides identified using SAPPHIREv1, indicating that the majority were unmodified. The error bars represent standard deviation. **b)** Volcano plot showing differential enrichment of all peptides identified in HEK293T cells between +HA and -HA conditions (Student's t-test; FDR = 0.05, S_0 = 0.1). Light blue points represent sites containing a phosphonated Cys modification (n=4 technical).

Despite these inefficiencies, the pilot experiment demonstrated the proof-of-principle for the use of phosphonate handles for targeted profiling of S-acylation sites. However, to improve specificity, minimise false positives, and enhance throughput, subsequent optimisation of IMAC enrichment efficiency, alkylation conditions, and overall workflow was necessary.

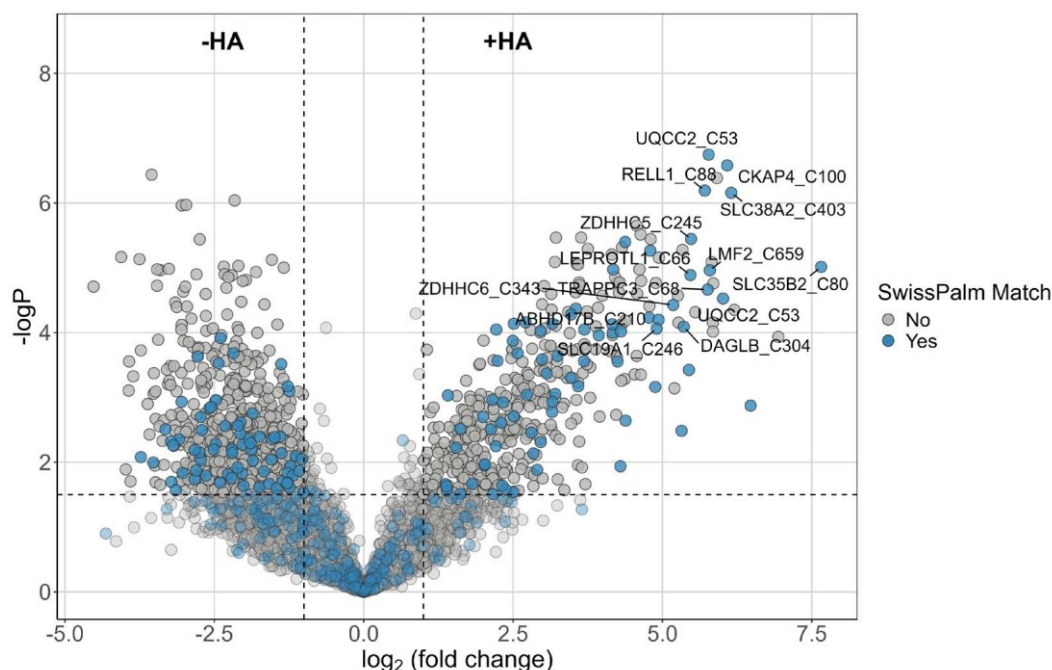


Figure 3.3. High rate of false-positive identifications in the SAPPHIREv1 dataset. Volcano plot illustrating the differential enrichment of phosphonated cysteine-containing peptides (representing putative S-acylation sites) between +HA and –HA conditions, based on a Student’s t-test (FDR = 0.05, S0 = 0.1). Sites previously reported in the SwissPalm database are highlighted in blue (n=4 technical).

3.3 Optimisation of the SAPPHIRE method

3.3.1 IMAC Enrichment

IMAC remains one of the most widely used techniques for phosphopeptide enrichment in phosphoproteomics and is based on the strong affinity of the metal cations such as Fe^{3+} , Ga^{3+} , Al^{3+} , Ti^{4+} and Zr^{4+} for the negatively charged phosphate groups on peptides¹⁹⁸. Recent reports have highlighted Zr-IMAC magnetic microparticles as an efficient and scalable phosphoenrichment strategy, demonstrating superior performance over the commonly

employed Ti-IMAC, TiO_2 and LC-based Fe(III)-IMAC^{199,200}. Moreover, automation of the Zr-IMAC workflow has been optimised on the KingFisher Duo and Flex systems, rendering it highly suitable for high-throughput proteomics^{199,200}. In developing SAPPHERE, the enhancement of scalability and specificity was prioritised, leading to the selection of Zr-IMAC for phosphonate peptide enrichment.

In the pilot experiment, manual enrichment with Zr-IMAC High performance (HP) beads from Resyn Biosciences, at a bead volume-to-peptide mass ratio of 1 μL :50 μg , exhibited low selectivity (11%) with substantial background noise (**Fig. 3.2a,b**). To enhance specificity and reproducibility, the enrichment was transitioned to an automated Kingfisher Flex (Thermo Instruments) system at a 96-well plate scale, as per Zr-IMAC manufacturer's instructions with the addition of an extra elution step (**Fig. 3.4a**). Using the automated setup, three bead-to-

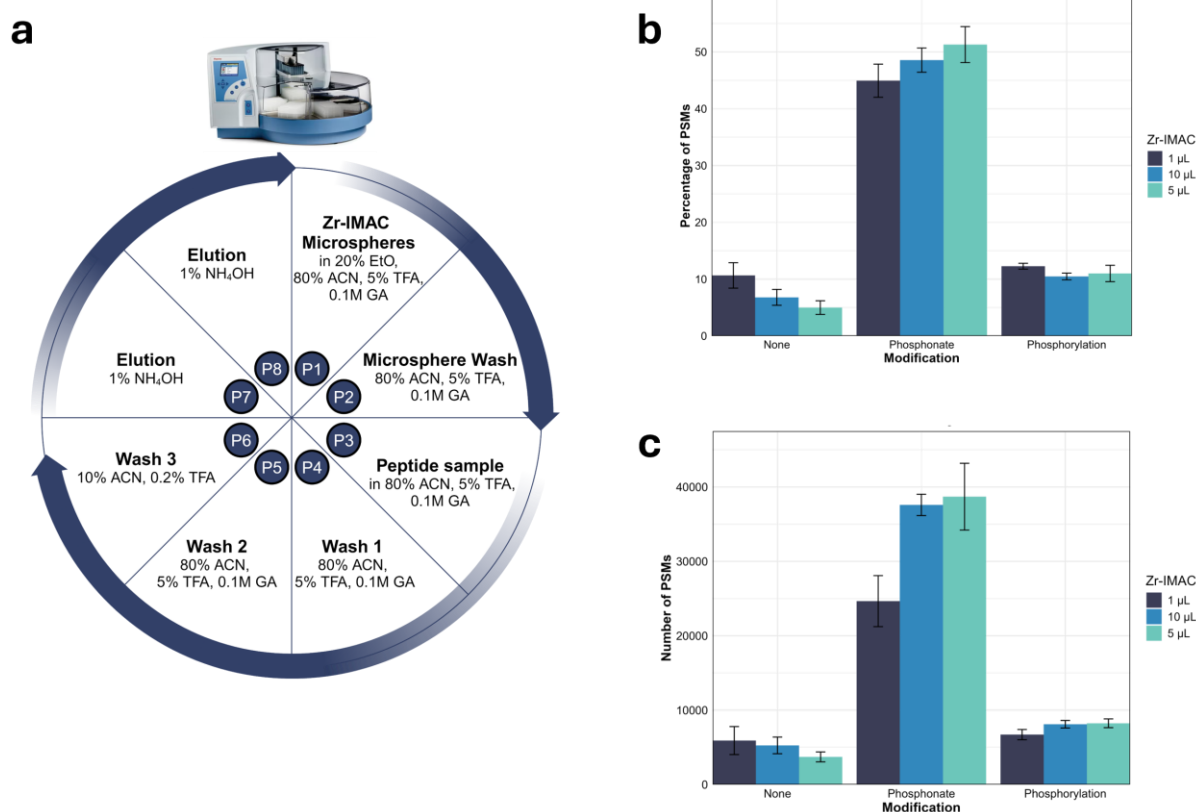


Figure 3.4. Optimisation and automation of Zr-IMAC enrichment. **a)** Schematic of the Zr-IMAC enrichment protocol automated on the KingFisher Flex platform. The eight plates and arrows represent sequential processing steps. **b)** Percentage of peptide-spectrum matches (PSMs) with different modifications. **c)** Total number of PSMs corresponding to each modification type. The error bars represent standard deviation.

peptide ratios (1 μ l, 1:200; 5 μ l, 1:40; and 10 μ l, 1:20) were systematically tested using 200 μ g digest generated by SAPPHIREv1 in three technical replicates. Automation significantly improved enrichment efficiency, consistently achieving >45% pCys recovery across all tested ratios (**Fig. 3.4b**). The optimal ratio was determined at 1:40, achieving the highest pCys recovery of ~52% and the lowest percentage of non-modified peptides (4.99%), without compromising the total number of identified peptide spectrum-matches (PSMs) (**Fig. 3.4c**). However, during this experiment, pipetting 1 μ l bead aliquots proved challenging due to the poor miscibility of the microparticles in the storage buffer, often requiring multiple attempts to accurately transfer the volume. To minimise pipetting errors, a minimum bead volume of 3 μ l was established for all subsequent experiments, even when the peptide input was below the optimal bead-to-peptide ratio.

Despite these improvements, the observed selectivity (~52%) remained markedly lower than the selectivity reported in the literature (97%)¹⁹⁹. Further optimisation experiments, such as testing different concentrations of glycolic acid (GA) in the loading buffer or sodium hydroxide in the elution buffer, and varying binding times, could have been pursued²⁰⁰. However, it is important to note that Zr-IMAC was originally designed and extensively optimised for deep profiling of endogenous phosphorylation, and its adaptation for less abundant PTMs, such as S-acylation has not been previously explored. Given that phosphorylation potentially affects 75% of the human proteome, whereas S-acylation, even at high estimates, occurs on 20%, it is possible that volumes below 1 μ l are sufficient for 100% pCys recovery^{63,201}. Consequently, the high background may result from non-specific binding of non-modified peptides to excess beads. As such, no further optimisations were undertaken and the automated workflow illustrated in Fig. 3.4a was used in all subsequent SAPPHIRE experiments. Importantly, at the time of this optimisation, blocking of free Cys was not optimised, which may have resulted in a higher proportion of phosphonate peptides. It was therefore anticipated that the enrichment efficiency might fall below 45% in later experiments where phosphonate labelling was highly specific for S-acylated Cys.

3.3.2 Free cysteine alkylation

Even minor incomplete blocking could lead to erroneous labelling of non-S-acylated Cys by CysPAT, resulting not only in false positive identifications, but also potential masking of true

lower-abundance S-acylated Cys, as well as ratio compression. Given the observed high sensitivity of SAPPHERE, optimisation of free and reduced disulfide Cys was undertaken (**Fig. 3.5**).

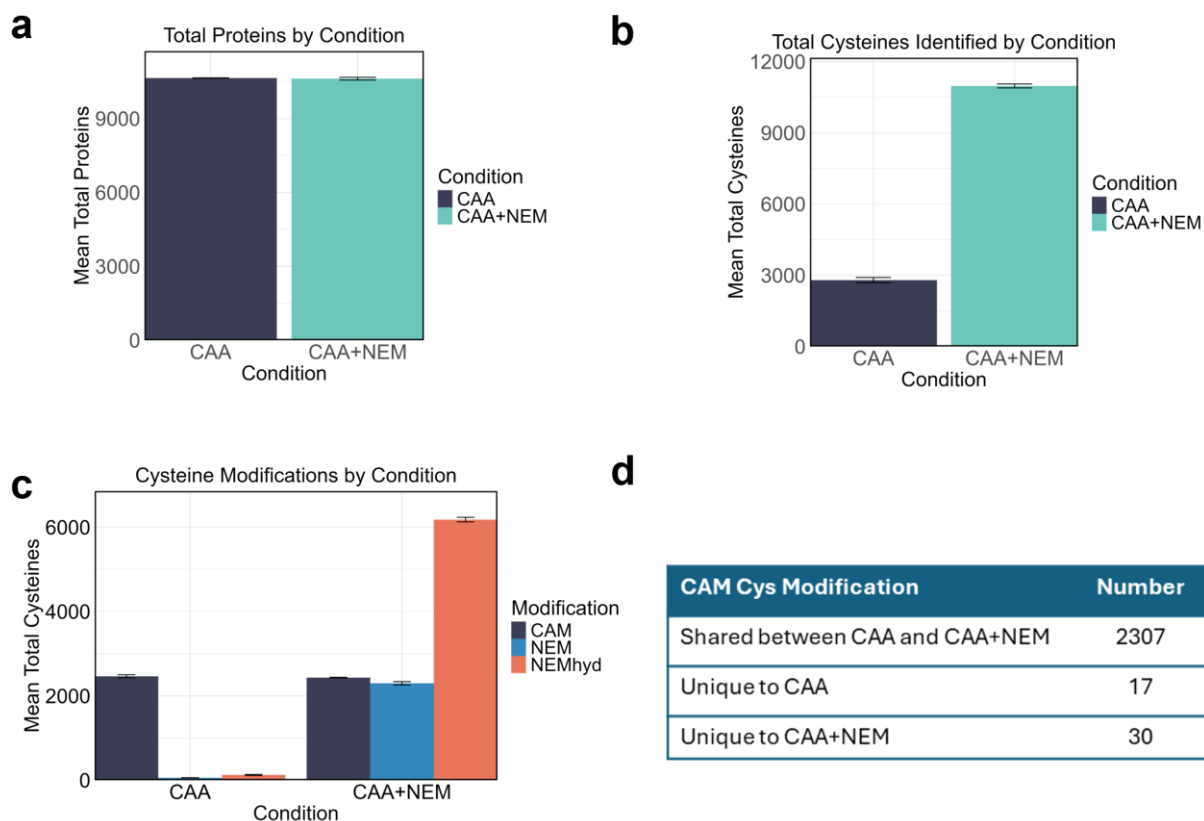


Figure 3.5. Optimisation of free cysteine alkylation conditions. Whole-proteome analysis of HEK293T cells treated either with 25 mM chloroacetamide (CAA) alone, or with 25 mM CAA followed by acetonitrile-based protein precipitation and subsequent blocking with TCEP and 25 mM N-ethylmaleimide (NEM). **a**) Total number of proteins identified under each condition. The error bars represent standard deviation. **b**) Total number of cysteines identified in each condition. **c**) Number of identified cysteines and their associated modifications. **d**) Number of carbamidomethylated cysteines (CAM) that are unique or shared between the two conditions.

In the pilot experiment, which included a single TCEP + NEM reduction/alkylation, many pCys were identified in the -HA condition, indicating low blocking efficiency. Although NEM remains the most widely employed alkylating agent in cysteine-centric methods, other reagents such as iodoacetamide (IAA) and methylmethanethiosulfonate (MMTS) – have been employed (**Table 3.1**). Moreover, enhanced strategies including combining NEM and 2,2’-

dithiodipyridine (DTDP), double IAA blocking, triple NEM blocking and double NEM have been proposed and demonstrated enhanced blocking efficiency^{175,181,182,193}. Despite these improvements, complete blocking has never been achieved.

To optimise the blocking efficiency of the SAPPHERE workflow, two blocking strategies were assessed using a HEK293T cell lysate. The first involved a single reduction and alkylation step using TCEP and chloroacetamide (CAA), whereas the second involved sequential double alkylation with the initial reduction and alkylation by TCEP and CAA, followed by TCEP + NEM treatment, separated by an MeCN-based precipitation (CAA+NEM condition). CAA has different chemical reactivity to NEM and was selected due to its wide use in proteomics and our group. Both conditions were performed in three technical replicates, subjected to another MeCN-precipitation, trypsin digestion, analysis using data-independent acquisition (DIA) and quantification by the DIA-NN software, with Cys modifications specified as NEM, NEMhyd and carbamidomethylation (CAM)²⁰².

Whilst the number of identified proteins remained the same, the number of total Cys was significantly higher for the CAA+NEM condition (**Fig. 3.5a,b**). As expected, only CAM Cys were identified in the CAA condition, whereas CAM, NEM and NEMhyd modified Cys were identified in the CAA+NEM condition (**Fig. 3.5c**). Surprisingly, most Cys were modified by NEMhyd, rather than intact NEM, which likely occurred due to the hydrolysis of the maleimide ring in a slightly basic tryptic digestion buffer (25 mM ammonium bicarbonate pH 8.0)^{203,204}. The number of CAM modified Cys was similar in both conditions, as they represented the same peptides (**Fig. 3.5c,d**). As CAA alkylation was performed prior to MeCN-based precipitation, it was hypothesised that the increase in Cys coverage in CAA+NEM condition was a result of MeCN precipitation-induced conformational changes that permitted accessibility to previously inaccessible Cys to the alkylating reagent, rather than the identity of the alkylating agent.

3.3.3 Increasing the throughput and ease of the workflow

With the establishment of a 96-well plate automated enrichment, further optimisations to scale up the remaining steps of the workflow were undertaken. It was found that following lysate preparation, the reduction, alkylation, CysPAT labelling, and protein precipitation steps could be performed in 2 mL Protein LoBind 96-well plates, with minimal adjustments to the existing protocol, unless more than 400 µl were to be used per sample. Otherwise, the

volumes and concentrations remained unchanged with only minor practical adjustments, such as increasing the centrifugation speed following MeCN addition from 5 to 10 mins and longer protein pellet resuspension times with gentle rotation (see **Methods 6.1.3**). Subsequent digestion and dephosphorylation could be performed in the same plate.

Next, to facilitate a 96-well scale of the peptide desalting step, alternative strategies were explored. At the time, a 96-well desalting protocol using OASIS®-HLB sorbent (Waters) and 96-well Durapore filter plate was set up in our lab by Matthew White. This protocol was integrated into the SAPPHIRE workflow with a recommended ratio of 100:1 Oasis HLB:peptide (100 µg Oasis per 1 µg peptide). Utilising the bulk material rather than pre-packed Oasis HLB µelution plates, enabled flexible adjustment to each experiment's scale, particularly beneficial for low input samples, while significantly reducing the overall costs of the workflow. Notably, SAPPHIREv1 included two desalting steps, one post-digestion and one post-dephosphorylation, based on published protocols utilising PATs and IMAC for non-phosphoproteomics⁴⁹⁻⁵¹. To test whether the workflow could be simplified, the pre-dephosphorylation desalting was omitted and a single desalting prior to IMAC enrichment was performed by directly adding the phosphatase and rCutSmart buffer to the digests. Pleasingly, a single desalting step was found sufficient for efficient dephosphorylation (below 15%), thereby streamlining the SAPPHIRE workflow **Fig. 3.4b**)¹⁹⁶.

3.4 Profiling S-acylation sites in HEK293T cells with SAPPHIRE

Cumulatively, the optimisations described above enabled informed design of the final version of the SAPPHIRE method (**Fig. 3.6**). After lysate generation, the workflow can be performed in a 96-well plate format and involves a double reduction/alkylation step, MeCN-based precipitation, HA treatment and CysPAT labelling at an increased concentration of 250 µM, MeCN-based precipitation, digestion, Oasis-HLB desalting, automated Zr-IMAC enrichment and LC-MS/MS analysis. Depending on the experiment, CAA and NEM were used interchangeably for Cys blocking.

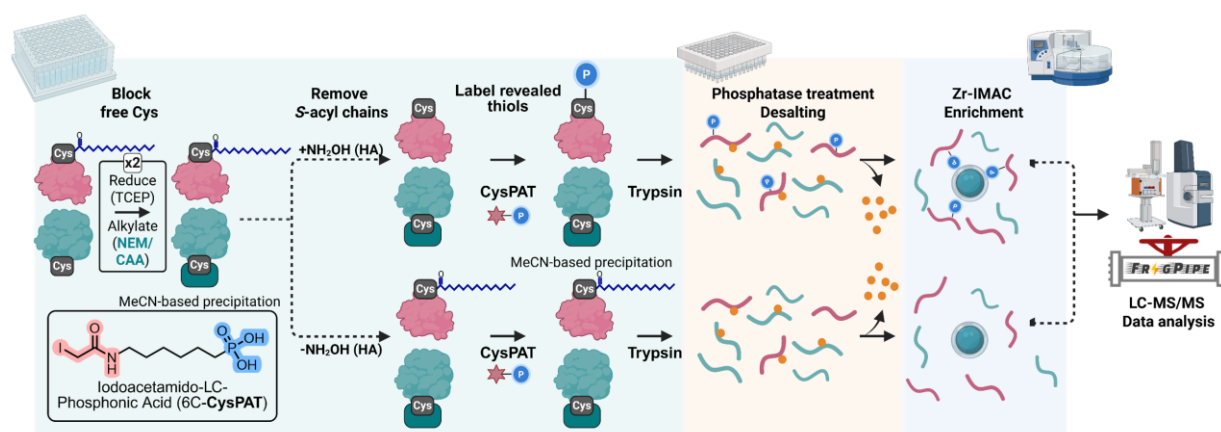


Figure 3.6. Schematic of the optimised SAPPHIRE S-acylproteomic workflow. Protein-containing lysates are initially blocked twice with N-ethylmaleimide (NEM) or chloroacetamide (CAA), followed by protein precipitation to remove residual NEM/CAA, which can be performed in 96-well plate format. Samples are then treated with hydroxylamine (NH₂OH) to cleave acyl chains, and the exposed free thiols are subsequently labelled with 250 μ M CysPAT. After a second precipitation and tryptic digestion, peptides are dephosphorylated using a phosphatase, then desalted using a 96-well plate desalting protocol prior to Zr-IMAC enrichment using the KingFisher Flex. Figure created with Biorender.com.

3.4.1 Results

The optimised SAPPHIRE workflow was then applied to HEK293T cells, identifying a total of 19,699 peptides (**Fig. 3.7a**). In comparison to the pilot experiment, the optimised workflow led to a marked reduction in background, with pCys enrichment efficiency increasing to ~30% in the +HA condition (**Fig. 3.7b**). Likewise, blocking efficiency was extremely high, as no pCys peptides were significantly enriched in the -HA control. Of the 5569 pCys peptides identified, 3085 unique S-acyl sites showed significant enrichment ($\log_2\text{FC} > 1$ and $-\log P > 1.5$) in the +HA condition (**Fig. 3.8a,b**). Among these, 185 (~8%) have been previously annotated on SwissPalm, including well-validated sites such as HRAS Cys181 and Cys184, ABHD17B Cys210, cytoskeleton-associated protein 4 (CKAP4) C100 and Flotillin-1 (FLOT1) C34 (**Fig. 3.8a**)^{177,205–208}. Additionally, many non-catalytic Cys of the zDHH family (zDHH17 Cys590, zDHH23 Cys66, zDHH24 Cys242, zDHH4 Cys298/C301, zDHH5 Cys245, zDHH6 Cys329/C343 and zDHH20 Cys263) were identified, some of which were previously validated (**Table 2.2**). Interestingly, of the 2865 sites lacking direct SwissPalm site annotation, 2262 had a SwissPalm annotation of their corresponding protein, potentially providing novel site characterisations of these proteins (**Fig. 3.8b**). Furthermore, among the 386 sites (~16%)

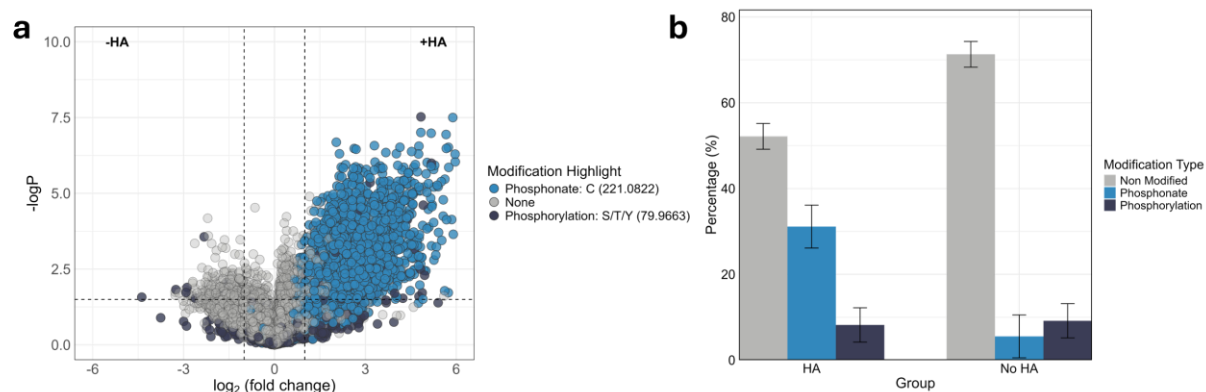


Figure 3.7. Site-specific S-acylated proteome profiling in HEK293T cells using the optimised SAPPHIRE protocol. a) Volcano plot showing differential enrichment of all peptides identified in HEK293T cells between +HA and –HA conditions (Student’s t-test; FDR = 0.05, S0 = 0.1). Light blue points represent sites containing a phosphonated Cys modification. **b)** Percentage breakdown of the unmodified, phosphonate and phospho-modified peptides identified using SAPPHIRE, revealing a 6-fold higher CysPAT labelling efficiency in +HA compared to -HA condition. The error bars represent standard deviation.

lacking both site and protein annotations, some have a strong literature validation record, such as p53 Cys135 and Cys182, which are important residues of one of the most extensively studied tumour suppressors^{81,209,210}. Of note, 1923 putative S-acylated Cys are associated with 597 SG proteins.

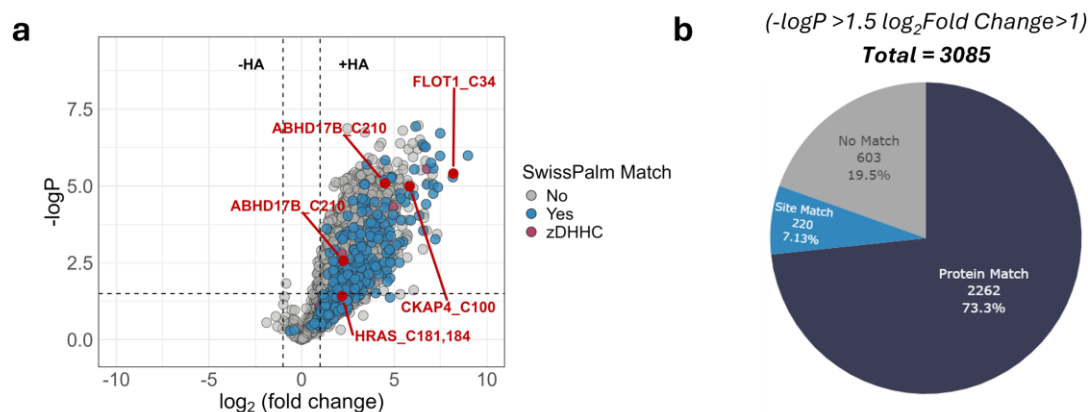


Figure 3.8. Site-specific S-acylated proteome profiling of HEK293T cells using SAPPHIRE. a) Volcano plot showing differential enrichment of phosphonated cysteine-containing peptides (i.e. putative S-acylation sites) between +HA and –HA samples (Student’s t-test; FDR = 0.05, S0 = 0.1) n=4 technical. **b)** Pie chart comparing S-acylated sites identified by SAPPHIRE ($-\log P > 1.5 \log_2 \text{Fold Change} > 1$) with those previously reported in the SwissPalm database.

Overall, SAPPHIRE enabled deep profiling of 3085 S-acylated sites, of which 2023 are novel, in HEK293 cells, constituting the largest single-experiment dataset to date. The high overlap with the existing database, combined with high specificity of the workflow, bring confidence to the true validity of these sites. However, further annotations and validation experiments were necessary.

Table 3.2. Summary of S-acylated sites of zDHHC enzymes identified in HEK293T cells with SAPPHIRE

zDHHC	Identified Cys	Active Site	Validation Techniques
zDHHC4	298, 301	179	Not found
zDHHC5	245	134	Validated with C236, C237. Identified in PalmPISC, ssMLCC and nGF
zDHHC6	329, 343	129	Site-directed mutagenesis and APE, radiolabelling
zDHHC17	590	467	Not found
zDHHC20	263	156	Identified by nGF, ssABE (Sardana <i>et al</i>) and bump-hole MLCC + NH ₂ OH CAA (Ocasio <i>et al</i>)
zDHHC23	66	289	Not found
zDHHC24	242	124	Identified by nGF

3.4.2 Functional and structural analysis of S-acylated sites identified by SAPPHIRE

The 3085 putative S-acyl sites identified by SAPPHIRE mapped to 1926 proteins. Functional enrichment analysis was then performed with Metascape, which integrates core ontologies, including Gene Ontology (GO) processes, KEGG pathways and Reactome gene sets, amongst others (**Fig. 3.9**)²¹¹. A strong enrichment for nuclear processes was observed, including RNA and DNA metabolism, RNA localisation and chromosome organisation (**Fig. 3.9**). However, given that this was global analysis of S-acylation and the broad range of processes regulated by S-acylation, it was challenging to draw definitive conclusions from this enrichment analysis.

To further assess the functional annotations on the site-specific level, rather than protein, the UniProt database was used to gain insight into functional and PTM annotations of the Cys residues (see **Methods 6.9.1**). Notably, a majority (2746 out of 3085 sites; 89%) lacked specific functional annotations. Among the annotated sites, 214 were palmitoylated, 73 were

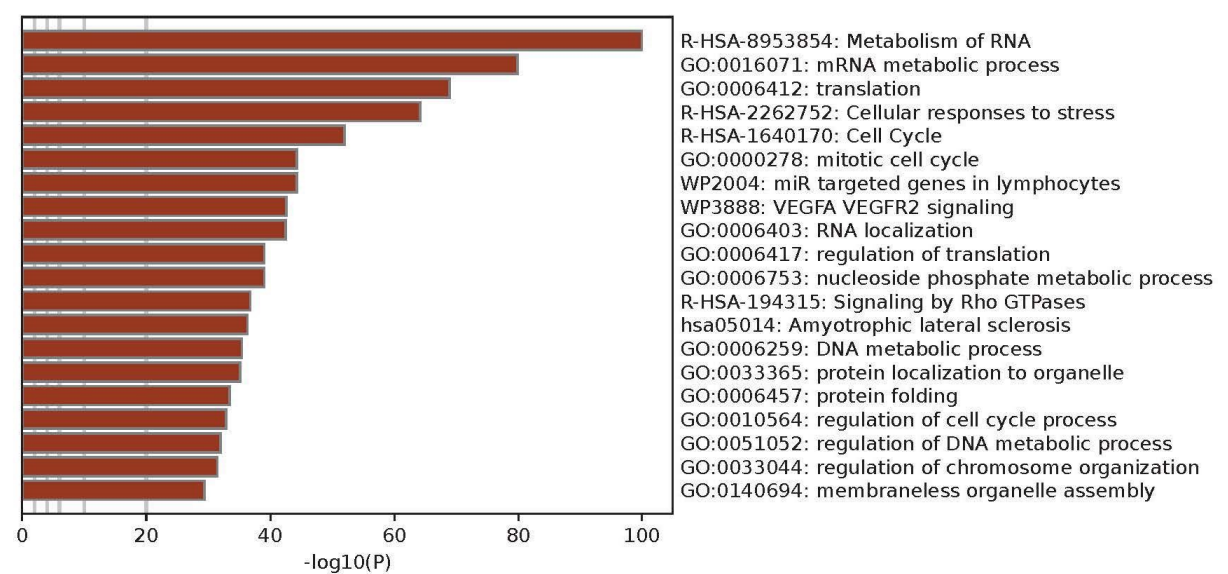


Figure 3.9. Functional enrichment analysis of putative S-acylated proteins identified by SAPPHIRE in HEK293T cells. Top enriched pathways among proteins identified as putative S-acylated using SAPPHIRE, as determined by functional enrichment analysis in Metascape (metascape.org).

associated with disulfide bond formation, and 52 sites had other types of functional annotations (**Fig. 3.10**). Importantly, while the lack of annotation strengthens confidence of these Cys being true S-acylation sites, additional annotations for some sites do not exclude the possibility of concurrent S-acylation. Further cross-referencing of the SAPPHIRE data against the residue-specific clinical annotations from ClinVar, revealed that although most Cys were not annotated, 14 sites corresponded to pathogenic variants (**Appendix Fig. 7.1, Appendix Table 7.2**)²¹². Collectively, these findings demonstrate the potential of the SAPPHIRE platform to provide new functional annotations and identify clinically relevant sites.

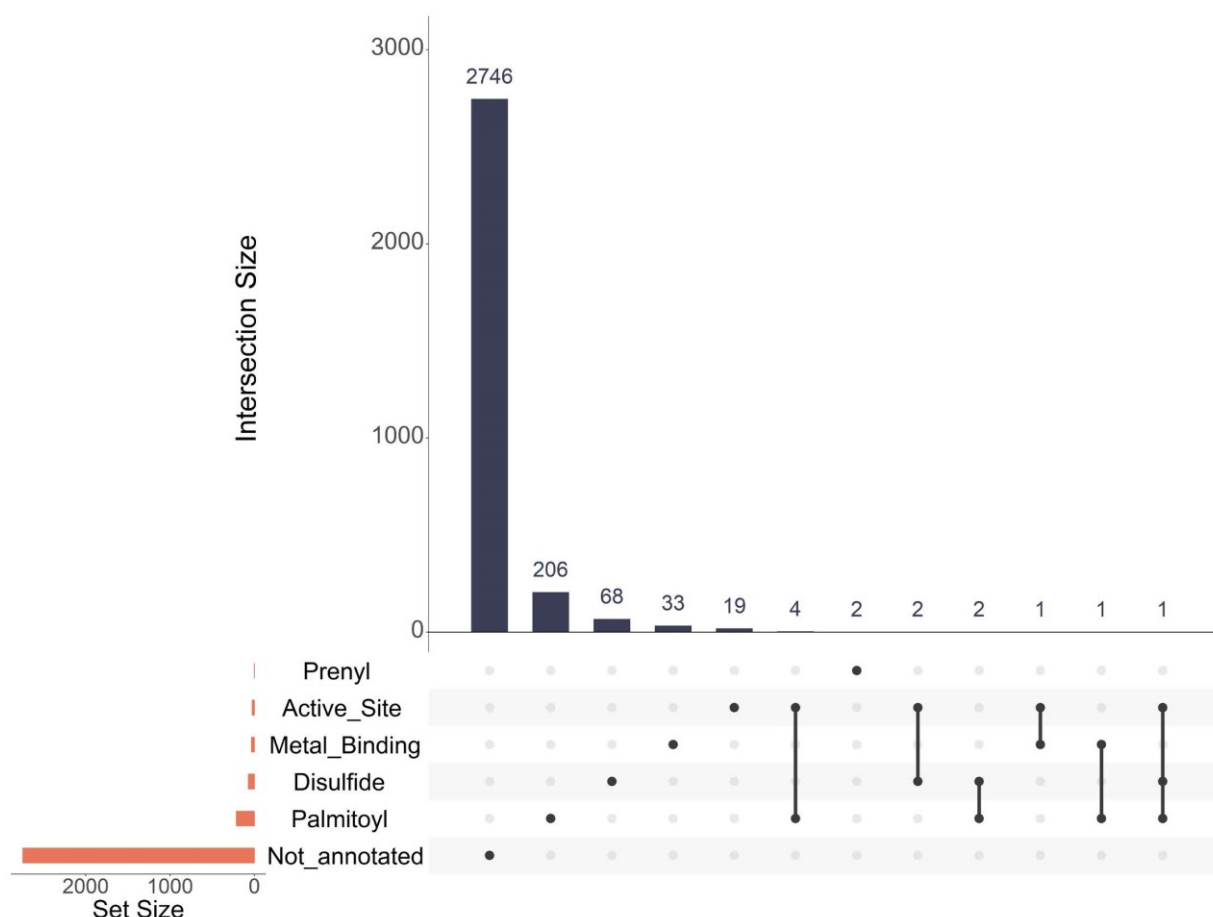


Figure 3.10. Post-translational modification (PTM) and functional annotation of putative S-acylation sites in HEK293T cells identified using SAPPHIRE. UpSet plot showing the overlap between 3,095 putative S-acylation sites and UniProt annotations for cysteines involved in functional roles or other PTMs.

3.4.3 S-acylation consensus sequence and cysteine accessibility analysis

Unlike other classes of protein lipidation, and indeed most PTMs, S-acylation lacks a defined consensus motif²¹³. Instead, the current paradigm emphasises Cys accessibility rather than a specific sequence, as the primary determinant of an acylation event²¹⁴. Recent advancements in sensitive methods have expanded the library of S-acylated sites, prompting renewed efforts in discovering potential sequence motifs. In 2017, ssABE analysis of mouse forebrain tissue (906 sites) reported an enrichment of a dicysteine motif in transmembrane proteins (CC* and C*C, where *denotes palmitoylation site) alongside motifs containing single Cys and hydrophobic amino acids within 4 residues for soluble proteins (C*xxL and C*F)¹⁸². In 2024,

S-acyl site profiling using the iodoTMT method of mouse brain, liver, and human kidney cell lines HK2 and HK11 cells (646, 767, 261 and 701 sites respectively) revealed only two statistically significant motifs of C*P and C*xxxxR from mouse liver from HK11 cells respectively¹⁸⁶. Notably, their mouse brain analysis did not corroborate neither the diCysteine nor the soluble protein motif previously reported by ssABE. Furthermore, in 2024 ssMLCC analysis of HeLa cells (606 sites) failed to identify a consensus motif but observed an overrepresentation of proline, glycine, lysine, and arginine residues around the S-acylated Cys²¹⁵. All three studies utilised the motif-x algorithm for their investigations.

