Understanding and Exploiting Peroxisomal Import Mechanisms

Sophie Louise Moul

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Astbury Centre for Structural Molecular Biology

March 2021

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

This work would not have been possible without the constant support and enthusiasm of my two supervisors Stuart Warriner and Alison Baker. Stuart, thank you for constant excitement and optimism about the project, and about life in general. Alison, thank you for your helpful feedback on all my work throughout the project and your sound advice for experimental protocol. The input from both of you has ensured that I have been able to produce a piece of work I can be very proud of.

Thank you to all the members of the Schrader lab at the University of Exeter for allowing me to visit and test my probes on peroxisomes in cells. You made me feel very welcome and stimulated great peroxisome discussion.

Thank you to Paul Beales, Darren Tomlinson, Richard Foster, Adrian Whitehouse and Thomas Edwards for allowing me to carry out rotations in your labs and gain a multitude of skills that have helped me throughout my PhD.

Thank you to the Wellcome Trust for funding the project and in particular extending that funding to allow me to complete my lab work after six months of mandatory working from home due to the Coronavirus pandemic.

Thanks to all the members of Lab 1.49, with whom I have crossed paths with during my time here. In particular, I must thank Gemma for her help and advice in the early stages of the project and Zoe, Holly and Devon for keeping things fun.

To my family, without whom I never would have gotten this far. Thank you for all your love and support. Although I'm not sure you understand what it is I have been doing, you were always there when I needed you. I can always rely on my sisters to keep me laughing.

Abstract

The peroxisome is a near-ubiquitous organelle in eukaryotes, which must import proteins to carry out its functions. Most peroxisomal matrix proteins are targeted by a peroxisome targeting signal (PTS1) at the C-terminus. This tripeptide is recognised by the import receptor protein PEX5, which shuttles the matrix protein across the peroxisomal membrane.

This work reports the design and synthesis of bifunctional chemical probes, which can bind to PEX5 via a PTS1 peptide and bind a second protein, which represents a cargo protein. Labelling a protein with a PTS1 in this way is a completely novel concept. These probes are used to investigate two different aspects of peroxisomal matrix protein import.

Chapter 2 addresses the question of whether a protein can be labelled with a PTS1 via the probes and subsequently be imported into the peroxisome by PEX5. The development of such a system could be used to trigger peroxisomal import of protein not native to the peroxisome as a tool to study protein-protein interactions and investigate the effects of protein knockdown. *In vitro* pulldown assays successfully demonstrated the interaction of probe-labelled proteins with PEX5 proteins. Cell-based assays investigated if using these probes can result in the labelled protein becoming resident in the peroxisomal matrix. Further optimisation is required to validate their function *in vivo*.

Chapter 3 develops a method to create a form of PEX5, which is permanently bound to a cargo protein. This would be a useful tool to elucidate the structure of peroxisomal matrix protein import complex and investigate the mechanisms of cargo unloading across the peroxisomal membrane. The probes interact with PEX5 via a PTS1, and subsequently become covalently bound to a cysteine on PEX5 to create a permanent PEX5-cargo complex. The probes were extended to include a cargo protein-binding motif and a protein complex of PEX5 with a covalently bound cargo protein was formed.

Together these investigations show the utility of chemical probes and how they can be used to investigate and manipulate peroxisome import mechanisms.

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations

Amino Acids

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

Organisms

At	*Arabidopsis thaliana*
Hs	*Homo sapiens*
Sc	*Saccharomyces Cerevisiae*
Tb	*Trypanosoma Brucei*

Units

'	Minutes
"	Seconds
°C	Degrees Celsius
Da	Daltons
G	Gravitational force
kDa	Kilodaltons
kpsi	Kilo Pounds per Square Inch
L	Litre
µL	Microliter
µM	Micromolar
mL	Millilitre
mM	Millimolar
nM	Nanomolar
V	Volts

General

2-HPCL	2-hydroxyphytanoyl-CoA Lyase
AAA-ATPase	ATPases associated with diverse cellular Activities
ABCB1	ATP binding cassette subfamily B Member 1
ABPP	Activity Based Protein Profiling
Ac	Acyl
ACOX	Acyl-CoA oxidase
AGT	Glyoxylate Amino Transferase
AIM	Auto Induction Media
AMACR	2-Methylacyl-CoA racemase

Ant1p	Adenine nucleotide transporter 1
ATP	Adenosine Triphosphate
BCFA	Branched Chain Fatty Acid
BIA	Benzylisoquinoline Alkaloid
BLAST	Basic Local Alignment Search Tool
Boc	tert-butyloxycarbonyl
CMV	Cytomegalovirus
Cl-Trt	Chlorotrityl
CoA	Coenzyme A
CPP	Cell-penetrating Peptide
DCM	Dichloromethane
Dde	1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl
DEAE-D	Diethyaminoethyl Dextran
DHA	Docosahexaenoic Acid
DhaA	Haloalkane Dehalogenase
DHAP	Dihydroxyacetone Phosphate
DHAPAT	Dihydroxyacetone Phosphate Acyl Transferase
DHCA	$3\alpha,7\alpha$-dihydroxycholestanoic acid
DIC	N,N′-Diisopropylcarbodiimide
DIPEA	N,N-Diisopropylethylamine
DMAP	4-Dimethylaminopyrimidine
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DNA	Deoxyribonucleic Acid
DTM	Docking/Translocation Module
DTT	Dithiothreitol
EC_{50}	Half Maximal Effective Concentration
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetracetic Acid
EGFP	Enhanced Green Fluorescent Protein

ER	Endoplasmic Reticulum
FA	Fluorescence Anisotropy
FAD	Flavin Adenine Dinucleotide
FALDH	Fatty Acid Aldehyde Dehydrogenase
Fmoc	Fluorenylmethoxycarbonyl
H_2O_2	Hydrogen Peroxide
HAT	Human African Trypanosomiasis
hAGT	Human O^6-alkylguanine-DNA-alkyltransferase
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-Performance Liquid Chromatography
HRP	Horseradish Peroxidase
HT	HaloTag Substrate
ICL	Isocitrate Lyase
IPTG	Isopropyl-ß-D-thiogalactopyranoside
k_D	Dissociation Constant
KOD	Thermococcus kodakaraensis KOD1
LB	Luria Broth
LC-MS	Liquid Chromatography- Mass Spectrometry
LCS	Long-Chain acyl-CoA Synthetase
MLS	Malate Synthase
NAD	Nicotinamide Adenine Dinucleotide
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCS	Norcoclaurine Synthase
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide Gel Electrophoresis
PAHX	Phytanoyl-CoA Hydroxylase
Pbf	Pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl
PCR	Polymerase Chain Reaction

PBD	Peroxisome Biogenesis Disorder
PDB	Protein Data Bank
PBS	Phosphate Buffered Saline
PED	Single Peroxisomal Enzyme Deficiency
PEG	Polyethylene Glycol
PHA	Polyhydroxyalkanoate
PEX	Peroxin
PMP	Peroxisomal Membrane Protein
PROTACs	Proteolysis Targeting Chimeras
PTS	Peroxisome Targeting Signal
PVDF	Polyvinylidene fluoride
QD	Quantum Dot
REM	Receptor Export Module
RING	Really Interesting New Gene
RNA	Ribonucleic Acid
RNAi	Ribonucleic Acid Interference
RNS	Reactive Nitrogen Species
ROS	Reactive Oxygen Species
SAXS	Small Angle X-ray Scattering
SCPx	Sterol Carrier Protein X
SCX	Strong Cation Exchange
SDS	Sodium Dodecyl Sulfate
S_NAr	Aromatic nucleophilic substitution
SRP	Signal Recognition Particle
TBS	Tris-Buffered Saline
TBS-T	Tris-Buffered Saline with Tween-20
tBu	Tertiary Butyl
TEA	Triethylamine
TFA	Trifluroacetic Acid
THA	Tetracosahexaenoic Acid

THCA	Trihydroxycholestanoic Acid
THF	Tetrahydrofuran
TIPS	Triisopropylsilane
TPR	Tetratricopeptide Repeat
Tris	Tris(hydroxymethyl)aminomethane
Trt	Trityl
USP	Ubiquitin Specific Protease
VLCFA	Very Long Chain Fatty Acid
WB	Western Blot
WT	Wild Type
ZSD	Zellweger Spectrum Disorder

Chapter 1

Introduction

1.1 The Peroxisome

Peroxisomes are organelles found in the large majority of eukaryotic cells and play an essential role in a variety of cellular processes (Lazarow and Fujiki, 1985). They are pleiomorphic organelles that range from 0.2–1 µM in diameter (Lazarow and Fujiki, 1985) and are bound by a single membrane bilayer. Peroxisomes are one of the most protein-dense compartments of the cell and can contain a diverse complement of proteins, which can vary based on the species and cellular environment in which they are present (Waterham et al., 2016).

Peroxisomes can arise *de novo* from the endoplasmic reticulum (ER) as well as multiply by growth and division of existing organelles (Agrawal and Subramani, 2016). Peroxisome biogenesis is a matter of continued debate in the field and a topic of ongoing research, though it is now accepted that both pathways play a role (see: 1.2). Some proteins and lipids that form the peroxisomal membrane bud from the ER to form empty pre-peroxisomal vesicles, which are then complemented with matrix proteins and some membrane proteins imported from the cytoplasm, which carry out the functions of the peroxisome (Gardner et al., 2018, van der Zand et al., 2012, Sugiura et al., 2017).

The term peroxisome actually encompasses a family of microbodies originally named for their specific enzymatic function but since found to share mechanisms of biogenesis (Pracharoenwattana and Smith, 2008). This means that although the enzymatic activities that take place within the peroxisomes from different organisms are not universal, the mechanisms employed to generate peroxisomes are largely conserved. Trypanosomes contain glycosomes, peroxisomes that perform glycolysis, a function not seen in the peroxisomes of other organisms, but also carry out metabolic functions common to other peroxisomes (Michels et al., 2006, Opperdoes, 1984). Glyoxysomes, found in plants and fungi but not in mammals, are specialised peroxisomes that contain enzymes to perform the glyoxylate cycle, a process which converts lipids to sugar and is particularly important during seed germination and early growth (Pracharoenwattana and Smith, 2008, Cooper and Beevers, 1969).

Proteins involved in the biogenesis of peroxisome membranes and the import of their contents are encoded by the *PEROXIN (PEX)* genes and are known as

peroxins. In both mammals and plants, the proteins are denoted as PEXN, whereas in yeast they are known as PexNp proteins.

Here, the focus of this work is largely on the mammalian peroxisome and their major metabolic functions are discussed in the following section. Due to the similarities in biogenesis mechanisms of peroxisomes across all organisms, it is likely that the chemical tools developed in this work could be more widely applied to peroxisomes from a diverse variety of organisms.

1.1.1 Mammalian Peroxisome Functions

The functions of peroxisomes are wide ranging, and it has been seen that peroxisomes will adapt their functions based on their environment. This does not only mean that the proteome composition varies based on cell type but also that they can dynamically modify their activities, size, and number based on the current growth conditions or if the cell becomes stressed. For example, when provided with the fatty acid oleate as an energy source, yeast cells will dramatically increase their peroxisome numbers to increase their capacity to metabolise oleate, an activity that occurs inside peroxisomes (Veenhuis et al., 1987). In a similar way, electron micrographs of certain yeasts grown on methanol, which requires the peroxisome for metabolism, show enlarged peroxisomes in comparison to those grown on ethanol, which does not require the peroxisome for metabolism (Fukui et al., 1975). Many so-called "peroxisome-proliferators" have been identified that stimulate the production of peroxisomes to meet an increasing demand (Schrader et al., 2013, Latruffe and Vamecq, 1997, Fukui et al., 1975, Manjithaya et al., 2010). When such stimuli are removed, the number of peroxisomes can be seen to decrease again in response to the reduced need for them by the cell (Manjithaya et al., 2010). The majority of peroxisomes metabolise reactive oxygen species but the specific functions in particular organism and cell types are wide-ranging (Wanders and Waterham, 2006a, Lazarow and Fujiki, 1985, del Río et al., 1992).

The peroxisome was named from the discovery of what was originally thought to be its main function: the metabolism of hydrogen peroxide (H_2O_2). It was found that H_2O_2-producing oxidases and catalase for H_2O_2 metabolism were found in peroxisomes that had been isolated by density gradient centrifugation (De Duve and Baudhuin, 1966). Since this initial finding, multiple avenues of investigation pointed to a much larger role for the peroxisome, that of lipid metabolism as well as lipid synthesis in the mammalian peroxisome (Hajra and Das, 1996, Lazarow and De Duve, 1976)

Mammalian peroxisomes are mainly responsible for the alpha-oxidation of branched chain fatty acids (BCFAs), the beta-oxidation of very long chain (more than 22 carbons) fatty acids (VLCFAs), the synthesis of ether-linked phospholipids such as those that coat the myelin sheath of neurons, the synthesis of bile acids (Ferdinandusse et al., 2009) and docosahexaenoic acid (DHA), an important compound in brain and retina development (Paker et al., 2010) and glyoxylate detoxification. It is the action of some of the enzymes involved in these processes that results in the production of reactive oxygen species (ROS), such as H_2O_2 , which are subsequently metabolised in the peroxisome (Fransen et al., 2012). The main functions of mammalian peroxisomes are summarised in Figure 1.1.
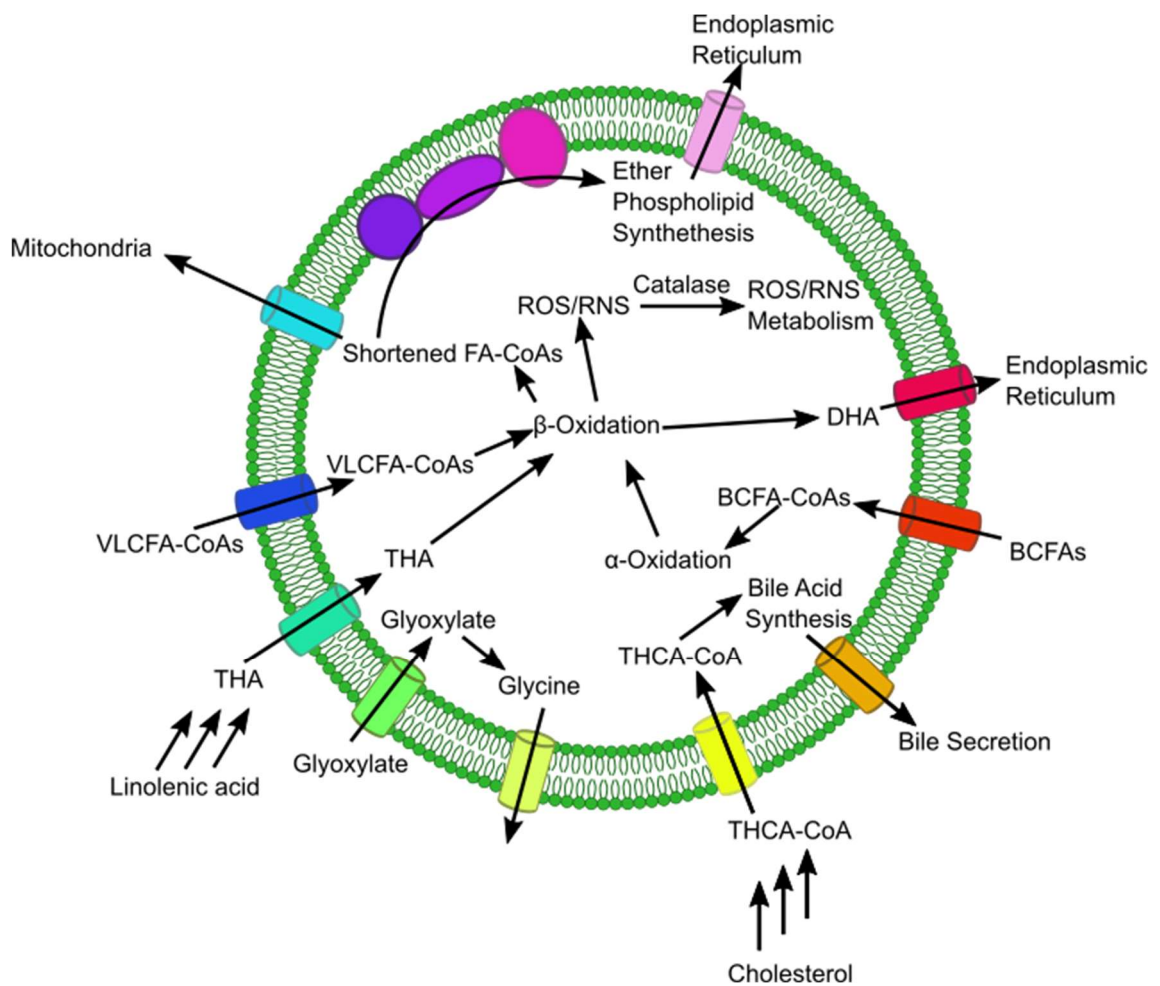
Figure 1.1 Major Functions of the Mammalian Peroxisome A schematic summary of the main functions carried out by the mammalian peroxisome. Abbreviations: ROS- Reactive Oxygen Species; RNS- Reactive Nitrogen Species; DHA- docosahexaenoic acid; BCFA- Branched Chain Fatty Acid; CoA-Coenzyme A; THCA- trihydroxycholestanoic acid; THA-tetracosahexaenoic acid; VLCFA- Very Long Chain Fatty Acid; FA- Fatty Acid

1.1.2 Metabolite Transport

The matrix proteins that carry out the functions of the peroxisome are imported by a specific mechanism (see section 1.3) but the substrates on which these enzymes act require different transport pathways. The peroxisome membrane is permeable to small metabolites of less than approximately 300-400 Da as they can passively diffuse through a pore (Van Veldhoven et al., 1987). In mammals, Pxmp2 has been identified as the protein in the peroxisomal membrane to form this size-excluding pore which allows the passage of small metabolites, such as H_2O_2 and glyoxylate. It is impermeable to larger species such as cofactors required for enzymatic reactions, which require their own transport systems (Rokka et al., 2009, Antonenkov and Hiltunen, 2006, Antonenkov et al., 2004, DeLoache et al., 2016).

For fatty acids to be metabolised in the peroxisome they must be transported across the peroxisomal membrane. This is facilitated by three half-ABC transporters in the D subfamily, which form homodimers with different substrate preferences (van Roermund et al., 2008, van Roermund et al., 2011, van Roermund et al., 2014). ABCD1 is the primary transporter for VLCFAs whereas ABCD2 is the transporter for shorter and polyunsaturated VLCFAs, such as tetracosahexaenoic acid (THA) for the synthesis of the structural fatty acid DHA (Fourcade et al., 2009, Paker et al., 2010, van Roermund et al., 2011). The intermediates for bile acid synthesis as well as other FAs with branched chains are imported by ABCD3 (van Roermund et al., 2014, Ferdinandusse et al., 2014). Although each transporter appears to have a preference for the substrates it imports into the peroxisomal matrix, there is some overlap. ABCD1-3 transport fatty acid cargoes as acyl-CoA esters with the CoA group hydrolysed as part of the import process, as shown by assaying thioesterase activity of the transporters (Okamoto et al., 2018, De Marcos Lousa et al., 2013, Kawaguchi et al., 2021). They must then be re-esterified to be able to undergo beta-oxidation. This process requires the presence of ATP and CoA (Scheme 1.1) within the peroxisomal matrix, which must also be imported. NAD and FAD are required for beta-oxidation reactions (see: 1.1.3) and so must also be imported into the peroxisome. Much investigation has gone into identifying the transporters of the cofactors and substrates required for the reactions that occur inside the peroxisome, and many are yet to be identified. Adenine nucleotide transporter 1 (Ant1p) has been identified in yeast as a transporter of ATP across the peroxisomal membrane (Palmieri et al., 2001, van Roermund et al., 2001). In mammals PMP34 shares a large amount of sequence homology with Ant1p but studies remain inconclusive as to which substrates PMP34 can

transport across the membrane of a mammalian peroxisome (van Roermund et al., 2001, Van Veldhoven et al., 2020, Visser et al., 2002, Agrimi et al., 2012, Kim et al., 2019).

It is known that intermediates metabolised by the enzymes within the peroxisome must cross the membrane, but the exact facilitators of this for many metabolites remains largely unknown (Kunze et al., 2006, Van Veldhoven et al., 2020).

1.1.3 Fatty Acid Beta-Oxidation

The beta-oxidation of fatty acids is a biochemical pathway that is carried out by almost all peroxisomes across different species and cells types (Waterham et al., 2016) and is a central activity for both the anabolic and catabolic functions performed by the peroxisome. In humans, beta-oxidation occurs in both peroxisomes and mitochondria (Wanders and Waterham, 2006a, Drysdale and Lardy, 1953), whereas the process is exclusively peroxisomal in some organisms such as yeast and plants. Beta oxidation of Very Long-Chain Fatty Acids (VLCFAs), phytanic acid (see 1.1.4) and intermediates in the synthesis of bile acids occurs exclusively in peroxisomes (Wanders et al., 2010). Beta oxidation is essential to produce docosahexaenoic acid (DHA) from dietary linolenic acid (Ferdinandusse et al., 2001a). As DHA is an abundant structural component in the brain, impairment of beta oxidation to produce DHA contributes to neurological problems seen in patients with peroxisomal diseases (see section 1.5). The production of DHA is a coordinated process between the peroxisome and endoplasmic reticulum, although the exact mechanisms of exchange of intermediates between the two organelles is not known (Ferdinandusse et al., 2001a, Wanders et al., 2016).

Beta-oxidation reduces the length of a fatty acid substrate by two carbons though a four-step mechanism of dehydrogenation, hydration, dehydrogenation, and thiolytic cleavage (Scheme 1.1).

1. Dehydrogenation   **2. Hydration**

3. Dehydrogenation   **4. Thiolytic Cleavage**
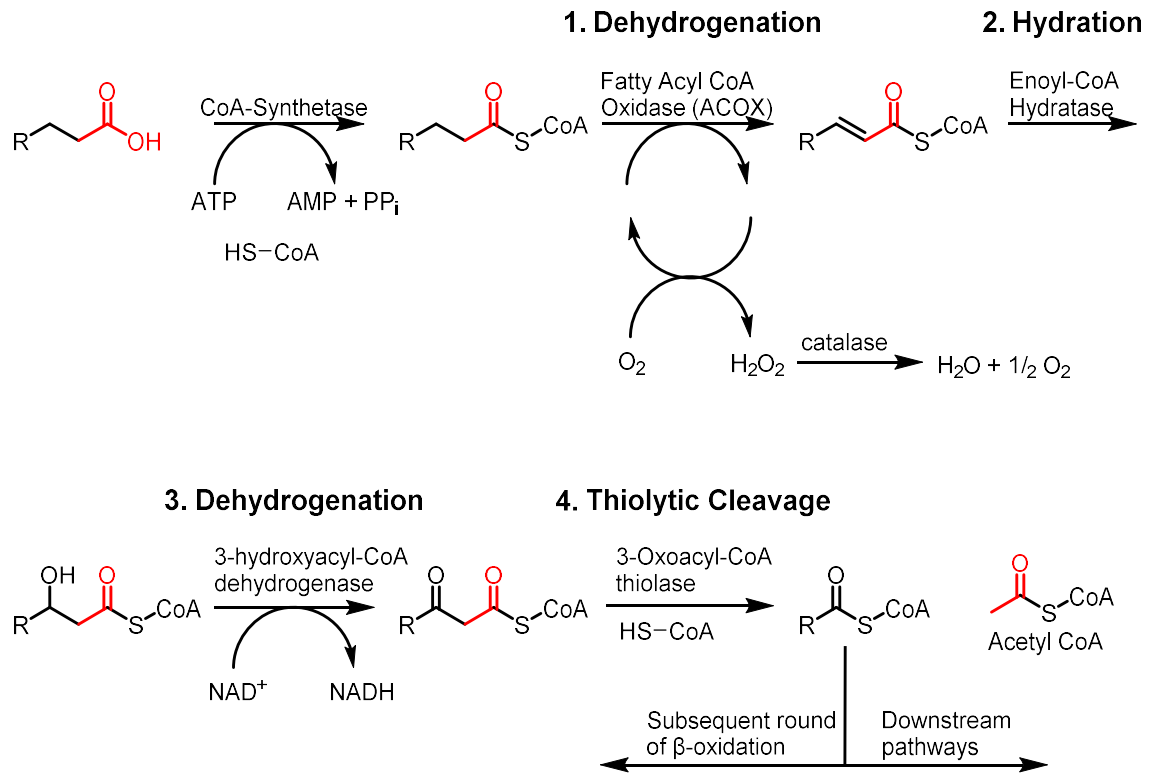
Scheme 1.1 Summary of reactions in peroxisomal beta-oxidation

The beta-oxidation pathway is important for the overall health of an organism as it acts to generate substances the cell requires as well as breaking down substances that would be detrimental to health if allowed to accumulate (see: 1.5). The peroxisome does not act alone in these functions and cross talk between other cellular compartments also plays an important part in the proper functioning of the cell.

1.1.4 Fatty Acid Alpha-Oxidation

In order to carry out beta-oxidation, some branch-chained fatty acids (BCFAs) must first undergo alpha-oxidation to produce compatible FAs with a methyl group at the 2 position (Wanders and Waterham, 2006a). Phytanic acid, which is obtained from the diet, has been shown to have detrimental effects on health if it accumulates (van den Brink and Wanders, 2006) and is the most prominently studied BCFA that undergoes alpha oxidation in the peroxisome. It is initially converted to phytanoyl-CoA by long-chain acyl-CoA synthetase (LCS) at the cytoplasmic face of the peroxisomal membrane (Miyazawa et al., 1985). Next phytanoyl-CoA is hydroxylated to 2-hydroxyphytanoyl-CoA by phytanoyl-CoA hydroxylase (PAHX) (Mihalik et al., 1995).

2-hydroxyphytanoyl-CoA lyase (2-HPCL) then cleaves 2-hydroxyphytanoyl-CoA into a 2-methyl aldehyde and formyl-CoA, which is ultimately converted to formic acid (Foulon et al., 1999). The 2 methyl aldehyde formed can then be

converted to a fatty acid, which in the case of phytanic acid alpha-oxidation, is pristanic acid (Verhoeven et al., 1997). This is catalysed by a fatty acid aldehyde dehydrogenase (FALDH), of which evidence suggests there is more than one capable of catalysing this reaction within the peroxisome (Jansen et al., 2001). This process is summarised in Scheme 1.2.
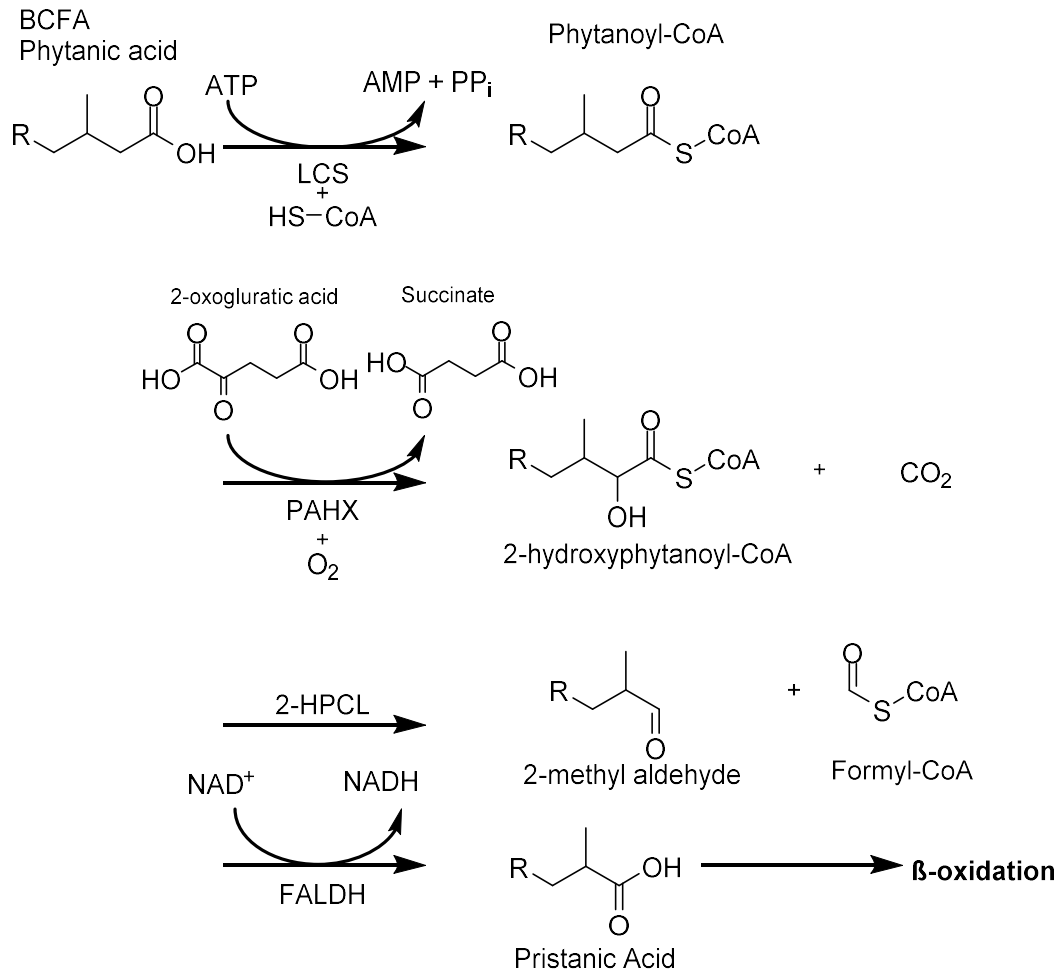
Scheme 1.2 Alpha Oxidation of Phytanic Acid to Pristanic Acid

1.1.5 Ether phospholipid Biosynthesis

In mammals, ether phospholipid biosynthesis is a vital function of peroxisomes. Ether phospholipids, the most abundant of which are the plasmalogens, play an important role in the development of bones, eyes and the brain and dysfunction of these organs is seen in many peroxisomal disorders (Wanders and Brites, 2010)(see: 1.5). Ether phospholipids are partially synthesised inside the peroxisome from acyl-CoAs and dihydroxyacetone phosphate (DHAP) (Hajra and Bishop, 1982). In this reaction the acyl group is replaced by an alkyl chain from a fatty acid, linked by an ether bond, hence the name ether phospholipid (Dean and Lodhi, 2018). The alkyl-DHAP is then reduced before being exported to the endoplasmic reticulum to complete the synthesis of the ether phospholipid. This process is summarised in Scheme 1.3.
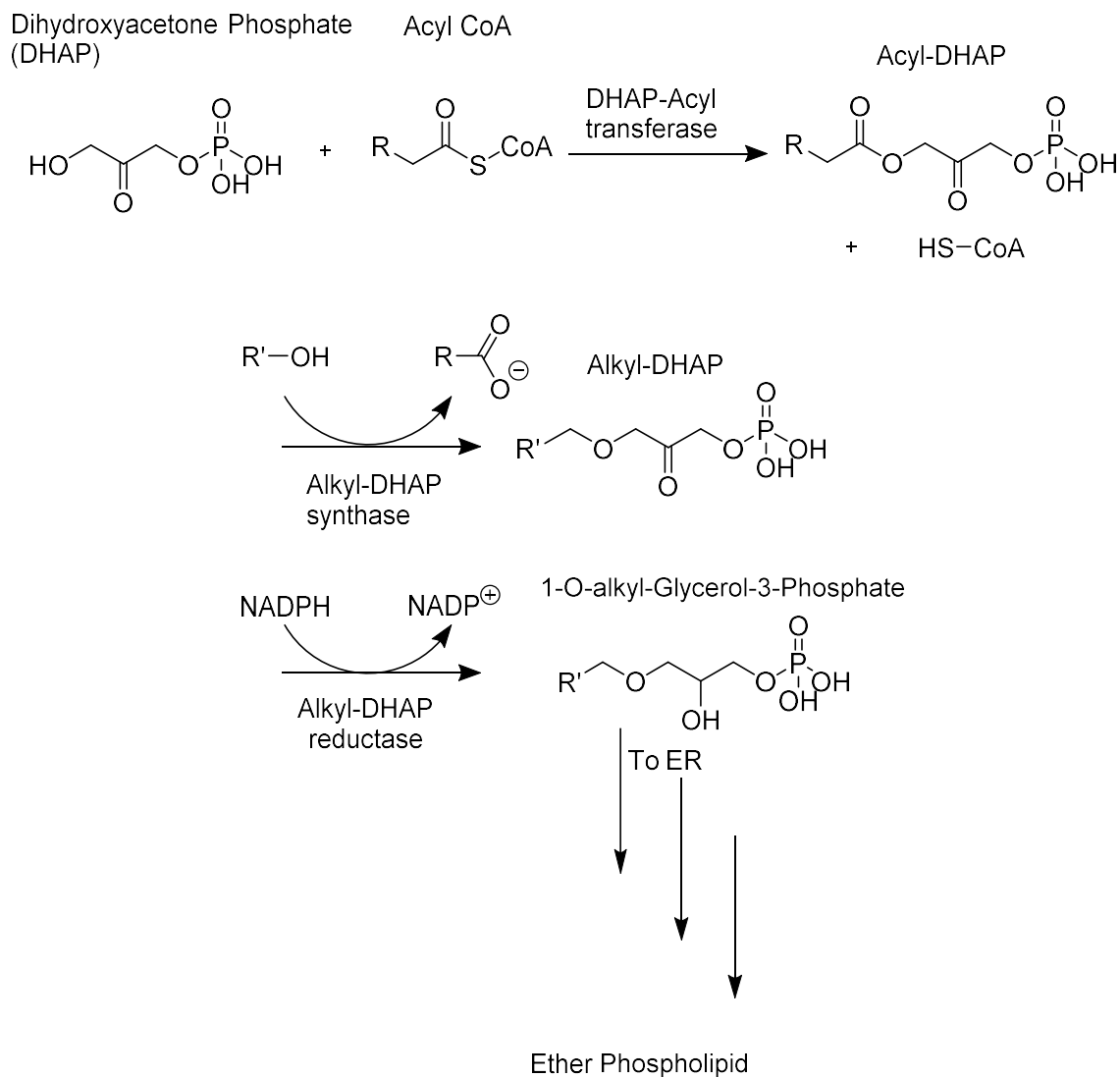
Scheme 1.3 Peroxisomal steps of ether phospholipid biosynthesis

1.1.6 Synthesis of Bile Acids

Bile acids are produced from the breakdown of cholesterol and are the main constituent of bile, which breaks down lipids in the small intestine (Russell, 2003). Intermediates from the breakdown of cholesterol in other parts of the cell undergo beta-oxidation in the peroxisome followed by the conjugation of either glycine or taurine to produce the final bile acid product, which is secreted out of the peroxisome and the cell and into bile (Ferdinandusse et al., 2009). Dysregulation of the enzymes in this process can contribute to problems observed in patients with peroxisomal disorders (see: 1.5).

1.1.7 Glyoxylate Detoxification

Peroxisomal enzymes facilitate the breakdown of glyoxylate into glycine in humans (Danpure and Jennings, 1986). If this process does not occur then glyoxylate is oxidised to oxalate, which ultimately precipitates as calcium oxalate crystals in the kidneys and urinary tract (Danpure, 2005). These deposits become lethal when they begin to accumulate in multiple tissues in diseases in which enzymes that breakdown glyoxylate are mutated (Williams et al., 2009). The precursors that result in the production of glyoxylate in the body are not yet fully understood (Knight et al., 2006), but its breakdown has been shown to be a vital function of peroxisomes found in liver cells (Danpure and Jennings, 1986).

1.2 Peroxisome Membrane Biogenesis

As stated earlier, the number and size of peroxisomes in any cell can alter in response to stimuli. This is controlled by biogenesis and degradation mechanisms. Peroxisomes form either through growth and division of existing peroxisomes or through *de novo* biogenesis.

The peroxins PEX3 and PEX19 are essential for peroxisome biogenesis across all organisms with PEX16 also playing a role in mammalian cells. This is demonstrated by experiments that show that patient cell lines from individuals with peroxisomal disorders (see: 1.5) usually contain so-called "peroxisome ghosts", peroxisomal membranes that do not contain the matrix proteins (Santos et al., 1988). However, in cell lines in which *PEX3, PEX16* or *PEX19* are the mutated genes, no peroxisomal membranes are formed (Shimozawa et al., 2004).

Through studying mammalian cells lines that lack peroxisomes, it has been found that some peroxisomal membrane proteins are localised to ER membranes (Toro et al., 2009). The current model of peroxisomal membrane import into the peroxisomal membrane postulates two pathways, PMPs that traffic to the peroxisome via the ER, and PMPs that traffic directly from the cytoplasm to the peroxisomal membrane (South et al., 2001, Titorenko and Rachubinski, 1998, Titorenko and Rachubinski, 2001)

Experimental evidence points to a number of different mechanisms which contribute to the formation of a mature and functional peroxisome. It is known that PEX19 and PEX3 are essential for the import of PMPs into the membrane but there has been much interest in elucidating how the membrane resident PEX3 reaches the peroxisomal membrane to begin with. Several different

routes have been demonstrated and it is possible that all these pathways play a role, particularly in *de novo* peroxisome biogenesis (Figure 1.2).

PEX19 acts as a shuttle to deliver PMPs from the cytoplasm to the peroxisomal membrane, where it interacts with PEX3 to deliver the PMP (Figure 1.2A). PEX19 binds to a membrane-targeting sequence (m-PTS) on the PMP and delivers it to the membrane where PEX3 facilitates its membrane import (Jones et al., 2004). In mammals, PEX16 plays an important role in two of the pathways thought to assist incorporation of PEX3 into the peroxisomal membrane. First, PEX3 has been shown to be directly imported into peroxisomal membranes from the cytoplasm in a PEX19- and PEX16-dependent manner in CHO-K1 and HEK293 cells (Matsuzaki and Fujiki, 2008) (Figure 1.2B). In this mechanism PEX19 appears to be necessary but does not bind to PEX3 in the same way that it binds to m-PTS sequences on other PMPs, this suggests that PEX19 plays a role in stabilising PEX3 to prevent it aggregating in the cytoplasm and acts as a chaperone in a way that has yet to be determined. PEX16 is thought to act as the membrane-resident receptor of PEX3 (Matsuzaki and Fujiki, 2008).

PEX16 is also found in the endoplasmic reticulum (ER) membrane and another mechanism that appears to contribute to some PMPs, including PEX3, becoming resident in peroxisomal membrane is a route in which they are trafficked via the ER (Kim et al., 2006) (Figure 1.2C). Evidence in mammalian cells has shown that PEX16 is inserted into the ER membrane co-translationally with the assistance of the signal recognition particle (SRP), which transports the nascent protein chain on a ribosome to the SEC61 translocon at the ER membrane (Kim et al., 2006) (Figure 1.2D). PEX16 can then recruit PMPs including PEX3 and PMP34, a peroxisomal ATP transporter (Visser et al., 2002), to the ER membrane (Kim et al., 2006) (Figure 1.2C). Evidence has shown that these PMPs use both pathways (direct and via-ER) to reach peroxisomal membranes and the only exception appears to be PEX16, which appears to always traffic to the peroxisome via the ER (Aranovich et al., 2014).

Of these two mechanisms of PEX3 import to peroxisomes, kinetic studies have shown that direct import is, perhaps unsurprisingly, faster than import via the ER (Aranovich et al., 2014). This suggests that it would be favourable for the direct import mechanism to dominate and this appears to be the case of both PEX3 and PMP34 when examined in the Aranovich et al. study. However, it also appears that the via-ER pathway may be preferred. It has been shown that, when overexpressed, the rate of import of PEX3 and PMP34 was dependent on the levels of PEX16. The authors propose that the pathway used is dependent

on the ratio of the levels of PMPs to PEX16. This means that the addition of exogenously expressed PEX16 caused an increase in PMP import via the ER, suggesting that this pathway is preferred, and direct import is only used when there is insufficient PEX16 relative to other PMPs. The reliance on PEX16 is a conundrum and at present it is not possible to say if these results truly represent what happens in cells in a mammalian organism as all the data has been gathered *in vitro* or in cell culture. Other organisms such as yeast do not have a PEX16 protein. Recently, a new peroxin was identified in *Komagataella phaffii* (formally *Pichia pastoris*) that appears to be a functional homolog of the mammalian PEX16 protein (Farré et al., 2017). The two proteins lack sequence homology, but their predicted structures are similar, and they can complement the function of one another in cell-based assays (Farré et al., 2017).

In addition to its recruitment to ER by PEX16, PEX3 has also been seen to be co-translationally imported into the ER by the SEC61 translocon (Mayerhofer et al., 2016) (Figure 1.2E). This represents another pathway for PEX3 to reach peroxisomes, though this pathway is possibly redundant in cells and has only been demonstrated in an *in vitro* system. It remains to be seen if this pathway is required for the proper maintenance of peroxisome biogenesis or if post-translational import of PEX3 to the ER or directly to peroxisomes is sufficient the majority of the time.

In the via ER pathway, PMPs traffic to a specialised ER subdomain which has been visualised by electron microscopy (Geuze et al., 2003) (Figure 1.2F). Similar ER subdomains have been identified in other species (Joshi et al., 2016), but how similar these are, considering the disparity on other aspects of peroxisome biogenesis, has not yet been determined. Mayerhofer et al. demonstrated that vesicles containing PEX3 budded from the ER and traffic to peroxisomes (Mayerhofer et al., 2016) and a separate study found pre-peroxisomes containing PEX3 and PEX16 which mature to subsequently contain other PMPs such as PEX14 over a number of days (Schmidt et al., 2012) providing evidence that vesicles bud from the ER to provide lipids and membrane proteins to pre-existing peroxisomes as well as generate new peroxisomes *de novo* (Figure 1.2H).
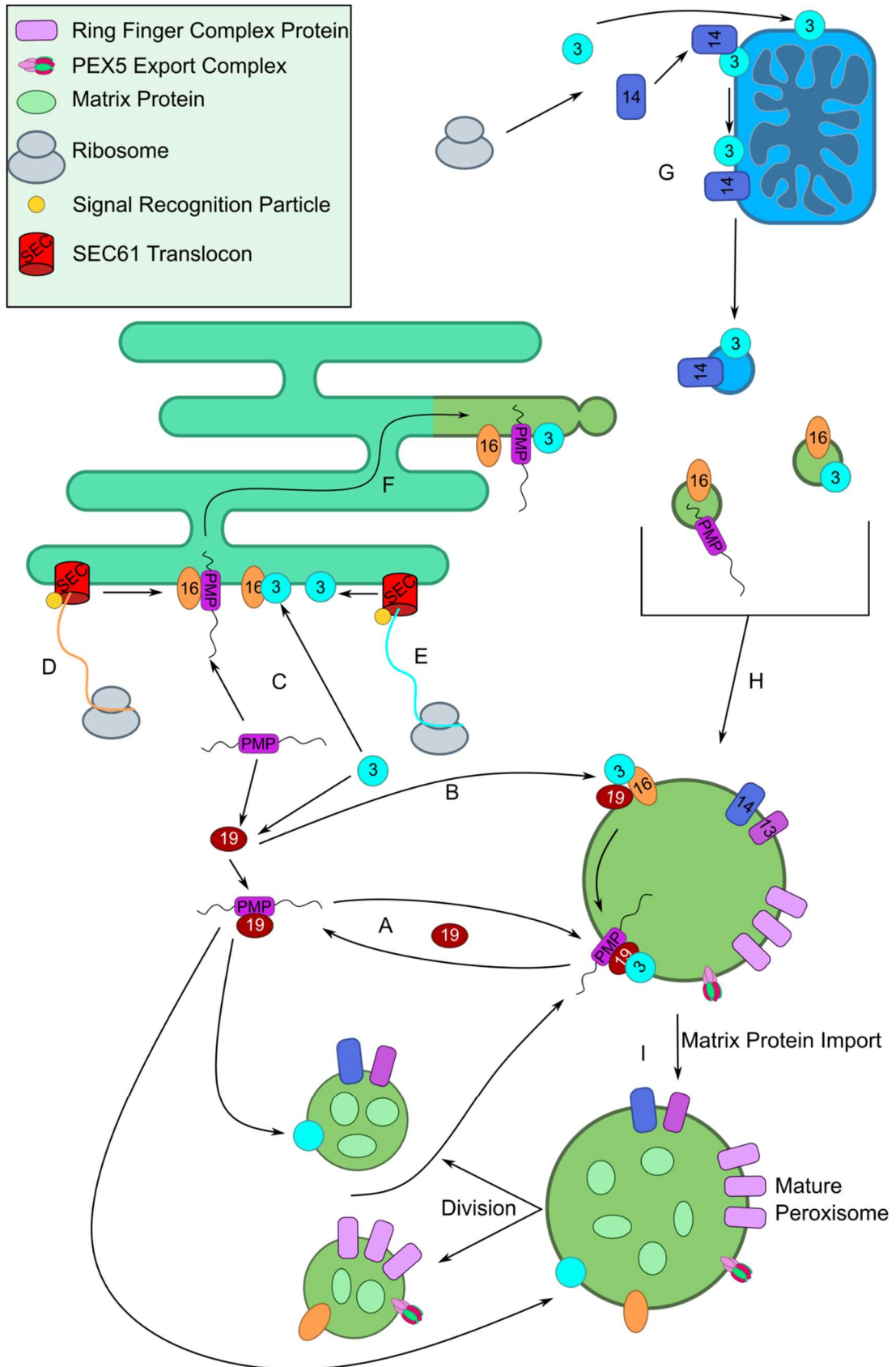
Figure 1.2 Summary of Peroxisome Biogenesis in Mammalian Cells The major stages in the biogenesis of peroxisomal membrane with their PMPs are depicted. Where a process has been discussed in the text, it is assigned a letter for clarity.

One other recent development in the import of PMPs to peroxisomes is the identification of a pathway in which they are imported via mitochondria (Sugiura et al., 2017). Peroxisomes and mitochondria are intimately connected in the biochemical pathways they are involved in (see: 1.1.1). Recent work has shown that in mutant cells lacking peroxisomes (lacking either PEX3 or PEX16), when infected with adenovirus that allowed the expression of the missing peroxin, both PEX3 and PEX14 inserted into the mitochondrial membrane (Sugiura et al., 2017) (Figure 1.2G). Vesicles containing PEX3 and PEX14 were seen to bud from the mitochondria and fuse with ER-derived vesicles containing PEX16. This mechanism only appears to be in *de novo* peroxisome biogenesis, as once import competent peroxisomes had been generated, PEX3 and PEX14 are no longer targeted to the mitochondria (Sugiura et al., 2017).

Once the peroxisomal membrane has been complemented with all the PMPs, it is then able to import matrix proteins and carry out its biochemical functions (Figure 1.2I).

A good understanding of the mechanisms involved in the formation of peroxisomes both *de novo* and through growth and division helps to better understand the disease phenotypes that occur when certain peroxins are mutated. A fuller picture of the processes involved across different organisms will help to better advance our ability to manipulate peroxisomes to treat disease.

1.3 Peroxisome Matrix Protein Import

Peroxisomes do not contain their own DNA and therefore proteins resident in the peroxisomal lumen (matrix proteins) must be imported post-translationally (Lazarow and Fujiki, 1985). These proteins are imported in most cases by one of two mechanisms through peroxisomal targeting sequences (PTSs) (see sections 1.3.1 and 1.3.2). In some yeast species, a third PTS has been proposed. The nature of this "PTS3" is yet to be unequivocally elucidated and is still an active area of research to understand the import of some matrix proteins (Kempiński et al., 2020).

The import mechanism of matrix proteins across the peroxisomal membrane is still not fully understood and structural data of the translocon has been difficult to obtain. Although there is heterogeneity in the PEX proteins between species, the import mechanisms are thought to be largely conserved. It has been shown that fully folded monomeric and oligomeric proteins (Freitas et al., 2011, Lanyon-Hogg et al., 2010) can be imported into peroxisomes and some proteins can "piggyback" their way into the peroxisome through interactions with proteins

tagged for import (Effelsberg et al., 2015, Islinger et al., 2009). This indicates that the import machinery must be able to adapt to various shapes and sizes in order to accommodate the range of cargo that could be brought to the membrane for import.

1.3.1 PTS1 Import

PTS1 sequences occur in the majority of the proteins found in the peroxisomal matrix in mammals and are by far the most common form of peroxisomal targeting sequence across all species. The PTS1s are a family of tripeptide sequences found at the protein's carboxyl-terminus fitting the consensus of [small]-[basic]-[hydrophobic] amino acid side chains, such as Serine-Lysine-Leucine (Brocard and Hartig, 2006, Gould et al., 1989). This peptide sequence is recognised and bound by the PTS1 receptor protein PEX5 (Fransen et al., 1995, Brocard et al., 1994). It is important that PTS1 sequences and therefore the C-terminus of the protein is sufficiently exposed to allow PEX5 to bind (Brocard and Hartig, 2006). The importance of the nature of the amino acid side chains is demonstrated in crystal structures of the PTS1 peptide bound to the PEX5 receptor protein (See section 1.3.3).

Despite these initial findings for the consensus sequence, peroxisomal matrix proteins do not always conform to this; the most notable of these is catalase, which is imported into peroxisomes via a non-canonical PTS1 sequence in all species in which it has been studied (Williams et al., 2012). In humans it is a C-terminal tetrapeptide KANL (Purdue and Lazarow, 1996) that was found to be vital for its peroxisomal targeting. Studies suggest that the use of a non-canonical PTS1 in catalase, which has a lower binding affinity and import efficiency for PEX5 than other peroxisomal matrix enzymes and is less efficiently imported into peroxisomes (Koepke et al., 2007, Legakis et al., 2002, Maynard et al., 2004), could confer an advantage to the cell, such as allowing catalase to properly fold before import (Williams et al., 2012) and to promote the retention of cytosolic catalase during times of cellular stress (Walton et al., 2017).

On the other hand, investigations into PTS1 signalling found the addition of a canonical –SKL peptide to the C-terminus of some proteins is not always sufficient to trigger import into the peroxisome (Distel et al., 1992), first suggesting that the PTS1-PEX5 interaction often relies on more than the tripeptide motif and that sequences that appear to act as PTS1 signals in one species are not necessarily transferrable to others (Kragler et al., 1993). Conversely, some proteins have been seen to be imported into the peroxisome

by PEX5 when expressed without their PTS1 (Parkes et al., 2003), though it is possible that such proteins could "piggyback" into peroxisomes with other PTS1 proteins.

Residues upstream of the PTS have also now been found to enhance binding to the PTS1 receptor protein PEX5 (Brocard and Hartig, 2006). There has been much work put into studying the binding affinities of longer synthetic peptides derived from the C-termini of peroxisomal matrix proteins (Maynard et al., 2004) as well as structural studies into further contacts with PEX5 made by full length PTS1 proteins (Stanley et al., 2006, Hagen et al., 2015). The PTS1 consensus, although still largely reliant on terminal tripeptide, can be extended to include the terminal 12 residues of the protein, with residues -12 to -8 acting as a linker to allow the tripeptide to be accessible to PEX5 (Neuberger et al., 2003). The residues upstream of terminal tripeptide have been seen to interact with PEX5 and help to enhance the cargo protein's interaction with the import protein (Fodor et al., 2012). It appears at present that there is no clear consensus for additional contacts of a PTS1 cargo protein with PEX5 as for some proteins the region immediately upstream of the PTS1 plays a role, whereas in others a more distant region of the sequence is critical for efficient peroxisomal import (Stanley et al., 2006). It has also been postulated that additional binding sites to PEX5 by the cargo protein help to ensure that it is properly folded before import into the peroxisome (Stanley et al., 2006). As improperly folded proteins can aggregate in the peroxisome to the detriment of the cell (Williams et al., 2012), such a quality control function in the PTS1-PEX5 interaction could confer an evolutionary advantage. The nature of these upstream residues seems to be variable between species and proteins that use a non-canonical PTS1 appear to rely more heavily on these additional PEX5 interactions (Lametschwandtner et al., 1998, Neuberger et al., 2003, Chowdhary et al., 2012).

Evidence points towards a system of "targeting priority" in the import of matrix proteins into the peroxisome. It has been shown that some proteins will outcompete rivals if PEX5 is scarce by using PTS1 sequences with stronger PEX5 affinity or utilise an additional binding site on PEX5 (Rosenthal et al., 2020). Additional binding sites of a cargo protein to PEX5 have been shown to enhance peroxisomal targeting as proteins with the same strong PTS1 of SKL do not have the same affinity for PEX5 (Rosenthal et al., 2020).

1.3.2 PTS2 Import

The second pathway for the import of peroxisomal matrix proteins uses PEX7 to interact with an N-terminal PTS2 sequence with the consensus nonapeptide [R/K]-[L/V/I]-X_5-[H/Q]-[L/A] (Kunze et al., 2011).The PTS2 signal is used by a much smaller proportion of proteins destined for the peroxisomal matrix with only a handful identified in mammals, most notably 3-Ketoacyl-CoA thiolase, an enzyme necessary for the final step of the β-oxidation pathway (Legakis and Terlecky, 2001, Otera et al., 1998). Interestingly, at least 20% of peroxisomal matrix proteins in *Arabidopsis thaliana* have been found to contain a PTS2 sequence (Reumann et al., 2009) and this could in part be due to the fact that the identified peroxisomal proteome in plants is much larger than that in mammals and other classes of organisms (Emanuelsson et al., 2003).

After PEX7 has identified and bound to a PTS2 sequence via a conserved groove that interacts with the PTS2 motif (Pan et al., 2013, Kunze et al., 2011), PEX7 requires co-receptors to dock with the peroxisomal membrane import complex (Cross et al., 2017). In *A. thaliana*, PEX5 contains an N-terminal PEX7 interaction domain, and through this, PTS2-containing proteins can be imported into the peroxisome in the same manner as PTS1 proteins (Woodward and Bartel, 2005). Similarly in mammals, PEX7 is recognised by the PEX5L isoform of PEX5 (Braverman et al., 1998) . In other species, a different protein, or proteins act as the PEX7 co-receptor (Kunze et al., 2011). *C. elegans* does not produce proteins with PTS2 targeting signals, its peroxisomal matrix proteins are only imported via PTS1 sequences (Motley et al., 2000).

1.3.3 Mechanisms of Import

The ability of the peroxisomal import system to import folded proteins, protein oligomers and heterogeneous complexes of proteins means that a unique mechanism of translocation must be required, different to that employed by other organelles in cell. The exact mechanisms employed by the proteins involved in the import of peroxisomal matrix proteins are still not conclusively defined and a number of models have been proposed. These models have been based on *in vitro* biochemical mechanistic studies, cell-based assays as well as some structural data.

There are several different proteins involved in the process of peroxisomal matrix protein import and there are species-species variations, which means that the mechanism becomes ever more complex to elucidate. The peroxins involved in the import of matrix proteins into peroxisomes in mammals are summarised in Table 1.1.

