

**Regulation of TMEM106B, a Gene
Implicated in Neurodegeneration, by
Neuronal Src Kinase**

Zoe Manser

MSc by Research

University of York
Biology

August 2023

Abstract

TMEM106B, a lysosomal membrane protein, has been identified as a risk factor for frontotemporal lobar degeneration with TDP-43 inclusions (FTLD-TDP). Although the cellular function of TMEM106B is poorly understood, it has been shown to affect lysosomal morphology and trafficking within neurons. Furthermore, in a preliminary high-throughput phosphoproteomics screen, TMEM106B was identified as a substrate for the tyrosine kinase C-Src, and its neuronal splice variant N1-Src. The identified phosphosite, Y50, lies within a putative YXX Φ sorting motif, which is typically bound by AP2 adaptor protein μ 1 subunit (AP2M1). AP2M1 is involved in clathrin mediated endocytosis and has been shown to bind TMEM106B in a co-immunoprecipitation study. Since TMEM106B phosphorylation has not been previously investigated, this project aimed to assess how phosphorylation of TMEM106B by N1-Src might regulate AP2M1 binding and how this might influence TMEM106B function.

This study utilised recombinant N-terminal TMEM106B (TMEM106B(1-96)) to show TMEM106B(1-96)-Y50 is phosphorylated by N1-Src *in vitro*. Assessment of the effect of phosphorylation on protein-protein interactions found TMEM106B(1-96) phosphorylation did not affect dimerisation of TMEM106B(1-96) but did reduce binding to AP2M1 *in vitro*. Since AP2M1 has been found to regulate trafficking of proteins to lysosomes, the localisation of wild-type and a Y50 phosphonull mutant TMEM106B-GFP was assessed in N1-Src inducible HeLa cells. However, fluorescent microscopy showed TMEM106B localisation was unaffected by the expression of N1-Src or Y50 mutation, but assessment of this in neurons is required. Despite this, mutation of TMEM106B-Y50 was found to increase cellular TMEM106B protein levels and expression of TMEM106B was found to decrease N1-Src protein levels, indicating TMEM106B phosphorylation may affect protein stability. Altogether, this study demonstrates that N1-Src phosphorylation of TMEM106B affects TMEM106B function through the regulation of binding partners such as AP2M1, providing a new avenue of investigation that could shed light on how TMEM106B dysfunction contributes to disease.

Table of contents

Abstract.....	2
Table of contents	3
List of figures.....	5
List of tables.....	6
Acknowledgements.....	7
Author's declaration	8
1. Introduction.....	9
1.1 Frontotemporal Lobar Degeneration (FTLD).....	9
1.2 Overview of <i>TMEM106B</i>	10
1.3 Structure and cleavage of <i>TMEM106B</i>	13
1.4 <i>TMEM106B</i> and lysosomes	14
1.4.1 <i>Overview of lysosomes</i>	14
1.4.2 <i>TMEM106B function in lysosomes</i>	16
1.5 <i>TMEM106B</i> and progranulin (PGRN)	17
1.6 Interactions of the <i>TMEM106B</i> C-terminus	18
1.7 Interactions of the <i>TMEM106B</i> N-terminus.....	19
1.8 Tyrosine phosphorylation of <i>TMEM106B</i> by Src family kinases.....	21
1.8.1 <i>C-Src and N1-Src</i>	22
1.8.2 <i>TMEM106B and phosphorylation</i>	24
1.9 Aims.....	26
2. Materials and methods	27
2.1 Materials	27
2.1.1 <i>Molecular biology reagents</i>	27
2.1.2 <i>Biochemistry reagents</i>	27
2.1.3 <i>Cell biology reagents</i>	28
2.2 Molecular biology.....	28
2.2.1 <i>Generation of TMEM106B(1-96)-GST plasmids</i>	28
2.2.2 <i>Generation of His-Δ80-N1-Src plasmid</i>	29
2.2.3 <i>Agarose gel electrophoresis</i>	30
2.2.4 <i>Bacterial transformation</i>	30
2.2.5 <i>Purification of plasmid DNA from E. coli XL-10 gold</i>	30
2.3 Protein biochemistry.....	31

2.3.1	<i>Protein induction (AP2M1, His-Δ80-N1-Src and TMEM106B(1-96))</i>	31
2.3.2	<i>Protein purification</i>	31
2.3.3	<i>Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)</i>	33
2.3.4	<i>Western blotting</i>	33
2.3.5	<i>In vitro N1-Src kinase assays</i>	34
2.3.6	<i>AP2M1 binding assays</i>	34
2.3.7	<i>Size-exclusion chromatography with multi-angle light scattering (SEC-MALS)</i>	35
2.4	<i>Cell biology</i>	36
2.4.1	<i>Culture and passage of N1-Src HeLas</i>	36
2.4.2	<i>Transient transfection of N1-Src HeLas</i>	37
2.4.3	<i>Lysis of N1-Src HeLas for western blotting</i>	37
2.4.4	<i>Immunocytochemistry and fluorescent microscopy</i>	37
2.5	<i>Data analysis</i>	38
3.	<i>Results</i>	39
3.1	<i>Y18 and Y50 are potential phosphosites for TMEM106B</i>	39
3.2	<i>Production and purification of -WT, -3F and TMEM106B(1-96)-Y50F</i>	41
3.3	<i>Successful expression and purification of N1-Src</i>	43
3.4	<i>N1-Src phosphorylates TMEM106B(1-96)</i>	45
3.5	<i>TMEM106B(1-96)-WT is a dimer in solution</i>	47
3.6	<i>A putative phospho-dependent interaction between TMEM106B(1-96) and AP2M1 50</i>	
3.7	<i>Expression of N1-Src does not affect TMEM106B localization in HeLa cells</i>	54
4.	<i>Discussion</i>	61
4.1	<i>TMEM106B(1-96)-GST and His-Δ80-N1-Src provide a reliable source of protein for in vitro studies</i>	61
4.2	<i>The TMEM106B N-terminus is a newly identified N1-Src substrate in vitro</i>	63
4.3	<i>Phosphorylation of TMEM106B(1-96) does not affect dimer formation</i>	64
4.4	<i>Phosphorylation reduces AP2M1 binding to TMEM106B(1-96) in vitro</i>	66
4.5	<i>TMEM106B localises to the endoplasmic reticulum and lysosomes</i>	69
4.6	<i>Conclusions and future directions</i>	74
	<i>List of abbreviations</i>	76
	<i>List of references</i>	79

List of figures

Figure 1 - Structure of TMEM106B, a lysosomal type-II membrane protein, containing a cytosolic N-terminus (1-96) and a luminal C-terminus (116-274)	11
Figure 2 - Regulated intramembrane proteolysis of TMEM106B.....	14
Figure 3 - Intracellular trafficking to the lysosome.....	15
Figure 4 - Binding of TMEM106B to MAP6 on microtubules acts as a molecular break to inhibit retrograde transport along dendrites	21
Figure 5 – The structure of Src.....	22
Figure 6 – TMEM106B N-terminus contains three tyrosine residues: Y18, Y50 and Y84	25
Figure 7 – TMEM106B has two potential phosphosites: Y18 and Y50	40
Figure 8 – Production and purification of TMEM106B(1-96)-GST.....	42
Figure 9 – Expression and purification of N1-Src protein in E. coli.....	44
Figure 10 - Characterisation of in vitro kinase assays with His-Δ80-N1-Src	46
Figure 11 – SEC-MALS analysis shows TMEM106B(1-96)-WT predominantly forms dimers in solution	48
Figure 12 – Alphafold prediction of TMEM106B dimers	49
Figure 13 – AP2M1 binds to TMEM106B(1-96)-WT	51
Figure 14 - Tyrosine phosphorylation or mutation of TMEM106B(1-96)-Y50 reduces AP2M1 binding	53
Figure 15 – Analysis of N1-Src HeLa cells transfected with TMEM106B	56
Figure 16 – TMEM106B is located predominantly in perinuclear regions	57
Figure 17 – LAMP1 staining is reduced in transfected cells	58
Figure 18 – TMEM106B localizes at lysosomes	59
Figure 19 – TMEM106B localizes to the endoplasmic reticulum (ER).....	60
Figure 20 – AP2M1 binding to TMEM106B-Y50 could promote endocytosis of TMEM106B at the plasma membrane.....	68
Figure 21 – Phosphorylation of TMEM106B-Y50 could inhibit binding to MAP6, promoting retrograde transport of lysosomes in neurons.....	71

List of tables

Table 1 - DNA sequences of -WT, -3F and TMEM106B(1-96)-Y50F GBlocks™ and the forward and reverse primers for PCR amplification.....	29
Table 2 – Forward and reverse primers for PCR amplification of His-Δ80-N1-Src and PTP1B for creation of His-Δ80-N1-Src plasmid.	30
Table 3 - Concentrations of primary and secondary antibodies used in western blotting.	34
Table 4 – Run details used for SEC-MALS analysis of TMEM106B(1-96)-WT samples.....	36
Table 5 – Concentrations of primary and secondary antibodies used for immunocytochemistry.....	38

Acknowledgements

I would like to thank my supervisor, Dr Gareth Evans, for his continuous support and guidance throughout the project. I am also eternally grateful for the help of Hannah Walker and Lubna Nuhu-Soso, who were always willing to answer my questions, and provided great company in the lab. I would also like to thank Dr Andrew Leech for his expertise regarding SEC-MALS and interesting insight into the different analytical techniques available.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other university. All sources are acknowledged as references. All experiments were performed by the author, except for the $\Delta 80$ -N1-Src bacterial plasmid cloning (Figure 9C), which was completed by Hannah Walker (Evans lab) and SEC-MALS (Figure 11), which was performed by Dr Andrew Leech (Molecular Interactions, Technology facility, University of York).

1. Introduction

1.1 Frontotemporal Lobar Degeneration (FTLD)

Dementia is an umbrella term that encompasses multiple diseases associated with a progressive decline in brain function and affects approximately 50 million people worldwide (1). Symptoms typically include loss of memory, language deficits and difficulties with carrying out daily tasks. Frontotemporal lobar degeneration (FTLD) is a syndrome characterised by neurodegeneration that occurs specifically in the frontal and temporal lobes. It is a predominantly pre-senile form of dementia, meaning onset typically occurs before the age of 65 (2) and affects approximately 15 per 100,000 people aged 45-64 (2). FTLD is often used interchangeably with the term frontotemporal dementia (FTD), but FTLD typically refers to a larger group of disorders, one of which is FTD. FTLD will be used henceforth for simplicity.

There are various phenotypic presentations of FTLD, making it difficult to distinguish from other forms of dementia, such as Alzheimer's disease (AD), particularly during the initial stages. Patients with behavioural variant dementia (bvFTD), approximately 40-60% of cases, typically present with personality changes that include increased lethargy and loss of social inhibition (3). This compares to those who suffer from the language variant, which is referred to as primary progressive aphasia (PPA). In this case, patients tend to have difficulties with speech and processing of information (4). However, in both cases, patients can also suffer from rare movement disorders called corticobasal syndrome and progressive supranuclear palsy, which occur due to atrophy and neuronal death in parts of the brain that control movement (5). Other movement related disorders within FTD include FTD-ALS, whereby patients suffer from the progressive muscle weakness typically associated with amyotrophic lateral sclerosis (ALS) (6), and FTLD with parkinsonism, caused by inherited mutations in tau, producing symptoms akin to that of Parkinson's disease (7).

FTLD is also characterised according to whether it is genetic (familial FTLD) or sporadic (sFTLD). In both instances, FTLD is associated with aggregations of proteins, typically either Tau or transactivation response DNA-binding protein (TDP-43). Tau accounts for 45% of FTLD cases and is involved in stabilising microtubules (8,9). Mutations can have an effect at both the mRNA and protein level, and either prevent binding, or promote the formation of hyperphosphorylated filaments. TDP-43 on the other hand is an RNA/DNA binding protein

encoded by the *TARDBP* gene, with various functions including transcriptional activation and mRNA splicing (10,11). Mutations in *TARDBP* account for <1% of FTLD cases, but TDP-43 proteinopathy occurs in ~43% of FTLD (11). Consequently, mutations in various proteins besides TDP-43 have also been discovered to result in TDP-43 aggregations. Most common mutations are those in progranulin (*GRN*) (12) or chromosome 9 open reading frame 72 (*C9ORF72*) (13). However, it is largely unknown how exactly these mutations lead to the formation of TDP-43 aggregates. Furthermore, phenotypic presentation and age of onset can vary vastly among individuals, even with the same disease type. This suggests that additional genetic and/or environmental factors are at play. Consequently, various studies have been conducted throughout the past decade with the aim of broadening knowledge of how FTLD may develop at the molecular level. One such breakthrough was the discovery of single nucleotide polymorphisms (SNPs) in *TMEM106B*, a gene that encodes a lysosomal membrane protein, and was found to increase the risk of FTLD-TDP for those with *GRN* mutations (14).

1.2 Overview of *TMEM106B*

In 2010, a genome-wide association study (GWAS) was conducted, which analysed 515 individuals with FTLD-TDP, with the aim of identifying additional genes that are implicated in FTLD-TDP (14). Van Deerlin *et al.* (14) identified three main SNPs that mapped to a region on 7p21.3 containing *TMEM106B*, with the risk-associated SNPs conferring increased *TMEM106B* expression. The top SNP, rs1990622, lies downstream of *TMEM106B*, with the major T allele conferring risk for FTLD-TDP with *GRN* mutations, whilst the minor C allele protects against FTLD development for those carrying the *GRN* mutation. The two additional SNPs, rs1020004 and rs6966915, are located within intron 3 and 5 of *TMEM106B* respectively (14–16). Furthermore, rs1990622 is in perfect linkage disequilibrium with an additional SNP, rs3173615. rs3173615 is a coding variant that causes an amino acid change at 185, with T185 forming the risk isoform and S185 forming the protective isoform (17–19). How the S185 *TMEM106B* isoform protects against FTLD is unclear, however studies have shown that the S185 isoform has lower cellular levels compared to the T185 isoform, likely due to increased protein degradation (19). Nonetheless, several other labs confirmed that *TMEM106B* expression is increased in patients with FTLD-TDP, supporting the link between *TMEM106B* expression and neurodegeneration (16,20,21).

Since the initial GWAS, TMEM106B has been characterised as a type-II lysosomal membrane protein (Figure 1) with a cytosolic N-terminus and luminal C-terminus and is expressed in various cell types including endothelial cells, neurons, and glia (22). Whilst the exact function of TMEM106B is unknown, multiple studies have demonstrated its involvement in lysosomal morphology (20,22–24), maintenance of lysosome pH (20,25) and lysosome trafficking within neurons (26). Additionally, *TMEM106B* has two paralogues: *TMEM106A* and *TMEM106C*, both of which are

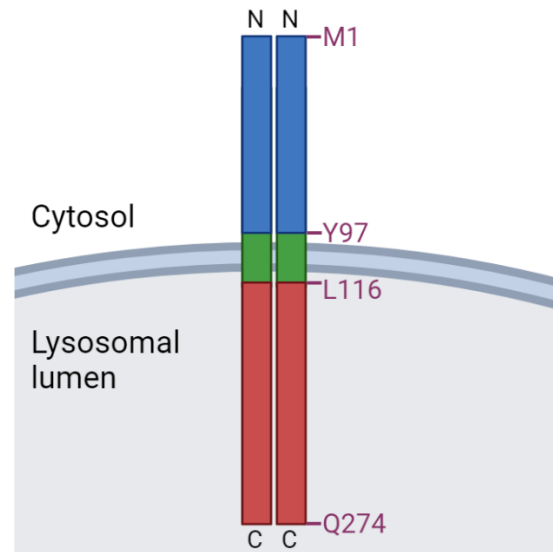


Figure 1 - Structure of TMEM106B, a lysosomal type-II membrane protein, containing a cytosolic N-terminus (1-96) and a luminal C-terminus (116-274).

transmembrane proteins of largely unknown function. However, both TMEM106A and TMEM106C have been detected at the lysosomal membrane, with TMEM106C shown to form heterodimers with TMEM106B (24).

How TMEM106B expression is increased in FTLD-TDP is unknown. Through global microRNA expression profiling, *TMEM106B* was shown to be regulated by microRNA-132 and microRNA-212 (20). MicroRNAs (miRs) are small, non-coding RNA molecules that repress gene expression by binding to complementary mRNAs and either promote their degradation or prevent their translation. In the case of *TMEM106B*, miR-132 and miR-212 have the same recognition sequence and can bind to two sites located within the 3' untranslated region (3' UTR) of *TMEM106B* (20). Sequential deletion of the two sites showed that site 1 resulted in a large increase in *TMEM106B* expression, whilst deletion of only site 2 led to a minimal increase in *TMEM106B* levels. However, deletion of both sites had the most significant impact on restoring *TMEM106B* expression, suggesting both sites are important in repressing *TMEM106B*. Chen-Plotkin *et al.* (20) also showed that miR-132 and miR-212 show significantly decreased expression in *GRN* positive FTLD-TDP patients and that the levels of *TMEM106B* are increased in these patients, corroborating the findings of the GWAS. This suggests that lack of repression by miR-132 and miR-212 may provide one mechanism for increased expression of *TMEM106B* in *GRN* positive FTLD-TDP.

In an additional mechanism, Gallagher *et al.* (27) found that the risk allele of rs1990622 can recruit the CCCTC-binding factor (CTCF), a conserved zinc finger protein that functions as both a transcriptional activator and a repressor (28). In the case of *TMEM106B*, CTCF binds 48 bp away from SNP rs1990622 and through long-range chromatin-looping interactions, can simultaneously bind an additional site at the *TMEM106B* promoter, resulting in enhanced expression of *TMEM106B*. Therefore, increased expression of *TMEM106B* may occur through multiple different mechanisms and include regulation at both the transcriptional and protein levels.

Besides FTL, *TMEM106B* has been identified to play a role in several other neurodegenerative diseases including AD, where the *TMEM106B* rs199622 risk variant was identified to contribute to increased TDP-43 pathology (29). Similarly, *TMEM106B* rs199622 has been implicated in limbic- predominant age-related TDP-43 encephalopathy (30) and has also been shown to affect brain ageing due to cognitive defects from neuronal loss in the frontal cerebral cortex (31). Furthermore, a D252N mutation in *TMEM106B* protein has been associated with hypomyelinating leukodystrophy, a neurodevelopmental disease whereby patients have myelination defects (32). In this case, it was found that the D252N mutation prevents the increased lysosome acidification typically seen with wild-type *TMEM106B* overexpression and leads to clustering of lysosomes near the nucleus (33).

Furthermore, through cryo-electron microscopy of sarkosyl extracts of patient brain tissue, the *TMEM106B* C-terminus was shown to form amyloid fibrils in patients with FTL-TDP (34–36). It was found that residues S120-G254 form a five layered core containing 18 β -strands, with *TMEM106B* remaining fully glycosylated at all five glycosylation sites (34–36). This fold was found to be conserved across four patients with FTL-TDP in one study (34), but three different protofilament folds were discovered when assessing *TMEM106B* filaments across different neurodegenerative conditions (35). *TMEM106B* amyloid fibrils were also found in older individuals considered to have normal neurology, but not in brains from younger individuals, suggesting age might be a contributing factor for the formation of *TMEM106B* fibrils (35). As such, the prevalence of *TMEM106B* in normal ageing, neurodevelopmental and neurodegenerative disorders, has further highlighted the importance of understanding its cellular function.

1.3 Structure and cleavage of TMEM106B

When TMEM106B was first implicated in FTL, knowledge of its function was very limited. Primary studies using differential membrane extraction and subsequent mutagenesis, revealed TMEM106B to be a single-pass transmembrane protein of 274 amino acids, with a cytoplasmic N-terminal domain and a C-terminal luminal domain (Figure 1) (22). Co-localisation studies with LAMP1 and GFP-Rab7, illustrated that TMEM106B localised to late-endosomes and lysosomes (20,22,23). Since targeting of lysosomal proteins typically occurs via N-glycosylation sites, the amino acid sequence of TMEM106B was analysed for the presence of N-glycosylation consensus sequences, leading to the identification of five sites within the C-terminus: N145, N151, N164, N183 and N256. Sequential mutation of each of the five sites, and treatment with endoglycosidase H showed that the first three sites contain non-complex glycans, whilst the fourth and fifth sites contained complex glycans. The authors went on to show that N4 (N183) and N5 (N256) are required for lysosomal sorting of TMEM106B (22).

Further biophysical characterisation of the N-terminus (amino acids 1-96) through circular dichroism (CD) and NMR spectroscopy, illustrated that the cytoplasmic domain is intrinsically disordered and lacks any stable three-dimensional structure (37). However, several small regions throughout the sequence are populated with weak helical conformations, suggesting that these segments may become more folded upon binding with interaction partners (37). As such, the overall disordered nature of TMEM106B N-terminus may prove advantageous in permitting interactions with a wide range of binding partners within the cytoplasm. This compares to the highly structured C-terminus, for which the crystal structure has recently been solved (38). This found that there are 7 β -sandwich folds, with a single α -helix present at residues 208-216, stabilised by a disulphide bond between residues C214 and C253.

Within the lysosome membrane, TMEM106B undergoes regulated intramembrane proteolysis (RIP), whereby the N- and C-terminus are cleaved from the transmembrane domain via two cleavage events (Figure 2) (39,40). The first is thought to occur at G127 via an unknown resident lysosome protease, resulting in luminal domain shedding that produces an N-terminal fragment (NTF). This NTF, containing the intracellular domain (ICD) and transmembrane domain, undergoes a second intramembrane cleavage event via the SPPL2a

protease, which cleaves close to L106 and releases the ICD (39,40). However, analysis of the ICD product on a gel reveals multiple bands, suggesting some degeneracy in SPPL2a cleavage sites (39). Furthermore, when, or why RIP occurs is largely unknown. On the one hand, it may regulate lysosomal trafficking, since the TMEM106B

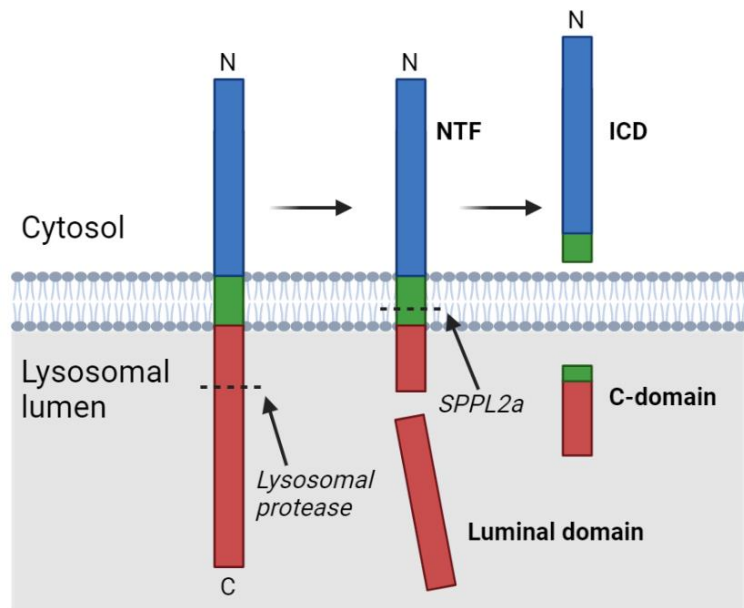


Figure 2 - Regulated intramembrane proteolysis of TMEM106B. The luminal C-terminal domain is cleaved by an unknown lysosomal protease. The remaining N-terminal fragment (NTF) is cleaved by SPPL2a protease, releasing the intracellular domain (ICD) and the remaining C-terminal domain.

N-terminus has been shown to interact with the C-terminus of microtubule associated protein 6

(MAP6) (26). Or alternatively, RIP processing may occur to prevent excess build-up of TMEM106B and therefore may promote degradation of the protein instead. Due to the short half-life of ICDs, it is particularly difficult to determine the exact function of proteolytic processing of TMEM106B.

The C-terminal fragment that forms amyloid fibrils also complicates things further, as it is formed from residues 120-254 (34,35). This suggests that there may be an alternative cleavage event that yields a 135-residue fragment. This could potentially occur through non-canonical shedding by SPPL2A, which has been observed for tumour necrosis factor alpha (TNF- α) (41,42). As such, which shedding process dominates could be determined by cellular conditions, with the non-canonical processing favoured in a disease state. However, whether the formation of these fibrils occurs within the lysosome or in the cytosol, and whether they even contribute to disease progression, is currently unknown.

1.4 TMEM106B and lysosomes

1.4.1 Overview of lysosomes

Lysosomes are dynamic organelles, mainly known for their role in the degradation of various cellular components including proteins, lipids, and nucleic acids, which are delivered via

endocytic, phagocytic and autophagic pathways (Figure 3) (43). Lysosomes have an acidic lumen, maintained at pH 5, enclosed by a single lipid bilayer containing 100+ membrane proteins. Maintaining the correct pH is essential for the correct functioning of approximately 50 lysosomal hydrolases that reside in the lysosome. This is achieved by the vacuolar ATPase that pumps protons across the membrane and into the lumen, via ATP hydrolysis (44,45). Lysosomal proteins are synthesised at the rough endoplasmic reticulum and are tagged with mannose-6-phosphate in the Golgi, allowing binding to mannose-6-phosphate receptors (46). These receptors facilitate transport, via clathrin-coated vesicles, from the trans-Golgi network to the lysosome, where the acidic pH of the lysosome promotes dissociation of the lysosomal proteins from the receptor (Figure 3).

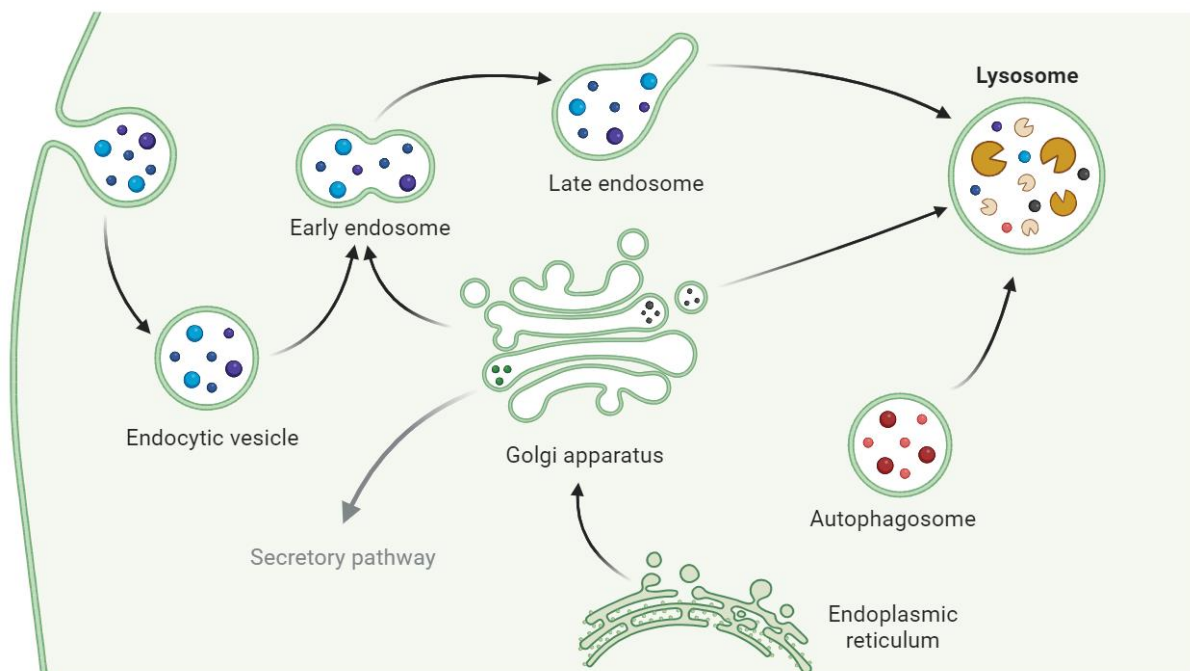


Figure 3 - Intracellular trafficking to the lysosome. Lysosomes are dynamic organelles, largely responsible for the degradation of cellular components received by endocytic (blue) and autophagic (red) pathways. Lysosomal proteins are synthesised at the endoplasmic reticulum and then transported to the Golgi apparatus, where they are tagged with mannose-6-phosphate. This directs their trafficking to the lysosome.

