A proteomics investigation into the role of zDHHC23 and MROH6 in neuroblastoma

Thesis submitted in accordance with the requirements of the

University of Liverpool

For the degree of Doctor in Philosophy by

Sally Olivia Oswald

February 2023

Acknowledgments

I am beyond indebted to my supervisors Prof. Claire Eyers and Prof. Ian Prior who kindly adopted me as a PhD student late into my studies. Thank you for your support and guidance whilst also having to quickly navigate and adapt to my project. I will miss our zoom sessions with the added bonus of the Eyers and Prior dogs joining in!

I am particularly in gratitude to Claire. Thank you for giving me the opportunity to work in your lab, to gain hands on experience with mass spectrometry (trusting me with expensive instruments!) and to attend an international conference. Despite this project being a massive learning curve and as I would always say, feeling 'thick', you were tremendously supportive, and I feel so fortunate to have had you as my supervisor. You have not only had a substantial positive impact on my development as a scientist, but also turned what was an arduous start to my PhD into an enjoyable experience; one I will cherish.

I would also like to thank past and present members of Lab A for welcoming me into the group and the mass spectrometry world. I will fondly remember our '10am time for coffee?' Costa/Nero trips. I also cannot forget Dom in Lab D, thank you for sharing your infinite wisdom and consumables. I would also like to thank Centre for Cell Imaging: particularly Jen, Marie and Tom for all their help and guidance in the mind-boggling and patience testing field of microscopy.

To my family, the Oswald's and the Hillerby's, living 150 miles away who would not see me for months at a time because I was too busy and stressed doing lab work. You may not have understood what I was doing, just that I wore a lab coat, but thank you for your love, support, patience and FaceTime sessions. My dear nephews, Ted and Seth, I have been envious of everyone else seeing you so regularly, you have grown up far too quickly.

Lastly, I cannot thank Leonard enough. From first meeting at my PhD interview, to becoming best friends, to now planning our wedding. You have supported me in every possible way that I simply could not list them all. Thank you for answering countless questions and being my anchor during a wild rollercoaster of emotions and experiments. A massive thank you for such memorable times in and out the lab. As we always say - we love science and adventures.

DEDICATION

I dedicate this thesis to the much loved and missed

Harry 'Haz' Oswald

Abstract

Neuroblastoma (NB) is the most common malignant solid tumour diagnosed in infants, accounting for ~15% of all childhood cancer-related deaths. Current patient risk stratification criteria are heavily reliant on the presence of a *MYCN* amplification, albeit only accounting for ~25% of patients. The inadequate prognostic risk stratification of patients results in children receiving either inefficient or excessive treatment with a myriad of severe lifelong side effects for survivors. Therefore, the identification and characterisation of novel biomarkers could not only identify new therapeutic targets but could also improve risk stratification and treatment planning. A comparative transcriptomic analysis of NB tumours (obtained from the chick embryo model) grown under normal oxygen tensions (normoxia, 21% O_2) or hypoxia (1% O_2), a model for aggressive NB tumours that correlates with poor patient prognosis, identified multiple significantly upregulated genes in aggressive (hypoxic) tumours specifically, with Zinc Finger DHHC-Type Palmitoyltransferase 23 (*zDHHC23*) and Maestro Heat Like Repeat Family Member 6 (*MROH6*) exhibiting the best correlation with poor prognosis. This thesis sought to validate these expressed gene products as potential biomarkers in NB. I also investigated the molecular function of these two proteins under normoxic and hypoxic conditions, supplementing the currently limited available knowledge.

Commercially available antibodies for these two proteins were unsuccessful for use in either immunostaining, a procedure currently used as the 'gold-standard' of clinical biomarker screening, or for immunoblotting of endogenous protein, with all of the antibodies evaluated lacking specificity. Although targeted mass spectrometry assays were successfully developed, they lacked the sensitivity to detect endogenous proteins, likely due to low levels of protein expression. Therefore, I focused on the biochemical characterisation of these two proteins, cloning dual reporter HA-mCherry-protein and protein-mCherry-HA plasmids to facilitate immunoprecipitation of exogenously expressed protein and evaluation of sub-cellular localisation. I developed and optimised a HA-tag based immunoprecipitation protocol for liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis to allow identification of post-translational modifications (PTMs) and protein interaction networks. These experiments revealed extensive hypoxia-induced regulation of protein binding partners, with ~70% of the interactome (from a total of 262 and 253 co-immunoprecipitated proteins for zDHHC23 and MROH6 respectively) changing as a function of O_2 tension. GO-term analysis of these interactomes suggests that zDHHC23 is a component of several potentially important malignancy pathways, including cytoskeletal reorganisation and adhesion. Label free quantification analysis of MROH6 identifies high stoichiometric binding to Breast Cancer Anti-oestrogen Resistance protein 1 (BCAR1), inferring potential roles in telomere maintenance and genetic stability. Additionally, PTM analysis identified one and three phosphorylation sites on MROH6 and zDHHC23 respectively, with zDHHC23 S252 predicted to be regulated by Cyclin dependent kinases. Finally, I developed, to my knowledge, the first reported click-chemistry based high-throughput LC-MS/MS pipeline for the unbiased identification of zDHHC23 palmitoylated substrates, concluding that the 'palmitome' is much more complex than currently understood and likely regulates localisation to membrane bound organelles and extracellular vesicles, as well as its established role in plasma membrane localisation.

Overall, using LC-MS/MS approaches, I explore and discuss how zDHHC23 and MROH6 overexpression may contribute to aggressive NB development and poor patient prognosis.

Abbreviations

2D	2 dimensional
3D	3 dimensional
Å	Angstrom
ACN	Acetonitrile
ALK	Anaplastic lymphoma kinase
AmBic	Ammonium bicarbonate
APS	Ammonium persulfate
ARPC5	Actin Related Protein 2/3 Complex Subunit 5
ARS	Aminoacyl-tRNA synthetase
BCAR1	Breast cancer anti-oestrogen resistance protein 1
BH	Bead homogeniser
BRIX1	Biogenesis of Ribosomes BRX1
CAM	Chorioallantoic membrane
CID	Collision-induced dissociation
CO-IP	Co-immunoprecipitation
COVID19	Coronavirus
CRC	Colorectal cancer
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cu(I)	Copper
CuAAC	Cu(I)-catalysed azide-alkyne cycloaddition
DC	Direct current
DDA	Data dependent-acquisition
DHHC	Aspartate-histidine-histidine-cysteine
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EMT	Epithethial-mesenchymal transition
ER	Endoplasmic reticulum
ERAD	Endoplasmic-reticulum-associated protein degradation
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FASN	Fatty acid synthase
FDR	False discovery rate
FFPE	Formalin-fixed, paraffin-embedded
GAR1	H/ACA ribonucleoprotein complex subunit 1
GD2	Anti disialoganglioside
GFP	Green fluorescent protein
GO	Gene ontology

GOI	Gene of interest
GPCR	G-protein-coupled receptors
H_2O	Water
HCD	Higher energy collisional dissociation
HIAR	Heat-induced antigen retrieval
HIF	Hypoxia-inducible factor
HIF-1α	Hypoxia Inducible Factor 1 Subunit Alpha
HIF-2α	Hypoxia Inducible Factor 2 Subunit Alpha
HPLC	High-performance liquid chromatography
HTP	High throughput
IAA	Iodoacetamide
ICC	Immunocytochemistry
IDRF	Image defined risk factors
IF	Immunofluorescent
IHC	Immunohistochemistry
INRG	International Neuroblastoma Risk Group
INRGSS	International Neuroblastoma Risk Group Staging System
IP	Immunoprecipitation
KCNMA1	Potassium Calcium-Activated Channel Subfamily M Alpha 1
LB	Luria-Bertani
LC	Liquid chromatography
LFQ	Label free quantification
m/z	mass to charge
MROH	Maestro heat like repeat
MROH6	Maestro heat like repeat family member 6
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
MYCN	MYCN Proto-Oncogene
NaCl	Sodium chloride
NB	Neuroblastoma
NCC	Neural crest cell
NCE	Normalised collision energy
NH4Cl	Ammonium chloride
nm	Excitation/emission wavelength
NOLC1	Nucleolar and Coiled-Body Phosphoprotein 1
O_2	Oxygen
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PD	Proteome Discoverer
PEI	Polyethyleneimine
PFA	Paraformaldehyde
PHD	Prolyl-4 hydroxylase domain
POI	Protein of interest

PRM	Parallel reaction monitoring
PSM	Peptide spectral match
PSMB6	Proteasome 20S Subunit Beta 6
PSNS	Postganglionic sympathetic nervous system
PTM	Post translational modifications
q-OT	quadrupole-Orbitrap
qRT-PCR	Real-Time Quantitative Reverse Transcription PCR
RF	Radio frequency
RFP	Red fluorescent protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT	Radiotherapy
SCA	Segmental chromosomal alteration
SCX	Strong cation exchange
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulfate
SHH	Sonic Hedgehog
siRNA	Small interfering RNA
SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1
SNRNP70	U1 small nuclear ribonucleoprotein 70 kDa
SRM	Selected reaction monitoring
TBS	Tris-buffered saline
TBST	Tris-buffered saline with tween
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TiO_2	Titanium Dioxide
TrK-A	Tropomyosin receptor kinase A
zDHHC23	Zinc finger DHHC – type containing 23
ZNF277	Zinc Finger Protein 277

Amino Acids

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

List of figures

List of tables

Contents

Chapter 1: Introduction

1.1. Cell signalling

For the evolution of multicellular organisms, it is important that all cells must be able to communicate with each other i.e. cells must be able to interpret and translate the desired functional response to a variety of extracellular signals that they are exposed to in order for the survival of the organism (Jordan et al., 2000). The ability to interact and adapt to environmental cues is an essential process in cell biology because homeostasis (equilibrium of physiological processes) of a living organism must be maintained while having the capacity to adjust to constantly changing and challenging environmental conditions. This responsiveness is primarily initiated at the cell surface with specific protein receptors that are sensitive to surrounding environmental signals. The receptors transduce the signal across the plasma membrane and activate intracellular signalling pathways to bring about the appropriate gene expression and functional responses (Jordan et al., 2000). There are various signals that trigger activation of protein receptors, including but not limited to: a ligand bound to another cell, the extracellular matrix and soluble factors (chemicals, proteins, sugars etc.). Consequently, the complexity of signalling pathways means intracellular translation is rarely a discrete entity, rather a complex integrated network with the ability to influence different signalling pathways through 'cross-talk' (Klipp and Liebermeister, 2006).

Additional to activation of a response, it is vital that cell signalling pathways have the ability to switch off once the stimulus has been responded to or is no longer present in order to prevent an unnecessary excessive response that may impact homeostasis. Therefore, cellular signalling networks are typically regulated by feedback loops, where the correct effector protein response inhibits its own signal transduction (Kholodenko, 2006). A common but highly metabolically inefficient method to regulate the effector proteins is the cyclic protein accumulation/degradation in response to stimuli. The requirement for complete effector protein transcription and translation inherently results in lag periods (length of time dependent on the system and organism) between detection and response, which may result in cellular/organism death if the environmental cue is too strong without adaption. As a result, cells have evolved a mechanism known as post translational modification (PTM) which allows effector proteins to be modified post-synthesis thereby rapidly altering its biochemical

properties. PTM can thus serve to alter protein function in a rapid and often reversible manner, reducing the time from stimuli-to-response.

1.1.1 Post translational modifications

PTM encompass a variety of modifications that can occur to a protein including; covalent attachments, chemical modifications and cleavages, to name a few. These occur on a specific amino acid residue within the protein sequence. These PTMs in turn alters the modified residues biochemical properties in a specific way that changes the protein's function, for example: altering sub-cellular localisation, stability, protein-protein or protein-ligand interactions and enzymatic activity. Due to the vast number of different environmental signals an organism must adapt to, over 400 PTMs have evolved (Consortium, 2019) inducing highly varied biochemical changes to a translated protein (Figure 1.1). Many covalent PTMs are reversible allowing rapid and cyclic signalling without the need for the highly energetic and long delays of accumulation/degradation cycles.

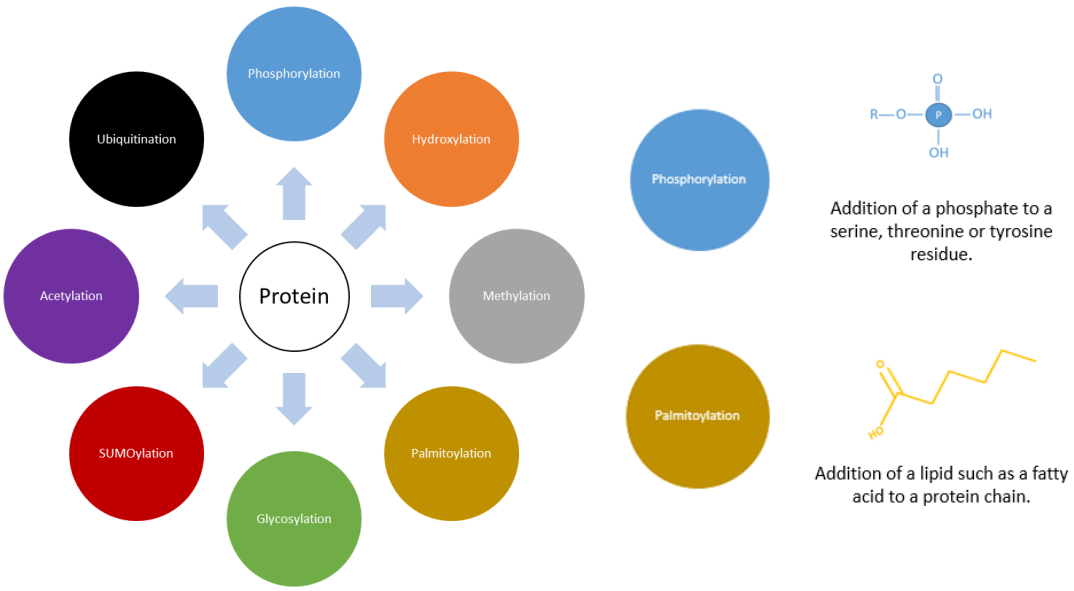

Figure 1.1 Schematic showing eight common covalent post-translational modifications.
The expanded PTMs – phosphorylation and palmitoylation will be investigated further in this thesis.

The effect of a specific PTM on protein function is dependent not only on the protein that is modified, but also the type of PTM and the site of modification. A given protein can also be

modified by multiple PTMs simultaneously, but often in a sub-stoichiometric manner, resulting in a heterogenous pool of partially 'decorated' proteins. In this manner, PTMs can finely tune signalling pathways and allows a single protein to participate in multiple regulatory pathways at once (Mann and Jensen, 2003). Therefore, a truer reflection of the cellular proteome must consider the multiple and separate 'proteoforms' (different PTM modified or mutated variants of a single protein), exponentially increasing the complexity of expressed proteins. Only a couple of covalent PTMs are relevant to this thesis and are discussed below.

1.1.2 Phosphorylation

Phosphorylation is the second most abundant signalling PTM employed by cells with approximately a third of all expressed proteins predicted to be phosphorylated at any one moment (Olsen et al., 2006). Protein phosphorylation is a reversible PTM, performed by a protein kinase and hydrolysed by a protein phosphatase, which involves the covalent attachment of a negatively charged phosphate group (PO_3-) to an amino acid side chain. The addition of this relatively large negatively charged phosphate group has been shown to have a variety of effects to a modified protein including, but not limited to: protein structural changes, promoting (or preventing) co-factor binding partner association, subcellular localisation and stability roles (Hunter and Karin, 1992). Considering the complexity involved with targeting a specific protein residue to different stimuli, it is therefore unsurprising that over 500 different human kinases are present, equating to approximately 2% of all proteins encoded by the human genome (Manning et al., 2002). Canonically, phosphorylation occurs on Serine (S), Threonine (T) and Tyrosine (Y) residues, at a ratio of around 86% : 12% : 2% (Olsen et al., 2006). However, more recent evidence identifies phosphorylation can occur on non-canonical residues including: Histidine (H), Arginine (R), Lysine (K), Aspartic acid (D), Glutamic acid (E) and Cysteine (C) in humans (Hardman et al., 2019). With phosphorylation having such vast roles in cellular transduction pathways, aberrant phosphorylation is linked to many diseases and therefore represents one of the largest markets for pharmaceutical discovery (Cohen, 2002).

1.1.3 Palmitoylation (S-acylation)

Multiple different lipids (fatty acids) can be conjugated onto different residues including Cys, Ser and Thr, known colloquiality as S-acylation or O-acylation respectively. Of importance for this thesis is palmitoylation - a lipid based PTM that results in the addition of palmitic acid (from palmitate, C16:0) to Cys, Ser or Thr residues through a labile thioester or hydroxyester bond, a process catalysed by palmitoyltransferases (Linder and Deschenes, 2007). Cells have not only developed methods to import palmitic acid across the cell membrane (Schwenk et al., 2010), they can also convert glucose into palmitic acid through a multi-step process involving glycolysis and fatty acid synthases (Chan and Vogel, 2010). Within this thesis I am specifically interested in S-palmitoylation of Cys residues.

Palmitoylation as a PTM was originally identified in 1980 (Schlesinger et al., 1980) and occurs in all eukaryotic organisms ranging from yeast to humans (Roth et al., 2006). Palmitoylation is exclusively post-translational (Smotrys and Linder, 2004) and therefore while most lipid based PTMs are irreversible, palmitoylation is reversible by the action of thioesterases and hydroxyesterases (Salaun et al., 2010). Two cytosolic acyl thioesterases (acyl-protein thioesterase 1 and 2) have been identified as palmitoyl thioesterases that depalmitoylate cytosolic cysteine residues of proteins (Dekker et al., 2010, Tomatis et al., 2010).

Palmitoylation has multiple effects on protein function arising due to the marked increase in localised hydrophobicity, with one of its key functions being to target proteins to membranes (with additional functions including regulating protein stability and protein-protein interactions) (Fukata and Fukata, 2010). The increased hydrophobicity facilitates the association of a protein within membrane structures, such as the plasma membrane and subcellular membrane organelles (endoplasmic reticulum/Golgi apparatus), acting as a focal point for localising signalling (Kaur et al., 2016). Palmitoylated proteins are currently thought to localise predominantly to the plasma membrane, and include receptor proteins such as G protein coupled receptors, ion channels, Ras proteins and G alpha subunits (Salaun et al., 2010). Historically, palmitoylation has been challenging to study due to the low sensitivity of available detection methods and the reliance on radiolabelling; our knowledge regarding the relevant enzymes responsible, target proteins and functional outcomes of lipid modification is thus sparse compared with e.g. phosphorylation or ubiquitination (Mitchell et al., 2006). To highlight this, despite having being identified in 1980, it was only in 2004 that the enzymes that mediate Cys-palmitoylation (known as the zDHHC family of mammalian s-

acyltransferases) were discovered (Fukata et al., 2004). These will be discussed further in Section 1.6.1.

1.2. Hypoxia

As discussed, cell to cell communication is essential for the correct coordination and response to environmental stimuli. Hence, cells constantly monitor and react to environmental cues, including extreme temperatures, changes in pH, exposure to toxins and of particular interest in this thesis, oxygen (O_2) levels. Monitoring O_2 availability is crucial due to it being a vital component for energy metabolism required for life. Consequently, cells have evolved a highly specific signalling pathway that is dedicated to detecting and facilitating the response to low O_2 (known as hypoxia).

1.2.1. Tumour hypoxia

Every tissue within an organism has different energy metabolism demands, and therefore has different requirements for O_2 availability. As a result, hypoxia (a lack of available O_2) is typically tissue dependent, with examples of median O_2 tensions of different tissue types and respective tumours depicted in Figure 1.2. The rapid cellular proliferation that occurs in the development of cancerous tumours results in large cell masses which have poor vasculature, which can result in hypoxic cores within the tumour (Figure 1.3). Tumour microenvironments, specifically in hypoxic tumour cores, can lead to the deregulation of several processes including: cell proliferation, survival, invasion and resistance to therapy, all of which are hallmarks of cancer (discussed further in Section 1.3). Furthermore, hypoxia has been widely reported to play roles in the deregulation of deoxyribonucleic acid (DNA) repair/damage response genes, cell cycle arrest, and DNA replication (Bristow and Hill, 2008).

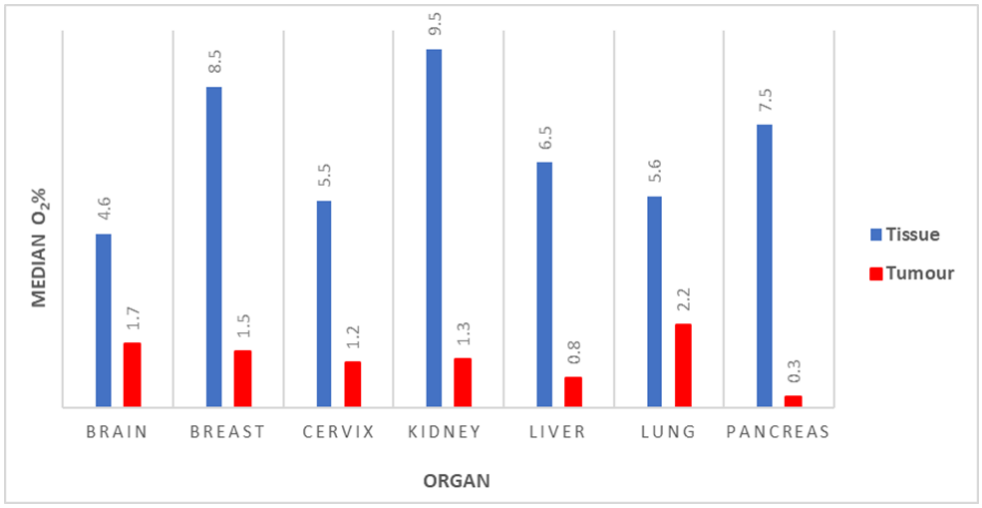

Figure 1.2 Examples of recorded median O_2 levels in organs and their associated tumours. Data collected from (Muz et al., 2015).

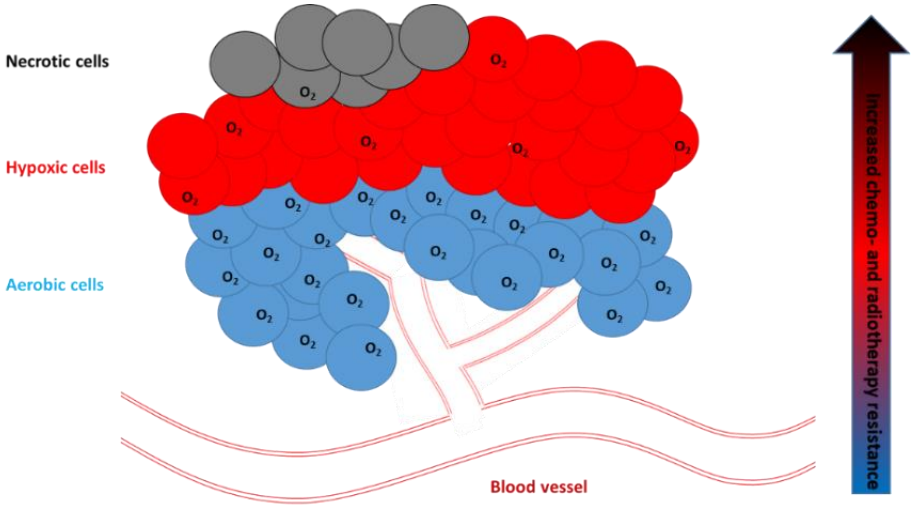

Figure 1.3 Hypoxia and the tumour microenvironment.
A large tumour mass results in inefficient vasculature causing hypoxic cells, with as low as 0.1% O_2 additionally these cells are increasingly resistant to the common cancer therapeutic strategies chemo- and radiotherapy.

1.3. Cancer

Despite decades of cancer research and a plethora of clinical trials investigating promising new therapies, cancer remains a major cause of mortality (1 in 5). Increasing lifespan also means that 1 in 2 people are now predicted to develop cancer in their lifetime (Ferlay et al., 2021). To highlight the prevalence of this non-communicable disease, in 2020 there were close to 20 million new cases and 10 million deaths worldwide, with the most common cancer death resulting from lung cancer (Ferlay et al., 2021). Exacerbating this, the coronavirus (COVID-19) pandemic has (and continues to) have a significant negative impact on a global scale on oncological management/treatment with scheduled operations and routine cancer treatments (including chemotherapy and radiotherapy) cancelled or postponed, to prioritise hospital beds for those seriously ill with COVID-19 (Lancet, 2020).

Tumour cells have developed multiple routes that enable their growth, survival and ability to metastasise, originally coined in 2000 as the 'Hallmarks of Cancer' (Hanahan and Weinberg, 2000), diagrammatically depicted in (Figure 1.4). Briefly, Hanahan and Weinberg stated that cancer cells must have the ability to evade cell death (apoptosis), sustain and induce blood vessel production (angiogenesis), invade healthy tissue (metastasis) and replicate uncontrollably (Hanahan and Weinberg, 2000). Following extensive advances in cancer research, the 'Hallmarks of Cancer' has been expanded to include: evading the immune system, genetic instability, metabolic reprogramming phenotypic plasticity, senescent cells, non-mutational epigenetic reprogramming and polymorphic microbiomes (Hanahan and Weinberg, 2011, Hanahan, 2022). Recent hallmarks also take into consideration that the tumour microenvironment can directly influence tumour progression and response to treatment.

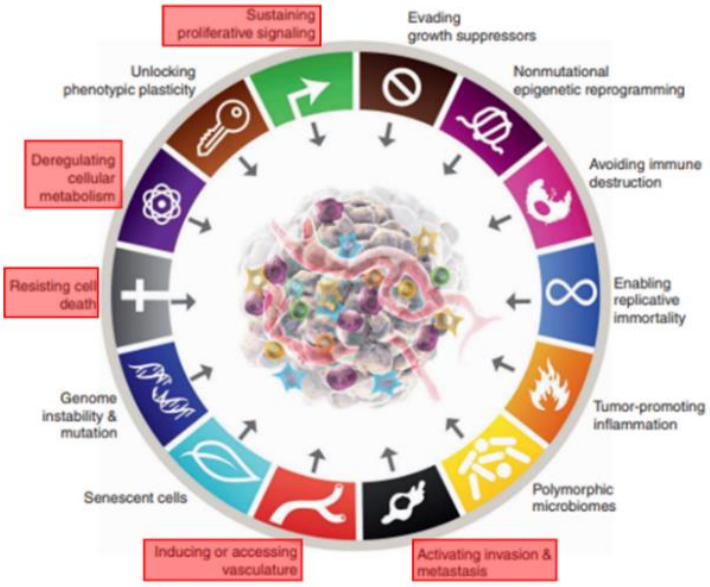

Figure 1.4 The Hallmarks of cancer survival and progression.
Each section of the wheel contains a hallmark trait which promotes cancer survival and progression. Highlighted section labels indicate targets of hypoxia regulation. Amended from (Hanahan, 2022).

Unfortunately, the COVID-19 pandemic also coincided with an unprecedented momentum in the field of paediatric oncology (area of medicine which diagnoses and treats cancer in children) which has severely hampered the 'Global Initiative for Childhood Cancer' (set by the World Health Organization) which aimed to improve the survival of children with cancer to more than 60% across the world by 2030 (Organization, 2021). Since the pandemic, reports have identified a decrease in out of hours/accident and emergency visits by paediatric patients with cancer, due to unfortunate delays in the timeliness of diagnoses (Lazzerini et al., 2020); whilst pre-pandemic cancer numbers were rising globally, post-pandemic there has been a 25-50% reduction in the numbers of new cancers identified from multiple countries including USA, Holland and Italy (Dinmohamed et al., 2020, Offenbacher et al., 2021, Ferrari et al., 2020). Reduced identification rates results in a substantial threat to the safe and effective care of children with cancer due to paediatric oncology relying heavily on early detection and evaluation. One particular childhood cancer, neuroblastoma (NB), is the focus of this thesis and is discussed below.

1.4. Neuroblastoma

NB is the most common malignant solid tumour diagnosed in infants, accounting for ~15% of all childhood cancer-related deaths, with 17 months being the median age of diagnosis (Maris, 2010, Colon and Chung, 2011). NBs differ from adult tumours as they arise early during embryonic development typically in the postganglionic sympathetic nervous system (PSNS), with ~65% of primary tumours arising in the adrenal medulla or lumbar sympathetic ganglia (Maris, 2010). The remaining tumours are generally present in the neck, chest or pelvis (Maris, 2010). Within the paediatric population, most adrenal masses are malignant, with NB accounting for ~85% of all adrenal mass cases (de Barros et al., 2012). Thoracic NB accounts for ~15% of NB cases and are associated with a better outcome than abdominal NB (Parikh et al., 2012). However, around half of all patients below the age of 1 year have metastatic NB at the time of diagnosis (DuBois et al., 1999).

Similar to many other cancers, there are ethnic differences in NB occurrence: there is a higher prevalence in those with a European ancestry whilst African-American NB patients tend to present as high-risk for mortality (Henderson et al., 2011). NB is not thought to be hereditary given that only 1% of patients have familial NB ties. The disease predominantly occurs as a result of spontaneous mutation during embryogenesis (Kushner et al., 1986), with mutations in the anaplastic lymphoma kinase (ALK) gene being a major contributing factor for familial predisposition to NB (Bourdeaut et al., 2012).

1.4.1. Origin

Due to NB being an embryonic cancer of the PSNS, the likely cell of origin is thought to be neural crest cells (NCC) that have failed to differentiate (Maris, 2010). NCCs are migratory multipotent embryonic cells derived from the developing neural tube where they undergo controlled epithethial-mesenchymal transition (EMT) to migrate and develop into a diverse cell lineage including melanocytes, cartilage, bone, smooth muscle, neurons and glia (Bronner-Fraser, 1995) (Figure 1.5).

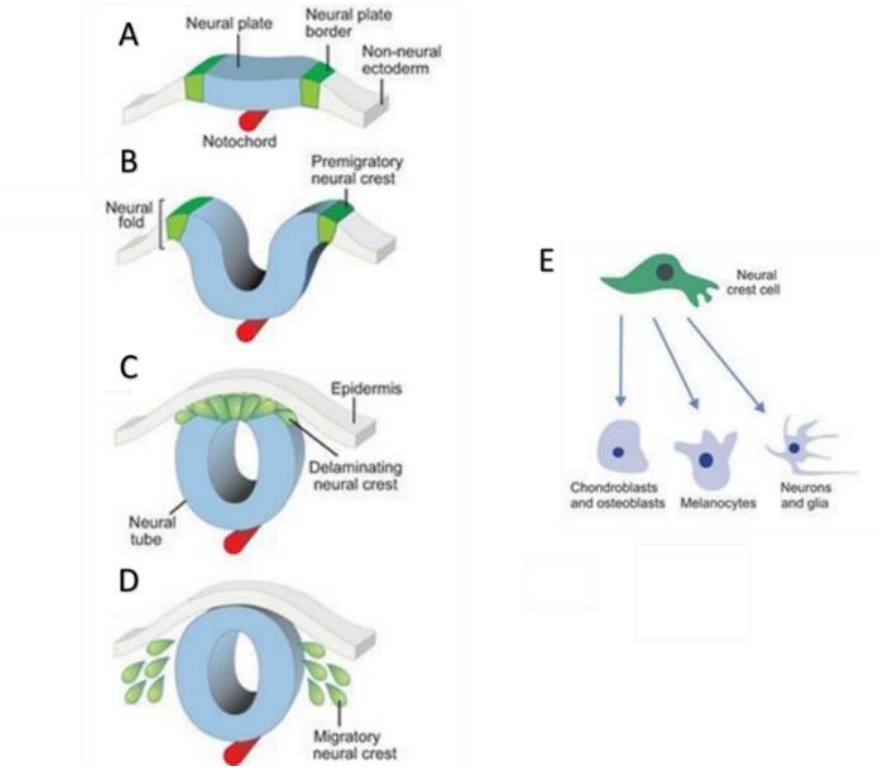

Figure 1.5 Neural crest development.
A) Neural crest development begins at the gastrula stage, with the specification of the neural plate border at the edges of the neural plate. B) The neural plate closes to form the neural tube; the neural crest progenitors are specified in the dorsal part of the neural folds. C) Neural crest cells undergo EMT and delaminate from the neural tube. D) Migratory neural crest cells follow pathways to diverse destinations. E) Possible derivatives of neural crest cells depending on their migratory pathways and external cues. Amended from (Simões-Costa and Bronner, 2015).

As NB tumours derive from NCCs of a sympathoadrenal lineage, they can therefore develop anywhere in the PSNS. Currently, the mechanisms that cause these embryonal cells to give rise to NB tumours are unknown, with suggestions that defects in embryonic genes that control NCC development are likely to become spontaneously mutated and cause the uncontrolled cellular proliferation and differentiation (Limpt et al., 2004). As briefly mentioned, ALK mutations can play a role in familial NB development. ALK is a receptor tyrosine kinase which has known roles in neuronal differentiation (Motegi et al., 2004), with high expression levels observed in the embryonic nervous system which eventually dissipates after birth. Hence, ALK mutations can perturb the normal genetic differentiation of NCCs, resulting in inappropriately timed differentiation (either too early or too late). It is these differences in differentiation, which is used in histopathology and diagnosis, and subsequently as a criterion for patient risk stratification.

1.4.2. Pathology

Neuroblastic tumours, including NB, belong to the 'small blue cell' neoplasms of childhood cancers (Whang-Peng et al., 1986). The neuroblastic tumour subtypes reflect the spectrum of cellular maturation, ranging from undifferentiated neuroblasts to fully differentiated neurons surrounded by Schwann cells and determined upon patient tumour biopsy (Figure 1.6).

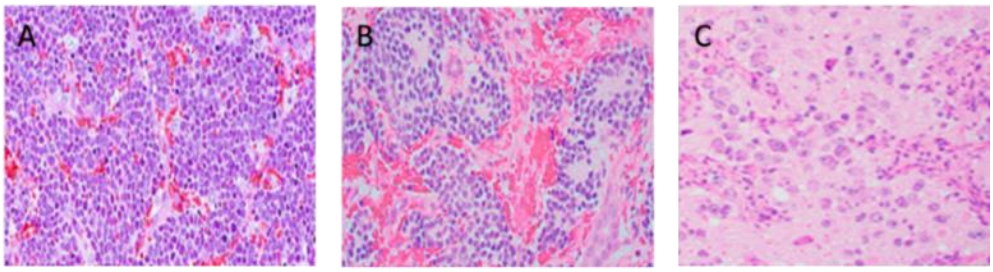

Figure 1.6 Types of neuroblastoma differentiation identified in histopathology.
Images A) Neuroblastoma - Undifferentiated B) Neuroblastoma - poorly differentiated C) Neuroblastoma – differentiating. Amended from (Ikegaki et al., 2019).

A NB tumour is categorised by its composition of neuroblastic cells which form groups (or nests) that are separated by stromal septa with no Schwannian proliferation. There are three NB subtypes: undifferentiated (Figure 1.6 A), poorly differentiated with <5% cells showing differentiation (Figure 1.6 B) and differentiated where >5% cells differentiate to form ganglions (Figure 1.6 C). The correlation between NB tumour differentiation and clinical outcome is well recognised, with a poor prognosis for patients with undifferentiated NB (Cohn et al., 2009). The vast heterogeneity of NB and the diverse clinical outcomes has a substantive impact on the treatments prescribed, discussed below.

1.4.3. Treatment

NB clinical behaviour and outcome is extremely diverse, ranging from a low-risk disease with spontaneous regression to a high risk, highly-lethal disease that is treatment resistant and has tendency to relapse (Cheung and Dyer, 2013). Therefore, there is highly varied treatment strategies that include simple patient observation, surgical resection of the tumour, or very

aggressive therapy protocols reliant on simultaneous multi-agent high-dose chemotherapy, radiation and immunotherapy (Simon et al., 2017). As NB is highly metastatic, surgical resection often has limited success alone and is used in combination with chemotherapy, where patients are typically given 4 to 8 cycles (~3-6 months) either before (aim of shrinking the tumour (Coughlan et al., 2017)) or after surgery. The chemotherapy drugs usually given include platinum compounds (carbo-platin, cisplatin), etoposide, cyclophosphamide and vincristine (Simon et al., 2007). These drugs typically work by damaging cells as they divide. As cancer cells divide much more rapidly and frequently than normal cells, these treatments preferentially affect tumour cells, resulting in an accumulation of DNA damage that is much less prevalent in 'normal' cells, and triggering cell cycle arrest and/or apoptosis. As a second stage treatment for advanced NB that has minimally responded to chemotherapy, radiation therapy (iodine-131 [131I]-8H9) is typically given prior to surgery in an attempt to shrink tumours, or alternatively post-surgery to kill any remaining cancer cells. Immunotherapy can also be given to patients which are deemed at higher risk of mortality (Section 1.4.4) this usually involves an anti disialoganglioside (GD2) monoclonal antibody which can help achieve a long-term overall survival of > 50% (Cheung et al., 2012). The overall aim of treatment is not only the remission of NB, but also minimising physical effects, and psychological complications induced by therapeutic toxicity for survivors (Wienke et al., 2021).

1.4.3.1. Late effects of treatment

The current treatment strategies discussed above have limited target specificity (Johnsen et al., 2019) and the systemic therapeutic application can also cause unwanted systemic side effects. Unfortunately, due to NB being a childhood cancer, tissues are still growing and developing. Chemotherapy that targets rapidly growing cells can thus result in long term side effects (Rombi et al., 2013). One of the most common issues for survivors is hearing loss, with 73% of individuals treated with platinum compounds affected (Cohen et al., 2014). Another common side effect of platinum agents is long-term renal toxicity, with 10-40% of survivors suffering from chronic renal failure (Laverdière et al., 2005). Treatment also has a substantial impact on patient development due to effects on the endocrine system resulting in growth hormone deficiency, delayed puberty, hypothyroidism and insulin resistance (Cohen et al., 2014). Other common side effects include: a 6-fold increase in the risk of blindness if cranial radiation was administered (Laverdière et al., 2005), neurological conditions (epilepsy),

musculoskeletal conditions (scoliosis and osteoporosis), learning disabilities and secondary malignant neoplasms (Cohen et al., 2014). The development of a secondary malignancy is rare (2%-7% of NB patients) with the most common secondary tumours including renal and thyroid carcinoma, lymphoma and acute myeloid leukaemia (Applebaum et al., 2015). Despite these risks, the spontaneous recurrence of NB is the most severe and likely secondary malignancy to form. Therefore, survival and improvement of the quality of life is a high priority for paediatric patient survivors. However, to achieve this, clinicians require efficient prognostic tools to help decide what is the appropriate and effective therapy for individual patients (Cohn et al., 2009, Liang et al., 2020).

1.4.4. Risk classification

Treatment strategies for NB patients has been tailored according to the predicted response to therapy and risk of relapse for more than 50 years (Zage et al., 2012). The current criteria for diagnosing and staging NB is based upon the International Neuroblastoma Risk Group Staging System (INRGSS) (Table 1.1) which was formulated in 1986 and revised in 1988 (Brodeur et al., 1988). This was formulated to provide consistency in staging of NB patients worldwide.

Table 1.1 The International Neuroblastoma Risk Group Staging System (INRGSS). Taken from (Monclair et al., 2009).

Stage	Description
L1	Localised tumour not involving vital structures as defined by the list of image-defined risk factors and confined to one body compartment
L2	Locoregional tumour with presence of one or more image-defined risk factors
M	Distant metastatic disease (except stage MS)
MS	Metastatic disease in children younger than 18 months with metastases confined to skin, liver, and/or bone marrow

Localised disease is classified as L1 or L2 while metastatic disease is classed as M or MS stage. It was in 1971, that the distinct subset of patients (MS) was identified who had extensive disseminated disease but who reliably underwent spontaneous regression (D'Angio et al., 1971). Radiological and nuclear medicine are essential for the staging of NB. Initial diagnostic

testing includes either a contrast-enhanced computed tomography or magnetic resonance imaging with the latter generally preferred due to the lack of radiation and higher soft tissue contrast resolution (Siegel and Jaju, 2008). This allows not only the tumour size to be determined but the tumours relationship with adjacent vital structures and the extent to which the disease has metastasised, thus the stage of the patient is determined. This is determined by image-defined risk factors (IDRF) to predict surgical risk and prognosis (Table 1.2).

Table 1.2 Description of imaging-defined risk factors for the International Neuroblastoma Risk Group Staging System (INRGSS)
Taken from (Brisse et al., 2011).

Anatomic region	Description
Multiple body compartments	Ipsilateral tumour extension within two body compartments
Neck	Tumour extending to skull base
	Tumour compressing trachea
	Tumour encasing carotid artery, vertebral artery or jugular vein
Cervicothoracic junction	Tumour encasing brachial plexus
	Tumour encasing subclavian vessels, vertebral artery or carotid artery
	Tumour compressing trachea
Thorax	Tumour encasing aorta or major branches
	Tumour compressing trachea or main bronchi
	Lower mediastinal tumour infiltrating costovertebral junction between T9 and T12 vertebral levels
Thoracoabdominal junction	Tumour encasing aorta or vena cava
Abdomen and pelvis	Tumour infiltrating porta hepatis or hepatoduodenal ligament
	Tumour encasing branches of superior mesenteric artery at mesenteric root
	Tumour encasing origin of celiac axis or superior mesenteric artery
	Tumour invading one or both renal pedicles
	Tumour encasing iliac vessels
	Pelvic tumour crossing sciatic notch
Intraspinal tumour extension	Intraspinal tumour extension (more than 1/3 of spinal canal in axial plane is invaded, perimedullary leptomeningeal spaces are not visible or spinal cord intensity is abnormal).
Infiltration of adjacent organs and structures	Pericardium, diaphragm, kidney, liver, duodenopancreatic block and mesentery

The risk of death for each NB patient is based upon not only the stage of the disease but also age at diagnosis, tumour histology, presence of a *MYCN* proto-oncogene amplification, tumour ploidy and segmental chromosome copy number alterations. These are compiled into the International Neuroblastoma Risk Group (INRG) classification system (Table 1.3).

Table 1.3 The International Neuroblastoma Risk Group (INRG) classification system Amended from (Cohn et al., 2009).

INRG Stage	Age (months)	Histologic Category	Grade of Tumour Differentiation	*MYCN*	11q Aberration	Ploidy	Pre-treatment Risk Group
L1/L2		GN maturing; GNB intermixed					Very Low
L1		Any, except GN maturing or GNB intermixed		NA			Very Low
				Amp			High
L2	<18	Any, except GN maturing or GNB intermixed		NA	No		Low
					Yes		Intermediate
	≥18	GNB nodular; neuroblastoma	Differentiating	NA	No		Low
					Yes		Intermediate
			Poorly differentiated or undifferentiated	NA			High
				Amp			
M	<18			NA		Hyperdiploid	Low
	<12			NA		Diploid	Intermediate
	12 to <18			NA		Diploid	Intermediate
	<18			Amp			High
	≥18						High
MS	<18			NA	No		Very low
					Yes		High
				Amp			High

Based upon the INRG classification system and the seven biological/clinical factors, NB patients are placed into one of four categories for five year survival rate: very low- (>85%), low- (<85-75%), intermediate- (<75-50%) or high-risk (<50%) (Table 1.3).

Patients that are low-risk have a highly favourable prognosis with a five year survival rate at more than 90% (Kholodenko et al., 2018) with metastases often not occurring despite incomplete resection of the tumour. High-risk disease in NB is defined by the presence of MYCN amplification and/or distinct metastatic spread, usually located in bone marrow, lymph nodes and liver (Pudela et al., 2020). Despite advances in treatment, high-risk NB patients will either experience refractory disease that does not respond sufficiently to therapy or disease recurrence.

1.4.4.1. MYCN Amplification

MYCN was identified in 1983 as an amplified gene homologous to *v-myc* with the ability to promote neoplastic transformation of mammalian cells (Schwab et al., 1985). Within two years of *MYCN*'s discovery in NB, it's amplification was shown to correlate with poor prognosis (Brodeur et al., 1984). *MYCN* amplification is highly associated with high-risk NB (accounting for 50% of high-risk cases), advanced stages of disease and subsequently poor prognosis (Seeger et al., 1985). However, *MYCN* amplification also accounts for only a quarter of all NB cases (Huang and Weiss, 2013).

MYCN amplification was the first identified clinically relevant molecular biomarker for NB and still remains a strong predictor for unfavourable patient outcome (Brodeur et al., 1984). However, a INRG report in 2019 identified that the prognostic impact of *MYCN* amplification is dependent on the context of other clinical and biological features (Campbell et al., 2019). High-risk patients with *MYCN* amplification that are over the age of 18 months generally present with widespread metastasis at diagnosis, with the five-year event free survival rate for children with high-risk disease remaining below 50% (Pearson et al., 2008). Approximately half of high-risk patients do not respond to first-line therapy or relapse within the first two years of treatment (Pearson et al., 2008). Prognosis for high-risk patients that relapse remains abysmal with a <10% 5 year survival rate (Basta et al., 2016). Therefore, risk assignment is critical for NB treatment alongside a greater need for the development of new treatment strategies for high-risk tumours.

1.4.4.2. Tumour ploidy and chromosomal aberrations

Tumour cell ploidy is also one of the factors used in the INRG, with ploidy associated with patients younger than 12 months of age at diagnosis. Diploid and hyperdiploidy NB tumours correlate with poor and highly favourable (>90% long-term survival) clinical outcomes, respectively (Kaneko and Knudson, 2000). These tumours also commonly involve chromosomal aberrations, with whole-chromosome copy number variations found to correlate with an excellent prognosis (Look et al., 1991). The highly variable clinical outcome for NB patients is also determined by different segmental chromosomal alterations (SCA) which has a heavy prognostic impact. These include: loss of chromosome 1p (seen in a third

of all cases (Caron et al., 1996, Gilbert et al., 1984)), 3p (Spitz et al., 2003), 4p (Caron et al., 1996), 6q (Ognibene et al., 2020), 11q (Carén et al., 2010) this is thought to express tumour suppressor genes, an aberration on 11q significantly increases risk of mortality in NB patients and therefore is part of the INRG criterion. Chromosomal gains have also been correlated with prognosis including at 1q (Gilbert et al., 1984), 2p (Szewczyk et al., 2019), and 17q which is seen in approximately half of all NB patients (Bown et al., 1999) and linked to *MYCN* amplification and very poor prognosis (Gilbert et al., 1984).

1.4.5. Hypoxia in neuroblastoma

As discussed in Section 1.2, hypoxia is a key characteristic of solid tumours. Deregulation of O_2 levels during embryonic development can interfere with neural crest-derived tissue maturation. Hypoxic events during normal embryonic and foetal development can affect placentation, angiogenesis and haematopoiesis (Bluhm et al., 2008). Local or systemic hypoxia has also been linked to NB development, including NCC migration and differentiation (Von Stedingk et al., 2019). Hypoxia is also known to trigger dedifferentiation of NB and associated with metastasis (Jögi et al., 2002). As such, hypoxia has been extensively reported as a poor prognostic marker. However, much remains to be discovered regarding the subsequent intracellular modifications and their link with a metastatic phenotype (Smith et al., 2016).

In NB cells, hypoxia promotes several genes to be differentially expressed, including the increase in expression of SERPINB9 and TRIO (Holmquist-Mengelbier et al., 2006), genes that are involved in metastasis in other cancers (Yoshizuka et al., 2004, ten Berge et al., 2002). Hypoxia seems to contribute to NB invasiveness and aggressiveness through HIF-1α (transcription factor which regulates hypoxic response) via the sonic hedgehog signalling (SHH) pathway (Chen et al., 2015b). It is HIF-2α (an isoform of HIF-1α) that has been directly linked to poor prognosis in NB, with HIF-2α inhibition leading to a reduction in vascular endothelial growth factors (Noguera et al., 2009). However, there are conflicting reports on using HIF-2α as a target for NB given that inhibition of HIF-2α has been linked to a decreased tumour response to treatment (Westerlund et al., 2017). Broadly speaking, hypoxia is represented by a decreased concentration of O_2 within tumour masses and associated with an aggressive cancer phenotype characterised by a higher metastatic potential and chemo-resistance (Brahimi-Horn et al., 2007), in NB cells specifically, there is HIF1α-dependent

resistance to etoposide and vincristine (Hussein et al., 2006). HIF-1α and HIF-2α differentially recruit target genes in NB in response to hypoxia, with HIF-2α preferentially promoting an aggressive phenotype. HIF-2α regulates the OCT4 gene which influences stem cell function, embryonic development and tumour growth (Covello et al., 2006).

1.5. Neuroblastoma research in chick embryo model

All preclinical trials of novel anti-cancer therapeutic modalities require thorough *in-vitro* and *in-vivo* investigation. Typically, at the *in-vivo* level, mouse rodent (*Mus musculus*) xenograft models are used to assess therapeutic efficiency/behaviour (Jung, 2014). In 1997, the first animal model for NB was developed by overexpressing *MYCN* under control of Tyrosine Hydroxylase transgenic mice, this is still by far the most popular model for NB research (Weiss et al., 1997). However, the rodent model is not without disadvantages: they are expensive, time-consuming and provoke ethical concerns. In contrast, the chick embryo model (*Gallus Gallus domesticus*) is an efficient alternative (Mapanao et al., 2021, Day et al., 2015) with a typical experiment concluding in 14 days.

The chick embryo is characterised by immunodeficiency, which is not typical for the rodent model which frequently do not allow growth of all human cancers (Mangir et al., 2018). Within the egg itself, the *in-ovo* chick is highly accessible, controlled and a relatively transparent model allowing any experimental manipulation to be easily visualised. The chorioallantoic membrane (CAM) is a well vascularised extraembryonic tissue, located just underneath the eggshell. It provides an excellent biological platform for the analysis of cancer development and growth (Cimpean et al., 2008), with the CAM shown to support murine sarcoma xenografts for more than 100 years (Murphy and Rous, 1912). Due to the chick embryo being immunodeficient in the early stages of development (Meirowsky et al., 1954) the CAM readily supports the engraftment of tumour tissues and is a valuable model for tumour angiogenesis and cancer metastasis (Zijlstra et al., 2002). The CAM assay has been utilised to study various cancers, including prostate (González-Chavarría et al., 2014), glioblastoma (Jaworski et al., 2013), osteosarcoma (Kunz et al., 2019), cervical cancer (Zhou et al., 2015), non-small cell lung cancer (Huang et al., 2017), thyroid (Pópulo et al., 2017) and of interest in this thesis neuroblastoma (Rodrigo et al., 2021).

The chick embryo model is well suited for NB research as this model mimics its natural pathophysiological conditions (embryonic). Studies in the chick have contributed

substantially to our knowledge on NCC. The first transcription factor regulating the EMT facilitated our understanding of neural crest emigration and subsequently cancer metastasis (Nieto et al., 1994). It has also been identified that despite NCCs being morphologically similar when they begin to migrate, not all subpopulations are capable of generating specific tissues (Bronner-Fraser, 1995). Additionally, NCCs migrate away from the neural tube, along the length which allows cells to populate different areas of the embryo and terminally differentiate to form different organs (Stern et al., 1991). In the Sée lab group, it was identified that NB SK-N-AS cells (non-*MYCN* amplified cell line) that were preconditioned in hypoxia (1% O_2) for three days prior to implantation into the chick embryo, triggered a metastatic phenotype in tumours (Herrmann et al., 2015). It was through this experimental technique that proteins of interest (POI) were identified, that could potentially be novel prognostic biomarkers for NB.

1.6. Proteins of interest

Preliminary work by the Sée lab has used transcriptomic profiling of SK-N-AS tumours grown in the chick embryo model to identify several significantly upregulated genes following three days of hypoxic (1% O_2) incubation, a model to mimic metastatic tumours that correlates with poor prognosis. Within the top 10 hypoxia upregulated genes, two are strongly associated with low patient survival (Figure 1.7). As a comparison, the clinical standard *MYCN* expression is included, showing that both of the identified genes are of similar (if not worse) indicators of patient prognosis. These two genes are *zDHHC23* (Zinc finger DHHC – type containing 23) and *MROH6* (Maestro heat like repeat family member 6). Consequently, and considering their identification in non-*MYCN* amplified cell lines (~75% of all NB cases), the purpose of this thesis was to investigate their putative use as NB biomarkers and to better characterise their molecular functions given the limited information currently available on these two proteins.

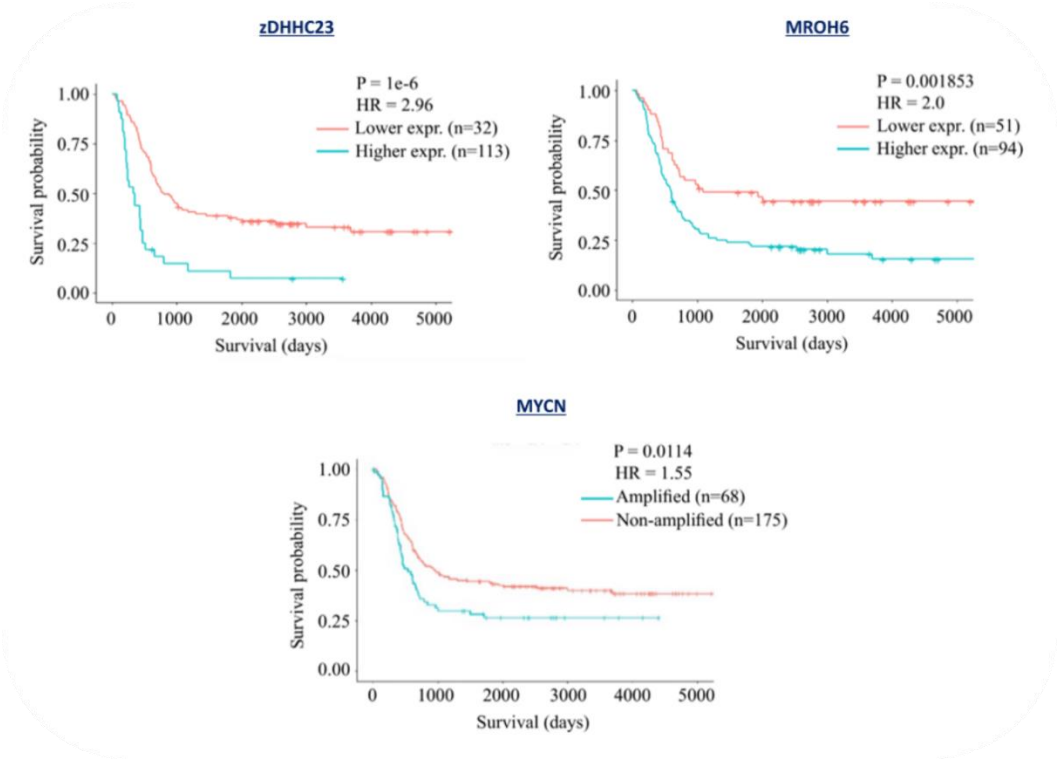

Figure 1.7 Survival analysis of the TARGET cohort.
Kaplan–Meier plots with patients separated into higher and lower expressing groups for the indicated genes and the standard amplified or non-amplified for MYCN that is currently used clinically. P-value from the log-rank test. HR = Hazard ratio.

1.6.1. DHHC-Palmitoyltransferases

As discussed in Section 1.1.1, palmitoylation is a PTM that is catalysed by palmitoyltransferases and can occur on Ser, Thr and Cys residues. zDHHC23 is one of the palmitoyltransferases from the DHHC (aspartate-histidine-histidine-cysteine tetra peptide) family (24 members) that are responsible for Cys-palmitoylation (Figure 1.8). The DHHC family characteristically encode a 51 residue, Cys rich domain in which the DHHC motif catalyses the transfer of palmitate to target proteins through the essential Cys residue (Greaves and Chamberlain, 2011, Jennings and Linder, 2012). Due to its vital function, the DHHC domain is conserved and located in the same location on every isoform, between the second and third transmembrane domains (De and Sadhukhan, 2018) (Figure 1.9).

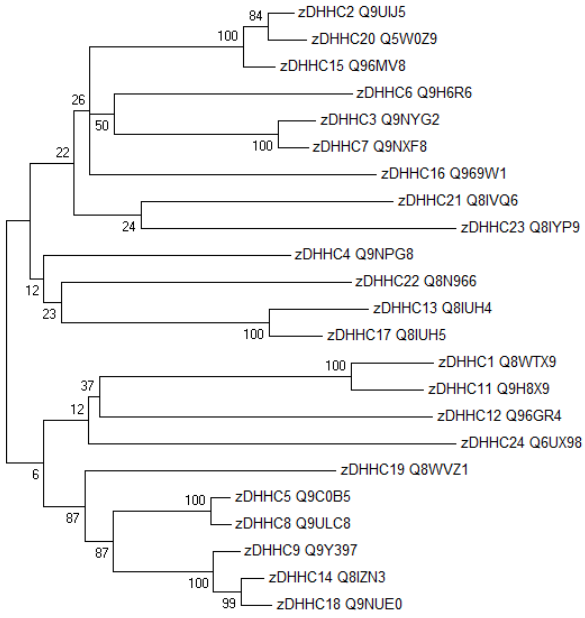

Figure 1.8 Molecular Phylogenetic analysis of DHHC family members by Maximum Likelihood method. The evolutionary history was inferred by using the Maximum Likelihood method based on the Le_Gascuel_2008 model (Le and Gascuel, 2008). The tree with the highest log likelihood (-9541.8090) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.6306)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 4.3849% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 23 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 200 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

```
                                  .*  *       *:**     * *  *   :
zDHHC1_Q8WTX9    SARSKHCSACNKCVCGFDHHCKWLNNCVGERNYRLFLHSVASALLGVLLL
zDHHC2_Q9UIJ5    PDRCHHCSVCDKCILKMDHHCPWVNNCVGFSNYKFFLLFLAYSLLYCLFI
zDHHC3_Q9NYG2    PDRAHHCSVCKRCIRKMDHHCPWVNNCVGENNQKYFVLFTMYIALISLHA
zDHHC4_Q9NPG8    PARSKHCSVCNWCVHRFDHHCVWVNNCIGAWNIRYFLIYVLTLTASAATV
zDHHC5_Q9C0B5    PPRCSHCSVCDNCVEEFDHHCPWVNNCIGRRNYRYFFLFLISLTAHIMGV
zDHHC6_Q9H6R6    APRSHHCRKCNRCVMKMDHHCPWINNCCGYQNHASFTLFLLLAPLGCIHA
zDHHC7_Q9NXF8    PERAHHCSICKRCIRKMDHHCPWVNNCVGEKNQRFFVLFTMYIALSSVHA
zDHHC8_Q9ULC8    PPRCSHCSVCDNCVEDFDHHCPWVNNCIGRRNYRYFFLFLISLSAHMVGV
zDHHC9_Q9Y397    PPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYLFILSISLLTIYV
zDHHC11_Q9H8X9   NKKTKHCISCNKCVSGFDHHCKWINNCVGSRNYWFFFSTVASATAGMLCL
zDHHC12_Q96GR4   PLRARHCRECRRCVRRYDHHCPWMENCVGERNHPLFVVYLALQLVVLLWG
zDHHC13_Q8IUH4   PLRSLHCHVCNCCVARYDQHCLWTGRCIGFGNHHYYIFFLFFLSMVCGWI
zDHHC14_Q8IZN3   PPRASHCSLCDNCVERFDHHCPWVGNCVGKRNYRFFYMFILSLSFLTVFI
zDHHC15_Q96MV8   PDRCHHCSVCAMCVLKMDHHCPWVNNCIGFSNYKFFLQFLAYSVLYCLYI
zDHHC16_Q969W1   PARTHHCSICNRCVLKMDHHCPWLNNCVGHYNHRYFFSFCFFMTLGCVYC
zDHHC17_Q8IUH5   PVRSKHCGVCNRCIAKFDHHCPWVGNCVGAGNHRYFMGYLFFLLFMICWM
zDHHC18_Q9NUE0   PPRTSHCSVCDNCVERFDHHCPWVGNCVGRRNYRFFYAFILSLSFLTAFI
zDHHC19_Q8WVZ1   PPRTYHCPWCNICVEDFDHHCKWVNNCIGHRNFRFFMLLVLSLCLYSGAM
zDHHC20_Q5W0Z9   PDRAHHCSACDSCILKMDHHCPWVNNCVGFSNYKFFLLFLLYSLLYCLFV
zDHHC21_Q8IVQ6   PKRSHHCSRCGHCVRRMDHHCPWINNCVGEDNHWLFLQLCFYTELLTCYA
zDHHC22_Q8N966   SPSTHFCRVCARVTLRHDHHCFFTGNCIGSRNMRNFVLFCLYTSLACLYS
zDHHC23_Q8IYP9   PARAWHCRICGICVRRMDHHCVWINSCVGESNHQAFILALLIFLLTSVYG
zDHHC24_Q6UX98   PPRSGHCSACRVCILRFDHHCRLLGRCVGFGNYRPFLCLLLHAAGVLLHV
```

Figure 1.9 Conserved catalytic DHHC cysteine-rich domain across DHHC family members.
*Shows the aligned zDHHC family member sequences from MUSCLE multiple sequence alignment and viewed in Clustal X. * indicates identical residues at a position whilst: identifies a strong conservation mutation. Blank spaces indicate at that position, there is no conservation mutation which was determined by the Clustal matrix. The red box highlights the DHHC region which contains the cysteine which is responsible for the protein's ability to palmitoylate.*

zDHHC palmitoyltransferases predominantly localise to the outer surface of the endoplasmic reticulum (ER) and Golgi, with some isoforms localise to the plasma membrane and endosomes (Gorleku et al., 2011). However, there is significant redundancy in the functions of these isoforms with most targets being capable of being modified by more than one enzyme (Greaves and Chamberlain, 2011). Evolutionary analysis of the 24 human zDHHC palmitoyltransferases shows two subfamilies of the zDHHC family (Figure 1.8). Interestingly for this thesis, zDHHC23 appears to be highly evolutionarily distinct with the common consensus suggesting it therefore may have non-redundant roles (Baker et al., 2021).

zDHHC-induced palmitoylation is known to be important for cancer progression, being identified as a key PTM modulating cell invasion and metastasis, cell proliferation, angiogenesis, and resistance to cell death (Yeste-Velasco et al., 2015). A small number of palmitoyltransferase roles and expression have so far been identified, including: zDHHC20 palmitoylation of the epidermal growth factor receptor (EGFR) aiding cell cycle progression (Kadry et al., 2021), increased expression of zDHHC9 observed in colorectal tumours (Birkenkamp-Demtroder et al., 2002), a downregulation of zDHHC14 in prostate cancer (Yeste-Velasco et al., 2014) and decreased zDHHC2 in gastric adenocarcinoma (Yan et al., 2013). Additionally, dysregulation of zDHHC family member enzymes has been linked with multiple conditions, including zDHHC17/13 in Huntington's disease (Young, 2012), zDHHC8 in schizophrenia, zDHHC9/15 in intellectual disability and zDHHC17 in type 1 diabetes (Greaves and Chamberlain, 2011).

1.6.1.1. zDHHC23

At the time of writing this thesis, little is known regarding the molecular mechanisms of action of zDHHC23 and consequently, its functional role in disease. In fact, its cellular localisation is currently debated with it initially having been identified in the Golgi apparatus (Tian et al., 2012), while a more recent study states that it is nuclear localised (Zhang et al., 2021b). Currently, only a single paper has identified a zDHHC23 mediated palmitoylation target, the Potassium Calcium-Activated Channel Subfamily M Alpha 1 channel (KCNMA1) (a BK channel protein), where palmitoylation was found to be essential for regulating it's cell surface expression/localisation (Jeffries et al., 2010). In regards to its role in cancer, zDHHC23 gene expression is reported to contribute to glioblastoma multiforme development and malignant progression through targeting (in an unknown manner) glioma stem cells for self-

renewal (Chen et al., 2019) and linked to regulating B cell malignancies (Edwards et al., 2016). Most recently, zDHHC23 gene expression has been suggested as a prognostic marker for osteosarcoma (Wan et al., 2021). These studies highlight that zDHHC23 may have important and tightly regulated cellular roles that cancer can manipulate for survival.

1.6.2. Maestro Heat Like Repeat Family Members

The second gene of interest, *MROH6*, is a member of the maestro heat like repeat (MROH) family (Mitchell et al., 2015). The MROH family of proteins remain a highly understudied group, with very little known about the structure or function of the proteins within it. Structurally, all MROH family members contain a HEAT repeat motif (Gray et al., 2012), which is similar to the much more studied armadillo repeat domain that forms rod-like helical structures and is found in several cytoplasmic proteins known to be involved with intracellular transport (Lipowsky et al., 2000). Currently there are nine members of the MROH family within humans, with evolutionary analysis shown in Figure 1.10.

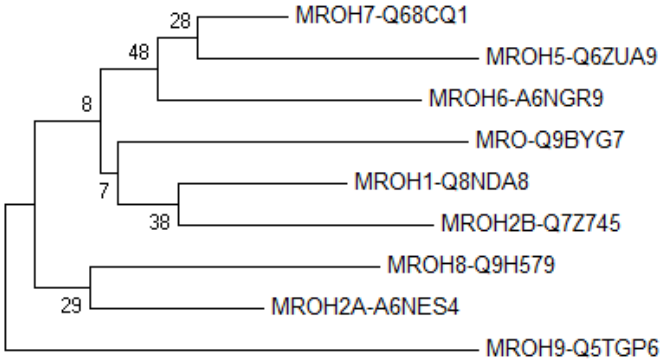

Figure 1.10 Molecular Phylogenetic analysis of MROH family members by Maximum Likelihood method.
The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-930.7355) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbour-Joining method to a matrix of pairwise distances estimated using a JTT model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 9 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 43 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

The human Maestro (*MRO*) gene was the first reported MROH family member, identified in a study comparing gene expression in human ovarian cumulus cells for cellular causes of progression into polycystic ovary syndrome (Kenigsberg et al., 2009). More recently, additional members of the MROH family have been identified as being associated with disease states, with MROH1 being upregulated in docetaxel-resistant metastatic prostate cancer (Cajigas-Du Ross et al., 2018). MROH2B acts as a mutational hot-spot with numerous mutations identified in skin cancer cases (Chen et al., 2015a), yet is also significantly downregulated in patient osteoblasts with Gorham-Stout disease (Rossi et al., 2020). MROH2A is among the top 30 significant genes which are induced in replicating B-cells (Klochendler et al., 2016). MROH5 is linked to smoking status and a chronic nicotine dependence brain disorder (Fan et al., 2021). MROH7 is overexpressed by over 10-fold in neonatal liver disease (biliary atresia) (Kyrönlahti et al., 2021), while MROH8 overexpression is associated with Alzheimer's disease (Potkin et al., 2009).

1.6.2.1. MROH6

Currently, MROH6 is annotated as a predicted protein in the human UniProt database (Consortium, 2019) and there is no known experimentally determined structure nor function. However, MROH6 has been identified in multiple studies investigating different disease states, including a study performed in rabbits identifying that MROH6 was upregulated in hypercholesterolemia/hypertension (Loke et al., 2017) and a study associating the MROH6 gene expression with lower levels of Indole3-propionic acid (tryptophan metabolite) in patients with liver fibrosis (Sehgal et al., 2021). The MROH6 gene also shows some cell type expression specificity to hepatocytes, with significant gene expression in liver injury in pre-adolescent children (Moylan et al., 2022). In regards to cancer, the MROH6 gene expression has been significantly associated with EGFR amplification in glioblastoma (López-Ginés et al., 2021) and has also been identified to be differentially expressed (mRNA level) and associated with metastasis in breast cancer (Qian et al., 2021). To date, no studies have investigated MROH6 at the protein level.

1.7. Proteomics

The term 'proteome' was originally used to describe all expressed proteins within a cell or tissue at a given time. Now, the term has evolved to reflect the fact that different PTM status' of proteins (proteoforms) is exponentially far more complex than simply the complement of expressed genes (Tyers and Mann, 2003). Advances in technology over the past 30 years has seen mass spectrometry (MS) become an integral part of proteomics analysis; now capable of going beyond just protein identification, this technology is used extensively for protein quantification (using suitable controls), determine PTM status and sub-cellular localisation, and binding partners be they proteins, nucleic acids, or small molecules. Structural MS can also be used to investigate the stoichiometry of complexes and protein conformational dynamics arising as a result of differential ligand binding (Han et al., 2008, Aebersold and Mann, 2016, Mallick and Kuster, 2010, Schneider et al., 2018, Lanucara et al., 2014).

1.7.1. Bottom-up proteomics

MS-based proteomic analysis can be performed on full length 'intact' proteins, known as Top-Down proteomics. However, there are multiple technical difficulties to Top-Down analysis including: 1) Inefficiency of current protein fragmentation technology (described further in Section 1.8.3) in a gaseous phase, required to identify the protein. 2) Lack of easily applicable fractionation techniques, required to decrease the complexity of samples at any one time, thus allowing much greater depth of information to be obtained from a sample. 3) Ionisation of proteins, required to move the protein in a gaseous state within the mass spectrometer, proves to be challenging and highly protein dependent (Lanucara and Eyers, 2013, Zhang et al., 2013). All of these difficulties can be circumvented through the known/specific proteolysis of proteins into peptides, which are then used to infer protein identification from peptide sequences; known as a 'Bottom-Up' or shotgun-proteomics approach (due to its similarity to shotgun genomic sequencing (Zhang et al., 2013)). Typically, proteins are extracted from the biological sample of interest, denatured and proteolyzed following reduction and alkylation of Cys disulphide bonds.

While there are several proteases that have proved useful for peptide-based proteomics analysis, the most commonly used is Trypsin, due to its highly specific and reliable cleavage

pattern at the C-terminus of Arginine (R) and Lysine (K) residues, expect where a Proline (P) residue is in the +1 relative to the site of cleavage. The advantage of tryptic peptides is that they typically generate doubly protonated peptide ions following electrospray ionisation (ESI, Section 1.8.1), due to the basic N-terminal amino group and the C-terminal K/R residue. These doubly protonated species are ideal for sequence determination using collision-mediated fragmentation strategies and are of an ideal size for separation by reversed-phase chromatography prior to tandem MS (LC-MS/MS), and fall within the *m/z* range of most mass analysers (Section 1.8.2) (Paizs, 2005, Michalski et al., 2012). However, proteolysis has the clear disadvantage of increasing sample complexity (one protein resulting in multiple peptides). Therefore, even with common liquid chromatography coupled to MS analysis (LC-MS/MS), complex samples (cell lysates) are likely subject to competitive/preferential ionisation of peptides resulting in the MS unable to analyse all peptides available at any one time and missing data. As a result, low abundant proteins and peptides containing PTMs are difficult to analyse without the application of fractionation, purification or enrichment strategies (Mann et al., 2002). One method which aids in analysis of specific proteins of interest is immunoprecipitation (IP).

1.7.2. IP-coupled proteomics

Proteins rarely function as separate entities but rather as highly complex, multi-protein complexes (i.e. binding partners/cofactors) (Alberts, 1998). Additionally, for a PTM to occur, a modifying enzyme must come into contact with the target protein, even if transiently; therefore, it is of biological relevance to investigate binding partners. One way of investigation protein interaction networks is to purify the protein of interest under conditions that retain these interactions, prior to LC-MS/MS analysis. The most common method of protein purification relies on using an antibody against the target protein, allowing purification not only of the protein of interest but also any interacting proteins at that specific time (termed Co-immunoprecipitation (Co-IP)) (Dunham et al., 2012). While it is possible to identify these protein interactors post IP through alternative techniques to LC-MS/MS, such as western blotting, these are mostly low throughput and inherently biased due to the requirement of prior knowledge on potential binding partners to investigate. Proteomics, however, provides an unbiased and High-Throughput (HTP) identification approach to

significantly improve the understanding of protein interaction networks (Mallick and Kuster, 2010, Dunham et al., 2012).

1.7.3. Phospho proteomics

As eluded to, the proteome is highly variable in nature, partly due to the dynamic nature of PTMs (Mann et al., 2002). As mentioned, PTMs are typically sub-stoichiometric. Combined with peptide-based proteomics approaches where protein sequence coverage may be lacking and the biophysical properties of the modification may influence its ability to be detected in an LC-MS/MS experiment, it can be difficult to identify and localise the sites of PTM, even after IP. PTM-specific enrichment strategies have therefore been developed that enrich a specific PTM in the background of unmodified peptides, vastly reducing the complexity of a sample. Several enrichment strategies for phosphorylation have been developed, including antibody-based techniques and strategies that exploit the negative charge introduced by the phosphate group. Generic anti phospho-tyrosine antibodies for IP can be a fairly efficient enrichment strategy for targeted investigation of this specific modified residue. However, such generic antibodies against phospho-serine and –threonine do not exist, being sequence or motif specific (Olsen et al., 2006, Fíla and Honys, 2012). Alternative phosphopeptide enrichment strategies have been developed and have much greater success at enriching the phosphoproteome in an unbiased manner including: targeting the negative charge introduced by phosphorylation (by using metal cations bound to a solid matrix, known as Immobilised Metal Ion Affinity Chromatography (Rainer and Bonn, 2015) or targeting the highly specific spatial coordination of the two hydroxyl groups introduced by phosphorylation (through hydrogen bonding to metal oxides) (Connor and McQuillan, 1999). The most commonly used, and utilised within this thesis is phospho-peptide enrichment with Titanium dioxide (TiO_2).

1.7.4. Liquid chromatography tandem mass spectrometry (LC-MS/MS)

MS is a non-recoverable, destructive analytical method that measures the mass to charge (*m/z*) ratio of ions and can thus be used to determine analyte structure and composition.

Consequently, MS can be used to determine the mass and amino acid content of peptides and thus be used to deduce the gene expression product from which it derived (Guilhaus, 1995). As mentioned, sample complexity is often too high for in-depth (particularly PTM analysis) by MS/MS analysis alone. To help deconvolute this complexity, a Liquid Chromatography (LC) separation step is commonly coupled prior to MS/MS analysis, in a Liquid chromatography–mass spectrometry (LC-MS/MS) pipeline. The most common form of LC is reverse phase, a method in which peptides are bound to a C18 (hydrophobic) column and eluted using an increasing concentration gradient of acetonitrile (ACN), thus separating peptides based on their hydrophobicity. The LC system is typically coupled in-line (but can also be performed offline) with the ESI (Section 1.8.1) source of the mass spectrometer, allowing eluted peptides to directly enter the MS for analysis. This reduction in sample complexity at any given point in time results in a dramatic increase in the depth of coverage of a sample in MS-based proteomic analysis.