Given the extensive coverage of 3085 putative S-acyl sites identified by SAPPHIRE – the largest single-experiment dataset to date - a conserved sequence motif was investigated. Peptide sequences of 13 residues, with the modified Cys at the centre, were extracted and analysed with classic and enhanced motif-x algorithms in MoMo software²¹⁶. Classic motif-x algorithm is run against background peptides derived from the entire human proteome, without adjusting for the difference in residue composition between the input (foreground) and background (BG) peptide sets, therefore discovering motifs with enriched residues rather than actual conserved motifs. Conversely, the enhanced motif-x algorithm generates background peptides by shuffling the foreground (FG) peptides while conserving the central residue, thus accounting for the relative residue composition of the FG and BG peptides. Using the classic motif-x algorithm, SAPPHIRE data revealed strong enrichment of Lys residues as well as acidic amino acids (D and E) around the acylated Cys, with the top motifs including C*xxxxxK with 8.2% frequency, DxxC*, ExC*, DxxxC* (**Appendix Fig 7.2/Table 3.2**). In contrast, the enhanced motif-x algorithm identified five significant motifs CxxC* (1.7%), C*xxxxR (6.9%), C*P (7.9%),

Table 3.3. S-acylation consensus motif analysis

Motif	Score	FG match	FG size	BG match	BG size	FG/BG	Unadjusted p-value	tests	Adjusted p-value
Classic motif-x (background derived from <i>H.sapiens</i> proteome)									
xxxxxx_C_xxxxxK	17.36	246	2986	22105	481785	1.8	5.80E-18		
xxxDxx_C_xxxxxx	13.73	206	2986	18398	481785	1.8	1.90E-15		

xxxxEx_C_xxxxxx	12.2	282	2986	27970	481785	1.6	3.50E-15		
xxDxxx_C_xxxxxx	12.05	197	2986	19729	481785	1.6	1.40E-10		
xxxExx_C_xxxxxx	11.25	234	2986	26266	481785	1.4	4.10E-08		
xxxxxD_C_xxxxxx	10.1	193	2986	20238	481785	1.5	6.70E-09		
xDxxx_C_xxxxxx	9.78	188	2986	20154	481785	1.5	4.80E-08		
xKxxx_C_xxxxxx	10.1	227	2986	22062	481785	1.7	3.20E-13		
xxxxx_C_xxxEx	7.9	271	2986	28597	481785	1.5	9.10E-12		
xxxVx_C_xxxxxx	8.42	235	2986	26648	481785	1.4	8.40E-08		
Enhanced motif-x (background derived from shuffled input/foreground peptides)									
xxxCxx_C_xxxxxx	10.25	50	2986	17	2986	2.9	3.10E-05	220	6.70E-03
xxxxx_C_xxxxR	5.65	207	2986	150	2986	1.4	1.10E-03	218	2.10E-01
xxxxx_C_Pxxxxx	4.88	236	2986	178	2986	1.3	1.80E-03	217	3.30E-01
xxxxxI_C_xxxxxx	4.58	166	2986	123	2986	1.3	5.60E-03	216	7.00E-01
xxxxx_C_xxxxC	4.27	34	2986	19	2986	1.8	2.60E-02	214	1.00E+00

IC* (5.6%) and C*xxxxxC (1.1%). The enrichment of Lys residues as well as the C*xxxxxR and C*P motifs corresponded to the literature. However, all identified motifs occurred with low frequency, insufficient to define a valid consensus sequence, consistent with other studies. Moreover, despite an adjusted p-value of <0.05 , all motifs discovered with the motif-x algorithm may not be significant²¹⁶. Further work, although beyond the scope of this thesis, may provide interesting insight into the validity and significance of these motifs. It is also

important to note that trypsin selectively cleaves at C-terminal of Lys and Arg residues, which may explain the overrepresentation of these residues using the classic motif-x algorithm.

Given the absence of a defined consensus sequence, the putative S-acyl sites were then assessed for accessibility using the amino acid side-chain prediction-aware part-sphere exposure (pPSE) values - a method previously developed in our group²¹⁷. The pPSE metric is based on AlphaFold2-predicted protein structural information across almost all human proteins and quantifies residue accessibility by counting proximal α -carbons within a conical volume projecting 12 Å along the C α -C β vector (or pseudo-vector in the case of glycine), with an internal angle of 70°, whereby a high value indicates a more buried site, and a low value indicates high accessibility. SAPPHIRE dataset showed enrichment for highly exposed Cys with ~59% of identified sites having a pPSE value ≤ 5 , in line with the accessibility paradigm (**Fig. 3.11**). It is important to note that no filtering was applied for AlphaFold2 prediction quality and may have lower pPSE annotation confidence.

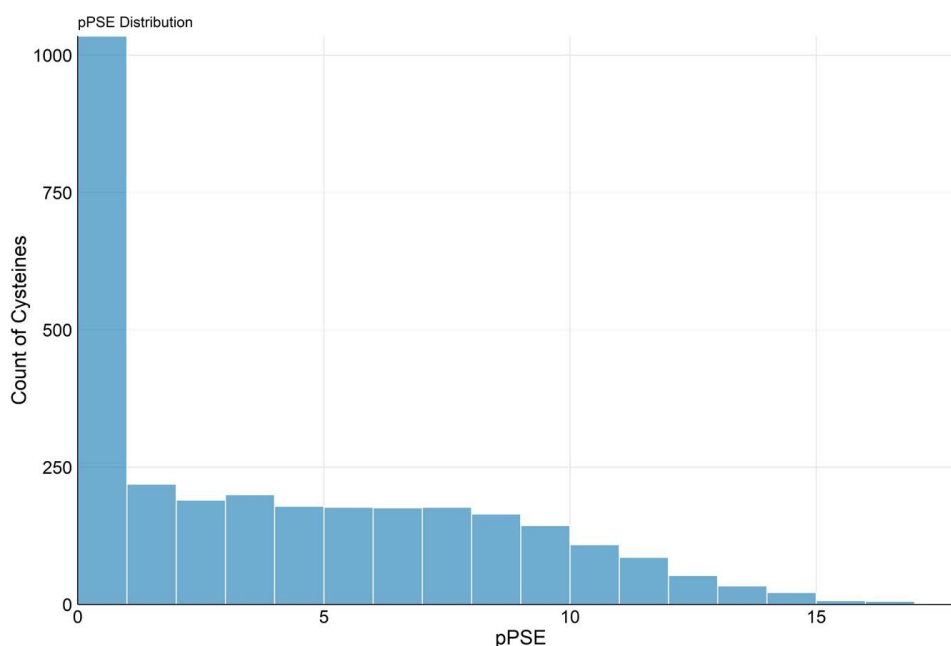


Figure 3.11. Surface exposure analysis of putative S-acylated cysteine sites identified in HEK293T cells using SAPPHIRE. Histogram showing the distribution of predicted prediction-aware part-sphere exposure (pPSE) values for identified cysteine sites, indicating that most are solvent-accessible (<5 pPSE).

3.5 Western blot validation of the identified sites

To validate sites identified by SAPPHERE, Western blot experiments using acyl-PEG exchange (APE) assay were performed. APE is a cysteine-centric, gel-based approach for validation of S-acylation and similarly involves free Cys blocking, HA aminolysis and subsequent labelling of the revealed thiols with functionalised methoxy polyethylene glycol (mPEG)-maleimide (mPEG-mal) of defined mass. As such, APE induces mobility shifts for S-acylated proteins amenable to Western blot analysis. HEK293T cells were lysed, double blocked with NEM, treated with HA and labelled with 10 kDa PEG maleimide. Novel candidate sites including PI4K2A Cys183, PCNA Cys81 and Cys135, FAM177A Cys114, G3BP1 Cys73 and RBM39 Cys303 were investigated, and well-validated S-acylated proteins such as Calnexin (CANX) and IFITM3 served as positive controls (**Fig. 3.12**). Unfortunately, APE demonstrated undetectable mass shifts for all tested proteins. Even CANX, which is dually S-acylated at Cys502 and 503 and typically observed only in acylated forms, appeared mostly unmodified¹⁵⁷. Similarly, the IFITM3 positive control, expected to exhibit three distinct mass shifts corresponding to known modification sites, displayed only two¹⁵⁷. Despite repeated attempts at the experimental protocol, with extra caution given to parameters such as the inclusion of EDTA during HA cleavage, that were previously shown to strongly affect the APE assay efficiency, the results remained unchanged.

The limited detection of mass shifts, even for positive controls, indicated low efficiency of the APE assay. One possible reason for this inefficiency is the use of the mPEG 10 kDa tag rather, which has higher steric hindrance potential than the more widely utilised 5 kDa tag. This is especially important for validation of dually S-acylated proteins where the two Cys are in close proximity, such as CANX, where this could lead to the lack of double labelling. In addition to that, these novel sites may be of low-stoichiometry, making them difficult to visualise with gel-based methods without increasing the loading protein input or specific treatments during cell culture. It is also possible that some proteins identified by SAPPHERE, are inherently resistant to detection by APE, similar to previous reports of resistance of certain proteins to identification by acyl-RAC but not ABE methods for unknown reasons²¹⁸. In line with this hypothesis, IFITM3 was absent from the SAPPHERE dataset and CANX Cys360/Cys366, but not Cys502/Cys503, were detected, although this could also have resulted from the lack of

optimisations to the SAPPHIRE workflow for maximising the recovery of highly hydrophobic transmembrane proteins. Nonetheless, it remains that the lack of validation by APE does not necessarily negate the validity of SAPPHIRE identifications. Further optimisation of APE and other gel-based approaches will be pursued, beyond the timeline of this PhD.

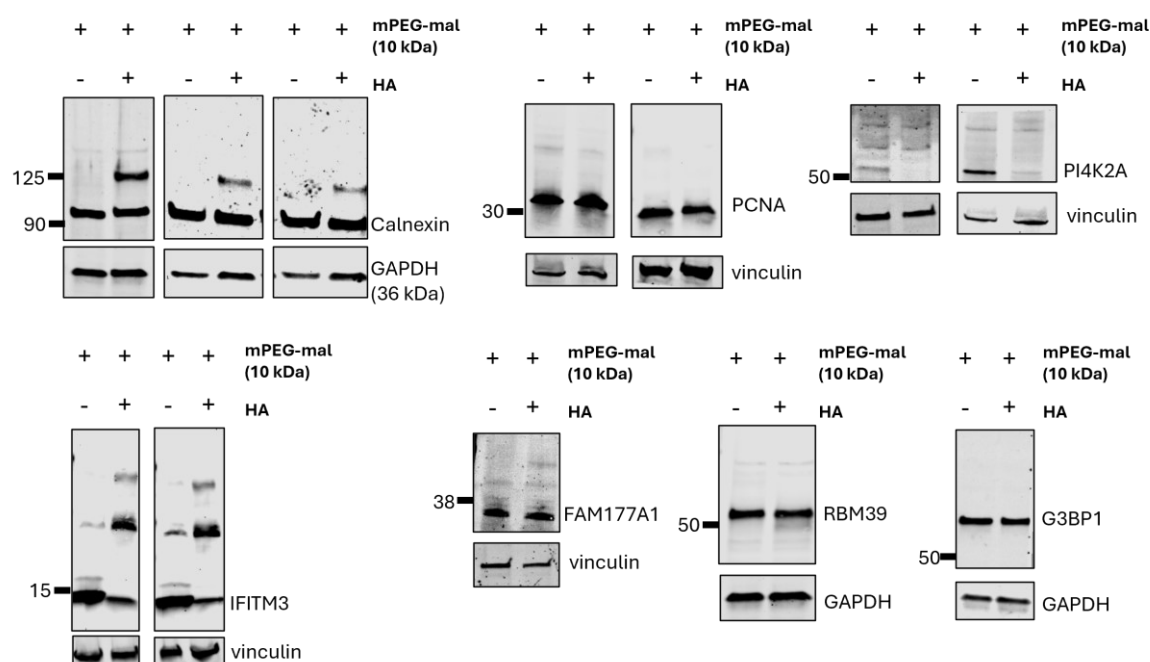


Figure 3.12. Western blot analysis of putative S-acylation sites using the acyl-mPEG exchange (APE) assay in HEK293T cells. Western blot showing validation of selected putative S-acylated proteins, the sites of which have been identified by SAPPHIRE in HEK293T cells. Positive controls include doubly S-acylated calnexin and triply S-acylated IFITM3.

3.6 Benchmarking SAPPHIRE against existing technologies

3.6.1 Site-specific ABE vs SAPPHIRE

Next, SAPPHIRE was benchmarked against ABE, the most widely employed proteomic and gel-based technique for S-acylation detection. Of the two proteome-scale site-specific ABE (ssABE) methods reported, the more recent approach was selected and applied to HEK293T cells, allowing direct comparison with SAPPHIRE sites from the same cell model¹⁷⁵. The ssABE workflow involved double capping with NEM, aminolysis and labelling of the previously

S-acylated Cys with biotin-HPDP, followed by enrichment with the Neutravidin beads. Captured S-acylated proteins were then released from the beads by TCEP reduction and the newly exposed Cys were labelled with CAA, tryptic digested and analysed on the LC-MS/MS (**Fig. 3.13**). The MSFragger software was then used to search for raw data with NEM and Carbamidomethylation (CAM) of Cys (+57.0512 Da) set as variable modifications. It is important to note that due to the column set-up at our proteomics facility available at the time, the ssABE experiment was conducted with a shorter LC gradient (60 SPD) compared to SAPPHERE profiling (30 SPD), which may have led to fewer identifications.

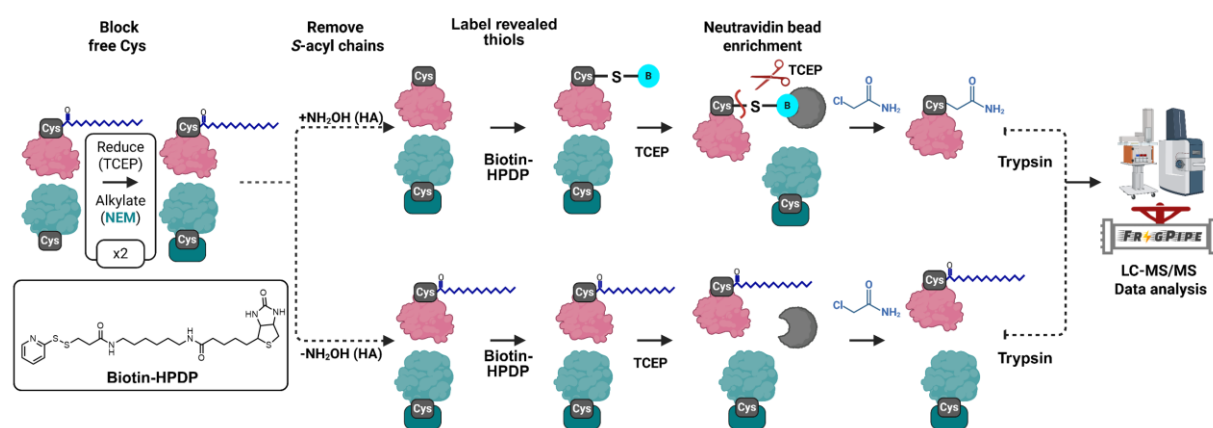
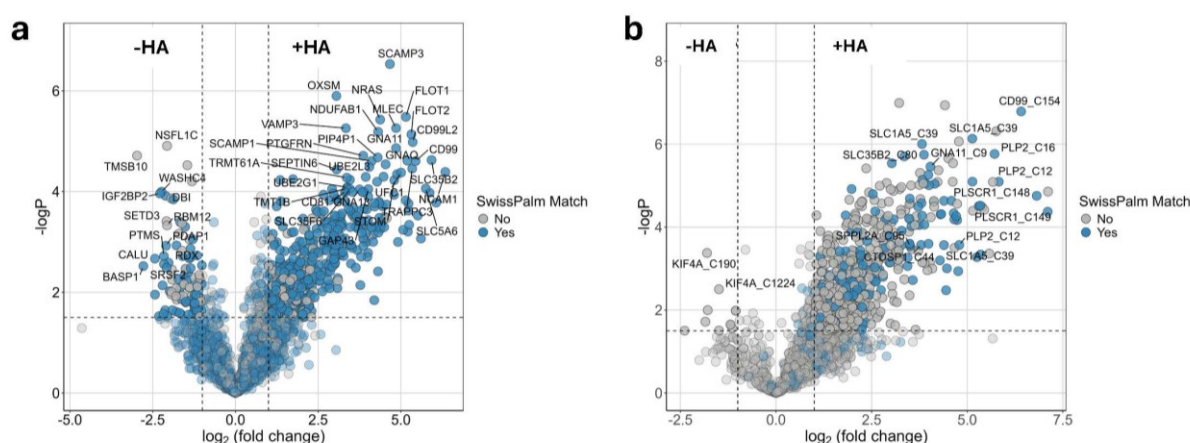


Figure 3.13. Schematic of the site-specific acyl-biotin exchange (ssABE) workflow. Protein-containing lysates are blocked twice with N-ethylmaleimide (NEM), followed by protein precipitation to remove residual NEM. Samples are then treated with hydroxylamine (NH_2OH) to cleave acyl chains, exposing free thiols that are labelled with HPDP-biotin and enriched using NeutrAvidin beads. Following enrichment, S-acylated proteins are eluted off the beads with TCEP, and the newly revealed thiols are labelled with chloroacetamide to enable site-specific identification. Samples are then digested with trypsin and analysed by LC-MS/MS. Figures created with Biorender.com.

Implementation of the ssABE workflow (hereinafter referred to as ssABE) identified 538 S-acylated proteins and 485 CAM Cys (putative S-acylated) sites (**Fig. 3.14**), compared to the 2962 proteins and 909 CAM Cys sites reported by ssABE authors in HEK293 cells (hereinafter referred to as literature ssABE) (**Fig. 3.14**)¹⁷⁵. Interestingly, a modest site overlap was observed, with only 133 of the 485 sites (~27%) matching those previously reported (**Fig. 3.15a**). However, reproducibility improved significantly at the protein level comparison, with 401 out of 538 (74.5%) also identified in literature ssABE and 425 of 538 (79%) found in both

literature ssABE and literature MLCC datasets (**Fig. 3.15a**). The observed discrepancies in identifications could partly be explained by the difference in the experimental and data analysis setup, whereby the published experiments were performed in duplicate and filtered for two valid values, whereas our protocol used four replicates and more stringent filtering of minimum three valid values and a significance cutoff of $\log_2\text{FC} > 1$.



tag, detection of some proteins can be inexplicably biased towards a specific technique, which may explain these discrepancies²¹⁸.

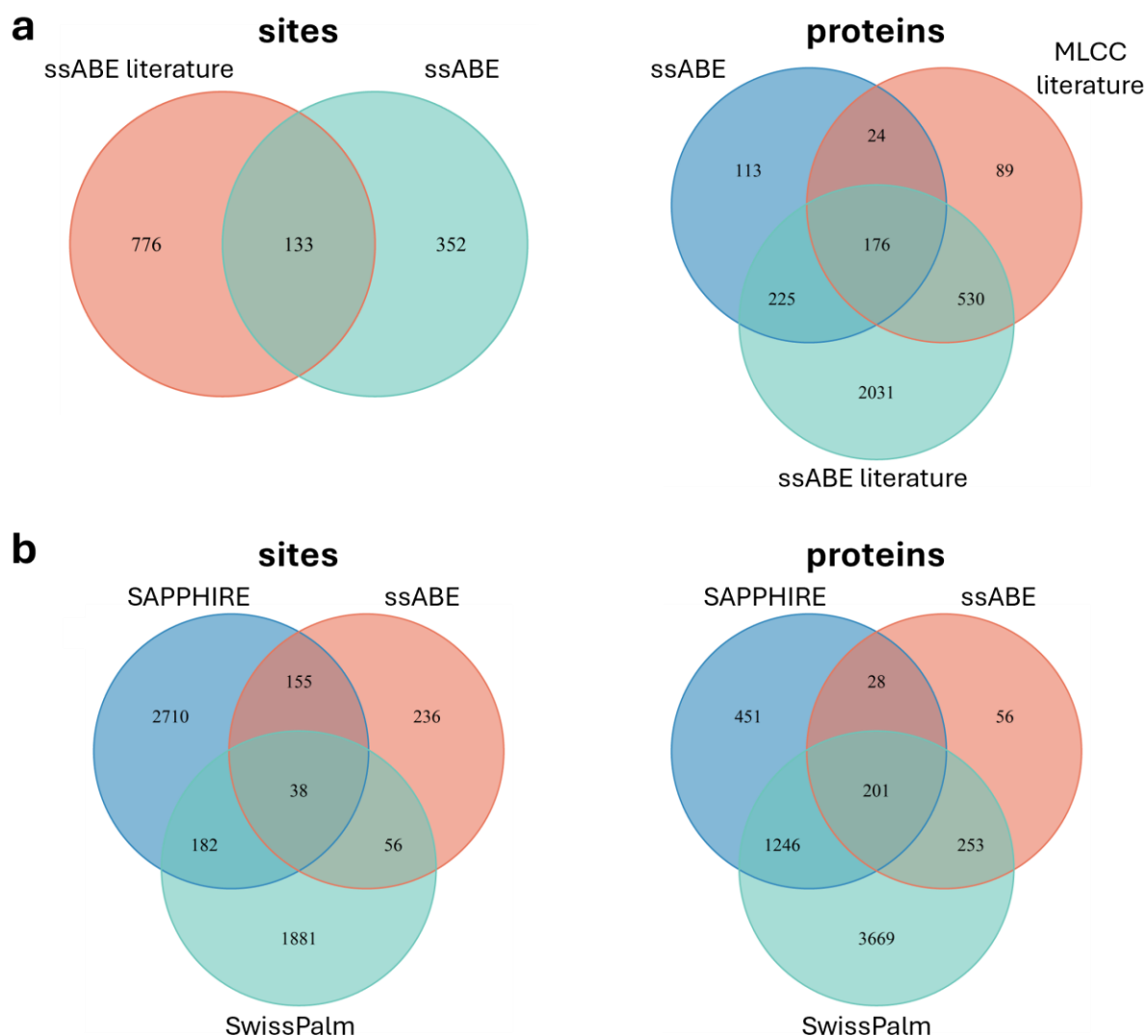


Figure 3.15. Comparison of putative S-acylated sites and proteins identified by SAPPHIRE, ssABE, and the SwissPalm database. a) Venn diagrams comparing S-acylated sites (left) and proteins (right) identified by our replication of the ssABE protocol to those reported in the original ssABE study. The same study also included a metabolic labelling click chemistry (MLCC) dataset, which is shown only at the protein level due to lack of site-specific resolution. All datasets, including ours, were generated in HEK293T cells. **b)** Venn diagrams comparing S-acylated sites (left) and proteins (right) identified by our ssABE replication, SAPPHIRE, and the SwissPalm database, all based on HEK293T data.

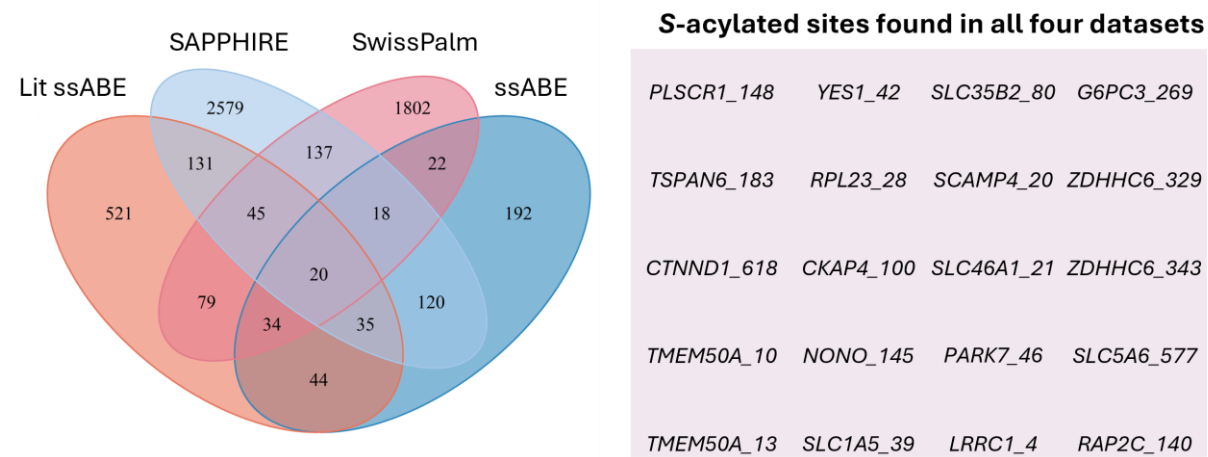


Figure 3.16. Limited S-acylated site overlap is observed between SAPPHIRE, ssABE, literature and SwissPalm datasets.

Overall, direct method comparison of SAPPHIRE against ssABE is difficult due to the differing enrichment levels (protein vs peptide) and LC times. Nevertheless, within the same cell model, SAPPHIRE exhibited enhanced sensitivity over ssABE, with a ~6.4-fold increase in (putative) S-acylated Cys site coverage, while requiring less protein input and precipitation steps, as well as facilitating high-throughput capability (**Table 3.4**).

Table 3.4. Comparison of the SAPPHIRE and ssABE experimental features

	SAPPHIRE	ssABE
Putative S-acylated site IDs	3085	485
Protein/Peptide input per sample	500 µg	1.6 mg
Protein precipitation steps	3	8

3.6.2 Nanographite fluoride-based solid-phase extraction (nGF-SPE) vs SAPPHIRE

The challenges surrounding direct analysis of intact lipidated peptides have been thoroughly established in the field and discussed throughout this thesis. Although impressive coverage of putative S-acyl sites is possible with SAPPHIRE and other methods, all cysteine-centric

workflows remain indirect detection approaches and lack information on the identity of the lipid moiety. On the other hand, MLCC approaches preferentially detect rapid-turnover S-acylation substrates, likewise lack information on the lipid species, and require challenging workflows to achieve site resolution. An ideal analytical workflow would directly identify intact S-acylated peptides, enabling unambiguous profiling of S-acylation sites.

In 2023, a nanographite fluoride-based solid-phase extraction (nGF-SPE) workflow was reported as the first method to achieve separation and detection of intact S-acylated peptides on a proteome-wide scale⁹⁶. This method identified 714 S-acylated Cys in HeLa cells, of which 124 sites were also found in the SAPPHIRE dataset, though differences between cell models likely affected the S-acylation substrate landscape (**Fig. 3.17**). It is important to note that these 714 sites were accumulated across multiple sample preparations and 77 distinct LC-MS/MS

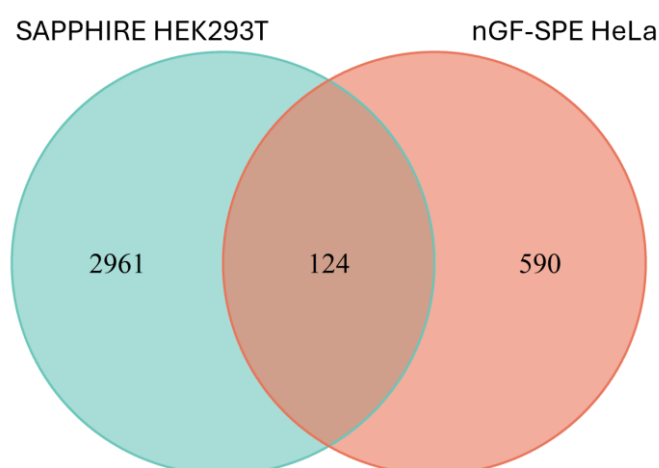


Figure 3.17. Comparison of S-acylated sites identified by nGF-SPE and SAPPHIRE methods.

runs, rather than a single experiment. The highest number of S-palmitoylated sites identified from a single sample preparation – performed in three replicates, each analysed in three LC-MS/MS runs (nine runs in total) – was 221. To corroborate these findings and establish a single sample preparation benchmark for maximal identification, their raw data (deposited on ProteomeXchange repository) was researched using MSFragger. This analysis yielded 303 S-palmitoylated sites. Given that our group had access to the same instrumentation used in the original nGF-SPE workflow, specifically the nanoElute chromatography system (Bruker) and timsTOF MS (Bruker), replication of the workflow that yielded the highest identification of S-

palmitoylation sites in HeLa cells, was undertaken in HEK293T cells. Practically, the nGF workflow is labour-intensive and requires large protein inputs of 7.5 mg of protein per sample (**Table 3.1**) and previous attempts at nGF-SPE in our group were unsuccessful (data not shown). Nonetheless, the workflow was replicated with minimal modifications (detailed in **Methods 6.6.3**), identifying a total of 7419 unique peptides. Using MSFragger closed search, with a specified palmitoylated Cys (+238.2297 Da) modification, 16 palmitoylated Cys were identified, of which FLOT1 Cys34, LDHB Cys164, EEF2 Cys41, RPS5 Cys172 and GNAI Cys139 were also found by SAPPHIRE (**Table 3.2**). Moreover, MSFragger open search mode (OSM) failed to identify any S-acylated Cys, compared to the 1119 sites reported in the publication using Peaks *De Novo* Sequencing method. Given the extremely low identifications observed and previous replication failure, the workflow was anticipated to involve extensive optimisation, which was beyond the immediate objectives of this thesis. Consequently, despite its potential as the least biased approach for detecting S-acyl sites, the nGF-SPE protocol was not successfully replicated and has yet to gain broader adoption within the field beyond the authors of this paper^{96,219}.

Table 3.5. S-palmitoylation sites identified by replication of the nGF-SPE workflow

UniProt ID	Gene	Cys Position	Found in SAPPHIRE?	Found in SwissPalm?
A0A075B6X5	TRAV18	C6	No	No
66O14828	SCAMP3	C142	No	No
O14828	SCAMP3	C147	No	No
O75955	FLOT1	C34*	Yes	Yes
P07195	LDHB	C164*	Yes	Yes
P13639	EEF2	C41*	Yes	No
P46782	RPS5	C172*	Yes	Yes
P63096	GNAI1	C139*	Yes	No
P68431	H3C1	C97	No	No
Q14974	KPNB1	C585	No	No
Q16695	H3-4	C97	No	No
Q5MIZ7	PPP4R3B	C218	No	No
Q5MIZ7	PPP4R3B	C225	No	No
Q7Z4Q2	HEATR3	C186	No	No
Q9Y2L6	FRMD4B	C7	No	No
Q9Y5L0	TNPO3	C419	No	No

3.7 Discussion

3.7.1 Strengths

Due to challenges associated with existing S-acylproteomic approaches, a novel method was developed through implementation of the CysPAT enrichment strategy. Through systematic optimisation of the SAPPHIRE workflow, the four main aims set for this method were achieved. Firstly, the high specificity and sensitivity of SAPPHIRE were demonstrated in HEK239T cells with zero pCys identifications in the -HA condition and detection of >3000 putative S-acylated Cys sites. Secondly, SAPPHIRE represents the first S-acylproteomic workflow amenable to a 96-well plate scale and is the only semi-automatable approach available to date. Next, in HEK239T profiling experiment 500 ug peptide per sample was used, which, although substantial, remained lower than most existing approaches. Further applications of SAPPHIRE will be discussed in Chapter 4, including profiling of human tumour samples, where efficient enrichment was observed at peptide input of ~10 µg, representing the lowest required input reported in any method to date. Lastly, the cost-efficiency was achieved by integrating a label-free quantification (LFQ) strategy, using minimal volumes of Zr-IMAC beads and flexible utilisation of the Oasis HLB desalting sorbent. Additionally, the CysPAT probe can be synthesised in-house, but can also be purchased commercially, making SAPPHIRE more accessible for research labs.

A major strength of SAPPHIRE is its streamlined identification of modified sites, with the highest site coverage in a single experiment to date. Given that some proteins may contain tens to hundreds of Cys, mutagenesis validation screens can become costly and laborious. Knowledge of the precise site can accelerate the functional annotations of S-acylation substrates and is crucial for drug discovery. Furthermore, SAPPHIRE effectively harnessed tools developed for phosphoproteomics, the most extensively studied PTM, for S-acylproteomics. Here, the optimised SAPPHIRE workflow included a dephosphorylation step to enable the sole enrichment of phosphonated Cys, and therefore deep profiling of the S-acylated proteome. Omitting this step would undoubtedly decrease the S-acylated site coverage, but would permit co-enrichment of phospho and phosphonate peptides, enabling unique analyses of the crosstalk between S-acylation and phosphorylation. Given the

dynamic nature of S-acylation and its established role in many signalling cascades, such integrated analyses could provide significant biological insights.

The generic advantages of cysteine-centric workflows over MLCC approaches, such as applicability to diverse biological samples and lack of bias toward high turnover substrates, likewise apply to SAPPHIRE. However, compared to existing cysteine-centric approaches, SAPPHIRE demonstrated superior specificity, sensitivity, reduced protein input requirements, high throughput, semi-automatability, and cost-efficiency. These attributes make SAPPHIRE a valuable addition to the S-acylproteomic toolbox, especially for validation of ongoing and future zDHHC and APT inhibitors screens due to its high throughput.

Beyond S-acylproteomics, significant improvements in phosphonate-based workflows were demonstrated. Specifically, our workflow was streamlined by reducing desalting steps and scaling up the desalting protocol to a 96-well plate. These optimisations may prove useful in activity-based protein profiling (ABPP) and PTM studies, given the growing popularity of PAT methods in these fields^{196,220–222}.

3.7.2 Limitations

Despite its advantages, SAPPHIRE suffers from inherent limitations of a cysteine-centric workflow. Firstly, the identification of any protein or site only directly indicates the presence of a thioester bond on the peptide at the time of cleavage but remains indirect evidence for bona fide S-acylation. Secondly, due to the HA cleavage of the acyl chain, the identity of the fatty acid species, which may provide functional insight, is lost^{96,97,223}. Similarly, linking the cognate zDHHC to specific acylation events is not possible. Lastly, some proteins, for currently unknown reason, may exhibit resistance to HA treatment or to the CysPAT probe, and therefore some physiologically important sites may be missed^{90,96,224}.

Aside from these general cysteine-centric method caveats, the high efficiency of blocking and enrichment steps observed in SAPPHIRE, resulted in many missing values in the -HA condition, thereby reducing statistical power. In the SAPPHIRE data analysis pipeline, missing values were imputed by random numbers drawn from a normal distribution (ND) (1.8 standard deviation down shift and with a width of 0.3 of each sample). Even though ND is suitable for handling missing not at random (MNAR) values, the overreliance on imputation complicates

confidence assignment for identified hits²²⁵. However, this problem is exacerbated in simple +/- HA experiments, where only a single condition (+HA) is expected to have many pCys identifications. Profiling differential S-acylation in response to any kind of treatment or stimulus, with multiple +HA conditions (without missing values), should in theory negate this effect.

Furthermore, despite achieving exceptional coverage of over 3000 sites in HEK293T cells, some well-validated proteins remained undetected by SAPPHIRE. One potential explanation for this could be incomplete dephosphorylation. Specifically, if an S-acylated Cys exists on the same proteotypic peptide as an uncleaved endogenously phosphorylated Ser/Thr/Tyr, this peptide's binding affinity to the Zr-IMAC beads would increase, preventing its recovery. This concern also applies to peptides with double or triple S-acylation sites and could potentially explain the absence of well-validated sites such as IFITM3 Cys71/72 and CANX Cys502/503 in our HEK293 dataset. It has been previously shown that multiply-phosphorylated peptides are better recovered with Zr-IMAC under competitive binding conditions, such as high GA concentration and low bead volumes²⁰⁰. As such, although some tryptic peptides with dual phosphonates have been recovered, such as HRAS Cys181/C184 and CANX Cys360/366, further optimisations specifically to the GA concentration may be crucial to maximise multiply-phosphonated (i.e. S-acylated) peptide recovery. Another possible explanation for this could be the lack of optimisations aimed at enhancing recovery of highly hydrophobic proteins. For example, previous work on site-specific MLCC has previously reported that the use of Rapigest, an MS-compatible surfactant, was necessary for proteomic identification of transmembrane proteins⁹⁰. Furthermore, alternative proteases, such as chymotrypsin and pepsin, have been demonstrated to enhance proteomic identification of hydrophobic proteins^{195,221}. Given the importance of S-acylation in membrane trafficking, the use of alternative protease alone or in combination with trypsin, as well as other sample preparation optimisations, may prove necessary. Although significant improvements have been demonstrated throughout this chapter, numerous opportunities remain to further enhance the recovery and identification of S-acylated proteins.

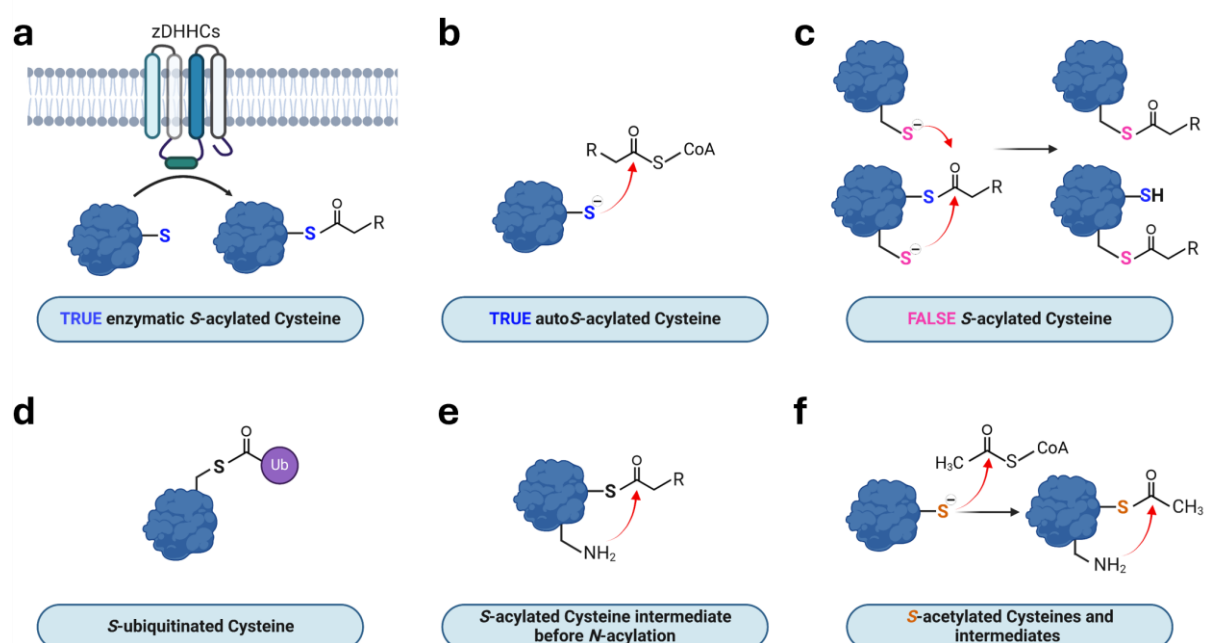
3.7.3 The complexity of hydroxylamine-sensitive *S*-acyl intermediates and non-enzymatic modifications

Bona fide *S*-acylation is catalysed by 23 zDHHC enzymes and reversed by APTs. While most SAPPHIRE hits are true enzymatic *S*-acylation events, the depth of coverage, vast substrate scope and lack of validation, raised concerns regarding the possibility of other cysteine-thioester bond-forming events being captured through HA-dependent phosphonate labelling (**Fig. 3.18**). Therefore, possible events beyond enzymatic *S*-acylation captured by SAPPHIRE were evaluated.