Additionally, lysosomes have been shown to be involved in homeostasis, cell signalling and the immune response. For example, the epidermal growth factor receptor (EGFR) is removed from the membrane via clathrin-mediated endocytosis and then targeted to the lysosome for degradation, reducing signalling through this pathway (47). This has a subsequent effect on cell growth, proliferation, and survival. As such, lysosomes are now regarded as a central signalling hub which have the capacity to respond to different environmental cues. As part of this, lysosomes can undergo fission and fusion to maintain correct size and number and are

also trafficked along microtubules in neuronal dendrites (48). Consequently, correct lysosome function is vital for healthy cell maintenance, with mutations in lysosomal proteins repeatedly associated with neurodegeneration (49).

1.4.2 *TMEM106B* function in lysosomes

Regarding lysosomal function, *TMEM106B* has been shown to influence various properties including size, morphology, and motility (23,24). In particular, overexpression of *TMEM106B* led to enlargement of lysosomes in neurons (23,24), which subsequently impaired degradation of endocytic cargo, such as the EGFR (23). It was also found that some of these enlarged lysosomes displayed abnormal morphology, which was typical of lysosomal fission or fusion intermediates. However, the same authors determined that lysosomal morphology did not differ between wild-type (WT) *TMEM106B* and the T185S coding variant (23). Similarly, knock-down of *TMEM106B* led to a reduction in lysosome size, however this was only visible with super resolution STED microscopy, suggesting the changes are less dramatic compared to *TMEM106B* overexpression (24).

Stagi *et al.* (24) also observed decreased motility of lysosomes upon overexpression of *TMEM106B* in cortical cultures but did not find a significant difference between retrograde or anterograde transport of lysosomes. This is in contrast to studies by Lüningschrör *et al.* (50) and Schwenk *et al.* (26) who showed that knockout or knockdown of *TMEM106B* increased retrograde transport in axons and dendrites of neurons respectively. This is further supported by the formation of large LAMP-1 positive vacuoles at the axon initial segment, since this suggests that anterograde transport of lysosomes has been inhibited (50). Similarly, it was shown that increased retrograde transport in dendrites resulted in decreased dendritic branching (26). Precisely how *TMEM106B* functions to regulate lysosome motility in dendrites and axons is unknown. In dendrites, it has been shown that the N-terminus of *TMEM106B* is able to interact with the C-terminus of MAP6 (26). MAP6 is a protein involved in the stabilisation of microtubules, as well as neuroreceptor homeostasis and endocytosis (51,52). Schwenk *et al.* demonstrated that enhanced retrograde motility of lysosomes along dendrites upon *TMEM106B* knockdown, could be rescued by MAP6 overexpression, suggesting the two are involved in a common pathway. But it is not known how the two proteins interact to modulate lysosomal transport or whether a similar mechanism occurs in axons.

Additionally, knockdown of TMEM106B affects lysosomal pH, with decreased acidification observed (25). Through co-immunoprecipitation assays, it was found that the C-terminus of TMEM106B interacts with V-ATPase via its accessory protein 1 (AP1), thereby aiding stabilisation of the protein (25). The same authors also showed that lack of TMEM106B reduces protein levels of the transmembrane V0 subunits of V-ATPase, further impairing lysosomal acidification. It would be expected that overexpression of TMEM106B would lead to increased acidification of lysosomes, however contradicting results have been obtained. In studies with lung cancer cell lines, it was found that overexpression of TMEM106B resulted in increased fluorescence of lysotracker (a pH sensitive dye), which is indicative of increased acidification (53). A similar result was also obtained in HEK293 cells, where overexpression of TMEM106B increased lysosome acidification, shown through the pH sensitive dye ApHID (33). However, in HeLa cells, overexpression of TMEM106B reduced acidification of lysosomes compared to control cells expressing endogenous levels of TMEM106B (20). These differences may be due to the different cell types employed, however a more thorough understanding of the molecular relationship between TMEM106B and V-ATPase is required.

1.5 TMEM106B and progranulin (PGRN)

The association between risk variants of TMEM106B and FTLD is most prominent in patients with *GRN* mutations. *GRN* encodes the protein progranulin (PGRN), a secreted glycoprotein involved in various cellular processes including growth, repair, and inflammation (54). Autosomal dominant mutations in *GRN* lead to haploinsufficiency of PGRN, resulting in reduced levels of the secreted protein. The formation of aggregated ubiquitinated TDP-43 is also observed in *GRN* patients, further contributing to FTLD (12). PGRN comprises 7.5 granulin domains, each of which are connected by short linker sequences (55). These granulin domains can be processed in the lysosome to form granulin peptides (56) and have been shown to interact with lysosomal enzymes, although their exact function remains unclear (57). Nonetheless, several studies have revealed the importance of PGRN in maintaining correct lysosome function. For example, PGRN knockout mice (*Grn*^{-/-}) have a similar phenotype to neuronal ceroid lipofuscinosis, a group of neurodegenerative lysosome storage disorders (58–60). Furthermore, *Grn*^{-/-} mice show increased accumulation of lysosomal proteins including cathepsin D and LAMP1 (61), as well as the pigment lipofuscin, which contains digested lipid

(62). This suggests that *GRN* mutations may contribute to FTLD due to deficits in lysosome function.

Despite there being a clear link between PGRN, lysosomes, and how this may contribute to neurodegeneration, there is a greater uncertainty of how PGRN and TMEM106B may be mechanistically linked. Various studies have shown that increased expression of TMEM106B increases intracellular levels of PGRN, but decreases the amount secreted (19,20,23). Similarly, overexpression of TMEM106B reduces the proteolytic processing of PGRN into granulin peptides, which in turn may contribute to lysosome dysfunction (63). This is supported by studies in *GRN*^{-/-} mice, which show that increased TMEM106B expression further promoted lysosome dysfunction compared to *GRN*^{-/-} mice with endogenous TMEM106B (64). Consequently, it was thought that knockdown of TMEM106B could rescue the lysosome deficits caused by *GRN* mutations. Initially, studies using a gene-trap approach, which reduces TMEM106B protein levels by 90-95%, had contradictory data. One study showed that TMEM106B knockdown restored lysosomal protein levels and rescued behavioural deficits (25), whilst a second study found that there was no statistically significant difference in phenotypes between mice with reduced and endogenous levels of TMEM106B (65). However, additional studies in double knockout mice (*GRN*^{-/-} *TMEM106B*^{-/-}) found that mice continued to display phenotypes associated with neurodegeneration and lysosome dysfunction, suggesting that lack of TMEM106B may also promote disease (66–69). As such, it is likely that levels of TMEM106B are carefully regulated in the cell and that imbalances in either direction can prevent correct lysosome function. However, how exactly the function of TMEM106B may intertwine with PGRN in the lysosome is unclear.

1.6 Interactions of the TMEM106B C-terminus

In addition to binding V-ATPase, the C-terminus of TMEM106B has also been shown to interact with several other lysosomal proteins including cathepsin D (33) and charged multivesicular body protein 2B (CHMP2B) (70,71). CHMP2B is a subunit of endosomal sorting complex required for transport III (ESCRT-III), a protein involved in the trafficking of molecules through the endo-lysosomal pathway (72). A mutation in CHMP2B, whereby the final 36 amino acids of the protein are replaced by a single valine residue (CHMP2B^{Intron5}), has been shown to contribute to FTLD by halting trafficking through the endolysosomal pathway (73).

Through co-immunoprecipitation studies, it was shown that the T185 risk variant of TMEM106B binds to CHMP2B more tightly than the S185 protective variant (71), suggesting one way in which TMEM106B variants might be implicated in disease. The authors also found that the T185 variant enhanced the endolysosomal defects caused by CHMP2B^{Intron5}, further contributing to neuronal toxicity. However, it was later found that knockdown of TMEM106B can rescue the trafficking and dendritic branching deficits caused by CHMP2B^{Intron5} (70). This suggests that TMEM106B could potentially act as a therapeutic target for FTLD patients with CHMP2B^{Intron5} mutations.

Cathepsin D, a lysosomal protease, also interacts with the C-terminus of TMEM106B in oligodendrocytes (33). A reduction in cathepsin D has shown to lead to myelination deficits due to a reduced trafficking of proteolipid protein (PLP), a transmembrane protein that forms the myelin sheath and is delivered to the membrane through lysosome exocytosis (74). More recently, it has been demonstrated that deletion of *TMEM106B* also reduces myelination, with significantly lower surface levels of the myelin proteins PLP and myelin oligodendrocyte glycoprotein (MOG) detected (33,75). This suggests that lack of TMEM106B prevents correct trafficking of lysosomes, which in turn leads to reduced insertion of PLP and MOG into the myelin sheath. This is also supported by increased clustering of lysosomes detected at the perinuclear region (33). However, the same authors found that the TMEM106B D252 mutation linked to hypomyelinating leukodystrophy does not affect its interaction with cathepsin D, nor does it affect the stability of TMEM106B protein. As such, it is likely that TMEM106B interacts with additional proteins that also modulate myelination. Currently it is unknown how TMEM106B might regulate cathepsin D and/or lysosome trafficking in oligodendrocytes.

1.7 Interactions of the TMEM106B N-terminus

Various binding partners have also been detected for the N-terminus of TMEM106B. Through co-immunoprecipitation studies, it was found that MAP6 (26), AP2 adaptor protein μ 1 subunit (AP2M1) (24) and clathrin heavy chain (CLTC) (24) interact with TMEM106B N-terminus. Both AP2M1 and CLTC are involved in clathrin-mediated endocytosis (76). AP2M1 is one subunit of AP2, an adaptor complex which binds receptors to be endocytosed at the plasma membrane. AP2 in turn recruits clathrin, a protein which consists of three heavy chains (CLTC)

and three light chains, forming a triskelion which interlock together to form a cage around the receptors bound by AP2 (77,78). This results in the formation of a transport vesicle that traffics the receptor to lysosomes, whilst AP2 and clathrin dissociate and are recycled back to the membrane. Furthermore, in a yeast-two hybrid screen, it was shown that vacuolar protein sorting-associated protein 11 (VPS11), VPS13D and TMEM106C also bind the cytosolic domain of TMEM106B (24). VPS11 forms part of the VPS-C core complex, which is a component of HOPS and CORVET (79). These are two protein complexes involved in the endo-lysosomal trafficking pathway, with CORVET thought to act as a Rab5 effector, aiding fusion of early endosomes, whilst HOPS regulates trafficking to late endosomes via Rab7 (79). VPS13D on the other hand is involved in lipid transfer between organelles and has also been shown to affect mitochondrial morphology and peroxisome biogenesis (80,81).

Consequently, the various binding partners of TMEM106B N-terminus suggest the cytosolic domain may be functionally involved in trafficking at various stages, both at the plasma membrane and then through the endo-lysosomal pathway. However, only one additional study has been conducted to further assess the functional relationship between TMEM106B and the above proteins. As previously mentioned, Schwenk and colleagues (26) conducted an in-depth study of TMEM106B and MAP6, which showed that the binding of these two proteins affected the trafficking of lysosomes within dendrites, with decreased TMEM106B expression resulting in increased retrograde movement of lysosomes. However, the authors also showed that knockdown of TMEM106B resulted in reduced dendritic branching in hippocampal neurons, which could also be rescued by knockdown of MAP6. This is likely due to the increased retrograde lysosome motility, since this could promote removal of lipids from the membrane, resulting in dendrite loss. Impaired trafficking of lysosomes could also inhibit signalling through growth factors, further leading to reduced dendritic branching. Consequently, the authors hypothesised that the binding of TMEM106B to microtubule bound MAP6 acts as a molecular break to prevent the retrograde trafficking of lysosomes. As such, reduced expression of TMEM106B prevents binding of lysosomes to MAP6, allowing them to be transported faster throughout dendrites (Figure 4). Additional knockdown of MAP6 could then restore correct lysosomal transport by promoting binding to microtubules through alternative proteins.

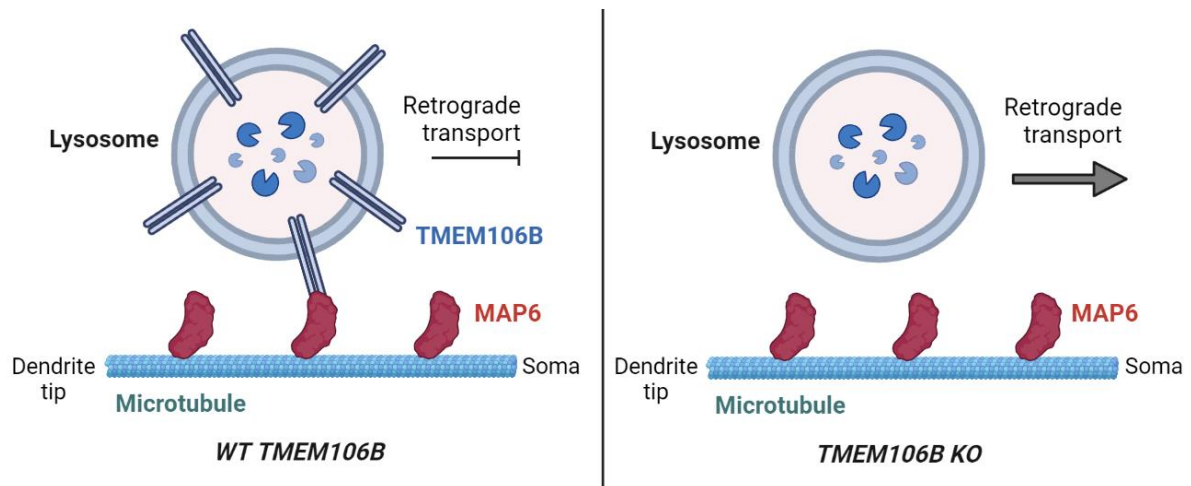


Figure 4 - Binding of TMEM106B to MAP6 on microtubules acts as a molecular break to inhibit retrograde transport along dendrites. Knockout of TMEM106B prevents binding to MAP6, promoting retrograde transport.

Although there are currently no additional studies which have confirmed binding, or assessed the functional importance of AP2M1 and TMEM106B, analysis of the amino acid sequence of TMEM106B reveals the presence of two AP2M1 YXX Φ consensus sequences and one extended dileucine motif (82). Both motifs, where Φ is an amino acid with a bulky hydrophobic side chain and X is any amino acid, typically act as an 'internalisation' signal, whereby AP2M1 recognises and binds the sequence on a membrane receptor, promoting subsequent endocytosis (83). The first sequence occurs at amino acids 18-21 (YDGV), the second occurs at amino acids 50-53 (YVEF), whilst the third occurs at the end of the N-terminus at E75 (ENQLVALI) supporting that these two proteins likely interact *in vivo*. However, the role of this interaction is unclear. One study found that mutation of the leucine and isoleucine residues within the dileucine motif prevented localisation of TMEM106B to the lysosome, but this trafficking was not attributed to AP2M1 (82). Another possibility may be that AP2M1 regulates lysosome exocytosis or lysosome fission, but there is currently no evidence in the literature to support this.

1.8 Tyrosine phosphorylation of TMEM106B by Src family kinases

There is limited knowledge regarding how TMEM106B is post-translationally modified or how this regulates protein function. Glycosylation of the C-terminus (N183 and N256) is required for targeting of TMEM106B to the lysosome (22), but currently there are no known modifications of the N-terminus in the literature. However, TMEM106B was identified as a substrate in a preliminary high-throughput phosphoproteomics screen for the non-receptor

tyrosine kinase C-Src and its neuronal splice variant N1-Src (84), suggesting TMEM106B function may be modulated via tyrosine phosphorylation.

1.8.1 C-Src and N1-Src

C-Src is part of the Src family of kinases (SFK), a group of non-receptor tyrosine kinases involved in a diverse set of signalling pathways that control various aspects of cellular function including proliferation, differentiation, and membrane trafficking (85). Several SFK members are highly expressed in the mammalian brain and have been implicated in neuronal functions including neurotransmitter release and learning and memory (86–88).

All SFK members share a conserved structure consisting of five domains, as summarised in Figure 5. The SH4 domain, located at the N-terminus, is responsible for localising the kinase to membranes via myristoylation (89). This is a process whereby myristic acid, an unsaturated C14 fatty acid, is added to the N-terminal glycine of proteins via the enzyme N-myristoyltransferase. Src associates both at the plasma membrane and intracellular membranes such as late endosomes and lysosomes, allowing Src to phosphorylate its substrates and promote signalling through various pathways (90). This contrasts with the SH3 and SH2 domains, which facilitate protein-protein interactions to mediate substrate binding and regulate kinase activity. The SH3 domain

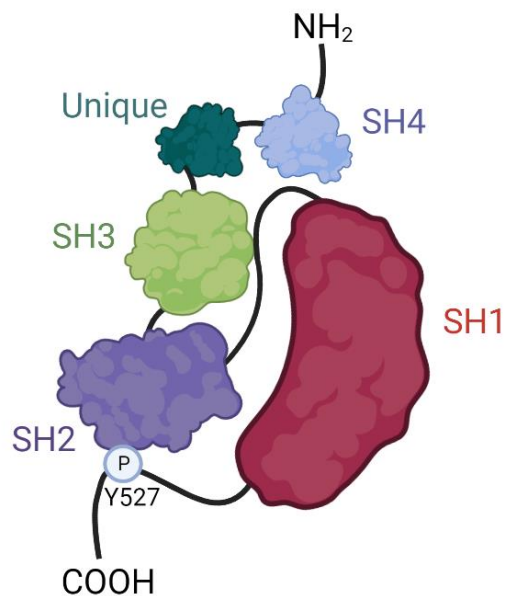


Figure 5 – The structure of Src. From the N to C-terminus, the domains are as follows: the myristoylated SH4 domain, the unique domain, the substrate binding SH3 and SH2 domains, and the catalytic SH1 domain. In the inactive state, the C-terminal tail is bound to SH2 via phosphorylation of Y527 and the SH2-kinase linker is bound to the SH3 domain.

recognises the proline rich motifs with a canonical core of PxxP, whilst the SH2 domain recognises phosphorylated tyrosines within various motifs, such as YEEL (91,92). The SH1 domain is responsible for the catalytic activity of the kinase and binds MgATP to facilitate phosphorylation of the substrate. However, catalytic activity can only occur once Src has been converted to an active state (93). In the inactive state, the Src C-terminus is bound to the SH2 domain through phosphorylation of Y527 and the SH2-kinase linker is bound to the SH3

domain (94). Conversion to the active state requires dephosphorylation of Y527 and/or SH3 binding to a PxxP motif, triggering a conformational change which exposes Y416 in the SH1 domain. Subsequent autophosphorylation of Y416 fully activates Src, leading to phosphorylation of bound substrates (95).

C-Src was discovered in 1979 through its homology to the oncogenic V-Src protein found in the chicken retrovirus Rous sarcoma virus (RSV) (96,97). C-Src is expressed throughout vertebrate cells and plays a significant role in the development of the nervous system. It has been shown that C-Src modulates axon guidance, such as through the phosphorylation of tubulin, as well as being linked to neurite outgrowth through phosphorylation of cytoskeletal substrates. These include cortactin, an F-actin binding protein that regulates cell movement (98) and Scar1, which binds the Arp2/3 complex to regulate actin polymerisation (99). Furthermore, C-Src regulates synaptic transmission through phosphorylation of NR2A, a subunit of the N-methyl-D-aspartate (NMDA) receptor (100), and the GluR2 subunit of the α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor (101). Both NMDA and AMPA receptors are excitatory glutamate receptors located on the post-synaptic membrane. AMPA receptors mediate a fast and short-lived response, whilst NMDA receptors mediate a slow and long-lived response. Consequently, C-Src is a multifunctional protein important for the developing and mature brain, although complete knockout of C-Src is not lethal (102). This is likely due to redundancy with other SFK members.

N1-Src was later discovered due to the presence of a slightly higher molecular weight band on an SDS-PAGE gel, which only occurred when analysing Src protein from neuronal tissue (103). Through the use of a cDNA library, it was found that this protein contained an additional 18 base pairs between exons 3 and 4 of C-Src. This translated to an additional six amino acids in the SH3 domain (RKVDVR) (104). In 1990, another splice variant was detected using PCR of human cDNAs. This isoform, later termed N2-Src, was found to contain the same 18 base pair insert of N1-Src, but with an additional 33 base pairs. This resulted in a total of 17 amino acids inserted between exons 3 and 4 (RKVDVSQTWFTFRWLQR) (105). These microexon insertions in the SH3 domain likely alter the structure of the domain, which suggests both N1-Src and N2-Src may have a different set of substrates compared to C-Src. Investigation of this via mass spectrometry and NMR found that N1-Src substrates did differ from the C-Src substrate set, as well as some substrates binding both proteins, but with differing affinity (84,106). Since

both N1-Src and N2-Src are expressed only in neurons, it suggests that these proteins may function as a more tightly regulated neuronal C-Src (107). Furthermore, the structural differences between N1-Src and C-Src are also likely the reason for the increased autophosphorylation of Y416 observed with N1-Src (103,106,108). It is thought that the additional amino acids weaken the interaction between the SH3 and SH2-kinase linker, improving access to Y416, allowing higher kinase activity (106).

N1-Src has recently been studied in relation to its role in neuronal development and differentiation (109). For example, in *Xenopus tropicalis*, the N1-Src expression profile during embryogenesis was investigated. It was found that N1-Src has high expression during primary and secondary neurogenesis, which compared to the lower and constant levels of C-Src throughout the various phases (109). Further analysis using *in situ* hybridisation illustrated enriched expression of N1-Src at the neural plate, a developmental structural key to the formation of the nervous system. This suggested that N1-Src may be involved in neurogenesis, a process through which neurons are formed from neural precursors. Consequently, splice site-specific antisense morpholino oligos were used to specifically knock-down N1-Src, preventing differentiation of neurons and leading to an abnormal dart response upon application of a touch stimulus (109).

N1-Src is also implicated in neurite outgrowth (108). It was found that overexpression of N1-Src, but not C-Src, in COS7 cells was found to increase the number of neurite-like processes per cell (108). The effect of N1-Src inhibition was then investigated in cerebellar granule neurons, where expression of PDN1, an N1-Src-specific inhibitor reduced neurite length compared to controls. These results were also replicated in hippocampal neurons, where N1-Src was knocked-down using N1-Src-specific shRNA, supporting a role of N1-Src in the regulation of neurite outgrowth (108).

1.8.2 TMEM106B and phosphorylation

Since identifying TMEM106B as a substrate of C-Src and N1-Src (84), no other studies have investigated phosphorylation as a mechanism underlying TMEM106B regulation. Crystal structures of TMEM106B amyloid fibrils suggest T185 is potentially phosphorylated, which may alter the glycosylation state of N183 (36), but the potential sites of phosphorylation for the cytosolic N-terminus are unknown. Analysis of the amino acid sequence of the N-terminus

reveals three tyrosine residues: Y18, Y50 and Y84, with each having the potential to be phosphorylated by Src. PhosphoSitePlus is a public repository of phosphoproteomics data and shows that Y50 has 383 high-throughput hits, Y18 has 261 and Y84 has zero, suggesting the TMEM106B N-terminus is phosphorylated on Y18 and Y50 in cells. S33 also has 16 high-throughput hits, suggesting this could be a potential phosphorylation site for a serine/threonine kinase. How phosphorylation might affect TMEM106B function is unclear. One possibility is that phosphorylation induces structural changes in TMEM106B, altering its function at the lysosomal membrane. Alternatively, phosphorylation could promote or inhibit interactions with a wide range of binding partners, such as MAP6 and AP2M1. Interestingly, both putative AP2M1 YXXΦ motifs are present at the potential Y18 and Y50 phosphosites, suggesting phosphorylation of TMEM106B may mediate AP2M1 binding (Figure 6).



Figure 6 – TMEM106B N-terminus contains three tyrosine residues: Y18, Y50 and Y84. Y18 and Y50 are most likely to be phosphorylated by the tyrosine kinase Src, according to the phosphoproteomics database PhosphoSitePlus. Putative AP2M1 binding motifs are present at both Y18 and Y50, suggesting phosphorylation of one or both residues may regulate the binding of AP2M1 to TMEM106B.

1.9 Aims

Although there is significantly more understanding of TMEM106B since the initial GWAS that highlighted its link to FTLD, there are still many questions regarding its function and particularly how TMEM106B is regulated. However, PhosphoSitePlus data demonstrate Y18/Y50 phosphorylation of TMEM106B is highly prevalent *in vivo*, and together with the results from the high-throughput C-Src and N1-Src phosphoproteomics screen, suggests TMEM106B undergoes tyrosine phosphorylation by C-Src and N1-Src on Y18 and/or Y50. These putative phosphosites also lie within AP2M1 sorting motifs, therefore it was hypothesised that phosphorylation of TMEM106B by C-Src and N1-Src regulates TMEM106B trafficking via AP2M1 binding, which might impact the function of TMEM106B in the healthy and diseased brain. Although both C-Src and N1-Src are predicted to phosphorylate TMEM106B, this study focused on N1-Src, due to its constitutive activation in neurons making it more likely to have a greater impact on TMEM106B function.

To achieve this, the study was subdivided into the following objectives:

- Confirm phosphorylation by N1-Src and validate phosphorylation sites using *in vitro* kinase assays
- Conduct *in vitro* binding assays with the known interactor AP2M1
- Determine whether localisation of TMEM106B is affected by phosphorylation in mammalian cell lines

This study largely utilised biochemical techniques, using recombinantly produced N-terminal TMEM106B, N1-Src and AP2M1 protein. N1-Src phosphorylation and AP2M1 binding was confirmed through SDS-PAGE and western blotting analysis. SEC-MALS was also conducted to determine if phosphorylation affects dimerisation of TMEM106B. Lastly, a doxycycline-inducible N1-Src HeLa cell line was employed to determine how expression of N1-Src affected subcellular localisation of TMEM106B, using fluorescent microscopy.

2. Materials and methods

2.1 Materials

2.1.1 *Molecular biology reagents*

TMEM106B(1-96) Gblocks™ and PCR primers were synthesised by Integrated DNA technologies. pCOLI_A_C_GST (#160245, a gift from John Weir (110)) and pHH0103_AP2M1_Adap-comp-sub (#109855, a gift from Sachdev Sidhu (unpublished)) plasmids for TMEM106B and AP2M1 respectively were purchased from Addgene. pEGFP-N1 plasmids containing either full length TMEM106B-WT or TMEM106B-Y50F, both of which were the T185 isoform, were synthesised and purchased from General Biosystems. DNA Hyperladder I and SYBR®Safe DNA stain was obtained from Bionline and Invitrogen respectively. The QIAquick gel extraction kit is from Qiagen and the miniprep/midiprep kits are from Machery-Nagel. BamHI was purchased from thermoscientific, and both HindIII and DNA ligase were from Promega. XL-10 Gold competent *E. coli* was obtained initially from Stratagene.