A mass spectrometer typically consists of three essential components: an ionisation source, a mass analyser and a detector, these components will vary depending on the instrument. In LC-MS/MS the ionisation source is coupled to the LC elution, allowing ionisation of peptides in a gaseous phase before entering the mass spectrometer. The mass analyser manipulates ions based on their *m/z* ratio whilst the detector records the ions signal (and intensity) when ejected from the mass analyser. Mass spectra then record the relative intensity of all ions detected and their *m/z* ratios. Following specific fragmentation (Section 1.8.3), MS/MS spectra can be used to determine the primary sequence of the peptide, which can then be searched against a database containing all potential proteins in a sample to identify the protein which the peptide ion comes from. Thus, a MS2 spectrum can be used to identify a peptide and the respective gene product it derived from.

To identify peptide primary sequence, required for protein inference, tandem MS (MS/MS) is generally implemented in a data-dependent acquisition mode (DDA). Initially, the *m/z* ratio of all peptide ions eluting at any given time is recorded in an MS1 scan. Subsequently, the most intense ions are sequentially isolated and fragmented, with fragments analysed by a second round of MS with the *m/z* ratio of each product ion recorded to generate a MS2 spectrum. DDA MS/MS is commonly combined with a *m/z*-based exclusion time window, thus preventing the same precursor *m/z* ion from being selected for multiple rounds of fragmentation in subsequent MS1 scans. Hence, a 'Top X' approach can be adopted to obtain MS/MS analysis of a defined number of the most intense precursor ions from a MS1 scan.

1.8. Thermo orbitrap fusion Tribrid

There are numerous types of mass spectrometer; for the purposes of this thesis, I will focus on those employed in these studies. The Thermo Orbitrap Fusion Tribrid mass spectrometer is an example of the advancements made in the development of MS instrumentation, incorporating three different types of mass analyser (a quadrupole, linear Ion trap and Orbitrap mass analyser). It is also capable of performing (at least) three different fragmentation strategies: higher collision-energy dissociation (HCD), Collisional induced dissociation (CID) and Electron transfer dissociation (ETD) (Figure 1.11), discussed further in Sections 1.8.3.1-1.8.3.2.

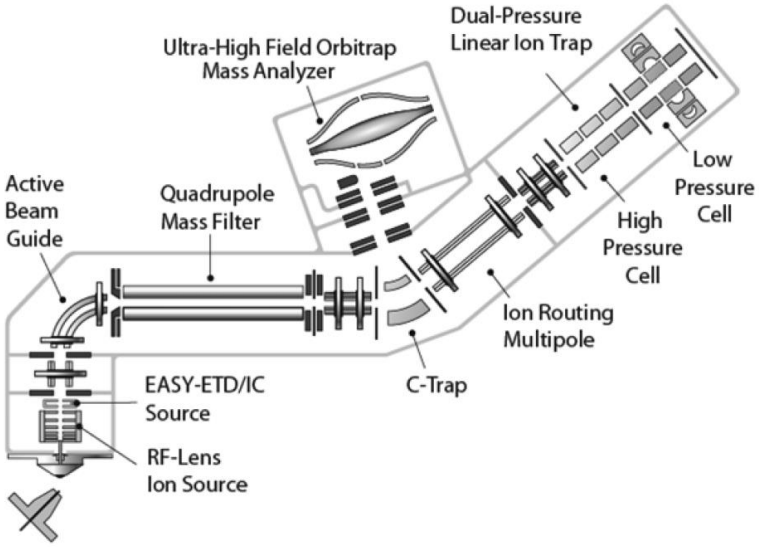

Figure 1.11 Schematic view of the Thermo Orbitrap Fusion Tribrid Mass Spectrometer.
Labelled areas identifying all major components of the MS, including: ion guides, mass analysers, ETD source and electrospray ionisation source. Taken from (Senko et al., 2013).

A significant advantage of the Thermo Orbitrap Fusion Tribrid mass spectrometer utilising multiple mass analysers, is that it is possible to simultaneously perform individual MS analyses (MS1, MS2 or MS^n) in different mass analysers, thus maximising efficiency by allowing MS/MS analysis to be performed in a time-dependent mode rather than a Top X mode (Senko et al., 2013). Additionally, the Thermo Orbitrap Fusion Tribrid mass spectrometer's ability to fragment ions using different approaches (discussed further in Section 1.8.3) allows for high flexibility dependent on purpose, for example application of ETD it is highly efficient for in-depth localisation of sites of phosphorylation.

1.8.1. Electrospray ionisation

In order to analyse peptide ions, 'soft' ionisation mechanisms are required to ensure that ions generated are not simultaneously fragmented. Electrospray ionisation (ESI) is the most common method for proteomics analysis. Briefly, a sample, usually in a volatile solvent (ACN) under acidic conditions is passed through a high voltage capillary which creates a highly charged droplet cloud which is of equal charge to the capillary (Fenn et al., 1989). The charged droplets are heated as they proceed through a high-low pressure gradient, resulting in the desolvation of droplets and creating multiply charged gaseous ions which can then be manipulated and analysed by MS (Fenn et al., 1989, Kebarle and Tang, 1993). The ions produced by ESI are in the form of $[Mass+nH]^{n+}$ where n = the number of protons.

1.8.2. Mass analysers

There are two types of mass analyser relevant to this thesis: Ion trap and Orbitrap and are discussed below. Upon entering the mass spectrometer, ions are initially guided into a mass analyser through ion optics. These mass analysers use a variable scanning magnetic (or electric field) in order to selectively (resonance-based) eject ions dependent on *m/z*, permitting all single *m/z* ratio ions to be detected. Mass analysers have their own inherent properties and advantages/disadvantages, including; allowing the analysis of different *m/z* ranges, resolution (precision of m/z measurement), sensitivity and mass accuracy (Haag, 2016, Yates et al., 2009). The previously discussed Thermo Orbitrap Fusion Tribrid consists of two main mass analysers (Ion trap and Orbitrap), with the quadrupole mass analyser typically being used in a *m/z* filtering mode only.

1.8.2.1. Ion trap

The ion trap mass analyser was introduced in the 1950s by Paul and Steinwedel and relies on a high frequency, oscillating electric field to spatially capture all ions within electrodes. Ions are trapped by an initial radio frequency (RF) and a fixed direct current voltage, hence all the ions within a given *m/z* range are trapped within the mass analyser. A scanning/varying RF voltage is subsequently applied to the electrodes, this results in resonance of a specific *m/z*

peptide ion and ejection from the mass analyser and detection by the detector; thus ejection/detection is dependent on the *m/z* ratio (Stafford Jr et al., 1984). The trajectory of trapped ions is dependent on the design of the ion trap used with two main types being the 3D ion trap (also known as Paul's trap) and 2D linear ion trap (Figure 1.12). 3D ion traps are not used in the instruments of relevance to this thesis and are therefore not discussed further.

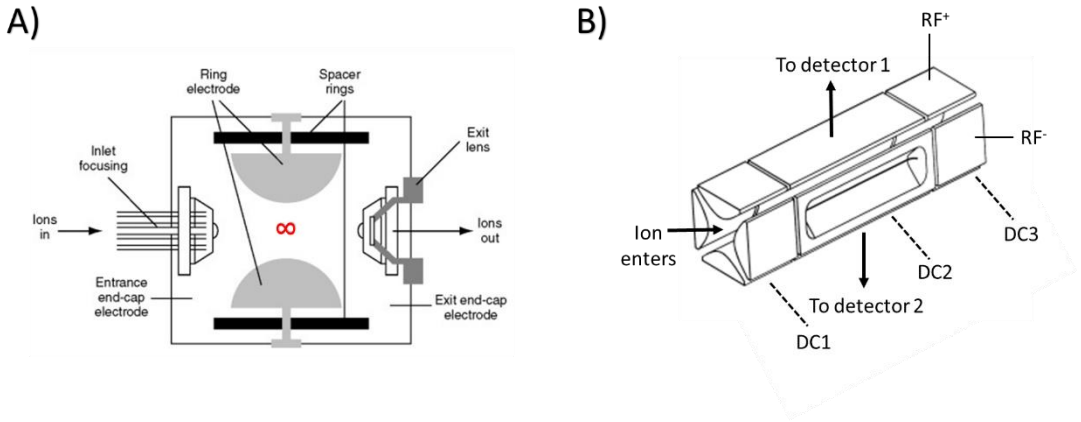

Figure 1.12 Schematic view of Ion trap mass analysers.
A) 3D (Paul) Ion trap, consists of two hyperbolic end cap electrodes (same AC voltage) surrounding a ring electrode to trap ions in a figure of 8. Amended from (Yinon, 2007). B) 2D Linear Ion trap, consisting of four parallel, oppositely charged, hyperbolic pole electrodes (varying RF frequencies) with capping electrodes (DC voltages) to trap ions within an internal cavity cloud. Detectors are perpendicular along the length of the Ion trap. Amended from (Schwartz et al., 2002).

The 2D linear ion trap consists of four parallel hyperbolic pole electrodes of adjacent opposite charges surrounded by end cap electrodes, resulting in ions trapped as a linear cloud (where ions spiral back-and-forth along the length of the Ion trap) (Figure 1.12 B). In this mass analyser, the RF voltage is applied to the end caps to result in selective resonance ejection of a particular *m/z* ratio perpendicularly to the Ion trap direction, enabling ion detection (Schwartz et al., 2002). Due to the structure, 2D linear Ion traps have a larger internal capacity compared to other mass analysers, allowing much larger accumulation of ions before space-charge ion effects occur, hence greatly increasing sensitivity. However, Ion trap's lack mass accuracy and resolving power compared to the more advanced mass analysers such as the Orbitrap (Douglas et al., 2005).

1.8.2.2. Orbitrap

The Orbitrap, which was invented by Makarov in the 1990's, consists of a spindle-like pole electrode that is encapsulated by concave electrodes to form a barrel-like structure (Figure 1.13) (Makarov and Scigelova, 2010). A linear electric field is applied to the electrodes causing ions to oscillate bi-spatially (both orbitally around and axially along the spindle), where the frequency of both types of oscillation are independent of each other (Makarov, 2000). The frequency at which an ion axially oscillates is directly proportional to the *m/z* ratio of the ion. Thus, detecting the frequency of all ions trapped generates a complex repeating wave which can be deconvoluted into individual ion wave component frequencies by Fourier Transformation, where each frequency is proportional to the *m/z* value and intensity. Transformation of this frequency data can thus be used to generate a mass spectrum (Scigelova et al., 2011). As frequency can be measured much more accurately than time, Orbitraps provide high mass accuracy and resolution. However, they are predisposed to space-charge effects, therefore resulting in decreased sensitivity compared with other types of Ion traps (Scigelova et al., 2011).

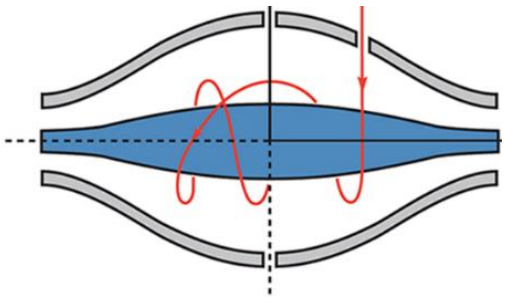

Figure 1.13 Schematic view of an Orbitrap mass analyser.
An Orbitrap consists of two concave electrodes forming a barrel around a spindle electrode. Ions are injected in 'packets' at time intervals through a hole in an outer electrode. Ions then oscillate both orbitally around and axially along the spindle. Taken from (Makarov and Scigelova, 2010).

1.8.3. Peptide fragmentation

In order to determine peptide sequence, peptide ions must be reliably fragmented in a known manner before being analysed by a second MS scan (MS2 or MS/MS). Multiple

fragmentation methods have been developed for peptide analysis typically resulting in fragmentation along the peptide backbone, including collisional-based and electron meditated mechanisms, as discussed below (Figure 1.14) (Hunt et al., 1986, Johnson et al., 1988, Zubarev, 2004, Syka et al., 2004, Frese et al., 2012).

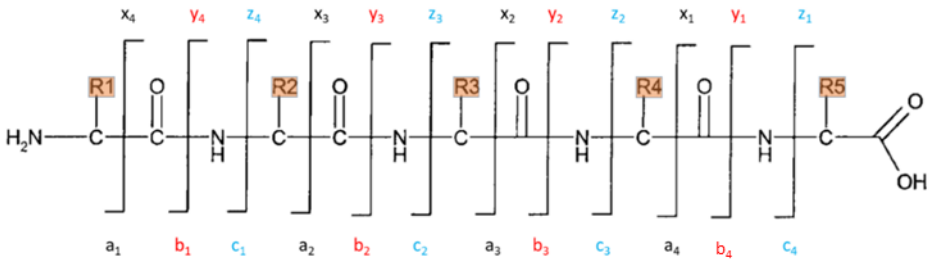

Figure 1.14 Nomenclature of peptide fragmentation.
a, b and c ions identify different fragmentations of the peptide backbone and numbered sequentially from the amino-terminus of the peptide. Their respective counterparts x, y and z ions are labelled from the carboxy- terminus of the peptide. Orange highlighted R groups represent unknown amino acid side chains. CID/HCD predominantly results in b, and y ions, whilst ECD/ETD predominantly results in c and z ions. EThcD fragmentation results in all four ion types. Amended from (Paizs, 2005).

1.8.3.1. Collisional induced dissociation

Collision-induced dissociation (CID) is the most commonly used fragmentation technique. There are two types of CID: resonance CID (performed in Ion trap's (Johnson et al., 1988)) and beam type CID, also known as higher energy collisional dissociation (HCD, performed in a specialised collision cell (Olsen et al., 2007)). Resonance CID involves the selective resonance excitation of a single *m/z* ratio, achieved by changing the RF voltage, in turn causing the ion to collide with an inert gas (typically helium) and subsequent fragmentation (Johnson et al., 1988). Thus, only a single fragmentation event can occur per peptide ion as resultant fragments have a different *m/z* ratio and are thus no longer excited. HCD relies on the same principles as resonance CID, however, typically uses higher energy and is not constrained by the requirement for resonant excitation. Firstly, HCD involves filtering a specific *m/z* ratio and accelerating peptide ions into a specialised collision cell (under a high pressure of inert gas, typically nitrogen) by use of a variable RF voltage (Olsen et al., 2007). Due to HCD not requiring resonance excitation, fragment ions can undergo secondary fragmentation events to provide greater depth of fragmentation, including the generation of

internal ions (Olsen et al., 2007). Independent of CID type, fragmentation typically occurs along the amide bond to produce a heterogeneous mixture of b- and y-ions, dependent on how the charge localises during fragmentation (Figure 1.14). The b-ions are derived when N-terminal peptide ion fragments maintain the charge, whilst y-ions are derived from C-terminal peptide ion fragments (Johnson et al., 1988). By knowing exact fragmentation position and associated mass shifts, using the exact mass of b/y ions determined by MS/MS, the primary sequence of the peptide can be determined by calculating the mass differences between identified *m/z* ratio peaks in the MS2 spectrum (Johnson et al., 1988, Hunt et al., 1986). Unfortunately, due to the collisional/vibrational energies associated, both CID and HCD methods are susceptible to inducing neutral loss, which is particularly problematic for site localisation e.g. of phosphopeptides (Zhang et al., 2009).

1.8.3.2. Electron transfer dissociation

Electron transfer dissociation (ETD) is an alternative, non-energetic fragmentation method that uses an external electron donor, a gaseous vehicle chemical that promotes the absorption of thermal electrons by multiply protonated peptide ions, resulting in the formation of hypervalent species that are unstable resulting in spontaneous fragmentation at the N-Cα bond and formation of c and z ions (Figure 1.14) (Zubarev, 2004, Zubarev et al., 1998, McLafferty et al., 2001, Syka et al., 2004). In ETD, c-ions are derived from N-terminal peptide ion fragments retaining charge while z-ions deriving from C-terminal peptide ion fragments retaining charge (Figure 1.14) (Zubarev et al., 1998). Similar to that of resonance CID/HCD, deducing the exact mass of c/z ions allows the determination of peptide primary sequence. However, ETD fragmentation is limited due a prerequisite of higher multiply charged species (3+), due to electron-induced charge-state reduction and highly inefficient fragmentation of singly charged (+1) peptide ions (Sobott et al., 2009, Good et al., 2007). Thus, recent developments have been to utilise a dual fragmentation approach, using both HCD and ETD fragmentation in a step wise manner (EThcD) (Frese et al., 2012). This fragmentation method results in a highly heterogeneous mixture of b, y, c and z ions (Figure 1.14), allowing more in-depth data to determine primary sequence whilst increasing the chance that any PTMs are not lost (Frese et al., 2012, Frese et al., 2013).

1.9. Data analysis

In order to interpret LC-MS/MS spectra to identify peptides and proteins from the recorded tandem mass spectra, search engines are used to automatedly match spectra against a database which contains all potential peptides from proteins in the sample. Two search engines are relevant to this thesis: MASCOT (Perkins et al., 1999) and PEAKS DB (Zhang et al., 2012). These search engines use complex algorithms to match MS2 spectra (also known as peptide spectral matches (PSMs)) against a database to determine primary peptide sequence and thus infer protein identity. However, while unlikely for random proteins, a peptide may not be unique to a single protein and can be found in multiple proteins (usually a family/isoforms of proteins). Therefore, a maximal parsimony approach to infer protein identification is adopted (where a non-unique peptide is considered to be from a protein that is identified by additional unique peptides), hence a score is applied to adjudge confidence in protein identification that is based on the sum of how confidently all PSMs match all peptides from a protein sequence match in an automated manner.

Due to the vast quantity of data obtained from an LC-MS/MS experiment, it is important to estimate and filter data based on the false-positive rate of PSMs randomly matching within the database (a false discovery rate - FDR). To evaluate FDR, all PSMs are subject to a second database search against a fictional database of equal size, usually the target database protein sequences reversed, and a comparison of the number of matching PSMs between databases gives an indication to FDR. Therefore, FDR infers the probability that any given PSM in the real dataset matches due to chance, allowing an FDR cut-off to be applied (typically 1%) and thereby statistically controlling the number of false positive proteins identified.

In PTM discovery, data analysis requires considering the known mass-shift change introduced by the PTM. PTM identification occurs by accounting for the mass shift at the peptide level in the MS1 spectrum, and localisation to a specific residue at the fragmentation MS2 spectrum level (Perkins et al., 1999). However, as previously discussed, ionisation and (particularly) fragmentation methods can result in the neutral loss of PTMs and difficulty in site localisation (Mann et al., 2002), both essential for biological interpretation and functional characterisation of the PTM. As a result, it is common practice in proteomics to provide a site localisation confidence score of identified PTMs. This is determined using a bioinformatics approach to analyse specific product ion peaks to predict PTM site localisation, in this project

the site localisation tool ptmRS (Taus et al., 2011) has been used with in conjunction with MASCOT (Perkins et al., 1999) as part of the Proteome Discoverer (PD) pipeline.

1.9.1. Quantitative proteomics

In order to understand the proteome, considering the various multiple and sub-stoichiometric complex nature of PTMs and protein interactions, it is essential to quantitatively measure these changes between conditions. LC-MS/MS is not an inherently quantitative technique due to different sequence (and PTM) dependent ionisation efficiencies of peptides (Bantscheff et al., 2007). However, proteins can be quantified by multiple strategies, either in relative terms (fold change comparisons between conditions) or absolute terms (the number of copies of a protein a cell expresses). Relative quantification of proteins can be divided into two distinct categories: labelled or label-free approaches (Bantscheff et al., 2012).

1.9.1.1. Label-based quantification

Label-based quantification approaches are considered more accurate than label-free as they permit individual labelling of samples prior to mixing together at an early stage in sample preparation, hence the samples experience near-identical experimental and analytical variance (Bantscheff et al., 2012). Generally, labelling requires the covalent incorporation of a known mass difference, with strategies either achieved by in-cellulo metabolic C13 labelling (Ong et al., 2002) or through post-digest isobaric tagging of peptides; either tandem mass tagging (TMT, (Thompson et al., 2003)) or Isobaric tags for relative and absolute quantitation (ITRAQ, (Ross et al., 2004)). However, these experiments require all samples to be obtained simultaneously, this is difficult with large scale experiments across multiple conditions. Additionally, these tagging methods are expensive and require optimal (near 100%) labelling efficiency, therefore label free quantification was utilised in this thesis.

1.9.1.2. Label free quantification

Relative quantification using a label free approach is often considered less accurate compared to label-based methods, due to samples being prepared and analysed independently (Bantscheff et al., 2012). However, variation in the precision of label free quantification is highly dependent on sample handling and the appropriate use of normalisation tools. Importantly, label free quantitation is both rapid and significantly cheaper method than label-based techniques, with effectively no limitations in the number of samples that can be sequentially analysed, whereas current labelling technology only allows for a maximum of 18 samples (Li et al., 2021). Due to the potential for increased variation, label free approaches require robust normalisation methods to compensate for any experimental error. Prior to normalisation, correction of LC elution times between runs is completed, aligning same m/z ratios within a time window. Normalisation is calculated against total intensity or against house-keeping proteins (albeit rarer) that are assumed to be of identical abundance in different samples (Gillet et al., 2016, Välikangas et al., 2018). Multiple strategies have been developed to infer protein abundance from peptides, including; spectral counting and peptide intensity averaging. Spectral counting is dependent on the fact that for a given protein, a greater number of PSMs will be identified for a protein of greater abundance when normalising against protein length (Washburn et al., 2001, Arike and Peil, 2014, Old et al., 2005). However, spectral counting is controversial as it does not measure any physical traits of the identified peptides and is fundamentally flawed by the inclusion of dynamic exclusion windows that prevent the same *m/z* ratio being used for multiple MS2 scans; limiting the total number of PSMs for a protein (Old et al., 2005).

Alternatively, a more computationally involved approach is available that determines the intensity-of every *m/z* peptide ion for comparison. The intensity of each *m/z* peak is calculated at the MS1 level by integrating ion intensity against the timescale of the LC chromatogram, allowing the peptide area intensity to be determined. The use of integration over time takes into consideration the different physiochemical properties of every peptide that can result in different elution width windows. Each PSM that is identified for a protein therefore has an associated intensity value which are averaged to infer protein intensity, which is then compared between experimental conditions (Old et al., 2005, Bantscheff et al., 2012, Bondarenko et al., 2002).

1.9.2. Software tools

There are various open source and commercial software packages that are available to analyse LC-MS/MS data. The ideal software tool is dependent on the type of analysis required and can be influenced by the experimental design and type of MS instrumentation used. In this thesis, Proteome Discoverer was used, in conjunction with the MASCOT search engine, for protein identification. Phosphorylation analysis was performed by application of the ptmRS node, as described in Section 1.9.

1.9.2.1. PEAKS PTM

PEAKS PTM was utilised to perform open PTM searches, a computationally heavy method to identify biological and artefactual PTMs. Unlike MASCOT, PEAKS performs an initial de-novo sequencing step where every PSM is directly matched to a peptide sequence (and hence a protein) (Zhang et al., 2012, Ma et al., 2003). For PEAKS-PTM, a second search is then performed on a new database that only contains protein identifications that were found in the first search, and applies *m/z* ion mass shifts and peptide spectral matching algorithms for all PTMs stored in the UniMod database to filter through all unidentified PSMs for possible PTM states of peptides (Han et al., 2011, Creasy and Cottrell, 2004). Due to the immensely greater search space introduced by searching for all PTMs, PEAKS PTM uses a much stricter FDR calculation that consists of using the initial large database and using it as a reversed fictional database to search against in an 'open' PTM search. These identifications are then compared against the identifications in the smaller correct orientation database, reducing the potential bias for underestimating FDR in these styles of data analyses (Bern et al., 2009). As such, PEAKS was used for 'open' PTM searching of all LC-MS/MS data collected during this thesis.

1.10. Research aims

NB stratification is currently heavily reliant on *MYCN* amplification as a criterion for NB treatment, despite only 25% of NB patients possessing such amplification. As a result, there is currently a 'gap' in the risk classification/treatment of non-*MYCN* amplified NB patients. Considering the extremely harsh nature of NB treatment required in young developing patients, with survivors experiencing extensive lifelong side effects, it is imperative that new biomarkers can be identified within the 75% of patients without *MYCN* amplification and stratify their associated risks for better treatment management. As zDHHC23 and MROH6 were identified in non-*MYCN* amplified cell lines and correlate with extremely poor prognosis, the aim of this research is two-fold:

- Firstly, investigate whether zDHHC23 and/or MROH6 could be used clinically as a novel prognostic biomarker for NB patients, through the development of a biomarker assay.

- Secondly, interrogate the molecular roles of zDHHC23 and MROH6 and their potential relevance in NB by exploring their signalling networks and PTM status. The characterisation of these proteins will hopefully provide mechanistic insight into the pathways via which zDHHC23 and MROH6 exert its role in NB, potentially revealing a novel target for therapeutics.

Chapter 2: Materials and methods

2.1. Chemicals and reagents

Tissue culture reagents were purchased from Gibco. Powdered chemical reagents and custom DNA primers were purchased from Sigma-Aldrich, unless otherwise stated. High-performance liquid chromatography grade MS solvents were purchased from Thermo Fisher.

2.2. Cell culture, transfection and treatment

2.2.1. Cell passaging

SK-N-AS neuroblastoma cells (ECACC No. 94092302) used within this thesis were maintained in Dulbecco's Modified Eagle Medium (DMEM) media with 10% Fetal Bovine Serum (FCS), 1% Non-Essential Amino Acids (NEAA) and maintained at 37 °C, 21% O_2 and 5% CO_2 and passaged when ~80-90% confluent. For passaging, culture medium was removed, cells were washed with phosphate buffered saline (PBS) and incubated for 5 min at 37 °C with 4 ml in a T175 flask (Corning) 1 x Trypsin/EDTA (~22.5 µl/cm^2). Detached cells were re-suspended in culture medium and counted using a Bio-Rad TC20 cell counter. Cells were seeded into either a 10 cm dish (1 x 10^6 cell count) or a 12 well pate (1 x 10^5 cell count) to maintain the cells for experimentation. Mycoplasma testing was carried out routinely using the MycoAlertTM Mycoplasma Detection Kit (Lonza) according to manufacturer ‘s instruction. SK-N-AS cells were stably labelled with Green Fluorescent Protein (GFP) using lentiviral particles produced with the transfer vector pLNT-SFFV-EGFP. Transduction efficiency was quantified by flow cytometry using a FACSCalibur Cytometer (BD Biosciences).

2.2.2. Transient transfection

Transient transfection was conducted 24 hr prior to experimental use. Transfection reagent (PEI 40K MAX, linear, Polyplus #24765-1) was used at the recommended Reagent: DNA ratio of 4:1, with 6 µg of DNA used for a 10 cm plate. The final transfection conditions for LC-MS/MS experiments (zDHHC23/MROH6) used throughout were as follows: reduced-serum medium (Opti-MEM, Thermo Fisher) was used to dilute DNA to a final concentration of 10

ng/µL, a ratio of 4 µL PEI: 1 µg of DNA was added. Transfection mixes were vortexed and left to stand at room temperature for 30 min. The volume of transfection mix added to cultured cells was equivalent to 5% of the total cell culture volume. All plasmids used are listed in Table 2.1. siRNA transfection was performed 72 hr prior to collection of samples using FuGENE HD. ON-TARGETplus SMARTPOOL siRNA (Dharmacon) was used for zDHHC23 (L-018381-01-0005) and MROH6 (L-034980-01-0005).

Table 2.1 Plasmids used and their availability.
N- and C- terminal constructs for both proteins with site directed mutagenesis (SDM) plasmid to create inactive zDHHC23 plasmid (zDHHS23).

Plasmid	Source
HA-MROH6	Synthesised (NovoPro Labs)
HA-zDHHC23	Synthesised (NovoPro Labs)
HA-mCherry	Cloned in-house
HA-mCherry zDHHC23	Cloned in-house
HA-mCherry zDHHS23	Cloned in-house
HA-mCherry MROH6	Cloned in-house
mCherry-HA	Cloned in-house
zDHHC23 mCherry-HA	Cloned in-house
zDHHS23 mCherry-HA	Cloned in-house
MROH6 mCherry-HA	Cloned in-house

2.2.3. Hypoxic incubation

Hypoxic incubation was at 1% O_2, performed in a Don Whitley H35 Hypoxystation. To ensure correct O_2 tensions, the hypoxic chamber was calibrated bi-weekly using a Microx 4 fibre Optic Oxygen Meter (PreSens). Cells were incubated in hypoxia for 72 hr before use.

2.3. Cell lysis and protein extraction

Post treatment (normoxia or hypoxia) cells were washed in PBS twice before adding 11.4 µL/cm^2 (equivalent to 600 µL / 10 cm plate) ice cold lysis buffer (50 mM Tris pH 8.0, 120 mM NaCl, 5 mM EDTA, 0.5% (v/v) NP-40, 1X EDTA-free cOmplete protease inhibitor (Roche) and 1X phosSTOP (Roche). Cells were scraped and supernatant collected into Ultra-High recovery

Eppendorf tubes (STARLAB). Lysates were rotated end-over-end for 30 min at 4 °C, before centrifugation at 10,000 *g* for 10 min at 4 °C with the cleared supernatant collected.

2.3.1. Protein concentration determination

Protein quantification was performed using the Pierce BCA Protein assay Kit (Thermo Fisher) following the manufacturer's recommended protocol. A reference standard was used, via serial dilution of bovine serum albumin (BSA) to 1.5, 1.0, 0.5, 0.25 mg/mL.

2.4. Immunoprecipitation (IP)

Post treatment by either normoxic/hypoxic incubation, $5x10^6$ cells (1x 10 cm plate) were lysed as previously described in Section 2.3 (600 µL lysis buffer). The Pierce CrossLink Magnetic IP/CO-IP Kit (#88805) was used, following manufacturer's recommended protocol. Briefly, using a magnetic stand for all bead collection steps, 25 µL of protein A/G bead slurry were washed in the provided lysis buffer (25 mM Tris pH 7.4, 120 mM NaCl, 1 mM EDTA, 1% (v/v) NP-40, 5% (v/v) glycerol) before addition of 2 µg of zDHHC23/MROH6 antibody (listed in Table 2.2) in 100 µL 1X coupling buffer (provided). Samples were agitated at 600 rpm for 15 min at room temperature to pre-bind the relevant antibody to the magnetic protein A/G beads. Antibody-bead complexes were washed in the supplied lysis buffer, 1 mg of cleared protein lysate was added and left to rotate end-over-end for 18 hr at 4 °C. IPs were washed 3 X in lysis buffer and eluted in 30 µL of 10% SDS with heating at 95 °C for 10 min, discarding the beads. Cleared supernatant was heated with 10 µL of 4 X Laemmli buffer.

2.4.1. HA-tag IP

Following transfection of SK-N-AS cells ($5x10^6$ cells; 1x 10 cm plate) with HA-tagged plasmid and subsequent 72 hr incubation in either normoxia or hypoxia, cells were processed for IP identically to the above section. HA-beads (Thermo Fisher) were added to lysate as a negative control for the IP experiments with precomplex-HA beads and left to rotate end-over-end for 18 hr at 4 °C prior to washing and elution using the same conditions as stated for CrossLink IP. Cleared supernatant was boiled with 10 µL of 4 X Laemmli.

2.4.2. Palmitoylation Copper-catalysed azide-alkyne cycloaddition (CuAAC) assay

Cells were transfected (Section 2.2.2) and incubated for 72 hr in either 21% or 1% O_2. Cells were dosed with 25 µM Azido Palmitic Acid (Click Chemistry Tools, #1346-5) 24 hr prior to use. Cells were lysed (Section 2.3) followed by the addition of: 100 um TBTA, 1 mM $CuSO^4$, 1 mM TCEP and 100 µM Alkyne Magnetic Beads (Click Chemistry Tools, #1035-1) and left to shake at 1500 rpm for 2 hr at room temperature. Using a magnetic stand for all bead collection steps, the magnetic beads were washed 3 X in 2% SDS and 5 X in 100 mM AmBic. Samples were alkylated and reduced (as described in Section 2.10.1) followed by digestion by 10:1 (w/w) Trypsin Gold (Promega) incubation at 37 °C for 18 hr with 900 rpm shaking. Post digestion, trifluoroacetic acid (TFA) and ACN was added to a final concentration of 1.5% (v/v) and 3% (v/v), respectively, and incubated at 37 °C with 600 rpm shaking for 1 hr. The sample was centrifuged at 13000 *g* for 10 min at 4 °C and the clear supernatant collected.

2.5. Gel based analysis

2.5.1. SDS-PAGE sample preparation

Cleared protein lysates (30 µL) were mixed with 4 X Laemmli buffer (10 µL) (250 mM Tris-HCl pH 6.8, 30% (v/v) glycerol, 10% (w/v) SDS, 500 mM DTT, 0.05% (w/v) bromophenol blue) and heated at 95 °C for 10 min.

2.5.2. SDS-PAGE

Mini-gels were made using ~1.5 mL of 4% stacking gel (4% (v/v) acrylamide, 125 mM Tris-HCl pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) APS, 0.1% (v/v) TEMED) and 5 mL of a lower 10% resolving gel (10% (v/v) acrylamide, 375 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) APS, 0.1% (v/v) TEMED). For IP efficiency determination, an equal percentage of protein sample was loaded per lane. For western blotting, all gels were loaded with 5 µL of Colour prestained Protein Standard, Broad Range (10-250 kDa, New England Biolabs, P7719). Electrophoretic separation was performed at 200 V until the bromophenol blue dye front reached the end of the gel (~45 min) in electrophoresis running buffer (25 mM Tris pH ~8.3, 192 mM Glycine, 1% (w/v) SDS).

2.5.3. Coomassie staining

Colloidal coomassie stain was prepared consisting of: 0.12% (w/v) Coomassie G250 dye, 10% (w/v) ammonium sulphate, 10% (v/v) phosphoric acid and 20% (v/v) methanol. Following electrophoretic separation of SDS-polyacrylamide gels, they were immersed in colloidal coomassie solution and allowed to stain for 18 hr at room temperature on an orbital shaker and destained in milliQ H_2O on an orbital shaker at room temperature. Water was changed frequently until the gel was clear. Images were captured using an Epson Scanner.

2.5.4. Western blotting

Protein transfer onto nitrocellulose membrane (0.2 µm, BIORAD) was performed in transfer buffer (electrophoresis buffer + 10% (v/v) methanol) for 2 hr at 300 mA, 4 °C. Membranes were blocked in a 5% (w/v) skimmed milk solution (Marvel) dissolved in TBST (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% (v/v) Tween-20) for 1 hr at room temperature on an orbital shaker. Membranes were washed 3 X 10 min in TBST with shaking before addition of primary antibody for overnight incubation at 4 °C rotating on a roller. After primary incubation, membranes were washed 3 X 10 min in TBST and incubated with the secondary antibody for 1 hr at room temperature on a roller. All antibodies used are listed in Table 2.2, and were diluted in 5% (w/v) skimmed milk solution in TBST. After secondary antibody incubation, membranes were washed 3 X 10 min in TBST and detected using ECL reagents: SuperSignal™ West Pico PLUS (Thermo Fisher). Images were captured using CL-XPosure Film (Thermo Fisher) and Protec ECOMAX X Ray Film Processor.

Table 2.2 Summary of antibodies used for western blotting and/or IP.

ANTIBODIES					
Target	**Company**	**Catalogue number**	**Dilution factor**	**Monoclonal vs Polyclonal**	**Host species**
ARPC5	Thermo Fisher	MA5-32019	1:500	Monoclonal	Rabbit
BCAR1	Thermo Fisher	TA502536	1:500	Monoclonal	Mouse
GAR1/NOLA1	Thermo Fisher	PA5-99801	1:500	Polyclonal	Rabbit
HA	Cell Signalling	3724S	1:1000	Monoclonal	Rabbit
MROH6	Thermo Fisher	PA5-64355	1:1000	Polyclonal	Rabbit
	Atlas Antibodies	HPA068049	1:1000	Polyclonal	Rabbit
PSMB6	Thermo Fisher	PA1-978	1:500	Polyclonal	Rabbit
Vinculin	Thermo Fisher	MA5-11690	1:1000	Monoclonal	Mouse
zDHHC23	Novus Biologicals	NBP1-82688	1:1000	Polyclonal	Rabbit
	Thermo Fisher	PA5-72909	1:1000	Polyclonal	Rabbit
ZNF277	Thermo Fisher	PA5-21561	1:500	Polyclonal	Rabbit
Mouse	Cell Signalling	7076S	1:3000	Not specified	Horse
Rabbit	Cell Signalling	7074S	1:3000	Not specified	Goat

2.6. Bacterial expression

2.6.1. LB (Luria-Bertani) broth

A 2.5% (w/v) solution of LB was made by diluting LB broth powder (Miller, Merck Millipore) in MilliQ H_2O, as per manufacturer's recommendations. LB solution was autoclaved at 121 °C for 15 min, allowing to cool to room temperature before use.

2.6.2. LB agar and antibiotic selection

A 3.7% (w/v) solution of LB agar was made by diluting LB agar powder (Miller, Merck Millipore) in MilliQ H_2O, as per manufacturer's recommendations. LB agar was autoclaved as above and allowed to cool before adding ampicillin or kanamycin (to a final concentration of 100 µg/mL) and pouring plates under flame. Plates were allowed to cool and either used instantly or stored at 4 °C for a maximum of two weeks (ampicillin) or four weeks (kanamycin).

2.6.3. Heat shock transformation

For plasmid propagation, DH5α heat shock competent cells were used, or following In-Fusion cloning, Stellar chemically competent cells (TaKaRa) were used. Bacteria were defrosted from

-80 °C on ice for 30 min, before adding 5 ng plasmid DNA per 50 µL bacterial cells and left on ice for a further 30 min in polypropylene tubes (Fisher Scientific #10384641). Using a water bath, tubes were heated to 42 °C for 45 sec before returning to ice for a further 2 min. Sterile S.O.C media (450 µL, Thermo Fisher #15544034) was added, and cells incubated at 37 °C on an orbital shaker for 1 hr. Under flame, bacteria were plated, with either ampicillin (100 µL) or kanamycin and (400 µL) agar plates and grown at 37 °C for 18 hr.

2.7. DNA based and cloning techniques

2.7.1. Plasmid visualisation and cloning design

SnapGene Viewer was used to visualise plasmids maps/sequences and aid in the design of sequencing/cloning primers (SnapGene software (from GSL Biotech; available at snapgene.com)).

2.7.2. Primer design and generation

The sequences of all primers used are available in Table 2.3 to Table 2.4. All primers were synthesised by Sigma-Aldrich as standard oligonucleotides in a desalted format. Primers were resuspended to 100 µM stock (information supplied by manufacturer) in sterile MilliQ water and stored at -20 °C. Primers were diluted to 5 µM stocks and vigorously vortexed before use.

Table 2.3 In-Fusion cloning primers.
Includes a description of how the primer was used, 5' -> 3' sequence and predicted melting temperature (Tm) for the gene specific portion of primer. For sequence, uppercase letters are vector specific, and bases inserted to correct reading frame, lower case letters are gene specific. For SDM, letters in red highlight mutated region.

Name	Sequence (5' ⟶ 3')	Tm (°C)
HA-mCherry into pcDNA3 (forward)	TATAGGGAGACCCAAGCTTACCATGGCCTACCCCTACGACGTGCCCGACTACGCCatggtgagcaagggcg	62.4
HA-mCherry into pcDNA3 (reverse)	CATGCTCGAGCGGCCttacttgtacagctcgtcc	59.6
MROH6 into HA-mCherry (forward)	CGGCATGGACGAGCTGTACAAGatggccgggggag	59.4
MROH6 into HA-mCherry (reverse)	CGGCCttaCTTGTACttatgcgcgtcttgg	56.5
zDHHC23 into HA-mCherry (forward)	CGGCATGGACGAGCTGTACAAGatgacgcagaaggg	54.4
zDHHC23 into HA-mCherry (reverse)	CGGCCttaCTTGTACttatcccagcaatcc	51.8
mCherry-HA into pcDNA3 (forward)	TATAGGGAGACCCAAGCTTACCatggtgagcaaggg	54.3
mCherry-HA into pcDNA3 (reverse)	CATGCTCGAGCGGCCTTAGGCGTAGTCGGGCACGTCGTAGGGGTAggccatcttgtacagctcgtcc	67.9
MROH6 into mCherry-HA (forward)	CTATAGGGAGACCCAAGCTTACCatggccgggggagtttgg	67.7
MROH6 into mCherry-HA (reverse)	TCACCATGGTAAGCTTtgcgcgtcttggtcc	62.1
zDHHC23 into mCherry-HA (forward)	CTATAGGGAGACCCAAGCTTACatgacgcagaaggg	54.4
zDHHC23 into mCherry-HA (reverse)	TCACCATGGTAAGCTTtcccagcaatccagtatc	59.1
SDM zDHHS23 (forward)	tcatcattccgtctggattaacag	58.4
SDM zDHHS23 (reverse)	tccagacggaatgatgatccatcc	59.4

Table 2.4 Quantitative Real Time PCR (qRT-PRC) primers.
Includes the gene name that primers were against and the sequence of the primer in the direction 5' -> 3'.

Gene	Direction	Sequence (5'⟶ 3')
CYCLO A	Forward	GCTTTGGGTCCGGAATGG
	Reverse	GTTGTCCACAGTCAGCAAT
MROH6	Forward	CTTCTTCACAGGGCTGTTGC
	Reverse	CACCTTCCTGCGATTCAGC
zDHHC23	Forward	TGTGTCTGGATAAATAGCTGCGT
	Reverse	ACACCGAGGTGAGCAAGAAG

2.7.3. Polymerase Chain Reaction (PCR)

CloneAmp HiFi PCR Premix (TaKaRa) was used following the manufacturer's guidelines. Briefly, for a 50 µL reaction mix: 25 µL HiFi PCR Premix, 3 µL of each diluted primer stock (5 µM), 20 ng plasmid were combined and made up to the final volume using milliQ H_2O. The KOD polymerase kit (Merck Millipore) PCR reactions were made using the supplied reagents and as described by manufacturer. For a 50 µL reaction mix: 5 µL 10X KOD polymerase buffer, 5 µL mixed dNTPs (2 mM each dNTP), 3 µL $MgSO_4$ (25 mM), 3 µL of each diluted primer stock (5 µM), 20 ng plasmid were combined and made up to the final volume in milliQ H_2O. For new primers, regardless of PCR reagent, a 50 µL reaction was split equally into 10 tubes to

test a gradient of temperatures at the extension phase (Stage 2 in Table 2.5). Once the optimal temperature was determined, a larger scale PCR reaction was set-up for in-gel DNA extraction for In-Fusion cloning. PCR reaction conditions are detailed in Table 2.5.

Table 2.5 PCR reaction conditions.
A gradient of 50-70 °C is used for new primers to determine optimal temperature. X = 1 min/kBp of amplified gene.

PCR step	# cycles	Temp (°c)	Time (min)
Stage 1	1	95	2
Stage 2	34	95	0.5
		50-70	0.5
		72	X
Stage 3	1	72	10
Stage 4	1	4	∞

2.7.4. Agarose gel electrophoresis

A 1% (w/v) high melting temperature agarose gel (Thermo Fisher) was made by dissolving agarose powder in TAE buffer (40 mM Tris-acetate pH ~8.5, 1 mM EDTA) and boiled until bubbling (~2 min/100mL) in a microwave. Molten agar was left to partially cool before a final concentration of 5% (v/v) Midori Green advanced DNA stain (Nippon Genetics) was added, mixed, and poured into the gel casting tray. DNA samples (30 µL) were mixed with 6 X DNA loading dye (6 µL) (New England BioLabs, B7025S). Samples ran alongside 10 µL of GeneRuler 1 kb DNA Ladder (Thermo Fisher). Samples were run at 100 V for 30 min. For extraction, gels were visualised using a UV box.

2.7.5. In-gel DNA extraction

DNA fragments were excised from 0.5-1.5% agarose gels using a clean scalpel and purified using the E.Z.N.A in gel extraction kit (Omega bio-tek) following manufacturer's protocol. Briefly, the extracted agarose slice was melted at 60 °C for 10 min, bound to a DNA binding membrane and washed in low-salt wash buffer (PDR044) before elution in 10 mM Tris-HCl pH 8.5. The elution volume used was dependent on downstream purposes: 20 µL (maximum concentration) was used for In-Fusion Cloning, 50 µL (maximum recovery) for mutagenic PCR and ligation.

2.7.6. In-Fusion HD Cloning

In-Fusion Cloning (Clontech) technology was used to create all gene of interest (GOI plasmids: HA-mCherry-GOI and GOI-mCherry-HA based plasmids (zDHHC23 and MROH6) following manufacturer's recommended protocols. The strategy developed for cloning is shown diagrammatically in Chapter 4. Briefly, the destination vector was linearized with restriction enzymes. PCR was performed on the GOI using In-Fusion cloning compatible primers which have a 5' 15 bp overhang complementary to the destination plasmid (described in Table 2.3). Primers had bases inserted as necessary to correct the reading frame and recover restriction digest sites. Both linearized plasmid and amplified inserts were purified using in gel extraction (Section 2.7.5). In-Fusion reaction mixes were performed in a total volume of 20 µL with a 3:1 molar ratio of insert: vector using the following equation:

$$\text{Insert mass (g)} = \frac{\text{ratio of insert:vector} \times \text{mass of vector (g)} \times \text{insert length (bp)}}{\text{vector length (bp)}}$$

For transformation, 5 µl of the infusion reaction mix was used with 50 µl Stellar competent cells (Clontech).

2.7.7. Site Directed Mutagenesis (SDM)

SDM was used to create point mutations within the zDHHC23 gene (Cys to Ser) using the PCR protocol (Section 2.7.3) and In-fusion cloning protocols (Section 2.7.6) described above with overlapping mutagenic primers (Table 2.3).

2.7.8. Plasmid amplification

For each plasmid, five colonies were picked from ampicillin plates and added into separate 50 mL falcon tubes containing 5 mL LB broth, maintaining antibiotic selection, and grown at 37 °C on an orbital shaker for 8 hr. The cell suspension was used to create glycerol stocks (1 ml + 1 ml 80% glycerol) and for DNA extraction (4 ml) using the GeneJET plasmid miniprep kit (Thermo Fisher). Briefly, cells were pelleted, lysed, neutralised to precipitate protein, and cleared by centrifugation (13,000 *g* for 5 min) before applying to a DNA binding membrane for washing and elution using centrifugation (13,000 *g* for 1 min). Plasmids were sequenced

and checked for errors (Section 2.7.9). For a plasmid containing the correct gene, 200 µL of respective glycerol stock bacteria was added into 200 mL LB broth, with antibiotic selection, and grown for 18 hr at 37 °C in an orbital shaker. Plasmid was then extracted from the bacteria using the PureLink HiPure Plasmid Maxiprep Kit (Invitrogen), following the manufacturer's protocol. The final plasmid maps are depicted in Figure 2.1-Figure 2.8.

2.7.9. Plasmidsaurus DNA sequencing

All sequencing was performed with the company Plasmidsaurus who sequenced the entire construct without the need for primers.

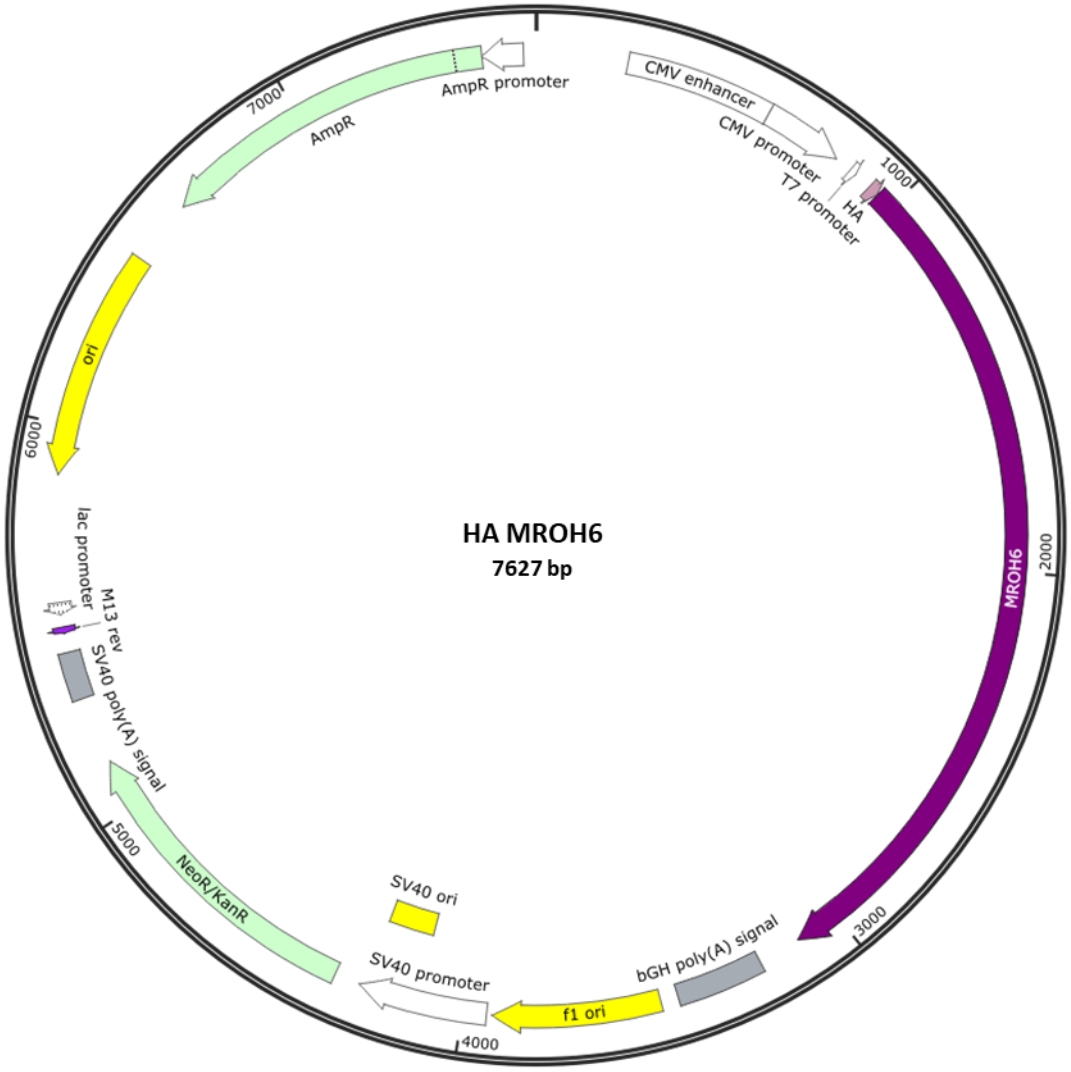

Figure 2.1 Plasmid map of HA-MROH6.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) MROH6 gene = purple, HA tag = pink.

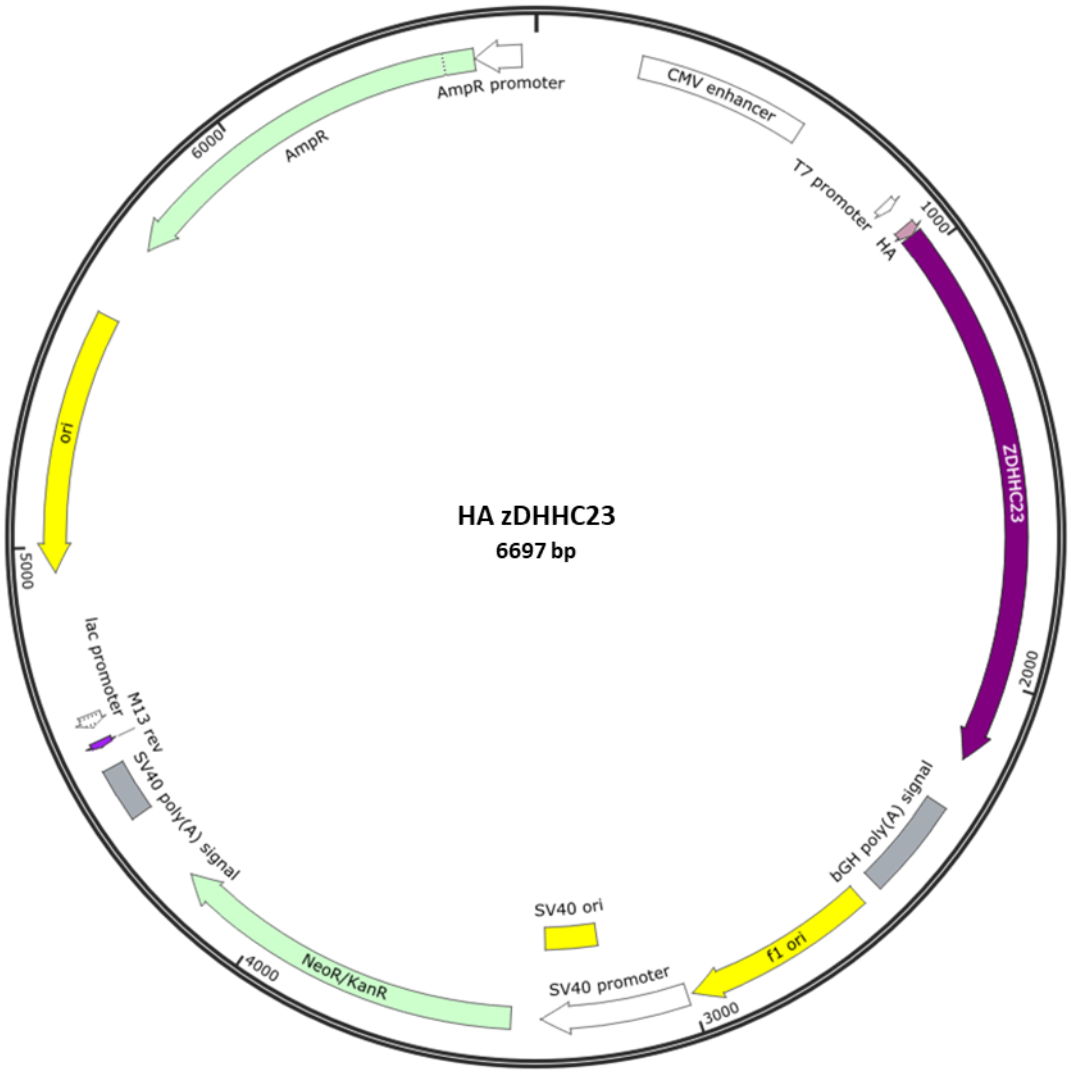

Figure 2.2 Plasmid map of HA-zDHHC23.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) zDHHC23 gene = purple, HA tag = pink.

Figure 2.3 Plasmid map of cloned HA-mCherry MROH6.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) MROH6 gene = purple, HA tag = pink, mCherry tag = red.

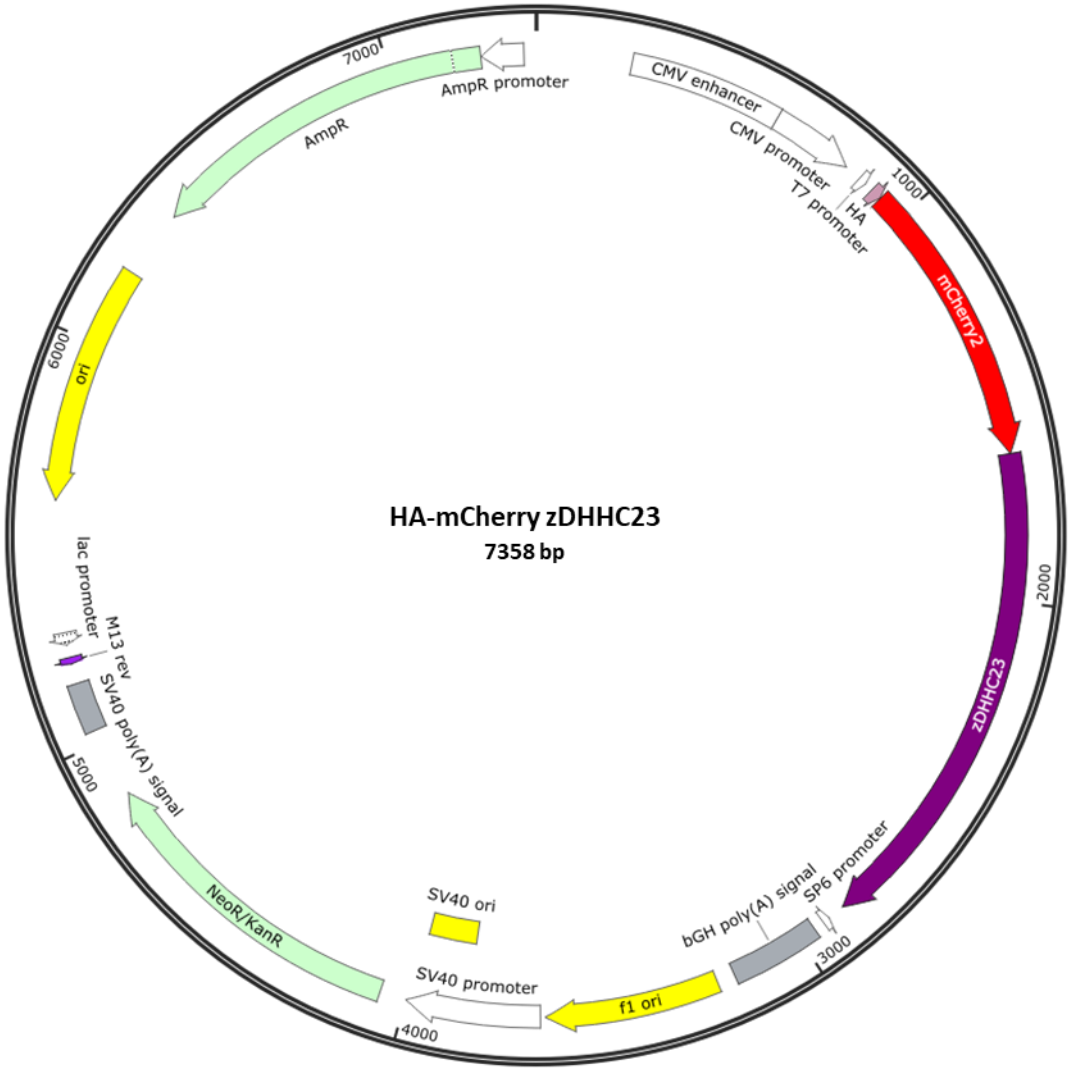

Figure 2.4 Plasmid map of cloned HA-mCherry zDHHC23.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) zDHHC23 gene = purple, HA tag = pink, mCherry tag = red.

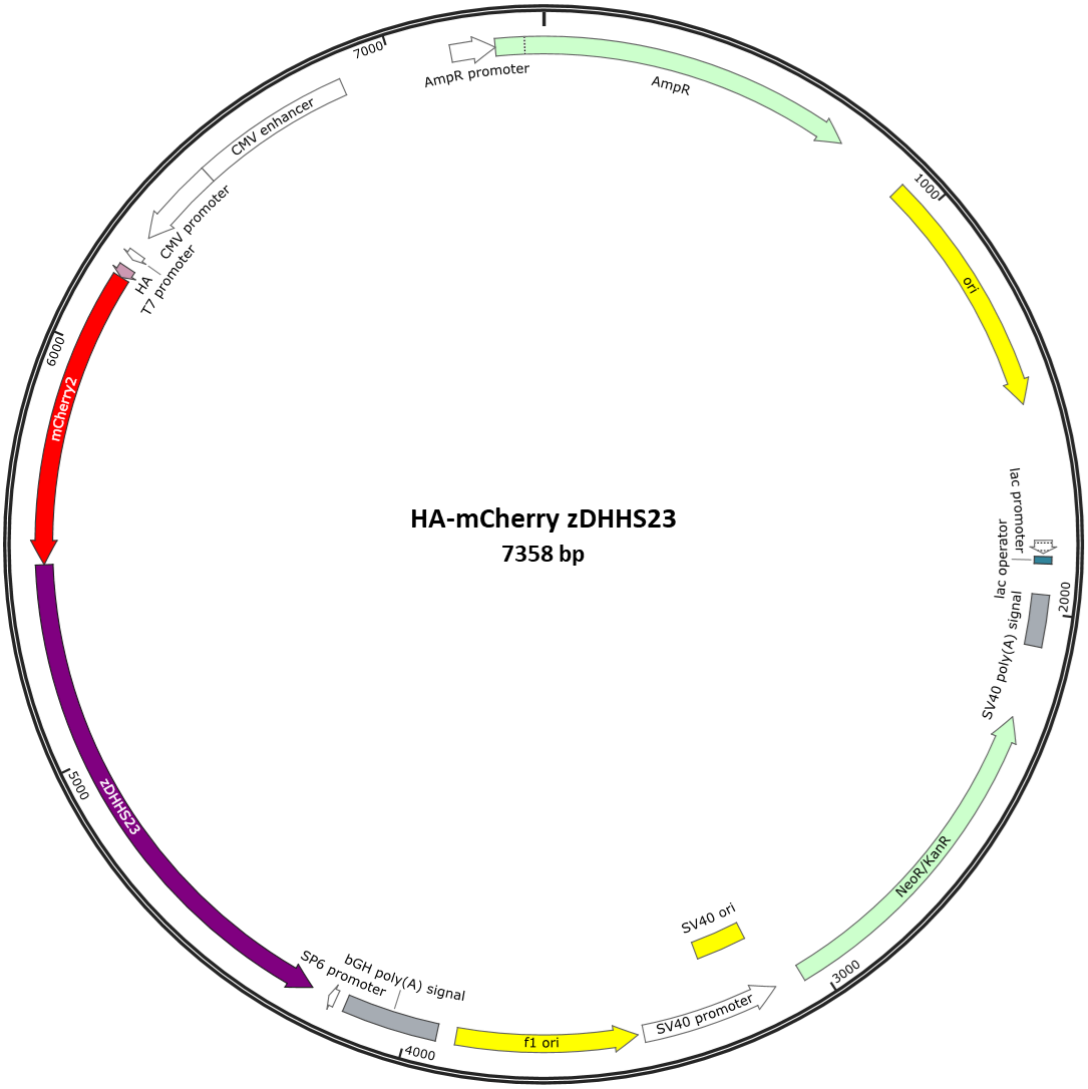

Figure 2.5 Plasmid map of cloned HA-mCherry zDHHS23.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) zDHHS23 gene = purple, HA tag = pink, mCherry tag = red.

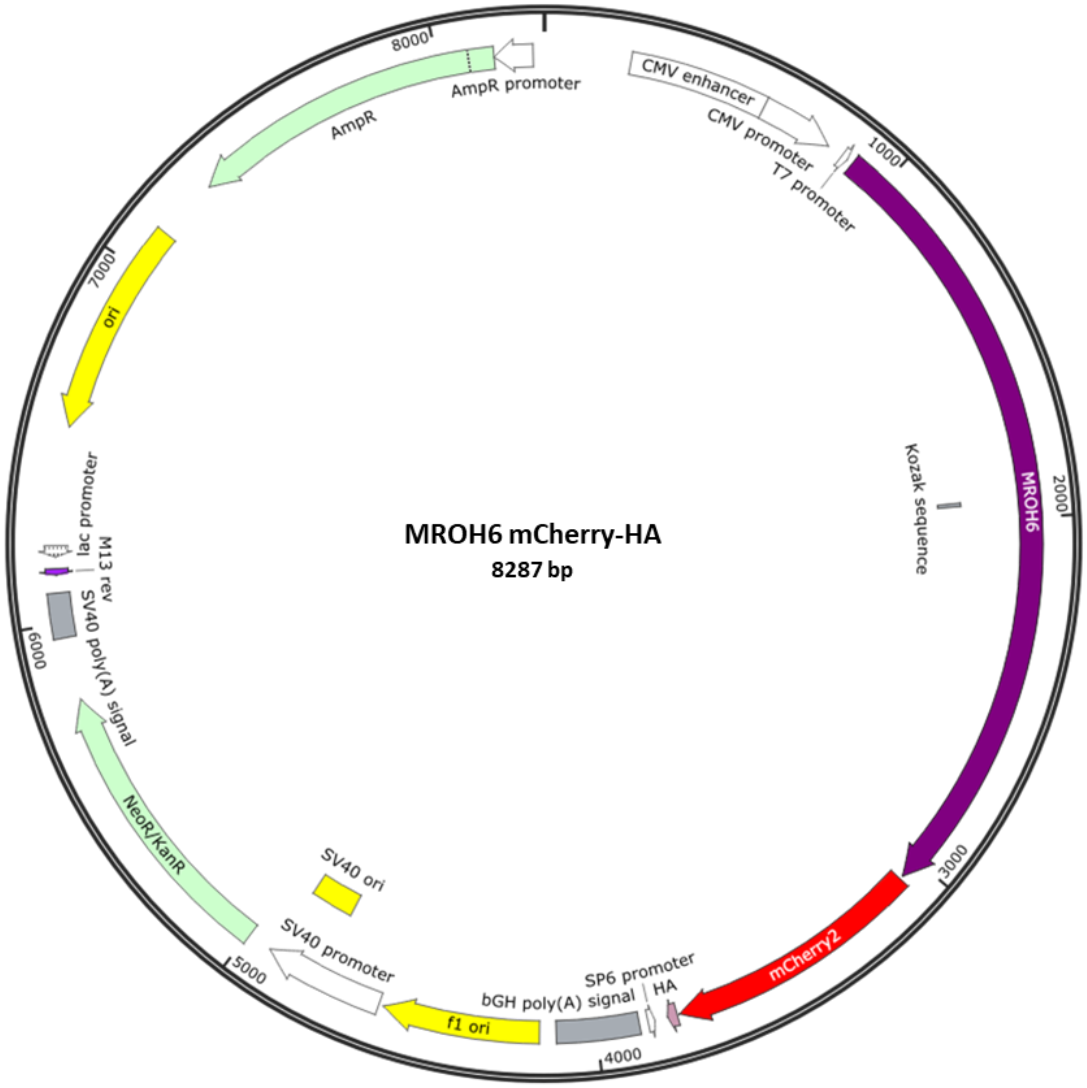

Figure 2.6 Plasmid map of cloned MROH6 mCherry-HA.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) MROH6 gene = purple, HA tag = pink, mCherry tag = red.

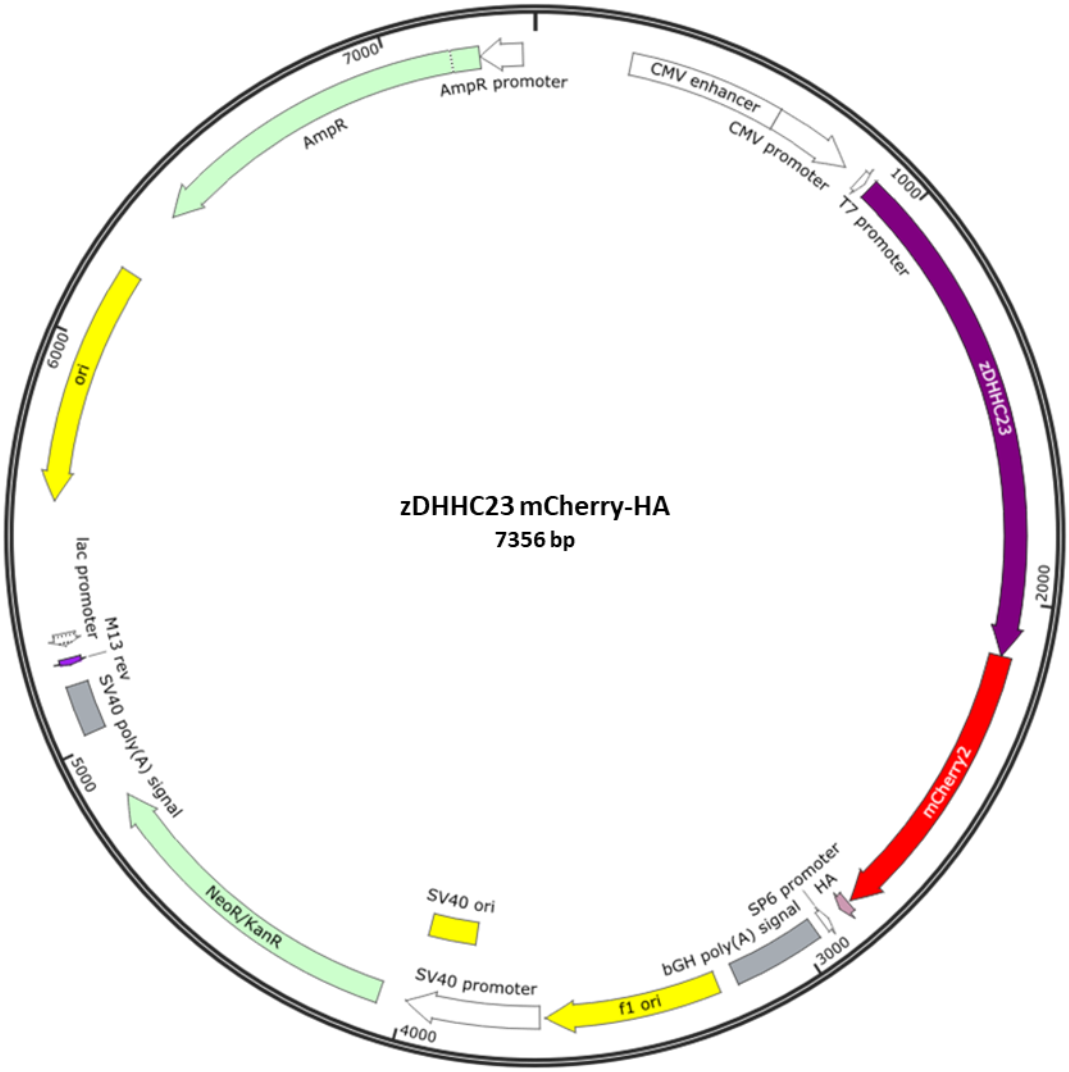

Figure 2.7 Plasmid map of cloned zDHHC23 mCherry-HA.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) zDHHC23 gene = purple, HA tag = pink, mCherry tag = red.

Figure 2.8 Plasmid map of cloned zDHHS23 mCherry-HA.
Viewed and annotated in Snapgene viewer (SnapGene software (from GSL Biotech; available at snapgene.com)) zDHHS23 gene = purple, HA tag = pink, mCherry tag = red.

2.8. Chick embryo model

Guidelines set by the Home Office of the United Kingdom were adhered to in chick embryo experiments, with all experiments terminated before embryonic day 14 (E14, two thirds of the gestation period) (Figure 2.9). Fertilised Bovans Brown eggs were stored in a humidified refrigerator (12-14 °C) for up to two weeks until use. During experiments, eggs were incubated (OvaEasy 190 advance Series II, Brinsea) at 40% at 37.8 °C, rolling every 20 min to mimic normal physiological settings.

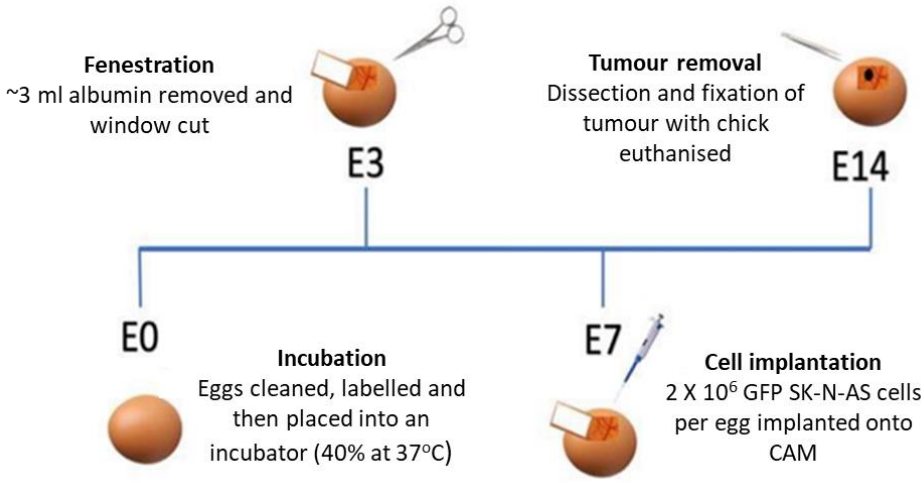

Figure 2.9 Schematic showing the chick embryonic timeline.
Initial incubation of the fertilised egg (E0), fenestration (E3), cell implantation (E7) to the collection of subsequent tumours (E14).

2.8.1. Egg Fenestration and cell implantation

After 72 h, on embryonic day 3 (E3) eggs were prepared for future cell implantation. Using a needle (20 G x 1, Terumo Neolus) and syringe (10 ml Terumo) ~3 ml albumin was removed from the wide base of the egg and sealed (Nev's Ink, Inc. Brand Labelling Tape) to prevent leakage. A 1x2 cm rectangular window was made using dissection scissors, with three sides separated from the shell, this window was kept in place using Scotch Magic adhesive tape. The eggs were then returned to the incubator until E7. For implantation of cultured cells onto the chorioallantoic membrane (discussed in Section 1.5), media in T175 flask was aspirated, and the cells washed with equal volume of PBS. PBS was then aspirated and 1 ml of 0.05% Trypsin-EDTA solution was added and incubated for 1 min at 37 °C. The trypsin solution

containing detached cells was neutralised by addition of 10 ml of media in a falcon tube. Cell counts were determined by pipetting 10 µl of cell suspension into a cell counter slide and counting using a Bio-Rad TC20 Automated Cell Counter. Total cell count was determined (multiplying cell count (cells/ml) by total volume) before cells were centrifuged (5 min, 100 *g*) to form a cell pellet. The media was aspirated, and the pellet transferred into an Eppendorf tube and placed on ice. The number of cells per egg to be implanted was 2×10^6 with the subsequent necessary volume calculated prior to implantation, the egg window would then be removed. To dry the CAM, tweezers (T5389, Dumont) were used to place sterile lens tissue paper (Fisher Scientific) directly onto the CAM and slowly removed followed by pipetting of cells. The window was then re-sealed as before, and the egg returned to the incubator for a further seven days until E14. Any dead or unfertilised eggs were discarded.

2.8.2. Tumour extraction, fixation and sectioning

On E14, egg windows were removed and the surrounding shell around the fenestrated area removed using forceps, in order to increase the visual field for microscopy and dissection. A small dissection scissor and tweezer (Springbow dissection scissors T5372, Dumont tweezers T5392) were used to dissect the tumours from the membranes and placed in a petri dish with a drop of PBS to remove any blood. Embryos were then quickly euthanised via decapitation. Each tumour was then placed in 4% paraformaldehyde (PFA) and stored at 4 °C for a maximum of 24 h. The PFA was aspirated, and the tumour suspended in sequentially increasing sucrose concentration for 24 h each (6%, 12%, 20%). The tumour was then placed in a plastic embedding mould and filled with embedding solution, Cryomatrix (Thermo Scientific). The mould was then placed on dry ice to freeze, with long term storage at -80 °C. A Leica CM1950 cryostat was used to slice the frozen tumours and was set at a temperature of -20 °C and slices set at 10 µm thickness. The slices of tumour were collected on Super-frost slides (Thermo Fisher) and stored at -20 °C.