Firstly, enzyme-independent auto acylation can occur through a nucleophilic attack between a deprotonated Cys thiol and the thioester of acyl-CoA (**Fig. 3.18b**). This mechanism has been extensively studied in TEA domain (TEAD) transcription factors^{226,227}. Several TEAD autoacylation sites, including TEAD1 Cys359, were identified by SAPPHIRE, demonstrating its capability to detect enzyme-independent events. Moreover, although TEAD1 Cys359 is a highly annotated and therapeutically targeted palmitoylation site, a recent study suggested this site is instead an intermediate for K366 fatty acylation through intramolecular *SN*-acyl transfer (**Fig. 3.18e**)^{228,229}. A similar concept involving *S*-acetyl intermediates has been established as a mechanism for non-enzymatic Lys acetylation, albeit only *in vitro* (**Fig. 3.18f**)^{179,230}. Furthermore, another recent study proposed that auto short chain acylation, specifically *S*-acetylation, is not only an intermediate for *N*-acetylation, but is also a widespread Cys modification(**Fig. 3.18f**)²³¹. It is therefore plausible that some sites identified by SAPPHIRE, are auto *S*-acyl or *S*-acetyl species or intermediates prior to *N*-acetylation. The latter hypothesis could partially explain the overrepresentation of Lys residues around the acylated Cys (**Table 3.3**) and previous analysis of mouse liver *S*-acylproteomic and Lys acetylation datasets identified 66 overlapping peptides (containing both acetylated Lys and an *S*-acylated Cys). For example, PCNA Cys81 which could not be validated by APE, was identified on the same proteotypic peptide as K77 (KxxxxxC*), a site previously shown to be acetylated²³².

Theoretically, it is possible that the thioester bond of the *S*-acyl/*S*-acetyl Cys could also be prone to nucleophilic attack not only by a nearby deprotonated Lys, but also by a deprotonated thiol Cys (**Fig. 3.18c**). Although this has not been previously explored, the extreme denaturing

conditions used in cysteine-centric workflows can facilitate proximity and accessibility of different Cys to endogenously S-acylated Cys and therefore enable long/short-chain acyl transfer between Cys prior to alkylation. This could potentially explain the identification of non-annotated sites for otherwise well-validated S-acylated proteins, such as PI4K2A Cys183



and CANX Cys360/366, as well as the lower site-level reproducibility between methods and

Figure 3.18. Overview of hydroxylamine-sensitive cysteine events that could be captured by SAPPHIRE. **a)** bona fide zDHHC-mediated S-acylation **b)** auto S-acylation through a nucleophilic attack between the thiolate nucleophile and thioester of the acyl-CoA **c)** S→S acyl transfer might occur in the lysate **d)** Cysteine ubiquitination intermediates **e)** S-acylated cysteine intermediates prior to lysine acylation **f)** Cysteine short-chain S-acetylation, and Cysteine acetylation intermediates

studies. These sites would be of no biological significance, but due to the missing values and imputation discussed above, may appear as high hits. Another notable cysteine-thioester intermediate is cysteine ubiquitination, occurring on active-site cysteines of E1, E2 and E3 ubiquitinating enzymes as well as on substrate proteins (**Fig. 3.18d**)^{233,234}.

Given the dynamic nature of S-acylation, deep S-acylproteomic profiling may be inherently prone to low reproducibility. This is compounded by the fact that cysteine-centric approaches, including SAPPHIRE, may be capturing several cysteine-thioester events beyond enzymatic S-acylation, including auto long-chain S-acylation, auto S-acetylation, SN-acyl transfer intermediates, SS-acyl transfer, and cysteine ubiquitination. Many of these events may be

transient or occur randomly but would be readily identified due to the extreme sensitivity of SAPPHIRE (**Fig. 3.18**). Whereas ubiquitination sites can be easily filtered out, currently no criteria or workflow modifications exist to differentiate between the other events. Experiments incorporating membrane fractionation could partially aid in distinguishing long-chain from acetyl modifications, although many nuclear and cytosolic proteins have validated S-acylation. These challenges are further exacerbated by the lack of robust and high-throughput validation tools and the discovery of specific zDHHC inhibitors would significantly aid in identifying true S-acylation sites. Ultimately, sensitivity of the SAPPHIRE method highlighted a critical yet often overlooked limitation of HA-dependent chemistry of all cysteine-centric workflows. However, while this sensitivity may increase the possibility of false positives in deep S-acylated site profiling (+/- HA) experiments, it may prove advantageous for stimulus-centric experiments. Further applications of SAPPHIRE will be discussed in subsequent chapters.

3.8 Chapter summary

Despite recent advancements in S-acylproteomics, current methodologies often face challenges such as low throughput, high protein input requirements, and limited site-specific resolution. This chapter reports SAPPHIRE, a novel proteomics platform for site-specific proteome-scale profiling of S-acylation in a 96-well plate format, to address existing limitations. Profiling of HEK293T cell lysate using SAPPHIRE resulted in the identification of over 3000 putative S-acylated sites, more than 2000 of which were not previously reported on SwissPalm. When benchmarked against ssABE, SAPPHIRE demonstrated enhanced sensitivity, while maintaining lower protein input requirements (<500 ug).

However, although most sites identified by SAPPHIRE are likely bona fide zDHHC enzymatic substrates, its sensitivity raises important considerations regarding specificity. SAPPHIRE potentially detects additional hydroxylamine sensitive events, such as Cys ubiquitination, auto S-acylation, S-acetylation, and other transient events, raising discussion on the limited interpretation of sites identified by all cysteine-centric workflows. Moreover, attempts at validating several novel sites identified by SAPPHIRE using a gel-based APE assay, were unsuccessful, although the efficiency of the workflow as well as the potentially varied stoichiometry of the sites may account for these results. This caveat however, is inherent to

all hydroxylamine-based approaches, and lies beyond direct optimisation of the SAPPHIRE method.

Further investigation into the full potential of SAPPHIRE, including its low input compatibility and ability to detect dynamic S-acylation events, will be explored and elaborated upon in Chapter 4.

Chapter 4.

Application of SAPPHIRE for S-acylated proteome profiling across disease models and organisms

4.1 Introduction

Chapter 3 outlined the development, optimisation, and initial implementation of the SAPPHIRE proteomics platform, demonstrating deep S-acylated site profiling in HEK293T cells. With the optimised method now in hand, its versatility and sensitivity across a broader range of experimental models and sample types will be assessed in this chapter. Given the central role of S-acylation in health and disease, SAPPHIRE will be applied to several selected systems, including human tissues, *Toxoplasma gondii* parasites, and pyroptosis- and stress-induced human cell line models. These systems were specifically selected based on their relevance to our group's interests, as well as to evaluate the three key advantages of SAPPHIRE highlighted previously - reduced protein input, increased throughput, and applicability to various biological samples.

4.2 Tumour tissue analysis

As discussed in **Chapter 1.5.2**, dysregulated S-acylation is implicated in many cancers and resistance to cancer therapies. This link has been established through extensive biochemical, bioinformatic, and chemoproteomic studies using human cell lines and murine models. However, analyses of human tumour tissues have so far been limited to profiling zDHHC expression at transcriptomic and proteomic levels, whereas global S-acylation patterns remained unexplored^{187–190}. One of the key objectives set for the SAPPHIRE method was enabling human tumour profiling by lowering the required protein input. These analyses could facilitate novel biomarker discovery, provide additional evidence for dysregulated S-acylation patterns in cancer, and potentially identify new therapeutic targets.

To enable comparative analyses of differential S-acylation in cancer, matched healthy biopsies -healthy tissue adjacent to malignant sites - were necessary alongside tumour tissues. However, obtaining such patient samples from the Imperial College Tissue Bank proved challenging. Fortunately, through collaborations established within our research group at the time of this PhD, matched tumour and healthy samples from one breast cancer and two colorectal cancer (CRC) patients were successfully acquired.

4.2.1 Breast cancer

The first application of SAPPHIRE to patient samples was conducted on breast cancer tumour and matched healthy tissue. Tissue lysis was done by Dr. Aidan Pidd, using a protocol optimised for a separate project focused on protease activity within our lab. Notably, this protocol employs a low detergent buffer to preserve protease activity, which may lead to a loss of the lipidated protein fraction (see **Methods 6.1.6**). Following lysis, each sample was split into three technical replicates and processed by SAPPHIRE. As the tumour samples were very small, low extracted protein yields were observed, resulting in the final peptide input ranging between 20-40 μ g per replicate for the Zr-IMAC enrichment.

Despite such low input, 2999 pCys were identified across the six samples. However, even with the double CAA alkylation step, many identifications appeared in the -HA fractions (**Fig. 4.1**). This likely resulted from deviations from the standard lysis protocol, which potentially limited Cys accessibility during the first alkylation step. Nevertheless, the specificity of labelling was higher for tumour samples (for unknown reasons) and further analysis revealed differential enrichment of 270 pCys in tumour compared to healthy tissue ($-\log P > 1.5$ and $\log_2 FC > 1$), representing 240 unique (putative) S-acyl sites, of which 26 have had previous annotations on SwissPalm (**Fig. 4.2a**).

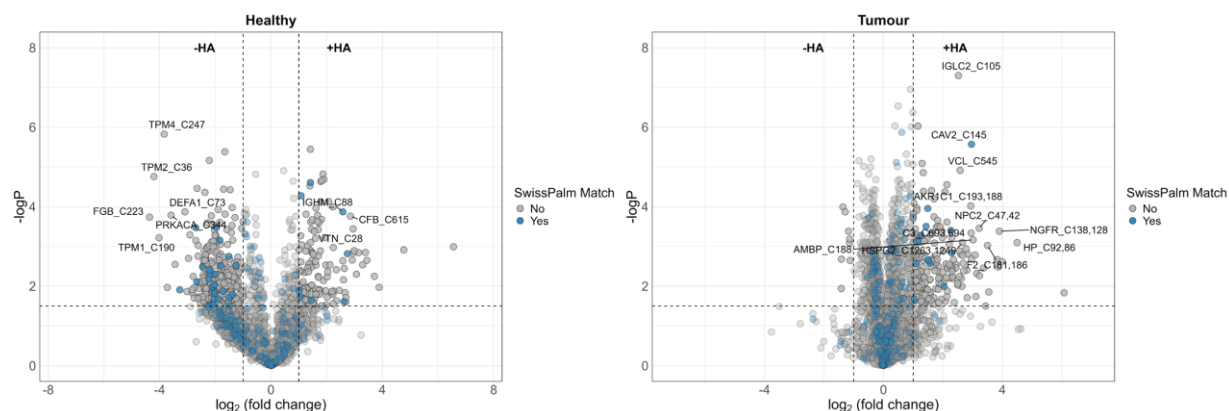


Figure 4.1. Site-specific S-acylated proteome profiling of breast cancer patient tissue with SAPPHIRE. Volcano plots depicting differential enrichment of phosphonated Cys peptides (i.e. putative S-acylated Cys) between +HA and -HA healthy (left) and tumour tissues (right) (student's t-test; FDR = 0.05, S0 = 0.1) n=3 technical. The blue points denote previous identification of the exact site in the SwissPalm database.

KEGG pathway enrichment analysis of the corresponding 174 proteins highlighted the enrichment of extracellular matrix (ECM)-receptor interaction and several cancer-related pathways (**Fig. 4.2b**). Validated S-acylation sites implicated in oncogenesis, such as β -catenin Cys466 (previously linked to CRC tumorigenesis), Caveolin-1 Cys143 and Caveolin-2 Cys145, were also identified^{235–237}. However, despite prior reports of several zDHHs upregulation in breast cancer, including zDHH3, zDHH7, zDHH9, zDHH17, zDHH20 and zDHH21, no zDHH peptides were identified in this experiment either due to the suboptimal lysis buffer or low peptide input^{190,238}.

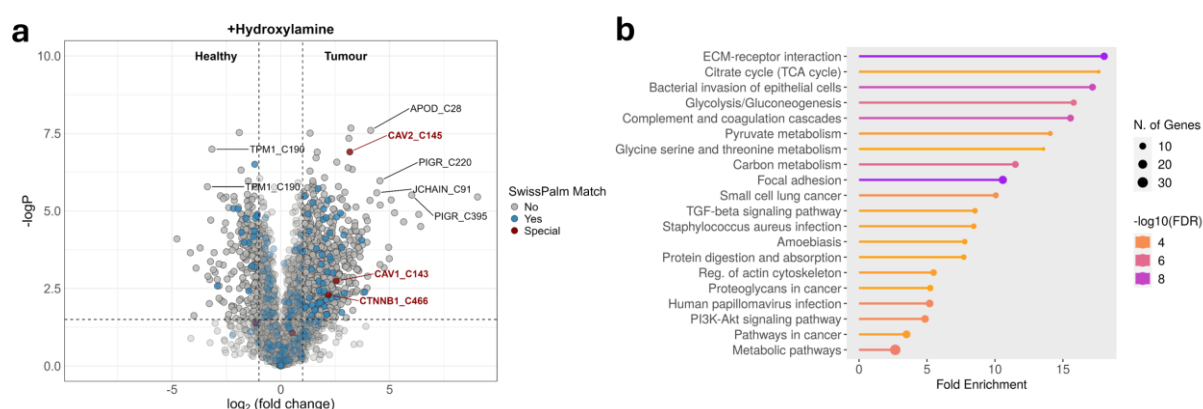


Figure 4.2. Identification of breast cancer tumour-enriched S-acylation sites. a) Volcano plot showing tumour and healthy-enriched S-acylated sites (student's t-test; FDR = 0.05, S0 = 0.1) n=3 technical. The blue points denote previous identification of the exact site in the SwissPalm database. **b)** KEGG pathway enrichment analysis showing the top 20 pathways enriched ($\log_2$ FC cut-off of +1) in tumour over healthy tissue

However, drawing definitive conclusions remains challenging due to inherent differences in protein expression expected between malignant and healthy tissues. An S-acylation event may appear highly upregulated in the tumour tissue simply due to low or absent expression of its corresponding protein in the healthy sample. Running a concurrent whole proteome analysis and subsequent normalisation of intensities would enhance the ability to discern true tumour-specific increases in S-acylation. However, due to the limited amount of sample, whole proteome analysis was omitted, and all available protein material was taken forward for the SAPPHERE protocol. Therefore, without normalisation to overall protein expression, interpreting the biological significance of identified differential S-acylation events remained

4.2.2 Colorectal cancer

The CRC and matched healthy samples from two patients were lysed by Dr. Aidan Pidd using the same protocol as described above and split into three technical replicates. These CRC samples were smaller than the previously analysed breast cancer samples, leading to even lower protein yields upon extraction. Peptide quantification before desalting revealed low quantities, particularly for tumour tissue from patient two, where peptide amounts were below detection limits for the fluorometric peptide quantification kit (Thermo) (**Appendix Table 7.3**). Nonetheless, small quantities of each of the untreated samples (-HA) were preserved for whole proteome analysis, given the challenges experienced in interpreting breast cancer profiling data. However, due to limited quantification and sample availability, suboptimal peptide quantities (<200 ng) were used for whole proteome analysis. Due to the same reasons, peptides were not normalised prior to Zr-IMAC enrichment, with inputs ranging between 0-20 µg.

Whole proteome analysis in DIA mode identified a total of 4797 proteins across all samples, revealing many differentially expressed proteins between healthy and tumour tissues (**Appendix Fig 7.3**). Unfortunately, despite previous reports of zDHHC enzyme dysregulation in CRC (downregulation of zDHHC1/2 and upregulation of zDHHC6/9), only zDHHC3 and zDHHC13 peptides were identified^{213,239,240}. Notably, patient one exhibited statistically significant overexpression of zDHHC3 in tumour tissue ($-\log P > 1.5$, $\log_2 FC > 1$) (**Appendix Fig. 7.3**). Proteins differentially expressed in tumour versus healthy tissues were compared between the two patients, revealing overlaps of 216 tumour-enriched and 471 healthy-enriched proteins, each exhibiting differential KEGG pathway enrichments (**Appendix Fig. 7.4**).

SAPPHIRE analysis identified a total of 1792 pCys peptides across all samples, 1556 of which were amenable to normalisation by protein abundance due to their presence in the whole proteome data (**Fig. 4.3, Appendix Fig. 7.5**). Similarly to the breast cancer samples, some non-specific CysPAT labelling was observed in the -HA fractions (**Appendix Fig. 7.6**). Normalisation was performed by subtracting $\log_2$ -transformed whole proteome intensities from $\log_2$ -transformed SAPPHIRE intensities for a given protein/peptide, as detailed in **Chapter 6.8.3.2**. Without normalisation to the whole proteome data, comparative analysis of tumour-enriched

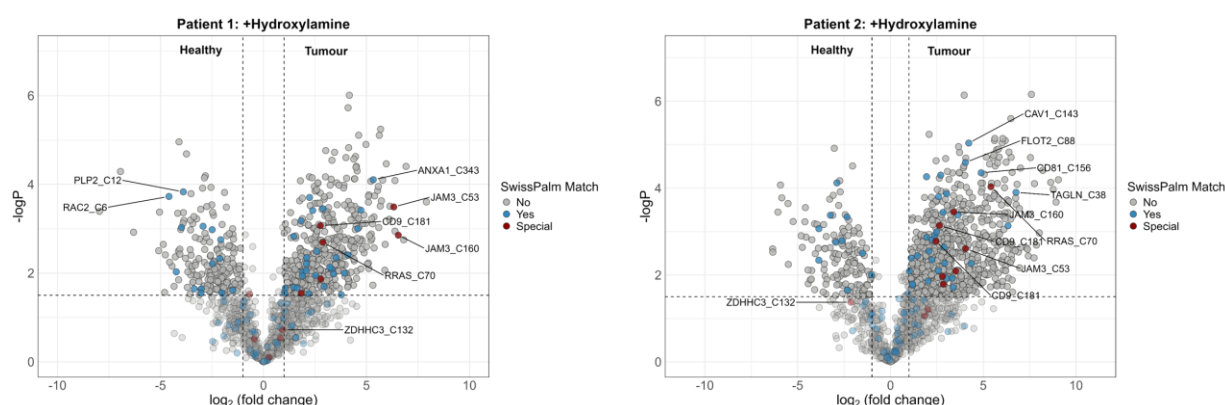


Figure 4.3. SAPPHIRE identification of colorectal cancer tumour-enriched S-acylation sites. Volcano plot depicting tumour-enriched S-acylated sites (student's t-test; FDR = 0.05, $S_0 = 0.1$) $n=3$ technical. The blue points denote previous identification of the exact site in the SwissPalm database. Data was normalised to protein expression.

unique (putative) S-acylated sites revealed a 26-site overlap between the two patients. After normalisation, this overlap increased to 209 sites, of which 11 were previously annotated on SwissPalm (**Fig. 4.4a**). KEGG pathway analysis of the overlapping proteins revealed significant associations with pathways related to ECM-receptor interaction, focal adhesion, and cancer-associated pathways, notably the PI3K-Akt signalling pathway (**Fig. 4.5**). Conversely, comparative analysis of S-acyl sites enriched in healthy tissue, revealed a modest overlap between the patients in both normalised and non-normalised datasets, insufficient for functional enrichment analysis (**Fig. 4.4b**). Given that a higher inter-patient overlap was observed in healthy tissues on the global protein expression level, suggests that healthy-enriched S-acylation patterns are more random, reflecting physiological variability (**Appendix Fig. 7.4**). Conversely, the substantial S-acylated site overlaps between tumour samples, in both normalised and non-normalised datasets, implies consistent S-acylation of specific

Chapter 4. Application of SAPPHIRE for S-acylated proteome profiling across disease models and organisms

sites and proteins linked to cancer-associated cellular reprogramming. PCA analysis of both normalised and non-normalised data revealed strong clustering between the technical replicates and a higher intra-patient separation (**Fig. 4.6**).

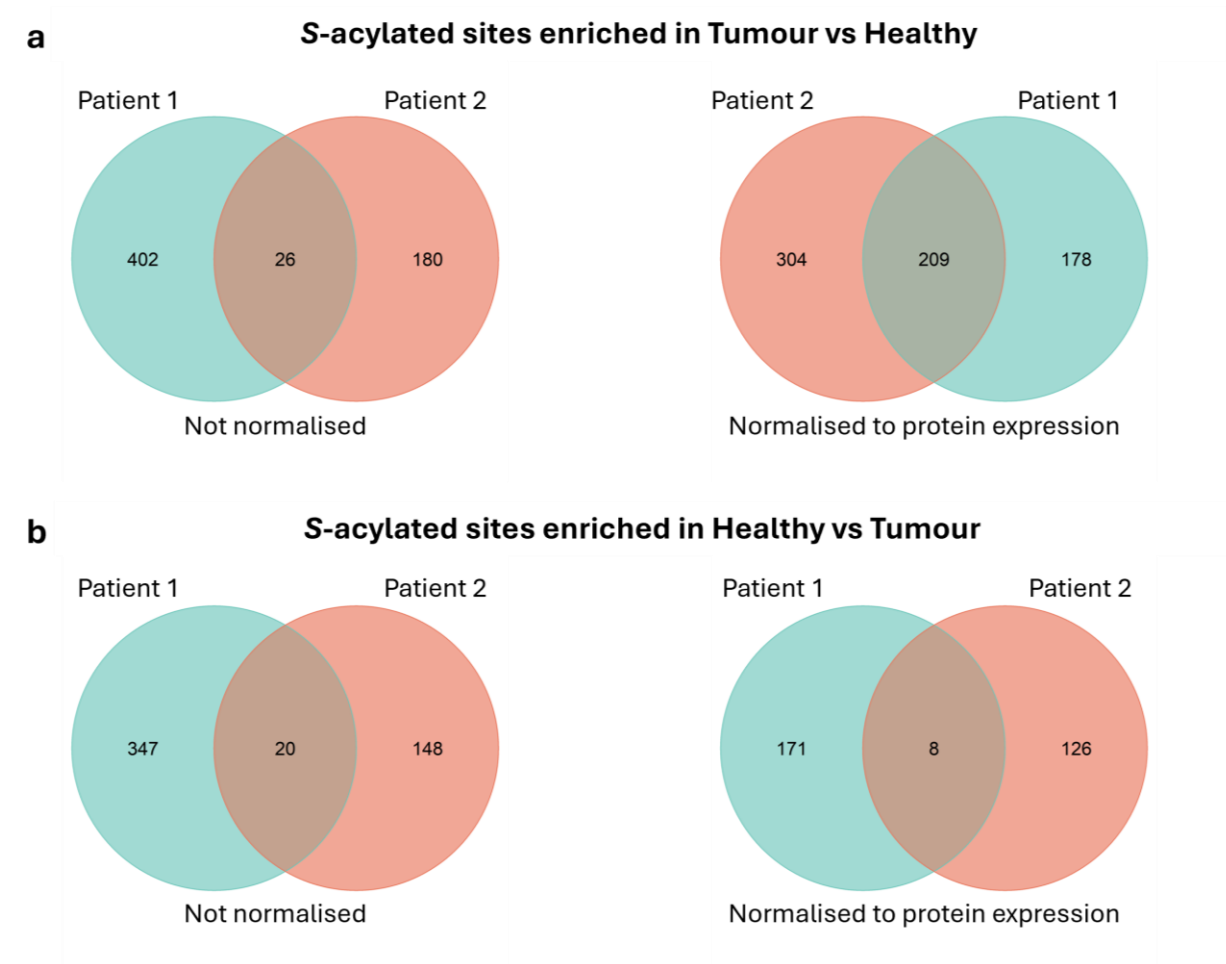


Figure 4.4. Comparison of colorectal cancer patients’ tumour-enriched vs healthy-enriched S-acylated sites. **a)** Venn diagram depicting overlap of significantly enriched S-acyl sites in tumour compared to healthy tissue ($\log_2FC > 1$, $-\log P > 1.5$) between Patient 1 (teal) and Patient 2 (orange) of non-normalised SAPPHIRE data (left) and SAPPHIRE data normalised to whole proteome analysis (right) **b)** Venn diagram depicting overlap of significantly enriched S-acyl sites in healthy compared to tumour tissue ($\log_2FC > 1$, $-\log P > 1.5$) between Patient 1 (teal) and Patient 2 (orange) of non-normalised SAPPHIRE data (left) and SAPPHIRE data normalised to whole proteome analysis (right)

The top ten tumour-enriched pCys peptides ranked by normalised mean $\log_2FC$ (tumour vs. healthy) from both patients are summarised in **Appendix Table 7.4**. Notably, an overrepresentation of integrin $\alpha 7$ (ITGA7) peptides was observed, with 16 Cys identified. ITGA7, in a heterodimeric complex with integrin $\beta 1$, acts as a transmembrane receptor for ECM

protein laminin, and has been characterised as a marker for poor prognosis, implicated in tumorigenesis, metastasis and resistance to therapy in various cancers^{241–243}. Although, no previous evidence exists for ITGA7 S-acylation, previous studies demonstrated the crucial role of S-acylation in integrin function^{244,245}.

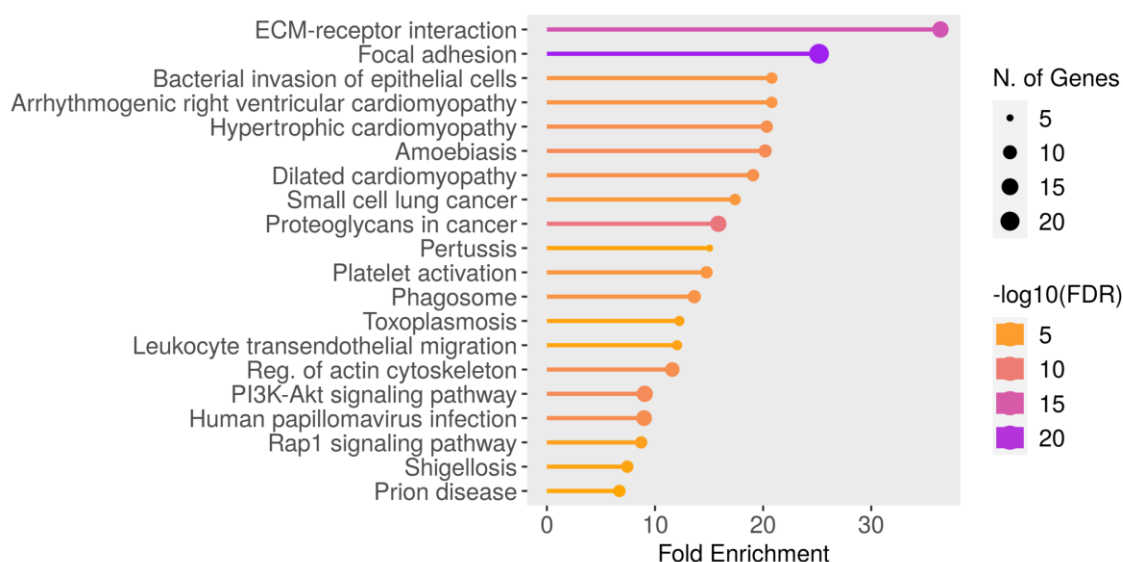


Figure 4.5. KEGG pathway enrichment analysis of colorectal cancer tumour-enriched S-acylated sites. Pathway enrichment was performed on S-acylation sites enriched in tumour over healthy (Log_2FC cut-off of +1) in both patients. The top 20 KEGG pathways significantly associated with these overlapping tumour-enriched sites are shown.

Thus, validation of ITGA7 S-acylation and its potential role in cancer cell models, would be an interesting follow-up. Another notable hit is junctional adhesion molecule C (JAM3) Cys53, the S-acylation of which on Cys264/265 by zDHHC7, was previously shown to be important for its localisation to tight junctions and cell migration (**Fig. 4.3**)²⁴⁶. Outside of the top ten, important cancer-associated S-acyl sites like GNAQ Cys9/Cys10, previously linked with uveal melanoma, was identified²⁴⁷. However, validated proteins with alternative sites, such as RRAS Cys70 but not validated Cys213, and CD9 Cys181 but not Cys9/78/79/87/218/219, were identified. Of note, only zDHHC3 Cys132 was identified in both datasets and zDHHC20 Cys131 was identified by SAPPHERE only.

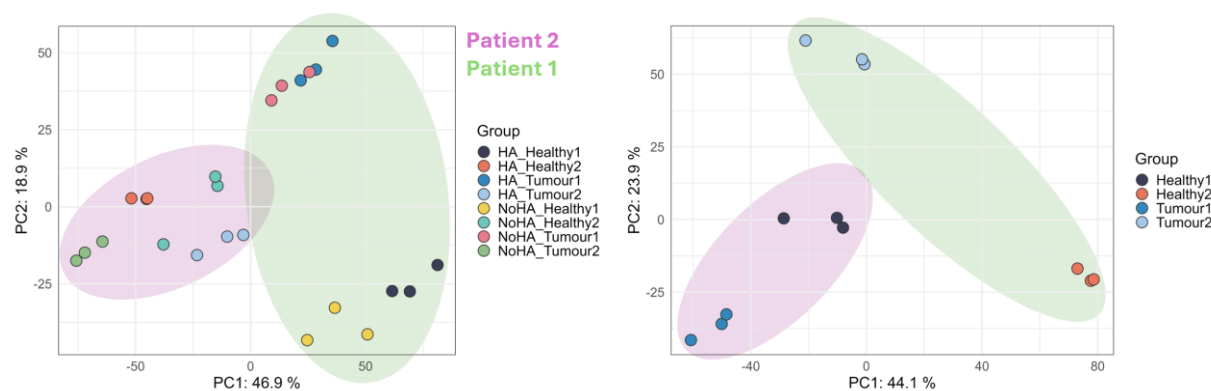


Figure 4.6. Principal component analysis of colorectal cancer patient samples profiling with SAPPHIRE. PCA of the 24 colorectal cancer patient samples, shown before (left) and after (right) normalisation to the corresponding whole proteome data.

4.2.3 Conclusions

These initial experiments demonstrate the proof-of-concept for the application of SAPPHIRE to tumour samples. The generated datasets of differential S-acylation but also protein expression in breast and CRC tumours will be a useful resource for the field and may uncover novel findings. Furthermore, the extremely low peptide inputs tested in these experiments underscore the sensitivity of the SAPPHIRE method, suggesting that lower protein input may be used in other experiments as well. However, further optimisations to the tissue extraction protocol, suitable for profiling of lipid PTMs will be necessary for studies with larger patient cohorts, to minimise false positive identifications. Moreover, many proteins were identified with sites distinct to those previously validated. It is not currently possible to discern between true S-acylation or $S \rightarrow S$ -acyl transfer and S-acetyl artefacts (see **Chapter 3.7.3**). Further validation of sites of interest through mutagenesis screens and orthogonal techniques in cell-based models is crucial for confirming bona fide S-acylation states.

4.3 *Toxoplasma gondii*

One of the biggest advantages of any cysteine-centric method is its applicability to a range of biological samples from different species. This is especially important given that S-acylation is highly conserved amongst all eukaryotic organisms, including protozoan parasites²⁴⁸. Eukaryotic parasites affect billions of people worldwide, causing significant health and socioeconomic burdens²⁴⁹. Consequently, understanding the biological roles and therapeutic

relevance of S-acylation and other lipid PTMs in these parasites has become an area of substantial interest. Particularly, for epidemiologically relevant parasites *Plasmodium falciparum*, *Toxoplasma gondii*, and species of *Trypanosoma* and *Leishmania*, responsible for malaria, toxoplasmosis, African Sleeping disease, and leishmaniasis respectively²⁴⁸. While inhibition of NMT has gained recognition as a promising antiparasitic strategy, further characterisation of S-acylation substrates is necessary to validate this PTM as a promising target in antiparasitic therapy²⁵⁰.

Thus far, SAPPHIRE demonstrated deep site coverage in a human HEK293T cell line model and human patient tumours. To test its versatility further, SAPPHIRE was applied to *T. gondii*, the causative agent of toxoplasmosis. Toxoplasmosis infection affects approximately one third of the world population, and although usually asymptomatic, can cause neurological disorders in immunocompromised individuals, as well as miscarriages and congenital disorders²⁵¹. To date, two global S-acylproteomic studies in *T. gondii* identified 282 and 401 S-acylated proteins using MLCC coupled with HA (MLCC-HA) and ABE approaches respectively^{92,252}. Although these numbers may be significant underestimations, with CSS-Palm 4.0, predicting that more than half of the *T. gondii* proteome is potentially S-acylated, though much caution should be given to interpreting computational tools for predicting S-acylation⁹². Despite these insights, no comprehensive site-specific analysis of the S-acyl proteome has been conducted in *T. gondii* or any other protozoan parasites.

4.3.1 Site-specific S-acylated proteome profiling

T. gondii parasites were cultured and lysed by Dr. Janessa Grech and then subjected to SAPPHIRE and whole proteome workflows using the respective DDA and DIA label free quantifications. SAPPHIRE identified a total of 6026 pCys, of which 5920 were enriched in +HA ($-\log P > 1.5$ and $\log_2 FC > 1$), mapping to 4930 unique S-acylation sites across 2053 proteins (**Fig. 4.7a**). Crucially, minimal changes in protein expression were observed between the +/- HA conditions, confirming that only increases in S-acylation are HA dependent (**Fig. 4.7b**, **Appendix Fig. 7.7**). Comparative analysis revealed that 167 (8.1%) of these proteins were previously found by the MLCC-HA approach (**Fig. 4.8a**). Of these, components critical to the parasite's glidesome machinery, such as Myosin A (Cys57/322/338/395/548/668), GAP45 (Cys201), GAP40 (Cys318), GAP50 (Cys308/335), and Myosin C (Cys542/567) were identified⁹². However, comparative analysis on the site level is hindered by the lack of

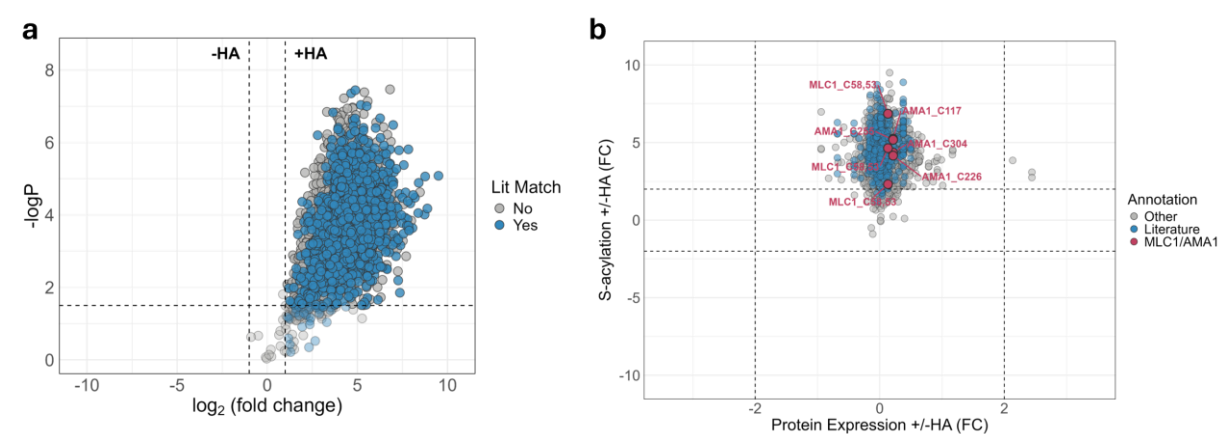


Figure 4.7. Site-specific S-acylated proteome profiling of *Toxoplasma gondii* with SAPPHIRE. **a)** Volcano plot showing differential enrichment of phosphonated cysteine-containing peptides (pCys) (i.e. putative S-acylated) between +HA and –HA *T. gondii* samples (Student’s t-test; FDR = 0.05, S_0 = 0.1) n=3 biological. Blue points indicate proteins previously identified by Foe et al. using the MLCC-HA approach. **b)** Quadrant plot comparing changes in S-acylation with corresponding changes in protein expression between + HA and –HA.

previously annotated sites, with the SwissPalm database containing only four annotated *T. gondii* sites - AMA1 Cys504 and HSP20 Cys3/4/160. Whereas no HSP20 peptides were identified in our dataset, AMA1 Cys117/226/255/304 but not Cys504 were phosphonate modified (**Fig. 4.7b**). Another study had previously identified MLC1 Cys8/11 as S-acylated through a gel-based MLCC approach. SAPPHIRE however identified a dually S-acylated MLC1

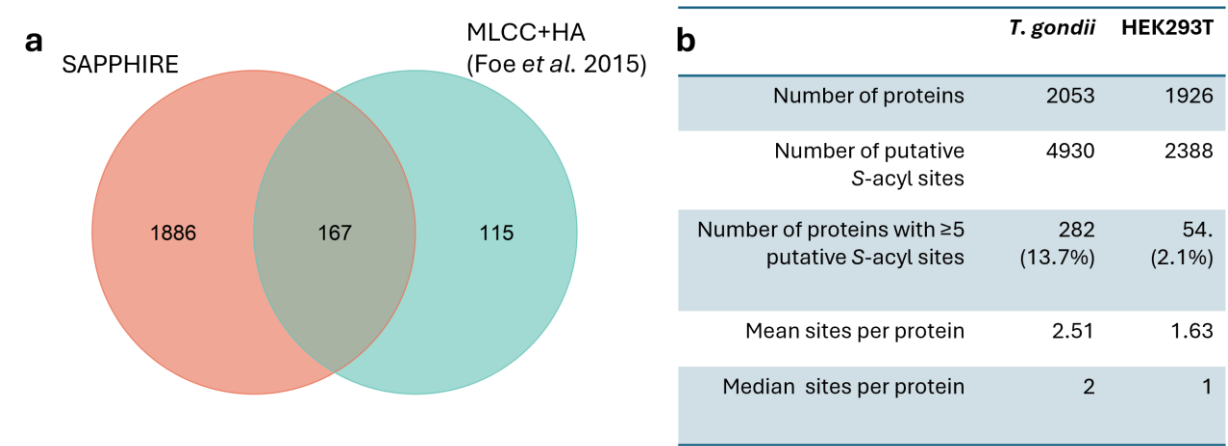


Figure 4.8. Comparative analysis of *T. gondii* S-acylation sites identified by SAPPHIRE. **a)** Venn diagram depicting the overlap between S-acylated proteins identified by SAPPHIRE and those reported by Foe et al. (2015) using the MLCC-HA method. **b)** Comparison of S-acylated site and protein counts, as well as their characteristics, between *T. gondii* and HEK293T datasets (see **Chapter 3.4.1**). A greater number of S-acylated cysteines per protein were identified in *T. gondii*, suggesting increased site density between the two species.