Table 1.1 Summary of the protein involved in peroxisomal matrix protein import

Mechanistic Step	Protein	Function
Cargo Binding	PEX5S	Binds to PTS1 sequences on proteins in the cytosol and facilitates their import across the peroxisomal membrane (Fransen et al., 1995)
	PEX5L	The long isoform of PEX5S containing an additional exon encoding the PEX7 binding domain (Braverman et al., 1998)
	PEX7	Binds to PTS2 sequences on proteins in the cytosol and facilitates their import across the peroxisomal membrane by binding to PEX5L (Kunze et al., 2011)
Membrane Interaction	PEX13	Interacts with both PEX5 and PEX14 to aid membrane docking and the formation of the peroxisomal translocon (Schell-Steven et al., 2005)
	PEX14	Interacts strongly with diaromatic motifs on PEX5 to aid membrane docking and the formation of the peroxisomal translocon (Saidowsky et al., 2001)
Receptor Ubiquitination	PEX2	Acts as E3 ubiquitin ligase on PEX5 (Platta et al., 2009)
	PEX10	Acts as E3 ubiquitin ligase on PEX5 (Platta et al., 2009)
	PEX12	Acts as E3 ubiquitin ligase on PEX5 (Platta et al., 2009)
	UbcH5a/b/c	E2 ubiquitin conjugation to ubiquitinate Cys-11 on PEX5 (Grou et al., 2008)
Receptor Export and Recycling	PEX1	Forms an AAA-ATPase with PEX6 to extract PEX5 from the peroxisomal membrane (Platta et al., 2005)

	PEX6	Forms an AAA-ATPase with PEX1 to extract PEX5 from the peroxisomal membrane (Platta et al., 2005)
	PEX26	Recruits and anchors PEX1/PEX6 AAA-ATPase in the peroxisomal membrane (Matsumoto et al., 2003)
	Usp9X	Removes ubiquitin from PEX5 after its export back to the cytosol (Grou et al., 2012)
	AWP1	Binds to PEX6 and ubiquitinated PEX5 in the export complex to aid export (Miyata et al., 2012)

In brief, PEX5S or PEX5L binds to a PTS1 cargo in the cytosol, or PEX7 binds to a PTS2 cargo with PEX7 subsequently bound by PEX5L. The protein complex is then transported to the peroxisome membrane. PEX5 interacts with PEX14 and PEX13 and inserts into the membrane. Cargo is transported across the membrane and unloaded into the matrix of the peroxisome. PEX5 is then recycled back to the cytosol. For this to happen, PEX5 is monoubiquitinated through the action of PEX2, PEX10 and PEX12 and exported via an AAA-ATPase complex formed by PEX1 and PEX6. PEX5 is then deubiquitinated by Usp9x and free to collect new cargo proteins.

Figure 1.3 summarises the process of PTS1 matrix protein import into peroxisomes. More details of the interactions made, and discussion of the mechanisms employed will be covered in the sections to follow. Here, the focus will remain primarily on the mechanisms found in mammalian cells as this is overall focus of the work presented. Much of the characterisation work has been carried out in other organisms such as yeast, which gives an insight into how the mechanism could work globally.
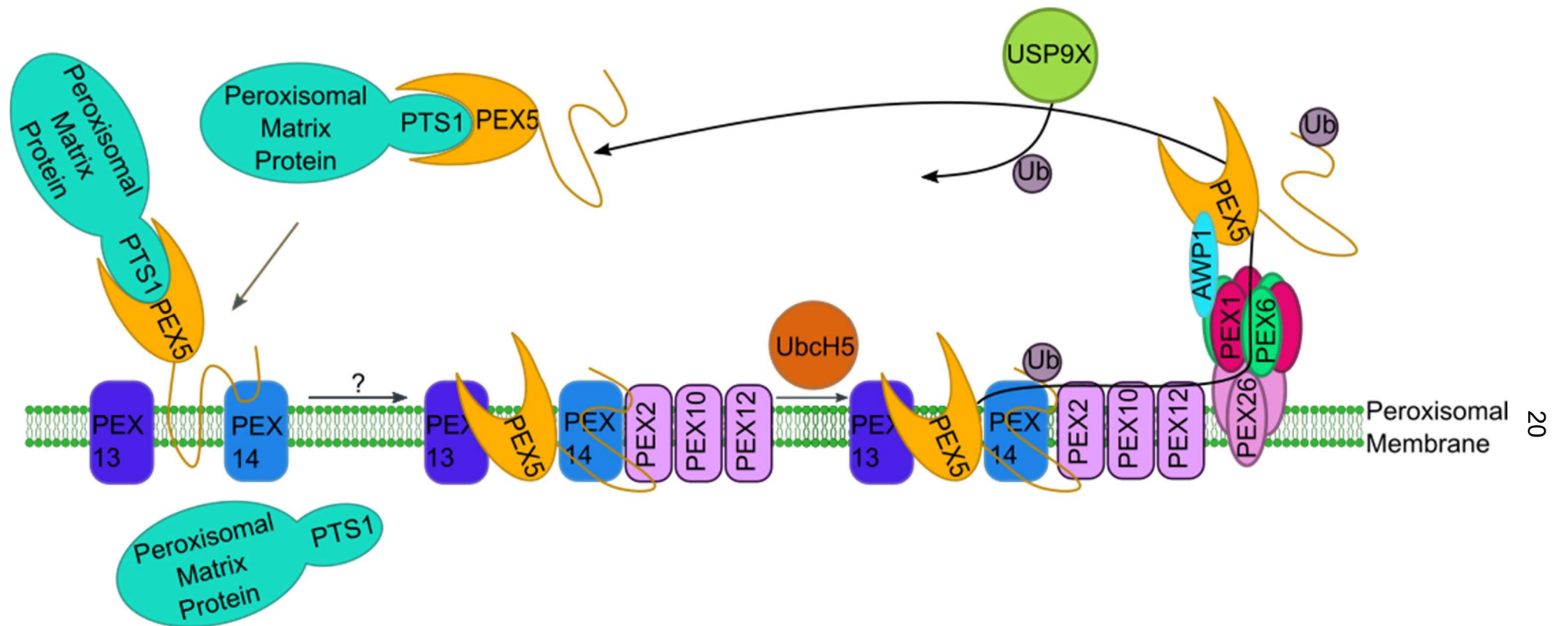
Figure 1.3 A summary of the mechanism used to import peroxisomal matrix proteins with a PTS1 As discussed in the text, a protein with a PTS1 is recognised in the cytosol by PEX5, which shuttles the protein to the membrane and several peroxins facilitate its import across the peroxisome membrane. PEX5 is then ubiquitinated, targeting it for export from the membrane and back to the cytosol, where it is deubiquitinated and able to collect further PTS1 cargo.

1.3.3.1 PEX5, the import receptor

PEX5 is the protein responsible for recognising PTS1 sequences on proteins and facilitates their transport to the peroxisomal membrane and subsequent import (Grou et al., 2009a, Wolf et al., 2010, Lanyon-Hogg et al., 2010). The full length protein is 70 kDa. PEX5 has two distinct domains that play different roles in its function. These are summarised in Figure 1.4.
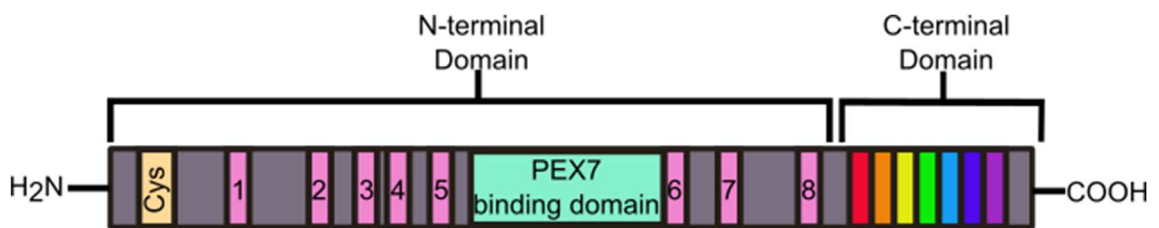
Figure 1.4 Schematic of domains and motifs in mammalian PEX5 The PEX5 protein consists of a highly conserved globular C-terminal domain made up of seven TPR domains (colouring consistent to Figure 1.5) and an intrinsically disordered N-terminal domain. In mammals there are eight diaromatic pentapeptide motifs (pink) which bind to PEX14 during peroxisomal matrix protein import by PEX5. Also, within the N-terminal domain in the PEX5L isoform is the PEX7 binding domain (mint green), which assists import of PTS2 proteins in mammals. A conserved cysteine residue at the N-terminus of PEX5 (peach) is ubiquitinated as part of the recycling of PEX5 back to the cytoplasm after protein import.

1.3.3.2 Cargo Binding

The C terminal domain of PEX5 contains seven tetratricopeptide repeat (TPR) domains. TPRs are a common structural feature that can mediate interactions between proteins. The PEX5 TPRs are composed of 34 amino acid repeats that form an alpha helix with clusters of TPRs having a helix-turn-helix arrangement (D'Andrea and Regan, 2003). The PEX5 TPRs form a funnel type structure with two sets of TPRs, linked by TPR4, with PTS1 peptides bound in the cavity between the two sets (Figure 1.5) (Gatto Jr et al., 2000). The crystal structure of the *Homo sapiens* PEX5 C-terminal domain bound to the PTS1 peptide YQSKL has been solved and shows the interactions made by PEX5 with the peptide (Figure 1.5). Sequence data was used to determine the position of each TPR (Dodt et al., 1995).
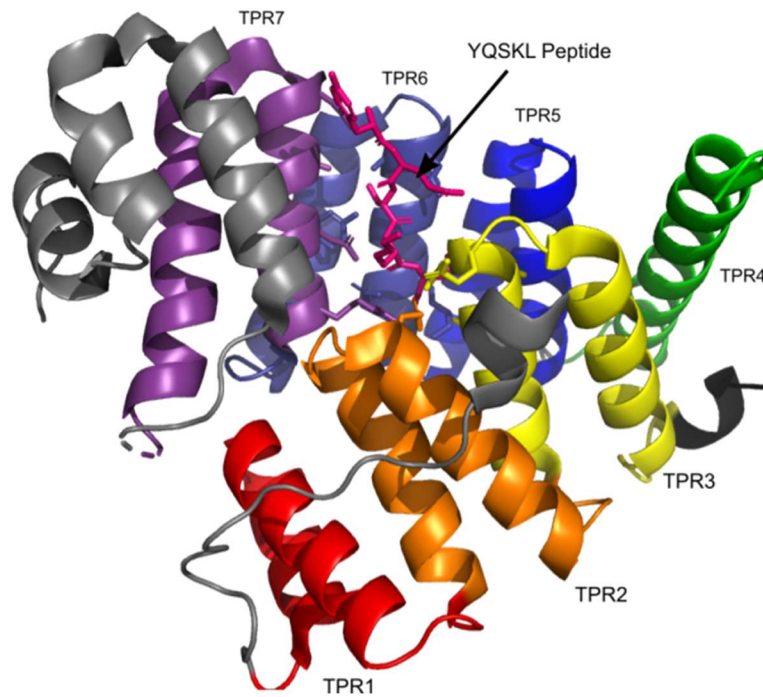
Figure 1.5 X-ray Crystal Structure of PEX5C from *Homo sapiens* interacting with PTS1 peptide The Protein Data Bank (PDB) file 1FCH crystal structure shows the C-terminal domain of PEX5 from *Homo sapiens* interacting with the peptide YQSKL, a PTS1 sequence. The seven TPR domain of the protein are coloured differently, and this shows the clusters of TPRs 1-3 and TPRs 5-7 connected by TPR4. TPR4 was not fully resolved in the crystal structure, represented by the black area and the lack of connection between TPRs 3 and 4. The PTS1 peptide YQSKL is shown in magenta and the residues it was found to interact with are displayed. The interacting residues are localised in the majority on TPRs 6 and 7, with some interactions on TPRs 2 and 3 (Gatto Jr et al., 2000).

Residues in the TPR domains interact with the peptide backbone and side chains of the PTS1 tripeptide motif and show the importance of the motif fitting the consensus of [small]-[basic]-[hydrophobic] side chains (Figure 1.6). The terminal carboxyl group of the PTS1 is also involved in making interactions with the TPR domain, explaining why the PTS1 sequence must be present at the C-terminus (Figure 1.6A) (Gatto Jr et al., 2000) PEX5 residues Thr377, Val374, Lys490 and Ala493 form a hydrophobic pocket around the leucine side chain (Figure 1.6A). Asn378 forms an interaction with the peptide backbone. Asn378 and Asn489 both form interactions with the carboxylate group at the C-terminus of the peptide, exemplifying that the PTS1 sequence must have the free carboxylate to interact, and therefore the sequence must be at the C-terminus of the protein destined for peroxisomal import. Arg520 and Lys490 also make interactions with the terminal carboxylate.

Asn524 interacts with the peptide backbone of the lysine residue at the -2 position (Figure 1.6B). The positively charged nitrogen on the lysine residue forms a hydrogen bond with a water molecule which allows interactions with Glu379 and Glu348.

Asn497 interacts with the peptide backbone of the serine residue (Figure 1.6C). The hydroxyl on the serine is hydrogen bonded to a water molecule which allows interactions with the hydroxyl oxygens on Tyr508 and Ser528 and the carbonyl oxygen in the backbone of
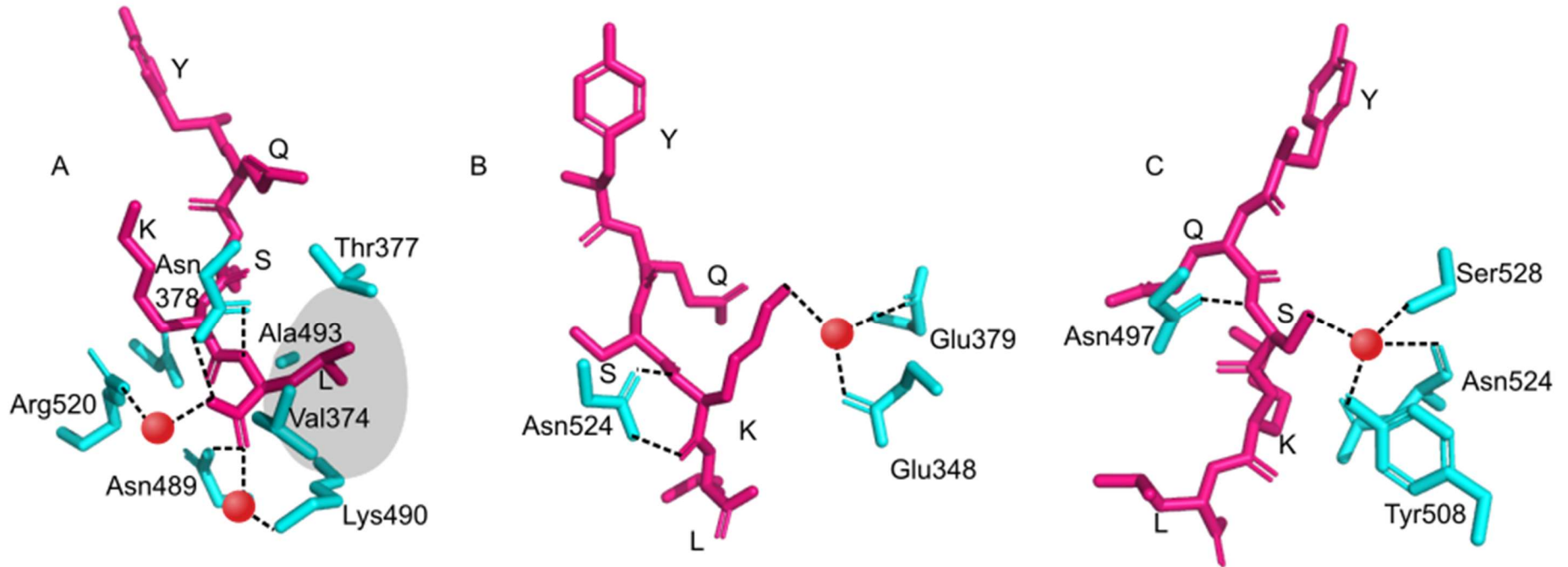
Figure 1.6 Interactions of amino acid residues from *Homo sapiens* PEX5C with the PTS1 peptide YQSKL A. The residues from *Homo sapiens* PEX5C (cyan) that interact with the leucine residue (L) of the PTS1 peptide are displayed. Thr377, Val374, Lys490 and Ala493 form a hydrophobic pocket around the leucine side chain (grey shaded area) (Gatto Jr et al., 2000), demonstrating the need for a hydrophobic residue at this position. B. The residues that interact with the lysine residue (K) of the PTS1 peptide are displayed. C. The residues that interact with the serine residue (S) of the PTS1 peptide are displayed.

Mammalian PEX5 exists as two alternately spliced forms known as PEX5L and PEX5S (Otera et al., 2002). PEX5L is 37 amino acids longer than PEX5S through the translation of an additional exon. The first 16 amino acids of this 37 amino acid internal insertion are responsible for binding the PTS2 receptor PEX7 (Braverman et al., 1998, Dodt et al., 2001) (see section 1.3.2). In *A. thaliana,* PEX5 also interacts with PEX7 to import PTS2 proteins, though there are not alternative splice variants of *A. thaliana* PEX5 (Nito et al., 2002). In other organisms different PEX proteins act as accessory proteins to enable the import of PTS2 proteins (Brown and Baker, 2008).

The N-terminal domain of PEX5 is intrinsically disordered and less conserved than the C-terminal domain but is essential for the import of cargo across the peroxisomal membrane into the peroxisome matrix (Barros-Barbosa et al., 2019b, Carvalho et al., 2006). The N-terminal domain has multiple different binding sites that allow it to perform the majority of the protein's functions such as membrane docking, binding to PEX7 to import PTS2 cargo proteins and contains the cysteine residue necessary for receptor recycling after cargo protein import (Figure 1.4) (Barros-Barbosa et al., 2019b, Dodt et al., 2001, Neuhaus et al., 2014, Gouveia et al., 2003b, Grou et al., 2012, Ma et al., 2013, Williams et al., 2007). This domain of PEX5 contains diaromatic penta-peptide motifs with the general consensus sequence WXXX(F/Y) (Figure 1.4) (Schliebs et al., 1999). These motifs are responsible for binding PEX14, an essential peroxisomal membrane protein (PMP) in the complex responsible for matrix protein import (see sections 1.3.3.3 and 1.3.3.4) (Reguenga et al., 2001). The human form of PEX5 contains eight of these motifs, with the most recently discovered having the non-canonical sequence of LVAEF (Schliebs et al., 1999, Neuhaus et al., 2014). Other species contain differing numbers of these motifs, for example *Saccharomyces cerevisiae* PEX5 contains only two WXXX(F/Y) domain and PEX5 from *Arabidopsis thaliana* contains nine (Schliebs et al., 1999, Nito et al., 2002).

The use of *in vitro* peroxisomal transport systems has shown that PEX5 will only integrate into the peroxisomal membrane once it is it cargo-bound (Gouveia et al., 2003b). Despite this finding, the N-terminal domain of the protein, when expressed without the C-terminal cargo-binding TPRs, will interact with the peroxisomal membrane, suggesting that the C-terminal domain interacts with the N-terminal domain to regulate its ability to bind at the peroxisomal membrane. This theory is supported by the fact that the two domains from yeast Pex5p have been shown to interact with one another *in vitro* (Harano et al.,

2001). This has led to studies examining whether the binding of cargo to PEX5 causes a conformational change in its structure to allow it to interact with the peroxisomal membrane. Studies using the *in vitro* peroxisomal import system developed by Azevedo et al. have shown that a PEX5 mutation found in a patient unable to effectively import proteins into peroxisomes (N526K) removed the necessity of cargo binding for membrane association of PEX5 (Carvalho et al., 2007a). The susceptibility of this mutant PEX5 protein to proteases suggested a difference in conformation in the C-terminal domain and the authors propose that this is mimicking of conformational changes seen upon cargo binding. However, this does not show that the conformational change causes a release of the N-terminal domain from internal repression and allow it to bind at the peroxisomal membrane, only that this mutation mimicking the cargo-bound state of PEX5 interacts with membrane components and differs in conformation to the wild-type protein when it is not bound to PTS1 cargo. Indeed SAXS data of PEX5 complexes with a cargo protein and peroxisomal membrane component PEX14 (see 1.3.3.3) indicated that no significant changes in the structure of PEX5 were seen upon forming cargo-bound complexes and also that cargo binding was not a requirement for the interaction of the N-terminal domain in the full length PEX5 protein with the peroxisomal membrane proteins (Shiozawa et al., 2009). This SAXS data is at odds with other structural data supporting a conformational change in the C-terminal region of PEX5 upon cargo binding (Stanley et al., 2006), it is possible that as SAXS is a solution-based technique, and the study largely focusses on the interactions of cargo-loaded and apo-PEX5 with the ordinarily membrane-bound PEX14, that the binding interactions seen may not truly represent the native system.

In the Stanley et al. study, the same cargo protein used by Shiowaza et al. is used in complex with PEX5C, a truncated from of the protein containing the C-terminal portion of the protein. Data comparing the X-ray crystal structures of cargo-bound and apo-PEX5 found that the so called "7C loop", which connects the final TPR to the extreme C-terminus, interacts with TPR1 when the receptor is cargo bound, forming a circular structure between the two bundles of TPRs which creates a tunnel in which the PTS1 cargo is bound. A rotation in the peptide backbone of TPRs 5 and 6 in the apo-PEX5 structure means that the 7C loop does not interact with TPR1 and PEX5C has a more open conformation when it is not bound to PTS1 cargo. Currently there are no X-ray crystal structures of full length PEX5, and this means that binding studies of this type only include either the N- or C-terminal domains. As this is the case, it is difficult to say with complete certainty if such conformational changes as seen here

would occur *in vivo*. However, Stanley et al. also carried out analysis of peroxisomal import in human fibroblasts transfected with PEX5 plasmids expressing PEX5 mutants with point mutations in the 7C loop of residues that appeared to make important interactions in their crystal structures. The results of these experiments demonstrated that the 7C loop was critical for proper peroxisomal protein import with PEX5, and the structural data suggest these residues in the 7C loop are likely to function through a conformational change in PEX5 upon cargo binding.

Upon further investigation, it was also found that a similar conformational change in PEX5C was seen when binding to the alternative PTS1 cargo protein alanine-glyoxylate aminotransferase (AGT) (Fodor et al., 2015). This further suggests that this change in conformation is important for efficient import of proteins into the peroxisomal matrix by PEX5.

1.3.3.3 Membrane Docking

After binding matrix protein cargo in the cytosol, PEX5 then needs to interact with the peroxisomal membrane to transport that cargo into the peroxisomal matrix. As stated previously, this is known to be mediated by the intrinsically disordered N-terminal domain of PEX5. It is at this point of the import cycle that the exact process of matrix protein import becomes less clear.

In mammals, the peroxisomal membrane proteins PEX13 and PEX14 are known to act as a docking complex for PEX5 at the peroxisomal membrane as they have been shown to interact with the N-terminal domain of PEX5 via the pentapeptide WXXXF/Y motifs. In yeasts, a third protein, Pex17p, is also part of the docking complex but its function is unclear, though it is known to interact with PEX14p and is vital for the formation of the import complex in yeast (Chan et al., 2016, Huhse et al., 1998). A homolog of PEX17p in mammals or plants has not been identified.

PEX14

Cargo loaded PEX5 makes strong interactions with PEX14 with all of the eight pentapeptide motifs in mammalian PEX5 (Saidowsky et al., 2001)..
Interestingly, there appears to be redundancy in the PEX5:PEX14 interaction as mutational analysis has shown that all the pentapeptide motifs need not be intact for the successful import of PEX5 cargo proteins. This throws support behind a fly-casting model for the interaction whereby the existence of multiple motifs capable interacting with PEX14 increases the likelihood of a successful binding event. The most recently discovered PEX14 binding motif in PEX5 does not fit the WXXXF/Y consensus and in fact has the sequence LVXEF (Neuhaus

et al., 2014). This study looked at the binding affinity and kinetics for the interaction of this motif with PEX14 compared to the WXXXF/Y motifs and found that although binding affinity was comparable between the motifs, LVXEF had a 3-fold faster k_{on} rate, suggesting that this motif could initiate the binding of PEX5 to PEX14 to dock it to the membrane (Neuhaus et al., 2014). It is also possible that these multiple binding sites for PEX14 allow the binding of multiple PEX14 proteins by one PEX5. This could help with the formation of a variably sized pore to allow the import of diverse cargo proteins. It can also be speculated that the existence of multiple PEX14 binding sites on PEX5 is necessary to help ensure that the wide variety of cargos of different shapes and size are able to be imported. It is possible that binding of certain cargos could block certain pentapeptide motifs and this steric hindrance is overcome by the existence of multiple possible binding sites for PEX5 with the membrane docking proteins. Although there is no evidence to support or refute this hypothesis and the fact the PEX5 proteins from other species do not have so many PEX14 binding motifs, makes the reasoning for multiple motifs with redundancy in the mammalian PEX5 more of a puzzle.

PEX5 interacts with the N-terminal domain of PEX14 (Schliebs et al., 1999, Saidowsky et al., 2001). In yeast, an additional binding site in the C-terminal domain of Pex14p has been identified, though a corresponding motif in mammalian PEX14 has not (Niederhoff et al., 2005, Williams et al., 2005). It is also important to note that in yeast the interaction of PEX14p with PEX5p in *S. cerevisiae* does not depend upon the WXXXF/Y motifs in the way that it does in other studied organisms (Bottger et al., 2000, Williams et al., 2005). This suggests that the interaction in this yeast is reliant on an alternative mechanism, and this difference could also mean that the import mechanism in yeast has other distinctions from other organisms. It also demonstrates how the research in this field is yet to give a clear picture of the mechanisms at work and it is also possible that the mechanisms are conserved across species, but that the techniques currently used to investigate this may be giving contradictory results.

PEX13

The role of PEX13 in PEX5 membrane docking is somewhat of an enigma and it appears that there is variation between species in this case. In *Saccharomyces cerevisiae*, Pex5p interacts with both Pex14p and Pex13p and has been shown to help mediate a second interaction between Pex13p and Pex14p (Schell-Steven et al., 2005). In mammals, PEX13 has been found to be essential for the import of catalase, which is targeted by a non-canonical PTS1 sequence (Otera et al., 2002), but the import of other matrix proteins was

unaffected by disruptions to interactions between PEX5 and PEX13, which suggests that, in mammals, PEX13 plays a much less essential role in matrix protein import than PEX14.

Three of the pentapeptide motifs in the mammalian PEX5 N-terminal domain interact with PEX13 (labelled 2-4 in Figure 1.4) (Otera et al., 2002). The affinity of these interactions appears weaker than those made with PEX14. The exact interaction sites on PEX13 to which PEX5 binds appears to be variable between species. PEX13 contains a SRC homology 3 (SH3) domain near its C-terminus which has been shown to interact with PEX14 and is also important for homooligomerisation and binding to the N-terminal domain of PEX5 in mammals and PEX5p in yeast (Costa-Rodrigues et al., 2005, Fransen et al., 1998, Douangamath et al., 2002, Girzalsky et al., 1999, Albertini et al., 1997). In mammals, the N-terminal domain of PEX13 also appears to be a vital domain for facilitating PEX5 interactions (Otera et al., 2002).

Although studies with recombinant proteins *in vitro* have allowed the identification of domains of PEX5, PEX14 and PEX13 that are able to bind one another, what remains elusive is how these proteins interact in the context of the peroxisomal membrane. As PEX5 cycles between the cytosol and the membrane, it is difficult to isolate the import complex, as it only forms transiently while the process of import is occurring. The exact architecture of the import complex is one of the most sought-after answers in the field of peroxisome biology.

Docking of the cargo loaded PEX5 with the peroxisomal membrane is a distinct, reversible step in the process of matrix protein import. Studies looking at the susceptibility of PEX5 to proteolysis have shown that PEX5 subsequently becomes irreversibly inserted into the peroxisomal membrane to deliver its cargo to the peroxisome matrix (Francisco et al., 2013).

1.3.3.4 Membrane Insertion

As PEX5 can be co-isolated with peroxisomes from rat liver (Gouveia et al., 2000), this shows that during the import cycle, PEX5 must become inserted into the peroxisomal membrane. Indeed, studies have shown that PEX5 is resistant to membrane extraction by alkaline pH, a method which is generally used to determine if a protein is interacting with lipids in the membrane (Gouveia et al., 2000). However, it has also been shown that *in vitro* the complex of PEX5 with PEX14 is also resistant to alkaline extraction (Dias et al., 2017). This sets the stage for one of the most widely discussed mysteries: what is the architecture of

the peroxisomal docking/translocation module (DTM) and does PEX5 interact directly with the membrane or only with membrane proteins embedded within it?

A number of different techniques have been employed to investigate the complex that enables peroxisomal matrix protein import. *In vitro* import systems have been designed to study the susceptibly of proteins to degradation by proteases, which sheds some light on the domains of proteins that are exposed to the cytosol and those that are protected from degradation by becoming embedded in the membrane or entering the peroxisomal matrix (Gouveia et al., 2003a) . This system has the advantage that it can accurately mimic the true environment that the peroxisome would find itself in within a cell, but conversely the results cannot give structural data, and this can only be inferred from the data obtained.

Small Angle X-ray Scattering (SAXS), Nuclear Magnetic Resonance (NMR) and X-ray crystallography have allowed some of the conformations of the proteins involved in the import complex to be visualised, but as yet none of these structures have been seen within their native membrane bound state. As cryo-electron microscopy (cryo-EM) methods begin to advance dramatically (Cheng, 2015), this could help to give a better picture of what the complex really looks like when reconstituted in a membrane. However, the fact still remains that the complex formed is only transient, making it difficult to isolate a homogenous sample that could be analysed using these cryo-EM techniques.

The known protein-protein interactions made by the proteins involved in the DTM *in vitro* and structures of these proteins in complexes with one another and in synthetic membranes help to paint a clearer picture of the complex.

As discussed earlier, both PEX14 and PEX13 are known to interact with PEX5 *in vitro*, although these interactions differ between organisms. Recent work using protease protection assays on mammalian PEX14 and PEX13 embedded within liposomes helped to better understand the topology of these proteins within a membrane (Barros-Barbosa et al., 2019a). They showed that PEX14 has N_{in}-C_{out} topology, meaning that the N-terminal domain that most strongly binds to PEX5, is inside the peroxisomal matrix. PEX13 was found to have N_{out}-C_{in} topology and potentially 3 membrane spanning segments. This puts the SH3 domain of PEX13, important for interacting with PEX5 and PEX14, inside the peroxisomal matrix. Structural data for PEX14p in yeast agrees with the topology found in the protease protection assays. Cryo-EM was used to visualise PEX14p in nanodiscs and liposomes and found the C-terminal domain was exposed on the outside and the N-terminal domain was on the opposite side of the membrane (Lill et al., 2020). These corroborating results using

orthogonal methods suggest that the PEX14 topology is truly N_{in}-C_{out}. Though it must be considered that liposomes and nanodiscs are not peroxisomes but a mimic of the membrane, the evidence of the mechanisms suggests that this topology makes logical sense under current models.

In placing the N-terminal domain of PEX14 and the SH3 domain of PEX13 inside the peroxisome matrix, this draws the question of how they make interactions with the PEX5 N-terminal domain. A model of these proteins in these orientations in the membrane suggests that PEX5 is initially docked by either the C-terminal domain of PEX14 or the N-terminal domain of PEX13, both of which would be cytosolic (Figure 1.7A). As the PEX14 N-terminal domain and the PEX13 SH3 domain are known to interact with one another, it makes sense for the topologies of these proteins to place these two domains on the same side of the membrane. In terms of the PEX5 interaction with these proteins, mammalian PEX5 has an interaction site with the N-terminal domain of PEX13 and PEX5p-PEX14p interactions have been shown to occur between the N-terminal domain of PEX5p and the C-terminal domain of PEX14p in yeast. In yeast there are several other interaction partners involved in the DTM, suggesting the involvement of a differing mechanism between yeast and mammals.

In this model, the PEX5 N-terminal domain is bound by one or both of the cytosolic domains of PEX13 and PEX14 and this guides PEX5 to the membrane to release its cargo. It is known that PEX5 becomes embedded in the membrane as it can obtain a protease-protected status. This would allow the PEX5 N-terminal domain to interact with its stronger PEX14 interaction site at the N-terminus of PEX14. It has also been proposed that the multiple WXXX(F/Y) motifs on PEX5 would allow it to be pulled into the membrane through sequential interactions with PEX14- much like pulling along a rope (Emmanouilidis et al., 2016). This may only be limited to PEX5 proteins with multiple interaction motifs with PEX14, yeast PEX5p only had 2 so the number of "handles on the rope" are more limited.

It has been shown that the PEX5 interaction with the SH3 domain of PEX13 is in fact stronger when PEX5 is not bound to its cargo (Otera et al., 2002), this suggests an involvement of this interaction in the release of cargo from PEX5.

What is not known is this exact composition of the pore formed that facilitates the movement of the cargo protein into the peroxisomal matrix. There are currently two main models for this. The first suggests that the pore is formed by the membrane-resident PEX13 and PEX14 upon initiation of the interaction with cargo-bound PEX5 (model 1) (Figure 1.7B). The second suggests that PEX5

itself forms the pore and it is stabilised in the membrane through its interactions with PEX13 and PEX14 (model 2) (Figure 1.7C). As this complex is formed so transiently in vivo it is difficult to isolate and study, hence only models based on what can be inferred from current data are available at present. Both models are feasible as they explain a way in which a wide variety of cargo of different sizes can be imported without the need to unfold proteins or break down protein complexes.
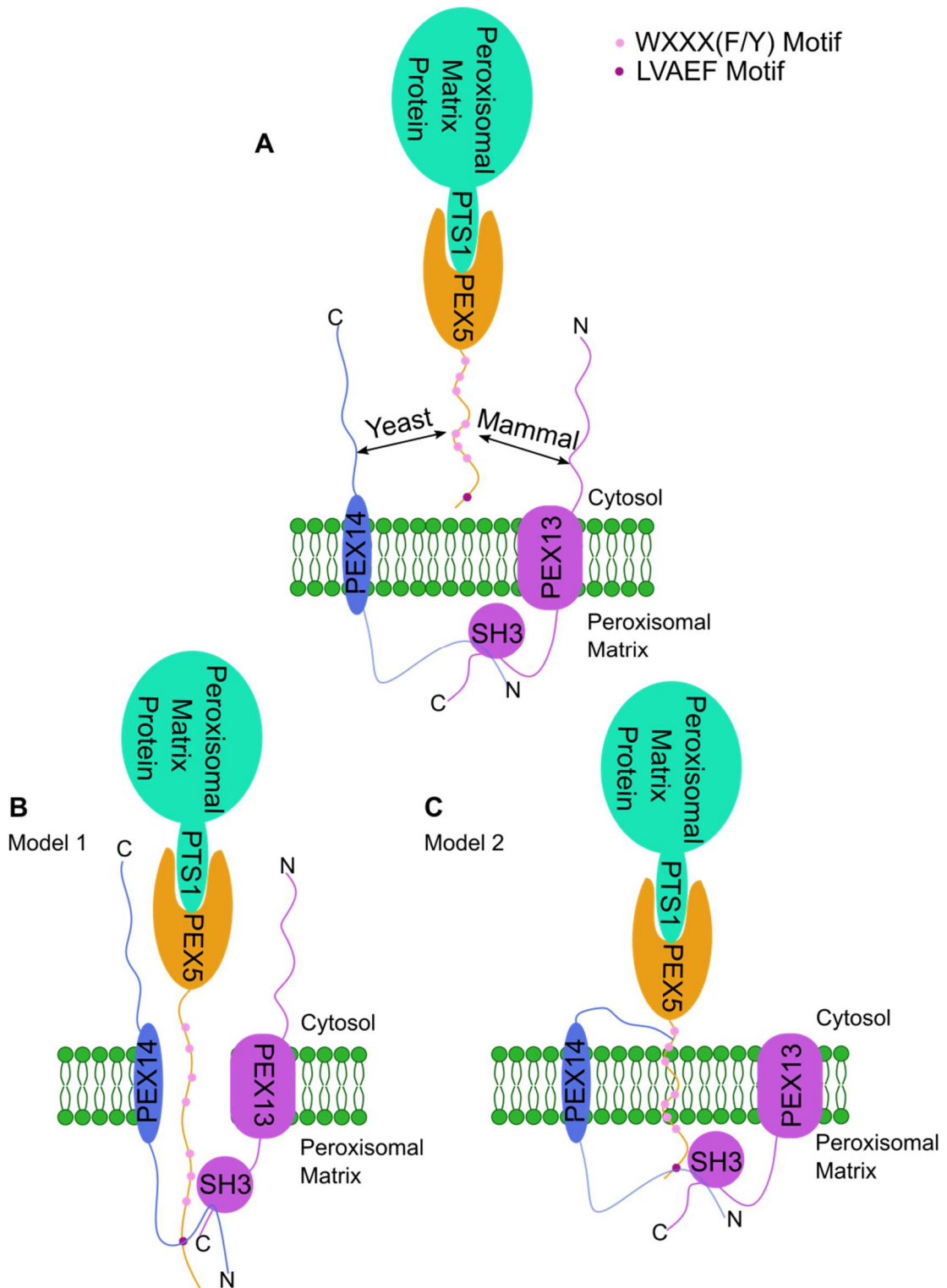
Figure 1.7 Models of PEX5 Membrane Docking and InsertionA. The recruitment of PEX5 to the membrane could be facilitated by interactions with the cytosolic domains of PEX14 and PEX13, which have interaction domains for the PEX5 N-terminal domain in some species. **B.** One model for the insertion of PEX5 into the membranes suggests the formation of a hydrophilic pore in the membrane into which PEX5 inserts to release its cargo. **C.** A different model suggests that PEX5 inserts directly into the membrane and is stabilised in its membrane interaction by PEX14 and PEX13 which act as scaffolding for the formation of a transient pore.

Model 1 is based on data gathered largely by the Azevedo group using *in vitro* studies on post nuclear supernatants from rat liver that contain peroxisomes and cytosolic components. Radiolabelled proteins are used to investigate their susceptibility to degradation by proteases to assess whether proteins or parts of proteins are protected from proteasomal degradation through insertion into the peroxisomal membrane and internal matrix. Liposomes have also been used to study the properties of some peroxisomal membrane proteins. This model depicts the formation of a hydrophilic transmembrane channel through the peroxisomal membrane into which cargo-loaded PEX5 is inserted (Gouveia et al., 2003a). PEX5 then becomes stably associated with this pore through interactions with the N-terminal domain of PEX14 and can pass its cargo into the peroxisomal matrix and associate with the SH3 domain of PEX13 in the process. For larger cargoes, homooligomers of PEX14 and PEX13 can form to create a larger pore which could also accommodate more copies of PEX5. As the PEX5 N-terminal domain is intrinsically disordered and soluble (Carvalho et al., 2006), the formation of hydrophilic pore for it to enter into to release cargo is a feasible mechanism of action, though there is currently no definitive evidence. Current knowledge of the characteristics of the N-terminal domain of PEX5 suggest that it would be more likely to form protein-protein interactions in a hydrophilic pore than insert into the hydrophobic lipid bilayer of the peroxisomal membrane, throwing weight behind this model as opposed to model 2.

If the import of proteins into the peroxisomal matrix did occur in this way, it would be similar to that used by Twin-arginine Protein Translocation (TaT) system found in bacteria, archaea and the thylakoid membranes of chloroplasts (Frain et al., 2019). The TaT system also imports fully folded proteins across a membrane (Léon et al., 2006). The TaT import system is made up entirely of membrane bound proteins (Hasan et al., 2013) and so this system does not have a PEX5 equivalent, which cycles between the cytosol and the membrane to collect and deliver cargo. Proteins that are transported by the TaT system are recognised by a peptide signal, like the PTS1 and PTS2 signals identify the peroxisomal matrix proteins. The canonical TaT signal peptide contains an S-R-R-x-F-L-K motif with the RR residues largely intolerant to substitution. The whole signal peptide will be around 30 residues long, and is cleaved from the protein after import, unlike the short PTS1 and PTS2 peptides (Frain et al., 2019).

Model 2 is supported by those working on the 'transient pore model' proposed by Erdmann and Schliebs (Erdmann and Schliebs, 2005). This model proposes that PEX5 itself becomes directly associated with the phospholipids of the peroxisomal membrane. This is supported by evidence that PEX5 can insert

into the membrane and successfully import cargo when PEX14 is absent or PEX5 is overexpressed *in vitro* (Kerssen et al., 2006, Salomons et al., 2000). The PEX14 structural data suggests it protrudes from the outer surface of the peroxisome to form an 'antennae-like' structure and this points to it having a role in the structural support of PEX5 in the membrane (Lill et al., 2020).

This model would mean that PEX5 behaves in a similar way to a pore forming toxin produced by bacteria and some higher organisms to disrupt cellular membrane and facilitate the import of fully folded cargos (Iacovache et al., 2008). These toxins, like PEX5, cycle between soluble and transmembrane forms. The initiation of pore forming by these toxins is triggered by high localised concentration of toxin monomers, which will assemble into a large oligomeric pore. It is postulated that the initial tethering of PEX5 to the membrane by PEX13 and/or PEX14 could allow PEX5 to behave similarly to these toxins and form a pore through oligomers of PEX5. This would agree with the Salomons et al. study showing insertion of PEX5 into the membrane when it is present in high concentrations.

Without a structure of the DTM complex while it is undergoing the process of matrix protein import, it is impossible to say which of these models, if either, represent the true mechanism used by the peroxisome to import its matrix proteins. Obtaining results from such a complex would require a way in which to freeze the transient importomer at a specific stage in the import cycle.

1.3.3.5 Cargo Unloading

The mechanism of cargo unloading is still under investigation and poorly understood. The process does not require ATP; it is thought to be entirely driven by protein-protein interactions (Francisco et al., 2013). Evidence from yeast has shown that the intraperoxisomal peroxin PEX8 is involved in the release of cargo from PEX5 and that a reducing environment is required (Ma et al., 2013). PEX8 has been shown to interact with both PEX14 (Johnson et al., 2001) and the N-terminal domain of PEX5 (Ma et al., 2013). There is no PEX8 equivalent protein in plants or mammals and so this raises the question of how the unloading mechanism can occur in other species without the action of PEX8. It is possible that a functional homolog of PEX8 may exist in these species that is yet to be identified. On the other hand, in these experiments, PEX8 has only been shown to interact with the docking complex at the peroxisomal membrane and its mechanistic role in the process has not been established, this suggests that it is not only PEX8 playing a role in the process and other proteins, which are present in all species could play a more vital role.

Some evidence has shown that the N-terminal domain of PEX14 is involved in the disruption of the interaction of PTS1 cargo with PEX5 to release it into the peroxisomal matrix (Freitas et al., 2011). This interaction appears to involve diaromatic motifs found in the N-terminal domain of PEX5. *In vitro* studies looking at PEX14 binding to cargo bound PEX5 found that the cargo was not released upon binding, demonstrating that this process is more complicated than this.

There is strong evidence to suggest that the reduction of disulphide bonds that hold together homooligomers of PEX5 also reduces the affinity of PTS1 cargo for PEX5 (Ma et al., 2013). At present these experiments have only been performed with proteins *in vitro* and so whether this truly happens at the peroxisomal membrane remains to be seen and could be a challenge to study.

As the process of cargo unloading is rapid and the PEX5 docking complex at the peroxisome membrane is formed only transiently, it is very difficult to study this part of the process. Tools developed in the work presented here seek to help give a better picture of the process of cargo unloading by creating cargo that cannot be unloaded from PEX5 as this could help to discover the key players involved in the process.

1.3.3.6 Ubiquitination

Once PEX5 has delivered its cargo to the peroxisomal matrix it must then be extracted from the peroxisomal membrane to be able to collect new cargo proteins. This process is facilitated first by the ubiquitination of a cysteine residue close to the N-terminus of PEX5, which is conserved across species (Williams et al., 2007, Carvalho et al., 2007b). This is unusual as ubiquitin modifications usually occur on NH_2 groups to form an amide bond as opposed to an SH group to form a thioester (Hershko and Ciechanover, 1998, McDowell and Philpott, 2013). In order to ubiquitinate a protein, an ubiquitin activating enzyme (E1) first conjugates to the C-terminal glycine residue of ubiquitin, a step which requires ATP. This E1 protein is present in the cytoplasm (Grou et al., 2008). This is then transferred to a ubiquitin carrier protein (E2). In yeast and plants, the peroxin PEX4p present in the peripheral membrane of the peroxisomes acts as the E2 ubiquitin-conjugating enzyme (van der Klei et al., 1998). In mammals however, there does not appear so far to be a peroxisome-specific E2 enzyme but the process is facilitated by UbcH5 enzymes in the cytosol (Grou et al., 2008). A ubiquitin ligase (E3) then links the ubiquitin to the target residue of the protein. The ubiquitin ligases responsible for the monoubiquitination of PEX5 are members of the RING (Really Interesting New Gene) family of proteins and are the peroxisomal membrane proteins PEX2,

PEX10 and PEX12 (Okumoto et al., 2014, Platta et al., 2009). These peroxins are conserved across species. These RING peroxins interact with PEX13 and PEX14 to form the so-called PEX5 "importomer" formed to import protein cargoes (Oeljeklaus et al., 2012, Reguenga et al., 2001).

1.3.3.7 PEX5 Export

It has been established that the initial steps of matrix protein import (from the binding of cargo to its unloading into the peroxisomal matrix) occur independently of ATP and the process is driven solely by protein-protein interactions (Francisco et al., 2013). ATP is only required for the ubiquitination of PEX5 (as described above) and the subsequent extraction of PEX5 from the membrane to allow it to collect further cargo. The Receptor Export Module (REM) is responsible for this step and consists of PEX1 and PEX6. PEX1 and PEX6 are ATPases Associated with diverse cellular Activities (AAA-ATPases), proteins that are part of a larger family of ATPases. PEX26 is a peroxisomal membrane protein that has been shown to interact with PEX1 and PEX6 to recruit them to the peroxisome (Matsumoto et al., 2003). Some evidence also suggests PEX26 acts as the link between membrane-embedded PEX5 and the REM and pulldown assays show interactions between ubiquitinated PEX5 and PEX26 (Hagmann et al., 2018). These same assays also show that PEX6 interacts with ubiquitinated PEX5 without the need for PEX26, the exact mechanism of this stage remains somewhat unclear. Structural studies have shown that PEX1 and PEX6 form a trimer of dimers and both proteins contain two domains for the binding and hydrolysis of ATP (Gardner et al., 2015, Ciniawsky et al., 2015, Blok et al., 2015). ATP hydrolysis by the PEX1/PEX6 complex has been found to be associated with the release of PEX5 from the membrane (Platta et al., 2005) but the mechanism for this process is still under investigation. Current evidence points towards the ATPase complex unfolding PEX5 to disassociate it from the membrane (Pedrosa et al., 2018, Gardner et al., 2018, Pedrosa et al., 2019) although how this occurs remains unknown and a topic for future work in the field.

Another player in the mechanism of PEX5 export from the peroxisome membrane is AWP1. AWP1 is a cytosolic protein that interacts with PEX6 and ubiquitinated PEX5 (Miyata et al., 2012). The evidence suggests that it is likely that AWP1 interacts with the ubiquitin groups on PEX5 (Miyata et al., 2012). This presents further data that could help to elucidate the mechanism by which PEX5 is extracted from the peroxisome membrane.

Once returned to the cytosol, PEX5 must be deubiquitinated to start the cycle of peroxisomal matrix protein import again. Ubiquitin conjugated to cysteine is

somewhat more unstable than when conjugated to lysine as the thiol ester is more susceptible to nucleophiles (Rose and Warms, 1983). This means they can be broken down in the cytosol both non-enzymatically (Grou et al., 2009b) and through the action of ubiquitin-specific protease 9X (USP9X) (Grou et al., 2012).

1.3.4 Export from Peroxisomes

The import of matrix proteins into the peroxisome is a widely studied topic. There has been less focus on the removal of proteins from the peroxisomal matrix in order to adapt the contents of the organelle to a changing environment (Williams, 2014). The export of PEX5 from the peroxisome after delivering its cargo (see 1.3.3.7) is the only peroxisomal export event that has been extensively studied.

The specialised peroxisomes known as glyoxysomes in plants gave the first indication that it was possible that some matrix proteins were exported. Isocitrate lyase (ICL) and malate synthase (MLS), enzymes involved in seed germination and early stage growth were found to be present in peroxisomes two days after germination but were no longer there after eight days, when a different set of enzymes were found to be present as the needs of the cells had now changed (Titus and Becker, 1985). At day four, both sets of enzymes were present. It was postulated that the ICL and MLS could have been degraded within the peroxisome, and therefore not exported. However, further experimentation showed that disruption of the peroxisomal Ubiquitin conjugating E2 enzyme (Zolman et al., 2005) or the AAA-ATPase PEX6 (Lingard et al., 2009) prevented degradation of ICL and MLS, suggesting that the peroxisomal export machinery were involved in the process. There is however also evidence indicating that the peroxisomal resident protease Lon2 could also play a role in the degradation of MLS (Farmer et al., 2013).

The export of the peroxisomal membrane protein PEX3 in *Hansenula. polymorpha* has been found to be necessary step for the degradation of peroxisomes when they are no longer required (see: 1.4 Pexophagy). The process requires PEX3 to be ubiquitinated, which targets it to the proteasome after removal from the membrane (Bellu et al., 2002). The same ubiquitin E3 ligases are involved in the ubiquitination of PEX3 that are responsible for PEX5 ubiquitination, but the AAA-ATPase PEX1 is not required for the process (Williams and van der Klei, 2013).

Although export of proteins from the peroxisome appears to occur, it currently seems that it is a protein-specific process to serve a targeted function. In

general, it appears peroxisomal matrix proteins are delivered to the peroxisomal matrix and will remain there for the life of the organelle.

1.4 Pexophagy

Much of the focus of research into the proper functioning of peroxisomes has focused on the formation of the membrane and the import of the matrix proteins. What is also important to consider, is the fate of the peroxisome when it reaches the end of its life. The predicted half-life of a peroxisome is approximately 2 days (Huybrechts et al., 2009, Poole et al., 1969) and early experiments indicated that the organelle is destroyed as a whole (Poole et al., 1969). Although the process is sometimes random, processes that selectively identify dysfunctional or surplus peroxisomes have been discovered. There are some matrix proteins that are specifically degraded through the action of proteases (Lingard et al., 2009, Kikuchi et al., 2004) and whole peroxisomes can undergo autolysis when the membrane is disrupted by the action of 15-lipoxygenase (Yokota et al., 2001). However, peroxisomes are predominantly turned over by selective macroautophagy, known as pexophagy.

Autophagy degrades cellular components by engulfing them within vesicles and delivering them to lysosomes, where they are broken down (Zientara-Rytter and Subramani, 2016). A double membrane vesicle, an autophagosome, forms around the peroxisome and it is transported to the lysosome via microtubules (Yang and Klionsky, 2010).

For the cell to know which peroxisomes need to be degraded, there must be markers on the peroxisome and receptors to recognise these markers and recruit autophagy components. Both the markers and receptors remain largely unknown. However, ubiquitination of proteins in the peroxisomal membrane is suspected to play a part in pexophagy in mammalian cells.

Artificial fusion of ubiquitin in PEX3 and PMP34 increases the rates of pexophagy (Kim et al., 2008). Although there is no evidence to show that this ubiquitination occurs endogenously, it suggests that ubiquitination of a protein in the peroxisomal membrane on the cytoplasmic side triggers the recruitment receptors to initiate the process of pexophagy. One candidate for this is PEX5. As discussed, PEX5 is monoubiquitinated at a conserved cysteine residue in the process of its recycling back to the cytoplasm to collect further cargo (see: 1.3.3.6). If the recycling process is impaired, this has been shown to trigger pexophagy (Nordgren et al., 2015). The monoubiquitination of PEX5 at a lysine residue has also been shown to promote pexophagy (Zhang et al., 2015).

In mammalian cells, the proteins NBR1 and p62 have been identified as receptors to facilitate pexophagy (Deosaran et al., 2013, Kim et al., 2008). Both of these proteins have binding domains for ubiquitin and LC3-II, a protein found on the surface of autophagosomes (Kabeya et al., 2000). NBR1 appears to be a necessary component for pexophagy, but when expressed at endogenous levels, p62 is also required. The two proteins interact co-operatively to target ubiquitinated peroxisomes for pexophagy (Deosaran et al., 2013).

1.5 Diseases Associated with Peroxisomes

One of the biggest driving forces for studying any biological process is to better understand the disease phenotypes seen when the process goes wrong. The peroxisome is no exception, and the effects of dysfunctional peroxisomes are far-reaching. The critical role of the peroxisome in maintaining a healthy cell is exemplified by the disease states seen in patients who lack functional peroxisomes or have impaired peroxisomal activities. These conditions are generally divided into two classes: peroxisome biogenesis disorders (PBDs), in which patients have mutations that prevent the formation of functional peroxisomes, and single peroxisomal enzyme deficiencies (PEDs), in which patients lack a particular enzyme that is responsible for a peroxisomal process, such as alpha- and beta-oxidation or the transport of metabolites (Wanders and Waterham, 2006b, Waterham et al., 2016).

The PBDs are caused by biallelic mutations in the PEX genes and result in the absence of functional peroxisomes in cells (Argyriou et al., 2016). The Zellweger Spectrum Disorders (ZSDs) are a heterogeneous group of more than 15 disorders with differing severity and clinical presentation, hence the term Zellweger Spectrum Disorders is used as the phenotypic presentation of these peroxisome disorders can vary greatly and cannot always be defined by separate disease terms (Poll-The et al., 1988). The genetic basis of peroxisomal disorders can determine the severity. In cases with more severe presentation, this is often due to a complete lack of a functional peroxin, whereas milder cases can be due to mutations that result in proteins retaining only partial functionality (Ebberink et al., 2011). In general, the earlier the onset of the disease, the more severe the presentation. The spectrum, in order of decreasing severity of phenotype, includes Zellweger syndrome, neonatal adrenoleukodystrophy, infantile Refsum disease and Heimler syndrome (Klouwer et al., 2015). Heimler syndrome has only recently been added to the spectrum as its causative genetic mutations were only recently identified as PEX1 and PEX6 (Ratbi et al., 2015) and its clinical phenotype is significantly

milder than other Zellweger spectrum disorders. The peroxisome biogenesis disorders are summarised in Table 1.2.

Table 1.2 The Peroxisome Biogenesis Disorders

Disorder	Cause	Clinical Presentation
Zellweger Syndrome	Mutations in any of *PEX1, PEX2, PEX3, PEX5, PEX6, PEX10, PEX11β, PEX12, PEX13, PEX14, PEX16, PEX19* and *PEX26* genes Non-functional peroxisomes result in the accumulation of VLCFAs, bile acid intermediates and reduced levels of plasmalogens found in membranes	Severe lack of muscle tissue, seizures, craniofacial dysmorphisms such as a high forehead and large anterior fontanel Liver disease Retinal damage Deafness Unlikely to survive longer than one year
Neonatal Adrenoleukodystrophy		Neonatal onset of hypotonia and seizures Delayed neurodevelopment (Benke et al., 1981) Few patients live to teenage years (Kelley et al., 1986)
Infantile Refsum Syndrome		Milder form of disease with some patients seen to live into adulthood. (Wanders et al., 1986)
Heimler Syndrome		Very mild form of Zellweger Spectrum Disorder. Hearing loss, nail and tooth abnormalities (Ratbi et al., 2015)

Rhizomelic Chondrodysplasia Punctata type 1 (RCDP1)	Mutation in PEX7 preventing import of PTS2-targeted proteins (Braverman et al., 1997) More deficient in plasmalogens than in ZSDs but normal levels of VLCFAs (Hoefler et al., 1988)	Shortening of limbs (rhizomelia) Facial abnormalities (e.g. high forehead) Growth and mental retardation (Waterham et al., 2016)
Rhizomelic Chondrodysplasia Punctata type 5 (RCDP5)	Loss of long isoform of PEX5, preventing import of PTS2-targeted proteins (Barøy et al., 2015)	Similar phenotype to RCDP1 due to lack of plasmalogen synthesis

The PEDs are usually subdivided into classes based on the specific peroxisomal function that is affected. These are summarised in Table 1.3.