2.9. Immunofluorescence

2.9.1. Immunohistochemistry (IHC)

Slides were thawed at room temperature for 10 min. A hydrophobic pen (SUPERPAP, Invitrogen) was used to circumscribe around each section of tissue. Each sample was then incubated with 100 µl of 50 mM Ammonium Chloride (NH4Cl) for 20 min at room temperature. Following the removal of NH4Cl, PBS was applied to rehydrate the tissue, with a 10 min incubation at room temperature. This was then removed and immediately followed by a 30 min incubation in blocking buffer (1% (w/v) BSA, 0.1% (v/v) Triton X-100, 0.4% (v/v) Tween 20 in PBS). Following this first blocking step, 100 µl of the required primary antibody dilution (Table 2.6) dissolved in blocking buffer was added per sample and incubated overnight at 4 °C in a humidified chamber.

Following primary antibody incubation, the sections were washed three times for 10 min each in blocking buffer with gentle agitation. Secondary antibodies were then added (100 µl per sample) in solution with blocking solution at dilutions indicated in Table 2.6. Sections were incubated for 1 h at room temperature, washed twice (10 min each) with PBS were and finally with deionised water. The slides were air-dried and then mounted with anti-fade mounting medium (DAKO, Agilent) with coverslips and stored at 4 °C overnight.

2.9.2. In-vitro immunocytochemistry (ICC)

Cells were seeded onto autoclaved glass coverslips in 6 well plates, 24 hr prior to fixation. Following a PBS wash, cells were fixed with 1 ml 4% (w/v) paraformaldehyde at room temperature for 10 min. After fixation, cells were again washed three times with PBS for 10 min. The cells were then blocked for 20 min at room temperature in blocking buffer (as before). The cells were incubated in primary antibody for 1 h, then washed in 3 x 10 min in blocking buffer before secondary fluorescent antibody (Table 2.6) incubation for a further 1 h at room temperature. During secondary antibody incubation coverslips were protected from light using foil. Cells were then washed and if required, a Hoechst nuclear stain added for 3 min before a final wash with PBS and mounted as previously described.

Table 2.6 Antibodies/stain used in immunofluorescence.

Antibody/Stain	Company	Catalogue number	Dilution factor (IHC)	Dilution factor (ICC)
Hoechst	Thermo Fisher	62249	N/A	1:1000
MROH6	Thermo Fisher	PA5-64355	1:200	1:200
zDHHC23	Novusbio	NBP1-82688	1:50	1:20
Alexa Fluor 555	Thermo Fisher	A-21428	1:500	1:500

2.10. Sample preparation for mass spectrometry (MS)

HPLC grade solvents and highest purity powdered chemical reagents were used throughout.

2.10.1. Reduction and alkylation

A final concentration of 3.33 mM DTT (Dithiothreitol) was added to each sample (from a stock of ~72 mM DTT dissolved in 25 mM ammonium bicarbonate (AmBic)) and heated at 60 °C for 10 min. The sample was cooled to room temperature and final concentration of 9.5 Mm of IAA (Iodoacetamide) (from a stock of ~63 mM IAA dissolved in 25 mM AmBic) and incubated at room temperature in the dark for 30 min.

2.10.2. Proteolytic digestion (SP3)

Reduced and alkylated samples were placed in low bind tubes for digestion via single-pot solid-phase-enhanced sample preparation (SP3) beads (Sera-Mag SpeedBeads A and B, GE Healthcare 45152105050250 and 65152105050250). Per IP, 10 µL of 20 ng/µL SP3 beads were added to each sample with the addition of 100% acetonitrile (ACN) to make the final volume 70% ACN. Samples were incubated at 25 °C for 30 min with 1400 rpm shaking on a Thermomixer (Eppendorf). Supernatant was removed and beads washed twice in 200 µL 70% HPLC grade ethanol and the same volume of 100% ACN. The beads were then subject to vacuum centrifugation (Centrifuge: UNIVAPO – 150 ECH, Cooling unit: UNICRYO MC2L -60 °C, Vacuum pump: UNIVAC DQ4) for 10 min. After drying, 200 µL of 100 mM AmBic was added

and samples sonicated for two mins (Cavitek 27 L professional ultrasonic cleaner, Allendale-ultrasonics). Samples were then digested by addition of 0.5 µg of Trypsin Gold (Promega) and incubated at 37 °C for 18 hr with 1400 rpm shaking. Post digestion, trifluoroacetic acid (TFA) and ACN was added to a final concentration of 1.5% (v/v) and 3% (v/v), respectively, and incubated at 37 °C with 600 rpm shaking for 1 hr. The sample was centrifuged at 13000 *g* for 10 min at 4 °C and the clear supernatant collected. Samples were split 5%:95% and dried to completion. For mass spectrometry analysis of binding partners, (5% samples (20 µL)) were solubilised in loading buffer (3% ACN, 0.1% TFA in H_2O), by sonication and centrifuged as before and clear supernatant collected into a total max recovery vial (WATERS) for injection. The 95% samples were used for Titanium dioxide (TiO_2) enrichment and PTM analysis.

2.10.3. Titanium dioxide (TiO_2) phospho-peptide enrichment

Dried peptides were dissolved in 100 µL of TiO_2 loading buffer (80% (v/v) ACN, 5% (v/v) TFA, 1 M glycolic acid in H_2O) and sonicated for 10 min before addition of 1 mg TiO_2 resin (GL Sciences) and mixed at 1750 rpm on a thermomixer (Eppendorf) at room temperature for 20 min before centrifugation at 2000 *g* for 1 min and supernatant discarded. Sequential wash steps equivalent to 1.2:1 (w/v) TiO_2 resin to wash buffer were performed subsequently with loading buffer, wash buffer 1 (80% (v/v) ACN, 1% (v/v) TFA in H_2O) and wash buffer 2 (10% (v/v) ACN, 0.2% (v/v) TFA in H_2O) with 1750 rpm shaking for 10 min each. All centrifugation steps were at 2000 *g* for 1 min with supernatant removed between washes. Peptides were eluted by sequential addition of 5% (v/v) ammonium hydroxide equivalent to 1.5:1 (w/v) TiO_2 resin to elution buffer and dried to completion by vacuum centrifugation. Dried peptides were resuspended as before, and samples were centrifuged at 13000 *g* for 15 min at 4 °C and 18 µL taken for LC-MS/MS analysis.

2.10.4. Strong Cation Exchange Stage Tip (SCX)

Strong cation exchange (SCX) was used to remove highly hydrophobic contaminants from peptide samples that had undergone CuAAC (Section 2.4.2). The dried peptides were dissolved in 200 µL of 1.5% (v/v) TFA in H_2O. SCX stage tips were prepared by packing 3 discs of SCX membrane (Empore™ Supelco 47mm Cation Exchange disc #2251) into a 200 µL pipette tip. Tips were equilibrated by sequential washing of 200 µL of each: acetone, methanol, H_2O, 5% (v/v) ammonium hydroxide (in H_2O) and H_2O. Each centrifugation step

was performed at 4000 *g* for 4 min, or until all liquid had passed through the stage tip. All samples were passed through the equilibrated tip twice and washed with H_2O before elution in 200 µL of 5% (v/v) ammonium hydroxide (in H_2O) followed by drying to completion by vacuum centrifugation (Section 2.10.2).

2.10.5. Heat Induced Antigen Retrieval (HIAR)

For optimisation, frozen SK-N-AS tumours (fixed in 4% PFA) were defrosted and submerged in H_2O. Tumours where then placed into 20mM Tris pH9, 2% SDS, 10mM DTT and 1mM EDTA. Using a bead-based tissue homogeniser (Minilys, P000673-MLYS0-A) the solution containing the tumour was placed into 2 ml tubes, pre-filled with stainless steel beads, 2.8 mm (Precellys, #432-3753) and shaken for 1.5 min at 3000 rpm. Following bead homogenisation, samples were moved into a low bind tube (without stainless steel beads) and placed in a heat block at 100 °C for 20 min followed by 60 °C for 2 hr before SP3 digestion (Section 2.10.2).

2.11. Thermo Orbitrap Fusion Tribrid Mass spectrometer

2.11.1. Liquid chromatography

Peptides were separated by reversed-phase capillary HPLC on an UltiMate 3000 nano system (Dionex) coupled in-line to a Fusion Lumos Tribrid mass spectrometer (Thermo Scientific, Bremen, Germany). Peptides were loaded onto the trapping column (PepMap 100, C18, 300um X 5mm) with 2% (v/v) ACN, 0.1% (v/v) TFA using partial loop injection at a flow rate of 9 µL/min for seven minutes then resolved on an analytical column (Easy-Spray C18 75um x 500mm, 2um bead diameter column) at a flow rate of 0.3 ul/min using a 60 minute gradient for both TiO_2 enriched and unenriched samples, from 96.2% A (0.1% (v/v) Formic acid): 3.8% B (80% (v/v) ACN with 0.1% (v/v) Formic acid) to 50% B.

2.11.2. High-low method for binding partner analysis

For unenriched samples, MS1 spectra were acquired in the Orbitrap (120K resolution at a *m/z* 200) over a *m/z* range of 350-2000, AGC target = 50%, maximum injection time = 50 ms. MS2 data were acquired in a data-dependent acquisition (DDA) mode, using a 'top speed' method with a cycle time of 3 sec. HCD fragmentation was set to 32% normalised collision energy (NCE) in the Ion trap. Ion trap MS2 settings were as follows: rapid mode (15K resolution at *m/z* 200), maximum injection time = dynamic, fragmentation intensity threshold = $1E^4$ for 2+ to 5+ charge states. A dynamic exclusion window of 60 s was applied at a 10 ppm mass tolerance.

2.11.3. High-high method for analysis of TiO_2 enriched samples

MS1 spectra were acquired in the Orbitrap (120K resolution at a *m/z* 200) over a *m/z* range of 350-2000, AGC target = 50%, maximum injection time = 50 ms. MS2 data was acquired in a DDA mode, using a top speed method with a cycle time of 3 s. HCD fragmentation was set to 32% NCE in the orbitrap. Orbitrap MS2 settings were: 30K resolution at *m/z* 200, maximum injection time = dynamic, fragmentation intensity threshold = $2.5E^4$ for 2+ to 5+ charge states. A dynamic exclusion window of 20 s was applied at a 10 ppm mass tolerance.

2.12. Thermo Q Exactive HF for PRM optimisation

A HA-tag IP was performed of both orientations of HA-mCherry constructs for zDHHC23 and MROH6 and analysed on a high-high method. Identified peptides were analysed via MASCOT before raw files and identities were imported into Skyline (Section 2.13.1). All peptides identified for either zDHHC23 or MROH6 were used to create a targeted mass list with retention times. The targeted mass list was imported into the Thermo Q Exactive HF for PRM mode analysis using the following settings: Full MS-resolution of 60K, AGC target $3E^6$, maximum injection time = 100 ms, scan range = 350 – 2000 *m/z*; MS2: resolution = 60K, AGC target = 1E5, maximum injection time 300 ms, isolation window = 1.0 *m/z*, with a RT window of +/- 1.5 min of stated time. PRM methods were validated to work on repeat IP injections and overexpression lysate from either construct, over multiple replicates/days.

2.13. Mass spectrometry Data analysis

2.13.1. Skyline

Skyline was accessed at: http://proteome.gs.washington.edu/software/skyline. An identification file generated by IP (fragment.csv) for the two targeted proteins zDHHC23/MROH6 was first imported into Skyline (MacLean et al., 2010) v21.2 to create a spectral library. The precursor *m/z* tolerance and retention time window was set at 2 *m/z* and 2 min, respectively. Skyline then performed peptide peak extraction, integration, and background subtraction before it exported a custom report that contained retention time and peak areas of different peptides. After data import, graphical displays of chromatographic traces (extracted ion chromatograms) were manually inspected for proper peak picking of MS1 filtered peptides. Using the listed *m/z* and RT, this was used as a PRM protocol for non-IP or enriched samples to identify if any peptides could be detected using a targeted approach.

2.13.2. Proteome Discoverer (PD)

Proteome Discoverer (Orsburn, 2021) v2.4 was used to process LC-MS/MS data from the Orbitrap Fusion Lumos Tribrid for binding partner identification and phospho-peptide identification and site localisation. Binding partner data was searched using the MASCOT search engine against the UniProt Human reviewed database (updated weekly); instrument type = ESI-FTICR with fixed modifications = cysteine carbamidomethylation and variable modifications = methionine oxidation, two miscleaves allowed and mass tolerances were MS1 = 10 ppm and MS2 = 0.5 Da. All files were searched separately to create a filtered list that was seen in two out of three replicates, before background subtraction using relevant controls (HA-mCherry/mCherry-HA). A second search combining all files was performed with the minora feature detector node set to calculate peptide abundance (area under the curve) and low abundance resampling imputation of missing values in order to allow label free quantification (LFQ) and statistical analysis. The background subtracted, seen two out of three replicate list was then used to filter relevant protein identifications in the LFQ and imputed dataset. This data was then inputted in R using a custom script (Appendix 1) to create volcano plots.

For TiO_2 enriched samples, data was searched as before with the following changes: variable modifications additionally included phosphorylation of S/T/Y and mass tolerances were MS1 = 10 ppm and MS2 = 0.01 Da. The ptmRS node in PD was used to localise the site of phosphorylation on peptides. All identified peptides were filtered to a 1% False Discovery Rate (FDR) and a ptmRS >0.99 on peptides at the Peptide Spectral Match (PSM) level (Ferries et al., 2017).

2.14. Bioinformatics analysis

2.14.1. STRING

STRING v11.5, accessed at https://string-db.org was used for network analysis, with proteins entered as a list using their UniProt accessions and searched against the *Homo sapiens* database. In all cases, data was filtered so that only interactions identified in experimental, or database-level evidence were included at a high (>0.7) confidence, and any unconnected nodes were removed for viewing. KMEANS clustering was used to identify clusters of proteins with shared functions, the number of clusters set was determined empirically for each network. Networks generated in STRING were exported as a tab-separated values (.tsv) file and imported into Cytoscape (Shannon et al., 2003) for editing and PowerPoint to customise the overall network presentation.

2.14.2. DAVID

Gene ontology functional annotation was performed using DAVID (Database for Annotation, Visualisation and Integrated Discovery, v6.8 (Dennis et al., 2003). Data was filtered so that GOterms: Molecular Function (MF), Cellular Compartment (CC) and Biological Process (BP) annotations were maintained. P-values were adjusted using the Benjamini-Hochberg method (stringent P-value correction that multiplies the initial P-value by the total number of annotations divided by the rank of a given annotation). Data visualisation was performed using a custom R script (provided by Dr Amy Campbell, Appendix 2) and labelling amended in InkScape v1.2.

2.15. Biochemical and *in*-vitro assays

2.15.1. Quantitative Real Time PCR (RT-qPCR)

Per sample, $1x10^6$ cells (in a 10 cm plate) were plated and incubated for 72 h at 21% or 1% O_2. RNA was extracted using the HiPure RNA isolation kit (Roche) following the manufacturer's recommended protocol. RNA concentration and purity were determined using a Nanodrop 2000 and 1 µg of RNA converted into cDNA using the SuperScript Vilo reverse 80 Transcription master mix (Invitrogen). cDNA samples were then incubated at 25 ^oC for 10 min followed by 1 h at 42 ^oC and terminated at 85 C for 5 min. cDNA was diluted 20-fold in RNAse free water before use. RT-PCR was performed in a LightCycler480. White 96-well plates (Roche) were used with the standard well component mix of 10 µl SYBR Green PCR Mix (Roche), 1 µl of 5 mM forward primer, 1 µl of 5 mM reverse primer (Table 2.4), 2 µl cDNA and 6 µl H_2O. Cycling parameters are described in Table 2.7. Subsequent data analysis was performed using the LightCycler 480 SW 1.8 software, values were normalised to the mean of the house-keeping gene cyclophilin A. Fold gene expression change was determined relative to normoxic controls and were calculated using the $2-(\Delta\Delta\mathrm{Ct})$ equation. Briefly, the difference between average cycle threshold (Ct values) for housekeeping gene and gene of interest being tested are determined (ΔCt value) and then the difference between the ΔCt value for housekeeping gene and gene of interest is calculated (ΔΔCt value). Finally, all calculations are in logarithm base 2 and therefore $2-(\Delta\Delta Ct)$ is used to determine fold change.

Table 2.7 General cycling parameters for RT-PCR

Step	Temperature (°C)	Time (s)	Number of Cycles
Pre-incubation	50 95	120 600	1
Amplification	95 60 72	10 15 15	55
Melt Curve	95 65 97	5 60 Continuous	1
Cooling	40	10	1

2.16. Microscopy and image analysis

2.16.1. IHC and ICC imaging

Slides prepared in Section 2.9 were imaged on a LSM710 confocal microscope (Zeiss) with Zen 2012 software, filters set to GFP (488 nm) and for red fluorescent protein (RFP) (561 nm). Image analysis was completed using machine learning-based pixel classification via LabKit in Fiji (Arzt et al., 2022). All values were recorded in GraphPad Prism 6, with T-test analysis between normoxic (21% O_2) and hypoxic (1% O_2) values.

2.16.2. Localisation imaging

SK-N-AS cells were grown and transfected as previously stated, in 12 well glass bottom dishes (Cellvis). Images were captured using a LSM780 Zeiss Microscope with Zen 2012 software, equipped with GFP (488 nm) RFP (561 nm) and Hoechst (405 nm) filters at a magnification of 63X at a temperature of 37 ^{o}C, CO_2 (5%) and 21% O_2.

Chapter 3: Investigation of zDHHC23 and MROH6 as potential prognostic biomarkers

3.1. Introduction

A cancer biomarker is a measurable and specific molecular level change indicative of a diseased state (Depledge et al., 1993), typically at the gene, transcript or protein expression levels. Technological advancements have resulted in large datasets and prolonged patient tracking, and as such biomarkers can be used for many clinical assessments. These include identifying pathological processes, development of the disease, clinical outcome, prognosis and therapeutic interventions, expanding the field into the age of personalised medicine (Robb et al., 2016). As such, biomarkers are classified into three categories; predictive (predicts a response to the chosen therapeutic interventions), diagnostic (identify whether a patient has a specific disease) and prognostic (used to inform physicians of the patient risk and clinical outcomes, such as recurrence or disease progression). The ultimate goal of biomarker development is to establish an easily accessible test that can be performed in the clinic and aid in decision making to improve patient outcomes (Parkinson et al., 2014).

As previously discussed, the focus of this thesis is the disease NB, a childhood cancer that has a lack of clinically relevant biomarkers; hence significant research is required to identify novel biomarkers with prognostic significance. The majority of NB patients (more than 50%) are diagnosed at a late/advanced stage of the disease where the cancer has extensively metastasised, primarily to the bone marrow, liver, skin and/or lymph nodes (DuBois et al., 1999). Diagnosis and treatment is further complicated by the heterogenous nature of NB, with clinical behaviour ranging from rapid tumour progression and death to spontaneous remission (Yáñez et al., 2015). Unfortunately, due to the scarcity in knowledge of how NB develops and progresses, there are few available therapeutic strategies for treatment of early stages of NB (Utnes et al., 2019). Novel biomarker identification could help facilitate disease stratification at an earlier stage, potentially improving treatment management and patient prognosis/outcome. An important factor to consider in biomarker utility is its application in a clinical setting i.e. cost, ease of assay/sample collection and technique required (Sanjay et al., 2015). The global “gold standard” of biomarker assays for cancer

diagnosis is immunohistochemistry (IHC, Figure 3.1). IHC is a method of immunostaining, reliant on the application of a specific antibody against a single protein to assess protein expression and subcellular localisation within tumour biopsy samples (for example, HER2 expression in breast cancer (Bartlett et al., 2001)). However, the success of IHC assays are fully dependent on two factors: reliable antibodies (Smith et al., 2010) and tissue sample integrity (Dunstan et al., 2011).

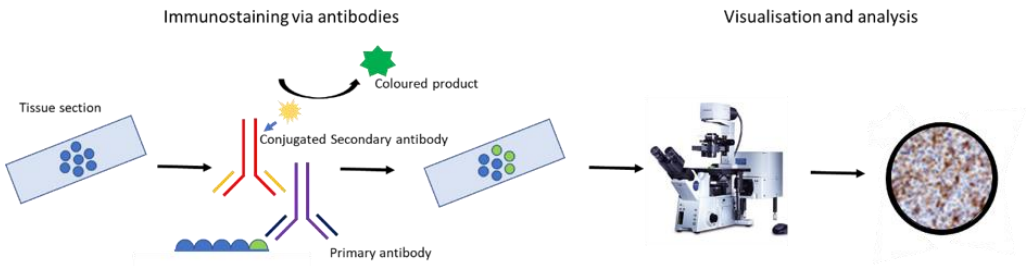

Figure 3.1 Schematic showing immunohistochemistry workflow.
Tissue is processed for immunohistochemistry and a primary antibody use to probe the desired target. A complimentary secondary antibody which is conjugated to a fluorescent fluorophore is then used to visualise the presence of the marker by fluorescent microscopy.

There are two types of antibodies commercially available: monoclonal or polyclonal antibodies, both with distinct advantages and disadvantages. Monoclonal antibodies are immortalised and derived from a single B cell clone, hence only recognise a single epitope (Nelson et al., 2000). This results in low background noise due to the antibody being less likely to cross react with other proteins. However, they are potentially confounded by antigens which have been modified (as the case with tissue fixation or PTM) or protein-protein interactions. On the other hand, polyclonal antibodies are generated from whole protein immunization of animals and as such multiple B cell clones that target multiple different epitopes on the same antigen (Stills, 2012). Hence, polyclonal antibodies can enhance signal to noise of detected proteins and are generally more resistant to epitope masking. However, targeting multiple epitopes does increase the likelihood of antibody off-target specificity (cross reactivity with non-target proteins), resulting in false positive data - a significant issue for clinical interpretation.

There have been significant technological advancements which circumvent antibody issues in biomarker detection assays, including real-time polymerase chain reaction (assessment at the gene expression/DNA mutations level) and fluorescent in situ hybridisation (assessment of DNA ploidy (copy number)). Of particular interest in the Eyers lab is a protein-centric

biomarker detection method that relies on the use of targeted and quantitative LC-MS/MS proteomic analysis, as is used routinely in clinical labs e.g. for the diagnosis of inherited metabolic disorders. MS has been invaluable in the growth of the proteomics research; hence LC-MS/MS has become a prominent technique in various multi-institute studies focusing on cancer biomarker discovery. The National Cancer Institute Clinical Proteomic Tumour Analysis Consortium has applied MS methods to successfully identify cancer-specific proteins and unique protein patterns which has subsequently led to the sub-classification of ovarian, breast and colorectal cancers (Rodriguez and Pennington, 2018). With advancements in MS technology and the development of different types of MS instrumentation, two predominant targeted proteomic methods have emerged as reliable and high throughput assays for cancer biomarker detection; Selected Reaction Monitoring (SRM) (on triple quadrupole instrumentation) and Parallel Reaction Monitoring (PRM) (on Orbitrap instrumentation) (Figure 3.2). Absolute quantification (determination of the number of moles/molecules of protein expressed) by SRM and PRM is achieved by comparing mass spectrometric signal intensity of target peptides with artificially spiked known amounts of heavy labelled variants of the same peptide following electrospray ionisation.

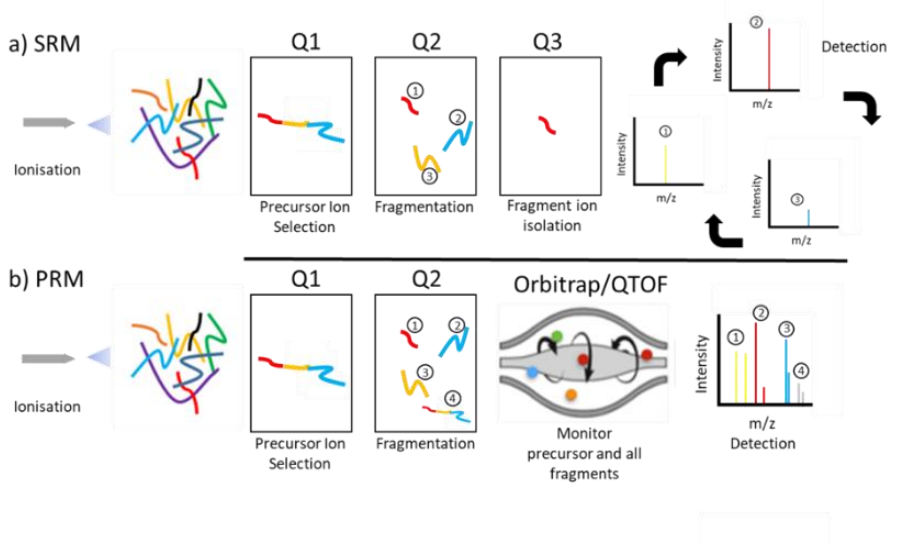

Figure 3.2 Schematic showing targeted LC-MS/MS assays.
A comparison between SRM and PRM targeted LC-MS/MS methods. A) SRM method which quantifies peptides individually. B) PRM method which allows the quantification of multiple peptides simultaneously.

An SRM experiment involves selecting a predefined precursor ion *m/z* (indicative of a specific peptide) which is isolated in the first quadrupole (Q1) and then fragmented in the second quadrupole (Q2) serving as a collision cell. Following fragmentation, a predefined set of

fragment ions (transitions) are filtered through the third quadrupole (Q3) and transmitted to the detector to confirm *m/z* ion identification. The peak area of each transition (defined precursor to fragment ion generation) allows for precise quantification (Lange et al., 2008). However, intensities of individual product ions derived from a single precursor ion can substantially differ, therefore in order to develop a sensitive assay it is vital that the most intense product ions are selected.

A PRM method provides a quicker alternative to SRM whilst having the additional advantage of being able to quantify multiple peptides in a single experiment and maintain increased sensitivity and specificity over SRM (Schiffmann et al., 2014). PRM experiments require high-resolution, hybrid quadrupole-Orbitrap (q-OT) or time-of-flight (QTOF). PRM utilises targeted tandem MS (MS/MS) which allows for the simultaneous monitoring of product ions of a target peptide with high resolution and mass accuracy, hence making prior selection of peptide transitions (fragment ions) redundant (Peterson et al., 2012). Q-OT based PRM is performed by identification of precursor *m/z* ions of interest in the quadrupole, fragmentation in the HCD cell and detection by the high resolution Orbitrap mass analyser (Gallien et al., 2012), i.e. all product ions are identified. Due to the C-trap's ability to store a high number of ions and fill for long periods of time, sensitivity and signal to noise ratio boosting can be achieved by increasing targeted ion counts and fill times (Domon and Gallien, 2015). The software Skyline aids assay design (transitions) and supports quantification of PRM-based targeted LC-MS/MS analysis (MacLean et al., 2010).

3.2. Chapter Aims

The aim of this chapter was to determine whether a biomarker detection method could be created and optimised to allow the detection of the proteins zDHHC23 and/or MROH6 within SK-N-AS cells and SK-N-AS tumours (obtained from the chick embryo). This would ultimately progress into detection of proteins from patient tumour samples, allowing investigation of their use as prognostic markers in combination with the associated clinical data.

3.3. Results

3.3.1. Immunostaining on SK-N-AS cells and tumours

To determine whether the dramatic changes in mRNA expression profiles observed in Kaplan-Meier plots from patient databases (Section 1.6) translated to protein level changes, I performed an adapted method to the clinical 'gold standard' of IHC staining; immunofluorescent (IF) staining. IF and IHC both rely on the same principles of antibody-based staining. However, IF generally refers to cellular staining with fluorescent dye conjugated antibodies whereas IHC generally refers to tissue/tumour staining with coloured chromogens; the two techniques can be used interchangeably on biological samples (Taube et al., 2020). IF allows the determination of the subcellular localisation of the proteins within cells and abundance levels to be quantified by image analysis (Section 2.16). Therefore, antibodies against zDHHC23 (NBP1-82688)and MROH6 (HPA068049) that were available in the Sée lab (purchased and tested prior to my PhD), were tested for IF staining of SK-N-AS cellular and tumour samples (Figure 3.3 - Figure 3.6).

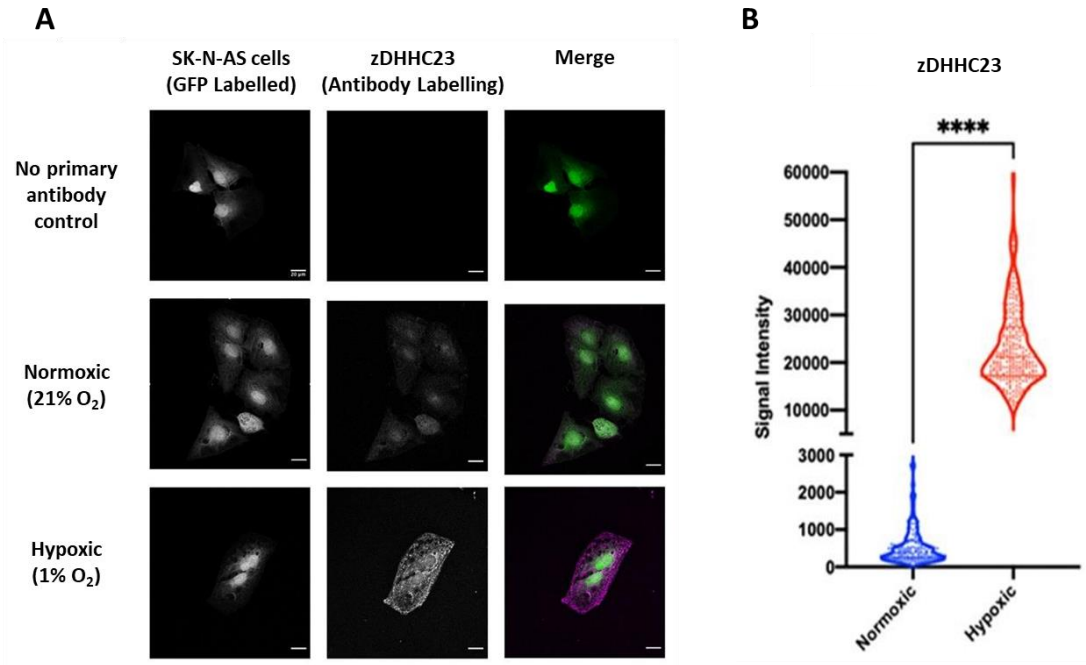

Figure 3.3 zDHHC23 staining of GFP SK-N-AS cells.
*GFP SK-N-AS cells incubated in either normoxia (21% O_2) or hypoxia (1% O_2) for 72 h before staining with zDHHC23 primary antibody (NBP1-82688, 1:20 dilution) and Alexa 555 secondary antibody (1:500 dilution). A) Confocal microscope images were obtained (40x objective), left to right: SK-N-AS cells, zDHHC23 signal, merge of both channels: Green (cells) Magenta (antibody labelling). Scale bar – 20 µm. (Images are representative of N=3 replicates, 5 positions) B) Quantification of images via LabKit in Fiji (Arzt et al., 2022) **** indicates p value <0.001, unpaired T-Test.*

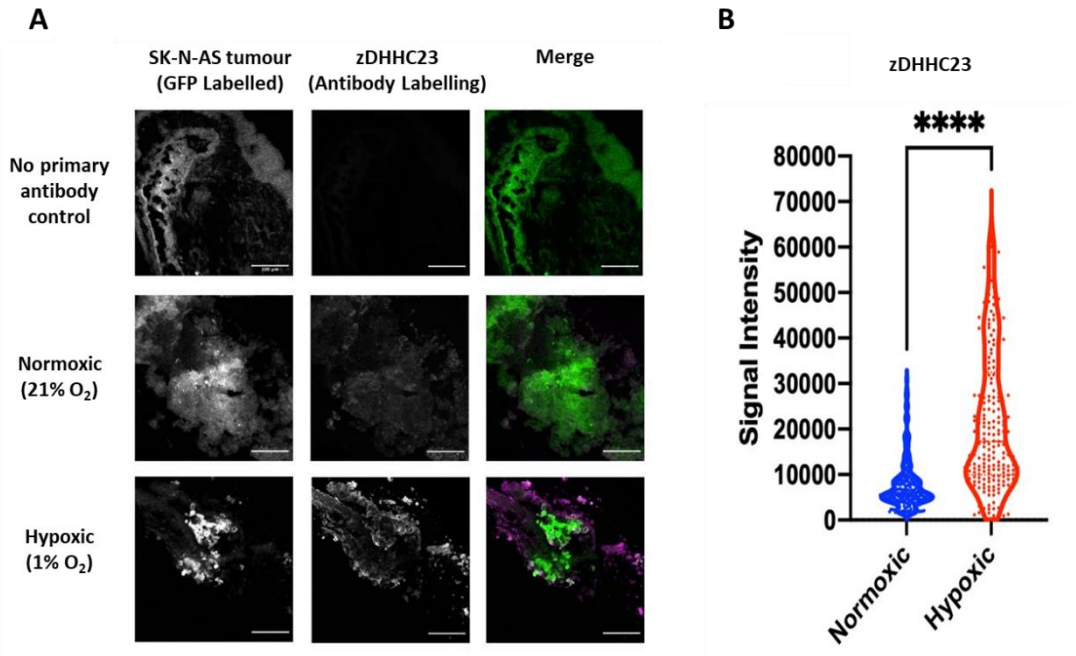

Figure 3.4 zDHHC23 staining of GFP SK-N-AS tumour.
*GFP SK-N-AS tumour (cells preconditioned in either normoxia (21% O_2) or hypoxia (1% O_2)) grown on the CAM of the chick embryo, stained with zDHHC23 primary antibody (NBP1-82688, 1:50 dilution) and Alexa 555 secondary antibody (1:500 dilution). A) Confocal microscope images were obtained (20x objective), left to right: SK-N-AS tumour, zDHHC23 signal, merge of both channels: Green (cells) Magenta (antibody labelling). Scale bar – 100 µm. (Images are representative of N=3, 5 positions) B) Quantification of images via LabKit in Fiji (Arzt et al., 2022) **** indicates p value <0.001, unpaired T-Test.*

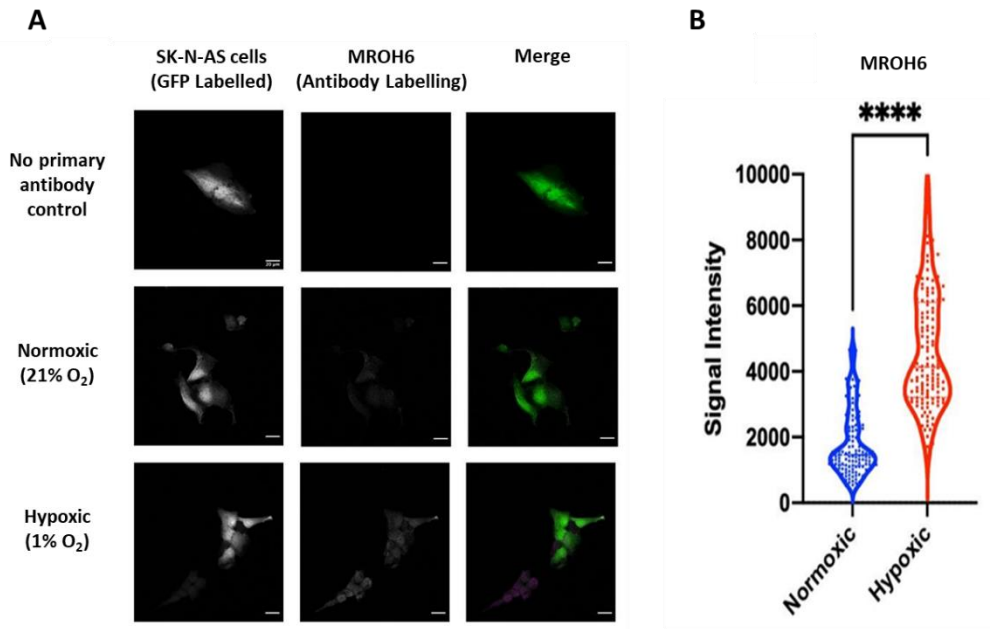

Figure 3.5 MROH6 staining of GFP SK-N-AS cells.
*GFP SK-N-AS cells incubated in either normoxia (21% O_2) or hypoxia (1% O_2) for 72 h before staining with MROH6 primary antibody (HPA068049, 1:200 dilution) and Alexa 555 secondary antibody (1:500 dilution). A) Confocal microscope images were obtained (40x objective), left to right: SK-N-AS cells, MROH6 signal, merge of both channels: Green (cells) Magenta (antibody labelling). Scale bar – 20 µm. (Images are representative of N=3, 5 positions). B) Quantification of images via LabKit in Fiji (Arzt et al., 2022) **** indicates p value <0.001, unpaired T-Test.*

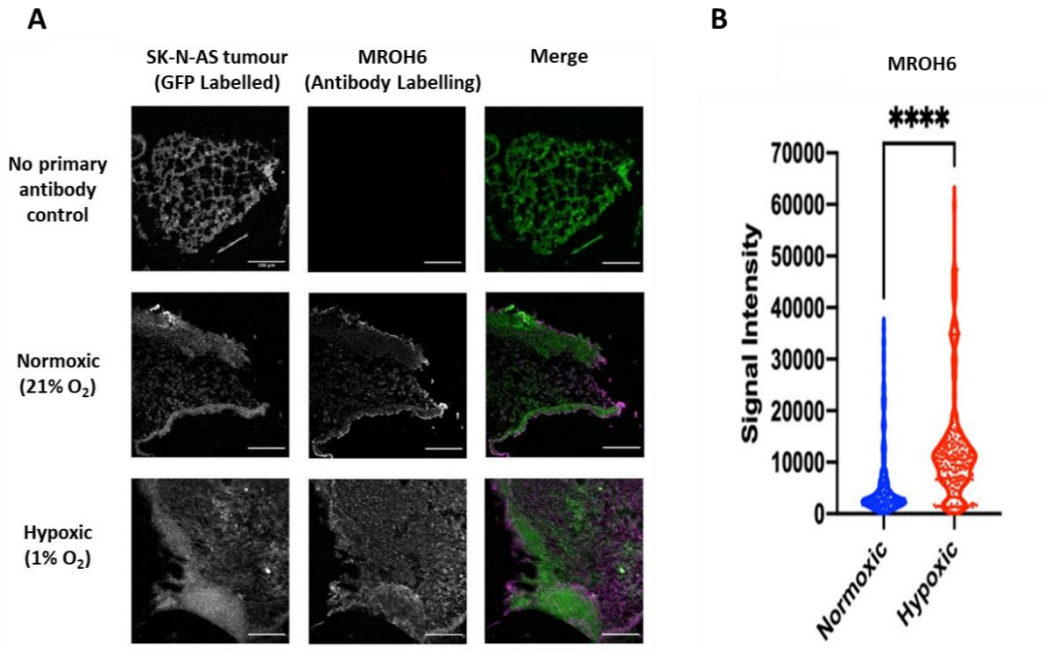

Figure 3.6 MROH6 staining of GFP SK-N-AS tumour.
*GFP SK-N-AS tumour (cells preconditioned in either normoxia (21% O_2) or hypoxia (1% O_2)) obtained via the CAM model, stained with MROH6 primary antibody (HPA068049, 1:200 dilution) and Alexa 555 secondary antibody (1:500 dilution). A) Confocal microscope images were obtained (20x objective), left to right: SK-N-AS tumour, MROH6 signal, merge of both channels: Green (cells) Magenta (antibody labelling). Scale bar – 100 µm. (Images are representative of N=3, 5 positions) B) Quantification of images via LabKit in Fiji (Arzt et al., 2022) **** indicates p value <0.001, unpaired T-Test.*

zDHHC23 protein expression was significantly up-regulated by hypoxia at both a cellular level (~80 fold increase, p-value < 0.001) and, to a lesser degree, in the tumour model (~3 fold increase, p-value < 0.001) (Figure 3.3 & Figure 3.4). The change in expression of zDHHC23 at the protein level in the tumours is comparable to the ~2 fold increase observed at the mRNA level in the presence of hypoxia in this system (data from the Sée lab, data not shown). Interestingly, while zDHHC23 is reported to predominantly localise to the Golgi in murine vascular smooth muscle cells (Duncan et al., 2019), I identified strong plasma membrane association and cytoplasmic localisation in these samples. MROH6 protein expression was also strongly and significantly up-regulated by hypoxia (to a similar level as that observed at the mRNA level, Sée lab, data not shown), with a ~3-fold increase (p-value < 0.001) observed under hypoxic conditions in both the cell and tumour model systems (Figure 3.5 and Figure 3.6). While there are no reported experimental localisation studies on MROH6, prediction tools suggest MROH6 to be cytoplasmic (Consortium, 2019), whereas my data predominantly indicates nuclear localisation. While these results are encouraging, before these antibodies can be used to draw potentially clinically relevant conclusions from patient samples, it is

essential to properly validate antibody specificity, particularly noting to differences in the cellular localisation data.

3.3.2. Validation of zDHHC23 and MROH6 antibodies

Small interfering RNA (siRNA) can be manufactured to a gene specific sequence, and thus suffer off-target effects to a lesser extent than antibody production (Jackson and Linsley, 2010). Target specific siRNA pools were used to reduce the expression of either zDHHC23 or MROH6, and western blotting with our antibody used to explore knockdown and antibody specificity. To confirm successful knockdown with these siRNAs, I also performed RT-qPCR to monitor mRNA levels in cells post knockdown (Figure 3.7).

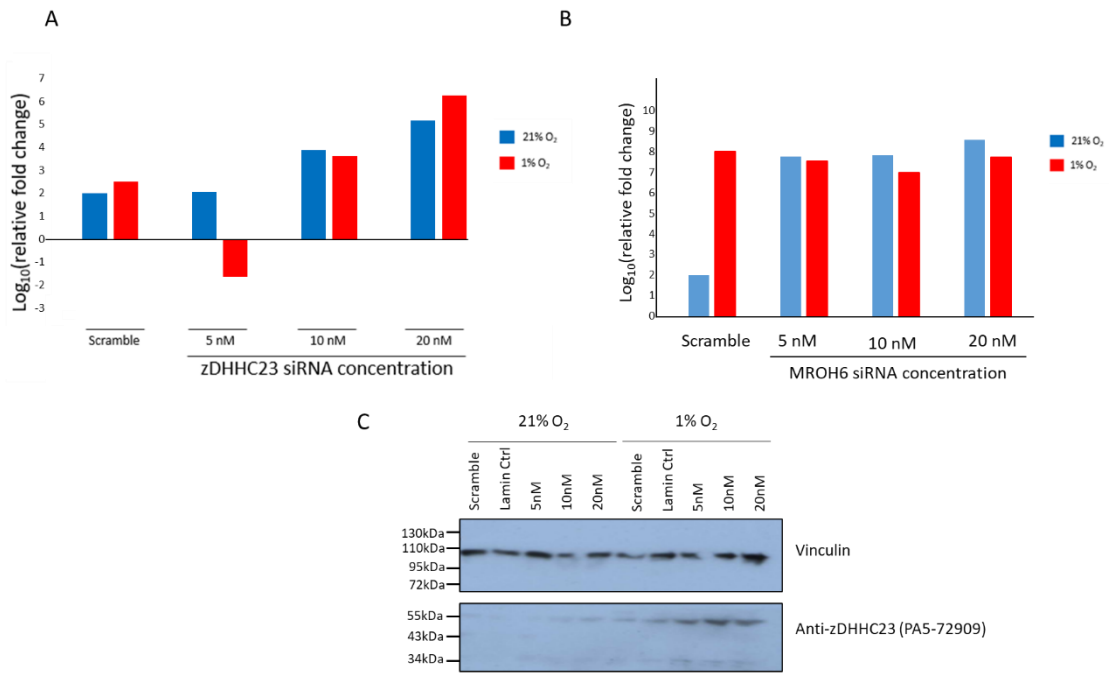

Figure 3.7 mRNA and protein expression of zDHHC23 and MROH6 following siRNA treatment. GFP SK-N-AS cells were transfected with varying concentrations of zDHHC23/MROH6 siRNA (Dharmacon) using FuGENE HD, followed by incubation in either 21% O_2 or 1% O_2 for 72 hr before collection. A) zDHHC23 mRNA levels were assessed by qPCR. Fold gene expression change was determined relative to normoxic (21% O_2) controls using the Delta Ct normalisation method. CYCLOA was used as a housekeeping gene. B) MROH6 mRNA levels were assessed by qPCR. Fold gene expression change was determined relative to normoxic (21% O_2) controls using the Delta Ct normalisation method. CYCLOA was used as a housekeeping gene. C) Western blot analysis of zDHHC23 protein levels with vinculin used as a loading control.

For zDHHC23, successful knockdown (in 1% O_2) was only achieved at 5 nM siRNA concentration (Figure 3.7A). It is also noted that an increase in siRNA concentration correlates with an increase in mRNA expression. Western blot analysis with the zDHHC23 antibody not only failed to identify a band at the expected molecular weight (~45 kDa), with antibody binding at ~55 kDa and 34 kDa, but also confirms our mRNA level observation, with a siRNA concentration-dependent increase in banding intensity, whilst the vinculin housekeeping protein shows similar protein loading (Figure 3.7C). Despite multiple attempts, MROH6 knockdown was not achieved (Figure 3.7B). With obvious differences between western blotting and IF staining (denatured versus folded protein state), it is possible that the lack of antibody specificity may be due to the technique used. Thus, I attempted to validate the antibodies by a technique conceptually more similar to IF (based on binding to a folded protein) namely antibody-based IP, using tandem MS to identify antibody-bound protein.

3.3.3. zDHHC23 and MROH6 antibody IPs

As eluded to, both of our proteins of interest are minimally researched, particularly MROH6 which at the time of writing is classified as a 'predicted protein' on UniProt (Consortium, 2019), therefore antibody selection was limited. At the time of conducting these experiments there were 12 available antibodies against zDHHC23, two of which claimed to be IP/IHC compatible – the antibody previously used by the Sée lab (Novusbio (NBP1-82688, now discontinued), and one from ThermoFisher (PA5-72909)) which I purchased for testing. Only four antibodies were commercially available for MROH6, none of which stated they were IP compatible, however two were IHC validated and were therefore purchased (ThermoFisher (PA5-64355) and Atlas Antibodies (HPA068049)). A protocol for the IP of HA-tagged proteins using a commercially available anti-HA antibody and magnetic bead crosslinking kit (Pierce Crosslink Magnetic IP/Co-IP Kit #88805) had previously been optimised within the Eyers lab. Therefore, I used this as a starting point for IP by substituting the anti-HA antibody with my zDHHC23 and MROH6 antibodies, analysing IP efficiency by western blot using the same antibody (Figure 3.8 and Figure 3.9).

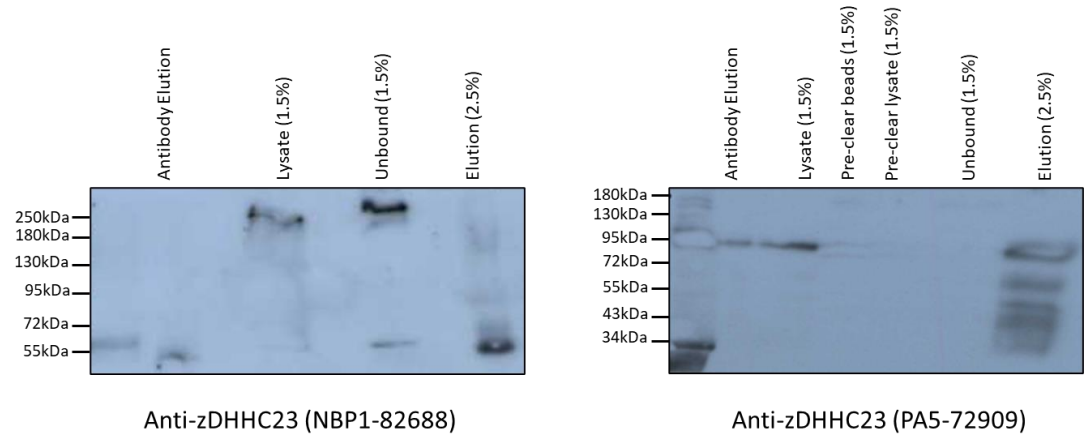

Figure 3.8 IP optimisation for endogenous zDHHC23 protein.
Western blot analysis using either the Novusbio (NBP1-82688) or the Thermofisher (PA5-72909) antibodies of total lysate (1.5% volume), unbound protein material (1.5% volume) and eluted fraction (2.5% volume) following IP using the Eyers lab protocol. GFP SK-N-AS cells were incubated at 21% O_2 for 72 hr before lysis. Lysis buffer composition: 50 mM Tris pH 8.0, 120 mM NaCl, 1 mM EDTA, 0.2% NP-40. Lysis buffer was supplemented with EDTA-free cOmplete protease inhibitors and phosSTOP phosphatase inhibitors (Roche).

Initial IP attempts with the anti-zDHHC23 antibody (NBP1-82688) failed to identify any band at the predicted molecular weight (MW) of ~45 kDa, instead showing a band at ~55 kDa (similar to Figure 3.7) (Figure 3.8). Additionally, this anti-zDHHC23 antibody (NBP1-82688) picked up a band at >250 kDa that was not IP'd (lack of band in elution lane) but is observed in lysate and unbound lanes (Figure 3.8). It is noted that the band at >250 kDa will not have been observed previously in Figure 3.7 as the same membrane was used to probe for the vinculin housekeeping protein at ~100 kDa. (Figure 3.8). The second anti-zDHHC23 antibody (PA5-72909) does identify a band at ~45 kDa (smearing to ~72 kDa which may arise due to extensive post-translational modification) post IP elution, but no such immunoreactivity is observed in the input lysate (Figure 3.8). It is also noted that a strong elution band is observed at ~95 kDa, however, since this band is also observed in the 'mock' IP lanes (antibody bead binding controls, i.e. beads that undergone IP using protein free lysis buffer) this band is likely from poor crosslinking efficiency of the antibody to the magnetic beads (Figure 3.8). As it is common for proteins to appear at a MW different to the expected (for example HIF-1α appears at ~120 kDa rather than the predicted 92 kDa likely due to extensive phosphorylation (Daly et al., 2021) and the possibility of falsely validating itself (by using the same antibody for IP and western blot), IP samples were prepared for mass spectrometry analysis for an unbiased analysis of the immunoprecipitated proteins. Unfortunately, no peptides deriving from zDHHC23 were identified following LC-MS/MS analysis of the material

immunoprecipitated with either of the zDHHC23 antibodies (Table 3.1), allowing me to conclude that these antibodies lacked suitable specificity for this protein.

Table 3.1 Top 20 LC-MS/MS hits for two zDHHC23 antibody based IPs.
List of proteins identified via LC-MS/MS analysis following IP with a zDHHC23 antibody. Left: NBP1-82688 Right: PA5-72909.

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P35579	Myosin-9	127	302	14138
2	P35580	Myosin-10	126	257	12401
3	P60709	Actin, cytoplasmic 1	10	115	5303
4	Q15149	Plectin	94	103	4209
5	P08670	Vimentin	42	100	3907
6	P21333	Filamin-A	60	83	3665
7	O95425	Supervillin	56	77	3637
8	Q9UM54	Unconventional myosin-VI	55	83	3247
9	P12814	Alpha-actinin-1	32	71	3223
10	Q9UMS6	Synaptopodin-2	47	69	3055
11	O43795	Unconventional myosin-Ib	45	67	2947
12	P48681	Nestin	52	64	2933
13	Q00839	Heterogeneous nuclear ribonucleoprotein U	39	67	2460
14	P13645	Keratin, type I cytoskeletal 10	27	44	2443
15	P52272	Heterogeneous nuclear ribonucleoprotein M	29	59	2404
16	Q08211	ATP-dependent RNA helicase A	44	58	2390
17	P08238	Heat shock protein HSP 90-beta	21	55	2315
18	P14618	Pyruvate kinase PKM	24	40	2171
19	Q9Y5B9	FACT complex subunit SPT16	36	50	2159
20	O43707	Alpha-actinin-4	21	51	2114

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P35579	Myosin-9	141	291	9268
2	P35580	Myosin-10	118	216	6571
3	P21333	Filamin-A	70	90	2834
4	P04264	Keratin, type II cytoskeletal 1	35	69	2611
5	P08670	Vimentin	45	84	2396
6	P60709	Actin, cytoplasmic 1	1	89	2085
7	P63261	Actin, cytoplasmic 2	1	89	2074
8	P35908	Keratin, type II cytoskeletal 2 epidermal	31	53	2071
9	Q15149	Plectin	79	83	1704
10	P13645	Keratin, type I cytoskeletal 10	28	44	1615
11	P35527	Keratin, type I cytoskeletal 9	22	37	1458
12	Q00839	Heterogeneous nuclear ribonucleoprotein U	31	49	1357
13	P02545	Prelamin-A/C	25	35	1330
14	P68032	Actin, alpha cardiac muscle 1	2	66	1315
15	Q9UMS6	Synaptopodin-2	43	52	1299
16	Q9UM54	Unconventional myosin-VI	25	40	1222
17	O95425	Supervillin	40	44	1186
18	P00558	Phosphoglycerate kinase 1	24	35	1090
19	P14618	Pyruvate kinase PKM	27	36	1015
20	P11142	Heat shock cognate 71 kDa protein	16	33	1014

Repeating the IP and LC-MS/MS analysis for both MROH6 antibodies resulted in equally unfruitful data (Figure 3.9 & Table 3.2). The initial anti-MROH6 antibody (HPA068049) revealed a band in the lysate and the IP samples at ~50 kDa (Figure 3.9, below the predicted ~78 kDa). However, MROH6 was not identified in subsequent LC-MS/MS analyses (Table 3.2).

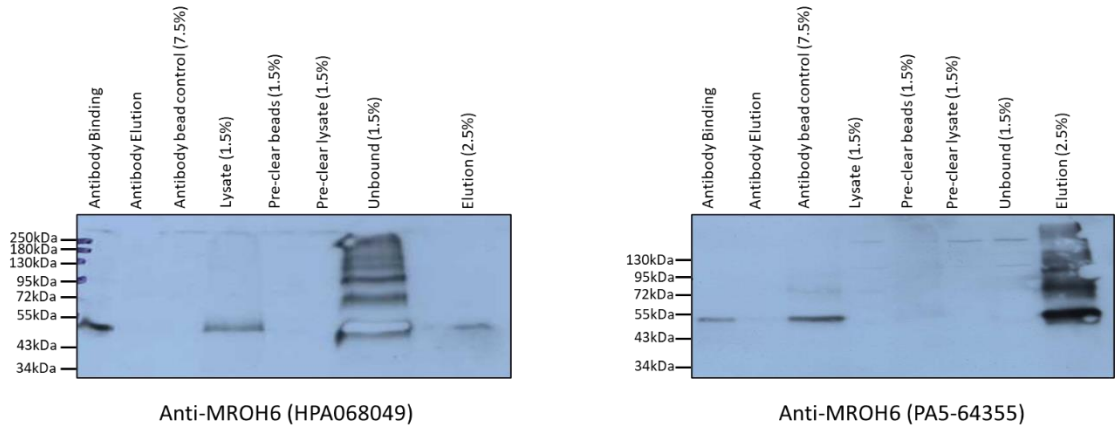

Figure 3.9 IP optimisation for endogenous MROH6 protein.
Western blot analysis using either the Atlas Antibodies (HPA068049) or the Thermofisher (PA5-64355) antibodies of total lysate (1.5% volume), unbound protein material (1.5% volume) and eluted fraction (2.5% volume) following IP using the Eyers lab protocol. GFP SK-N-AS cells were incubated at 21% O_2 for 72 hr before lysis. Lysis buffer composition: 50 mM Tris pH 8.0, 120 mM NaCl, 1 mM EDTA, 0.2% NP-40. Lysis buffer was supplemented with EDTA-free cOmplete protease inhibitors and phosSTOP phosphatase inhibitors (Roche).

Table 3.2 Top 20 LC-MS/MS hits for two MROH6 antibody based IPs.
List of proteins identified via LC-MS/MS analysis following IP with a MROH6 antibody. Left: HPA068049 Right: PA5-64355.

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P60709	Actin, cytoplasmic 1	9	283	7574
2	P68133	Actin, alpha skeletal muscle	1	224	5960
3	P04264	Keratin, type II cytoskeletal 1	29	183	5683
4	P35527	Keratin, type I cytoskeletal 9	18	131	4570
5	P21333	Filamin-A	50	160	3658
6	P08670	Vimentin	38	141	3268
7	P35579	Myosin-9	56	115	2846
8	P07437	Tubulin beta chain	3	116	2555
9	P68371	Tubulin beta-4B chain	3	96	1795
10	P04406	Glyceraldehyde-3-phosphate dehydrogenase	11	48	1737
11	Q562R1	Beta-actin-like protein 2	3	71	1719
12	P35908	Keratin, type II cytoskeletal 2 epidermal	9	49	1714
13	P62979	Ubiquitin-40S ribosomal protein S27a	9	60	1602
14	P35580	Myosin-10	36	84	1526
15	P08238	Heat shock protein HSP 90-beta	18	81	1493
16	P13645	Keratin, type I cytoskeletal 10	15	75	1486
17	P14618	Pyruvate kinase PKM	19	44	1247
18	P42766	60S ribosomal protein L35	11	54	1236
19	P02768	Albumin	5	57	1234
20	P68363	Tubulin alpha-1B chain	1	56	1164

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P35579	Myosin-9	102	208	5086
2	P35580	Myosin-10	87	152	3220
3	P60709	Actin, cytoplasmic 1	7	76	2062
4	P21333	Filamin-A	50	66	1517
5	Q15149	Plectin	66	71	1344
6	Q9UHD8	Septin-9	27	45	1382
7	P62736	Actin, aortic smooth muscle	1	65	1656
8	Q9UMS6	Synaptopodin-2	34	50	1387
9	O95425	Supervillin	33	43	1048
10	P04264	Keratin, type II cytoskeletal 1	23	35	1340
11	P08670	Vimentin	31	56	908
12	P35908	Keratin, type II cytoskeletal 2 epidermal	20	32	1181
13	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	18	35	962
14	P09493	Tropomyosin alpha-1 chain	8	32	749
15	Q9UPQ0	LIM and calponin homology domains-containing protein 1	26	33	644
16	P46821	Microtubule-associated protein 1B	22	25	682
17	Q00839	Heterogeneous nuclear ribonucleoprotein U	17	28	541
18	Q86V48	Leucine zipper protein 1	22	26	670
19	P10809	60 kDa heat shock protein, mitochondrial	16	23	847
20	P35527	Keratin, type I cytoskeletal 9	15	19	853

Interestingly, yet perplexing, while the IP is performed at 4 °C in the presence of phosphatase and protease inhibitors, it is clear that in the unbound fraction there is a laddered banding pattern that is not present in the lysate or elution lanes (Figure 3.9). While it may be assumed the laddering may be due to PTM states that occurred due to overnight 4 °C incubation for IP, it is hard to conceive that the PTMs completely mask the polyclonal antibody epitopes

recognised by IP but not western blot (due to banding in unbound and lack of in elution), or result in complete protein degradation (or PTM removal) within the ~15 min required to wash and elute into 4X SDS loading buffer. Similar to the anti-zDHHC23 antibody PA5-72909, the second anti-MROH6 antibody (PA5-64355) appeared to inefficiently crosslink to the magnetic beads, with the presence of banding in the 'mock' IPs antibody-bead control lane at ~50 kDa which is also observed in IP elution lanes (Figure 3.9). While a band did appear closer to the predicted MW in the elution lane at ~72 kDa (and smearing above), LC-MS/MS-based analysis failed to identify MROH6 (Table 3.2).

To confirm that LC-MS/MS analysis could in fact identify zDHHC23 and MROH6 tryptic peptides and validate our MS results was not due to the stochastic nature of tryptic cleavage generating poor LC-MS compatible peptides, I purchased HA-tagged plasmids of zDHHC23 (HA-zDHHC23) and MROH6 (HA-MROH6) from NovoPro Labs and performed the IP and subsequent LC-MS/MS analysis using the anti-HA antibody (Cell Signalling #3724S) (Figure 3.10 and Table 3.3). Exogenous expression of HA-zDHHC23 and HA-MROH6 and probing with anti-HA antibody shows that both proteins predominantly appear at their expected MWs of ~45 kDa and ~78 kDa (Figure 3.10). Additionally, both zDHHC23 and MROH6 are one of the most confidently identified proteins in their respective by LC-MS/MS analysis post IP, showing that their lack of identification following IP on endogenous protein is not due to LC-MS/MS related issues (Table 3.3).

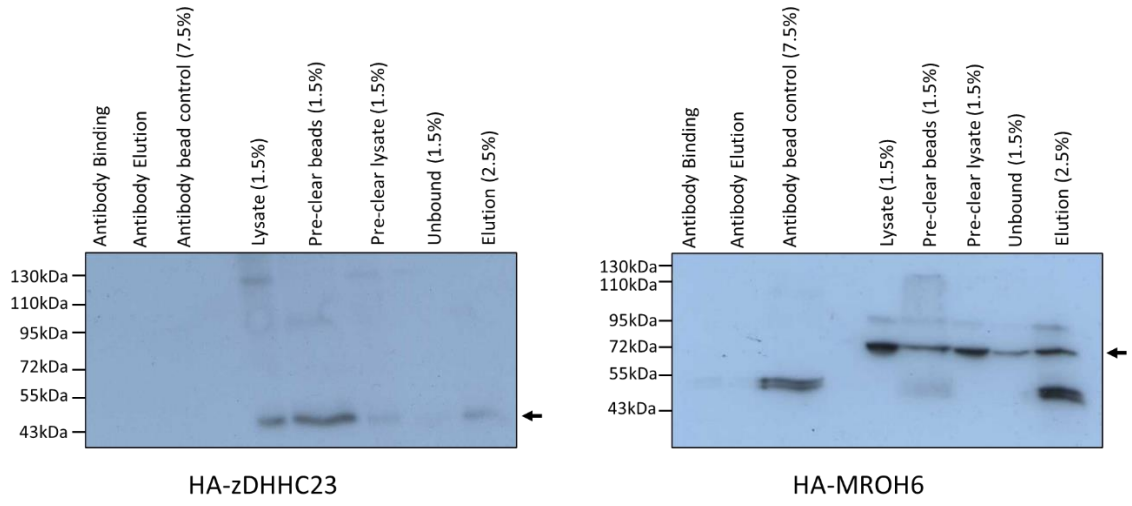

Figure 3.10 Determination of the suitability of the HA tag for zDHHC23 and MROH6 IP.
Western blot analysis of GFP SK-N-AS cells transfected with either the HA-zDHHC23 or HA-MROH6 plasmid 16 hr in 21% O_2 prior to lysis. The mass spectrometry compatible HA tag protocol was followed and probed with an anti-HA primary antibody. Arrows identify potential protein of interest.

Table 3.3 Top LC-MS/MS hits for HA-zDHHC23 and HA-MROH6 IP.
List of proteins identified via LC-MS/MS analysis following IP with HA-zDHHC23 (left) or HA-MROH6 (right) IP.

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P04264	Keratin, type II cytoskeletal 1	31	187	5466
2	P35527	Keratin, type I cytoskeletal 9	31	150	5253
3	P62805	Histone H4	7	170	4368
4	P13645	Keratin, type I cytoskeletal 10	24	113	3658
5	P35908	Keratin, type II cytoskeletal 2	26	83	2529
6	P60709	Actin, cytoplasmic 1	6	82	1977
7	P08670	Vimentin	29	66	1726
8	Q5QNW6	Histone H2B type 2-F	2	95	1708
9	Q16778	Histone H2B type 2-E	2	72	1521
10	P13647	Keratin, type II cytoskeletal 5	20	60	1421
11	P08779	Keratin, type I cytoskeletal 16	10	43	1401
12	P35579	Myosin-9	47	64	1399
13	P02533	Keratin, type I cytoskeletal 14	3	40	1205
14	P63267	Actin, gamma-enteric smooth muscle	2	52	1121
15	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	14	41	1036
16	Q9NR30	Nucleolar RNA helicase 2	27	46	1029
17	P02538	Keratin, type II cytoskeletal 6A	3	43	992
18	Q00839	Heterogeneous nuclear ribonucleoprotein U	20	44	967
19	Q08211	ATP-dependent RNA helicase A	23	35	957
20	P04259	Keratin, type II cytoskeletal 6B	1	42	957
21	P35580	Myosin-10	28	38	844
22	Q8IYP9	Palmitoyltransferase zDHHC23	13	27	835

Score Position	Accession	Description	#Unique Peptides	#PSMs	Mascot Score
1	P35579	Myosin-9	79	158	5276
2	P35908	Keratin, type II cytoskeletal 2	30	77	3132
3	P04264	Keratin, type II cytoskeletal 1	29	79	2972
4	P60709	Actin, cytoplasmic 1	11	96	2769
5	P13645	Keratin, type I cytoskeletal 10	23	69	2656
6	P35580	Myosin-10	50	84	2594
7	A6NGR9	Maestro heat-like repeat-containing protein family member 6	29	67	2322
8	P35527	Keratin, type I cytoskeletal 9	25	53	2285
9	P08670	Vimentin	30	64	2241
10	P56945	Breast cancer anti-oestrogen resistance protein 1	27	69	2060

Overall, I have little confidence that the antibodies are capable of binding to zDHHC23 or MROH6 in a folded (IP) or unfolded (western blot) state, thus it is difficult to draw meaningful conclusions from IHC/IF data from Figure 3.3 to Figure 3.6. With associated cost and timing issues to validate and the test additional antibodies, I decided to explore the use of LC-MS/MS-based PRM analysis as a method of identifying whether zDHHC23 and MROH6 can be used as NB biomarkers.

3.3.4. Parallel Reaction Monitoring

As previously discussed, PRM is a method in which an Orbitrap MS is used to selectively and sensitively identify a few peptides of interest and can be used quantitatively to determine the number of molecules of a protein expressed in a given sample (Schiffmann et al., 2014). In order to develop the PRM assay, data from the HA-tagged IP LC-MS/MS experiments were used to define peptide precursor ion *m/z* values and retention time information for targeting of zDHHC23 and MROH6, choosing 6 and 11 peptides respectively (listed in Table 3.4). The PRM assay was validated on the IP'd material (Figure 3.11) and in SK-N-AS cells that were exogenously expressing the HA-tagged zDHHC23 or MROH6.

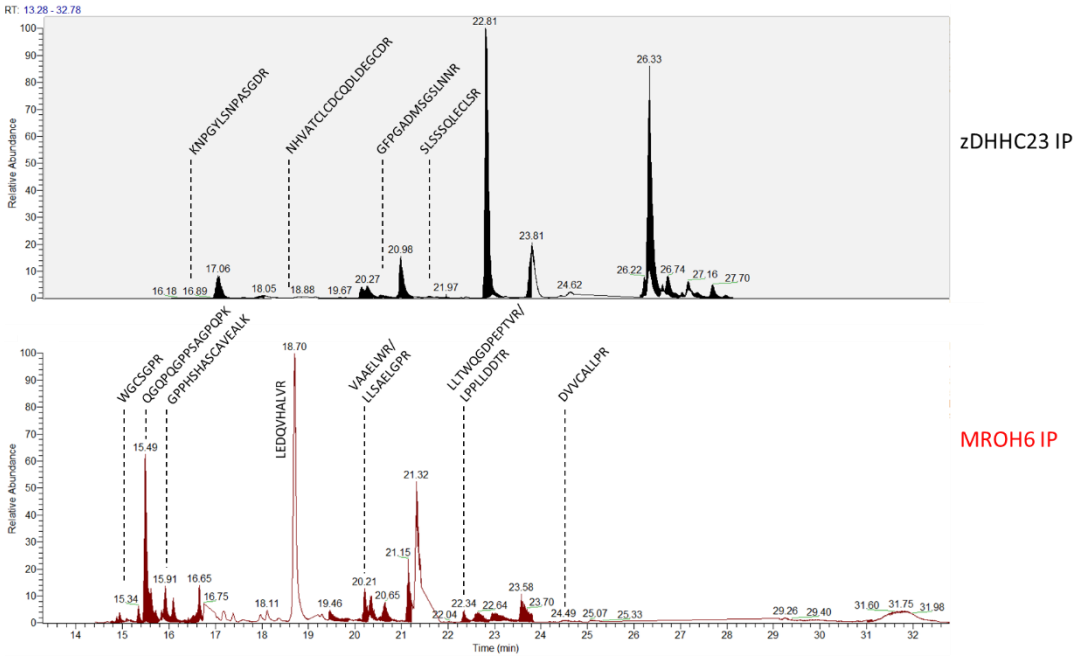

Figure 3.11 Extracted ion chromatogram of PRM on zDHHC23 and MROH6 IP'd SK-N-AS lysate.
A representative example of LC/MS-MS chromatogram results from PRM method on zDHHC23 IP lysate (black) and MROH6 IP lysate (red). Peptides seen are labelled at their elution time. Chromatograms obtained from Xcalibur.

Table 3.4 Peptide inclusion list for PRM.
List of peptides identified in at least one of three biological replicates which were then used as the PRM inclusion list for zDHHC23 and MROH6. RT and m/z for each peptide is listed, m/z was searched with a +/- 10 ppm parameter.

zDHHC23			
m/z	**Start RT (Min)**	**End RT (Min)**	**Sequence**
504.2344	13.64	15.64	AKEDWCAK
560.2586	14.59	16.59	SLQPETCER
738.3655	15.54	17.54	KNPGYLSNPASGDR
1169.4512	17.44	19.44	NHVATCLCDCQDLDEGCDR
638.8352	19.28	21.28	SLSSSQLECLSR
711.8252	20.10	22.10	GFPGADMSGSLNNR

MROH6			
m/z	**Start RT (Min)**	**End RT (Min)**	**Sequence**
737.3759	14.41	16.41	QGQPQGPPSAGPQPK
780.8830	14.74	16.74	GPPHSHASCAVEALK
410.1820	14.75	16.75	WGCSGPR
590.3277	17.00	19.00	LEDQVHALVR
478.2822	19.21	21.21	LLSAELGPR
422.7374	19.43	21.43	VAAELWR
520.2928	21.82	23.82	LPPLLDDTR
756.3963	21.83	23.83	LLTWQGDPEPTVR
521.7893	23.37	25.37	DVVCALLPR
578.8561	29.12	31.12	ASAVGLLGTLVR
856.4831	31.01	33.01	EAPVGALTLTALTEGIR

Unfortunately, we were unable to identify any of our target peptides in endogenous SK-N-AS cells (data not shown). While lack of observation of these peptides may be due to a lack of

protein expression (or very low expression levels) it is possible that these proteins are differentially expressed in the cell versus the tumour models, particularly considering the difference in transcript levels obtained from the tumours. Hence, I continued with the PRM assay for tumour analysis.

3.3.1. Optimisation of SK-N-AS tumour preparation for PRM analysis

The ultimate aim for the application of the PRM assay is for screening of clinical samples, therefore it is important to consider the standard clinical processing techniques of tumours that involves formalin fixation and paraffin embedding (FFPE) (Gaffney et al., 2018). While FFPE results in minimal interference with tissue integrity/architecture for IHC staining and being far cheaper in comparison to cryopreservation (Buczak et al., 2020), its application to proteomics MS-based workflows is significantly more difficult. FFPE relies upon formalin-induced crosslinking of proteins, lipids and nucleic acids, hence reducing the efficiency of protein extraction, and inducing multiple artefactual protein modifications that need to be accounted for in LC-MS/MS data analysis (Shi et al., 2001, Auerbach et al., 1977, Coscia et al., 2020).

Methods such as Heat Induced Antigen Retrieval (HIAR), have been developed to reverse the crosslinking of proteins and facilitate protein analysis (Shi et al., 1991). Variations to HIAR protocols that enhance protein recovery and/or reduce harsh treatment conditions are common practise, however all require high temperatures (Ikeda et al., 1998). The mechanism of HIAR for the reversal of crosslinking is unknown, with predictions involving: the hydrolytic cleavage of methylene bridges, unfolding of proteins to reveal potentially non-crosslinked epitopes and the potential extraction of calcium ions from protein complexes (Morgan et al., 1994, Taylor et al., 1996, Morgan et al., 1997, Yamashita and Okada, 2005).

As such, multiple variations to the original HIAR protocol exist, each claiming to result in better protein identification numbers by LC-MS/MS. Some common variations include pH, metal ion chelation, detergents, reduction agents and salt (NaCl) concentration (Shi et al., 1995, Beckstead, 1994, Krenacs et al., 2010, Yamashita and Okada, 2005, Yamashita et al., 2013, Larsson, 2020). However, they all suggest a need for sample-dependent optimisation. Hence, I began by optimising the HIAR protocol on paraformaldehyde treated, SK-N-AS tumours grown in the chick embryo model (conditions outlined in Table 3.5). I additionally

investigated the aid of mechanical perturbations by use of bead beating (Minilys, P000673-MLYS0-A) with stainless steel beads (Precellys, #432-3753). All conditions tested were analysed by SDS-PAGE with coomassie staining and typical shotgun-proteomics LC-MS/MS analysis pipelines (Figure 3.12 and Figure 3.13).

Table 3.5 Buffer conditions used for testing HIAR in SK-N-AS tumours.

Condition Number	HIAR Buffer
1	20mM Tris and 2% SDS
2	20mM Tris, 2% SDS and 10mM DTT
3	20mM Tris, 2% SDS, 10mM DTT and bead homogeniser (BH)
4	20mM Tris, 2% SDS, 10mM DTT, 1 mM EDTA and BH
5	20mM Tris, 10% SDS, 10mM DTT, 1 mM EDTA and BH
6	20mM Tris, 10% SDS, 10 mM DTT, 1 mM EDTA, 120 Mm NaCl and BH

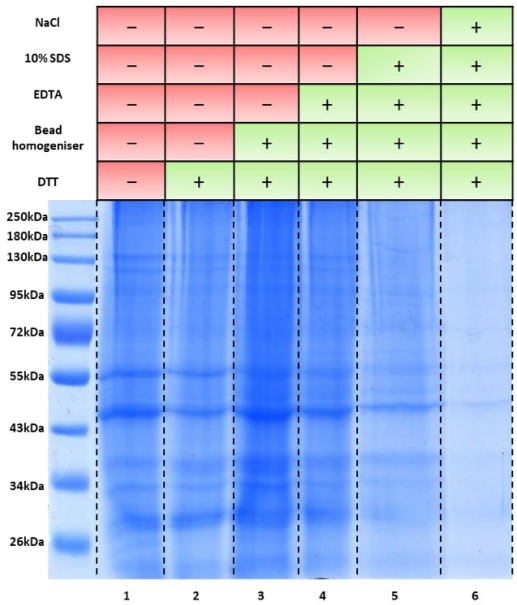

Figure 3.12 Coomassie gel of SK-N-AS tumour HIAR methods.
Colloidal coomassie stained SDS-polyacrylamide gel of tumour lysates processed by various lysis solutions/methods (listed in Table 3.5), 10% of the sample volume was loaded and the gel stained with Colloidal coomassie.