Cys 53/58 tryptic peptide (**Fig. 4.7b**). Interestingly, many proteins were S-acylated on multiple distinct Cys sites. Of the 2053 identified proteins, 282 (13.7%) had five or more phosphonated (i.e. putative S-acylated) Cys, with a mean average of 2.51 S-acylated Cys per protein (median 2) and a maximum identification of 21 Cys in Cysteine Repeat Modular Protein B (CRMPB). This is in direct contrast to human HEK293T cells, where a mean average of 1.63 sites per protein and 54 (2.1%) have more than five S-acylated Cys (**Fig. 4.7b**).

4.3.2 Conclusions

This experiment constitutes the first site-specific S-acylated proteome analysis for a protozoan parasite, further expanding the scope of putative S-acylation substrates in *T. gondii*. Consistent with previous reports of the abundance and functional significance of S-acylation in *T. gondii* biology, our findings highlight that inhibiting S-acylation may potentially illicit pleiotropic effects, similar to NMT inhibitors. Given that the *T. gondii* genome encodes for 18 zDHHCs, renders their selective or pan inhibition as a promising therapeutic avenue²⁴⁸. For example, prior work showed that 2-BP treatment disrupted the gliding motility, host cell invasion and overall parasite morphology. Moreover, inhibition of a *T. gondii* APT (TgPPT1) increased parasite motility and invasion^{253,254}. Even though some discrepancies were found between the validated sites in the literature and the sites identified by SAPPHIRE, limited site annotation currently exists for *T. gondii* and other parasites, therefore further validation studies are needed.

Overall, this site-specific dataset for the *T. gondii* S-acylated proteome demonstrated the robustness and broad applicability of the SAPPHIRE platform across diverse species and provided a useful resource for the field. Further application of SAPPHIRE across different protozoan parasites will aid in elucidating the overlapping and distinct S-acylated proteins and sites amongst these organisms. Moreover, SAPPHIRE will be a useful tool for profiling global S-acylation changes in response to novel antiparasitic agents.

4.4 Pyroptosis

To test the sensitivity of SAPPHIRE in identifying dynamic and stimulus-driven S-acylation events, a model of pyroptosis was selected. Pyroptosis is a highly inflammatory form of programmed cell death, crucial for host defence against pathogen infection and other danger signals²⁵⁵. The Gasdermin (GSDM) family, particularly Gasdermin D (GSDMD), are the key

mediators of pyroptosis. Typically, activation of the inflammasome triggers caspase-1 mediated cleavage of full-length GSDMD (FL-GSDMD) into a 31 kDa N-terminal and a 22 kDa C-terminal fragments. The N-terminal fragment (GSDMD-NT) then undergoes oligomerisation and translocates to the cell membrane to assemble pores of 10-33 nm in diameter. These GSDMD pores result in lytic cell death, release of inflammatory cytokines, notably interleukin 1 β (IL-1 β), and downstream immune responses. GSDMD-mediated pyroptosis is also induced by the non-canonical inflammasome pathway, through caspase 4, 5 and murine caspase 11-mediated cleavage²⁵⁵.

Until recently, cleavage of FL-GSDMD was considered the sole trigger in its pyroptotic activation. However, the recent discovery of GSDMD Cys191 (murine Cys 192) S-acylation, has shifted this paradigm, with five independent studies establishing it as the key event for GSDMD membrane trafficking, pore formation and pyroptosis^{256–260}. The current model proposes that FL-GSDMD Cys191 is S-acylated at low levels during the priming stage. Upon inflammasome activation and subsequent cleavage, S-acylated GSDMD-NT (and to some extent FL-GSDMD) interacts with mitochondrial lipid cardiolipin to form mitochondria damaging pores, which in turn increases reactive oxygen species (ROS) generation. Elevated ROS subsequently lead to the upregulation of specific zDHHC enzymes, resulting in higher levels of GSDMD S-acylation, inducing its translocation to the membrane, oligomerisation and pore formation (**Fig. 4.9**)^{256,257}. Additionally, one study has demonstrated that GSDMD Cys191 is deacylated by APT2 at the membrane, which promotes its oligomerisation²⁵⁹. Importantly, this S-acylation event is crucial for both canonical and non-canonical inflammasome activation pathways, making it an attractive therapeutic target for inflammatory disorders. Unlike specific NLRP3 inflammasome inhibitors, such as MCC950, inhibitors for GSDMD S-acylation would broadly block pyroptosis and downstream inflammation in response to any stimulus. Notably, small molecules like dimethyl fumarate and disulfiram, have been demonstrated to elicit anti-pyroptotic and anti-inflammatory effects by targeting GSDMD Cys191^{260–262}.

However, despite the consensus on GSDMD Cys191 S-acylation, the assignment of the cognate zDHHC remains contradictory. Two studies assigned zDHHC5 and zDHHC9 through overexpression and siRNA knockdown (KD) screens in HEK293 and Tohoku Hospital Pediatrics-1 (THP1) cells and CRISPR-Cas9 KO in THP1s^{256,257}. Two other studies reported zDHHC7 through zDHHC overexpression in HEK293 cells as well as shRNA KD and KO in

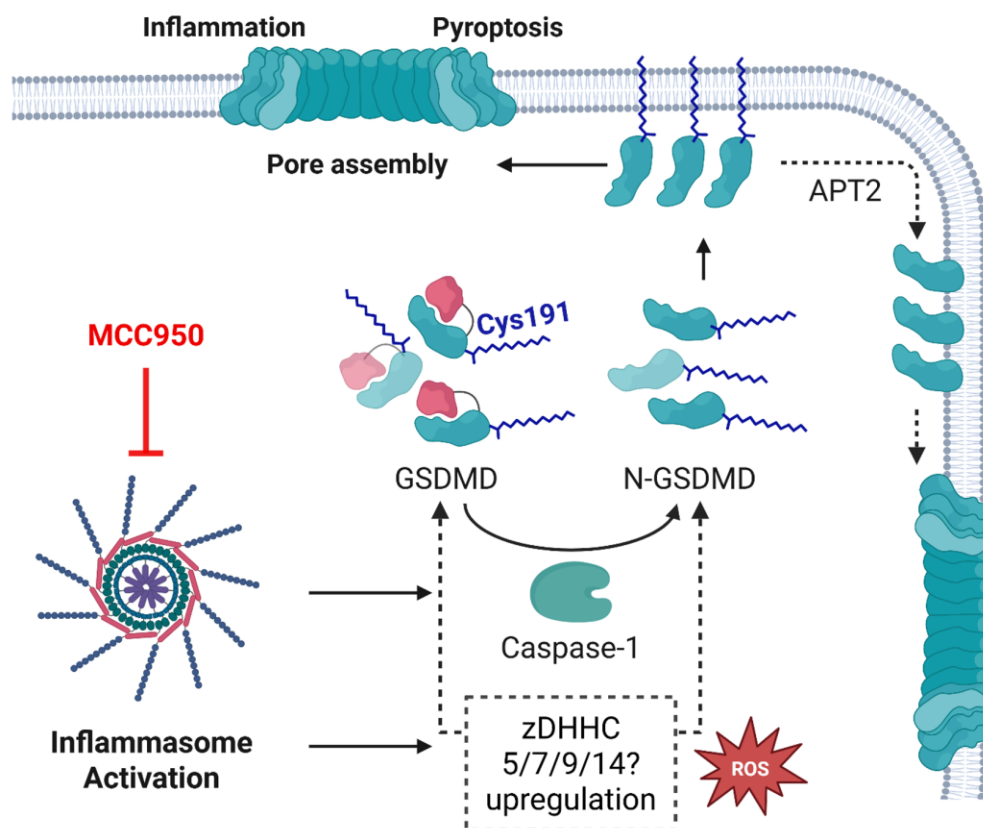


Figure 4.9. GSDMD Cys191 S-acylation is the key checkpoint in pyroptosis. The current proposed model for how S-acylation mediates N-terminal GSDMD (N-GSDMD) and full-length GSDMD (FL-GSDMD) pore formation. Upon priming with LPS, full-length GSDMD is slightly S-acylated at Cys191. Upon full activation of the inflammasome and subsequent activation of caspase-1, FL-GSDMD gets cleaved and N-GSDMD is S-acylated at high levels, although the cognate zDHHC remains ambiguous. This increases GSDMD's membrane affinity, leading to its membrane trafficking, pore formation, and pyroptosis. One study proposed that further deacylation by APT2 at the membrane is required for oligomerisation. Figure created with Biorender.com.

murine bone marrow derived macrophages (BMDMs)^{258,259}. Another study assigned zDHHC14 through siRNA KD in cardiomyocytes²⁶⁰. These studies predominantly relied on targeted gel-based and microscopy validation methods without conducting a comprehensive global S-acylproteomic analysis. Only one study reported a modified ABE proteomics approach involving overexpression of FLAG-tagged GSDMD, anti-FLAG- enrichment and subsequent ABE for targeted identification of the modified GSDMD Cys191²⁵⁷. It was therefore hypothesised that the first global profiling of S-acylated sites in pyroptosis at an endogenous level with SAPPHIRE, may aid in deciphering these inconsistencies and provide novel insights into this biologically and therapeutically relevant event. Furthermore, the effects of

inflammasome inhibition on the downstream GSDMD Cys191 S-acylation or S-acylation patterns in general have not been previously characterised but could provide interesting insights into their anti-pyrototic mode of action. Given the availability of MCC950, a small molecular inhibitor of NOD-like receptor protein 3 (NLRP3) inflammasome, in our lab, its effects on global S-acylation patterns were also sought to be investigated.

4.4.1 GSDMD Cys191 S-acylation

To profile changes in pyroptosis, THP1 cell line was selected as it endogenously expresses the inflammasome pathway proteins and was previously used in three studies reporting GSDMD Cys191 palmitoylation. For the pyroptosis condition, THP1 cells were differentiated by pretreatment with phorbol-12-myristate-13-acetate (PMA), primed with lipopolysaccharide (LPS) and stimulated with 10 μ M Nigericin to induce pyroptosis. For inflammasome inhibition, 10 μ M MCC950 was added at the same time as LPS. A control condition of differentiated but otherwise untreated cells, as well as differentiated and LPS-primed only cells were also included in this experiment, generating four experimental conditions: Untreated, LPS primed,

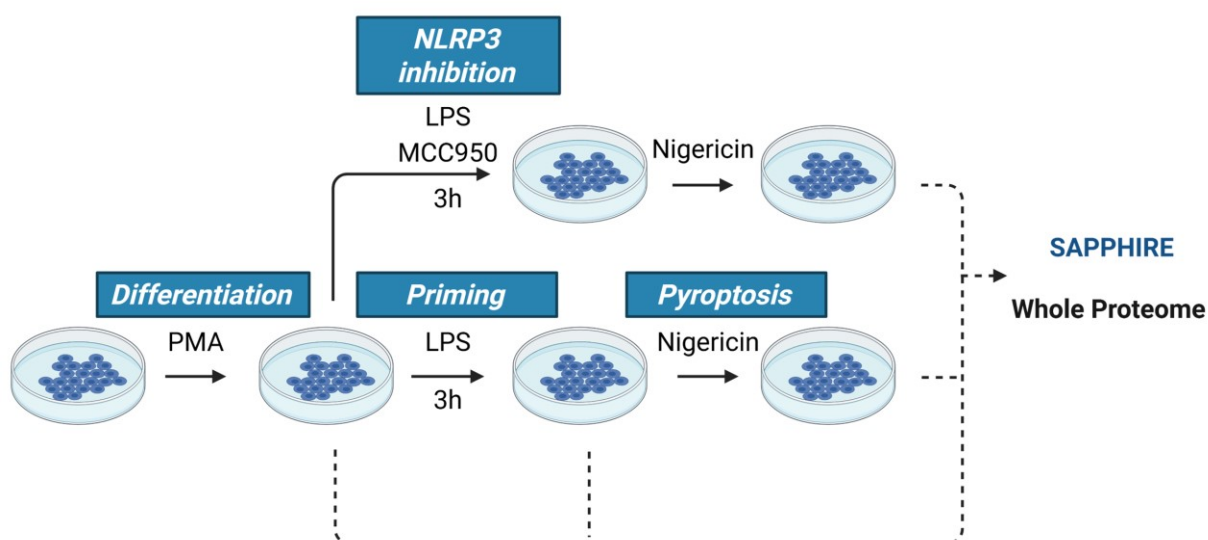


Figure 4.10. Schematic of the experimental conditions for profiling global changes in S-acylation in response to pyroptosis. THP1 cells were differentiated with phorbol-12-myristate-13-acetate (PMA). Then treated 4 conditions: untreated (only control treatments), priming only (only LPS), pyroptosis (LPS + 10 μ M nigericin), pyroptosis + inhibitor (LPS + 10 μ M nigericin + 10 μ M MCC950). LPS priming only, LPS primed and added MCC950 10 μ M at the same time and subsequently treated with 10 μ M Nigericin. Pyroptosis condition after treatment with nigericin. Also not treated but differentiated as untreated. All conditions were n=4 and processed by the SAPPHIRE protocol and submitted for whole proteome analysis. Figure created with Biorender.com.

Pyroptosis (LPS/Nigericin), and 10 μ M MCC950 treatment (LPS/Nigericin with MCC950 pretreatment) (**Fig. 4.10**). All treatments were performed in four biological replicates and subjected to the optimised SAPPHIRE protocol, as well as whole proteome analysis. Notably, the Zr-IMAC enrichment input in this experiment was below 100 μ g peptide. All cell treatments and lysate generation were performed by Dr. Anne Schuhmacher.

SAPPHIRE analysis identified 594 pCys peptides across all conditions. As expected with the optimised protocol, extreme specificity was observed with no pCys peptides identified in all four -HA conditions tested (**Fig. 4.11**). S-acylation of GSDMD Cys191 was markedly upregulated during pyroptosis, compared to both untreated and LPS-primed cells, supporting existing literature (**Fig. 4.12a,b,c**). Additional S-acylated sites enriched in pyroptosis in comparison to both untreated and LPS primed conditions included DExH-Box Helicase 9

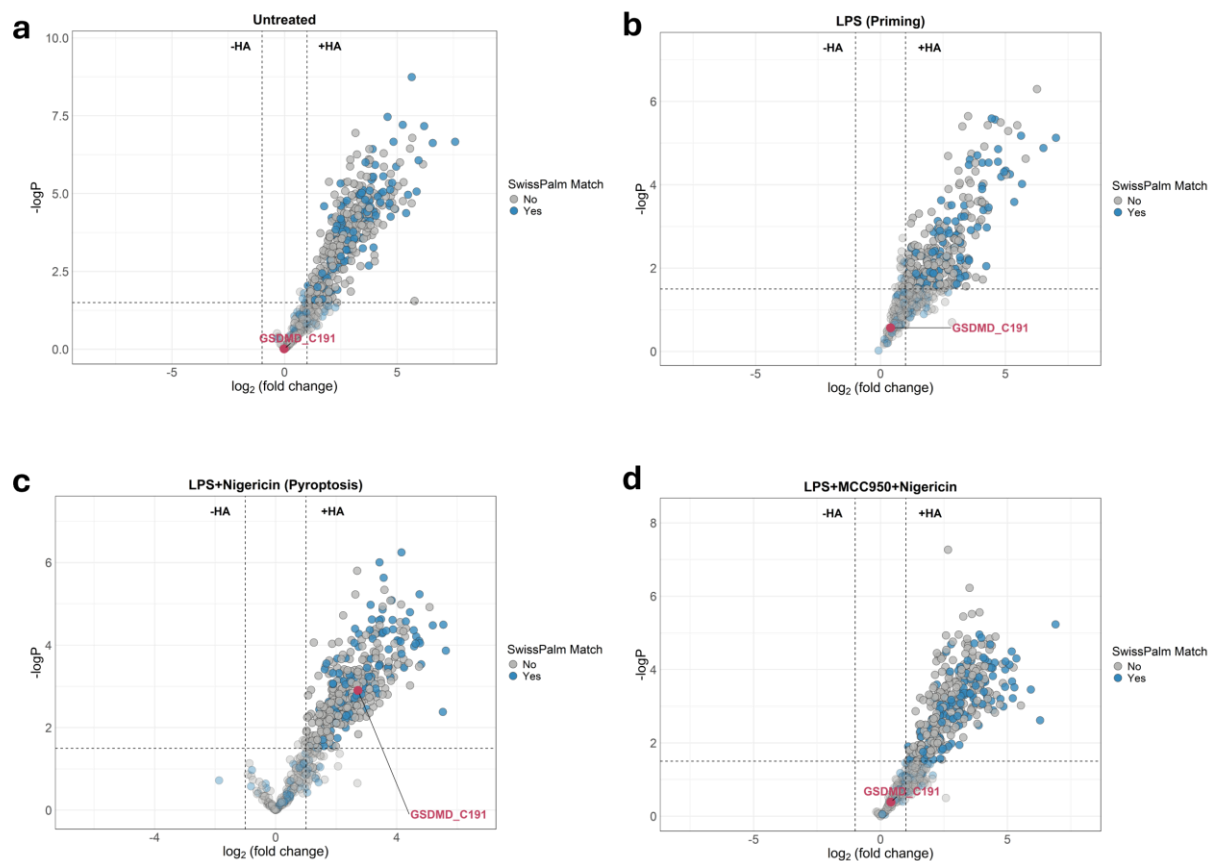


Figure 4.11. Site-specific S-acylated proteome profiling of THP1 cells during pyroptosis induction with SAPPHIRE. Volcano plots depicting +HA-enriched phosphonated Cys peptides (i.e. putative S-acylated Cys) compared to the -HA control for PMA-differentiated THP1 cells that were **a)** untreated **b)** LPS-primed **c)** LPS-primed and treated with Nigericin for pyroptosis induction and **d)** LPS-primed and treated with 10 μ M MCC950 inhibitor (student's t-test; FDR = 0.05, S0 = 0.1) n=4 biological. The blue points denote previous identification of the exact site in the SwissPalm database.

(DHX9) Cys140/242, thioredoxin-related transmembrane protein (TMX1) Cys106 and Acyl carrier protein, mitochondrial (NDUFAB1) Cys140 (**Fig. 4.12b,c**). Interestingly, TMX1 protein has two validated S-acylation sites at Cys206/207, whereas Cys106 had also been previously found by ssMLCC in HeLa cells. Apart from TMX1, all upregulated sites are novel annotations. No significant differences in global S-acylation, including GSDMD, were observed between LPS-primed and untreated cells (**Fig. 4.12a**). However, comparison of FCs between +/-HA for each condition revealed a slightly higher FC in LPS-primed cells compared to untreated cells (**Fig. 4.12d**). This data is consistent with the published studies where LPS priming alone induces a slight increase in GSDMD Cys191 S-acylation, whereas strong S-acylation and cleavage is triggered by LPS+Nigericin treatment^{256,259}.

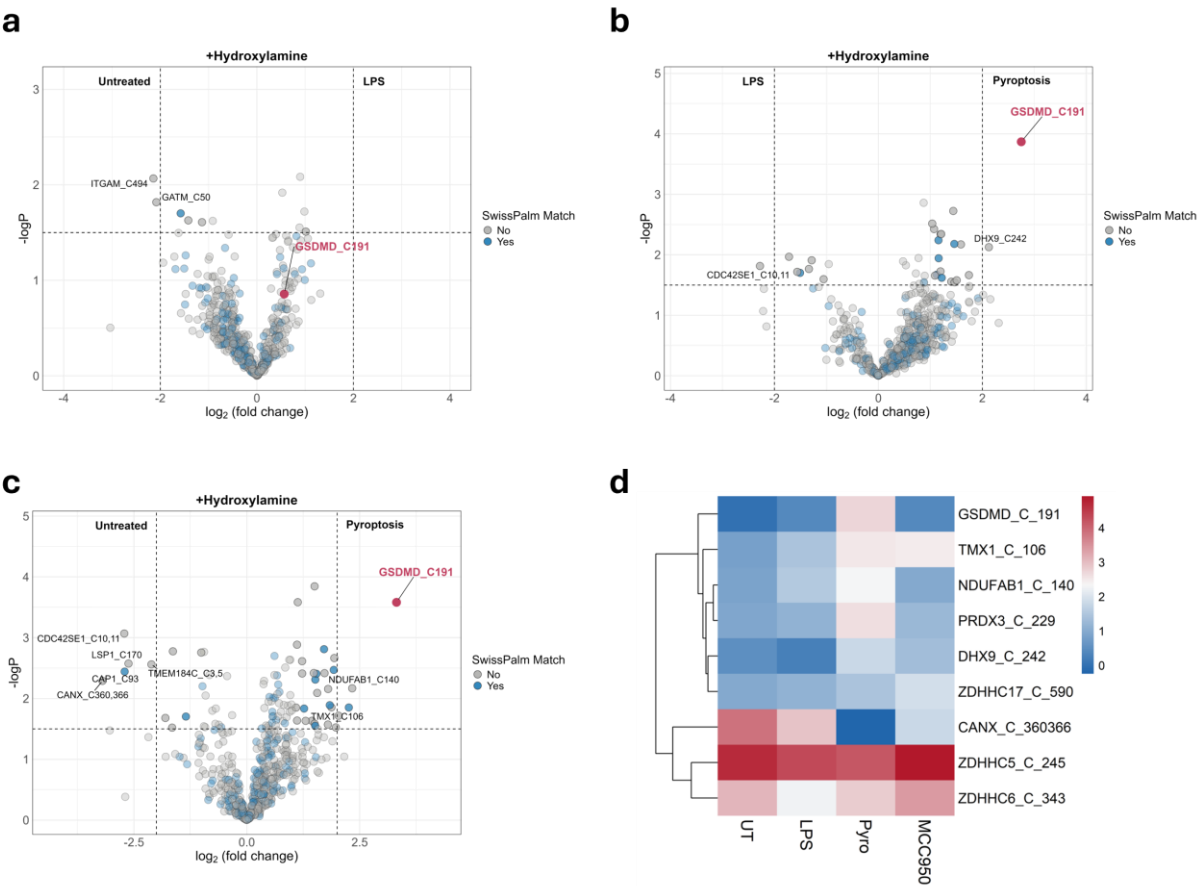


Figure 4.12. GSDMD Cys191 is S-acylated in response to pyroptosis. Volcano plots depicting the differential enrichment of phosphonated cysteine-containing peptides (i.e. putative S-acylated cysteines) in PMA-differentiated THP1 cells for the following comparisons: **a**) LPS-primed vs untreated, **b**) LPS-primed and treated with Nigericin (Pyroptosis) vs LPS-primed, and **c**) Pyroptosis (LPS + Nigericin) vs untreated. Statistical analysis was performed using Student's t-test (FDR = 0.05, $S_0 = 0.1$) $n=4$ biological, and blue points indicate S-acylation sites previously reported in the SwissPalm database. **d**) Heatmap displaying the $\log_2$ fold-change values for individual S-acylation sites, calculated between +HA and -HA samples for each experimental condition.

Furthermore, SAPPHIRE identified three non-catalytic Cys sites belonging to the zDHHC family - zDHHC5 Cys245, zDHHC6 Cys343 and zDHHC17 Cys590 - of which two have been previously reported as S-acylated (**Table 3.2**). Of them, zDHH5 exhibited the highest and constitutive S-acylation in all conditions, followed by zDHHC6 and zDHHC17 (**Fig. 4.12d**). These findings indirectly support the role of zDHHC5 in GSDMD Cys191 S-acylation, however further investigations are required.

Next, analysis of the whole proteome data revealed that GSDMD protein expression remained relatively unchanged in all conditions, in line with other studies (**Fig. 4.13, Appendix Fig. 7.8**). In fact, most changes in S-acylation identified by SAPPHIRE were independent of protein expression. Unfortunately, a calculation error during the preparation of the whole proteome samples led to a ten-fold higher peptide dilution, causing injection of 20 ng rather than 200 ng during LC-MS/MS analysis. As such, a total of 2806 proteins were identified, which is markedly lower than the usual identification rate of human cell samples (8-10k proteins) using the Evosep-timsTOF DIA set up in our proteomics facility. Therefore, the expression of many proteins, including zDHHC enzymes was not detected in this experiment. Future replication of this experiment with correct loading amount would shed light on to which zDHHCs are upregulated in response to both priming and pyroptosis, providing support for GSDMD zDHHC assignment.

Recent studies pointed to the dependency of GSDMD Cys191 S-acylation on Reactive Oxygen Species (ROS), which upregulate specific zDHHCs by inhibiting their proteasomal degradation²⁵⁶. This in turn drives more GSDMD S-acylation, increasing mitochondrial damage and ROS generation in a feed-forward loop mechanism. Interestingly, a dually phosphonated CANX peptide (Cys360/366) exhibited a dramatic loss of S-acylation in pyroptosis, while its protein expression remained unchanged (**Fig. 4.12d, Fig. 4.13**). In contrast to this, S-acylation of TMX1 Cys106 increased in pyroptosis (**Fig. 4.12c,d**). Both CANX and TMX1 are known to localise to mitochondria-associated membranes (MAMs) in an S-acylation-dependent manner, and interact with sarco/endoplasmic reticulum Ca^{2+} ATPase (SERCA)^{263,264}. However, whereas TMX1 inhibits SERCA, decreasing ER calcium uptake and promoting mitochondrial Ca^{2+} transfer and metabolism, calnexin maintains SERCA activity, allowing higher ER Ca^{2+} retention. Moreover, increased TMX1 MAM binding has been shown to compensate for CANX depletion, affecting ER-mitochondria Ca^{2+} flux and ER-mitochondria positioning. Given the crucial roles that these proteins play in mitochondrial metabolism, it is plausible that their

regulation by S-acylation may contribute to this feedforward mechanism and this pyroptotic cascade. Further investigation into this regulation would offer novel insight into this form of cell death.

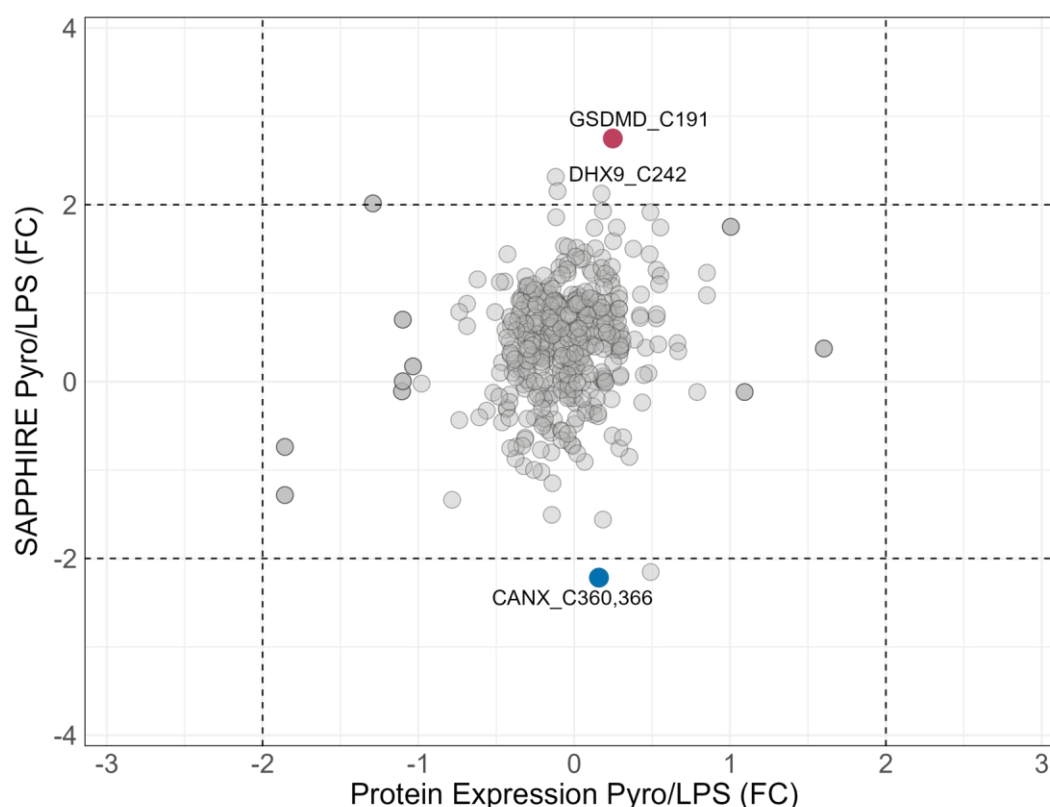


Figure 4.13. GSDMD Cys191 S-acylation during pyroptosis occurs independently of GSDMD expression. Quadrant plot illustrating global changes in S-acylation relative to protein expression between pyroptosis-induced and LPS-primed THP1 cells. GSDMD Cys191 and DHX9 Cys242 show increased S-acylation without corresponding changes in protein abundance, while Calnexin Cys360/366 exhibit a loss of S-acylation despite stable protein expression.

4.4.2 NLRP3 S-acylation

NLRP3 itself is highly regulated by S-acylation. However, similarly to GSDMD, conflicting evidence exists on the precise site and cognate zDHHCs involved in this regulation. Previously, zDHHC5-mediated S-acylation of NLRP3 Cys837, sequential S-acylation of Cys130 and Cys958 by zDHHC1, and Cys130 (murine Cys126) S-acylation by zDHHC7 were demonstrated to promote inflammasome assembly and activation in three independent studies^{265–267}. Conversely, zDHHC12-mediated Cys844 S-acylation prevented inflammasome activation and initiated its degradation in the lysosome⁷⁹. Additionally, as discussed in Chapter 1, the most

recent study demonstrated S-acylation of Cys130 and Cys261 by zDHHC7 induced NLRP3 LLPS, licensing its activation¹⁰⁸. It is not clear whether these discrepancies are due to the differences in experimental design, redundancy between the enzymes or coordinated regulation by multiple zDHHCs. Therefore, insight into the S-acylation patterns of NLRP3 during pyroptosis, but also in response to MCC950 inhibition, would aid in deconvoluting these observations.

Unfortunately, neither SAPPHIRE nor whole proteome analyses were able to detect any NLRP3 peptides, despite the many well characterised S-acylated sites. This is somewhat unsurprising as despite the predicted proteotypic tryptic peptides protein coverage of up to 90%, NLRP3 remains infamously challenging to identify by LC-MS/MS based proteomics for unknown reasons²⁶⁸. Further attempts with higher enrichment input, as well as the use of alternative proteases and quantification methods, could enable NLRP3 inhibition. Such optimisation may be necessary to enable global profiling of future inhibitors targeting NLRP3 S-acylation sites but are outside the scope of this PhD.

4.4.3 MCC950 blocks GSDMD palmitoylation

MCC950, a diaryl sulfonylurea-containing compound, is a potent and selective inhibitor of the NLRP3 inflammasome (**Fig. 4.14a**). Direct binding of MCC950 to NLRP3 prevents its active conformation and subsequent downstream events²⁶⁹. SAPPHIRE analysis revealed that 10 μ M MCC950 treatment prior to Nigericin, almost completely blocked GSDMD Cys191 S-acylation, to levels matching those observed in LPS-primed cells (**Fig. 4.12d, Fig. 4.14b,c,d**). These results are in line with the current model, whereby only after the full activation of the inflammasome does GSDMD get S-acylated at high levels. Moreover, no significant changes were observed between MCC950 treated, untreated and LPS-primed cells (**Fig. 4.14b,c**).

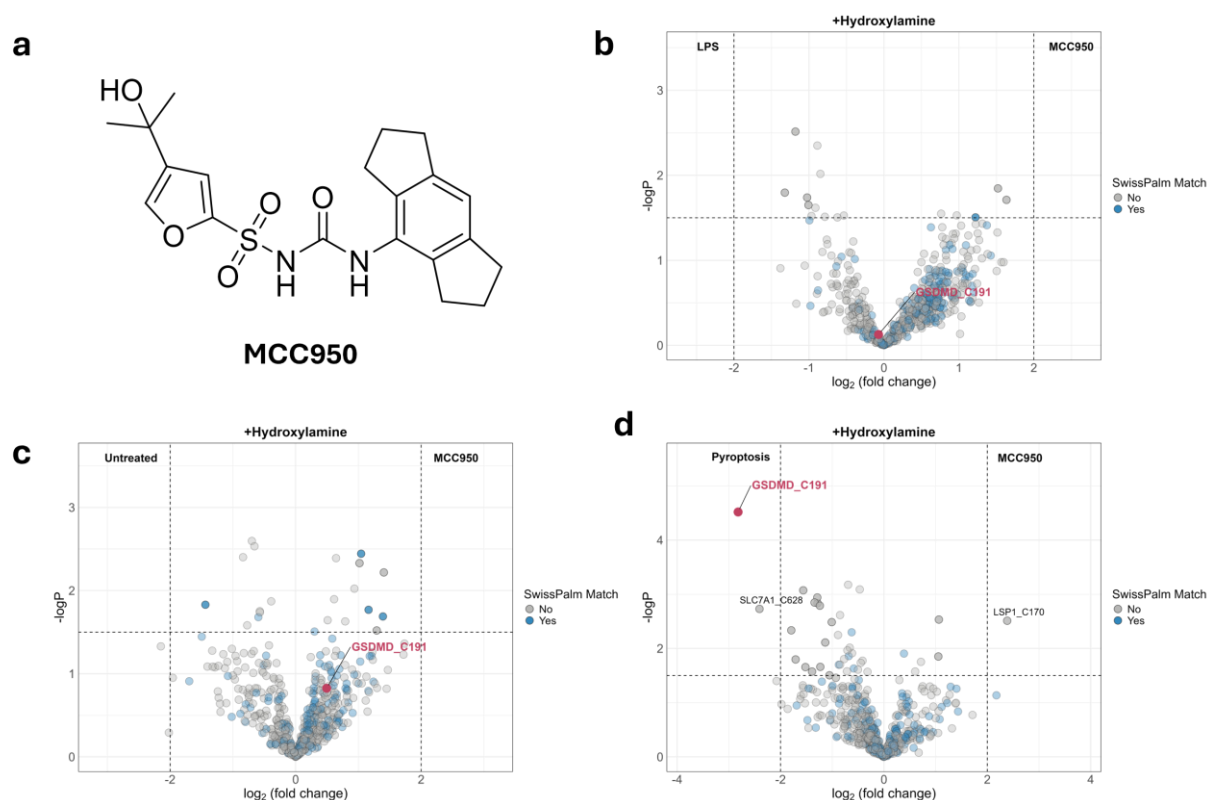


Figure 4.14. MCC950 blocks S-acylation of GSDMD Cys191. **a)** Chemical structure of MCC950, a selective NLRP3 inflammasome inhibitor. **b–d)** Volcano plots showing differential enrichment of phosphonated cysteine-containing peptides (i.e. putative S-acylated cysteines) in PMA-differentiated THP1 cells for the following comparisons: **b)** MCC950-treated vs LPS-primed, **c)** MCC950-treated vs untreated, and **d)** MCC950-treated vs pyroptosis-induced. Statistical analysis was performed using Student's t-test (FDR = 0.05, $S_0 = 0.1$) $n=4$ biological. MCC950 treatment refers to co-treatment of THP1 cells with LPS and 10 μ M MCC950, followed by addition of 10 μ M Nigericin after 3 hours.

4.4.4 Conclusions

Overall, SAPPHIRE enabled the first analysis of endogenous S-acylation on a proteome-wide scale during priming and pyroptosis. The precise identification of pyroptosis-induced S-acylation of GSDMD, further underscores sensitivity of the SAPPHIRE workflow for catching dynamic and biologically significant events. Moreover, this experiment constitutes the first report of the effects of the MCC950 inflammasome inhibitor for blockage of downstream GSDMD S-acylation, providing a proof-of-concept for the use of SAPPHIRE for profiling changes in global S-acylation patterns in response to pharmacological agents. However, one of the limitations of this experiment is the absence of quantitative data for pyroptosis and inflammation. Even though THP1 cells were treated with LPS+Nigericin according to

established protocols and cell death was observed by eye in cell culture, this experiment would benefit from additional assays such as lactate dehydrogenase and IL-1B release to confirm the respective pyroptosis and inflammation were induced.