Table 1.3 The Single Peroxisomal Enzyme Deficiencies

Affected Process	Disorder	Cause	Clinical Presentation
Fatty Acid Alpha Oxidation	Refsum Disease	Mutation in *PAHX* gene encoding PAHX essential for alpha oxidation of phytanic acid.(Jansen et al., 1997) Phytanic acid accumulates (Steinberg et al., 1967)	Progressive sensory loss. Cardiac arrhythmias (Wierzbicki et al., 2002)
Fatty Acid Beta Oxidation	X-linked adrenoleukodystrophy(Wanders et al., 1988)	Mutation in *ABCD1* gene VLCFAs accumulate (Aubourg et al., 1993)	Variable phenotypes. VLCFA accumulation causes damage to adrenal glands and myelin sheaths. Results in progressive neurological and behavioural decline (Moser et al., 2000)

	Acyl-CoA oxidase deficiency (Poll-The et al., 1988)	Mutation in *ACOX1* gene VLCFAs can accumulate Results in enlarged peroxisomes. (Poll-The et al., 1988)	VLCFA accumulation causes similar damage to cells as described in X-ALD. Usually earlier onset than X-ALD (Suzuki et al., 2002)
	D-bifunctional protein deficiency (Wanders et al., 1992b)	Mutation in *HSD17B4* gene encoding D-bifunctional protein/multifunctional protein 2 (Ferdinandusse et al., 2006a) VLCFAs, pristanic acid and bile acid intermediates can all be seen to accumulate (Wanders and Waterham, 2006b)	Lack muscle tone and a lack of motor skills. Frequent seizures. Impaired vision and hearing. Liver failure (Ferdinandusse et al., 2006a)
	2-Methylacyl-CoA racemase (AMACR) deficiency	Mutation in *AMACR* gene encoding AMACR, an enzyme that converts BCFAs to the appropriate stereochemistry for further metabolism(Ferdinandusse et al., 2000b) Accumulation of pristanic acid and bile acid intermediates DHCA and THCA (Ferdinandusse et al., 2001b)	Adult-onset motor and sensory impairment (Ferdinandusse et al., 2000a)

	SCPx deficiency	Mutation in *SCP2* gene encoding SCPx, the thiolase for branched chain FAs. Accumulation of pristanic acid and bile acid intermediates occurs (Ferdinandusse et al., 2006b)	Reduced muscle tone and motor control (only observed in a single patient) (Ferdinandusse et al., 2006b)
Ether Phospholipid Biosynthesis	Rhizomelic chondrodysplasia punctata type 2	Mutation in *GNPAT* gene encoding DHAPAT enzyme essential for synthesis of ether phospholipids (Ofman et al., 1998) Deficiency in plasmalogens seen	Facial abnormalities from birth, dwarfism, lack of muscle tone , bone abnormalities (Wanders et al., 1992a)
	Rhizomelic chondrodysplasia Type 3	Mutation in *ADHAPS* gene encoding alkyl-DHAP synthase enzyme essential for synthesis of ether phospholipids (de Vet et al., 1998) Deficiency in plasmalogens seen	Facial abnormalities from birth, dwarfism, lack of muscle tone , bone abnormalities (Wanders et al., 1994)

| Glyoxylate Detoxification | Primary hyperoxaluria type 1 | Mutation in *AGXT* gene causing deficiency in alanine glyoxylate aminotransferase (AGT) in the peroxisome (Williams et al., 2009). AGT is responsible for the detoxification of glyoxylate in the liver. Causes build-up of calcium glyoxylate, particularly in the kidneys and urinary tract (Fodor et al., 2012) | Heterogeneous in symptoms and onset. In severe cases kidney failure occurs and glyoxylate deposits in other tissues can cause multiple tissue and organ failures. (Danpure, 1989) |

1.6 Chemical Biology to Study and Target Peroxisomes

Peroxisomes have been of great interest in biotechnology and synthetic biology. They act as containment vessels for enzymatic reactions and the specific import mechanisms of this organelle allow tight control over their contents (Baker et al., 2016). Their ability to import only proteins with specific peroxisome targeting signals (PTSs) and proteins in their fully folded state (Walton et al., 1995) as well as oligomeric protein complexes (McNew and Goodman, 1994) makes them a unique compartment ideally suited to customisation for biotechnology and synthetic biology studies.

Studying peroxisomes has revealed a number of characteristics that could be exploited to help further our knowledge of cellular functions as well as target disease. There has also been much interest in the use of peroxisomes as compartments for the synthesis of useful products. In this section, examples of how chemical tools can be used to manipulate peroxisomes and the development of molecules to target peroxisomal components will be discussed.

1.6.1 PEX14-PEX5 Inhibitors in Trypanosomatids

As stated previously, trypanosomatids contain glycosomes, which perform glycolysis, an essential reaction to metabolise glucose and release energy in trypanosomes. It has been found that selective degradation of mRNA encoding the PEX14 protein in *Trypanosoma brucei* (*Tb*PEX14) using RNAi resulted in glucose having a toxic effect on the parasite (Haanstra et al., 2008, Furuya et al., 2002).

Trypanosoma brucei causes Human African trypanosomiasis (HAT, sleeping sickness) and bovine nagana, and *Trypanosoma cruzi* causes Chagas disease, parasitic diseases transmitted by the tsetse fly (Alsford et al., 2013). Both diseases are a major threat to health of both humans and livestock in Africa and South America (Dawidowski et al., 2017, Kristjanson et al., 1999). Trypanosome species are also becoming resistant to current therapies used to treat these diseases (Barrett et al., 1995, Geerts et al., 2001). For these reasons, it is important to develop new therapies to effectively target these parasites.

As the RNAi knockdown of *Tb*PEX14 was shown to be lethal to *T. brucei*, small molecules are in development to have a similar effect (Dawidowski et al., 2017). These molecules were designed to disrupt the interaction between *Tb*PEX14 and *Tb*PEX5 specifically and not interfere with the PEX14-PEX5 interaction in human cells. To do this, the binding pocket on *Tb*PEX14 was studied using NMR and the chemical shifts upon binding *Tb*PEX5 compared with the shifts of

the *Hs*-PEX14-*Hs*-PEX5 interaction. There was found to be distinction between the two which could be exploited to create a specific binder for *Tb*-PEX14 that would disrupt its interaction with *Tb*-PEX5. The structure of the inhibitor mimics a WXXXF motif on PEX5 that would bind to PEX14 via hydrophobic interactions (Figure 1.8). The specificity for *Tb*-PEX14 is conveyed through a vital interaction of an NH_2 group with the carboxyl side chain of Glu34 in *Tb*-PEX14 via a water molecule. In *Hs*-PEX14 the equivalent residue is a lysine, which has a positively charged lysine that would repel the NH_2 group. These inhibitors were also found to effective against other trypanosomatids with the same conserved binding region in PEX14.
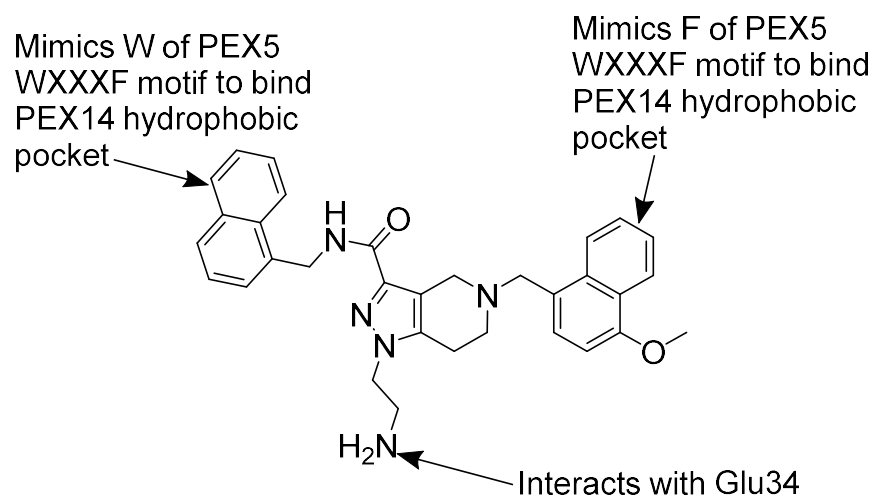
Figure 1.8 Inhibitor of PEX14-PEX5 interaction in Trypanosomes The inhibitor developed by Dawidowski et al. to specifically inhibit interactions between PEX14 and PEX5 in trypanosomes.

The development of inhibitors such of these demonstrates the importance of good understanding of peroxisomal mechanisms and the protein-protein interactions required for functional peroxisomes.

1.6.2 Exploiting Peroxisomes as Reaction Vessels

Much work has focused on using the peroxisome as a specialised compartment to contain chemical reactions and metabolic pathways. To do this, research has found ways to import non-native cargo proteins in to the peroxisomal matrix (DeLoache et al., 2016). Engineering peroxisomes in this way is being studied as a method to produce fatty-acid derived biofuels (Zhou et al., 2016). Peroxisomes can be used to compartmentalise pathways where side reactions may interfere with efficient production of desirable substances.

What also makes the peroxisome a good compartment for redirecting heterologous metabolic pathways is the fact that cell growth on glucose in yeast is not adversely affected by disrupting peroxisomes (Purdue and Lazarow, 2001). Yeast peroxisomes have already been utilised to contain the enzymes for chemicals that have proven difficult to synthesise chemically such as the pigmentation chemical lycopene, the fragrant sesquiterpenoid (+)-nootkatone used in the food, cosmetic and pharmaceutical industries and polyhydroxyalkanoates (PHAs), which can be made into biodegradable polymers (Bhataya et al., 2009, Wriessnegger et al., 2014, Poirier et al., 2002).

Most recently, peroxisomal compartmentalisation has been used to produce highly valuable plant natural products (Grewal et al., 2021). When yeast has been used to make benzylisoquinoline alkaloid (BIA) family natural products, the process is limited by the fact that the enzyme required, norcoclaurine synthase (NCS), is toxic to *S. cerevisiae* when expressed in the cytosol. By targeting NCS to the peroxisome with a PTS1 sequence, the substrates for the reaction (which were less than 300Da and therefore able to permeate the peroxisomal membrane) could be metabolised by NCS without toxic side effects on the yeast cell. This ultimately resulted in a significantly improved yield of product.

These examples show that the peroxisome provide a unique compartment in which to produce desirable products with increased yields. This approach could be more widely applied to help to develop more sustainable manufacturing methods.

Previous work in our laboratory has taken this concept further to instead create "designer" peroxisomes, which preferentially import the desired cargoes (Cross et al., 2017).

1.6.3 Studying Protein Interactions in Peroxisomes

Peroxisomes have also been used to study protein-protein interactions. Work in the Dueber and DeLoache labs have utilised peroxisomal import in this way by creating a "bait" protein with a PTS tag for peroxisomal import and a fluorescently tagged "prey" protein. By monitoring subsequent fluorescence within peroxisomes they were able to detect whether the two proteins interacted in the cytosol, even if the interaction was weak (Chen et al., 2015).

Similarly, other groups have looked at the ability to study protein interactions with small RNAs; these interactions are generally quite labile. Using the "piggybacking" process (McNew and Goodman, 1994), which is unique to peroxisomes in their import mechanism, allowed protein-RNA interactions not observable by immunoprecipitation (IP) to be detected by subsequent isolation of peroxisomes (Incarbone et al., 2018).

1.7 Objectives of the study

The aim of this work is to synthesise chemical probes containing a PTS1 peptide and use these probes to investigate the import of proteins across the peroxisomal membrane.

In the first study (Chapter 2), probes are synthesised to investigate if the attachment of a PTS1 peptide to a protein via a probe is sufficient for that labelled protein to be recognised for peroxisomal import by PEX5. Success of these probes would allow non-peroxisomal matrix proteins to be directed to the peroxisomal matrix under the control of an external reagent. The proteins tested contain a tag that allow the covalent attachment of the probe, and therefore a PTS1 to the protein as a branch from the main protein chain. Mass spectrometry is used to assess these labelling reactions. Pulldown assays are used to evaluate the interaction of PEX5 with these probe-labelled proteins. These pulldown assays are optimised through several alterations to materials and methods. Some investigation into the ability of PEX5 to import probe-labelled proteins across the peroxisomal membrane in a cellular environment is also carried out.

Specific Goals:

- Synthesise PTS1 peptides containing tag-reactive motifs
- Assess the ability of PEX5 C-terminal domain to bind to proteins labelled with PTS1 via a covalent tag
- Assess the ability of endogenous PEX5 to import proteins labelled with PTS1 via a covalent tag in cells

The success of this study would allow targeted re-localisation of non-peroxisomal proteins to the peroxisome matrix. This could allow for the study of interactions and downstream targets of proteins to be studied by isolating them from the cytoplasm.

In the second study (Chapter 3), tools are generated to allow the study of PEX5 in its usually transient cargo-bound state. The probes in this instance contain a PTS1 peptide and a tag-reactive motif as before. Additionally, they contain a cysteine-reactive group to allow covalent attachment of the probe to PEX5 in the PTS1 binding pocket of the protein. The covalent attachment of the probes is assessed using mass spectrometry. The ability of a protein labelled with a PTS1 via a tag to become covalently attached to PEX5 is analysed by mass spectrometry and SDS-PAGE.

Specific Goals

- Synthesise PTS1 peptides containing cysteine-reactive motifs
- Introduce reactive cysteine residues into the PEX5 cargo-binding pocket using site-directed mutagenesis
- Assess the ability of reactive probes to bind to PEX5 cysteine mutants
- Synthesise PTS1 peptides containing both tag-reactive and cysteine-reactive motifs
- Assess the ability of PEX5 cysteine mutants to bind to proteins labelled with a PTS1 via tag-reactive probes

The success of this study could allow full length PEX5 to be studied in its cargo-bound state, in the peroxisomal membrane, something that has until now not been possible. Probes that allow a protein cargo to be bound by PEX5 but not released give the potential for structural studies to elucidate a mechanism for the insertion of PEX5 into the peroxisomal membrane and a better understanding of the interactions PEX5 makes with peroxisomal membrane proteins and how cargo proteins destined for the peroxisomal matrix are transferred from PEX5 to the matrix.

Chapter 2

Covalent Attachment of PTS1 peptides to re-localise proteins to the peroxisome

2.1 Introduction

As described in Chapter 1, the most common way peroxisomal matrix proteins are imported into the peroxisome is through a C-terminal tripeptide sequence (a PTS1) that is recognised by the protein PEX5. It has been shown that the presence of a sequence fitting the PTS1 consensus of [short side chain]-[basic side chain]-[hydrophobic side chain] at the C-terminus of any protein can allow it to be recognised by PEX5 and trigger cargo import into the peroxisome (Gould et al., 1989).

In this chapter we ask if such import can be triggered by addition of a chemical probe. We envisioned that post-translational modification of a protein with a PTS1 sequence may be sufficient to enable peroxisomal re-localisation (Figure 2.1).
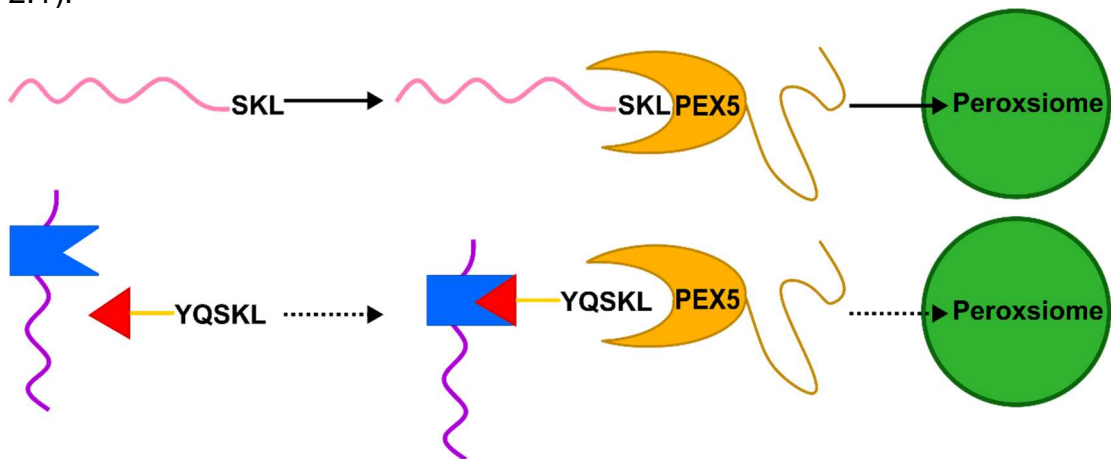
Figure 2.1 A Cartoon Schematic of the project aim The canonical import of a protein into the peroxisomal matrix is with a C-terminal PTS1 sequence, encoded in the gene for the protein. This is recognised by PEX5 which facilitates the import of the protein across the peroxisomal membrane (top). In this study the possibility of the attachment of a PTS1 peptide (YQSKL) to a protein post-translationally has been investigated (bottom).

There are a number of reasons it might be desirable to re-locate a protein to the peroxisomal matrix. By creating a chemical probe that can trigger proteins to become imported into the peroxisome on demand, it could help to better understand the effects the protein has on the cell. In addition to this, the peroxisome could potentially be used as a containment vessel to selectively knock-down the activity of a particular protein through its re-localisation to the peroxisome preventing it from interacting with binding partners and causing downstream biological effects. This concept is inspired by PROTACs, which target proteins for degradation by the proteasome as a therapeutic strategy (Toure and Crews, 2016).

This concept has recently been shown to be a mechanism the cell might use to regulate the activities of certain proteins. In nature, there have been examples of proteins which only exist in the peroxisome under certain conditions, but do not play a functional role inside the peroxisomal matrix (Reglinski et al., 2015). One such example is the deubiquitinating enzyme USP2. The role the four isoforms of USP2 play on the fate of the cell is controversial as it appears to behave in both tumorigenic (e.g. (Priolo et al., 2006)) and apoptotic (e.g. (Gewies and Grimm, 2003)) ways. Each isoform contains the same C-terminus, which terminates in the potential PTS1 sequence SRM. This appears to be sufficient to translocate USP2 isoforms to the peroxisome, despite USP2 not having any known targets in the peroxisomal matrix (Reglinski et al., 2015). In studying the import of USP2 isoforms into peroxisomes in mammalian cells, it was found that the proapoptotic downstream effects of USP2 were reduced when the efficiency of peroxisomal import was increased. This suggests that the cytosolic activities of USP2 are being regulated by its peroxisomal import. What is not known is how the cell might trigger an increased import of USP2 proteins to the peroxisome, this could be regulated by modifications to the protein under certain conditions to increase its affinity for PEX5. The regulation of protein activities by removing them from the cytosol into the peroxisome is the overall aim of this project.

To investigate if a PTS1 attached to a protein via a branch from the nascent protein chain is sufficient for peroxisomal import, the labelling peptides were used to react with their substrates both *in vitro* and in mammalian cells.

In vitro the recognition of labelled proteins by PEX5 was investigated using recombinant expressed C-terminal domains of PEX5 from *Arabidopsis thaliana* and *Homo sapiens* (Figure 2.2). The C-terminal domains each contain an N-terminal His$_6$-tag for purification purposes; the His$_6$-tag must be at the N-terminus so as not to interfere with the binding activities of the C-terminal TPR

domains. These constructs are referred to as *At*-His$_6$-PEX5C and *Hs*–His$_6$-PEX5C. The cargo proteins to be used here were based on SNAP-Tag and HaloTag proteins, which can be covalently modified with chemical probes (in this case containing a PTS1 motif) both *in vitro* and *in vivo*.
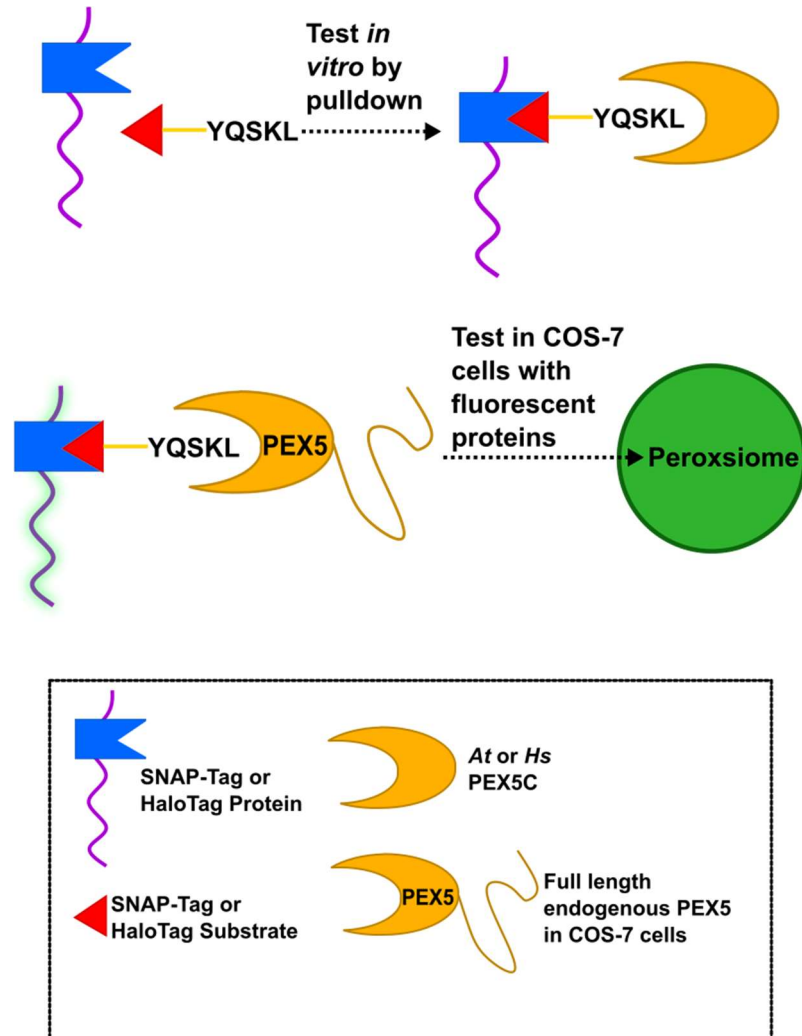
Figure 2.2 Summary of strategy used The strategy used to test the functionality of the labelling of proteins with PTS1 peptides showing the reagents required to be synthesised and the proposed methods of investigation

2.1.1 The SNAP-Tag

The SNAP-Tag is a 20kDa modified form of the human DNA repair enzyme O^6-alkylguanine-DNA-alkyltransferase (hAGT). The endogenous enzyme acts to remove alkyl groups from guanine residues in DNA (Figure 2.3A) (Keppler et al., 2004). The SNAP-Tag was developed by directed evolution of hAGT to give an enzyme with increased activity and higher stability in cells (Juillerat et al., 2003, Mollwitz et al., 2012). The SNAP-Tag-reactive substrate is commonly an O^6-benzylguanine derivative (Cole, 2013). It reacts with hAGT at a cysteine residue, forming a covalent bond through loss of the O^6-pyrimidine group (Keppler et al., 2004). The O^6-benzyl-4-chloropyrimidine is an alternative SNAP-Tag substrate to the benzylguanine derivatives (Figure 2.3B). The chloropyrimidine substrate was chosen to be used in this case and has been found to have better solubility and cell permeability than benzylguanine alternatives (Correa et al., 2013, Cole, 2013). The commercial reagent SNAP-Cell® TMR-Star (NEB) uses chloropyrimidine as the SNAP-Tag substrate and has been demonstrated to be effective when for imaging in living systems (Erdmann et al., 2019).
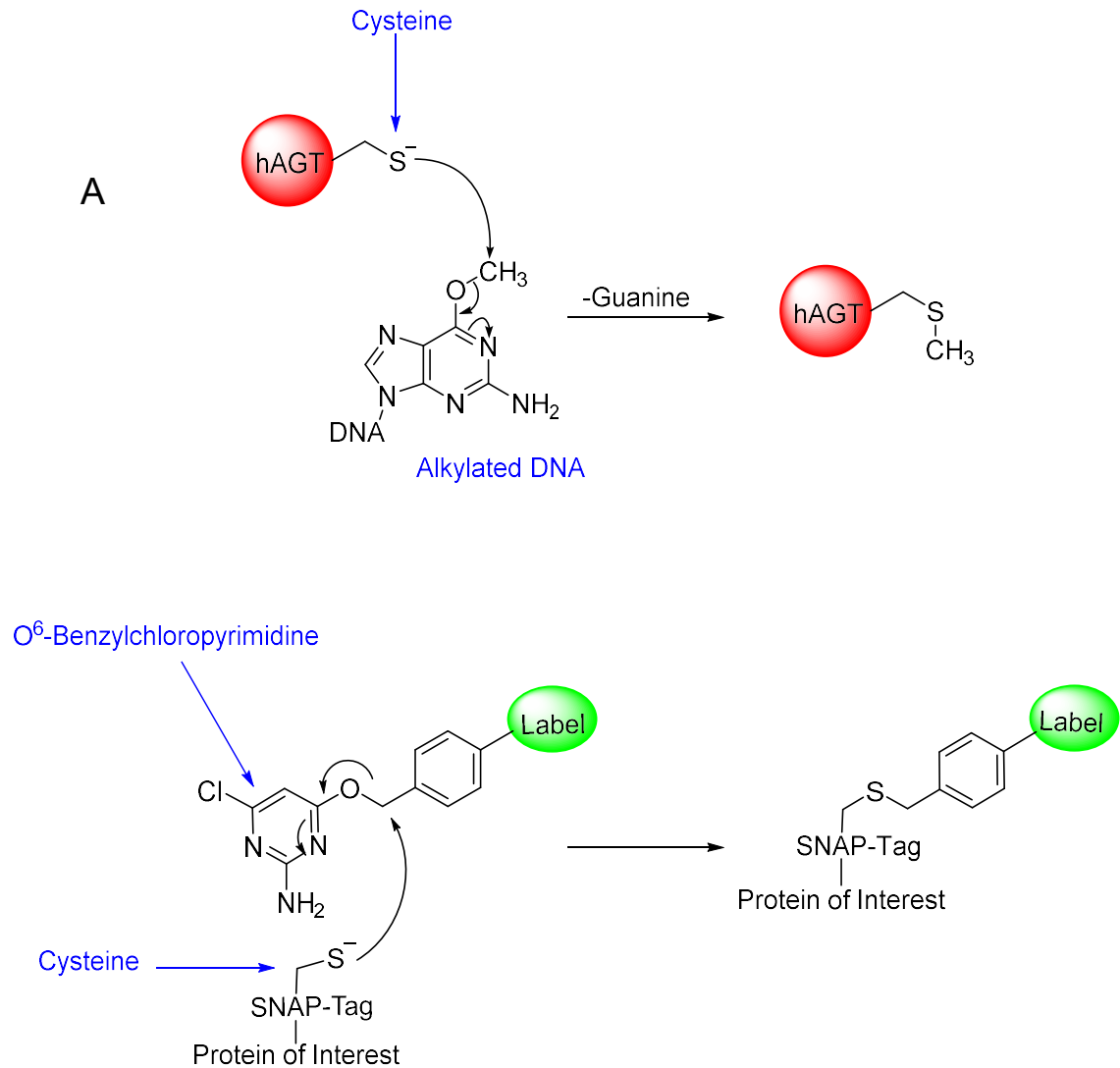
Figure 2.3 Mechanism of Action of the SNAP-Tag A. The mechanism used by hAGT to remove an alkyl group from guanine residues in DNA. **B.** The mechanism the SNAP-Tag uses to form a covalent bond with molecules containing a O^6-Benzylguanine or O^6-Benzylchloropyrimidine group. A cysteine forms a covalent thioether bond with the molecule irreversibly.

2.1.2 The HaloTag

The HaloTag is an alternative to the SNAP-Tag protein that can also be selectively reacted with suitable chemical probes both *in vitro* and *in vivo*. The HaloTag protein is derived from bacterial haloalkane dehalogenases (Janssen, 2004). The haloalkane dehalogenase enzymes catalyse the substitution of a terminal chlorine or bromine for the corresponding alcohol molecule (Janssen, 2004). This is catalysed by first an aspartic acid residue (Asp106) in the enzyme performing a nucleophilic substitution of the halogen and forming an intermediate where the enzyme is covalently linked to the substrate (Figure 2.4A). A histidine residue (His272) then acts as a base to generate the alcohol product and release the enzyme to catalyse further reactions.

To re-engineer this protein to make it a useful protein tag, the His272 catalysing the release of the enzyme was substituted for a phenylalanine (Los et al., 2008). As phenylalanine cannot catalyse the hydrolysis of the enzyme-alkane intermediate, the mutated enzyme, the HaloTag, can irreversibly form a covalent bond with its substrate (Figure 2.4B). Synthetic HaloTag substrates all contain a chlorohexyl motif. As with the SNAP-Tag, the HaloTag can be fused to a protein of interest to allow the labelling of the protein with the desired label.
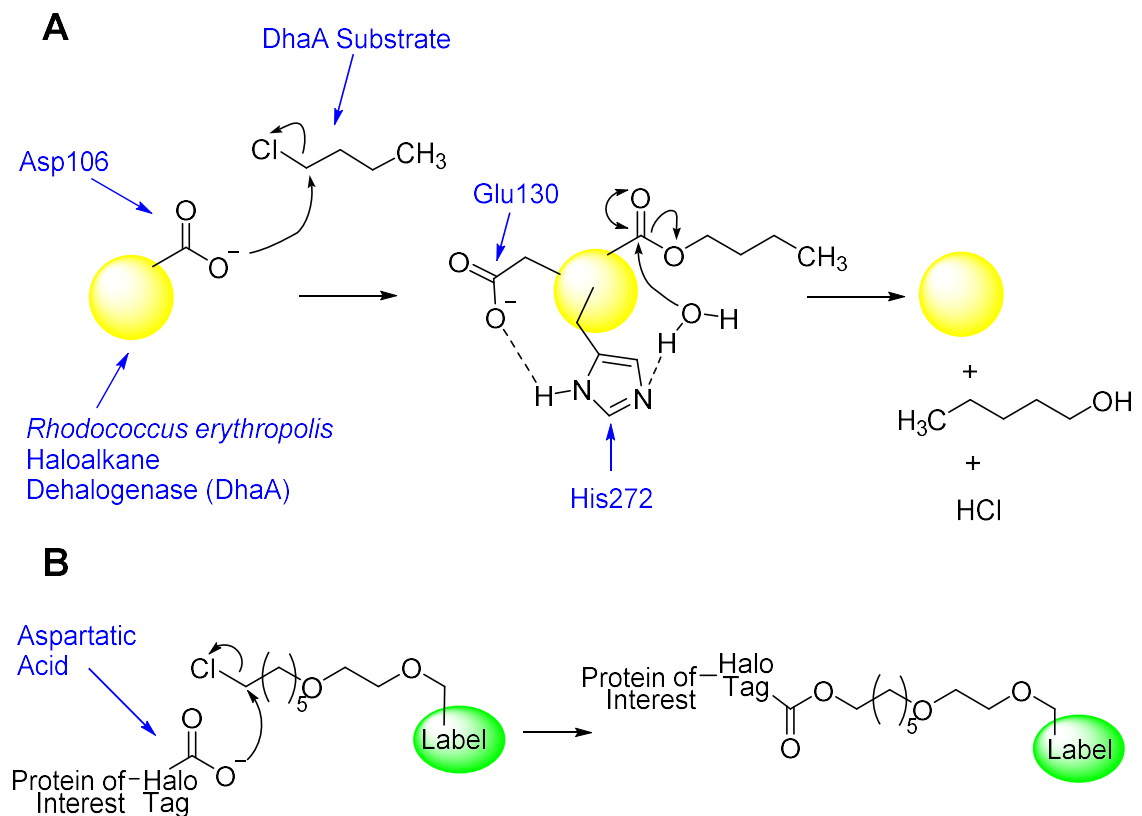
Figure 2.4 Mechanism of Action of the HaloTag

A. The mechanism used by haloalkane dehalogenase found in *Rhodococcus erthytopolis* (DhaA). **B.** The mechanism for the reengineered HaloTag protein from DhaA. This results in the irreversible formation of a covalent bond between the HaloTag and its substrate. The substrate contains the desired label for the protein on interest, which has been fused to the HaloTag protein.

2.2 Synthesis of the SNAP-Tag substrate

The SNAP-Tag substrate molecule was synthesised in 5 steps from methyl-4-cyanobenzoate using a route based on literature precedent (Hoffer et al., 2018, Keppler et al., 2003, Srikun et al., 2010). The first step required the reduction of methyl-4-cyanobenzoate with $LiAlH_4$. Due to the highly pyrophoric nature of this reaction, it was only be performed on a maximum of a 1g scale. The reaction was however very efficient with complete reduction seen to occur and no further purification needed. The reaction was quenched using a 50:50 (v/v) mixture of $NaSO_4.H_2O_{10}$ and Celite 577. A small amount of H_2O is then added to ensure the $LiAlH_4$ is completely quenched; this prevents the formation of large amounts of aluminium salts observed using literature methods (Song et al., 2013, Hoffer et al., 2018). The optimised quenching conditions used here avoid the formation of oily emulsions and only required the THF-soluble product to be filtered from the solids before proceeding with the subsequent step.

In the second step of the reaction the amine group must be protected in order to prevent it from reacting in place of the alcohol in subsequent steps. Initially, a tert-butyloxycarbonyl (Boc) group was used for this and gave a good yield. In subsequent steps of the scheme however, products became difficult to purify and consequently low yielding. The protecting group was then changed to a trifluoroacetamide group. This was preferable as no purification was required before continuing with the next step in the scheme.

Once the amine was protected, the chloropyrimidine group could then be added to the molecule to make the molecule reactive with a SNAP-Tag-containing protein. This was achieved using an S_NAr reaction on a dichloropyrimidine. The dichloropyrimidine only becomes monosubstituted as the pyrimidine ring is significantly less electrophilic after the first substitution reaction. It is therefore a much more energetically favourable reaction for the nucleophilic O^- group to attack the dichloropyrimidine. As a slight excess (10%) of dichloropyrimidine was added, this favoured the desired product over the disubstituted pyrimidine. This was the only intermediate step that required purification before further use.

The chloropyrimidine containing molecule (Scheme 2.1, Compound 3) was then deprotected using the standard protocol of $MeNH_2$ in ethanol to expose the amine group (Kinderman and Schwab, 2006, Srikun et al., 2010).
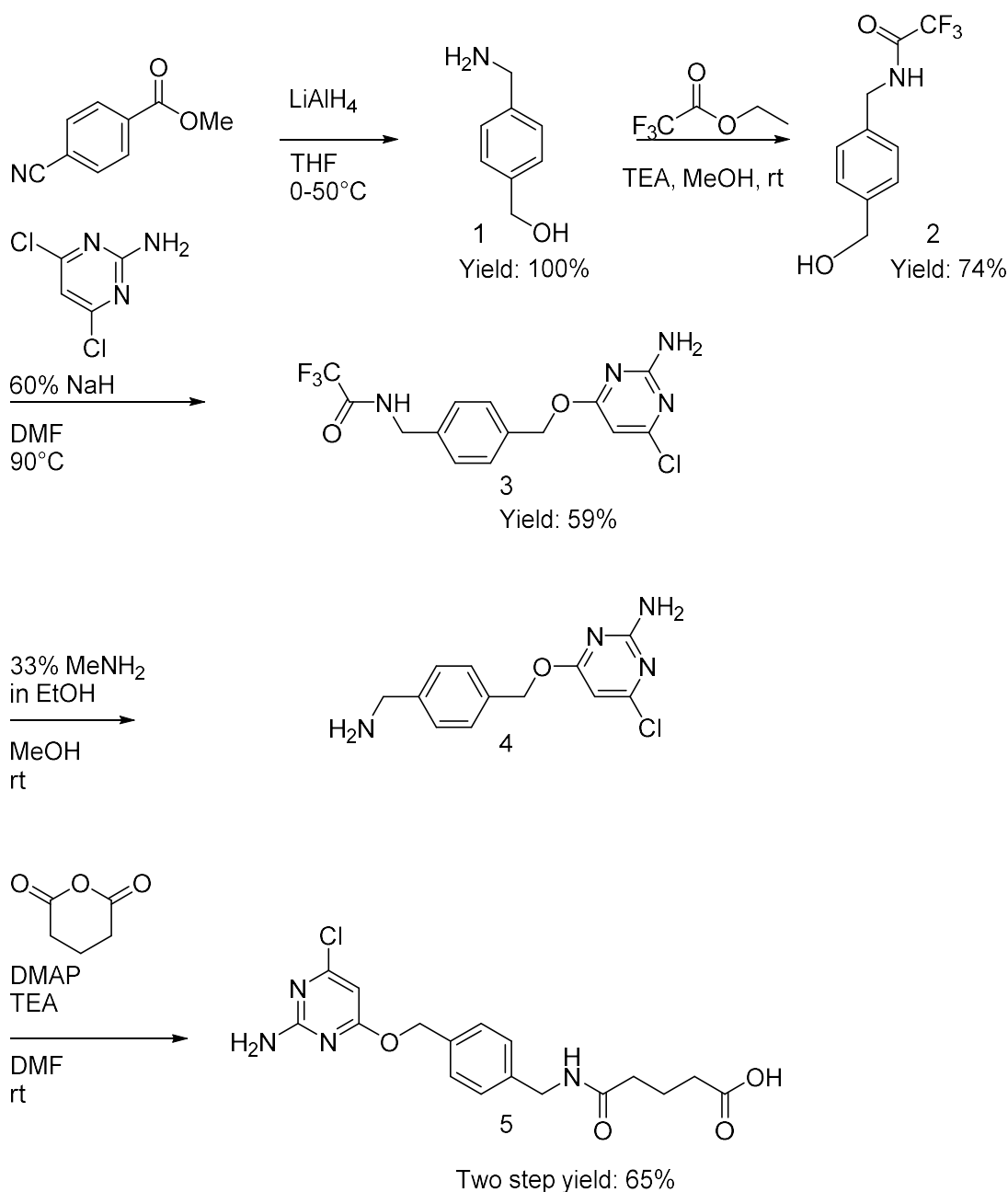
Scheme 2.1 Synthesis of SNAP-Tag substrate

To link the SNAP-Tag substrate chloropyrimidine to the PTS1 peptide a free carboxylic acid group was needed to react with the N-terminal Tyrosine of the PTS1 sequence. This was installed by reacting the free amine (Scheme 2.1, compound 4) with glutaric anhydride, using DMAP as a catalyst. In the optimisation of this reaction, the relative amount of DMAP used were decreased from 2.0 equivalents to 0.5 equivalents as analysis of the product found traces of DMAP still associated with the product as a salt. As DMAP is only required in catalytic amounts, decreasing it in the reaction allowed for better yields. The product from this reaction (Scheme 2.1, Compound 5) was purified using

reverse phase chromatography, which was found to be the best way to remove DMAP and other impurities from the crude product.

Compound 5 was then coupled to the appropriate peptide to make the desired final product.

2.3 Synthesis of the HaloTag Substrate

The HaloTag substrate molecule is, like the SNAP-Tag, a known compound up to the point of the addition of the glutaric anhydride. The first three steps were based upon a published synthetic route (Singh et al., 2013). To synthesise the substrate molecule for the HaloTag protein 2-(2-aminoethoxy)ethanol was first Boc protected. The alcohol was then reacted with 6-chloro-1-iodo-hexane to attach the chlorohexane motif that is recognised by the HaloTag protein.

The Boc group is then removed from the amine using acidic conditions. The free amine formed in this step initially proved difficult to purify by normal phase chromatography. To overcome this, a method of purification using a cation exchange column was developed. A strong cation exchange (SCX) column can be used to capture cationic species with an acidic group bonded to silica, the columns used here contained a benzenesulfonic acid functionality. The desired product could then be eluted through the addition of 2M Ammonia in methanol. Initially, this methodology was not as high yielding as predicted; this was due to the fact that the acidic deprotection reaction was quenched with K_2CO_3. The K^+ ions appeared to preferentially bind to the SCX column, meaning that the majority of the amine was found in the washes from the column. By removing the quenching step from the reaction and instead applying the crude reaction mixture to the column, a high yield of amine was obtained in the elutions from the column.

The next steps in the synthesis were identical to those used to generate the SNAP-Tag reagent (see 2.2). The amine was reacted with glutaric anhydride, purified by reverse phase chromatography and coupled to the appropriate peptide to make the desired final product.
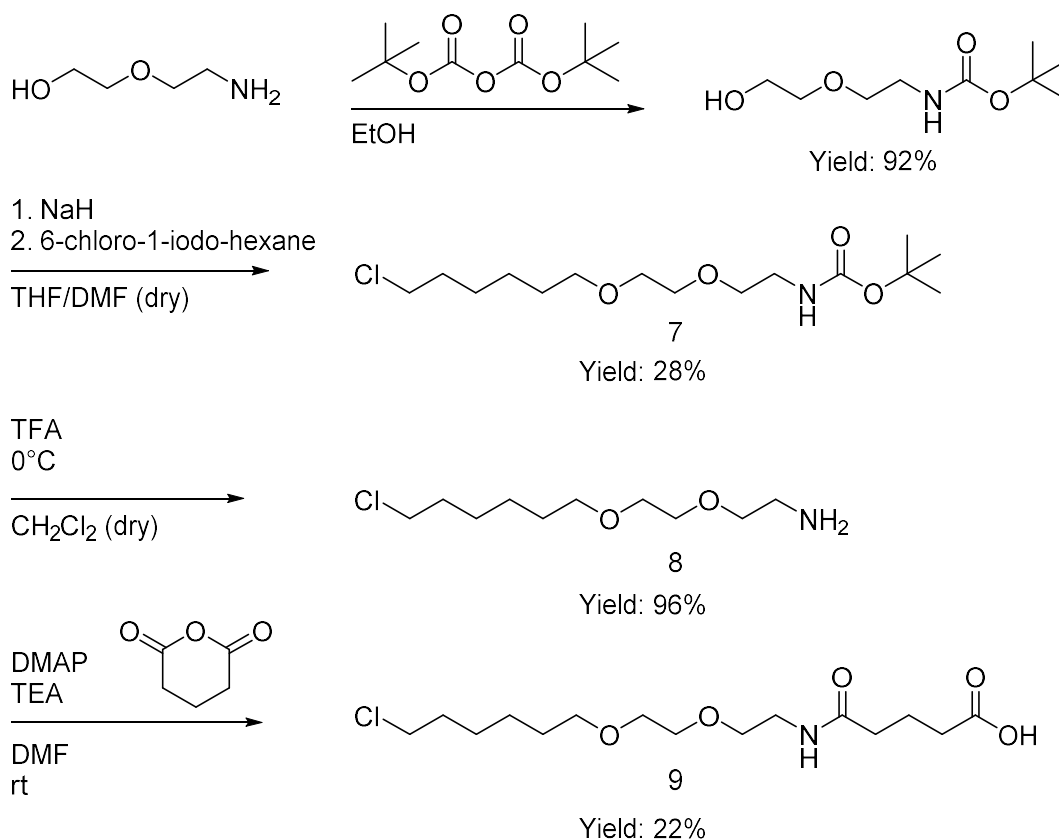
Scheme 2.2 Synthesis of HaloTag Substrate

2.4 Synthesis of PTS1 Peptide and Control Peptide

The peptide (YQSKL) was synthesised by solid phase peptide synthesis on H-Leu-2-ClTrt resin (Novabiochem®) using DIC and OxymaPure as coupling agents. The control peptide YQLKS, which is not a PTS1 sequence, was also synthesised in a similar way starting from Ser-2-ClTrt resin.

After deprotection of the N-terminal Fmoc group of Tyrosine the SNAP-Tag (Scheme 2.1, Compound 5) or HaloTag (Scheme 2.2, Compound 9) substrate was coupled to the peptide using OxymaPure and DIC (Scheme 2.3). After coupling, the final compound needed to be cleaved from the resin and all protecting groups removed from amino acid side chains. The side chains were all protected with acid-labile groups and so the deprotection and cleavage couple be carried out in one step with TFA. The final compound was found to precipitate in ice cold Et_2O allowing for any remaining TFA to be removed before purifying the compound.

Reverse phase chromatography was again found to be the best way to purify the final compound and prevent loss of yield that was seen when purifying by other methods.
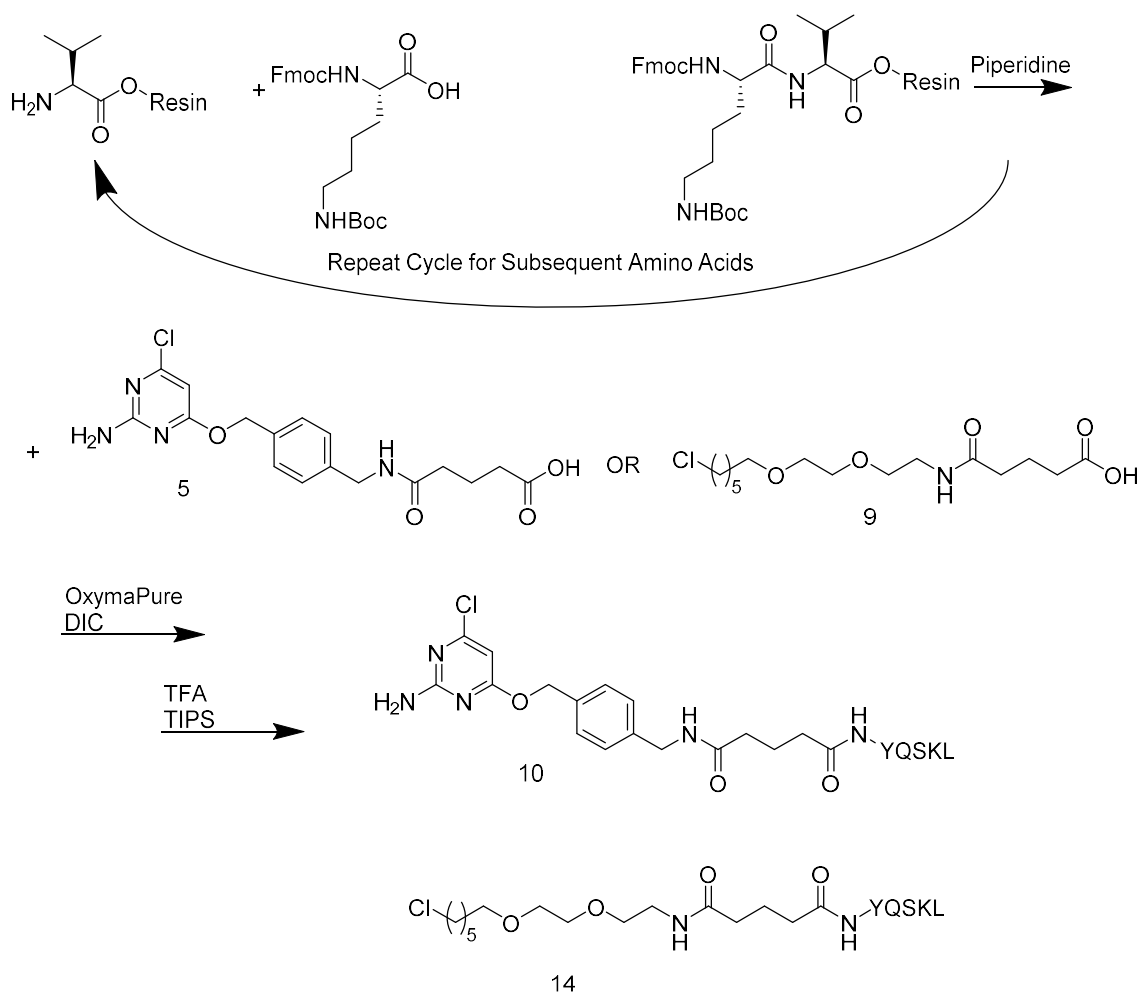
Scheme 2.3 Solid Phase Peptide Synthesis and Coupling to Substrate Molecules

With the required probes in hand, attention then turned to the preparation of the required protein constructs.

2.5 Synthesis of Recombinant Proteins

The proteins used here were encoded by vectors with a promotor recognised by T7 RNA polymerase upstream of the gene. The protocol for the expression and purification for these proteins is detailed in section 5.2.5. The proteins needed in this study were the SNAP-Tag protein, the HaloTag proteins and two different constructs of PEX5, one based on the *Arabidopsis thaliana* PEX5 protein (*At*–His6-PEX5C) and the other on the human PEX5 protein (*Hs*–His6-PEX5C).

2.5.1 Expression and Purification of SNAP-Tag Protein

The SNAP-Tag protein was encoded on a pET12b vector and contained a C-terminal Twin-Strep-tag® and ordered ready to use from GenScript (Appendix A). The Twin-Strep-tag® is a 28 amino acid sequence containing two copies of the short 8 amino acid sequence WSHPQFEK, which was found to bind to the

biotin binding site on streptavidin (Schmidt et al., 2013). The protein was purified using Strep-Tactin® resin, an engineered streptavidin with enhanced binding capability, immobilised on sepharose beads (Schmidt and Skerra, 2007, Voss and Skerra, 1997). The addition of desthiobiotin allows the removal of a protein with a Twin-Strep-tag® from the resin.

The *E.* coli cells used to express this protein contained the gene of T7 RNA polymerase under the control of a lac promoter. This means that the expression of the T7 RNA polymerase can be induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG), which relieves the promoter from the lac repressor bound to it. T7 RNA polymerase is then expressed, which in turn initiates the expression of the gene under the T7 promoter (Studier and Moffatt, 1986). The expressed protein can then be isolated from the other proteins expressed in *E. coli* using its affinity tag (the Twin-Strep-tag® in this case).

Samples were taken during each stage of protein purification to analyse the purification process on an SDS-PAGE gel (Figure 2.5A). The gel shows the protein was successfully induced and the majority of the yield found to be soluble after disruption of *E. coli* cells. Some protein was lost in the 'unbound' suggesting the Strep-Tactin® resin had reached saturation point. The wash fractions show that any untagged proteins were washed off the resin before the addition of elution buffer with desthiobiotin. The elutions show a high yield of pure protein. The protein purified had its identity verified by mass spectrometry (Figure 2.5B). This showed that the protein was the expected mass of 22.8 kDa.
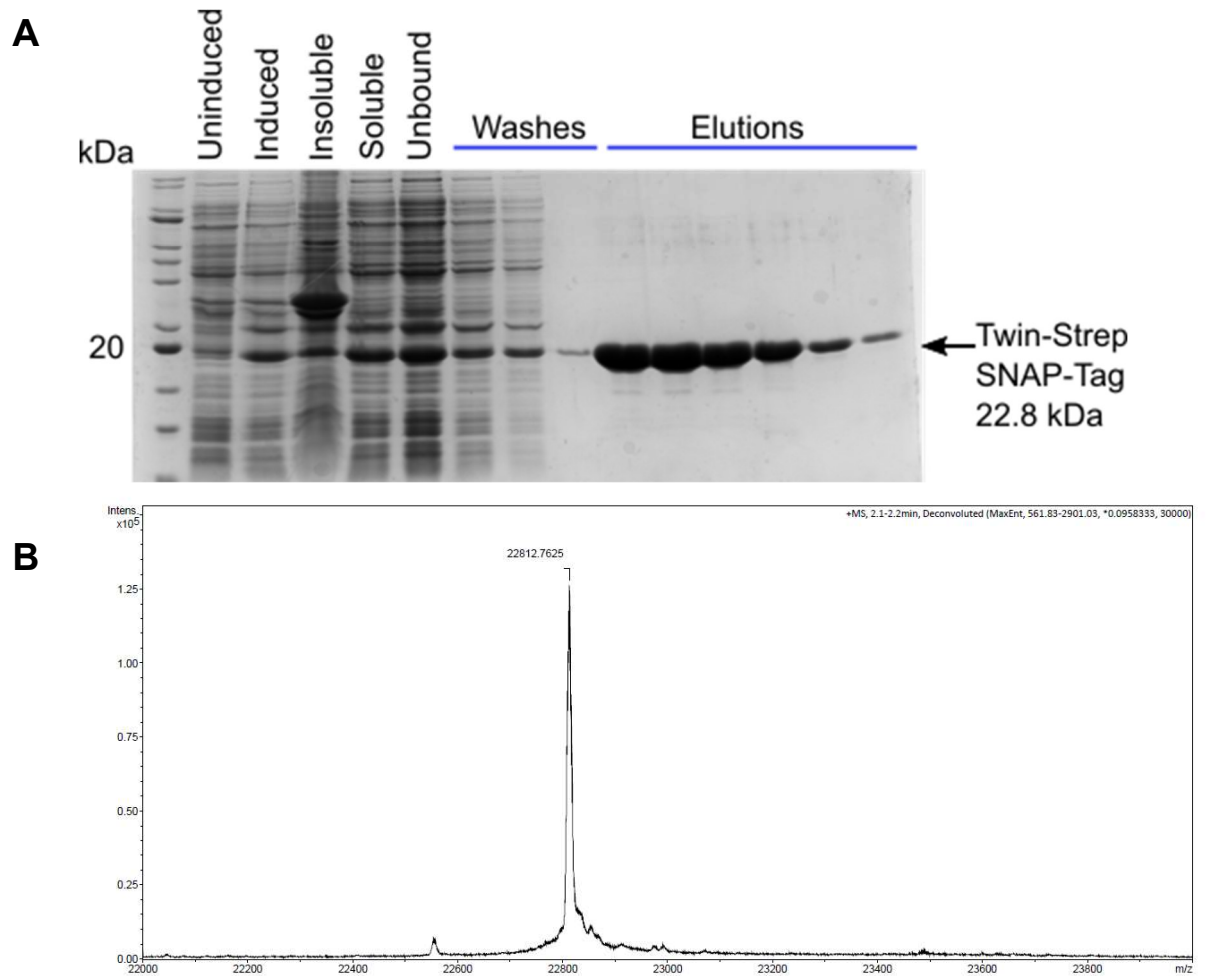
Figure 2.5 Expression, Purification and Verification of SNAP-Tag protein
A. Coomassie stained SDS-PAGE gel showing stages of SNAP-Tag protein purification. Samples were taken at various points during the purification of the protein and analysed by SDS-PAGE. **B.** The combined elutions from the purification were concentrated and the identity verified by mass spectrometry. The deconvoluted mass spectrum shows a single pure peak at the expected mass for the protein of 22.8 kDa.

2.5.2 Generation of HaloTag Protein Expression Vector

The vector to express the HaloTag protein with a C-terminal Twin-Strep-tag® was generated using the StarGate Cloning system (Selmer and Pinkenburg, 2008). This system allows for a gene to be cloned into a vector of choice in a single step and uses blue-white colony selection to determine successful cloning (see Appendix A.4.3).