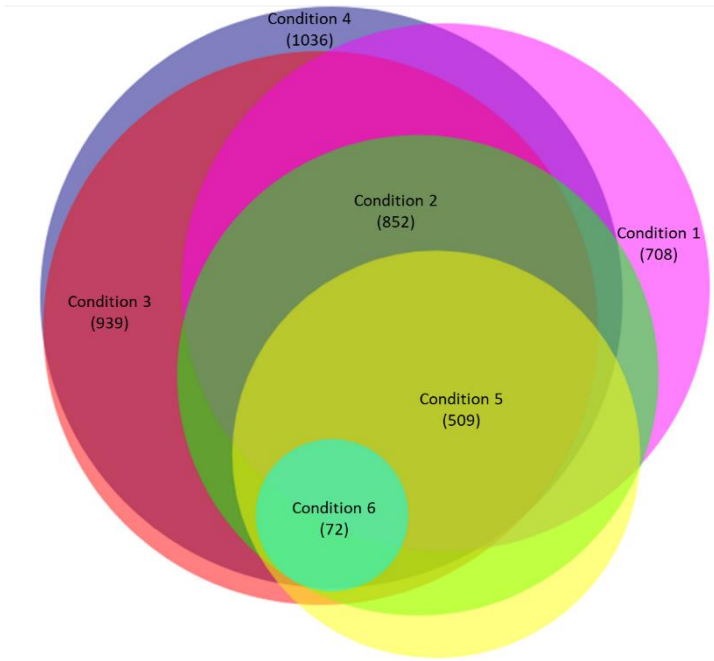

Figure 3.13 Venn diagram showing number of proteins identified in 21% O_2 SK-N-AS tumours obtained from the chick embryo model.
Purple (Condition 1) = tumour processed in 20mM Tris, 2% SDS. Green (Condition 2) = tumour processed in 20mM Tris, 2% SDS and DTT. Red (Condition 3) = tumour processed in 20mM Tris, 2% SDS, DTT and BH used. Navy (Condition 4) = tumour processed in 20mM Tris, 2% SDS, DTT, EDTA and BH used. Yellow (Condition 5) = tumour processed in 20mM Tris, 10% SDS, DTT, EDTA and BH used. Aqua (Condition 6) = tumour processed in 20mM Tris, 10% SDS, DTT, EDTA, NaCl and BH used. Circles are scaled respectively to the number of proteins per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

All tumours were simultaneously grown from identical numbers of cells and weighed prior to HIAR treatment, resulting in six tumours that were a maximum of +/- 0.25mg apart. Coomassie staining of all proteins shows fairly consistent staining intensity between conditions 1-4, with condition 3 appearing the most intense (Figure 3.12). Surpisingly, I identify the harsher treatment protocols (5 & 6) significantly reduced protein staining, and presumed protein recovery (Figure 3.12 and Figure 3.13). LC-MS/MS analysis of equal volume injections identified HIAR protocol 4 was the overall best method identifying 1036 proteins, ~100 proteins more than the second best method; method 3 (Figure 3.13). Corroberating Figure 3.12, protocol 6 was significantly worse than any other protocol only identifying 72 proteins, hence suggesting the high concentration of NaCl somhow inhibits the crosslink reversal.

Having optimised the PRM assay for zDHHC23 and MROH6 detection and HIAR protocol for SK-N-AS tumours grown in the chick embryo model, I applied the PRM technique to 3x SK-N-

AS CAM tumours grown in normoxic and hypoxic conditions (representative data from one of each tumour type (Figure 3.14) and (Figure 3.15). Exogenously expressed HA-tagged zDHHC23 or MROH6 was used as a positive control for the PRM assay. While peaks with a similar *m/z* value and retention time were observed for most of the selected zDHHC23 and MROH6 peptides, upon MASCOT data analysis the tumour peptide ions were identified as being specific to Serine palmitoyltransferase 2 and Nestin (in zDHHC23 PRM) and Cytoplasmic dynein 1 heavy chain (in MROH6 PRM), two proteins that were also identified following overexpression of the relevant target proteins. Thus, currently I am unable to determine whether zDHHC23 or MROH6 may be viable biomarkers for NB.

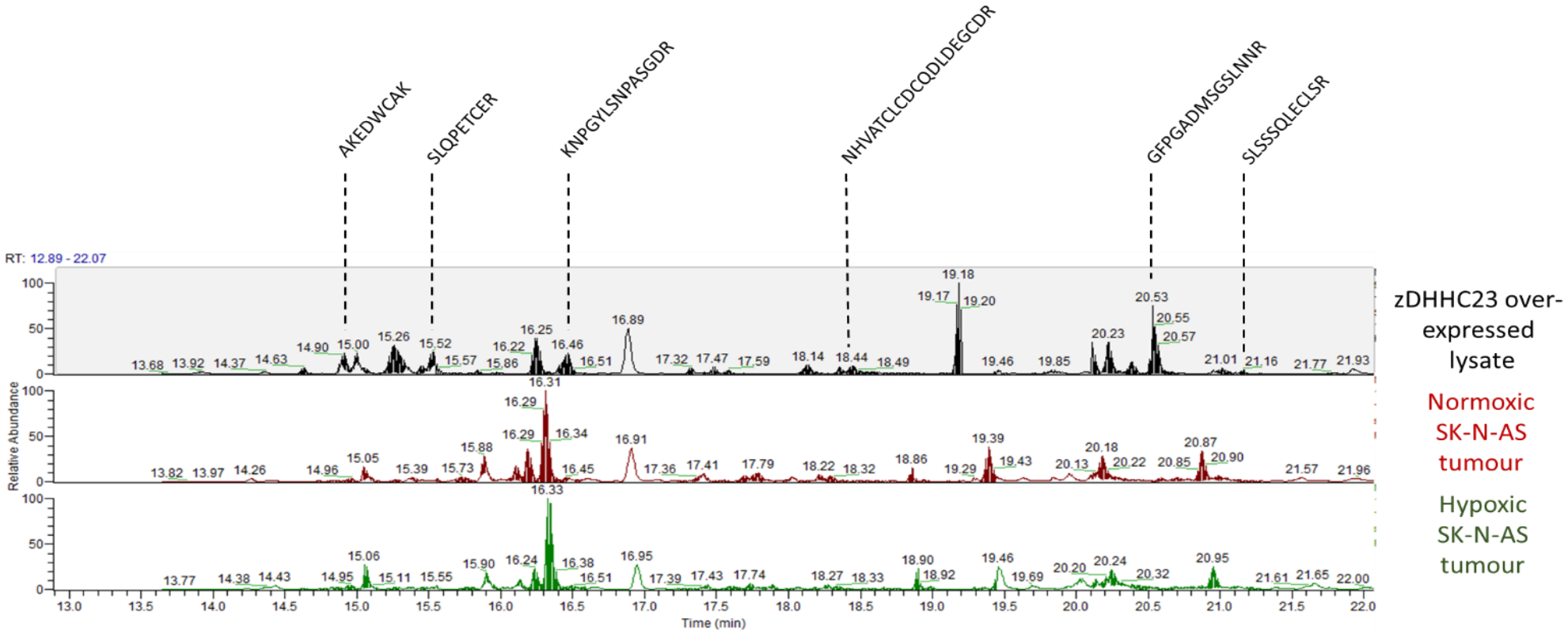

Figure 3.14 Extracted ion chromatogram of zDHHC23 PRM on 21% and 1% O_2 SK-N-AS tumours.
LC/MS-MS chromatogram results from zDHHC23 PRM method on zDHHC23 over-expressed lysate (black), normoxic (21% O_2) SK-N-AS tumour (red) and hypoxic (1% O_2) SK-N-AS tumour (green). Peptides seen in over-expressed lysate are labelled at their elution time. Chromatograms obtained from Xcalibur. Tumours N = 3.

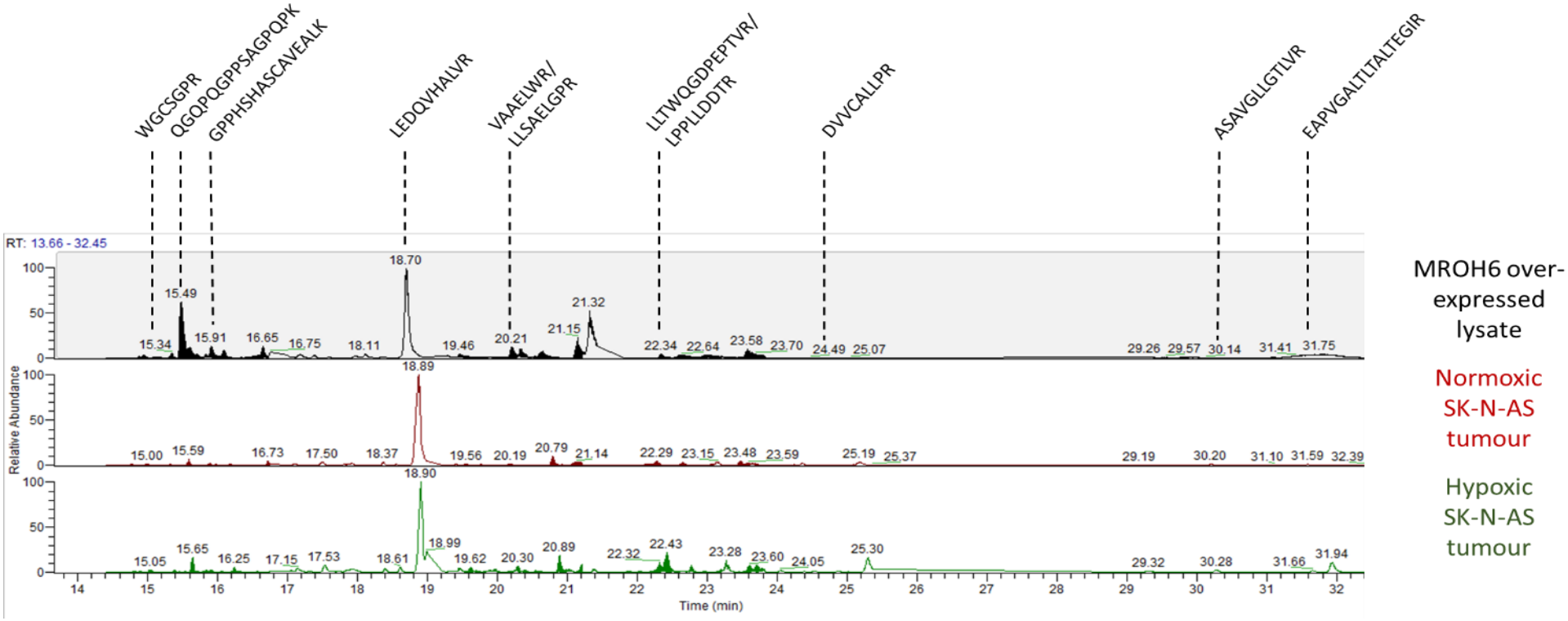

Figure 3.15 Extracted ion chromatogram of MROH6 PRM on 21% and 1% O_2 SK-N-AS tumours.
LC/MS-MS chromatogram results from MROH6 PRM method on MROH6 over-expressed lysate (black), normoxic (21% O_2) SK-N-AS tumour (red) and hypoxic (1% O_2) SK-N-AS tumour (green). Peptides seen in over-expressed lysate are labelled at their elution time. Chromatograms obtained from Xcalibur. Tumours N = 3.

3.4. Discussion

This chapter aimed to develop a clinically applicable assay to determine whether zDHHC23 or MROH6 could be used as biomarkers for NB. Due to a lack of success through the traditional 'gold standard' immunostaining of clinical biomarker screening, PRM LC-MS/MS analysis was investigated. However, the targeted assay also proved unfruitful for detecting endogenous proteins in the tumour model. Hence, I did not proceed further with obtaining valuable patient samples for biomarker research. The limitations of the methods applied in this chapter are discussed below.

3.4.1. Immunostaining

This chapter highlights the necessity for self-validating commercial antibodies. While initial IF staining of SK-N-AS cells and tumours looked promising for confirmation of zDHHC23 and MROH6 proteins as markers of aggressive NB phenotypes, validation of the antibodies using siRNA knockdown and western blotting, or IP and LC-MS/MS analysis revealed the antibodies to be non-specific. While additional antibodies are commercially available, the associated cost of purchasing, optimising and validating these exceeds what is possible within the realm of PhD funding and time restraints. If it was possible, exploring the route of custom antibody production could have been useful. However, it was not possible to generate these with the time and financial constraints of the project.

3.4.2. PRM

Despite advances in MS instrumentation and the increased sensitivity of targeted proteomics over global proteomics analysis, the PRM assay created here was unable to detect endogenous peptides of interest from the two target proteins within complex samples, although they were able to detect exogenously expressed HA-tagged proteins used as a positive control. While further steps could be taken to try and improve the sensitivity of the PRM assay, such as prior off-line fractionation to simplify the mixtures (Horvatovich et al., 2010, Cox and Emili, 2006) assays that transition to the clinic need to be relatively simple and

there is a risk that such an assay would be too complex and beyond the capabilities for application into clinical laboratories due to the resultant increased sample handling and dramatic expansion of the number (thus time) of samples for LC-MS/MS analysis.

3.4.3. Concluding remarks

Unfortunately, I was not able to develop an assay to confirm zDHHC23 and MROH6 proteins as novel biomarkers for aggressive NB. While I attempted to use antibody-based and LC-MS/MS-based techniques to determine potential biomarker applicability, I am unable to make any claims for their use. Due to the increased mRNA expression profiles for these two proteins, and poor patient prognosis seen in patient database Kaplain-Meier plots (Figure 1.7) I decided to continue the investigation into zDHHC23 and MROH6 to better understand their physiological and disease-relevant functions using a biochemical approach.

Chapter 4: Development of a mass spectrometry compatible immunoprecipitation protocol for zDHHC23 and MROH6

4.1. Introduction

In order to determine the physiological function of a protein, it is necessary to be able to confidently identify, and for the purpose of *in-vitro* experiments, purify the protein of interest. This can be achieved using immunoprecipitation (IP), and when there are no suitable protein-specific antibodies (as seen in Chapter 3), IPs can be conducted using an external epitope tag, where a known amino acid sequence is fused to a target protein at the DNA/plasmid level. The HA tag (YPYDVPDYA) derived from the human influenza hemagglutinin protein (Field et al., 1988) was used in this thesis. The same technology for protein tagging can also be used for other functions, such as the case of green-fluorescent protein (GFP) which is useful for cell imaging. The significant discovery of fluorescent protein reporters (Tsien, 1998) has made it possible to directly detect proteins intracellularly in live cells, removing the need for harsh fixing and permeabilising steps (which can potentially confound results), and overcoming antibody specific issues as discussed in Chapter 3. As a result, fluorescence-based reporters continue to be valuable tools for visualizing signalling dynamics in living systems (Grynkiewicz et al., 1985). Since the identification of GFP in 1992, direct mutagenesis has led to a major expansion of fluorescent reporters available, ranging across the whole visible spectrum of wavelengths: cyan/green (400-500 nm), orange/yellow (mBanana, 500-550 nm), red (mCherry, 550-650 nm - used in this thesis), and far red (mRaspberry, 650+ nm). The development of a wide range of excitation and emission spectra allows the use of multiple fluorescent reporters simultaneously to investigate protein-protein interactions (Bansal et al., 2013). These fluorescence reporters are commonly combined with an epitope tag on a gene of interest to allow both analytical and visual experiments in a single construct (Xu et al., 2016). However, a fundamental problem of external epitope tagging is the unknown affects these tags may have on a protein function, as well as the potential implication that exogenous (and elevated) expression may have on endogenous regulatory mechanisms.

4.2. Aims

The aim of this chapter was to create a plasmid with both an epitope tag and fluorescent reporter to allow IP and visualisation of zDHHC23 and MROH6, and optimise a protocol for analysis of the target proteins and binding partners by LC-MS/MS.

4.3. Results

As discussed in Chapter 3, HA tagged plasmids were purchased from NovoPro Labs, and using manufacturer recommended protocols, successful HA antibody IPs were performed allowing PRM analysis (Figure 3.10). To enable binding partner discovery by LC-MS/MS, I optimised this IP protocol, focusing on protein recovery and minimal 'noise' co-purification. When using IP to identify binding partners of a target protein, it is imperative to consider and ultimately minimise the co-purification of non-specific (background/noise) proteins. These are typically unintentionally introduced (environmentally) during sample preparation (serum albumin and keratins), or high cellular abundance and 'sticky' proteins such as actin, myosin and heat shock proteins. Two of these proteins (serum albumin and myosin) are predicted to appear in up to 95% of all IP coupled LC-MS/MS experiments (Trinkle-Mulcahy, 2012). Therefore, optimisation of IP conditions is essential to limit non-specific interactions whilst also maintaining weaker interacting proteins of interest.

4.3.1. HA Bead ratio and NaCl optimisations

The first step of optimising IP of the HA-proteins from cell extracts was to determine the ratio of pre-conjugated HA beads: quantity (µg) protein lysate overexpressing the HA-protein of interest. I defined the optimal volume of HA-beads as the amount required to capture as much HA tagged protein as possible, i.e. no further decrease in banding observed by western blotting in unbound samples in the presence of additional beads. This definition allows maximal protein recovery whilst minimising excess beads that may increase non-specific binding. Therefore, I varied the volume of HA beads added to identical 200 µg quantities of HA-zDHHC23 and HA-MROH6 expression lysate and analysed the unbound material by western blot in comparison with the total lysate (Figure 4.1).

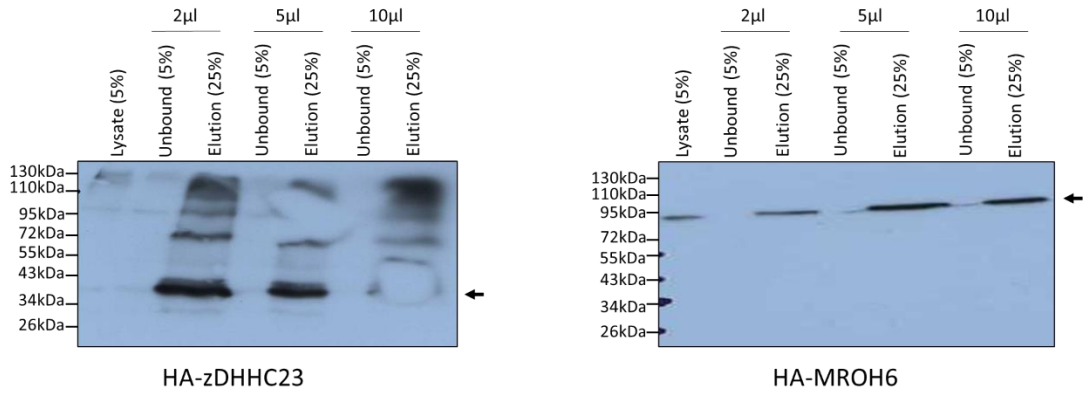

Figure 4.1 Testing pre-complex HA bead binding efficiency.
Western blot analysis with GFP SK-N-AS cell lysate incubated at 21% O_2 for 16 hr. Cells were transfected using optimal PEI:DNA ratio with either HA-zDHHC23 or HA-MROH6 plasmid. Following lysis, IP using pre-complex HA beads (either 2, 5 or 10 µl). The membrane was probed with HA antibody (Cell Signalling, 37245) at a 1:5000 dilution. Arrow identifies potential protein of interest.

Figure 4.1 shows, as before (Figure 3.10), the predominant banding appears at the correct molecular weight of both HA-zDHHC23 and HA-MROH6, with heavier molecular weight banding for HA-zDHHC23 additionally observed. Unfortunately, the HA-zDHHC23 membrane appears to have regions of 'poor transfer' (likely due to bubble formation between gel and membrane) in the 5 µl and 10 µl elution lanes. Importantly, minimal differences were observed for HA-zDHHC23 when comparing IP with 2 µl and 5 µl HA-beads for either the unbound material, or the predominant ~35 kDa elution band (Figure 4.1). As such, a ratio of 2 µl of HA beads per 200 µg HA-zDHHC23 expression cell lysate was used in all subsequent zDHHC23 IP experiments. For HA-MROH6, while no band was observed in any unbound lane, there is a clear increase in banding intensity between the 5 µl and 2 µl elution lanes (Figure 4.1). With no observable difference between the elution lanes of 5 µl and 10 µl bead quantities (Figure 4.1), I opted to use a ratio of 5 µl of HA beads per 200 µg HA-MROH6 expression cell lysate for all subsequent MROH6 IP experiments.

Another crucial optimisation step for IP of potential bindings partners is the concentration of NaCl used in the lysis and wash buffers. Protein-protein binding is achieved through hydrogen bonding and electrostatic interactions (both reliant on charge affects (Sheinerman et al., 2000)). The addition of NaCl into buffers results in free moving charged (negative and positive) ions which disrupt both hydrogen and electrostatic bonding between proteins. Therefore, higher concentration of NaCl prevents weaker non-specific interactions of background proteins with either the protein of interest or the HA-agarose beads. However, NaCl concentration acts as a 'double-edged' sword as all protein-protein interactions are

equally affected (including antibody-target interactions). As a result, weak and transient interactions (as is often the case for enzymes-substrate binding) can be lost, as well as potentially reducing the quantity of target protein recovered. To evaluate the effect of NaCl concentration on target binding and protein co-IP, an equal number of transfected cells were lysed and IP complexes washed in a similar lysis buffer, and analysed for target protein recovery by western blotting (Figure 4.2A) and interactome differences by coomassie staining (Figure 4.2B).

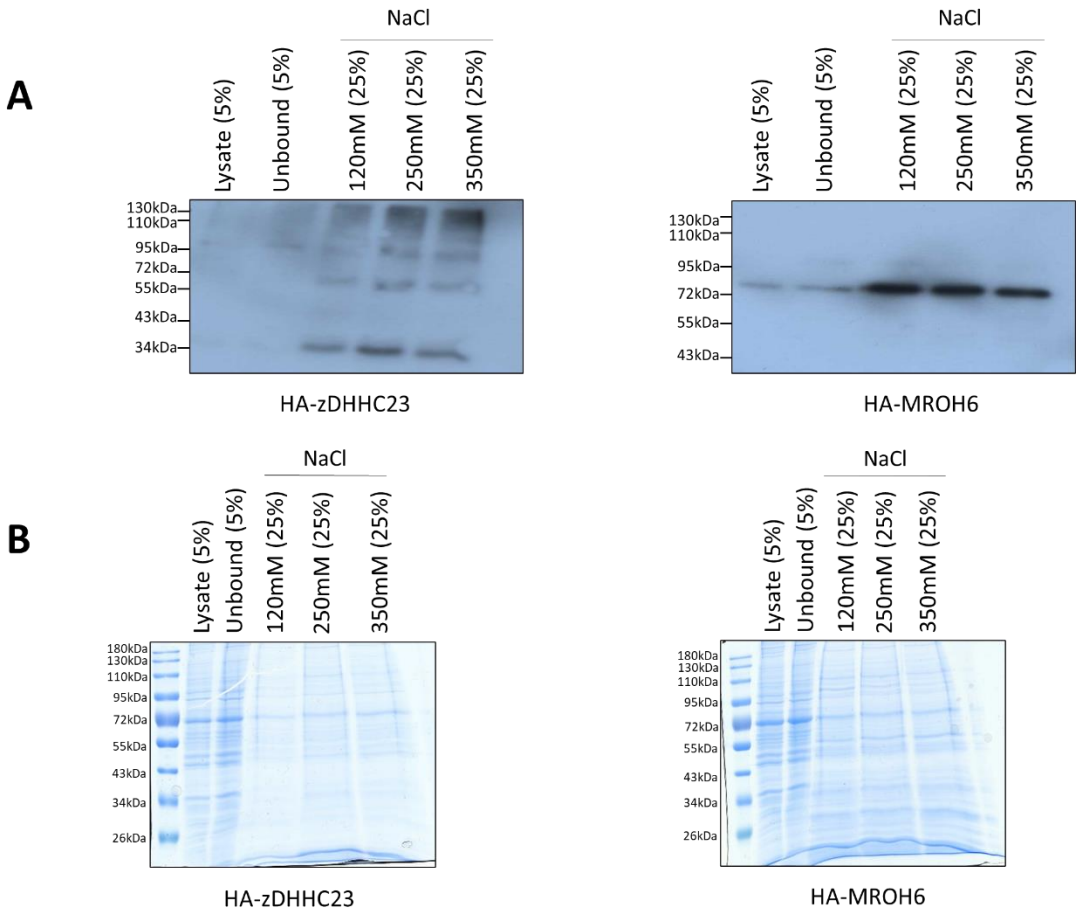

Figure 4.2 Optimisation of NaCl concentration for IP wash steps.
A) Western blot analysis GFP SK-N-AS cell lysate incubated at 21% O_2 for 16 hr. Cells were transfected using optimal PEI:DNA ratio with either HA-zDHHC23 or HA-MROH6 plasmid, probed with HA antibody (Cell Signalling, 37245) at a 1:5000 dilution. B) Colloidal coomassie stained SDS-PAGE gel of IPs using the optimised volume of pre-complex HA beads. IPs were performed following the optimised protocol from the Eyers lab (see Chapter 2). Lysis performed in 5X Laemmli's buffer.

For HA-zDHHC23 I identify minimal differences in the total recovery of target protein (Figure 4.2A) and interactomes co-purified (Figure 4.2B), therefore without significant reasoning I opted to use 250 mM NaCl concentrations for all subsequent zDHHC23 IPs. For HA-MROH6 I

identify small, yet sequential, decreases in total recovery of target protein with increasing NaCl concentration (120 mM vs 250 mM and 250 mM vs 350 mM, Figure 4.2A). However, minimal differences were observed for the total proteins in the eluent indicating relatively small changes to the interactomes co-purified (Figure 4.2B). Therefore, I opted to use 120 mM NaCl for all subsequent MROH6 IPs.

4.3.2. Maximising HA-tagged protein coverage identified by LC-MS/MS through using multiple proteases

An additional opportunity IP coupled LC-MS/MS allows for and is a focus point of this thesis, is the in depth PTM discovery of target proteins zDHHC23 and MROH6. As discussed in Chapter 1, PTMs typical exist as small sub-populations of <1%. As such, comprehensive PTM mapping generally requires multiple different proteases to maximise the identifiable protein coverage. Standard mass spectrometry workflows utilise trypsin for digestion, however these pipelines may not be optimal for all protein sequences depending on to the distribution of Lys (K) and Arg (R) residues, with preventing identification of PTMs in regions of the protein sequence not identified. A method to try and mitigate poor sequence coverage is through the application of multiple proteases during protein digestion. While there is a large complement of proteases, each cleaving at different amino acids or recognition motifs (and theoretically should improve protein coverage identified). The issue of this strategy is scalability; for the addition of one protease to the workflow this requires double the amount of input, beads etc. therefore can result in very large experimental designs. As the Eyers lab has had previous success with this type of optimisation using chymotrypsin (cleaves at C-terminal of F, W, Y, L) and elastase (cleaves at the C-terminal of A, V, S, G, L, I) for the in-depth PTM characterisation of HIF-1α and HIF-2α (Daly et al., 2021), I started by testing these proteases. Additionally, using EXPASY protease cutter tool (Gasteiger et al., 2005), Glu-C (cleaves at C-terminal of D, E) appeared to be a potential good candidate for generating reasonably large peptides (7-15 residues) which suit standard LC-MS/MS workflows.

Miscleavage (failure to cleave an intended site) results in the reduced abundance of any singular peptide ion species, but also needs to be accounted for in search engine parameters thus, dramatically increasing the database search space and potential for false positive identifications and a high FDR. Another caveat to proteases with different specificity is the fact that they may contain fewer (or no) basic residues, reducing their ionisation efficiency (in positive mode) and compromising proton-driven collision-mediated fragmentation.

Additionally, most LC-MS/MS-based proteomics experiments typically filter out singly charged ions given the propensity for these to arise preferentially from contaminants, for example polyethyleneglycol, plasticisers and buffer components (Hodge et al., 2013).

To identify whether the application of multiple proteases would increase sequence coverage (and potentially PTMs identified) of zDHHC23 and MROH6, a large-scale IP (500 µg per protease) was performed, and the eluent divided into equal fractions for differential proteolysis. LC-MS/MS analysis was then used to determine protease-dependent sequence coverage for either zDHHC23 (Figure 4.3 & Table 4.1) or MROH6 (Figure 4.4 & Table 4.2). As different proteases will result in peptides which span the same regions of the protein, a hierarchical analysis was applied to the sequence coverage map of: Trypsin, Chymotrypsin, Glu-C, Elastase (i.e. if a peptide is identified by both trypsin and chymotrypsin it is labelled as trypsin identified, any additional residues observed beyond the trypsin peptide is then labelled as chymotrypsin, (Figure 4.3 & Figure 4.4) Actual protein coverage (without hierarchical analysis), number of peptides specific to that protein, and total number of proteins identified are found in Table 4.1 & Table 4.2.

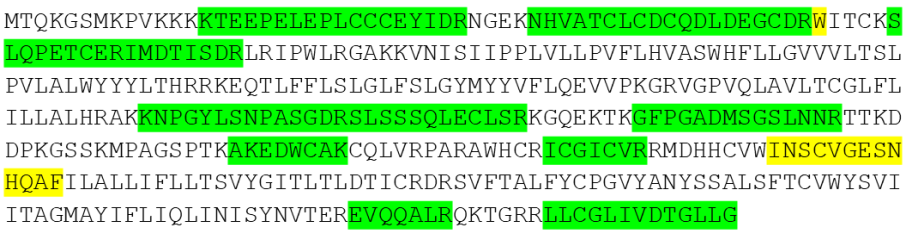

Figure 4.3 Identified protease sequence map of zDHHC23.
zDHHC23 protein sequence, with peptide identifications at a 1% FDR highlighted. Data from the digest stage LC-MS/MS analysis using a High-High MS/MS method, LC gradient 1 hr. Green = Trypsin digested (2 miscleaves), Yellow = Chymotrypsin digested (3 miscleaves).

Table 4.1 Determination of the best protease for maximal sequence coverage for zDHHC23.
Table includes stated sequence coverage as a percentage and subsequent number of unique peptides and the total number of proteins identified for zDHHC23. LC-MS/MS data on a 1 hr gradient was acquired following digestion using Trypsin, Elastase, Chymotrypsin and Glu-C.

Protease	% coverage	# peptides	# proteins
Trypsin	32%	13	754
Elastase	0%	0	75
Chymotrypsin	14%	6	359
Glu-C	0%	0	0

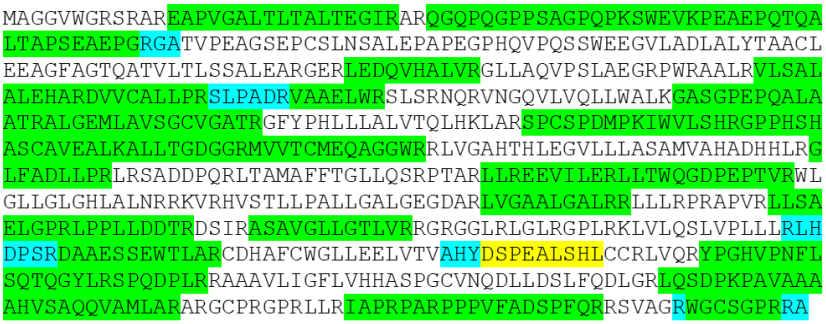

Figure 4.4 Identified protease sequence map of MROH6.
MROH6 protein sequence, with peptide identifications at a 1% FDR highlighted. Data from the digest stage LC-MS/MS analysis using a High-High MS/MS method, LC gradient 1 hr. Green = Trypsin digested (2 miscleaves), Yellow = Chymotrypsin digested (3 miscleaves), Blue = Elastase digested (5 miscleaves).

Table 4.2 Determination of the best protease for maximal sequence coverage for MROH6.
Table includes stated sequence coverage as a percentage and subsequent number of unique peptides and the total number of proteins identified for MROH6. LC-MS/MS data on a 1 hr gradient was acquired following digestion using Trypsin, Elastase, Chymotrypsin and Glu-C.

Protease	% coverage	# peptides	# proteins
Trypsin	47%	32	1032
Elastase	12%	11	371
Chymotrypsin	1%	1	171
Glu – C	0%	0	0

These analyses identify, as may be expected, Trypsin to be superior protease for maximal sequence coverage, identifying 32% and 47% of zDHHC23 and MROH6 respectively (Figure 4.3 & Table 4.1, Figure 4.4 & Table 4.2). While 14% sequence coverage was observed for zDHHC23 with chymotrypsin, this was only a single 13 residue peptide additionally identified to trypsin (Figure 4.3), yielding an improvement in sequence coverage of only ~3% (13 residues/409 residues). Neither elastase nor Glu-C identified any peptides of zDHHC23 post FDR filtering. For MROH6, elastase identified 12% sequence coverage (Table 4.2), equating to only 22 additional residues (Figure 4.4), a ~3% improvement in total sequence coverage. Chymotrypsin identified a single peptide of MROH6, while Glu-C failed to identify any MROH6 peptides. As I did not observe any peptides when using Glu-C for either POI, this would

suggest that this protease was inactive in my digest conditions. Due to the very minimal improvements in total sequence coverage when considering doubling experimental time and cost (for two proteases), I decided to only use trypsin protease treatment for both zDHHC23 and MROH6 IPs.

4.3.3. Creation of HA-mCherry and mCherry-HA constructs

While I originally purchased HA-tagged zDHHC23 and MROH6 constructs, one of the aims of this thesis was to investigate cellular localisation of these proteins by confocal microscopy. As discussed earlier, the discovery of fluorescent reporter proteins allows for easier localisation analysis in live cells (circumventing potential experimental induced variation/artefacts through fixing and permeabilising cells). Therefore, I wanted to incorporate mCherry (a red fluorescent protein) into our plasmids, additional to the HA-tag for IP, generating a multi-functional plasmid. However, as these proteins are unstudied, it is unknown to what degree adding large protein tags may affect the target protein in terms of function (binding partners & PTMs) and localisation. Therefore, I decided to create both N- and C- terminally tagged constructs for both zDHHC23 and MROH6 (HA-mCherry-POI and POI-mCherry-HA) in order to minimise positional effects of an exogenous tag. I chose mCherry as the fluorescent tag due to its excitation/emission wavelengths (587 nm, 610 nm) which minimally overlaps with the stably expressed GFP in the SK-N-AS cells (with an excitation/emission wavelength of 488 nm/510 nm) during confocal microscopy imaging. This large difference in excitation/emission wavelength of fluorescent proteins reduces cross-activation of the GFP signal into the mCherry signal, which would result in false interpretation of protein localisation, i.e. GFP signal 'bleaching' into (appearing) mCherry filters applied, thus interpreted as signal from our mCherry tagged proteins.

I strategized a protocol to develop the four required constructs (HA-mCherry-zDHHC23, zDHHC23-mCherry-HA, and MROH6 equivalents) as well as background subtraction plasmids for confident binding partner identification (HA-mCherry & mCherry-HA) (Figure 4.5-Figure 4.12). In-Fusion cloning technology was used throughout due to its ease, high success rate and use of PCR to generate products with overlapping 5' and 3' base pair regions complementary to the site of insertion on the backbone. Inherent advantages of this method include: no limitations to restriction enzyme sites; the gene of interest is always inserted in the correct orientation and can also be used to incorporate new restriction enzyme sites/amend the reading frame with additional DNA bases in the primer, if necessary. The

first step of this strategy was to create HA-mCherry and mCherry-HA constructs, which later provide the backbone of the final zDHHC23 and MROH6 plasmid constructs. Within the Sée lab a pcDNA3-mCherry (AddGene #128744) construct was available. However, this lacked any N-terminal restriction enzyme sites and therefore would be difficult to create POI-mCherry-HA constructs. Therefore, to create HA-mCherry and mCherry-HA constructs (depicted in Figure 4.5) my strategy used the available HA-Clover construct (AddGene #163366) in the Sée lab, which was created with restriction enzyme sites positioned to remove the HA-Clover gene using HINDIII and NOTI. I then used the pcDNA3-mCherry plasmid as a template for PCR amplification of HA-mCherry or mCherry-HA (HA-tag incorporated into the primer design).

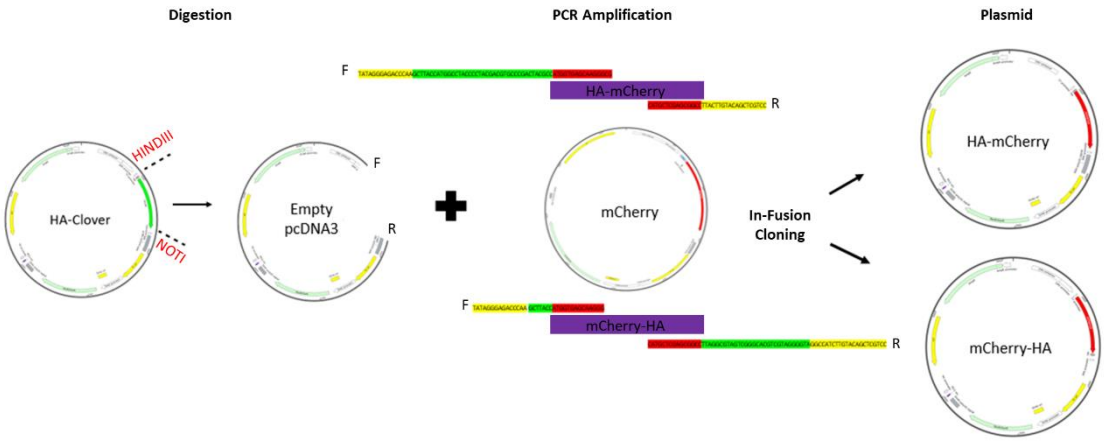

Figure 4.5 Schematic view of the In-Fusion cloning strategy used for the creation of HA-mCherry and mCherry-HA plasmids.
Restriction enzyme sites labelled in red (HindIII and NotI) and forward and reverse primer designs (Table 2.6) for each plasmid are included: Red = gene specific, Green = inserts to correct reading frame/digest sites, Yellow = vector overlapping region. Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com)).

As In-Fusion cloning is reliant on PCR, I first tested different PCR temperatures to ensure I could achieve optimal amplification of the desired PCR product. I additionally tested two commercially available PCR kits: CloneAmp HIFI (ClonTech) and KOD DNA polymerase (Merck), where CloneAmp HIFI appeared to be superior for the amplification of a band at the desired molecular weight of the predicted PCR product across all varying temperatures tested, analysed by a DNA gel (Figure 4.6). The brightest band for both regions appeared to

be in the region of 57.5 ˚C and therefore was used with the CloneAmp HIFI kit for large scale PCR amplification.

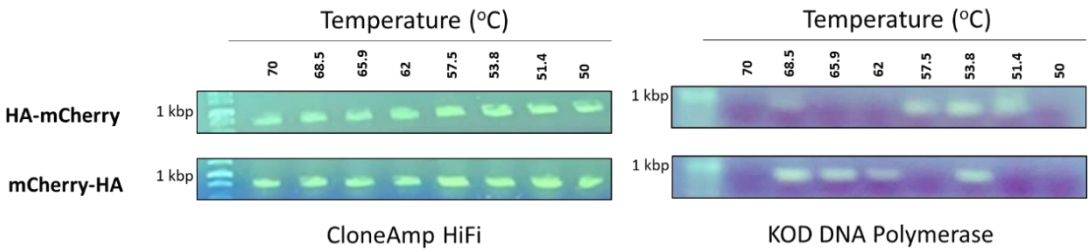

Figure 4.6 Identifying optimal PCR conditions for HA-mCherry and mCherry-HA regions.
A comparison of CloneAmp HiFi (ClonTech) and KOD DNA polymerase (Merck) in PCR temperature gradient test (70-50 °C) for In-Fusion cloning of HA-mCherry and mCherry-HA.

As a relatively quick test to determine if the In-Fusion cloning was successful (completed prior to sequencing), I performed a restriction enzyme digestion test on both plasmids with the DNA fragments ran on a gel and imaged (Figure 4.7). The strategy involved identifying restriction enzyme sites (HINDIII and BmgBI, which cleave in the N-terminal T7 promoter and HA-tag respectively) that would create different banding patterns to differentiate between HA-mCherry and mCherry-HA constructs, depicted and approximate bands produced in Figure 4.7. Restriction enzyme double digests shows that, as predicted, the mCherry-HA and HA-mCherry plasmids are easily differentiated due to the formation of an 800 bp band (mCherry gene) in mCherry-HA lanes and a lack of which in HA-mCherry constructs (predicted 10 bp band) (Figure 4.7). As our constructs contain the mCherry protein, to ensure plasmids were in the correct reading frame I analysed (and confirmed the expression of) them by confocal microscopy for mCherry excitation/emission (Figure 4.8). Therefore, our cloning was likely successful, and confirmed by whole plasmid sequencing with Plasmidsaurus (Section 2.7.9).

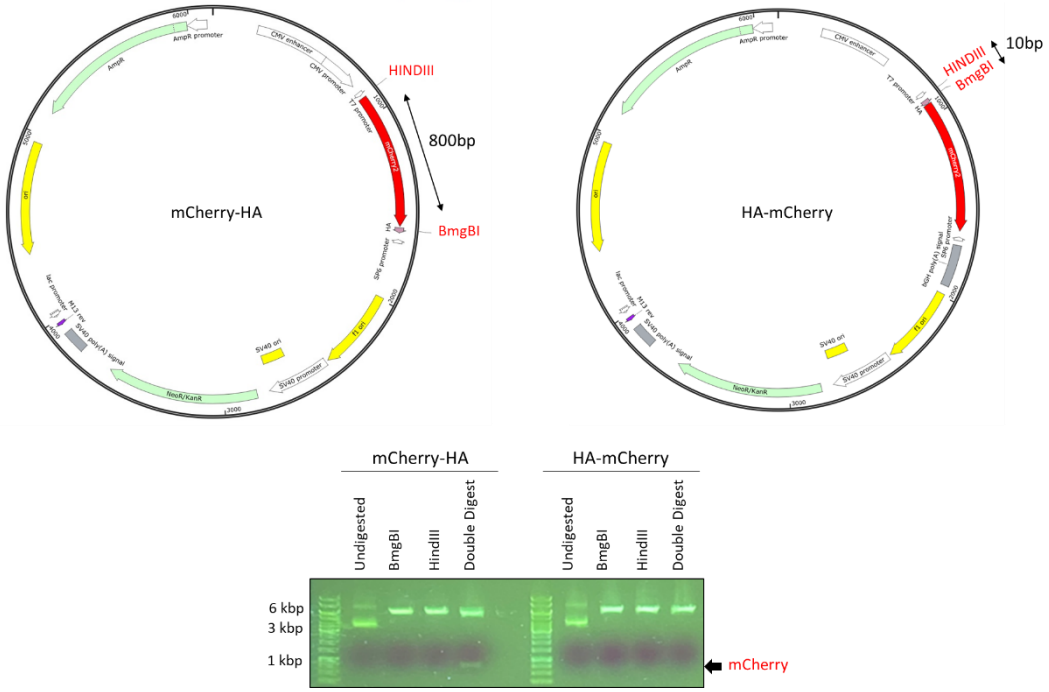

Figure 4.7 Digestion test of mCherry-HA and HA-mCherry plasmids.
Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com)). Restriction enzyme sites labelled (HindIII and BmgBI) and resulted digested fragment size shown. Agarose gel (0.5%) with undigested and digested plasmid (0.5µg) restriction enzymes labelled.

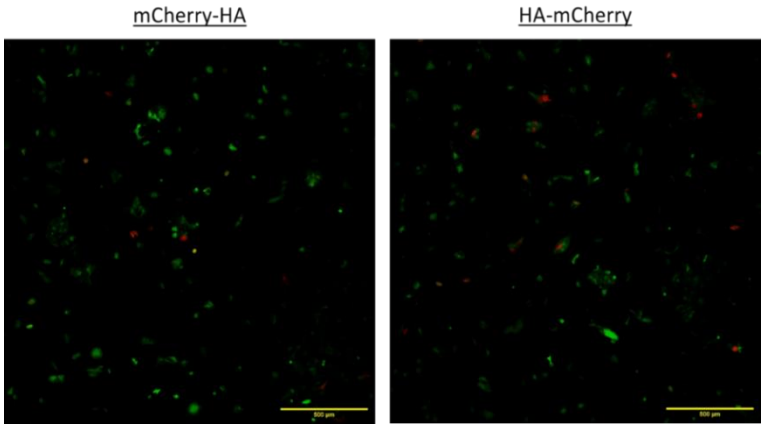

Figure 4.8 Microscopy validation of successful In-Fusion cloning of mCherry-HA and HA-mCherry plasmids.
GFP SK-N-AS cells transfected using PEI with mCherry-HA or HA-mCherry plasmid (2µg) at 21% O_2 incubated for 16 hrs prior to confocal microscopy. Fluorescence images (488nm, 587nm) obtained on LSM780. Scale bar = 100 µm.

4.3.4. Creation of HA-mCherry/mCherry-HA zDHHC23 and MROH6 constructs

Due to the capabilities of In-Fusion cloning, I engineered a single N-terminal HINDIII and C-terminal BsrGI site into both HA-mCherry and mCherry-HA constructs generated. This allowed me to use restriction enzymes and In-Fusion cloning to insert the zDHHC23 and MROH6 genes site specifically into required plasmids with correct primer design (depicted in Figure 4.9).

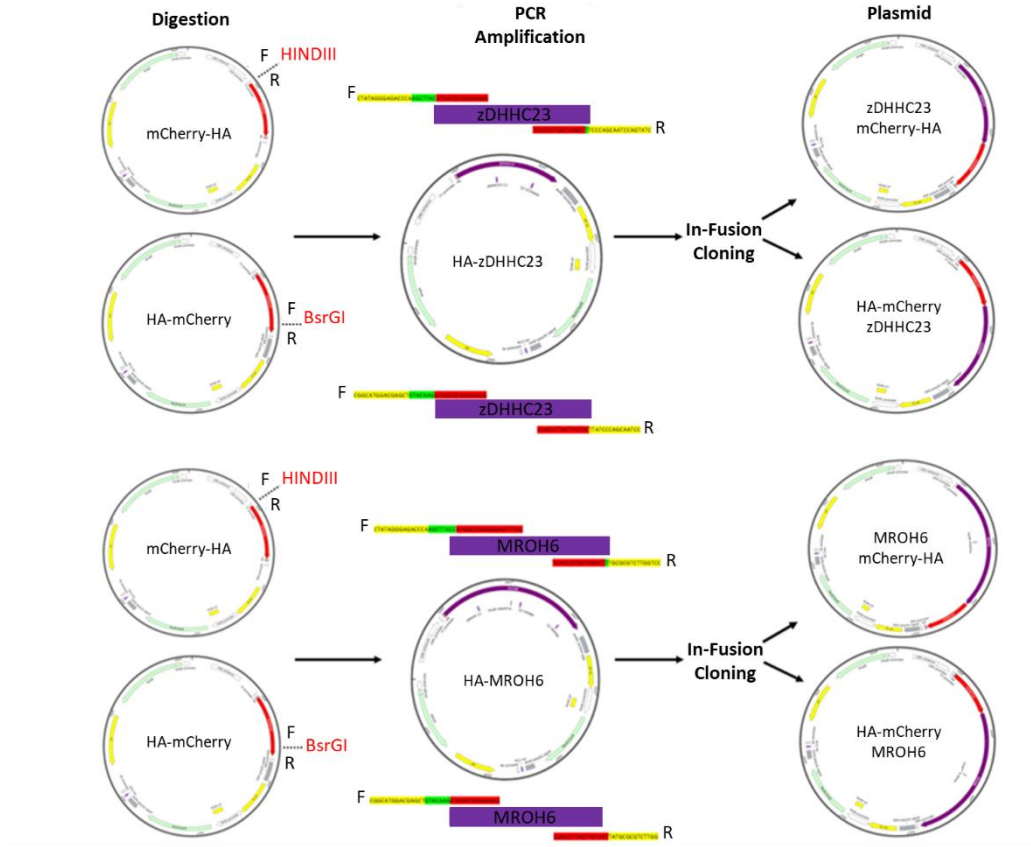

Figure 4.9 Schematic view of the In-Fusion cloning strategy used for the creation of zDHHC23 and MROH6 HA-mCherry and mCherry-HA plasmids.
Restriction enzyme sites labelled in red (HindIII and BsrGI) and forward and reverse primer designs (Table 2.6) for each plasmid are included: Red = gene specific, Green = inserts to correct reading frame/digest sites, Yellow = vector overlapping region. Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com)).

Thus, as with the HA-mCherry and mCherry-HA constructs, I tested both PCR kits at varying temperatures to idenitfy the optimal temperature for amplifiation prior to In-Fusion cloning. As before, CloneAmp HIFI was the superior kit and was much less sensitive to temperature alterations than the KOD DNA polymerase kit (Figure 4.10). With minimal difference in brightness (PCR efficiency) in the CloneAmp HiFi samples across all four primer sets, I opted to perform PCR at 57.5°C for all inserts on a large scale.

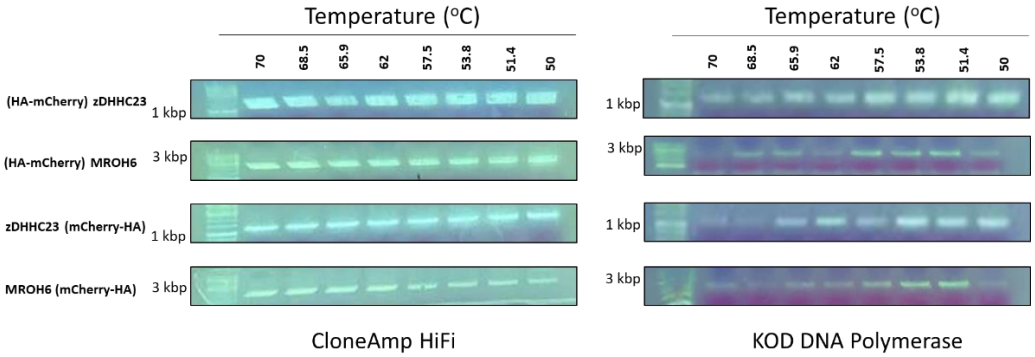

Figure 4.10 Identifying optimal PCR conditions for In-Fusion cloning of zDHHC23 and MROH6.
A comparison of CloneAmp HIFI (ClonTech) and KOD DNA polymerase (Merck) in PCR temperature gradient test (70-50 °C) for all In-Fusion cloning inserts of zDHHC23 and MROH6.

As with the HA-mCherry and mCherry-HA constructs, restriction enzyme digestion confirmed successful cloning for both zDHHC23 constructs (Figure 4.11) and both MROH6 constructs (Figure 4.12), generating bands at the predicted molecular weights for zDHHC23 (HA-mCherry zDHHC23 a single band at ~6 kbp, zDHHC23 mCherry-HA as two bands at ~6 kbp and ~1.2 kbp (Figure 4.11) and MROH6 (HA-mCherry MROH6 as two bands at ~6kbp and 2.3kbp, MROH6 mCherry-HA as three bands at ~6kbp, 1.5 kbp and 580 bp Figure 4.12). Successful cloning, reading frame and tag orientation was confirmed through whole plasmid sequencing.

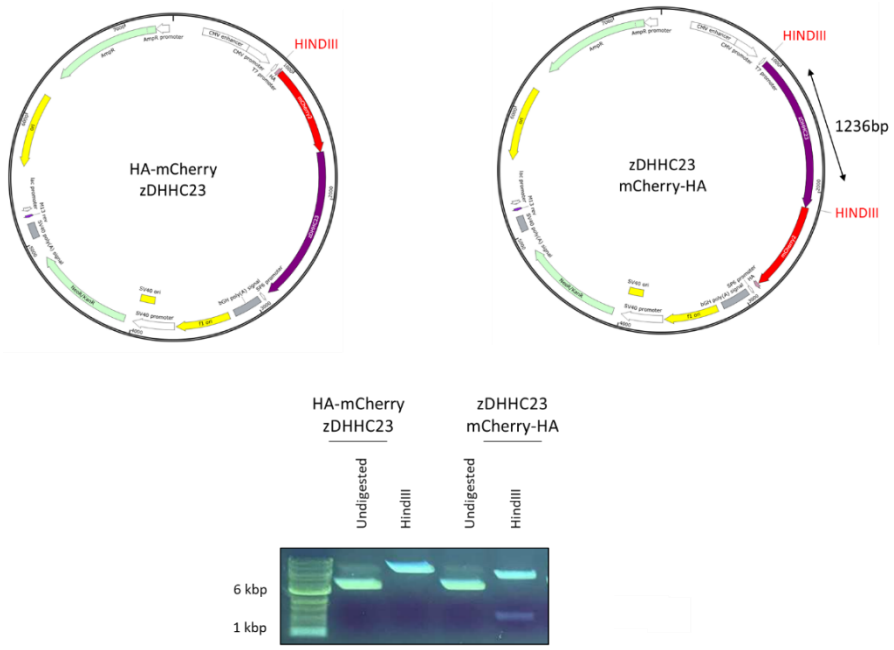

Figure 4.11 Digestion test of mCherry-HA zDHHC23 and HA-mCherry zDHHC23.
Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com)). Restriction enzyme sites labelled (HindIII) and resulted digested fragment size shown. Agarose gel (0.5%) with undigested and digested plasmid (0.5 µg) restriction enzymes labelled.

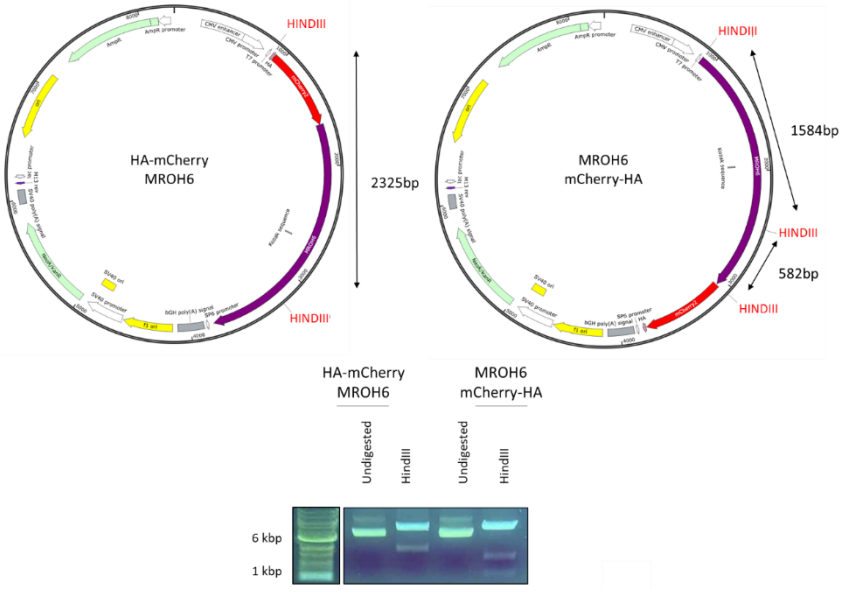

Figure 4.12 Digestion test of mCherry-HA MROH6 and HA-mCherry MROH6.
Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com)). Restriction enzyme sites labelled (HindIII) and resulted digested fragment size shown. Agarose gel (0.5%) with undigested and digested plasmid (0.5µg) restriction enzymes labelled.

4.3.5. Suitability of the HA-mCherry/mCherry-HA tag for IP and MS

As previously discussed, without prior knowledge on how the N-terminal HA-mCherry or C-terminal mCherry-HA tag may affect protein expression and IP efficiency, I tested these new constructs for expression by western blotting, and their ability to be IP'd (Figure 4.13 & Figure 4.14). SK-N-AS cells were transfected as before (during IP optimisation) with either the HA-mCherry or mCherry-HA zDHHC23 or MROH6 constructs and using the optimised IP protocol described above. An HA-tag based IP was conducted and analysed by western blot (Figure 4.13 & Figure 4.14), confirming that addition of the mCherry tag (N- or C- terminal) had minimal effect on IP efficiency for either proteins (Figure 4.13 and Figure 4.14), with minimal protein lost (relatively empty pre-clear bead eluent & unbound lanes) with the majority of protein appearing in the eluent. As expected with the addition of a large fluorescent tag (28 kDa) the bands present in these westerns appear at a higher molecular weight than previously observed (zDHHC23 at ~60 kDa versus ~35 kDa and MROH6 at ~95 kDa versus ~70 kDa). These results also confirm that the optimised volume of pre-complex HA beads and NaCl concentration remain suitable for efficient IP. There is also marginally more protein seen in the unbound material for the C-terminally tagged MROH6 suggesting an orientation specific effect of this tag (Figure 4.14).

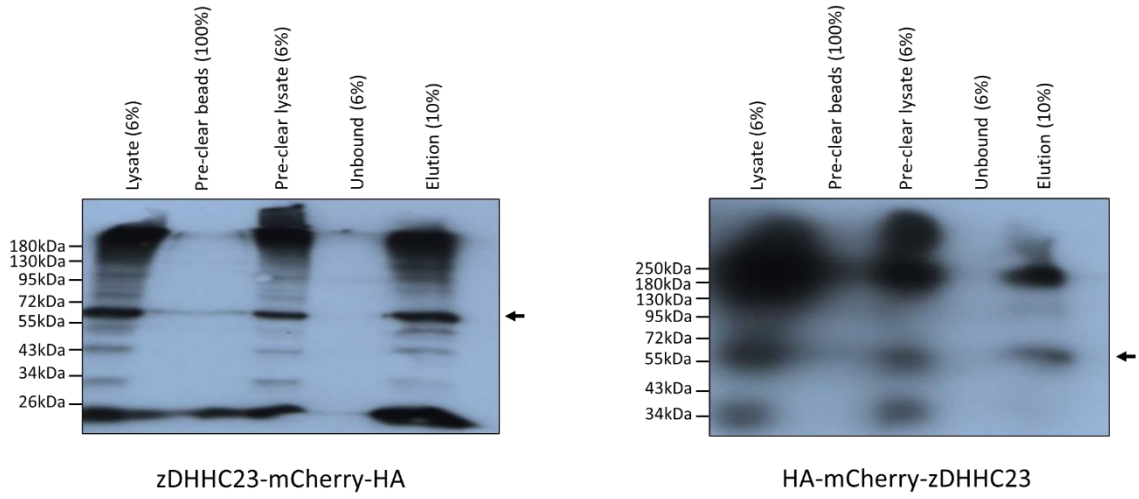

Figure 4.13 Determination of the suitability of the HA-mCherry/mCherry-HA tag for zDHHC23 IP. Western blot analysis of GFP SK-N-AS cells transfected with either the HA-mCherry zDHHC23 or zDHHC23-mCherry-HA plasmid 16 hr in 21% O_2 prior to lysis. The mass spectrometry compatible HA tag protocol was followed and probed with an anti-HA primary antibody. Arrow identifies potential protein of interest.

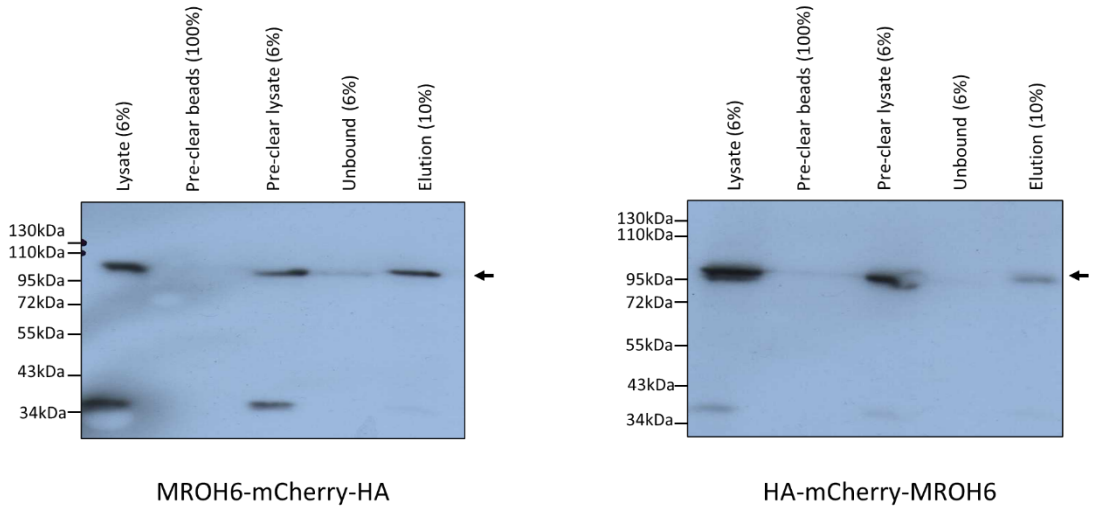

Figure 4.14 Determination of the suitability of the HA-mCherry/mCherry-HA tag for MROH6 IP. Western blot analysis of GFP SK-N-AS cells transfected with either the HA-mCherry MROH6 or MROH6-mCherry-HA plasmid 16 hr in 21% O_2 prior to lysis. The mass spectrometry compatible HA tag protocol was followed and probed with an anti-HA primary antibody. Arrow identifies potential protein of interest. Arrow identifies potential protein of interest.

To confirm the results of the IP western blot experiments, these IP samples were simultaneously processed for LC-MS/MS analysis, which confirmed the presence of zDHHC23 and MROH6 as some highest scoring proteins identified, irrespective of tag orientation.

4.3.6. Optimisation of transfection levels

As stated previously, over-expression of proteins can alter their sub-cellular localisation, regulation or function as a result of differential complex formation and an imbalance in cellular binding partners. The last optimisation step was thus to determine the best quantity of plasmid (and thus protein expression level) to use to minimise the likelihood of identification of artefactual protein interaction networks. Without an available technique to compare endogenous protein of interest expression levels with exogenous expression levels, I decided to investigate expression levels by microscopy, using the least quantity (µg) of transfected DNA sufficient to maintain confocal microscopy detection and reasonable transfection efficiencies while avoiding obvious cellular stress and viability. I therefore transfected 200,000 GFP SK-N-AS cells with varying levels of my plasmids of interest (ranging from 1 µg to 0.2 µg in sequential decreases of 0.2 µg), maintaining the overall 1 µg quantity of DNA per transfection by including empty pcDNA3 vector (Figure 4.15 & Figure 4.16).

At the lowest concentration of 0.2 µg, both N- and C- terminal mCherry constructs and subsequently the zDHHC23 protein could be observed (Figure 4.15). The increase in ratio of construct to empty pcDNA, had little effect on the mCherry signal for HA-mCherry zDHHC23, whilst for zDHHC23 mCherry-HA the amount of specific plasmid transfected correlated with an increase in mCherry signal. Therefore, I decided for all future experiments to transfect 0.2 µg of zDHHC23 plasmid per 200,000 cells. Figure 4.16 shows that for the HA-mCherry-MROH6 construct, mCherry signal intensity increased up to 0.6 µg but decreased for 0.8 and 1 µg. This suggests that >0.6 µg results in levels of protein expression that induces cellular stress. For MROH6-mCherry-HA there was no discernible difference as a function of increasing plasmid concentration. Therefore, I proceeded in future experiments to transfect 0.2 µg for both N- and C- terminal constructs of MROH6.

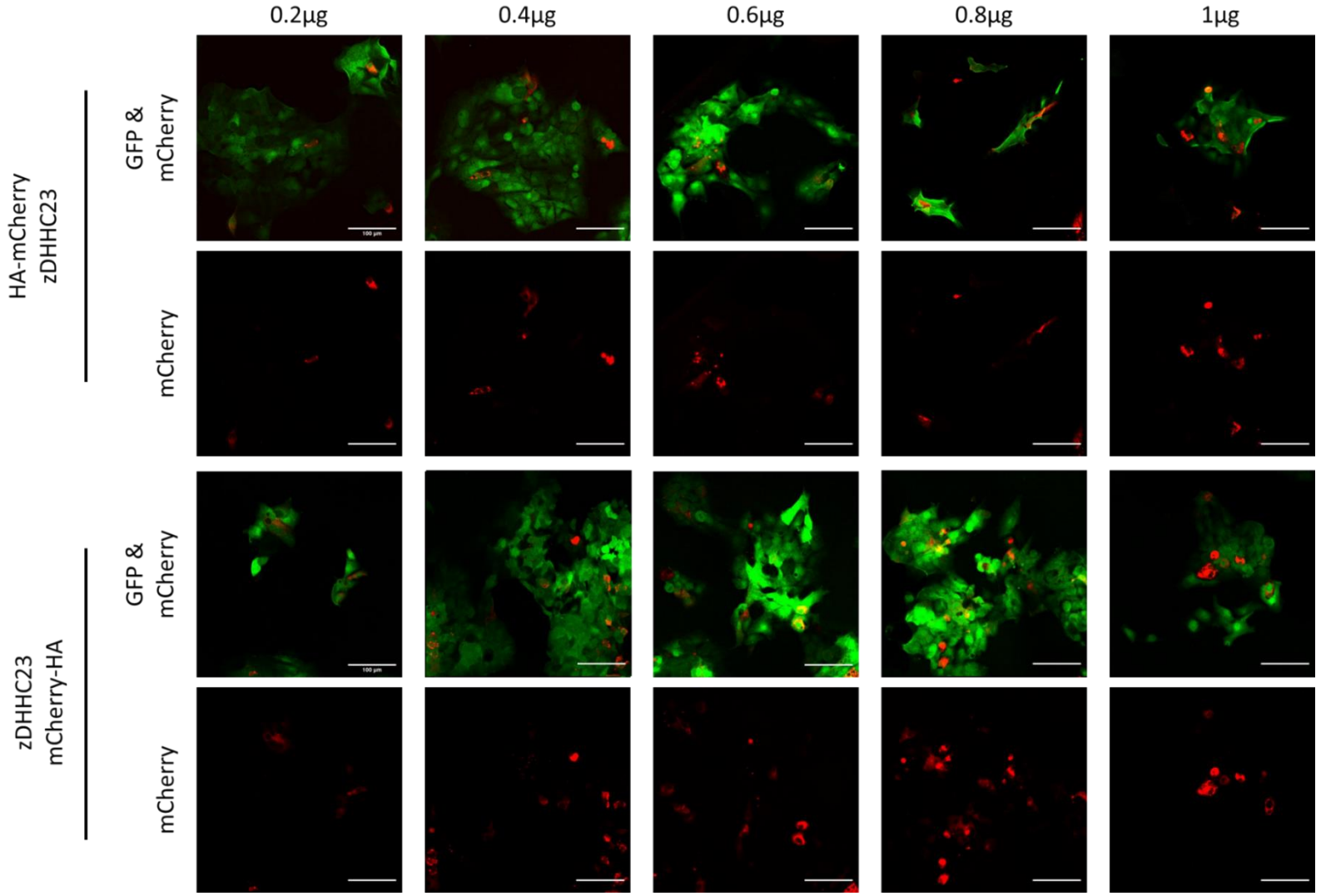

Figure 4.15 Investigation of mCherry-HA and HA-mCherry zDHHC23 plasmid concentration.
GFP SK-N-AS cells transfected using PEI with mCherry-HA or HA-mCherry zDHHC23 plasmid at varying concentrations. Incubated at 21% O_2 for 72 hrs post transfection prior to confocal microscopy. Fluorescence images (488nm, 587nm) 20x magnification obtained on Zeiss LSM780. Scale bar = 100µm.

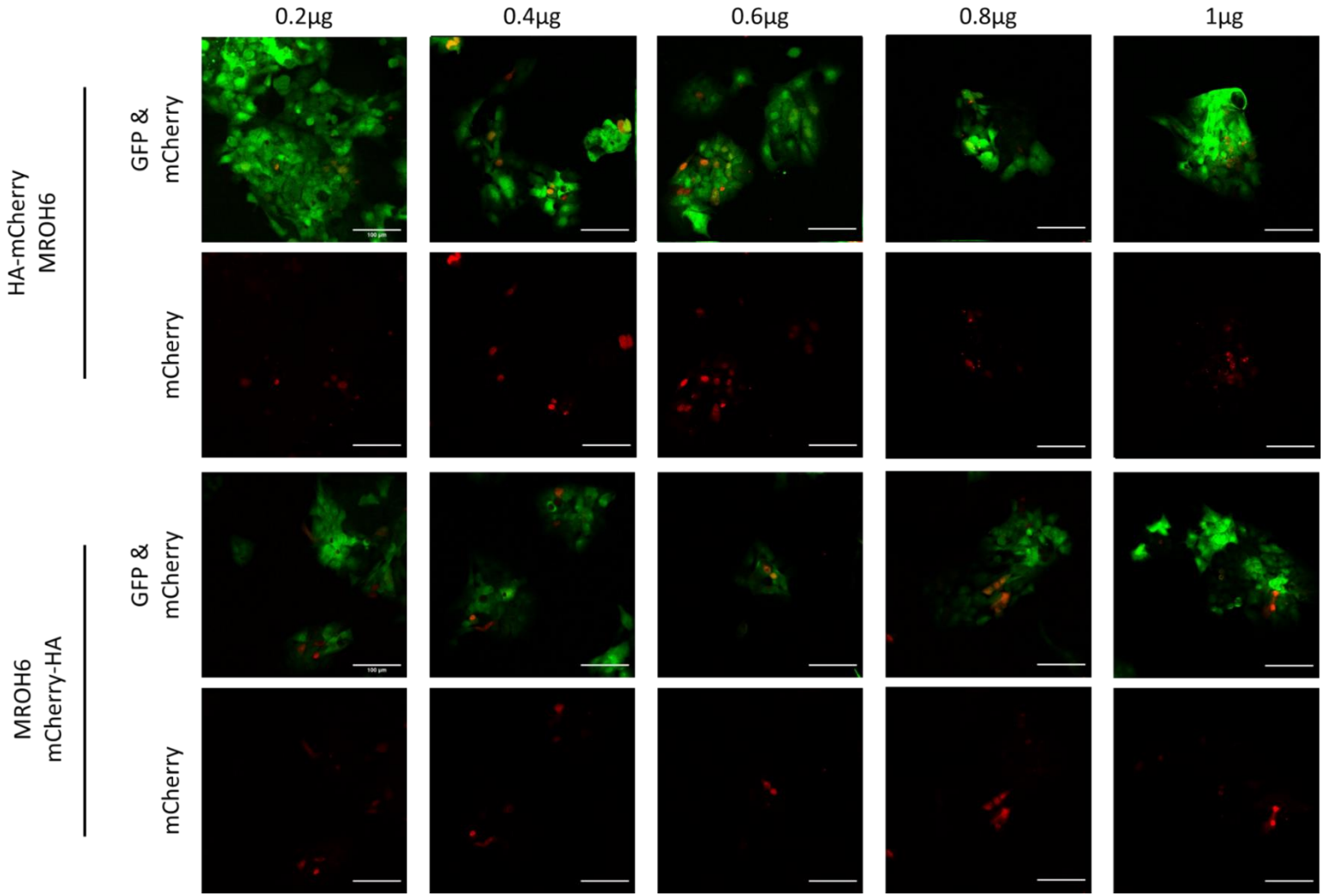

Figure 4.16 Investigation of mCherry-HA and HA-mCherry MROH6 plasmid concentration.
GFP SK-N-AS cells transfected using PEI with mCherry-HA or HA-mCherry MROH6 plasmid at varying concentrations. Incubated at 21% O_2 for 72 hrs post transfection prior to confocal microscopy. Fluorescence images (488nm, 587nm) 20x magnification obtained on Zeiss LSM780. Scale bar = 100µm.

4.4. Discussion

This chapter details the design and cloning of multi-functional plasmids allowing both the purification of zDHHC23 and MROH6 from cell extracts using an HA-tag, and protein visualisation by confocal microscopy with mCherry in a single construct. I have attempted to optimise the expression levels of these plasmids to minimise the unwanted effects of protein over-expression. Unfortunately, given the absence of tools with which to evaluate endogenous proteins levels, no comparator information was available. To mitigate against effects of the presence of covalent tag at one or other end of these two proteins, I generated both N- and C-terminally tagged protein constructs (and the tagged empty vectors) to allow averaging and background subtraction across all conditions. I also optimised IP conditions in terms of volume of beads used and NaCl concentrations to maximise target protein recovery and minimise non-specific binding. I also evaluated the effect of different proteases in an attempt to improve sequence coverage of zDHHC23 and MROH6 and thereby increase PTM mapping, however trypsin yielded the best sequence coverage under these conditions.

4.4.1. Concluding remarks

This chapter describes an optimised exogenous expression-based protocol able to IP full length zDHHC23 and MROH6 from NB SK-N-AS cells, establishing a system for the future identification of binding partners and PTMs by LC-MS/MS analysis.

Chapter 5: Identification of zDHHC23 and MROH6 PTMs and binding partners

5.1. Introduction

Protein-protein interactions are a requirement for a wide range of biological processes, including cell-to-cell interactions, metabolic and developmental control (Braun and Gingras, 2012). It is a rare occurrence that a protein will act as an isolated species to perform its function (Yanagida, 2002), with over 80% of proteins operating as protein complexes (Berggård et al., 2007). Hence, it is an estimated that at any one time ~375,000 protein interactions are present within the human cell (Ramani et al., 2005). Proteins with a large number of interactions generally belong to families of transcription factors, enzymes and disordered proteins to list a few (Dunker et al., 2005, Sarmady et al., 2011).

Protein-protein interactions can be separated into multiple groups, including: homo- (same protein) / hetero- (different proteins) oligomeric, obligate/non-obligate (requirement of a protein interaction) and transient/permanent (length of interaction). All of these protein interactions have a multitude of roles, some common roles include; modifying the kinetic properties of enzymes, supporting substrate channelling, creating a binding site for small molecules, inactivating or suppressing additional regulatory binding proteins, or changing the specificity of a protein's substrate site. The discovery of protein-protein interactions has opened a field for clinical research and drug targeting through identification of drugs that manipulate disease-relevant interactions (Pedamallu and Posfai, 2010). As discussed in Section 1.1, endogenous protein-protein interactions can be regulated by the presence (or absence) of PTMs, a mechanism which can rapidly, and typically reversibly, modify the proteins biochemical properties such as conformation and localised pI (isoelectric point). Consequently, a change in PTM status can alter the availability of docking sites for protein binding partners or other biomolecules *e.g.* DNA, RNA, co-factors (Hunter and Karin, 1992).