2.1.2 *Biochemistry reagents*

BL21 DE3 competent *E. coli* were purchased from Sigma-Aldrich. Lysozyme, protease inhibitor cocktail, Ni²⁺-agarose beads, Pur-A-lyzer (MWCO 3.5 kDa), PVDF, and ATP were all obtained from Sigma-Aldrich. Glutathione agarose beads were from Cytivia and Spin-X columns (0.45 µM, 2.0 ml tubes) were from Corning Inc. Bovine serum albumin (BSA) was purchased from Fisher Scientific and Precision Plus all blue protein standard (161-0373) was from Bio-Rad. Laemmli buffer was obtained from Bioworld and enhanced chemiluminescence (ECL) reagent from Amersham.

Antibodies: mouse anti-phosphotyrosine (clone PY20, BD Bioscience, 610000), mouse anti-Calnexin (BD Bioscience, 610523), mouse anti-TMEM106B (Proteintech, 60333-1-Ig), mouse anti-FLAG M2 (Sigma, F1804), mouse anti-LAMP1 (DSHB, H4A3-S), mouse anti-GAPDH (Sigma, MAB374), anti-mouse HRP conjugate (Sigma, A4416), anti-mouse Alexa Fluor™-594 (Invitrogen, A11020), rabbit anti-AP2M1 (Abcam, ab75995), rabbit anti-Myc (Proteintech, 16286-1-AP) and anti-rabbit HRP conjugate (Sigma, A6154).

2.1.3 Cell biology reagents

Tetracycline-inducible Flp-in HeLa cells, expressing N1-Src (N1-Src HeLas), were produced in the Evans Lab by Dr Philip Lewis (111). Dulbecco's Modified Eagle Medium (DMEM) was from Gibco (41966-029), hygromycin-B from TOKU-E, Polyjet reagent from SignaGen laboratories, trypsin/EDTA and penicillin-streptomycin (Penstrep) from Invitrogen, fetal bovine serum (FBS) from Hyclone (HYC85).

2.2 Molecular biology

2.2.1 Generation of TMEM106B(1-96)-GST plasmids

Double-stranded DNA templates (Gblocks™) containing N-terminal WT, Y50F or 3F TMEM106B (AA 1-96), were synthesised by Integrated DNA Technologies (sequences stated in Table 1). Each Gblock™ was amplified via PCR using primers stated in Table 1 and purified using QIAquick gel extraction kit. Gblocks™ and 9 pCOLI_A_C_GST plasmid were subjected to a restriction enzyme digest using BamHI and HindIII at 37 °C for 1 h. Digest products were separated from the reaction mixture using agarose gel electrophoresis and purified according to the manufactures protocol (QIAquick gel extraction kit). Successful purification was confirmed by a second agarose gel electrophoresis, with concentrations of digest products determined by comparison to the BIOLINE Hyperladder I. 25 ng of each Gblock™ was ligated into 100 ng of plasmid (3:1 molar ratio of insert to vector), using 0.3 units/μl ligase and incubated at 4 °C overnight.

Table 1 - DNA sequences of -WT, -3F and TMEM106B(1-96)-Y50F GBLOCKS™ and the forward and reverse primers for PCR amplification.

Oligonucleotide	DNA Sequence
TMEM106B(1-96)-WT	ATGGGTAAATCCCTGTCCCACCTGCCACTGCACTCCTCTAAAGAG GACGCATATGACGGTGTGACCTCTGAAAACATGCGTAACGGTCTG GTGAATAGCGAAGTACACAACGAAGACGGTCGTAACGGCGACGT ATCCCAGTTTCCGTACGTTGAGTTCACCGGTCGTGATTCCGTCACC TGTCGACATGTCAGGGTACCGGCCGTATCCCGCGTGGCCAGGAA AACCAGCTGGTAGCGCTGATCCCGTACTCTGATCAGCGTCTGCGT CCACGCCGTACTAAGCTG
TMEM106B(1-96)-3F	ATGGGTAAATCCCTGTCCCACCTGCCACTGCACTCCTCTAAAGAGG ACGCATTCGACGGTGTGACCTCTGAAAACATGCGTAACGGTCTGGT GAATAGCGAAGTACACAACGAAGACGGTCGTAACGGCGACGTATC CCAGTTTCCGTTTCGTTGAGTTCACCGGTCGTGATTCCGTCACCTGTC CGACATGTCAGGGTACCGGCCGTATCCCGCGTGGCCAGGAAAACC AGCTGGTAGCGCTGATCCCGTTCTCTGATCAGCGTCTGCGTCCACG CCGTACTAAGCTG
TMEM106B(1-96)-Y50F	ATGGGTAAATCCCTGTCCCACCTGCCACTGCACTCCTCTAAAGAGG ACGCATATGACGGTGTGACCTCTGAAAACATGCGTAACGGTCTGGT GAATAGCGAAGTACACAACGAAGACGGTCGTAACGGCGACGTATC CCAGTTTCCGTTTCGTTGAGTTCACCGGTCGTGATTCCGTCACCTGTC CGACATGTCAGGGTACCGGCCGTATCCCGCGTGGCCAGGAAAACCA GCTGGTAGCGCTGATCCCGTACTCTGATCAGCGTCTGCGTCCACGCC GTACTAAGCTG
Forward primer	5'-GGTGGTGGATCCATGGGTAAATCCCTGTCCCACCTG-3'
Reverse primer	5'-ACCACCAAGCTTCAGCTTAGTACGGCGTGGACG-3'

2.2.2 Generation of His-Δ80-N1-Src plasmid

The pCDF-Duet-1-His-Δ80-N1-Src was prepared by Hannah Walker. Briefly, the His-Δ80-N1-Src cDNA sequence was subcloned from pGEX4T-1-PTP1B-His-Δ80-N1-Src (33) (Table 2) into the first multiple cloning site of pCDF-Duet-1 with 5' BamHI and 3' Sall restriction sites. The PTP1B cDNA sequence from pGEX4T-1-PTP1B-His-Δ80-N1-Src was cloned into the second multiple cloning site of pCDF-Duet-1-N1-Src plasmid, using 5' NdeI and 3' KpnI restriction sites.

Table 2 – Forward and reverse primers for PCR amplification of His- Δ 80-N1-Src and PTP1B for creation of His- Δ 80-N1-Src plasmid.

Oligonucleotide	DNA Sequence
N1-Src forward primer	5'-GGTCATGGATCCAGGTGGGGTGACTACCTTTGTG-3'
N1-Src reverse primer	5'-ACCACCGTCGACCTATAGTTCTCCCCGGGCTG-3'
PTP1B forward primer	5'-GGTCATCATATGATGGAGATGGAAAAGGAGTTC-3'
PTP1B reverse primer	5'-ACCACCGGTACCCCCATTGTGTGGCTCCAGGAT-3'

2.2.3 Agarose gel electrophoresis

0.7-2 % (w/v) agarose gels were prepared by adding agarose to 90 ml TAE buffer (20 mM acetic acid, 1 mM EDTA, 40 mM Tris) and heated in the microwave. After cooling, SYBR[®]Safe DNA stain was added (1:18,000) and solution poured into the cassette. Once set, DNA samples were loaded using Orange G buffer and DNA was resolved at 70 V in TAE buffer. DNA bands were visualised using blue safelight and imaged using Bio-Rad gel doc EZ imager.

2.2.4 Bacterial transformation

5 μ l of ligation product or 50-100 ng plasmid DNA was added to 50 μ l *E. coli* XL-10 gold cells and incubated on ice for 30 min. *E. coli* were heat shocked at 42 °C for 45 s and incubated on ice for a further 2 min. 400 μ l of lysogeny broth (LB) (1 % NaCl, 1 % tryptone, 0.5 % yeast extract) was added to the bacteria and subsequently incubated at 37 °C for 1 h. 50-200 μ l of transformed bacteria were plated onto agar plates containing the required antibiotic (100 μ g/ml ampicillin, 50 μ g/ml spectinomycin or 50 μ g/ml kanamycin) and incubated overnight at 37 °C.

2.2.5 Purification of plasmid DNA from *E. coli* XL-10 gold

10 ml (mini) or 100 ml (midi) of LB containing the relevant antibiotic was inoculated with a single colony of the relevant cells and incubated overnight at 37 °C, 225 rpm. Bacteria were pelleted via centrifugation at 2880 RCF for 10 min. The Machery miniprep or midiprep kit was

used to purify each plasmid according to the manufacturer's instructions and eluted using 50 μ l (mini) or 300 μ l (midi) elution buffer. DNA concentration was determined using the NanoDrop spectrophotometer by measuring absorbance at A260. Plasmid sequences were confirmed via sequencing (Eurofins genomics) and stored at -20 °C.

2.3 Protein biochemistry

2.3.1 Protein induction (AP2M1, His- Δ 80-N1-Src and TMEM106B(1-96))

Competent *E. coli* BL21 or BL21 DE3 cells were chemically transformed with the relevant plasmid (as stated in section 2.2.4). 50 mL of LB, containing required antibiotic (100 μ g/ml ampicillin or 50 μ g/ml spectinomycin) was inoculated with a single colony of transformed bacteria and incubated overnight at 37 °C, 200 RPM. Overnight culture, centrifuged at 2800 RCF for 10 min, was added to 1 L LB (containing 100 μ g/ml ampicillin, 50 μ g/ml spectinomycin or 12.5 μ g/ml tetracycline), to produce a starting OD₆₀₀ equal to 0.05. The culture was incubated at 37 °C (225 RPM), until an OD₆₀₀ of 0.6-1.0 was achieved, at which point isopropyl β -d-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM or 0.5 mM if incubating overnight (N1-Src only). Each culture was further incubated either at 37 °C (200 RPM) for 3 h, or overnight at 20 °C. Centrifugation at 5000 RCF for 15 min was used to obtain a cell pellet, and this was stored at -70 °C until required.

2.3.2 Protein purification

Bacterial pellets containing GST-tagged proteins (TMEM106B(1-96)) were resuspended in phosphate-buffered saline (PBS; 142 mM NaCl, 2 mM KCl, 8 mM Na₂HPO₄, 1.5 mM NaH₂PO₄) containing 1 mM phenylmethylsulfonyl fluoride (PMSF), 13.3 mg/l lysozyme and protease inhibitor cocktail. Each suspension was incubated on ice for 30 min, followed by addition of Triton X-100 (to 1%) and 4.5 mM dithiothreitol (DTT). Sonification in 30 s intervals for a total of 5 min 30 s, was used to lyse cells. Cell lysates were clarified via centrifugation at 17,000 RCF for 30 min (4 °C). Supernatant was added to glutathione agarose beads and incubated at 4 °C for 1 h with agitation. Beads were pelleted at 500 RCF for 10 min (4 °C), and supernatant removed. Beads were washed five times in at least 10 vols of PBS via centrifugation at 500 RCF for 5 min (4 °C). This was followed by a wash with 1.2 M NaCl in PBS and then two additional washes in PBS. Beads were either stored long term in 50% glycerol in PBS at -20 °C

for use in binding assays or resuspended in PBS to generate a 50% slurry for glutathione elution.

For His-tagged proteins (AP2M1 and His- Δ 80-N1-Src), purification was as above, except bacterial pellets were resuspended in breaking buffer (100 mM HEPES pH 7.0, 500 mM KCl and 2 mM β -mercaptoethanol) and loaded onto Ni²⁺-agarose beads. All washes occurred in buffer A containing 20 mM HEPES (pH 7.0), 200 mM KCl, 50 mM imidazole, 2 mM β -mercaptoethanol and 10% (v/v) glycerol. His- Δ 80-N1-Src was eluted and dialysed, whilst AP2M1 was incubated with HRV 3C-protease for His-GST cleavage.

For elution of tagged protein, the 50% slurry was transferred to a spin-X column and pulsed at 15,700 RCF for 15 s to remove PBS or buffer A. Beads were incubated with elution buffer (GST-tagged proteins: 100 mM Tris (pH 8), 20 mM glutathione and 100 mM NaCl; His-tagged proteins: buffer A containing 300 mM imidazole) at room temperature with agitation for 20 min. Beads were pulsed for 15 s at 15,700 RCF to elute protein. This was repeated three times to give three elutions. Suitable elutions were either dialysed (His- Δ 80-N1-Src only) or aliquoted.

For dialysis of His- Δ 80-N1-Src, the required elutions were transferred into a midi Pur-a-lyzer™ and dialysed in buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 0.05% NP-40. After 1 h, the Pur-a-lyzer™ was swapped into fresh buffer and dialysed overnight. Protein was aliquoted in the same buffer, but with the addition of 1 mg/ml BSA and 10% glycerol (kinase storage buffer).

For cleavage of His-GST tag (AP2M1 only), the Ni²⁺-agarose beads were washed three times with cleavage buffer (50 mM Tris-HCl, pH 7.0, 150 mM NaCl, 1 mM EDTA and 1 mM DTT) and incubated overnight at 4 °C with Acro HRV 3C-protease (1 unit cleaves 100 μ g of protein), with agitation. The beads were transferred to a Spin-X column and centrifuged at 15,700 RCF for 15 s to elute cleaved protein.

Protein concentration was determined by comparison to BSA standards either through a Bradford assay or an SDS-PAGE gel quantified using densitometry. All proteins were aliquoted accordingly and frozen at -70 °C.

2.3.3 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The resolving gel was prepared using 0.1 % (w/v) SDS, 375 mM Tris pH 8.8, 10-12.5% (v/v) acrylamide, 0.05% (w/v) ammonium persulphate (APS) and 0.1% tetramethylethylenediamine (TEMED) (v/v). This was allowed to set, before adding the stacking gel (0.1 % SDS, 125 mM Tris pH 6.8, 4 % acrylamide, 0.05 % APS and 0.1 % TEMED). Protein samples were prepared by addition of Laemmli buffer and boiled at 90 °C for 10 min. These were loaded onto the gel, and resolved in running buffer (25 mM Tris, 192 mM glycine and 0.1% SDS) at 150 V for 15 min, followed by 200 V for 45 min. Gels were either transferred for western blotting or stained for 30 min using Coomassie (0.3% (w/v) Brilliant Blue R, 54% (v/v) methanol and 9% (v/v) acetic acid). After staining, gels were destained (40% (v/v) methanol and 10% (v/v) acetic acid) to visualise protein bands for subsequent quantification via densitometry.

2.3.4 Western blotting

For immunostaining, proteins were transferred onto a polyvinylidene fluoride (PVDF) membrane. The membrane was first activated by submerging in 100% methanol for 1 min, followed by 1 min in dH₂O and then in transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol). The activated membrane and gel were assembled onto a cassette and transferred at 66 V for 1 h in transfer buffer. The membrane was blocked in blocking buffer (5% (w/v) skimmed milk powder in PBS or tris-buffered saline (TBS; 20 mM Tris, pH 7.5 and 137 mM NaCl) for 1 h and incubated overnight in primary antibody (Table 3) at 4 °C with agitation. The membrane was washed (0.5% (v/v) Tween-20 in PBS or TBS) 3x for 10 min each and incubated for 1 h at room temperature in secondary antibody diluted in 5% (w/v) skimmed milk powder in PBS or TBS with 0.5% (v/v) Tween-20 (Table 3). The membrane was washed an additional 3x for 10 min each in PBS or TBS containing 0.5% Tween-20. ECL reagent was added directly to the membrane and immunoreactivity visualised using iBright imaging systems.

Table 3 - Concentrations of primary and secondary antibodies used in western blotting.

Primary antibody		Secondary antibody		
Antibody	Concentration	Antibody	Concentration	PBS/TBS
PY20	1:1000	Anti-mouse HRP	1:5000	TBS
TMEM106B	1:1000	Anti-mouse HRP	1:5000	PBS
FLAG	1:1000	Anti-mouse HRP	1:5000	PBS
GAPDH	1:10,000	Anti-mouse HRP	1:5000	TBS
AP2M1	1:10,000	Anti-rabbit HRP	1:5000	PBS
Myc	1:1000	Anti-Rabbit HRP	1:5000	TBS

2.3.5 In vitro N1-Src kinase assays

For dose-response phosphorylation of YA (0.2 $\mu\text{g}/\mu\text{l}$ final concentration) by N1-Src, reactions were prepared using 100 mM Tris-HCl (pH 7.5), 10 mM MgCl_2 , 0.25 mM vanadate and various concentrations of N1-Src stated in Figure 10A. Mixtures were prewarmed to 30 °C and reaction initiated by addition of pre-warmed ATP (0.5 mM final concentration). Samples were incubated for 3 h at 30 °C and reaction quenched by addition of Laemmli buffer. All samples were run on 12.5% SDS-PAGE gel and transferred for western blotting. Phosphorylation of substrates was determined by addition of mouse anti-phosphotyrosine (PY20) and total protein content was determined through Coomassie staining of a separate gel.

For phosphorylation of TMEM106B(1-96)-WT/3F/Y50F (0.2 $\mu\text{g}/\mu\text{l}$ final concentration), reactions were prepared as above using N1-Src with a final concentration of 1.2 ng/ μl . Control samples used 0.2 $\mu\text{g}/\mu\text{l}$ of GST or YA peptide. Samples were taken at time points stated in the Figure 10, and reaction quenched by addition of Laemmli buffer. Phosphorylation was determined through immunoblotting with PY20 and quantified via densitometry.

2.3.6 AP2M1 binding assays

For preparation of phosphorylated TMEM106B(1-96)-WT, an N1-Src kinase assay was conducted directly prior to the binding assay. For this, glutathione beads containing bound TMEM106B(1-96)-WT, TMEM106B(1-96)-Y50F, or GST only, stored as aliquots in 50% glycerol,

were washed 5x in 100 mM Tris-HCl (pH 7.5) at 13,000 RCF for 15 seconds each to remove glycerol. Reactions were prepared as in Section 2.3.5, with final concentrations of 0.2 µg/µl TMEM106B(1-96) or GST. 1.2 ng/µl N1-Src was used for one aliquot of TMEM106B(1-96)-WT, with a corresponding volume of kinase storage buffer used for all other proteins. Reactions were initiated by addition of 0.5 mM ATP and incubated at 37 °C for 3 h. Reactions were quenched by washing 5x in binding buffer (150 mM potassium acetate, 20 mM HEPES, pH 7.4, 1 mM MgCl₂, 0.5 mM DTT and 0.05% Tween-20) to remove ATP.

Following this, each set of beads were diluted in Sepharose beads to achieve 1 µM concentration. These were incubated with varying concentrations (0-1 µM) of AP2M1 in binding buffer at 4 °C for 1 h with end-over-end rotation. Solutions were centrifuged at 13,000 RCF for 1 min and then transferred to a Spin-X column. Unbound protein was eluted by centrifuging at 13,000 RCF for 15 s. Beads were washed 5x in binding buffer by centrifugation at 13,000 RCF for 15 s. Bound protein was eluted from the beads by addition of 2x Laemmli buffer. Samples from the bound elution were analysed via immunoblotting with anti-AP2M1 and quantified via densitometry in ImageJ.

In a second set of experiments, AP2M1 binding to TMEM106B(1-96)-3F was tested. For this, the binding assay was conducted as stated above, but TMEM106B(1-96)-WT and TMEM106B(1-96)-3F beads were not subjected to the kinase assay.

2.3.7 Size-exclusion chromatography with multi-angle light scattering (SEC-MALS)

Directly prior to SEC-MALS, a kinase assay was conducted to produce phosphorylated TMEM106B(1-96)-WT. Reactions were prepared as stated in section 2.3.5, but with 7.4 ng/µl N1-Src and 1.23 µg/µl TMEM106B(1-96)-WT. Control TMEM106B(1-96)-WT was subjected to the kinase assay, but in the absence of N1-Src. Native TMEM106B(1-96)-WT was thawed on ice and used directly for SEC-MALS, without the kinase assay. Final concentrations of each sample are as follows: 2.5 mg/ml native TMEM106B(1-96)-WT, 1.1 mg/ml phosphorylated TMEM106B(1-96)-WT and 1.1 mg/ml control TMEM106B(1-96)-WT.

The following method for SEC-MALS analysis was conducted and provided by Dr Andrew Leech (Molecular interactions, Bioscience Technology facility, University of York):

Experiments were conducted on a system comprising a Wyatt HELEOS-II multi-angle light scattering detector and a Wyatt rEX refractive index detector linked to a Shimadzu HPLC system (SPD-20A UV detector, LC20-AD isocratic pump system, DGU-20A3 degasser and SIL-20A autosampler). Work was conducted at room temperature ($20 \pm 2^\circ\text{C}$). Solvent (PBS, Gibco) was $0.2 \mu\text{m}$ filtered before use and a further $0.1 \mu\text{m}$ filter was present in the flow path. The column was equilibrated with at least 2 column volumes of solvent before use and flow was continued at the working flow rate until baselines for UV, light scattering and refractive index detectors were all stable.

Sample injection volume was $100 \mu\text{L}$ except where stated; Shimadzu LabSolutions software was used to control the HPLC and Astra 7 software for the HELEOS-II and rEX detectors. The Astra data collection was 1 minute shorter than the LC solutions run to maintain synchronisation. Blank buffer injections were used as appropriate to check for carry-over between sample runs. Data were analysed using the Astra 7 software. MWs were estimated using the Zimm fit method with degree 1. A value of 0.182 was used for protein refractive index increment (dn/dc).

The run details are summarised below (Table 4):

Table 4 – Run details used for SEC-MALS analysis of TMEM106B(1-96)-WT samples.

Column	Superdex S200-2, 10/300 GL #10245604 (G.E. Healthcare)
Flow rate	0.5 ml/min
Temperature	Ambient ($\sim 20^\circ\text{C}$)
Duration	60 min (59 min data collection)
UV detection	280 nm
Injection volume	$100 \mu\text{l}$
Injection mass	$\sim 100 \mu\text{g}$ (nominal)

2.4 Cell biology

2.4.1 Culture and passage of N1-Src HeLas

Cells were cultured in T75 cm^2 flasks containing DMEM (with 4.5 g/L D-glucose, L-glutamine, and pyruvate), supplemented with 10 % FBS and 1 % Penstep. Cells were incubated at 37°C

with 5 % CO₂ in a humidified atmosphere and passaged when approaching 100 % confluency. For this, media was removed, and cells washed 1x in sterile PBS. 1 ml prewarmed trypsin/EDTA was applied and incubated for 3 min at 37 °C. Cells were dislodged and resuspended in fresh media. Following centrifugation (720 RCF, 4 °C), the cell pellet was resuspended in 1 ml fresh media and split into new T75 cm² as required. Hygromycin B (0.5 mg/ml) was used to maintain selectivity for N1-Src.

2.4.2 Transient transfection of N1-Src HeLas

The total number of cells was estimated using a haemocytometer, and cells were plated onto either 6-well or 24-well plates, with densities of 200,000 and 35,000 respectively. 6-well plates were used for the generation of cell lysates and analysed via western blotting. 24-well plates, containing 13 mm coverslips, were used for fluorescent microscopy. Plates were incubated for 24 h prior to transfection.

Directly before transfection, cells were incubated in fresh media for 1 h at 37 °C (5 % CO₂), with 1 µg/ml doxycycline added to relevant wells to induce expression of N1-Src. Following this, cells were transfected with 2 µg midi-prepped DNA for 6-well plates and 0.5 µg DNA for 24-well plates, using PolyJet reagent (1 µg: 3 µl DNA:Polyjet ratio). DNA and PolyJet were prepared individually using serum-free DMEM, and then mixed and incubated for 10 min at room temperature. The DNA/PolyJet mixture was added to each well and incubated for 48 h at 37 °C, 5 % CO₂.

2.4.3 Lysis of N1-Src HeLas for western blotting

48 h post-transfection, 6-well plates were removed from the incubator and immediately placed on ice. Media was removed and cells were washed 2x in ice-cold PBS. 150 µl 2x Laemmli was added to each well to lyse cells. Cells were scraped into Eppendorf tubes and analysed via western blotting.

2.4.4 Immunocytochemistry and fluorescent microscopy

48 h post-transfection, 24-well plates were removed from the incubator and immediately placed on ice. Media was removed and cells washed 3x in ice-cold PBS. Cells were fixed in 4 % paraformaldehyde for 20 min at room temperature. Cells were washed 3x in PBS and then incubated in PBS containing 0.1 % Triton-X-100 (permeabilises) and 1 % BSA (blocks) for 30

min at room temperature. Buffer was removed and cells incubated in primary antibody (PBS containing 1 % BSA, Table 5) for 2 h at 37 °C in a humidified atmosphere. Cells were washed 3x in PBS and incubated in secondary antibody (PBS containing 1 % BSA, Table 5) for 1 h in the dark. Cells were washed 3x in PBS, followed by 1x in 50 mM ammonium chloride to de-acidify lysosomes. Coverslips were removed from wells and airdried before mounting onto microscope slides using Mowial containing 4',6-diamidino-2-phenylindole (DAPI). Slides were imaged using a 100x objective (oil-emersion) on a Leica DM IL fluorescent microscope. All image analysis was conducted using ImageJ.

Table 5 – Concentrations of primary and secondary antibodies used for immunocytochemistry.

Primary antibody		Secondary antibody	
Antibody	Concentration	Antibody	Concentration
FLAG-M2	1:1000	Anti-mouse Alexa Fluor 594	1:500
LAMP1	1:20	Anti-mouse Alexa Fluor 594	1:500
Calnexin	1:500	Anti-mouse Alexa Fluor 594	1:500

2.5 Data analysis

Densitometry analysis was conducted using ImageJ for relevant SDS-PAGE gels and Western blots. For the *in vitro* kinase assays and binding assays, bands for the western blot were quantified and normalised to the bands in the SDS-PAGE. For N1-Src HeLa cell lysate samples, bands for each antibody were normalised to GAPDH levels for each condition. These normalised values were used for statistical analysis, which was conducted for all triplicate data sets in GraphPad Prism 10. One-way ANOVA, followed by Tukey post-hoc test was used for the mutant kinase assay. Two-way ANOVA with Tukey post-hoc was used for quantification of N1-Src and TMEM106B from cell lysates. Pairwise t-tests were performed for the AP2M1 binding assay. All graphs were created in RStudio.

3. Results

3.1 Y18 and Y50 are potential phosphosites for TMEM106B

The potential sites of phosphorylation of TMEM106B are currently unknown. To determine these sites, analysis of the N-terminus of TMEM106B (TMEM106B(1-96)) was conducted and revealed three tyrosine residues that could potentially be phosphorylated by the tyrosine kinase N1-Src; Y18, Y50 and Y84, as depicted in Figure 7A. Further analysis on PhosphoSite Plus (112), a repository of phosphoproteomics data, shows Y18 has 261 high-throughput (HTP) hits, Y50 has 383 HTP hits, whilst Y84 has zero HTP hits (Figure 7B). This suggests phosphorylation of TMEM106B(1-96) is more likely to occur on either Y18 or Y50. However, in a high-throughput phosphoproteomics screen with recombinant N1-Src, Y50 was the only residue identified to be phosphorylated (84). For this assay, cells were treated with 5 mM 5'-4-Fluorosulfonylbenzoyl adenosine (FSBA) to irreversibly inhibit endogenous kinases, and then lysed to generate cell lysates. These cell lysates were incubated with either C-Src, N1-Src or neither (FSBA only control). The resulting phosphopeptides were enriched by phosphotyrosine immunoprecipitation and subjected to tandem mass-spectrometry to identify substrates and the specific phosphosites. From this, TMEM106B-Y50 was shown to be phosphorylated by both C-Src and N1-Src, with quantitative area scores (which is proportional to the amount of peptide present) similar to that of EWSR1-Y278 (Figure 7C), a previously identified C-Src substrate (113).