2.5.3 Expression and Purification of the HaloTag protein

The HaloTag protein also contained a Twin-Strep-tag and was expressed and purified in an identical manner to the SNAP-Tag protein (see 2.5.1). Analysis of the protein purification by SDS-PAGE and mass spectrometry showed that a high yield of pure protein of the expected mass of 37.2 kDa was observed.
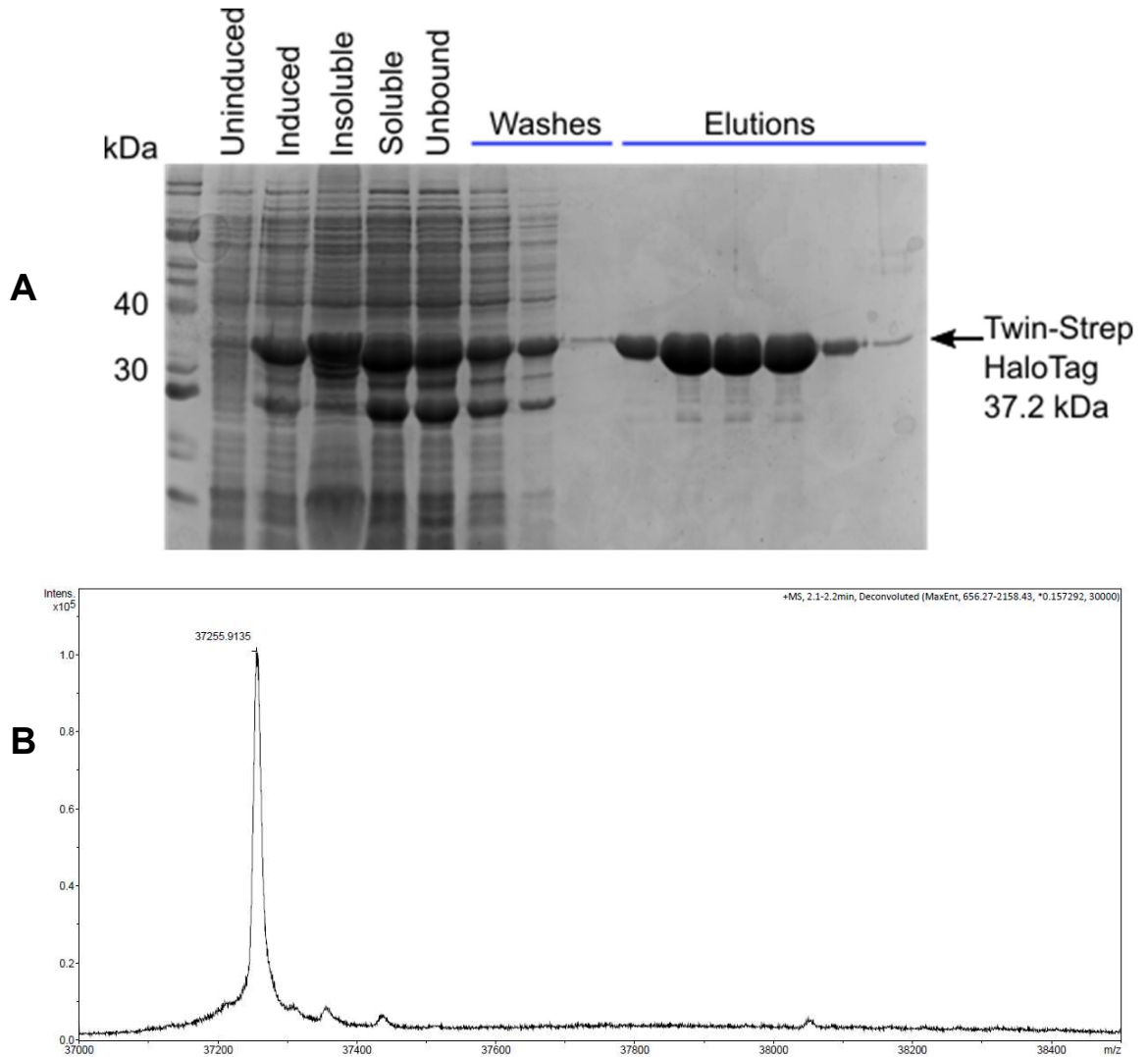
Figure 2.6 Expression, Purification and Verification of the HaloTag Protein
A. Coomassie stained SDS-PAGE gel showing stages of HaloTag protein purification. Samples were taken at various points during the purification of the protein and analysed by SDS-PAGE. **B.** The combined elutions from the purification were concentrated and the identity verified by mass spectrometry. The deconvoluted mass spectrum shows a single pure peak at the expected mass for the protein of 37.2 kDa.

2.5.4 *H. sapiens* and *A. thaliana* PEX5 constructs

The plasmids for the expression of the two PEX5 proteins used in this study were designed to include the TPRs of the C terminal domain, as these are responsible for binding the PTS1 sequence. Previous work in the group found the full-length *Arabidopsis thaliana* PEX5 protein to be less stable for *in vitro* studies than the *At*-His$_6$-PEX5C construct. The *At*-His$_6$-PEX5C construct also bound PTS1 peptides with a similar affinity to the full length protein (Skoulding et al., 2015). The *Hs*-His$_6$-PEX5C construct used in this project was designed based on the construct used by Gatto Jr. et al. when producing the crystal structure of *Hs*-His$_6$-PEX5C bound to the PTS1 peptide (PBD ID: 1FCH) with the addition of an N-terminal His$_6$-tag to aid purification. The *At*-His$_6$-PEX5C construct was cloned previously for use in similar such projects at the University of Leeds (Lanyon-Hogg et al., 2014). A schematic of both PEX5C constructs used in this study are shown in Figure 2.7.
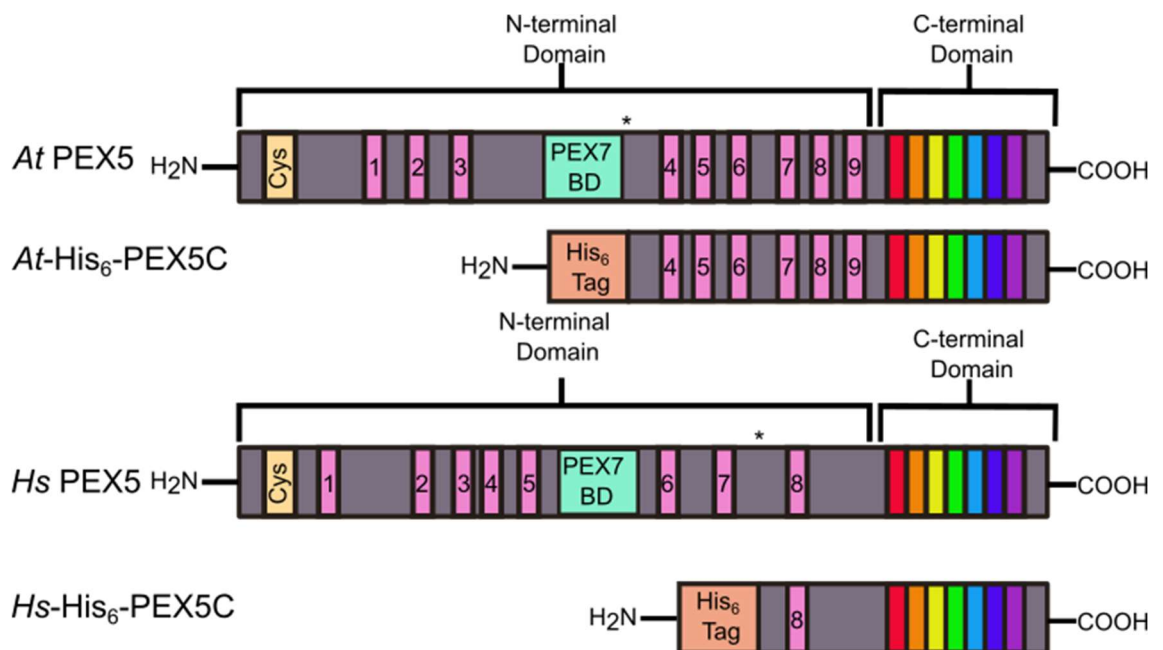
Figure 2.7 Domains found in PEX5C proteins from *A. thaliana* and *H. sapiens* A schematic showing the C-terminal domain present in the PEX5C proteins from *Homo sapiens* and *Arabidopsis thaliana*. WXXXF motifs are shown in pink. The TPR domains are shown in colour corresponding to those used in the crystal structure (Figure 1.6). The asterisk on each structure indicates the position at which the protein is truncated in the PEX5C constructs.

2.5.5 Expression and purification of PEX5 constructs

The genes encoding the PEX5C constructs were encoded in pET28b plasmids fused to a His$_6$-tag at the N-terminus.

Expression was performed by autoinduction. Purification of proteins was then achieved by affinity chromatography using cobalt-agarose resin. The stages of purification of *Hs*-His$_6$-PEX5C (Figure 2.8) and *At*–His$_6$-PEX5C (Figure 2.9) were analysed by SDS-PAGE and the protein identity verified by mass spectrometry. The SDS-PAGE analysis showed that the protein elutions for both proteins were pure and did not require any further purification steps. The mass spectrum of *At*-His$_6$-PEX5C showed the main protein peak at 45.5792 kDa, as was expected for this protein. The mass spectrum for *Hs*-His$_6$-PEX5C showed the main protein peak to be 42.9778 kDa, as would be expected for this protein.
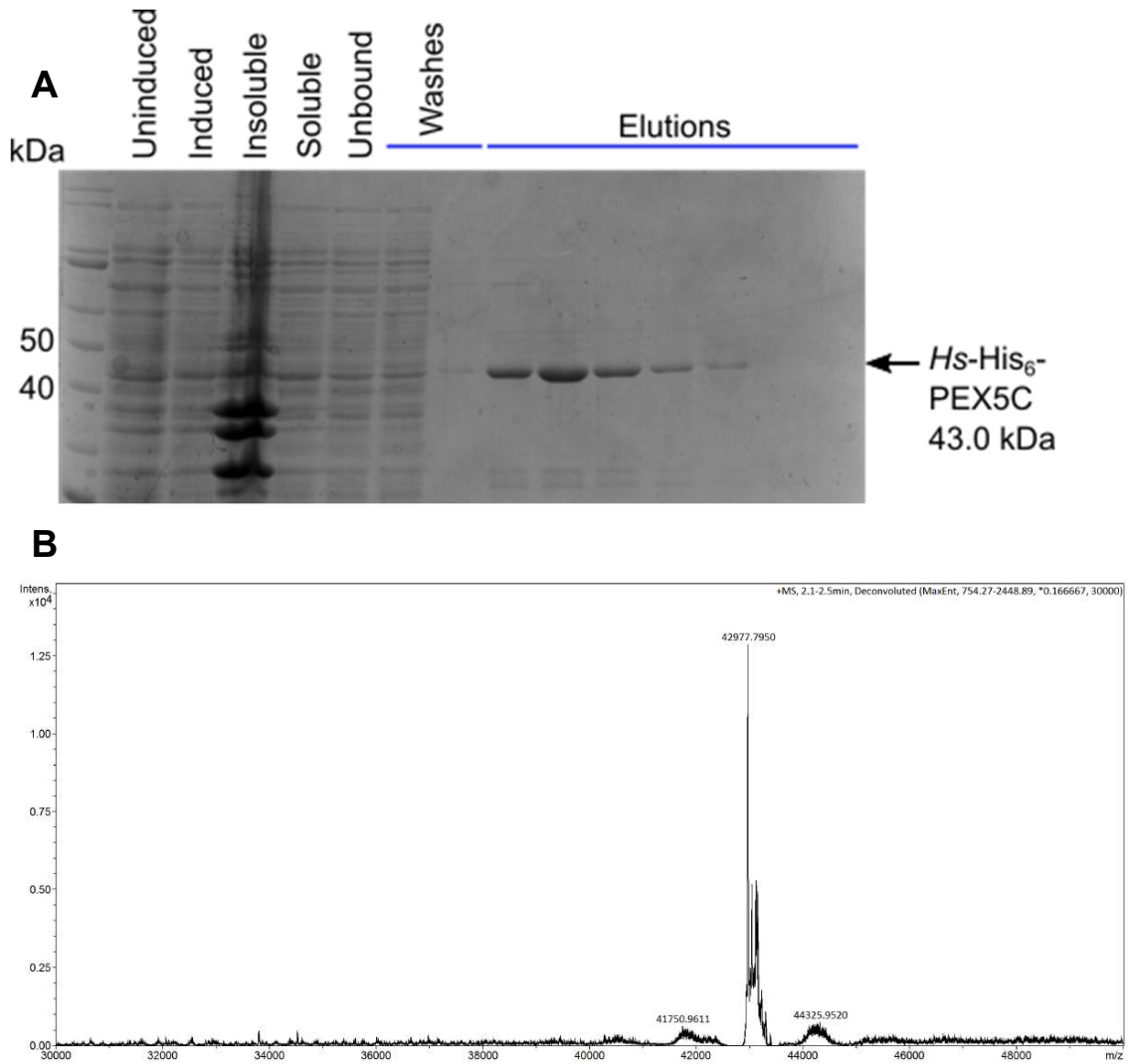
Figure 2.8 Purification and Verification of *Hs*-His₆-PEX5C protein A.
Coomassie stained SDS-PAGE gel showing stages of *Hs*-His₆-PEX5C protein purification. Samples were taken at various points during the purification of the protein and analysed by SDS-PAGE. **B.** The combined elutions from the purification were concentrated and the identity verified by mass spectrometry. The deconvoluted mass spectrum shows a single pure peak at the expected mass for the protein of 43.0 kDa.
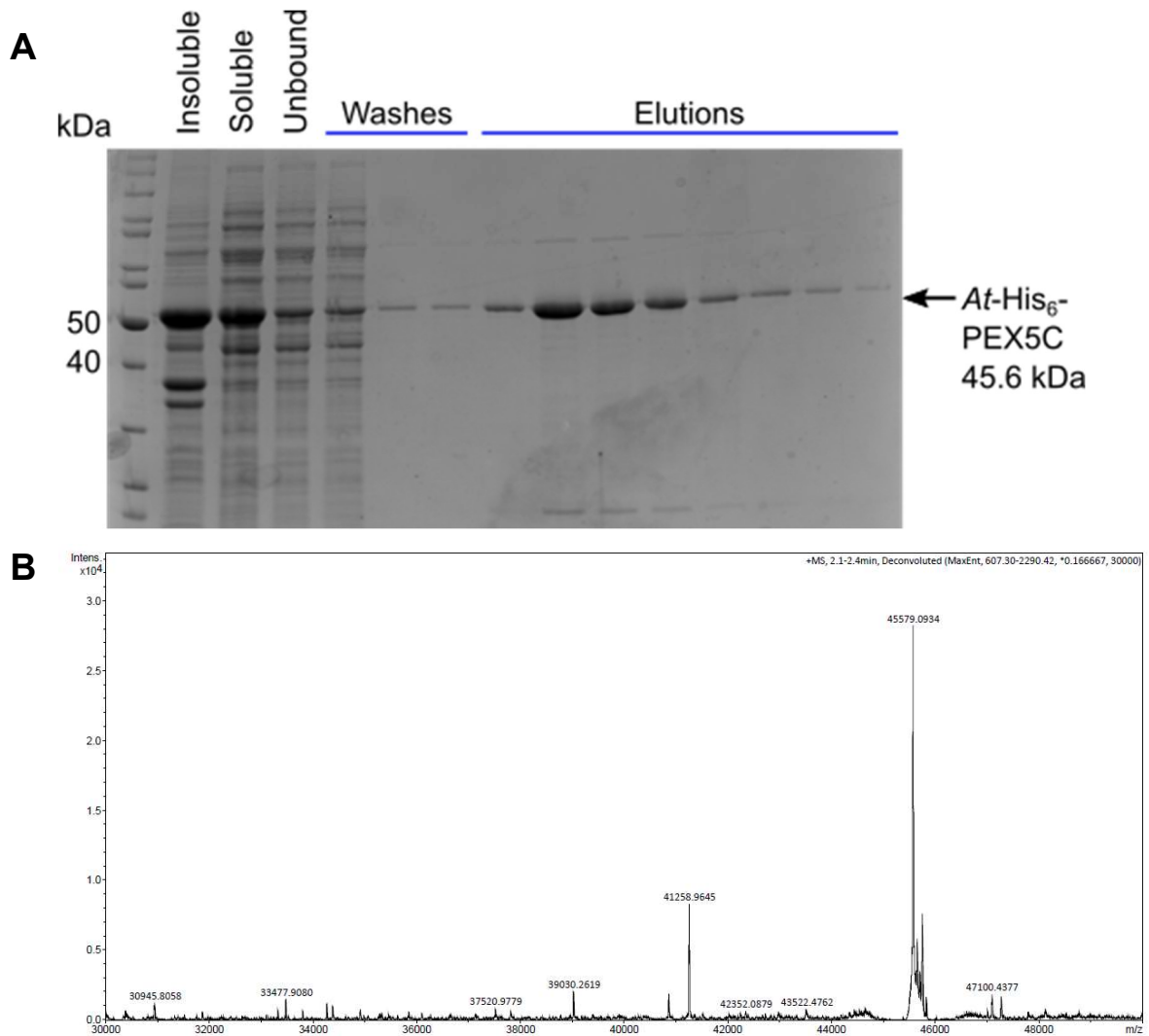
Figure 2.9 Purification and Verification of *At*-His₆-PEX5C protein A. Coomassie stained SDS-PAGE gel showing stages of *At*-His₆-PEX5C protein purification. Samples were taken at various points during the purification of the protein and analysed by SDS-PAGE. **B.** The combined elutions from the purification were concentrated and the identity verified by mass spectrometry. The deconvoluted mass spectrum shows a single pure peak at the expected mass for the protein of 45.6 kDa.

2.6 Labelling of SNAP-Tag and HaloTag Proteins with Reactive PTS1 Peptide

Once purified recombinant proteins and tag-reactive peptides had been synthesised, their interactions with one another could be studied. First, the reactivity of the peptide substrates with their relative tag proteins was studied by mass spectrometry. To do this, the SNAP-Tag and HaloTag proteins were incubated with the peptide containing the corresponding reactive motif. In the case of the SNAP-Tag, DTT was also added to the reaction mixture to ensure the reactive cysteine was available and not forming a disulphide bond. After the proteins had been incubated with a three times excess of peptide for 1 hour on ice, the resulting protein species were analysed by mass spectrometry. For brevity, the PTS1 peptide of the reactive peptides is represented here by the one letter amino acid codes for the residues (Figure 2.10A). It was found in these reactions that the peptides were able to efficiently label their respective proteins shown by the increase in protein mass seen in the deconvoluted mass spectra for the proteins before and after incubation with the peptides (Figure 2.10B and C). The expected mass increase was calculated as the mass of the reactive peptide, minus the mass of the groups lost in the reaction with the reactive residue in the protein.

Reactive peptides with the sequence YQLKS (compounds 11 and 15) were also tested to act as negative controls as YQLKS is not a PTS1 sequence but otherwise is similar to the YQSKL peptides.

This next stage was to determine if the PTS1 peptides, now covalently attached to the proteins at a site other than the C-terminus, could be recognised and bound by PEX5 proteins.

After it had been confirmed that the SNAP-Tag and HaloTag proteins had been successfully labelled by the peptides, the excess peptide was removed from the reaction. This was achieved using a centrifugal filter with a molecular weight cut-off of 10,000 Da. This allowed the reactive peptide reagent to pass through the filter but retained the protein, therefore separating the two species. This was necessary as the free reactive peptide labelling reagent would also be a substrate for PEX5 binding and therefore would interfere with the binding of the labelled SNAP-Tag and HaloTag proteins to PEX5.
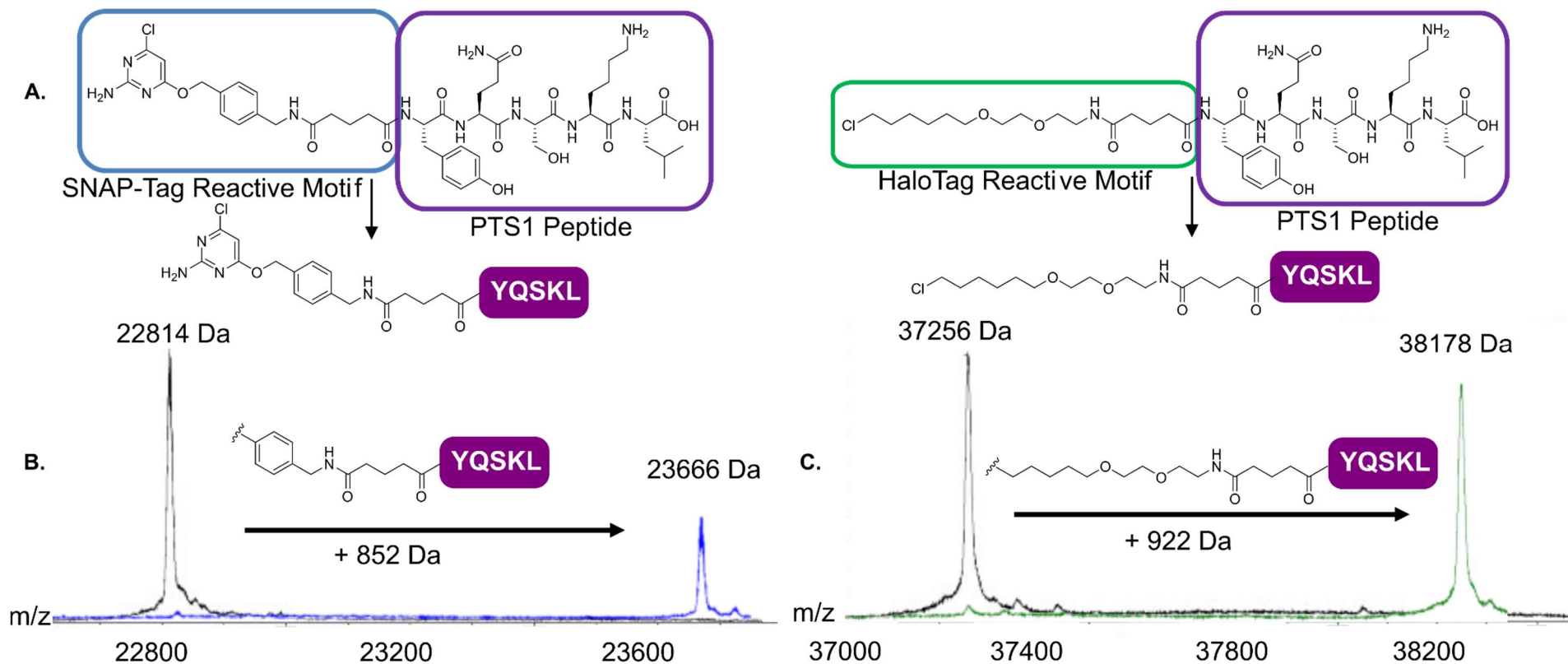
Figure 2.10 Labelling of SNAP-Tag and HaloTag proteins with PTS1 peptides

A. The PTS1 peptide YQSKL was attached to SNAP-Tag or HaloTag reactive motifs. The PTS1 peptide portion of the reagents has been abbreviated to the one letter amino acid code to simplify their appearance. The equivalent YQLKS peptides were also synthesised and tested as controls **B.** The SNAP-Tag protein increased in mass by 852 Da, the mass of the SNAP-Tag reactive peptide, minus the O^6-chloropyrimidine group **C.** The HaloTag protein increased in mass by 922 Da, the mass of the HaloTag reactive peptide minus the chloride

2.7 Pulldown assays of PEX5C proteins with PTS1 peptide labelled SNAP-Tag and HaloTag Proteins

To assess if the PEX5 proteins could recognise the PTS1 peptide attached to proteins via a SNAP-Tag or HaloTag, the Twin-Strep-tag on the SNAP-Tag and HaloTag proteins was utilised.

The PTS1- and control-labelled SNAP-Tag and HaloTag proteins were incubated with either Hs-His$_6$-PEX5C or At-His$_6$-PEX5C in a 1:2 ratio for one hour on ice. MagStrep XT beads, magnetic beads coated in Strep-Tactin® XT, a form of Strep-Tactin® with higher binding capacity, were equilibrated and the protein mixture applied to the beads.

The use of magnetic beads was chosen here as it was found that the use of sepharose Strep-Tactin® resin gave some contamination of the washes and elutions with resin as it was difficult not to disturb the resin pellet when removing fractions. The magnetic beads do not require pelleting by centrifugation but instead a magnet can be used to separate the beads from the surrounding buffer.

The protein mixture was incubated with the beads for one hour with periodic gentle agitation to bring the beads back into suspension. After this time, the beads were washed and then proteins bound to the beads eluted (see Materials and Methods).

The washes and elutions from the beads were then analysed by SDS-PAGE. As the SNAP-Tag and HaloTag proteins both contained the Twin-strep-tag, it was expected that they would be present in elution fractions. If the PEX5C protein could also be found in these elution fractions, the two proteins would need to be interacting in some way as the PEX5C proteins do not contain a tag bound by Strep-Tactin®.

The At-His$_6$-PEX5C protein was found to co-elute from the MagStrep XT beads when in the presence of the YQSKL-labelled SNAP-Tag and HaloTag proteins (Figure 2.11A), but importantly this interaction appeared to require the PTS1 peptide, as At-His$_6$-PEX5C did not co-elute with the control peptide (YQLKS). This suggests that the TPR domains of At-His$_6$-PEX5C are recognising the PTS1 peptide attached to the SNAP-Tag and HaloTag proteins, even though the peptide is not at the protein's C-terminus.

However, the Hs-His$_6$-PEX5C protein did not give the same result. This protein was not seen to co-elute with the YQSKL-labelled SNAP-Tag and HaloTag proteins (Figure 2.11B). This was an unexpected result, particularly as the At-

His$_6$-PEX5C protein was seen to bind, and the YQSKL peptide has been shown to bind a similar truncated form of *Homo sapiens* PEX5 in crystal structures (Gatto Jr et al., 2000). Further investigation was then carried out to discover the reasons behind the lack of binding of *Hs*-His$_6$-PEX5C to the peptide when *At*-His$_6$-PEX5C bound well.
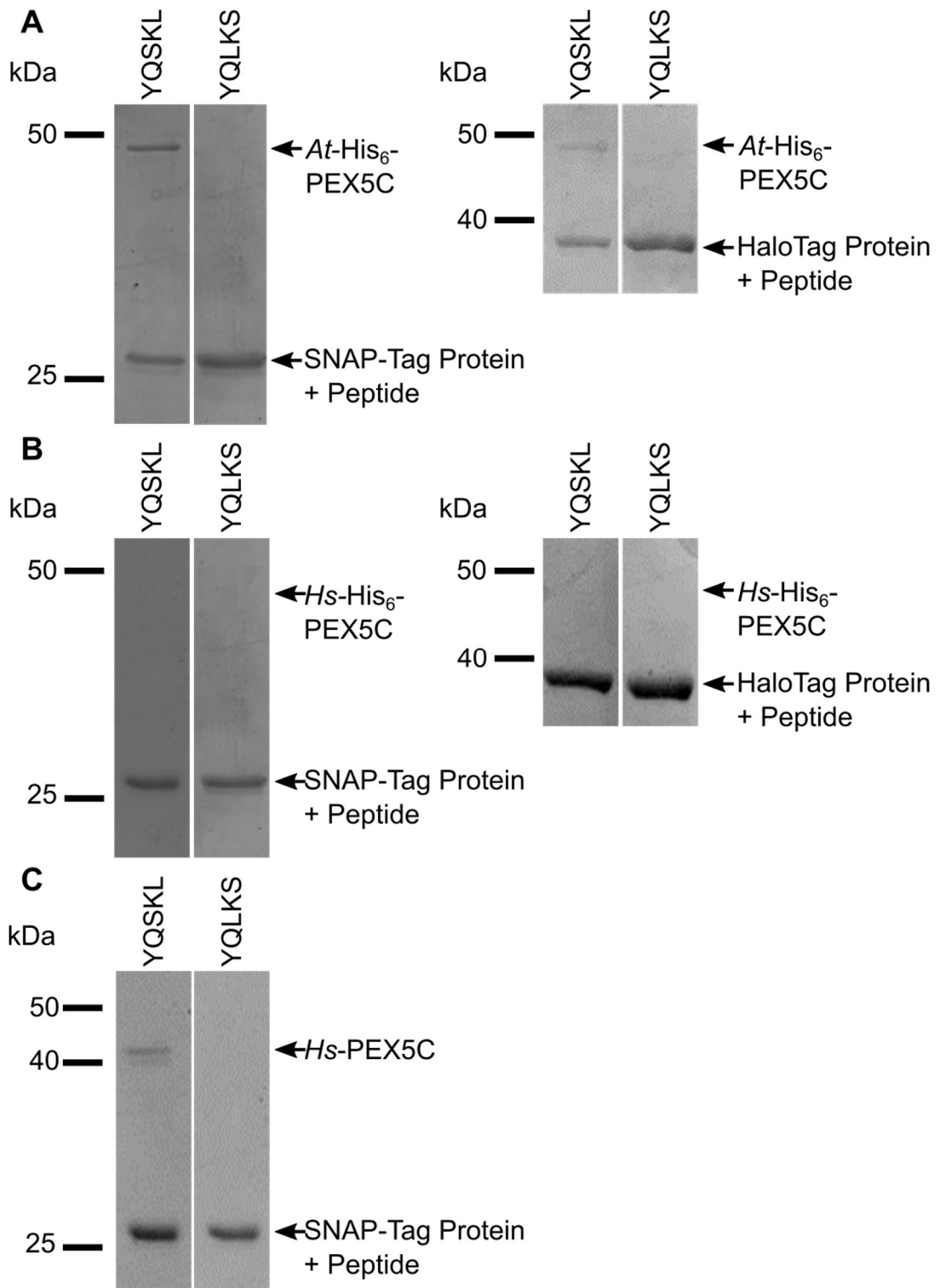
Figure 2.11 Coomassie Stained Elution Fractions from Pulldown Assays Analysed by SDS-PAGE A. *At*-His$_6$-PEX5C was seen to co-elute with the YQSKL-labelled SNAP-Tag (left) and HaloTag (right) proteins, but not with YQLKS-labelled proteins. **B.** *Hs*-His$_6$-PEX5C did not co-elute with SNAP-Tag or HaloTag proteins labelled with either the PTS1 peptide YQSKL or the control peptide (YQLKS).

2.8 Fluorescence Anisotropy of PEX5C proteins with PTS1 Peptide

The *Hs*-His$_6$-PEX5C protein was shown to be alpha helical by circular dichroism with characteristic absorbance at 208 and 222 nm and a calculated alpha helical character of >90% (Andrade et al., 1993, Whitmore and Wallace, 2004). Therefore, it was correctly folded to be able to bind to PTS1 peptides. It also bore significant sequence similarity to the *Hs*-PEX5C protein used by Gatto Jr et al., which has been shown to bind to a YQSKL peptide. And yet it appeared that the *Hs*-His$_6$-PEX5C used here did not bind to the YQSKL peptide in the pulldown assay.
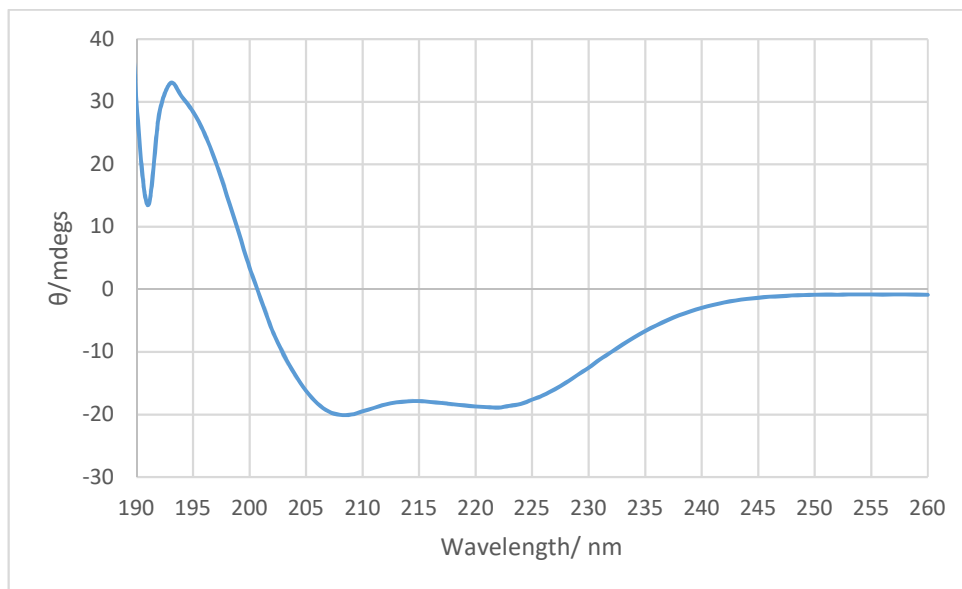
Figure 2.12 Analysis of *Hs*-His$_6$-PEX5C protein using circular dichroism
The secondary structure of the *Hs*-His$_6$-PEX5C protein is analysed using circular dichroism and shows a typical spectrum indicative of an alpha helical protein.

To investigate the binding affinity of the different PEX5 constructs for the peptide YQSKL, fluorescence anisotropy was used with a fluorescent form of the peptide, Lissamine-YQSKL.

We found that the *At*-His$_6$-PEX5C protein bound with a EC$_{50}$ of approximately 40 nM, (Figure 2.13). The *Hs*-His$_6$-PEX5C however, showed a binding affinity of >5000 nM. (Figure 2.13).

When taken together with the results on the pulldown assay, this result corroborates the fact that the *At*-His$_6$-PEX5C protein bound to the YQSKL-labelled proteins and *Hs*-His$_6$-PEX5C did not.
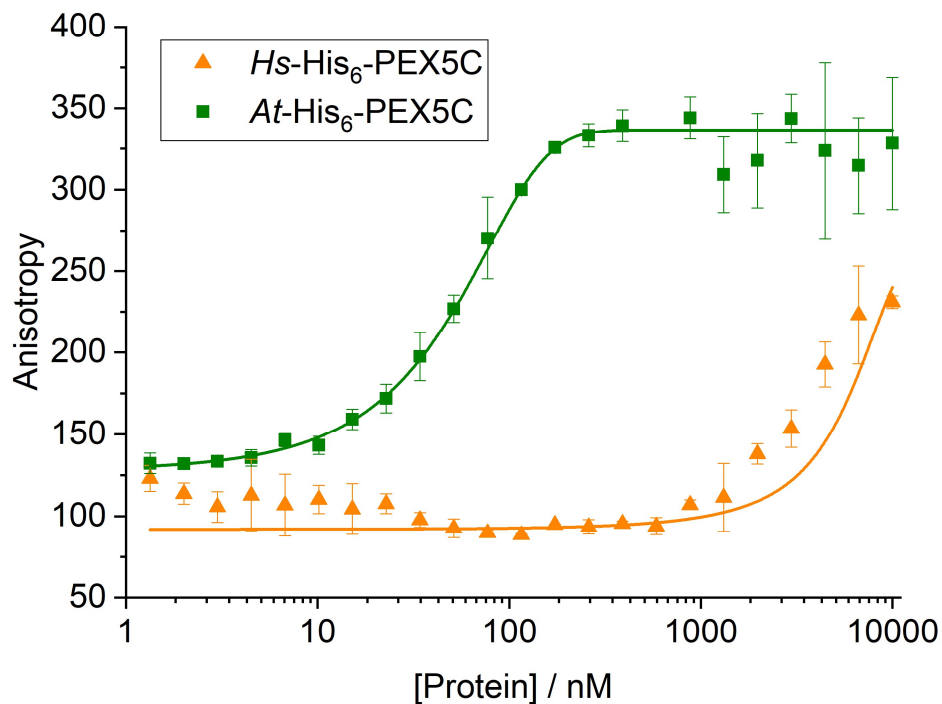
Figure 2.13 Fluorescence Anisotropy of *At*-His₆-PEX5C and *Hs*-His₆-PEX5C with the Lissamine-YQSKL peptide *At*-His6-PEX5C and *Hs*-His6-PEX5C proteins ranging in concentration for 0 – 10000 nM were incubated with the 500 μM fluorescent peptide Lissamine-YQSKL and the change in the anisotropy of the peptide measured for each protein concentration.

The difference between the *Hs*-His6-PEX5C protein and the *Hs*-PEX5C used by Gatto Jr et al. was then studied more closely. The 368 amino acid residues taken from the *Hs* PEX5 protein were identical in both constructs, excepting a threonine to isoleucine base change at residue 388 and the construct used in this project was the addition of an N-terminal His6-tag and a thrombin cleavage site between the tag and PEX5 sequence an addition of 20 amino acids and 2.181 kDa. It was then hypothesised that this tag could be interfering with the binding of the *Hs*-His6-PEX5C protein to the YQSKL peptide. For this reason, the effect of removing the His6-tag from the *Hs*-His6-PEX5C construct was investigated.

2.9 Cleavage of N-terminal His₆-tag from *Hs*-His₆-PEX5C

As a thrombin cleavage site was included in the *Hs*-His6-PEX5C when designing its expression plasmid, this made the removal of the His6-tag from the construct a straightforward process. Purified *Hs*-His6-PEX5C was incubated with thrombin sepharose beads (BioVision Inc.) and samples removed from the reaction at a number of time points to monitor the progression of the cleavage. Near 100% cleavage was achieved after 4 hours and the beads were then

removed through pelleting them by centrifugation. To remove the cleaved tag and any remaining uncleaved protein, the resulting mixture was incubated with Co-NTA resin, which binds to His_6-tags with high affinity. The cleaved protein was then analysed by SDS-PAGE and mass spectrometry to confirm its identity. It was found that the protein mass after the cleavage was the expected mass of 41.2 kDa, a loss of 1.8 kDa, consistent with the loss of the 17 N-terminal amino acids before the thrombin cleavage site (see A.1.2).
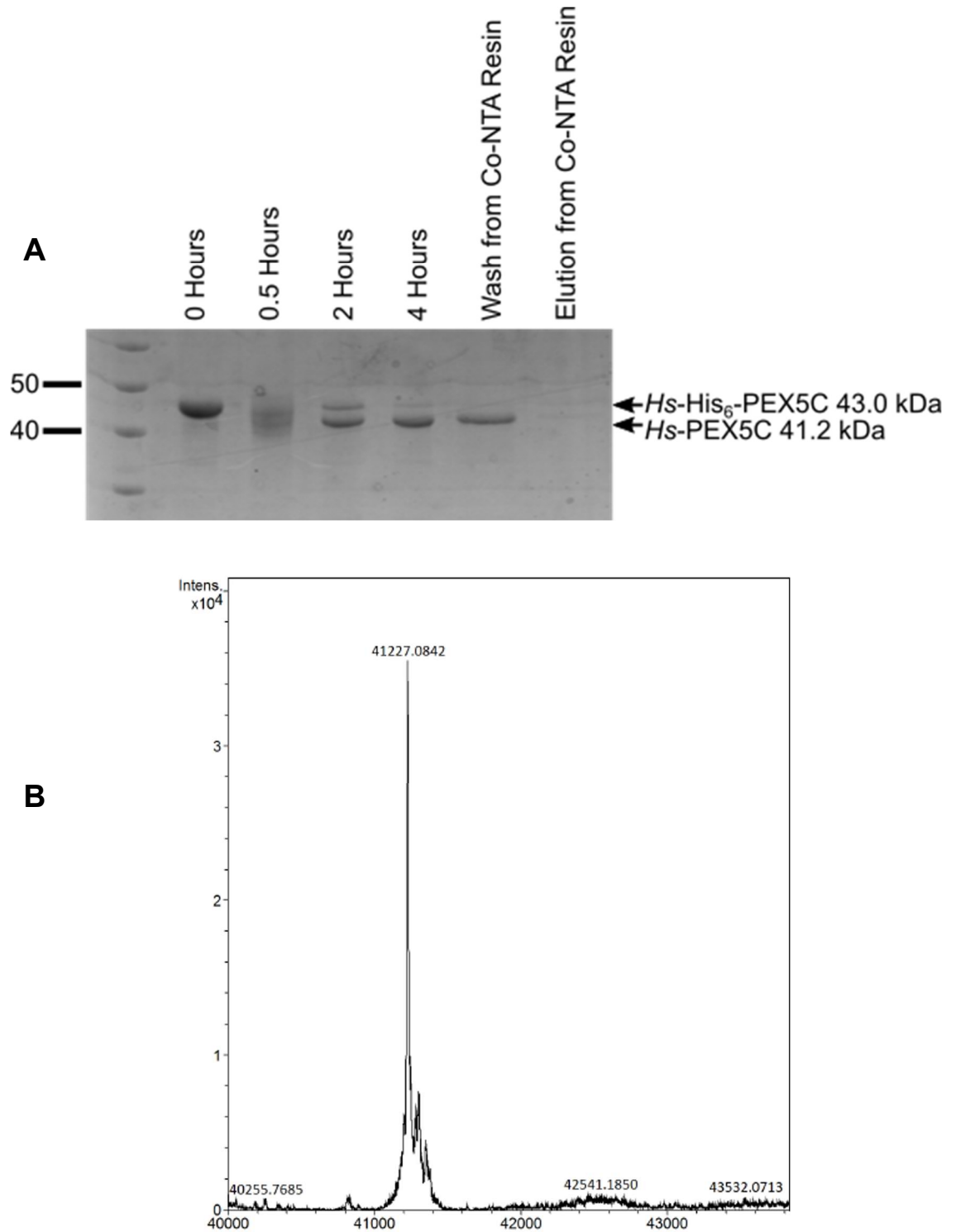
Figure 2.14 Cleavage of His$_6$-tag from *Hs*-His$_6$-PEX5C protein A. Coomassie Stained SDS PAGE gel monitoring the progress of cleacage of His$_6$-tag from *Hs*-His$_6$-PEX5C protein **B.** The cleaved protein was analysed by mass spectrometry and a protein of the predicted mass (41.2 kDa) was observed.

2.10 Fluorescence Anisotropy of Cleaved *Hs*-PEX5C proteins with PTS1 Peptide

After the cleavage of the His$_6$-tag from the *Hs*-His$_6$-PEX5C protein, the fluorescence anisotropy experiments were repeated as before. In this case the binding of *At*-His$_6$-PEX5C, *Hs*-His$_6$-PEX5C and *Hs*-PEX5C (the His$_6$-tag cleaved protein) to Lissamine-YQSKL was measured (Figure 2.15). In this instance, it was found that the binding of *Hs*-PEX5C to the peptide was significantly improved compared to *Hs*-His$_6$-PEX5C. The *Hs*-PEX5C protein gave a EC$_{50}$ of 232 nM in contrast to the >5000 nM seen for the *Hs*-PEX5C protein.
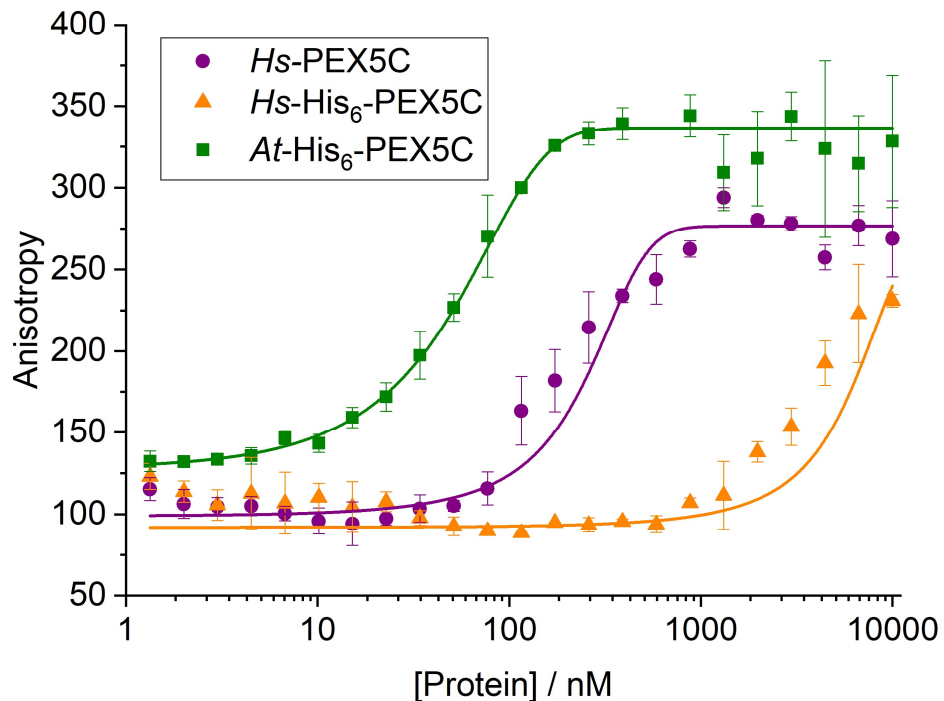
Figure 2.15 Fluorescence Anisotropy of *At*-His$_6$-PEX5C, *Hs*-His$_6$-PEX5C and *Hs*-PEX5C with the Lissamine-YQSKL peptide *At*-His$_6$-PEX5C and *Hs*-His$_6$-PEX5C and *Hs*-PEX5C proteins ranging in concentration for 0 – 10000 nM were incubated with the 500 µM fluorescent peptide Lissamine-YQSKL and the change in the anisotropy of the peptide measured for each protein concentration.

2.11 Pulldown Assays of cleaved *Hs*-PEX5C with PTS1 labelled SNAP-Tag and HaloTag Proteins

As it had now been found that the cleavage of the His_6-tag from the *Hs*-PEX5C protein improved its binding to YQSKL in fluorescence anisotropy assays, it then needed to be investigated whether that His_6-tag on the *Hs*-His_6-PEX5C protein was also preventing the protein from binding the PTS1-labelled SNAP-Tag and HaloTag proteins in the pulldown assays. The protocol for the pulldown assays was repeated as before but instead incubating the labelled SNAP-Tag and HaloTag proteins with the *Hs*-PEX5C protein after the cleavage had been performed. These pulldown assays (Figure 2.16) showed that the *Hs*-PEX5C protein co-eluted with the YQSKL-labelled SNAP-Tag and HaloTag proteins, but importantly not with the YQLKS control labelled proteins. This result reinforced the importance of the specific recombinant protein construct for the PEX5 binding of a PTS1 peptide.
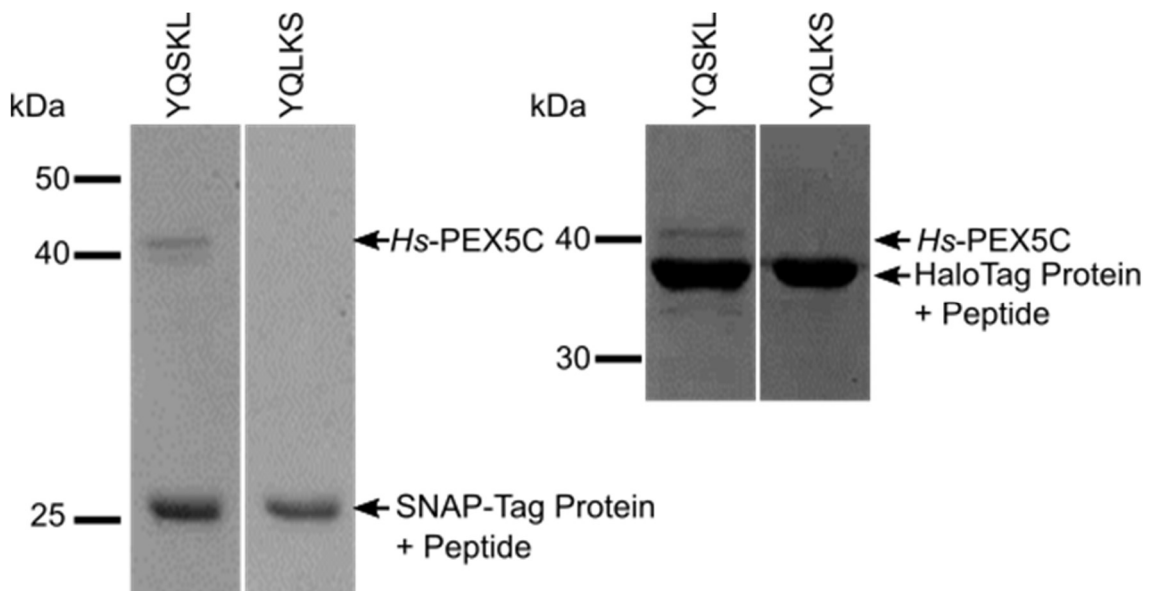
Figure 2.16 Pulldown assays with *Hs*-PEX5C After cleavage of the His_6-tag from *Hs*-His_6-PEX5C its ability to interact with YQSKL-labelled SNAP-Tag and HaloTag proteins was assessed. It was found that *Hs*-PEX5C co-eluted with the YQSKL-labelled SNAP-Tag and HaloTag proteins, but not with the YQLKS-labelled proteins, demonstrating the specificity of the interaction.

2.12 Using SNAP-Tag and HaloTag-Reactive PTS1 Peptides in Mammalian Cells

In parallel to investigations into the *in vitro* interactions of PEX5 proteins with proteins labelled with PTS1 peptides via SNAP-Tag or HaloTag, the ability to assess the interaction in live cells was also being developed. The same SNAP-Tag and HaloTag-reactive peptide reagents could be used in this case, with the expectation that they would label SNAP-Tag and HaloTag proteins being expressed inside cells. PEX5 is expressed endogenously in eukaryotic cells and the aim of this part of the project was to find out if the peptide reagents could enter cells, label their target tag protein, and facilitate the import of that protein into the peroxisomes present in the cells. The cells used in this study are COS-7 cells, which have previously been used in a variety of studies surrounding peroxisomes, particularly in the Schrader research group at the University of Exeter. For this reason, the cell-based experiments were carried out at the University of Exeter under the guidance of Michael Schrader and his team of researchers.

2.12.1   Generation of SNAP-Tag-GFP and HaloTag-GFP mammalian cell expression vectors

To express protein with SNAP-Tags and HaloTags inside mammalian cells, the genes for the SNAP-Tag and HaloTag proteins were cloned into an expression vector compatible for mammalian cell expression. In order to be able to visualise the proteins within the cells, the tag proteins were fused to Enhanced Green Fluorescent Protein (EGFP). This was achieved by cloning the SNAP-Tag and HaloTag protein genes into the expression vector pEGFP-C1 which contains the gene for EGFP under the control of a human cytomegalovirus (CMV) promotor, which allows constitutive expression in mammalian cells (Morgan, 2014). The plasmid contains a multiple cloning site at the C-terminus of the EGFP gene, to allow the cloning of a gene of interest to create a vector to express the protein of interest as a fusion protein with EGFP (Appendix A).

To generate the mammalian cell expression vectors, the genes for the Twin-Strep-tagged SNAP-Tag and HaloTag proteins were amplified from their vectors for expression in *E. coli*. The primers used in this amplification also included DNA encoding restriction sites for endonucleases that would allow the gene to be cloned into the pEGFP-C1 acceptor vector. The amplified genes and acceptor vector were both cut with the same two restriction endonucleases to create compatible sticky ends. The SNAP-Tag or HaloTag gene was then mixed with the cut vector and DNA ligase to create the final expression vector. The

ligated vectors were transformed into XL-10 cells to amplify them and the purified plasmid had its sequence verified by Sanger DNA sequencing.

2.12.2   Using SNAP-Tag and HaloTag-Reactive PTS1 Peptides in COS-7 cells

Once the plasmids for the expression of EGFP-SNAP-Tag and EGFP-HaloTag proteins had been generated, they could be used in mammalian cells. The cells used here were COS-7 cells. COS-7 are kidney cells from the African green monkey. They are adherent, fibroblast-like cells that grow well under standard cell culturing conditions and have a doubling time of 35-48 hours. Importantly COS-7 cells have been used for the immunofluorescent imaging of peroxisomes, as well as a number of different organelles in other studies (Valm et al., 2017, Schrader, 2001). This makes them appropriate for use in this study.

2.12.2.1   Transfection of Cells

The cells needed first to be transfected with the plasmids for the expression of either EGFP-SNAP-Tag or EGFP-HaloTag proteins. To do this, the transfection reagent diethylaminoethyl-Dextran (DEAE-D) was used. This is a commonly used transfection reagent as it allows for the rapid and simple transfection of plasmid DNA into cells at a low cost relative to alternatives (Gulick, 2003) and it has been previously established in the Schrader laboratory to give the desired result in a reproducible manner. The DEAE-D is a cationic polymer that forms a complex with the negatively charged plasmid DNA to form particles with an overall positive charge. This complex is adsorbed to the negatively charged cell membrane and endocytosed across the membrane (Schenborn and Goiffon, 2000).

As the transfected plasmids should result in the constitutive expression of a protein with EGFP fluorescence, successful transfection could be determined by the presence of fluorescence in the cells. Proteins containing EGFP were observed both in cells transfected with the pEGFP-SNAP-Tag and pEGFP-HaloTag plasmids. As the plasmids had been verified as encoding EGFP-SNAP-Tag and EGFP-HaloTag by DNA sequencing, it could be concluded that these proteins were being expressed in these cells.
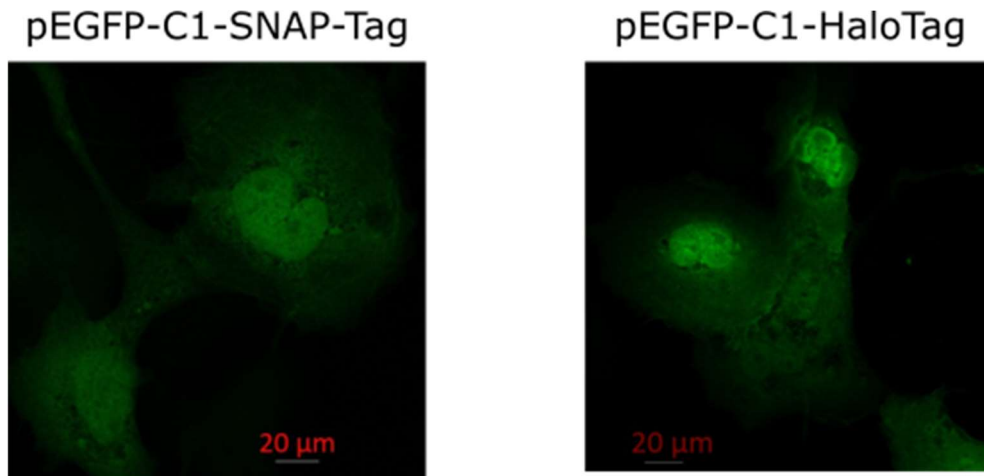
Figure 2.17 COS-7 cells transfected with pEGFP-C1-SNAP-Tag or pEGFP-C1-HaloTag plasmids Cells were transfected with plasmid DNA encoding for the expression of either EGFP-SNAP-Tag or EGFP-HaloTag proteins. The successful transfections of cell and expression of the proteins is observed here by the green fluorescence present in cells.

2.12.2.2   Incubation with Peptides

Once cells had been successfully transfected with the SNAP-Tag or HaloTag-encoding plasmids, they were incubated with the reactive peptide reagents. The strategy behind the experiment was that the peptide would be taken up into the cells and the SNAP-Tag or HaloTag proteins would react with their substrates, as they had been seen to do *in vitro*. Endogenous PEX5 in the cells would then recognise the PTS1 peptide attached to the EGFP-SNAP-Tag or EGFP-HaloTag protein and transport the protein to the peroxisome. The success of the transport would be monitored by the re-localisation of the green fluorescent from being dispersed in the cytoplasm to existing in punctate dots, where these punctate dots can be identified as peroxisomes. The reactive peptides were incubated at 7.5 and 12.5 µM with the cells by pipetting the peptides directly into the media surrounding a coverslip of transfected cells. These were incubated for 24 hours to observe any differences to the pattern of green fluorescence in these cells after this time period. After the incubation period cells were fixed and immunofluorescently stained for peroxisomes using a primary antibody against PEX14 and a secondary antibody to give a red fluorescent stain.