High Throughput (HTP) mass spectrometry methodologies are commonly used for the investigation of protein binding partners and PTMs, often following protein immunoprecipitation. Chapter 4 details optimisation of a specific and efficient IP protocol for zDHHC32 and MROH6 which can be used for binding partner analysis, simultaneously permitting investigation of PTM states of these two proteins. However, comprehensive PTM

discovery can be challenging due to: 1) the relatively low stoichiometry of PTMs which confounds detection in complex mixtures; 2) the ability to generate peptides of a suitable size for LC-MS/MS using standard procedures that may prohibit complete protein sequence coverage; 3) the higher lability of covalent PTMs compared to the amide bond during collision-induced dissociation (CID) which can cause uncertainties in site localisation.

Low PTM stoichiometry can be overcome in part by targeted PTM enrichment, as performed for phosphopeptides using TiO_2 in this chapter. While algorithms such as phosphoRS (Taus et al., 2011) and A Score (Han et al., 2011) have been designed to aid in the site specific localisation of PTMs, preferential loss of covalent modifications such as phosphorylation during collision-induced dissociation can be overcome by application of alternative fragmentation strategies: non-vibrational, electron mediated fragmentation such as Electron Transfer Dissociation (ETD) significantly limits the degree of neutral loss observed from phosphopeptides (Syka et al., 2004). However, ETD is slower than collisional fragmentation strategies and is highly inefficient for fragmentation of low charge state peptide ions (<3+) (Good et al., 2007), hence is largely unsuitable for analysis of the predominantly doubly protonated tryptic peptide ions.

To shed some light on the potential cellular functions of zDHHC23 and MROH6 (neither proteins role in NB is currently known) and their differential regulation in patients with poor prognosis, here I employ a Co-IP coupled LC-MS/MS analysis pipeline to explore the interaction networks and PTM-mediated changes under normoxic and hypoxic conditions.

5.2. Aims

Using the HA-mCherry-POI and POI-mCherry-HA tagged constructs and optimised methods described in Chapter 4, I performed Co-IP coupled LC-MS/MS based discovery experiments. The aim of this chapter can be split in three: 1) Identify PTMs that occur in varying O_2 tensions for zDHHC23 and MROH6, applying evolutionary and cancer database searching strategies to hypothesis as to biological function; 2) Identify binding partners of zDHHC23 and MROH6 in response to either normoxic (21% O_2) or hypoxic (1% O_2) incubation to explore (changes in) signalling networks; 3) Identify the cellular localisation of zDHHC23 and MROH6 in response to 21% and 1% O_2.

5.3. Results

5.3.1. Phosphorylation site analysis

As discussed in Chapter 4, the application of different proteases failed to improve the sequence coverage for zDHHC23 and MROH6 beyond the ~45-50% obtained with trypsin (Figure 5.1). Using TiO_2 enrichment for phosphopeptide enrichment prior to LC-MS/MS, I was able to confidently and repeatably identify three O_2-independent phosphorylation sites for zDHHC23 (S206, S232 & S252) and one O_2-independent site for MROH6 (S40) (Table 5.1, Figure 5.1 & Appendix 3). While zDHHC23 S232 and MROH6 S40 was only observed in a single replicate in only one construct, we argue that a single replicate is an average of millions of cells and these sites may be of lower abundance, thus scaling up would identify these sites in more replicates. Hence, I maintain these phosphorylation sites for further analysis, although I do acknowledge the requirement for validation before continued biological investigation, as such these are depicted with blue lines in Figure 5.1.

Table 5.1 Characteristics of phosphorylation sites identified for zDHHC23 and MROH6.
Table includes site of phosphorylation, peptide sequence analysed, best ptmRS score, MASCOT score, the O_2 tension the PhosphoSite was observed in (21% O_2, 1% O_2 or both) and which plasmid construct the PTM was observed in (N or C terminal construct or both).

Site	Peptide	PtmRS Score (Best)	MASCOT score	O_2 observed in	Plasmid observed in
S206	SLSSSQLECLSR	100	85	Both	Both
S232	GFPGADMSGSLNNR	99.7	33	Both	zDHHC23 mCherry-HA
S252	MPAGSPTKAKEDWCAK	100	72	Both	Both

Site	Peptide	PtmRS Score (Best)	MASCOT score	O_2 observed in	Plasmid observed in
S40	QGQPQGPPSAGPQPK	100	47	Both	HA-mCherry MROH6

The single MROH6 phosphorylation site was specifically identified in N-terminally tagged constructs, thus suggesting that tag orientation may have an important role on MROH6 regulation/function (Figure 5.1, Table 5.1).

To determine novelty of these phosphorylation sites identified, I compared my data to previous HTP-large scale global phospho-proteomics studies with PhosphoSitePlus, an online repository of these experiments (Hornbeck et al., 2004). This analysis identified that

zDHHC23 phosphorylation sites S206 and S252 had been previously identified in four and five HTP (although lack any LTP) studies respectively, supporting my findings. There was no reported evidence of observation of zDHHC23 S232 and MROH6 S40. Another feature of PhosphoSitePlus is the kinase prediction tool, which takes sequence motifs from known kinase targets and analyses the surrounding PTM sequence to rank the likelihood of an individual site to be catalysed by a kinase versus the whole proteome (Johnson et al., 2023) therefore all four sites I identified were analysed for putative regulatory kinase (Table 5.2 and Appendix 4).

Table 5.2 Predicted top ranked kinases responsible for identified phosphorylation PTMs.
Table includes site of PTM, predicated kinase responsible for phosphorylation and the kinase family it belongs to, the log2score which predicts how accurate the motif prediction is to the sequence and site percentile showing how likely the kinase would interact with the PTM site in comparison to all other sites in the sequence.

Site	Kinase	Kinase family	Log2(score)	Site percentile
S206	MARK2	CAMK	2.884	97.91%
S232	ULK2	Other	2.531	99.87%
S252	CDK3	CMGC	9.046	99.87%

Site	Kinase	Kinase family	Log2(score)	Site percentile
S40	ERK5	CMGC	3.986	99.22%

Kinase prediction suggests that zDHHC23 S206 phosphorylation is catalysed by a protein kinase from the CAMK family (10 top kinases are all CAMK members), with MARK2 being the highest scoring kinase. Similarly, zDHHC23 S252 phosphorylation is predicted to be a substrate of a member of the CMGC kinase family, with 9 out of 10 of the top kinases from the CMGC family, where CDK3 is the highest scoring kinase. Interestingly, 8 out of 9 of these CMGC kinases are cyclin-dependent kinases (CDKs) possibly suggesting important cell cycle roles. The kinase family catalysing zDHHC23 S232 phosphorylation is less clear with four different kinase families identified (Other, TKL, STE & PIKK), with ULK2 (closely followed by ULK1) classified from the 'Other' family the highest scoring kinases. Similarly, the kinase family responsible for MROH6 S40 phosphorylation is unclear, with four families identified in the top 10 kinases (CMGC, TKL, CAMK & Other) with ERK5 the highest scoring kinase.

Using multiple sequence analysis and phylogeny analysis to interrogate the evolutionary conservation of phosphorylation is commonly used (Macek et al., 2008). I therefore performed this specifically focusing on the most common in-vivo models used in research (Figure 5.2 & Figure 5.3). Evolutionary analysis identifies that zDHHC23 S206 is highly conserved, only missing in chicken, which may be due to alignment or sequencing errors, while conservation is lower for S232 and S252. Interestingly, zDHHC23 S252 is specifically identified as a Glu variant in zebrafish and chicken (most divergent species to humans). Glu as an alternative to a putative site of phosphorylation is interesting from a biochemical perspective given the net negative charge of the side chain under physiological conditions. Therefore, this could be potentially highlighting an evolutionary developed activation mechanism, further exciting by its likely CDK mediated phosphorylation. Evolutionary analysis of MROH6 S40 identifies the surrounding sequence is highly conserved, while a specific Pro variant is observed in rabbits. It is difficult to draw conclusion around this considering potential sequencing and alignment errors with only one species exhibiting this variation, however, it is important to note Pro has unique folding properties and is an unusual mutation.

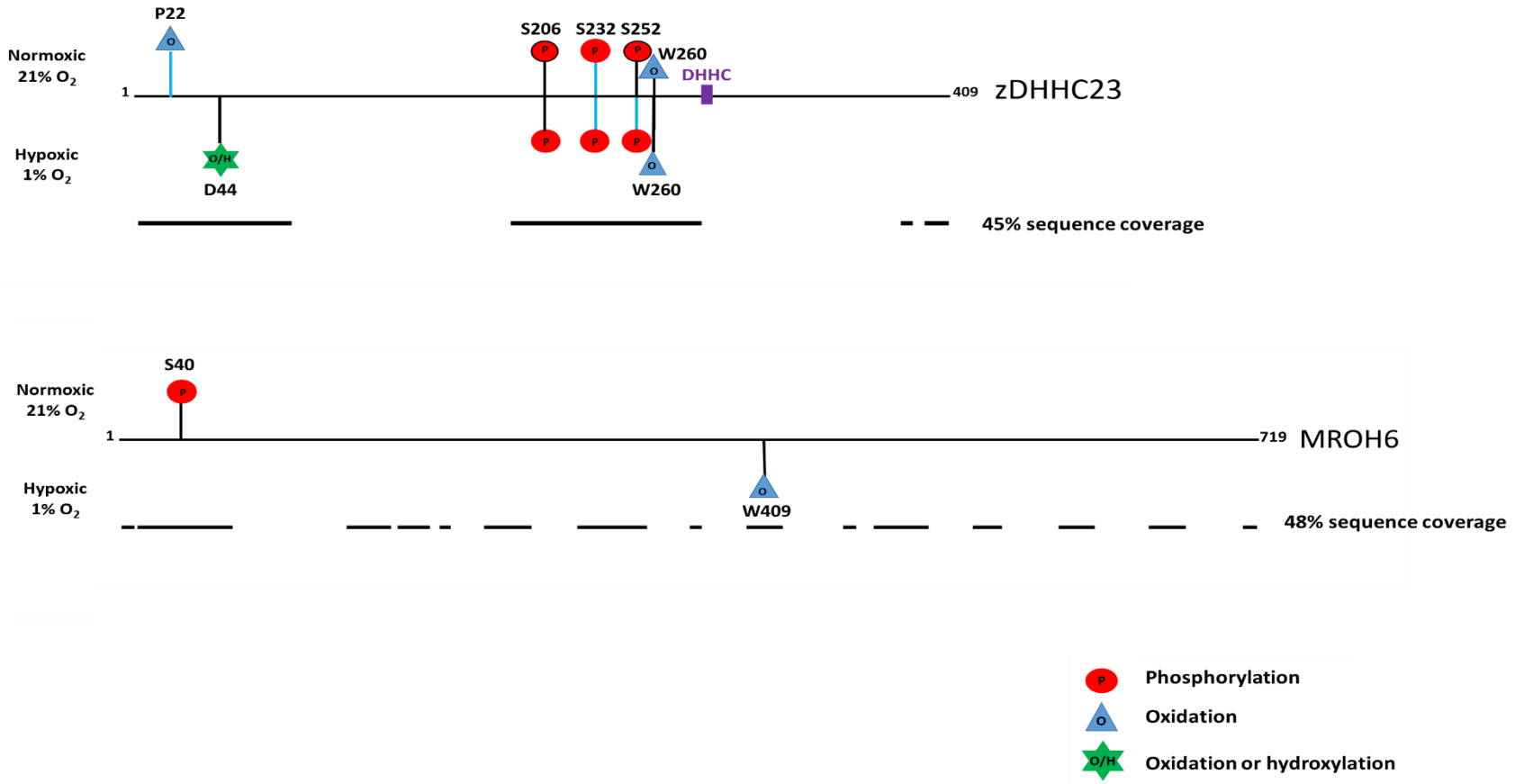

Figure 5.1 Schematic view of the PTM status of zDHHC23 and MROH6.
Confidently identified PTMs (phosphorylation 1% FDR, observed in two of three replicates using Proteome Discoverer 2.4 with oxidation/hydroxylation 1% FDR, observed in two of three replicates and a peptide confidence score of –Log10p over 30 using PEAKS PTM) are mapped to the protein and colour dependent on modification. The sequence coverage seen from mass spectrometry analysis is depicted below the protein schematic. PTMs with a black outline identify modifications observed at least four times in PhosphoSitePlus. PTMs labelled with a blue line were only seen in one out of three replicates. DHHC domain of zDHHC23 is highlighted.

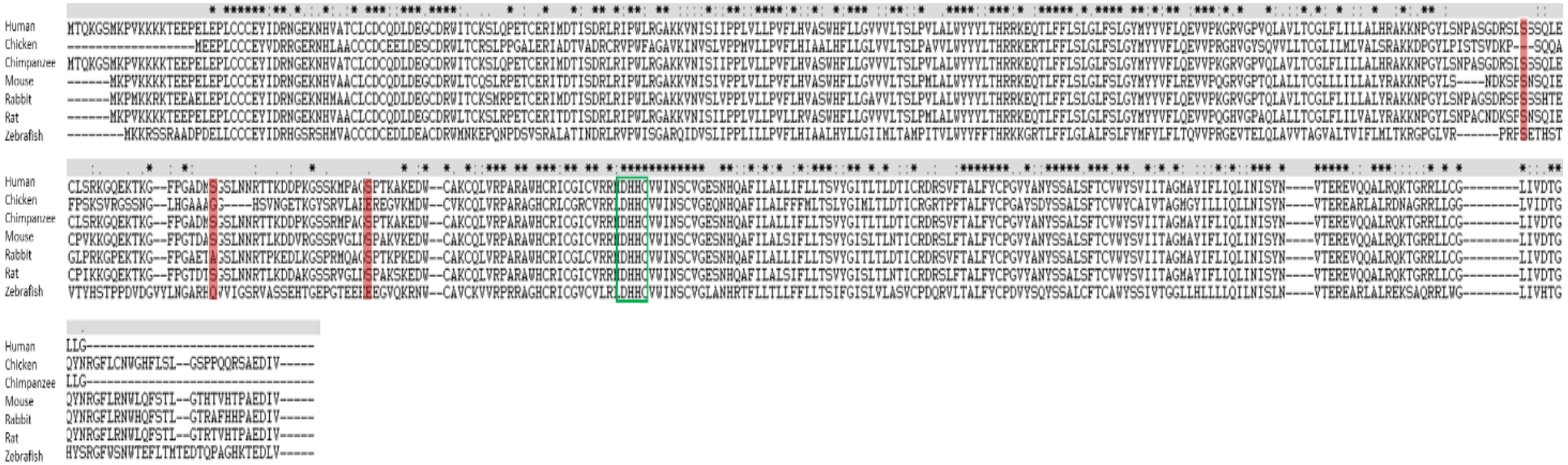

Figure 5.2 Phosphorylated residue sites observed in zDHHC23.

*Shows the aligned zDHHC23 sequences from common in vivo research models in MUSCLE multiple sequence alignment and viewed in Clustal X. * indicates identical residues at a position, : identifies a strong conservation mutation, . identifies a weak conservation mutation and blank spaces indicate no conservation mutation (determined by Clustal matrix). Red boxes highlight the phosphorylation sites with the green outlined region highlighting the DHHC domain.*

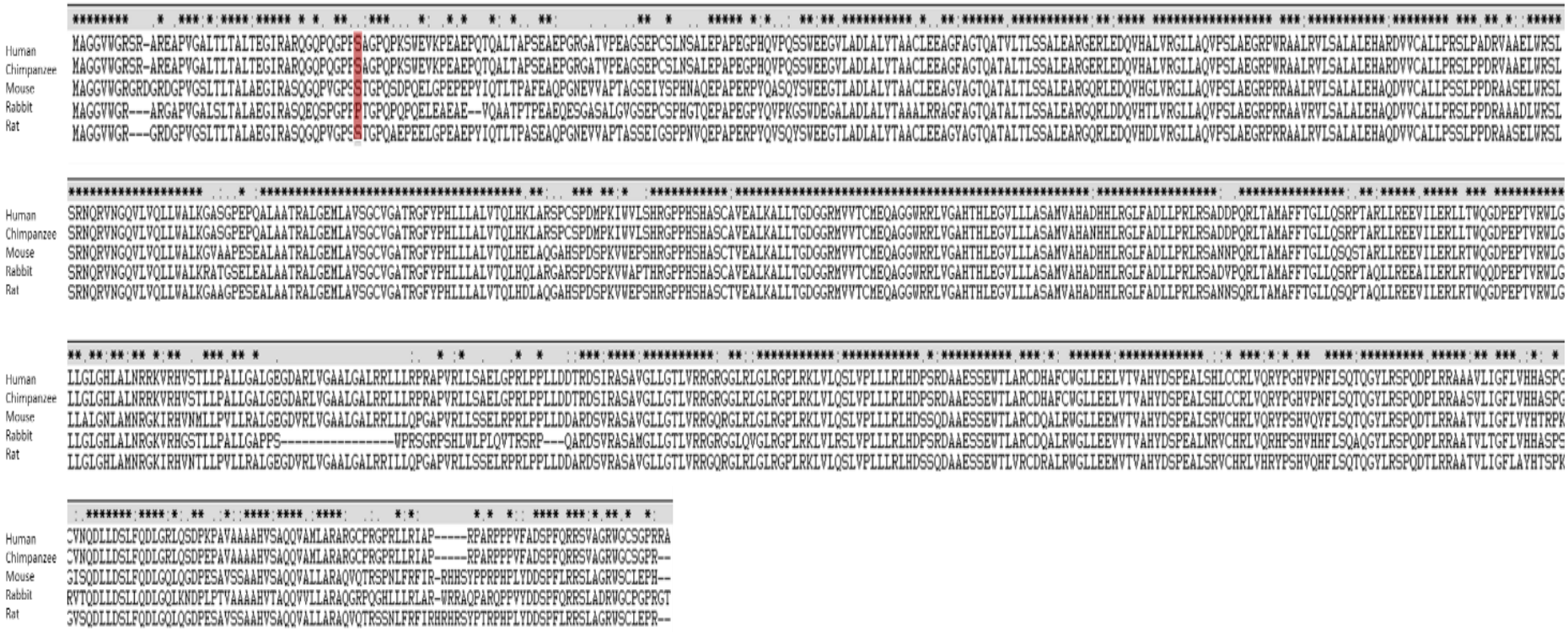

Figure 5.3 Phosphorylated residue sites observed in MROH6.

*Shows the aligned MROH6 sequences from common in vivo research models in MUSCLE multiple sequence alignment and viewed in Clustal X. * indicates identical residues at a position, : identifies a strong conservation mutation, . identifies a weak conservation mutation and blank spaces indicate no conservation mutation (determined by Clustal matrix). Red box highlights the phosphorylation site.*

As there is no crystal structure available for zDHHC23 or MROH6, I used AlphaFold to predict/model the structures of zDHHC23 and MROH6 based on their sequences (Bagdonas et al., 2021). Using PyMOL, a structure visualisation tool (DeLano, 2002) and a specifically designed plug-in for mapping different PTMs to the structure (PyTMsP, (Warnecke et al., 2014), I attempted structure based, functional predictions that identified phosphorylation sites may play (Figure 5.4 & Figure 5.5).

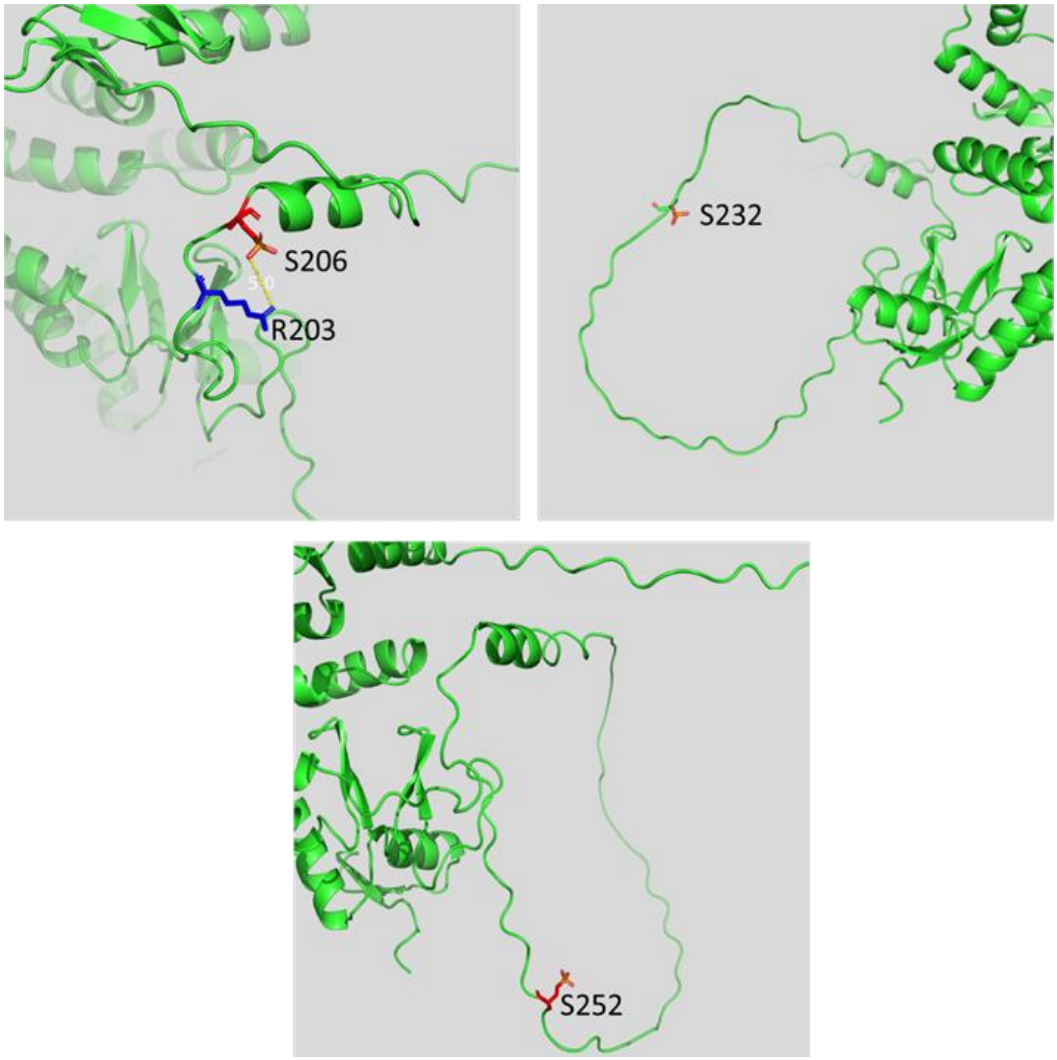

Figure 5.4 Predicted structure of zDHHC23 highlighting phosphorylation sites S206, S232 and S252. AlphaFold structure prediction modelled in PyMOL (DeLano, 2002) and phosphorylated using the PyTMs plugin (Warnecke et al., 2014). Green: zDHHC23, Red: phospho S206/S232/S252, Blue: possible S206 interacting residue R203. Distance measured using PyMOL distance wizard.

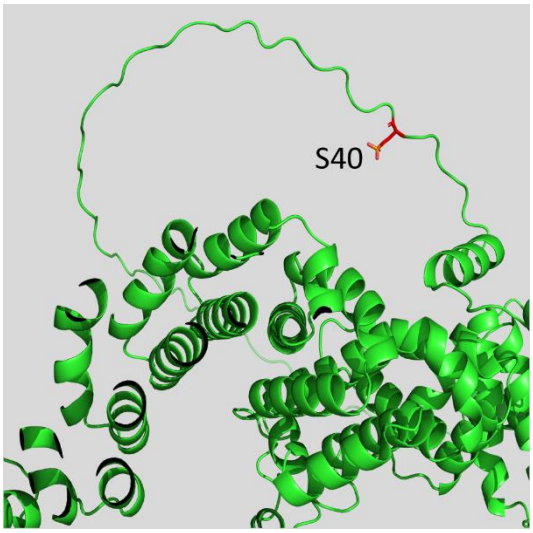

Figure 5.5 Predicted structure of MROH6 highlighting phosphorylation site S40.
AlphaFold structure prediction modelled in PyMOL (DeLano, 2002) and phosphorylated using the PyTMs plugin (Warnecke et al., 2014) Green: MROH6, Red: phospho S40.

Structure predictions suggest that zDHHC23 S206 phosphorylation is very close (5 Å apart) to R203 (Arg). The proximity of S206 and R203 may result in a strong electrostatic bond (salt bridge), strengthening the predicted structure in Figure 5.4, a possible indication that this PTM may be involved in stabilising protein conformation. Interestingly, zDHHC23 S252 and S232 are localised in a disordered loop region (generally these loop regions are known for important regulatory roles due to high flexibility and ease of binding partner 'docking'/association) that would face the cell membrane when anchored through the C-terminal transmembrane domain motif (Figure 5.4).

MROH6 is predicted to be a highly α-helical structure (a known property of HEAT repeat properties), therefore it is interesting that S40 is located in the only disordered loop region of the model (Figure 5.5). Due to the large loop that S40 is located in, it is hard to speculate on likely function. However, we can hypothesis that phosphorylation of this flexible region at S40 site serves to regulate docking of protein binding partners to MROH6.

5.3.2. Search for other PTMs (Open PTM searching)

In addition to phosphorylation analysis following phosphopeptide enrichment, I performed an open PTM search (a peptide spectral matching approach that relies on the highly specific monoisotopic mass shift that a PTM incurs) on all LC-MS/MS data following IP of either

MROH6 or zDHHC23 (Figure 5.1). For open PTM searching I used PEAKS (Zhang et al., 2012), a software tool that can search for over 300 different PTMs (biological and artefactual). Interestingly, PEAKS open PTM searching identified two types of oxygen-based PTM, oxidation and hydroxylation for both zDHHC23 and MROH6 (Figure 5.1) (Table 5.3).

Table 5.3 Characteristics of PTMs identified for zDHHC23 and MROH6.
Table includes site of PTM, peptide sequence analysed, Log10p, A score, the O_2 tension the PTM was observed in (21% O_2, 1% O_2 or both) and which plasmid construct the PTM was observed in (N or C terminal construct or both).

Site	Peptide	Log10p	A Score	O_2 observed in	Plasmid observed in
P22	KTEEPELEPLCCC	32.60	23.13	21%	HA-mCherry zDHHC23
D44	NHVATCLCDCQDLDEGCDR	130.93	38.16	1%	Both
W260	AKEDWCAK	34.14	1000	Both	Both

Site	Peptide	Log10p	A Score	O_2 observed in	Plasmid observed in
W409	LLTWQGDPEPTVR	34.17	1000	1%	HA-mCherry MROH6

5.3.2.1. Oxidative modifications

Oxidation (and the isobaric hydroxylation) modifications occur due to the addition of a singular oxygen atom to an amino acid side chain. This modification can occur on the following amino acids: methionine, phenylalanine, proline, tryptophan, histidine and cysteine. While these can occur artefactually due to typical LC-MS/MS processing techniques, these oxidative PTMs can also occur by either an enzymatic reaction or interaction with reactive oxygen species (ROS) (Sharma and Graham, 2010). For both zDHHC23 and MROH6 I observe tryptophan oxidation (W260 for zDHHC23, W409 for MROH6, Figure 5.1). Deciphering whether oxidative PTMs are artefactual or biological is difficult due to the propensity to spontaneously form during cellular lysis and probe sonication protocols. However, tryptophan oxidation is observed much less frequently than methionine oxidation (a known LC-MS/MS preparation inducible artefact) (Perdivara et al., 2010), due to the need for much stronger oxidising conditions and is thus seldom observed in global proteomics-based studies. The role of tryptophan oxidation is controversial, however, its importance is highlighted by the dysregulation of tryptophan oxidative enzymes being associated with Amyotrophic lateral sclerosis (Zhang et al., 2004) and the development of cataracts (Hains and Truscott, 2007). Aspartic acid hydroxylation is additionally observed on zDHHC23 at D44 (Figure 5.1). Hydroxylation of Asp residues results in 3-hydroxyaspartic acid and is thought to

have roles in Ca^{2+} binding (Sunnerhagen et al., 1993). Of particular interest considering the O_2 tension dependent basis of these experiments, although only confidently identified in one replicate at 21% O_2 (Figure 5.1, Appendix 3), is proline hydroxylation. Proline hydroxylation is currently only characterised for HIFα (a known hypoxia regulated protein) and results in rapid degradation of the protein (Ivan et al., 2001) to the point that a study investigating HIFα PTMs failed to identify proline hydroxylation as a PTM and argued its rapid degradation lowered its available abundance below MS-detectable levels (Daly et al., 2021). Therefore, considering the basis of these experiments of zDHHC23 over-expression linked to hypoxia, it is not unforeseeable that there may exist similar degradation pathways.

5.3.3. COSMIC Database

Advancements in genome sequencing technology has resulted in the production of vast quantities and high quality of cancer genomes sequenced. As a consequence, a wealth of data has been obtianed from various cancer samples which have undergone sequencing, allowing interrogation to identify gene specific mutational patterns which can result in an amino acid mutation (missense mutation) aiding in cancer progression and survivability. Web-based portals have been created for the ease of searching, aligning and annotating genes of interest. The COSMIC database (Catalogue Of Somatic Mutations In Cancer (Forbes et al., 2006) is the largest of genome sequencing repositories, with approximately six million coding mutations identified across all genes in 1.4 million tumour samples (Tate et al., 2019). Therefore, I used the COSMIC database to identify observed mutations in zDHHC23 and MROH6, specifically searching for missense mutations with the aim of identifying any mutational hotspots that may align with the PTMs I have identified (Figure 5.6 and Figure 5.7). I also searched COSMIC for additional phosphosites reportes in HTP studies on PhosphoSitePlus (Hornbeck et al., 2004).

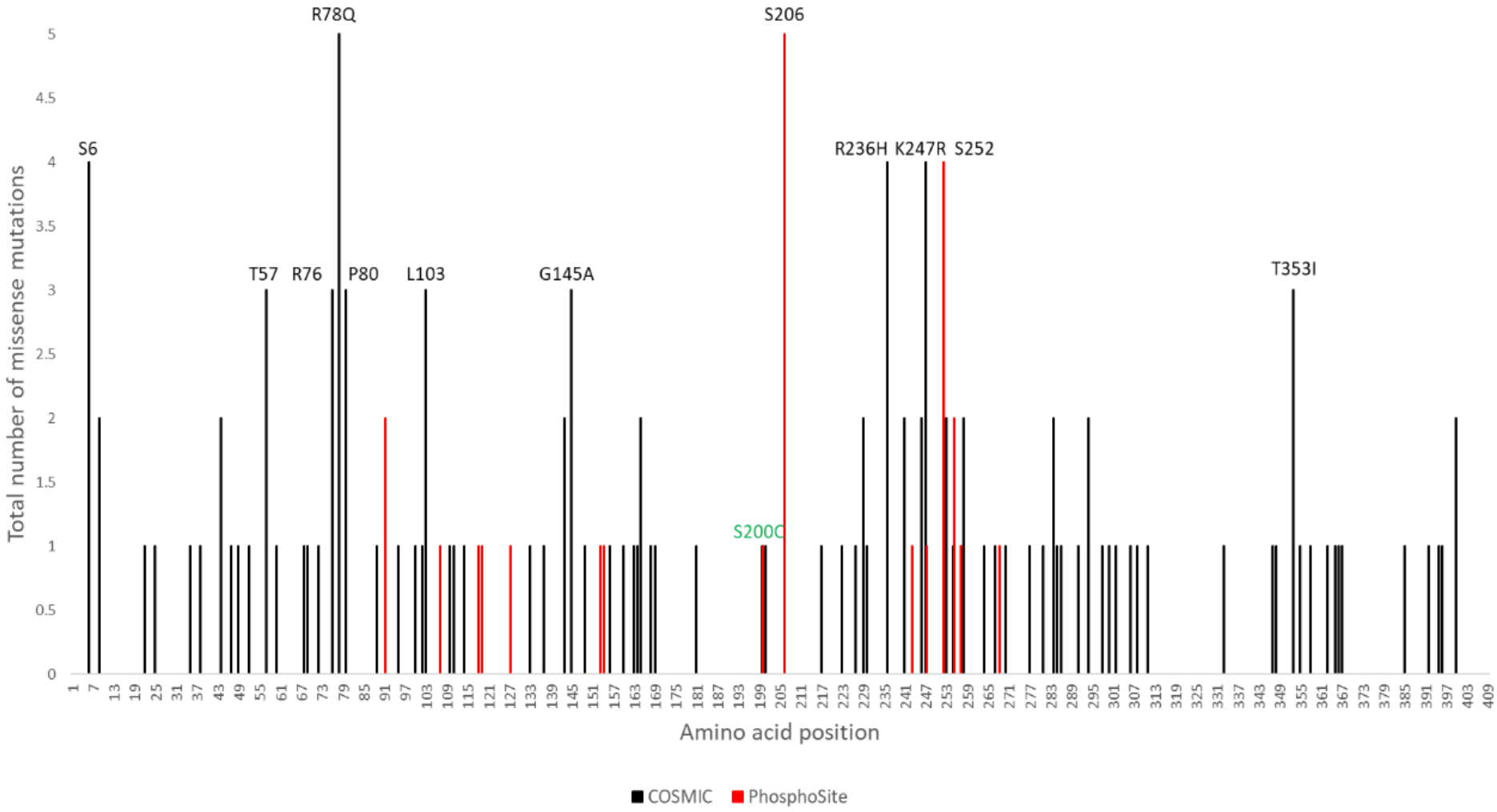

Figure 5.6 Missense mutational hotspot mapping of zDHHC23.

Data collected from COSMIC database from 39,616 unique tumour samples sequenced containing the zDHHC23 gene (Black), alongside data from PhosphoSite (Red). The mutations which were identified on both databases are labelled in green. Cumulative missense mutation count is plotted against protein length and sites. Total mutations = 112.

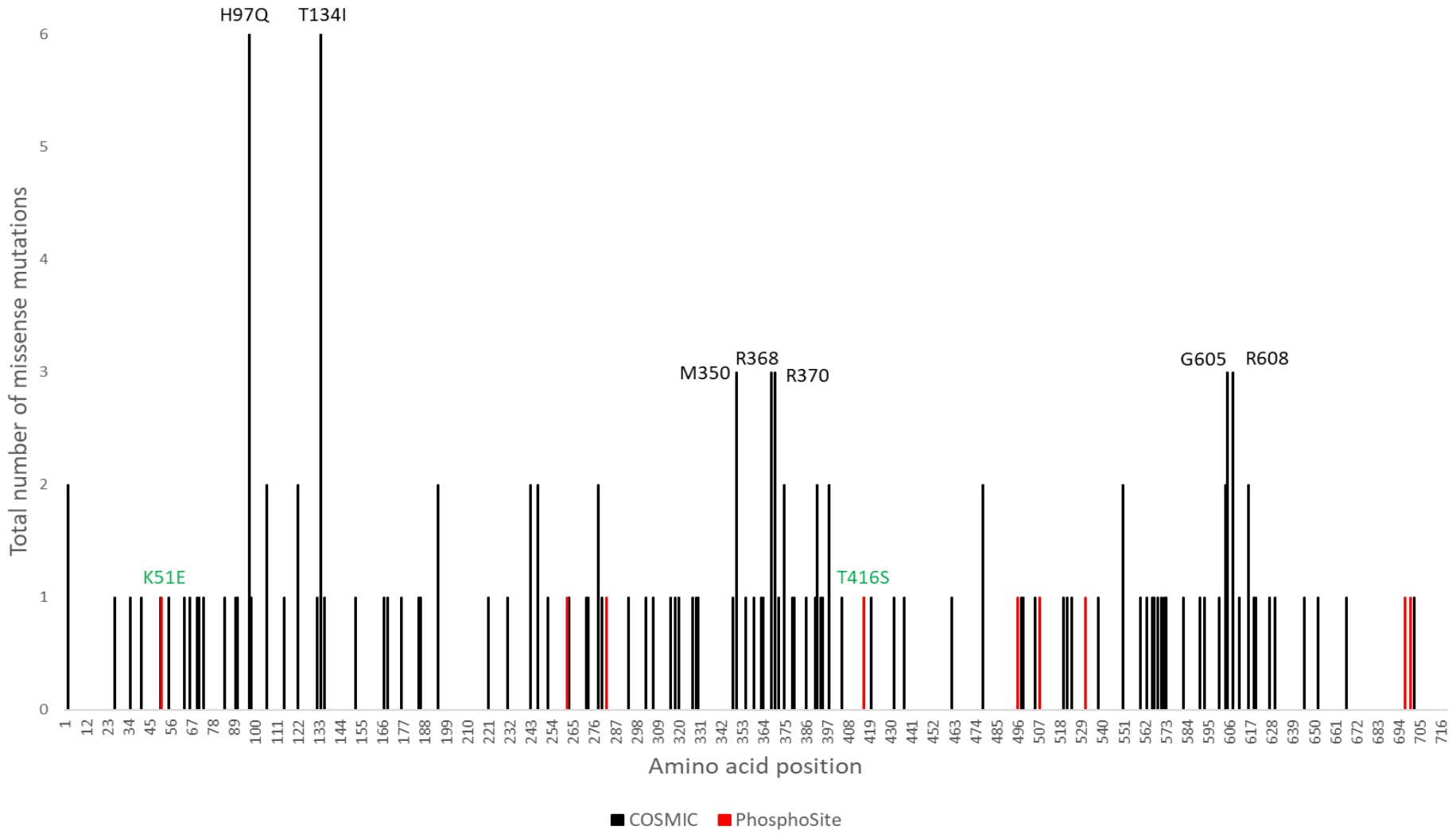

Figure 5.7 Missense mutational hotspot mapping of MROH6.

Data collected from COSMIC database from 39,530 unique tumour samples sequenced containing the MROH6 gene (Black), alongside data from PhosphoSite (Red). The mutations which were identified on both databases are labelled in green. Cumulative missense mutation count is plotted against protein length and sites. Total mutations = 134.

A total of 39,616 and 39,530 unique tumour samples had been sequenced containing the zDHHC23 and MROH6 genes respectively. As the COSMIC database stores different DNA mutations as separate records, a highly mutated protein site with no specific DNA mutation may not be identified as significantly mutated. Therefore, all mutations for a single protein site were cumulated and plotted against amino acid position for both zDHHC23 and MROH6 protein (Figure 5.6 and Figure 5.7). Filtering out synonymous, nonsense, frameshift insertions/deletion and in-frame insertion/deletion mutations from the dataset, a total of 112 and 134 missense mutations were identified for zDHHC23 and MROH6 respectively. It is clear from Figure 5.6 and Figure 5.7, neither zDHHC23 or MROH6 experience prevalent mutational hotspots, with a maximum of five and six protein coding mutations observed respectively, equating to ~0.01% of all samples mutated at a specific residue. For comparison, p53 (TP53 gene) a renowned tumour suppressor been sequenced in 178,596 unique tumours and was found mutated in 50,443 tumours (28% of all samples), equating to a mutation frequency ~3000X greater than zDHHC23 and MROH6. These data suggest that they do not have roles as key oncogenic players. Interestingly, zDHHC23 phosphorylation sites S206 and S252 are identified as mutated five and four times respectively, suggesting potentially important roles that can, although infrequently, be manipulated by cancer to aid its progression.

5.3.4. Construct induced protein identification variability and selection of proteins for analysis

Initial data analysis of binding partners for both zDHHC23 & MROH6 involved analysing N- and C- terminal tag orientations independently, filtering for 2 of 3 replicates at 1% FDR. This allowed me to identify proteins (irrespective of O_2) for either the N- or C- terminal constructs identifying the overlap and the unique proteins identified for zDHHC23 and MROH6 (Figure 5.8 and Figure 5.9).

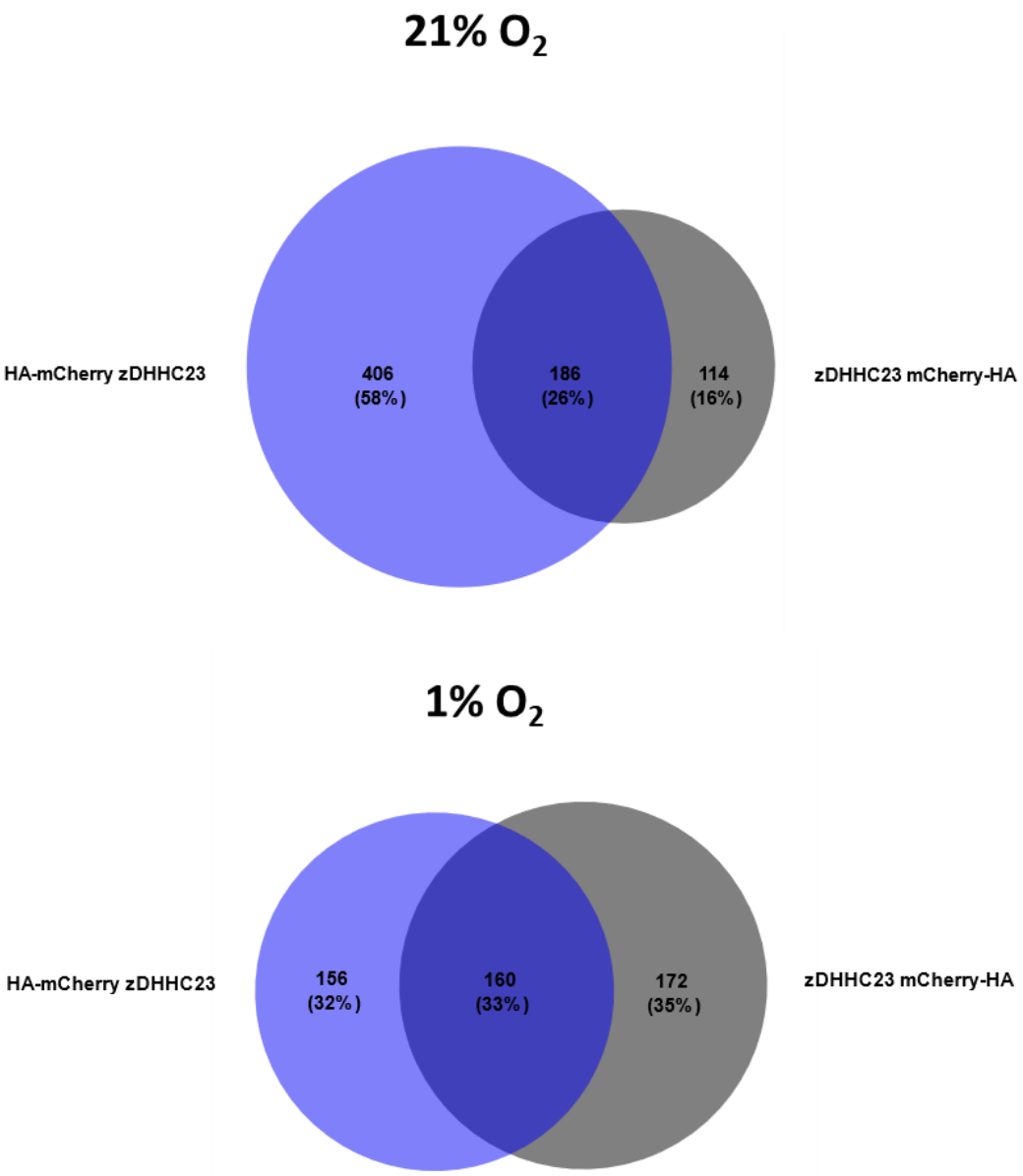

Figure 5.8 Number of binding partners identified for zDHHC23 in 21% and 1% O_2, comparing identifications seen in N- and C- terminal constructs.
Venn diagrams of proteins identified in 21% O_2 or 1% O_2 for zDHHC23 post background subtraction. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % values reflect the number of binding partners identified per section, relative to the total protein count. Circles are scaled respectively to the number of binding partners per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

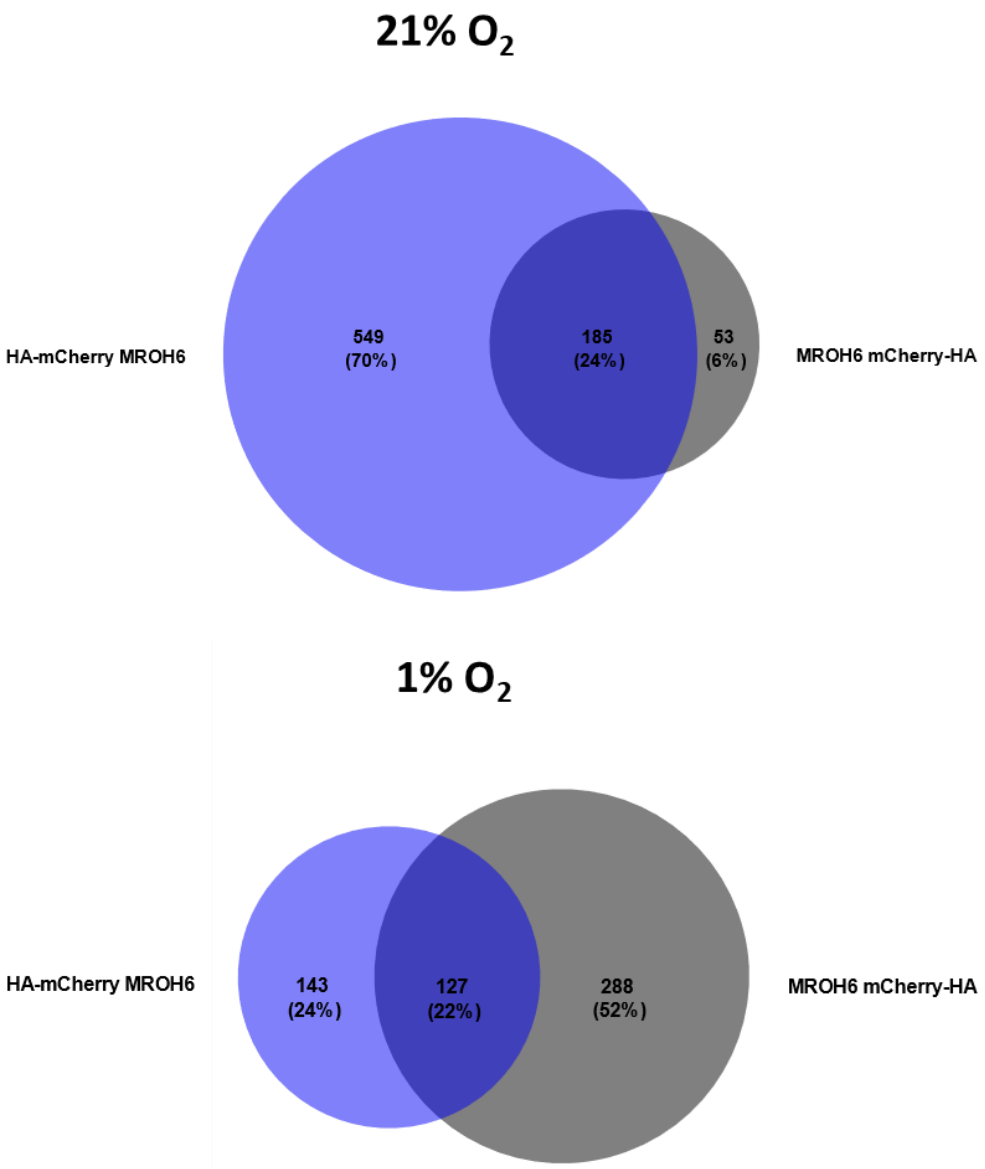

Figure 5.9 Number of binding partners identified for MROH6 in 21% and 1% O_2, comparing identifications seen in N- and C- terminal constructs.
Venn diagrams of proteins identified in 21% O_2 or 1% O_2 for zDHHC23 post background subtraction. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % values reflect the number of binding partners identified per section, relative to the total protein count. Circles are scaled respectively to the number of binding partners per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

Focusing on zDHHC23, in 21% O_2, the N- terminally tagged construct (HA-mCherry zDHHC23) identified a total of 592 proteins, compared to the C- terminally tagged construct (zDHHC23-mCherry HA) that identified 300 proteins, with only 186 proteins (24%) seen in both constructs (Figure 5.8); inferring the interactome changes by 76% depending on tag orientation. Similar results are observed in 1% O_2 with the interactome changing by 67% (Figure 5.8). This discrepancy in the number of binding partners identified due to tag orientation is also observed for MROH6. The largest of which is seen in 21% O_2 with the N-terminally tagged construct (HA-mCherry MROH6) identifying a total of 734 proteins, compared to the C- terminally tagged construct (MROH6-mCherry HA) that identified 238 proteins, with only 185 proteins ~24% seen in both constructs (Figure 5.9). However, the opposite effect of tag orientation is seen in 1% O_2, with the C- terminal tagged construct identifying the most proteins (415) whilst the N- terminal tagged construct identified 145 less proteins (270) with only 127 proteins (22%) seen in both constructs (Figure 5.9).

I also investigated how the orientation of the tag effected the binding partners identified in an O_2-dependent manner (Figure 5.10 and Figure 5.11). There was a marked contrast in the proteins identified with the two different constructs for each protein showing that tag orientation had a substantive influence on the protein interaction networks of both proteins.

For zDHHC23, the N- terminally tagged construct (HA mCherry-zDHHC23) identified a total of 720 proteins, of which 403 were 21% O_2 specific (~56%), 127 1% O_2 specific (~18%) and 190 O_2-independent (~26%). Compared to the C- terminally tagged construct (zDHHC23-mCherry HA) that identifies a considerable reduction in the total number of proteins identified at 487 (~33% reduction versus N-terminal identifications) split as, 154 21% O_2 specific (~38%), 186 1% O_2 specific (~46%) and 147 O_2-independent (~36%). This data not only highlights the vast interactome differences but also the large O_2 tension inducible discrepancies observed with N-terminal tagging, with 21% and 1% O_2 specific identification equating to ~56% and 18% respectively of all identifications. Whilst C-terminal tagging equates to ~38% (~38% reduction in the proportion of 21% O_2 specific identifications) and 46% (~156% increase in the proportion of 1% O_2 specific identifications) respectively (Figure 5.10).

Similarly, MROH6 experiences the same tag orientation and O_2 tension induced interactome differences (Figure 5.11). For MROH6, the N- terminally tagged construct (HA mCherry-MROH6) identifies a total of 827 proteins, of which 556 were 21% O_2 specific (~67%), 92 1% O_2 specific (~11%) and 179 O_2-independent (~22%). Again, compared to the C-terminally tagged construct (MROH6-mCherry-HA) there is a considerable reduction in the total number

of proteins identified at 387 (~53% reduction versus N-terminal identifications) equating to 79 21% O_2 specific (~20%), 249 1% O_2 specific (~64%) and 59 O_2-independent (~15%). Hence, MROH6 tag orientation resulted in an ~70% reduction in 21% O_2 specific and an ~482% increase in 1% O_2 specific binding partners (N- terminal versus C-terminal) (Figure 5.10 and Figure 5.11).

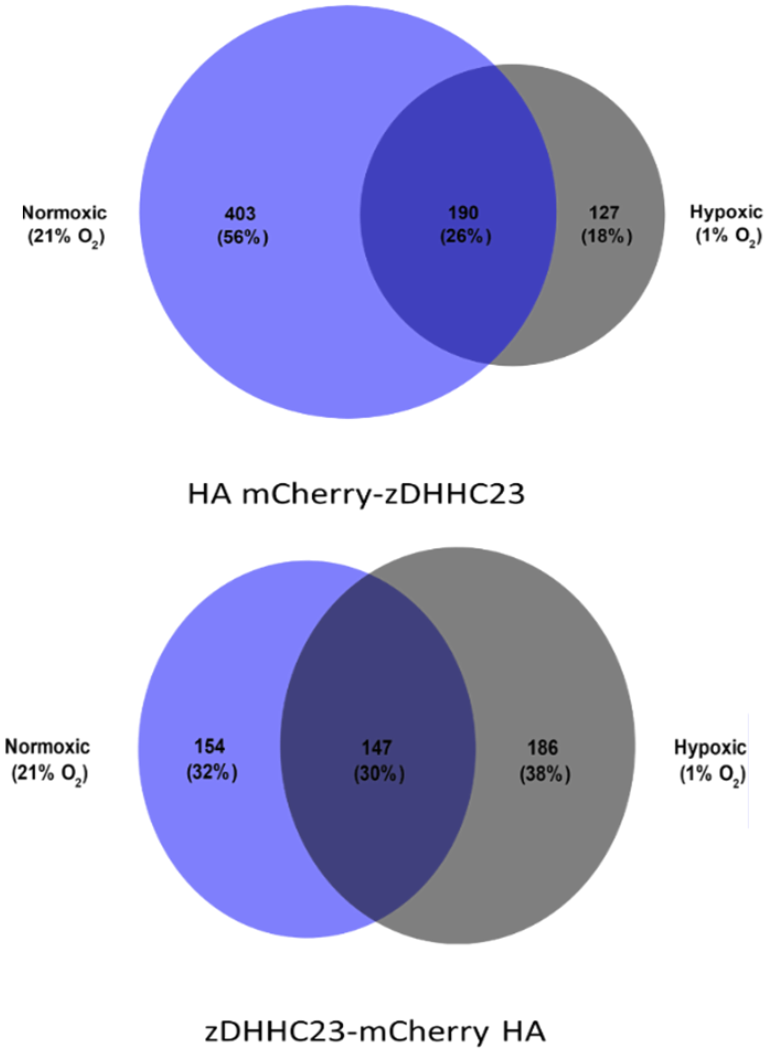

Figure 5.10 Number of binding partners identified for zDHHC23 in all O_2 tensions.
Venn diagrams of proteins identified in 21% O_2 versus 1% O_2 for zDHHC23 post background subtraction. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % values reflect the number of binding partners identified per section, relative to the total protein count. Circles are scaled respectively to the number of binding partners per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

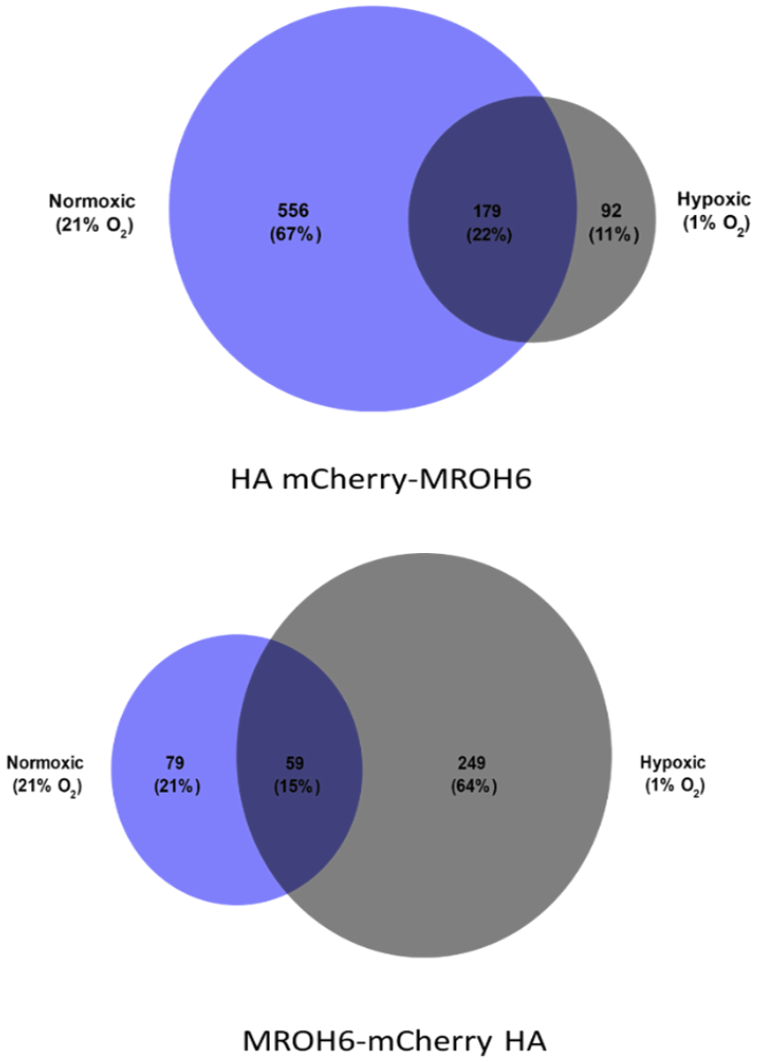

Figure 5.11 Number of binding partners identified for MROH6 in all O_2 tensions.
Venn diagrams of proteins identified in 21% O_2 versus 1% O_2 for MROH6 post background subtraction. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % values reflect the number of binding partners identified per section, relative to the total protein count. Circles are scaled respectively to the number of binding partners per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

It is known that the addition of relatively bulky protein tags (HA-mCherry tags are ~30 kDa) can interfere with endogenous functions (Ramage et al., 2002). Therefore, combined with the earlier motif analysis, I can postulate that for zDHHC23, tagging the C-terminus where the transmembrane domain lies could influence interactors in an unknown manner, potentially causing mis-localisation or prohibiting binding due to change in conformation. Similarly, the MROH6 N-terminal is predicted to be a large, disordered region while the C-terminal is the large α-helical repeat domains of HEAT proteins. Therefore, it is foreseeable

that tagging of either terminus of MROH6 may influence functional roles by affecting likely regulatory regions of functional domains. Regardless of these hypothesises, it is difficult to reason why these would result in such extensive O_2 tension inducible interactome shifts. Combining this with the lack of knowledge of which tag orientation is better suited, I decided to merge the tag-orientation datasets (similar to (Byrne et al., 2022)) to create an interactome that is orientation independent (with the previous filtering parameters applied, now creating a pseudo 4 out of 6 replicates filter) (Figure 5.12 & Appendix 5). This method obviously has the caveat of inevitably reducing the total number of proteins retained for further investigations, with the corollary that the confidence of these being genuine binding partners (either direct or indirect) is improved.

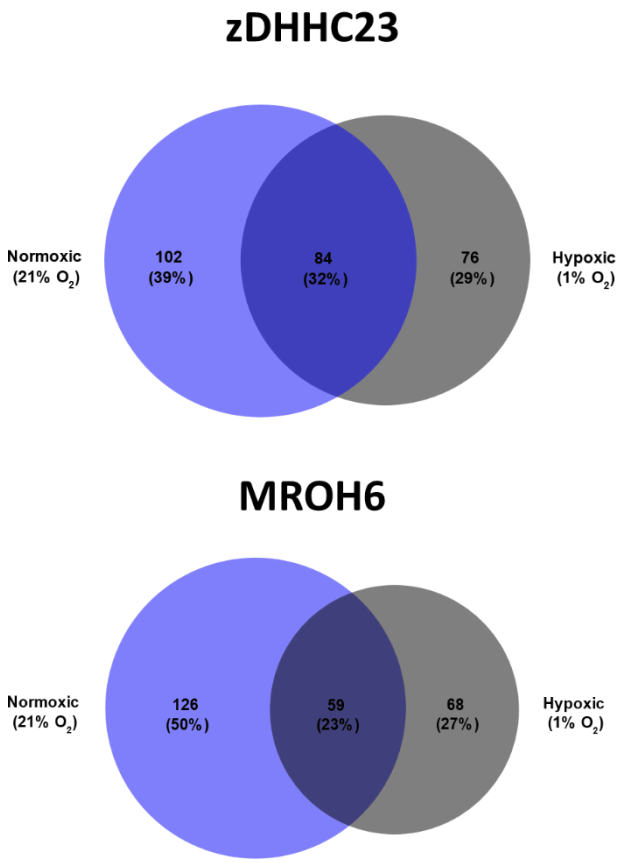

Figure 5.12 Number of binding partners identified for zDHHC23 and MROH6 with both C and N terminal construct datasets combined.
Venn diagrams of proteins identified in 21% O_2 versus 1% O_2 zDHHC23 and MROH6 with C and N terminal construct data combined with post background subtraction. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % values reflect the number of binding partners identified per section, relative to the total protein count. Circles are scaled respectively to the number of binding partners per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008). Top = zDHHC23 Bottom = MROH6.

Using these filtering parameters, I identified 262 proteins in the zDHHC23 interactome, of which 102 (~39%) are 21% O_2 specific, 76 (~29%) are 1% O_2 specific and 84 (~32%) are O_2-independent interactors. The MROH6 interactome is a total of 253 proteins, of which 126 (~50%) are 21% O_2 specific, 69 (~27%) are 1% O_2 specific and 59 (~23%) are O_2-independent interactors. These data highlights that both zDHHC23 and MROH6 have highly O_2 tension inducible interactomes with 68% and 77% of all identifications being specific to a single O_2 tension, hence having ample opportunity for being associated with aggressive NB and poor patient prognosis.

5.3.5. Gene Ontology (GO) analysis

To explore the interactomes of zDHHC23 (262 proteins) and MROH6 (253 proteins) I performed gene ontology (GO) enrichment analysis to identify which biological processes, molecular functions, cellular compartments and disease pathways (KEGG) were significantly enriched by interaction with either zDHHC23 or MROH6. GO term enrichment analysis was performed in an O_2-dependent and independent manner to differentiate essential roles (O_2-independent) and adaptive roles (O_2-dependent) in an attempt to identify biochemical reasoning as to why zDHHC23 and MROH6 are associated with aggressive hypoxic NB and poor patient prognosis (Figure 5.13, Figure 5.14 & Appendix 6). GO annotation enrichment analysis was performed using the DAVID (Database for Annotation, Visualisation and Integrated Discovery) tool (Dennis et al., 2003) and visualised as volcano plots of statistical significance (q-value) versus GO term enrichment in the background of the human proteome database, plotted using custom R scripts available in Appendix 2.

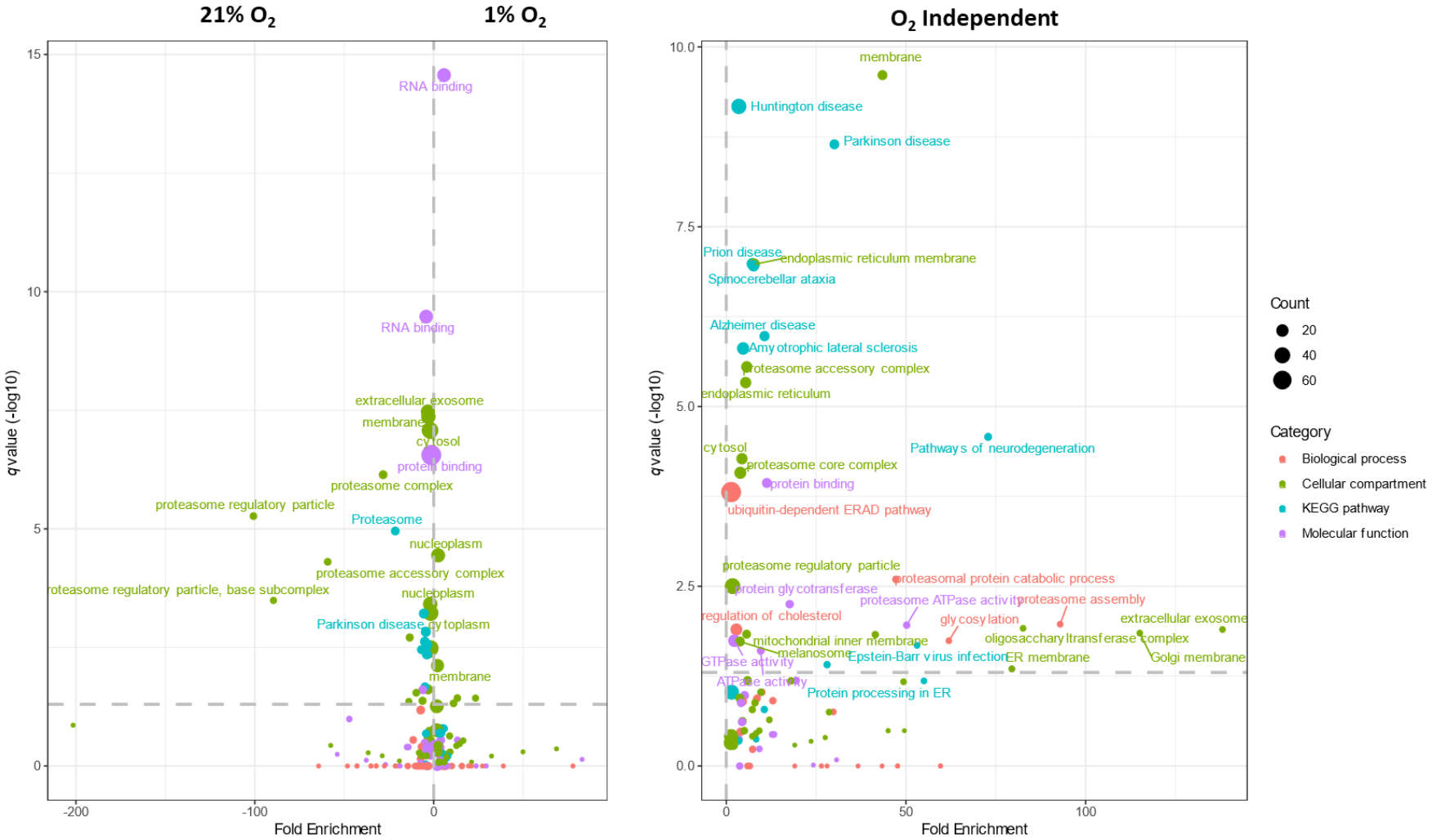

Figure 5.13 GO Enrichment Analysis of 21%, 1% and O_2-independent zDHHC23 binding partners.
GO enrichment analysis was performed in DAVID using the list of interactors identified in two or more replicates of zDHHC23 with C and N terminal construct data combined. P values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg correction and are plotted against fold enrichment. The horizontal line represents an adjusted p value of 0.05. Points are sized according to the number of proteins within each enriched term and coloured according to the enrichment category. Left = 21% and 1% O_2 Right = O_2-independent.

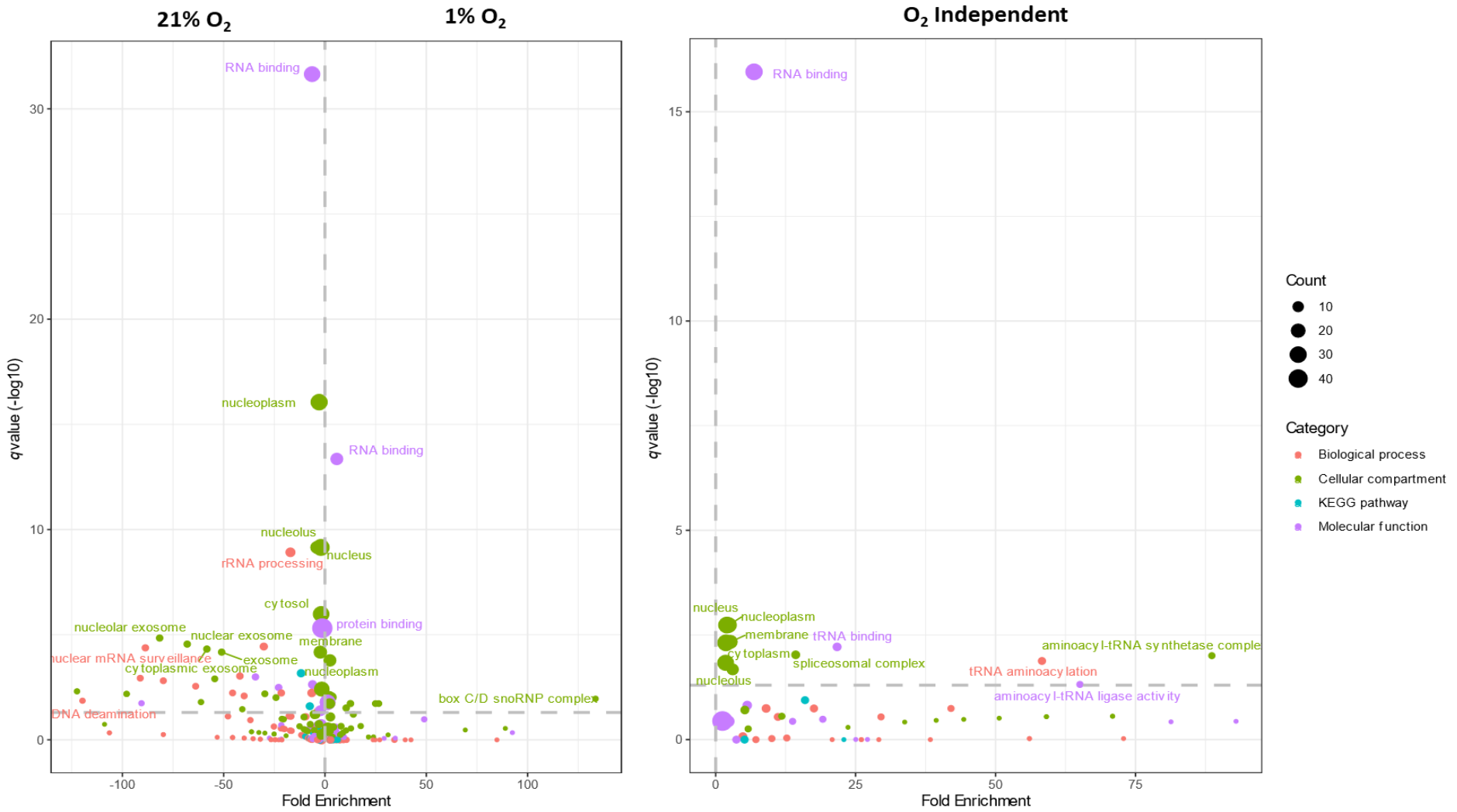

Figure 5.14 GO Enrichment Analysis of 21%, 1% and O_2-independent MROH6 binding partners.
GO enrichment analysis was performed in DAVID using the list of interactors identified in two or more replicates of MROH6 with C and N terminal construct data combined. P values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg correction and are plotted against fold enrichment. The horizontal line represents an adjusted p value of 0.05. Points are sized according to the number of proteins within each enriched term and coloured according to the enrichment category. Left = 21% and 1% O_2 Right = O_2-independent.

Reassuringly, the cellular compartment term 'membrane' is present in all O_2 tension conditions for zDHHC23; 21% O_2 (Benjamini-Hochberg q-value) = 4.29E-08, Count 38 = 37%), 1% O_2 (Benjamini-Hochberg q-value) = 7.63E-03, Count 29 = 38 %) and O_2-independent (Benjamini-Hochberg q-value) = 2.46E-10, Count 10 = 12 %), equating to the third and first most significant term identified respectively (Figure 5.13). These data suggest that while a large proportion of zDHHC23 interacting proteins identified are cell membrane localised and are unaffected by O_2 tension, there is a subpopulation of interactors that are highly O_2 sensitive. Additionally, multiple membranous-organelle terms are present in the O_2-independent binding partners including; ER membrane (Benjamini-Hochberg q-value) = 1.05E-07, Count 15 = 18 %), Golgi membrane (Benjamini-Hochberg q-value) = 0.014, Count 3 = 4 %) and mitochondrial inner membrane (Benjamini-Hochberg q-value) = 0.015, Count 7 = 8 %). The identification of these terms suggests that the function of zDHHC23 as a palmitoyltransferase is not prohibited by tagging, given current knowledge that palmitoylation targets strongly membrane associate, albeit currently thought to be primarily the plasma cell membrane (Tian et al., 2012). Therefore, these data suggest zDHHC23 may have important O_2-independent roles localising target proteins to different membrane organelles.

The most significant GO-term for the molecular function of zDHHC23 in 21% O_2 and 1% O_2 datasets is RNA binding; 21% O_2 (Benjamini-Hochberg q-value) = 3.35E-10, Count 33 = 32 %) and 1% O_2 (Benjamini-Hochberg q-value) = 2.72E-15, Count 34 = 45 %). This data, unexpectedly, suggests that zDHHC23 may have a role in protein translation. However, the RNA binding GO-term is absent from the zDHHC23 O_2-independent dataset, suggesting that there are highly distinct and different RNA binding roles dependent on O_2 tension. Interestingly, it is clear that multiple significant proteasome terms are identified as 21% O_2 specific, including: proteasome complex (Benjamini-Hochberg q-value) = 7.22E-07 , Count 8 = 8 %), proteasome regulatory particle (Benjamini-Hochberg q-value) = 5.38E-06 , Count 5 = 5%), proteasome (Benjamini-Hochberg q-value) = 1.11E-05, Count 8 = 8%) and proteasome accessory complex (Benjamini-Hochberg q-value) = 4.95E-05, Count 5 = 5%). It is difficult to interpret the proteasome identifications due to being suggestive of multiple processes: 1) zDHHC23 is degraded in 21% O_2 (similar to HIF-1α (Moroz et al., 2009)). 2) zDHHC23 binds to or regulates (potentially through PTM) proteins being degraded. 3) zDHHC23 somehow regulates the proteasome. However, in combination with IP, western blots where zDHHC23 laddering is observed (indicative of polyubiquitination and proteolytic degradation (Pickart and Cohen, 2004) (Figure 4.1).

GO term analysis identifies the cellular compartment term 'nucleus' as an O_2-independent nuclear related term for MROH6 (Benjamini-Hochberg q-value = 0.132, Count 31 = 53%), the first indication of MROH6 subcellular localisation reported (Figure 5.14). This conclusion is supported with the knowledge that many HEAT repeat protein family members are nuclear, including: Huntingtin (Hackam et al., 1999) PP2A (Seshacharyulu et al., 2013) and ATR (Kidiyoor et al., 2020). Interestingly, the term 'nucleus' is also identified as enriched in 21% O_2 (Benjamini-Hochberg q-value = 3.28E-03, Count 74 = 59%) and 1% O_2 (Benjamini-Hochberg q-value = 1.75E-01, Count 28 = 41%) specific. While MROH6 is thus likely nuclear, the cellular pathways/interactors it is involved with change substantially dependent on O_2 tension. The most significant molecular function GO-term observed for MROH6 in all O_2 tensions is RNA binding; 21% O_2 (Benjamini-Hochberg q-value = 6.91E-61, Count 62 = 49%), 1% O_2 (Benjamini-Hochberg q-value) = 9.01E-78, Count 31 = 46%) and O_2-independent (Benjamini-Hochberg q-value) = 1.11E-16, Count 31 = 53%). Considering current knowledge of other HEAT repeat proteins (including: Sda1p, Noc1p, Noc2p, and Noc3p) and their requirement for pre-rRNA processing of the 60S subunit (Ferreira-Cerca et al., 2005), this finding is less surprising, albeit no-less important, for understanding MROH6 functions.