To determine if either Y18 or Y50 are conserved across species, sequence alignment of TMEM106B was conducted in ClustalOmega. Figure 7D illustrates that Y50 is evolutionarily conserved, supporting that this residue is likely important for mediating TMEM106B function, with phosphorylation acting as a regulatory mechanism. Y18 is conserved in chicken and mouse, but not *Xenopus tropicalis* or zebrafish, suggesting this may be required for higher-order function. Similarly, Y18 is not present in either paralogues TMEM106A or TMEM106C (Figure 7E), but Y50 is present in TMEM106C and Y84 is present in TMEM106A. These residues could potentially aid differentiation of binding partners, resulting in differing functions of the paralogues. Consequently, to confirm that TMEM106B(1-96) is phosphorylated directly by N1-Src and to identify the target residue(s), three bacterial expression plasmids were generated: TMEM106B(1-96)-WT, TMEM106B(1-96)-Y50F, and TMEM106B(1-96)-3F containing Y18F, Y50F and Y84F mutations.

independent experiments. Statistical significance determined by PEAKS Q significance model in PEAKS studio software. *** $p < 0.001$. D, E) Sequence alignment of TMEM106B N-terminus across different species (D) and the known paralogues TMEM106A and TMEM106C (E). Indicated residues correspond to that of human TMEM106B. * Indicates positions which are conserved across all sequences, whilst : indicates conservation between amino acids with strongly similar properties and . indicates conservation between amino acids with weakly similar properties according to the Gonnet PAM 250 matrix.

3.2 Production and purification of -WT, -3F and TMEM106B(1-96)-Y50F

Double stranded DNA templates (Gblocks™) containing sequences for TMEM106B(1-96)-WT, TMEM106B(1-96)-3F and TMEM106B(1-96)-Y50F were amplified via PCR, digested with restriction enzymes (BamHI and HindIII) and ligated into the 9 pCOLI_A_C_GST plasmid (110), containing a C-terminal GST tag (Figure 8A). A C terminal tag was chosen to reflect *in vivo* anchorage of TMEM106B to the lysosomal membrane. *E. coli* XL-10 gold were transformed with these plasmids, with three colonies selected for DNA purification of each genotype. Presence of the insert was confirmed via double restriction digest, followed by agarose gel electrophoresis. All plasmids contained the insert (Figure 8B) except one TMEM106B(1-96)-3F plasmid (3F 1) and one TMEM106B(1-96)-Y50F plasmid (Y50F 2).

E. coli BL21 DE3 were subsequently transformed with one plasmid prep for each genotype, for the recombinant overexpression of TMEM106B(1-96)-WT, TMEM106B(1-96)-3F and TMEM106B(1-96)-Y50F. Protein was purified from clarified cell lysate using glutathione agarose beads, and successfully yielded TMEM106B(1-96)-GST for each WT and mutant protein at the expected molecular weight. As illustrated in Figure 8C, monomeric TMEM106B is observed at ~36 kDa in each elution (E1-E4) and a putative dimeric form at ~72 kDa that aligns with previous reports (20,23), as well as GST at ~26 kDa. The concentrations of each TMEM106B protein, determined via Bradford assay, were 3.6 mg/ml for -WT, 3.2 mg/ml for -3F and 5.1 mg/ml for TMEM106B(1-96)-Y50F.

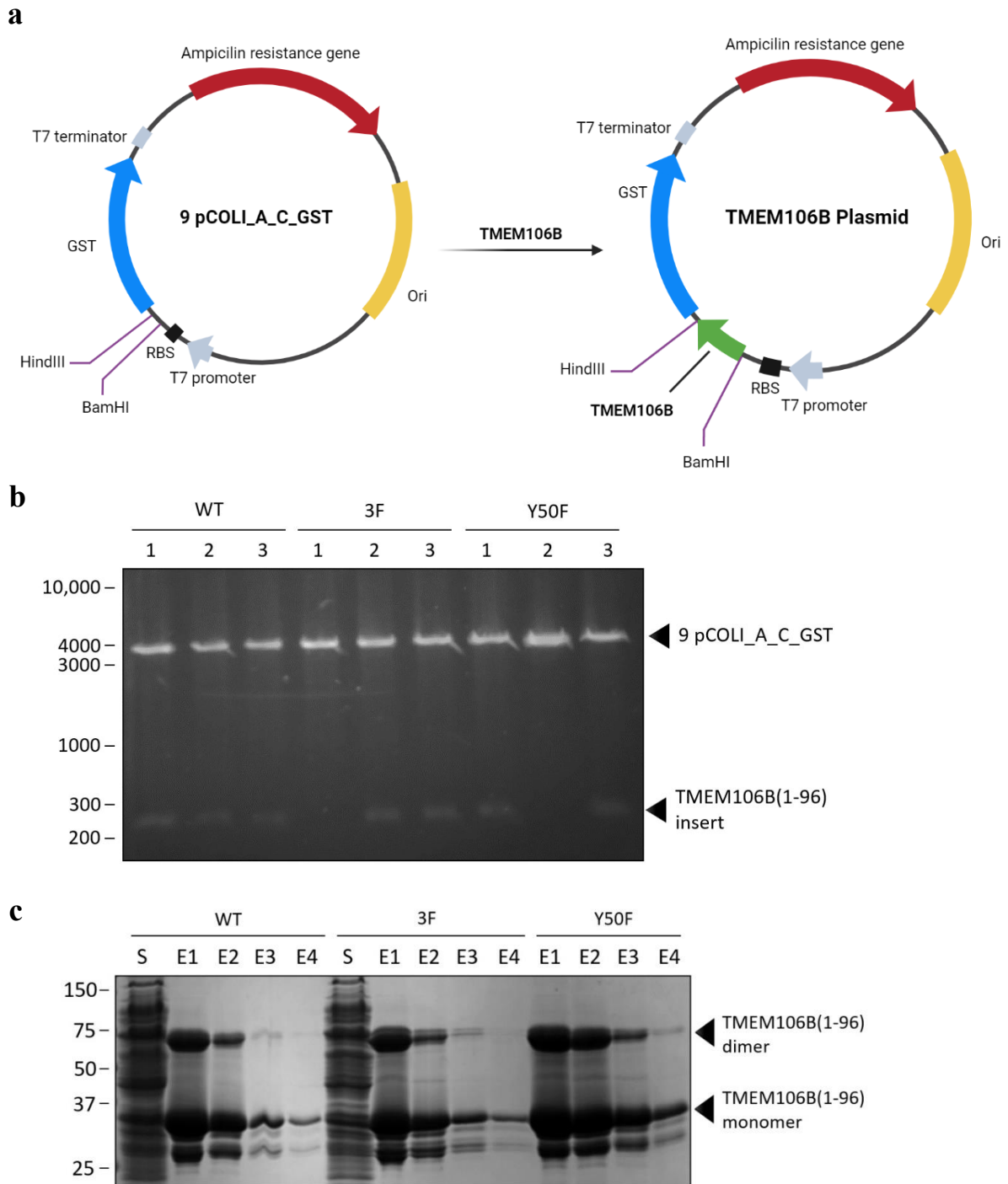


Figure 8 – Production and purification of TMEM106B(1-96)-GST. A) Schematic of N-terminal TMEM106B GBlock™ insertion into 9 pCOLI_A_C_GST plasmid, through restriction enzyme digest with BamHI and HindIII. This results in a TMEM106B(1-96) construct containing a C-terminal GST tag, controlled by a T7 promoter. B) Successful ligation of the insert into the plasmid was confirmed by a second double digest using BamHI and HindIII, with the digest products analysed using agarose gel electrophoresis. Three plasmid preps were tested for each genotype, with a band at 288 bp confirming the presence of the TMEM106B(1-96) insert. C) SDS-PAGE analysis of purified TMEM106B(1-96)-GST proteins. *E. coli* BL21 DE3 were transformed with TMEM106B(1-96), with induction occurring following addition of 1 mM IPTG and incubation at 37°C for 3 h. Glutathione agarose beads were used to purify -WT, -3F and TMEM106B(1-96)-Y50F proteins from the cell lysates. Four elutions (E1-E4) using elution buffer (100 mM Tris (pH 8), 20 mM glutathione and 100 mM NaCl) successfully yielded protein at the expected molecular weight. A sample of supernatant from after incubation with beads (S) was also run to determine the presence of any unbound protein.

3.3 Successful expression and purification of N1-Src

To prepare recombinant N1-Src kinase, a bacterial expression plasmid previously generated in the lab (106) was employed. The plasmid encoded a 120 kDa fusion protein that contained GST-tagged protein tyrosine phosphatase 1B (PTP1B) catalytic domain fused to His-tagged N1-Src that lacked the first 80 amino acids (Δ 80 N1-Src) (Figure 9A). These first 80 amino acids in N1-Src contain an N-terminal membrane localisation signal within the SH4 domain that is not required for activity *in vitro*, whilst PTP1B is essential in preventing the toxic build-up of tyrosine phosphorylated protein in *E. coli* since bacteria do not undergo this form of post-translational modification. At the C-terminus of PTP1B, there is a 3C protease cleavable linker, enabling release of His-tagged N1-Src (53 kDa) upon purification of the protein. However, despite this method producing suitable yields in previous reports (106,108,114), recently the majority of expressed protein was found to be insoluble and was consequently lost after clarification of cell lysate during the purification process (Figure 9B).

To overcome this issue, Hannah Walker in the Evans lab subcloned His-tagged Δ 80-N1-Src and PTP1B into the two ribosomal binding sites (RBS) respectively of pCDF-Duet-1 (Figure 9C). This enabled expression of the two separate proteins simultaneously, preventing the need for the 3C cleavage site to release His-tagged Δ 80 N1-Src. As shown in Figure 9D, an intense band at ~50 kDa for N1-Src was observed upon induction of *E. coli* BL21 DE3, as well as a band at 37 kDa for PTP1B, suggesting the proteins were being expressed as expected. Ni²⁺-agarose beads were used to purify His-tagged Δ 80 N1-Src, with 300 mM imidazole used to elute the protein. The majority of the protein was present in the first elution at the expected molecular weight of ~50 kDa (Figure 9E), which was then dialysed to remove imidazole and stored at -70 °C in kinase storage buffer. Approximately 20 μ g of protein was obtained using this method; a substantial improvement on the first method, suggesting independent expression of PTP1B and N1-Src improved protein solubility.

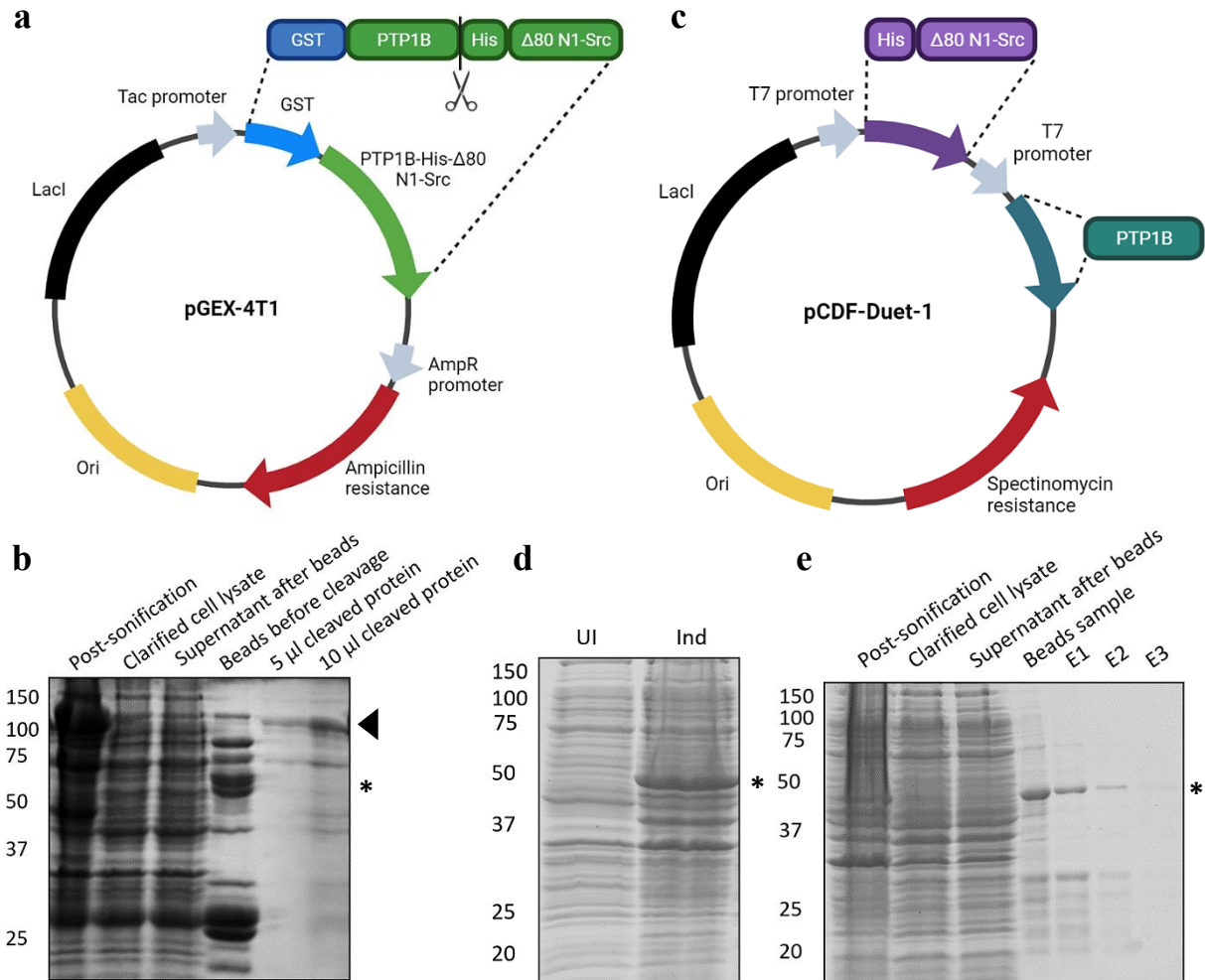


Figure 9 – Expression and purification of N1-Src protein in *E. coli*. A) Schematic of pGEX-4T1 plasmid containing GST-PTP1B-His-Δ80-N1-Src controlled by the tac promoter. Addition of IPTG induces this protein as a single construct. B) Purification of GST-PTP1B-His-Δ80-N1-Src (arrow) using glutathione agarose beads. GST-PTP1B-His-Δ80-N1-Src was cleaved at PTP1B C-terminus using 3C protease, resulting in release of His-Δ80-N1-Src upon elution. Expected mass of His-Δ80-N1-Src after cleavage is indicated with *. C) Schematic of pCDF-Duet-1 plasmid which contains two multiple cloning sites. His-Δ80-N1-Src was cloned into the first site (BamHI and Sall), with PTP1B cloned into the second site using restriction enzyme digest (NdeI and KpnI). Both constructs are controlled by the T7 promoter, with addition of IPTG inducing separate expression of the two constructs. D) Uninduced (UI) and induced (Ind) samples of His-Δ80-N1-Src (*) expressed in *E. coli* BL21 DE3 cell. Induction occurred following addition of 0.5 mM IPTG, with culture incubated overnight at 20 °C. E) Purification of His-Δ80-N1-Src (*) using Ni²⁺-agarose beads, followed by elution with 300 mM imidazole to give three elutions (E1-E3). E1 was dialysed overnight to remove imidazole.

3.4 N1-Src phosphorylates TMEM106B(1-96)

Having successfully obtained purified His- Δ 80-N1-Src, its activity was validated using a 3 h kinase assay with GST-tagged YA, a canonical Src substrate peptide (106). Immunoblotting with anti-PY20, which recognises phosphorylated tyrosine residues, demonstrated strong phosphorylation of YA with 15 nM N1-Src (Figure 10A), illustrating that the kinase is active. A further increase in N1-Src concentration did not increase YA phosphorylation (Figure 10B), suggesting maximal phosphorylation occurred with 15 nM N1-Src.

Following confirmation of active N1-Src, phosphorylation of TMEM106B(1-96)-WT was tested using 22.6 nM N1-Src in a time-course assay. Small amounts of phosphorylation were observed at 5 min in Figure 10C, confirming that TMEM106B(1-96)-WT is a substrate of N1-Src. A further increase in phosphorylation was observed at later time points (Figure 10D), with a 3 h incubation time selected for all future kinase assays. This phosphorylation was also apparent on the Coomassie gel, due to the clear phosphoshift of the TMEM106B(1-96)-WT band (Figure 10C). Phosphorylation of TMEM106B(1-96)-Y50F, containing a Y50F mutation and TMEM106B(1-96)-3F, containing Y18F, Y50F and Y84F mutations, were also tested in a kinase assay (Figure 10E). Unexpectedly, weak phosphorylation of the 3F mutant was observed (Figure 10F), suggesting phosphorylation of GST may be occurring in the fusion proteins, since TMEM106B(1-96)-3F is phosphoresistant. Lack of phosphorylation of the GST control may be due to the different GST proteins used. The GST control protein was expressed using the pGEX-4T1 plasmid, whilst all TMEM106B(1-96) proteins were expressed from pCOLI_A_C_GST. Lack of a start codon for GST means pCOLI_A_C_GST cannot be used for the control. Furthermore, the amount of phosphorylation of the two mutants is likely an overestimation, due to overexposure of the blot, which is also supported by the barely visible phosphoshift for the two mutants on the Coomassie gel. This, coupled with the significantly higher level of phosphorylation for TMEM106B(1-96)-WT, means it is highly likely that phosphorylation of TMEM106B(1-96) is also occurring alongside any phosphorylation of GST. As such, TMEM106B(1-96)-WT can still be confirmed as an N1-Src substrate *in vitro*. Since there is no difference in the level of phosphorylation between TMEM106B(1-96)-Y50 and TMEM106B(1-96)-3F, this suggests the majority of phosphorylation is occurring on Y50 and hence is the main phosphosite for TMEM106B.

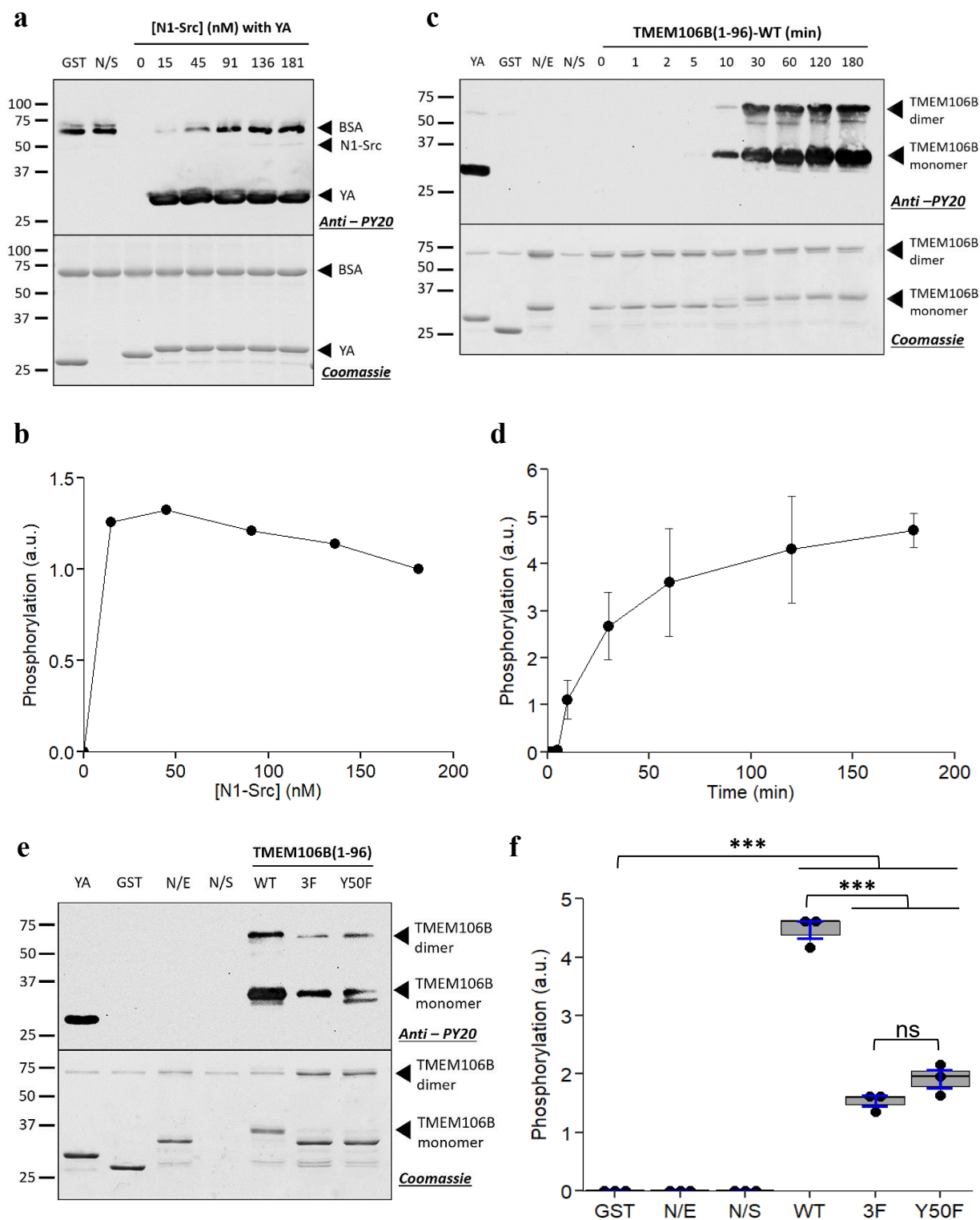


Figure 10 - Characterisation of *in vitro* kinase assays with His- Δ 80-N1-Src. All assays were conducted at 30 °C for 3 h, initiated by the addition of 0.5 mM ATP and halted by the addition of Laemmli buffer. N/E = no enzyme and N/S = no substrate. A) Increasing concentrations of N1-Src were incubated with 7 μ M YA, a GST-peptide which contains a single phosphorylatable tyrosine. Immunoblotting with anti-PY20 demonstrates phosphorylation of YA by N1-Src (top panel), with total protein shown via Coomassie staining (bottom panel). B) Quantification of (A) obtained via densitometry in ImageJ, where anti-PY20 bands were normalised to amount of protein shown by Coomassie. C) 5.4 μ M TMEM106B(1-96)-WT was incubated with N1-Src, with samples taken at time points indicated and treated as stated in (A). D) Quantification of monomeric TMEM106B(1-96)-WT in (C), analysed as described in (B). Data are plotted \pm SEM, from $n=3$ independent experiments. E) 5.4 μ M TMEM106B(1-96)-WT, TMEM106B(1-96)-Y50F and TMEM106B(1-96)-3F were incubated with N1-Src. Samples were analysed as stated in (A). F) Quantification of monomeric TMEM106B(1-96) in (E), as described in (B). Error bars are plotted \pm SEM (blue), from $n=3$ independent experiments. Statistical significance was determined by one-way ANOVA with Tukey post-hoc. *** $p < 0.001$.

3.5 TMEM106B(1-96)-WT is a dimer in solution

Due to the consistent presence of TMEM106B(1-96) dimers on SDS-PAGE gels, it was hypothesised that TMEM106B dimer is the predominant multimeric form of the soluble protein *in vitro* and potentially for the full length protein in cells. Indeed, previous immunoprecipitation studies in transfected cell lines have confirmed TMEM106B can dimerise (24). To assess the multimerisation of soluble TMEM106B(1-96)-WT, size exclusion chromatography with multi-angle light scattering (SEC-MALS) was employed (conducted by Dr Andrew Leech). This technique separates species along a column according to size, with larger species eluted first and smaller species eluted last. Refractive index (RI) measurements, combined with UV absorbance and light scattering signal, provide accurate molecular weight estimates, alongside relative amounts of individual species. In this experiment, purified TMEM106B(1-96)-WT was tested under three different conditions; native TMEM106B(1-96)-WT not subjected to any assay conditions, TMEM106B(1-96)-WT phosphorylated by N1-Src in a standard kinase assay (phosphorylated TMEM106B(1-96)-WT) and 'control' TMEM106B(1-96)-WT subjected to the conditions of a kinase assay, but in the absence of N1-Src. Successful phosphorylation of the phosphorylated TMEM106B(1-96) sample was confirmed via western blotting with anti-PY20 (Figure 11A).

Figure 11B-D show that there are three major peaks for each condition at approximately 21 min, 23 min and 27 min. Since the molecular weight of monomeric TMEM106B(1-96)-WT is 37 kDa, these peaks correspond to hexameric, tetrameric and dimeric forms of TMEM106B(1-96)-WT respectively. The largest peak for all conditions is the dimer, suggesting TMEM106B(1-96)-WT predominantly forms dimers in solution. Furthermore, since all peaks largely overlap between native, phosphorylated and the control TMEM106B(1-96)-WT (Figure 11E), the formation of dimers is unaffected by phosphorylation or the conditions of the kinase assay. The slight shift to the left for the phosphorylated TMEM106B(1-96) trace (blue) is indicative of a slightly bulkier conformation, concomitant with repulsion between the two phosphate groups on the dimer, resulting in a slightly shorter elution time. Furthermore, comparison of the amount of species present for each condition shows the kinase assay affects the overall composition of total TMEM106B(1-96)-WT (Figure 11G), with a higher percentage of dimer and a lower percentage of tetramer and hexamer for the phosphorylated and control TMEM106B(1-96)-WT conditions compared to the native TMEM106B(1-96)-WT. This is likely

due to the kinase assay being conducted at 30 °C for 3 h, resulting in breakdown of higher molecular weight species.