Together, these findings underscore the capability of SAPPHIRE for detecting biologically and therapeutically relevant dynamic events, opening interesting avenues to explore. For example, SAPPHIRE profiling during canonical and non-canonical inflammasome activation may shed light on some gaps in the current model. Moreover, as APT2 has been shown to deacylate GSDMD to promote pyroptosis, analysis of ML349, an APT2-specific inhibitor, treated cells, would aid in characterisation of this model.

4.5 Stress Response

The latter part of this PhD was primarily centred around the development, optimisation and application of the SAPPHIRE approach. Thus, further optimisation to the APEX-AX or establishment of other PL-based double tagging approaches fell outside the timeline of this project. Nevertheless, although the initial objective, to profile stress-dependent S-acylation changes in SGs *in situ*, remains to be explored in future studies, a global S-acylation analysis in response to NaAsO₂ stress was conducted in HEK293T cells to partially address this aim.

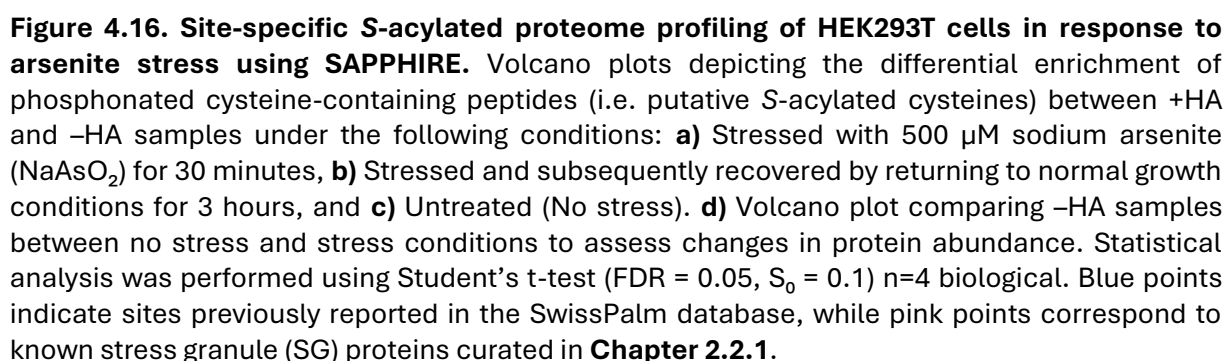
Based on our hypothesis, certain SG proteins would undergo dynamic S-acylation in response to stress, facilitating nucleation and/or formation of SGs. These changes can still be captured by SAPPHIRE, without the requirement of a PL step. As such, HEK293T cells were either untreated (control), stressed with 500 μ M NaAsO₂ for 30 minutes, or stressed and subsequently allowed to recover by removing NaAsO₂ containing media, washing the cells with PBS and allowing them to recover in fresh media for 3 hours (n=4 biological replicates per condition) (**Fig. 4.15a**). These samples were then processed with SAPPHIRE, utilising double NEM alkylation. It was anticipated that in the recovery condition, S-acylation patterns would mostly resemble those in the unstressed condition.

Cys112/120, with all four of these sites reported previously using ssMLCC and acyl-RAC approaches^{195,270}. Interestingly, IGF2BP1 was found to be S-acylated on Cys336/Cys337. IGF2BP1 is a well-established SG component and was previously identified as a stress-dependent G3BP1 interactor (**Chapter 2.6.2**). A previous study has shown that IGF2BP1 is S-acylated on Cys337 by zDHC1 using mutational gel-based validation experiments, functioning in RNA stability²³⁹. Thus, it is possible that IGF2BP1 is S-acylated in response to stress, promoting its association with SGs. However, further experiments are needed to probe this hypothesis.

Additionally, 16 Cys of RanBP2, also known as NUP358, were identified. RanBP2, a nucleoporin and key component of the nuclear pore complex (NPC), also functions as an E3 SUMO ligase with SUMOylating activity previously implicated in the disassembly of SGs¹⁴⁷. Another nucleoporin, NUP153, was identified as S-acylated on Cys678/681/871/874. Although the role of S-acylation in nucleoporin biology remains largely unexplored, it is unlikely that a single protein is S-acylated at 16 distinct Cys. These identifications may have arisen from non-S-acylation events described in **Chapter 3.7.3**. At the same time, incomplete blocking observed in this experiment, as discussed below, may have contributed to these observations.

4.5.1 Limitations

A major limitation of this experiment was the unexpected identification of many pCys in the -HA condition (**Fig 4.16a,b,c**). Blocking conditions remained unchanged (in regard to time and concentration) from previous experiments where no peptides were observed in the -HA condition. However, one significant difference was that this experiment was performed entirely in a 96-well plate format (**Table 6.2**). PCA revealed that all -HA control samples clustered tightly together, whereas the +HA samples showed greater variability depending on the condition (**Fig. 4.17**). Moreover, comparison of the -HA fractions of unstressed and stressed cells revealed non-significant changes, compared to the +HA counterparts (**Fig. 4.16d**). This suggests that the same Cys are being labelled in the -HA controls, reducing the likelihood of random contamination between +/-HA samples. Although this possibility still exists, if cross-contamination had occurred due to incomplete sealing of the plate or other practical reasons, the control replicates would exhibit higher variations and closer clustering with the HA samples. One possibility is that the low shaking speed used for all plate-based



131

background observed in the -HA controls, much caution should be given to any results observed.

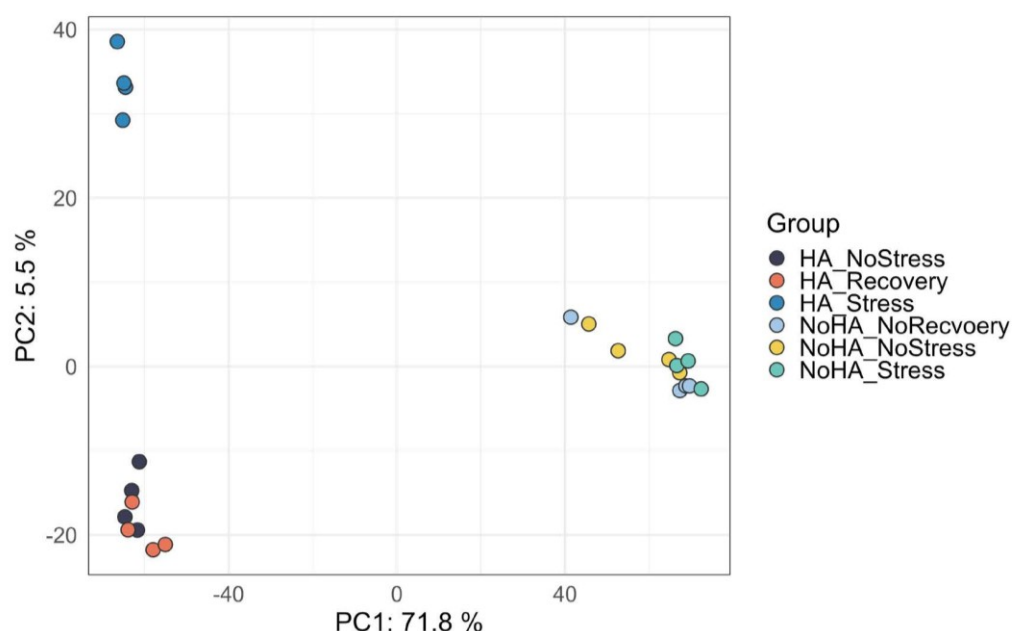


Figure 4.17. Principal component analysis (PCA) of HEK293T SAPPHERE profiling in response to stress. PCA shows clear separation between +HA and -HA samples across all conditions, indicating distinct S-acylation profiles. Among the +HA samples, the No stress and Recovery conditions cluster closely, while the Stress condition forms a distinct cluster, suggesting dynamic and stress-specific changes in S-acylation.

4.5.2 Conclusions

Despite these limitations, this is the first dataset providing insight into global S-acylation changes in response to arsenite-induced stress. Although the -HA control anomalies require cautious interpretation, the identification of SG proteins, such as PTGES3, RAN and IGF2BP1, supports our hypothesis of stress-responsive S-acylation. However, repetition of this experiment without background CysPAT labelling is necessary to corroborate these findings. Furthermore, given the global nature of SAPPHERE analysis and pleiotropic cellular effects induced by NaAsO₂ treatment, interpretation remains complex. It is important to note that identification of an SG protein with SAPPHERE, does not necessarily mean that the observed change in S-acylation is related to its SG composition, during SG formation or its SG localisation.

Given the technical challenges experienced with APEX-AX and double enrichment proteomics in general, global profiling demonstrated here may be a useful tool for hypothesis generation. However, targeted validation of specific hits for their S-acylation status via site-directed mutagenesis approaches and further interrogation of SG dynamics by microscopy approaches, such as FRAP and live-cell imaging, will be necessary to draw definitive conclusions. Without these experiments, global stress changes remain an indirect interrogation of SGs.

4.6 Discussion

4.6.1 Site discrepancy and enzyme redundancy

Throughout this chapter, a recurring observation for all experiments is site discrepancy, that is the identification of some Cys with SAPPHIRE that belong to known S-acylated proteins but differ from previously validated sites. This is particularly evident for calnexin, as this is the most commonly used positive control for gel-based S-acylation experiments. Thus far, CANX Cys360/366 sites were identified as phosphonate modified with SAPPHIRE in HEK293T cells (**Chapter 3.4.1**), pyroptosis experiment THP1 cells and stress response HEK293T cells, as well as enriched as CAA-modified in our and literature ssABE profiling (**Chapter 3.6.1**). This consistent identification of the same sites suggests that HA-dependent modification of these residues is not random. Given that the validated residues are the transmembrane Cys502/503, it is possible that upon lysis, specific conformational changes are induced whereby the two S-acyl chains are transferred to the 360/366 residues. As discussed in **Chapter 3.7.3**, these S→S acyl transfers are theoretically possible, but have not been investigated to date. Alternatively, these sites may represent S-acetylation or transient auto long-chain acylation. However, as discussed throughout this thesis, no tools exist that would allow to differentiate between these events.

This phenomenon of site discrepancy is commonly observed between proteomics-based studies and techniques, as discussed in **Chapter 3.7.3**. This chapter has additionally highlighted the inconsistencies between the site and zDHHC assignment in key pyroptosis proteins – NLRP3 and GSDMD. In addition to pyroptosis, similar discrepancies have emerged in studies on another form of cell death - ferroptosis. Ferroptosis is an iron-dependent programmed cell death, characterised by lipid peroxidation, and is a promising target for

improving cancer immunotherapy¹⁰². Recent studies revealed that glutathione peroxidase 4 (GPX4), a key enzyme in ferroptosis, is regulated by S-acylation. However, whilst one study reported zDHHC20-mediated S-acylation of Cys66, another study identified Cys75 S-acylation by zDHHC8, despite both studies employing GPX4 point mutation screens^{93,102}. Notably, both Cys66 and Cys75 were identified in SAPPHERE profiling of HEK293T cells, with no other GPX4 Cys identified. Moreover, another ferroptosis regulator upstream of GPX4, SLC7A11, has also been reported to be S-acylated by two independent studies at differing residues of Cys327 and Cys414, using point mutation and ABE or MLCC gel-based approaches^{271,272}.

These four cases of targeted investigations of GSDMD, NLRP3, GPX4 and SLC7A11 underscore the profound challenges in the field of S-acylation. Even with the use of identical techniques, cell lines and point mutants, different sites and zDHHCs can be assigned. Several factors could contribute to the zDHHC assignment variability. The inherent redundancy of the 23-member zDHHC family may lead to differential compensation when using KD approaches and therefore inconsistent assignments¹⁵². Furthermore, it is possible that several zDHHC enzymes may regulate the same or different sites on the same protein in a coordinated manner. This regulation may be both time and context-dependent, i.e. can occur simultaneously or sequentially, acting as either positive or negative regulators. However, these hypotheses remain largely unexplored.

Whilst the zDHHC enzyme redundancy is widely acknowledged in the field, the possibility of site redundancy has not been investigated previously. Given the dynamic and reversible nature of S-acylation, it is possible that some proteins with multiple accessible Cys may undergo S-acylation at alternate or multiple sites to achieve the same molecular outcome, i.e. localisation to the membrane. This hypothesis would explain the observed site discrepancy in both proteomic and biochemically targeted studies, and the higher overlap on the protein but not site level between studies (**Fig. 3.15**). Moreover, given the high sensitivity of SAPPHERE observed throughout this chapter, the identification of more transient and low stoichiometry sites may be increased compared to other methods. Nonetheless, GSDMD possesses nine Cys in its FL form and three Cys (Cys38, Cys56, Cys191) in its NT- form²⁷³. The precise identification of GSDMD Cys191 in a pyroptosis-induced condition only, highlights the ability of SAPPHERE for dynamic site identification.

4.6.2 Site summary

Collectively, application of SAPPHIRE across diverse cell types (HEK293T, THP1), tissues (breast and CRC tumours), and organisms (*H.sapiens*, *T. gondii*) has highlighted both the strengths and limitations of this method, as well as areas for further optimisation. One of the biggest advantages observed was the exceptional tolerance of the enrichment for low peptide input. As such, with further optimisations, cell culture and treatments could be adapted to 48 and 96-well plate formats, facilitating high scalability for the entire workflow. Furthermore, the sensitivity of SAPPHIRE for capturing dynamic S-acylation in response to cellular stimulus has been demonstrated by the identification of pyroptosis induced S-acylation of GSDMD Cys191. However, as shown in patient tissue and stress response experiments, further optimisations to the blocking conditions for tissue- and plate-based experiments will be necessary to minimise false positives. Another important area for optimisation is within the data analysis pipeline. All experiments presented in this chapter, apart from pyroptosis profiling, were “open-ended” in nature, that is profiling S-acylation without focusing on a specific protein or having a clear positive control. Given the observed site discrepancy and results of the pyroptosis experiment (**Fig. 4.12**), a more stringent fold change cut-off (>2) may be adapted for stimulus-driven S-acylation experiments. Furthermore, experiments proposed in **Chapter 3.7.3**, such as SAPPHIRE profiling in membrane vs soluble fractions, may aid in establishing criteria or curating a SAPPHIRE specific contaminants list.

Cumulatively for humans, five comprehensive site-specific S-acylated proteome profiling was performed using SAPPHIRE, +/- HA in HEK293T profiling, healthy vs tumour in breast cancer and CRC, pyroptosis in THP1, and stress response in HEK293T. Given the challenges observed with the tumour tissues analysis, sites identified in the three human cell line experiments were curated together for a total of 4271 unique putative S-acylated sites across 2404 proteins. Sites were classified as S-acylated if they have a $\log_2FC > 1$ and $-\log P > 1.5$. Of these, 139 sites were associated with the ubiquitin pathway proteins and filtered out of the database, resulting in a total of 4132 sites. Furthermore, given the non-specific labelling observed in the stress response experiment, any sites that was only identified in that experiment is classified as low confidence, whereas any other sites are classified as high confidence for a total of 3146 high confidence and 986 low confidence hits. Each site is additionally cross referenced against the newest release of the SwissPalm database - Release 5 (09/04/2025) that was released during

the preparation of this PhD thesis. This annotated site database is available to download electronically at

https://www.dropbox.com/scl/fi/u949yq546pi8ilx40va5t/Combined_Binary_Acylation_Matrix_Final.tsv?rlkey=rmxto9b3r26l28pf0n0smp96s&st=xqk6lfda&dl=0).

4.7 Chapter summary

This chapter demonstrates the application of the SAPPHIRE method, developed in Chapter 3, for site-specific profiling of S-acylated proteome across various sample types and experimental models. Notably, SAPPHIRE enabled several pioneering analyses, including the first-ever profiling of S-acylation patterns in human patient samples, site-specific profiling in *T. gondii* parasites, and global profiling of THP1 cells undergoing pyroptosis, supporting existing findings on the critical role of GSDMD Cys191 S-acylation. Additionally, SAPPHIRE effectively captured global S-acylation dynamics in response to NaAsO₂ stress in HEK293T cells, revealing increased S-acylation under stress conditions, in part addressing the initial hypotheses of this PhD

This chapter highlights the method's versatility, especially its tolerance for low peptide input and ability to capture dynamic site-specific events. However, it also identifies key limitations, including challenges related to the identification of many putative S-acylated Cys on a single protein, often alternative to the validated site, a phenomenon of site discrepancy. While these discrepancies could partly result from the identifications of other hydroxylamine-sensitive events, their frequent occurrence in mutant-based targeted approaches in other published studies remains unexplained. SAPPHIRE, together with existing literature, shed light on the profound challenges in studies S-acylation. Furthermore, practical challenges such as incomplete free Cys blocking in a 96-well format, highlights the need for further optimisation.

Overall, this chapter illustrates the potential of SAPPHIRE as a powerful, semi-automatable, high-throughput tool for dynamic and sensitive profiling of S-acylation, while acknowledging areas requiring further optimisation.

Chapter 5.

Conclusions and Future Perspectives

S-acylation is a crucial PTM in eukaryotes, involved in numerous cellular processes and frequently dysregulated in diseases. This thesis initially aimed to investigate the role of S-acylation in biomolecular condensates, specifically SGs, and subsequently developed and optimised a novel proteomics method – SAPPHIRE - to address existing methodological limitations. Initial experiments using APEX-AX revealed its inherent incompatibility for investigating S-acylation, primarily due to the destabilisation of the labile S-acyl thioester bond by the H₂O₂ treatment required for the activity of the APEX2 enzyme. Crucially, the development of APEX-AX underscored the critical need for streamlined, site-specific S-acylproteomic approaches. Recognising these gaps, the development of the SAPPHIRE methodology was pursued.

The SAPPHIRE work presented in this thesis highlighted significant technical challenges inherent to S-acylproteomics, necessitating extensive optimisation at each step of the workflow. Despite these challenges, SAPPHIRE successfully identified over 4000 putative S-acylation sites across multiple human cell lines, making it the most sensitive and high-throughput S-acylproteomic platform currently available. However, the high sensitivity of SAPPHIRE also revealed critical limitations common to cysteine-centric workflows, notably the potential identification of non-S-acylation hydroxylamine-sensitive events such as S-acetylation and S-ubiquitination. Presently, methods to differentiate these events from bona fide S-acylation are lacking, and validation approaches, like the gel-based Acyl-PEG Exchange (APE) assay, are complicated by the varied stoichiometry of the identified sites. Additionally, the frequent observation of site discrepancy phenomena in both proteomic and targeted mutant-based approaches (e.g., GPX4, NLRP3) in published studies calls for further investigation, particularly regarding potential lipid S→S transfer events as discussed in **Chapter 3.7.3**.

While SAPPHIRE represents an important advancement, its underlying chemistry relies on hydroxylamine treatment, which is an indirect method for identifying S-acylation. Therefore, progress in the S-acylproteomics field will greatly benefit from the development of streamlined intact S-acylation identification methods and more specific inhibitors for zDHHC enzymes to distinguish true substrates from false positives. Although methods like nGF-SPE allow intact S-acylation detection, their high protein input requirements and reproducibility issues significantly limit widespread adoption in the field.

Overall, this thesis underscores the complexity of studying S-acylation and highlights the urgent need for enhanced analytical methods and tools. Despite current challenges, SAPPHIRE has generated datasets of substantial biological interest and serves as a powerful tool in the field. Future directions must prioritise developing robust, intact S-acylation identification techniques and specific zDHHC enzyme inhibitors to advance the utility and reliability of S-acylproteomic approaches. Although ongoing efforts toward these developments continue, their timelines remain uncertain. Consequently, utilising the current methodological toolkit to curate datasets of potential false positives, for instance through replicating the recently described S-acetylation workflows or employing membrane fractionation strategies, will be critical. Additionally, further optimisation of the 96-well plate-based workflow, particularly regarding cysteine blocking efficiency, will significantly enhance the robustness of the SAPPHIRE approach.

Chapter 6.

Materials and Methods

6.1 General Procedures

6.1.1 Cell culture

Human Bone Osteosarcoma Epithelial Cells U2OS lacking G3BP1/G3BP2 and stably expressing G3BP1-APEX2 (U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1) under a tetracycline-inducible promoter and U2OS cells stably expressing APEX2-NES (U2OS-APEX2-NES), were a kind gift from the Eran Hornstein lab at Weizmann Institute of Science. Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific) supplemented with 10% foetal bovine serum (FBS), at 37 °C, with 5% CO₂.

Human embryonic kidney (HEK293T) and human myeloid leukaemia mononuclear Tohoku Hospital Pediatrics-1 (THP1) cells were acquired through the Francis Crick Institute after STR (short tandem repeat) profile authentication. HEK293T cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS, at 37°C, with 5% CO₂. THP1 cells were maintained in RPMI1640 supplemented with 1 mM sodium pyruvate, GlutaMAX and 10% FBS and maintained at 0.1 – 1.0 Mio cells / mL. THP1 cell culture was performed by Dr. Anne Schuhmacher.

Human foreskin fibroblasts (HFF) were obtained from ATCC® (ATCC SCRC-1041). HFF cells were grown in DMEM supplemented with 10% FBS, and 2 mM L-glutamine. Cell culture was maintained at 37 °C and 5% CO₂. Fully confluent cells were washed with pre-warmed PBS prior to the addition of trypsin/EDTA. These were then allowed to detach during a 10-minute incubation at 37 °C and 5% CO₂, after which fully supplemented media was added, and the culture split into new culture vessels. *T. gondii* tachyzoites were grown on a monolayer of HFF cells under the same conditions until complete lysis of the monolayer, after which the parasites were transferred onto a new fresh monolayer for maintenance. HFF and *T. gondii* cell culture was performed by Dr. Janessa Grech.

6.1.2 Acetonitrile (MeCN)-based protein precipitation

For 1 volume of sample, proteins were precipitated by the addition of 4 volumes MeCN and gently vortexed (400 rpm) for 5 s. Samples were centrifuged for 5 min at 1000-4000 × g at RT to obtain a protein pellet. The supernatant was then removed, and the pellet was washed with 80% (v/v) EtOH >5x original sample volume without disturbing the pellet. Samples were then

centrifuged ($3,220-17,000 \times g$, 4 min, RT) and supernatants removed. The 80% EtOH wash was repeated for a total of three washes. After the final wash, samples were resuspended in the appropriate buffer or topped off with trypsin digestion buffer depending on downstream application. The pellets were then resolubilised either using a water bath sonicator (low-input, 1.5 mL Eppendorf Protein LoBind tube-based samples) or a Fisherbrand 505 probe sonicator (Fisher Scientific) when appropriate (high-input, 15 mL Protein LoBind Falcon tube-based samples). For probe sonication, samples were placed in a beaker containing ice to prevent heat-mediated protein degradation, and sonicated using 5 second on/off pulses at 20% amplitude for 2 min. Probe was washed with MeCN between each sample.

6.1.3 MeCN-based protein precipitation: 96-well plate format

Plate-based protein precipitation was performed as described above with slight modifications. For 1 volume of sample in 96-well 2mL Protein LoBind plates, 4 volumes MeCN were added, and the plate was sealed with Aluminium StarSeal (PCR) plate seal (Starlab E2796-9792) and gently vortexed (400 rpm) for 5 min at RT. Samples were then centrifuged for 10 min at $3,220 \times g$ at RT to obtain a protein pellet. To remove the supernatant, the plate was quickly inverted upside down over a waste container and put down on lint-free Kimtech wipes (Kimberly-Clark #7558) for 1-2 min. The pellets were then washed with 80% (v/v) EtOH $>5\times$ original sample volume without disturbing the pellet, and centrifuged ($3,220 \times g$, 5 min, RT). The supernatant was removed as described above and the 80% EtOH wash step was repeated for a total of three washes. After the final wash, the plate was left upside down on the Kimtech wipe for 5 min before addition of the appropriate buffer. For samples were resuspended in the appropriate buffer or topped off with trypsin digestion buffer depending on downstream application.

6.1.4 SDS-PAGE

Samples were prepared using 4x Laemmli sample buffer (BioRad 1610747) supplemented with 10% β -mercaptoethanol, boiled for 5 min at 95°C , allowed to cool to RT and 10-14 μL (10-20 μg protein) loaded onto or 4-15% (w/v) SDS-PAGE pre-cast gels (BioRad 4561086DC). Proteins were separated at 100 V for 30 mins and 150V for 0.5-1 h in Tris-glycine-SDS buffer (Bio-Rad, 1610732).

6.1.5 Western blot

After SDS-PAGE, proteins were transferred to nitrocellulose membranes (0.45 μ m, Amersham Protran, GE Healthcare) using the “mixed molecular weight” protocol (25 V, 2.5 A, 7 min) on the Trans-Blot Turbo system (Bio-Rad, 1704150). Following transfer, membranes were stained with Ponceau S (Sigma, P7170) for 10 min to assess transfer efficiency. Membranes were washed using Tris-buffer (50 mM Tris pH 7.4, 150 mM NaCl) supplemented with 1.0% Tween-20 (TBST). Prior to Western blot with the primary antibody, membranes were blocked with 5% bovine serum albumin (BSA) (Sigma-Aldrich) or 5% semi-skimmed milk (Sigma-Aldrich) in TBST for 1 h at RT or at 4°C overnight. After blocking, membranes were incubated with primary antibodies in blocking solution for 1-2 h at RT and subsequently washed three times (10 min each) with TBST. Membranes were then incubated with secondary antibodies in blocking solution for 1 h at RT and washed three times (10 min each) in TBST and imaged. For all Western blots in **Chapter 2**, HRP-conjugated secondary antibodies for used. As such, membranes were incubated with Immobilon Crescendo Western HRP substrate (Merck, WBLUR0100) and chemiluminescence imaged using an ImageQuant LAS4000 (GE Healthcare). For all Western Blots in **Chapter 3**, fluorophore-conjugated secondary antibodies were used and were imaged on LI-COR Odyssey (LiCor). Images were processed using Image J (NIH Bethesda) and Empiria Studio software (LiCor). A list of all antibodies used and their dilutions is given in **Table 6.1**.

6.1.6 Patient tissue preparation

Patient samples were provided by existing Imperial BioBank collaborators. Patient tissues were thawed on ice and transferred to 2 mL BeadBug™ tubes prefilled with 1.5 mm zirconium beads (Merck, Z763799) and weighed by difference. To each sample, 4 μ l of lysis buffer (PBS + 0.1% NP40) for every 1 mg of sample was added, or a minimum of 350 μ l. The samples were then homogenised using FastPrep 24 5GS (MP Biomedicals) with the breast tissue sample processing program. Following homogenisation, the lysates were transferred to 1.5 mL Protein LoBind Eppendorf tubes. To maximise protein recovery, the beads were additionally washed with 100 μ l lysis buffer, and the supernatant was combined with the original lysates. Samples were subsequently centrifuged (17 000 x g for 5 min at 4 °C) and the supernatant was transferred using an insulin needle and syringe to the upper chamber of a 0.22 μ m spin filter centrifuge tubes (Corning, #8161) taking care not to disturb the lipid layer. Samples were

centrifuged (16 000 x g for 5 min at 4 °C), and the resulting flow through was collected and taken forward for the SAPPHERE protocol. All tissue handling and homogenisation was performed by Dr. Aidan Pidd.

6.2 Cell Treatments

6.2.1 Pyroptosis

4 conditions: untreated (only control treatments), priming only (only LPS), pyroptosis (LPS + nigericin), pyroptosis + inhibitor (LPS + nigericin + MCC950) (**Fig. 4.11**). Monocyte differentiation: 35 Mio cells are seeded in a 15 cm dish at 1 Mio/mL concentration and treated with 100 nM PMA for 48 hours. The media was replaced with normal growth media and the cells were incubated for 24 hours. Cells were treated with 250 ng/mL LPS or PBS as control and 10 µM MCC950 (pyroptosis inhibitor) or DMSO as control and incubated for 3 hours. 10 µM nigericin or methanol as control was added and the cells were incubated for 1 hour. Cells were washed twice with 30 mL ice-cold PBS, dislodged in 5 mL ice-cold PBS with a cell scraper and transferred to a falcon tube. Cells were pelleted at 200 xg for 5 min, PBS was removed and 2 mL lysis buffer (4% SDS, 150 mM NaCl in 50 mM TEA, pH 7.3) supplemented with Benzonase and Roche EDTA-free protease inhibitor was added. The lysates were incubated at 37 °C for 30 mins then EDTA was added to a final concentration of 5 mM. The protein concentration was determined using the Biorad DC assay according to manufacturer instructions. The protein concentration was adjusted to 2 mg/mL using lysis buffer.

6.2.2 *Toxoplasma gondii*

Parasites were used to infect and replicate in fresh HFFs for a period of 36-42 hours, after which the parasites were egressed manually and filtered using 3 µm filters on ice. The egressed parasites were centrifuged at 1500 x g for 5 minutes at 4°C and washed with cold PBS. The washing step was repeated thrice on ice, after which the parasites were pelleted, snap-frozen, and stored at -80°C until use.

6.2.3 Stress Response

To trigger SG formation, a solution of 500 µM NaAsO₂ in complete DMEM was applied to cells for 30 min. For the recovery conditions, cells were rinsed three times with PBS to remove the

stressor and fresh and pre-warmed DMEM was added and then left to recover for three hours at 37 °C, 5% CO₂.

6.3 SAPPHIRE Proteomics

6.3.1 Buffers

Buffer A - 4% (wt/vol) SDS 4% (wt/vol) SDS (Sigma Aldrich) in triethanolamine (TEA) buffer (50 mM TEA, 150 mM NaCl, 5 mM EDTA, pH 7.3)

Buffer B - 0.2% Triton X-100 in 50 mM TEA, 150 mM NaCl pH 7.3

6.3.2 Cell Lysis

Cells were lysed with 4% (wt/vol) SDS (Sigma Aldrich) in TEA buffer (50 mM TEA, 150 mM NaCl, pH 7.3) supplemented with 1,500 units/mL benzonase (Millipore) and x1 EDTA-free cOmplete protease inhibitor cocktail (Roche). Brief sonication was sometimes used to reduce viscosity. When a clear solution was obtained, 0.5 M ethylenediaminetetraacetic acid (EDTA) (Sigma Aldrich) was added for a final concentration of 5 mM. Protein concentration was determined using the DC protein assay (Bio-Rad) according to manufacturer instructions, and all samples were normalised to 2-4 mg/mL per sample, depending on the experiment. Each lysate was split into four technical replicates.

6.3.3 Free Cysteine alkylation and acyl-phosphonate exchange

Samples were treated with 10 mM neutralised TCEP and 25 mM of either NEM or CAA (as specified for each experiment in **Table 6.2**) for 1.5 h at RT with nutation. All samples were then precipitated using MeCN-based protein precipitation. The resulting protein pellets were resuspended in Buffer A at their original volume and subjected to a second alkylation step under the same conditions (10 mM neutralised TCEP and 25 mM either NEM or CAA for 1.5 h at RT with nutation). Samples were precipitated again using MeCN-based precipitation and the resulting protein pellets were resuspended in ¼ of the original volume in buffer A and split into two aliquots. One aliquots was incubated with 0.75 M HA (50% by weight in H₂O, Sigma-Aldrich, 438227-50ML, x3 volumes of 1 M HA, diluted to 1 M in buffer B) and 100 µM 6C-CysPAT (ThermoFisher, A52285) in 100 mM triethylammonium bicarbonate (TEAB) buffer, to facilitate S-acyl cleavage and phosphonate labelling of the newly revealed Cys. The second aliquot

served as the negative control, where x3 volumes of the aliquot's original volume of Buffer B supplemented with 100 μ M CysPAT were added. The mixtures were thoroughly vortexed and incubated for 1 h at RT in the dark with mixing by nutation. Subsequently, samples were then precipitated using MeCN-based precipitation. After the final wash, sequence grade trypsin (Promega) in 50 mM ammonium bicarbonate (pH 8.0) was added in a 1:100 trypsin:protein ratio, on top of the protein pellets for 18-24 h at 37 °C at 1000 rpm on a thermoshaker, typically yielding peptide concentration of 1-2 μ g/ μ L.

6.3.4 Dephosphorylation

The next day, trypsin was quenched by the addition of 1 x cOmplete Protease Inhibitor (Roche) and peptides were quantified using the (ThermoFisher, #23290). Peptides were treated with 1× rCutSmart buffer (pH 8, New England BioLabs) and 5 units of alkaline phosphatase (calf intestinal, QuickCIP, New England BioLabs, 10,000 units/mL) per 100 μ g of peptide material or a minimum of 40 units (4 μ L) for 6 h or overnight at 37 °C while shaking. After dephosphorylation, all peptides were desalted using plate-based Oasis-HLB desalting sorbent.

6.3.5 96-well plate-based Oasis-HLB desalting

After dephosphorylation, peptides were acidified to pH 2-3 by addition of trifluoroacetic acid (TFA). Oasis HLB (Waters, #186007549) was prepared at 100 g/L in MeCN and an appropriate volume (100 μ g Oasis per 1 μ g peptide) was added to each well of a 96 well filter plate from either OROCHEM (OF1100-7PE) or Millipore (MSHVS4510, Multiscreen® hydrophilic PVDF membrane, pore size: 0.45 μ m) using a 12-channel pipette. The sorbent was washed twice with 100 μ L MeCN, twice 125 μ L H₂O and twice 125 μ L 0.1% TFA with a centrifugation 50 x g for 1 min after each wash. Acidified peptides were mixed with Oasis-HLB by pipetting up and down 20 times and centrifuged for 2 min 50g. The samples were then washed twice with 125 μ L 0.1% TFA, once with 125 μ L H₂O and eluted with 95 μ L MeCN followed by a second elution of 95 μ L 70% MeCN by centrifugation (50 x g, 1 min). To each sample, 11 μ L TFA, 11 μ L 2M glycolic acid and 8 μ L 70% MeCN was added for a final volume of 220 μ L (5% TFA, 0.1 M GA, 70% MeCN).

6.3.6 Automated Zirconium (IV) Immobilised Metal Affinity Chromatography

Phosphopeptide enrichment was performed on a KingFisher Flex System (Thermo Scientific) using MagReSyn zirconium-based immobilized metal affinity chromatography high performance (Zr-IMAC HP) beads (ReSyn Biosciences) as per the manufacturer's instructions, as depicted in **Fig. 3.4a**. The enrichment protocol utilised a 1:40 bead-to-peptide weight ratio. Briefly, the magnetic particles were first equilibrated with loading/binding solvent (80% MeCN, 5% TFA, 0.1M Glycolic Acid) for 5 min, followed by 20 min incubation with the peptide mixture. In most experiments, apart from the pilot (10 μ L) and Zr-IMAC optimisation (1 μ L, 5 μ L, 10 μ L), 3 μ L of beads was used. After binding, beads were then washed for 2 min each, first with the binding solvent, then wash solvent 1 (80% MeCN, 1% TFA, 500 μ L) and lastly by wash solvent 2 (10% MeCN, 0.2% TFA, 500 μ L). Phosphopeptides were eluted from the microparticles by 10 min incubation in 1% NH_4OH (100 μ L), followed by another elution. Total program time was 58 mins and all incubation steps were done with continuous mixing to keep a homogenous bead suspension. The two elutions were then quickly combined and acidified to pH 2-3 with 50 μ L of 20% formic acid (FA) for a total of 250 μ L. Half of the available peptides (125 μ L) was loaded onto EvoTips for LC-MS/MS analysis.

6.4 SAPPHIRE version 1 (SAPPHIREv1) and optimisation

HEK293T cell lysis was performed as described in **6.3.2**. The samples were treated with 10 mM neutralised TCEP and 25 mM NEM for 1.5 h at RT with nutation. All samples were then precipitated using MeCN-based protein precipitation. The resulting protein pellets were resuspended in $\frac{1}{4}$ of original sample volume in buffer A and split into two aliquots. To one aliquot, 0.75 M HA (50% by weight in H_2O , Sigma-Aldrich, 438227-50ML, x3 volumes of 1 M HA, diluted to 1 M in buffer B) and 100 μ M 6C-CysPAT (in 100 mM triethylammonium bicarbonate (TEAB) buffer) were added to facilitate S-acyl cleavage and phosphonate labelling of the newly revealed Cys. The second aliquot served as the negative control, where x3 volumes of the aliquot's original volume of Buffer B supplemented with 100 μ M CysPAT were added. The mixtures were thoroughly vortexed and incubated for 1 h at RT in the dark with mixing by nutation. Subsequently The samples were then precipitated using MeCN-based precipitation. After the final wash, sequence grade trypsin (Promega) in 50 mM ammonium bicarbonate (pH

8.0) was added in a 1:100 trypsin:protein ratio, on top of the protein pellets for 18-24 h at 37 °C at 1000 rpm on a thermoshaker.

The next day peptides were desalted using the Pierce™ peptide desalting spin columns (ThermoFisher, #89852), according to the manufacturer's instructions. Samples were evaporated (without heating) overnight in a Savant SPD1010 vacuum concentrator (ThermoFisher). Desalted peptides were dissolved in 1× rCutSmart buffer (pH 8, New England BioLabs) and incubated with 50 units of alkaline phosphatase (calf intestinal, QuickCIP, New England BioLabs) overnight at 37 °C in a thermomixer at 1000 rpm. After dephosphorylation, all peptides were again desalted as described above and air-dried using a vacuum centrifuge.