In these experiments, it was found that there was no alteration of the green fluorescence in any of the test conditions (Figure 2.18). The peroxisomes were successfully stained with the peroxisomal marker antibody anti-PEX14 but there was no evidence of co-localisation between the red and green fluorescence. This meant that the strategy was then revaluated to try and develop reactive peptides that could cause the desired peroxisomal re-localisation of the peptide labelled protein.
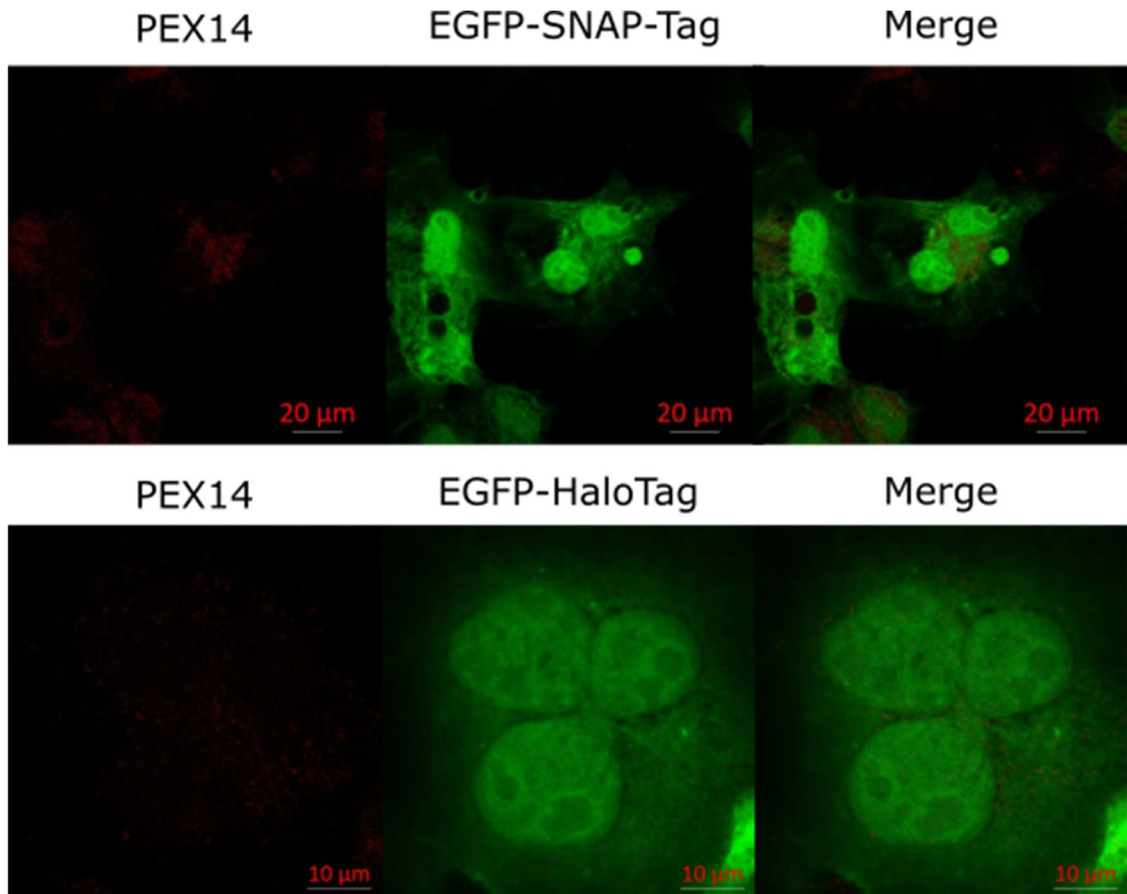
Figure 2.18 Confocal imaging of COS-7 cells treated with reactive peptides
COS-7 cells transfected with plasmids to express either EGFP-SNAP-Tag (top) or EGFP-HaloTag protein (bottom). Cells were incubated with 12.5 μM of a YQSKL peptide containing either a SNAP-Tag-reactive (top) or HaloTag-reactive (bottom) group. After 24 hours cells were fixed and immunofluorescently stained with a primary antibody for the peroxisomal membrane protein PEX14. Cells were imaged to determine localisation of red and green staining. The PEX14 staining showed punctate dots representative of peroxisomes (left panel). The green staining was seen to be cytoplasmic.

2.13 Development of new Reactive PTS1 Peptides

A number of different factors could have impaired the reactive peptides from causing re-localisation of the SNAP-Tag and HaloTag proteins to the peroxisome. All of these factors were considered when designing a new set of reactive peptide probes in order to repeat the experiments.

New reactive peptides were designed to help better understand the mechanisms at play in these cell-based experiments and potentially give the desired result of peroxisomal re-localisation of the EGFP-tag proteins. The second generation of reactive peptides added in two extra motifs. The first was the addition of a fluorophore. Pacific Blue was chosen as its excitation and emission wavelengths could be used in tandem with both EGFP and the AlexaFluor594 antibody used to label PEX14. The addition of the fluorophore to the peptides serves a number of roles. SNAP-Tag and HaloTag ligands conjugated to fluorophores have been shown to be taken up into cells and interact with their target proteins (Erdmann et al., 2019, Los et al., 2008, Cole, 2013, Bosch et al., 2014, Fransen, 2014).

Significantly, the ability to visualise the cellular location of the reactive peptides if they are indeed taken up into cells, will help to determine the mechanisms that affect the peptides and their ultimate cellular location. If the blue fluorescence stays cytoplasmic, then its co-localisation with the green fluorescence would be of interest. It is also possible that the blue fluorescence will be seen as the punctate dots that represent peroxisomes. In this instance co-staining for PEX14 would help to confirm this and support the theory that the reactive peptide reagent is imported by PEX5 before it has the opportunity to react with its substrate tag protein.

The other moiety added into the reactive peptides is a Polyethylene Glycol (PEG) group. This group extended the length of the peptide to create a larger distance between the PTS1 peptide and the SNAP-Tag or HaloTag reactive group. If the issue with import is a result of PEX5 not being able to bind to the peptide after it has reacted with it EGFP-tag protein due to steric hindrance, the addition of the linker should make a contribution to overcoming this problem.

The structure of the newly designed reactive peptides incorporating these features is shown in Figure 2.19 and Figure 2.20. These indicate the four different moieties present in the peptides.
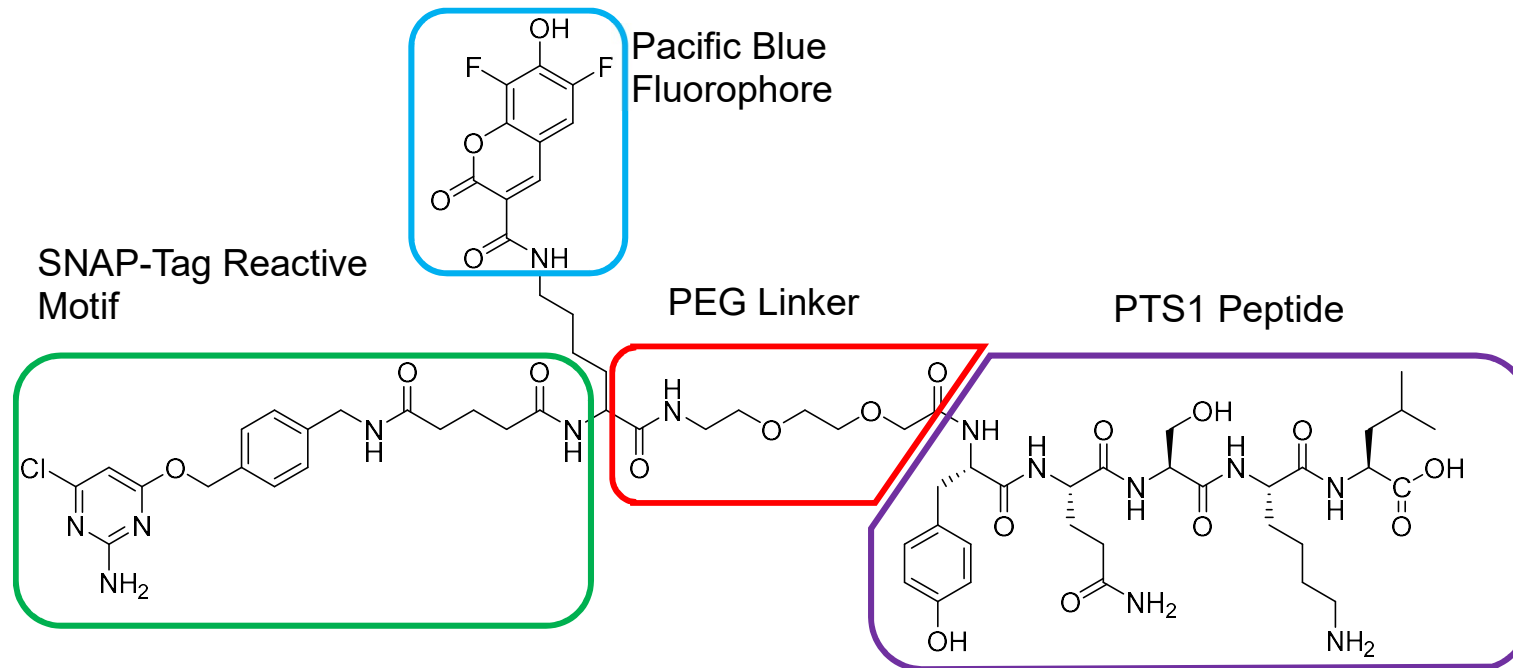
Figure 2.19 Second generation reactive peptide to attach the PTS1 peptide SNAP-Tag proteins

The second generation SNAP-Tag reactive peptide contained the same SNAP-Tag-reactive motif as the original reactive peptide and YQSKL as the PTS1 peptide. Additionally, a PEG linker is attached to the N-terminus of the Y residue, followed by a further lysine residue to which the Pacific Blue fluorophore has been attached.
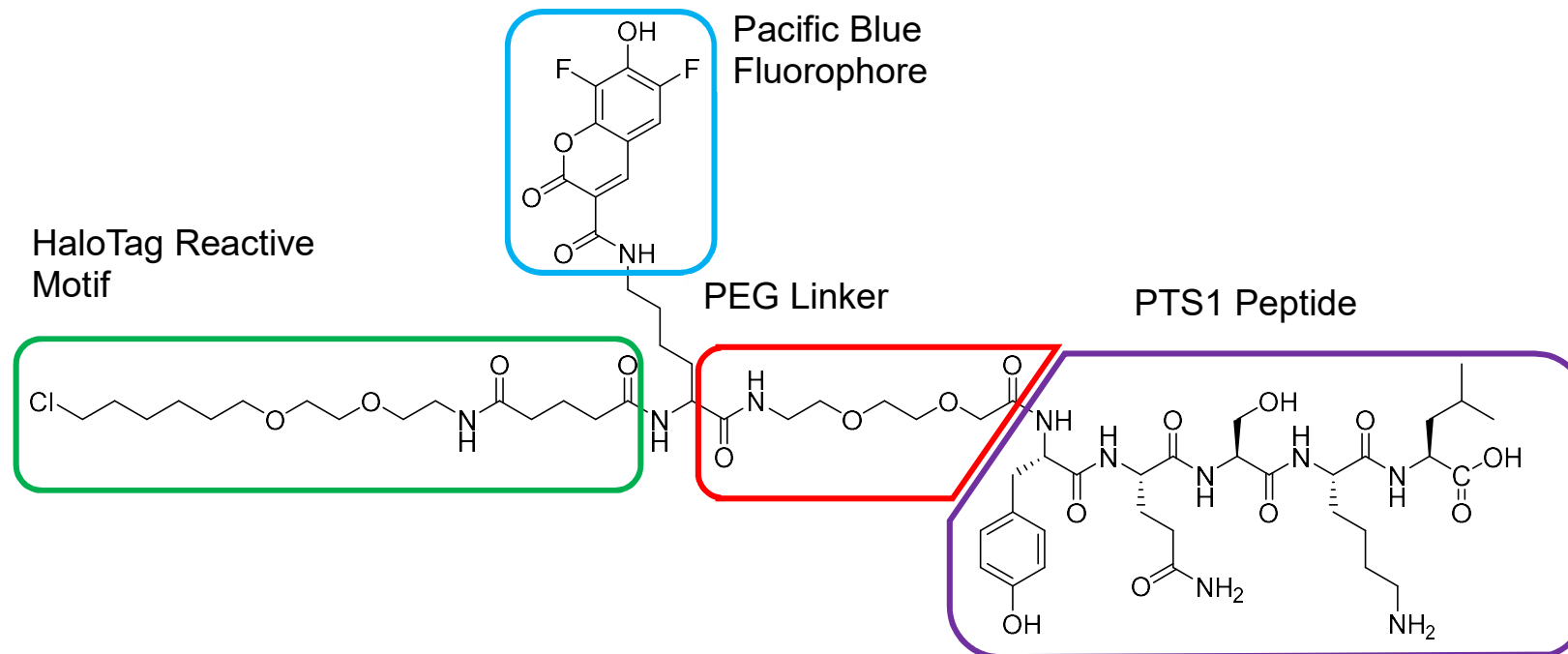
Figure 2.20 Second generation reactive peptide to attach the PTS1 peptide HaloTag proteins

The second generation HaloTag reactive peptide contained the same HaloTag-reactive motif as the original reactive peptide and YQSKL as the PTS1 peptide. Additionally, a PEG linker is attached to the N-terminus of the Tyrosine residue, followed by a further lysine residue to which the Pacific Blue fluorophore has been attached.

2.13.1 Synthesis of Redesigned SNAP-Tag and HaloTag-Reactive PTS1 Peptides

To synthesise the new peptides, the tag-reactive and PTS1 peptide motifs were produced as discussed previously (see: 2.2 and 2.3). The PEG linker was purchased as Fmoc-NH-(PEG)-COOH, which allowed it to be coupled to the N-terminus of the Tyrosine residue in the same manner as the amino acids in solid phase peptide synthesis. After Fmoc-deprotection of the PEG, an Fmoc-L-Lys(Dde)-OH amino acid was added. The lysine in the PTS1 peptide has its side chain protected by an acid-labile Boc group. This second lysine instead has its side chain protected by a N-[1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl] (Dde) group, which can be removed orthogonally with the addition of 2% (v/v) hydrazine in DMF (Scheme 2.4). This means that the Lysine-Dde can be specifically deprotected without unmasking the reactive side chains of any other amino acids.
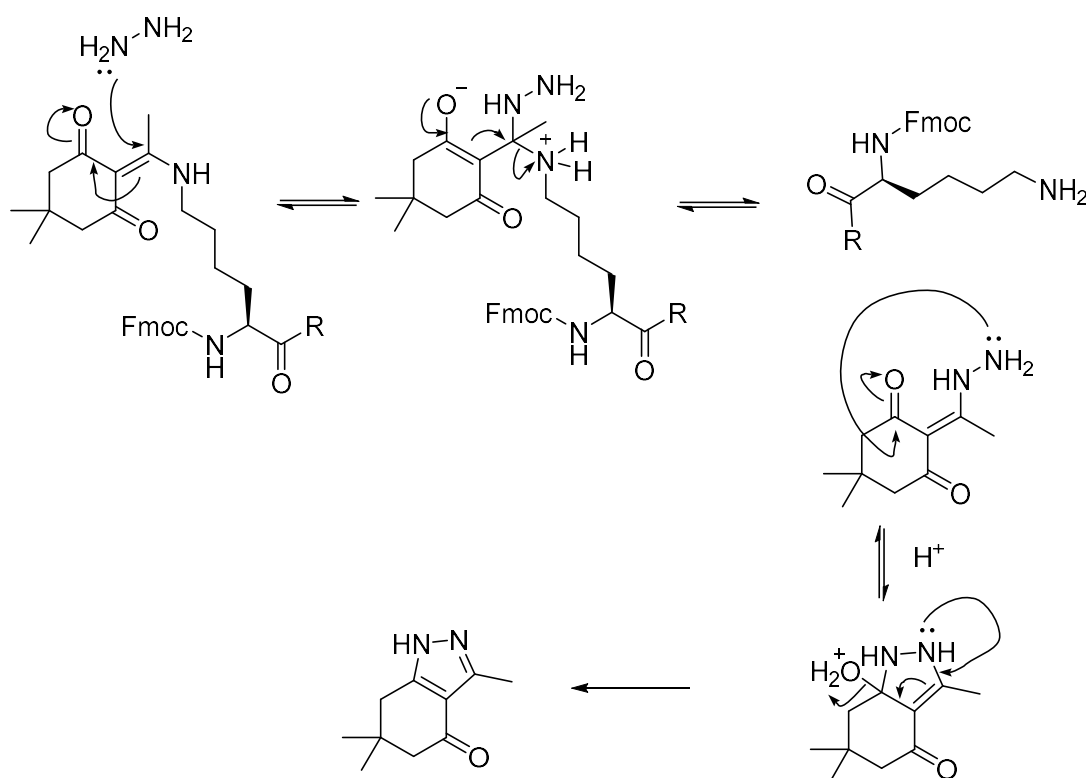
Scheme 2.4 Mechanism for the deprotection of Dde protected lysine residue

The hydrazine acts as a nucleophile to attack the more electron-deficient carbon of the alkene. The double bound is then reformed and the lysine removed as it is a better leaving group than the hydrazine. An intramolecular rearrangement then occurs, in which H_2O is lost and a 5-membered ring is formed.

The Pacific blue fluorophore was purchased as 6,8-Difluoro-7-hydroxy-2-oxo-2H-chromene-3-carboxylic acid. The carboxylic acid group could then react with the free amine on the Dde deprotected lysine to form an amine bond.

The Fmoc from the Lysine can then be removed with the standard basic piperidine and the SNAP-Tag or HaloTag substrate group coupled as before.

The peptides were purified as before with the C18 reverse phase column in a gradient of MeCN/H_2O (10→90%).

The peptides were tested in the labelling reactions of recombinant proteins as before and found to label the recombinant proteins similarly well to the first generation peptides.

2.13.2   Using new Reactive Peptides in COS-7 Cells

Once the second generation peptides had been synthesised and purified, they were tested on the COS-7 cells in a similar manner as previously, with some alterations.

The cells were transfected with EGFP-tag protein expressing plasmids as before. The peptides were added to the media surrounding the cells; however, lower concentrations were used on this occasion. Looking at studies using HaloTag-fluorophore substrates to study peroxisome dynamics performed by Marc Fransen (Fransen, 2014), a concentration range of 100nM -1µM was used.

When imaging the coverslips after an incubation time of 24 hours, no change in localisation of the GFP in the cells was observed upon incubated with the peptides. There was no evidence of any blue fluorescence in the cells (Figure 2.21).

This result suggested that the reactive peptides were not able to cross the plasma membrane of the COS-7 cells.

EGFP-SNAP-Tag + 100nM SNAP-PacificBlue-PEG-YQSKL 48 Hrs
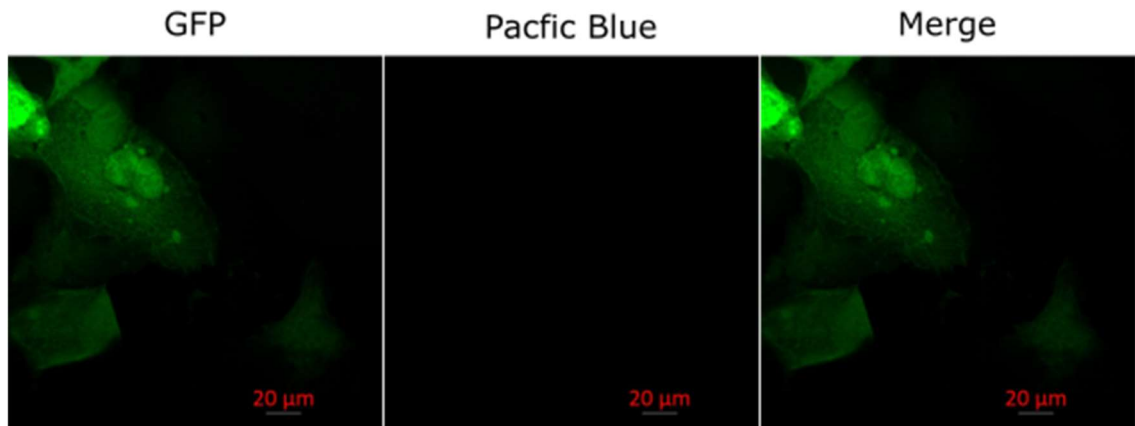
EGFP-HaloTag + 100nM Halo-PacificBlue-PEG-YQSKL 48 Hrs
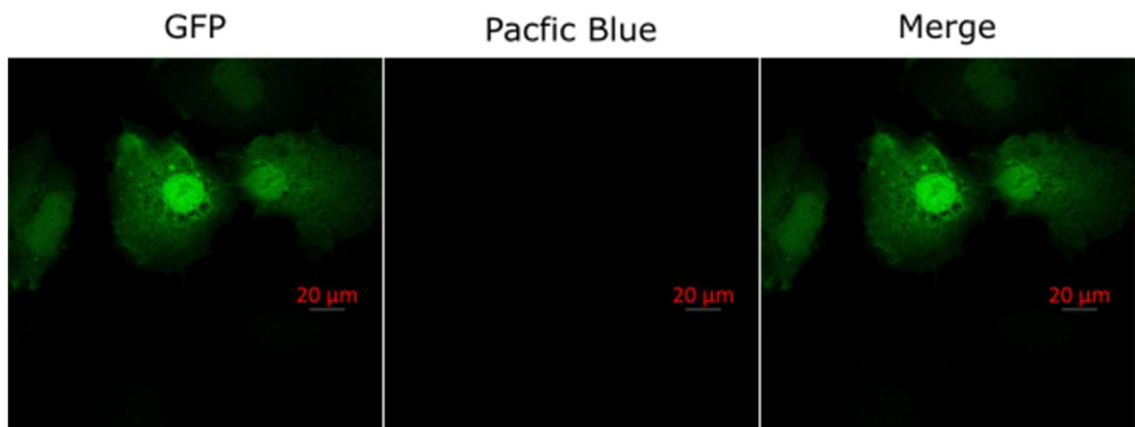
Figure 2.21 Treatment of Transfected COS-7 cells with second generation peptides COS-7 cells incubated with different concentrations on the tag-reactive Pacific Blue-PEG-YQSKL peptides.

2.14 Discussion

The aim of this study was to investigate if the attachment of a PTS1 peptide to a protein via a probe is sufficient for that labelled protein to be recognised for peroxisomal import by PEX5. To do this, the two main objectives were to (1) show binding of a protein labelled with a PTS1 via a tag can bind to the TPR domains of PEX5 *in vitro* and (2) show PEX5 mediated import of a protein labelled with a PTS1 via a tag into the peroxisome in cells. Of these objectives, we have clearly shown using the pulldown assays that the C-terminal domains of both *A. thaliana* and *H. sapiens* can bind to proteins labelled with a PTS1 via a tag. The second objective was investigated with two generations of probes and still requires further investigation and optimisation to create probes that can be used to fully determine if import across the peroxisomal membrane is possible for proteins labelled with a PTS1 from the side of the protein as opposed to the canonical C-terminally encoded PTS1 sequence.

Throughout this study there were a number of hurdles that had to be overcome to optimise experimental protocol and progress towards the final results.

2.14.1 Synthesis of Materials

The chemical synthesis of substrates for two different protein tags was necessary to produce reactive species with a bifunctional mode of action, something that has not previously been investigated. Although SNAP-Tags and HaloTags are well established as protein tags that will react with their substrates to introduce labels post-translationally, this has largely been restricted to using fluorescently labelled substrates that are commercially available. The attachment of a peptide to these tag substrates is a new area that required the synthesis reaction and importantly the purification strategy to be optimised. However, the reactive peptides were successfully synthesised and purified to give sufficient yield and purity for their use in both *in vitro* and cell-based assays. The purification strategy could be improved as it did not consistently produce peptide reagents of high purity (see Appendix C for HPLC analysis). Although reverse phase chromatography allowed the purification of some of the peptides to >95% purity, others were used at a purity of only 44%. In the future purification conditions could be further optimised to give more consistent results. However, mass spectrometry analysis of all of the reagents used suggested that the impurities were not of similar mass to the desired product. This suggests that the impurities present were small remnants of reagents used in the synthesis and would be unlikely to interfere with the reactivity of the reagents. Furthermore, this was observed *in vitro* that the

reagents reacted with their targets efficiently and effectively when used in two times molar excess, demonstrating that the purity was not affecting the final experimental outcome.

The synthesis of recombinant proteins for *in vitro* studies was largely straightforward with proteins able to be purified only by affinity chromatography and no requirement for further purification steps.

2.14.2   *In Vitro* Investigations

The pulldown assays used here demonstrate the sensitivity of PEX5 in its ability to bind to a PTS1 cargo. The human PEX5 protein was the desired target for use with the reactive peptides as moving forward such bifunctional molecules could have clinical relevance. However, up until this point all focus in the laboratory had been on the *A. thaliana* PEX5 protein. It had initially been expected that the C-terminal domains of the two proteins with N-terminal His_6-tags would behave in a similar manner in assays to bind PTS1 peptides. This expectation was not unfounded as the two proteins share similar domains, sequence homology (see Appendix A for sequence alignment) and function *in vivo.* The binding assays performed by the Berg laboratory (Gatto Jr et al., 2000) indicated that the truncation of *Hs*-PEX5 at the amino acid position used in this study would not affect its binding to the PTS1 peptide as they were able to show good binding of this portion of the protein to a YQSKL peptide. What could not be predicted however, was the impact the addition of a His_6 tag to aid in the purification of this protein, would have on its ability to bind to the PTS1 peptide. The fluorescence anisotropy assays showed that the *Hs*-His_6-PEX5C construct was unable to bind the PTS1 peptide YQSKL with the affinities seen for the *Hs*-PEX5C protein used by Gatto Jr et al. The cleavage of the His_6-tag from the protein however appeared to be the solution to alleviating the lack of binding of this protein construct to its PTS1 target as shown in both the FA and pulldown assays. This showed that the specific construct of *Hs*-PEX5C used is very important for its ability to bind cargo and regions upstream of the TPR domains have a significant impact on PEX5's ability to bind PTS1 peptides.

As it has been established that it is more than just the PTS1 that plays a role in PEX5 affinity (Fodor et al., 2012, Hagen et al., 2015, Maynard et al., 2004, Rosenthal et al., 2020), it seems likely that other regions of PEX5 make interactions with cargo proteins to increase the affinity. As the N-terminal domain of PEX5 is intrinsically disordered, this makes it more difficult to determine how cargo binding might be enhanced by regions outside the globular C-terminal domain containing the TPRs. However, in terms of the

PTS1-based probes used here, as there is not an upstream region on the cargo to make additional interactions with PEX5, it is unlikely that additional regions of PEX5 would make a significant difference to the binding interaction in this case

When investigating the interactions of the *Homo sapiens* PEX5 C-terminal domain with the YQSKL peptide, Jeremy Berg's groups found that their *Hs*-PEX5C protein bound to a lissamine labelled YQSKL peptide with a K_D of 70 ± 20 nM (Gatto Jr et al., 2000). As the design of our *Hs*-His$_6$-PEX5C protein was based upon this construct, a similar binding affinity would be expected. We tested this by synthesising the same Lissamine-YQSKL peptide and titrating both the *Hs*-His$_6$-PEX5C and the *At*-His$_6$-PEX5C proteins against this a measuring the change in anisotropy (Figure 2.13).

Although this improvement in binding the YQSKL was seen by fluorescnece anisotropy, the affinity of the binding did not appear to be as strong as that of *At*-His$_6$-PEX5C, or the binding found by Gatto Jr et al. of 70 ±20 nM. This could be due to the additional 3 amino acid residues remaining on the C-terminal side of the thrombin cleavage site.

By cleaving the His$_6$-tag used for the purification of the protein, its binding for the PTS1 peptide substrate was seen to markedly increase in both fluorescence anisotropy and pulldown assays. This suggests that in the *Hs*-His$_6$-PEX5C construct, the tag may block the TPR domains from binding the PTS1 peptide in some way. This appears not to be a problem for the *At*-His$_6$-PEX5C. Assessment of the specific constructs (Figure 2.7) shows that in *At*-His$_6$-PEX5C the PEX5 protein is truncated at an earlier point in the sequence so a larger portion of the region upstream of the TPR domains is included in this construct. The *At*-His$_6$-PEX5C construct is 52 amino acids longer than the *Hs*-His$_6$-PEX5C construct. These 52 amino acids might contain elements that enhance PTS1 binding, although it more likely that that this would enhance binding to protein cargos rather than isolated PTS1 peptides. This also means that the His$_6$-tag in *At*-His$_6$-PEX5C is further spaced from the TPR domains and therefore potentially less able to impact upon their binding activity. This may also suggest that the TPR domains bind to PTS1 peptides better in the presence of these upstream regions. In these experiments, a fine balance between recombinant protein stability and activity must be struck. In this case, the *Hs*-PEX5C construct used may bind to the PTS1 peptide better if more upstream regions were included. This could be further investigated by testing different sized portions of the *Hs* PEX5 C-terminal domain. But what must also be noted is that the fluorescence anisotropy assays performed by Gatto at al. showed that *Hs*-PEX5C of this length bound to YQSKL peptides with high affinity.

2.14.3 Cell-based assays

When testing the initial peptide probes in COS-7 cells the addition of the peptide probes did not cause any change in the localisation of the EGFP-SNAP-Tag and EGFP-HaloTag proteins. In these cell-based assays, truncated PEX5 proteins were no longer the target for the PTS1 peptide, but the endogenous PEX5 expressed in these mammalian cells. PEX5 in the cells was expected to not only bind to the peptide but also facilitate its import into the peroxisome. However, it is still unknown from the results shown here whether this endogenous PEX5 would have been able to import proteins labelled with the PTS1 peptide. There were a number of reasons why we were unable to see a change in the localisation of green fluorescence and the second generation of probes were designed to try and assess and alleviate the different options.

The first possibility considered was that the peptides were unable to cross the cell membrane to enter the cells. The experiments performed relied on the reactive peptide SNAP-Tag and HaloTag substrates to be taken up from the cell growth media. Previous work with SNAP-Tag and HaloTag reagents have shown that this strategy can result in the successful labelling of both SNAP-Tag and HaloTag proteins in live cells (Erdmann et al., 2019, Los et al., 2008, Cole, 2013, Bosch et al., 2014) and even in the labelling of peroxisomes in COS-7 cells (Fransen, 2014). In all cases however, commercial SNAP-Tag and HaloTag ligands were used that generally contained a fluorescent group attached to the SNAP-Tag or HaloTag-reactive moiety. The difference in this instance is that the "label" attached to the SNAP-Tag and HaloTag substrates is instead the PTS1 peptide. There is one example of fluorescent PTS1 peptides being taken up into living mammalian cells (Pap et al., 2001). This example indicates that the strategy used here to incubate cells with the peptides and expect spontaneous uptake is not unfounded. As this is a novel reagent, it was not possible to predict its cell permeability.

The addition of PEG groups has been shown to help promote cellular uptake of therapeutics in nanoparticle and liposomes (Sadzuka et al., 2003, Osman et al., 2018). The addition of PEG-like polymers to peptides has also been shown to help increase their cellular uptake in comparison to the peptide without the polymer linker (Sakuma et al., 2012). The addition of PEG groups to a peptide have also been shown to increase solubility (Li et al., 2015). It is for these reasons that a PEG linker was included in the second generation of peptides synthesised here. It was found that this second generation of reactive peptides were more soluble in aqueous buffers than those without the PEG linker, but its inclusion did not appear to enhance cellular uptake in this case.

In designing the second generation of probes it also had to be considered that it was not uptake of the probes across the plasma membrane that was the main barrier to success of peroxisomal import. If it were assumed that the reactive peptide substrates were able to cross the plasma membrane, there were a number of points to consider that could have prevented the import of the SNAP-Tag or HaloTag proteins into the peroxisome. The first of these was that if we were to assume that the peptides entered cells and reacted with the SNAP-Tag and HaloTag proteins, as they were seen to do *in vitro*, we do not know if the attached PTS1 peptide protruded far enough from the tag protein for the endogenous PEX5 to be able to bind to it. As we do not know the conformation of the EGFP-SNAP-Tag and EGFP-HaloTag proteins in the cellular environment, it is possible that a complex forming between PEX5 and the EGFP-tag proteins is not sterically possible. It is also possible that the endogenous PEX5 recognised the PTS1 peptide immediately upon its entry into the cell and the reaction between the tag proteins and their substrate does not occur before the reactive peptide is imported to the peroxisome- this would also result in the green fluorescence remaining in the cytoplasm. Therefore, addition of the PEG linker might also help by providing a spacer to allow the interaction of both PEX5 and the GFP-tag protein simultaneously.

A further possibility related to this is that the concentration of peptide reagent that reached the cytoplasm of the cell may be very high. This would mean there would be an excess of reactive peptide over tag protein for it to react with. It is likely that PEX5 would preferentially bind to the free reactive peptide and import this in preference to the peptide that has reacted with the tag protein as the smaller cargo would be more kinetically favourable. A study indicated that the cytoplasmic concentration of PEX5 per cell in rat liver cells was approximately 0.75µM (Freitas et al., 2011), a concentration significantly lower than that of the peptides added to the media. If all of the peptides were taken up into the cells, PEX5 would likely become overloaded with cargo and therefore there would unlikely be any available PEX5 to bind to PTS1 peptide reacted with tag proteins.

An alternative possibility to the peptides not being able to cross the plasma membrane, is that they are able to enter cells but are rapidly exported again.

In the Pap et al. study, it was found that although their fluorescent PTS1 peptides were spontaneously incorporated, in some cell lines efflux pumps needed to be inhibited by the addition of verapamil to the cells in order to observe peroxisomal relocalisation of the probes (Pap et al., 2001).

Experiments carried out using fluorescent HaloTag ligands by Fransen et al. indicate that those ligands were able to cross the plasma membrane to react with peroxisomal matrix and membrane proteins expressed with a HaloTag (Huybrechts et al., 2009). In their study they found that the HaloTag ligands could not cross the peroxisomal membrane unless reacted with the PTS1-HaloTag protein in the cytoplasm and also that excess ligand was removed when washing cells, suggesting it is rapidly exported across the plasma membrane. The authors suggested that it was likely that the multidrug resistance protein ABCB1 was responsible for this export, as the HaloTag ligands used in their case were rhodamine based and there has been evidence to show that rhodamine related compounds can be exported by ABCB1 (Saengkhae et al., 2003).

Efflux pumps could very well play a role in the removal of the probes used in this project from cells. This could mean that the SNAP-Tag and HaloTag ligands used in this work could have been too rapidly exported from the cells to be able to react with their corresponding tag protein and hence no blue fluorescence from the Pacific Blue fluorophore on the ligands was observed in the imaging (Figure 2.19 and Figure 2.20). If this is the case, this would discount that theory that the ligands were able to react but unable to import, as in this scenario, blue fluorescence would have been seen in the cytoplasm. Initially, a logical next step would be to study the effect of the addition of verapamil to the cells to see if this impacts upon the results seen.

However, in the Huybrechts study, some HaloTag ligand was able to react with HaloTag-PTS1 protein, in this study this did not appear to be the case. This makes it more likely that the probes used in this study did not make it across the plasma membrane at a concentration sufficient that their subcellular localisation could be observed using the Pacific Blue fluorescent label included in the second generation of peptides.

The next steps to investigate this avenue further could be the use of electroporation or microinjection of the peptides into the cells to ensure that they are in the cytoplasm of the cells to react with their target proteins. Electroporation has been used to allow the uptake of HaloTag-conjugated quantum dots (QDs) into cells when incubation of the cells with the QDs proved unsuccessful to facilitate their uptake (Hatakeyama et al., 2017). Microinjection has also been used to allow HaloTag substrates into mammalian cells (Koike and Jahn, 2017). Cell-impermeable SNAP-Tag ligands also use microinjection as a method of entry into cell, these were however used a very high

concentrations (30 M), so this may not have been practical for our purposes (Keppler et al., 2004).

Another option could be to modify the reactive peptides again to increase cell permeability. Some peptides have been seen to be able to cross the cell plasma membrane. The first example of these cell-penetrating-peptides (CPPs) was the trans-activating transcriptional activator (Tat) peptide derived from the protein of the same name from Human Immunodeficiency Virus (HIV) (Vivès et al., 1997). Various other CPPs have been developed since this and the common feature across all of these CPPs is the presence of multiple basic residues, sometimes interspersed with hydrophobic residues (Patel et al., 2019, Jones and Sayers, 2012).

The PTS1 peptide used in the reactive molecules synthesised here does not fit the consensus of the peptides that have been used to facilitate cellular uptake, and therefore it is unlikely that this peptide would enhance the process. As the results showed it was found that the second generation of probes were unable to cross the plasma membrane, it would be advisable with future probes to include residues that enhance cell penetration such as multiple basic and hydrophobic residues. This would however need to be balanced with aqueous solubility of the probes, as they would still need to be water soluble to be able to be added to the cell growth media.

These properties may be difficult to predict and so it may take many rounds of further optimisation to develop such peptides. What also must be considered, is that including more cell-penetrating peptide properties in the probe could have an adverse effect on the binding of the probe to both PEX5 and the protein it is intended to label, further complicating the process of optimising the probes.

Ultimately, the PTS1 peptide in the molecules could potentially be replaced with a non-peptide moiety that is still capable of being bound by the TPR domains of PEX5. Such a molecule could be developed through the high-throughput screen of compound libraries and further FA assays to compete new compounds with the PTS1 peptide for the PEX5 binding site. It is possible that replacing the peptide in the molecule with a molecule that shows more optimal pharmacokinetic properties could improve its uptake into cells. Such strategies have been used to improve drug candidates. No drugs have previously been developed to specifically target the PTS1 binding site of PEX5, though some work has been done to look at inhibiting peroxisomal protein import through a small molecule inhibiting interactions between PEX14 and PEX5 (Dawidowski et al., 2017). Such work could act as a basis for looking into the development of a small molecule mimic of the PTS1 peptide.

The idea of generating heterobifunctional molecules to relocate proteins to the peroxisome and knockdown their cellular activity is inspired by the work of Craig Crews developing Proteolysis Targeting Chimeras (PROTACs). PROTACS also include two different targeting groups, similar to the molecules generated here. One half of a PROTAC binds to a target protein to be removed, the other recruits an E3 ligase, which together causes the target protein to be targeted for proteasomal degradation (Sakamoto et al., 2001). The benefits of PROTACs over conventional small molecule inhibitors also hold true the peroxisomal re-localisation strategy. First, all of the functions of the protein can be inhibited simultaneously, including protein scaffolding functions that cannot be easily targeted by a specific protein inhibitor, be this through degradation by the proteosome or peroxisomal imprisonment. These peroxisome-targeting molecules also have a use as biological probes to help better understand proteins functions. By sending a target protein to the peroxisome or targeting it for proteasomal degradation by addition of the bifunctional molecule, the target protein can be removed from its endogenous cellular environment at a precise moment. When compared to knocking down proteins by altering them at a genetic level, the use of protein degrading or eliminating probes allows for better temporal control of the process. As peroxisomes are capable of importing complexes of proteins through piggybacking (Islinger et al., 2009, McNew and Goodman, 1994, Effelsberg et al., 2015), it is possible that by using heterobifunctional probes such as those developed here, new protein-protein interactions could be discovered. By importing a target protein, and then analysing the content of the peroxisome, interacting partners of the target protein could be found. This could be used as an assay to uncover a multitude of otherwise unknown interactions. Importantly, it could be the case that these protein-protein interactions have not previously been uncovered because they only occur transiently under certain conditions or are very weak. The use of a peroxisomal re-localisation assay where probes are added under specific cellular conditions could be a novel approach to discover these interaction partners. This would work well as the next step on from the work mentioned previously by Dueber and DeLoache (Chen et al., 2015) and Dunoyer (Incarbone et al., 2018) where the peroxisome was used to discover weak protein-protein and protein-RNA interactions respectively.

The initial studies into developing PROTACs also utilised the HaloTag technology to develop probes (Buckley et al., 2015). They demonstrated the ability of their probes to degrade GFP-HaloTag proteins in HEK293 cells and found that linker length and orientations of the two functional groups played an important role in the functionality of the "HaloPROTACs". There did not appear

to be issues with cell penetration for the HaloPROTACs and this could be due to the E3 ligase ligand portion of the probe being largely hydrophobic. This hydrophobicity could have enhanced the ability of the probes to cross the plasma membrane. This suggests that the inclusion of more hydrophobic regions in the probes developed in this project could help enhance their cellular uptake.

2.14.4   Conclusions

At the present time, it remains inconclusive as to whether the functional molecules developed here can trigger the peroxisomal import of a protein by attachment of the PTS1 peptide at a site that is not the C-terminus. The initial findings however indicate that the proteins labelled with the PTS1 peptide via a SNAP-Tag or HaloTag can both be bound by the TPR domain of PEX5 from both *A. thaliana* and *H. sapiens*. At this point the optimisation of the molecules involved and modifications to the methods used to facilitate their uptake across the cell membrane could shed more light on the feasibility of this mechanism to introduce proteins into the peroxisome that are not found there natively.

There are a number of next steps to be taken to build on what has been found in this project. First, the different methods to ensure the probes cross the cell membrane should be investigated. This should be the priority as once it can be proven that the probes are inside a cell the viability of this probes to induce peroxisomal relocalisation can be truly evaluated. I believe that the electroporation would be the best method to try initially as this could be set up quickly and wouldn't involve the synthesis of new probes. If these experiments show that the probes cause the relocalisation of the GFP-tag proteins as hypothesised, then the redesign of the probes to improve their cellular uptake a retention could be investigated.

In relation to this project, work looking into developing a small molecule mimetic of the PTS1 peptide is now in progress. If a successful small molecule is identified, this could be incorporated into probes in place of the PTS1 peptide. At this point, the redesign of the probes to make them more cell permeable could be revisited.

Chapter 3

Covalent binding of PTS1 probes to PEX5 to study cargo unloading

3.1 Introduction

As discussed previously, the exact mechanism of import of peroxisomal matrix proteins across the peroxisomal membrane is yet to be fully understood. The process is rapid and therefore the import complex formed at the membrane with PEX5 and other proteins is transient and difficult to study.

In the previous chapter, probes were developed which allowed the labelling of a protein with a PTS1 peptide via a SNAP-Tag or a HaloTag. Proteins labelled in this way were seen to interact with the peroxisomal matrix protein import receptor PEX5. After showing that proteins labelled with a PTS1 peptide in this way could interact with PEX5, attention turned to the use of similar such probes that could be used to better understand the process of import of PEX5 cargo protein, which are destined to the peroxisomal matrix, across the peroxisomal membrane.

The strategy used here is to create a PTS1 peptide probe that will bind to PEX5 through the PTS1 peptide and subsequently, irreversibly bind to PEX5 via a reactive group installed into the probe. The hypothesis is that PEX5 will then attempt to import this cargo into the peroxisome, but the import complex formed will be stalled as the cargo is unable to be unloaded (Figure 3.1). It is hoped that this would aid better understanding of the cargo unloading step of peroxisomal protein import.

To do this, the structure of the PTS1 binding site of *Hs*-PEX5C (Gatto Jr et al., 2000) was examined and residues on the surface of the binding funnel selected. These residues were altered to cysteines to create a series of mutants. This was initially carried out on the *At*-PEX5C protein as this had previously performed better in *in vitro* assays. These mutants would have the potential capability of reacting with chloroacetamide groups in peptide probes. Once a cysteine mutant with the desired reactivity was found, the probes were developed further to contain a HaloTag-reactive group. By adding these probes to the HaloTag protein and subsequently reacting the labelled protein with the PEX5 cysteine mutant, PEX5 could be irreversibly loaded with a protein cargo. This could allow further studies into the process of cargo unloading at the peroxisomal membrane.
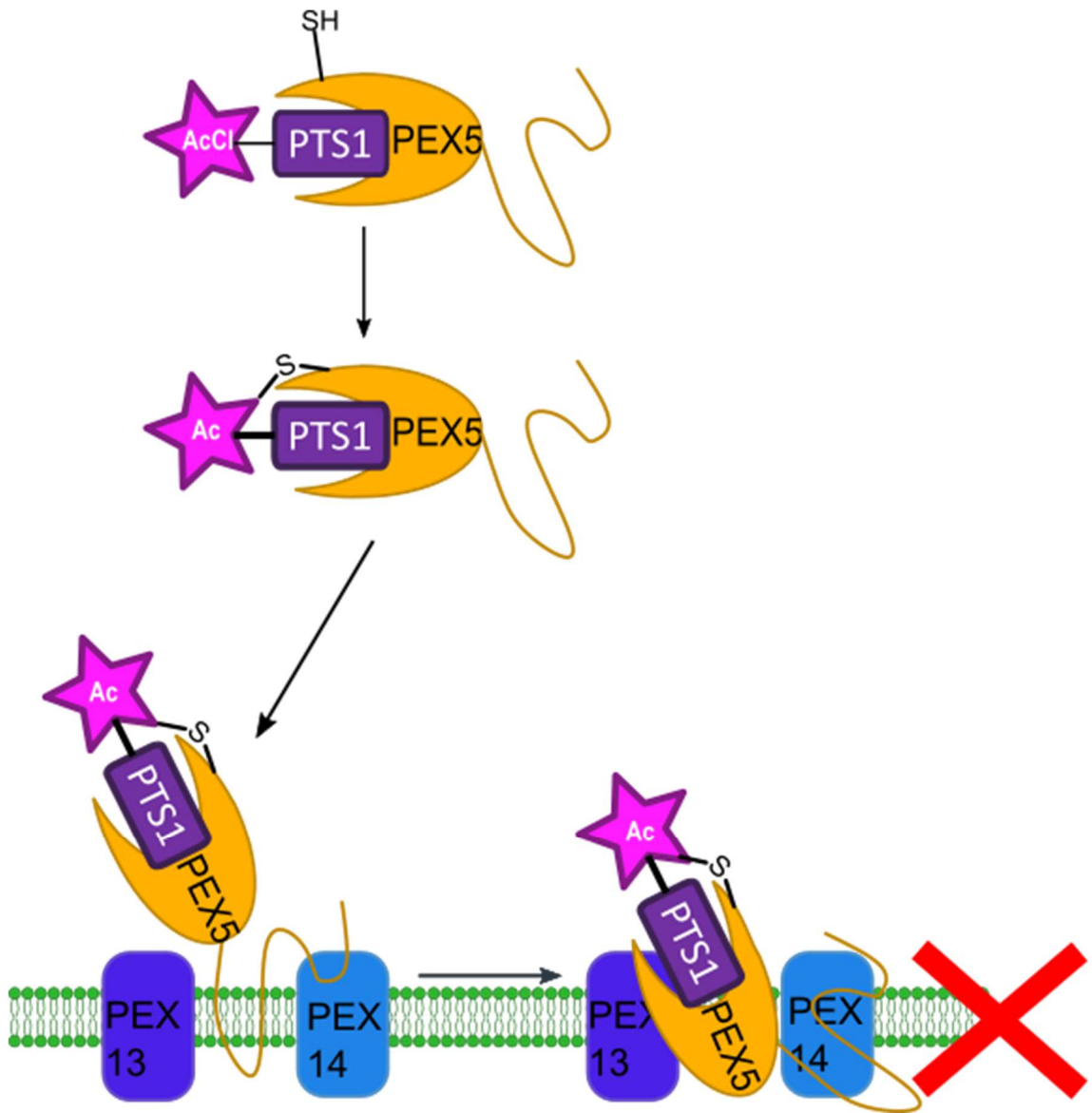
Figure 3.1 A Cartoon Summary of Project Aim An overview of the aim of this project is shown. A reactive PTS1 probe that binds irreversibly to PEX5 could cause PEX5 to stall in its import competent conformation. This could then be studied to better understand the import mechanism.

3.2 Design of Reactive PTS1 peptide probes

To make the cysteine-reactive PTS1 peptide probes, the first consideration was the specific PTS1 peptide to introduce. In the previous work (see Chapter 2), YQSKL was used at the PEX5-binding PTS1 peptide. Here, the peptide used was YQSRL. The lysine residue was replaced with an arginine as the guanidine in the side chain of arginine is less nucleophilic than the primary amine group in the side chain of lysine. Many peroxisomal matrix proteins are targeted using -SRL at the C-terminus (Reumann, 2004) and although binding to SRL by mammalian PEX5 confers a slight reduction in affinity compared to -SKL (Gatto et al., 2003), -SRL is still a "strong" PTS1 with high peroxisomal import efficiency (Gould et al., 1989, Elgersma et al., 1996, Swinkels et al., 1992). As the probe was to contain a cysteine-reactive group, reducing the nucleophilicity of the peptide by replacing lysine with arginine guarded against possible reaction of the probe with itself. As arginine is still a basic residue, it interacts with the TPR domains of PEX5 similarly to the lysine residue and so it is unlikely that this change in the PTS1 peptide would affect the interaction of the PTS1 peptide with PEX5.

To introduce a cysteine-reactive group to the probe, the same Lysine-Dde residue as was used to create the pacific blue labelled peptides was introduced during the synthesis (see 5.3.3.3). This allowed the incorporation of a cysteine reactive handle via the amine of the lysine side chain (Figure 3.2). As cysteine is a nucleophilic amino acid, the thiol side chain will react in a nucleophilic substitution reaction with electrophilic carbons to irreversibly form a covalent bond. Haloacetic acid groups have long been established as effective cysteine-reactive probes (Cole et al., 1958). These have since been developed and are most widely used as haloacetamide groups (Hoch et al., 2018). Such probes have been used for activity-based protein profiling (ABPP). ABPP uses chemical probes to specifically modify proteins. The probes consist generally of a reactive group to specifically react with the target and a tag group to give the ability to profile where the reaction has a occurred (Barglow and Cravatt, 2004). The probes used here contain a group that promotes binding to the protein of interest (the PTS1 peptide in this case) and a reactive group to covalently label proteins bound by the probe (the chloroacetamide here). This is a similar approach to ABPP, although in this case the dependency of a reaction is based on the affinity of the peptide for the PTS1 binding site for the reaction as opposed to a specific protein activity.

The labelling was initially studied using mass spectrometry to determine a change in mass upon protein labelling. Later, a fluorophore may be used to help

establish probe specificity for PEX5. One of the key aspects to determine with these probes is whether or not any reaction seen with PEX5 is specific to the interaction of PEX5 with the PTS1 peptide and not a general reaction of a cysteine residue with the probe without PTS1 binding. For this, probes containing the non-PTS1 peptide sequence YQLRS were used.

The PTS1 sequence in the peptide probe must be at the C-terminus in order to interact with the PEX5 TPR domains. Therefore, the reactive chloroacetamide group was introduced via the Lysine-Dde residue at the N-terminus of the peptide (Figure 3.2). What was not known was how much distance would be needed between the PTS1 peptide and the reactive group in order for the cysteine in PEX5 to be able to access the chloroacetamide group. For this reason, three different probes were generated with different linker lengths. The first contained no linker between the PTS1 and the chloroacetamide-modified lysine as it was possible that the probe would not have to protrude far from the PEX5 binding pocket to be available to interact with the introduced cysteine residue and the butyl chain of the lysine side chain would give enough distance for the nucleophilic substitution reaction to occur. The second probe contained a short PEG linker of 9 atoms in length and the third a longer linker of 20 atoms in length. By utilising these three different lengths of probe it was hoped that at least one would be of optimum length to carry out the desired labelling reaction.
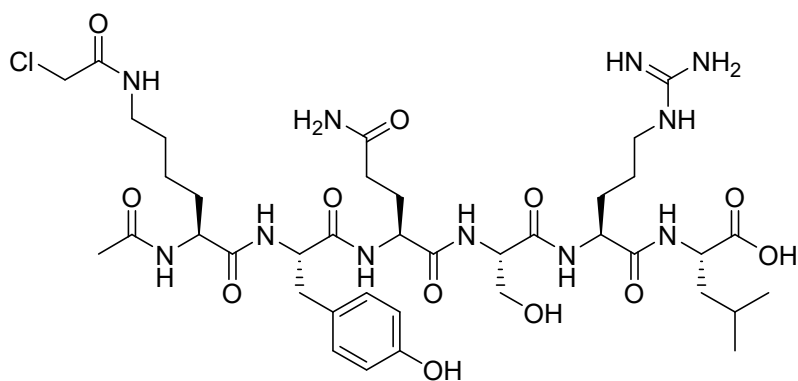
AcCl-no linker-YQSRL
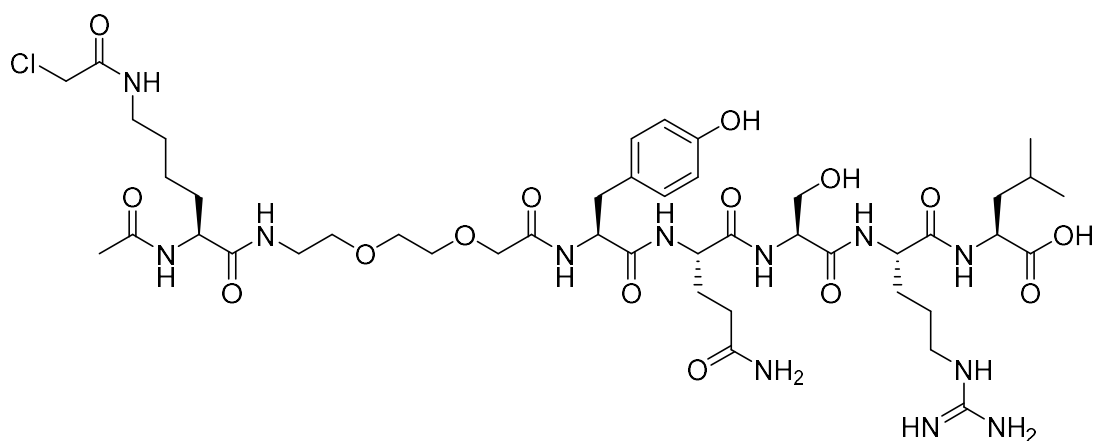
AcCl-PEG-YQSRL
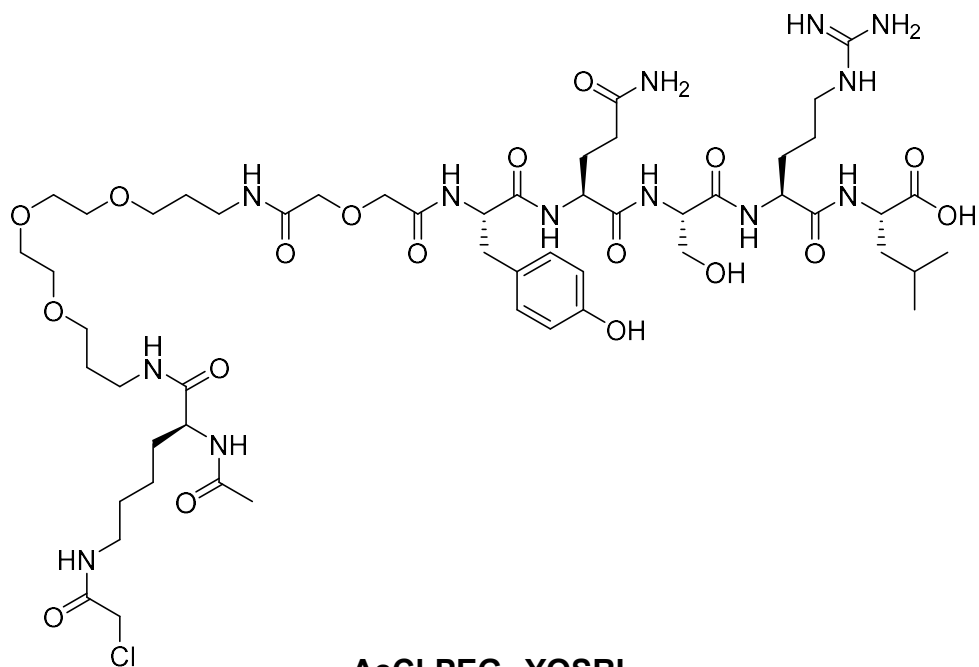
AcCl-PEG2-YQSRL

Figure 3.2 Structure of PTS1 Peptides with Cysteine-reactive group The three peptides initially synthesised to test for reactivity with PEX5 cysteine mutants, and their names are they are referred to in the text are shown.

3.3 Synthesis of Reactive PTS1 peptide probes

The peptides were synthesised similarly to the those detailed in Chapter 2 and in Materials and Methods (5.3.3). To add the PEG linker groups, Fmoc-protected amino acids were used and therefore could be coupled onto the growing peptide chain in the same way as any amino acid. The final amino acid added was the Fmoc-Lysine -(Dde)-OH. As the chloroacetamide group is reactive, it was found that better yields of peptide were obtained if the reactive groups were added last. This meant that the Fmoc group was removed before the Dde. When making the peptides for work in the previous chapter, SNAP-Tag or HaloTag substrates were coupled to the N-terminus after Fmoc deprotection. In this instance the free amine group at the N-terminus needed to be "capped" to render it unreactive. To do this the amino group was reacted with acetic anhydride to generate an amide that could not act as a nucleophile in any side reactions.

Once the peptide N-terminus had been capped, the Dde was removed from the amine in the lysine side chain as previously described (5.3.3.3). Chloroacetyl chloride was then added to react with this free amine, resulting in the addition of a cysteine-reactive chloroacetamide group into the probe.