Interestingly, the highest fold enrichment of any term (disregarding O_2 dependency) was observed in 1% O_2 for the term 'box C/D snoRNP complex' (Fold enrichment = 133, Benjamini-Hochberg q-value = 0.011, Count 3 = 4%), which is an absent term from 21% O_2 and O_2-independent datasets. The 'box C/D snoRNP complex' is a ribonucleoprotein complex assembly which is predominantly localised in the nucleolus and primarily catalyses PTMs of ribosomal RNAs2 (rRNAs) (NEWMAN et al., 2000). The third highest fold enrichment terms in O_2-independent proteins identified is 'aminoacyl-tRNA synthase complex' (Fold enrichment = 89, Benjamini-Hochberg q-value = 5.75E-02, Count 3 = 5%) 'amino-acyl tRNA ligase activity' (Fold enrichment = 65, Benjamini-Hochberg q-value = 4.80E-02, Count 3 = 5%) and 'tRNA amino-acylation' (Fold enrichment 58, Benjamini-Hochberg q-value =1.31E-02, Count 4 = 7%). Aminoacyl-tRNA synthetases (ARSs) and their catalytic activity (catalysing the ligation of amino acids to their cognate tRNAs) is an essential role in maintenance of cell viability (Gomez and Ibba, 2020). It has also been suggested that some ARSs have evolutionary acquired additional functions including apoptosis and angiogenesis (Lee et al., 2004, Park et al., 2005). Interestingly, Tryptophanyl-tRNA synthetase (TrpRS) expression, a key enzyme of protein biosynthesis belonging to the family of ARS, has shown to correlate with NB growth rate in SH-SY5Y cells (Paley et al., 2007). The enrichment of these proteins is therefore

intriguing when considering the poor survivability of MROH6 overexpression in the Kaplan-Meier survival plots (Figure 1.7).

5.3.6. STRING analysis

Following on from GO-term enrichment, I analysed my confidently identified interactors of zDHHC23 (262 proteins) and MROH6 (253 proteins) by STRING analysis (version 11.5, http://string.embl.de). STRING analysis generates interaction networks from known published interactions for a list of proteins, allowing the identification of potential molecular pathways/cell signalling pathways at play. Considering protein interactions identified by Co-IP can be indirect, hence O_2-dependent changes may be from direct binding partners that may not change by O_2 tension, I performed STRING analysis on all proteins identified in a specific O_2 tension (i.e., 21% O_2 specific and O_2-independent data (zDHHC23 = 186, MROH6 = 185) or 1% O_2 specific and O_2-independent data (zDHHC23 = 160, MROH6 = 127)). STRING analysis was filtered for interactions stored at the level of either experimental or database evidence with high confidence (>0.7). STRING interactomes for zDHHC23 (Figure 5.15) and for MROH6 (Figure 5.16) are below.

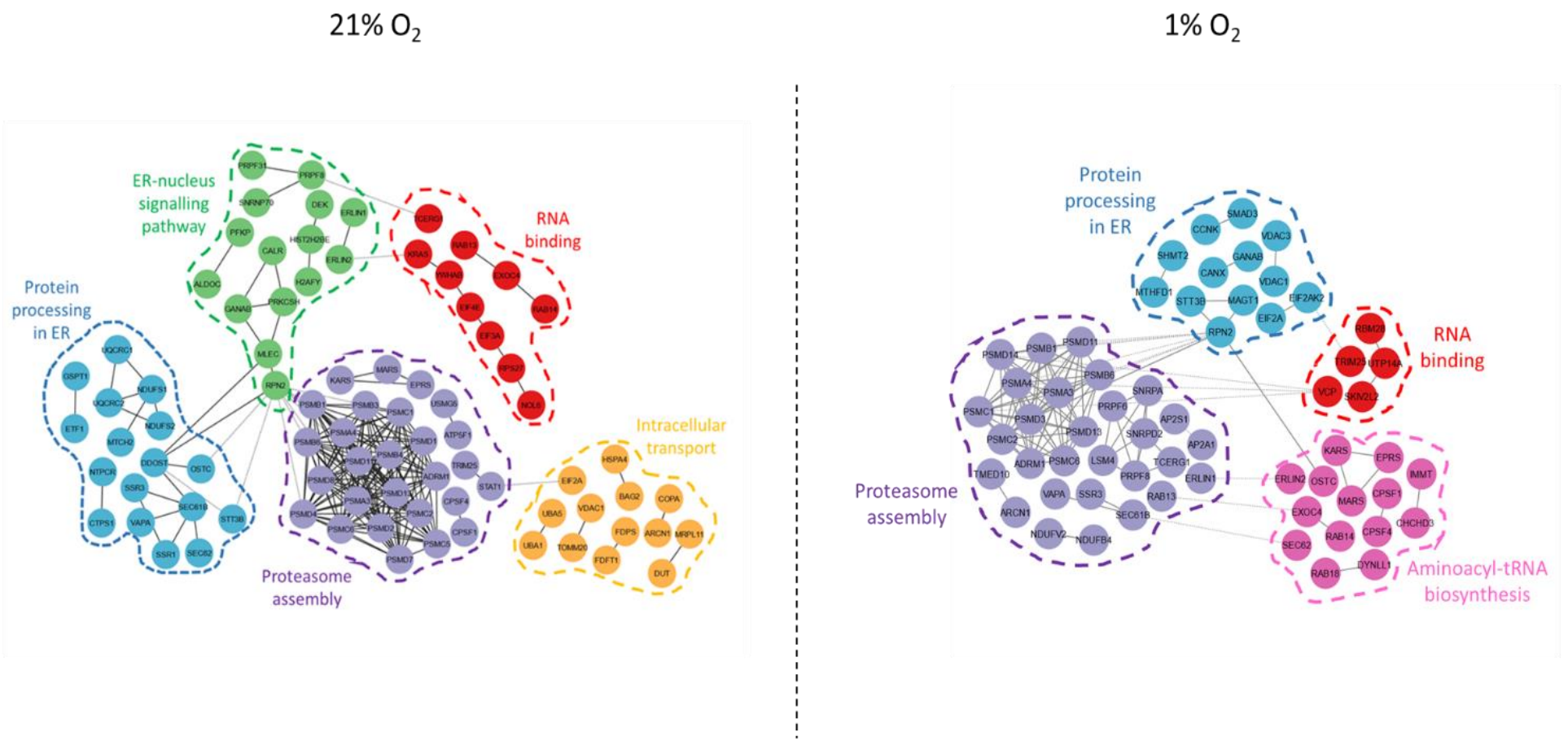

Figure 5.15 Network of proteins identified in two or more 21% and 1% O_2 exogenous zDHHC23 IP.
The original network was generated in STRING to include interactors with experimental or database evidence of high (interaction score >0.7, dark grey lines) or medium (interaction score >0.4, light grey lines) confidence. Any nodes with no connections of medium confidence or lower were removed from the network. The network was re-arranged in Cytoscape and re-coloured to group nodes into functional groups/complexes. Left = 21% O_2, Right = 1% O_2.

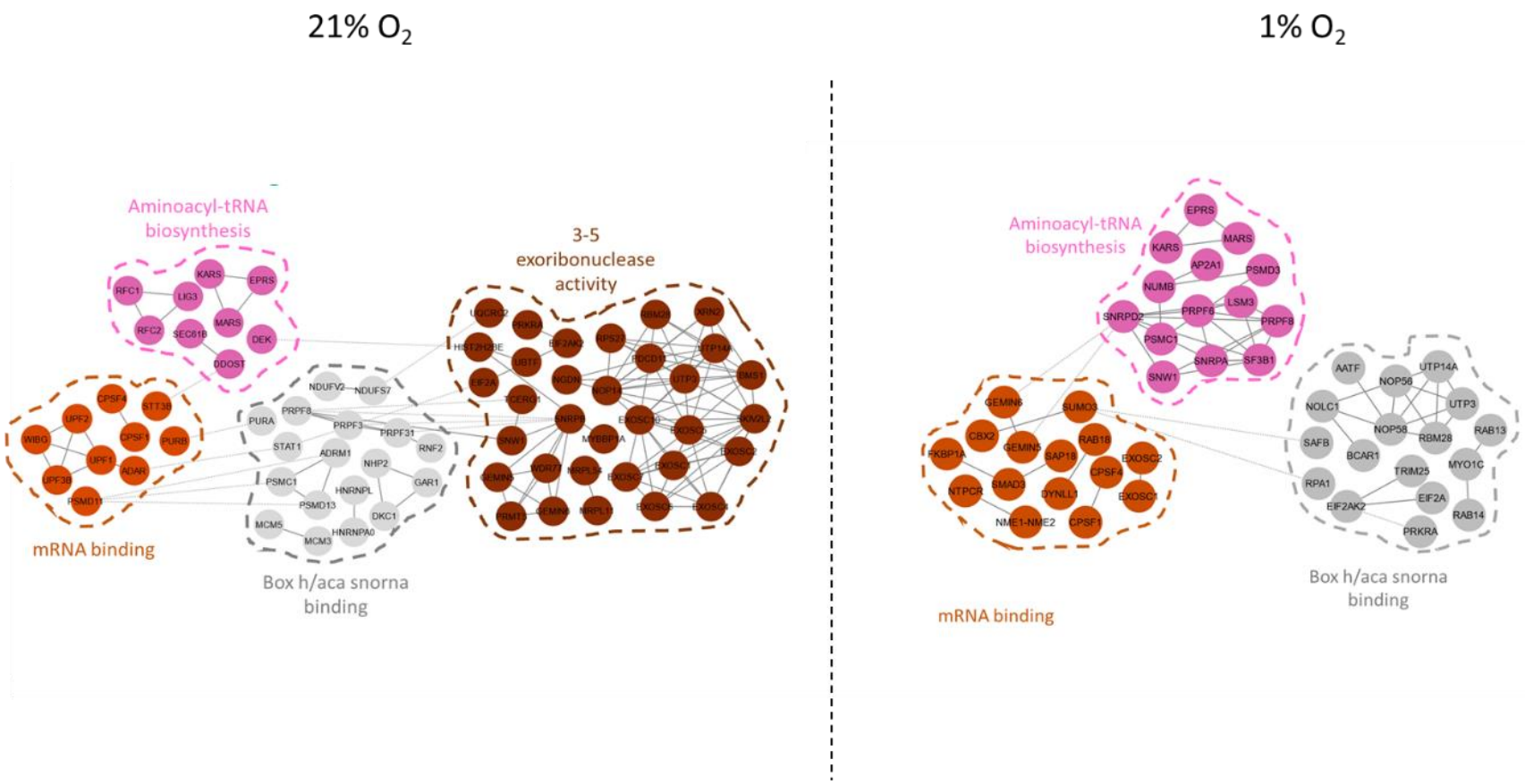

Figure 5.16 Network of proteins identified in two or more 21% and 1% O_2 exogenous MROH6 IP.
The original network was generated in STRING to include interactors with experimental or database evidence of high (interaction score >0.7, dark grey lines) or medium (interaction score >0.4, light grey lines) confidence. Any nodes with no connections of medium confidence or lower were removed from the network. The network was re-arranged in Cytoscape and re-coloured to group nodes into functional groups/complexes. Left = 21% O_2, Right = 1% O_2.

Reiterating previous observation in my GO analysis, zDHHC23 has a large proteasomal network in 21% O_2 (Figure 5.15). While many of these proteasomal proteins are identified in 1% O_2, this can be explained due to using all proteins identified in 1% O_2 (i.e. 1% only and O_2-independent) where I identify some proteasomal GO-terms (proteasome accessory complex, proteasome core complex and proteasome regulatory complex). Multiple proteins comprising the different subunits of the proteasome complex are identified by STRING analysis (Figure 5.15), including: structural (PSMD1, PSMD3, PSMD6, PSMD7, PSMD8, PSMD11, PSMD13), ubiquitin binding (PSMD2, PSMD4) or ATPase activity (PSMC1, PSMD2, PSMD5) (Bard et al., 2018).

The additional co-identification of 'Protein processing in the ER' in 21% and 1% O_2 datasets supports another O_2-independent GO-term identified: ubiquitin dependent endoplasmic reticulum associated protein degradation (ERAD) pathway (Benjamini-Hochberg q-value = 0.437), Count 3.80 74 = 88%). The ERAD pathway is a quality control mechanism, that utilises the 26S proteasome, for ensuring correct protein folding of newly synthesised membrane bound proteins, a considerable threat due to proteins entering the ER in an unfolded manner (Wu and Rapoport, 2018, Braakman and Hebert, 2013). The identification of such proteasomal terms and ERAD signalling can be suggestive of either two explanations, which our current data cannot decipher: 1) zDHHC23 interacts with proteins that are targeted for intracellular trafficking and thus ERAD quality control mechanisms or 2) exogenous expression of tagged-zDHHC23 (a predicted membrane associated protein, thus likely undergoes ER shuttling) traverses the ER and undergoes ERAD quality control. Unfortunately, it is not possible to determine from this data whether exogenous tagged-zDHHC23 is prone to incorrect folding and thus ERAD mediated degradation, or the larger quantity of protein is more likely to result in mixed populations of folded and newly synthesised protein (i.e. undergoing quality control) leading to enrichment of the ERAD proteins. Alternatively, it is also possible that zDHHC23 is ER localised and may interact with proteins of the ERAD pathway.

An interaction network involved with 'intracellular transport' is specifically identified in 21% O_2. Investigating the proteins that are part of this network, many have previously been reported as having roles in various cancers, including: HSPA4 (aids proliferation in CRC (Zhang et al., 2021a)) and hepatocellular cancer (Shang et al., 2021), while increased expression in breast cancer is linked to metastasis (Gu et al., 2019), BAG2 (linked to gastric cancer progression (Sun et al., 2020)) and public database analysis showing overexpression is generally associated with poor cancer patient prognosis (Yue et al., 2015), FDSP (promotes

prostate cancer (Seshacharyulu et al., 2019) and glioblastoma (Abate et al., 2017) cell growth), TOMM20 (associated with CRC proliferation (Park et al., 2005) and high-grade chondrosarcoma (Roche et al., 2020)) and VDAC1 (where silencing has been shown to inhibit tumour growth (Koren et al., 2010, Amsalem et al., 2020)). Therefore, zDHHC23 association with these proteins with known oncogenic roles may provide some insight into how zDHHC23 may influence NB progression. Additionally, STRING analysis identifies 'RNA binding' networks, thus cementing earlier GO-term analysis observations.

For MROH6, while the major network classifications identified are near identical for 21% and 1% O_2 interactors, the constituent proteins are highly different, suggesting that while the overall function of MROH6 may not significantly alter, its direct targets/roles is significantly influenced by O_2 tension (Figure 5.16). The three functional networks identified as O_2-independent are: mRNA binding, aminoacyl-tRNA biosynthesis and box H/ACA snorna binding; strongly supporting previous identifications through GO-term analysis. Additionally, multiple of these mRNA binding proteins in the networks are described as nuclear localised by UniProt (Consortium, 2019) (DYNLL1, SMAD3, SAP18, SUMO3, EXOSC2/1, GEMIN5/6, CBX2, WIBG, UPF3B, PSMBD11, ADAR, PURB and CPSF1/4), reiterating GO-term analysis identifying multiple nucleus terms (Figure 5.14).

As alluded to, in the mRNA binding networks only two proteins (CPSF1 and CPSF4) are observed in both 21% and 1% O_2 datasets. Both CPSF1 and CPSF4 are involved with polyadenylation of mRNA (Misra and Green, 2016) and have been implicated in various cancers, where CPSF1 has been shown to promote ovarian cancer, (Zhang et al., 2017), gastric cancer (Kang et al., 2022) and head and neck squamous cell carcinomas (Sakai et al., 2020) and CPSF4 has been shown to play a role in lung adenocarcinoma (Chen et al., 2013), triple negative breast cancer (Lee et al., 2021b) and CRC (Zhang et al., 2016). The O_2-independent identification of these proteins as binding partners suggests MROH6 may have roles in mRNA polyadenylation/lifespan and provide some insight into why high expression of MROH6 results in a poor outcome for NB patients, with different co-interacting proteins potentially targeting different mRNA sequences. Further reiterating GO-term analysis, another network identified as O_2-independent is 'box H/ACA snorna binding'. This network involves non-coding RNAs that localise within the nucleolus and function predominantly for nucleotide-based PTMs (Massenet et al., 2017, Kiss et al., 2010). Within this network I identify two minichromosome maintenance (MCM) proteins – MCM3 and MCM5. These proteins are fundamental for the initiation of DNA replication (Forsburg, 2004) and are considered as indicators of poor prognosis with MCM5 observed in cervical (Wang et al.,

2018) and thyroid cancer (Mio et al., 2016) and MCM3 in leukaemia, lymphoma and lung cancer to name few (Ha et al., 2004). These protein identifications are of particular interest considering that NB patients have a high occurrence of SCA with the aberration of 11q chromosome used in current risk stratification (Table 1.3). It is worth noting that an essential core protein of this 'box H/ACA snorna binding' complex is observed in a 21% O_2 dependent manner– GAR1 (discussed further in Section 5.3.7.2).

The only network identified specifically to a single O_2 tension (21% O_2), also has the greatest number of protein identification of any network at 33, is '3-5 exoribonuclease activity'. I identify multiple proteins known to form the core exosome complex, including: EXOSC1, EXOSC2, EXOSC4, EXOSC5, EXOSC6, EXOSC7 and EXOSC10 (Figure 5.16). Due to the data analysis pipeline employed, it is likely these proteins are 21% O_2 specific. Exonucleases are known to play a role in RNA decay (Ibrahim et al., 2008), with high expression associated with poor survival in mantle cell lymphoma patients (Zhang et al., 2019).

5.3.7. Label Free Quantification (LFQ) analysis of binding partners

Current analysis methods have relied upon identification and discrete categorisation of proteins under different O_2 tensions. However, particularly for O_2-independent binding partners, this ideology is potentially removing interesting data due to the heterogeneity associated with complexes under different conditions. I therefore used label free quantification (LFQ) to determine fold of proteins Co-IP'd by zDHHC23 and MROH6 in response to O_2 tension. Due to cost and time constraints, I did not explore the feasibility of a label based quantification strategy.

A caveat of LFQ approaches is that samples are prepared and analysed independently of each other, hence prone to greater experimental error and variation between runs which can potentially impact the accuracy of protein quantification (Bantscheff et al., 2012). Therefore, LFQ approaches are heavily reliant on normalisation strategies, typically against 'housekeeping proteins' (Välikangas et al., 2018). However, this strategy is fundamentally flawed here due to IP specifically enriching a subset of proteins from the proteome in an unknown manner, combined with the large gene expression changes in many house keeping genes known to be induced in hypoxia (Caradec et al., 2010). Thus, this normalisation strategy is not adequate for these experiments. Instead, since IP targets a single protein, I

normalised against the peptides from zDHHC23 or MROH6 per replicate in order to investigate binding partners as a proportion of the level of target protein and analyse relative fold change between O_2 tensions.

LFQ analysis was conducted using Proteome Discoverer v2.4, with low abundance imputation of missing data performed, with abundance determined by calculating the area under the curve of a peptide ion peak with a minimum of 1 peptide identified (at a 1% FDR) and quantified (Appendix 7). Proteome Discoverer v2.4 applies a match between runs parameter, in which, if a peptide ion *m/*z is observed at a similar retention time but has no associated PSM data, then the abundance is calculated from this ion due to the likelihood of being identical. Low abundance resampling is a technique in which a normally distributed 'fake' dataset is created 5% below the lowest quantified values and missing values are randomly generated from this distribution. Hence, while imputation allows for statistical analysis (three versus three for all protiens), low abundance proteins or high variability proteins (e.g. observed in two replicates but not the third) can result in identifications not being statistically significantly different yet classified as O_2 tension specific by previous analysis. Using the filtered dataset of identifications above, a second dataset with imputation active was filtered to identify these proteins only (zDHHC23 262 proteins & MROH6 252 proteins), where quantification was calculated using the Hi3 (TOP 3) method (Silva et al., 2006). Briefly HI3 involves using the sum of the signal intensities of the three most, non-modified (no Met residues which can be oxidated), abundant peptides of each protein. Protein abundance per replicate was normalised to the respective protein of interest (either zDHHC23 or MROH6) before average fold changes and statistical analysis performed and plotting using a custom R script (Figure 5.17, Figure 5.18 & Appendix 1).

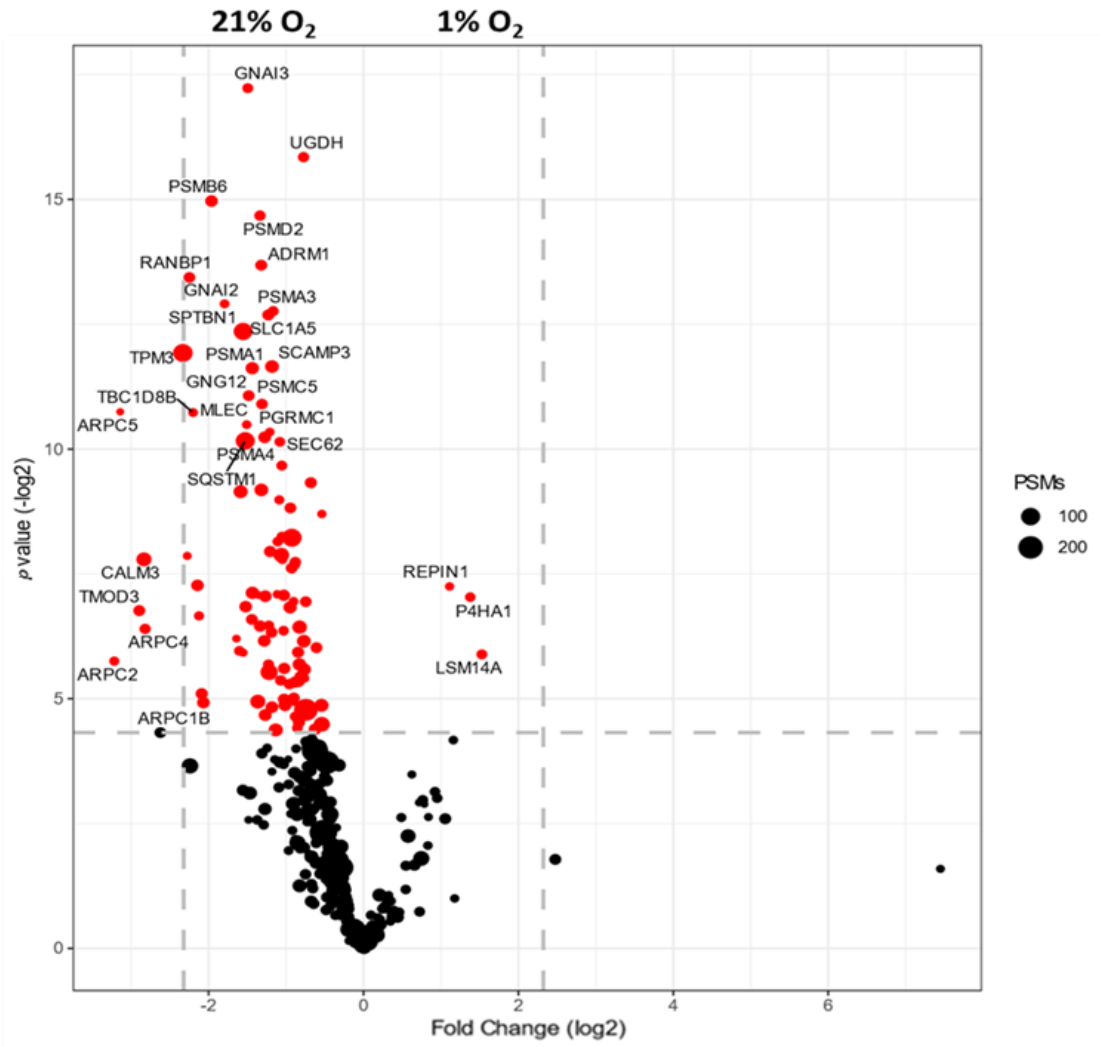

Figure 5.17 LFQ analysis of zDHHC23 binding partners in response to O_2 tension.
LFQ was performed using Proteome Discoverer 2.4 using the minora feature detector node set up for using Hi3 protein quantification method. Missing intensities were imputed by low abundance resampling and a must be identified in at least 2-of-3 replicates filter applied. Protein intensities were normalised to zDHHC23 protein intensity, per replicate and fold change calculated by 1% O_2 average/21% O_2 average. P-values were determined using a Student's 2-sample t-test. Plot is of –Log2 (P-value) against Log2 (fold change), where positive values are enriched in 1% O_2 and negative values enriched in 21% O_2. Dotted lines are at a p-value of <0.05 and 5-Fold enrichment. Points coloured red are significant at p-value <0.05. Point size is scaled to the total number of PSMs identifying a protein.

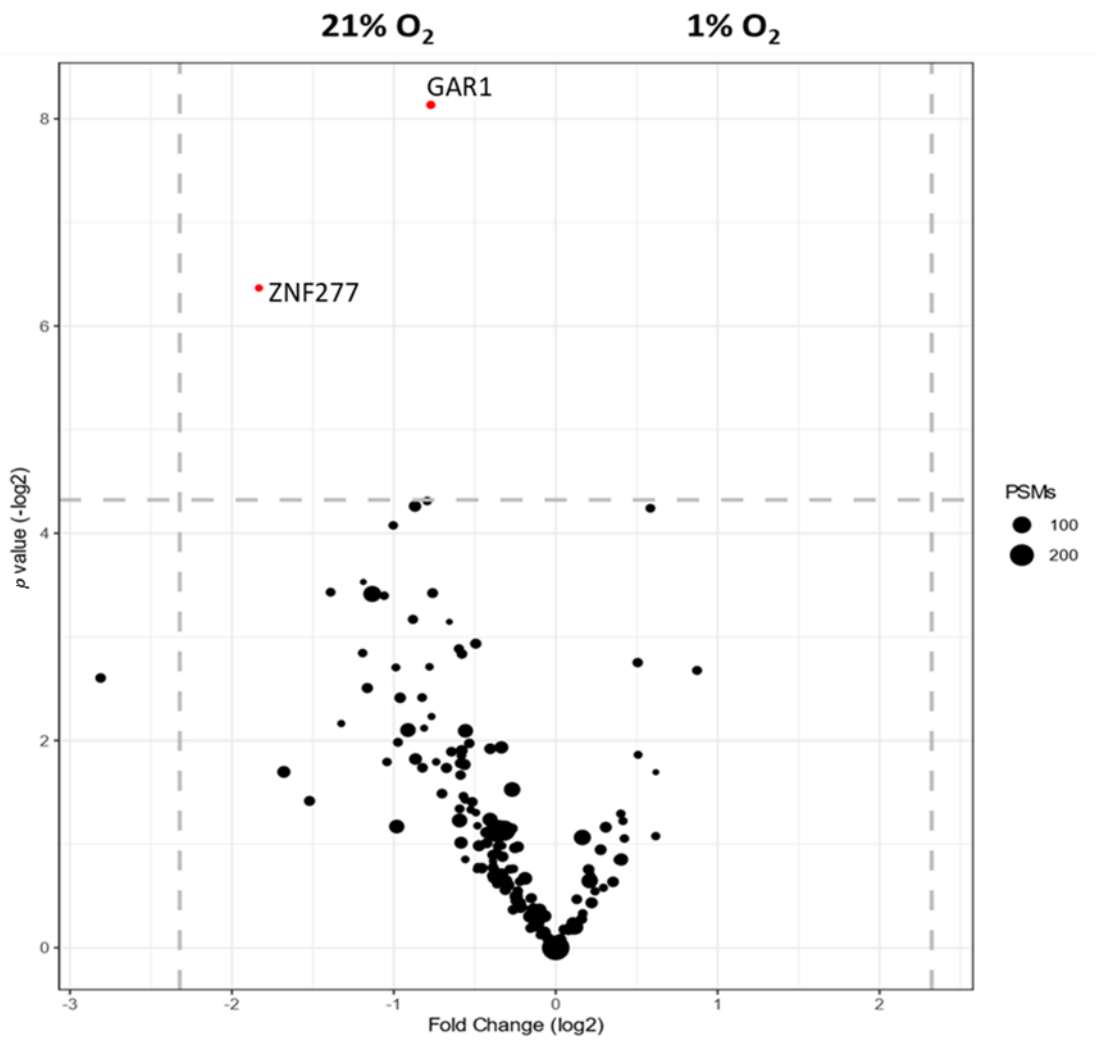

Figure 5.18 LFQ analysis of MROH6 binding partners in response to O_2 tension.
LFQ was performed using Proteome Discoverer 2.4 using the minora feature detector node set up for using Hi3 protein quantification method. Missing intensities were imputed by low abundance resampling and a must be identified in at least 2-of-3 replicates filter applied. Protein intensities were normalised to MROH6 protein intensity, per replicate and fold change calculated by 1% O_2 average/21% O_2 average. P-values were determined using a Student's 2-sample t-test. Plot is of –Log2 (P-value) against Log2 (fold change), where positive values are enriched in 1% O_2 and negative values enriched in 21% O_2. Dotted lines are at a p-value of <0.05 and 5-Fold enrichment. Points coloured red are significant at p-value <0.05. Point size is scaled to the total number of PSMs identifying a protein.

5.3.7.1. zDHHC23 binding partners

Somewhat surprisingly, although discrete categorization resulted in relatively equal numbers of proteins as 21% or 1% (102 and 76 proteins respectively) O_2 specific, LFQ analysis clearly identifies the vast majority of proteins as significantly more abundant in 21% O_2 (Figure 5.17) highlighting the importance of the two different analysis approaches applied. Interestingly, the two largest (and significant) fold changes for 21% O_2 are proteins part of the Actin Related Protein 2/3 complex (ARPC); ARPC2 (log2(Fold Change) -3.21, log2(t-test p-value) 5.75) and ARPC5 (log2(Fold Change) -3.14, log2 (t-test p-value) 10.74). Additional to these proteins, another two ARPC proteins are also significantly more abundant in 21% O_2; ARPC4 (log2(Fold Change) -2.82, log2(t-test p-value) 6.40) and ARPC1B (log2(Fold Change) -2.62, log2(t-test p-value) 4.32). These four proteins account for over half of the known ARP 2/3 complex (7 subunits, ARPC 1-7) which is shown to have major roles in cell migration of different cancers due to its actin nucleation function (Rauhala et al., 2013, Goley and Welch, 2006, Bretschneider et al., 2004, Kazazian et al., 2017). Hence, it is foreseeable, zDHHC23 may have important roles in regulating actin polymerisation that are lost by 1% O_2 incubation.

LFQ analysis of zDHHC23 also identifies two key clusters of interest that are more abundant in 21% O_2. Firstly, two G proteins (guanine nucleotide binding subunits) are identified in the top five most significantly differentially regulated interactors by O_2 tension, including: GNAI3 (log2(Fold Change) -1.49, log2(t-test p-value) 17.22) and GNAI2 (log2(Fold Change) 1.80, log2(t-test p-value) 12.90). Other G-protein are also identified, including GNG12 (log2(Fold Change) 1.48, log2(t-test p-value) 11.06) and GNA12 (log2(Fold Change) 1.08, log2(t-test p-value) 8.98). While the identification of G-proteins is reassuring that our exogenous expression system is functional (due to G-proteins and GPCRs (G-protein coupled receptors) being common targets of palmitoylation (Qanbar and Bouvier, 2003) it is intriguing to observe their preferential binding in 21% O_2. Additionally, these identifications are of potential importance due to being commonly observed as fundamental regulators of tumour growth and metastasis (Dorsam and Gutkind, 2007).

The second cluster of proteins I identify are related to the proteasomal complex, supporting previous observations with GO-term and STRING analysis (Figure 5.13 and Figure 5.15). Proteasomal proteins identified include (in order of highest significance); PSMB6 (log2(Fold Change) -1.96, log2(t-test p-value) 15), PSMD2 (log2(Fold Change) -1.33, log2(t-test p-value) 14.68), PSMA3 (log2(Fold Change) -1.17, log2(t-test p-value) 12.76), PSMA1 (log2(Fold

Change) -1.44, log2(t-test p-value) 11.62), PSMA4 (log2(Fold Change) -1.28, log2(t-test p-value) 10.24) and PSMC5 (log2(Fold Change) -1.31, log2(t-test p-value) 10.90). These proteasomal proteins are subunits of both the 20S or 26S proteasome complexes, particularly interesting with the 20S proteasome linked to NB through mediating neuronal communication through the co-translational cleavage of nascent proteins (Olshina et al., 2018, Ramachandran et al., 2018).

Interestingly, only three interactors were identified to be more abundant by 1% O_2 incubation: REPIN1 (log2(Fold Change) 1.11, log2(t-test p-value) 7.25), P4HA1 (log2(Fold Change) 1.38, log2(t-test p-value) 7.03) and LSM14A (log2(Fold Change) 1.53, log2(t-test p-value) 5.88). While these are significant and all >2.2-fold changes, it is difficult to identify any known pathways or functions which connect these together. It is worth commenting that a 1% O_2 interactor Solute carrier family 2, facilitated glucose transporter member 1 (SLC2A1) has, by far, the greatest fold change difference of any protein being ~175X more abundant (Fold enrichment = 174). However, SLC2A1 is not significantly different, interrogating the data further there appears a tag orientation specific effect with zDHHC23-mCherry-HA having an ~270-fold induction of binding by 1% O_2 while HA-mCherry-zDHHC23 only experiences a fivefold induction, although statistical analysis also does not identify this as significant; likely due to a single replicate being >100 times greater than the other two replicates (Appendix 7).

5.3.7.2. MROH6 binding partners

LFQ analysis of MROH6 identifies that only two interactors were significant, both of which were more abundant in 21% O_2: GAR1 (log2(Fold Change) -0.77, log2(t-test p-value) 8.13) & Zinc Finger Protein 277 (ZNF277) (log2(Fold Change) 1.83, log2(t-test p-value) 6.37) (Figure 5.18). Interestingly, both GAR1 and ZNF277 can be functionally linked through telomere maintenance, where GAR1 is crucial for extending telomeres and involved in the biogenesis of telomerases (Gall et al., 1999, Egan and Collins, 2010). Additionally, ZNF227 is important for regulating cellular senescence, potentially, induced by telomere shortening and is observed upregulated in chronic lymphocytic leukaemia, colon neoplasia, and endocrine tumours (Negishi et al., 2010, Cheng et al., 2014). These findings are of particular interest due to a phenotypic hallmark of high-risk NB being the activation of telomere maintenance mechanisms (Ackermann et al., 2018).

Using the LFQ data, I was able to investigate the relative abundance of each binding partner versus the amount of target protein, in this case MROH6. The top three proteins which bind most commonly to MROH6 were SPLATLS, BCAR1 & BCAR3 (Appendix 7). These three proteins were O_2-independent in binding intensity (highlighting why they were not identified by previous analyses comparing O_2 tensions). SPLATLS was highly affected depending on construct used, with an average abundance value of 1.5 (greater than MROH6) for MROH6-MHA and 0.7 for HAM-MROH6, suggesting a tag orientation effect and therefore was not analysed further. BCAR1 & BCAR3 were tag orientation independent, with BCAR1 average abundance value of 0.4 for MROH6-MHA and 0.3 for HAM-MROH6 and BCAR3 average abundance being 0.2 for both N- and C- terminal constructs. All other proteins had an average abundance <0.1 (Appendix 7). Due to the high abundance of BCAR1/3 these were investigated further.

BCAR1 and BCAR3 are known to regulate cell adhesion, motility and migration ((Mao et al., 2020, Wilson et al., 2013) with the protein abundance of the BCAR1-BCAR3 complex used an indicator of malignancy in breast cancer cell lines (Near et al., 2007). It is therefore interesting to see such a high abundance of these proteins Co-IP'd with MROH6. This could suggest that MROH6 affects NB cell growth and metastasis, providing insight into why a high expression of MROH6 correlates with poor prognosis in NB patients. It is noted that BCAR1 also appeared in our STRING analysis as interacting with the Nucleolar and Coiled-Body Phosphoprotein (NOLC1). Firstly, due to it being observed specifically 1% O_2 (Figure 5.16); NOLC1 is a 1% O_2 specific interactor of MROH6 while BCAR1/3 are identified O_2-independently. The identification of NOLC1 as 1% O_2 (aggressive phenotype) specific is pertinent to NB with literature stating that ubiquitination of NOLC1 is crucial for neural crest cell (NCC) specification (Zlotorynski, 2015). As discussed in Section 1.4.1, NB is hypothesised to occur due to NCCs failing to differentiate during embryonic development (Maris, 2010), therefore the presence of NOLC1, combined with the interaction with BCAR1/3 (known to be influence tumour growth and invasion) suggests that MROH6 may have vital roles in more aggressive NB tumours.

5.3.8. Validation of key binding partners

Following LC-MS/MS-based identification and subsequent bioinformatics characterisation of the interaction networks of zDHHC23 or MROH6, key proteins deemed to be of particularly interest were selected for validation as interaction partners by Co-IP and immunoblotting

validated. For zDHHC23, the proteins of interest included: PSMB6 (subunit of the proteasome cluster with the most significant 21% O_2 increase) and ARPC5 (subunit of the ARP2/3 complex with the biggest fold change compared to 1% O_2) (Figure 5.19), and for MROH6: GAR1 & ZNF277 (the only two significant regulated by O_2 tension, both more abundant in 21% O_2) and BCAR1 (due to its high abundance relative to MROH6 abundance) (Figure 5.20). HA immunoblotting was used as a positive control for all IPs It is important to note that 21% O_2 and 1% O_2 samples were analysed as separate membranes, hence no conclusions can be drawn between O_2 tensions due to possible transfer issues and different imaging parameters. Of note, the input lysate samples account for 2.5% of the total, while elution equates to 25%, i.e. a 10-fold increase in amount was loaded for interrogation.

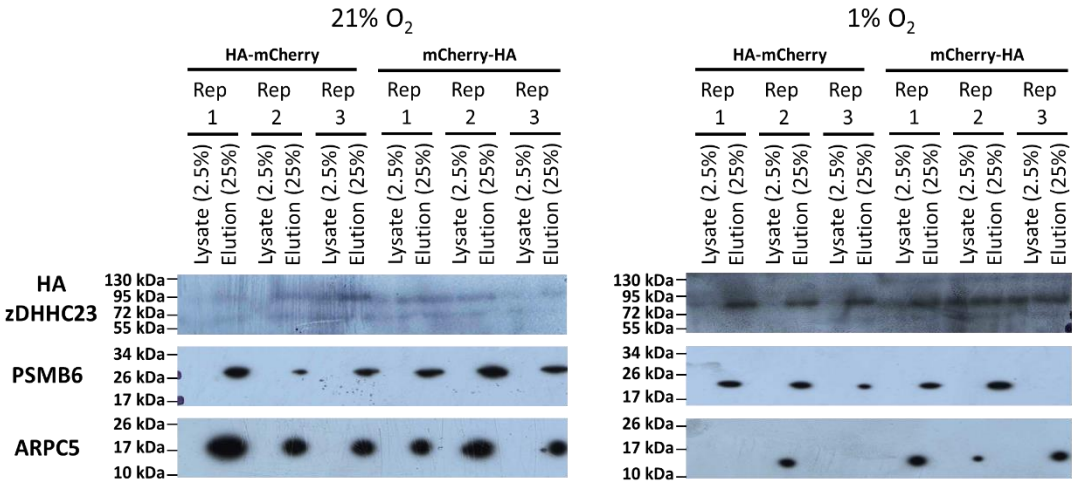

Figure 5.19 Western blot validation of key zDHHC23 binding partners in 21% and 1% O_2.
Western blot analysis on triple replicate of GFP-SK-N-AS cells transfected with either HA-mCherry zDHHC23 or zDHHC23 mCherry-HA plasmid 72 hr in 21% or 1% O_2 prior to lysis and IP from 500ug. Membranes were probed with corresponding antibodies (Table 1.4); HA antibody (1:1000), ARPC5 (1:500), PSMB6 (1:500).

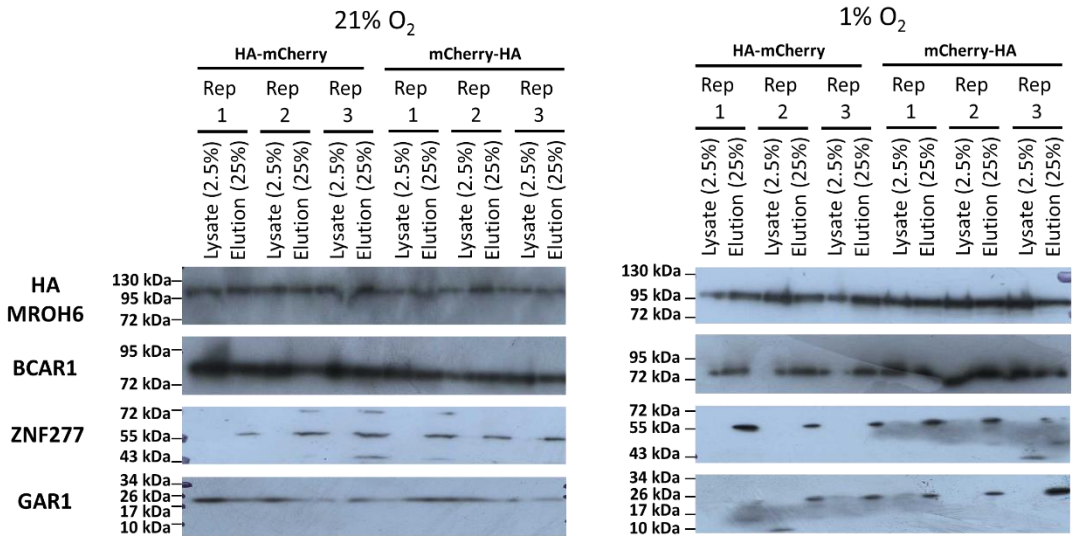

Figure 5.20 Western blot validation of key MROH6 binding partners in 21% and 1% O_2.
Western blot analysis on triple replicate of GFP-SK-N-AS cells transfected with either HA-mCherry MROH6 or MROH6 mCherry-HA plasmid 72 hr in 21% or 1% O_2 prior to lysis and IP from 500ug. Membranes were probed with corresponding antibodies (Table 1.4); HA antibody (1:1000), GAR1 (1:500), ZNF277 (1:500) and BCAR1 (1:500).

For zDHHC23, western blotting shows that PSMB6 (~25 kDa) was consistently identified in all elution lanes in an O2-independent manner (Figure 5.19). However, ARPC5 (~16 kDa) appears to be a 21% O_2 specific interactor (validating the LFQ data to some degree) due to its abundant and consistent banding pattern in 21% O_2 elution lanes whereas in 1% O_2 it is inconsistent (four out of six) and lysate specific banding (Figure 5.19). For MROH6, BCAR1 (93 kDa) and ZNF277 (53 kDa) are clear O_2-independent interactors due to strong and consistent binding in all elution lanes (Figure 5.20). While GAR1 is an obvious 1% O_2 binding partner with consistent elution bands, the 21% O_2 data is more difficult to interpret with similar (if not greater) banding intensities generally observed in lysate input lanes than elution. While this may suggest GAR1 binds preferentially under 1% O_2 (contradicting my LFQ analysis), it is important to remember that we cannot draw conclusions on banding intensities between 21% and 1% O_2 due to being analysed on different gels; hence it is equally valid that in 21% O_2 GAR1 is substantially more abundant than in 1% O_2 and only a subpopulation, that is maintained in 1% O_2, interacts with MROH6. However, overall, we can validate GAR1 as an interactor of MROH6.

5.3.9. Investigating cell localisation of zDHHC23 and MROH6

zDHHC23 subcellular localisation has been investigated, with contradicting reports, one identifying localisation within the Golgi (Tian et al., 2012) and another identifying the nucleus as the localisation site (Zhang et al., 2021b). I decided to utilise the fluorescence reporter gene mCherry I cloned into plasmids to allow for fluorescence-based cell imaging experiments to unequivocally determine cellular localisation. MROH6 subcellular localisation has never been investigated. Therefore, I explored sub-cellular localisation as a function of O_2 tension. Fluorescent confocal microscopy was performed using the nuclear dye Hoescht and cells stably expressing GFP to aid in cell compartmentalisation for cells expressing either N- and C- terminally tagged HA-mCherry- zDHHC23 (Figure 5.21) and MROH6 (Figure 5.22).

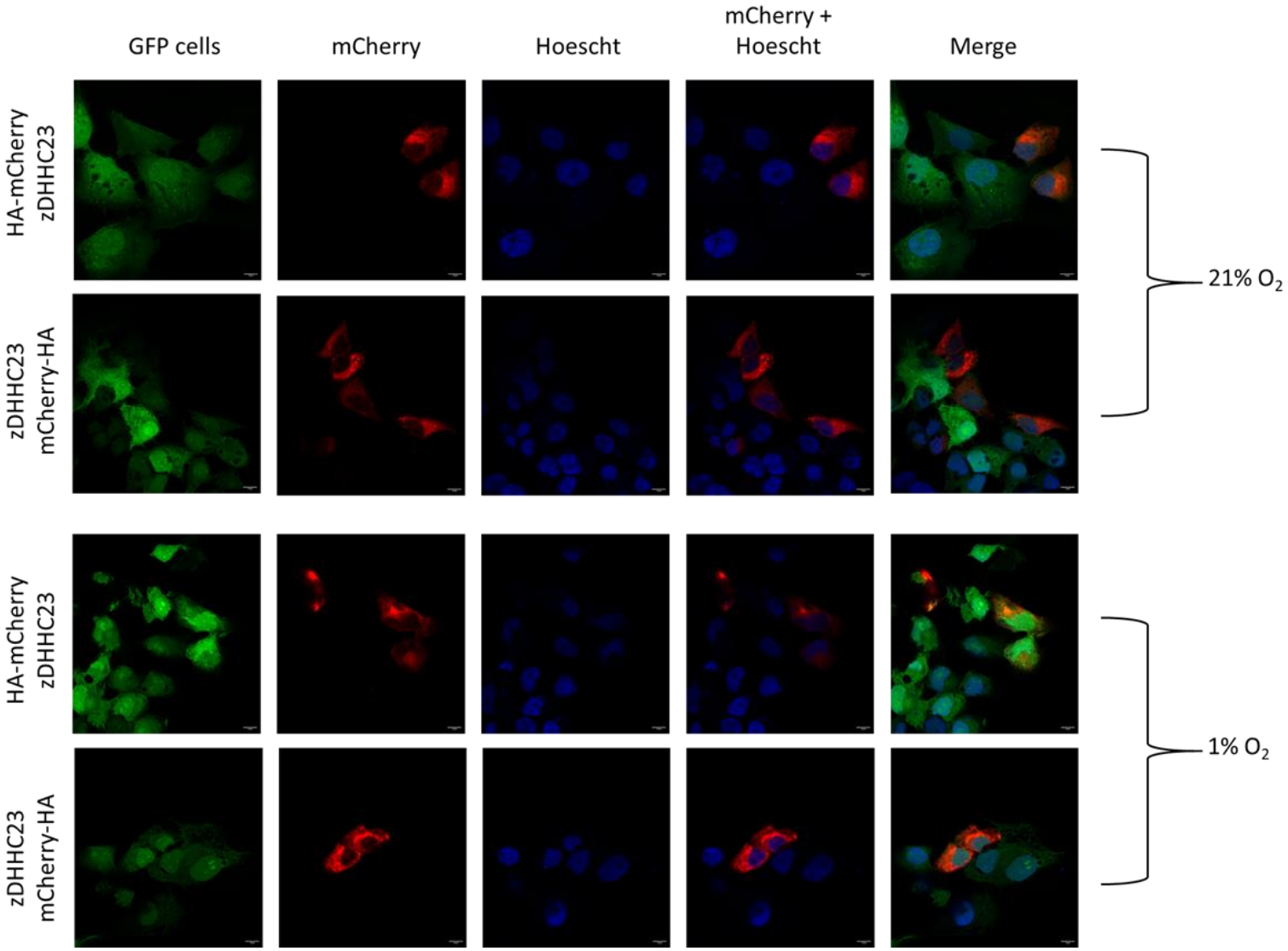

Figure 5.21 Exogenous zDHHC23 cell localisation in 21% and 1% O_2.

GFP SK-N-AS cells transfected using PEI with mCherry-HA or HA-mCherry zDHHC23 plasmid and incubated at 21% or 1% O_2 for 72 hrs post transfection. Cells were fixed with 4% PFA before a Hoescht nuclear stain prior to confocal microscopy. Fluorescence images (405 nm ,488nm, 587nm) 63x magnification obtained on Zeiss LSM780. Scale bar = 10um.

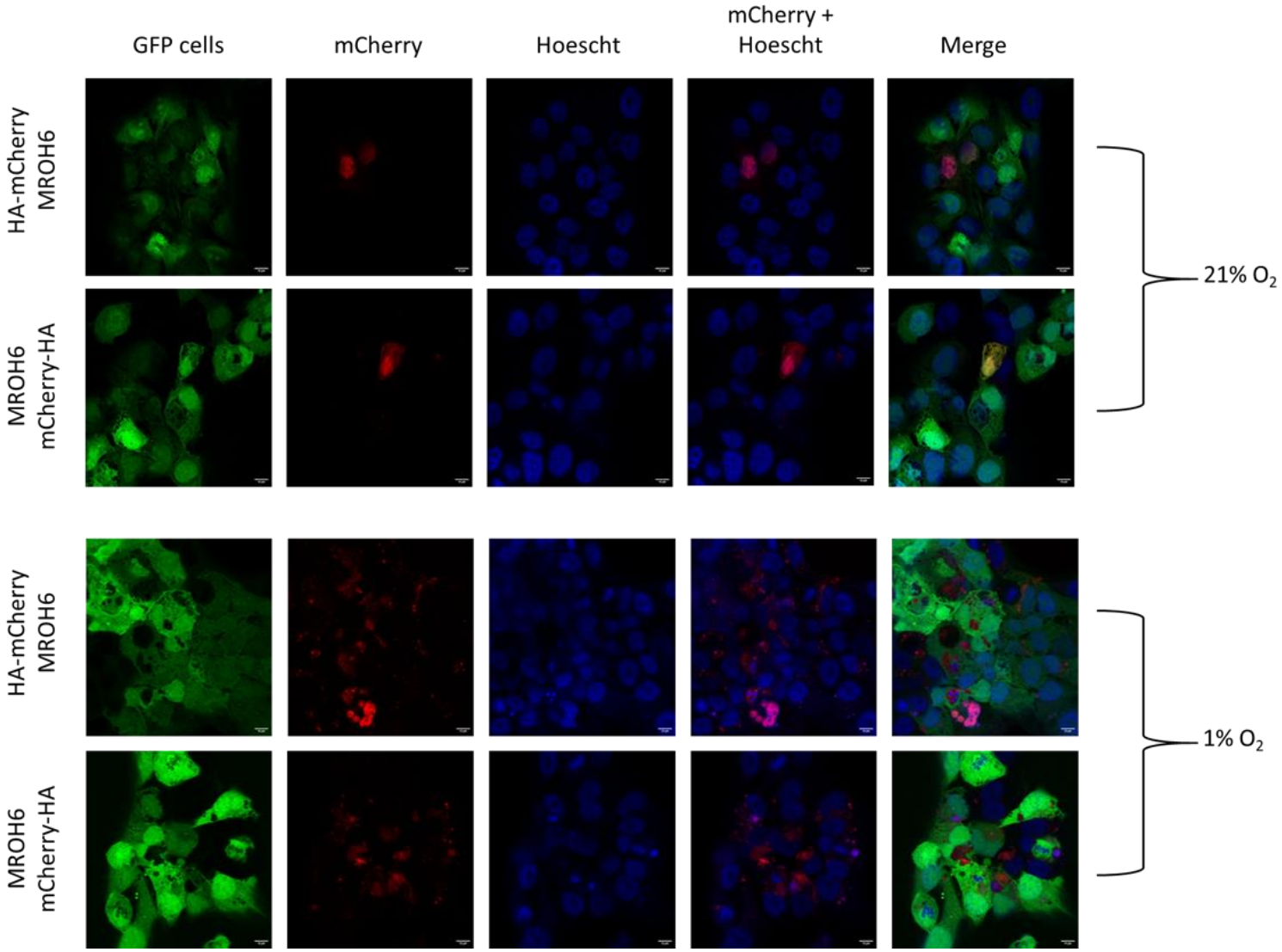

Figure 5.22 Exogenous MROH6 cell localisation in 21% and 1% O_2.
GFP SK-N-AS cells transfected using PEI with mCherry-HA or HA-mCherry MROH6 plasmid and incubated at 21% or 1% O_2 for 72 hrs post transfection. Cells were fixed with 4% PFA before a Hoescht nuclear stain prior to confocal microscopy. Fluorescence images (405 nm,488nm,587nm) 63x magnification obtained on Zeiss LSM780. Scale bar = 10um.

IF-confocal live cell imaging of zDHHC23 clearly shows mCherry (i.e. zDHHC23) signal is excluded from the nucleus regardless of tag-orientation and O_2 tension, contradicting (Zhang et al., 2021b). I identify that zDHHC23 is strongly peri-nuclear with signal dispersing further into the cytoplasm (Figure 5.21), commonly associated with Golgi associated proteins (Lipsky and Pagano, 1985); thus supporting (Tian et al., 2012). To validate zDHHC23 Golgi localisation attempts were made to stain Giantin, a known Golgi specific protein (Linstedt and Hauri, 1993), however I was unsuccessful with this approach and due to time restraints lacked time to optimise staining. For MROH6, IF cell imaging identifies that, independently of tag orientation, O_2 tension plays a significant role in MROH6 subcellular localisation; where MROH6 is essentially nuclear only in 21% O_2 (overlapping mCherry and Hoescht signal) yet has substantial cytoplasmic signal induced by 1% O_2 (Figure 5.22). Additionally, there appears to be increased signal (indicative of abundance) in 1% O_2 versus 21% O_2 in a tag-independent manner (Figure 5.22). While it is possible that increased expression may be due to differential transfection efficiency between 21% and 1% O_2, this observation is not observed for zDHHC23 which were in identical plasmid backbones and transfected and imaged simultaneously to MROH6, reducing the likelihood of this being an explanatory factor. Due to time restraints, the O_2 tension MROH6 protein expression was no further investigated.

5.4. Discussion

This chapter describes the most in detail investigation into the molecular characterisation of pathways that may regulate, or be regulated by, zDHHC23 and MROH6 to my knowledge. Importantly, it is also the only investigation that considers O_2 tension as a variable. Using extensive bioinformatic characterisation, some of which I have validated now opens the path for further investigation of physiological function and their roles in NB. Therefore, it is exciting to contemplate how this data may influence future more targeted investigations. While as many experimental conditions and controls were employed as possible with the technological and time restraints present, there are considerations that require acknowledgement and are discussed separately below.

5.4.1. PTM data

While I was able to identify some PTMs on zDHHC23 (five in total; three phosphorylation and two oxidation) and MROH6 (two in total; one phosphorylation and one oxidation) it is likely there are multiple other PTMs that I have failed to identify. The reasoning of this could be due to two reasons: 1) although protease optimisation was performed (Section 4.3.2) I was only able to obtain ~50% sequence coverage for either protein i.e. 50% un-analysed that can contain PTMs. 2) Due to its ease of application and being the second most abundant PTM, I opted to specifically investigate phosphorylation by TiO_2 enrichment.

Interestingly, the high number of oxidation (molecular O_2-dependent) PTMs observed are generally identified in 1% O_2. While this observation can be partially explained by the increased presence of ROS due to oxidative stress in 1% O_2, this does not explain the observation (only in a single replicate) of proline hydroxylation at 21% O_2 in zDHHC23 P22. This is particularly interesting considering its only known role in HIFα degradation (Ivan et al., 2001) and initial identification of zDHHC23 through increased (albeit at the mRNA level) levels in 1% O_2. A recent study that generates the most in-depth HIFα PTM maps to date in the Eyers lab (Daly et al., 2021) fails to identify the highly characterised proline hydroxylation as a PTM, arguing the rapid induced degradation likely prevents its accumulation to MS-detectable levels without any enrichment. Therefore, identifying it here warrants further investigation. If more time was available, site-directed mutational analysis combined with functional assays (such as proliferation, migration flow cytometry for cell cycle analysis) would be a priority for investigation these PTMs, particularly with the cell cycle association of zDHHC23 S252.

5.4.2. Binding partners

As discussed in Chapter 3, unfortunately current technology prevails my attempts to investigate zDHHC23 and MROH6 endogenously by IP. Hence, I opted for an exogenous HA-based IP protocol to identify binding partners for both zDHHC23 and MROH6 in both 21% and 1% O_2. While this research significantly expands our (the fields) current knowledge of zDHHC23 and MROH6 functions and roles specifically in NB, the underpinning issue of exogenous expression means biological relevance can be questioned until validated at the

endogenous level. Therefore, I applied highly strict filtering parameters to specifically identify interactors that are tag orientation independent, however this results in dramatic losses to the number of proteins identified (e.g. for zDHHC23 the final data used for all analysis consisted of between 50-65% less identifications than either tag orientation); hence while it would have been interesting to compare tag-orientation specific effects to increase the potential interactome, I decided this would be semi-irrelevant with such a dataset already generated. It is pertinent to mention that I fail to identify the only zDHHC23 interacting protein published, KCNMA1 (Duncan et al., 2019), which may be a result of cell-type specific differences.

Overall, I identify multiple different putative pathways that may somewhat explain the aggressive phenotype observed by high expression of zDHHC23 or MROH6 in NB, including: cytoskeletal reorganisation, cell adherence and cellular senescence.

5.4.3. Cell localisation

While I identified that zDHHC23 is likely O_2 tension independently localised to the Golgi (supporting published data, (Tian et al., 2012)), it is also possible that zDHHC23 is ER localised, therefore additional counter staining with Golgi/ER specific antibodies is required before any final conclusion is drawn. However, the subcellular localisation of MROH6 is potentially more intriguing with O_2 tension dependent localisation to the nucleus and cytoplasm observed. While many proteins are known to shuttle between the nucleus and cytoplasm, these generally have different functions/outcomes of the protein (Jeffery, 1999). Therefore, it would be interesting to validate this 1% O_2 induced tendency for cytoplasmic localisation and secondary use subcellular fractionation protocols to specifically investigate the differences between cytoplasmic and nuclear localised MROH6 to decipher if this differential localisation may aid to explain the hypoxia induced aggressive phenotype of MROH6.

5.4.4. Future experimentation

As discussed, all experimentation within this chapter relied on the exogenous expression of HA-mCherry tagged zDHHC23 and MROH6. To circumvent potential issues associated with exogenous expression, it would be useful in future to generate custom antibodies to facilitate

immunoblotting, IF and IP of endogenous protein. Alternatively, the creation of CRISPR (clustered regularly interspaced short palindromic repeats) cell lines could be utilised to knock-in purification tags (allowing endogenous purification and staining) or knock-out the zDHHC23/MROH6 gene to investigate molecular effects on an endogenous scale. These were not explored here due to cost associated and the extensive characterisation of cells required when creating CRISPR cell lines to ensure no off-target effects are accidentally introduced.

5.4.5. Concluding remarks

This study is the largest investigation into zDHHC23 and MROH6 and has successfully identified multiple novel molecular pathways that integrate with zDHHC23 and MROH6 either by being regulated by or regulating the proteins of interest. While data of interest likely needs validating, it is exciting to consider how these datasets may unlock new avenues of research into zDHHC23 and MROH6 functions, potential roles in cancer generally and the possibility of therapeutic translation in the future.

Chapter 6: Elucidating zDHHC23 role as a palmitoyltransferase in neuroblastoma

6.1 Introduction

As discussed in Section 1.1.3, protein palmitoylation is a reversible lipid modification whereby palmitate, a C16:0 palmitic acid, is added to either a Ser residue via a hydroxy-ester bond or a Cys residue via a thioester bond (Linder and Deschenes, 2007). Palmitoylation is primarily thought to localise proteins to membranes (plasma or organelle). However, more recent evidence suggests potential roles for palmitoylation in regulating protein–protein interactions (Huang et al., 2022). In the past decade, the importance of protein palmitoylation in tumorigenesis has started to be highlighted, with palmitoylation occurring on many key cancer-related proteins including those involved in cell proliferation, survival, invasion, and metastasis. As a result, palmitoylation and the role of ZDHHCs in a wide variety of cancers has been investigated (reviewed in (Zhou et al., 2022).

The importance of palmitoylation signalling can be observed in influencing pro- and anti-oncogenic properties, with some examples including:

1) epidermal growth factor receptor (EGFR) signalling, where palmitoylation has been described to play a vital role in the progression of breast cancer (Bollu et al., 2014), prostate cancer (Thomas et al., 2019), and non-small cell lung cancer (Ali et al., 2018).

2) Claudin 3 is palmitoylated (by zDHHC12 at Cys103, Cys106, Cys181, Cys182, and Cys184) which regulates the proteins stability, localisation and subsequently determines its role in ovarian cancer progression (Yuan et al., 2020).

3) CD82 (a known cancer metastasis suppressor (Liu and Zhang, 2006)) is palmitoylated by zDHHC4 resulting in localisation to the cell membrane, which induces the production of leukaemia inhibitory factor thereby preventing angiogenesis and tumour survival (Lee et al., 2021a). Additionally, CD82 has also been shown to be palmitoylated at multiple sites (Cys5, Cys74, Cys83, Cys251 and Cys253) in prostate cancer cells which resulted in the inhibition of cell migration and invasion (Zhou et al., 2004).

Breast cancer models have been extensively used to investigate protein-palmitoylation function and truly highlight the “double-edged sword” effect of palmitoylation signalling in

the regulation of metastasis. (Babina et al., 2014) showed that palmitoylation of CD44 at Cys286 and Cys295 inhibited cell migration. In contrast, palmitoylation of integrin b4 (ITGb4) at Cys732, Cys736, Cys738, Cys739, and Cys742 by zDHHC3 promotes the invasiveness of breast cancer cells (Coleman et al., 2015). Additionally, palmitoylation of phosphatidylinositol 4- kinase IIa (PI4KIIa) accelerates tumour growth (via regulation of subcellular localisation and catalytic activity) which has led to development of a small molecule inhibitor which targets the site of palmitoylation and results in significant inhibition of cell growth (Barylko et al., 2009) & (Li et al., 2017).

At the time of writing this thesis, only one known substrate has been identified that is palmitoylated by zDHHC23. Using small-interfering RNAs to supress zDHHC23 expression, Tian et al. demonstrated a reduction in palmitoylation of the pore-forming alpha subunit of large-conductance calcium- and voltage-activated potassium (BK) channels (Tian et al., 2012). In this chapter, I utilised a click chemistry reaction approach to identify additional zDHHC23-mediated palmitoylated proteins, enhancing our understanding of zDHHC23-mediated signalling.

Click chemistry as a concept was first introduced in 2001 by K. B. Sharpless (who was subsequently awarded the 2022 Nobel Prize in Chemistry) and illustrated how organic and bioorganic molecules could be coupled (Kolb et al., 2001). Click chemistry refers to a group of highly reliable, efficient and selective reactions that can be utilised to develop constructs by simplistic reaction conditions (Hou et al., 2012). There are four such reactions: nucleophilic substitution reaction (electron rich nucleophile reacts with a positively charged electrophile to replace a leaving group); carbonyl condensation (a nucleophile reacts with a electrophile of a second carbonyl compound (1,2-addition reaction) forming a new C-C bond); addition reactions (addition to carbon-carbon multiple bonds such as amine-ene and thiol-ene) and of interest and implicated in this chapter - cycloaddition reaction (Kaur et al., 2021). The foundation of cycloaddition reactions is an azide-alkyne Huisgen 1,3-dipolar cycloaddition. The relatively simplistic reaction conditions results in high reaction efficiency and is highly specific; it is therefore a popular choice in experiments coupling different types of organic building blocks (Hong et al., 2019).

While the original design of Huisgen 1,3-cycloaddition reaction was slow (due to the lack of a catalyst) and required high temperatures (resulting in non-regioselective reactions), the method was optimised and ultimately modified by the addition of Cu(I) that acts as a catalyst allowing room temperature reactions and selectively synthesising the 1,4-regioisomer -

termed Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC, (Rostovtsev et al., 2002) (Figure 6.1). The proposed mechanism of action in CuAAC was described by Maarseveen and colleagues in 2006 (Bock et al., 2006), depicted in Figure 6.2. Typically this catalytic reaction requires a reducing agent like sodium ascorbate (Spiteri and Moses, 2010) or tris(2-carboxyethyl) phosphine (TCEP) (Zhong et al., 2016, Baskin et al., 2007), the latter employed in the results described in this chapter.

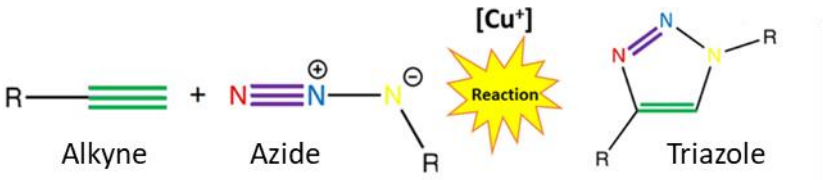

Figure 6.1 Schematic view of a CuAAC reaction.
Azide and alkyne analogues catalysed by copper to form a triazole.

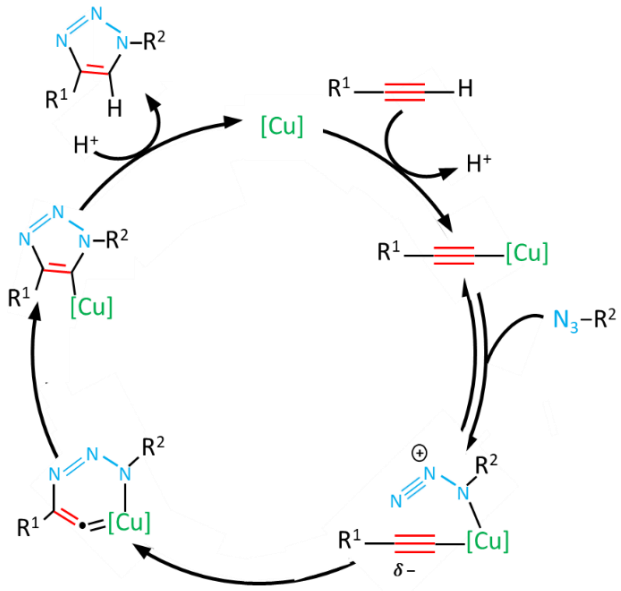

Figure 6.2 Proposed mechanism of CuAAC reaction.
Cycle begins with the formation of a Cu-alkyne π complex which is followed by the deprotonation of the alkyne proton which forms a copper acetylide. This copper increases the acidity of the acetylenic proton resulting in facilitated deprotonation within aqueous media. The copper acetylide is equilibriously distributed between monomeric and a dimeric species. A copper ion from the dimeric species interacts with azide nitrogen, activating it toward the terminal nitrogen of the azide group on the alkyne carbon resulting in the synthesis of the metallacycle. The metallacycle undergoes a ring contraction (transcellular interaction between lone pair of electrons on azide nitrogen and the carbon-copper bond), this results in a Cu triazolide species being formed which undergoes protonation to generate a 1,4-disubstituted triazole and Cu(I) catalyst (Zhu et al., 2016).

Due to the mild and highly specific reaction kinetics, CuAAC is the most potent, reliable, and fast acting synthetic tool for the synthesis of substituted triazoles, resulting in CuAACs being successfully exploited as pharmaceutical agents for medicinal chemistry (Moussa et al., 2018), biomedical imaging and drug discovery (Wu et al., 2018). However, there are some limitations to this method, which include Cu(I) saturation resulting in unwanted alkyne-alkyne side reactions, uncertainty of in-vivo safety requiring a Cu(I) catalyst (where copper is a known toxic). Additionally, some of the heavy metal azides are explosive resulting in safety concerns (Kluba and Mindt, 2013).

In this thesis, I have used a CuAAC based approach for enrichment zDHHC23-mediated palmitoylated proteins for identification by LC-MS/MS. SK-N-AS cells were dosed with an azide-tagged palmitic acid analogue and lysed by addition of alkyne functional group-linked magnetic beads in the presence of overexpressed active or catalytically dead zDHHC23. The workflow is diagrammatically depicted in Figure 6.3.

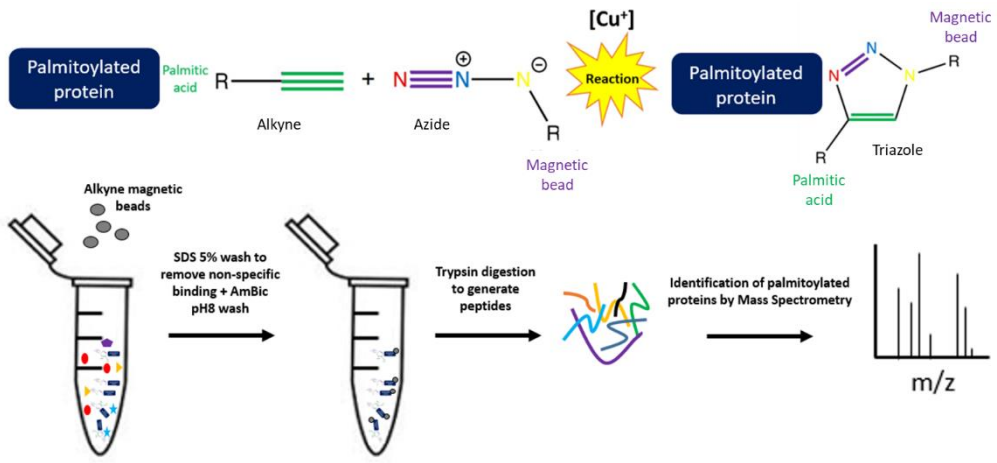

Figure 6.3 Schematic view of the palmitoyl-azide/alkyne assay for zDHHC23.
SK-N-AS cells were dosed with azide tagged palmitic acid, resulting in palmitoylated proteins containing the azide tag. Following cell lysis, magnetic beads with an alkyne tag were added with in the presence of $CuSO_4$ which catalyses the reaction, resulting in palmitoylated proteins covalently bound to the magnetic beads. Beads underwent high-stringency washing in 5% SDS, prior to on-bead digestions with trypsin and LC-MS/MS analysis of the released peptides.

6.2 Chapter Aims

The aim of this chapter was to optimise and ultimately develop a CuAAC protocol that was LC-MS/MS compatible to enable the identification of zDHHC23 palmitoylated proteins and thus further interrogate its role within NB.

6.3 Results

6.3.1 Cloning of enzymatically inactive zDHHS23 plasmid constructs

The creation of enzymatically inactive zDHHS23 plasmid constructs was performed to generate a suitable control for specific identification of zDHHC23 palmitoylation targets, considering that all endogenous palmitoyltransferases will also use the azide tagged palmitic acid and consequently be identified simultaneously. Figure 6.4 shows the schematic workflow and primers used in the mutagenesis (mutating the catalytically active Cys to an inactive Ser (Malgapo and Linder, 2021) via In-Fusion cloning to both the N- and C- terminal tagged constructs created in Chapter 4.

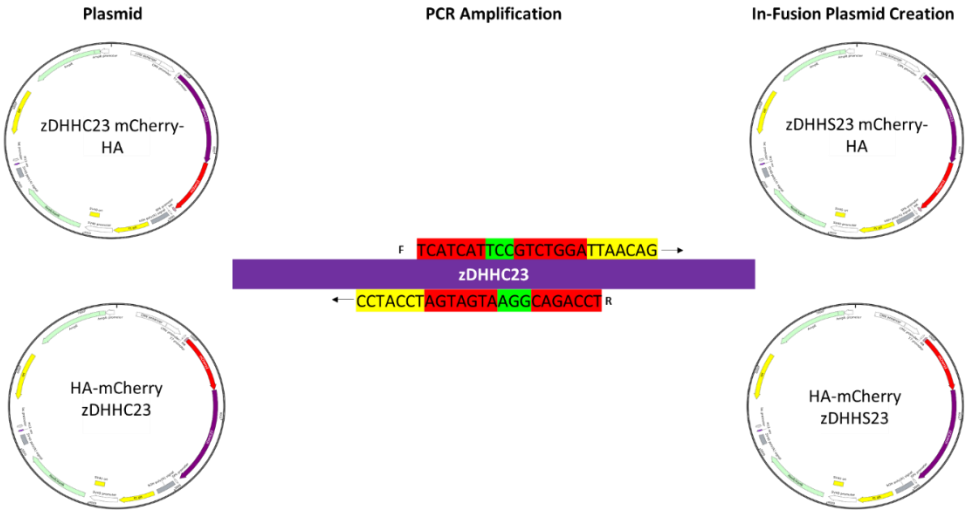

Figure 6.4 Schematic view of the Mutagenesis via In-Fusion cloning strategy used for the creation of zDHHS23 HA-mCherry and mCherry-HA plasmids.
Forward and reverse primer designs included: Red = overlapping region, Green = DHHC to DHHS mutation, Yellow = primer specific overhanging region. Plasmid maps viewed in SnapGene software (SnapGene software (from GSL Biotech; available at snapgene.com))

As described in Chapter 4, PCR temperatures were first optimised, using the CloneAmp HIFI PCR kit (Figure 6.5). For both the N- and C- terminal constructs, amplification appeared to be most efficient (brightest band on gel) at 62 °C and therefore subsequent large-scale PCR amplification was performed at this temperature and the PCR product in-gel purified (Section 2.7.3). Purified PCR products were inserted into the required plasmid using the In-Fusion cloning method, amplified and sequence verified (Sections 2.7.6 - 2.7.9).

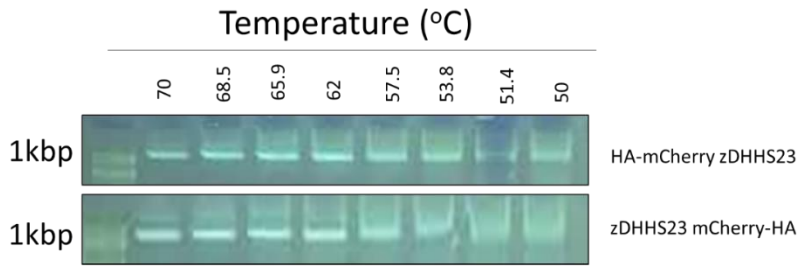

Figure 6.5 Identifying optimal PCR conditions for zDHHS23 plasmid coding region.
CloneAmp HIFI (ClonTech) PCR temperature gradient test (70-50°C) for both HA-mCherry and mCherry-HA zDHHS23 plasmids.