Peptides were then dissolved in loading buffer (80% MeCN, 5% TFA, 0.1M GA) and enriched manually, with the Zr-IMAC HP beads following the manufacturer's instructions. The enrichment protocol utilised a 1:50 bead-to-peptide weight ratio. Briefly, the magnetic particles were first equilibrated with loading/binding solvent (80% MeCN, 5% TFA, 0.1M Glycolic Acid) for 5 min. The peptides were added on top of the beads and pipetted up and down for 30 times, followed by incubation on a plate shaker (750 rpm). After binding, beads were then washed for 2 min each, first with the binding solvent, then wash solvent 1 (80% MeCN, 1% TFA, 500 µL) and lastly by wash solvent 2 (10% MeCN, 0.2% TFA, 500 µL). Phosphopeptides were eluted from the microparticles by 10 min incubation in 1% NH₄OH (100 µL), followed by another elution. The two elutions were then quickly combined and acidified to pH 2-3 with 50 µL of 20% formic acid (FA) and stored at -20 before LC-MS/MS loading.

6.5 APEX2 proximity labelling

6.5.1 APEX2 in-cell biotinylation and cell lysis

U2OS-ΔΔG3BP1/2-APEX2-G3BP1 and APEX2-NES cells were plated in appropriate plates (specified below for each experiment) at 80% confluency using media supplemented with 50 ng/mL doxycycline to induce expression of the APEX2 construct. The next day, cells were incubated with NaAsO₂ (500 µM, Sigma-Aldrich, 1062771000) to induce stress granule (SG) formation, and biotin-phenol (500 µM, Iris Biotech GmbH, LS-3500) for 30 mins at 37°C. To initiate proximity labelling H₂O₂ (1 mM, Sigma-Aldrich) was added for exactly 1 min. APEX2 activity was quenched by three washes with ice-cold DPBS supplemented with sodium azide (10mM, Mallinckrodt, 1953-57), sodium ascorbate (10mM, Sigma-Aldrich, A7631) and Trolox

(5mM, Sigma-Aldrich, 238813). Cells were then lysed on-plate, by the addition of supplemented RIPA buffer (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox, 1 mM PMSF, 1x cOmplete Protease Inhibitor (Roche) in RIPA, pH 7.5). Lysis proceeded for 20 mins at 4 °C and the lysates were clarified by centrifugation (17 000 x g for 10 min at 4°C).

6.5.2 Western Blotting

For Western Blot shown in **Fig. 2.5**, U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1 and U2OS- $\Delta\Delta\Delta$ FXR1/2, FMR1-APEX2-FXR1 that have an endogenous expression of G3BP1, were plated at 70-80% confluency using media supplemented with 50 ng/mL doxycycline to induce expression of the APEX2 construct. Cells were lysed in RIPA buffer supplemented with cOmplete Protease Inhibitor cocktail (Roche). The concentration of the isolated proteins was determined using with DC Protein Assay (Bio-Rad), and 20 μ g protein subjected to SDS-PAGE and Western blotting as described in sections **6.1.4** and **6.1.5**.

For Western Blot shown in **Fig. 2.6**, U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1 and U2OS APEX2-NES cells were plated in 6-well plates and APEX2 induction, labelling and lysis were performed as described in **6.5.1**. The concentration of the isolated proteins was determined using with DC Protein Assay (Bio-Rad), and 20 μ g protein subjected to SDS-PAGE and Western blotting as described in sections **6.1.4** and **6.1.5**.

6.5.3 Confocal Microscopy

U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1 cells were seeded on 13 mm glass coverslips (placed in 48-well plate) and APEX2 labelling was performed as described above. After quenching, cells were fixed with 4% formaldehyde in X1 PBS for 15 mins and permeabilised with 0.5% Triton, X1 PBS for 10 min. Cells were then rinsed with PBS and blocked with blocking buffer (2% BSA, 0.02% Triton, X1 PBS) overnight at 4°C. Cells were incubated with the primary antibody against TIA1 diluted in blocking buffer for 1 h at RT, subsequently washed three times for 5 mins each with 1X DPBS, and incubated with AlexaFluor488-conjugated goat anti-mouse IgG antibody and Neutravidin-TexasRed conjugate diluted in blocking buffer for 2 h at RT. Cells were washed three times with X1 PBS and coverslips were mounted on slides with Fluoromount-G™ Mounting Medium containing DAPI (Thermo Fisher). Slides with coverslips were stored in the dark at 4°C before imaging on the Leica STELLARIS 8 confocal microscope (63 × oil objective; excitation wavelengths: GFP - 488 nm; mCherry - 560 nm. Images were cropped and

processed using Image J (NIH Bethesda) software. All antibody information is available in Table 6.1.

6.5.4 APEX-AX Proteomics

6.5.4.1 *First attempt: Double enrichment*

U2OS- $\Delta\Delta$ G3BP1/2-APEX2-G3BP1 cells were plated in T175s and APEX2 induction, labelling and lysis were performed as described in 6.5.1. The lysates were then split into four technical replicates and treated with 10 mM neutralised TCEP and 25 mM NEM for 1.5 h at RT with nutation. All samples were then precipitated using MeCN-based protein precipitation. The resulting protein pellets were resuspended in ¼ of original sample volume in buffer A and split into two aliquots. To one aliquot, 0.75 M HA (50% by weight in H₂O, Sigma-Aldrich, 438227-50ML, x3 volumes of 1 M HA, diluted to 1 M in buffer B) and 100 μ M 6C-CysPAT (in 100 mM triethylammonium bicarbonate (TEAB) buffer) were added to facilitate S-acyl cleavage and phosphonate labelling of the newly revealed Cys. The second aliquot served as the negative control, where x3 volumes of the aliquot's original volume of Buffer B supplemented with 100 μ M CysPAT were added. The mixtures were thoroughly vortexed and incubated for 1 h at RT in the dark with mixing by nutation. Subsequently The samples were then precipitated using MeCN-based precipitation.

Protein pellets were then resuspended in 0.2% SDS, 50 mM HEPES, pH 7.5, and incubated with pre-washed (x3 times: 0.2% SDS, 50 mM HEPES, pH 7.5) Pierce NeutrAvidin beads (100 μ L beads per 500 μ g protein, ThermoFisher) for 3 h at RT nutation. Following enrichment, the beads were washed x3 times with 1.0% (w/v) SDS in HEPES pH 7.5 (500 μ L), and x3 times with 50 mM HEPES pH 8.0 (500 μ L). TCEP (5 mM in HEPES pH 8.0). Sequence grade trypsin (0.2 μ g, Promega) in 50 mM ammonium bicarbonate (pH 8.0) was added to the bead suspension and incubated overnight at 37 °C in a thermomixer (1200 rpm). The following day, beads were pelleted by centrifugation (5,000 $\times$ g, 5 min), and the supernatant was transferred to a separate 1.5 mL Protein LoBind Eppendorf tubes. To maximise peptide recovery, the beads were additionally washed with 100 μ L 50 mM ammonium bicarbonate (pH 8.0), and the supernatant was combined with the peptides.

Peptides were then desalted using the Pierce™ peptide desalting spin columns (ThermoFisher, #89852), according to the manufacturer's instructions. Samples were evaporated (without

heating) overnight in a Savant SPD1010 vacuum concentrator (ThermoFisher). Desalted peptides were dissolved in 1× rCutSmart buffer (pH 8, New England BioLabs) and incubated with 50 units of alkaline phosphatase (calf intestinal, QuickCIP, New England BioLabs) overnight at 37 °C in a thermomixer at 1000 rpm. After dephosphorylation, all peptides were again desalted as described above and air-dried using a vacuum centrifuge.

Peptides were then subjected to High-Select™ Fe-NTA Phosphopeptide Enrichment (ThermoFisher, A32992) following the manufacturer's instructions. After the final elution, peptides were air-dried overnight and dissolved in 0.1% FA for LC-MS/MS analysis.

6.5.4.2 Second attempt: Phosphonate as a mass tag

U2OS-ΔΔG3BP1/2-APEX2-G3BP1 cells were plated in T75s and APEX2 induction, labelling were performed as described in **6.5.1** (n=4) and lysis was performed as described in **6.3.2**. Samples were subjected to first half of the SAPPHIRE protocol described above (**6.3.1 - 6.3.3**). Briefly, samples were treated with 10 mM neutralised TCEP and 25 mM NEM for 1.5 h at RT with nutation. All samples were then precipitated using MeCN-based protein precipitation. The resulting protein pellets were resuspended in Buffer A at their original volume and subjected to a second alkylation step under the same conditions (10 mM neutralised TCEP and 25 mM NEM for 1.5 h at RT with nutation). Samples were precipitated again using MeCN-based precipitation and the resulting protein pellets were resuspended in ¼ of the original volume in buffer A and split into two aliquots. To one aliquot, 0.75 M HA (50% by weight in H₂O, Sigma-Aldrich, 438227-50ML, x3 volumes of 1 M HA, diluted to 1 M in buffer B) and 100 uM 6C-CysPAT (in 100 mM triethylammonium bicarbonate (TEAB) buffer) were added to facilitate S-acyl cleavage and phosphonate labelling of the newly revealed Cys. The second aliquot served as the negative control, where x3 volumes of the aliquot's original volume of Buffer B supplemented with 100 uM CysPAT were added. The mixtures were thoroughly vortexed and incubated for 1 h at RT in the dark with mixing by nutation. Subsequently, samples were then precipitated using MeCN-based precipitation.

Protein pellets were then resuspended in 0.2% SDS, 50 mM HEPES, pH 7.5, and quantified using the DC protein assay. 600 µg of each sample was then incubated with pre-washed (x3 times: 0.2% SDS, 50 mM HEPES, pH 7.5) Pierce NeutrAvidin beads (100 µL beads per 500 µg protein, ThermoFisher) for 3 h at RT nutation. The supernatant was removed and kept for analysis. Following enrichment, the beads were washed x3 times with 1.0% (w/v) SDS in

HEPES pH 7.5 (500 μ L), and x3 times with 50 mM HEPES pH 8.0 (500 μ L). TCEP (5 mM in HEPES pH 8.0). Sequence grade trypsin (0.2 μ g, Promega) in 50 mM ammonium bicarbonate (pH 8.0) was added to the bead suspension and incubated overnight at 37 °C in a thermomixer (1200 rpm). The following day, beads were pelleted by centrifugation (5,000 $\times$ g, 5 min), and the supernatant was transferred to a separate 1.5 mL Protein LoBind Eppendorf tubes. To maximise peptide recovery, the beads were additionally washed with 100 μ L 50 mM ammonium bicarbonate (pH 8.0), and the supernatant was combined with the peptides. Peptides were acidified and stored at -80 prior to submission for LC-MS/MS analysis. Half of the available peptides was loaded for LC-MS/MS analysis.

6.5.4.2.1 Neutravidin-HRP Western Blot

Roughly 10 μ g of protein was kept from the input lysate (pre Neutravidin enrichment) for the Neutravidin-HRP Western blot analysis to assess pulldown efficiency. After the enrichment as described above the supernatant was saved. After all the washes, 9 μ L of the bead slurry was kept and enriched proteins were eluted by boiling in 1x Laemmli sample loading buffer supplemented with 5% (v/v) β -mercaptoethanol (95 °C, 10 min).

6.6 Cysteine-centric and intact *S*-acylation ID

6.6.1 Acyl-PEG Exchange (APE)

APE was adapted from the previously published protocol¹⁵⁷. HEK293T cell lysates were prepared as described in **6.3.2**. Approximately 200 μ g of lysate was carried forward for each condition and samples were subject to double alkylation as described in **6.3.3**. After the double blocking, samples were precipitated with MeCN-based precipitation and the resulting protein pellets were resuspended in ¼ of the original volume in buffer A and split into two aliquots. One aliquots was incubated with 0.75 M HA (50% by weight in H₂O, Sigma-Aldrich, 438227-50ML, x3 volumes of 1 M HA, diluted to 1 M in buffer B; or neutralised hydroxylamine hydrochloride (SigmaAldrich) in Buffer B) and 1 mM 10 kDa methoxypolyethylene glycol maleimide (mPEG-maleimide) (Sigma-Aldrich #712469) in Buffer B, to facilitate *S*-acyl cleavage and mPEG labelling of the newly revealed Cys. The second aliquot served as the negative control, where x3 volumes (of the aliquot's original volume) of Buffer B supplemented

with 1 mM mPEG-mal were added. The mixtures were thoroughly vortexed and incubated for 2 h at RT. All samples were subjected to another MeCN-based precipitation and the protein pellets were dissolved in 1x Laemmli loading buffer. Subsequent SDS-PAGE and Western Blots were performed as described in **6.1.4** and **6.1.5**.

6.6.2 Site-specific Acyl-Biotin Exchange (ssABE)

ssABE was performed as described previously with slight modifications¹⁷⁵. The same HEK293T cell lysate was used as for the experiment described in **Chapter 3.4**, and split into four technical replicates. The protein concentration was determined by the DC protein assay (Bio-Rad) and 1.6 mg (1.6 mg/mL) of protein per replicate was taken forward for the ssABE protocol. The lysates were incubated with 112 μ L 500 mM TCEP pH 7.5 for 30 min at RT, followed by the addition of 35.5 μ L 1 M NEM for 1 h at RT. Again, 35.5 μ L 1 M NEM was added and 568 mg urea (8 M) were added to the samples and incubated for 3 h at RT. The samples were then precipitated once by MeCN-based precipitation and redissolved in 400 μ L 2% (w/v) SDS, 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA. To the samples, 21 μ L 500 mM TCEP pH 7.5 was added for 30 min at RT, and subsequently, 22 μ L 1 M NEM was added for 1 h at RT. The samples were then precipitated x4 times by MeCN-based precipitation and redissolved in 400 μ L 2% (w/v) SDS, 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA. Each sample was then split in two aliquots: + and - HA treatment. To the +HA sample, 200 μ L 5 mM HPDP-biotin (in DMSO, ThermoFisher #21341) and 62 μ L 7.5 M HA pH 7.5. To the negative -HA control, 200 μ L 5 mM HPDP-biotin and H₂O (proteomics grade) were added. All samples were then incubated for 90 min at RT.

The samples were precipitated three times using the MeCN-based precipitation and redissolved in 75 μ L 2% (w/v) SDS, 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA. Samples were diluted with 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA to a final SDS concentration of 0.1%. The samples were added to 25 μ L Pierce NeutrAvidin beads (pre-washed with 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% SDS) and incubated for 3 h at RT with end-to-end rotation. Following enrichment, the beads were washed x3 times with 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% SDS and x4 times with 50 mM TEA pH 7.5, 150 mM NaCl, 5 mM EDTA. To the beads, 70 μ L 50 mM TEA pH 7.5, 50 mM TCEP, 5 mM EDTA, 0.1% (w/v) Rapigest was added. After incubation at RT for 30 min with nutation, the supernatant was collected and diluted with 230 μ L 50 mM HEPES pH 8. The enriched samples were reacted with 25 mM CAA

at RT for 20 min. Subsequently, the HEK293T samples were digested with LysC (0.1 μ g) at 37 °C for 30 min and trypsin (0.1 μ g) at 37 °C overnight. The next day, peptides were acidified with FA to pH 2-3 and stored at -20 °C until LC-MS/MS analysis. All available peptides were loaded onto EvoTips for LC-MS/MS analysis.

6.6.3 nanographite fluoride-based solid-phase extraction (nGF-SPE)

The nGF-SPE was performed as described previously with slight modifications⁹⁶. The most notable modifications include a lower input (6 mg compared to their 7.5 mg per sample), desalting with Oasis-HLB sorbent (compared to their use of C18 columns), peptides were dried overnight in SPD1010 vacuum concentrator (ThermoFisher) (compared to freeze-drying). HEK293T cells were cultured in T175s (n=4) were collected and lysed as described in. Lysates were subjected to MeCN-based precipitation in 50 mL falcon tubes. The resulting protein pellets were redissolved in 4% SDS, 50 mM Tris, 5 mM EDTA, pH = 7.4, and the concentration was determined using DC protein assay. 6 mg of protein was taken per sample and diluted 1:1 with 150 mM NaCl, 50 mM Tris, 5 mM EDTA, pH = 7.4, and reduced with 10 mM TCEP and alkylated with 50 mM NEM for 1 h at RT with nutation. Samples were then moved to a 96-well plate, with each sample equally divided across a single row (12 wells) and precipitated with MeCN-based precipitation. Protein pellets were redissolved with 50 mM Tris-HCl (pH = 7.4) and digested at 37 °C overnight at a trypsin:protein ratio of 1:100 (w/w). The next day, the tryptic digests were desalted using the Oasis-HLB sorbent in a 96-well plate format as described in and eluted with 95 μ L MeCN containing 0.1% TFA followed by a second elution of 95 μ L 70% MeCN by centrifugation (50 x g, 1 min). All eluates from one row (corresponding to the previously divided single sample) were combined together in a 5 mL Protein LoBind falcon tube and subsequently aliquoted into x2 1.5 mL Protein LoBind Eppendorf tubes and dried using a Savant SPD1010 vacuum concentrator (ThermoFisher) overnight without heat. Peptides were then dissolved with 90% MeOH at 1 mg/mL, followed by vortexing and water bath sonication to aid dissolving. 10% MeOH was added on top in a 1:1 ratio for a final concentration of 50% MeOH at 0.5 mg/mL per sample. The two aliquots from the same sample were combined and peptides were then incubated with 20 mg nGF (prewashed with MeOH) for 2 h at RT with nutation. Samples were then centrifuged (17,000 x g, 5 min). The resting nGF was sequentially washed with 20% MeCN, 50% MeOH and 60% MeOH. The peptides were eluted sequentially with 500 μ L 90% MeOH and 100% MeOH and the two eluates were pooled

together and dried using Savant vacuum concentrator. Then the lyophilized eluate was dissolved with 50% MeOH (first 90% then 10% as described above) and subjected to the second enrichment with 3 mg nGF for 2 h at RT with nutation, for a total of two enrichments. The supernatant was removed by centrifugation (17,000 x g, 5 min) and nGF was sequentially washed with 50% MeOH and 60% MeOH. Peptides were then sequentially eluted with 500 µL 90% MeOH and 100% MeOH. The two eluates were combined and dried down overnight as describe before. Dried peptides were stored at -80 prior to LC-MS/MS analysis.

6.7 General proteomics methods

Peptides for proteomics analysis were loaded onto Evotips (Evosep) by Daria Berezina, as per the manufacturer's instructions. Samples were run on the LC-MS/MS by Dr. Jianan Lu or Dr. Jack Houghton. All enrichment proteomics samples (SAPPHIRE, APEX2, ssABE, nGF-SPE) were filtered through 0.45 µm hydrophilic Durapore® filters (Millipore) before LC-MS/MS loading.

6.7.1 Whole Proteome sample preparation

Protein lysates were collected after the double reduction and alkylation as part of the SAPPHIRE protocol and ~10 µg of each sample was taken forward for MeCN-based precipitation. The samples were then tryptic digested with 25 µL of freshly prepared trypsin (Promega) at 14 ng/µL in 25 mM ammonium bicarbonate (pH 8.0) overnight at 37 °C on a thermoshaker (1000 rpm). The peptides were then quantified using the NanoDrop (NanoDrop One, ThermoFisher) and normalised to 10 ng/µL. The peptides were then acidified to pH 2-3 with FA and stored at -20 before loading of 20 µL (200 ng) into the EvoTips.

6.7.2 Evosep-timsTOF LC-MS/MS: DDA enrichment proteomics

Peptides were analysed by nanoLC-MS/MS using an Evosep One (Evosep) coupled with a timsTOF HT (Bruker) equipped with an 8 cm x 150 µm, 1.5 µm analytical column (Evosep). Peptides were separated using the Evosep 60 SPD workflow (Analytical solvents A: 0.1% FA and B: acetonitrile plus 0.1% FA). The only experiment where 30 SPD workflow was used was in HEK293T profiling in **Chapter 3.4**. Column was held at 40 °C. Data was acquired in DDA PASEFmode, with the following settings: m/z range from 100 m/z to 1700 m/z, ion mobility range from $1/K_0 = 1.30$ to 0.85 Vs/cm² using equal ion accumulation and ramp times in the

dual TIMS analyser of 100 ms each. The collision energy was lowered as a function of increasing ion mobility from 59 eV at $1/K_0 = 1.6$ Vs/cm² to 20 eV at $1/K_0 = 0.6$ Vs/cm². Isolation width was lowered as a function of decreasing m/z from 3 m/z at 800 m/z to 2 m/z at 700 m/z . Active exclusion was applied for 0.4 min. Each cycle consistent of 4 PASEF ramps (total cycle time 0.53 sec) with 2.75 ms measuring time allowed for each selected precursor.

6.7.3 Evosep-timsTOF LC-MS/MS: DIA whole proteome

Peptides were analysed by nanoLC-MS/MS using an Evosep One (Evosep) coupled with a timsTOF HT (Bruker) equipped with an 8 cm x 150 μ m, 1.5 μ m analytical column (Evosep). 200ng peptides (or less for experiments in **Chapter 4.2** and **4.4**) were separated by the Evosep 60 SPD workflow (Analytical solvents A: 0.1% FA and B: acetonitrile plus 0.1% FA). Column was held at 40°C. Data were acquired in DIA PASEF mode with the following settings: m/z range from 100 m/z to 1700 m/z , ion mobility range from $1/K_0 = 1.30$ to 0.85 Vs/cm² using equal ion accumulation and ramp times in the dual TIMS analyser of 100 ms each. Each cycle consisted of 8 PASEF ramps covering 21 mass steps each with 25 Da windows each with 2/3 non-overlapping ion mobility windows covering the 475 to 1000 m/z range and 0.85 and 1.26 Vs/cm² ion mobility range. The collision energy was lowered as a function of increasing ion mobility from 59 eV at $1/K_0 = 1.6$ Vs/cm² to 20 eV at $1/K_0 = 0.6$ Vs/cm².

6.7.4 nanoElute2-timsTOF LC-MS/MS: DDA nGF-SPE

Dried peptides were dissolved with 1 μ L 35% MeCN, aided by water bath sonication and 2 h shaking at RT, and 2 μ L were injected. Peptides were analysed by nanoLC-MS/MS with a nanoElute 2 (Bruker) coupled with a timsTOF HT (Bruker) equipped with a 25 cm, 150 μ m, 1.5 μ m analytical column (PepSep). Peptides were separated by a 57-minute analytical gradient which was 35%–65% solvent B over 30 minutes, rising to 95% solvent B over 5 minutes, followed by a 22-minute wash at 95% solvent B (Analytical solvents A: 0.1% FA and B: acetonitrile plus 0.1% FA). Data were acquired in data-dependent acquisition PASEF mode with the following settings: m/z range from 100 m/z to 1700 m/z , ion mobility range from $1/K_0 = 1.30$ to 0.70 Vs/cm² using equal ion accumulation and ramp times in the dual TIMS analyser of 100 ms each. The collision energy was lowered as a function of increasing ion mobility from 59 eV at $1/K_0 = 1.6$ Vs/cm² to 20 eV at $1/K_0 = 0.6$ Vs/cm². Isolation width was lowered as a function of decreasing m/z from 3 m/z at 800 m/z to 2 m/z at 700 m/z . Active exclusion was

applied for 0.4 min. Each cycle consistent of 10 PASEF ramps (total cycle time 1.17 sec) with 2.75 ms measuring time allowed for each selected precursor.

6.8 Proteomics data analysis

6.8.1 Fragpipe

The data obtained as ddaPASEF Bruker .d files was processed using Fragpipe version 22.0 (Nesvilab)²⁷⁴. Data was searched against a human reference proteome with isoforms (Uniprot, UP000005640, 13 July 2023 containing 20552 proteins) with 50% decoys and contaminants added. The built-in label-free quantification-match between runs (LFQ-MBR) workflow was used with slight modifications. The Bruker .d files were searched using MSFragger (version 4.1). The following parameters were applied: strict trypsin digestion; maximum of 2 missed cleavages allowed; precursor ion tolerance of 20 ppm. The variable modifications were set as follows for all **SAPPHIRE** experiments: trimming of protein N-terminal Met; Met oxidation (mass delta 15.9949), N-terminal acetylation (mass delta 42.0106), Ser/Thr/Tyr (mass delta 79.96633), Cys phosphonate modification (mass delta 221.0822), Cys hydrolysed NEM modification (mass delta 143.1406), Cys NEM modification (mass delta 125.0477), with a maximum of three variable modifications allowed. For the **nGF-SPE** experiments the following variable modifications were set: trimming of protein N-terminal Met; Met oxidation (mass delta 15.9949), N-terminal acetylation (mass delta 42.0106), Cys palmitoylation (mass delta 238.2297), Cys NEM modification (mass delta 125.0477). For the **ssABE** experiment: trimming of protein N-terminal Met; Met oxidation (mass delta 15.9949), N-terminal acetylation (mass delta 42.0106), Cys carbamylmethylation (mass delta 57.02146), Cys NEM modification (mass delta 125.0477). For **APEX2** experiments: trimming of protein N-terminal Met; Met oxidation (mass delta 15.9949), N-terminal acetylation (mass delta 42.0106), Ser/Thr/Tyr (mass delta 79.96633), Cys phosphonate modification (mass delta 221.0822), Cys hydrolysed NEM modification (mass delta 143.1406), Cys NEM modification (mass delta 125.0477) No fixed modifications were set in all searches. All other default parameters were used for processing. MSFragger search results were processed using Percolator (version 3.6.5) for PSM validation, followed by Philosopher (version 5.1.1) for protein and FDR filtering. LFQ values were calculated using the MaxLFQ algorithm through IonQuant (version 1.10.27) with MBR enabled and minimum ions set as 1.

6.8.2 DIA-NN: Whole proteome data

The data was obtained as diaPASEF Bruker .d files, which were processed using library-free analysis in DIA-NN (version 1.8.1)²⁷⁵. The following parameters were used: human protein database (downloaded from UniProt 13 July 2023 containing 20552 proteins and 246 common contaminants) “deep learning-based spectra and RTs prediction” enabled; digestion using trypsin with a maximum of 1 missed cleavage; N-terminal excision and as fixed modifications; UniMod:108 with mass delta 125.048 at C (NEM), UniMod:320 with mass delta 143.058 at C (NEMhyd), UniMod:4 with mass delta 57.0215 at C (carbamedomethylation) as variable modifications, with a maximum of two variable modifications allowed; match between runs enabled; quantification strategy set to “robust LC (high precision)””; heuristic protein inference disabled; mass and MS1 accuracy set to 0.

6.8.3 Statistical analysis and visualisation

All volcano plots, quadrant plots, PCA plots, Venn diagrams and Upset plots were created in RStudio (version 2023.6.2.561) using ggplot2, Eulerr and UpsetR packages^{276–278}. All cross-referencing to the SwissPalm database was likewise performed in RStudio with the data downloaded from <https://swisspalm.org/> (Release 4 through this thesis and Release 5 only in **Chapter 4.6.2**)⁶³.

6.8.3.1 Whole proteome

The “report.pg_matrix.tsv” generated by DIA-NN was uploaded to Perseus (version 1.6.2.3). Replicates for each condition were grouped together, and MaxLFQ intensity values were log₂ transformed. Rows were filtered for contaminants and the presence of at least 3 replicates in at least one condition group. Missing values were imputed from a downshifted normal distribution (width = 0.3, downshift = 1.8). A two-tailed student’s t-test was then performed to determine abundance changes and significance (permutation based using 250 randomisations, FDR=0.05, s0=1). Fold-change values were plotted against adjusted p-values in a volcano plot.

6.8.3.2 *SAPPHIRE*

The “combined_modified_peptides.tsv” files generated by Fragpipe were uploaded to the Perseus platform workspace (version 1.6.2.3)²⁷⁹. Replicates for each condition were grouped together, and Intensity values were $\log_2$ transformed. Peptides were then filtered for contaminants, presence of at least one phosphonated Cys (pCys) within the peptide sequence, and the presence of at least 3 replicates in at least one condition group. Missing values were imputed from a downshifted normal distribution (width = 0.3, downshift = 1.8). A two-tailed student’s t-test was then performed to determine abundance changes and significance (permutation based using 250 randomisations, FDR=0.05, s0=1). Fold-change values were plotted against adjusted p-values in a volcano plot. For data presented in **Fig. 4.4**, the whole proteome and SAPPHIRE datasets were processed as described above, and the final $\log_2$ transformed whole proteome intensities were subtracted from each $\log_2$ transformed peptide intensities corresponding to that protein. The resulting values were plotted and a two-tailed student’s t-test was then performed to determine abundance changes and significance (permutation based using 250 randomisations, FDR=0.05, s0=1)

6.8.3.3 *ssABE and APEX-AX (phosphonate as a mass tag)*

The “combined_protein.tsv” files generated by Fragpipe were uploaded to the Perseus platform workspace (version 1.6.2.3)²⁷⁹. Replicates for each condition were grouped together, and Intensity values were $\log_2$ transformed. MaxLFQ intensities were then filtered for contaminants and the presence of at least 3 replicates in at least one condition group. Missing values were imputed from a downshifted normal distribution (width = 0.3, downshift = 1.8). A two-tailed student’s t-test was then performed to determine abundance changes and significance (permutation based using 250 randomisations, FDR=0.05, s0=1). Fold-change values were plotted against adjusted p-values in a volcano plot.

6.9 Bioinformatic analysis

6.9.1 Cysteine PTM, functional and pPSE annotation

Residue-specific annotations for PTMs, functions and pPSE values were obtained from published data²¹⁷. Residue-specific annotations for functions (metal binding, active site nucleophile) and PTMs (prenyl, disulfide) were downloaded from UniProt, and matched to

unique UniProt accession IDs and residue positions. Palmitoyl modifications were downloaded from SwissPalm (<https://swisspalm.org/>). No filtering for prediction quality was applied across the pPSE distributions. For all mappings, the UniProt canonical human FASTA (downloaded 26/10/2022) was used. The reference proteome and formatted databases used for these analyses are provided at: <https://zenodo.org/record/7752842>.

6.9.2 Identification of S-acylation Motifs

Sequences of width 13 with identified phosphonated Cys (pCys) at the centre were used. The motif-x algorithm in MoMo software was used to identify the S-acylation motifs²¹⁶. The “classic” motif-x algorithm was run with the background extracted from the human proteome. The enhanced motif-x was ran with default settings where background peptides are generated by shuffling foreground peptides, while conserving the central residues. Motifs were required to occur at least 5% of the entries for all samples, and the *p*-value threshold was set to 0.0001 (α -value of 0.025).

6.9.3 Clinical variants annotations

Residue-specific annotations for clinical variants were downloaded from (https://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/) and reformatted by Matthew White.

6.9.4 Functional enrichment analysis

Functional enrichment analysis presented in **Fig.3.9** was performed using Metascape and the KEGG pathway enrichment analyses throughout this were performed in ShinyGO^{211,280}.

Table 6.1. List of antibodies used in this study.

Target	Species	Use	Clone	Dilution	Supplier	Code
G3BP1	Mouse	Western Blot	Monoclonal	1:1000	Santacruz	sc-365338
Neutravidin-HRP	N/A	Western Blot	N/A	1:1000	Invitrogen	A2664
PI4K2A	Rabbit	Western Blot	Polyclonal	1:1000	Proteintech	15318-1-AP
PCNA	Rabbit	Western Blot	Polyclonal	1:50,000	Proteintech	10205-2-AP
RBM39	Rabbit	Western Blot	Polyclonal	1:1000	Proteintech	21339-1-AP
FAM177A	Rabbit	Western Blot	Polyclonal	1:10,000	Atla antibodies	HPA055440
Calnexin	Rabbit	Western Blot	Monoclonal	1:1000	Cell Signalling Technology	2679
IFITM3	Rabbit	Western Blot	Polyclonal	1:5000	Proteintech	11714-1-AP
Vinculin	Rabbit	Western Blot	Polyclonal	1:10,000	Abcam	ab91459
GAPDH	Mouse	Western Blot	Monoclonal	1:1000	Abcam	ab8245
IRDye® 800CW anti-Rabbit IgG Secondary Antibody	Goat	Western Blot	Polyclonal	1:10,000	LiCor Bio	926-32211
IRDye® 800CW anti-Mouse IgG Secondary Antibody	Goat	Western Blot	Polyclonal	1:10,000	LiCor Bio	926-32210
anti-rabbit IgG (H+L), HRP conjugate	Goat	Western Blot	Polyclonal	1:10,000	Advansta	R-05072-500
anti-mouse IgG (H+L), HRP conjugate	Goat	Western Blot	Polyclonal	1:10,000	Advansta	R-05071-500
TIA1	Mouse	Microscopy	Monoclonal	1:200	Thermo Fisher	OTI1D7)
Neutravidin-TexasRed	N/A	Microscopy	N/A	1:500	Thermo Fisher	A2665
AlexaFluor488-conjugated goat anti-mouse IgG antibody	Goat	Microscopy	Polyclonal	1:500	Abcam	ab150113

Table 6.2. SAPPHERE experiments experimental conditions

Experiment	Free Cys alkylation First/Second	Peptide input for Zr-IMAC enrichment	Steps conducted in 96- well plate
HEK293T (Chapter 3.4)	CAA/NEM	500 µg	Last precipitation, digestion, dephosphorylation, desalting, enrichment
Breast Cancer/Colorectal Cancer patient tissues (Chapter 4.2)	CAA/CAA	0-40 µg	Last precipitation, digestion, dephosphorylation, desalting, enrichment
<i>T. Gondii</i> (Chapter 4.3)	CAA/CAA	200 µg	Last precipitation, digestion, dephosphorylation, desalting, enrichment
Pyroptosis (Chapter 4.4)	NEM/NEM	~75 µg	Last precipitation, digestion, dephosphorylation, desalting, enrichment
Stress Response (Chapter 4.5)	NEM/NEM	µg	Full protocol

References

1. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 2017 185 **18**, 285–298 (2017).
2. Wagner, R. Einige bemerkungen und fragen über das keimbläschen (vesicular germinativa). *Müller's Arch. Anat Physiol Wiss. Med* **268**, 373–377 (1835).
3. Pederson, T. The Nucleolus. *Cold Spring Harb. Perspect. Biol.* **3**, a000638 (2011).
4. Sabari, B. R., Dall'Agnese, A. & Young, R. A. Biomolecular Condensates in the Nucleus. *Trends Biochem. Sci.* **45**, 961–977 (2020).
5. Mitrea, D. M., Mittasch, M., Gomes, B. F., Klein, I. A. & Murcko, M. A. Modulating biomolecular condensates: a novel approach to drug discovery. *Nat. Rev. Drug Discov.* 2022 1–22 (2022) doi:10.1038/s41573-022-00505-4.
6. Lyon, A. S., Peeples, W. B. & Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* 2020 223 **22**, 215–235 (2020).
7. Fujioka, Y. & Noda, N. N. Biomolecular condensates in autophagy regulation. *Curr. Opin. Cell Biol.* **69**, 23–29 (2021).
8. Boija, A., Klein, I. A. & Young, R. A. Biomolecular Condensates and Cancer. *Cancer Cell* **39**, 174–192 (2021).
9. Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688-699.e16 (2018).
10. Das, S., Lin, Y. H., Vernon, R. M., Forman-Kay, J. D. & Chan, H. S. Comparative roles of charge, π , and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 28795–28805 (2020).
11. Krainer, G. *et al.* Reentrant liquid condensate phase of proteins is stabilized by

- hydrophobic and non-ionic interactions. *Nat. Commun.* 2021 121 **12**, 1–14 (2021).
12. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* vol. 357 (2017).
 13. Ditlev, J. A., Case, L. B. & Rosen, M. K. Who's In and Who's Out—Compositional Control of Biomolecular Condensates. *J. Mol. Biol.* **430**, 4666–4684 (2018).
 14. Snead, W. T. & Gladfelter, A. S. The Control Centers of Biomolecular Phase Separation: How Membrane Surfaces, PTMs, and Active Processes Regulate Condensation. *Mol. Cell* **76**, 295–305 (2019).
 15. Li, J. *et al.* Post-translational modifications in liquid-liquid phase separation: a comprehensive review. *Mol. Biomed.* 2022 31 **3**, 1–17 (2022).
 16. Hofweber, M. & Dormann, D. Friend or foe—Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J. Biol. Chem.* **294**, 7137–7150 (2019).
 17. Ambadipudi, S., Biernat, J., Riedel, D., Mandelkow, E. & Zweckstetter, M. Liquid-liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat. Commun.* **8**, 1–13 (2017).
 18. Li, J. *et al.* Post-translational modifications in liquid-liquid phase separation: a comprehensive review. *Mol. Biomed.* 2022 31 **3**, 1–17 (2022).
 19. Tulpule, A. *et al.* Kinase-mediated RAS signaling via membraneless cytoplasmic protein granules. *Cell* **184**, 2649-2664.e18 (2021).
 20. Huang, C. *et al.* UBAP2L arginine methylation by PRMT1 modulates stress granule assembly. *Cell Death Differ.* 2019 271 **27**, 227–241 (2019).
 21. Hofweber, M. *et al.* Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173**, 706-719.e13 (2018).
 22. Saito, M. *et al.* Acetylation of intrinsically disordered regions regulates phase separation. *Nat. Chem. Biol.* **15**, 51–61 (2019).
 23. Gibson, B. A. *et al.* Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**, 470-484.e21 (2019).