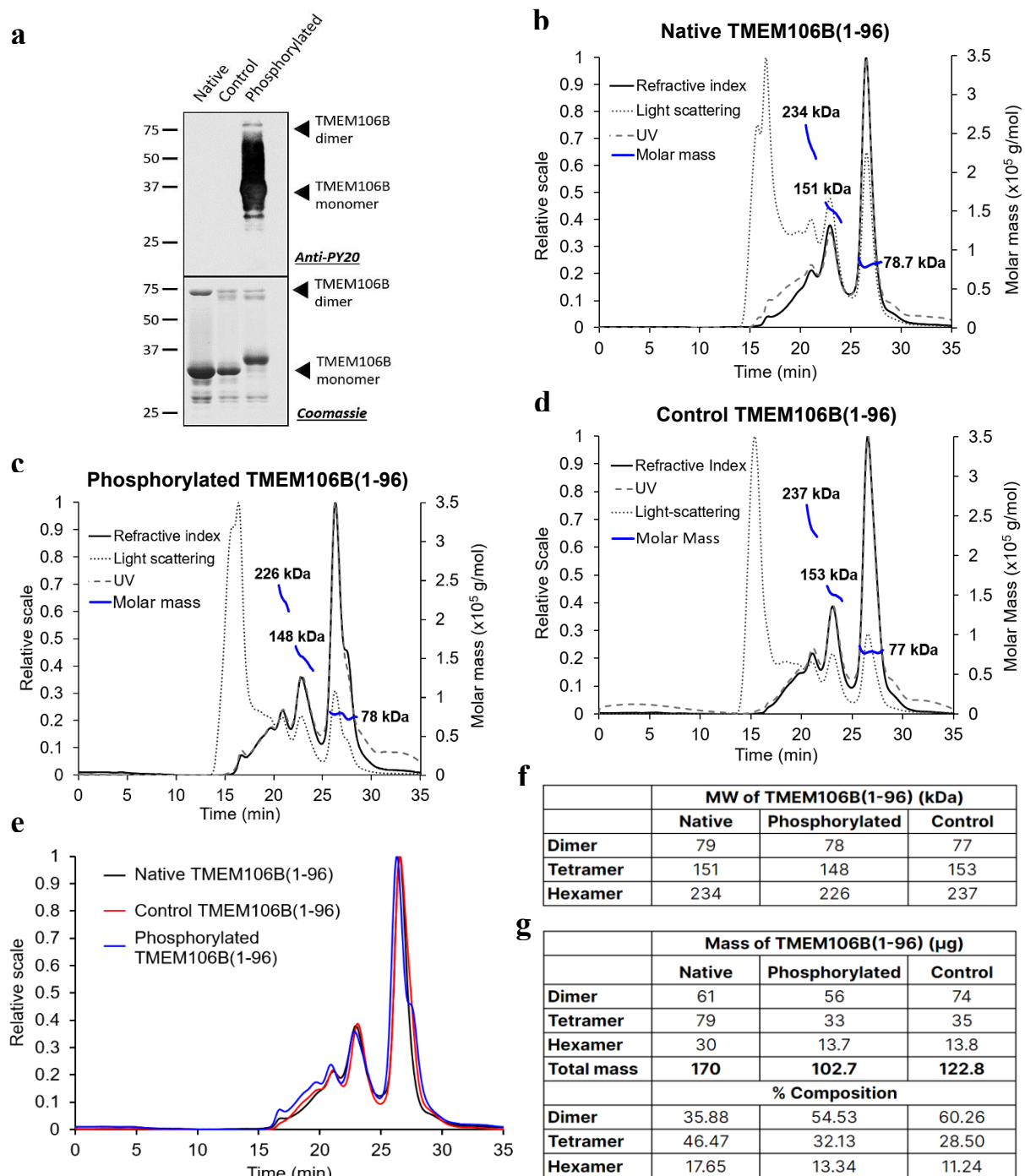


Figure 11 – SEC-MALS analysis shows TMEM106B(1-96)-WT predominantly forms dimers in solution. A) SEC-MALS samples were run on SDS-PAGE, with subsequent anti-PY20 immunoblotting demonstrating phosphorylation of the phosphorylated TMEM106B(1-96) sample by N1-Src (top panel). Total protein content is shown via Coomassie staining (bottom panel). B-D) Size exclusion chromatography (SEC) separates protein according to size, with larger proteins eluted first and smaller proteins eluted last. The refractive index (RI), UV absorption and light scattering signals are indicative of the weight concentration of protein present. Combination of all three signals provides a reliable estimation of the molar mass of protein present within each peak. TMEM106B(1-96)-WT was tested under three conditions: (B) ‘native’ TMEM106B(1-96)-WT not subjected to any assay conditions, (C) TMEM106B(1-96)-WT phosphorylated by N1-Src in a standard kinase assay

(phosphorylated TMEM106B(1-96)-WT) and (D) 'control' TMEM106B(1-96)-WT subjected to the conditions of a kinase assay, but in the absence of N1-Src. E) Overlay of the RI traces from each condition. F) Table summarising the molecular weights of each species present for each condition. G) Table summarising the % composition of TMEM106B(1-96)-WT under each condition.

There are no published structures of the TMEM106B N-terminus and so to shed light on how a dimer might be formed, a prediction was conducted using AlphaFold Multimer (115) (Figure 12). This showed that the N-terminus, including the YXXΦ motifs at Y18 and Y50, are predicted to be disordered, suggesting availability for binding by both N1-Src and AP2M1. This compares to the highly structured transmembrane and C-terminal domains, which contain α-helices and β-sheets respectively and aligns well with the recently solved crystal structure of the TMEM106B C-terminus (38). This prediction also supports the presence of a CxxCxGxG motif at amino acids 61-68, which was previously identified using the Multiple Expectation Maximization Algorithms for Motif Elucidation (MEME) tool (116), and is indicative of a Zn²⁺ binding domain. This could promote formation of the dimer, as well as providing structural support to the disordered N-terminus.

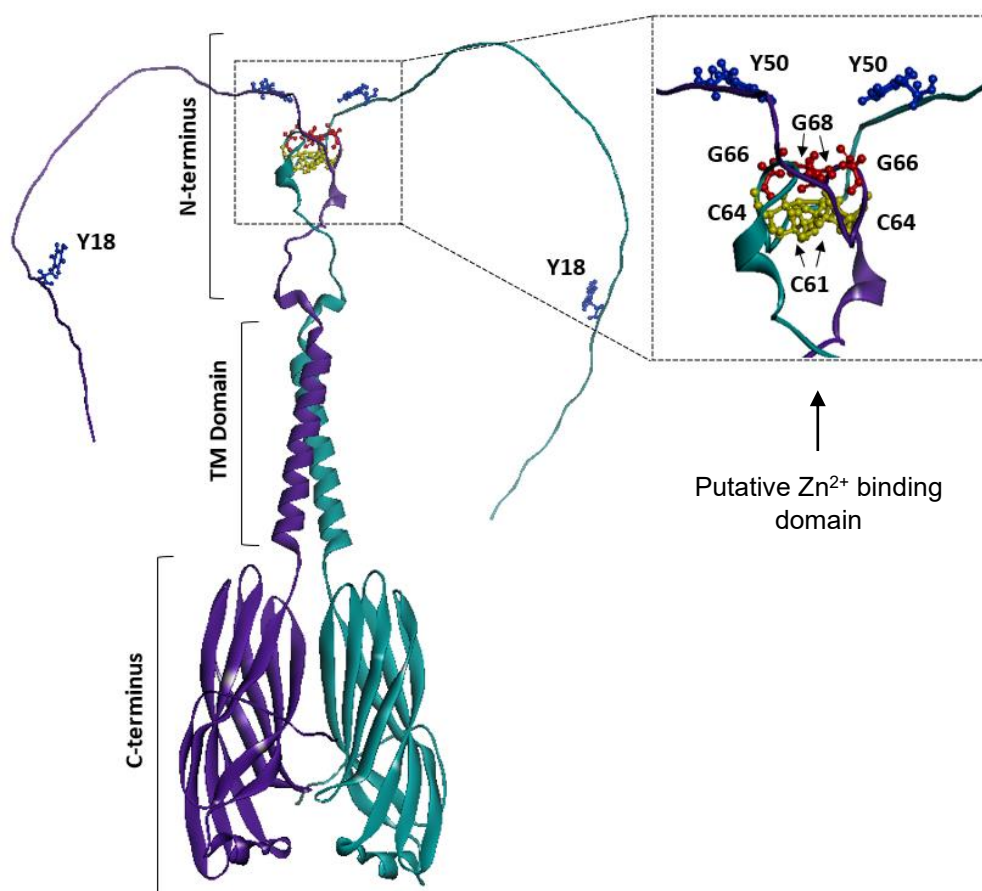


Figure 12 – AlphaFold prediction of TMEM106B dimers. The N-terminus is disordered, whilst the transmembrane and C-terminal domains are structured, containing α-helices and β-sheets respectively. Putative YXXΦ motifs at Y18 and Y50, predicted to bind the AP2 subunit AP2M1, are indicated. A putative Zn²⁺ binding domain is also indicated at residues C61-G68.

3.6 A putative phospho-dependent interaction between TMEM106B(1-96) and AP2M1

AP2M1 is a subunit of the adaptor complex AP2, and binds membrane receptors through sorting signals such as YXX Φ and dileucine motifs, subsequently triggering endocytosis of the receptor. AP2M1 was identified as a TMEM106B binding partner in a yeast 2 hybrid screen, validated by co-immunoprecipitation (24), with three potential AP2M1 binding sites identified in the N-terminus of TMEM106B. The first two are YXX Φ consensus sequences at Y18 and Y50, whilst the third is an extended dileucine motif with the sequence 'ENQLVALI' at E75. Since the TMEM106B(1-96)-Y50 phosphosite is present within the second YXX Φ motif, it was hypothesised that N1-Src phosphorylation of Y50 might regulate AP2M1 binding and impact the trafficking of TMEM106B to the lysosome.

To investigate this, *in vitro* binding assays were conducted with recombinant TMEM106B(1-96) and AP2M1. AP2M1 containing only the adaptor complex medium subunit (amino acids 167-435) was used, as amino acids 1-167 of AP2M1 form the clathrin adaptor complex subunit, which is not required for activity *in vitro*. His-GST-tagged AP2M1 was expressed using *E. coli* BL21 and purified using Ni²⁺-agarose beads. The His-GST tag was cleaved using 3C protease, enabling elution of un-tagged AP2M1. Although binding of AP2M1 to TMEM106B has not been assessed previously, analysis of AP2M1 binding to various YXX Φ peptides show AP2M1 has a k_D of 0.49 μ M (117), suggesting AP2M1 binding is a low affinity interaction. For this reason, AP2M1 binding using a concentration range of 0-1 μ M was initially trialled with 1 μ M GST and 1 μ M TMEM106B(1-96). For this, increasing concentrations of AP2M1 (0-1 μ M) were incubated with glutathione beads containing either GST or TMEM106B(1-96)-WT. Unbound AP2M1 was eluted, and following five wash steps, Laemmli buffer was used to elute the bound protein, with samples from the bound elution analysed via SDS-PAGE. Figure 13 shows AP2M1 binds TMEM106B(1-96)-WT in a concentration dependent manner, with increased binding seen with increased AP2M1 concentration. There was very little binding to GST, showing this interaction was specific. However, saturation was not achieved with TMEM106B(1-96)-WT, suggesting higher concentrations of AP2M1 were required. Based on a k_D of 0.49 μ M, 5 μ M AP2M1 would have likely been needed to ensure saturation. However, due to difficulties with producing AP2M1, this was not possible for this study. For this reason, a concentration range of 0-1 μ M AP2M1 was selected for all additional experiments, with 1

μM bait protein also chosen to ensure the protein was visible on the Coomassie gel, removing the need to strip and re-probe AP2M1 blots.

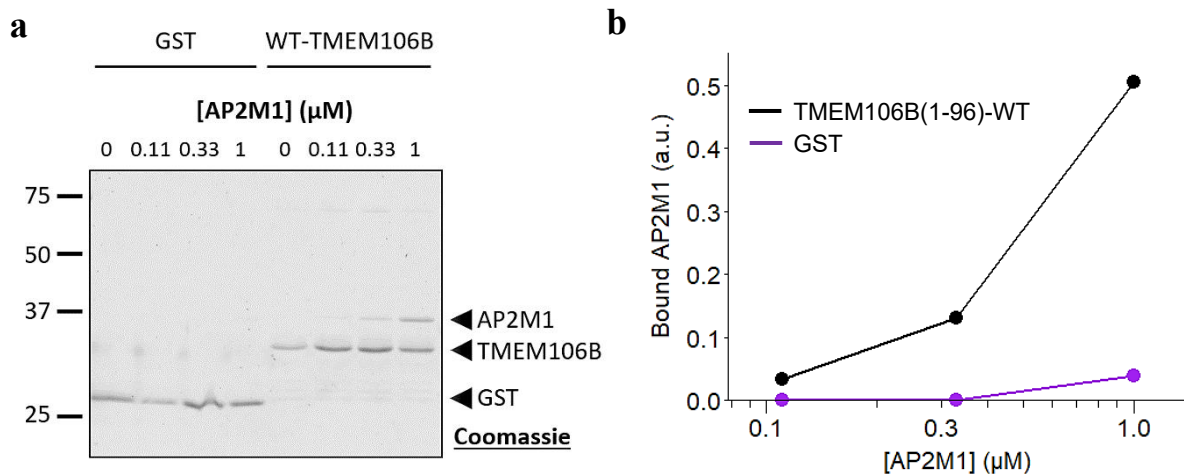


Figure 13 – AP2M1 binds to TMEM106B(1-96)-WT. Increasing concentrations of AP2M1 (0-1 μM) were incubated with glutathione-agarose beads containing either GST or TMEM106B(1-96)-WT (1 μM) for 1 h at 4 $^{\circ}\text{C}$. Any unbound protein was eluted and following five wash steps, bound protein was eluted from beads using Laemmli buffer. Samples were analysed via SDS-PAGE (A) and quantified using densitometry analysis in ImageJ (B).

The next set of experiments (Figure 14A and B) compared AP2M1 binding between TMEM106B(1-96)-WT, phosphorylated TMEM106B(1-96)-WT and TMEM106B(1-96)-Y50F. TMEM106B(1-96)-WT was phosphorylated using N1-Src in a kinase assay, with the other bait proteins subjected to the same conditions of the kinase assay, but in the absence of N1-Src. Phosphorylation was confirmed via immunoblotting with anti-PY20 (Figure 14A). Figure 14B shows there was little difference in AP2M1 binding to TMEM106B(1-96)-WT, phosphorylated TMEM106B(1-96)-WT or TMEM106B(1-96)-Y50F at 0.11 μM and 0.33 μM AP2M1 concentrations, but a large difference was observed upon phosphorylation or mutation of Y50 at 1 μM AP2M1. The reduced binding for phosphorylated TMEM106B(1-96)-WT and TMEM106B(1-96)-Y50F was not found to be statistically significant when compared to TMEM106B(1-96)-WT ($p = 0.10$ and $p = 0.06$ respectively), but this is likely due to the low AP2M1 concentration. Increasing the concentration of AP2M1 to where saturation is reached with TMEM106B(1-96)-WT, would allow the effects of phosphorylation and Y50F mutation to be better explored. Despite this, these results demonstrate that AP2M1 binds TMEM106B(1-96)-WT, with some binding occurring on Y50. To test whether binding also occurs at Y18 and/or the extended dileucine motif at E75, TMEM106B(1-96)-3F was employed in a third set of experiments. If binding occurred only at Y18 and Y50, then you would expect no AP2M1 binding to TMEM106B(1-96)-3F. If binding occurred at Y50 and E75, then you would expect a

similar reduction in binding as seen with TMEM106B(1-96)-Y50F, as the dileucine motif is not mutated. Surprisingly, AP2M1 binding to TMEM106B(1-96)-3F was the same as that to TMEM106B(1-96)-WT (Figure 14C and D). One possibility for this may be that the Y84F mutation strengthens the dileucine motif, promoting binding at this site.

3.7 Expression of N1-Src does not affect TMEM106B localization in HeLa cells

To assess how N1-Src phosphorylation of TMEM106B affects its cellular localisation, an N1-Src (FLAG-tagged) doxycycline-inducible HeLa cell line was transfected with plasmids containing either full length TMEM106B-WT or TMEM106B-Y50F, both of which were the T185 isoform and contained two consecutive Myc tags, followed by GFP at the C-terminus (TMEM106B-Myc-GFP). Immunoblotting with anti-FLAG and anti-Myc confirmed expression of N1-Src and TMEM106B respectively (Figure 15A-C). This showed that the level of N1-Src is decreased with expression of either TMEM106B-WT or TMEM106B-Y50F compared to GFP only. Previous studies suggest that increased activation of Src, increases its degradation through the proteasomal pathway due to ubiquitination by the E3 ubiquitin ligase protein c-Cbl (118,119), and might be what was occurring here. Although not statistically significant for this study, the increased protein level for TMEM106B-Y50F compared to TMEM106B-WT, is suggestive of a role of Y50 in TMEM106B stability. This is in keeping with a previous study that found Y50A mutation increased the amount of TMEM106B protein compared to WT (82). Furthermore, immunoblotting with anti-TMEM106B in Figure 15D shows there is an increased level of endogenous TMEM106B upon expression of transfected TMEM106B-WT and TMEM106B-Y50F, suggesting transfection of TMEM106B is stabilising the endogenous TMEM106B.

Fluorescent imaging of fixed N1-Src HeLa cells, showed TMEM106B (GFP fluorescence) is predominantly perinuclear, with no differences between TMEM106B-WT and TMEM106B-Y50F (Figure 16-19). Staining with anti-FLAG indicates N1-Src is present mainly at either the plasma membrane or perinuclear regions (Figure 16). Co-localisation between N1-Src and TMEM106B-WT/Y50F supports TMEM106B is a substrate of N1-Src in cells, however this co-localisation is unaffected by mutation of Y50.

To determine whether TMEM106B-WT localises to lysosomes, as found previously (20,22–24), cells were stained with the late-endosomal/lysosomal marker LAMP1. Surprisingly, LAMP1 staining appeared reduced in cells that were successfully transfected with either TMEM106B or eGFP control plasmids, compared to cells that were not (Figure 17). This suggests that potentially the uptake of the plasmid and transfection agent is affecting lysosome function. However, in the few cells that did have sufficient LAMP1 staining, co-localisation between LAMP1 and TMEM106B-WT was observed, but this was unaffected by

the presence of N1-Src (Figure 18). Furthermore, it was also apparent that these cells lacked the swollen vacuoles that are typically observed with TMEM106B overexpression, as found in other published studies (20,23,82).

Due to the diffuse TMEM106B staining pattern observed, staining with the ER-marker anti-Calnexin was also conducted (Figure 19). This showed a high level of co-localisation between calnexin and both TMEM106B-WT and TMEM106B-Y50F, suggesting the majority of TMEM106B was present in the ER. However, again this was unaffected by the presence of N1-Src or mutation of Y50. It is possible that the heterogeneity of the cell population is making it difficult to see any differences between the conditions. As such serum starvation to induce cell cycle synchronization would be useful to determine if this is the case. Furthermore, since N1-Src is a neuronal-specific kinase, the effect of TMEM106B phosphorylation on trafficking is likely better analysed in neurons, particularly since TMEM106B deficiency was found to increase retrograde transport of lysosomes in neurons (26,50). It was also found that GFP fluorescence was quenched by the acidity of lysosomes, making it difficult to identify cells which had been successfully transfected. Staining with anti-Myc, alongside GFP fluorescence, was also trialled but found to have a high background. Removal of the GFP tag and staining with anti-Myc only might be more useful.

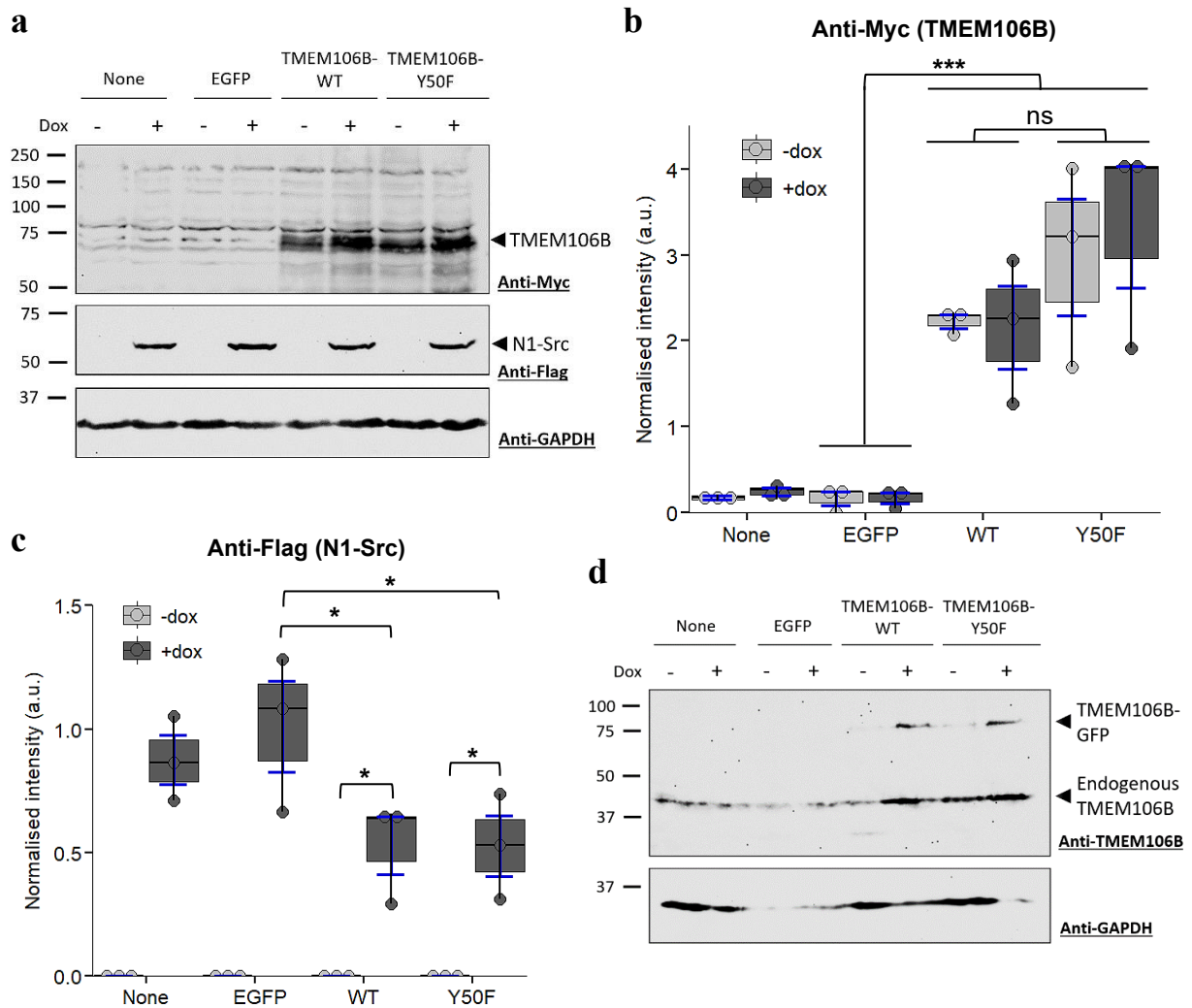


Figure 15 – Analysis of N1-Src HeLa cells transfected with TMEM106B. N1-Src HeLas were plated at a density of 200,000 cells per well and transfected 24 h later with 2 μ g plasmid containing either TMEM106B-WT, TMEM106B-Y50F or EGFP only. ‘None’ refers to cells not transfected with a plasmid. After 48 h incubation, cells were lysed using Laemmli buffer. Cell lysate samples were subjected to SDS-PAGE and transferred onto PVDF for immunoblotting. A) Immunoblotting with anti-Myc indicates total TMEM106B-Myc-GFP, whilst anti-Flag indicates level of N1-Src. B,C) Quantification of (A) obtained via densitometry in ImageJ, where anti-Myc (B) and anti-Flag (C) bands were normalised to GAPDH. Error bars (blue) are plotted \pm SEM, from $n=3$ independent experiments. Statistical significance was determined via two-way ANOVA with Tukey post-hoc. * $p < 0.05$, *** $p < 0.001$. D) Immunoblotting with anti-TMEM106B indicates level of endogenous TMEM106B and transfected TMEM106B.

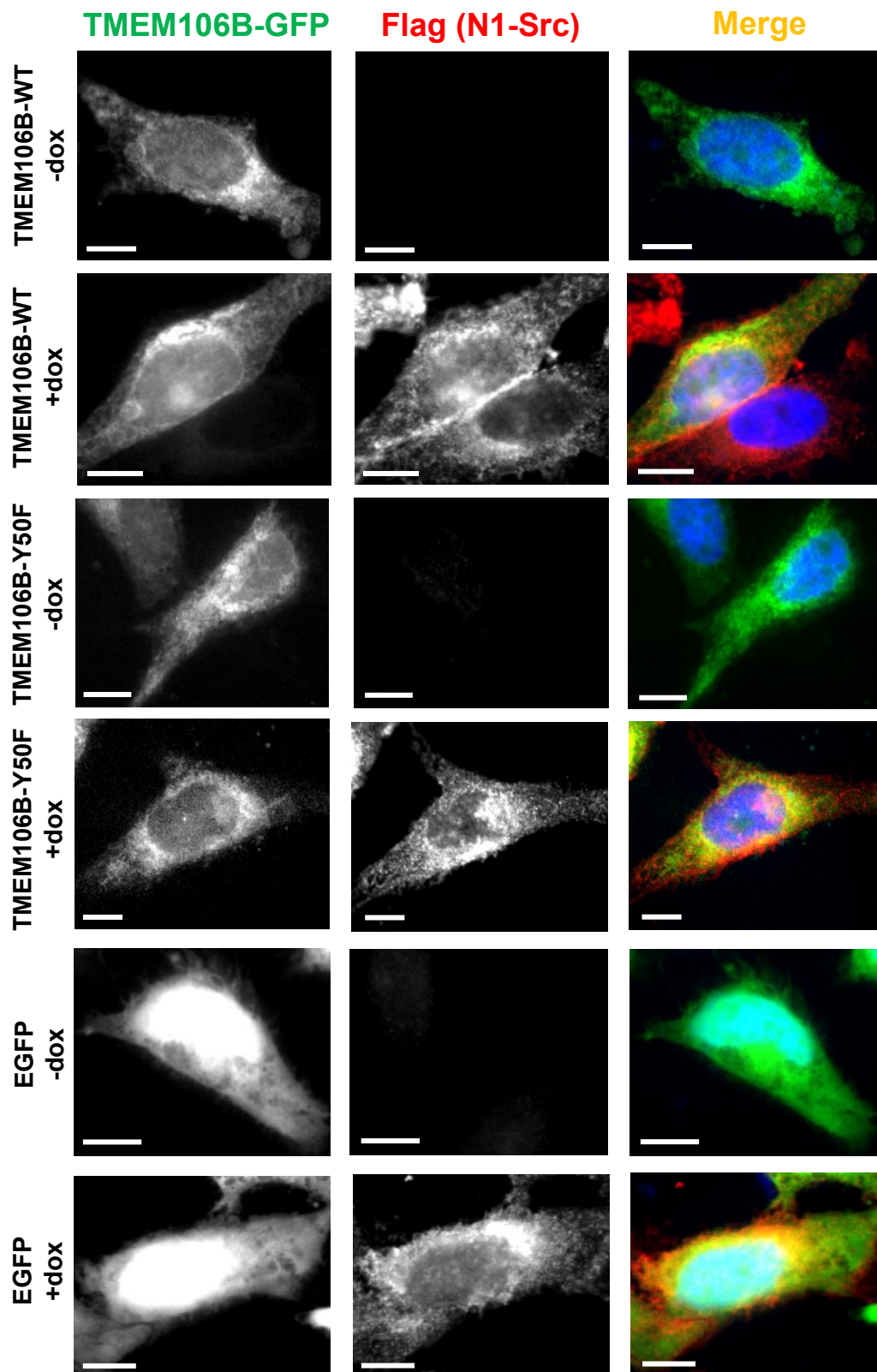


Figure 16 – TMEM106B is located predominantly in perinuclear regions. An N1-Src (FLAG tagged) inducible HeLa cell line was transfected with either -WT or TMEM106B-Y50F tagged with two consecutive Myc tags and GFP (TMEM106B-Myc-GFP) at the C-terminus (green), or with empty pEGFP-N1 as a negative control (green). Cells were either treated with 1 $\mu\text{g}/\text{ml}$ doxycycline to induce N1-Src expression (+dox) or left untreated (-dox). After 48 h post-transfection, cells were fixed and stained with anti-FLAG (red). Images were captured at 100x magnification, with the scale bar indicating 10 μm .