In order to determine if any reaction of these peptides is specific, control peptides were also synthesised. These were based on the middle length AcCl-PEG-YQSRL peptide. A control of AcCl-PEG-YQLRS was synthesised to determine if the reaction with PEX5 was templated to the binding of a PTS1 sequence, as YQLRS is not a PTS1. A second control that did not contain the AcCl group, unreactive-PEG-YQSRL, was also synthesised to determine if the probe was covalently reacting through the chloroacetamide group and not by some other means (Figure 3.3).
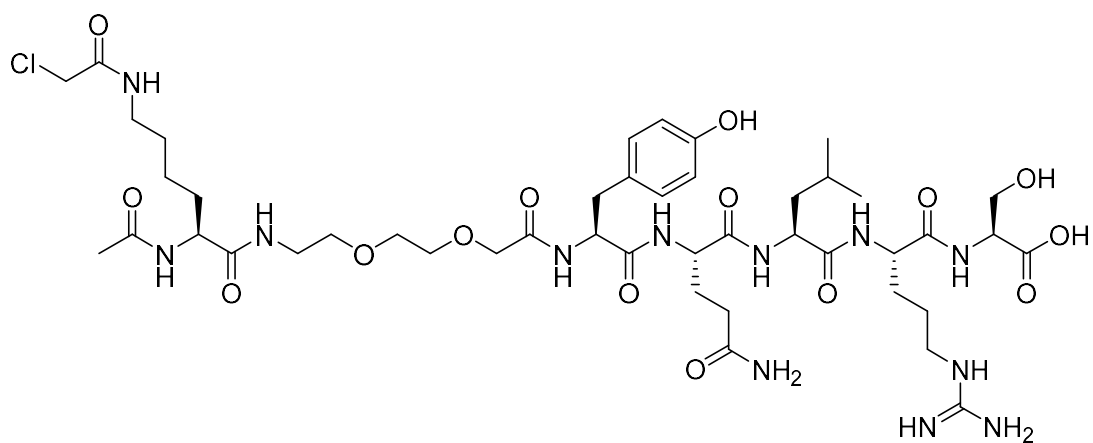
AcCl-PEG-YQLRS
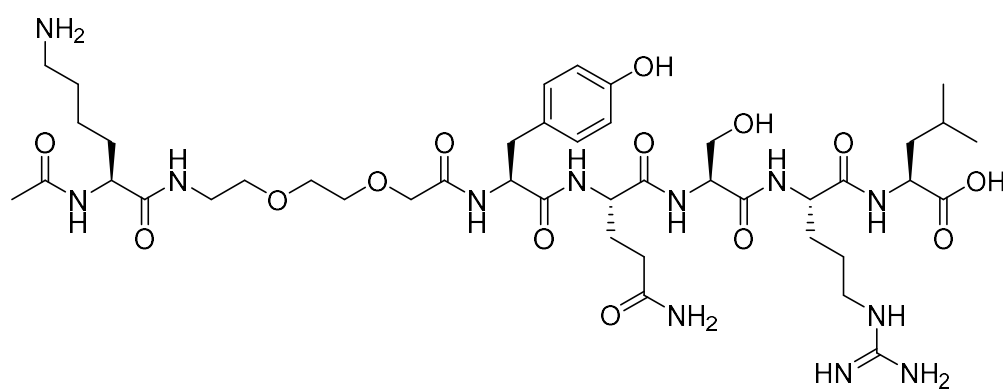
Unreactive-PEG-YQSRL

Figure 3.3 Structures of Control Peptides The two control peptides generated to test if both a PTS1 and a cysteine-reactive group were both required for the reaction with PEX5

3.4 Selection of Mutant Residues in PEX5

In order for the synthesised probes to react and form a covalent bond with PEX5, a cysteine residue needed to be introduced at a position where it would be able to react with the chloroacetamide when the PTS1 of the probe was bound by the PEX5 TPR domains. To do this, residues needed to be selected to be altered to cysteine.

The first stage was to identify which residues might be well positioned to react. Such residues would need to be on the surface of the binding pocket on PEX5. As the crystal structure of the C-terminal domain of PEX5 from *Homo sapiens* bound to the peptide YQSKL has been solved (PDB:1FCH) this was used as the starting point to identify potential residues. Residues were first selected by eye through examining the 1FCH structure and selecting residues that appeared to be in the surface of the binding pocket (Figure 3.4A). When selecting residues, those with more similarities to a cysteine in size and/or chemical properties were preferred. This meant that smaller residues such as serine or alanine were preferred as well as other non-polar residues such as leucine or glycine. Twenty residues that surrounded the top of the PTS1 binding funnel were initially selected (Table 3.1).

The human and *Arabidopsis thaliana* PEX5 sequences were then aligned using the Basic Local Alignment Search Tool (BLAST). These sequence alignments can be found in Appendix B. From this alignment, the equivalent residues from the human proteins were found in the *A. thaliana* protein (Table 3.1). The nature of the residues were compared with one another and ranked using a traffic light system (Table 3.1). Green represented an exact match i.e. the residue is a leucine in both proteins, amber for mismatched residues that were similar in size and/or chemical properties (e.g. alanine and glycine) and red where the two residues were not similar. It was found that one of the selected residues did not have a counterpart in the *A. thaliana* protein and therefore could not be selected to be mutated.

The initial twenty residues were also assessed on their predicted solvent accessibility. To do this, the 1FCH crystal structure was analysed using STRIDE. STRIDE uses a knowledge-based algorithm to define protein secondary structure as well as other properties of the protein query (Heinig and Frishman, 2004). One of the parameters obtained in the STRIDE output is the solvent accessible area of each residue. The output from this analysis can be found in Appendix B. The STRIDE solvent accessible area percentage values for each of the selected twenty residues was recorded and ranked, again using a traffic light system (Table 3.1). If the solvent accessible area was greater than

75%, the residue was designated green, if it was between 50 and 75% the residue was amber and residues with a score of less than 50% were designated red.

Table 3.1 Initial residues considered for mutation and their properties

Hs-PEX5 Residue	Location	*At*-PEX5 Residue	% Solvent Exposed	Selected for Mutation?
S280	Pre-TPRs	A438	84.3	✓
A281	Pre-TPRs	G439	80.3	✓
Y283	Pre-TPRs	Q441	22.3	
L349	TPR2	Q508	85.7	✓
E379	TPR3	E538	71.0	
S380	TPR3	L539	86.2	✓
L381	TPR3	E540	36.0	
S424	TPR4	-	91.4	
L426	TPR4	L566	87.9	✓
L463	TPR5	L602	32.2	
S464	TPR5	S603	46.4	✓
G465	TPR5	R604	37.5	
Y467	TPR5	F606	62.8	✓
E502	TPR6	A641	64.2	✓
L532	TPR7	Q671	92.9	✓
S568	C-Term	Q696	72.9	
L572	C-Term	L700	59.1	✓
S575	C-Term	S703	53.9	
M576	C-Term	C704	31.9	
S580	C-Term	Q708	77.4	
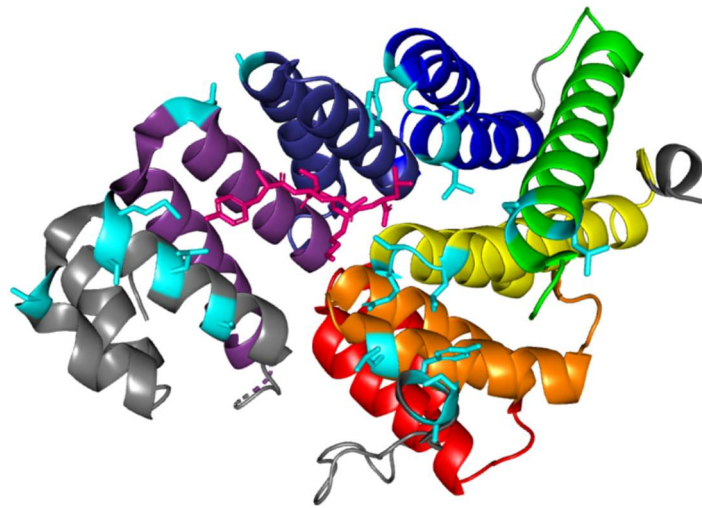
20 residues intitially selected
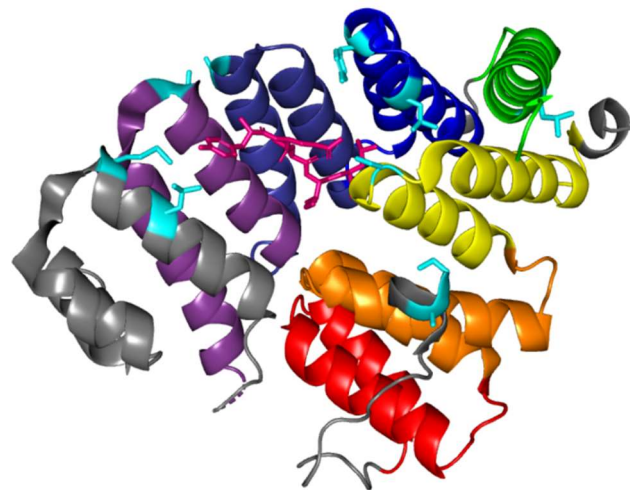
10 residues chosen for mutation

Figure 3.4 Selection of amino acid residues to be mutated to cysteine to generate probe-reactive *At*-PEX5C mutants The 20 residues initially selected to have the potential for mutation to cysteine to react with a YQSKL (shown in magenta) based probe were selected on the structure of *Hs*-PEX5C (cyan). After analysis of these residues, 10 were chosen for the initial round of site-directed mutagenesis on the *At*-PEX5C protein.

Both of these rankings were considered when selecting ten residues from the initial twenty for which to design mutagenesis primers. The ten residues selected based on these two parameters and also the desire to mutate residues across the whole surface of the binding funnel. This meant that some of the selected residues did not necessarily have good predicted solvent accessibility or a similar counterpart in the human protein. It was hoped that by keeping the location of the residue diverse, it would be more likely to introduce a probe-reactive cysteine residue. The positions of the selected residues are shown in cyan in Figure 3.4.

3.4.1 Mutagenesis

Cysteine residues were introduced in the *At*-PEX5C protein that had been used previously. This protein contains the C-terminal domain of *A. thaliana* PEX5 from residue 340 onwards.

Site directed mutagenesis was used to alter the selected residues in the wild type protein to cysteines. The QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) was used for this and primers were designed using the primer design guidelines stated in the manual. The primers used can be found in Appendix B.

After the mutagenesis reactions, the plasmids were amplified by transformation into XL-10 cells and plasmid DNA isolated from the colonies grown were verified for the introduction of the mutation by Sanger sequencing. Seven of the ten mutants were successfully generated. Of these seven mutants, six were found to express well in BL21-Gold (DE3) cells and were able to be taken forward to test their reactivity with the probes. The outcome of the mutagenesis and expression of each At-PEX5C mutant is summarised in Table 3.2.

Table 3.2 Summary of outcome of mutation and expression for At-PEX5C Mutants

Mutant	Generated?	Expressed?
A438C	✓	✗
G439C	✓	✓
Q508C	✗	✗
L539C	✓	✓
L566C	✓	✓
S603C	✓	✓
F606C	✓	✓
A641C	✗	✗
Q671	✗	✗
L700C	✓	✓

Once both the reactive probes and PEX5 mutants had been generated, testing to find a compatible combination could commence.

3.5 Testing of Probes with *At*-PEX5C

The six cysteine mutants that were found to be successfully expressed in BL21-Gold (DE3) cells (G439C, L539C, L566C, F606C, A641C and L700C) were all expressed and purified in the same manner as the wild type *At*-PEX5C protein (Materials and Methods). The wild type protein was also assessed for its ability to react with the probes. Residue 704 in *At*-PEX5C is a cysteine that corresponded to M576 in the human protein and was one of the twenty residues initially selected as it is on the correct face of the protein structure to possibly interact with a YQSKL-based probe. If the probes were able to react with the wild type protein, it would be expected that this cysteine residue would be responsible, and that we would also see reactivity with all the mutants.

To test the reactivity of the probes with *At*-PEX5C and its cysteine mutants, the proteins were incubated with each probe separately in a 1:2 molar ratio of protein:probe. The reaction mixtures were incubated for 1 hour on ice before analysing by mass spectrometry. The reaction mixtures were then analysed again after overnight incubation at 4°C.

The wild type protein did not react with any of the probes, therefore validating the need to introduce new cysteine residues into the protein to allow the probes to react.

One mutant, F606C, was found to show some reaction with all three of the probes after one hour. After the overnight incubation, the peaks of increased mass became more dominant in the mass spectrum (Figure 3.5).

All three of the probes labelled the F606C mutant protein to some extent. The probes containing the shorter PEG linker (Figure 3.5C) appeared to label best as its peak relative to the unlabelled protein was larger than seen with the other probes.

The results from the reactions with the control peptides show that the *At*-PEX5C F606C mutant protein does not react with either of the control probes demonstrating that the probe is reliant on both the presence of the reactive chloroacetamide group (Figure 3.5D) and a PTS1 sequence (Figure 3.5E) to orient the probe in the correct position in the PEX5 TPR domains to access the cysteine residue.
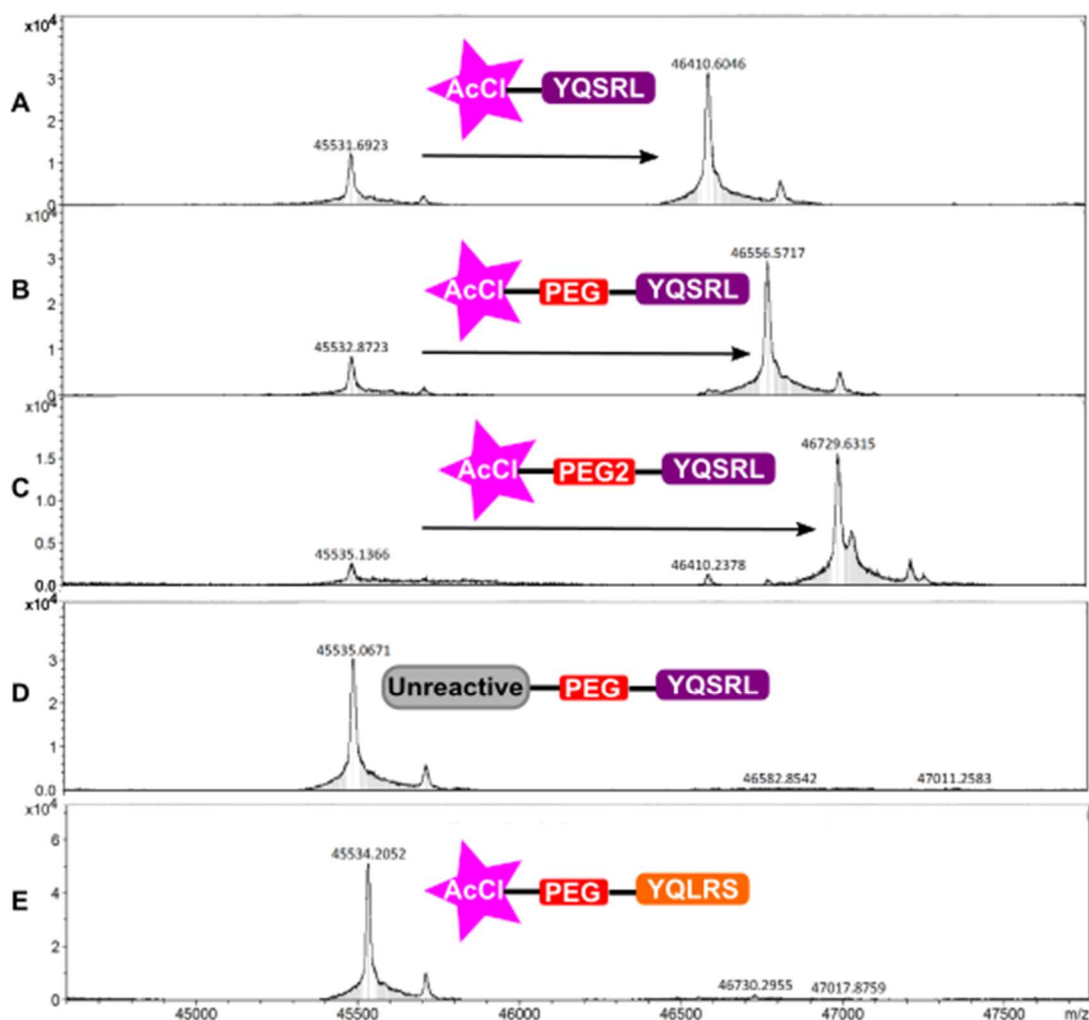
Figure 3.5 Labelling of F606C *At*-PEX5C mutant with chloroacetamide probes Deconvoluted mass spectra of the F606C mutant of *At*.PEX5C are shown after the protein was incubated with the three different reactive probes overnight at 4°C. The spectra show that the protein increases in mass by the mass that would be expected with a successful reaction with the probe. The control spectra (D and E) show the protein mass remaining the same.

3.6 Testing of Probes with *Hs*-PEX5C

After the *At*-PEX5C F606C mutant was found to react with the PTS1 reactive probe, the equivalent mutant was generated from the *Hs*-PEX5C protein. This mutant was Y467C (Table 3.1). As shown in chapter 2, the *Hs*-PEX5C protein binds to PTS1 peptides with higher affinity after the removal of the N-terminal His$_6$-tag. For this reason, the His$_6$-tag was cleaved from the Y467C mutant protein before attempting to react it with any of the PTS1 probes.

When incubating the His$_6$-tag-cleaved Y467C mutant with the probes in the same way as for the *At*-PEX5C F606C mutant, the labelling reaction was not seen to occur by mass spectrometry.

For this reason, the ability of the F606C and Y467C mutants to bind a PTS1 peptide was assessed using fluorescence anisotropy (Figure 3.6). The F606C mutant of At-PEX5C appeared to bind to Lissamine-YQSKL with identical affinity to that seen in the wild type protein. The Y467C mutant, cleaved of its His$_6$-tag, bound to the Lissamine-YQSKL peptide with an affinity of approximately 600 nM, higher affinity than the Wild Type *Hs*-His$_6$-PEX5 protein (which was >5000 nM), but not as well as the wild type protein without the His$_6$-Tag, (232 nM). This could explain the lack of a reaction between the Y467C mutant and the probes; its affinity for the peptide is not sufficient for the reaction to occur. From this point, only the F606C *At*-PEX5C mutant was take forward into further experiments.
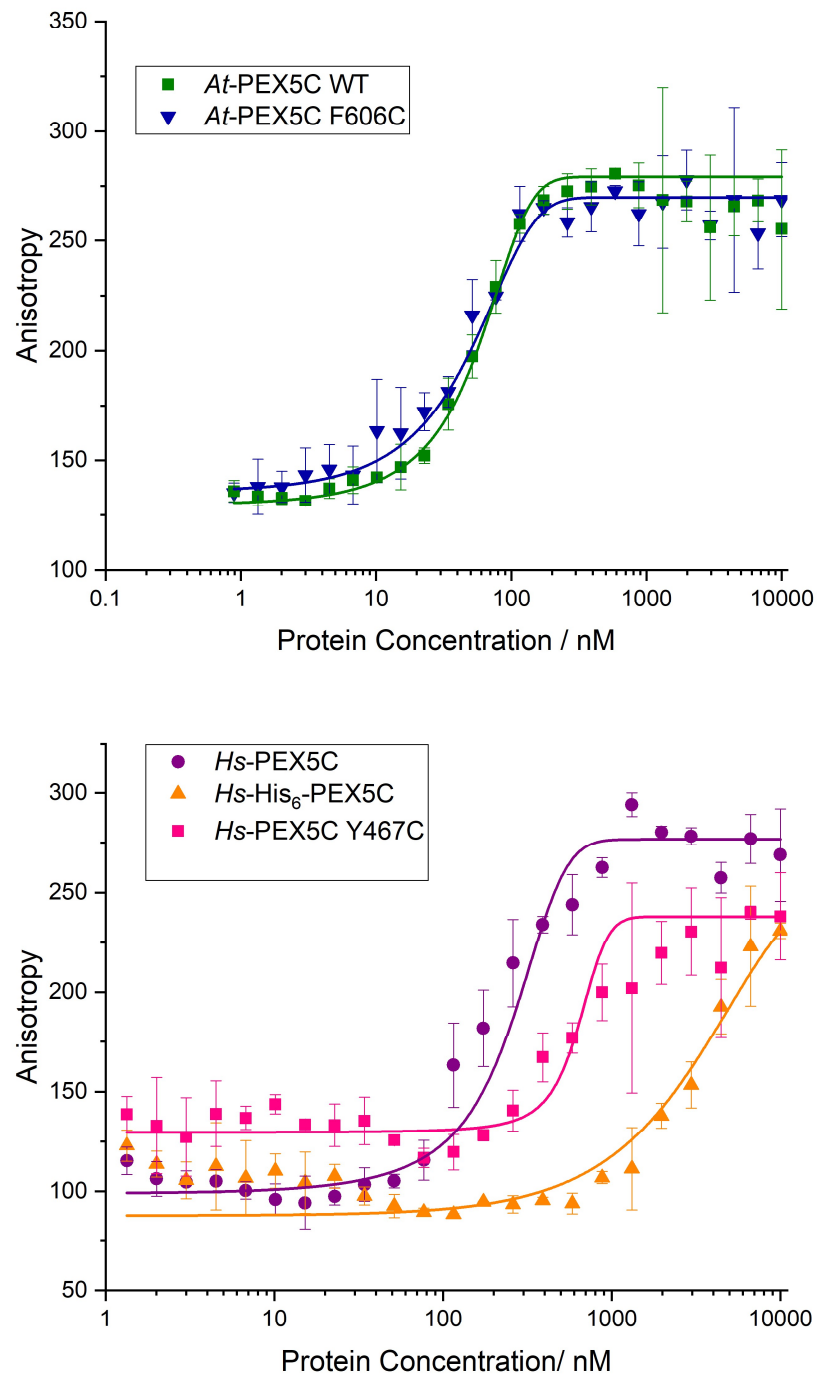
Figure 3.6 Fluorescence Anisotropy Analysis of Wild-Type and Mutant PEX5 proteins A. Binding of the F606C mutant and the Wild Type *At*-PEX5C protein to the Lissamine-YQSKL peptide. B. Binding of Y467C mutant, Wild Type *Hs*-PEX5C and *Hs*-His$_6$-PEX5C proteins to the Lissamine-YQSKL peptide

3.7 Covalent binding of HaloTag Protein Cargo to PEX5

Once it was established that the *At*-PEX5C F606C mutant was able to covalently bind to the probe after binding to the PTS1 peptide, the next stage was to extend the probe to include a HaloTag-reactive motif. In this instance, the AcCl-PEG-YQSRL probe was chosen as the probe to be extended as it showed overall to have the most reliable labelling of *At*-PEX5C F606C. To incorporate the HaloTag-reactive motif, this was synthesised as detailed previously (see Chapter 2). A second 9 atom PEG linker was included after the Lys(AcCl) residue to create space between the two reactive motifs in the probe. Two other control probes were also synthesised, one with the control non-PTS1 peptide YQLRS and a second lacking the reactive chloride to form the covalent bind via cysteine in *At*-PEX5C F6060C. The structure of the probes is shown in Figure 3.7.
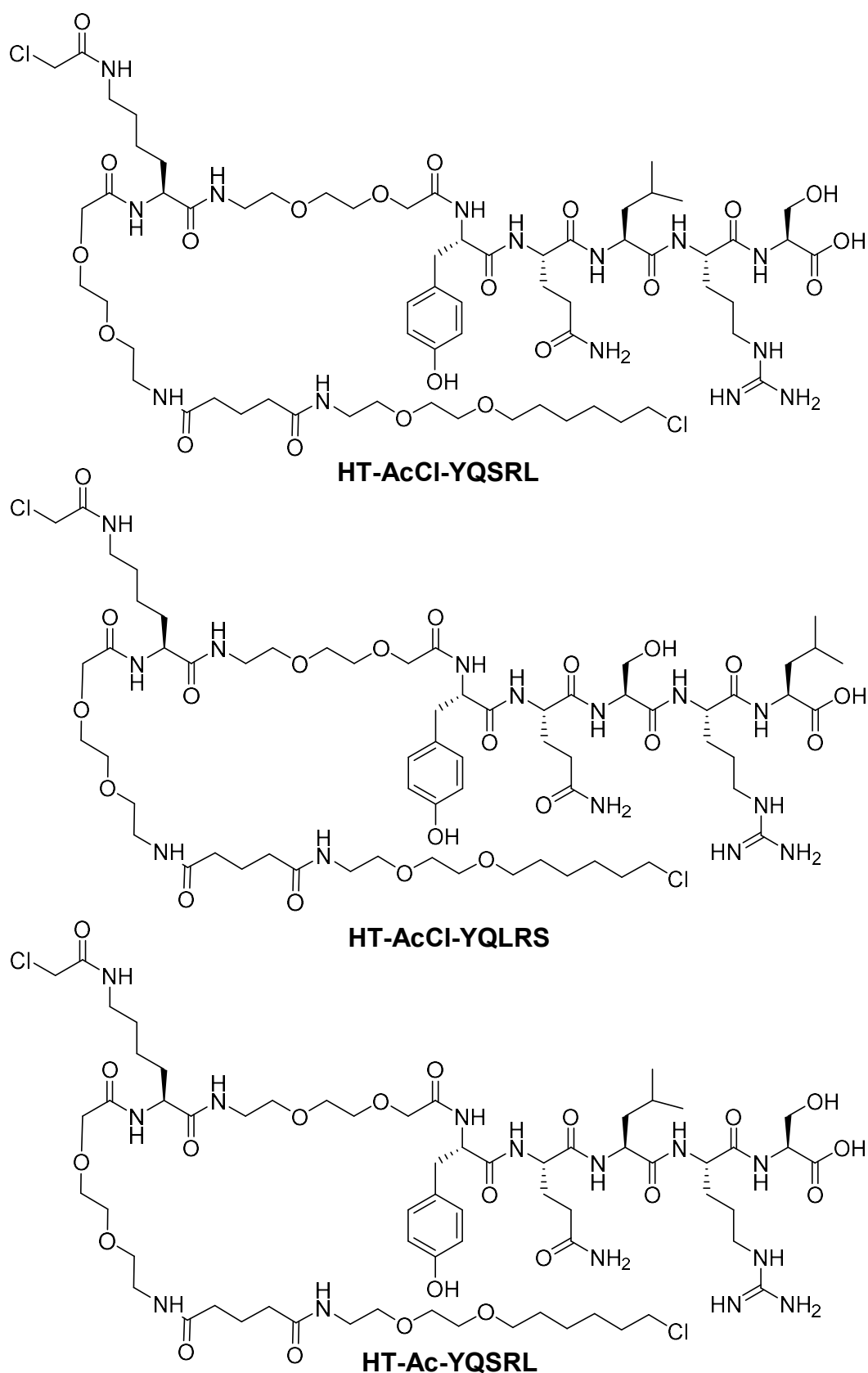
Figure 3.7 Bifunctional Probes with HaloTag and Cysteine Reactivity The three probes synthesised to test if HaloTag protein cargo can be covalently bound to PEX5 as a result of PTS1 binding

These probes were first tested for their reactivity with the HaloTag protein used in the previous experiments and were found to react and covalently label the HaloTag protein (Figure 3.8). Excess probe was removed from the reaction mixture by filtering through a 30 kDa molecular weight cut off spin concentrator. The labelled HaloTag protein was then mixed 1:1 with the At-PEX5C F606C protein and allowed to label overnight. The formation of a covalent complex between the PTS1 labelled HaloTag protein and At-PEX5C F606C would be expected to form a protein complex of 84.2 kDa. Initially, the labelling reaction was analysed by mass spectrometry, however, no labelling could be seen in this way. As it is rare for the instrument to be required to detect such large masses, it is possible that instead of the reaction not occurring and the complex not being present, the protein complex could be formed, just undetected in this way. For this reason, samples from the labelling reactions were run on 8% SDS-PAGE gels to detect if any larger complexes of proteins were present. This analysis showed that there were two distinct larger bands after overnight incubation of At-PEX5C F606C with the HaloTag protein labelled with the HT-AcCl-YQSRL probe (Figure 3.9). One of these bands appears to be at approximately 84 kDa, the expected mass for the protein complex that should be formed. There is also a second band at approximately 110 kDa, which cannot be accounted for but appears to be specific to the presence of the HT-AcCl-YQSRL labelled HaloTag protein.
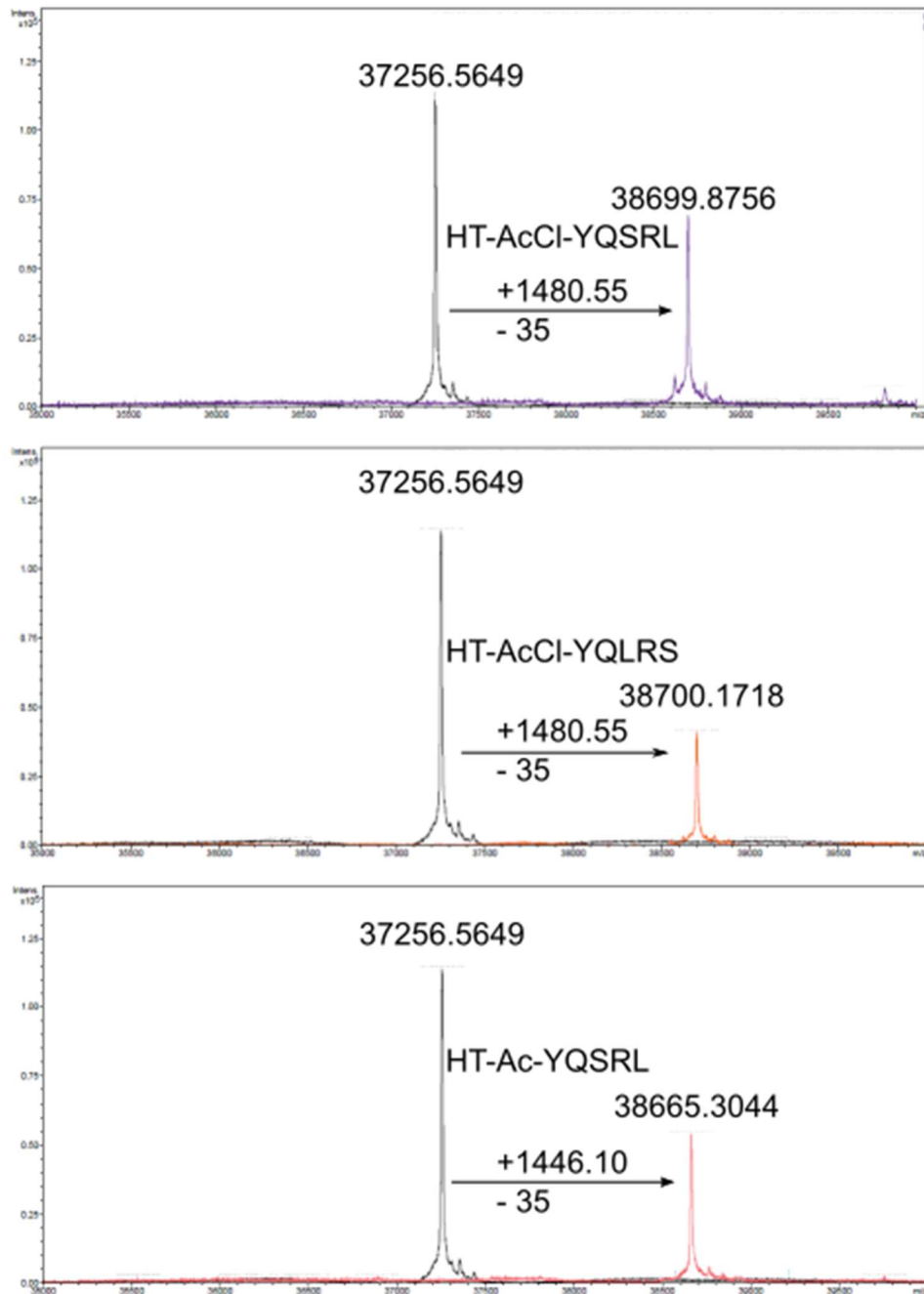
Figure 3.8 Mass Spectrometry Data for HaloTag Protein labelling with HaloTag-reactive probes Initial mass of HaloTag protein (black) is seen to increase by the expected value when incubated with each probe, all containing the HaloTag-reactive substrate motif (HT)
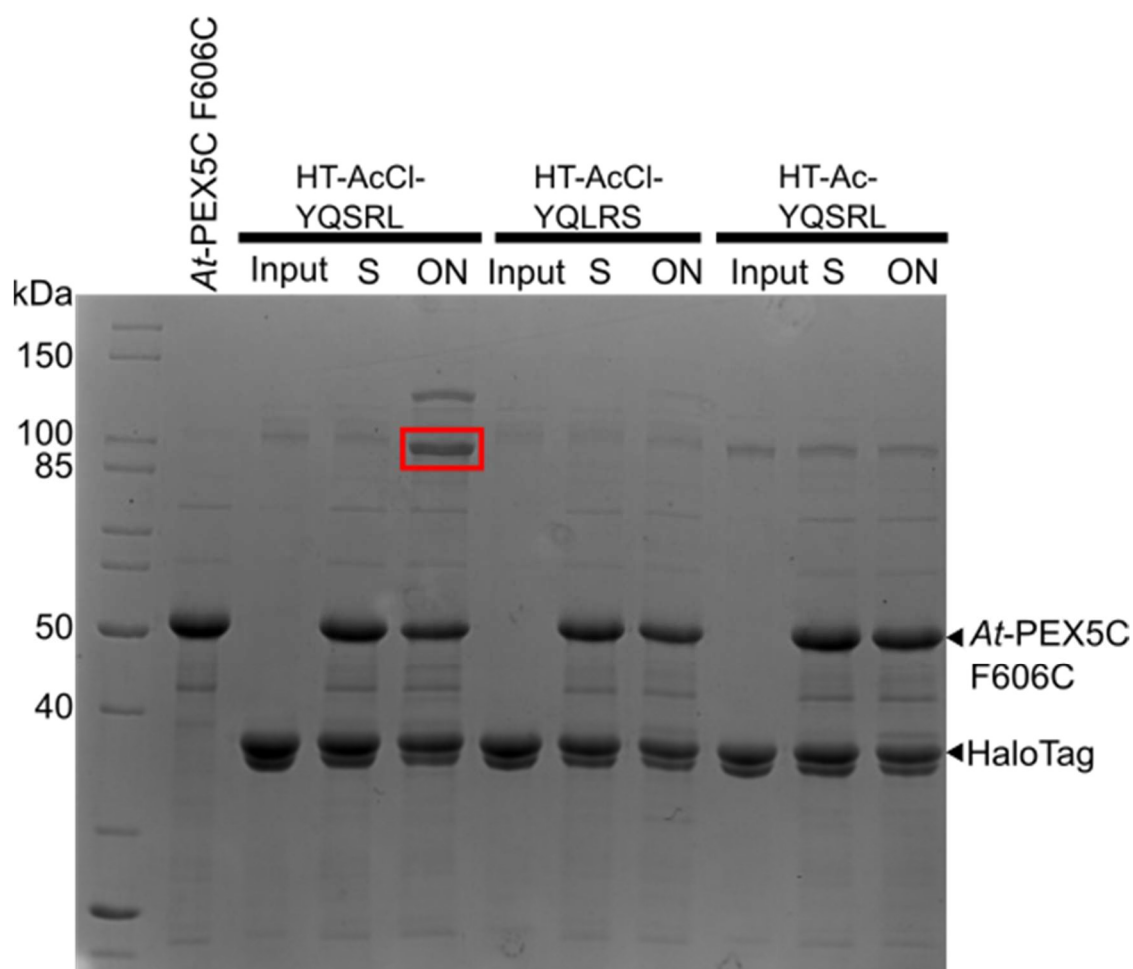
Figure 3.9 Coomassie Stained SDS-PAGE gel analysing protein labelling reactions The HaloTag protein labelled with one of the three probes shown was loaded into the lanes labelled input. Once mixed with *At*-PEX5C F606C (shown in lane 2) a portion of the starting reaction mixture was immediately boiled and prepared for loading in the lanes labelled S. After overnight (ON) incubation of the HaloTag proteins with *At*-PEX5C F606C, a sample from the reaction mixture was taken to be run in the ON lanes. The band highlighted in the red box shows that a complex of approximately 84 kDa has been formed specifically in the presence of the HaloTag protein labelled with the HT-AcCl-YQSRL probe.

To try and identify the protein types present in these bands, attempts were made to use antibodies against both PEX5, HaloTag and the Twin-Strep-tag® present in the HaloTag protein. The experiment shown in Figure 3.9 was repeated and samples were run on 8% SDS-PAGE gels and blotted onto Polyvinylidene fluoride (PVDF) membranes. Previous work in the group required the development of a polyclonal antibody raised against the Δ340 *At*-PEX5C protein. After testing to confirm that this antibody was able to detect the F606C mutant of this protein, it was used in a Western Blot to detect the presence of that protein in any bands containing the *At*-PEX5C F606C protein.

A HaloTag antibody and an antibody to detect the presence of the Twin-Strep-tag® were also used.

The blots (although somewhat unclear at lower molecular weights due to unresolved membrane transfer problems) again showed two bands of higher molecular weight that only formed in the presence of the HaloTag protein labelled that the PEX5 protein was present. However, in this instance the bands appeared to run at masses of approximately 70 and 80 kDa respectively- lower than was seen on the Coomassie gel. As these blots were of poor quality it is difficult to interpret this result with certainty and the work would need to be repeated to see if these differing masses require more investigation (Figure 3.10). The HaloTag antibody indicated that the HaloTag was also present in these bands. However, the Anti-Strep antibody only very weakly detected the Strep-Tag epitope in these larger protein bands, despite detecting it strongly in the starting HaloTag-Strep protein (Figure 3.11).

The HaloTag antibody also detected that the impurities seen in the labelled HaloTag protein samples on the Coomassie stained gel (Figure 3.9) contained the HaloTag epitope, which suggested the formation of protein complexes of the HaloTag protein with itself.
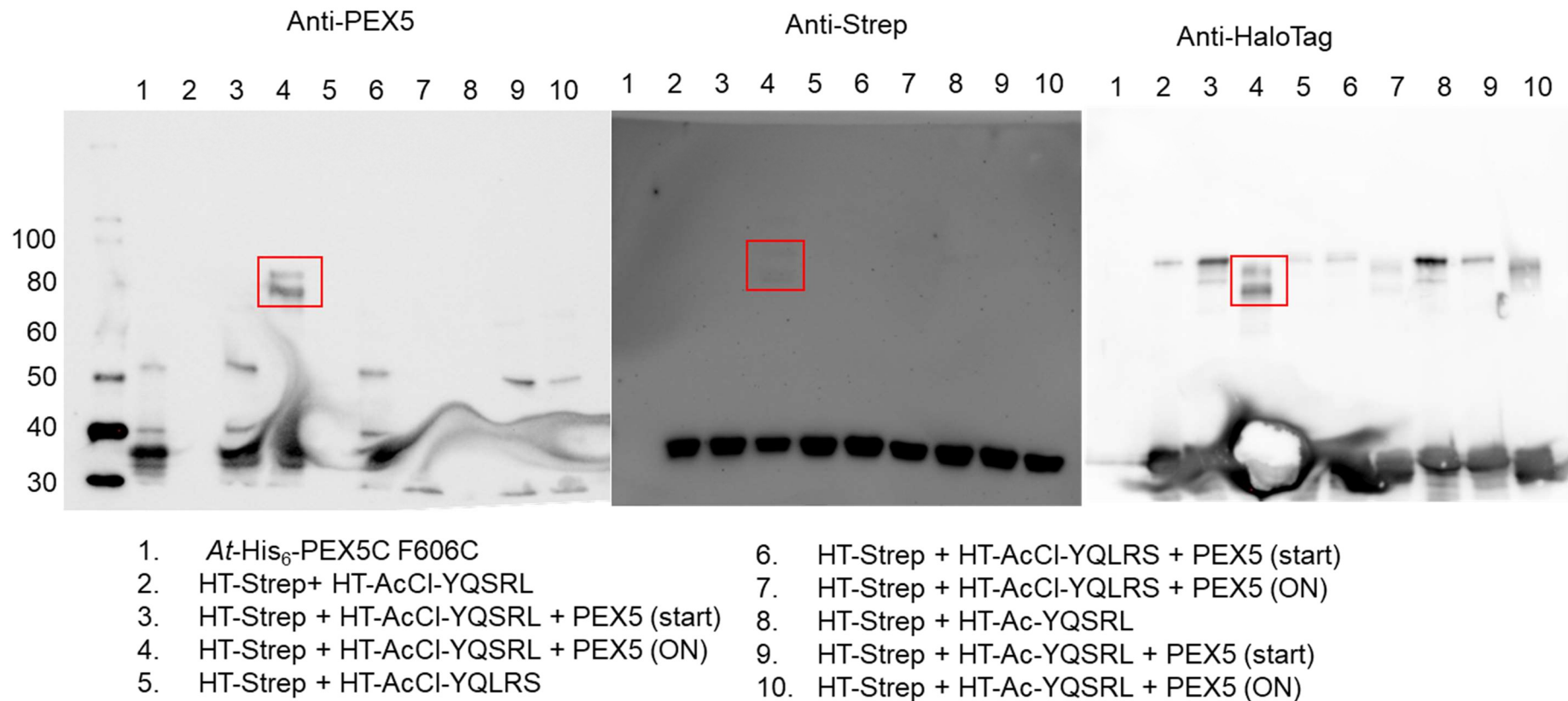
Figure 3.10 Western Blots Identifying species present in protein labelling bands SDS-PAGE gels containing samples from the reaction were blotted with Anti-PEX5, Anti-Strep-Tag and Anti-HaloTag antibodies. The blots show that after overnight incubation of the PEX5 protein with the HT-Strep protein labelled with HT-AcCl-YQSRL, the two larger protein bands detected contained both PEX5 and HaloTag epitopes.

To further investigate the labelling event, the labelling reaction was carried out in reverse. The *At*-PEX5C F606C protein was first incubated overnight with the HT-AcCl-YQSRL peptide, the excess peptide probe was removed from the reaction mixture by filtering through a 30 kDa molecular weight cut off spin concentrator. The mixture was then incubated 1:1 with the HaloTag-Strep protein overnight before analysing by Western blotting. This showed, as in the Coomassie stained gel (Figure 3.9), that there were two protein bands of approximately 85 and 110 kDa that contained both PEX5 and HaloTag epitopes but little to no evidence of the Strep-Tag epitope (Figure 3.11).

1. At-PEX5C F606C + HT-AcCl-YQSRL
2. HaloTag-Strep
3. 1 + 2 Start
4. 1 + 2 16 Hours
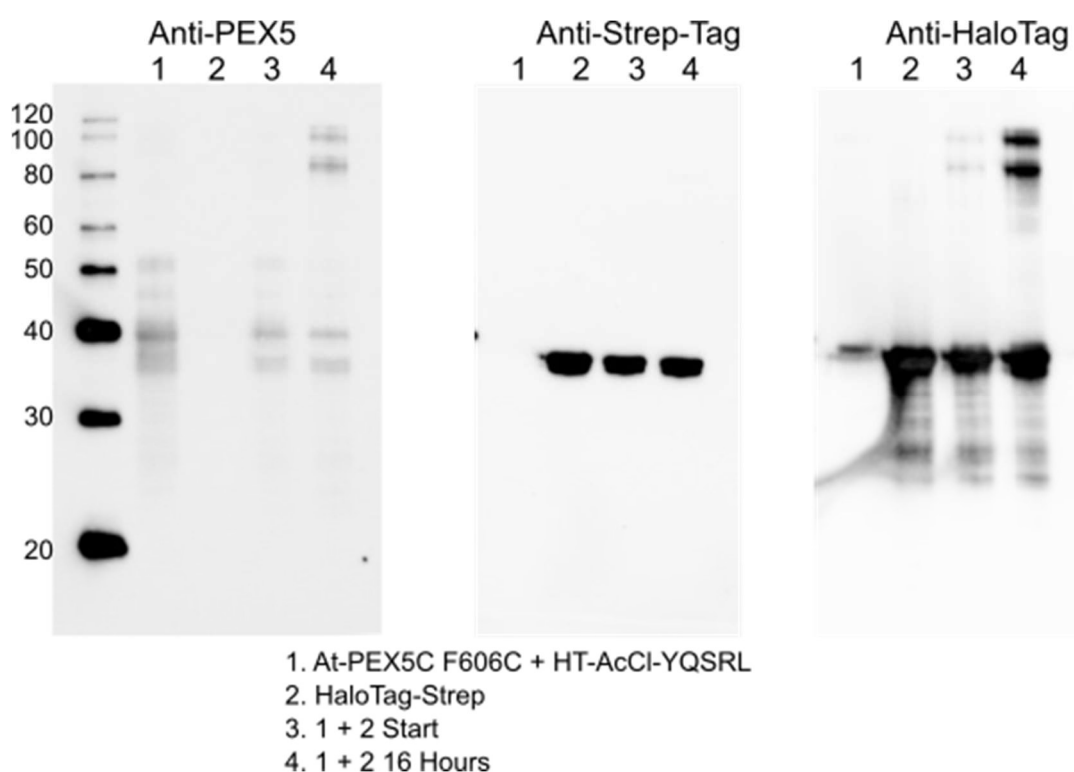
Figure 3.11 Western blots identifying species present in protein labelling bands Samples from labelling reactions were blotted with Anti-PEX5, Anti-Strep- and Tag and Anti-HaloTag antibodies to identify the species present in the 84 and 110 kDa bands seen on the Coomassie stained gel.

3.8 *At*-PEX5C F606C protein degradation

During the course of these labelling experiments, it was found that the *At*-PEX5C F606C mutant protein appeared to be prone to degradation. Many steps were taken to optimise the purification to prevent this. The degradation consistently produced a cleavage product of 36.2 kDa, a reduction in mass of 9.3 kDa. Despite identical conditions, there appeared to be variation in the amount of cleavage occurring to the protein between purifications. The source of this cleavage remains undetermined.

3.9 Discussion

This work has successfully developed a method of generating a PEX5 that can be irreversibly bound to a PTS1 cargo. By introducing a cysteine residue in the vicinity of the PTS1 binding pocket in PEX5, this enabled the protein to react with a cysteine-reactive group in a probe to irreversibly form a covalent bond with the probe. This was shown initially using probes that contained the PTS1 peptide YQSRL and a chloroacetamide group, with varying linker lengths between them. Six different cysteine mutants were all tested, in addition to the wild type protein, and of these only one, the F606C mutant, reacted with the probes. The structure of *At*-PEX5 bound to a PTS1 has not been solved, but the *Hs*-PEX5 C-terminal domain bound to YQSKL has been. The equivalent residue to F606C in *Hs*-PEX5C, Y467C, is highlighted in Figure 3.12.

Looking at the position of the residue (shown in cyan) in the crystal structure, and the position in which the PTS1 peptide (shown in magenta), the residue is present on TPR5, is close to the position of the peptide and appears to be exposed on the surface of the binding pocket (Figure 3.12). All of these factors add up to allow the covalent binding of our probe to this residue when it has been mutated to cysteine.

Further to this, the predicted structure of the Arabidopsis thaliana PEX5 protein was taken from the recently developed protein prediction tool AlphaFold (https://alphafold.ebi.ac.uk/entry/Q9FMA3). This structure was aligned with the 1FCH crystal structure of Homo sapiens PEX5 using the align function in PyMOL. This showed that the structures of the C-terminal domains of the two PEX5 proteins are predicted to be largely similar. In Figure 3.13 the 1FCH structure is showed in green and the *At* PEX5 predicted structure in orange. There is a clear structural overlap between the two proteins in spite of a sequence similarity between the two of only 44% (Appendix B.1). When looking specifically at the F606 residue, which was mutated in the Arabidopsis PEX5 protein, its location in the structure overlaps with that of the Y467 residue in the 1FCH structure. This validates the methodology used here whereby the residues to mutate were selected from the 1FCH *Hs* PEX5 structure but the mutations made on the equivalent residues in the *At* PEX5 protein.
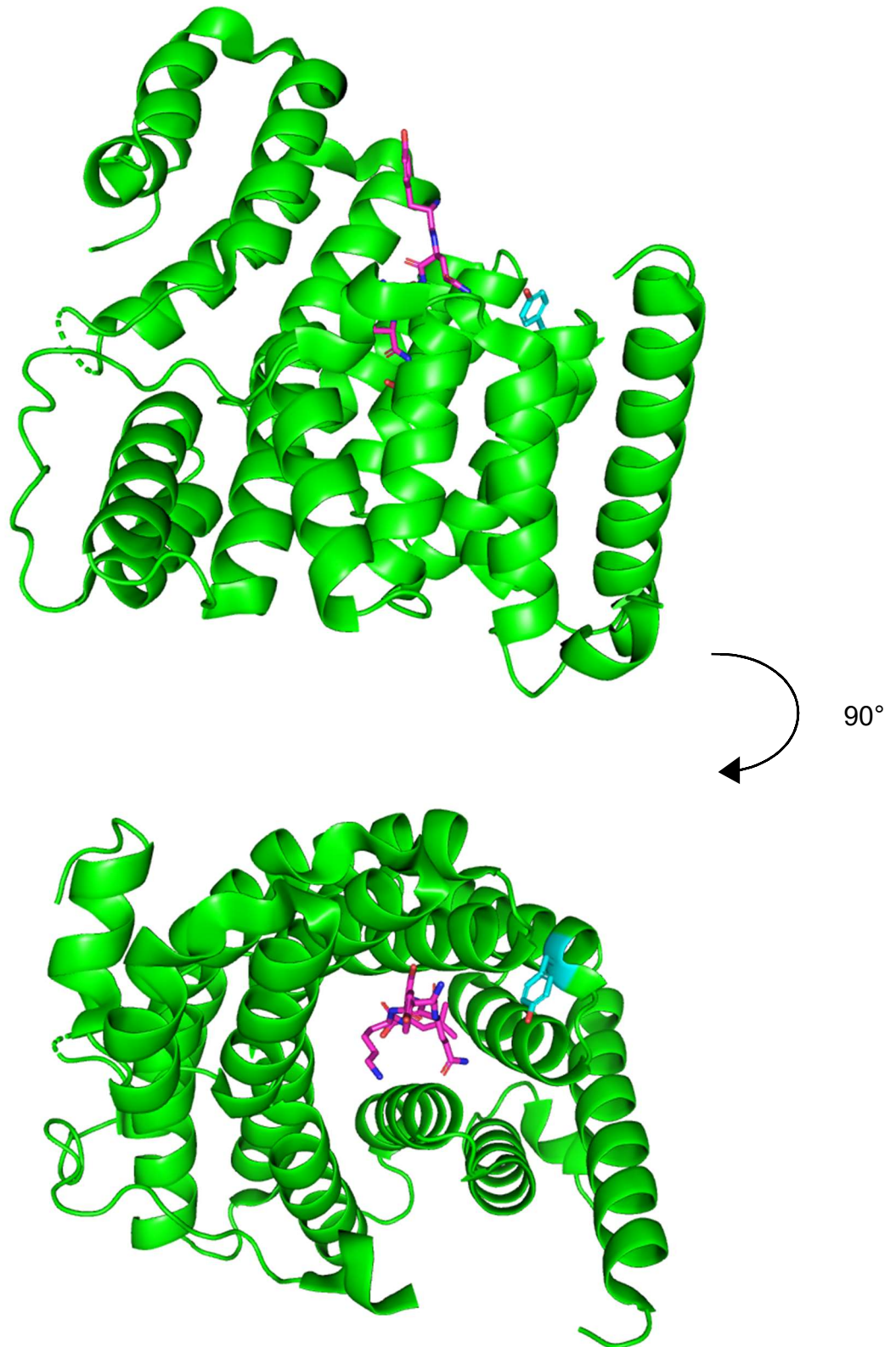
Figure 3.12 1FCH crystal structure of *Hs*-PEX5C bound to YQSKL peptide in two different views The crystal structure shows the TPRs of *Hs*-PEX5C. The Y467 residue, which is the equivalent of the F606 residue in *At*-PEX5, is shown in cyan. The YQSKL peptide is shown in magenta.
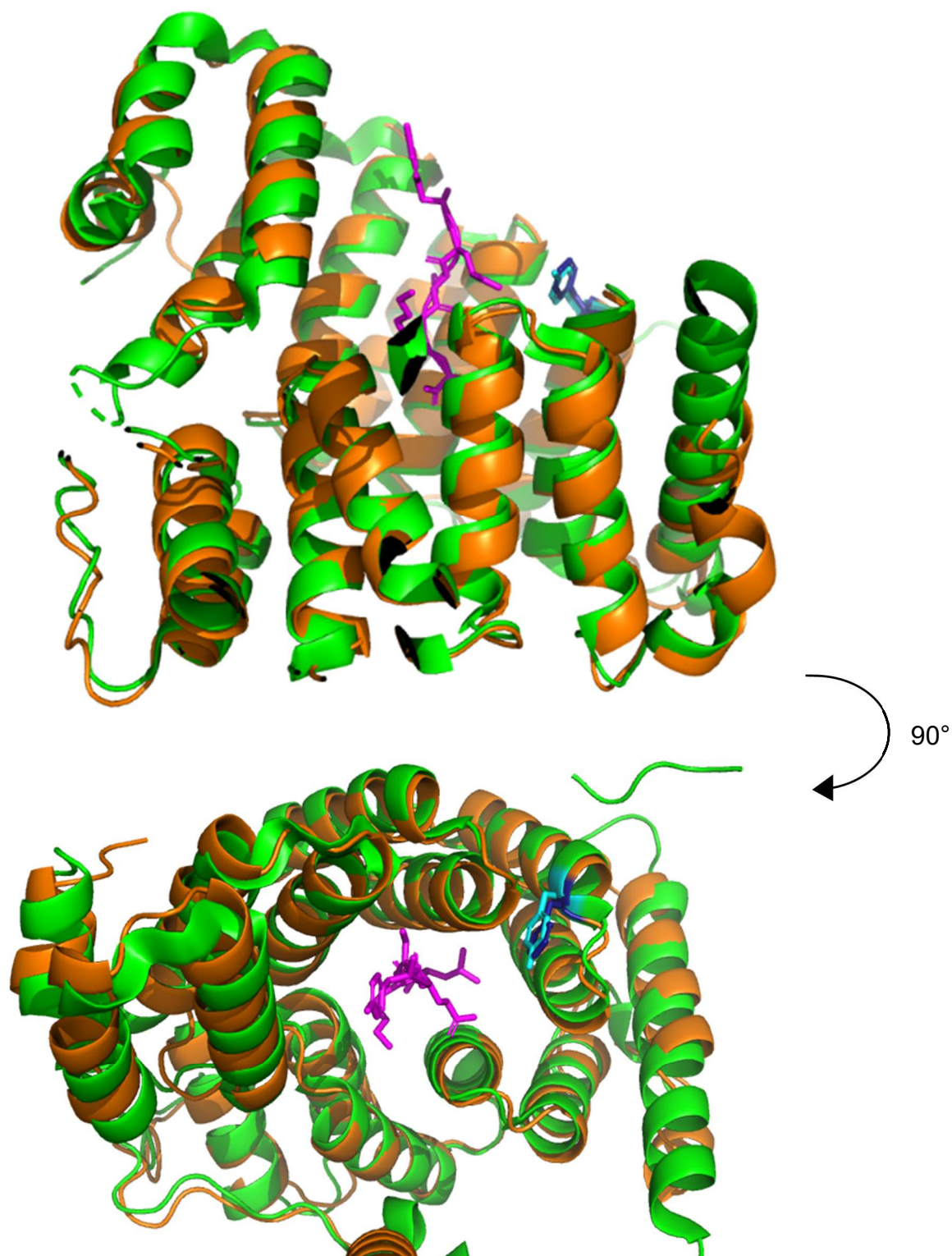
Figure 3.13 Alignment of Homo sapiens and Arabidopsis Thaliana PEX5 structures Alignment of the 1FCH crystal structure of Homo sapiens PEX5 (green) with predicted structure of Arabidopsis thaliana PEX5 (orange) generated using AlphaFold. The Y467 residue of *Hs* PEX5 is shown in cyan and the F606 residue of *At* PEX5 in navy blue. The YQSKL peptide in the 1FCH structure is shown in magenta.