6.3.2 Optimisation of CuAAC protocol for efficient labelling

Following the creation of the inactive zDHHS23 control plasmids, the next step was to optimise the number of times cells needed to be dosed with azide tagged palmitic acid. A number of factors have the potential to influence the efficiency of the click-chemistry palmitoylation labelling including: 1) how long it takes for the incorporation of azide-linked palmitic acid, 2) whether azide-tagged palmitic acid may have adverse or long-term effects when incorporated into cells, 3) the turnover rate of palmitoylated proteins (i.e. would a single dose be sufficient for the 72 hr incubation) and 4) how hypoxia may influence all of the above. Therefore, SK-N-AS cells were dosed at time point 0 h, and then either left untreated, or re-dosed once at 24 h with or without an additional dose 24 h later. These cells were lysed and processed through the CuAAc protocol and LC-MS/MS analysis pipeline (Section 2.4.2, Figure 6.3), investigating the total number of proteins identified as a proxy for defining optimal dosing parameters (Figure 6.6).

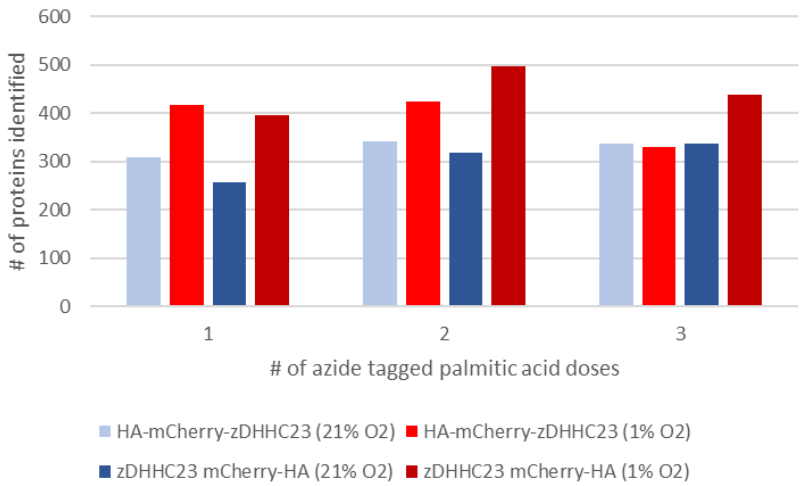

Figure 6.6 Number of proteins identified via LC/MS-MS following CuAAC with varying doses of azide-palmitic acid.
N or C terminal zDHHC23 constructs in normoxia and hypoxia were dosed with either one, two or three doses of azide-palmitic acid during 72 hr incubation. Peptides were analysed by LC-MS/MS using ThermoFisher Fusion Tribrid instrument (as detailed in Section 1.8) and searched against UniHumanReviewed (updated weekly) database using Proteome Discoverer. Bar chart details the number of proteins (identified applying a 1% FDR) under each condition (n=1).

Figure 6.6 shows that there was no substantive change in the number of proteins identified as a function of the number of doses applied. Interestingly, across all conditions (with the exception of the 3-dose experiment with HA-mCherry-zDHHC23) hypoxia resulted in a ~25% (~100) increase in the number of proteins identified compared with normoxia. Given these data and considering the potential unknown consequences of azide-linking toxicity and large excesses of free palmitic acid, we decided to use only one dose of palmitic acid for the final experiment. A single dose also limits the likelihood that excess azide-tagged palmitic acid could be used by other endogenous palmitoyltransferases and thus decreases the possibility of increasing false positives (noise). It must be noted that these results are from an n=1 due to time constraints and as such there is variability in the data. However, cell death was examined using a confocal microscope before processing for LC/MS-MS analysis, with increased cell death observed in multiple doses.

Having defined the number of palmitic acid doses, it was important to ensure the control zDHHS23 plasmids worked as intended. Therefore, I transfected cells with the active and inactive plasmids and compared the number of proteins identified. Additionally, a second control was included where SK-N-AS cells were transfected with active plasmid but was not dosed with any azide tagged palmitic acid - identifying 'noise' created by non-specific binding to the alkyne tagged beads (Figure 6.7).

Figure 6.7 shows that both orientations of the inactive enzyme resulted in ~200 fewer proteins identified by LC-MS/MS compared to the active zDHHC23 samples, indicating not only that the over-expressed zDHHC23 is active in both orientations, but also that zDHHS23 serve as appropriate controls for identification of zDHHC23 palmitoylated targets. This data also reinforces the importance in using suitable controls to understand background levels of palmitoylation in this type of assay given the considerable non-specific binding in the palmitic acid free controls.

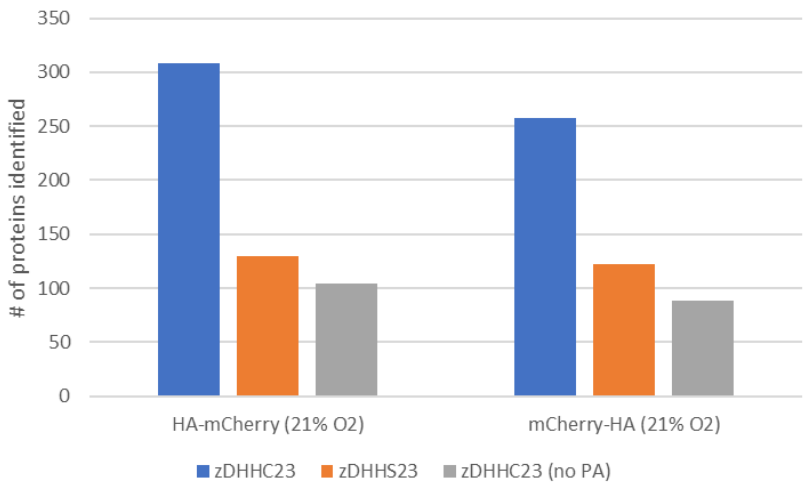

Figure 6.7 Background control testing for CuAAC.
SK-N-AS cells in normoxia (21% O_2) dosed once with azide-palmitic acid in 72 hr incubation, with transfection of N or C terminal zDHHC23 constructs or background controls (N or C terminal zDHHS23/zDHHC23 without azide-palmitic dosing). Peptides were analysed by LC-MS/MS using ThermoFisher Fusion Tribrid instrument (as detailed in Chapter 2) and searched against UniHumanReviewed (updated weekly) database using Proteome Discoverer. Bar chart details the number of proteins (identified applying a 1% FDR) under each condition (n=1).

6.3.3 Identification of zDHHC23 targeted palmitoylated proteins

Following CuAAC protocol optimisation, a three replicate experiment was conducted in the presence of either 21% or 1% O_2. Proteins identified in two out of three replicates for each orientation of zDHHC23, and not appearing in the control datasets (zDHHS23 or zDHHC23 in the absence of PA) were considered as zDHHC23 targets (Appendix 7). As discussed in chapter 5, both N- and C- terminal construct data were combined, providing a pseudo 4 out of 6 replicate datasets. Identifications under the two different oxygen tensions were then compared to identify if hypoxia-regulated palmitoylation targets (Figure 6.8).

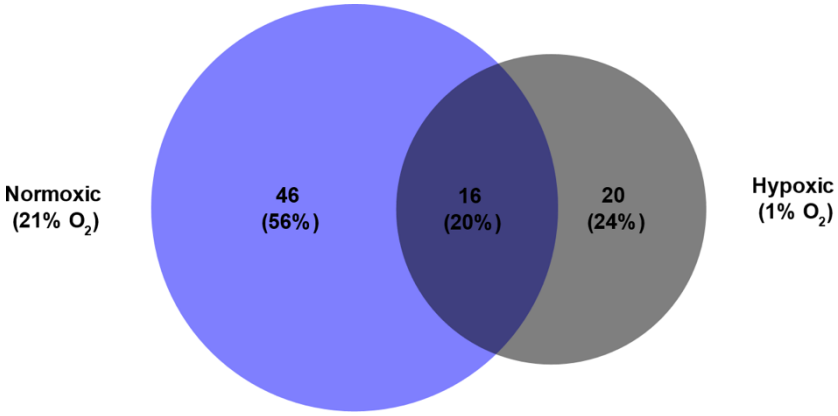

Figure 6.8 zDHHC23 mediated palmitoylated targets with both C and N terminal constructs combined in 21% and 1% O_2.
Venn diagrams of proteins identified in 21% O_2 versus 1% O_2 zDHHC23 with C and N terminal construct data combined with post background subtraction of enzymatically inactive zDHHS23 C and N terminal construct and non-azide palmitic acid treated samples. Purple = proteins specific to 21% O_2, grey = proteins specific to 1% O_2, dark purple = seen in both O_2 tensions. % Values reflect the proportion of palmitoylated targets identified per section, relative to the total protein count. Circles are scaled respectively to the number of palmitoylated targets per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

Figure 6.8 shows that the total number of putative zDHHC23 palmitoylated proteins identified using this assay is 82, with the majority (46 identifications, ~56%) identified as 21% O_2 specific, whilst 20 (~24%) were identified as 1% O_2 specific and 16 (~20%) as O_2 tension independent. It is worth noting that the number of proteins identified as zDHHC23 palmitoylated targets in this experiment is much lower than the optimisation studies (62 versus ~150 from Figure 6.7 - number obtained by subtracting the identifications of zDHHS23 from zDHHC23). As discussed in Chapter 5, this result is likely due to tag orientation specific differences and without knowledge on which is more suitable for investigation it is more reliable to use both to make conclusions. As described in Chapter 5, Gene Ontology (GO) enrichment analysis was performed using DAVID to identify the cellular pathways and cellular functions which zDHHC23 regulates and whether there are O_2-dependent differences (Figure 6.9, Figure 6.10 and Figure 6.11) (Appendix 8).

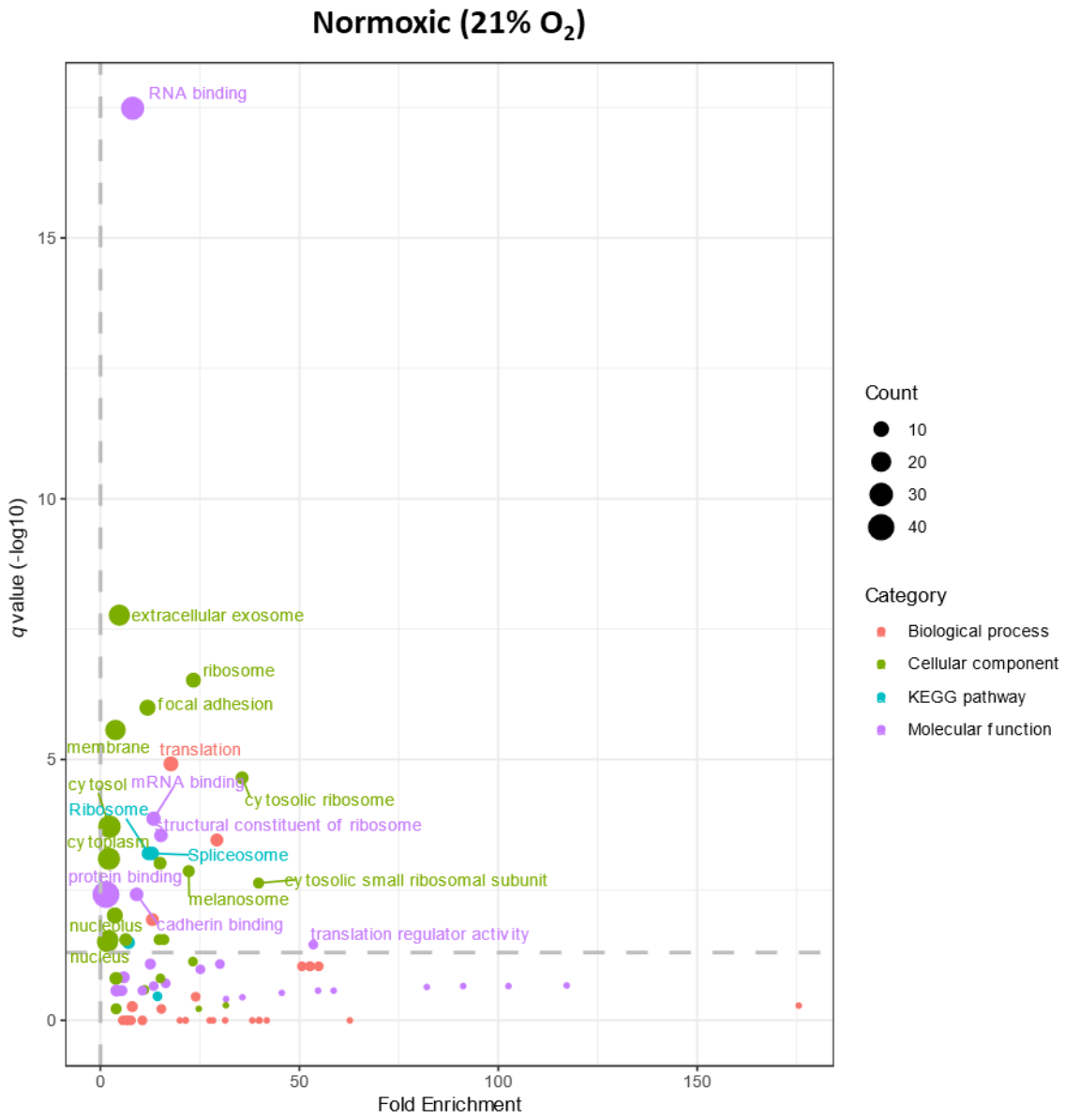

Figure 6.9 GO Enrichment Analysis of 21% O_2 zDHHC23 palmitoylated targets.
GO enrichment analysis was performed in DAVID using the list of interactors identified in two or more replicates of zDHHC23 with C and N terminal construct data combined. p values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg correction and are plotted against fold enrichment. The horizontal line represents an adjusted p value of 0.05. Points are sized according to the number of proteins within each enriched term and coloured according to the enrichment category.

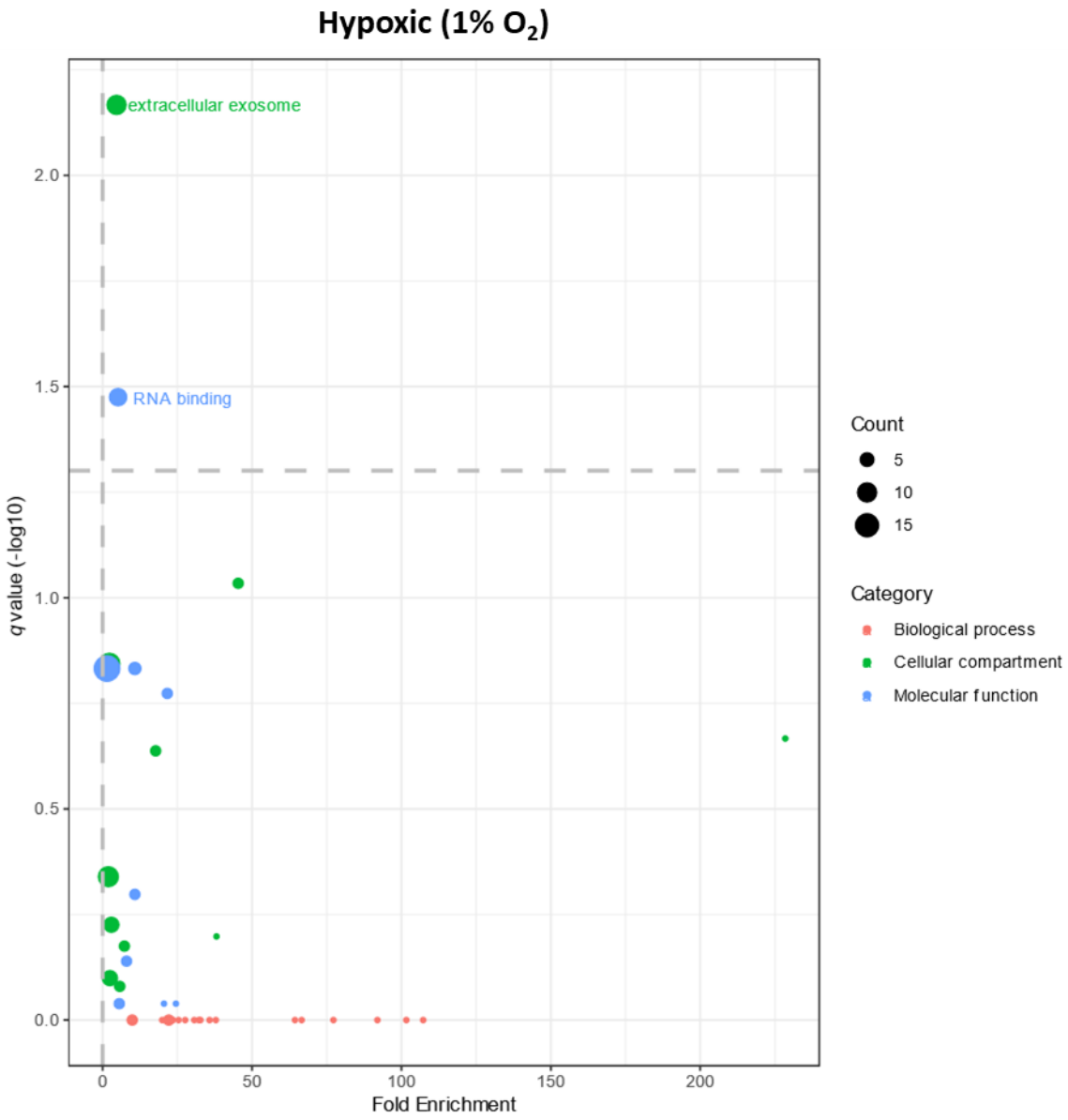

Figure 6.10 GO Enrichment Analysis of 1% O_2 zDHHC23 palmitoylated targets.
GO enrichment analysis was performed in DAVID using the list of interactors identified in two or more replicates of zDHHC23 with C and N terminal construct data combined. p values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg correction and are plotted against fold enrichment. The horizontal line represents an adjusted p value of 0.05. Points are sized according to the number of proteins within each enriched term and coloured according to the enrichment category.

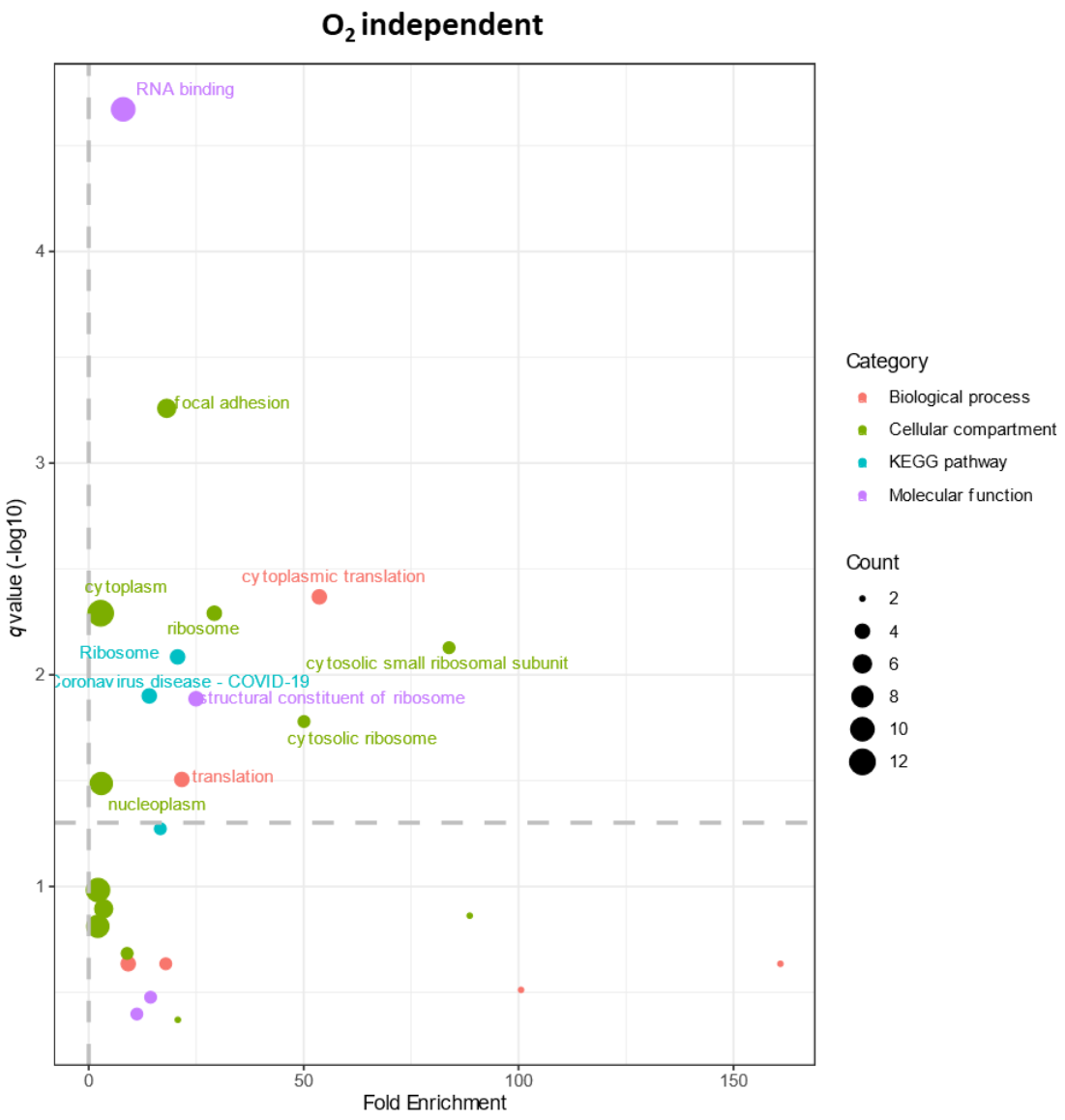

Figure 6.11 GO Enrichment Analysis of O_2-independent zDHHC23 palmitoylated targets.
GO enrichment analysis was performed in DAVID using the list of interactors identified in two or more replicates of zDHHC23 with C and N terminal construct data combined. p values were adjusted for multiple hypothesis testing using the Benjamini-Hochberg correction and are plotted against fold enrichment. The horizontal line represents an adjusted p value of 0.05. Points are sized according to the number of proteins within each enriched term and coloured according to the enrichment category.

Comparing GO-terms for proteins identified under either normoxic (Figure 6.9) or hypoxic (Figure 6.10) conditions, it is clear that far fewer terms were significantly enriched in hypoxia, likely due to the lower numbers of proteins identified. However, it is interesting that extracellular exosome is enriched for both normoxic specific (Benjamini-Hochberg q-value = 1.70E-08) and hypoxic specific (Benjamini-Hochberg q-value = 6.80E-03) identifications, yet not for O_2-independent identifications (Figure 6.11) indicating that a major role of zDHHC23

is in extracellular trafficking and that the protein targets significantly differ in an oxygen dependent manner.

Another point of interest is the lack of GO-term identifications related to cellular structure/motility under hypoxic conditions (Figure 6.10). While O_2-independent analysis identifies focal adhesion as a significant term (Figure 6.11) (Benjamini-Hochberg q-value = 5.50E-04), normoxic specific identifies multiple GO-terms including focal adhesion (Benjamini-Hochberg q-value= 1.01E-06), membrane (Benjamini-Hochberg q-value = 2.71E-06) and cadherin binding (Benjamini-Hochberg q-value = 3.83E-03); overall suggesting hypoxic-induced loss of zDHHC23-dependent cellular motility. However, surprisingly, the most common O_2-independent and significant GO-terms involve ribosomal proteins, RNA binding: 21% O_2 (Benjamini-Hochberg q-value = 3.26E-18, Count 29 = 63%), 1% O_2 (Benjamini-Hochberg q-value = 3.35E-02, Count 8 = 40%), O_2-independent (Benjamini-Hochberg q-value = 2.13E-05, Count 10 = 63%), Ribosome: 21% O_2 (Benjamini-Hochberg q-value = 3.00E-07, Count 9 = 19.6%), O_2-independent (Benjamini-Hochberg q-value = 5.11E-03, Count 4 = 25%). While ribosomal subunits are often considered as noise in these types of proteomics experiments, our strict filtering conditions, background subtraction controls and strong washing conditions, combined with the number of identifications suggest these as genuine palmitoylation targets. Supporting these observations, the 40S ribosomal protein S11 as previously been identified as being Cys-palmitoylated (Forrester et al., 2011), suggesting that zDHHC23 may have important roles in regulating ribosomal function.

6.3.4 Cell localisation of palmitoylated proteins

Due to the GO Enrichment analysis of zDHHC23 interactors showing a wide variety of cellular compartments/localisation, I decided to compare zDHHC23 dependent palmitoylation protein organellar localisation to published protein localisation (independent of palmitoyl-transferase, hence includes O-palmitoylation). (Consortium, 2019) was used searching for 'palmitoylation' in the human database, to identify all known human palmitoylated proteins (302) compared to our list of zDHHC23 palmitoylated targets (82) and manually searched to extract all proteins in the lists that had published cellular localisation information (35 vs 33) (Figure 6.12).

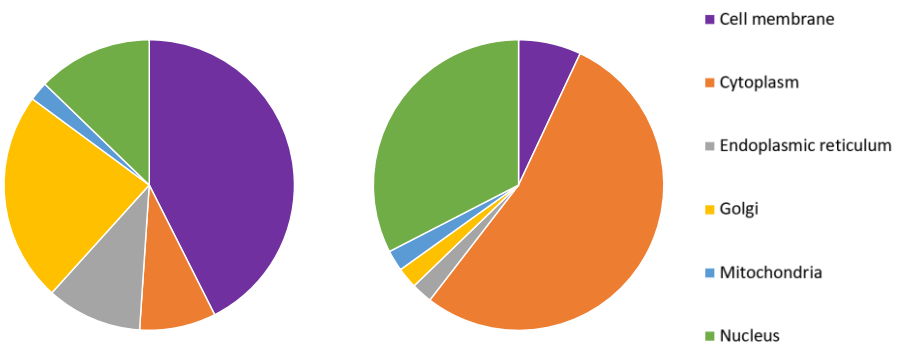

Figure 6.12 Cell localisation of proteins known to be palmitoylated vs zDHHC23 palmitoylated target proteins.
Comparison on the abundance of different cell localisation of palmitoylated proteins. Left = published cell localisation of known palmitoylated proteins (UniProt), Right = published cell localisation of zDHHC23 palmitoylation proteins.

Unsurprisingly, as the focus of palmitoylation has to date been on proteins localising to the cell membrane, Figure 6.12 shows that most proteins already known to be palmitoylated are localised to the cell membrane (57%) and Golgi (31%), with a smaller proportion of proteins localising to other membranous organelles including endoplasmic reticulum (14%), nucleus (17%) and mitochondria (3%). Interestingly, this is in stark contrast to our data, where over 69% of identifications are localised to the cytoplasm and nucleus (42%), with much smaller numbers of proteins being preferentially localised to the cell membrane (9%) and Golgi (3%) comparatively. However, our results are reliant on UniProt Annotation of protein localisation, this will not take into consideration proteins that have more than one localisation site, for example; Pyruvate kinase M2 mainly localises within the cytosol, but translocates to the nucleus to promote cell proliferation (Yang and Lu, 2015). While this comparison further reiterates the GO analysis results that palmitoylated proteins extend far beyond the canonical assumption of cellular membrane localisation, the fact that these experiments were focussed on targets of zDHHC23 makes it difficult to determine if this is a general phenomenon or specific to this particular enzyme but is worthy of further investigation.

6.3.5 STRING analysis

Following filtering for non-specific binding, the 82 remaining zDHHC23 palmitoylation targets were analysed in STRING (version 11.5, http://string.embl.de) to generate an interaction

network of putative substrates. Data for analysis was based on all proteins identified either in normoxia or hypoxia, filtered for data either stored as experimental or database evidence at medium confidence (>0.4) (Figure 6.13).

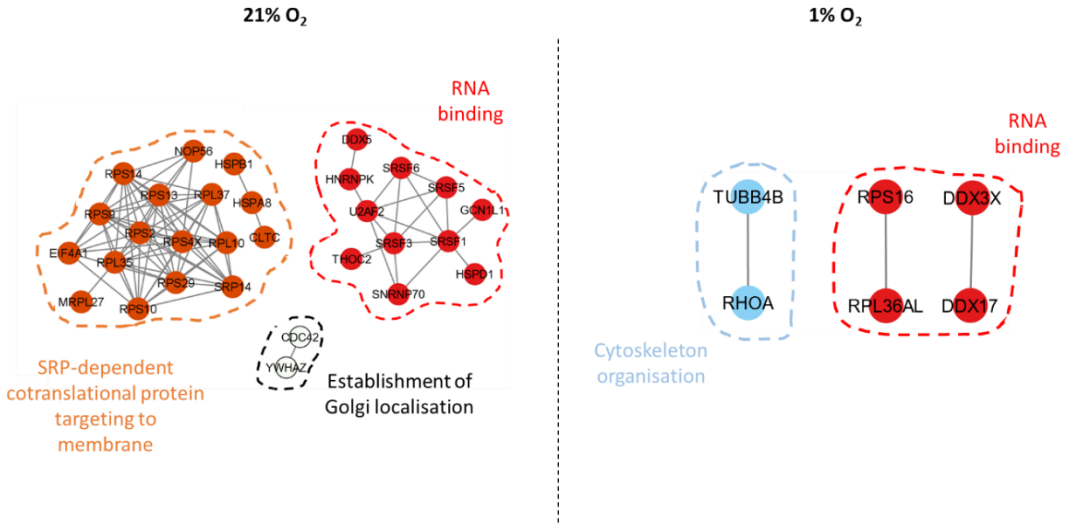

Figure 6.13 Network of proteins identified in two or more replicates of 21% and 1% O_2 zDHHC23 CuAAC experiment.
The original network was generated in STRING to include interactors with experimental or database evidence of medium confidence (interaction score >0.4). Any nodes with no connections of medium confidence or lower were removed from the network. The network was re-arranged in Cytoscape and re-coloured to group nodes into functional groups/complexes. Left = 21% O_2, Right = 1% O_2. Any disconnected nodes were removed from the network.

Reiterating observations in the binding partner analysis, zDHHC23 palmitoylated targets are involved in RNA binding regardless of O_2 tension, although many more identifications are found in 21% O_2. Additionally, a significant network specific to 21% O_2 (consisting predominantly of ribosomal proteins) are responsible for Signal Recognition Particle (SRP)-dependent co-translational protein targeting to membrane thus, further confirming current literature and knowledge of palmitoylation serving to aid protein movement to the cell membrane (Dunphy and Linder, 1998). Furthermore, two proteins, CDC42 and YWHAZ, that are 21% O_2 specific, are responsible for protein localisation to the Golgi, supporting GO enrichment analysis. Interestingly, CDC42 and YWHAZ are known to influence cell cycle progression and growth (Gan et al., 2020, Olson et al., 1995). Importantly, two proteins that are 1% O_2 specific are heavily responsible for cytoskeletal re-organisation: RHOA and TUBB4B, which help form the cytoskeleton and regulation cell cycle control (Dharmapal et al., 2021, O'Connor and Chen, 2013) potentially explaining the aggressive phenotype seen in some NB patients.

6.3.6 Comparing binding partner data with palmitoylated targets

Using the binding partner analysis completed in Chapter 5, a comparison was made to identify if any zDHHC23 binding partners previously identified were also zDHHC23 palmitoylated targets in either normoxic or hypoxic conditions (Figure 6.14). While this analysis method is limited due to weak transient interactions between the enzyme and target generally are, Figure 6.14 shows there is minimal overlap under 21% O_2 conditions in binding partners and palmitoylated targets data with 2 proteins identified in both: BRIX1 (Biogenesis of Ribosomes BRX1) and SNRNP70 (Small Nuclear Ribonucleoprotein U1 Subunit 70). Figure 6.15 also shows that the was no overlap under 1% O_2 conditions.

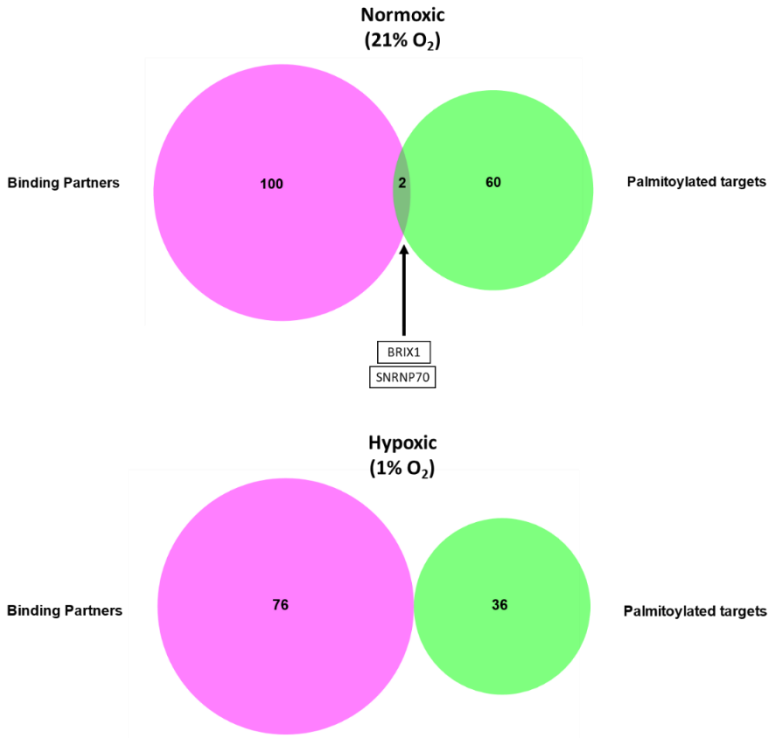

Figure 6.14 Comparison of binding partners and palmitoylated targets for zDHHC23 with both C and N terminal constructs combined.
Venn diagrams of proteins identified in 21% O_2 and 1% O_2 zDHHC23 with C and N terminal construct data combined with post background subtraction. Pink = proteins specific to binding partner data, green = proteins specific palmitoylation target data, dark green = proteins seen in both binding partner and palmitoylation target data. Circles are scaled respectively to the number of binding partners/palmitoylated targets per section. Venn diagrams created using BioVenn tool (Hulsen et al., 2008).

Both BRIX1 and SNRNP70 proteins are ribosomal proteins, where BRIX1 is involved in the synthesis of ribosomal 60S subunits (Eisenhaber et al., 2001) and SNRNP70 is a component of the small nuclear ribonuclear protein complex a core component of the spliceosome (Spritz et al., 1987). While the role of BRIX1 in cancer has not been investigated, increased expression has been observed in a variety of cancers (Jung et al., 2009) and there are increasing observations that upregulated ribosome biogenesis is associated with an increased risk of cancer onset (Pelletier et al., 2018). The reduction of the splicing ratio of SNRNP70 is a known hallmark of a subtype of Amyotrophic lateral sclerosis (Nakaya, 2022). Cancer cells are known to use alternative splicing to produce proteins that help contribute to growth and survival (Chen and Weiss, 2015).

6.4 Discussion

The application of CuAAC combined with LC-MS/MS analysis has been crucial to identify putative zDHHC23 mediated palmitoylated proteins. However, despite optimisation there are limitations to this experimental method and must be considered when drawing conclusions. All analysis was performed on exogenously expressed tagged proteins and are thus susceptible to the same pitfalls that are discussed in Chapter 5. Also, as discussed in Chapter 5, to determine the most valid results by combining the data seen in both N and C terminal constructs, inevitably a caveat of this method results in the filtering out of true palmitoylation targets that may be influenced by a tag at one, but not both termini, potentially masking the full range of zDHHC23 substrates. Additionally, due to the dosage of azide tagged palmitic acid occurring over a short period of time (24 hr) the results will be biased towards short half-live proteins and therefore we will miss identification of any palmitoylated proteins which have a longer half-life where their synthesis in the presence of azide will be reduced.

GO analysis results identify palmitoylation as a potential modulator of ribosomes and RNA binding functions in an O_2-independent manner. This is an interesting result, due to palmitoylation typically being associated with membrane proteins whilst ribosomes are classically cytoplasmic with some bound to the ER and nuclear envelope. Furthermore, this data suggests that palmitoylation plays a role in protein translation, this could be either aiding association of the multi-conglomerate ribosome species or subcellular localisation into discreet cellular compartments (such as polysome formation (Neidermyer Jr and Whelan,

2019)). Whilst this was not the aim of this project, this would be an interesting avenue for future research. Interestingly, the GO term 'extracellular exosome' is prominent in both 21% and 1% O_2 datasets but not O_2-independent inferring the role of intercellular trafficking differs depending on O_2 tension. Intercellular trafficking can also be associated with zDHHC23 role in aggressive NB. This is due to evidence that extracellular vesicles play a vital role in preparing and inducing cancer metastasis (Chin and Wang, 2016) suggestive of the correlating poor survival rate of NB with a high expression of zDHHC23.

As briefly explained, the palmitoylated residue and peptide is covalently bound to the bead. Due to time constraints, we did not attempt to optimise this any further. One route to elucidating the site of modification would be the application of hydroxylamine to chemically elute palmitoylated proteins by hydrolysing the thio-ester bond. This approach would have two advantages: 1) chemical elution will further reduce noise contaminants occurring from nonspecific binding to the bead matrix, and 2) The chemical reaction eluting the cysteine containing peptide additionally results in the formation of dehydro-alanine, which would then be indicative of palmitoylation site. Thus, this protocol adaption would allow the palmitoylated site to be specifically determined by LC/MS-MS analysis and may aid in motif identification.

To date research has focused on palmitoylation having a direct effect on proteins localising to the cell membrane. However, this study highlights that we should consider organelle membranes as being affected by palmitoylation, a PTM known to modify several proteins regulating oncogenes and tumour suppressors (Zhou et al., 2022). Palmitoylation as a PTM is significantly understudied in comparison to other PTMs, however with the application of click chemistry we could greatly increase our knowledge of palmitoylation function not only for zDHHC23 but other palmitoyltransferases.

6.4.1 Concluding remarks

The zDHHC23 palmitoylated targets of ribosomal proteins, RNA binding proteins and extracellular vesicles localised proteins implies not only a broad range functionality but also that zDHHC23 role in aggressive neuroblastoma is a mediator rather than an elicitor.

Chapter 7: Final Discussion

At the start of this PhD project, zDHHC23 and MROH6 had been identified as potential markers of NB, with high mRNA expression correlating significantly with poor prognosis, to a degree that was much more significant than *MYCN*-amplification which is currently used in risk stratification (Table 1.3). In an attempt to understand the functions of zDHHC23 and MROH6 and their roles in the context of NB, I thus investigated and developed tools to explore their cellular roles.

7.1. Putative biomarker validation

The initial aim of this thesis was to identify whether the observed mRNA level increase in hypoxic tumours (mimicking an aggressive phenotype) correlated with protein level expression for my two proteins of interest, zDHHC23 and MROH6. As 75% of NB patients lack *MYCN* amplification, I used a cell line lacking this amplification (SK-N-AS) to explore zDHHC23 and MROH6 protein levels. While I attempted multiple strategies to detect these endogenous proteins including the 'gold standard' IHC/IF staining (Figure 3.3 – Figure 3.6) and MS-based PRM methods (Figure 3.14 and Figure 3.15), these experiments were unsuccessful. At the time of these studies, very few commercially available antibodies were available for either zDHHC23 or MROH6. While attempts were made to validate a number of these, I determined they were unsuitable for the specific detection of their putative target protein in either native or denatures states. Indeed, the sale of some of these antibodies has now been discontinued. Newer antibodies for both of these proteins have since been made commercially available and it would be interesting to determine their specificity and characterise endogenous protein levels in hypoxic tumours and in NB samples. Successful PRM assays were developed against both zDHHC23 and MROH6 using immunoprecipitated material. Unfortunately, I was not able to use these assays to detect endogenous protein from (normoxic and hypoxic) tumours or total cell extracts. While not practicable in a clinical setting for screening purposes, it may have been useful to use off-line peptide-based fractionation prior to PRM to reduce sample complexity. In theory, this approach would increase PRM sensitivity, and potentially have allowed me to address a main tenet of these investigations, namely whether mRNA expression correlates with protein levels and if zDHHC23 and MROH6 proteins could be used as putative biomarkers for NB. However, these experiments were constrained by our

inability to obtain biobanked tissue samples for investigation. Assuming that we could develop either a more sensitive targeted MS assay, or validate target specific antibodies, it would also be interesting to investigate both major sub-populations of NB (+/- *MYCN* amplification) as there are a couple of possibilities: 1) zDHHC23 and MROH6 may express differently in the presence of *MYCN* amplification and 2) Co-expression of *MYCN* with zDHHC23 and MROH6 could significantly alter patient prognosis.

7.2. zDHHC23 and MROH6 cellular signalling

Beyond previously discussed caveats of exogenous expression systems, I generated constructs with tags at the N- and C- terminus to overcome potential constraints in protein binding and thus signalling due to the tag. Depending on the construct used, different data was yielded and therefore, it would be intriguing to explore this further from a structural perspective. These experiments yielded a vast quantity of information, describing extensive signalling networks, producing O_2 tension dependent, unbiased data (both at the interactor and PTM level) which substantially furthers our current understanding of both zDHHC23 and MROH6. Some of the pertinent points are discussed below.

7.2.1. zDHHC23

While zDHHC23 has been studied previously, the literature is sparse. At the time of writing there was only one identified binding partner for zDHHC23, the calcium activated potassium channel KCNMA1 (Tian et al., 2012) which was also reported to be a substrate for palmitoylation. While I failed to identify KCNMA1 as a binding partner (or zDHHC23 palmitoylation target) in the cohort of 262 proteins that I confidently identified in the zDHHC23 interaction network post filtering, its absence may be due to a number of reasons including the MS-detectability of KCNMA1, the likely transient nature of the substrate/enzyme interaction, cell line dependency or the fact the publication expressed mouse zDHHC23 in human cells. However, searching my dataset for other metal-ion channel proteins, I identified multiple Ca^{2+} ion transporting protein channels including VDAC1, VDAC3, TMCO1 and CLIC1. Interestingly, all of these Ca^{2+} transporter proteins are reported to have roles in tumour cell survival (Shoshan-Barmatz et al., 2013, Zheng et al., 2022, Li et al., 2018).

Another important avenue I would prioritise for investigation would be the identified sites of post-translational modification. For example, S252 which I identify as an O_2-independent phosphorylation site, lies in a disordered loop region and is predicted to be regulated by CDKs. It would thus be interesting to explore the dynamics of this phosphosite at different stages of the cell cycle and explore the effects of phosphonull (S252A) or phosphomimetic (S252D) constructs on cell cycle progression.

7.2.2. MROH6

At the time of writing, MROH6 remains a 'predicted protein' within the UniProt database (Consortium, 2019) with all data obtained in this thesis pertaining to MROH6 localisation, binding partners and PTMs being novel. The scale of the interaction network identified, with 253 confidently identified binding partners, means that it will take a while to explore this in any depth. However, bioinformatics characterisation of these data did suggest its involvement in mRNA binding, telomere maintenance and cellular senescence. However, a striking observation is the high relative abundance of MROH6 with BCAR1 (Section 5.3.7). This observation is potentially very important given recent identification of BCAR1 as a binding partner of a mutant form of p53 (TP53) and its direct involvement in promoting cellular invasion (Guo et al., 2021). It is noted that I do not identify p53 as a binding partner of MROH6 here, but this may be due the very low abundance of p53 (~13.5 ppm in cell lines, (Wang et al., 2015)) and/or that it may not be a direct or stable interaction. It would thus be interesting to specifically evaluate the ability of MROH5 to interact with p53 either directly or indirectly, and to evaluate the role of MROH6 on the ability of BCAR1 to promote cellular invasion. The identification of GAR1 and ZNF227 as interactors of MROH6 suggests that this protein may be involved in maintaining telomere length, a typical cellular mechanism to cause cellular senescence after a set number of cell divisions (Martínez and Blasco, 2015). Therefore, it is possible MROH6 overexpression prevents telomere shortening and escapes the fundamental 'time-of-life' pathway that triggers cellular senescence, thereby allowing tumour cells to continually proliferate (Shay and Wright, 2011). Possible techniques to investigate genetic instability and telomere shortening could be Giesma staining and karyotyping for different chromosomal aberrations using light microscopy (Garini et al., 1996) and quantitative fluorescence in-situ hybridisation (FISH) for telomere length analysis (Lai et al., 2018).

7.3. Expanding the regulatory network

The initial transcriptomic experiment performed by the Sée lab which identified zDHHC23 and MROH6 upregulated in hypoxic SK-N-AS tumours, used a 72 hr 1% O_2 incubation for pre-culturing of NB cells as this resulted in metastasis in the chick embryo model (Herrmann et al., 2015). There are great variations in literature for *in-vitro* hypoxia exposure times ranging from as short as 4 hr to several weeks (Hsieh et al., 2010, Sweet et al., 2010). As such, hypoxia is subcategorized into two types: acute (temporary O_2 disruption, e.g. red blood cell fluctuations in microvascular O_2 supply (Bayer et al., 2011) and chronic (prolonged disruption, e.g. limitations in O_2 diffusion from surrounding tumour micro vessels and tissue (Brown, 1979). As acute and chronic hypoxia tend to represent different pathophysiology's, there are different therapeutic implications (Vaupel, 2009). It would thus be interesting to investigate how the levels and interactions of zDHHC23 and MROH6 in SK-N-AS cells differ as a response to varying lengths of time pre-cultured in 1% O_2.

7.4. Functional *in-vitro* investigations

Following successful identification and validation of some (five proteins) binding partners for both zDHHC23 and MROH6 there are experiments that could follow on from our proteomic investigation. Due to the presumed role that zDHHC23 and MROH6 have in aggressive NB based on the elevated transcript levels and correlation with poor prognosis, functional assays should investigate the cell cycle progression (e.g. FACS), migration (e.g. wound healing assays) and metabolic activity (e.g. MTT assay) to identify how the knockdown/overexpression of these proteins may promote the aggressive phenotype observed.

7.5. Click chemistry

CuAAC has been successfully used to identify zDHHC enzyme fatty acid selectivity (Greaves et al., 2017). However, to my knowledge, we describe the first use of the technique in a global HTP discovery pipeline to identify the zDHHC23 'palmitome'. While I suggest possible

improvements to the protocol for better LC-MS/MS analysis and potential PTM-site localisation (Chapter 6), this investigation significantly highlights that the palmitome is far more complex than initially believed with the majority of our 82 identified proteins to be localised to various organelle membranes, rather than the primary role of plasma membrane targeting.

7.6. Future perspective

I have discovered multiple PTMs, binding partners and palmitoylated targets in this work and as such it has posed many new questions for the future, including: 1) The function behind the PTMs and interactions we have identified. 2) Whether PTMs and binding partners are transient or essential for function, and potential cell cycle dependent regulation of them. 3) PTM and binding partner robustness in response to different hypoxia protocols and cell lines used. 4) The real time effects of zDHHC23 and MROH6 expression (either over-expressed or knocked-out) and roles NB cell migration and proliferation. 5) The potential role of palmitoylation in localising proteins to sub-cellular membranous organelle compartments. 6) The application of my CuAAC-coupled LC-MS/MS protocol to other organic biological PTMs. Whilst I was unable to determine whether zDHHC23 and MROH6 could be used as novel biomarkers in NB.

Finally, our data suggests that zDHHC23 may have roles in regulating cell migration and survival, whilst MROH6, a protein currently with no published function, may have direct roles by either promoting genetic instability or telomere lengthening to allow cancer cell survivability. While I was unable to investigate the actual functional/phenotypic roles of zDHHC23 and MROH6 to determine why their expression is associated with such poor NB patient prognosis, the data discussed in this thesis could ultimately help guide future experimentation resulting in therapeutic investigations and translation – an area of research that NB patients are in dire need of.

References

ABATE, M., LAEZZA, C., PISANTI, S., TORELLI, G., SENECA, V., CATAPANO, G., MONTELLA, F., RANIERI, R., NOTARNICOLA, M. & GAZZERRO, P. 2017. Deregulated expression and activity of Farnesyl Diphosphate Synthase (FDPS) in Glioblastoma. *Scientific reports,* 7**,** 1-10.

ACKERMANN, S., CARTOLANO, M., HERO, B., WELTE, A., KAHLERT, Y., RODERWIESER, A., BARTENHAGEN, C., WALTER, E., GECHT, J. & KERSCHKE, L. 2018. A mechanistic classification of clinical phenotypes in neuroblastoma. *Science,* 362**,** 1165-1170.

AEBERSOLD, R. & MANN, M. 2016. Mass-spectrometric exploration of proteome structure and function. *Nature,* 537**,** 347-355.

ALBERTS, B. 1998. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *cell,* 92**,** 291-294.

ALI, A., LEVANTINI, E., TEO, J. T., GOGGI, J., CLOHESSY, J. G., WU, C. S., CHEN, L., YANG, H., KRISHNAN, I. & KOCHER, O. 2018. Fatty acid synthase mediates EGFR palmitoylation in EGFR mutated non-small cell lung cancer. *EMBO molecular medicine,* 10**,** e8313.

AMSALEM, Z., ARIF, T., SHTEINFER-KUZMINE, A., CHALIFA-CASPI, V. & SHOSHAN-BARMATZ, V. 2020. The mitochondrial protein VDAC1 at the crossroads of cancer cell metabolism: the epigenetic link. *Cancers,* 12**,** 1031.

APPLEBAUM, M. A., HENDERSON, T. O., LEE, S. M., PINTO, N., VOLCHENBOUM, S. L. & COHN, S. L. 2015. Second malignancies in patients with neuroblastoma: The effects of risk-based therapy. *Pediatric blood & cancer,* 62**,** 128-133.

ARIKE, L. & PEIL, L. 2014. Spectral counting label-free proteomics. *Shotgun proteomics.* Springer.

ARZT, M., DESCHAMPS, J., SCHMIED, C., PIETZSCH, T., SCHMIDT, D., TOMANCAK, P., HAASE, R. & JUG, F. 2022. LABKIT: labeling and segmentation toolkit for big image data. *Frontiers in Computer Science***,** 10.

AUERBACH, C., MOUTSCHEN-DAHMEN, M. & MOUTSCHEN, J. 1977. Genetic and cytogenetical effects of formaldehyde and related compounds. *Mutation Research/Reviews in Genetic Toxicology,* 39**,** 317-361.

BABINA, I. S., MCSHERRY, E. A., DONATELLO, S., HILL, A. D. & HOPKINS, A. M. 2014. A novel mechanism of regulating breast cancer cell migration via palmitoylation-dependent alterations in the lipid raft affiliation of CD44. *Breast Cancer Research,* 16**,** 1-14.

BAGDONAS, H., FOGARTY, C. A., FADDA, E. & AGIRRE, J. 2021. The case for post-predictional modifications in the AlphaFold Protein Structure Database. *Nature Structural & Molecular Biology,* 28**,** 869-870.

BAKER, E., GILBERT, S., SHIMELD, S. & WOOLLARD, A. 2021. Extensive non-redundancy in a recently duplicated developmental gene family. *BMC Ecology and Evolution,* 21**,** 1-17.

BANSAL, L., NELSON, R., YANG, E., JAYARAMAN, A. & HAHN, J. 2013. Experimental design of systems involving multiple fluorescent protein reporters. *Chemical Engineering Science,* 101**,** 191-198.

BANTSCHEFF, M., LEMEER, S., SAVITSKI, M. M. & KUSTER, B. 2012. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Analytical and bioanalytical chemistry,* 404**,** 939-965.

BANTSCHEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry,* 389**,** 1017-1031.

BARD, J. A., GOODALL, E. A., GREENE, E. R., JONSSON, E., DONG, K. C. & MARTIN, A. 2018. Structure and function of the 26S proteasome. *Annual review of biochemistry,* 87**,** 697.

BARTLETT, J. M., GOING, J. J., MALLON, E. A., WATTERS, A. D., REEVES, J. R., STANTON, P., RICHMOND, J., DONALD, B., FERRIER, R. & COOKE, T. G. 2001. Evaluating HER2 amplification and overexpression in breast cancer. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland,* 195**,** 422-428.

BARYLKO, B., MAO, Y. S., WLODARSKI, P., JUNG, G., BINNS, D. D., SUN, H.-Q., YIN, H. L. & ALBANESI, J. P. 2009. Palmitoylation controls the catalytic activity and subcellular distribution of phosphatidylinositol 4-kinase IIα. *Journal of Biological Chemistry,* 284**,** 9994-10003.

BASKIN, J. M., PRESCHER, J. A., LAUGHLIN, S. T., AGARD, N. J., CHANG, P. V., MILLER, I. A., LO, A., CODELLI, J. A. & BERTOZZI, C. R. 2007. Copper-free click chemistry for dynamic in vivo imaging. *Proceedings of the National Academy of Sciences,* 104**,** 16793-16797.

BASTA, N. O., HALLIDAY, G. C., MAKIN, G., BIRCH, J., FELTBOWER, R., BOWN, N., ELLIOTT, M., MORENO, L., BARONE, G. & PEARSON, A. D. 2016. Factors associated with recurrence and survival length following relapse in patients with neuroblastoma. *British journal of cancer,* 115**,** 1048-1057.

BAYER, C., SHI, K., ASTNER, S. T., MAFTEI, C.-A. & VAUPEL, P. 2011. Acute versus chronic hypoxia: why a simplified classification is simply not enough. *International journal of radiation oncology, biology, physics,* 80**,** 965-968.

BECKSTEAD, J. H. 1994. Improved antigen retrieval in formalin-fixed, paraffin-embedded tissues. *Applied Immunohistochemistry+ Molecular Morphology,* 2**,** 274-281.

BERGGÅRD, T., LINSE, S. & JAMES, P. 2007. Methods for the detection and analysis of protein–protein interactions. *Proteomics,* 7**,** 2833-2842.

BERN, M., PHINNEY, B. S. & GOLDBERG, D. 2009. Reanalysis of Tyrannosaurus rex mass spectra. *Journal of proteome research,* 8**,** 4328-4332.

BIRKENKAMP-DEMTRODER, K., CHRISTENSEN, L. L., OLESEN, S. H., FREDERIKSEN, C. M., LAIHO, P., AALTONEN, L. A., LAURBERG, S., SORENSEN, F. & HAGEMANN, R. 2002. ORntoft. *Gene expression in colorectal cancer. Cancer Res,* 62**,** 4352-4363.

BLUHM, E., MCNEIL, D. E., CNATTINGIUS, S., GRIDLEY, G., EL GHORMLI, L. & FRAUMENI JR, J. F. 2008. Prenatal and perinatal risk factors for neuroblastoma. *International journal of cancer,* 123**,** 2885-2890.

BOCK, V. D., HIEMSTRA, H. & VAN MAARSEVEEN, J. H. 2006. CuI-catalyzed alkyne–azide "click" cycloadditions from a mechanistic and synthetic perspective. *European Journal of Organic Chemistry,* 2006**,** 51-68.

BOLLU, L. R., REN, J., BLESSING, A. M., KATREDDY, R. R., GAO, G., XU, L., WANG, J., SU, F. & WEIHUA, Z. 2014. Involvement of de novo synthesized palmitate and mitochondrial EGFR in EGF induced mitochondrial fusion of cancer cells. *Cell cycle,* 13**,** 2415-2430.

BONDARENKO, P. V., CHELIUS, D. & SHALER, T. A. 2002. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography− tandem mass spectrometry. *Analytical chemistry,* 74**,** 4741-4749.

BOURDEAUT, F., FERRAND, S., BRUGIÈRES, L., HILBERT, M., RIBEIRO, A., LACROIX, L., BÉNARD, J., COMBARET, V., MICHON, J. & VALTEAU-COUANET, D. 2012. ALK germline mutations in patients with neuroblastoma: a rare and weakly penetrant syndrome. *European journal of human genetics,* 20**,** 291-297.

BOWN, N., COTTERILL, S., ŁASTOWSKA, M., O'NEILL, S., PEARSON, A. D., PLANTAZ, D., MEDDEB, M., DANGLOT, G., BRINKSCHMIDT, C. & CHRISTIANSEN, H. 1999. Gain of

chromosome arm 17q and adverse outcome in patients with neuroblastoma. *New England Journal of Medicine,* 340**,** 1954-1961.
BRAAKMAN, I. & HEBERT, D. N. 2013. Protein folding in the endoplasmic reticulum. *Cold Spring Harbor perspectives in biology,* 5**,** a013201.
BRAHIMI-HORN, M. C., CHICHE, J. & POUYSSÉGUR, J. 2007. Hypoxia and cancer. *Journal of molecular medicine,* 85**,** 1301-1307.
BRAUN, P. & GINGRAS, A. C. 2012. History of protein–protein interactions: From egg-white to complex networks. *Proteomics,* 12**,** 1478-1498.
BRETSCHNEIDER, T., DIEZ, S., ANDERSON, K., HEUSER, J., CLARKE, M., MÜLLER-TAUBENBERGER, A., KÖHLER, J. & GERISCH, G. 2004. Dynamic actin patterns and Arp2/3 assembly at the substrate-attached surface of motile cells. *Current Biology,* 14**,** 1-10.
BRISSE, H. J., MCCARVILLE, M. B., GRANATA, C., KRUG, K. B., WOOTTON-GORGES, S. L., KANEGAWA, K., GIAMMARILE, F., SCHMIDT, M., SHULKIN, B. L. & MATTHAY, K. K. 2011. Guidelines for imaging and staging of neuroblastic tumors: consensus report from the International Neuroblastoma Risk Group Project. *Radiology,* 261**,** 243-257.
BRISTOW, R. G. & HILL, R. P. 2008. Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. *Nature Reviews Cancer,* 8.
BRODEUR, G. M., SEEGER, R. C., BARRETT, A., BERTHOLD, F., CASTLEBERRY, R. P., D'ANGIO, G., DE BERNARDI, B., EVANS, A. E., FAVROT, M. & FREEMAN, A. I. 1988. International criteria for diagnosis, staging, and response to treatment in patients with neuroblastoma. *Journal of Clinical Oncology,* 6**,** 1874-1881.
BRODEUR, G. M., SEEGER, R. C., SCHWAB, M., VARMUS, H. E. & BISHOP, J. M. 1984. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science,* 224**,** 1121-1124.
BRONNER-FRASER, M. 1995. Origins and developmental potential of the neural crest. *Experimental cell research,* 218**,** 405-417.
BROWN, J. 1979. Evidence for acutely hypoxic cells in mouse tumours, and a possible mechanism of reoxygenation. *The British journal of radiology,* 52**,** 650-656.
BUCZAK, K., KIRKPATRICK, J. M., TRUCKENMUELLER, F., SANTINHA, D., FERREIRA, L., ROESSLER, S., SINGER, S., BECK, M. & ORI, A. 2020. Spatially resolved analysis of FFPE tissue proteomes by quantitative mass spectrometry. *Nature protocols,* 15**,** 2956-2979.
BYRNE, D., SHRESTHA, S., DALY, L. A., MARENSI, V., RAMAKRISHNAN, K., EYERS, C. E., KANNAN, N. & EYERS, P. A. 2022. Evolutionary and cellular analysis of the'dark'pseudokinase PSKH2. *Biochemical Journal*.
CAJIGAS-DU ROSS, C. K., MARTINEZ, S. R., WOODS-BURNHAM, L., DURÁN, A. M., ROY, S., BASU, A., RAMIREZ, J. A., ORTIZ-HERNÁNDEZ, G. L., RÍOS-COLÓN, L. & CHIRSHEV, E. 2018. RNA sequencing reveals upregulation of a transcriptomic program associated with stemness in metastatic prostate cancer cells selected for taxane resistance. *Oncotarget,* 9**,** 30363.
CAMPBELL, K., SHYR, D., BAGATELL, R., FISCHER, M., NAKAGAWARA, A., NIETO, A. C., BRODEUR, G. M., MATTHAY, K. K., LONDON, W. B. & DUBOIS, S. G. 2019. Comprehensive evaluation of context dependence of the prognostic impact of MYCN amplification in neuroblastoma: A report from the International Neuroblastoma Risk Group (INRG) project. *Pediatric blood & cancer,* 66**,** e27819.
CARADEC, J., SIRAB, N., KEUMEUGNI, C., MOUTEREAU, S., CHIMINGQI, M., MATAR, C., REVAUD, D., BAH, M., MANIVET, P. & CONTI, M. 2010. 'Desperate house genes': the dramatic example of hypoxia. *British journal of cancer,* 102**,** 1037-1043.
CARÉN, H., KRYH, H., NETHANDER, M., SJÖBERG, R.-M., TRÄGER, C., NILSSON, S., ABRAHAMSSON, J., KOGNER, P. & MARTINSSON, T. 2010. High-risk neuroblastoma

tumors with 11q-deletion display a poor prognostic, chromosome instability phenotype with later onset. *Proceedings of the National Academy of Sciences,* 107**,** 4323-4328.

CARON, H., VAN SLUIS, P., DE KRAKER, J., BÖKKERINK, J., EGELER, M., LAUREYS, G., SLATER, R., WESTERVELD, A., VOUTE, P. & VERSTEEG, R. 1996. Allelic loss of chromosome 1p as a predictor of unfavorable outcome in patients with neuroblastoma. *New England Journal of Medicine,* 334**,** 225-230.

CHAN, D. I. & VOGEL, H. J. 2010. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochemical Journal,* 430**,** 1-19.

CHEN, J. & WEISS, W. 2015. Alternative splicing in cancer: implications for biology and therapy. *Oncogene,* 34**,** 1-14.

CHEN, K., MERIC-BERNSTAM, F., ZHAO, H., ZHANG, Q., EZZEDDINE, N., TANG, L.-Y., QI, Y., MAO, Y., CHEN, T. & CHONG, Z. 2015a. Clinical actionability enhanced through deep targeted sequencing of solid tumors. *Clinical chemistry,* 61**,** 544-553.

CHEN, S., ZHANG, M., XING, L., WANG, Y., XIAO, Y. & WU, Y. 2015b. HIF-1α contributes to proliferation and invasiveness of neuroblastoma cells via SHH signaling. *PLoS One,* 10**,** e0121115.

CHEN, W., GUO, W., LI, M., SHI, D., TIAN, Y., LI, Z., WANG, J., FU, L., XIAO, X. & LIU, Q. Q. 2013. Upregulation of cleavage and polyadenylation specific factor 4 in lung adenocarcinoma and its critical role for cancer cell survival and proliferation. *PloS one,* 8**,** e82728.

CHEN, X., YANG, F., ZHANG, T., WANG, W., XI, W., LI, Y., ZHANG, D., HUO, Y., ZHANG, J. & YANG, A. 2019. MiR-9 promotes tumorigenesis and angiogenesis and is activated by MYC and OCT4 in human glioma. *Journal of Experimental & Clinical Cancer Research,* 38**,** 1-16.

CHENG, K., XIE, G., KHURANA, S., HEATH, J., DRACHENBERG, C. B., TIMMONS, J., SHAH, N. & RAUFMAN, J.-P. 2014. Divergent effects of muscarinic receptor subtype gene ablation on murine colon tumorigenesis reveals association of M3R and zinc finger protein 277 expression in colon neoplasia. *Molecular cancer,* 13**,** 1-13.

CHEUNG, N.-K. V., CHEUNG, I. Y., KUSHNER, B. H., OSTROVNAYA, I., CHAMBERLAIN, E., KRAMER, K. & MODAK, S. 2012. Murine anti-GD2 monoclonal antibody 3F8 combined with granulocyte-macrophage colony-stimulating factor and 13-cis-retinoic acid in high-risk patients with stage 4 neuroblastoma in first remission. *Journal of clinical oncology,* 30**,** 3264.

CHEUNG, N.-K. V. & DYER, M. A. 2013. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nature Reviews Cancer,* 13**,** 397-411.

CHIN, A. R. & WANG, S. E. 2016. Cancer-derived extracellular vesicles: the 'soil conditioner'in breast cancer metastasis? *Cancer and Metastasis Reviews,* 35**,** 669-676.

CIMPEAN, A. M., RIBATTI, D. & RAICA, M. 2008. The chick embryo chorioallantoic membrane as a model to study tumor metastasis. *Angiogenesis,* 11**,** 311-319.

COHEN, L., GORDON, J., POPOVSKY, E., GUNAWARDENE, S., DUFFEY-LIND, E., LEHMANN, L. & DILLER, L. 2014. Late effects in children treated with intensive multimodal therapy for high-risk neuroblastoma: high incidence of endocrine and growth problems. *Bone marrow transplantation,* 49**,** 502-508.

COHEN, P. 2002. The origins of protein phosphorylation. *Nature cell biology,* 4**,** E127-E130.

COHN, S. L., PEARSON, A. D., LONDON, W. B., MONCLAIR, T., AMBROS, P. F., BRODEUR, G. M., FALDUM, A., HERO, B., IEHARA, T. & MACHIN, D. 2009. The International Neuroblastoma Risk Group (INRG) classification system: an INRG task force report. *Journal of clinical oncology,* 27**,** 289.

COLEMAN, D. T., SOUNG, Y. H., SURH, Y.-J., CARDELLI, J. A. & CHUNG, J. 2015. Curcumin prevents palmitoylation of integrin β4 in breast cancer cells. *PLoS One,* 10**,** e0125399.

COLON, N. C. & CHUNG, D. H. 2011. Journal: Advances in Pediatrics, 2011, № 1, p. 297-311. *Advances in Pediatrics***,** 297-311.

CONNOR, P. & MCQUILLAN, A. J. 1999. Phosphate adsorption onto TiO2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study. *Langmuir,* 15**,** 2916-2921.

CONSORTIUM, U. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic acids research,* 47**,** D506-D515.

COSCIA, F., DOLL, S., BECH, J. M., SCHWEIZER, L., MUND, A., LENGYEL, E., LINDEBJERG, J., MADSEN, G. I., MOREIRA, J. M. & MANN, M. 2020. A streamlined mass spectrometry–based proteomics workflow for large-scale FFPE tissue analysis. *The Journal of pathology,* 251**,** 100-112.

COUGHLAN, D., GIANFERANTE, M., LYNCH, C. F., STEVENS, J. L. & HARLAN, L. C. 2017. Treatment and survival of childhood neuroblastoma: evidence from a population-based study in the United States. *Pediatric hematology and oncology,* 34**,** 320-330.

COVELLO, K. L., KEHLER, J., YU, H., GORDAN, J. D., ARSHAM, A. M., HU, C.-J., LABOSKY, P. A., SIMON, M. C. & KEITH, B. 2006. HIF-2α regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes & development,* 20**,** 557-570.

COX, B. & EMILI, A. 2006. Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nature protocols,* 1**,** 1872-1878.

CREASY, D. M. & COTTRELL, J. S. 2004. Unimod: Protein modifications for mass spectrometry. *Proteomics,* 4**,** 1534-1536.

D'ANGIO, G., EVANS, A. & KOOP, C. E. 1971. Special pattern of widespread neuroblastoma with a favourable prognosis. *The Lancet,* 297**,** 1046-1049.

DALY, L. A., BROWNRIDGE, P. J., BATIE, M., ROCHA, S., SÉE, V. & EYERS, C. E. 2021. Oxygen-dependent changes in binding partners and post-translational modifications regulate the abundance and activity of HIF-1α/2α. *Science Signaling,* 14**,** eabf6685.

DAY, C.-P., MERLINO, G. & VAN DYKE, T. 2015. Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell,* 163**,** 39-53.

DE BARROS, F., ROMÃO, R. L. P., DE PINHO-APEZZATO, M. L., VELHOTE, M. C. P., RICARDI, L. R. S., LEAL, A. J. G., TANNURI, A. C. A., CARVALHO, B., ODONE-FILHO, V. & TANNURI, U. 2012. Laparoscopic adrenalectomy in children for neuroblastoma: report of case series. *Surgical Laparoscopy Endoscopy & Percutaneous Techniques,* 22**,** 79-81.

DE, I. & SADHUKHAN, S. 2018. Emerging roles of DHHC-mediated protein S-palmitoylation in physiological and pathophysiological context. *European journal of cell biology,* 97**,** 319-338.

DEKKER, F. J., ROCKS, O., VARTAK, N., MENNINGER, S., HEDBERG, C., BALAMURUGAN, R., WETZEL, S., RENNER, S., GERAUER, M. & SCHÖLERMANN, B. 2010. Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nature chemical biology,* 6**,** 449-456.

DELANO, W. L. 2002. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr,* 40**,** 82-92.

DENNIS, G., SHERMAN, B. T., HOSACK, D. A., YANG, J., GAO, W., LANE, H. C. & LEMPICKI, R. A. 2003. DAVID: database for annotation, visualization, and integrated discovery. *Genome biology,* 4**,** 1-11.

DEPLEDGE, M., AMARAL-MENDES, J., DANIEL, B., HALBROOK, R., KLOEPPER-SAMS, P., MOORE, M. & PEAKALL, D. The conceptual basis of the biomarker approach.

Biomarkers: Research and Application in the assessment of Environmental Health, 1993. Springer, 15-29.
DHARMAPAL, D., JYOTHY, A., MOHAN, A., BALAGOPAL, P., GEORGE, N. A., SEBASTIAN, P., MALIEKAL, T. T. & SENGUPTA, S. 2021. β-Tubulin Isotype, TUBB4B, Regulates The Maintenance of Cancer Stem Cells. *Frontiers in oncology,* 11.
DINMOHAMED, A. G., VISSER, O., VERHOEVEN, R. H., LOUWMAN, M. W., VAN NEDERVEEN, F. H., WILLEMS, S. M., MERKX, M. A., LEMMENS, V. E., NAGTEGAAL, I. D. & SIESLING, S. 2020. Fewer cancer diagnoses during the COVID-19 epidemic in the Netherlands. *The Lancet Oncology,* 21**,** 750-751.
DOMON, B. & GALLIEN, S. 2015. Recent advances in targeted proteomics for clinical applications. *PROTEOMICS–Clinical Applications,* 9**,** 423-431.
DORSAM, R. T. & GUTKIND, J. S. 2007. G-protein-coupled receptors and cancer. *Nature reviews cancer,* 7**,** 79-94.
DOUGLAS, D. J., FRANK, A. J. & MAO, D. 2005. Linear ion traps in mass spectrometry. *Mass spectrometry reviews,* 24**,** 1-29.
DUBOIS, S. G., KALIKA, Y., LUKENS, J. N., BRODEUR, G. M., SEEGER, R. C., ATKINSON, J. B., HAASE, G. M., BLACK, C. T., PEREZ, C. & SHIMADA, H. 1999. Metastatic sites in stage IV and IVS neuroblastoma correlate with age, tumor biology, and survival. *Journal of pediatric hematology/oncology,* 21**,** 181-189.
DUNCAN, P. J., BI, D., MCCLAFFERTY, H., CHEN, L., TIAN, L. & SHIPSTON, M. J. 2019. S-Acylation controls functional coupling of BK channel pore-forming α-subunits and β1-subunits. *Journal of Biological Chemistry,* 294**,** 12066-12076.
DUNHAM, W. H., MULLIN, M. & GINGRAS, A. C. 2012. Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *Proteomics,* 12**,** 1576-1590.
DUNKER, A. K., CORTESE, M. S., ROMERO, P., IAKOUCHEVA, L. M. & UVERSKY, V. N. 2005. Flexible nets: the roles of intrinsic disorder in protein interaction networks. *The FEBS journal,* 272**,** 5129-5148.
DUNPHY, J. T. & LINDER, M. E. 1998. Signalling functions of protein palmitoylation. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids,* 1436**,** 245-261.
DUNSTAN, R. W., WHARTON JR, K. A., QUIGLEY, C. & LOWE, A. 2011. The use of immunohistochemistry for biomarker assessment—can it compete with other technologies? *Toxicologic pathology,* 39**,** 988-1002.
EDWARDS, H., RUBENSTEIN, M., DOMBKOWSKI, A. A., CALDWELL, J. T., CHU, R., XAVIER, A. C., THUMMEL, R., NEELY, M., MATHERLY, L. H. & GE, Y. 2016. Gene signature of high white blood cell count in b-precursor acute lymphoblastic leukemia. *PLoS One,* 11**,** e0161539.
EGAN, E. D. & COLLINS, K. 2010. Specificity and stoichiometry of subunit interactions in the human telomerase holoenzyme assembled in vivo. *Molecular and cellular biology,* 30**,** 2775-2786.
EISENHABER, F., WECHSELBERGER, C. & KREIL, G. 2001. The Brix domain protein family–a key to the ribosomal biogenesis pathway? *Trends in biochemical sciences,* 26**,** 345-347.
FAN, R., CUI, W., CHEN, J., MA, Y., YANG, Z., PAYNE, T. J., MA, J. Z. & LI, M. D. 2021. Gene-based association analysis reveals involvement of LAMA5 and cell adhesion pathways in nicotine dependence in African-and European-American samples. *Addiction Biology,* 26**,** e12898.
FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science,* 246**,** 64-71.

FERLAY, J., COLOMBET, M., SOERJOMATARAM, I., PARKIN, D. M., PIÑEROS, M., ZNAOR, A. & BRAY, F. 2021. Cancer statistics for the year 2020: An overview. *International journal of cancer,* 149**,** 778-789.

FERRARI, A., ZECCA, M., RIZZARI, C., PORTA, F., PROVENZI, M., MARINONI, M., SCHUMACHER, R. F., LUKSCH, R., TERENZIANI, M. & CASANOVA, M. 2020. Children with cancer in the time of COVID-19: an 8-week report from the six pediatric onco-hematology centers in Lombardia, Italy. *Pediatric blood & cancer*.

FERREIRA-CERCA, S., PÖLL, G., GLEIZES, P.-E., TSCHOCHNER, H. & MILKEREIT, P. 2005. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Molecular cell,* 20**,** 263-275.

FERRIES, S., PERKINS, S., BROWNRIDGE, P. J., CAMPBELL, A., EYERS, P. A., JONES, A. R. & EYERS, C. E. 2017. Evaluation of parameters for confident phosphorylation site localization using an orbitrap fusion tribrid mass spectrometer. *Journal of proteome research,* 16**,** 3448-3459.

FIELD, J., NIKAWA, J.-L., BROEK, D., MACDONALD, B., RODGERS, L., WILSON, I., LERNER, R. & WIGLER, M. 1988. Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. *Molecular and cellular biology,* 8**,** 2159-2165.

FÍLA, J. & HONYS, D. 2012. Enrichment techniques employed in phosphoproteomics. *Amino acids,* 43**,** 1025-1047.

FORBES, S., CLEMENTS, J., DAWSON, E., BAMFORD, S., WEBB, T., DOGAN, A., FLANAGAN, A., TEAGUE, J., WOOSTER, R. & FUTREAL, P. 2006. COSMIC 2005. *British journal of cancer,* 94**,** 318-322.