24. Wang, Z., Zhang, C., Fan, C. & Liu, Y. Post-translational modifications in stress granule and their implications in neurodegenerative diseases. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1866**, 194989 (2023).
25. Sidibé, H., Dubinski, A. & Vande Velde, C. The multi-functional RNA-binding protein G3BP1 and its potential implication in neurodegenerative disease. *J. Neurochem.* **157**, 944–962 (2021).
26. Kim, N., Yun, H., Lee, H. & Yoo, J. Y. Interplay between membranes and biomolecular condensates in the regulation of membrane-associated cellular processes. *Exp. Mol. Med.* 2024 5611 **56**, 2357–2364 (2024).
27. Zhao, Y. G. & Zhang, H. Phase Separation in Membrane Biology: The Interplay between Membrane-Bound Organelles and Membraneless Condensates. *Dev. Cell* **55**, 30–44 (2020).
28. Case, L. B., Ditlev, J. A. & Rosen, M. K. Regulation of Transmembrane Signaling by Phase Separation. *Annu. Rev. Biophys.* **48**, 465–494 (2019).
29. Banjade, S. & Rosen, M. K. Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife* **3**, (2014).
30. Su, X. *et al.* Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science (80-.)*. **352**, 595–599 (2016).
31. Lin, C. C. *et al.* Receptor tyrosine kinases regulate signal transduction through a liquid-liquid phase separated state. *Mol. Cell* **82**, 1089-1106.e12 (2022).
32. Ma, W. & Mayr, C. A Membraneless Organelle Associated with the Endoplasmic Reticulum Enables 3'UTR-Mediated Protein-Protein Interactions. *Cell* **175**, 1492-1506.e19 (2018).
33. Liao, Y. C. *et al.* RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. *Cell* **179**, 147-164.e20 (2019).
34. Yu, X. *et al.* The STING phase-separator suppresses innate immune signalling. *Nat. Cell Biol.* 2021 234 **23**, 330–340 (2021).
35. Bussi, C. *et al.* Stress granules plug and stabilize damaged endolysosomal

- membranes. *Nat. 2023* 6237989 **623**, 1062–1069 (2023).
36. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science (80-.).* **361**, (2018).
 37. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171**, 163-178.e19 (2017).
 38. Zanconato, F. *et al.* Transcriptional addiction in cancer cells is mediated by YAP/TAZ through BRD4. *Nat. Med.* **24**, 1599–1610 (2018).
 39. Yang, J. *et al.* MYC phase separation selectively modulates the transcriptome. *Nat. Struct. Mol. Biol. 2024* 3110 **31**, 1567–1579 (2024).
 40. Yu, M. *et al.* Interferon- γ induces tumor resistance to anti-PD-1 immunotherapy by promoting YAP phase separation. *Mol. Cell* **81**, 1216-1230.e9 (2021).
 41. Tong, X. *et al.* Liquid–liquid phase separation in tumor biology. *Signal Transduct. Target. Ther. 2022* 71 **7**, 1–22 (2022).
 42. Grabocka, E. & Bar-Sagi, D. Mutant KRAS Enhances Tumor Cell Fitness by Upregulating Stress Granules. *Cell* **167**, 1803-1813.e12 (2016).
 43. Spann, S., Tereshchenko, M., Mastromarco, G. J., Ihn, S. J. & Lee, H. O. Biomolecular condensates in neurodegeneration and cancer. *Traffic* **20**, 890–911 (2019).
 44. Li, Y. R., King, O. D., Shorter, J. & Gitler, A. D. Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.* **201**, 361–372 (2013).
 45. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).
 46. Ryan, V. H. *et al.* Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **69**, 465-479.e7 (2018).
 47. Wolozin, B. & Ivanov, P. Stress granules and neurodegeneration. *Nat. Rev. Neurosci. 2019* 2011 **20**, 649–666 (2019).

48. Mann, J. R. *et al.* RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron* **102**, 321–338.e8 (2019).
49. Cohen, T. J. *et al.* An acetylation switch controls TDP-43 function and aggregation propensity. *Nat. Commun.* 2015 61 **6**, 1–13 (2015).
50. Kilgore, H. R. & Young, R. A. Learning the chemical grammar of biomolecular condensates. *Nat. Chem. Biol.* 2022 1–9 (2022) doi:10.1038/s41589-022-01046-y.
51. Klein, I. A. *et al.* Partitioning of cancer therapeutics in nuclear condensates. *Science* (80-.). **368**, 1386–1392 (2020).
52. Yuan, Y. *et al.* Protein lipidation in health and disease: molecular basis, physiological function and pathological implication. *Signal Transduct. Target. Ther.* 2024 91 **9**, 1–43 (2024).
53. Kallemeyn, W. W. *et al.* Proteome-wide analysis of protein lipidation using chemical probes: in-gel fluorescence visualization, identification and quantification of N-myristoylation, N- and S-acylation, O-cholesterylation, S-farnesylation and S-geranylgeranylation. *Nat. Protoc.* 2021 1611 **16**, 5083–5122 (2021).
54. Tate, E. W., Soday, L., de la Lastra, A. L., Wang, M. & Lin, H. Protein lipidation in cancer: mechanisms, dysregulation and emerging drug targets. *Nat. Rev. Cancer* 2024 244 **24**, 240–260 (2024).
55. Odeniyide, P. *et al.* Targeting farnesylation as a novel therapeutic approach in HRAS-mutant rhabdomyosarcoma. *Oncogene* 2022 4121 **41**, 2973–2983 (2022).
56. Cheng, S. *et al.* Farnesyltransferase haploinsufficiency reduces neuropathology and rescues cognitive function in a mouse model of Alzheimer disease. *J. Biol. Chem.* **288**, 35952–35960 (2013).
57. Liu, Z. *et al.* Membrane-associated farnesylated UCH-L1 promotes α -synuclein neurotoxicity and is a therapeutic target for Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4635–4640 (2009).
58. Ho, A. L. *et al.* Tipifarnib in head and neck squamous cell carcinoma with HRAS mutations. *J. Clin. Oncol.* **39**, 1856–1864 (2021).

59. Gilardi, M. *et al.* Tipifarnib as a precision therapy for HRAS-mutant head and neck squamous cell carcinomas. *Mol. Cancer Ther.* **19**, 1784–1796 (2020).
60. Dhillon, S. Lonafarnib: First Approval. *Drugs* **81**, 283–289 (2021).
61. Chamberlain, L. H. & Shipston, M. J. The physiology of protein s-acylation. *Physiol. Rev.* **95**, 341–376 (2015).
62. Busquets-Hernandez, C. *et al.* Quantitative analysis of protein lipidation and acyl-CoAs reveals substrate preferences of the S -acylation machinery. *Chem. Sci.* **15**, 12845–12855 (2024).
63. Blanc, M., David, F. P. A. & van der Goot, F. G. SwissPalm 2: Protein S-palmitoylation database. *Methods Mol. Biol.* **2009**, 203–214 (2019).
64. Ohno, Y., Kihara, A., Sano, T. & Igarashi, Y. Intracellular localization and tissue-specific distribution of human and yeast DHHC cysteine-rich domain-containing proteins. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1761**, 474–483 (2006).
65. Rana, M. S. *et al.* Fatty acyl recognition and transfer by an integral membrane S-acyltransferase. *Science (80-.).* **359**, (2018).
66. Rana, M. S. *et al.* Fatty acyl recognition and transfer by an integral membrane S-acyltransferase. *Science (80-.).* **359**, (2018).
67. Verardi, R., Kim, J. S., Ghirlando, R. & Banerjee, A. Structural Basis for Substrate Recognition by the Ankyrin Repeat Domain of Human DHHC17 Palmitoyltransferase. *Structure* **25**, 1337-1347.e6 (2017).
68. Ko, P. J. *et al.* A ZDHHC5-GOLGA7 Protein Acyltransferase Complex Promotes Nonapoptotic Cell Death. *Cell Chem. Biol.* **26**, 1716-1724.e9 (2019).
69. Woodley, K. T. & Collins, M. O. S-acylated Golga7b stabilises DHHC 5 at the plasma membrane to regulate cell adhesion . *EMBO Rep.* **20**, (2019).
70. Nguyen, P. L., Greentree, W. K., Kawate, T. & Linder, M. E. GCP16 stabilizes the DHHC9 subfamily of protein acyltransferases through a conserved C-terminal cysteine motif. *Front. Physiol.* **14**, 1167094 (2023).
71. Abrami, L. *et al.* Palmitoylated acyl protein thioesterase APT2 deforms membranes

- to extract substrate acyl chains. *Nat. Chem. Biol.* **17**, 438–447 (2021).
72. Yang, A. *et al.* Regulation of RAS palmitoyltransferases by accessory proteins and palmitoylation. *Nat. Struct. Mol. Biol.* **31**, 436–446 (2024).
 73. Dubois, F., Leroy, C., Simon, V., Benistant, C. & Roche, S. YES oncogenic activity is specified by its SH4 domain and regulates RAS/MAPK signaling in colon carcinoma cells. *Am. J. Cancer Res.* **5**, 1972 (2015).
 74. Kharbanda, A. *et al.* Blocking EGFR palmitoylation suppresses PI3K signaling and mutant KRAS lung tumorigenesis. *Sci. Signal.* **13**, 2364 (2020).
 75. Runkle, K. B. *et al.* Inhibition of DHHC20-Mediated EGFR Palmitoylation Creates a Dependence on EGFR Signaling. *Mol. Cell* **62**, 385–396 (2016).
 76. Misaki, R. *et al.* Palmitoylated Ras proteins traffic through recycling endosomes to the plasma membrane during exocytosis. *J. Cell Biol.* **191**, 23–29 (2010).
 77. Yao, H. *et al.* Inhibiting PD-L1 palmitoylation enhances T-cell immune responses against tumours. *Nat. Biomed. Eng.* 2019 34 **3**, 306–317 (2019).
 78. Yang, Y. *et al.* Palmitoylation stabilizes PD-L1 to promote breast tumor growth. *Cell Res.* 2018 291 **29**, 83–86 (2018).
 79. Wang, L. *et al.* Palmitoylation prevents sustained inflammation by limiting NLRP3 inflammasome activation through chaperone-mediated autophagy. *Mol. Cell* **83**, 281-297.e10 (2023).
 80. Zhang, M. *et al.* A STAT3 palmitoylation cycle promotes TH17 differentiation and colitis. *Nat.* 2020 5867829 **586**, 434–439 (2020).
 81. Tang, J. *et al.* Cancer cells escape p53's tumor suppression through ablation of ZDHHC1-mediated p53 palmitoylation. *Oncogene* 2021 4035 **40**, 5416–5426 (2021).
 82. Sun, Y. *et al.* ZDHHC2-Mediated AGK Palmitoylation Activates AKT–mTOR Signaling to Reduce Sunitinib Sensitivity in Renal Cell Carcinoma. *Cancer Res.* **83**, 34–51 (2023).
 83. Mao, S., Zhao, X., Wang, L., Man, Y. & Li, K. Palmitoylation-related gene ZDHHC22 as a potential diagnostic and immunomodulatory target in Alzheimer's disease:

- insights from machine learning analyses and WGCNA. *Eur. J. Med. Res.* **30**, 46 (2025).
84. Martin, D. D. O. & Sanders, S. S. Let's get fat: emergence of S-acylation as a therapeutic target in Huntington disease. *Biochem. Soc. Trans.* **52**, 1385–1392 (2024).
 85. S. Mesquita, F. *et al.* Mechanisms and functions of protein S-acylation. *Nat. Rev. Mol. Cell Biol.* 2024 256 **25**, 488–509 (2024).
 86. Schmidt, M. F., Bracha, M. & Schlesinger, M. J. Evidence for covalent attachment of fatty acids to Sindbis virus glycoproteins. *Proc. Natl. Acad. Sci.* **76**, 1687–1691 (1979).
 87. Wang, Y. & Yang, W. Proteome-Scale Analysis of Protein S-Acylation Comes of Age. *J. Proteome Res.* **20**, 14–26 (2021).
 88. Ji, Y. *et al.* Direct detection of S-palmitoylation by mass spectrometry. *Anal. Chem.* **85**, 11952–11959 (2013).
 89. Forrester, M. T. *et al.* Site-specific analysis of protein S-acylation by resin-assisted capture. *J. Lipid Res.* **52**, 393–398 (2011).
 90. Thinon, E., Fernandez, J. P., Molina, H. & Hang, H. C. Selective Enrichment and Direct Analysis of Protein S-Palmitoylation Sites. *J. Proteome Res.* **17**, 1907–1922 (2018).
 91. Martin, B. R. & Cravatt, B. F. Large-scale profiling of protein palmitoylation in mammalian cells. *Nat. Methods* 2009 62 **6**, 135–138 (2009).
 92. Foe, I. T. *et al.* Global Analysis of Palmitoylated Proteins in *Toxoplasma gondii*. *Cell Host Microbe* **18**, 501–511 (2015).
 93. Huang, B. *et al.* Palmitoylation-dependent regulation of GPX4 suppresses ferroptosis. *Nat. Commun.* 2025 161 **16**, 1–19 (2025).
 94. Ocasio, C. A. *et al.* A palmitoyl transferase chemical-genetic system to map ZDHHC-specific S-acylation. *Nat. Biotechnol.* (2024) doi:10.1038/S41587-023-02030-0.

95. Puthenveetil, R. *et al.* Orthogonal Enzyme-Substrate Design Strategy for Discovery of Human Protein Palmitoyltransferase Substrates. *J. Am. Chem. Soc.* **145**, 22287–22292 (2023).
96. Ji, G. *et al.* Global Analysis of Endogenously Intact S-Acylated Peptides Reveals Localization Differentiation of Heterogeneous Lipid Chains in Mammalian Cells. *Anal. Chem.* **95**, 13055–13063 (2023).
97. Zhang, B. *et al.* Amino acid and protein specificity of protein fatty acylation in *C. elegans*. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2307515121 (2024).
98. Davda, D. *et al.* Profiling targets of the irreversible palmitoylation inhibitor 2-bromopalmitate. *ACS Chem. Biol.* **8**, 1912–1917 (2013).
99. Pedro, M. P. *et al.* 2-Bromopalmitate Reduces Protein Deacylation by Inhibition of Acyl-Protein Thioesterase Enzymatic Activities. *PLoS One* **8**, e75232 (2013).
100. Azizi, S. A. *et al.* Development of an Acrylamide-Based Inhibitor of Protein S-Acylation. *ACS Chem. Biol.* **16**, 1546–1556 (2021).
101. Lanyon-Hogg, T., Faronato, M., Serwa, R. A. & Tate, E. W. Dynamic Protein Acylation: New Substrates, Mechanisms, and Drug Targets. *Trends Biochem. Sci.* **42**, 566–581 (2017).
102. Zhou, L. *et al.* Palmitoylation of GPX4 via the targetable ZDHHC8 determines ferroptosis sensitivity and antitumor immunity. *Nat. Cancer* **2025** 1–18 (2025) doi:10.1038/s43018-025-00937-y.
103. Wu, M. *et al.* ZDHHC3-mediated SCAP S-acylation promotes cholesterol biosynthesis and tumor immune escape in hepatocellular carcinoma. *Cell Rep.* **43**, (2024).
104. Lee, J.-Y. *et al.* A selective S-acyltransferase inhibitor suppresses tumor growth. *bioRxiv* 2024.07.18.604152 (2024) doi:10.1101/2024.07.18.604152.
105. Dekker, F. J. *et al.* Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nat. Chem. Biol.* **6**, 449–456 (2010).
106. Won, S. J. *et al.* Molecular Mechanism for Isoform-Selective Inhibition of Acyl Protein

- Thioesterases 1 and 2 (APT1 and APT2). *ACS Chem. Biol.* **11**, 3374–3382 (2016).
107. Huang, X. *et al.* S-acylation of p62 promotes p62 droplet recruitment into autophagosomes in mammalian autophagy. *Mol. Cell* **83**, 3485–3501.e11 (2023).
 108. Zou, G. *et al.* Signal-induced NLRP3 phase separation initiates inflammasome activation. *Cell Res.* 2025 1–16 (2025) doi:10.1038/s41422-025-01096-6.
 109. Hofmann, S., Kedersha, N., Anderson, P. & Ivanov, P. Molecular mechanisms of stress granule assembly and disassembly. *Biochim. Biophys. Acta - Mol. Cell Res.* **1868**, 118876 (2021).
 110. Yang, P. *et al.* G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* **181**, 325–345.e28 (2020).
 111. Glauninger, H., Wong Hickernell, C. J., Bard, J. A. M. & Drummond, D. A. Stressful steps: Progress and challenges in understanding stress-induced mRNA condensation and accumulation in stress granules. *Mol. Cell* **82**, 2544–2556 (2022).
 112. Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H. & Takekawa, M. Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nat. Cell Biol.* **10**, 1324–1332 (2008).
 113. Takahara, T. & Maeda, T. Transient Sequestration of TORC1 into Stress Granules during Heat Stress. *Mol. Cell* **47**, 242–252 (2012).
 114. Kedersha, N. & Anderson, P. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem. Soc. Trans.* **30**, 963–969 (2002).
 115. Mateju, D. *et al.* Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules. *Cell* **183**, 1801–1812.e13 (2020).
 116. Yoshioka, D., Nakamura, T., Kubota, Y. & Takekawa, M. Formation of the NLRP3 inflammasome inhibits stress granule assembly by multiple mechanisms. *J. Biochem.* **175**, 629–641 (2024).
 117. Szczerba, M. *et al.* Canonical cellular stress granules are required for arsenite-induced necroptosis mediated by Z-DNA-binding protein 1. *Sci. Signal.* **16**, (2023).
 118. Amen, T. & Kaganovich, D. Stress granules inhibit fatty acid oxidation by modulating

- mitochondrial permeability. *Cell Rep.* **35**, 109237 (2021).
119. Aulas, A. *et al.* Stress-specific differences in assembly and composition of stress granules and related foci. *J. Cell Sci.* **130**, 927–937 (2017).
 120. Markmiller, S. *et al.* Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell* **172**, 590-604.e13 (2018).
 121. Ditlev, J. A., Case, L. B. & Rosen, M. K. Who's In and Who's Out—Compositional Control of Biomolecular Condensates. *J. Mol. Biol.* **430**, 4666–4684 (2018).
 122. Gao, X. *et al.* Stress granule: A promising target for cancer treatment. *Br. J. Pharmacol.* **176**, 4421–4433 (2019).
 123. Uechi, H. *et al.* Small-molecule dissolution of stress granules by redox modulation benefits ALS models. *Nat. Chem. Biol.* 2025 1–12 (2025) doi:10.1038/s41589-025-01893-5.
 124. Wolozin, B. & Ivanov, P. Stress granules and neurodegeneration. *Nat. Rev. Neurosci.* 2019 2011 **20**, 649–666 (2019).
 125. Mackenzie, I. R. *et al.* TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* **95**, 808-816.e9 (2017).
 126. Ling, S. C. *et al.* ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13318–13323 (2010).
 127. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: The FUS about TDP-43. *Cell* **136**, 1001–1004 (2009).
 128. Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by α -Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537–1549 (2016).
 129. Song, M.-S. & Grabocka, E. Stress Granules in Cancer. 1–28 (2020) doi:10.1007/112_2020_37.
 130. Somasekharan, S. P. *et al.* YB-1 regulates stress granule formation and tumor

- progression by translationally activating G3BP1. *J. Cell Biol.* **208**, 913–929 (2015).
131. Sim, E., Irollo, E. & Grabocka, E. Evaluating stress granules in pancreatic cancer in vitro and in vivo. *Methods Mol. Biol.* **1882**, 183–195 (2019).
 132. Vilas-Boas, F. de A. S. *et al.* Impairment of stress granule assembly via inhibition of the eIF2 α phosphorylation sensitizes glioma cells to chemotherapeutic agents. *J. Neurooncol.* **127**, 253–260 (2016).
 133. Zhou, H. *et al.* Stress granules: functions and mechanisms in cancer. *Cell Biosci.* **13**, 1–13 (2023).
 134. Adjibade, P. *et al.* Sorafenib, a multikinase inhibitor, induces formation of stress granules in hepatocarcinoma cells. *Oncotarget* **6**, 43927–43943 (2015).
 135. Kaehler, C., Isensee, J., Hucho, T., Lehrach, H. & Krobitsch, S. 5-Fluorouracil affects assembly of stress granules based on RNA incorporation. *Nucleic Acids Res.* **42**, 6436–6447 (2014).
 136. McEwen, E. *et al.* Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. *J. Biol. Chem.* **280**, 16925–16933 (2005).
 137. Christopher, J. A. *et al.* Subcellular proteomics. *Nat. Rev. Methods Prim.* **2021 11** **1**, 1–24 (2021).
 138. Qin, W., Cho, K. F., Cavanagh, P. E. & Ting, A. Y. Deciphering molecular interactions by proximity labeling. *Nat. Methods* **2021 182** **18**, 133–143 (2021).
 139. Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **2016 113** **11**, 456–475 (2016).
 140. Han, Y. *et al.* Directed Evolution of Split APEX2 Peroxidase. *ACS Chem. Biol.* **14**, 619–635 (2019).
 141. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).

142. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* 2018 369 **36**, 880–887 (2018).
143. Knutson, S. D., Buksh, B. F., Huth, S. W., Morgan, D. C. & MacMillan, D. W. C. Current advances in photocatalytic proximity labeling. *Cell Chem. Biol.* **31**, 1145–1161 (2024).
144. Geri, J. B. *et al.* Microenvironment mapping via Dexter energy transfer on immune cells. *Science (80-.).* **367**, 1091–1097 (2020).
145. Nunes, C. *et al.* MSGP: the first database of the protein components of the mammalian stress granules. *Database* **2019**, 31 (2019).
146. Youn, J. Y. *et al.* High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol. Cell* **69**, 517-532.e11 (2018).
147. Marmor-Kollet, H. *et al.* Spatiotemporal Proteomic Analysis of Stress Granule Disassembly Using APEX Reveals Regulation by SUMOylation and Links to ALS Pathogenesis. *Mol. Cell* **80**, 876-891.e6 (2020).
148. Pan, C., Knutson, S. D., Huth, S. W. & MacMillan, D. W. C. μ Map proximity labeling in living cells reveals stress granule disassembly mechanisms. *Nat. Chem. Biol.* **21**, 490–500 (2024).
149. Jain, S. *et al.* ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**, 487–498 (2016).
150. Hu, S. *et al.* Time-resolved proteomic profiling reveals compositional and functional transitions across the stress granule life cycle. *Nat. Commun.* 2023 141 **14**, 1–17 (2023).
151. Marmor-Kollet, H. *et al.* Spatio-temporal Proteomic Analysis of Stress Granule Disassembly Using APEX Reveals Regulation by SUMOylation and Links to ALS Pathogenesis. *bioRxiv* 2020.01.29.830133 (2020) doi:10.1101/2020.01.29.830133.
152. Ocasio, C. A. *et al.* A palmitoyl transferase chemical–genetic system to map ZDHHC-specific S-acylation. *Nat. Biotechnol.* 2024 4210 **42**, 1548–1558 (2024).

153. Martin, B. R. & Cravatt, B. F. Large-scale profiling of protein palmitoylation in mammalian cells. *Nat. Methods* 2009 62 **6**, 135–138 (2009).
154. Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R. & Parker, R. Distinct stages in stress granule assembly and disassembly. *Elife* **5**, (2016).
155. Liu, X., Rossio, V., Gygi, S. P. & Paulo, J. A. Enriching Cysteine-Containing Peptides Using a Sulfhydryl-Reactive Alkylating Reagent with a Phosphonic Acid Group and Immobilized Metal Affinity Chromatography. *J. Proteome Res.* **22**, 1270–1279 (2023).
156. Huang, H. *et al.* Simultaneous enrichment of cysteinecontaining peptides and phosphopeptides using a cysteine-specific phosphonate adaptable tag (CysPAT) in combination with titanium dioxide (TiO₂) chromatography. *Mol. Cell. Proteomics* **15**, 3282–3296 (2016).
157. Percher, A. *et al.* Mass-tag labeling reveals site-specific and endogenous levels of protein S-fatty acylation. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 4302–4307 (2016).
158. Liu, Y. *et al.* Spatiotemporally resolved subcellular phosphoproteomics. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2025299118 (2021).
159. Yan, T. *et al.* Proximity-labeling chemoproteomics defines the subcellular cysteinome and inflammation-responsive mitochondrial redoxome. *Cell Chem. Biol.* **30**, 811-827.e7 (2023).
160. Yan, T., Boatner, L. M., Cui, L., Tontono, P. J. & Backus, K. M. Defining the Cell Surface Cysteinome Using Two-Step Enrichment Proteomics. *JACS Au* **3**, 3506–3523 (2023).
161. Qin, W. *et al.* Dynamic mapping of proteome trafficking within and between living cells by TransitID. *Cell* **186**, 3307-3324.e30 (2023).
162. Duan, Y. *et al.* PARylation regulates stress granule dynamics, phase separation, and neurotoxicity of disease-related RNA-binding proteins. *Cell Res.* 2019 293 **29**, 233–247 (2019).
163. Wang, B. *et al.* ULK1 and ULK2 Regulate Stress Granule Disassembly Through

- Phosphorylation and Activation of VCP/p97. *Mol. Cell* **74**, 742–757.e8 (2019).
164. Markmiller, S. *et al.* Active Protein Neddylation or Ubiquitylation Is Dispensable for Stress Granule Dynamics. *Cell Rep.* **27**, 1356–1363.e3 (2019).
 165. S. Mesquita, F. *et al.* Mechanisms and functions of protein S-acylation. *Nat. Rev. Mol. Cell Biol.* 2024 256 **25**, 488–509 (2024).
 166. Ernst, A. M. *et al.* S-Palmitoylation Sorts Membrane Cargo for Anterograde Transport in the Golgi. *Dev. Cell* **47**, 479–493.e7 (2018).
 167. Hancock, J. F., Magee, A. I., Childs, J. E. & Marshall, C. J. All ras Proteins Are Polyisoprenylated but Only Some Are Palmitoylated. *Cell* **57**, 1167–177 (1989).
 168. Chamberlain, L. H. & Shipston, M. J. The physiology of protein s-acylation. *Physiol. Rev.* **95**, 341–376 (2015).
 169. Chen, B. *et al.* ZDHHC7-mediated S-palmitoylation of Scribble regulates cell polarity. *Nat. Chem. Biol.* 2016 129 **12**, 686–693 (2016).
 170. Won, S. J. & Martin, B. R. Temporal Profiling Establishes a Dynamic S-Palmitoylation Cycle. *ACS Chem. Biol.* **13**, 1560–1568 (2018).
 171. Anderson, A. M. & Ragan, M. A. Palmitoylation: a protein S-acylation with implications for breast cancer. *npj Breast Cancer* 2016 21 **2**, 1–10 (2016).
 172. Sun, Y. *et al.* S-palmitoylation of PCSK9 induces sorafenib resistance in liver cancer by activating the PI3K/AKT pathway. *Cell Rep.* **40**, 111194 (2022).
 173. Vora, H. D., Johnson, M., Brea, R. J., Rudd, A. K. & Devaraj, N. K. Inhibition of NRAS Signaling in Melanoma through Direct Depalmitoylation Using Amphiphilic Nucleophiles. *ACS Chem. Biol.* **15**, 2079–2086 (2020).
 174. Ko, P.-J. & Dixon, S. J. Protein palmitoylation and cancer. *EMBO Rep.* **19**, e46666 (2018).
 175. Sardana, S., Nederstigt, A. E. & Baggelaar, M. P. S-Palmitoylation during Retinoic Acid-Induced Neuronal Differentiation of SH-SY5Y Neuroblastoma Cells. *J. Proteome Res.* **22**, 2421–2435 (2023).

176. Cooper, D. A., Jha, B. K., Silverman, R. H., Hesselberth, J. R. & Barton, D. J. Ribonuclease L and metal-ion-independent endoribonuclease cleavage sites in host and viral RNAs. *Nucleic Acids Res.* **42**, 5202–5216 (2014).
177. McClellan, B., Wilson, C. N., Brenner, A. J., Jolly, C. A. & deGraffenried, L. Flotillin-1 palmitoylation is essential for its stability and subsequent tumor promoting capabilities. *Oncogene* 2024 4314 **43**, 1063–1074 (2024).
178. Zhou, B., An, M., Freeman, M. R. & Yang, W. Technologies and Challenges in Proteomic Analysis of Protein S-acylation. *J. Proteomics Bioinform.* **7**, 256 (2014).
179. James, A. M. *et al.* Non-enzymatic N-acetylation of Lysine Residues by AcetylCoA Often Occurs via a Proximal S-acetylated Thiol Intermediate Sensitive to Glyoxalase II. *Cell Rep.* **18**, 2105–2112 (2017).
180. Tsumagari, K., Isobe, Y., Arita, M. & Imami, K. Exploring protein lipidation by mass spectrometry-based proteomics. *J. Biochem.* **175**, 225–233 (2024).
181. Zhou, B. *et al.* Low-background acyl-biotinyl exchange largely eliminates the coisolation of non- s-acylated proteins and enables deep s-acylproteomic analysis. *Anal. Chem.* **91**, 9858–9866 (2019).
182. Collins, M. O., Woodley, K. T. & Choudhary, J. S. Global, site-specific analysis of neuronal protein S-acylation. *Sci. Reports* 2017 71 **7**, 1–14 (2017).
183. Jones, M. L., Collins, M. O., Goulding, D., Choudhary, J. S. & Rayner, J. C. Analysis of protein palmitoylation reveals a pervasive role in Plasmodium development and pathogenesis. *Cell Host Microbe* **12**, 246–258 (2012).
184. Kang, R. *et al.* Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation. *Nat.* 2008 4567224 **456**, 904–909 (2008).
185. Wan, J. *et al.* Tracking brain palmitoylation change: Predominance of glial change in a mouse model of Huntington’s disease. *Chem. Biol.* **20**, 1421–1434 (2013).
186. Cai, J. *et al.* Site-Specific Identification of Protein S-Acylation by IodoTMT0 Labeling and Immobilized Anti-TMT Antibody Resin Enrichment. *J. Proteome Res.* **23**, 673–683 (2024).

187. Zhang, H. *et al.* ZDHHC20-mediated S-palmitoylation of YTHDF3 stabilizes MYC mRNA to promote pancreatic cancer progression. *Nat. Commun.* 2024 151 **15**, 1–21 (2024).
188. Tomić, G. *et al.* Palmitoyl transferase ZDHHC20 promotes pancreatic cancer metastasis. *Cell Rep.* **43**, (2024).
189. Chaturvedi, S. & Sonawane, A. Recapitulating the potential contribution of protein S-palmitoylation in cancer. *Cancer Metastasis Rev.* 2024 441 **44**, 1–26 (2024).
190. Sharma, C. *et al.* Protein acyltransferase DHHC3 regulates breast tumor growth, oxidative stress, and senescence. *Cancer Res.* **77**, 6880–6890 (2017).
191. Roth, A. F. *et al.* Global Analysis of Protein Palmitoylation in Yeast. *Cell* **125**, 1003–1013 (2006).
192. Yang, W., Di Vizio, D., Kirchner, M., Steen, H. & Freeman, M. R. Proteome Scale Characterization of Human S-Acylated Proteins in Lipid Raft-enriched and Non-raft Membranes. *Mol. Cell. Proteomics* **9**, 54–70 (2010).
193. Zhang, X. *et al.* Ultradeep Palmitoylomics Enabled by Dithiodipyridine-Functionalized Magnetic Nanoparticles. *Anal. Chem.* **90**, 6161–6168 (2018).
194. Forrester, M. T., Egol, J. R., Tata, A., Tata, P. R. & Foster, M. W. Analysis of Protein Cysteine Acylation Using a Modified Suspension Trap (Acyl-Trap). *J. Proteome Res.* **23**, 3716–3725 (2024).
195. Thinon, E., Fernandez, J. P., Molina, H. & Hang, H. C. Selective Enrichment and Direct Analysis of Protein S-Palmitoylation Sites. *J. Proteome Res.* **17**, 1907–1922 (2018).
196. Kleinpenning, F., Steigenberger, B., Wu, W. & Heck, A. J. R. Fishing for newly synthesized proteins with phosphonate-handles. *Nat. Commun.* 2020 111 **11**, 1–10 (2020).
197. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* 2017 145 **14**, 513–520 (2017).

198. Thingholm, T. E., Jensen, O. N. & Larsen, M. R. Analytical strategies for phosphoproteomics. *Proteomics* **9**, 1451–1468 (2009).
199. Arribas Diez, I. *et al.* Zirconium(IV)-IMAC Revisited: Improved Performance and Phosphoproteome Coverage by Magnetic Microparticles for Phosphopeptide Affinity Enrichment. *J. Proteome Res.* **20**, 453–462 (2021).
200. Bortel, P. *et al.* Systematic Optimization of Automated Phosphopeptide Enrichment for High-Sensitivity Phosphoproteomics. *Mol. Cell. Proteomics* **23**, 100754 (2024).
201. Sharma, K. *et al.* Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* **8**, 1583–1594 (2014).
202. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* 2019 171 **17**, 41–44 (2019).
203. Hung, C. W. *et al.* Characterization of post-translational modifications in full-length human BMP-1 confirms the presence of a rare vicinal disulfide linkage in the catalytic domain and highlights novel features of the EGF domain. *J. Proteomics* **138**, 136–145 (2016).
204. Kleinberg, A., Mao, Y. & Li, N. Practical solutions for overcoming artificial disulfide scrambling in the non-reduced peptide mapping characterization of monoclonal antibodies. *MAbs* **16**, 2420805 (2024).
205. Sada, R. *et al.* Dynamic palmitoylation controls the microdomain localization of the DKK1 receptors CKAP4 and LRP6. *Sci. Signal.* **12**, 9519 (2019).
206. Morrow, I. C. *et al.* Flotillin-1/reggie-2 traffics to surface raft domains via a novel Golgi-independent pathway. Identification of a novel membrane targeting domain and a role for palmitoylation. *J. Biol. Chem.* **277**, 48834–48841 (2002).
207. Jang, D., Kwon, H., Jeong, K., Lee, J. & Pak, Y. Essential role of flotillin-1 palmitoylation in the intracellular localization and signaling function of IGF-1 receptor. *J. Cell Sci.* **128**, 2179–2190 (2015).
208. Lin, W. C. *et al.* H-Ras forms dimers on membrane surfaces via a protein - Protein

- interface. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2996–3001 (2014).
209. Chen, X. *et al.* EZH2 Palmitoylation Mediated by ZDHHC5 in p53-Mutant Glioma Drives Malignant Development and Progression. *Cancer Res.* **77**, 4998–5010 (2017).
 210. Liu, J. *et al.* Lipogenic enzyme FASN promotes mutant p53 accumulation and gain-of-function through palmitoylation. *Nat. Commun.* **2025 161 16**, 1–18 (2025).
 211. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019 101 10**, 1–10 (2019).
 212. Landrum, M. J. *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980–D985 (2014).
 213. Chamberlain, L. H. & Shipston, M. J. The physiology of protein s-acylation. *Physiol. Rev.* **95**, 341–376 (2015).
 214. Rodenburg, R. N. P. *et al.* Stochastic palmitoylation of accessible cysteines in membrane proteins revealed by native mass spectrometry. *Nat. Commun.* **2017 81 8**, 1–9 (2017).
 215. Wu, R. *et al.* Simultaneous and site-specific profiling of heterogeneity and turnover in protein S-acylation by intact S-acylated peptide analysis with a cleavable bioorthogonal tag. *Analyst* **149**, 1111–1120 (2024).
 216. Cheng, A., Grant, C. E., Noble, W. S. & Bailey, T. L. MoMo: discovery of statistically significant post-translational modification motifs. *Bioinformatics* **35**, 2774–2782 (2019).
 217. White, M. E. H., Gil, J. & Tate, E. W. Proteome-wide structural analysis identifies warhead- and coverage-specific biases in cysteine-focused chemoproteomics. *Cell Chem. Biol.* **30**, 828-838.e4 (2023).
 218. Edmonds, M. J., Geary, B., Doherty, M. K. & Morgan, A. Analysis of the brain palmitoyl-proteome using both acyl-biotin exchange and acyl-resin-assisted capture methods. *Sci. Reports* **2017 71 7**, 1–13 (2017).
 219. Ji, G. *et al.* Global analysis of N-myristoylation and its heterogeneity by combining N-terminomics and nanographite fluoride-based solid-phase extraction. *Talanta*

- 276**, 126300 (2024).
220. van Bergen, W., Nederstigt, A. E., Heck, A. J. R. & Baggelaar, M. P. Site-Specific Competitive Kinase Inhibitor Target Profiling Using Phosphonate Affinity Tags. *Mol. Cell. Proteomics* **24**, 100906 (2025).
 221. van Bergen, W., Hevler, J. F., Wu, W., Baggelaar, M. P. & Heck, A. J. R. Site-Specific Activity-Based Protein Profiling Using Phosphonate Handles. *Mol. Cell. Proteomics* **22**, 100455 (2023).
 222. Ibáñez-Vea, M. *et al.* Characterization of Macrophage Endogenous S-Nitrosoproteome Using a Cysteine-Specific Phosphonate Adaptable Tag in Combination with TiO₂ Chromatography. *J. Proteome Res.* **17**, 1172–1182 (2018).
 223. Greaves, J. *et al.* Molecular basis of fatty acid selectivity in the zDHHC family of S-acyltransferases revealed by click chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1365–E1374 (2017).
 224. Aramsangtienchai, P., Spiegelman, N. A., Cao, J. & Lin, H. S-Palmitoylation of Junctional Adhesion Molecule C Regulates Its Tight Junction Localization and Cell Migration. *J. Biol. Chem.* **292**, 5325–5334 (2017).
 225. Jin, L. *et al.* A comparative study of evaluating missing value imputation methods in label-free proteomics. *Sci. Reports 2021 111* **11**, 1–11 (2021).
 226. Noland, C. L. *et al.* Palmitoylation of TEAD Transcription Factors Is Required for Their Stability and Function in Hippo Pathway Signaling. *Structure* **24**, 179–186 (2016).
 227. Chan, P. *et al.* Autopalmitoylation of TEAD proteins regulates transcriptional output of the Hippo pathway. *Nat. Chem. Biol.* 2016 124 **12**, 282–289 (2016).
 228. Noritsugu, K. *et al.* Lysine long-chain fatty acylation regulates the TEAD transcription factor. *Cell Rep.* **42**, (2023).
 229. Zagiel, B., Melnyk, P. & Cotellet, P. Progress with YAP/TAZ-TEAD inhibitors: a patent review (2018-present). *Expert Opin. Ther. Pat.* **32**, 899–912 (2022).
 230. James, A. M. *et al.* Nucleotide-binding sites can enhance N-acylation of nearby protein lysine residues. *Sci. Reports 2020 101* **10**, 1–13 (2020).