The fluorescence anisotropy analysis of the F606C mutant binding to lissamine-YQSKL showed that the mutant and the wild type protein bound to the peptide near-identically. This shows that by mutating the phenylalanine to cysteine, binding to PTS1 peptides was unaffected. This makes this mutant very suitable for future studies using probes such as those synthesised in this work.

The next step for this project would look into the identification of the complex formed in the reaction between *At*-PEX5C F606C and the Halotagged protein reacted with the probe. To do this, the bands from the Coomassie gel (Figure 3.9) could be cut from the and digested into smaller peptide fragments. These fragments could then be analysed by LC-MS/MS to identify them and ultimately identify the proteins that are present in the complex. To then identify how the proteins presents in the complex are interacting with each other, an isolated complex could be used in structural studies to visualise the complex using Cryo-EM. Initially, the structure of the complex would only serve to determine the best conditions in which to obtain a high-resolution structure of PEX5 bound to a protein cargo, which is currently not a structure that has been solved. However, this work could lead to using these probes to solve the structure of PEX5 in the peroxisomal membrane. Such a structure could answer major questions in the field about how the transient pore for the translocation of PEX5 cargo proteins is formed and the interactions PEX5 makes with other peroxisomal membrane components in order to unload its cargo.

Although all three of the linker lengths tested appeared to be capable of labelling the F606C mutant of *At*-PEX5C, the AcCl-PEG-YQSRL proved to be the most efficient and reliable at doing so. Further probes generated by a Masters student in the group used the shorter peptide SRL in place of YQSRL to make a second set of probes. These probes showed that the AcCl-no linker-SRL did not react with *At*-PEX5C F606C. This demonstrates that the length of the linker between the PTS1 and the reactive group is important for a successful reaction with the F606C cysteine residue. Looking back at the crystal structure, if the reactive group was in the position where the "Q" residue is in our probes, it is likely not close enough to the reactive cysteine residue to react (Figure 3.12). By testing these other probes, we have demonstrated that the length of the linker is crucial to a successful reaction. It was for these reasons that that the subsequent HaloTag probes used the middle length of linker.

The HaloTag (or HT) probes were synthesised to demonstrate that the *At*-PEX5C F606C mutant could interact with and bind to a protein cargo, instead of only a PTS1 peptide. The HaloTag had been used previously (see: Chapter 2) to attach a PTS1 sequence to a HaloTag protein. This attachment was shown to

allow the HaloTag protein to interact with PEX5 using pulldown assays. In this instance, the probe would allow the HaloTag protein to interact with *At*-PEX5C F606C and become covalently bound to it. In doing this, *At*-PEX5C F606C would be expected to be in its cargo-binding conformation permanently, rather than transiently, as is the usual case. The results showed that on an SDS-PAGE gel, incubation of *At*-PEX5C F606C with the HaloTag protein labelled with the HT-AcCl-YQSRL probe, two bands of approximately 84 and 110 kDa were present. Western Blotting of these bands showed that they contained both PEX5 and HaloTag epitopes, but interestingly, the Twin-Strep-Tag that is present in the HaloTag protein could not be detected. This suggested that in the formation of protein complex, the epitope recognised by the Strep-Tag antibody becomes masked in some way. To investigate this further, different Strep-Tag antibodies could be used to see if a different result is obtained.

However, the fact that both the HaloTag and PEX5 antibodies bound to both of the 84 and 110 kDa bands, suggests that the complex that we would expect to form, does indeed form here. The expected complex (Figure 3.14) would be 84.2 kDa in mass and as the band at approximately 84 kDa seen on the Coomassie stained SDS PAGE gel is reliant on the presence of all three elements (*At*-PEX5C, HT-AcCl-YQSRL probe and HaloTag-Strep protein), the evidence suggests that this complex is formed (Figure 3.14).
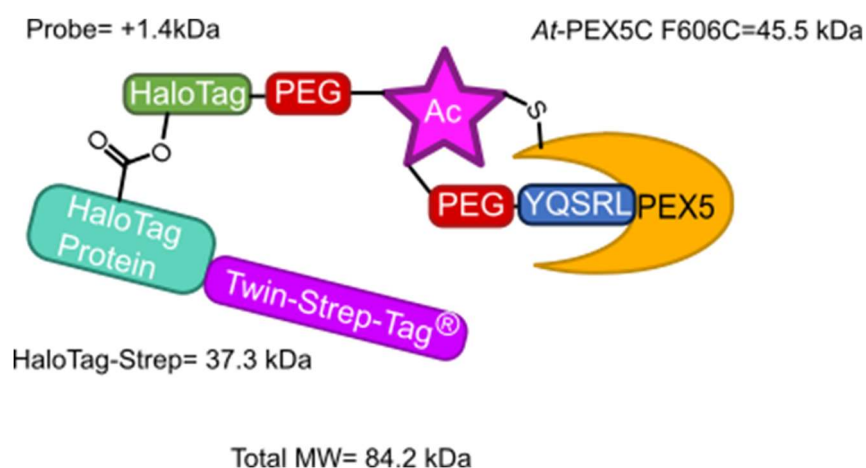
Figure 3.14 Schematic of protein complex formed in labelling reaction The complex that would be expected to form in these experiments is shown. The molecular weight of each element of the complex is also indicated.

What remains unknown, is the identity of the larger protein seen at approximately 110 kDa. The experiments here have shown that it is only formed in the presence of both proteins and the reactive probe with a PTS1 (Figure 3.9) and further experiments where *At*-PEX5C F606C was incubated with the probe alone without the HaloTag protein did not form any protein complexes that could

be detected by SDS-PAGE. To investigate this further, it would be informative to determine the exact mass of these bands. Use of a mass spectrometer that is able to detect these higher masses is being investigated and by determining the exact mass, we may be able to deduce more about the elements that are present in this band. It is possible that the band that appears to be the higher mass could represent the expected 84.2 kDa complex and the lower band the result of the formation of a similar complex with one of the degradation products of PEX5 that remains able to bind PTS1 peptides. An exact mass of these protein bands would help in determining the constituents of the complex.

The second unresolved element of these experiments is the lack of binding of the Strep-Tag antibody to the protein complex. It seems likely from these results that the formation of the complex masks the Twin-Strep-tag® on the HaloTag protein, preventing the antibody from interacting with it. Another possibility is that in the process of forming the protein complex, the Twin-Strep-tag® is somehow cleaved from the HaloTag-Strep protein, and hence would not be detected. The ability to analyse the reaction by mass spectrometry would also give greater insight into the identity of these bands. Determining the exact mass of the protein complexes would help in differentiating between these possibilities.

At present the use of these probes has not been transferrable to the *Hs*-PEX5C protein. The work in the previous chapter demonstrated that the *Hs* form of the protein bound to the PTS1 peptide less well than the *At* protein, even after cleavage of the His_6-tag from the *Hs* protein. When introducing the Y467C mutation into *Hs*-PEX5C, the equivalent of the F606C mutation in *At*-PEX5C, it did not bind as well to the PTS1 peptide as the wild type, even after cleavage of the His_6-tag. In order to be able to transfer this work into mammalian systems, more work would be required to look at finding a *Hs*-PEX5C that had a higher affinity of binding to PTS1 peptides.

Using probes such as those developed here, it would be expected that in the presence of peroxisomal membranes, PEX5 would attempt to import its cargo across the membrane but as it would be unable to unload the process would become frozen at the point of cargo-unloading. The work presented here has shown that it is possible to bind a PTS1 peptide to the C-terminal domain of PEX5 irreversibly through the formation of a covalent bond between a cysteine residue in the protein binding site and a cysteine-reactive chemical group on the PTS1 peptide. This initial finding opens up the possibility to study the intricacies of the peroxisomal protein import of PTS1 peptides. If used in a system containing peroxisomal membranes, such as the *in vitro* import assays used by

Jorge Azevedo et al (Gouveia et al., 2003a), similar such probes could be used to obtain samples of PEX5 and its associated proteins in the peroxisomal membrane, frozen at the stage of cargo release. At present, it is unclear whether PEX5 is associated with the lipids in the membrane or only with membrane proteins (see: 1.3.3.3). The process by which cargo proteins are released into the peroxisomal matrix, and proteins that may aid this process are also unknown. If these probes are able to help gain better insight into the precise mechanism of this process, by studying protein-protein interactions made by PEX5 in its cargo-bound state, it may become easier manipulate the process to use the peroxisome in biotechnology applications (see:1.6.2) and develop methods to restore proper peroxisomal import in cases where it is impaired.

An informative next step on from the experiments shown here would be to study the structure of the protein complexes formed in the 84 and 110 kDa bands. Initial steps have been discussed to be able to look at the protein complex using cryo-electron microscopy. To do this, the protein complex would need to be isolated form the reaction mixture, most likely using size exclusion chromatography. As we have shown in this work that both the PEX5 and HaloTag antibodies can bind to the complex, these could be utilised to increase the mass of the complex, which could help to solve its structure. Through the optimisation of visualising the structure of the protein complexes formed between PEX5 and its probe-labelled cargo this could help gain insight into the conformation of PEX5 when it is cargo bound. The probes developed here allow for PEX5 to be permanently cargo-bound, as opposed to only transiently as it is endogenously and in studies to look at cargo binding previously. Importantly, by stalling PEX5 in its cargo-bound state, this could give structural information about how it interacts with the peroxisomal membrane. The ability to visualise the complexes formed in this project would be the first step towards obtained a structure of cargo-bound PEX5 in a peroxisomal membrane environment. In looking at the interactions a cargo-bound PEX5 makes with proteins such as PEX13 and PEX14, this could help to uncover the unsolved mystery of exactly how the process of cargo translocation into the peroxisomal matrix occurs.

Chapter 4 General Discussion

The work presented here has made substantial progress towards the development of chemical probes to investigate and manipulate the peroxisomal import system. Two different types of probes have been developed, both containing a PTS1 peptide, to investigate different possible uses of such probes.

In Chapter 2, it has been shown that a PTS1 does not necessarily have to be present at the C-terminus for a protein to interact with PEX5. By using a probe that could attach a PTS1 peptide to a protein as a branch out from the main protein chain, the peptide was available to bind to PEX5, as if it were at the protein's C-terminus. The probes developed here covalently labelled the SNAP-Tag or HaloTag protein with a PTS1 peptide, and using *in vitro* pulldown assays it was shown that the labelled proteins were able to interact with the C-terminal TPR domains of PEX5 proteins from both humans and Arabidopsis. This is novel, as although it has been shown that by adding a PTS1 to the C-terminus of a protein genetically allows it to interact with PEX5 (Gould et al., 1989, Swinkels et al., 1992), and that PTS1 peptides alone can interact with PEX5 (Gatto Jr et al., 2000, Maynard et al., 2004), this is the first time a protein labelled with a PTS1 at an internal site using a probe has been shown to be a specific ligand for the PEX5 protein.

The development of these probes opens up the opportunity for a number of uses. Chemical probes can be used to trigger their downstream effect at a specific moment in the cell. When trying to discover new drug targets, traditionally protein-coding genes are altered or knocked down at the transcriptional or translational level (Boettcher and McManus, 2015), and if this is effective to alleviate a disease phenotype, small molecule drugs are developed to inhibit the disease-causing protein. Although these technologies have proven to be useful tools, they cannot always truly represent the mechanisms that occur *in vivo*. The flaw in this methodology, is that the protein is being eliminated by two different strategies and altering its expression at the genetic level may not translate to the same outcome as inhibiting it post-translationally (Campbell and Bennett, 2016). For example, when investigating the interactions of glycoprotein CD44 with receptor tyrosine kinase c-Met, it was found that when CD44 was knocked out genetically in mice, only mild abnormalities were seen (Olaku et al., 2011). Whereas, in cell-based assays it was shown that the interaction of CD44 with c-Met was crucial for cell survival when the interaction was prevented by use of antibodies (Olaku et al., 2011). These differences show how different methods of protein activity modulation

give a better picture of overall mechanisms. A similar idea has looked at inducing the cleavage of a specific protein through induction of expression of a protease to give temporal control of the presence of a protein (Harder et al., 2008). However, the use of probes garners the advantage that genetic manipulation is not required to be able to initiate protein knockdown temporally. Using chemical probes can help to uncover more translatable therapeutic target proteins as the action of the probes is likely to be more relevant to the action of a therapeutic and this in turn could help to develop drugs with greater efficacy (Schreiber et al., 2015).

Another way in which chemical probes can be utilised has been demonstrated in the use of PROTACs, which bind to target proteins and target them for proteasomal degradation. The relocalisation probes use here could have a similar mode of action in that they bind to a target protein and remove it from the cytoplasm, but instead of causing it to become degraded by the proteosome it is "imprisoned" within a peroxisome and will remain there until the eventual degradation of the peroxisome. The PROTAC probes have been used to demonstrate the effectiveness of inhibiting a target protein as a therapeutic strategy (Galdeano and Barril, 2021, Chessum et al., 2018).

Probes such as these could be used to better understand the actions of disease-causing proteins and therefore develop more effective therapies for a disease. Phenotypic screens test large libraries of small molecules or fragments in cells to find compounds which give the desired phenotypic changes to treat a disease (Moffat et al., 2014). One such screen identified that a ligand for pirin as an effective modulator for transcription driven by transcription factor Heat Shock Factor 1 (HSF1), which is implicated in many cancers (Cheeseman et al., 2017). To validate pirin as the true target of the probe found in this screen, protein degradation probes (PDPs) were used. The PDP was designed to contain the probe from the phenotypic screen and an E3 ligase recruitment moity. This bound to pirin and targeted it for degradation, demonstrating that pirin was the true target for the probe (Chessum et al., 2018). This example demonstrates how probes that can remove or relocalise a specific protein can be used in orthogonal assays to help validate the discovery of new drug targets.

The added benefit of these peroxisomal re-localisation probes, is that the protein-binding portion of the probe does not need to be able to both bind and inhibit the activity of the protein, as is expected of a traditional protein inhibitor. The peroxisomal re-localisation probes developed in this work need only to bind to a protein present in the cytosol and then its activity is abrogated by the fact it is removed from its usual cellular environment into the peroxisome.

Development of probes for specific protein targets would then be accelerated by the fact that screens for protein binding fragments or small molecules would only need to initially find moieties that bound the protein and not that both bound and inhibited the protein. This also opens up the use of probes such as these to be able to target a larger proportion of the proteome. At present a large proportion of disease-causing protein are considered "undruggable" as they do not have the well-defined binding pockets which can be bound by small molecules to inhibit their mode of action (Crews, 2010). Proteins which act in regulatory or scaffolding functions, for example, might be better targeted by PROTACs or relocalisation probes which only need to bind to the protein at any site, rather than a specific active site pocket. For example, PROTACs have been developed to target the Tau protein (Chu et al., 2016), found to aggregate in Alzheimer's disease, which could in the future be used as a treatment to break down protein aggregates. In a similar way to which PROTACs are being developed into therapeutics, these peroxisomal targeting chimeras could also be used as an alternative method to inhibit proteins.

These probes could also be used to investigate the peroxisomal targeting of endogenous peroxisomal matrix proteins. To better understand the mechanisms at play in peroxisomal enzyme deficiencies, probes such as those developed here could be used to study the effects of preventing a protein from being imported into the peroxisomal matrix by removing its PTS and subsequently causing it to be imported using a probe and analysing whether this is effective in altering the cell phenotype. This could be integrated onto the work being carried out by Schuldiner's lab looking at targeting priority of peroxisomal matrix protein import (Rosenthal et al., 2020). By using probes with differing "strengths" of PTS1, the peroxisomal proteome could be altered and the effects of importing matrix proteins in different proportions studied. This could give insight into the exact mechanisms at play that cause peroxisomal dysfunction and disease and may uncover new strategies to alleviate disease symptoms.

After establishing that a protein can be made into a PEX5 cargo protein by adding a PTS1-containing probe, Chapter 3 looked to build on this by developing a probe to irreversibly bind a protein cargo to PEX5 by introducing a covalent interaction between the probe and a cysteine residue at the PEX5 binding site. The work presented here showed the successful development of a PEX5 cysteine mutant which bound to a PTS1 peptide with the same affinity as the wild type PEX5 and subsequently reacted with the probe to create a form of PEX5 that is permanently cargo-bound.

A permanently bound form of PEX5 has the potential to give a better understanding of how PEX5 delivers its cargo across the peroxisomal membrane and answer some of the most prominent questions in the peroxisomal field. As the mutant protein generated was in the Arabidopsis PEX5 protein, it could be used in well-established *in vitro* import assays (Bhogal et al., 2016, Rodrigues et al., 2016). These assays use PEX5 or cargo proteins, which have been radio-labelled, with either isolated peroxisomes (Arabidopsis) or postnuclear supernatant (PNS) (mammalian cells). After treating with the radio-labelled reporter, proteases are used to degrade any proteins that are not associated with peroxisomes. By incubating with a radio-labelled version of the F606C mutant, and a protein cargo labelled with a cysteine-reactive PTS1 probe, the interactions made by PEX5 at the initial formation of the membrane docking and translocation complex (DTM) before the unloading of cargo could be investigated.

These assays have long been used to investigate peroxisomal import mechanisms and more recently have been used extensively in the work studying the molecular mechanisms of peroxisome biogenesis and matrix protein import in mammalian cells by the Azevedo group at the University of Porto (Gouveia et al., 2003). In order to use probes such as those developed here in mammalian peroxisome import assays, a mammalian PEX5 mutant protein that reacts well with the probe would need to be found. In the work carried out here the Y467C mammalian PEX5 mutant, the equivalent of the F606C mutant of Arabidopsis PEX5 mutant, was investigated but found to have poor binding to PTS1 peptides and did not react with the probes. The work carried out in Chapter 2 showed that this human PEX5 construct does not bind well to PTS1 peptides *in vitro* and optimisation of the construct used could help to develop a better reactive mutant.

At present, PEX5 has only been able to be stalled at the stage of its monoubiquitination after cargo unloading by removing the N-terminal conserved cysteine residue (Francisco et al., 2013). It has been shown that docking of PEX5 to the peroxisomal membrane is cargo dependent (Gouveia et al., 2003b) and that there are two distinct docking and membrane insertion steps (Francisco et al., 2013). The PEX5-reactive probes and mutant protein pairing developed here could be used to help get a better picture of how the transition from docking to insertion occurs as all that is currently known is that these steps occur before the monoubiquitination of PEX5.

It is postulated that the binding of these probes to PEX5 irreversibly would generate a form of PEX5 that can interact with the peroxisomal membrane to

form the import complex but would not be able to proceed to subsequent steps of the cycle. This form of PEX5 has the potential to aid in structural studies to elucidate the structure of the importomer for proteins across the peroxisomal membrane. Recent work at the Max Planck Institute made the first steps towards obtaining a structure of the complex formed at the peroxisomal membrane by characterising the topology and rod-like structure of the yeast Pex14p/Pex17p complex in nanodiscs (Lill et al., 2020). Use of the probes developed here could be the next step to obtaining the structure of higher order complexes that are formed at the peroxisomal membrane and solve the debate as to how exactly a pore is formed through which to translocate proteins destined for the peroxisomal matrix.

The importance of chemical probes to investigate biological systems is becoming more widely recognised as we seek to be able to modulate the activities of more challenging targets. In this work it has been shown that the peroxisome presents a unique opportunity to use such probes to broaden our knowledge of protein activities as well as better understand mechanisms that until now have been difficult to definitively elucidate and visualise.

Chapter 5

Materials and Methods

5.1 Materials

5.1.1 Bacterial Strains

XL10-Gold Ultracompetent Cells (Agilent Technologies)

Genotype: TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F′ proAB lacIqZDM15 Tn10 (Tetr) Amy Camr]

BL21-Gold (DE3) (Agilent Technologies)

Genotype: E. coli B F– ompT hsdS(rB– mB–) dcm+ Tetr gal λ(DE3) endA The

Rosetta™ 2 (DE3) (Novagen)

Genotype: F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE2 (CamR)

BL21-CodonPlus(DE3)-RIL-X strain (Agilent Technologies)

Genotype: F– ompT HsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA metA::Tn5(kanr) Hte [argU ileY leuW Camr]

5.1.2 Plasmids

The SNAP-Tag gene was cloned into the pET 12b vector downstream of a Twin-Strep-tag® by Genscript.

The HaloTag7 gene was cloned into the pPSGIBA103 vector using StarGate cloning as described in Appendix A.

The *Hs*-PEX5(235-602) gene was cloned downstream of a thrombin cleavage site and a His_6 tag in a pET 28b vector by Genscript

The *At*-His_6-PEX5(340-728) gene was originally cloned into pET-28b by S. Gunn (former student of University of Leeds).

The SNAP-tag and HaloTag7 genes were cloned into pEGFP-C1 vector as part of this work.

For plasmid maps, see Appendix A

5.1.3 Bacterial Growth Media

All bacterial media was made up in double distilled water and autoclaved at 121°C for 20 minutes. Sugars for AIM were filter-sterilised before adding to autoclaved media.

LB Media

Tryptone (1% w/v), yeast extract (0.5% w/v), NaCl (1% w/v)

LB-agar plates

Tryptone (1% w/v), yeast extract (0.5% w/v), NaCl (1% w/v), agar (1.5% w/v)

Autoinduction Media (AIM)

Tryptone (1% w/v), yeast extract (0.5% w/v), NaOH (1 mM), $MgSO_4$ (1 mM), $(NH_4)_2SO_4$ (25 mM), KH_2PO_4 (50 mM), Na_2HPO_4 (50 mM), glycerol (0.5% w/v), glucose (0.05% w/v), α-lactose (0.2% w/v)

Antibiotic Stock Solutions

All solutions were made up in a 1000X stock and filter sterilised by passing through a 0.2 μM syringe filter.

Antibiotic	Stock Concentration	Solvent
Kanamycin	50 mg/mL	ddH_2O
Ampicillin	100 mg/mL	ddH_2O
Chloramphenicol	30 mg/mL	EtOH

5.1.4 Buffers

Phosphate Buffered Saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10mM Na_2PO_4, 1.8 mM KH_2PO_4, pH 7.4

SDS-PAGE Running Buffer: 25 mM Tris, 192 mM Glycine and 0.1 % (w/v) SDS.

2 x SDS-PAGE Sample Buffer: 100 mM Tris/HCl, pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 10 mM DTT, 0.25% (w/v) bromophenol blue

Coomassie Brilliant Blue Stain Solution

0.1% (w/v) Coomassie Brilliant Blue (Bio-Rad), 50% (v/v) methanol, 10% (v/v) Acetic Acid, 40% (v/v) H_2O

Coomassie Destain Solution

50% (v/v) methanol, 10% (v/v) Acetic Acid, 40% (v/v) H_2O

HisPur™ Cobalt Resin buffers

Lysis buffer: 50mM NaH_2PO_4, 400 mM NaCl, pH 8.0

Wash Buffer: 50 mM sodium phosphate, 400 mM sodium chloride, 20 mM imidazole; pH 7.4

Elution Buffer: 50 mM sodium phosphate, 400 mM sodium chloride, 250 mM imidazole; pH 7.4

Strep-Tactin Superflow high-capacity resin buffers

Buffer W: 100 mM Tris/HCl, pH 8.0 150 mM NaCl 1 mM EDTA

Buffer E: 100 mM Tris/HCl, pH 8.0 150 mM NaCl 1 mM EDTA 2.5 mM desthiobiotin

MagStrep "type3" XT beads

Buffer W: 100 mM Tris/HCl, pH 8.0 150 mM NaCl 1 mM EDTA

Buffer BXT: 100 mM Tris/HCl, pH 8.0 150 mM NaCl 1 mM EDTA 50 mM Biotin

Protease inhibitors

COmplete EDTA-free, Roche were used at a final concentration of 1 tablet per 50 mL buffer in lysis buffer

FA buffer

HEPES (20 mM), NaCl (150 mM), pH 7.5

Plate blocking buffer

Gelatin from porcine skin, type A (Sigma) (0.32 mg/mL) in FA buffer

Western Blot Transfer Buffer: 48 mM Tris, 39 mM glycine, 20% methanol (v/v), 0.04% SDS (w/v)

10x Tris-buffered Saline (TBS): 200mM Tris base, 1500 mM NaCl, pH 7.6

TBS-T: 10% 10x TBS (v/v), 0.1% TWEEN 20 (v/v)

Western Blot Blocking Buffer: 5% skimmed milk powder (w/v) in TBS-T

5.1.5 Antibodies

5.1.5.1 Primary Antibodies

Raised Against	Manufacturer	Cat. Number	Species Raised In	WB Concentration
At-PEX5C	GenoSphere	Custom Order	Rabbit	1:10000
Strep-Tag	GeneTex	GTX628900	Mouse	1:5000
HaloTag	Promega	G9211	Mouse	1:10000

5.1.5.2 Secondary Antibodies

Name	Manufacturer	Cat. Number	WB Concentration
Anti-Rabbit IgG HRP	Jackson ImmunoReasearch	AB_2313567	1:5000
Anti-Mouse IgG HRP	GeneTex	GTX221667-01-S	1:5000

5.2 Methods

5.2.1 Classical Cloning

To generate plasmids expressing EGFP-SNAP-Tag or EGFP-HaloTag7 proteins in mammalian cells, the genes were cloned into the pEGFP-C1 vector. To do this, primers were designed to introduce *HindIII* and *KpnI* sites (Appendix B) at the beginning and end of the gene-encoding sequence. These primers were used to amplify the gene from their bacterial expression plasmids by PCR. These inserts were purified from the PCR reaction mixture using the QIAquick PCR purification kit (QIAGEN).

The PCR reaction was set up as follows:

Component	Concentration	Volume/ µL
KOD Hot Start DNA Buffer	10x	10
dNTPs	2 mM each	10

MgSO₄	25 mM	6
Forward Primer	10 µM	3
Reverse Primer	10 µM	3
Template DNA	20 ng/µL	1
KOD Hot Start DNA Polymerase	1U/ µL	2
Sterile autoclaved ddH₂O	N/A	65 µL

PCR conditions were: 95°C 2' + 40x(95°C 20", 55°C 10", 70°C 20")

An aliquot from the reaction (5 µL) was mixed with 5X loading dye and run on 2% agarose gel with 1X SYBRSafe (30 minutes, 100V) and visualised using a ChemiDoc MP (Bio-Rad). The PCR reaction could be frozen until ligation was performed.

DpnI enzyme (1 µL) was added to the PCR reaction mixture and incubated for 2 hours at 37°C to degrade template DNA. Product was purified using QIAquick PCR purification kit (QIAGEN) according to manufacturer's instructions. PCR product was eluted in 50 µL sterile autoclaved ddH₂O. This was digested with *KpnI* and *HindIII* in the following reaction mixture:

PCR Reaction Product	43 µL
NEB Buffer	5 µL
KpnI	1 µL
HindIII	1 µL

The digestion mixture was purified using the QIAquick PCR purification kit (QIAGEN) according to manufacturer's instructions.

Insert DNA was ligated into pEGFP-C1 vector through a ligation reaction. Reactions were prepared in 250 µL PCR tubes thus:

Reaction	Vector (40ng/ µL)/ µL	Insert (110ng/ µL)/ µL	H₂O/ µL	T4 Buffer/ µL	T4 Ligase/ µL
Vector Control	0.1	0	7.9	1	1
Insert Control	0	0.2	7.8	1	1

3X	0.1	0.1	7.8	1	1
6X	0.1	0.2	7.7	1	1

Tubes were incubated at 16°C overnight and then heated to 65°C for 20 minutes to inactivate the ligase.

The pEGFP-C1 plasmid was digested using *HindIII* and *KpnI* restriction enzymes. Reactions were incubated at 37°C with shaking for 3 hours. Calf Intestinal Alkaline Phosphatase (CIP) (1 µL) was added to the vector reaction mixture to dephosphorylate cut ends and prevent relegation of the vector. Completed reactions were run on a 1% agarose gel (45 minutes, 100V). Vector DNA was cut from the gel and purified using the QIAprep Gel Extraction Kit (QIAGEN) according to manufacturer's instructions.

The digested and purified plasmid was incubated with one of the amplified genes and DNA ligase overnight at 16°C. It was then heated to 65°C for 20 minutes to inactivate the ligase.

To propagate the plasmid, ligation reactions were transformed in to XL10-Gold Ultracompetent cells and plated on to LB agar with ampicillin selection. Colonies from the 6X plate were grown in LB media with ampicillin selection overnight and plasmid DNA extracted using the QIAprep spin Miniprep kit (QIAGEN) according to manufacturer's instructions. Samples were sent for sequencing to ensure successful ligation of insert DNA.

5.2.2 StarGate Cloning

A gBlock® gene fragment was designed to encode the HaloTag protein known as HaloTag7 (Ohana et al., 2009) and cut sites for the Type IIS restriction endonuclease enzyme Esp3I at each end and purchased (Integrated DNA Technologies). The plasmid pPSG-IBA103 was purchased from IBA Life Sciences and amplified in XL-10 *E. coli* cells and purified by Miniprep (Qiagen).

The reaction mixture was prepared as follows:

Reagent	Concentration	Volume
pPSG-IBA103 Plasmid	0.66 ng/ µL	7.5 µL
HaloTag7 gBlock®	2 ng/ µL	12.5 µL
ESP3I	10U/ µL	0.5 µL
T4 DNA ligase	1U/ µL	1 µL
CutSmart Buffer	N/A	2.5 µL

DTT	250 mM	0.5 µL
ATP	12.5 mM	0.5 µL

The reaction mixture was incubated at 30°C for 1 hour. This allowed the gBlock® and the acceptor plasmid to both be cut with Esp3I to reveal compatible ends to then be ligated by the ligase enzyme.

Colony PCR was performed on colonies that appeared white to determine those that contained the HaloTag-encoding plasmid.

5.2.3 Colony PCR

A single colony was picked from a fresh agar plate containing cells transformed with the desired plasmid. The colony was dissolved in ddH$_2$O (50 µL). A reaction mixture was prepared as follows:

Reagent	Concentration	Volume
Colony in ddH$_2$O	1 colony/ 50 µL	2 µL
10 Taq reaction Buffer	10X	2.5 µL
Forward Primer	10 µM	0.5 µL
Reverse Primer	10 µM	0.5 µL
Hot Start Taq DNA Polymerase	5 U/ µL	4 µL
ddH$_2$O	N/A	15.5 µL

The reaction was run under the following cycle:

95°C 2' + 30x(95°C 30", 55°C 50", 72°C 2'30")

After the cycle had completed a sample from the reaction was run on a 2% agarose gel with 1X SYBRSafe and a with a molecular weight marker (30 minutes, 100V) and visualised using a the ChemiDocMP (Bio-Rad) to determine if the target fragments was amplified and therefore was present in the plasmid.

5.2.4 Site-directed mutagenesis

A QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) was used according to manufacturer's instructions. Transformation of each DpnI-treated mutagenesis reaction into XL10-Gold cells was performed, followed by DNA purification and confirmation of mutagenesis by Sanger sequencing (Eurofins Genomics).

5.2.5 Expression and Purification of recombinant proteins

Protein	Expression Vector	Antibiotic Resistance	Expression Cell Line	Expression Induction Method	Cell Lysis Method
Hs-His$_6$-PEX5C	pET-28b	Kanamycin	Rosetta 2 (DE3)	IPTG	Cell Disrupter
At-His$_6$-PEX5C (and mutants)	pET-28b	Kanamycin	BL21 Gold (DE3)	IPTG	Sonicator
SNAP-Strep	pET-12b	Ampicillin	BL21-Gold (DE3)	Autoinduction	Cell Disruptor
HaloTag-Strep	pPSG-IBA103	Ampicillin	BL21-Gold (DE3)	Autoinduction	Cell Disruptor

5.2.6 Transformations

Competent *E. coli* expression cells (50µL) were aliquoted into 1.5 mL Eppendorf tubes containing the desired plasmid (2µL) to be transformed and incubated on ice for 30 minutes. Cells were heat shocked in a 42°C water bath for 30 seconds. Cells were incubated on ice for 1 minute before the addition of 450µL LB media at 42°C. Tubes were incubated at 37°C for 1 hour with shaking before plating on LB-Agar plates with appropriate antibiotic selection. LB-agar plates with no antibiotic were used as a positive control, untransformed cells were plated with antibiotic as a negative control. All plates were incubated ay 37°C for 16-18 hours.

5.2.7 Autoinduction

A single colony from an LB-agar plate of transformed cells was grown for 16-18 hours in LB media (5 mL) with the appropriate antibiotic selection at 37°C. The overnight culture (20 μL) was diluted in 1 mL LB media to make the day culture with antibiotic selection and cultured for a further 8 hours at 37°C. Flasks with sterile AIM (500 mL per 2L flask) and appropriate antibiotic selection were inoculated with day culture (1:2000) and incubated at 28°C for 18 hours. Cells were then pelleted by centrifugation at 4,500 x g for 20 minutes and either immediately lysed or stored at -80°C until required.

5.2.8 IPTG induced expression

A single colony from an LB-agar plate of transformed cells was grown for 18 hours in LB media (50 mL) with the appropriate antibiotic selection at 37°C. Overnight culture (50 mL) was diluted in 1 L LB media with antibiotic selection and cultured at 37°C until OD_{600}=0.6-0.8. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4mM and the flask incubated at 18°C overnight. Cells were then pelleted by centrifugation at 4,500 x g for 20 minutes and either immediately lysed or stored at -80°C until required.

5.2.9 Cell Lysis

5.2.9.1 Cell disruption

Pellets from bacterial cell culture were resuspended in chilled lysis buffer (30mL per 500mL culture pellet) supplemented with protease inhibitors. Once a smooth suspension was achieved the cells were transferred to the inlet of the cell disrupter (Constant Systems) and broken apart by applying 20 kpsi. Cells were chased with an equal volume of chilled lysis buffer with protease inhibitors. The slurry was passed through the disrupter a second time. Lysed cells were then cleared by centrifugation at 25,000xg for 30 minutes and protein purified from the supernatant.

5.2.9.2 Sonication

Pellets from bacterial cell culture were resuspended in chilled lysis buffer supplemented with protease inhibitors (20 mL per 1L culture). Once fully resuspended cells were sonicated for 1 x 5 minutes with the Bandelin SONPULS ultrasonic homogeniser with a MS 73 micro tip probe. Lysed cells were then cleared by centrifugation at 25,000 x g for 30 minutes and protein purified from the supernatant.

5.2.10 Purification of His$_6$-tagged proteins

HisPur™ Cobalt Resin (Fisher Scientific) (1mL per 500 mL culture) was equilibrated in lysis buffer in a polypropylene column. Cell lysate was passed over the resin twice. Resin was washed with two resin bed volumes of wash buffer and the flow through collected. This step was repeated until the absorbance at 280 nm measured <0.090 using the NanoDrop™ 2000 spectrophotometer. His-tagged protein was eluted from the column through addition of one resin bed volume of elution buffer until the absorbance at 280 nm measured <0.090 using the NanoDrop™ 2000 spectrophotometer.

5.2.11 Purification of Twin-Strep-tagged proteins

Strep-Tactin® Superflow® high capacity 50% suspension (IBA Life Sciences) (1mL per 500 mL culture) was equilibrated in 2 column volumes buffer W in a polypropylene column. Cell lysate was passed over the resin twice. Resin was washed with one resin bed volume of buffer W and the flow through collected. This step was repeated 4 times. Twin-Strep-tagged protein was eluted from the column through addition of one resin bed volume of elution buffer (Buffer E) 6 times.

5.2.12 Size exclusion chromatography

Where necessary, proteins were further purified after affinity chromatography by fast protein liquid chromatography on an Superdex™ 200 column. The column was equilibrated in 1.5 column volumes of wash buffer before sample was injected via a loop and buffer flowed through the column at a rate of 1 mL/minute until a steady baseline of absorbance at 280nm was reached. 4 mL fractions were collected and analysed by SDS-PAGE.

5.2.13 Buffer exchange and concentration

To remove small molecules used in elution buffers from eluted protein samples, proteins were buffer-exchanged into PBS using PD-10 desalting columns (GE Healthcare) according to manufacturer's instructions. After buffer exchange, proteins were concentrated to the desired concentration using Amicon® Ultra-15 Centrifugal Filter Units (Millipore). Protein was then used immediately or flash-frozen with liquid N_2 and stored at -80°C.

5.2.14 Protein concentration determinations

Protein concentration was determined by the Beer-Lambert law using the absorbance at 280 nm, measured using a NanoDrop™ 2000 spectrophotometer. Extinction coefficients were calculated using ExPASy ProtParam.

5.2.15 SDS-PAGE gels

Protein samples were separated by their molecular weight using sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE).

Depending on the size of protein required to be separated, the appropriate percentage SDS PAGE gels were poured. Separating gels were poured first and allowed to set with isopropanol in place of the stacking gel. This was then poured off and the stacking gel poured on top. A well comb is inserted, and the stacking gel allowed to set.

Recipe for separating gels (yields 2 gels):

	Volume/ mL		
	8%	12%	15%
1.5M Tris.HCl pH 8.8	2.533	2.533	2.533
40% Acrylamide	2.000	3.000	3.750
H_2O	5.257	4.257	3.507
10% SDS	0.100	0.100	0.100
10 %Ammonium Persulfate	0.100	0.100	0.100
TEMED	0.010	0.010	0.010

Recipe for stacking gel:

	Volume/ mL
0.5M Tris.HCl pH 6.8	0.945
40% Acrylamide	0.625
H_2O	3.320
10% SDS	0.050
10 %Ammonium Persulfate	0.050
TEMED	0.010

Protein samples were mixed 1:1 with 2 x SDS loading dye subsequently boiled at 95°C for 5 min and loaded alongside the Unstained Protein Standard Broad Range (10-200 kDa) (New England Biolabs Inc.) onto polyacrylamide gels. Gels were run in SDS running buffer at 180 V for 45 min or until the dye front reached the end of the separating gel.

To visualise protein bands, gels were removed from glass plates and incubated with Coomassie Brilliant Blue Stain Solution for 1 hour with agitation. Stain solution was removed and replaced with destain solution. Gels were incubated with destain solution until protein bands could be clearly imaged using the ChemiDocMP (Bio-Rad).

5.2.16 In vitro labelling of SNAP-Tag-Strep protein with peptide-functionalised SNAP-Tag substrate

To label the SNAP-Tag-Strep protein with SNAP-Tag substrate-containing peptides, the following components were combined in a tapered HPLC vial:

	Concentration	Volume	Final Concentration
DTT	50 mM	1 µl	1 mM
SNAP-Tag-Strep protein	50 µM	25 µl	25 µM
Peptide Substrate	100 µM	24 µl	48 µM

Once the peptide substrate was added and the tube mixed by vortexing it was allowed to stand on ice for 1 hour. The mixture was analysed by mass spectrometry.

5.2.17 In vitro labelling of HaloTag-Strep protein with peptide-functionalised HaloTag substrate

To label the HaloTag-Strep protein with HaloTag substrate-containing peptides, the following components were combined in a tapered HPLC vial:

	Concentration	Volume	Final Concentration
HaloTag-Strep protein	50 µM	25 µl	25 µM
Peptide Substrate	100 µM	25 µl	50 µM

Once the peptide substrate was added and the tube mixed by vortexing it was allowed to stand on ice for 1 hour. The mixture was analysed by mass spectrometry.

5.2.18 Pull-down assays for detection of protein-protein interactions

After labelling of the SNAP-Tag-Strep or HaloTag-Strep protein with the peptide, excess peptide was removed by applying the reaction mixture to a Amicon Ultra-0.5 Centrifugal Filter Unit with a molecular weight cut off of 10 kDa (Millipore). This was centrifuged for 10 minutes at 10,000 x g at 4°C.

PEX5C proteins (100 μM, 25 μL) were then added to the mixture and the two proteins incubated on ice for 1 hour.

MagStrep XT magnetic bead suspension (30 μL) was equilibrated according to manufacturer's instructions. The protein mixture was incubated on 30 μL magnetic bead suspension on ice for 30 minutes with occasional vortexing to bring the beads into suspension. The tube was then placed on a magnetic separator and supernatant collected. Beads were then washed 4 times with buffer W (150 μL) by adding the buffer, vortexing and replacing the tube in the magnetic separator and collecting the supernatant. Proteins were then eluted from the beads by adding Buffer BXT (37.5 μL) and incubating for 10 minutes on ice with occasional vortexing. The tube was replaced in the magnetic separator and supernatant collected. This was repeated 3 times. A sample of the beads after elutions was also taken to analyse by SDS-PAGE.

5.2.19 Cell-based assay for peroxisomal re-localisation probes

COS-7 cells were obtained from the American Type Culture Collection (ATCC; Rockville, MD) and were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 2 g/litre sodium bicarbonate, 2 mM glutamine, 100 U/ml penicillin, 100 μg/mL streptomycin, and 10% foetal calf serum (FCS) (all from Gibco BRL; Gaithersburg, MD) at 37°C in a humidified atmosphere containing 5% CO2.

For DEAE-Dextran transfection 2.5×10^4 cells/cm^2 were seeded on coverslips (in 6-cm dishes) 24 hours prior to transfection. Plasmids to be transfected (5-10 μg) were mixed DMEM supplemented with 100 U/ml penicillin, 100 μg/ml streptomycin (1.5 mL) and DEAE-Dextran (9 μL) and incubated for 15 minutes at room temperature. Cells were washed twice with PBS before adding the DNA-DEAE mix to the dish. Cells were incubated for 90 minutes with gentle agitation every 15 minutes to prevent drying out of coverslips. DNA-DEAE mix was then removed and replaced with DMEM with chloroquine (10 μL) and incubated for 3 hours at 37°C. DMEM with chloroquine was removed, cells

washed with PBS and then changed for complete DMEM fresh media. Cells were incubated overnight.

Probes were added to cells at concentrations of 0-20 µM through direct pipetting into media surrounding the coverslip. Cells were incubated with probes for 8-24 hours. After incubation cells were stained for peroxisomes by immunofluorescence.

Cells were fixed using 4% paraformaldehyde, pH 7.4 for 20 min at room temperature. Following fixation, cells were washed 3 times in PBS, membranes were permeabilised with 0.2% Triton X-100 for 10 min at room temperature. Cells were washed again in PBS. Cells were blocked with 2% bovine serum albumin (BSA) for 10 min at room temperature. Cells were then incubated for 1 hour with rabbit polyclonal antibody against PEX14 diluted in PBS (1:1400), washed three times in PBS and incubated for 1 h in the dark with Goat anti-Rabbit IgG AlexaFluor 594 secondary antibodies diluted in PBS (1:500).

To mount coverslips on slides, cells were washed three times in PBS, dipped in distilled H_2O and mounted (after removal of excess water) in Mowiol 4-88 containing n-propyl galate for anti-fading.

5.2.20   Fluorescence Anisotropy Assays

Lissamine-YQSKL peptide was used to assess peptide binding ability of different PEX5 proteins used in this work. The Lissamine-YQSKL peptide was synthesised and purified as detailed (5.3.20).

Proteins for analysis were buffer exchanged into FA buffer and concentrated to 30 µM.

Plate blocking buffer (80 µL) was added to all wells of a Black Optiplate™-384 F (PerkinElmer) 384-well plate. The plate was sealed with aluminium sealing film (StarLab) and refrigerated for at least 14 hours.

To carry out the assay, 60 µL of plate blocking buffer was removed from each of the required wells. Protein solution (40 µL) was added to rows A-F in column 1. This allowed for 3 repeats of test and control wells. The solution was agitated by pipetting to mix the protein solution with the remaining blocking buffer. 40 µL was removed from each well and transferred to column 2. The process of mixing and transferring solution to the subsequent column was repeated to dilute the protein across 23 columns. 40 µL of FA buffer was added, mixed and removed from column 24. This gave a range of concentrations form 0-10 µM.

A solution of Lissamine-YQSKL in FA buffer (200 nM) was prepared. Lissamine-YQSKL (20 µL) was added to all wells in rows A-C. FA buffer (20 µL) was added to all wells in rows D-F.

The plate was read using the EnVision™ 2103 multilabel plate reader (Perkin Elmer) using a method developed by previous student Laura Cross (Cross, 2016) as detailed below.

The values obtained by reading using a BODIPY TMR dichroic mirror (555 nm) and the following filters:

Excitation: BODIPY TMR FP 531 (Wavelength 531 nm, bandwidth 25 nm)

Emission 1: BODIPY TMR FP P-pol 595 (Wavelength 595 nm, bandwidth 60 nm)

Emission 2: BODIPY TMR FP S-pol 595 (Wavelength 595 nm, bandwidth 60 nm)

Plates were read at a measurement height of 7.4 mm with a g-factor of 1. Each well received 30 flashes per measurement. Anisotropy values were obtained by applying the following formula to the blank corrected P-values and S-values:

Anisotropy (r) = 1000*(S-G*P)/(S + 2*G*P)

The mean values and standard deviations for each triplicate were plotted using OriginPro 9.1, with protein concentration plotted along the x axis (logarithmic scale). The values were fitted to a logistic curve , fixing the maxima and minima of the curves if necessary

5.2.21   Testing of PEX5 mutants for reactivity with chloroacetamide probes

Each *At*-PEX5C mutant was buffer-exchanged into PBS using Amicon® Ultra-15 Centrifugal Filter Units (molecular weight cut off 30 kDa). It was then concentrated to 50 µM before combining with 100 µM of probe. The mixture was incubated at 4°C for between 1 and 18 hours. The mixture was analysed for labelling by mass spectrometry.

5.2.22   Testing HaloTag probes for dual labelling of HaloTag proteins and PEX5 mutant

HaloTag protein (50 µL, 50 µM in PBS) was combined with probes containing the HaloTag ligand motif (50 µL, 100 µM in PBS) and incubated a 4°C for 18 hours. After confirmation of successful HaloTag labelling by mass spectrometry, excess peptide was removed using Amicon Ultra-0.5 Centrifugal Filter Units (molecular weight cut off 10 kDa). The protein was concentrated back its original volume (50 µL) and *At*-PEX5C F606C mutant was added and incubated for 18 hours at 4°C. A sample of the reaction mixture was taken before and after incubation and analysed by SDS-PAGE.

5.2.23   Western Blotting

Samples were first run on an SDS-Page gel (see: 5.2.15) before being placed in transfer buffer to equilibrate for 5 minutes. The proteins on the gel were transferred onto a nitrocellulose membrane using the Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad) for 30 minutes at 15V according to the manufacturer's instructions. Membranes were blocked for 1 hour in 5% skimmed milk in TBS-T and incubated with primary antibody in blocking buffer for 1 hour. Membranes were then washed in TBS-T (3 x % minutes) and incubated with secondary antibody in blocking buffer for 1 hour. The membrane was washed with PBS-T (3 x 5 minutes). Membranes were then visualised using the ChemiDocMP (Bio-Rad) after incubation with Clarity ECL substrate (Bio-Rad).

5.3 Experimental for Chemical Synthesis

5.3.1 Synthesis of SNAP-Tag substrate motifs

4-(aminomethyl)-benzyl alcohol (1) (Hoffer et al., 2018)
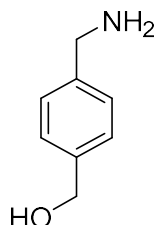
Lithium aluminium hydride (1.28g, 33.85mmol) was added to anhydrous THF(50 mL) under N_2 at 0°C and stirred for 15 minutes to give a grey suspension. A solution of methyl-4-cyanobenzoate in anhydrous THF (30 mL) was added to the flask dropwise over a period of 10 minutes. The reaction was allowed to warm to room temperature and stirred for 30 minutes. The reaction mixture was then heated to 50°C and allowed to react for 18 hours. After this time, the reaction was quenched through gradual addition of a homogenous mixture of $NaSO_4.(H_2O)_{10}$ and Celite 577 (2:1, w/w) until no further effervescence was observed upon addition. Water (3 mL) was added dropwise, and the mixture stirred for 30 minutes. The reaction mixture was filtered to remove solids and washed with DCM. The clear filtrate was concentrated under reduced pressure to give a white solid product which required no further purification (850mg, 6.2mmol, 100%).

^{1}H-NMR (400 MHz, CDCl3) δ 3.89 (2H, s, -CH2-N), 4.71 (2H, s, -CH2-O), 7.28-7.38 (4H, m, Ar H) ppm.

^{13}C-NMR (100.6 MHz, CDCl3) δ 46.1 (-N-CH2), 64.8 (HO-CH2), 127.3 (CH), 127.28 (CH), 139.85 (Ar C), 142.27 (Ar C) ppm.

Trifluoro-N-[[4(hydroxymethyl)phenyl]methyl]acetamide (2) (Keppler et al., 2003)
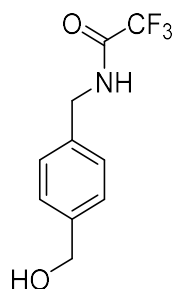
4-(aminomethyl)-benzyl alcohol (1) (0.5g, 3.64 mmol) was dissolved in MeOH (5 mL) and stirred under N_2. Triethylamine (711 μL, 5.1 mmol) and trifluoroacetic acid ethyl ester (433 μL, 3.64 mmol) were added and the reaction mixture stirred for 24 hours at room temperature under N_2. The reaction mixture was partitioned between EtOAc (30 mL) and brine (20 mL) and the aqueous layer extracted with further EtOAc (2 x 30 mL) and the combined organic layers were dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a crude yellow solid. The crude solid was pre-absorbed on to silica from MeOH and purified by flash chromatography in $CHCl_3$/MeOH (20:1). to yield the product as a pale yellow solid (632 mg, 2.71 mmol, 74%).

^{1}H-NMR (500 MHz, DMSO-d6) δ 4.37 (2H, s CH_2-N), 4.48 (2H, d, J = 6.0 Hz, -CH_2-O), 5.16 (1H, s, -OH), 7.23 (2H, d, J = 8.0 Hz, Ar H), 7.30 (2H, d, J = 8.0 Hz, Ar H), 9.98 (1H, br s, -NH-).

^{13}C-NMR (125.8 MHz, DMSO-d6) δ 42.9 (-N-CH_2), 63.1 (-CH_2-O), 116.5 (q, J= 288.1 Hz CF_3), 127.1 (Ar CH), 127.7 (Ar CH), 136.3 (Ar C), 142.1 (Ar C), 156.8 (q, J= 36.2 Hz, C=O) ppm.

HRMS:

$C_{10}H_{10}F_3NO_2$ requires $[M+Na]^+$ 256.0556, found: $[M+Na]^+$ 256.0551

N-[[4-[[(2-amino-6-chloro-4-pyrimidinyl)oxy]methyl]phenyl]methyl]-2,2,2-trifluoroacetamide (3) (Srikun et al., 2010)
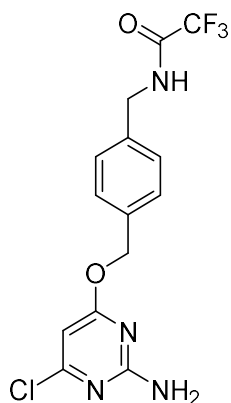
60% NaH in mineral oil (43.2 mg, 1.08 mmol) was suspended in dry DMF (1 mL). Trifluoro-N-[[4(hydroxymethyl)phenyl]methyl]acetamide (2) (120 mg, 0.51 mmol) was dissolved in dry DMF (3 mL) and added dropwise to the suspension. The flask was stirred under N_2 for 15 minutes. 2-amino-4,6-dichloropyrimidine was dissolved in dry DMF (3 mL) and added to the reaction mixture. The reaction was heated to 90°C and allowed to react for 18 hours. The reaction was cooled to room temperature and water (1 mL) was added and the mixture stirred for 30 minutes. The reaction mixture was poured into 0.5 M HCl (50 mL) and partitioned between EtOAc (3 x 25 mL) and brine (30 mL). The combined organic phases were dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude solid was preabsorbed on the silica from EtOAc and purified by flash column chromatography using a gradient of 40-50% EtOAc in cyclohexane to yield the title compound as a white solid.(108 mg, 0.30 mmol, 59%)

^{1}H-NMR (400 MHz, DMSO-d6) δ 4.39 (2H, d, J = 6.0 Hz, $-CH_2-N$), 5.30 (2H, s, -CH_2-O), 6.14 (1H, s, pyrimidine Ar H), 7.11 (2H, br s, pyrimidine Ar-NH_2), 7.29 (2H, d, J = 8.1 Hz, Ar H), 7.43 (2H, d, J = 8.1 Hz, Ar H), 10.00 (1H, t, J = 6.0 Hz, -NH-).

^{13}C-NMR (101 MHz, DMSO-d6) δ 42.8 ($-CH_2-N$), 67.6 ($-CH_2-O-$), 94.9 (pyrimidine Ar C), 116.5 (q, J=288.2 Hz, CF_3),128.0 (Ar C), 129.0 (Ar C), 135.8 (Ar C), 137.9 (Ar C),156.8 (q, J = 36.2 Hz, -C=O), 160.5 (pyrimidine Ar C-Cl), 163.3 (pyrimidine Ar C), 170.8 (pyrimidine Ar C)

HRMS:

$C_{12}H_{12}ClN_4O$ requires [M+H]$^+$ 361.0674, found [M+H]$^+$ 361.0670

4-[[4-(aminomethyl)phenyl]methoxy]-6-chloro-2-pyrimidinamine (4) (Srikun et al., 2010)
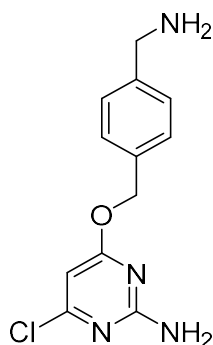
N-[[4-[[(2-amino-6-chloro-4-pyrimidinyl)oxy]methyl]phenyl]methyl]-2,2,2-trifluoroacetamide (3) was dissolved in ethanol (1 mL) and 33% $MeNH_2$ in EtOH (2 mL) added. The reaction was stirred under N_2 at room temperature and the progress of the reaction monitored using LC-MS. After 8 hours, a further 1 mL of 33% $MeNH_2$ in EtOH was added and the reaction continued to be stirred under N_2 overnight. After 18 hours, the reaction mixture was concentrated *in vacuo* to yield a crude yellow solid (**4**). which was used without further purification (125mg, 0.47 mmol)

^{1}H-NMR (400 MHz, MeOD) δ 4.02 (2H, s, $-CH_2-N$), 5.30 (2H, s, $-CH_2-O$), 6.14 (1H, s, pyrimidine Ar H), 7.10 (2H, br s, pyrimidine Ar $-NH_2$), 7.36 (2H, d, J = 8.0 Hz, Ar H), 7.42 (2H, d, J = 8.0 Hz, Ar H), 9.37 (2H, br s, $-CH_2-NH_2$).

^{13}C-NMR (100.1 MHz, DMSO-d$_6$) δ 45.1 ($-CH_2-N$), 67.8 ($-CH_2-O-$), 94.9 (pyrimidine Ar C), 127.9 (Ar C), 128.8 (Ar C), 135.0 (Ar C), 142.6 (Ar C), 160.5 (pyrimidine Ar C-Cl), 163.3 (pyrimidine Ar $C-NH_2$), 170.8 (pyrimidine Ar C-O-).