FORRESTER, M. T., HESS, D. T., THOMPSON, J. W., HULTMAN, R., MOSELEY, M. A., STAMLER, J. S. & CASEY, P. J. 2011. Site-specific analysis of protein S-acylation by resin-assisted capture [S]. *Journal of lipid research,* 52**,** 393-398.

FORSBURG, S. L. 2004. Eukaryotic MCM proteins: beyond replication initiation. *Microbiology and Molecular Biology Reviews,* 68**,** 109-131.

FRESE, C. K., ALTELAAR, A. M., VAN DEN TOORN, H., NOLTING, D., GRIEP-RAMING, J., HECK, A. J. & MOHAMMED, S. 2012. Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Analytical chemistry,* 84**,** 9668-9673.

FRESE, C. K., ZHOU, H., TAUS, T., ALTELAAR, A. M., MECHTLER, K., HECK, A. J. & MOHAMMED, S. 2013. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD). *Journal of proteome research,* 12**,** 1520-1525.

FUKATA, M., FUKATA, Y., ADESNIK, H., NICOLL, R. A. & BREDT, D. S. 2004. Identification of PSD-95 palmitoylating enzymes. *Neuron,* 44**,** 987-996.

FUKATA, Y. & FUKATA, M. 2010. Protein palmitoylation in neuronal development and synaptic plasticity. *Nature Reviews Neuroscience,* 11**,** 161-175.

GAFFNEY, E., RIEGMAN, P., GRIZZLE, W. & WATSON, P. 2018. Factors that drive the increasing use of FFPE tissue in basic and translational cancer research. *Biotechnic & Histochemistry,* 93**,** 373-386.

GALL, J. G., BELLINI, M., WU, Z. A. & MURPHY, C. 1999. Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Molecular biology of the cell,* 10**,** 4385-4402.

GALLIEN, S., DURIEZ, E., CRONE, C., KELLMANN, M., MOEHRING, T. & DOMON, B. 2012. Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Molecular & cellular proteomics,* 11**,** 1709-1723.

GAN, Y., YE, F. & HE, X.-X. 2020. The role of YWHAZ in cancer: A maze of opportunities and challenges. *Journal of Cancer,* 11**,** 2252.

GARINI, Y., MACVILLE, M., DU MANOIR, S., BUCKWALD, R. A., LAVI, M., KATZIR, N., WINE, D., BAR-AM, I., SCHRÖCK, E. & CABIB, D. 1996. Spectral karyotyping. *Bioimaging,* 4**,** 65-72.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S. E., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. *Protein identification and analysis tools on the ExPASy server*, Springer.

GILBERT, F., FEDER, M., BALABAN, G., BRANGMAN, D., LURIE, D. K., PODOLSKY, R., RINALDT, V., VINIKOOR, N. & WEISBAND, J. 1984. Human neuroblastomas and abnormalities of chromosomes 1 and 17. *Cancer research,* 44**,** 5444-5449.

GILLET, L. C., LEITNER, A. & AEBERSOLD, R. 2016. Mass spectrometry applied to bottom-up proteomics: entering the high-throughput era for hypothesis testing. *Annual review of analytical chemistry,* 9**,** 449-472.

GOLEY, E. D. & WELCH, M. D. 2006. The ARP2/3 complex: an actin nucleator comes of age. *Nature reviews Molecular cell biology,* 7**,** 713-726.

GOMEZ, M. A. R. & IBBA, M. 2020. Aminoacyl-tRNA synthetases. *Rna,* 26**,** 910-936.

GONZÁLEZ-CHAVARRÍA, I., CERRO, R. P., PARRA, N. P., SANDOVAL, F. A., ZUÑIGA, F. A., OMAZÁBAL, V. A., LAMPERTI, L. I., JIMÉNEZ, S. P., FERNANDEZ, E. A. & GUTIÉRREZ, N. A. 2014. Lectin-like oxidized LDL receptor-1 is an enhancer of tumor angiogenesis in human prostate cancer cells. *PloS one,* 9**,** e106219.

GOOD, D. M., WIRTALA, M., MCALISTER, G. C. & COON, J. J. 2007. Performance characteristics of electron transfer dissociation mass spectrometry. *Molecular & Cellular Proteomics,* 6**,** 1942-1951.

GORLEKU, O. A., BARNS, A.-M., PRESCOTT, G. R., GREAVES, J. & CHAMBERLAIN, L. H. 2011. Endoplasmic reticulum localization of DHHC palmitoyltransferases mediated by lysine-based sorting signals. *Journal of Biological Chemistry,* 286**,** 39573-39584.

GRAY, K. A., DAUGHERTY, L. C., GORDON, S. M., SEAL, R. L., WRIGHT, M. W. & BRUFORD, E. A. 2012. Genenames. org: the HGNC resources in 2013. *Nucleic acids research,* 41**,** D545-D552.

GREAVES, J. & CHAMBERLAIN, L. H. 2011. DHHC palmitoyl transferases: substrate interactions and (patho) physiology. *Trends in biochemical sciences,* 36**,** 245-253.

GREAVES, J., MUNRO, K. R., DAVIDSON, S. C., RIVIERE, M., WOJNO, J., SMITH, T. K., TOMKINSON, N. C. & CHAMBERLAIN, L. H. 2017. Molecular basis of fatty acid selectivity in the zDHHC family of S-acyltransferases revealed by click chemistry. *Proceedings of the National Academy of Sciences,* 114**,** E1365-E1374.

GRYNKIEWICZ, G., POENIE, M. & TSIEN, R. Y. 1985. A new generation of Ca2+ indicators with greatly improved fluorescence properties. *Journal of biological chemistry,* 260**,** 3440-3450.

GU, Y., LIU, Y., FU, L., ZHAI, L., ZHU, J., HAN, Y., JIANG, Y., ZHANG, Y., ZHANG, P. & JIANG, Z. 2019. Tumor-educated B cells selectively promote breast cancer lymph node metastasis by HSPA4-targeting IgG. *Nature medicine,* 25**,** 312-322.

GUILHAUS, M. 1995. Special feature: Tutorial. Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts. *Journal of mass spectrometry,* 30**,** 1519-1532.

GUO, A. K., ITAHANA, Y., SESHACHALAM, V. P., CHOW, H. Y., GHOSH, S. & ITAHANA, K. 2021. Mutant TP53 interacts with BCAR1 to contribute to cancer cell invasion. *British Journal of Cancer,* 124**,** 299-312.

HA, S.-A., SHIN, S. M., NAMKOONG, H., LEE, H., CHO, G. W., HUR, S. Y., KIM, T. E. & KIM, J. W. 2004. Cancer-associated expression of minichromosome maintenance 3 gene in several human cancers and its involvement in tumorigenesis. *Clinical cancer research,* 10**,** 8386-8395.

HAAG, A. M. 2016. Mass analyzers and mass spectrometers. *Modern Proteomics–Sample Preparation, Analysis and Practical Applications*, 157-169.

HACKAM, A. S., HODGSON, J. G., SINGARAJA, R., ZHANG, T., GAN, L., GUTEKUNST, C.-A., HERSCH, S. M. & HAYDEN, M. R. 1999. Evidence for both nucleus and cytoplasm as subcellular sites of pathogenesis in Huntington'sdisease in cell culture and in transgenic mice expressing mutant huntingtin. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 354**,** 1047-1055.

HAINS, P. G. & TRUSCOTT, R. J. 2007. Post-translational modifications in the nuclear region of young, aged, and cataract human lenses. *Journal of proteome research,* 6**,** 3935-3943.

HAN, X., ASLANIAN, A. & YATES III, J. R. 2008. Mass spectrometry for proteomics. *Current opinion in chemical biology,* 12**,** 483-490.

HAN, X., HE, L., XIN, L., SHAN, B. & MA, B. 2011. PeaksPTM: mass spectrometry-based identification of peptides with unspecified modifications. *Journal of proteome research,* 10**,** 2930-2936.

HANAHAN, D. 2022. Hallmarks of cancer: new dimensions. *Cancer discovery,* 12**,** 31-46.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *cell,* 100**,** 57-70.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *cell,* 144**,** 646-674.

HARDMAN, G., PERKINS, S., BROWNRIDGE, P. J., CLARKE, C. J., BYRNE, D. P., CAMPBELL, A. E., KALYUZHNYY, A., MYALL, A., EYERS, P. A. & JONES, A. R. 2019. Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation. *The EMBO journal,* 38**,** e100847.

HENDERSON, T. O., BHATIA, S., PINTO, N., LONDON, W. B., MCGRADY, P., CROTTY, C., SUN, C.-L. & COHN, S. L. 2011. Racial and ethnic disparities in risk and survival in children with neuroblastoma: a Children's Oncology Group study. *Journal of clinical oncology,* 29**,** 76.

HERRMANN, A., RICE, M., LEVY, R., PIZER, B., LOSTY, P., MOSS, D. & SÉE, V. 2015. Cellular memory of hypoxia elicits neuroblastoma metastasis and enables invasion by non-aggressive neighbouring cells. *Oncogenesis,* 4**,** e138-e138.

HODGE, K., TEN HAVE, S., HUTTON, L. & LAMOND, A. I. 2013. Cleaning up the masses: exclusion lists to reduce contamination with HPLC-MS/MS. *Journal of proteomics,* 88**,** 92-103.

HOLMQUIST-MENGELBIER, L., FREDLUND, E., LÖFSTEDT, T., NOGUERA, R., NAVARRO, S., NILSSON, H., PIETRAS, A., VALLON-CHRISTERSSON, J., BORG, Å. & GRADIN, K. 2006. Recruitment of HIF-1α and HIF-2α to common target genes is differentially regulated in neuroblastoma: HIF-2α promotes an aggressive phenotype. *Cancer cell,* 10**,** 413-423.

HONG, T., LIU, W., LI, M. & CHEN, C. 2019. Click chemistry at the microscale. *Analyst,* 144**,** 1492-1512.

HORNBECK, P. V., CHABRA, I., KORNHAUSER, J. M., SKRZYPEK, E. & ZHANG, B. 2004. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics,* 4**,** 1551-1561.

HORVATOVICH, P., HOEKMAN, B., GOVORUKHINA, N. & BISCHOFF, R. 2010. Multidimensional chromatography coupled to mass spectrometry in analysing complex proteomics samples. *Journal of separation science,* 33**,** 1421-1437.

HOU, J., LIU, X., SHEN, J., ZHAO, G. & WANG, P. G. 2012. The impact of click chemistry in medicinal chemistry. *Expert opinion on drug discovery,* 7**,** 489-501.

HSIEH, C.-H., LEE, C.-H., LIANG, J.-A., YU, C.-Y. & SHYU, W.-C. 2010. Cycling hypoxia increases U87 glioma cell radioresistance via ROS induced higher and long-term HIF-1 signal transduction activity. *Oncology reports,* 24**,** 1629-1636.

HUANG, J., LI, J., TANG, J., WU, Y., DAI, F., YI, Z., WANG, Y., LI, Y., WU, Y. & REN, G. 2022. ZDHHC22-mediated mTOR palmitoylation restrains breast cancer growth and endocrine therapy resistance. *International journal of biological sciences,* 18**,** 2833.
HUANG, M. & WEISS, W. A. 2013. Neuroblastoma and MYCN. *Cold Spring Harbor perspectives in medicine,* 3**,** a014415.
HUANG, W. T., CEN, W. L., HE, R. Q., XIE, Y., ZHANG, Y., LI, P., GAN, T. Q., CHEN, G. & HU, X. H. 2017. Effect of miR-146a-5p on tumor growth in NSCLC using chick chorioallantoic membrane assay and bioinformatics investigation. *Molecular Medicine Reports,* 16**,** 8781-8792.
HUNT, D. F., YATES 3RD, J., SHABANOWITZ, J., WINSTON, S. & HAUER, C. R. 1986. Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences,* 83**,** 6233-6237.
HUNTER, T. & KARIN, M. 1992. The regulation of transcription by phosphorylation. *Cell,* 70**,** 375-387.
HUSSEIN, D., ESTLIN, E. J., DIVE, C. & MAKIN, G. W. 2006. Chronic hypoxia promotes hypoxia-inducible factor-1α–dependent resistance to etoposide and vincristine in neuroblastoma cells. *Molecular cancer therapeutics,* 5**,** 2241-2250.
IBRAHIM, H., WILUSZ, J. & WILUSZ, C. J. 2008. RNA recognition by 3′-to-5′ exonucleases: the substrate perspective. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1779**,** 256-265.
IKEDA, K., MONDEN, T., KANOH, T., TSUJIE, M., IZAWA, H., HABA, A., OHNISHI, T., SEKIMOTO, M., TOMITA, N. & SHIOZAKI, H. 1998. Extraction and analysis of diagnostically useful proteins from formalin-fixed, paraffin-embedded tissue sections. *Journal of Histochemistry & Cytochemistry,* 46**,** 397-403.
IKEGAKI, N., SHIMADA, H. & COMMITTEE, I. N. P. 2019. Subgrouping of unfavorable histology neuroblastomas with immunohistochemistry toward precision prognosis and therapy stratification. *JCO precision oncology,* 3.
IVAN, M., KONDO, K., YANG, H., KIM, W., VALIANDO, J., OHH, M., SALIC, A., ASARA, J. M., LANE, W. S. & KAELIN JR, W. G. 2001. HIFα targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science,* 292**,** 464-468.
JACKSON, A. L. & LINSLEY, P. S. 2010. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature reviews Drug discovery,* 9**,** 57-67.
JAWORSKI, S., SAWOSZ, E., GRODZIK, M., KUTWIN, M., WIERZBICKI, M., WŁODYGA, K., JASIK, A., REICHERT, M. & CHWALIBOG, A. 2013. Comparison of tumour morphology and structure from U87 and U118 glioma cells cultured on chicken embryo chorioallantoic membrane. *Bull Vet Inst Pulawy,* 57**,** 593-598.
JEFFERY, C. J. 1999. Moonlighting proteins. *Trends in biochemical sciences,* 24**,** 8-11.
JEFFRIES, O., GEIGER, N., ROWE, I. C., TIAN, L., MCCLAFFERTY, H., CHEN, L., BI, D., KNAUS, H. G., RUTH, P. & SHIPSTON, M. J. 2010. Palmitoylation of the S0-S1 linker regulates cell surface expression of voltage-and calcium-activated potassium (BK) channels. *Journal of Biological Chemistry,* 285**,** 33307-33314.
JENNINGS, B. C. & LINDER, M. E. 2012. DHHC protein S-acyltransferases use similar ping-pong kinetic mechanisms but display different acyl-CoA specificities. *Journal of Biological Chemistry,* 287**,** 7236-7245.
JÖGI, A., ØRA, I., NILSSON, H., LINDEHEIM, Å., MAKINO, Y., POELLINGER, L., AXELSON, H. & PÅHLMAN, S. 2002. Hypoxia alters gene expression in human neuroblastoma cells toward an immature and neural crest-like phenotype. *Proceedings of the National Academy of Sciences,* 99**,** 7021-7026.
JOHNSEN, J. I., DYBERG, C. & WICKSTRÖM, M. 2019. Neuroblastoma—A neural crest derived embryonal malignancy. *Frontiers in molecular neuroscience,* 12**,** 9.

JOHNSON, J. L., YARON, T. M., HUNTSMAN, E. M., KERELSKY, A., SONG, J., REGEV, A., LIN, T.-Y., LIBERATORE, K., CIZIN, D. M. & COHEN, B. M. 2023. An atlas of substrate specificities for the human serine/threonine kinome. *Nature*, 1-8.

JOHNSON, R. S., MARTIN, S. A. & BIEMANN, K. 1988. Collision-induced fragmentation of (M+ H)+ ions of peptides. Side chain specific sequence ions. *International Journal of Mass Spectrometry and Ion Processes,* 86**,** 137-154.

JONES, D. T., TAYLOR, W. R. & THORNTON, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics,* 8**,** 275-282.

JORDAN, J. D., LANDAU, E. M. & IYENGAR, R. 2000. Signaling networks: the origins of cellular multitasking. *Cell,* 103**,** 193-200.

JUNG, H. M., CHOI, S. J. & KIM, J. K. 2009. Expression profiles of SV40-immortalization-associated genes upregulated in various human cancers. *Journal of cellular biochemistry,* 106**,** 703-713.

JUNG, J. 2014. Human tumor xenograft models for preclinical assessment of anticancer drug development. *Toxicological research,* 30**,** 1-5.

KADRY, Y. A., LEE, J.-Y. & WITZE, E. S. 2021. Regulation of EGFR signalling by palmitoylation and its role in tumorigenesis. *Open Biology,* 11**,** 210033.

KANEKO, Y. & KNUDSON, A. G. 2000. Mechanism and relevance of ploidy in neuroblastoma. *Genes, Chromosomes and Cancer,* 29**,** 89-95.

KANG, W., YANG, Y., CHEN, C. & YU, C. 2022. CPSF1 positively regulates NSDHL by alternative polyadenylation and promotes gastric cancer progression. *American Journal of Cancer Research,* 12**,** 4566.

KAUR, I., YAROV-YAROVOY, V., KIRK, L. M., PLAMBECK, K. E., BARRAGAN, E. V., ONTIVEROS, E. S. & DÍAZ, E. 2016. Activity-dependent palmitoylation controls SynDIG1 stability, localization, and function. *Journal of Neuroscience,* 36**,** 7562-7568.

KAUR, J., SAXENA, M. & RISHI, N. 2021. An overview of recent advances in biomedical applications of click chemistry. *Bioconjugate Chemistry,* 32**,** 1455-1471.

KAZAZIAN, K., GO, C., WU, H., BRASHAVITSKAYA, O., XU, R., DENNIS, J. W., GINGRAS, A.-C. & SWALLOW, C. J. 2017. Plk4 Promotes Cancer Invasion and Metastasis through Arp2/3 Complex Regulation of the Actin CytoskeletonPlk4 and Arp2/3 in Invasion and Metastasis. *Cancer research,* 77**,** 434-447.

KEBARLE, P. & TANG, L. 1993. From ions in solution to ions in the gas phase-the mechanism of electrospray mass spectrometry. *Analytical chemistry,* 65**,** 972A-986A.

KENIGSBERG, S., BENTOV, Y., CHALIFA-CASPI, V., POTASHNIK, G., OFIR, R. & BIRK, O. S. 2009. Gene expression microarray profiles of cumulus cells in lean and overweight-obese polycystic ovary syndrome patients. *Molecular human reproduction,* 15**,** 89-103.

KHOLODENKO, B. N. 2006. Cell-signalling dynamics in time and space. *Nature reviews Molecular cell biology,* 7**,** 165-176.

KHOLODENKO, I. V., KALINOVSKY, D. V., DORONIN, I. I., DEYEV, S. M. & KHOLODENKO, R. V. 2018. Neuroblastoma origin and therapeutic targets for immunotherapy. *Journal of immunology research,* 2018.

KIDIYOOR, G. R., LI, Q., BASTIANELLO, G., BRUHN, C., GIOVANNETTI, I., MOHAMOOD, A., BEZNOUSSENKO, G. V., MIRONOV, A., RAAB, M. & PIEL, M. 2020. ATR is essential for preservation of cell mechanics and nuclear integrity during interstitial migration. *Nature communications,* 11**,** 1-16.

KISS, T., FAYET-LEBARON, E. & JÁDY, B. E. 2010. Box H/ACA small ribonucleoproteins. *Molecular cell,* 37**,** 597-606.

KLIPP, E. & LIEBERMEISTER, W. 2006. Mathematical modeling of intracellular signaling pathways. *BMC neuroscience,* 7**,** 1-16.

KLOCHENDLER, A., CASPI, I., COREM, N., MORAN, M., FRIEDLICH, O., ELGAVISH, S., NEVO, Y., HELMAN, A., GLASER, B. & EDEN, A. 2016. The genetic program of pancreatic β-cell replication in vivo. *Diabetes,* 65**,** 2081-2093.

KLUBA, C. A. & MINDT, T. L. 2013. Click-to-chelate: development of technetium and rhenium-tricarbonyl labeled radiopharmaceuticals. *Molecules,* 18**,** 3206-3226.

KOLB, H. C., FINN, M. & SHARPLESS, K. B. 2001. Click chemistry: diverse chemical function from a few good reactions. *Angewandte Chemie International Edition,* 40**,** 2004-2021.

KOREN, I., RAVIV, Z. & SHOSHAN-BARMATZ, V. 2010. Downregulation of voltage-dependent anion channel-1 expression by RNA interference prevents cancer cell growth in vivo. *Cancer biology & therapy,* 9**,** 1046-1052.

KRENACS, L., KRENACS, T., STELKOVICS, E. & RAFFELD, M. 2010. Heat-induced antigen retrieval for immunohistochemical reactions in routinely processed paraffin sections. *Immunocytochemical Methods and Protocols.* Springer.

KUNZ, P., SCHENKER, A., SÄHR, H., LEHNER, B. & FELLENBERG, J. 2019. Optimization of the chicken chorioallantoic membrane assay as reliable in vivo model for the analysis of osteosarcoma. *PLoS One,* 14**,** e0215312.

KUSHNER, B. H., GILBERT, F. & HELSON, L. 1986. Familial neuroblastoma. Case reports, literature review, and etiologic considerations. *Cancer,* 57**,** 1887-1893.

KYRÖNLAHTI, A., GODBOLE, N., AKINRINADE, O., SOINI, T., NYHOLM, I., ANDERSSON, N., HUKKINEN, M., LOHI, J., WILSON, D. B. & PIHLAJOKI, M. 2021. Evolving Up-regulation of Biliary Fibrosis–Related Extracellular Matrix Molecules After Successful Portoenterostomy. *Hepatology communications,* 5**,** 1036-1050.

LAI, T.-P., WRIGHT, W. E. & SHAY, J. W. 2018. Comparison of telomere length measurement methods. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 373**,** 20160451.

LANCET, T. 2020. COVID-19: protecting health-care workers. *Lancet (London, England),* 395**,** 922.

LANGE, V., PICOTTI, P., DOMON, B. & AEBERSOLD, R. 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular systems biology,* 4**,** 222.

LANUCARA, F. & EYERS, C. E. 2013. Top-down mass spectrometry for the analysis of combinatorial post-translational modifications. *Mass spectrometry reviews,* 32**,** 27-42.

LANUCARA, F., HOLMAN, S. W., GRAY, C. J. & EYERS, C. E. 2014. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature chemistry,* 6**,** 281-294.

LARSSON, L.-I. 2020. Section pretreatment, epitope demasking, and methods for dealing with unwanted staining. *Immunocytochemistry: theory and practice.* CRC Press.

LAVERDIÈRE, C., CHEUNG, N. K. V., KUSHNER, B. H., KRAMER, K., MODAK, S., LAQUAGLIA, M. P., WOLDEN, S., NESS, K. K., GURNEY, J. G. & SKLAR, C. A. 2005. Long-term complications in survivors of advanced stage neuroblastoma. *Pediatric blood & cancer,* 45**,** 324-332.

LAZZERINI, M., BARBI, E., APICELLA, A., MARCHETTI, F., CARDINALE, F. & TROBIA, G. 2020. Delayed access or provision of care in Italy resulting from fear of COVID-19. *The Lancet Child & Adolescent Health,* 4**,** e10-e11.

LE, S. Q. & GASCUEL, O. 2008. An improved general amino acid replacement matrix. *Molecular biology and evolution,* 25**,** 1307-1320.

LEE, J.-W., HUR, J., KWON, Y.-W., CHAE, C.-W., CHOI, J.-I., HWANG, I., YUN, J.-Y., KANG, J.-A., CHOI, Y.-E. & KIM, Y. H. 2021a. KAI1 (CD82) is a key molecule to control angiogenesis and switch angiogenic milieu to quiescent state. *Journal of Hematology & Oncology,* 14**,** 1-21.

LEE, K., ZHENG, Q., LU, Q., XU, F., QIN, G., ZHAI, Q., HONG, R., CHEN, M., DENG, W. & WANG, S. 2021b. CPSF4 promotes triple negative breast cancer metastasis by upregulating MDM4. *Signal transduction and targeted therapy,* 6**,** 1-3.

LEE, S. W., CHO, B. H., PARK, S. G. & KIM, S. 2004. Aminoacyl-tRNA synthetase complexes: beyond translation. *Journal of cell science,* 117**,** 3725-3734.

LI, B.-P., MAO, Y.-T., WANG, Z., CHEN, Y.-Y., WANG, Y., ZHAI, C.-Y., SHI, B., LIU, S.-Y., LIU, J.-L. & CHEN, J.-Q. 2018. CLIC1 promotes the progression of gastric cancer by regulating the MAPK/AKT pathways. *Cellular Physiology and Biochemistry,* 46**,** 907-924.

LI, J., CAI, Z., BOMGARDEN, R. D., PIKE, I., KUHN, K., ROGERS, J. C., ROBERTS, T. M., GYGI, S. P. & PAULO, J. A. 2021. TMTpro-18plex: the expanded and complete set of TMTpro reagents for sample multiplexing. *Journal of proteome research,* 20**,** 2964-2972.

LI, J., GAO, Z., ZHAO, D., ZHANG, L., QIAO, X., ZHAO, Y., DING, H., ZHANG, P., LU, J. & LIU, J. 2017. PI-273, a Substrate-Competitive, Specific Small-Molecule Inhibitor of PI4KIIα, Inhibits the Growth of Breast Cancer CellsThe First PI4KIIα Substrate-Competitive Specific Inhibitor. *Cancer Research,* 77**,** 6253-6266.

LIANG, W. H., FEDERICO, S. M., LONDON, W. B., NARANJO, A., IRWIN, M. S., VOLCHENBOUM, S. L. & COHN, S. L. 2020. Tailoring therapy for children with neuroblastoma on the basis of risk group classification: past, present, and future. *JCO Clinical Cancer Informatics,* 4**,** 895-905.

LIMPT, V. V., SCHRAMM, A., LAKEMAN, A., SLUIS, P. V., CHAN, A., NOESEL, M. V., BAAS, F., CARON, H., EGGERT, A. & VERSTEEG, R. 2004. The Phox2B homeobox gene is mutated in sporadic neuroblastomas. *Oncogene,* 23**,** 9280-9288.

LINDER, M. E. & DESCHENES, R. J. 2007. Palmitoylation: policing protein stability and traffic. *Nature reviews Molecular cell biology,* 8**,** 74-84.

LINSTEDT, A. D. & HAURI, H.-P. 1993. Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Molecular biology of the cell,* 4**,** 679-693.

LIPOWSKY, G., BISCHOFF, F. R., SCHWARZMAIER, P., KRAFT, R., KOSTKA, S., HARTMANN, E., KUTAY, U. & GÖRLICH, D. 2000. Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *The EMBO journal,* 19**,** 4362-4371.

LIPSKY, N. G. & PAGANO, R. E. 1985. A vital stain for the Golgi apparatus. *Science,* 228**,** 745-747.

LIU, W. M. & ZHANG, X. A. 2006. KAI1/CD82, a tumor metastasis suppressor. *Cancer letters,* 240**,** 183-194.

LOKE, S.-Y., WONG, P. T.-H. & ONG, W.-Y. 2017. Global gene expression changes in the prefrontal cortex of rabbits with hypercholesterolemia and/or hypertension. *Neurochemistry international,* 102**,** 33-56.

LOOK, A. T., HAYES, F. A., SHUSTER, J. J., DOUGLASS, E. C., CASTLEBERRY, R. P., BOWMAN, L., SMITH, E. & BRODEUR, G. 1991. Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastoma: a Pediatric Oncology Group study. *Journal of Clinical Oncology,* 9**,** 581-591.

LÓPEZ-GINÉS, C., MUÑOZ-HIDALGO, L., SAN-MIGUEL, T., MEGÍAS, J., TRIVIÑO, J. C., CALABUIG, S., ROLDÁN, P., CERDÁ-NICOLÁS, M. & MONLEÓN, D. 2021. Whole-exome sequencing, EGFR amplification and infiltration patterns in human glioblastoma. *American journal of cancer research,* 11**,** 5543.

MA, B., ZHANG, K., HENDRIE, C., LIANG, C., LI, M., DOHERTY-KIRBY, A. & LAJOIE, G. 2003. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid communications in mass spectrometry,* 17**,** 2337-2342.

MACEK, B., GNAD, F., SOUFI, B., KUMAR, C., OLSEN, J. V., MIJAKOVIC, I. & MANN, M. 2008. Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Molecular & cellular proteomics,* 7**,** 299-307.

MACLEAN, B., TOMAZELA, D. M., SHULMAN, N., CHAMBERS, M., FINNEY, G. L., FREWEN, B., KERN, R., TABB, D. L., LIEBLER, D. C. & MACCOSS, M. J. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics,* 26**,** 966-968.
MAKAROV, A. 2000. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical chemistry,* 72**,** 1156-1162.
MAKAROV, A. & SCIGELOVA, M. 2010. Coupling liquid chromatography to Orbitrap mass spectrometry. *Journal of Chromatography A,* 1217**,** 3938-3945.
MALGAPO, M. I. P. & LINDER, M. E. 2021. Substrate recruitment by zDHHC protein acyltransferases. *Open biology,* 11**,** 210026.
MALLICK, P. & KUSTER, B. 2010. Proteomics: a pragmatic perspective. *Nature biotechnology,* 28**,** 695-709.
MANGIR, N., RAZA, A., HAYCOCK, J. W., CHAPPLE, C. & MACNEIL, S. 2018. An improved in vivo methodology to visualise tumour induced changes in vasculature using the chick chorionic allantoic membrane assay. *in vivo,* 32**,** 461-472.
MANN, M. & JENSEN, O. N. 2003. Proteomic analysis of post-translational modifications. *Nature biotechnology,* 21**,** 255-261.
MANN, M., ONG, S.-E., GRØNBORG, M., STEEN, H., JENSEN, O. N. & PANDEY, A. 2002. Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends in biotechnology,* 20**,** 261-268.
MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-1934.
MAO, C. G., JIANG, S. S., SHEN, C., LONG, T., JIN, H., TAN, Q. Y. & DENG, B. 2020. BCAR1 promotes proliferation and cell growth in lung adenocarcinoma via upregulation of POLR2A. *Thoracic cancer,* 11**,** 3326-3336.
MAPANAO, A. K., CHE, P. P., SAROGNI, P., SMINIA, P., GIOVANNETTI, E. & VOLIANI, V. 2021. Tumor grafted–chick chorioallantoic membrane as an alternative model for biological cancer research and conventional/nanomaterial-based theranostics evaluation. *Expert Opinion on Drug Metabolism & Toxicology,* 17**,** 947-968.
MARIS, J. M. 2010. Recent advances in neuroblastoma. *New England Journal of Medicine,* 362**,** 2202-2211.
MARTÍNEZ, P. & BLASCO, M. A. 2015. Replicating through telomeres: a means to an end. *Trends in biochemical sciences,* 40**,** 504-515.
MASSENET, S., BERTRAND, E. & VERHEGGEN, C. 2017. Assembly and trafficking of box C/D and H/ACA snoRNPs. *RNA biology,* 14**,** 680-692.
MCLAFFERTY, F. W., HORN, D. M., BREUKER, K., GE, Y., LEWIS, M. A., CERDA, B., ZUBAREV, R. A. & CARPENTER, B. K. 2001. Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. *Journal of the American Society for Mass Spectrometry,* 12**,** 245-249.
MEIROWSKY, E., FREEMAN, L. & WOODARD, J. 1954. The response of embryonic chick membrane to Bowen's intraepithelial cancer. *Journal of Investigative Dermatology,* 22**,** 417-429.
MICHALSKI, A., DAMOC, E., LANGE, O., DENISOV, E., NOLTING, D., MÜLLER, M., VINER, R., SCHWARTZ, J., REMES, P. & BELFORD, M. 2012. Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Molecular & Cellular Proteomics,* 11.
MIO, C., LAVARONE, E., CONZATTI, K., BALDAN, F., TOFFOLETTO, B., PUPPIN, C., FILETTI, S., DURANTE, C., RUSSO, D. & ORLACCHIO, A. 2016. MCM5 as a target of BET inhibitors in thyroid cancer cells. *Endocrine-related cancer,* 23**,** 335-347.
MISRA, A. & GREEN, M. R. 2016. From polyadenylation to splicing: Dual role for mRNA 3'end formation factors. *RNA biology,* 13**,** 259-264.

MITCHELL, A., CHANG, H.-Y., DAUGHERTY, L., FRASER, M., HUNTER, S., LOPEZ, R., MCANULLA, C., MCMENAMIN, C., NUKA, G. & PESSEAT, S. 2015. The InterPro protein families database: the classification resource after 15 years. *Nucleic acids research,* 43**,** D213-D221.

MITCHELL, D. A., VASUDEVAN, A., LINDER, M. E. & DESCHENES, R. J. 2006. Thematic review series: lipid posttranslational modifications. Protein palmitoylation by a family of DHHC protein S-acyltransferases. *Journal of lipid research,* 47**,** 1118-1127.

MONCLAIR, T., BRODEUR, G. M., AMBROS, P. F., BRISSE, H. J., CECCHETTO, G., HOLMES, K., KANEKO, M., LONDON, W. B., MATTHAY, K. K. & NUCHTERN, J. G. 2009. The international neuroblastoma risk group (INRG) staging system: an INRG task force report. *Journal of clinical oncology,* 27**,** 298.

MORGAN, J., JANSANI, B. & NAVABI, H. 1997. A mechanism for high temperature antigen retrieval involving calcium complexes produced by formalin fixation. *Journal of Cellular Pathology,* 2**,** 89-92.

MORGAN, J., NAVABI, H., SCHMID, K. & JASANI, B. 1994. Possible role of tissue-bound calcium ions in citrate-mediated high-temperature antigen retrieval. *The Journal of pathology,* 174**,** 301-307.

MOROZ, E., CARLIN, S., DYOMINA, K., BURKE, S., THALER, H. T., BLASBERG, R. & SERGANOVA, I. 2009. Real-time imaging of HIF-1α stabilization and degradation. *PloS one,* 4**,** e5077.

MOTEGI, A., FUJIMOTO, J., KOTANI, M., SAKURABA, H. & YAMAMOTO, T. 2004. ALK receptor tyrosine kinase promotes cell growth and neurite outgrowth. *Journal of cell science,* 117**,** 3319-3329.

MOUSSA, G., ALAAEDDINE, R., ALAEDDINE, L. M., NASSRA, R., BELAL, A. S., ISMAIL, A., EL-YAZBI, A. F., ABDEL-GHANY, Y. S. & HAZZAA, A. 2018. Novel click modifiable thioquinazolinones as anti-inflammatory agents: Design, synthesis, biological evaluation and docking study. *European journal of medicinal chemistry,* 144**,** 635-650.

MOYLAN, C. A., MAVIS, A. M., JIMA, D., MAGUIRE, R., BASHIR, M., HYUN, J., CABEZAS, M. N., PARISH, A., NIEDZWIECKI, D. & DIEHL, A. M. 2022. Alterations in DNA methylation associate with fatty liver and metabolic abnormalities in a multi-ethnic cohort of pre-teenage children. *Epigenetics***,** 1-16.

MURPHY, J. B. & ROUS, P. 1912. The behavior of chicken sarcoma implanted in the developing embryo. *The Journal of experimental medicine,* 15**,** 119-132.

MUZ, B., DE LA PUENTE, P., AZAB, F. & AZAB, A. K. 2015. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia,* 3**,** 83.

NAKAYA, T. 2022. A specific gene-splicing alteration in the SNRNP70 gene as a hallmark of an ALS subtype. *Gene,* 818**,** 146203.

NEAR, R. I., ZHANG, Y., MAKKINJE, A., BORRE, P. V. & LERNER, A. 2007. AND-34/BCAR3 differs from other NSP homologs in induction of anti-estrogen resistance, cyclin D1 promoter activation and altered breast cancer cell morphology. *Journal of cellular physiology,* 212**,** 655-665.

NEGISHI, M., SARAYA, A., MOCHIZUKI, S., HELIN, K., KOSEKI, H. & IWAMA, A. 2010. A novel zinc finger protein Zfp277 mediates transcriptional repression of the Ink4a/arf locus through polycomb repressive complex 1. *PloS one,* 5**,** e12373.

NEIDERMYER JR, W. J. & WHELAN, S. P. 2019. Global analysis of polysome-associated mRNA in vesicular stomatitis virus infected cells. *PLoS pathogens,* 15**,** e1007875.

NELSON, P., REYNOLDS, G., WALDRON, E., WARD, E., GIANNOPOULOS, K. & MURRAY, P. 2000. Monoclonal antibodies. *Molecular pathology: MP,* 53**,** 111-117.

NEWMAN, D. R., KUHN, J. F., SHANAB, G. M. & MAXWELL, E. S. 2000. Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *Rna,* 6**,** 861-879.

NIETO, M. A., SARGENT, M. G., WILKINSON, D. G. & COOKE, J. 1994. Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science,* 264**,** 835-839.

NOGUERA, R., FREDLUND, E., PIQUERAS, M., PIETRAS, A., BECKMAN, S., NAVARRO, S. & PÅHLMAN, S. 2009. HIF-1α and HIF-2α Are Differentially Regulated In vivo in Neuroblastoma: High HIF-1α Correlates Negatively to Advanced Clinical Stage and Tumor VascularizationHIFs, VEGF, and Vascularization in Neuroblastoma. *Clinical cancer research,* 15**,** 7130-7136.

O'CONNOR, K. & CHEN, M. 2013. Dynamic functions of RhoA in tumor cell migration and invasion. *Small GTPases,* 4**,** 141-147.

OFFENBACHER, R., KNOLL, M. A. & LOEB, D. M. 2021. Delayed presentations of pediatric solid tumors at a tertiary care hospital in the Bronx due to COVID-19. *Pediatric Blood & Cancer,* 68**,** e28615.

OGNIBENE, M., MORINI, M., GARAVENTA, A., PODESTÀ, M. & PEZZOLO, A. 2020. Identification of a minimal region of loss on chromosome 6q27 associated with poor survival of high-risk neuroblastoma patients. *Cancer Biology & Therapy,* 21**,** 391-399.

OLD, W. M., MEYER-ARENDT, K., AVELINE-WOLF, L., PIERCE, K. G., MENDOZA, A., SEVINSKY, J. R., RESING, K. A. & AHN, N. G. 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics* S. *Molecular & cellular proteomics,* 4**,** 1487-1502.

OLSEN, J. V., BLAGOEV, B., GNAD, F., MACEK, B., KUMAR, C., MORTENSEN, P. & MANN, M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell,* 127**,** 635-648.

OLSEN, J. V., MACEK, B., LANGE, O., MAKAROV, A., HORNING, S. & MANN, M. 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods,* 4**,** 709-712.

OLSHINA, M. A., BEN-NISSAN, G. & SHARON, M. 2018. Functional regulation of proteins by 20S proteasome proteolytic processing. *Cell Cycle,* 17**,** 393-394.

OLSON, M. F., ASHWORTH, A. & HALL, A. 1995. An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. *Science,* 269**,** 1270-1272.

ONG, S.-E., BLAGOEV, B., KRATCHMAROVA, I., KRISTENSEN, D. B., STEEN, H., PANDEY, A. & MANN, M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics,* 1**,** 376-386.

ORGANIZATION, W. H. 2021. CureAll framework: WHO global initiative for childhood cancer: increasing access, advancing quality, saving lives.

ORSBURN, B. C. 2021. Proteome Discoverer—A Community Enhanced Data Processing Suite for Protein Informatics. *Proteomes,* 9**,** 15.

PAIZS, B. A. S., S. 2005. Fragmentation pathways of protonated peptides. *Mass spectrometry reviews,* 24**,** 508-548.

PALEY, E. L., SMELYANSKI, L., MALINOVSKII, V., SUBBARAYAN, P. R., BERDICHEVSKY, Y., POSTERNAK, N., GERSHONI, J. M., SOKOLOVA, O. & DENISOVA, G. 2007. Mapping and molecular characterization of novel monoclonal antibodies to conformational epitopes on NH2 and COOH termini of mammalian tryptophanyl-tRNA synthetase reveal link of the epitopes to aggregation and Alzheimer's disease. *Molecular immunology,* 44**,** 541-557.

PARIKH, D., SHORT, M., ESHMAWY, M. & BROWN, R. 2012. Surgical outcome analysis of paediatric thoracic and cervical neuroblastoma. *European journal of cardio-thoracic surgery,* 41**,** 630-634.

PARK, S. G., EWALT, K. L. & KIM, S. 2005. Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends in biochemical sciences,* 30**,** 569-574.

PARKINSON, D. R., MCCORMACK, R. T., KEATING, S. M., GUTMAN, S. I., HAMILTON, S. R., MANSFIELD, E. A., PIPER, M. A., DEVERKA, P., FRUEH, F. W. & JESSUP, J. M. 2014. Evidence of clinical utility: an unmet need in molecular diagnostics for patients with cancer. *Clinical Cancer Research,* 20**,** 1428-1444.

PEARSON, A. D., PINKERTON, C. R., LEWIS, I. J., IMESON, J., ELLERSHAW, C., MACHIN, D., GROUP, E. N. S., CANCER, C. S. & GROUP, L. 2008. High-dose rapid and standard induction chemotherapy for patients aged over 1 year with stage 4 neuroblastoma: a randomised trial. *The lancet oncology,* 9**,** 247-256.

PEDAMALLU, C. S. & POSFAI, J. 2010. Open source tool for prediction of genome wide protein-protein interaction network based on ortholog information. *Source code for biology and medicine,* 5**,** 1-6.

PELLETIER, J., THOMAS, G. & VOLAREVIĆ, S. 2018. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nature Reviews Cancer,* 18**,** 51-63.

PERDIVARA, I., DETERDING, L. J., PRZYBYLSKI, M. & TOMER, K. B. 2010. Mass spectrometric identification of oxidative modifications of tryptophan residues in proteins: chemical artifact or post-translational modification? *Journal of the American Society for Mass Spectrometry,* 21**,** 1114-1117.

PERKINS, D. N., PAPPIN, D. J., CREASY, D. M. & COTTRELL, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS: An International Journal,* 20**,** 3551-3567.

PETERSON, A. C., RUSSELL, J. D., BAILEY, D. J., WESTPHALL, M. S. & COON, J. J. 2012. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular & Cellular Proteomics,* 11**,** 1475-1488.

PICKART, C. M. & COHEN, R. E. 2004. Proteasomes and their kin: proteases in the machine age. *Nature reviews Molecular cell biology,* 5**,** 177-187.

PÓPULO, H., NUNES, B., SAMPAIO, C., BATISTA, R., PINTO, M. T., GASPAR, T. B., MIRANDA-ALVES, L., CAI, R.-Z., ZHANG, X. Y. & SCHALLY, A. V. 2017. Inhibitory effects of antagonists of growth hormone-releasing hormone (GHRH) in thyroid cancer. *Hormones and Cancer,* 8**,** 314-324.

POTKIN, S. G., GUFFANTI, G., LAKATOS, A., TURNER, J. A., KRUGGEL, F., FALLON, J. H., SAYKIN, A. J., ORRO, A., LUPOLI, S. & SALVI, E. 2009. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PloS one,* 4**,** e6501.

PUDELA, C., BALYASNY, S. & APPLEBAUM, M. A. 2020. Nervous system: embryonal tumors: neuroblastoma. *Atlas of genetics and cytogenetics in oncology and haematology,* 24**,** 284.

QANBAR, R. & BOUVIER, M. 2003. Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function. *Pharmacology & therapeutics,* 97**,** 1-33.

QIAN, L., XIE, H., ZHANG, L., ZHAO, Q., LÜ, J. & YU, Z. 2021. Piwi-interacting RNAs: A new class of regulator in human breast cancer. *Frontiers in Oncology,* 11**,** 695077.

RAINER, M. & BONN, G. K. 2015. Enrichment of phosphorylated peptides and proteins by selective precipitation methods. *Bioanalysis,* 7**,** 243-252.

RAMACHANDRAN, K. V., FU, J. M., SCHAFFER, T. B., NA, C. H., DELANNOY, M. & MARGOLIS, S. S. 2018. Activity-dependent degradation of the nascentome by the neuronal membrane proteasome. *Molecular cell,* 71**,** 169-177. e6.

RAMAGE, P., HEMMIG, R., MATHIS, B., COWAN-JACOB, S., RONDEAU, J., KALLEN, J., BLOMMERS, M., ZURINI, M. & RÜDISSER, S. 2002. Snags with tags: Some observations made with (His) 6-tagged proteins. *Life Sci. News,* 11**,** 1-4.

RAMANI, A. K., BUNESCU, R. C., MOONEY, R. J. & MARCOTTE, E. M. 2005. Consolidating the set of known human protein-protein interactions in preparation for large-scale mapping of the human interactome. *Genome biology,* 6**,** 1-12.

RAUHALA, H. E., TEPPO, S., NIEMELÄ, S. & KALLIONIEMI, A. 2013. Silencing of the ARP2/3 complex disturbs pancreatic cancer cell migration. *Anticancer research,* 33**,** 45-52.

ROBB, M. A., MCINNES, P. M. & CALIFF, R. M. 2016. Biomarkers and surrogate endpoints: developing common terminology and definitions. *Jama,* 315**,** 1107-1108.

ROCHE, M. E., LIN, Z., WHITAKER-MENEZES, D., ZHAN, T., SZUHAI, K., BOVEE, J. V., ABRAHAM, J. A., JIANG, W., MARTINEZ-OUTSCHOORN, U. & BASU-MALLICK, A. 2020. Translocase of the outer mitochondrial membrane complex subunit 20 (TOMM20) facilitates cancer aggressiveness and therapeutic resistance in chondrosarcoma. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1866**,** 165962.

RODRIGO, M. A. M., MICHALKOVA, H., STRMISKA, V., CASAR, B., CRESPO, P., DE LOS RIOS, V., IGNACIO CASAL, J., HADDAD, Y., GURAN, R. & ECKSCHLAGER, T. 2021. Metallothionein-3 promotes cisplatin chemoresistance remodelling in neuroblastoma. *Scientific reports,* 11**,** 1-14.

RODRIGUEZ, H. & PENNINGTON, S. R. 2018. Revolutionizing precision oncology through collaborative proteogenomics and data sharing. *Cell,* 173**,** 535-539.

ROMBI, B., MACDONALD, S., MAURIZIO, A., TARBELL, N. & YOCK, T. 2013. Proton radiotherapy for childhood tumors: an overview of early clinical results. *J Nucl Med Radiat Ther,* 4**,** 1-9.

ROSS, P. L., HUANG, Y. N., MARCHESE, J. N., WILLIAMSON, B., PARKER, K., HATTAN, S., KHAINOVSKI, N., PILLAI, S., DEY, S. & DANIELS, S. 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics,* 3**,** 1154-1169.

ROSSI, M., BUONUOMO, P. S., BATTAFARANO, G., CONFORTI, A., MARIANI, E., ALGERI, M., PELLE, S., D'AGOSTINI, M., MACCHIAIOLO, M. & DE VITO, R. 2020. Dissecting the mechanisms of bone loss in Gorham-Stout disease. *Bone,* 130**,** 115068.

ROSTOVTSEV, V. V., GREEN, L. G., FOKIN, V. V. & SHARPLESS, K. B. 2002. A stepwise huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie,* 114**,** 2708-2711.

ROTH, A. F., WAN, J., BAILEY, A. O., SUN, B., KUCHAR, J. A., GREEN, W. N., PHINNEY, B. S., YATES III, J. R. & DAVIS, N. G. 2006. Global analysis of protein palmitoylation in yeast. *Cell,* 125**,** 1003-1013.

SAKAI, A., ANDO, M., FUKUSUMI, T., REN, S., LIU, C., QUALLIOTINE, J., HAFT, S., SADAT, S., SAITO, Y. & GUO, T. W. 2020. Aberrant expression of CPSF1 promotes head and neck squamous cell carcinoma via regulating alternative splicing. *PloS one,* 15**,** e0233380.

SALAUN, C., GREAVES, J. & CHAMBERLAIN, L. H. 2010. The intracellular dynamic of protein palmitoylation. *Journal of Cell Biology,* 191**,** 1229-1238.

SANJAY, S. T., FU, G., DOU, M., XU, F., LIU, R., QI, H. & LI, X. 2015. Biomarker detection for disease diagnosis using cost-effective microfluidic platforms. *Analyst,* 140**,** 7062-7081.

SARMADY, M., DAMPIER, W. & TOZEREN, A. 2011. HIV protein sequence hotspots for crosstalk with host hub proteins. *PLoS One,* 6**,** e23293.

SCHIFFMANN, C., HANSEN, R., BAUMANN, S., KUBLIK, A., NIELSEN, P. H., ADRIAN, L., VON BERGEN, M., JEHMLICH, N. & SEIFERT, J. 2014. Comparison of targeted peptide

quantification assays for reductive dehalogenases by selective reaction monitoring (SRM) and precursor reaction monitoring (PRM). *Analytical and bioanalytical chemistry,* 406**,** 283-291.
SCHLESINGER, M., MAGEE, A. & SCHMIDT, M. 1980. Fatty acid acylation of proteins in cultured cells. *Journal of Biological Chemistry,* 255**,** 10021-10024.
SCHNEIDER, M., BELSOM, A. & RAPPSILBER, J. 2018. Protein tertiary structure by crosslinking/mass spectrometry. *Trends in biochemical sciences,* 43**,** 157-169.
SCHWAB, M., VARMUS, H. E. & BISHOP, J. M. 1985. Human N-myc gene contributes to neoplastic transformation of mammalian cells in culture. *Nature,* 316**,** 160-162.
SCHWARTZ, J. C., SENKO, M. W. & SYKA, J. E. 2002. A two-dimensional quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry,* 13**,** 659-669.
SCHWENK, R. W., HOLLOWAY, G. P., LUIKEN, J. J., BONEN, A. & GLATZ, J. F. 2010. Fatty acid transport across the cell membrane: regulation by fatty acid transporters. *Prostaglandins, Leukotrienes and Essential Fatty Acids (PLEFA),* 82**,** 149-154.
SCIGELOVA, M., HORNSHAW, M., GIANNAKOPULOS, A. & MAKAROV, A. 2011. Fourier transform mass spectrometry. *Molecular & Cellular Proteomics,* 10.
SEEGER, R. C., BRODEUR, G. M., SATHER, H., DALTON, A., SIEGEL, S. E., WONG, K. Y. & HAMMOND, D. 1985. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *New England Journal of Medicine,* 313**,** 1111-1116.
SEHGAL, R., ILHA, M., VAITTINEN, M., KAMINSKA, D., MÄNNISTÖ, V., KÄRJÄ, V., TUOMAINEN, M., HANHINEVA, K., ROMEO, S. & PAJUKANTA, P. 2021. Indole-3-Propionic Acid, a Gut-Derived Tryptophan Metabolite, Associates with Hepatic Fibrosis. *Nutrients,* 13**,** 3509.
SENKO, M. W., REMES, P. M., CANTERBURY, J. D., MATHUR, R., SONG, Q., ELIUK, S. M., MULLEN, C., EARLEY, L., HARDMAN, M. & BLETHROW, J. D. 2013. Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Analytical chemistry,* 85**,** 11710-11714.
SESHACHARYULU, P., PANDEY, P., DATTA, K. & BATRA, S. K. 2013. Phosphatase: PP2A structural importance, regulation and its aberrant expression in cancer. *Cancer letters,* 335**,** 9-18.
SESHACHARYULU, P., RACHAGANI, S., MUNIYAN, S., SIDDIQUI, J. A., CRUZ, E., SHARMA, S., KRISHNAN, R., KILLIPS, B. J., SHEININ, Y. & LELE, S. M. 2019. FDPS cooperates with PTEN loss to promote prostate cancer progression through modulation of small GTPases/AKT axis. *Oncogene,* 38**,** 5265-5280.
SHANG, B.-B., CHEN, J., WANG, Z.-G. & LIU, H. 2021. Significant correlation between HSPA4 and prognosis and immune regulation in hepatocellular carcinoma. *PeerJ,* 9**,** e12315.
SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research,* 13**,** 2498-2504.
SHARMA, V. K. & GRAHAM, N. J. 2010. Oxidation of amino acids, peptides and proteins by ozone: a review. *Ozone: Science & Engineering,* 32**,** 81-90.
SHAY, J. W. & WRIGHT, W. E. Role of telomeres and telomerase in cancer. Seminars in cancer biology, 2011. Elsevier, 349-353.
SHEINERMAN, F. B., NOREL, R. & HONIG, B. 2000. Electrostatic aspects of protein–protein interactions. *Current opinion in structural biology,* 10**,** 153-159.
SHI, S.-R., COTE, R. J. & TAYLOR, C. R. 2001. Antigen retrieval techniques: current perspectives. *Journal of Histochemistry & Cytochemistry,* 49**,** 931-937.

SHI, S.-R., IMAM, S. A., YOUNG, L., COTE, R. J. & TAYLOR, C. R. 1995. Antigen retrieval immunohistochemistry under the influence of pH using monoclonal antibodies. *Journal of Histochemistry & Cytochemistry,* 43**,** 193-201.

SHI, S.-R., KEY, M. E. & KALRA, K. L. 1991. Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *Journal of Histochemistry & Cytochemistry,* 39**,** 741-748.

SHOSHAN-BARMATZ, V., MIZRACHI, D. & KEINAN, N. 2013. Oligomerization of the mitochondrial protein VDAC1: from structure to function and cancer therapy. *Progress in molecular biology and translational science,* 117**,** 303-334.

SIEGEL, M. J. & JAJU, A. 2008. MR imaging of neuroblastic masses. *Magnetic Resonance Imaging Clinics of North America,* 16**,** 499-513.

SILVA, J. C., GORENSTEIN, M. V., LI, G.-Z., VISSERS, J. P. & GEROMANOS, S. J. 2006. Absolute quantification of proteins by LCMSE: A virtue of parallel MS acquisition* S. *Molecular & cellular proteomics,* 5**,** 144-156.

SIMÕES-COSTA, M. & BRONNER, M. E. 2015. Establishing neural crest identity: a gene regulatory recipe. *Development,* 142**,** 242-257.

SIMON, T., HERO, B., SCHULTE, J. H., DEUBZER, H., HUNDSDOERFER, P., VON SCHWEINITZ, D., FUCHS, J., SCHMIDT, M., PRASAD, V. & KRUG, B. 2017. 2017 GPOH guidelines for diagnosis and treatment of patients with neuroblastic tumors. *Klinische Pädiatrie,* 229**,** 147-167.

SIMON, T., LÄNGLER, A., HARNISCHMACHER, U., FRÜHWALD, M. C., JORCH, N., CLAVIEZ, A., BERTHOLD, F. & HERO, B. 2007. Topotecan, cyclophosphamide, and etoposide (TCE) in the treatment of high-risk neuroblastoma. Results of a phase-II trial. *Journal of cancer research and clinical oncology,* 133**,** 653-661.

SMITH, H., BOARD, M., PELLAGATTI, A., TURLEY, H., BOULTWOOD, J. & CALLAGHAN, R. 2016. The effects of severe hypoxia on glycolytic flux and enzyme activity in a model of solid tumors. *Journal of cellular biochemistry,* 117**,** 1890-1901.

SMITH, N. R., BAKER, D., JAMES, N. H., RATCLIFFE, K., JENKINS, M., ASHTON, S. E., SPROAT, G., SWANN, R., GRAY, N. & RYAN, A. 2010. Vascular Endothelial Growth Factor Receptors VEGFR-2 and VEGFR-3 Are Localized Primarily to the Vasculature in Human Primary Solid CancersVEGFR-2 and VEGFR-3 Status of Human Solid Tumors. *Clinical Cancer Research,* 16**,** 3548-3561.

SMOTRYS, J. E. & LINDER, M. E. 2004. Palmitoylation of intracellular signaling proteins: regulation and function. *Annual review of biochemistry,* 73**,** 559.

SOBOTT, F., WATT, S. J., SMITH, J., EDELMANN, M. J., KRAMER, H. B. & KESSLER, B. M. 2009. Comparison of CID versus ETD based MS/MS fragmentation for the analysis of protein ubiquitination. *Journal of the American Society for Mass Spectrometry,* 20**,** 1652-1659.

SPITERI, C. & MOSES, J. E. 2010. Copper-catalyzed azide–alkyne cycloaddition: regioselective synthesis of 1, 4, 5-trisubstituted 1, 2, 3-triazoles. *Angewandte Chemie International Edition,* 49**,** 31-33.

SPITZ, R., HERO, B., ERNESTUS, K. & BERTHOLD, F. 2003. Deletions in chromosome arms 3p and 11q are new prognostic markers in localized and 4s neuroblastoma. *Clinical cancer research,* 9**,** 52-58.

SPRITZ, R. A., STRUNK, K., SUROWY, C. S., HOCH, S. O., BARTON, D. E. & FRANCKE, U. 1987. The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing and RNA-binding. *Nucleic acids research,* 15**,** 10373-10391.

STAFFORD JR, G., KELLEY, P., SYKA, J., REYNOLDS, W. & TODD, J. 1984. Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes,* 60**,** 85-98.

STERN, C. D., ARTINGER, K. B. & BRONNER-FRASER, M. 1991. Tissue interactions affecting the migration and differentiation of neural crest cells in the chick embryo. *Development,* 113**,** 207-216.

STILLS, H. F. 2012. Polyclonal antibody production. *The laboratory rabbit, Guinea pig, Hamster, and other rodents.* Elsevier.

SUN, L., CHEN, G., SUN, A., WANG, Z., HUANG, H., GAO, Z., LIANG, W., LIU, C. & LI, K. 2020. BAG2 promotes proliferation and metastasis of gastric cancer via ERK1/2 signaling and partially regulated by miR186. *Frontiers in oncology,* 10**,** 31.

SUNNERHAGEN, M. S., PERSSON, E., DAHLQVIST, I., DRAKENBERG, T., STENFLO, J., MAYHEW, M., ROBIN, M., HANDFORD, P., TILLEY, J. & CAMPBELL, I. 1993. The effect of aspartate hydroxylation on calcium binding to epidermal growth factor-like modules in coagulation factors IX and X. *Journal of Biological Chemistry,* 268**,** 23339-23344.

SWEET, R., PAUL, A. & ZASTRE, J. 2010. Hypoxia induced upregulation and function of the thiamine transporter, SLC19A3 in a breast cancer cell line. *Cancer biology & therapy,* 10**,** 1101-1111.

SYKA, J. E., COON, J. J., SCHROEDER, M. J., SHABANOWITZ, J. & HUNT, D. F. 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences,* 101**,** 9528-9533.

SZEWCZYK, K., WIECZOREK, A., MŁYNARSKI, W., JANCZAR, S., WOSZCZYK, M., GAMROT, Z., CHABER, R., WYSOCKI, M., POGORZAŁA, M. & BIK-MULTANOWSKI, M. 2019. Unfavorable outcome of neuroblastoma in patients with 2p gain. *Frontiers in oncology,* 9**,** 1018.

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution,* 30**,** 2725-2729.

TATE, J. G., BAMFORD, S., JUBB, H. C., SONDKA, Z., BEARE, D. M., BINDAL, N., BOUTSELAKIS, H., COLE, C. G., CREATORE, C. & DAWSON, E. 2019. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic acids research,* 47**,** D941-D947.

TAUBE, J. M., AKTURK, G., ANGELO, M., ENGLE, E. L., GNJATIC, S., GREENBAUM, S., GREENWALD, N. F., HEDVAT, C. V., HOLLMANN, T. J. & JUCO, J. 2020. The Society for Immunotherapy in Cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. *Journal for immunotherapy of cancer,* 8.

TAUS, T., KÖCHER, T., PICHLER, P., PASCHKE, C., SCHMIDT, A., HENRICH, C. & MECHTLER, K. 2011. Universal and confident phosphorylation site localization using phosphoRS. *Journal of proteome research,* 10**,** 5354-5362.

TAYLOR, C. R., SHI, S. R. & COTE, R. J. 1996. Antigen retrieval for immunohistochemistry status and need for greater standardization. *Applied Immunohistochemistry and Molecular Morphology,* 4**,** 144-166.

TEN BERGE, R. L., MEIJER, C. J., DUKERS, D. F., KUMMER, J. A., BLADERGROEN, B. A., VOS, W., HACK, C. E., OSSENKOPPELE, G. J. & OUDEJANS, J. J. 2002. Expression levels of apoptosis-related proteins predict clinical outcome in anaplastic large cell lymphoma. *Blood, The Journal of the American Society of Hematology,* 99**,** 4540-4546.

THOMAS, R., SRIVASTAVA, S., KATREDDY, R. R., SOBIESKI, J. & WEIHUA, Z. 2019. Kinase-inactivated EGFR is required for the survival of wild-type EGFR-expressing cancer

cells treated with tyrosine kinase inhibitors. *International journal of molecular sciences,* 20**,** 2515.

THOMPSON, A., SCHÄFER, J., KUHN, K., KIENLE, S., SCHWARZ, J., SCHMIDT, G., NEUMANN, T. & HAMON, C. 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry,* 75**,** 1895-1904.

TIAN, L., MCCLAFFERTY, H., KNAUS, H.-G., RUTH, P. & SHIPSTON, M. J. 2012. Distinct acyl protein transferases and thioesterases control surface expression of calcium-activated potassium channels. *Journal of Biological Chemistry,* 287**,** 14718-14725.

TOMATIS, V. M., TRENCHI, A., GOMEZ, G. A. & DANIOTTI, J. L. 2010. Acyl-protein thioesterase 2 catalizes the deacylation of peripheral membrane-associated GAP-43. *PloS one,* 5**,** e15045.

TRINKLE-MULCAHY, L. 2012. Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry. *Proteomics,* 12**,** 1623-1638.

TSIEN, R. Y. 1998. The green fluorescent protein. *Annual review of biochemistry,* 67**,** 509-544.

TYERS, M. & MANN, M. 2003. From genomics to proteomics. *Nature,* 422**,** 193-197.

UTNES, P., LØKKE, C., FLÆGSTAD, T. & EINVIK, C. 2019. Clinically relevant biomarker discovery in high-risk recurrent neuroblastoma. *Cancer informatics,* 18**,** 1176935119832910.

VÄLIKANGAS, T., SUOMI, T. & ELO, L. L. 2018. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Briefings in bioinformatics,* 19**,** 1-11.

VAUPEL, P. 2009. Physiological mechanisms of treatment resistance. *The impact of tumor biology on cancer treatment and multidisciplinary strategies***,** 273-290.

VON STEDINGK, K., GISSELSSON, D. & BEXELL, D. 2019. Multidimensional intratumour heterogeneity in neuroblastoma. *Oncotarget,* 10**,** 3.

WAN, Y., QU, N., YANG, Y., MA, J., LI, Z. & ZHANG, Z. 2021. Identification of a 3-gene signature based on differentially expressed invasion genes related to cancer molecular subtypes to predict the prognosis of osteosarcoma patients. *Bioengineered,* 12**,** 5916-5931.

WANG, D., LI, Q., LI, Y. & WANG, H. 2018. The role of MCM5 expression in cervical cancer: Correlation with progression and prognosis. *Biomedicine & Pharmacotherapy,* 98**,** 165-172.

WANG, M., HERRMANN, C. J., SIMONOVIC, M., SZKLARCZYK, D. & VON MERING, C. 2015. Version 4.0 of PaxDb: protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics,* 15**,** 3163-3168.

WARNECKE, A., SANDALOVA, T., ACHOUR, A. & HARRIS, R. A. 2014. PyTMs: a useful PyMOL plugin for modeling common post-translational modifications. *BMC bioinformatics,* 15**,** 1-12.

WASHBURN, M. P., WOLTERS, D. & YATES, J. R. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature biotechnology,* 19**,** 242-247.

WEISS, W. A., ALDAPE, K., MOHAPATRA, G., FEUERSTEIN, B. G. & BISHOP, J. M. 1997. Targeted expression of MYCN causes neuroblastoma in transgenic mice. *The EMBO journal,* 16**,** 2985-2995.

WESTERLUND, I., SHI, Y., TOSKAS, K., FELL, S. M., LI, S., SUROVA, O., SÖDERSTEN, E., KOGNER, P., NYMAN, U. & SCHLISIO, S. 2017. Combined epigenetic and differentiation-based treatment inhibits neuroblastoma tumor growth and links HIF2α to tumor suppression. *Proceedings of the National Academy of Sciences,* 114**,** E6137-E6146.

WHANG-PENG, J., TRICHE, T., KNUTSEN, T., MISER, J., KAO-SHAN, S., TSAI, S. & ISRAEL, M. 1986. Cytogenetic characterization of selected small round cell tumors of childhood. *Cancer genetics and cytogenetics,* 21**,** 185-208.

WIENKE, J., DIERSELHUIS, M. P., TYTGAT, G. A., KÜNKELE, A., NIERKENS, S. & MOLENAAR, J. J. 2021. The immune landscape of neuroblastoma: challenges and opportunities for novel therapeutic strategies in pediatric oncology. *European journal of cancer,* 144**,** 123-150.

WILSON, A. L., SCHRECENGOST, R. S., GUERRERO, M. S., THOMAS, K. S. & BOUTON, A. H. 2013. Breast cancer antiestrogen resistance 3 (BCAR3) promotes cell motility by regulating actin cytoskeletal and adhesion remodeling in invasive breast cancer cells. *PloS one,* 8**,** e65678.

WU, X. & RAPOPORT, T. A. 2018. Mechanistic insights into ER-associated protein degradation. *Current opinion in cell biology,* 53**,** 22-28.

WU, Y., JIANG, Z., LI, Z., GU, J., YOU, Q. & ZHANG, X. 2018. Click chemistry-based discovery of [3-Hydroxy-5-(1 H-1, 2, 3-triazol-4-yl) picolinoyl] glycines as orally active hypoxia-inducing factor prolyl hydroxylase inhibitors with favorable safety profiles for the treatment of anemia. *Journal of Medicinal Chemistry,* 61**,** 5332-5349.

XU, Y., ANDERSON, D. E. & YE, Y. 2016. The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. *Cell discovery,* 2**,** 1-16.

YAMASHITA, S., KUDO, A., KAWAKAMI, H. & OKADA, Y. 2013. Mechanisms of angiogenic suppression in uteri exposed to diethylstilbestrol neonatally in the mouse. *Biology of reproduction,* 88**,** 116, 1-13.

YAMASHITA, S. & OKADA, Y. 2005. Mechanisms of heat-induced antigen retrieval: analyses in vitro employing SDS-PAGE and immunohistochemistry. *Journal of Histochemistry & Cytochemistry,* 53**,** 13-21.

YAN, S.-M., TANG, J.-J., HUANG, C.-Y., XI, S.-Y., HUANG, M.-Y., LIANG, J.-Z., JIANG, Y.-X., LI, Y.-H., ZHOU, Z.-W. & ERNBERG, I. 2013. Reduced expression of ZDHHC2 is associated with lymph node metastasis and poor prognosis in gastric adenocarcinoma. *PloS one,* 8**,** e56366.

YANAGIDA, M. 2002. Functional proteomics; current achievements. *Journal of Chromatography B,* 771**,** 89-106.

YÁÑEZ, Y., GRAU, E., RODRÍGUEZ-CORTEZ, V. C., HERVÁS, D., VIDAL, E., NOGUERA, R., HERNÁNDEZ, M., SEGURA, V., CAÑETE, A. & CONESA, A. 2015. Two independent epigenetic biomarkers predict survival in neuroblastoma. *Clinical epigenetics,* 7**,** 1-14.

YANG, W. & LU, Z. 2015. Pyruvate kinase M2 at a glance. *Journal of cell science,* 128**,** 1655-1660.

YATES, J. R., RUSE, C. I. & NAKORCHEVSKY, A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering,* 11**,** 49-79.

YESTE-VELASCO, M., LINDER, M. E. & LU, Y.-J. 2015. Protein S-palmitoylation and cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* 1856**,** 107-120.

YESTE-VELASCO, M., MAO, X., GROSE, R., KUDAHETTI, S. C., LIN, D., MARZEC, J., VASILJEVIĆ, N., CHAPLIN, T., XUE, L. & XU, M. 2014. Identification of ZDHHC14 as a novel human tumour suppressor gene. *The Journal of pathology,* 232**,** 566-577.

YINON, J. 2007. Detection of explosives by mass spectrometry. *Counterterrorist Detection Techniques of Explosives.* Elsevier.

YOSHIZUKA, N., MORIUCHI, R., MORI, T., YAMADA, K., HASEGAWA, S., MAEDA, T., SHIMADA, T., YAMADA, Y., KAMIHIRA, S. & TOMONAGA, M. 2004. An alternative transcript derived from the trio locus encodes a guanosine nucleotide exchange

factor with mouse cell-transforming potential. *Journal of Biological Chemistry,* 279**,** 43998-44004.

YOUNG, F. B., BUTLAND, S.L., SANDERS, S.S., SUTTON, L.M. AND HAYDEN, M.R. 2012. Putting proteins in their place: palmitoylation in Huntington disease and other neuropsychiatric diseases. . *Progress in neurobiology,* 97**,** 220-238.

YUAN, M., CHEN, X., SUN, Y., JIANG, L., XIA, Z., YE, K., JIANG, H., YANG, B., YING, M. & CAO, J. 2020. ZDHHC12-mediated claudin-3 S-palmitoylation determines ovarian cancer progression. *Acta Pharmaceutica Sinica B,* 10**,** 1426-1439.

YUE, X., ZHAO, Y., LIU, J., ZHANG, C., YU, H., WANG, J., ZHENG, T., LIU, L., LI, J. & FENG, Z. 2015. BAG2 promotes tumorigenesis through enhancing mutant p53 protein levels and function. *Elife,* 4**,** e08401.

ZAGE, P. E., LOUIS, C. U. & COHN, S. L. 2012. New aspects of neuroblastoma treatment: ASPHO 2011 symposium review. *Pediatric blood & cancer,* 58**,** 1099-1105.

ZHANG, B., LIU, Y., LIU, D. & YANG, L. 2017. Targeting cleavage and polyadenylation specific factor 1 via shRNA inhibits cell proliferation in human ovarian cancer. *Journal of biosciences,* 42**,** 417-425.

ZHANG, H., JOSEPH, J., CROW, J. & KALYANARAMAN, B. 2004. Mass spectral evidence for carbonate-anion-radical-induced posttranslational modification of tryptophan to kynurenine in human Cu, Zn superoxide dismutase. *Free Radical Biology and Medicine,* 37**,** 2018-2026.

ZHANG, J., XIN, L., SHAN, B., CHEN, W., XIE, M., YUEN, D., ZHANG, W., ZHANG, Z., LAJOIE, G. A. & MA, B. 2012. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Molecular & cellular proteomics,* 11.

ZHANG, M., DAI, W., LI, Z., TANG, L., CHEN, J. & CHEN, C. 2021a. HSPA4 knockdown retarded progression and development of colorectal cancer. *Cancer Management and Research,* 13**,** 4679.

ZHANG, N., XIE, Y., TAI, Y., GAO, Y., GUO, W., YU, W., LI, J., FENG, X., HAO, J. & GAO, Y. 2016. Bufalin inhibits hTERT expression and colorectal cancer cell growth by targeting CPSF4. *Cellular Physiology and Biochemistry,* 40**,** 1559-1569.

ZHANG, W., ZHU, J., HE, X., LIU, X., LI, J., LI, W., YANG, P., WANG, J., HU, K. & ZHANG, X. 2019. Exosome complex genes mediate RNA degradation and predict survival in mantle cell lymphoma. *Oncology letters,* 18**,** 5119-5128.

ZHANG, Y., FICARRO, S. B., LI, S. & MARTO, J. A. 2009. Optimized Orbitrap HCD for quantitative analysis of phosphopeptides. *Journal of the American Society for Mass Spectrometry,* 20**,** 1425-1434.

ZHANG, Y., FONSLOW, B. R., SHAN, B., BAEK, M.-C. & YATES III, J. R. 2013. Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews,* 113**,** 2343-2394.

ZHANG, Y., LI, F., FU, K., LIU, X., LIEN, I.-C. & LI, H. 2021b. Potential Role of S-Palmitoylation in Cancer Stem Cells of Lung Adenocarcinoma. *Frontiers in cell and developmental biology,* 9.

ZHENG, S., ZHAO, D., HOU, G., ZHAO, S., ZHANG, W., WANG, X., LI, L., LIN, L., TANG, T.-S. & HU, Y. 2022. iASPP suppresses Gp78-mediated TMCO1 degradation to maintain Ca 2+ homeostasis and control tumor growth and drug resistance. *Proceedings of the National Academy of Sciences,* 119**,** e2111380119.

ZHONG, W., SUN, B., LU, C., YU, H., WANG, C., HE, L., GU, J., CHEN, S., LIU, Y. & JING, X. 2016. Problems and solutions in click chemistry applied to drug probes. *Scientific reports,* 6**,** 1-6.

ZHOU, B., HAO, Q., LIANG, Y. & KONG, E. 2022. Protein palmitoylation in cancer: molecular functions and therapeutic potential. *Molecular oncology*.

ZHOU, B., LIU, L., REDDIVARI, M. & ZHANG, X. A. 2004. The palmitoylation of metastasis suppressor KAI1/CD82 is important for its motility-and invasiveness-inhibitory activity. *Cancer research,* 64**,** 7455-7463.

ZHOU, X., XU, C.-J., WANG, J.-X., DAI, T., YE, Y.-P., CUI, Y.-M., LIAO, W.-T., WU, X.-L. & OU, J.-P. 2015. Metastasis-associated in colon cancer-1 associates with poor prognosis and promotes cell invasion and angiogenesis in human cervical cancer. *International Journal of Gynecologic Cancer,* 25.

ZHU, L., BRASSARD, C. J., ZHANG, X., GUHA, P. M. & CLARK, R. J. 2016. On the Mechanism of Copper (I)-Catalyzed Azide–Alkyne Cycloaddition. *The Chemical Record,* 16**,** 1501-1517.

ZIJLSTRA, A., MELLOR, R., PANZARELLA, G., AIMES, R. T., HOOPER, J. D., MARCHENKO, N. D. & QUIGLEY, J. P. 2002. A quantitative analysis of rate-limiting steps in the metastatic cascade using human-specific real-time polymerase chain reaction. *Cancer research,* 62**,** 7083-7092.

ZLOTORYNSKI, E. 2015. Ubiquitylation determines cell fate. *Nature Reviews Molecular Cell Biology,* 16**,** 640-641.

ZUBAREV, R. A. 2004. Electron-capture dissociation tandem mass spectrometry. *Current opinion in biotechnology,* 15**,** 12-16.

ZUBAREV, R. A., KELLEHER, N. L. & MCLAFFERTY, F. W. 1998. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society,* 120**,** 3265-3266.