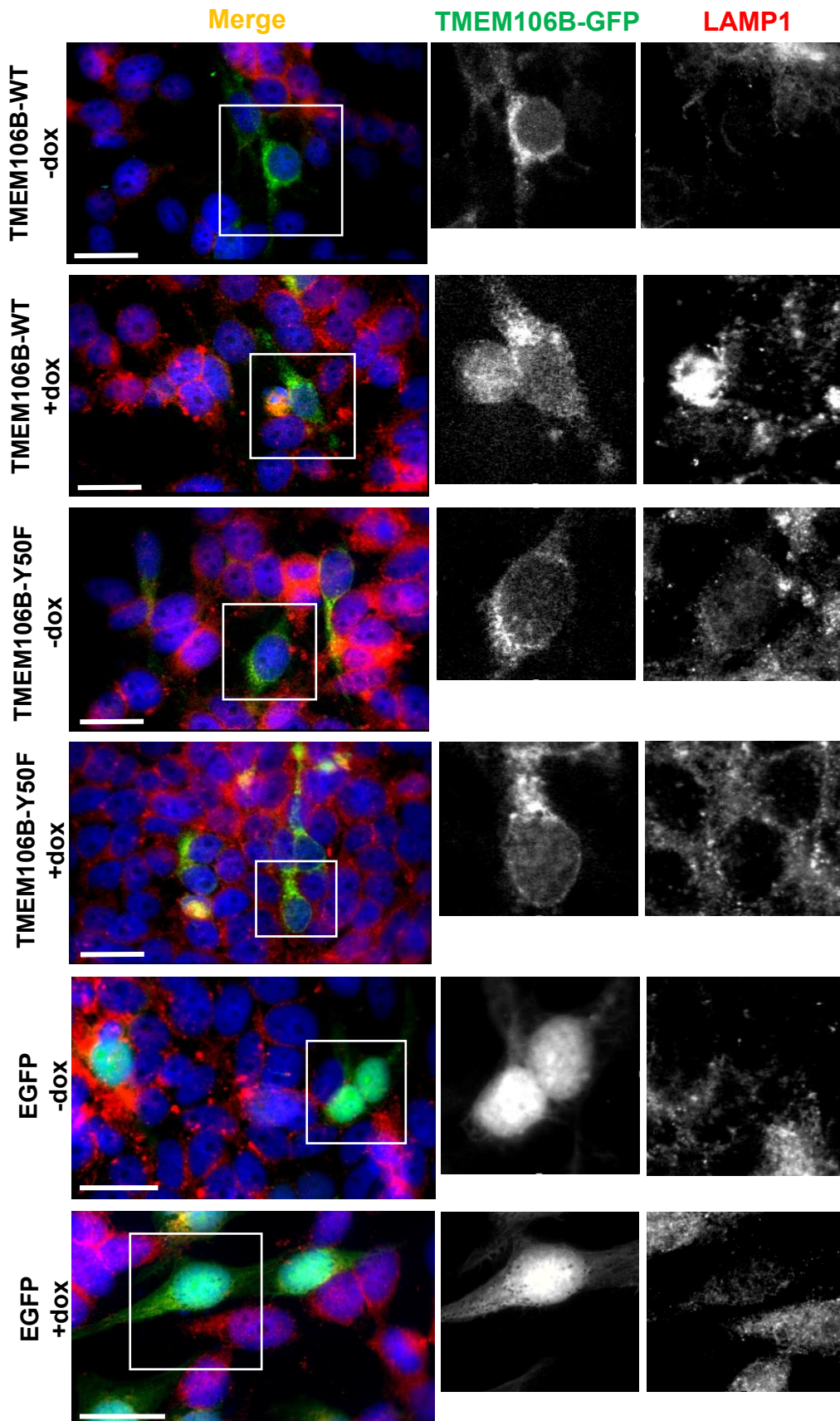


Figure 17 – LAMP1 staining is reduced in transfected cells. An N1-Src (FLAG tagged) inducible HeLa cell line was transfected with either -WT or TMEM106B-Y50F tagged with two consecutive Myc tags and GFP (TMEM106B-Myc-GFP) at the C-terminus (green), or with empty pEGFP-N1 as a negative control (green). Cells were either treated with 1 μ g/ml doxycycline to induce N1-Src expression (+dox) or left untreated (-dox). After 48 h post-transfection, cells were fixed and stained with the lysosomal marker anti-LAMP1 (red). Images were captured at 100x magnification, with the scale bar indicating 25 μ m.

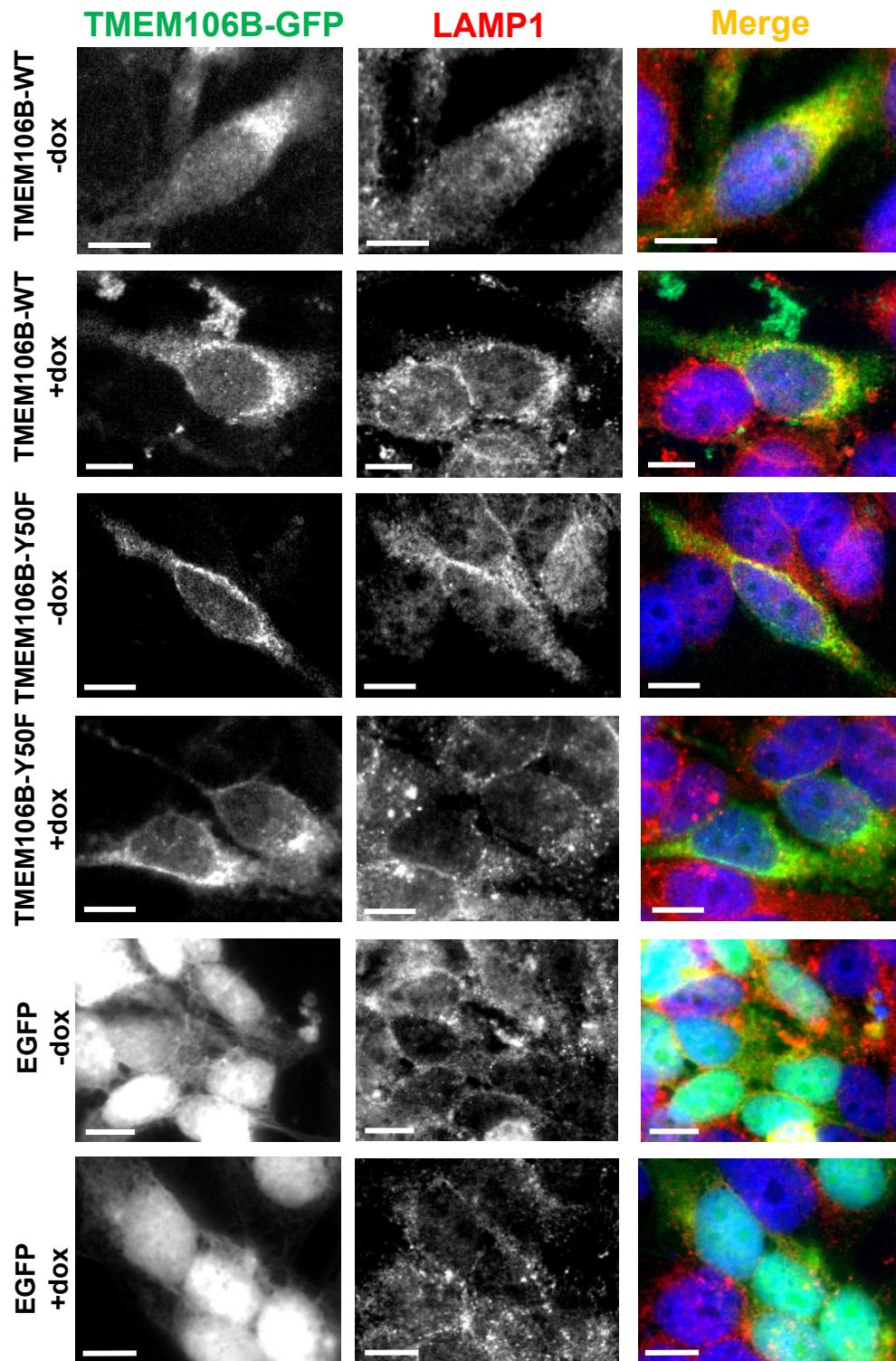


Figure 18 – TMEM106B localizes at lysosomes. An N1-Src (FLAG tagged) inducible HeLa cell line was transfected with either -WT or TMEM106B-Y50F tagged with two consecutive Myc tags and GFP (TMEM106B-Myc-GFP) at the C-terminus (green), or with empty pEGFP-N1 as a negative control (green). Cells were either treated with 1 $\mu\text{g/ml}$ doxycycline to induce N1-Src expression (+dox) or left untreated (-dox). After 48 h post-transfection, cells were fixed and stained with the lysosomal marker anti-LAMP1 (red). Images were captured at 100x magnification, with the scale bar indicating 10 μm .

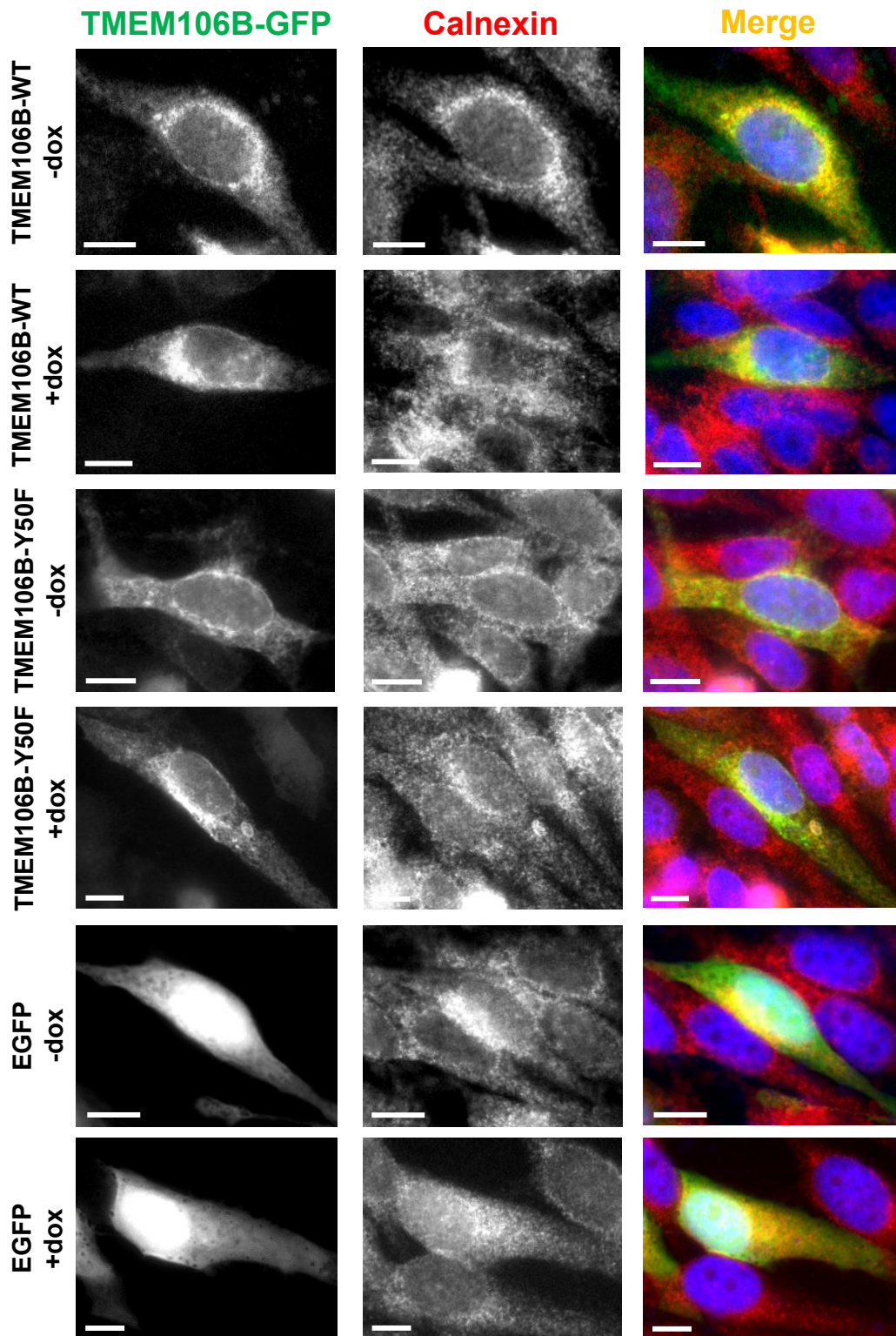


Figure 19 – TMEM106B localizes to the endoplasmic reticulum (ER). An N1-Src (FLAG tagged) inducible HeLa cell line was transfected with either -WT or TMEM106B-Y50F tagged with two consecutive Myc tags and GFP (TMEM106B-Myc-GFP) at the C-terminus (green), or with empty pEGFP-N1 as a negative control (green). Cells were either treated with 1 μ g/ml doxycycline to induce N1-Src expression (+dox) or left untreated (-dox). After 48 h post-transfection, cells were fixed and stained with the ER marker anti-calnexin (red). Images were captured at 100x magnification, with the scale bar indicating 10 μ m.

4. Discussion

TMEM106B is a lysosomal membrane protein (22–24) that has been implicated in FTLD-TDP, with the C-terminus (amino acids 120-254) shown to form amyloid fibrils in patients with FTLD-TDP and other neurodegenerative diseases (34–36). However, very little is known regarding the post-translational modifications of the protein, particularly how they might regulate its function or how it might relate to the development of FTLD-TDP. Recently, TMEM106B-Y50 was identified as a substrate for N1-Src, a neuronal nonreceptor tyrosine kinase, in a high-throughput phosphoproteomics screen (84). This suggests TMEM106B undergoes tyrosine phosphorylation in cells either on one or both residues. In this study, TMEM106B(1-96)-Y50 was confirmed as a substrate of N1-Src in an *in vitro* kinase assay. Following this, the study aimed to characterise the role of TMEM106B tyrosine phosphorylation using biochemical and cell biological approaches. Testing the effect of N-terminal TMEM106B phosphorylation on its protein-protein interactions revealed no influence on homodimerisation, but a reduction in its interaction with the cargo adapter AP2M1. Subsequent assessment of whether reduced AP2M1 binding affects sorting of phosphorylated TMEM106B in HeLa cells found TMEM106B localisation is unaffected by phosphorylation. However, quantification of TMEM106B protein levels revealed TMEM106B protein was increased upon Y50F mutation, suggesting this residue may regulate protein stability. Conversely, TMEM106B expression was found to reduce N1-Src protein levels, indicating N1-Src activity might regulate its degradation. Together, these results demonstrate TMEM106B is phosphorylated by N1-Src at the N-terminus and that this regulates the function of TMEM106B through the modulation of binding partners such as AP2M1. However, further investigation is required to understand how phosphorylation of TMEM106B might affect stability of both TMEM106B and N1-Src.

4.1 **TMEM106B(1-96)-GST and His-Δ80-N1-Src provide a reliable source of protein for *in vitro* studies**

Currently, no bacterial N-terminal TMEM106B expression plasmids are publicly available to purchase. However, previous groups have generated mammalian full length *TMEM106B* expression constructs with either a HA (22) or FLAG (23) tag for use in mammalian cell lines. For this study, a C-terminal GST tag was chosen to reflect *in vivo* anchorage of the N-terminus to the lysosomal membrane, thereby enabling a free N-terminus that can interact with

proteins. GST should also improve solubility, since the N-terminus of TMEM106B is largely disordered (37), as well as aiding purification. The mutations Y18F, Y50F and Y84F were chosen as N1-Src phosphorylates tyrosine residues and these are the only tyrosines present within the N-terminus. The Y84F mutation has not been previously tested in any capacity, but Y18A and Y50A mutations have been assessed in mammalian cells, where it was shown trafficking of TMEM106B to lysosomes was unaffected by these mutations (82). However, these residues have not been analysed in relation to phosphorylation.

The three TMEM106B(1-96) plasmids were used to transform *E. coli* BL21 DE3 cells for subsequent protein expression. Induction conditions of 1.0 mM IPTG with a 3-hour incubation at 37 °C were employed and replicated those used by Schweighauser *et al.* (35) to produce TMEM106B C-terminal fragment (AA 120-274). As illustrated in Figure 8C, TMEM106B(1-96)-GST proteins were successfully produced, with TMEM106B(1-96) present at the expected molecular weight of ~36 kD. However, an additional species was present at ~72 kD, which corresponds to a TMEM106B(1-96) dimer and is in-line with previous observations that showed TMEM106B can interact both with itself and with its homolog TMEM106C to form homodimers and heterodimers respectively (24). Since generation of recombinant TMEM106B(1-96) has not been recorded previously, this study provides the first reliable method of producing TMEM106B(1-96) *in vitro*.

Recombinant expression of N1-Src proved more difficult than TMEM106B(1-96) since bacterial genomes lack kinases and thereby do not have mechanisms to support phosphorylation. As such, expression of N1-Src in bacteria can trigger aberrant signalling pathways resulting in toxic build-up of phosphorylated protein (120). Various methods have been developed to enable successful expression and purification of both active N1-Src and C-Src, including ways to increase protein solubility (121) and to counteract kinase function (122). In this study, a new construct was engineered using the pCDF-Duet-1 vector. This plasmid contains two RBS, enabling simultaneous expression of PTP1B and His- Δ 80-N1-Src independently from the same plasmid. PTP1B is vital for dephosphorylation of the N1-Src autophosphorylation site (Y416), preventing Src activity, as well as dephosphorylating any bacterial proteins phosphorylated by Src. A similar construct has been used to express the serine/threonine kinase tousled-like kinase 1B (TLK1B), alongside the protein phosphatase from bacteriophage λ (123), but has not yet been used for Src expression. This method

dramatically increased the yield of soluble N1-Src protein (Figure 9), suggesting independent expression of PTP1B may improve phosphatase function. Further alterations to induction conditions, such as length and temperature of induction, might improve this yield.

4.2 The TMEM106B N-terminus is a newly identified N1-Src substrate *in vitro*

With TMEM106B(1-96) and His- Δ 80-N1-Src in hand, phosphorylation of TMEM106B(1-96)-WT was tested in an *in vitro* kinase assay (Figure 10). This confirmed TMEM106B(1-96) is a substrate of N1-Src *in vitro*, providing the first evidence that TMEM106B function may be regulated by tyrosine phosphorylation. To determine whether phosphorylation occurs predominantly on Y50, as indicated by the phosphoproteomics screen (84), TMEM106B(1-96)-Y50F and TMEM106B(1-96)-3F were also tested in an N1-Src kinase assay. Unexpectedly, a low level of non-specific phosphorylation of TMEM106B(1-96)-3F was observed, despite being phosphoresistant (Figure 10E). Analysis of GST reveals a sequence (EYLEEKYE) at the N-terminus that resembles a Src kinase substrate motif, suggesting the residual phosphorylation may be due to phosphorylation of GST. Lack of phosphorylation with the negative GST control may be due to protein conformation, where the N-terminus is buried and inaccessible to N1-Src in this state. Addition of a peptide at the N-terminus may lead to a more open conformation, enabling phosphorylation by N1-Src. Phosphorylation of GST could be confirmed using trypsin digest followed by mass-spectrometry. Nonetheless, the increased phosphorylation for TMEM106B(1-96)-WT confirms TMEM106B(1-96) is an N1-Src substrate, with Y50 acting as the main phosphosite due to similar phosphorylation levels between TMEM106B(1-96)-Y50F and TMEM106B(1-96)-3F.

The main caveat with kinase assays is that you are creating an artificial environment that does not reflect physiological conditions. As such, confirming that this phosphorylation also occurs in cells would be an important next step. However, pinpointing phosphorylation of a substrate to one kinase is difficult, due to the large expanse of kinases present in the cell. Phosphorylation also typically occurs at very low basal levels, making it harder to detect through processes such as western blotting. Despite this, PhosphoSitePlus shows TMEM106B-Y50 phosphorylation is readily detectable in cells, due to the high number of HTP recorded for this residue. As such, mass spectrometry could be used to compare TMEM106B-Y50 phosphorylation in a neuronal cell line, in the presence and absence of N1-Src siRNA. Upon

knockdown of N1-Src, it would be hypothesised that TMEM106B-Y50 phosphorylation would be reduced. Since phosphorylation of TMEM106B has not been investigated previously, this post-translational modification could provide a new avenue for identifying additional functions of TMEM106B. The potential implications of TMEM106B phosphorylation are discussed in the following sections.

4.3 Phosphorylation of TMEM106B(1-96) does not affect dimer formation

TMEM106B dimers have been frequently detected on SDS-PAGE gels and western blots, both in this study and previous studies (20,23), suggesting TMEM106B functions as a dimer at the lysosome membrane. Stagi *et al.* (24) also went one step further, by showing that two TMEM106B constructs (TMEM106B-Cherry and TMEM106B-eGFP) can co-immunoprecipitate with each other, further supporting that TMEM106B is able to dimerise in cells. Often, phosphorylation can induce structural changes that either promote or inhibit dimerisation of proteins. As such, it was hypothesised that phosphorylation of TMEM106B(1-96) may affect dimer formation.

Analysis using SEC-MALS supported that TMEM106B(1-96)-WT predominantly forms dimers in solution (Figure 11), however N1-Src phosphorylation did not affect this, suggesting TMEM106B structure is unaffected by phosphorylation. Higher molecular weight species were also observed for both non-phosphorylated and phosphorylated TMEM106B(1-96)-WT that corresponded to TMEM106B(1-96)-WT tetramers and hexamers. Currently, there is no report of TMEM106B tetramers or hexamers elsewhere in the literature, and as such it is unknown whether these multimers have a physiological basis. Conducting a cell-based crosslinking experiment, followed by immunoblotting with anti-TMEM106B could reveal whether these higher order multimers are present in cells. Nonetheless, using only the N-terminus of TMEM106B shows that this alone is sufficient for dimerisation. However, it is possible that the GST-tag is promoting dimerisation, as GST alone can dimerise, and as such the analysis should be repeated without the GST tag to confirm this is not the case.

Currently, there is little information regarding the structure of TMEM106B or how dimerisation might occur, as the crystal structure of full length TMEM106B has not yet been solved. To provide some insight into this, an AlphaFold multimer prediction of full length TMEM106B was conducted (Figure 12). This clearly highlighted the disordered nature of the

cytoplasmic N-terminus, supporting previous findings (37). Having an intrinsically disordered domain is thought to be advantageous, as it enables interactions with a diverse set of binding partners, such as N1-Src. Whether binding to these partners induces a more well-defined and folded structure, or if TMEM106B remains disordered, is currently unknown. The AlphaFold prediction also supports the presence of a putative Zn²⁺ binding domain at amino acids 61-68, with the sequence 'CPTCQGTG'(116). This, alongside a potential 'S/AxxxCxxxSG/S' motif in the transmembrane helix, is thought to drive dimerisation of TMEM106B (14). The importance of the Zn²⁺ binding domain in N-terminal dimerisation could be tested using SEC-MALS in the presence or absence of a Zn²⁺ chelator. In the absence of Zn²⁺, it could be hypothesised that TMEM106B(1-96) is unable to dimerise. Furthermore, the fact that Y50 is sufficiently far enough away from the Zn²⁺ binding domain, might explain why dimerisation is unaffected by phosphorylation.

Contrastingly, the AlphaFold prediction depicts the TMEM106B luminal domain to be a highly ordered structure, consisting mainly of β -sheets with a single α -helix. This aligns with a recent crystal structure solved for TMEM106B C-terminus, which spans residues 118-261 (38). This crystal structure showed there is a total of 7 β -sandwich folds, reminiscent of an immunoglobulin domain. The single α -helix is present at residues 208-216 and is stabilised by a disulphide bond between residues C214 and C253 (38). Furthermore, it has been shown through bioinformatic analysis, that this domain has homology to the late embryogenesis abundant-2 (LEA-2) domain superfamily (116). Although the LEA-2 domain is not well-understood, analysis of its structure reveals a groove or a series of cavities that are lined with hydrophobic residues, suggesting the domain binds hydrophobic molecules such as lipids. In the same study, TMEM106B was shown to have homology to the yeast vacuole proteins Vac7 and Tag1, both of which contain luminal LEA-2 domains (116). The vacuole is the yeast equivalent of a lysosome, and it's been shown that deletion of Vac7 leads to vacuole enlargement (124). Vac7 is also known to increase the lysosome inositide lipid PI(3,5)P₂ through an unknown mechanism, in response to stress (124), further supporting the link between LEA-2 domains and lipid signalling. As such, T. Levine (116) predicted TMEM106B to be a lipid transfer protein that functioned in the lumen of late endosomes and lysosomes. However, this has yet to be confirmed experimentally, and the crystal structure of TMEM106B luminal domain did not reveal a lipid-binding groove (38).

TMEM106B has also been found to form heterodimers with the paralogue TMEM106C (24). The function of TMEM06C is also unknown but has been found to localise to the lysosomal membrane (24), as well as the endoplasmic reticulum (39). Analysis of the TMEM106C sequence, shows that residues 51-58 contains the same putative Zn²⁺ binding motif as TMEM106B (Figure 7E), which is likely why the two paralogues can heterodimerise. TMEM106A also contains the same binding motif and has been shown to localise to the lysosome (39), suggesting TMEM106B might be able to form dimers with TMEM106A, but this has yet to be confirmed experimentally. Although phosphorylation did not affect homodimerisation of TMEM106B, it may affect heterodimerisation with either TMEM106A or TMEM106C, as the paralogues have slight differences in their N-terminal amino acid sequences, which might be more greatly affected by the presence of a negatively charged phosphate group. Interestingly, TMEM106B-Y50 and its surrounding sequence is conserved in TMEM106C (but not TMEM106A) at Y40, suggesting N1-Src may also be able to phosphorylate TMEM106C.

4.4 Phosphorylation reduces AP2M1 binding to TMEM106B(1-96) *in vitro*

Besides affecting structure, phosphorylation can also regulate protein-protein interactions. There are only a handful of known binding partners of the TMEM106B N-terminus, including AP2M1, MAP6 and CLTC (24,26). For this study, AP2M1 binding to TMEM106B(1-96) was investigated since it is currently unknown how AP2M1 may influence the function of TMEM106B and more importantly, an AP2M1 binding motif (YVEF) is present at the Y50 phosphorylation site. It was hypothesised that phosphorylation is more likely to prevent the interaction rather than strengthen it, as addition of a large negative charge to tyrosine will likely inhibit binding to the hydrophobic pocket of AP2M1 (125). There are also numerous examples in the literature that support this, such as phosphorylation of the YXXΦ motif in GABA_A receptors, which was found to inhibit AP2 binding and prevent subsequent internalisation (126). This resulted in an increase of surface receptors, increasing synaptic inhibition. Similarly, a Src family member was shown to phosphorylate Tyr-365 of K⁺-dependent Na⁺/Ca²⁺ exchanger 2 (NCKX2), present on the neuronal plasma membrane (127) This reduced endocytosis of NCKX2 and was found to enhance activity in dentate granule cells, supporting a link between Src proteins and trafficking.

Consequently, recombinant AP2M1 was incubated with either TMEM106B(1-96)-WT, phosphorylated TMEM106B(1-96)-WT, or TMEM106B(1-96)-Y50F in an *in vitro* binding assay. Increasing the concentration of AP2M1, increased the amount of AP2M1 bound to non-phosphorylated TMEM106B(1-96)-WT (Figure 14B), corroborating the co-immunoprecipitation assay that first identified AP2M1 as a binding partner (24). This binding was reduced for phosphorylated TMEM106B(1-96)-WT and TMEM106B(1-96)-Y50F, supporting the initial hypothesis that AP2M1 binds TMEM106B-Y50. However, AP2M1 could also be binding alternative sites within TMEM106B N-terminus - either the 'YDGV' motif at Y18, or the 'ENQLVALI' sequence at E75. Consequently, TMEM106B(1-96)-3F was tested in the same binding assay, meaning both Y18 and Y50 motifs are mutated, but not E75. However, binding to TMEM106B(1-96)-3F was the same as binding to TMEM106B(1-96)-WT (Figure 14D). This was unexpected, as it was thought that binding would either be reduced to similar levels as seen with TMEM106B(1-96)-Y50F, or almost abolished. It is possible that the Y84F mutation strengthens binding at E75, as Y84 is directly adjacent to the dileucine motif.