231. Keenan, E. K., Bareja, A., Lam, Y., Grimsrud, P. A. & Hirschey, M. D. Cysteine S-acetylation is a post-translational modification involved in metabolic regulation. *bioRxiv* 2024.05.21.595030 (2024) doi:10.1101/2024.05.21.595030.
232. Cazzalini, O. *et al.* CBP and p300 acetylate PCNA to link its degradation with nucleotide excision repair synthesis. *Nucleic Acids Res.* **42**, 8433–8448 (2014).
233. Squair, D. R. & Virdee, S. A new dawn beyond lysine ubiquitination. *Nat. Chem. Biol.* **2022 188 18**, 802–811 (2022).
234. Gao, P. *et al.* E3 ligase Nedd4l promotes antiviral innate immunity by catalyzing K29-linked cysteine ubiquitination of TRAF3. *Nat. Commun.* **2021 121 12**, 1–13 (2021).
235. Zhang, Q. *et al.* Reprogramming of palmitic acid induced by dephosphorylation of ACOX1 promotes β -catenin palmitoylation to drive colorectal cancer progression. *Cell Discov.* **2023 91 9**, 1–19 (2023).
236. Di Vizio, D. *et al.* Caveolin-1 interacts with a lipid raft-associated population of fatty acid synthase. *Cell Cycle* **7**, 2257–2267 (2008).
237. Choi, M., Lee, J., Jeong, K. & Pak, Y. Caveolin-2 palmitoylation turnover facilitates insulin receptor substrate-1-directed lipid metabolism by insulin receptor tyrosine kinase. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1870**, 167173 (2024).
238. Anderson, A. M. & Ragan, M. A. Palmitoylation: a protein S-acylation with implications for breast cancer. *npj Breast Cancer* **2016 21 2**, 1–10 (2016).
239. Zhang, Q. *et al.* ZDHHC1 downregulates LIPG and inhibits colorectal cancer growth via IGF2BP1 Palmitoylation. *Cancer Gene Ther.* **2024 319 31**, 1427–1437 (2024).
240. Shan, J. *et al.* Palmitoyltransferase ZDHHC6 promotes colon tumorigenesis by targeting PPAR γ -driven lipid biosynthesis via regulating lipidome metabolic reprogramming. *J. Exp. Clin. Cancer Res.* **43**, 1–27 (2024).
241. Liu, J. *et al.* Forkhead box C1 promotes colorectal cancer metastasis through transactivating ITGA7 and FGFR4 expression. *Oncogene* **2018 3741 37**, 5477–5491 (2018).
242. Han, H., Feng, P. & Yuan, G. Molecular typing of gliomas on the basis of integrin

- family genes and a functional study of ITGA7. *Sci. Reports* 2025 151 **15**, 1–16 (2025).
243. Ming, X. Y. *et al.* Integrin $\alpha 7$ is a functional cancer stem cell surface marker in oesophageal squamous cell carcinoma. *Nat. Commun.* 2016 71 **7**, 1–14 (2016).
 244. Sharma, C., Rabinovitz, I. & Hemler, M. E. Palmitoylation by DHHC3 is critical for the function, expression, and stability of integrin $\alpha 6\beta 4$. *Cell. Mol. Life Sci.* **69**, 2233–2244 (2012).
 245. Yang, X. *et al.* Palmitoylation supports assembly and function of integrin–tetraspanin complexes. *J. Cell Biol.* **167**, 1231–1240 (2004).
 246. Aramsangtienchai, P., Spiegelman, N. A., Cao, J. & Lin, H. S-palmitoylation of junctional adhesion molecule C regulates its tight junction localization and cell migration. *J. Biol. Chem.* **292**, 5325–5334 (2017).
 247. Zhang, Y. *et al.* Palmitoylation of GNAQ/11 is critical for tumor cell proliferation and survival in GNAQ/11-mutant uveal melanoma. *Front. Med.* **16**, 784–798 (2022).
 248. Brown, R. W. B., Sharma, A. I. & Engman, D. M. Dynamic protein S-palmitoylation mediates parasite life cycle progression and diverse mechanisms of virulence. *Crit. Rev. Biochem. Mol. Biol.* **52**, 145–162 (2017).
 249. Ritzeveld, M., Wright, M. H. & Tate, E. W. New developments in probing and targeting protein acylation in malaria, leishmaniasis and African sleeping sickness. *Parasitology* **145**, 157–174 (2018).
 250. Schlott, A. C., Holder, A. A. & Tate, E. W. N -Myristoylation as a Drug Target in Malaria: Exploring the Role of N -Myristoyltransferase Substrates in the Inhibitor Mode of Action. *ACS Infect. Dis.* **4**, 449–457 (2018).
 251. Robert-Gangneux, F. & Dardé, M. L. Epidemiology of and diagnostic strategies for toxoplasmosis. *Clin. Microbiol. Rev.* **25**, 264–296 (2012).
 252. Caballero, M. C. *et al.* Identification of new palmitoylated proteins in *Toxoplasma gondii*. *Biochim. Biophys. Acta - Proteins Proteomics* **1864**, 400–408 (2016).
 253. Alonso, A. M. *et al.* Protein palmitoylation inhibition by 2-bromopalmitate alters gliding, host cell invasion and parasite morphology in *Toxoplasma gondii*. *Mol.*

- Biochem. Parasitol.* **184**, 39–43 (2012).
254. Child, M. A. *et al.* Small-molecule inhibition of a depalmitoylase enhances Toxoplasma host-cell invasion. *Nat. Chem. Biol.* **2013 910** **9**, 651–656 (2013).
 255. Yu, P. *et al.* Pyroptosis: mechanisms and diseases. *Signal Transduct. Target. Ther.* **2021 61** **6**, 1–21 (2021).
 256. Du, G. *et al.* ROS-dependent S-palmitoylation activates cleaved and intact gasdermin D. *Nat.* **2024 6308016** **630**, 437–446 (2024).
 257. Balasubramanian, A. *et al.* The palmitoylation of gasdermin D directs its membrane translocation and pore formation during pyroptosis. *Sci. Immunol.* **9**, (2024).
 258. Liu, Z. *et al.* Palmitoylation at a conserved cysteine residue facilitates gasdermin D-mediated pyroptosis and cytokine release. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2400883121 (2024).
 259. Zhang, N. *et al.* A palmitoylation–depalmitoylation relay spatiotemporally controls GSDMD activation in pyroptosis. *Nat. Cell Biol.* **2024 265** **26**, 757–769 (2024).
 260. ZHUANG, Z., GU, J., LI, B. O. & YANG, L. I. N. G. Inhibition of gasdermin D palmitoylation by disulfiram is crucial for the treatment of myocardial infarction. *Transl. Res.* **264**, 66–75 (2024).
 261. Hu, J. J. *et al.* FDA-approved disulfiram inhibits pyroptosis by blocking gasdermin D pore formation. *Nat. Immunol.* **2020 217** **21**, 736–745 (2020).
 262. Humphries, F. *et al.* Succination inactivates gasdermin D and blocks pyroptosis. *Science (80-.).* **369**, (2020).
 263. Gutiérrez, T. *et al.* The ER chaperone calnexin controls mitochondrial positioning and respiration. *Sci. Signal.* **13**, 6660 (2020).
 264. Lakkaraju, A. K. K. *et al.* Palmitoylated calnexin is a key component of the ribosome-translocon complex. *EMBO J.* **31**, 1823–1835 (2012).
 265. Zheng, S. *et al.* ZDHHC5-mediated NLRP3 palmitoylation promotes NLRP3-NEK7 interaction and inflammasome activation. *Mol. Cell* **83**, 4570-4585.e7 (2023).

266. Nie, L. *et al.* Consecutive palmitoylation and phosphorylation orchestrates NLRP3 membrane trafficking and inflammasome activation. *Mol. Cell* **84**, 3336–3353.e7 (2024).
267. Yu, T. *et al.* NLRP3 Cys126 palmitoylation by ZDHHC7 promotes inflammasome activation. *Cell Rep.* **43**, (2024).
268. Kennedy, C. R. *et al.* A Probe for NLRP3 Inflammasome Inhibitor MCC950 Identifies Carbonic Anhydrase 2 as a Novel Target. *ACS Chem. Biol.* **16**, 982–990 (2021).
269. Coll, R. C. *et al.* MCC950 directly targets the NLRP3 ATP-hydrolysis motif for inflammasome inhibition. *Nat. Chem. Biol.* **15**, 556–559 (2019).
270. Fang, C. *et al.* Identification of Palmitoylated Transitional Endoplasmic Reticulum ATPase by Proteomic Technique and Pan Antipalmitoylation Antibody. *J. Proteome Res.* **15**, 956–962 (2016).
271. Wang, Z. *et al.* AMPK α 1-mediated ZDHHC8 phosphorylation promotes the palmitoylation of SLC7A11 to facilitate ferroptosis resistance in glioblastoma. *Cancer Lett.* **584**, 216619 (2024).
272. Shi, Z. *et al.* Loss of LncRNA DUXAP8 synergistically enhanced sorafenib induced ferroptosis in hepatocellular carcinoma via SLC7A11 de-palmitoylation. *Clin. Transl. Med.* **13**, e1300 (2023).
273. Liu, Z. *et al.* Palmitoylation at a conserved cysteine residue facilitates gasdermin D-mediated pyroptosis and cytokine release. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2400883121 (2024).
274. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520 (2017).
275. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2019).
276. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011).

277. Chen, H. & Boutros, P. C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 1–7 (2011).
278. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. doi:10.1093/bioinformatics/btx364.
279. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 2016 139 **13**, 731–740 (2016).
280. Ge, S. X., Jung, D., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).

Chapter 7.

Appendix

Table 7.1. Stress-dependent G3BP1 interactors identified by APEX-AX

Gene	Protein.ID	Difference	-LogP	SG Match	SwissPalm Match	Stress-dependent? (MarkMiller et al 2018 ¹²⁰)
IGF2BP3	O00425	1.680973	6.07541	SG protein	Yes	No
HNRNPDL	O14979	1.208448	9.33041	SG protein	Yes	Yes
MSI1	O43347	1.534221	3.976915	SG protein	Yes	No
EIF4G3	O43432	1.112834	3.392473	SG protein	No	No
SYNCRIP	O60506	1.118699	7.991499	SG protein	Yes	No
PTBP3	O95758	1.600644	4.839144	SG protein	Yes	Yes
PTBP1	P26599	1.015034	8.224742	SG protein	Yes	Yes
RBMS1	P29558	1.072577	2.676062	SG protein	No	Yes
TIA1	P31483	1.195672	1.950682	SG protein	Yes	No
HNRNPH3	P31942	1.038869	5.387319	SG protein	Yes	No
FUS	P35637	1.011611	7.233278	SG protein	Yes	Yes
EIF4A3	P38919	1.453755	4.272636	SG protein	Yes	No
YARS1	P54577	1.098808	1.507783	SG protein	Yes	No
EIF4A1	P60842	1.279309	10.24598	SG protein	Yes	Yes
SRSF1	Q07955	1.603371	5.105894	SG protein	Yes	Yes
TRIM26	Q12899	1.74006	2.072446	No	Yes	No
ELAVL2	Q12926	2.59143	11.54669	SG protein	Yes	No
SRSF9	Q13242	1.91124	5.49022	SG protein	Yes	Yes
HNRNPD	Q14103	1.300386	8.017828	SG protein	Yes	No
EIF4A2	Q14240	1.29854	8.878475	SG protein	Yes	Yes
EIF4H	Q15056	1.121165	7.925508	SG protein	Yes	Yes
ELAVL1	Q15717	1.729414	8.391606	SG protein	Yes	Yes
LSM12	Q3MHD2	1.189031	1.952887	SG protein	No	Yes
EIF3M	Q7L2H7	1.086063	1.647533	SG protein	Yes	No
DDX1	Q92499	1.099256	8.441535	SG protein	Yes	Yes
CELF1	Q92879	1.470348	10.40402	SG protein	Yes	Yes
WBP2	Q969T9	1.135276	2.565014	SG protein	No	No
HNRNPAB	Q99729	1.249681	6.545543	SG protein	Yes	Yes
LARP6	Q9BRS8	1.231988	2.390258	SG protein	No	No
RBM24	Q9BX46	1.114734	2.499205	No	No	No
C11orf68	Q9H3H3	1.115286	2.229003	No	No	No
UPF2	Q9HAU5	1.40321	1.721565	SG protein	Yes	No
XPO5	Q9HAV4	1.487537	3.24482	No	Yes	No
MOV10	Q9HCE1	1.256741	10.64582	SG protein	Yes	No
IGF2BP1	Q9NZI8	1.528167	8.826955	SG protein	Yes	No
NXF1	Q9UBU9	1.941952	4.17365	No	Yes	No
G3BP2	Q9UN86	2.854427	2.556313	SG protein	Yes	No
YBX2	Q9Y2T7	1.171951	3.967259	No	No	No
IGF2BP2	Q9Y6M1	1.745778	12.42276	SG protein	Yes	Yes

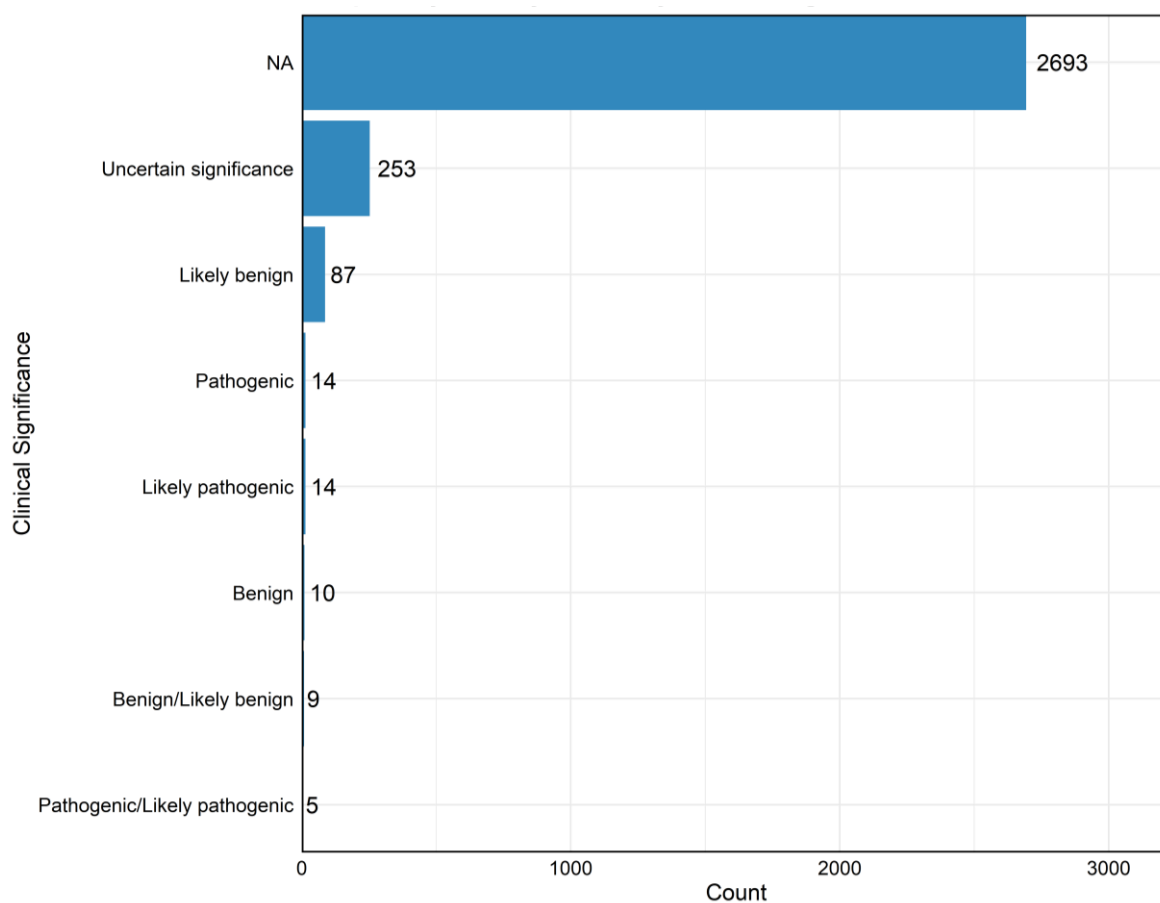


Figure 7.1. Clinical variant annotation of putative *S*-acylation sites in HEK293T cells identified using SAPPHIRE. Most identified Cys are not annotated (NA), however 14 Cys are associated with pathogenic variants. Residue-specific clinical annotations were obtained from ClinVar²¹².

Table 7.2. Putative S-acylated Cysteines identified by SAPPHIRE in HEK293T cells associated with pathogenic clinical variants

UniProt ID	Gene	Type	Clinical Significance	Residue Canonical	Residue position	Residue mutant
P51648	ALDH3A2	single nucleotide variant	Pathogenic	C	214	Y
P35520	CBS	single nucleotide variant	Pathogenic	C	165	Y
Q6PI48	DARS2	single nucleotide variant	Pathogenic	C	152	F
P26196	DDX6	single nucleotide variant	Pathogenic	C	390	R
Q86SX6	GLRX5	single nucleotide variant	Pathogenic	C	67	Y
P49773	HINT1	single nucleotide variant	Pathogenic	C	84	R
Q9BX67	JAM3	single nucleotide variant	Pathogenic	C	219	Y
P43246	MSH2	single nucleotide variant	Pathogenic	C	199	R
P61916	NPC2	single nucleotide variant	Pathogenic	C	99	R
P04181	OAT	single nucleotide variant	Pathogenic	C	93	F
P00558	PGK1	single nucleotide variant	Pathogenic	C	316	R
O15269	SPTLC1	single nucleotide variant	Pathogenic	C	133	Y
P60174	TPI1	single nucleotide variant	Pathogenic	C	42	Y
P07437	TUBB	single nucleotide variant	Pathogenic	C	239	S

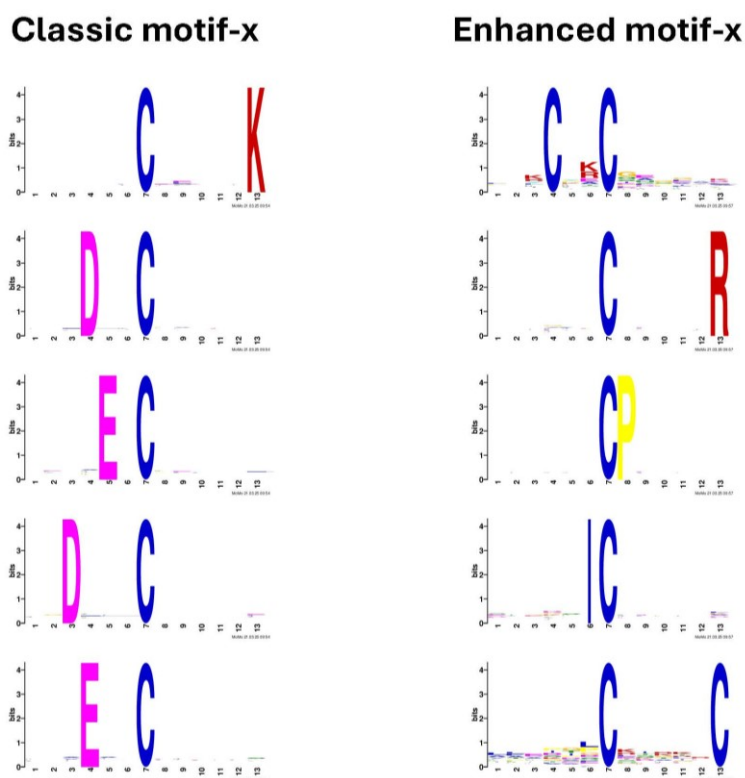


Figure 7.2. S-acylation conserved sequence motif analysis. Using the putative S-acylated Cysteine sites identified in HEK293T cells by SAPPHERE. Consensus sequences was analysed using MoMo Software using both the classic motif-x algorithm, whereby the human proteome is used as the background and the enhanced motif-x algorithm where the foreground peptides are shuffled and used as background.

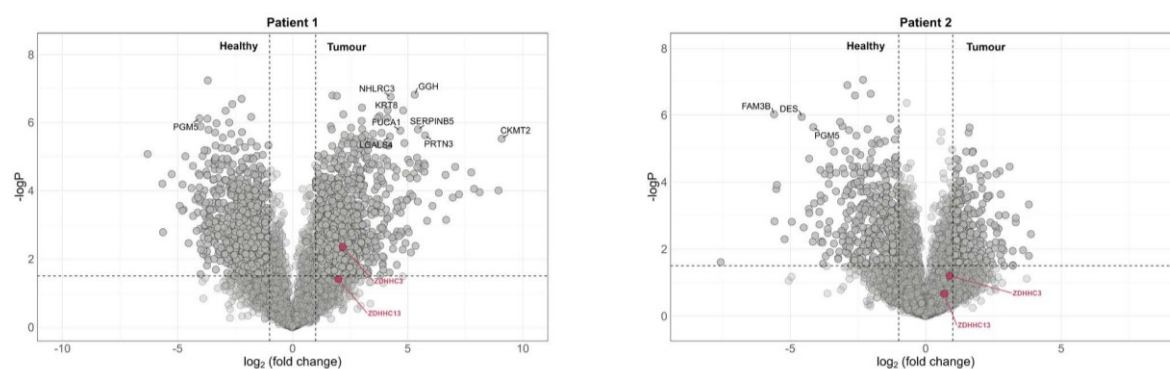
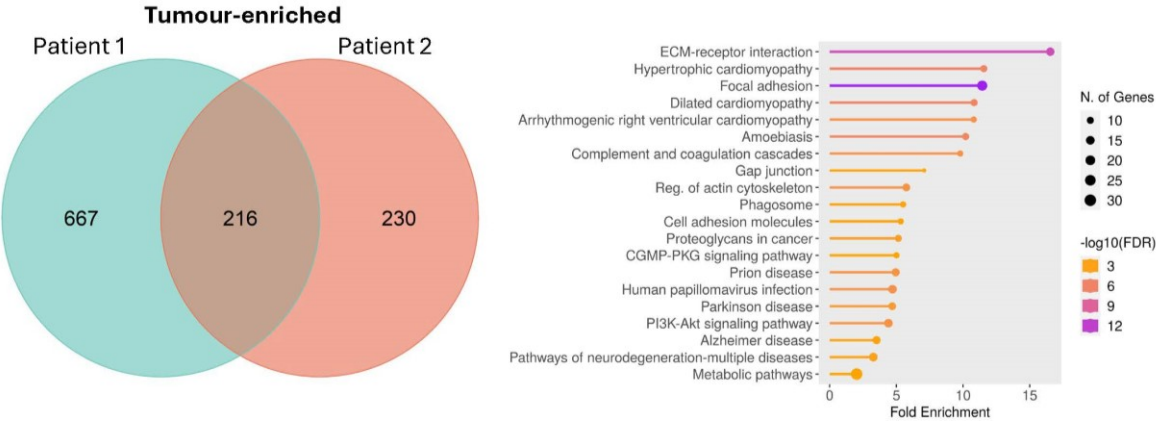


Figure 7.3. Whole proteome analysis of colorectal cancer patient tissues. Volcano plot showing tumour and healthy-enriched proteins (student's t-test; FDR = 0.05, S0 = 0.1)

a



b

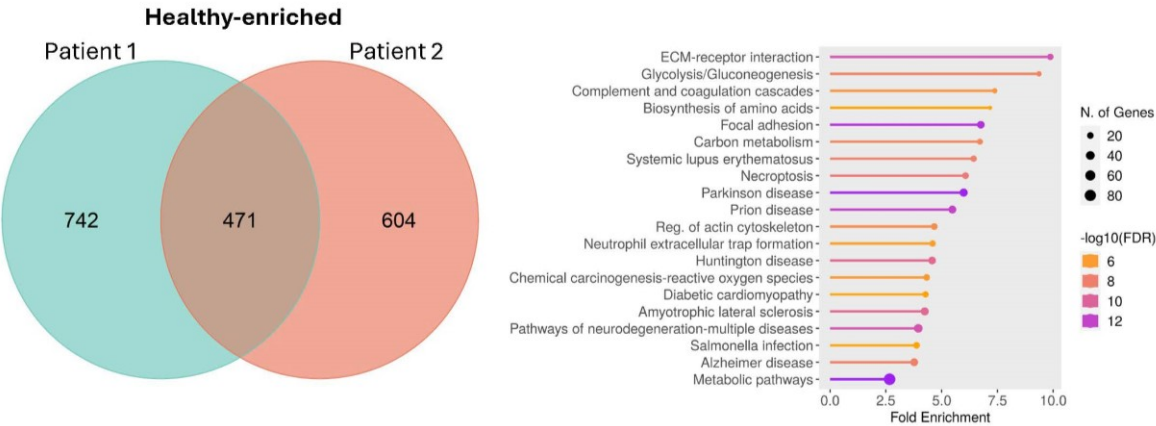


Figure 7.4 Global protein expression changes between healthy and tumour tissues of colorectal cancer patients. a) Venn diagram showing overlap between the tumour-enriched S-acylated sites between the two patients (left). KEGG pathway enrichment analysis of proteins corresponding to the overlapped sites. **b)** Venn diagram showing overlap between the healthy tissue-enriched S-acylated sites between the two patients (left). KEGG pathway enrichment analysis of proteins corresponding to the overlapped sites.

Table 7.3. Peptide quantification for CRC samples

Sample	+HA1	+HA2	+HA3	-HA1	-HA2	-HA3
Patient 1 Healthy	11.926	13.228	18.673	12.742	11.887	13.87
Patient 1 Tumour	-0.848	-0.869	-1.715	-1.226	-0.617	1.486
Patient 2 Healthy	6.628	9.109	4.507	13.552	12.4	14.893
Patient 2 Tumour	16.051	21.013	10.438	5.755	2.623	5.827

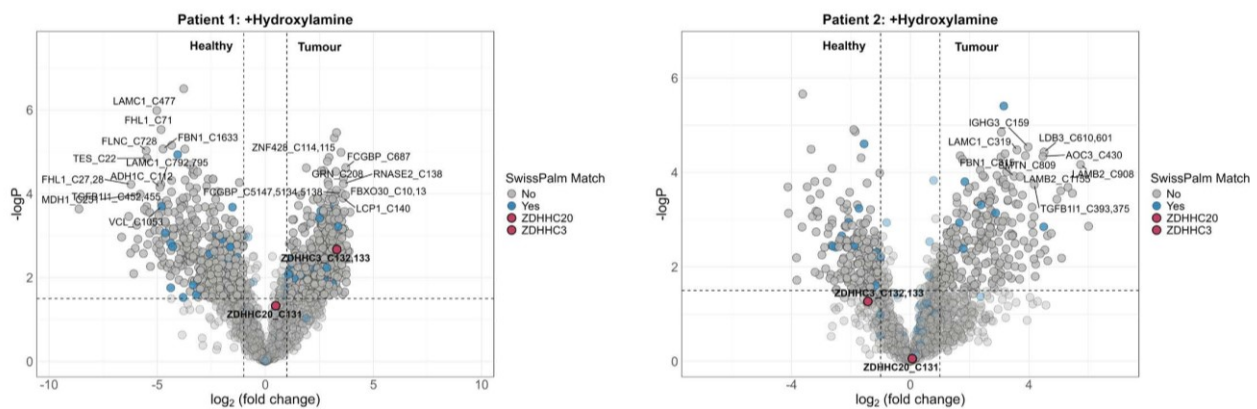


Figure 7.5 SAPPHERE identification of colorectal cancer tumour-enriched S-acylation sites. Volcano plot depicting tumour-enriched S-acylated sites (student's t-test; FDR = 0.05, S0 = 0.1) The blue points denote previous identification of the exact site in the SwissPalm database. Data was not normalised to protein expression.

Table 7.4 Top 10 colorectal cancer tumour-enriched S-acylated sites ranked by mean $\log_2$ FoldChange

Protein ID	Gene	Cys Position	$\log_2$ FC(Tumour/Healthy) Patient 1	$\log_2$ FC (Tumour/Healthy) Patient 2	Mean FC
Q13683	ITGA7	615	7.907076677	5.591089924	6.749083
P11047	LAMC1	884,886	4.415199866	8.205833197	6.310517
Q13683	ITGA7	546	5.909143329	6.380723198	6.144933
Q13683	ITGA7	546,539	6.136376739	5.830489953	5.983433
Q13683	ITGA7	1006	5.540376544	6.180656672	5.860517
Q14315	FLNC	438	6.387033502	5.186466654	5.78675
Q9BX67	JAM3	53	6.325530052	4.044723312	5.185127
Q13683	ITGA7	1001	5.294443329	5.045823306	5.170133
P35443	THBS4	403	6.919933399	3.277356716	5.098645
Q14315	FLNC	1639	5.232899964	4.940733393	5.086817

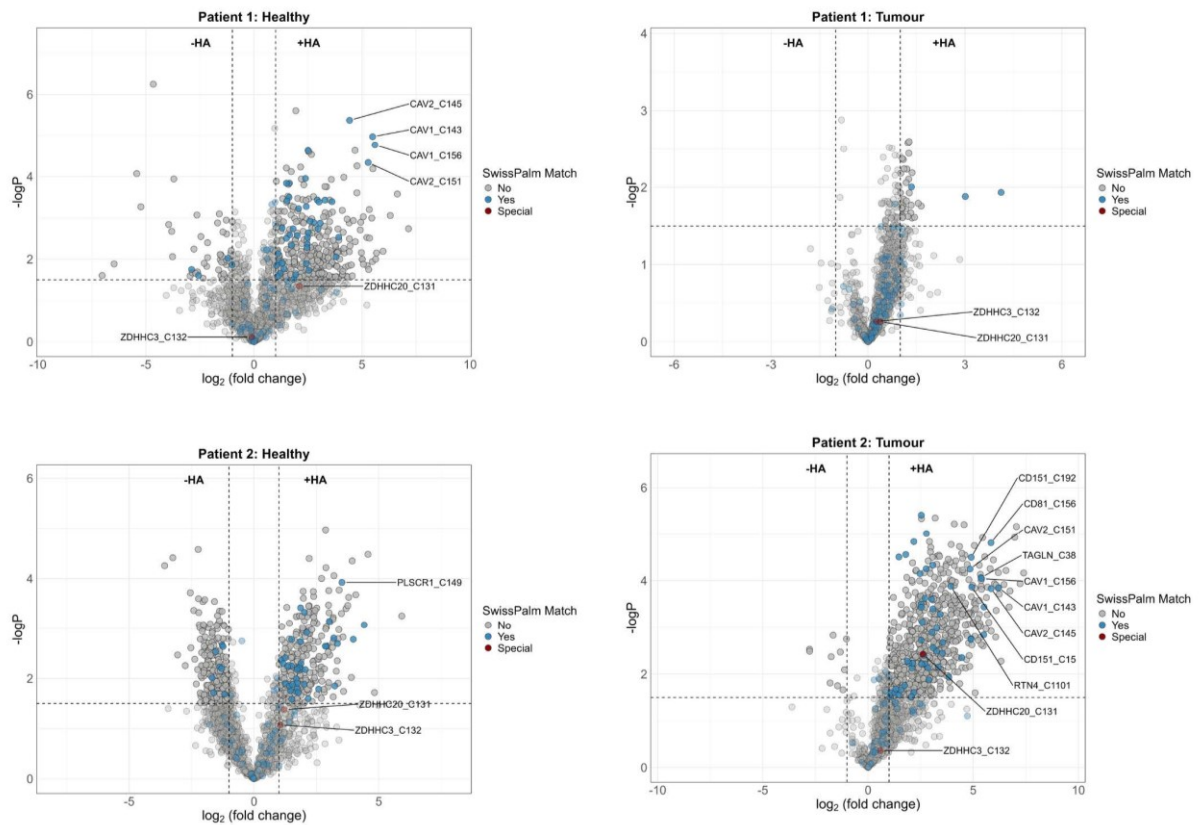


Figure 7.6. Site-specific S-acylated proteome profiling of colorectal cancer patient tissue with SAPPHERE. Volcano plots depicting differential enrichment of phosphonated Cys peptides (i.e. putative S-acylated Cys) between +HA and -HA healthy and tumour tissues (student's t-test; FDR = 0.05, S0 = 0.1). The blue points denote previous identification of the exact site in the SwissPalm database.

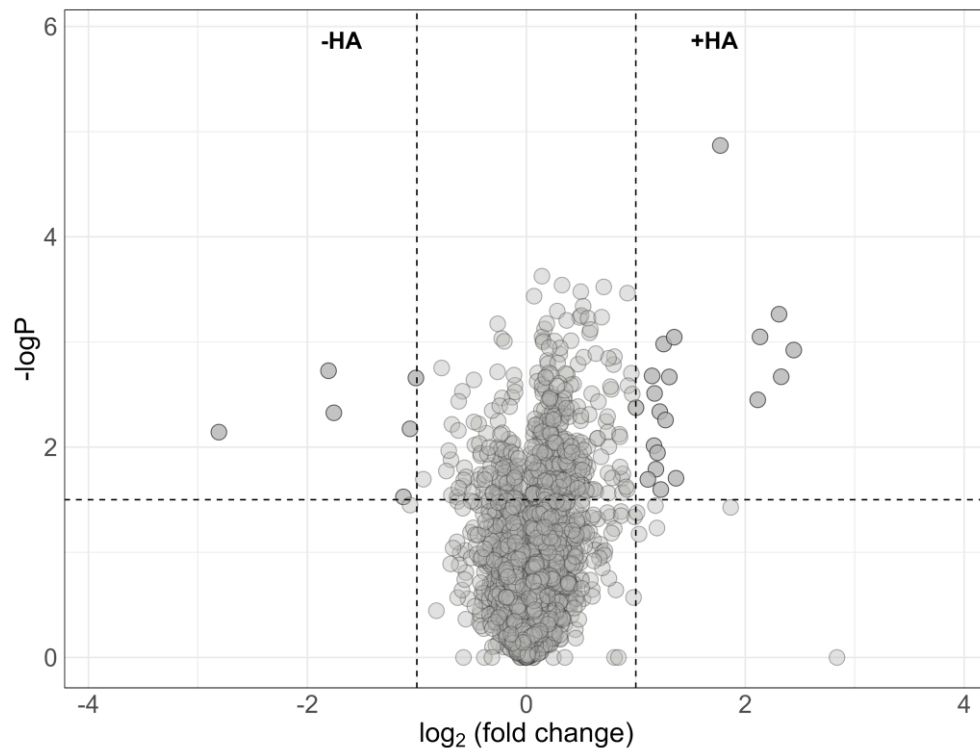


Figure 7.7. Whole proteome analysis of *Toxoplasma gondii*. Volcano plot showing differentially enriched proteins between +HA and -HA treated lysates (student's t-test; FDR = 0.05, S0 = 0.1)

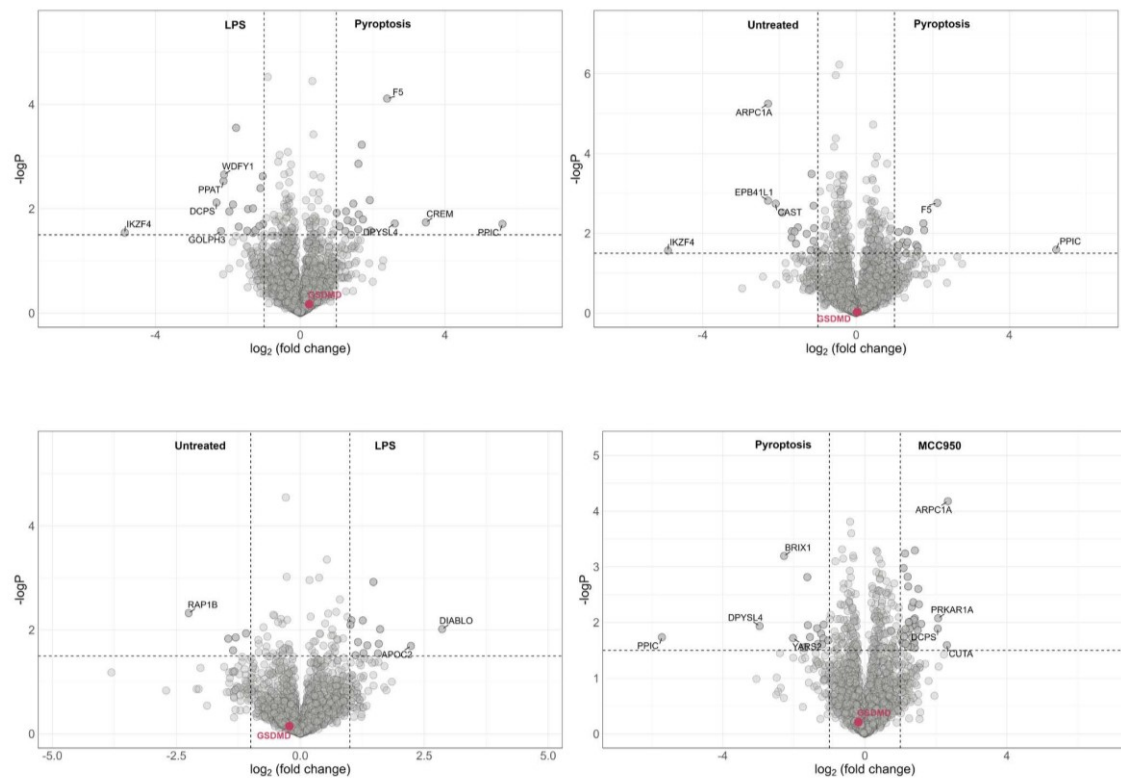


Figure 7.8. Whole proteome analysis of THP1 cells in response to pyroptosis. Volcano plot showing differentially enriched proteins between THP1 cells that were left untreated, LPS-primed only, LPS+nigericin treated to induced pyroptosis and LPS+MCC950+Nigericin treated (student's t-test; FDR = 0.05, S0 = 0.1).