It should be noted that *in vivo*, AP2M1 is phosphorylated by AAK1, a serine/threonine kinase that phosphorylates AP2M1 on T156 and has been shown to dramatically increase AP2M1 binding affinity for sorting motifs (117,128). Consequently, it would be useful to conduct the binding assay in the presence of AAK1, as this would better reflect physiological conditions and AP2M1 phosphorylation could better discern differences in binding between TMEM106B(1-96)-WT and TMEM106B(1-96)-3F. Alternatively, a pull-down assay from cell lysates using WT, phosphorylated and mutant TMEM106B(1-96) could also be performed. Any endogenous AP2M1 detected will likely be phosphorylated and thereby provide a more robust link between AP2M1 and TMEM106B(1-96) binding. This was trialled in this study, but due to insufficient protein stock, the concentration of bait protein was too low, and the pulldown was unsuccessful. Furthermore, it is not possible to determine accurate K_d values from the current binding assay data, as plateau is not achieved. For this to occur, a larger ratio between AP2M1 and TMEM106B(1-96) concentrations is required. This would allow the binding curve to extend beyond what was achieved in this study, enabling differences between TMEM106B(1-96)-WT, TMEM106B(1-96)-Y50F and TMEM106B(1-96)-3F to be better assessed. Alternatively, a previous study used surface plasmon resonance (SPR) to determine the binding affinity between AP2M1 and YXX Φ peptides (117). This could be

trialled with TMEM106B, however, this can be expensive and requires large amounts of protein.

Nonetheless, the data generated in this study supports an interaction between AP2M1 and TMEM106B, with a function that is modulated by phosphorylation. Whether the various sorting motifs act cooperatively, or mediate independent functions is currently unclear. In a previous study, it was shown that only the E75 dileucine motif is responsible for localisation of TMEM106B to lysosomes, and that mutation of either Y18 or Y50 to alanine did not affect lysosome sorting (82). This suggests that the YXXΦ motifs may have an alternative function, and instead may be involved in internalisation of the protein from the plasma membrane, as a small proportion of TMEM106B has been detected on the plasma membrane in non-permeabilised cells (38). This was found to be true for the glycoprotein B (gB) of the varicella-zoster virus (VZV), which contains two YXXΦ motifs. One motif was required for endocytosis of the protein, whilst the other was not required for internalisation but promoted transport to the Golgi in infected cells (129). Endocytosis of TMEM106B could act as a recycling

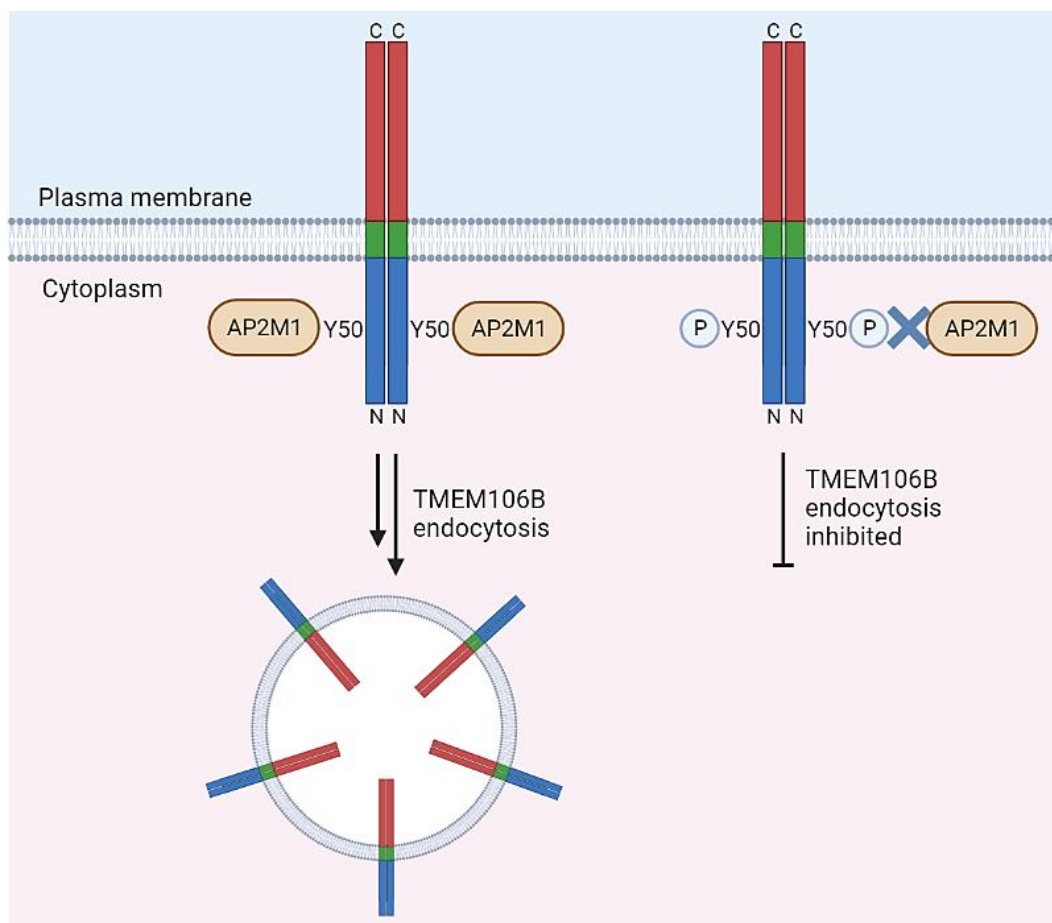


Figure 20 – AP2M1 binding to TMEM106B-Y50 could promote endocytosis of TMEM106B at the plasma membrane. Phosphorylation of TMEM106B-Y50 may prevent this interaction, inhibiting TMEM106B internalisation.

mechanism to retrieve TMEM106B from the plasma membrane after lysosome exocytosis has occurred (Figure 20). This hypothesis could be tested using an antibody-feeder assay (130–132). This would involve incubating cells expressing TMEM106B in media containing monoclonal antibodies (mAB) against the luminal domain of TMEM106B, thereby labelling any TMEM106B present on the plasma membrane. Endocytosis is subsequently triggered by incubating the cells at 37 °C, resulting in internalisation of the TMEM106B-mAB complex. Addition of a fluorophore-conjugated secondary antibody to non-permeabilised cells would allow detection of TMEM106B remaining on the plasma membrane. Permeabilization and use of a different fluorophore-conjugated secondary antibody would allow quantification of internalised TMEM106B. Comparison between cells expressing TMEM106B-WT and TMEM106B-Y50F might reveal if Y50 YXXΦ motif is required for TMEM106B internalisation.

Furthermore, from this study it is not clear whether AP2M1 binds at Y18. Sequence alignment revealed that the Y50 YXXΦ motif is conserved across all species (Figure 7D) but is only present in human and mouse at Y18, suggesting the Y18 YXXΦ motif is required for higher order function. It is possible that Y18 binds LC3 instead of AP2M1, as searching on the ELM database reveals a putative LC3 interacting region (LIR) overlapping the YXXΦ motif at amino acids 16-21 (DAYDGV), as well as at amino acids 46-51 (SQFPYV). LIRs typically have the sequence [W/F/Y]₀-X₁-X₂-[L/V/I]₃ and are mainly located in regions that are intrinsically disordered (133). LC3 are ubiquitin-like proteins conserved from the Atg8 family in yeast and are involved in autophagy and intracellular trafficking (134). Although binding of LC3 to TMEM106B has not been previously investigated, loss of TMEM106B has been shown to increase protein levels of LC3 and thereby might regulate autophagy through its lysosome function (68).

4.5 TMEM106B localises to the endoplasmic reticulum and lysosomes

N1-Src doxycycline-inducible HeLas were employed to investigate how N1-Src phosphorylation might influence the localisation of TMEM106B in cells. These cells were transiently transfected with plasmids containing TMEM106B (T185 isoform) C-terminally tagged with Myc and GFP. A C-terminal tag was chosen to prevent interaction with binding partners at the N-terminus and to avoid affecting potential myristylation of TMEM106B, as the start of the N-terminus contains a putative myristylation signal (MGXXXS) that might direct anchoring of the N-terminus to the lysosomal membrane (116). However, this meant

that GFP was exposed to the acidic lysosomal lumen resulting in the quenching of fluorescence, which made it more difficult to identify cells which had been successfully transfected.

It was hypothesised that expression of N1-Src would alter the lysosomal localisation of TMEM106B, and this was tested through the staining of the lysosome marker LAMP1. However, it was found that majority of transfected cells (both TMEM106B and eGFP only control) lacked LAMP1 staining, for reasons that are unknown (Figure 17). One study found that there was reduced lysosome number with TMEM106B overexpression in N2A cells (23), but this has not been noted elsewhere in the literature. It is possible that the transfection reagent was affecting lysosome function, but you would expect to see this across all cells. Immunoblotting of cell lysate samples with LAMP1 would have been useful to determine if total LAMP1 content differed between the different conditions. If the levels were similar between transfected and non-transfected cells, then this could indicate that LAMP1 was relocated to the plasma membrane. Staining of TMEM106B transfected, non-permeabilised HeLa cells with anti-LAMP1 might reveal if this is the case. However, when LAMP1 was present, co-localisation with TMEM106B was observed, but was unaffected by expression of N1-Src (Figure 18).

Due to the reticulated pattern seen with TMEM106B-GFP fluorescence, staining with the ER-marker Calnexin was also assessed (Figure 19). This showed strong co-localisation, indicating TMEM106B also localises to the ER. This has not been recorded previously for WT TMEM106B, although mutation of Y125 resulted in retention of TMEM106B in the ER (unpublished observations (39)). Again, phosphorylation did not affect ER localisation of TMEM106B, neither did mutation of Y50, suggesting phosphorylation does not affect steady state localisation of TMEM106B in HeLa cells. This is perhaps not surprising, as TMEM106B-Y50A was found to not affect localisation of TMEM106B in a previous study (82). It is possible that the Myc-GFP tag is interfering with C-terminal glycosylation, preventing the trafficking of TMEM106B from the ER to the lysosome. However, GFP-tagged TMEM106B has been successfully used previously, both with GFP tagged at the N-terminus (23,82) and C-terminus (24). Alternative C-terminal tags that have been used include cherry (24) and HA (22). The transfected cells also lacked the swollen vacuole phenotype that is common with TMEM106B

overexpression (20,23,82), but this could be due to majority of TMEM106B being trapped in the ER.

Furthermore, since N1-Src is a neuronal-specific tyrosine kinase, it is likely that the effects of phosphorylation are better examined in neurons, rather than HeLa cells. It has been shown that TMEM106B is implicated in lysosome trafficking along primary motorneurons (50). More specifically, it was found that TMEM106B deficient mice have swollen LAMP1-positive vacuoles which accumulate at the distal end of the axon initial segment, due to increased retrograde transport of lysosomes (50). Similarly, in primary hippocampal neurons increased retrograde transport of vesicles along dendrites was observed upon knockdown of TMEM106B (26). The authors proposed this was due to TMEM106B interacting with MAP6, where MAP6 acts as a molecular break, thereby inhibiting retrograde transport. In the absence of TMEM106B, this interaction cannot occur, and retrograde transport is disinhibited (26). As such, phosphorylation may have a more profound effect on trafficking in neurons, and could potentially prevent interaction with MAP6, thereby mimicking TMEM106B knockdown (Figure 21). This could be tested using live imaging of either Rab7a-GFP (26) or LAMP1-GFP (24) labelled lysosomes in neurons transfected with TMEM106B. Comparison of neurons with and without N1-Src knockdown could reveal differences in lysosome trafficking. Improved understanding of how TMEM106B regulates lysosomal trafficking could have implications for disease, as TMEM106B knockdown has been shown to rescue trafficking

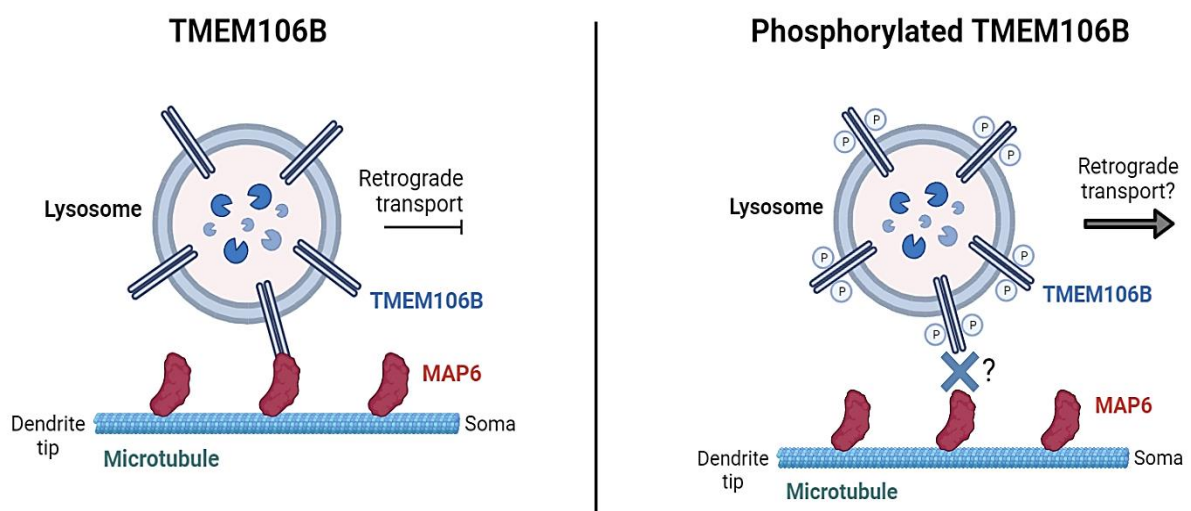


Figure 21 – Phosphorylation of TMEM106B-Y50 could inhibit binding to MAP6, promoting retrograde transport of lysosomes in neurons. TMEM106B, a lysosomal membrane protein was found to bind the microtubule stabilisation protein MAP6, and this interaction is thought to inhibit retrograde transport of lysosomes in dendrites. N-terminal phosphorylation of TMEM106B by the neuronal tyrosine kinase N1-Src, could prevent this interaction, thereby promoting retrograde transport along microtubules in neurons.

defects caused by CHMP2B^{Intron5} mutation in dendrites (70). CHMP2B is a protein involved in the trafficking of protein through the endo-lysosomal pathway, with mutations found to contribute to FTL. Incorporation of CHMP2B^{Intron5} into neurons was found to reduce endolysosomal trafficking in dendrites, but subsequent knockdown of TMEM106B restored trafficking to the correct levels (70). As such, in FTL patients with CHMP2B^{Intron5} mutation, TMEM106B could act as a therapeutic target.

Lysosome movement is also important for lysosome exocytosis, a process where lysosomes fuse with the plasma membrane and secrete lysosomal content into the extracellular matrix (44). In lung cancer cells, it was found that overexpression of TMEM106B was found to increase lysosome exocytosis, determined by increased levels of LAMP1 on the plasma membrane (53). The opposite effect was observed in Oli-Neu cells (oligodendrocyte precursors) where TMEM106B knockout resulted in 49.6% reduction in LAMP1 levels on the cell surface, compared to controls (33). Together, these results support a role of TMEM106B in lysosome exocytosis, likely linked to its role in trafficking. As such, LAMP1 cell surface levels could also be compared between neurons overexpressing TMEM106B, with and without N1-Src to further determine whether phosphorylation affects lysosome trafficking and exocytosis. Quantification of surface LAMP1 could be achieved through staining of non-permeabilised fixed cells with an antibody that recognises the extracellular domain of LAMP1 (33). Alternatively, cell fractionation assays could also be conducted (53), with the level of LAMP1 compared between the plasma membrane and cytoplasmic fractions for the different conditions. Mutation of Y18 and Y50 could also be assessed, to determine whether the TMEM106B YXXΦ motifs are required for lysosome exocytosis.

Besides a role in trafficking, Y50 may also mediate the degradation of TMEM106B. The degradation of TMEM106B is not well understood, but TMEM106B is known to undergo RIP at the lysosomal membrane. However, there are two theories regarding how TMEM106B is cleaved. The first is referred to as canonical shedding and involves the release of the TMEM106B C-terminal domain by a resident lysosomal protease at G127 (39). Following this, the N-terminal domain is released by an intramembrane cleavage event close to L106 by the SPPL2a protease. The second theory is referred to as non-canonical shedding, where SPPL2a directs shedding of the C-terminus at S120, which is subsequently followed by intramembrane cleavage to release the N-terminus (35). It is hypothesised that these cleavage events

promote degradation of TMEM106B, with dysregulation of non-canonical shedding potentially leading to the formation of TMEM106B amyloid fibrils (42). In this study, quantification of TMEM106B-WT and TMEM106B-Y50F protein levels in HeLa cells (Figure 15B) found an increase in TMEM106B-Y50F protein compared to TMEM106B-WT, suggesting Y50 may direct protein degradation. Although this increase was not statistically significant, mutation of Y50 was found to increase the amount of cellular TMEM106B in a previous study (82). Similarly, fluorescent imaging in this study suggested TMEM106B-Y50F was brighter than TMEM106B-WT, supporting an increase in TMEM106B-Y50F protein compared to TMEM106B-WT, although quantification of fluorescent intensity was not conducted. Inhibition of protein synthesis in cells, such as via cycloheximide treatment (19), and subsequent determination of TMEM106B protein levels over time, could reveal if Y50 is important in mediating TMEM106B degradation. How this may relate to TMEM106B RIP is unclear, but Y50 might promote recruitment of SPPL2a, or it may be involved in an alternative degradation pathway. Conducting these experiments in the presence and absence of N1-Src could also reveal whether phosphorylation affects TMEM106B degradation.

Conversely, in this study, it was found that TMEM106B expression reduced N1-Src protein levels (Figure 15C), suggesting that increased activation of N1-Src triggered by TMEM106B expression, might promote its degradation. This is thought to occur through the E3 ubiquitin ligase, c-Cbl, where active Src phosphorylates c-Cbl (119). This activates c-Cbl, leading to ubiquitination of both Src and c-Cbl itself (119). Ubiquitination of Src is thought to promote degradation of the protein through the proteasomal pathway, resulting in a feedback loop that is regulated by Src activity. Since active Src is autophosphorylated on Y416, increased N1-Src activity could have been tested by immunoblotting with antibodies against phospho-Src(pY416). If TMEM106B overexpression increases N1-Src activity, a higher level of Y416 phosphorylation would be expected compared to cells without TMEM106B overexpression. Knockdown of c-Cbl coupled with the degradation studies discussed above, could also be conducted to determine if N1-Src degradation is reduced and whether this has a subsequent effect on TMEM106B levels.

4.6 Conclusions and future directions

Within the past decade, TMEM106B has been transformed from a previously unknown protein to one at the forefront of neurological research. Various studies have showcased its involvement in a variety of neurological diseases, yet its cellular function largely remains a mystery. This study has successfully shown that TMEM106B N-terminus is phosphorylated by N1-Src, providing an alternative route for understanding how TMEM106B dysfunction might contribute to disease. Aberrant phosphorylation events have often been implicated in disease, and neurodegeneration is no exception (135). For example, increased activation of Src is thought to promote disease progression in AD (136). In AD, the formation of amyloid-beta ($A\beta$) plaques and increased microglial activation are two hallmarks of the disease. It has been shown that microglia, the immune cell of the brain, can interact with $A\beta$ fibrils triggering increased tyrosine phosphorylation events, through downstream activation of C-Src and other SFK members (137,138). Increased C-Src activation in turn promotes microglial activity, resulting in increased inflammatory responses that contribute to further neurodegeneration (136). As such, it is possible that changes in TMEM106B phosphorylation through aberrant N1-Src signalling, could trigger a cascade of events that ultimately result in diseases such as FTL. However, what these events are specifically, remain the ultimate question.

This study provides a strong starting point, illustrating that phosphorylation does not affect TMEM106B dimerisation, but does reduce binding with AP2M1, suggesting phosphorylation is vital in regulating TMEM106B protein-protein interactions. Further characterisation of AP2M1 and TMEM106B binding in neurons could be achieved through live cell imaging with fluorescently tagged AP2M1 and TMEM106B. Tracking of TMEM106B within the cell body and along axons could reveal when AP2M1 binding occurs, shown through colocalization of the two proteins. Mutation of TMEM106B-Y18 and -Y50, alongside the assessment of phosphorylation via N1-Src knock-down could provide further insight into how this affects the function of TMEM106B. Similar experiments could also be conducted with MAP6 to determine if this interaction is affected by TMEM106B phosphorylation. Investigation of alternative binding partners, such as TMEM106C and potentially LC3 could also reveal new functions of TMEM106B. Similarly, improved understanding of how Y50 regulates protein stability and TMEM106B degradation could provide a link to how TMEM106B amyloid fibrils

are formed in a diseased brain. Altogether, this could shed more light on how TMEM106B functions in the cell and how its dysfunction might contribute to neurodegeneration.

List of abbreviations

µg	Microgram(s)
µM	Micromolar
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid
AP2M1	AP2 adaptor protein µ1 subunit
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
bvFTD	Behavioural variant dementia
C9ORF72	Chromosome 9 open reading frame 72
CHMP2B	Charged multivesicular body protein 2B
CLTC	Clathrin heavy chain
CTCF	CCCTC-binding factor
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle's medium
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EGFR	Epidermal growth factor receptor
ESCRT-III	Endosomal sorting complex required for transport III
FBS	Fetal bovine serum
FSBA	Fluorosulfonylbenzoyl adenosine
FTD	Frontotemporal dementia
FTLD	Frontotemporal lobar degeneration
GFP	Green fluorescent protein
<i>GRN</i>	Progranulin (gene)
GST	Glutathione S-transferase
GWAS	Genome-wide association study

h	Hour(s)
HTP	High-throughput
ICD	Intracellular domain
LB	Lysogeny broth
LEA-2	Late embryogenesis abundant-2
LIR	LC3 interacting region
M	Molar
MAP6	Microtubule associated protein 6
MEME	Multiple Expectation Maximization Algorithms for Motif Elucidation
Min	Minute(s)
miR	MicroRNA
ml	Millilitre(s)
mM	Millimolar
MOG	Myelin oligodendrocyte glycoprotein
N1-Src	Neuronal Src kinase 1
N2-Src	Neuronal Src kinase 2
NMDA	N-methyl-D-aspartate
NTF	N-terminal fragment
°C	Degrees centigrade
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PGRN	Progranulin (protein)
PLP	Proteolipid protein
PMSF	Phenylmethylsulfonyl fluoride
PPA	Primary progressive aphasia
PTP1B	Tyrosine phosphatase 1B
PVDF	Polyvinylidene fluoride
RBS	Ribosomal binding site
RCF	Relative centrifugal force

RI	Refractive index
RIP	Regulated intramembrane proteolysis
RSV	Rous sarcoma virus
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC-MALS	Size-exclusion chromatography with multi-angle light scattering
SFK	Src family of kinases
SNP	Single nucleotide polymorphism
SPR	Surface plasmon resonance
TBS	Tris-buffered saline
TDP-43	Transactivation response DNA-binding protein
TEMED	Tetramethylethylenediamine
TLK1B	Tousled-like kinase 1B
TMEM106B	Transmembrane protein 106B
TNF- α	Tumour necrosis factor alpha
UTR	Untranslated region
UV	Ultraviolet
v/v	Volume per volume
VPS11	Vacuolar protein sorting-associated protein 11
w/v	Weight per volume
WT	Wild-type

List of references

1. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*. 2020 Aug 8;396(10248):413–46.
2. Ratnavalli E, Brayne C, Dawson K, Hodges JR. The prevalence of frontotemporal dementia. *Neurology*. 2002 Jun 11;58(11):1615–21.
3. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;134(Pt 9):2456–77.
4. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011 Mar 15;76(11):1006–14.
5. Ljubenkov PA, Miller BL. A Clinical Guide to Frontotemporal Dementias. *Focus (Am Psychiatr Publ)*. 2016 Oct 13;14(4):448–64.
6. Bennion Callister J, Pickering-Brown SM. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Exp Neurol*. 2014 Dec 1;262(Part B):84–90.
7. Deuschländer AB, Ross OA, Dickson DW, Wszolek ZK. Atypical parkinsonian syndromes: a general neurologist's perspective. *Eur J Neurol*. 2018 Jan 1;25(1):41–58.
8. Halliday G, Bigio EH, Cairns NJ, Neumann M, MacKenzie IRA, Mann DMA. Mechanisms of disease in frontotemporal lobar degeneration: Gain of function versus loss of function effects. *Acta Neuropathol*. 2012 Sep 10;124(3):373–82.
9. Murphy DB, Johnson KA, Borisy GG. Role of tubulin-associated proteins in microtubule nucleation and elongation. *J Mol Biol*. 1977 Nov 25;117(1):33–52.
10. Ayala YM, Pantano S, D'Ambrogio A, Buratti E, Brindisi A, Marchetti C, et al. Human, *Drosophila*, and *C. elegans* TDP43: Nucleic Acid Binding Properties and Splicing Regulatory Function. *J Mol Biol*. 2005 May 6;348(3):575–88.
11. Gao J, Wang L, Huntley ML, Perry G, Wang X. Pathomechanisms of TDP-43 in neurodegeneration. *J Neurochem*. 2018 Jul 1;146(1):7–20.
12. Gass J, Cannon A, Mackenzie IR, Boeve B, Baker M, Adamson J, et al. Mutations in progranulin are a major cause of ubiquitin-positive frontotemporal lobar degeneration. *Hum Mol Genet*. 2006;15(20):2988–3001.
13. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*. 2011 Oct 20;72(2):245–56.
14. van Deerlin VM, Sleiman PMA, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet*. 2010 Mar;42(3):234.

15. van der Zee J, van Langenhove T, Kleinberger G, Slegers K, Engelborghs S, Vandenberghe R, et al. TMEM106B is associated with frontotemporal lobar degeneration in a clinically diagnosed patient cohort. *Brain*. 2011;134(3):808–15.
16. Finch N, Carrasquillo MM, Baker M, Rutherford NJ, Coppola G, DeJesus-Hernandez M, et al. TMEM106B regulates progranulin levels and the penetrance of FTL in GRN mutation carriers. *Neurology*. 2011 Feb 1;76(5):467–74.
17. van Blitterswijk M, Mullen B, Nicholson AM, Bieniek KF, Heckman MG, Baker MC, et al. TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. *Acta Neuropathol*. 2014 Mar 3;127(3):397–406.
18. Cruchaga C, Graff C, Chiang HH, Wang J, Hinrichs AL, Spiegel N, et al. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. *Arch Neurol*. 2011 May;68(5):581–6.
19. Nicholson AM, Finch NA, Wojtas A, Baker MC, Perkerson RB, Castanedes-Casey M, et al. TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia. *J Neurochem*. 2013 Sep 1;126(6):781–91.
20. Chen-Plotkin AS, Unger TL, Gallagher MD, Bill E, Kwong LK, Volpicelli-Daley L, et al. TMEM106B, the Risk Gene for Frontotemporal Dementia, Is Regulated by the microRNA-132/212 Cluster and Affects Progranulin Pathways. *The Journal of Neuroscience*. 2012 Aug 8;32(33):11213.
21. Götzl JK, Mori K, Damme M, Fellerer K, Tahirovic S, Kleinberger G, et al. Common pathobiochemical hallmarks of progranulin-associated frontotemporal lobar degeneration and neuronal ceroid lipofuscinosis. *Acta Neuropathol*. 2014;127(6):845–60.
22. Lang CM, Fellerer K, Schwenk BM, Kuhn PH, Kremmer E, Edbauer D, et al. Membrane Orientation and Subcellular Localization of Transmembrane Protein 106B (TMEM106B), a Major Risk Factor for Frontotemporal Lobar Degeneration. *Journal of Biological Chemistry*. 2012 Jun 1;287(23):19355–65.
23. Brady OA, Zheng Y, Murphy K, Huang M, Hu F. The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function. *Hum Mol Genet*. 2013 Feb 15;22(4):685–95.
24. Stagi M, Klein ZA, Gould TJ, Bewersdorf J, Strittmatter SM. Lysosome size, motility and stress response regulated by fronto-temporal dementia modifier TMEM106B. *Molecular and Cellular Neuroscience*. 2014 Jul 1;61:226–40.
25. Klein ZA, Takahashi H, Ma M, Stagi M, Zhou M, Lam TKT, et al. Loss of TMEM106B Ameliorates Lysosomal and Frontotemporal Dementia-Related Phenotypes in Progranulin-Deficient Mice. *Neuron*. 2017 Jul 19;95(2):281-296.e6.
26. Schwenk BM, Lang CM, Hogg S, Tahirovic S, Orozco D, Rentzsch K, et al. The FTL risk factor TMEM106B and MAP6 control dendritic trafficking of lysosomes. *EMBO J*. 2014 Mar 3;33(5):450–67.
27. Gallagher MD, Posavi M, Huang P, Unger TL, Berlyand Y, Gruenewald AL, et al. A Dementia-Associated Risk Variant near TMEM106B Alters Chromatin Architecture and Gene Expression. *The American Journal of Human Genetics*. 2017 Nov 2;101(5):643–63.

28. Kim S, Yu NK, Kaang BK. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med*. 2015 Jun 5;47(6):e166–e166.
29. Rutherford NJ, Carrasquillo MM, Li M, Bisceglia G, Menke J, Josephs KA, et al. TMEM106B risk variant is implicated in the pathologic presentation of Alzheimer disease. *Neurology*. 2012 Aug 14;79(7):717–8.
30. Hokkanen SRK, Kero M, Kaivola K, Hunter S, Keage HAD, Kiviharju A, et al. Putative risk alleles for LATE-NC with hippocampal sclerosis in population-representative autopsy cohorts. *Brain Pathology*. 2020 Mar 1;30(2):364–72.
31. Rhinn H, Abeliovich A. Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes. *Cell Syst*. 2017 Apr 26;4(4):404–415.e5.
32. Simons C, Dymont D, Bent SJ, Crawford J, D’Hooghe M, Kohlschütter A, et al. A recurrent de novo mutation in TMEM106B causes hypomyelinating leukodystrophy. *Brain*. 2017 Dec 1;140(12):3105–11.
33. Feng T, Sheng RR, Solé-Domènech S, Ullah M, Zhou X, Mendoza CS, et al. A role of the frontotemporal lobar degeneration risk factor TMEM106B in myelination. *Brain*. 2020 Jul 1;143(7):2255–71.
34. Jiang YX, Cao Q, Sawaya MR, Abskharon R, Ge P, DeTure M, et al. Amyloid fibrils in FTL-D-TDP are composed of TMEM106B and not TDP-43. *Nature* 2022 605:7909. 2022 Mar 28;605(7909):304–9.
35. Schweighauser M, Arseni D, Bacioglu M, Huang M, Lövestam S, Shi Y, et al. Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature* 2022 605:7909. 2022 Mar 28;605(7909):310–4.
36. Chang A, Xiang X, Wang J, Lee C, Arakhamia T, Simjanoska M, et al. Homotypic fibrillization of TMEM106B across diverse neurodegenerative diseases. *Cell*. 2022 Apr 14;185(8):1355.
37. Kang J, Lim L, Song J. TMEM106B, a risk factor for FTL and aging, has an intrinsically disordered cytoplasmic domain. *PLoS One*. 2018 Oct 1;13(10):e0205856.
38. Baggen J, Jacquemyn M, Persoons L, Vanstreels E, Pye VE, Wrobel AG, et al. TMEM106B is a receptor mediating ACE2-independent SARS-CoV-2 cell entry. *Cell*. 2023 Jul;186(1):1–16.
39. Brady OA, Zhou X, Hu F. Regulated Intramembrane Proteolysis of the Frontotemporal Lobar Degeneration Risk Factor, TMEM106B, by Signal Peptide Peptidase-like 2a (SPPL2a). *Journal of Biological Chemistry*. 2014 Jul 11;289(28):19670–80.
40. Suzuki H, Matsuoka M. The lysosomal trafficking transmembrane protein 106B is linked to cell death. *Journal of Biological Chemistry*. 2016 Oct 7;291(41):21448–60.
41. Spitz C, Schlosser C, Guschtschin-Schmidt N, Stelzer W, Menig S, Götz A, et al. Non-canonical Shedding of TNF α by SPPL2a Is Determined by the Conformational Flexibility of Its Transmembrane Helix. *iScience*. 2020 Dec 18;23(12):101775.
42. Perneel J, Rademakers R. Identification of TMEM106B amyloid fibrils provides an updated view of TMEM106B biology in health and disease. *Acta Neuropathologica* 2022. 2022 Sep 2;1:1–13.

43. Saftig P, Puertollano R. How Lysosomes Sense, Integrate, and Cope with Stress. *Trends Biochem Sci.* 2021 Feb 1;46(2):97–112.
44. Trivedi PC, Bartlett JJ, Pulinilkunnil T. cells Lysosomal Biology and Function: Modern View of Cellular Debris Bin. *Cells.* 2020;9(5):1131.
45. Ohkuma S, Moriyama Y, Takano T. Identification and characterization of a proton pump on lysosomes by fluorescein isothiocyanate-dextran fluorescence. *Proc Natl Acad Sci U S A.* 1982;79(9 I):2758–62.
46. Braulke T, Bonifacino JS. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research.* 2009 Apr 1;1793(4):605–14.
47. Tomas A, Futter CE, Eden ER. EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol.* 2014 Jan 1;24(1):26–34.
48. Pu J, Guardia CM, Keren-Kaplan T, Bonifacino JS. Mechanisms and functions of lysosome positioning. *J Cell Sci.* 2016 Dec 1;129(23):4329–39.
49. Root J, Merino P, Nuckols A, Johnson M, Kukar T. Lysosome dysfunction as a cause of neurodegenerative diseases: Lessons from frontotemporal dementia and amyotrophic lateral sclerosis. *Neurobiol Dis.* 2021 Jul 1;154:105360.
50. Lüningschrör P, Werner G, Stroobants S, Kakuta S, Dombert B, Sinske D, et al. The FTLD Risk Factor TMEM106B Regulates the Transport of Lysosomes at the Axon Initial Segment of Motoneurons. *Cell Rep.* 2020 Mar 10;30(10):3506–19.
51. Cuveillier C, Delaroche J, Seggio M, Gory-Fauré S, Bosc C, Denarier E, et al. MAP6 is an intraluminal protein that induces neuronal microtubules to coil. *Sci Adv.* 2020 Apr 1;6(14).
52. Gory-Fauré S, Windscheid V, Brocard J, Montessuit S, Tsutsumi R, Denarier E, et al. Non-microtubular localizations of microtubule-associated protein 6 (MAP6). *PLoS One.* 2014 Dec 19;9(12).
53. Kundu ST, Grzeskowiak CL, Fradette JJ, Gibson LA, Rodriguez LB, Creighton CJ, et al. TMEM106B drives lung cancer metastasis by inducing TFEB-dependent lysosome synthesis and secretion of cathepsins. *Nat Commun.* 2018;9(1):2731.
54. Bateman A, Cheung ST, Bennett HPJ. A brief overview of progranulin in health and disease. *Methods in Molecular Biology.* 2018;1806(1):3–15.
55. Tolkatchev D, Malik S, Vinogradova A, Wang P, Chen Z, Xu P, et al. Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities. *Protein Science.* 2008 Apr 1;17(4):711–24.
56. Holler CJ, Taylor G, Deng Q, Kukar T. Intracellular Proteolysis of Progranulin Generates Stable, Lysosomal Granulins that Are Haploinsufficient in Patients with Frontotemporal Dementia Caused by GRN Mutations. *eNeuro.* 2017 Jul 1;4(4):100–17.
57. Paushter DH, Du H, Feng T, Hu F. The lysosomal function of progranulin, a guardian against neurodegeneration. *Acta Neuropathol.* 2018 May 9;136(1):1–17.

58. Wils H, Kleinberger G, Pereson S, Janssens J, Capell A, van Dam D, et al. Cellular ageing, increased mortality and FTLN-TDP-associated neuropathology in progranulin knockout mice. *J Pathol.* 2012 Sep 1;228(1):67–76.
59. Petkau TL, Neal SJ, Milnerwood A, Mew A, Hill AM, Orban P, et al. Synaptic dysfunction in progranulin-deficient mice. *Neurobiol Dis.* 2012 Feb 1;45(2):711–22.
60. Ahmed Z, Sheng H, Xu YF, Lin WL, Innes AE, Gass J, et al. Accelerated Lipofuscinosis and Ubiquitination in Granulin Knockout Mice Suggest a Role for Progranulin in Successful Aging. *Am J Pathol.* 2010 Jul 1;177(1):311–24.
61. Götzl JK, Mori K, Damme M, Fellerer K, Tahirovic S, Kleinberger G, et al. Common pathobiochemical hallmarks of progranulin-associated frontotemporal lobar degeneration and neuronal ceroid lipofuscinosis. *Acta Neuropathol.* 2014;127(6):845–60.
62. Ward ME, Chen R, Huang HY, Ludwig C, Telpoukhovskaia M, Taubes A, et al. Individuals with progranulin haploinsufficiency exhibit features of neuronal ceroid lipofuscinosis. *Sci Transl Med.* 2017 Apr 12;9(385):5642.
63. Holler CJ, Taylor G, Deng Q, Kukar T. Intracellular Proteolysis of Progranulin Generates Stable, Lysosomal Granulins that Are Haploinsufficient in Patients with Frontotemporal Dementia Caused by GRN Mutations. *eNeuro.* 2017 Jul 1;4(4):100–17.
64. Zhou X, Sun L, Brady OA, Murphy KA, Hu F. Elevated TMEM106B levels exaggerate lipofuscin accumulation and lysosomal dysfunction in aged mice with progranulin deficiency. *Acta Neuropathol Commun.* 2017 Jan 26;5(1):9.
65. Arrant AE, Nicholson AM, Zhou X, Rademakers R, Roberson ED. Partial Tmem106b reduction does not correct abnormalities due to progranulin haploinsufficiency. *Mol Neurodegener.* 2018 Jun 22;13(1):1–12.
66. Zhou X, Brooks M, Jiang P, Koga S, Zuberi AR, Baker MC, et al. Loss of Tmem106b exacerbates FTLN pathologies and causes motor deficits in progranulin-deficient mice. *EMBO Rep.* 2020 Oct 5;21(10):e50197.
67. Feng T, Mai S, Roscoe JM, Sheng RR, Ullah M, Zhang J, et al. Loss of TMEM 106B and PGRN leads to severe lysosomal abnormalities and neurodegeneration in mice. *EMBO Rep.* 2020 Oct 5;21(10):e50219.
68. Werner G, Damme M, Schludi M, Gnörich J, Wind K, Fellerer K, et al. Loss of TMEM106B potentiates lysosomal and FTLN-like pathology in progranulin-deficient mice. *EMBO Rep.* 2020 Oct 10;21(10).
69. Clayton EL, Isaacs AM. Progranulin and TMEM106B: when two become wan. *EMBO Rep.* 2020 Oct 5;21(10):e51668.
70. Clayton EL, Milioto C, Muralidharan B, Norona FE, Edgar JR, Soriano A, et al. Frontotemporal dementia causative CHMP2B impairs neuronal endolysosomal traffic-rescue by TMEM106B knockdown. *Brain.* 2018 Dec 1;141(12):3428–42.
71. Jun MH, Han JH, Lee YK, Jang DJ, Kaang BK, Lee JA. TMEM106B, a frontotemporal lobar dementia (FTLN) modifier, associates with FTD-3-linked CHMP2B, a complex of ESCRT-III. *Mol Brain.* 2015 Dec 10;8(1):1–10.

72. Henne WM, Stenmark H, Emr SD. Molecular Mechanisms of the Membrane Sculpting ESCRT Pathway. *Cold Spring Harb Perspect Biol.* 2013 Sep 1;5(9):a016766.
73. Ghazi-Noori S, Froud KE, Mizielińska S, Powell C, Smidak M, Fernandez De Marco M, et al. Progressive neuronal inclusion formation and axonal degeneration in CHMP2B mutant transgenic mice. *Brain.* 2012 Mar 1;135(3):819–32.
74. Guo DZ, Xiao L, Liu YJ, Shen C, Lou HF, Lv Y, et al. Cathepsin D deficiency delays central nervous system myelination by inhibiting proteolipid protein trafficking from late endosome/lysosome to plasma membrane. *Exp Mol Med.* 2018;50(3):457.
75. Zhou X, Nicholson AM, Ren Y, Brooks M, Jiang P, Zuberi A, et al. Loss of TMEM106B leads to myelination deficits: implications for frontotemporal dementia treatment strategies. *Brain.* 2020 Jun 1;143(6):1905–19.
76. Mettlen M, Chen PH, Srinivasan S, Danuser G, Schmid SL. Regulation of Clathrin-Mediated Endocytosis. *Annu Rev Biochem.* 2018 Jun 20;87(1):871–96.
77. Fotin A, Cheng Y, Sliz P, Grigorieff N, Harrison SC, Kirchhausen T, et al. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature.* 2004 Oct 24;432(7017):573–9.
78. Robinson MS. Forty Years of Clathrin-coated Vesicles. *Traffic.* 2015 Dec 1;16(12):1210–38.
79. Nickerson DP, Brett CL, Merz AJ. Vps-C complexes: gatekeepers of endolysosomal traffic. *Curr Opin Cell Biol.* 2009 Aug;21(4):543–51.
80. Baldwin HA, Wang C, Kanfer G, Shah H v, Velayos-Baeza A, Dulovic-Mahlow M, et al. VPS13D promotes peroxisome biogenesis. *Journal of Cell Biology.* 2021;220(5):e202001188.
81. Guillén-Samander A, Leonzino M, Hanna IV MG, Tang N, Shen H, de Camilli P. VPS13D bridges the ER to mitochondria and peroxisomes via Miro. *Journal of Cell Biologycell.* 2021;220(5):e202010004.
82. Busch JI, Unger TL, Jain N, Skrinak RT, Charan RA, Chen-Plotkin AS. Increased expression of the frontotemporal dementia risk factor TMEM106B causes C9orf72-dependent alterations in lysosomes. *Hum Mol Genet.* 2016;25(13):2681–97.
83. Ohno H, Stewart J, Fournier MC, Bosshart H, Rhee I, Miyatake S, et al. Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science.* 1995;269(5232):1872–5.
84. West L. Proteomic and structural analysis of the Neuronal Src SH3 domain. PhD thesis. [York]: University of York; 2019.
85. Thomas SM, Brugge JS. Cellular functions regulated by Src family kinases. *Annu Rev Cell Dev Biol.* 1997;13(1):513–609.
86. Kalia L v., Gingrich JR, Salter MW. Src in synaptic transmission and plasticity. *Oncogene.* 2004 Oct 18;23(48):8007–16.
87. Grant SGN, O’Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER. Impaired long-term potentiation, spatial learning, and hippocampal development in fyn mutant mice. *Science.* 1992;258(5090):1903–10.

88. Ohnishi H, Yamamori S, Ono K, Aoyagi K, Kondo S, Takahashi M. A src family tyrosine kinase inhibits neurotransmitter release from neuronal cells. *Proc Natl Acad Sci U S A*. 2001 Sep 11;98(19):10930–5.
89. Resh MD. Myristylation and palmitoylation of Src family members: The fats of the matter. *Cell*. 1994 Feb 11;76(3):411–3.
90. Resh MD. Interaction of tyrosine kinase oncoproteins with cellular membranes. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 1993 Dec 23;1155(3):307–22.
91. Feng S, Chen JK, Yu H, Simon JA, Schreiber SL. Two Binding Orientations for Peptides to the Src SH3 Domain: Development of a General Model for SH3-Ligand Interactions. *Science* (1979). 1994 Nov 18;266(5188):1241–7.
92. Zhou S, Shoelson SE, Chaudhuri M, Gish G, Pawson T, Haser WG, et al. SH2 domains recognize specific phosphopeptide sequences. *Cell*. 1993 Mar 12;72(5):767–78.
93. Boggan TJ, Eck MJ. Structure and regulation of Src family kinases. *Oncogene*. 2004 Oct 18;23(48):7918–27.
94. Cooper JA, Gould KL, Cartwright CA, Hunter T. Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science*. 1986;231(4744):1431–4.
95. Smart JE, Oppermann H, Czernilofsky AP, Purchio AF, Erikson RL, Bishop JM. Characterization of sites for tyrosine phosphorylation in the transforming protein of Rous sarcoma virus (pp60vsrc) and its normal cellular homologue (pp60c src). *Proc Natl Acad Sci U S A*. 1981;78(10):6013–7.
96. Oppermann H, Levinson AD, Varmus HE, Levintow L, Bishop JM. Uninfected vertebrate cells contain a protein that is closely related to the product of the avian sarcoma virus transforming gene (src). *Proc Natl Acad Sci U S A*. 1979;76(4):1804–8.
97. Collett MS, Erikson E, Purchio AF, Brugge JS, Erikson RL. A normal cell protein similar in structure and function to the avian sarcoma virus transforming gene product. *Proceedings of the National Academy of Sciences*. 1979 Jul 1;76(7):3159–63.
98. Amanchy R, Zhong J, Molina H, Chaerkady R, Iwahori A, Kalume DE, et al. Identification of c-Src Tyrosine Kinase Substrates Using Mass Spectrometry and Peptide Microarrays. *J Proteome Res*. 2008;7(9):3900–10.
99. Ardern H, Sandilands E, Machesky LM, Timpson P, Frame MC, Brunton VG. Src-dependent phosphorylation of Scar1 promotes its association with the Arp2/3 complex. *Cell Motil Cytoskeleton*. 2006 Jan 1;63(1):6–13.
100. Yu XM, Askalan R, Keil GJ, Salter MW. NMDA channel regulation by channel-associated protein tyrosine kinase Src. *Science*. 1997 Jan 31;275(5300):674–8.
101. Scholz R, Berberich S, Rathgeber L, Kollek A, Köhr G, Kornau HC. AMPA Receptor Signaling through BRAG2 and Arf6 Critical for Long-Term Synaptic Depression. *Neuron*. 2010 Jun 10;66(5):768–80.
102. Soriano P, Montgomery C, Geske R, Bradley A. Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. *Cell*. 1991 Feb 22;64(4):693–702.

103. Brugge JS, Cotton PC, Queral AE, Barrett JN, Nonner D, Keane RW. Neurones express high levels of a structurally modified, activated form of pp60c-src. *Nature*. 1985 Aug 1;316(6028):554–7.
104. Levy JB, Dorai T, Wang LH, Brugge' JS. The structurally distinct form of pp60c-src detected in neuronal cells is encoded by a unique c-src mRNA. *Mol Cell Biol*. 1987 Nov;7(11):4142–5.
105. Pyper JM, Bolen JB. Identification of a novel neuronal C-SRC exon expressed in human brain. *Mol Cell Biol*. 1990 May;10(5):2035–40.
106. Keenan S, Lewis PA, Wetherill SJ, Dunning CJR, Evans GJO. The N2-Src neuronal splice variant of C-Src has altered SH3 domain ligand specificity and a higher constitutive activity than N1-Src. *FEBS Lett*. 2015 Dec 31;589(15):1995–2000.
107. Brugge J, Cotton P, Lustig A, Yonemoto W, Lipsich L, Coussens P, et al. Characterization of the altered form of the c-src gene product in neuronal cells. *Genes Dev*. 1987 May 1;1(3):287–96.
108. Keenan S, Wetherill SJ, Ugbode CI, Chawla S, Brackenbury WJ, Evans GJO. Inhibition of N1-Src kinase by a specific SH3 peptide ligand reveals a role for N1-Src in neurite elongation by L1-CAM. *Sci Rep*. 2017 Feb 21;7(1):1–9.
109. Lewis PA, Bradley IC, Pizzey AR, Isaacs H v., Evans GJO. N1-Src Kinase Is Required for Primary Neurogenesis in *Xenopus tropicalis*. *Journal of Neuroscience*. 2017 Aug 30;37(35):8477–85.
110. Altmannova V, Blaha A, Astrinidis S, Reichle H, Weir JR. InteBac: An integrated bacterial and baculovirus expression vector suite. *Protein Sci*. 2021 Jan 1;30(1):108–14.
111. Lewis PA. The role of N-Src kinases in neuronal differentiation. [York]: University of York; 2014.
112. Hornbeck P V., Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res*. 2015 Jan 28;43(D1):D512–20.
113. Amanchy R, Zhong J, Molina H, Chaerkady R, Iwahori A, Kalume DE, et al. Identification of c-Src tyrosine kinase substrates using mass spectrometry and peptide microarrays. *J Proteome Res*. 2008 Sep;7(9):3900–10.
114. Dunning CJR, Black HL, Andrews KL, Davenport EC, Conboy M, Chawla S, et al. Multisite tyrosine phosphorylation of the N-terminus of Mint1/X11 α by Src kinase regulates the trafficking of amyloid precursor protein. *J Neurochem*. 2016 May 6;137(4):518–27.
115. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021 596:7873. 2021 Jul 15;596(7873):583–9.
116. Levine TP. TMEM106B in humans and Vac7 and Tag1 in yeast are predicted to be lipid transfer proteins. *Proteins: Structure, Function and Bioinformatics*. 2022 Jan 1;90(1):164–75.
117. Ricotta D, Conner SD, Schmid SL, Von Figura K, Höning S. Phosphorylation of the AP2 mu subunit by AAK1 mediates high affinity binding to membrane protein sorting signals. *J Cell Biol*. 2002 Mar 4;156(5):791–5.
118. Harris KF, Shoji I, Cooper EM, Kumar S, Oda H, Howley PM. Ubiquitin-mediated degradation of active Src tyrosine kinase. *Proc Natl Acad Sci U S A*. 1999 Nov 11;96(24):13743.

119. Yokouchi M, Kondo T, Sanjay A, Houghton A, Yoshimura A, Komiya S, et al. Src-catalyzed Phosphorylation of c-Cbl Leads to the Interdependent Ubiquitination of Both Proteins. *Journal of Biological Chemistry*. 2001 Sep 14;276(37):35185–93.
120. Kemble DJ, Wang YH, Sun G. Bacterial Expression and Characterization of Catalytic Loop Mutants of Src Protein Tyrosine Kinase †. *ACS Publications*. 2006 Nov;45(49):14749–54.
121. Saya H, Lee PSY, Nishi T, Izawa I, Nakajima M, Gallick GE, et al. Bacterial expression of an active tyrosine kinase from a protein A/truncated c-src fusion protein. *FEBS Lett*. 1993 Jul 26;327(2):224–30.
122. Wang YH, Ayrapetov MK, Lin X, Sun G. A new strategy to produce active human Src from bacteria for biochemical study of its regulation. *Biochem Biophys Res Commun*. 2006 Jul 28;346(2):606–11.
123. Bhoir S, Shaik A, Thiruvengatam V, Kirubakaran S. High yield bacterial expression, purification and characterisation of bioactive Human Trosled-like Kinase 1B involved in cancer. *Scientific Reports* . 2018 Mar;8(1):1–9.
124. Bonangelino CJ, Catlett NL, Weisman LS. Vac7p, a Novel Vacuolar Protein, Is Required for Normal Vacuole Inheritance and Morphology. *Mol Cell Biol*. 2023 Dec 1;17(12):6847–58.
125. Bonifacino JS, Traub LM. Signals for Sorting of Transmembrane Proteins to Endosomes and Lysosomes*. *Annu Rev Biochem*. 2003 Nov 28;72(1):395–447.
126. Kittler JT, Chen G, Kukhtina V, Vahedi-Faridi A, Gu Z, Tretter V, et al. Regulation of synaptic inhibition by phosphodependent binding of the AP2 complex to a YECL motif in the GABAA receptor $\gamma 2$ subunit. *Proc Natl Acad Sci U S A*. 2008 Mar 4;105(9):3616–21.
127. Lee KH, Ho WK, Lee SH. Endocytosis of somatodendritic NCKX2 is regulated by Src family kinasedependent tyrosine phosphorylation. *Front Cell Neurosci*. 2013 Feb 5;7(1):14.
128. Conner SD, Schmid SL. Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J Cell Biol*. 2002 Mar 4;156(5):921–9.
129. Heineman TC, Hall SL. VZV gB Endocytosis and Golgi Localization Are Mediated by YXX ϕ Motifs in Its Cytoplasmic Domain. *Virology*. 2001 Jun 20;285(1):42–9.
130. Lonic A, Onglao W, Khew-Goodall Y. Quantifying EGFR endosomal recycling via immunofluorescence in breast cancer cells. *STAR Protoc*. 2022 Jun 17;3(2):101305.
131. Blagojević Zagorac G, Mahmutefendić H, Maćešić S, Karleuša L, Lučin P. Quantitative Analysis of Endocytic Recycling of Membrane Proteins by Monoclonal Antibody-Based Recycling Assays. *J Cell Physiol*. 2017 Mar 1;232(3):463–76.
132. Arancibia-Cárcamo IL, Fairfax BP, Moss SJ, Kittler JT. Studying the Localization, Surface Stability and Endocytosis of Neurotransmitter Receptors by Antibody Labeling and Biotinylation Approaches. *The Dynamic Synapse: Molecular Methods in Ionotropic Receptor Biology*. 2006 Jan 1;1:91–118.
133. Johansen T, Lamark T. Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors. *J Mol Biol*. 2020 Jan 3;432(1):80–103.

134. Shpilka T, Weidberg H, Pietrokovski S, Elazar Z. Atg8: An autophagy-related ubiquitin-like protein family. *Genome Biol.* 2011 Jul 27;12(7):1–11.
135. Tenreiro S, Eckermann K, Outeiro TF. Protein phosphorylation in neurodegeneration: friend or foe? *Front Mol Neurosci.* 2014 May 13;7(42).
136. Portugal CC, Almeida TO, Socodato R, Relvas ao B, C Portugal CC, Socodato R, et al. Src family kinases (SFks): critical regulators of microglial homeostatic functions and neurodegeneration in Parkinson's and Alzheimer's diseases. *FEBS J.* 2022 Dec 1;289(24):7760–75.
137. Wood JG, Zinsmeister P. Tyrosine phosphorylation systems in Alzheimer's disease pathology. *Neurosci Lett.* 1991 Jan 2;121(1–2):12–6.
138. Koenigsknecht J, Landreth G. Microglial Phagocytosis of Fibrillar β -Amyloid through a β 1 Integrin-Dependent Mechanism. *Journal of Neuroscience.* 2004 Nov 3;24(44):9838–46.