HRMS: $C_{12}H_{13}ClN_4O$ requires $[M+H]^+$ 265.0851, found $[M+H]^+$ 265.0847

5-[[[4-[[(2-Amino-6-chloro-4-pyrimidinyl)oxy]methyl]phenyl]methyl]amino]-5-oxopentanoic acid (5)
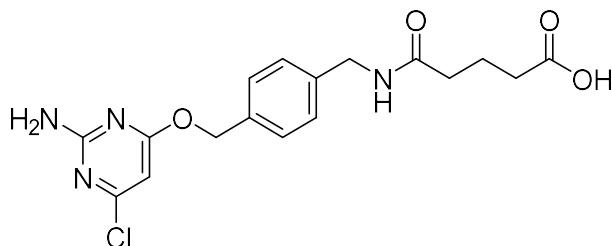
Crude 4-[[4-(aminomethyl)phenyl]methoxy]-6-chloro-2-pyrimidinamine (4) (150mg) was combined with glutaric anhydride (71.1mg, 0.62mmol) and DMAP (127.8mg, 1.14mmol). The components were dissolved in DMF (5 mL) before the addition of TEA (159 μL, 1.14mmol) to initiate the reaction. The reaction was stirred under N_2 for 18 hours. The reaction mixture was concentrated under reduced pressure. The crude product in DMF (1 mL) was purified by reverse phase chromatography on a RediSep Gold C18 column (Teldyne Isco) using a gradient of 10 - 90% H_2O/MeCN. To yield the product (130 mg, 0.34 mmol, 65%) as a white solid (5).

^{1}H-NMR (500 MHz, DMSO-d6) δ 1.74 (2H, quin, J= 7.4 Hz, -CH2-C$\underline{H}$2-CH2-), 2.17 (2H, t, J= 7.4 Hz, -C$\underline{H}$2-CH2-CH2-COOH), 2.22 (2H, t, J= 7.4 Hz, -CH2-CH2-C$\underline{H}$2-COOH), 4.26 (2H, d, J= 5.0 Hz, -C$\underline{H}$2-NH-), 5.29 (2H, s, -CH2-O -), 6.14 (1H, s, pyrimdine ArH), 7.10 (1H, br s, pyrimidine Ar-NH2), 7.25 (2H, d, J = 8.0 Hz, Ar H), 7.38 (2H, d, J = 8.0 Hz, Ar H), 8.11 (2H, s, -NH2), 8.32 (1H, t, J = 5.0 Hz, -NH-C=O) ppm

^{13}C-NMR (100.6 MHz, DMSO-d6) δ 21.3 (-$\underline{C}$H2-CH2-COOH), 33.7 (-$\underline{C}$H2-COOH), 34.9 (-$\underline{C}$H2-C=O-NH-), 42.3 (-CH2-N), 67.7 (-CH2-O-), 94.9 (pyrimidine Ar C), 127.7 (Ar C), 128.8 (Ar C), 135.1 (Ar C), 140.2 (Ar C), 160.5 (pyrimidine Ar C-Cl), 163.3 (pyrimidine Ar C-O-), 170.8 (carbonyl C), 172.1 (carbonyl C), 174.1 (pyrimidine Ar C-NH2).

HRMS: $C_{17}H_{19}ClN_4O_4$ requires [M+H]$^+$ 379.1168, found [M+H]$^+$ 379.1163
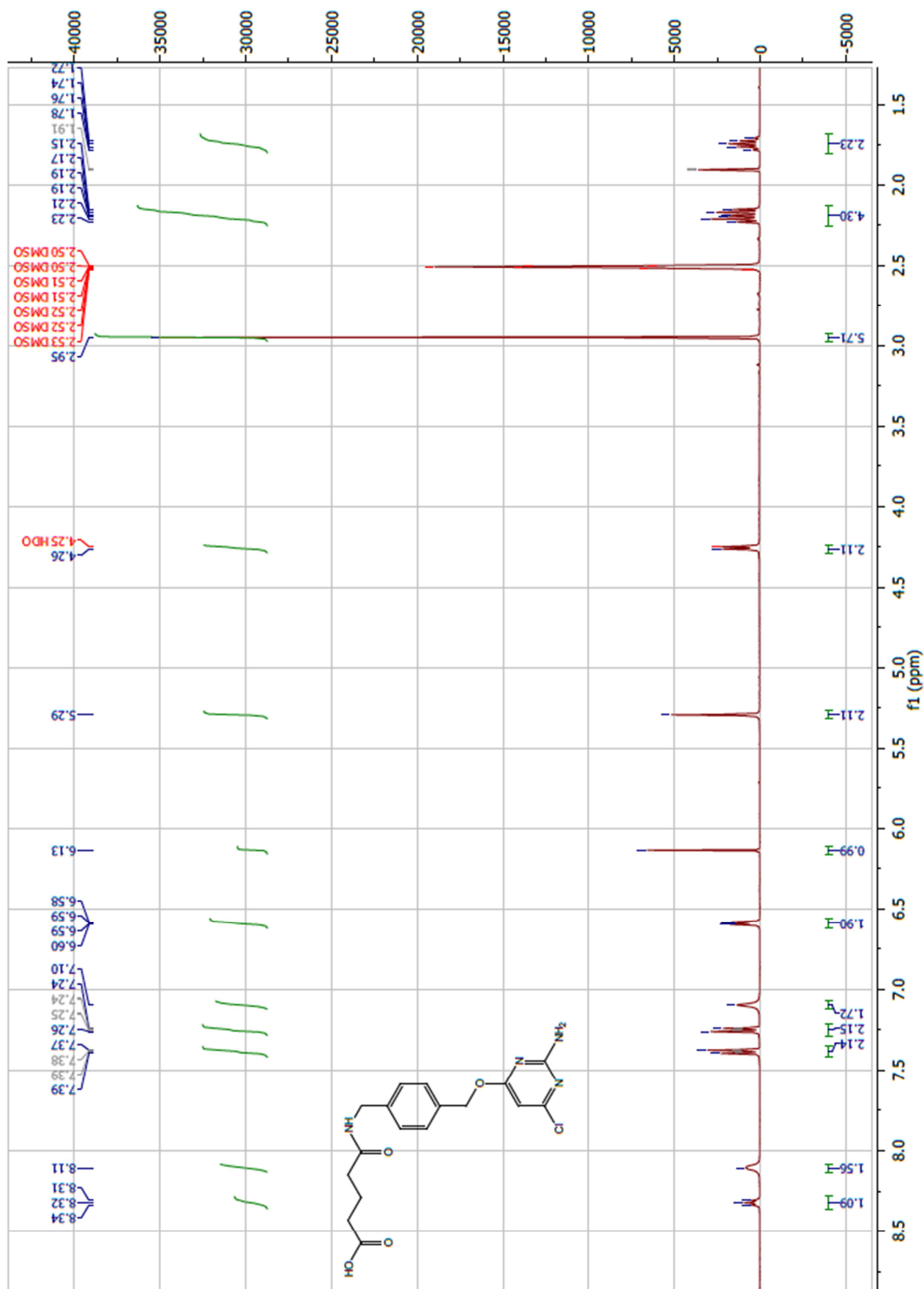
Spectrum 1 ¹H NMR Spectrum for compound 5
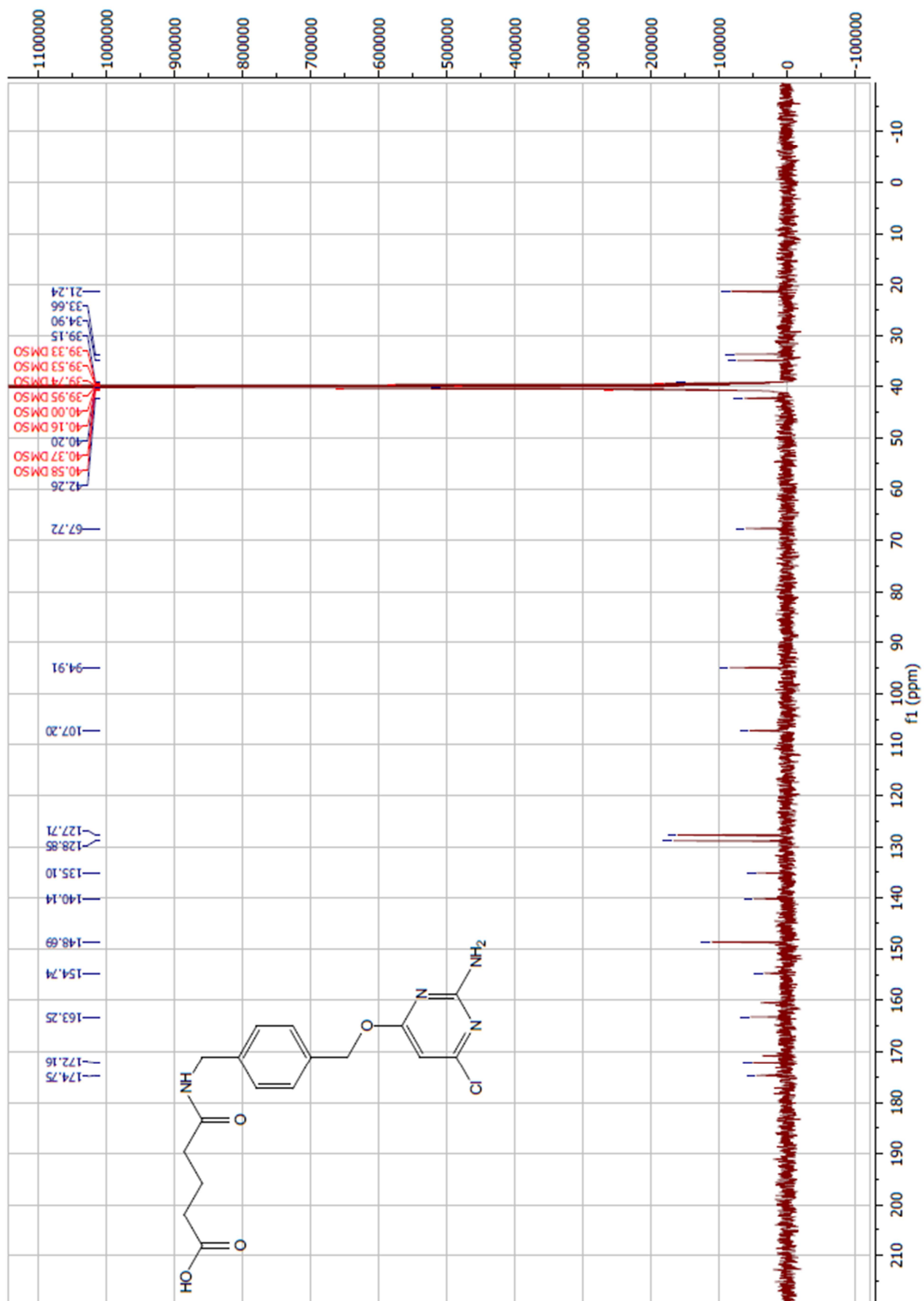
Spectrum 2 13C NMR spectrum for compound 5

5.3.2 Synthesis of HaloTag substrate motifs

tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate (6) (Singh et al., 2013)
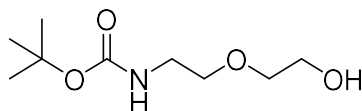
2-(2-aminoethoxy)ethanol (10 g, 99.7 mmol) was dissolved in dry EtOH at 0°C. Di-tert-butyl dicarbonate (21.8 g, 99.7 mmol) was added. The reaction was warmed to room temperature and stirred for 2 hours. The reaction mixture was concentrated under reduced pressure and then partitioned between DCM (3 × 100 mL) and water (100mL). Combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo* to yield the product as a colourless oil which was used without further purification. (18.9 g, 92.2 mmol, 92%)

^{1}H-NMR (600 MHz, CDCl$_3$) δ 1.42 (s, 9H, ^{t}Bu), 3.39-3.29 (m, 2H, N-CH$_2$), 3.67-3.45 (m, 4H, O-CH$_2$-CH$_2$-OH), 3.81-3.71 (m, 2H, NH-CH$_2$-C$\underline{H}_2$O), 5.21 (br, 1H, NH) ppm

^{13}C-NMR (125.9 MHz, CDCl$_3$) δ 28.39 (^{t}Bu), 40.35 (**C**-^{t}Bu), 61.60 (C-OH), 70.30(O-CH$_2$), 72.26 (O-CH$_2$), 79.35 (C-NH-), 156.19 (C=O) ppm

HRMS $C_9H_{19}NO_4$ [M+Na]$^+$ requires 228.1206, found [M+Na]$^+$ 228.1206

tert-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (7) (Singh et al., 2013)
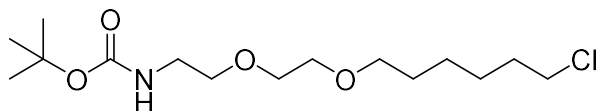
tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate (6) (2.0 g, 9.74 mmol) was dissolved in dry THF (14 mL) and dry DMF (7 mL) at 0°C. NaH (468mg, 60% in mineral oil, 11.69 mmol) was added and the mixture stirred for 30 minutes. 6-chloro-1-iodo-hexane (3.4 g, 13.64 mmol) was added to the reaction vessel and the reaction mixture stirred overnight at room temperature. The reaction was quenched with saturated NH_4Cl and extracted with EtOAc (3 x 30 mL) and washed with H_2O and brine. The combined organic layers were dried over $MgSO_4$ and under reduced pressure. The product was purified by flash column chromatography using a gradient of EtOAc and Hexane (20:80 → 50:50). (883mg, 2.73 mmol, 28%)

^{1}H-NMR (600 MHz, CDCl$_3$) δ 1.34-1.54 (m, 13H, ^{t}Bu + (CH$_2$)$_2$), 1.57-1.67 (m, 2H, CH$_2$), 1.73-1.85 (m, 2H, CH$_2$), 3.33 (quintet, J=5.5Hz, 2H CH$_2$NH), 3.48 (t, J=6.7, 2H, CH$_2$Cl), 3.60-3.52 (m, 6H, CH$_2$O), 3.61-3.63 (m, 2H, NHCH$_2$CH$_2$O), 5.04 (s, 1H, NH) ppm

^{13}C-NMR (125.9 MHz, CDCl$_3$) δ 25.3 (CH$_2$), 26.6 (CH$_2$), 28.4 (^{t}Bu) 29.4 (CH$_2$), 32.5 (CH$_2$), 40.3(CH$_2$-NH), 44.9 (CH$_2$), 70.0 (O-CH$_2$-C$_5$H$_{10}$-Cl), 70.1 (O-CH$_2$), 70.2 (O-CH$_2$), 71.2 (O-CH$_2$), 78.9 (C-^{t}Bu), 155.9 (C=O) ppm

HRMS C$_{15}$H$_{30}$ClNO$_4$ [M+Na]$^+$ requires 346.1756, found [M+Na]$^+$ 346.1753

2-(2-((6-chlorohexyl)oxy)ethoxy)ethanamine (8) (Singh et al., 2013)
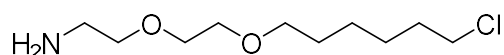
tert-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (7) (1.1g, 3.40 mmol) was dissolved in dry CH_2Cl_2 at 0°C and trifluoroacetic acid (3.65 mL, 47.31 mmol) added portionwise. The reaction was warmed to room temperature and allowed to react for 2 hours. The solvent was removed under reduced pressure. The residue was dissolved in MeOH and purifed using a Bond Elut-SCX, 2gm 6mL column (Agilient Techologies). The column was equilibrated with 20 column volumes of acetic acid (10% in methanol). The crude sample was disolved in methanol and loaded onto the column which was then washed with 5 column volumes of methanol. The sample was eluted by addition of 3 column volumes ammonia (2M in methanol). The eluates were immediately concentrated under reduced pressure to remove ammonia.

^{1}H-NMR (600 MHz, MeOD) δ 1.37-1.55 (4H, m, 2 x CH_2) 1.65 (2H, quintet, J= 6.8 Hz, CH_2) 1.82 (4H, m 2 x CH_2) 2.91 (2H, t, J= 5.2 Hz) 3.51 (t, J= 6.7 Hz, CH_2Cl), 2H, 3.55-3.59 (m, 4H, 2 x CH_2O) 3.61-3.67 (m, 4H, 2 x CH_2O)

^{13}C-NMR (125.9 MHz, MeOD) δ 25.1 ($CH_2\underline{C}H_2CH_2O$), 26.33 ($\underline{C}H_2CH_2CH_2O$), 29.1 ($CH-NH_2$), 32.4 ($\underline{C}H_2CH_2CH_2Cl$), 40.4 ($\underline{C}H_2CH_2Cl$), 44.3 ($CH_2Cl$), 69.8 (O-$CH_2\underline{C}H_2$-O), 69.9 ($NH_2CH_2\underline{C}H_2O$), 70.7 (O-$CH_2$), 70.8 (O-$\underline{C}H_2$-$C_5H_{10}$-Cl)

HRMS $C_{10}H_{22}NO_2Cl$ [M+H]$^+$ requires 224.1412, found [M+H]$^+$224.1414

2-(2-((6-chlorohexyl)oxy)ethoxy)ethanamino-5-oxopentanoic acid (9)
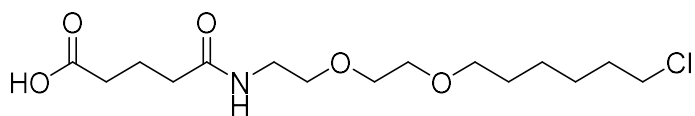
2-(2-((6-chlorohexyl)oxy)ethoxy)ethanamine (8) (197 mg, 0.84 mmol) was dissolved in dry DMF. Glutaric anhydride (96 mg, 0.84 mmol) and DMAP (9 mg, 0.08 mmol) were dissolved in dry DMF and added to the reaction vessel. DIPEA (295 µL, 1.69 mmol) was added to the reaction. The reaction mixture was stirred at room temperature for 18 hours. The reaction mixture was concentrated under reduced pressure and the crude product purified by reverse phase chromatography on a RediSep Gold C18 column with elution using a gradient of MeCN:H_2O (10:90 → 90:10). Fractions were analysed by LC-MS and appropriate fractions were combined and concentrated under reduced pressure. (63 mg, 0.19 mmol, 22%)

^{1}H-NMR (600 MHz, CDCl$_3$) δ 1.26-1.34 (m, 2H, CH$_2$), 1.35-1.44 (m, 2H, CH$_2$), 1.54 (quintet, 6.8 Hz, 2H, CH$_2$), 1.71 (quintet, 6.8 Hz, 2H, CH$_2$), 1.91 (quintet, 6.8 Hz, 2H, CH$_2$NH), 2.34 (m, 4H, 2 x CH$_2$C=O), 3.41 (t, J= 6.8 Hz, CH$_2$Cl), 3.47 (t, J= 6.8 Hz, CH$_2$O), 3.51-3.57 (m, 6H, 3 x CH$_2$O), 6.90 (s, 1H, OH), 7.09 (s, 1H, NH)

^{13}C-NMR (125.9 MHz, CDCl$_3$) δ 20.9 (CH$_2$CH$_2$COOH), 25.3 (CH$_2$CH$_2$CH$_2$O), 26.6 (CH$_2$CH$_2$CH$_2$O), 29.3 (CH$_2$CH$_2$CH$_2$Cl), 32.5 (CH$_2$CH$_2$CH$_2$Cl), 33.0 (CH$_2$-C=O-NH), 34.9 (CH$_2$COOH), 39.6 (CH$_2$-NH), 45.1 (CH$_2$Cl), 69.4 (CH$_2$-O), 69.9 (O-CH$_2$-C$_5$H$_{10}$-Cl), 70.1 (O-CH$_2$), 71.3 (NH-CH$_2$-CH$_2$-O), 173.9 (HN-C=O), 176.3 (COOH)

HRMS C$_{17}$H$_{19}$NO$_5$Cl [M+Na]$^+$ requires 360.1548, found [M+Na]$^+$ 360.1553

Spectrum 3 ^{1}H NMR Spectrum for compound 9

165
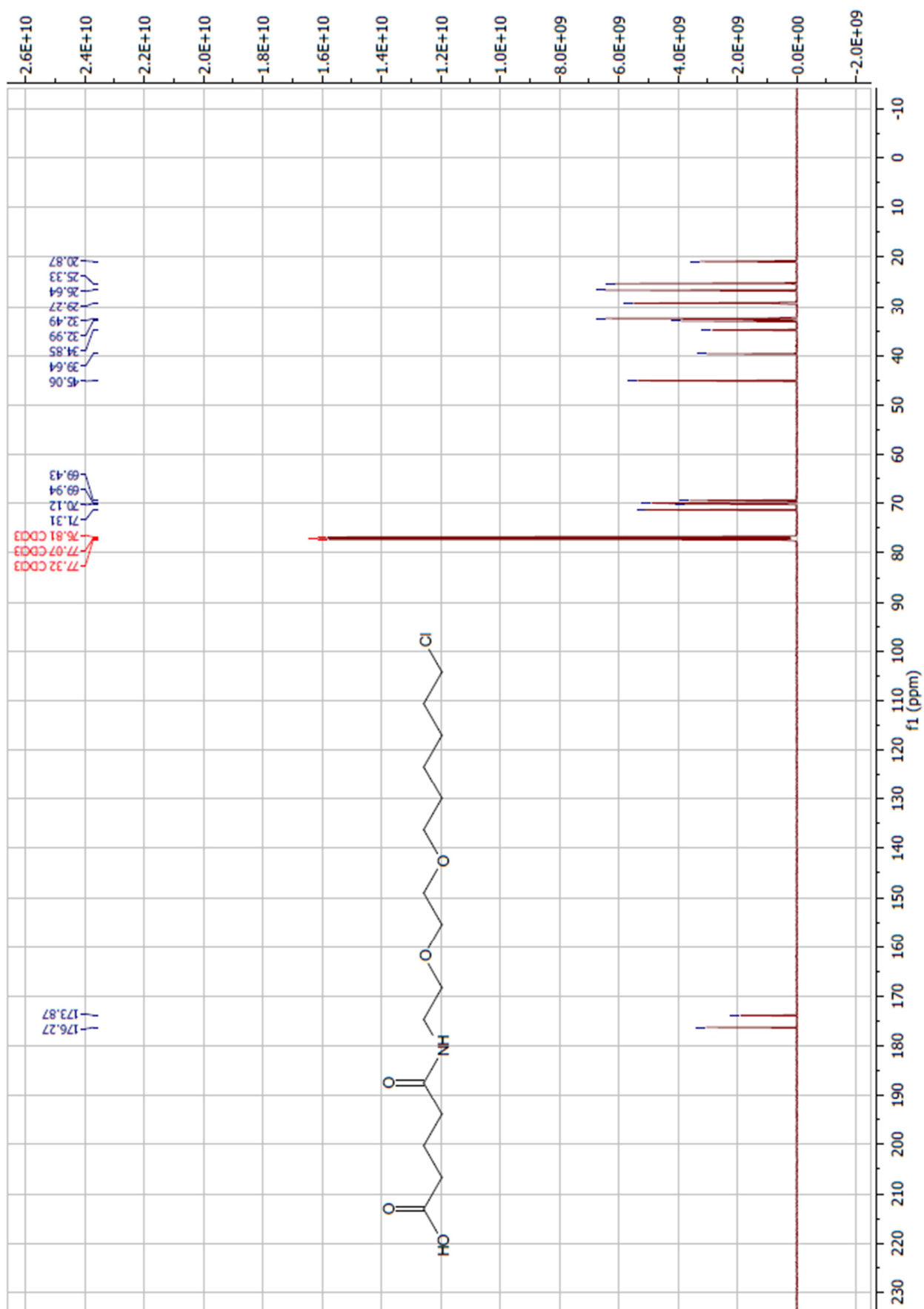
Spectrum 4 ^{13}C NMR spectrum for compound 9

5.3.3 Peptide Synthesis

5.3.3.1 Coupling of Fmoc-protected amino acids (Method A)

Peptides were prepared on 2 Cl-Trt resin preloaded with the C-terminal amino acid (0.154mmol, 0.77 mmol/g) using the following amino acid building blocks (Novabiochem):
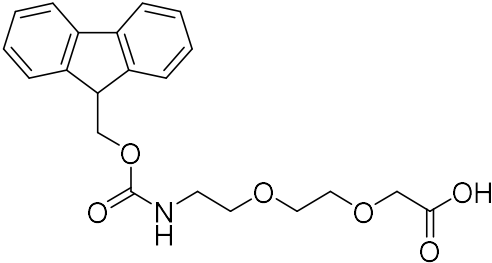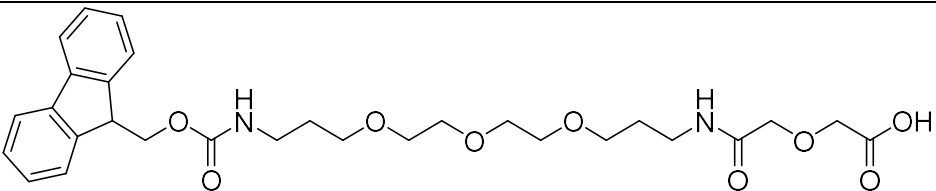
Lysine	K	Fmoc-Lys-OH
Arginine	R	Fmoc-Arg(Pbf)-OH
Leucine	L	Fmoc-Leu-OH
Serine	S	Fmoc-Ser(tBu)-OH
Glutamine	Q	Fmoc-Gln(Trt)-OH
Tyrosine	Y	Fmoc-Tyr(tBu)-OH

The resin was swollen in DMF (2 mL) for 10 min then washed with DMF (5 x 2 mL, 2 mins). Fmoc amino acid (0.77 mmol) and Oxyma Pure (0.77 mmol) were dissolved in the DMF (1 mL) and DIC (0.77 mmol) was added. This was transferred to the tube containing the resin. The mixture was gently agitated for 1 hour on a laboratory rotator and the resin was isolated by filtration and washed with DMF (5 x 2 mL, 2 min). The resin was treated with a solution of 20 % piperidine in DMF (3 x 2 mL, 5 min) followed by further washes in DMF (5 x 2 mL, 2 min). Couplings were repeated until the desired sequence had been assembled. The functionalised resin was washed with DMF (5 x 2 mL, 2 mins), DCM (2 x 2 mL) and dried overnight.

5.3.3.2 Coupling of PEG linkers (Method B)

To incorporate the PEG or PEG_2 linkers into peptides, Fmoc protected PEG amino acids were coupled using the same general procedure as for coupling amino acids detailed above. Fmoc-NH-(PEG)-COOH or Fmoc-NH-(PEG_2)-COOH (5 equivalents) and OxymaPure (5 equivalents) were dissolved in DMF (2 mL) and DIC (5 equivalents) added. The solution was added to DMF swelled resin with appropriate peptide attached and rotated overnight. The Fmoc group was deprotected with piperidine as before.

Table 5.1 Structures of PEG linkers

Fmoc-NH-(PEG)-COOH	
Fmoc-NH-(PEG)₂-COOH	

5.3.3.3 Lysine Dde Deprotection (Method C)

Dde-protected lysines were used in some of the probes to give the ability to incorporate the Pacific Blue fluorophore and AcCl reactive groups. When coupling groups to the lysine chain, the Dde group was removed by addition of 2% hydrazine in DMF (3 x 3 mL x 3 mins) followed by washes with DMF (5 x 2mL x 2 mins).

5.3.3.4 Pacific Blue Coupling (Method D)

To couple Pacific Blue to the lysine chain after Dde deprotection, 6,8-Difluoro-7-hydroxy-2-oxo-2H-chromene-3-carboxylic acid (2.5 equivalents) was combined with DIPEA (5 equivalents) and HCTU (2.5 equivalents) in DCM (2 mL). The reaction mixture was added to the peptide on the resin and incubated for 1 hour with end-over-end mixing. The resin was isolated by filtration and washed with DMF (5 x 2 mL, 2 min) before continuing with further steps.

5.3.3.5 Lissamine Coupling (Method E)

Lissamine Rhodamine B sulfonyl chloride (3 equivalents) was combined with DIPEA (6 equivalents) in DMF (3mL) and incubated with peptide on resin for 1 hour with end-over-end mixing. The resin was isolated by filtration and washed with DMF (5 x 2 mL, 2 min) before continuing with further steps.

5.3.3.6 Chloroacetyl Coupling (Method F)

To couple a chloroacetyl group to the lysine chain after Dde deprotection, chloroacetyl chloride (10 equivalents) was combined with DIPEA (20 equivalents) in DCM (2 mL). The reaction mixture was added to the peptide on the resin and incubated for 1 hour with end-over-end mixing. The resin was

isolated by filtration and washed with DMF (5 x 2 mL, 2 min) before continuing with further steps.

5.3.3.7 Acetyl end capping Coupling (Method G)

To add an acetyl group to the end of a peptide chain, or to a Dde deprotected lysine when making control peptides, acetic anhydride (10 equivalents) was combined with DIPEA (20 equivalents) in DCM (2 mL). The reaction mixture was added to the peptide on the resin and incubated for 1 hour with end-over-end mixing. The resin was isolated by filtration and washed with DMF (5 x 2 mL, 2 min) before continuing with further steps.

5.3.3.8 SNAP-Tag and HaloTag Substrate Coupling (Method H)

Either SNAP-Tag or HaloTag substrate compounds (3 equivalents) were dissolved in DMF (2 mL) with OxymaPure (3 equivalents) and DIC (3 equivalents). This mixture was attached to the peptide by overnight incubation with the Fmoc-deprotected peptide on resin. Resin was then washed with DMF (5 x 2mL x 2 mins) and DCM (2 x 2 mL x 2mins) and dried.

5.3.3.9 Cleavage of peptides from the resin (Method I)

Once synthesis of the peptide compound on the resin was complete, it was cleaved from the resin. A cleavage mix of TFA/H_2O/TIPS (95:2.5:2.5, 1 mL / 25 mg resin) was added to the resin and the mixture gently agitated for 1 hour. The resin beads were removed from the mixture by filtration and washed with TFA. Peptide products were precipitated with chilled Et_2O and the pelleted by centrifugation (4,000 x g, 5 mins). The Et_2O supernatant was removed and the process repeated. Precipitated peptide products were then dried under N_2, dissolved in H_2O/ 1,4-dioxane and freeze-dried overnight. Resulting products were purified by reverse phase chromatography on a RediSep Gold C18 column (Teldyne Isco) using a gradient of 10 - 90% H_2O/MeCN.

5.3.4 SNAP-YQSKL (10)
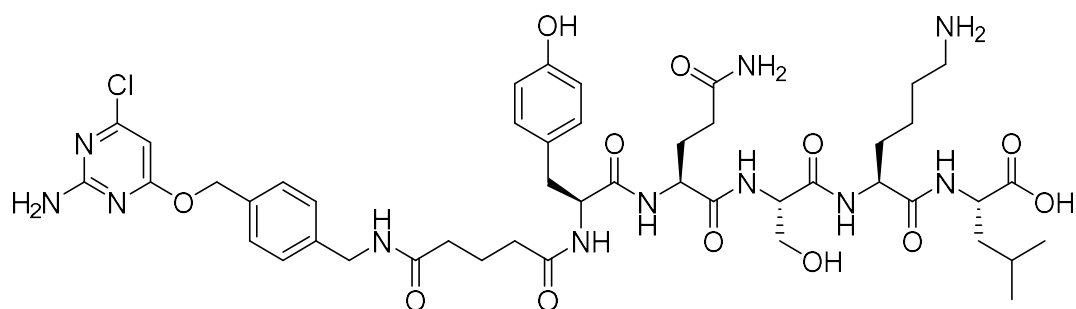
SNAP-Tag substrate (5) was synthesised as detailed (5.3.1) and YQSKL peptide synthesised on 50 mg pre-loaded Leu resin (NovaBiochem) using Method A. SNAP-Tag substrate was coupled using Method H and cleaved from the resin using Method I (Yield 6.5mg, 17.1%)

HRMS:

$C_{46}H_{64}ClN_{11}O_{12}$ requires $[M+H]^+$ 998.4497, found: $[M+H]^+$ 998.4529

5.3.5 SNAP-YQLKS (11)
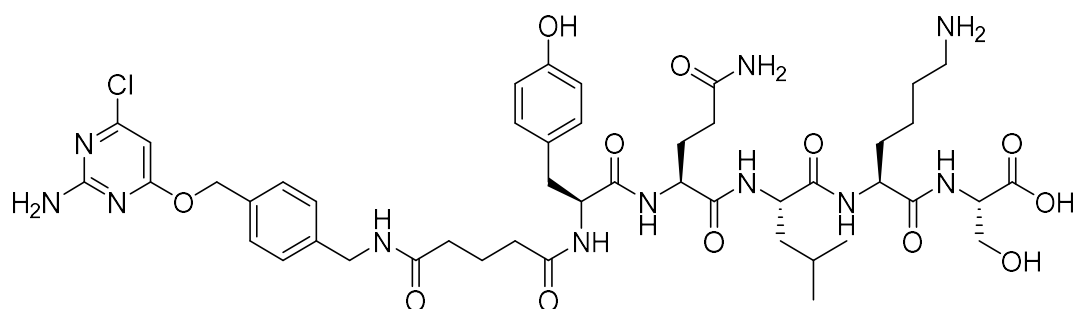
SNAP-Tag substrate (5) was synthesised as detailed (5.3.1) and YQLKS peptide synthesised on 50 mg pre-loaded Ser resin (NovaBiochem) using Method A. SNAP-Tag substrate was coupled using Method H and cleaved from the resin using Method I. (Yield 6.5mg, 11.3 %)

HRMS:

$C_{46}H_{64}ClN_{11}O_{12}$ requires [M+H]$^+$ 998.4497, found: [M+H]$^+$ 998.4504

5.3.6 SNAP-PacificBlue-PEG-YQSKL (12)

SNAP-Tag substrate (5) was synthesised as detailed (5.3.1) and YQSKL peptide synthesised on 25 mg pre-loaded Leu resin (NovaBiochem) using Method A. PEG linker was added using method B, lysine Dde residue was added using method A and SNAP-Tag substrate was coupled using Method H. Lysine Dde was then deprotected using method C and Pacific Blue added using Method D. The peptide was cleaved from the resin using Method I. (Yield: 0.6 mg, 2.1 %)

HRMS:

$C_{68}H_{89}ClF_2N_{14}O_{20}$ requires [M+H]$^+$ 1495.6107, found: [M+H]$^+$ 1495.6110

5.3.7 SNAP-PacificBlue-PEG-YQLKS (13)
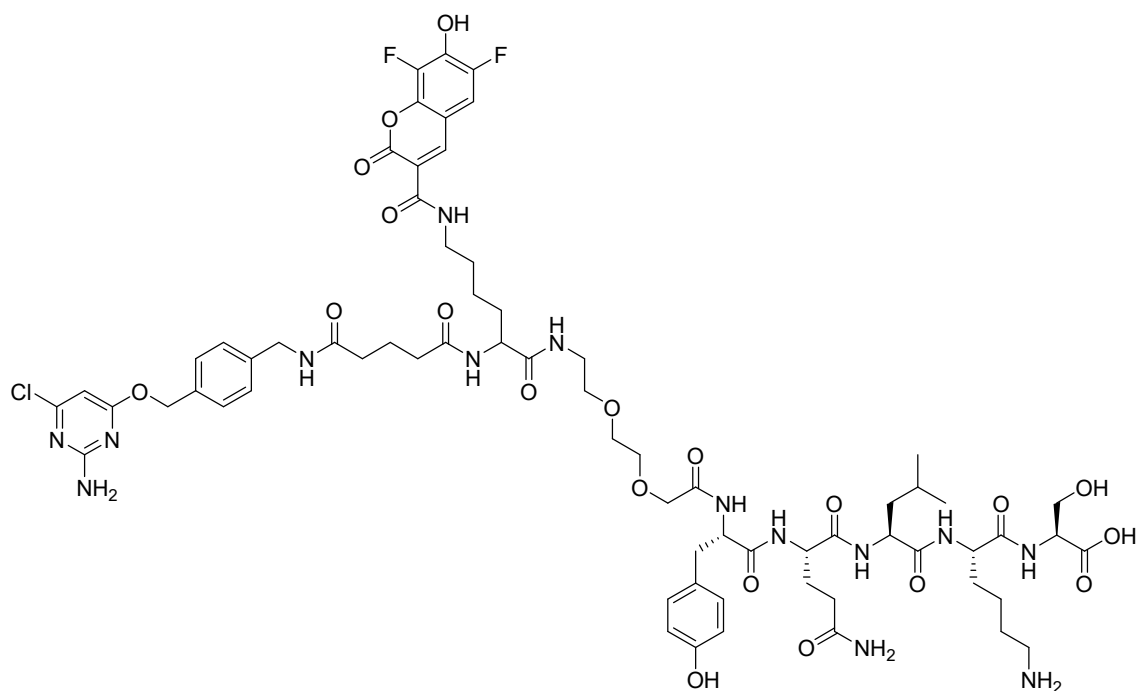
SNAP-Tag substrate (5) was synthesised as detailed (5.3.1) and YQLKS peptide synthesised on 25 mg pre-loaded Ser resin (NovaBiochem) using Method A. PEG linker was added using method B, lysine Dde residue was added using method A and SNAP-Tag substrate was coupled using Method H. Lysine Dde was then deprotected using method C and Pacific Blue added using Method D. The peptide was cleaved from the resin using Method I. (Yield: 0.8mg, 2.8%).

HRMS:

$C_{68}H_{89}ClF_2N_{14}O_{20}$ requires [M+H]$^+$ 1495.6107, found: [M+H]$^+$ 1495.6141

5.3.8 HaloTag-YQSKL (14)
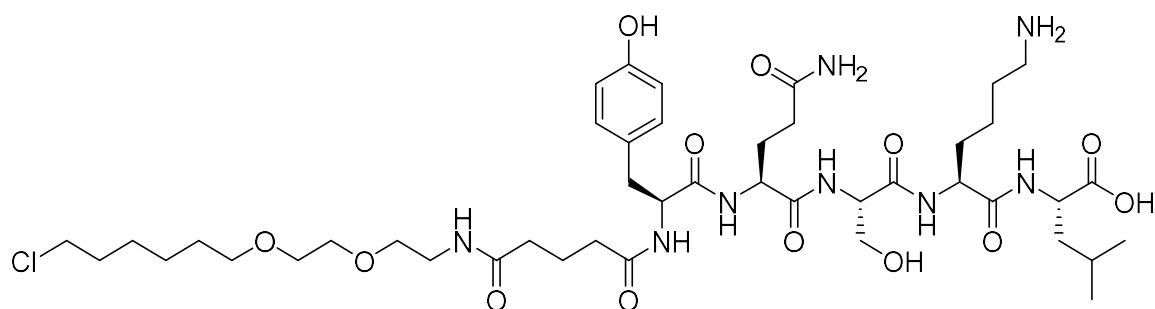
HaloTag substrate was synthesised as detailed (5.3.2) and YQSKL peptide synthesised on 50 mg pre-loaded Leu resin (NovaBiochem) using Method A. HaloTag substrate was coupled using Method H and cleaved from the resin using Method I (Yield 5.2mg, 14.3%)

HRMS:

$C_{44}H_{73}ClN_8O_1$ requires [M+H]$^+$ 957.5058, found: [M+H]$^+$ 957.5078

174

5.3.9 HaloTag-YQLKS (15)
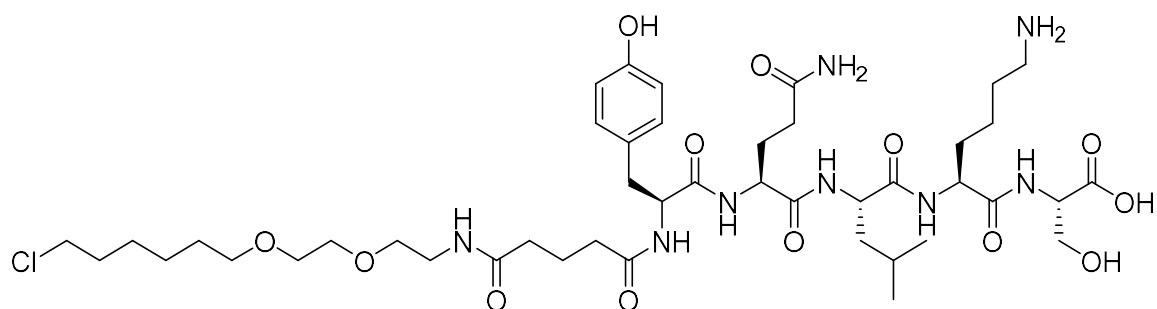
HaloTag substrate was synthesised as detailed (5.3.2) and YQLKS peptide synthesised on 50 mg pre-loaded Ser resin (NovaBiochem) using Method A. HaloTag substrate was coupled using Method H and cleaved from the resin using Method I (Yield 4.2mg, 11.5%)

HRMS:

$C_{44}H_{73}ClN_8O_1$ requires [M+H]$^+$ 957.5058, found: [M+H]$^+$ 957.5064

5.3.10 HaloTag-PacificBlue-PEG-YQSKL (16)
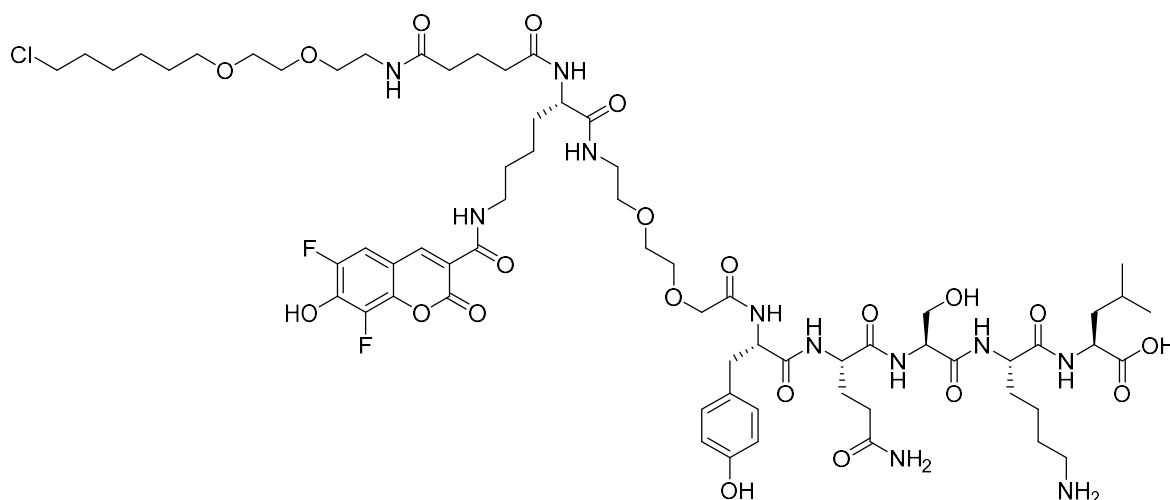
HaloTag substrate was synthesised as detailed (5.3.2) and YQSKL peptide synthesised on 25 mg pre-loaded Leu resin (NovaBiochem) using Method A. PEG linker was added using method B, lysine Dde residue was added using method A and SNAP-Tag substrate was coupled using Method H. Lysine Dde was then deprotected using method C and Pacific Blue added using Method D. The peptide was cleaved from the resin using Method I. (Yield: 0.9 mg, 2.1 %)

HRMS:

$C_{66}H_{98}ClF_2N_{11}O_{21}$ requires [M+H]$^+$ 1454.6668, found: [M+H]$^+$ 1454.6661
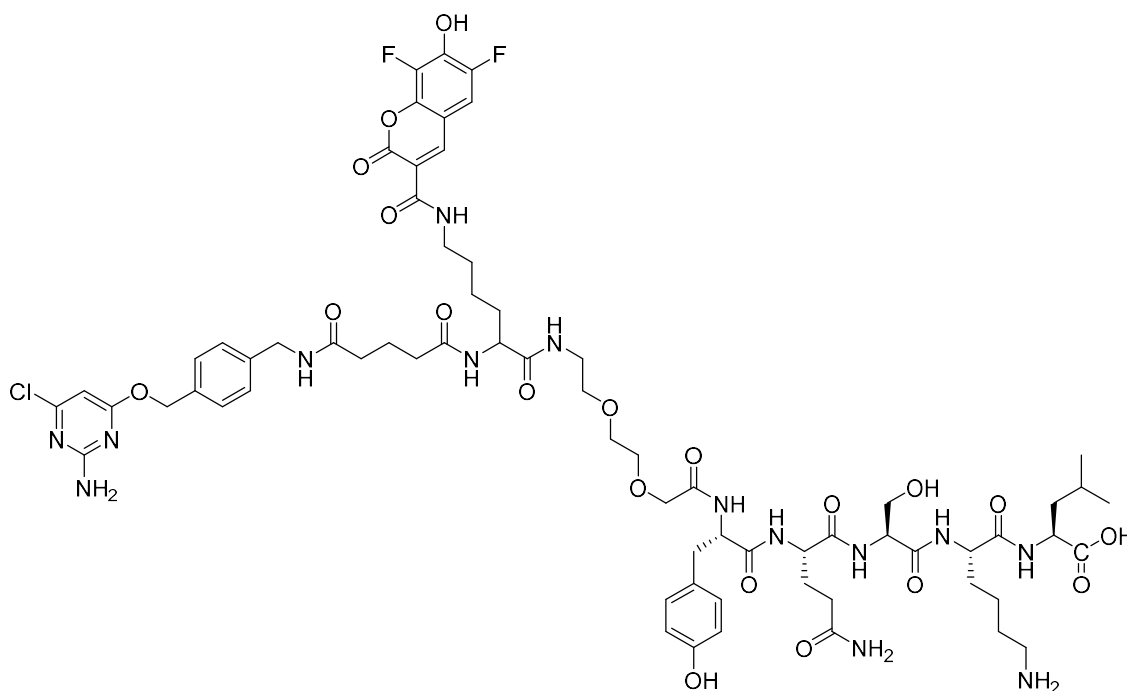
5.3.11 HaloTag-PacificBlue-PEG-YQLKS (17)
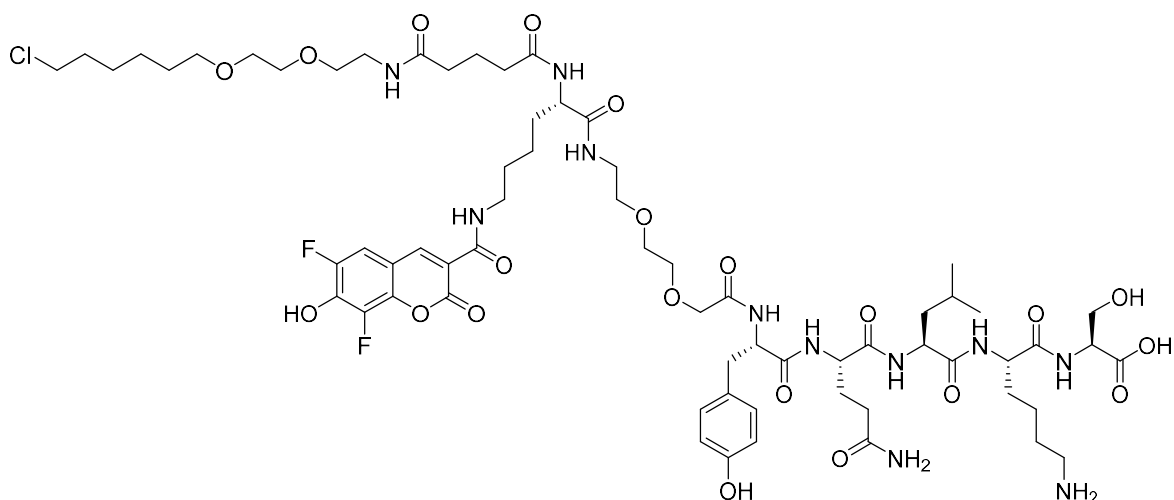
HaloTag substrate was synthesised as detailed (5.3.2) and YQLKS peptide synthesised on 25 mg pre-loaded Ser resin (NovaBiochem) using Method A. PEG linker was added using method B, lysine Dde residue was added using method A and SNAP-Tag substrate was coupled using Method H. Lysine Dde was then deprotected using method C and Pacific Blue added using Method D. The peptide was cleaved from the resin using Method I. (Yield: 1.3 mg, 2.4 %)

HRMS:

$C_{66}H_{98}ClF_2N_{11}O_{21}$ requires $[M+H]^+$ 1454.6668, found: $[M+H]^+$ 1454.6690

5.3.12 AcCl- No linker-YQSRL (18)
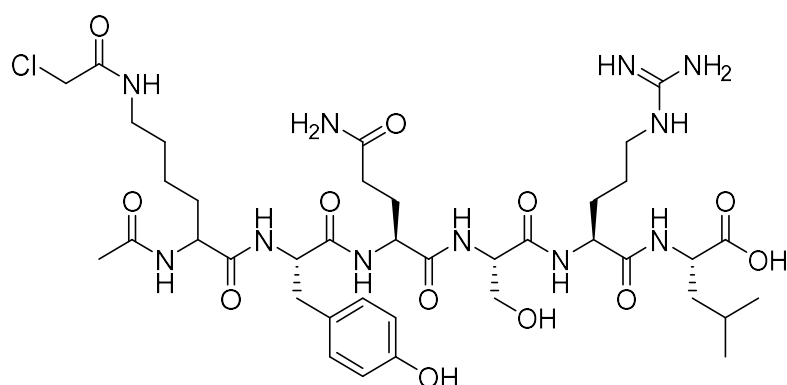
YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-Lysine Dde was coupled using Method A. After Fmoc deprotection, the N-terminus was capped with an acetyl group using Method G. Lysine Dde was then deprotected using Method C and chloroacetyl chloride coupled using Method F. The peptide was immediately cleaved from the resin and purified using method I.(Yield: 4.1mg, 11.8%)

HRMS:

$C_{39}H_{62}ClN_{11}O_{12}$ requires $[M+H]^+$ 912.4341, found: $[M+H]^+$ 912.4375

5.3.13 AcCl- PEG-YQSRL (19)

YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A. After Fmoc deprotection, the N-terminus was capped with an acetyl group using Method G. Lysine Dde was then deprotected using Method C and chloroacetyl chloride coupled using Method F. The peptide was immediately cleaved from the resin and purified using Method H.(Yield: 3.5mg, 8.7%)

HRMS:

$C_{45}H_{73}ClN_{12}O_{15}$ requires [M+H]+ 1057.5080, found: [M+H]+ 1057.5099

5.3.14   AcCl-PEG$_2$-YQSRL (20)
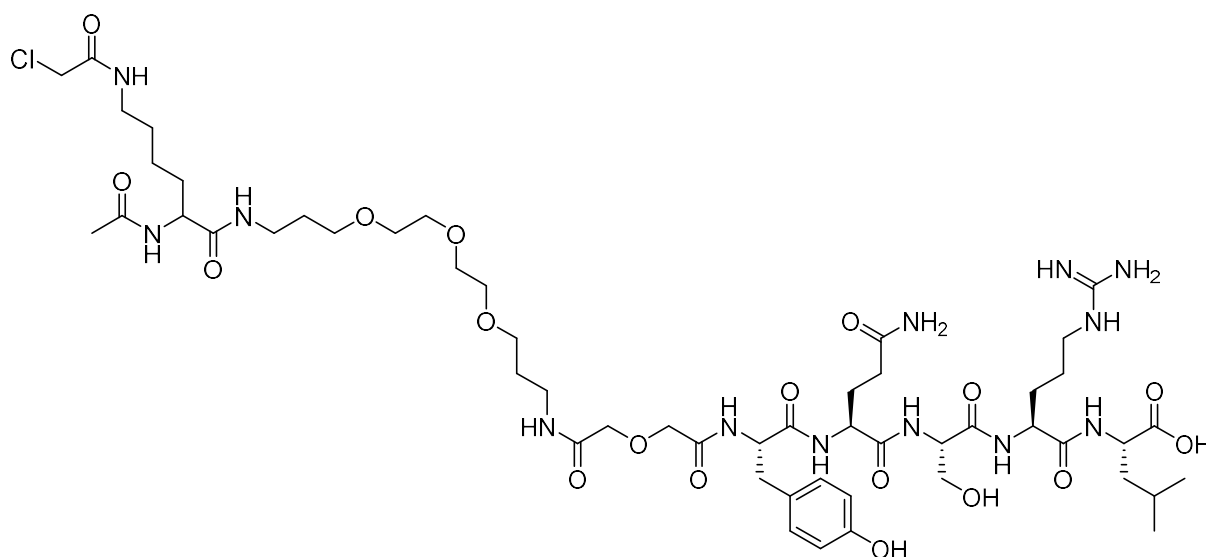
YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG$_2$-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A. After Fmoc deprotection, the N-terminus was capped with an acetyl group using Method G. Lysine Dde was then deprotected using Method C and chloroacetyl chloride coupled using Method F. The peptide was immediately cleaved from the resin and purified using Method I.(Yield: 2.8mg, 6.0%)

HRMS:

C$_{53}$H$_{88}$ClN$_{13}$O$_{18}$ requires [M+H]$^+$ 1230.6132, found: [M+H]$^+$ 1230.6186

5.3.15 AcCl-PEG-YQLRS (21)
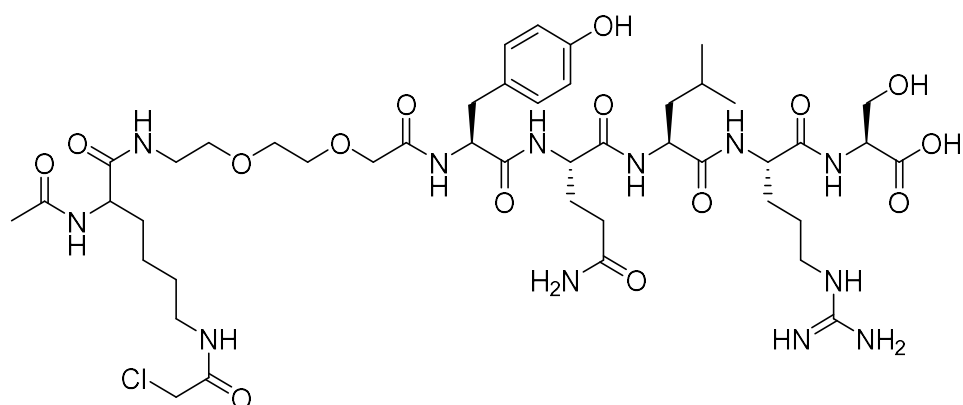
YQLRS peptide was synthesised on 50mg pre-loaded Ser resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A. After Fmoc deprotection, the N-terminus was capped with an acetyl group using Method G. Lysine Dde was then deprotected using Method C and chloroacetyl chloride coupled using Method F. The peptide was immediately cleaved from the resin and purified using method I.(Yield: 3.5mg, 8.7%)

HRMS:

$C_{45}H_{73}ClN_{12}O_{15}$ requires [M+H]$^+$ 1057.5080, found: [M+H]$^+$ 1057.5112

5.3.16   Unreactive-PEG-YQSRL (22)
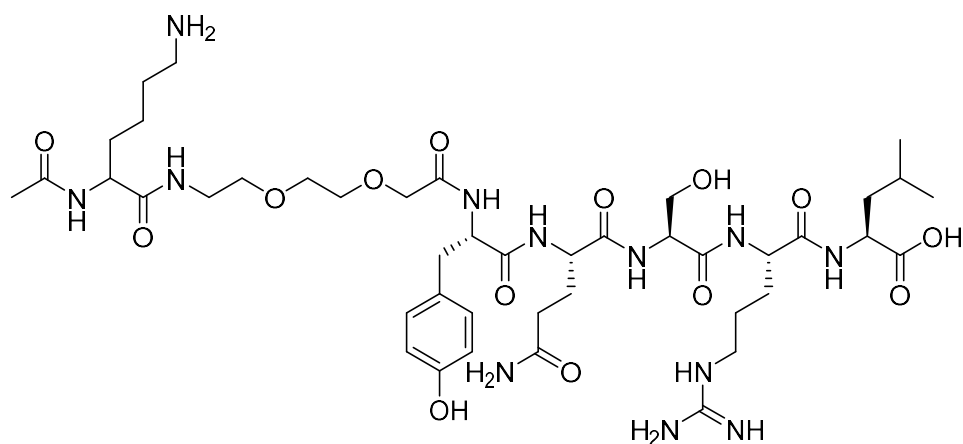
YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A. After Fmoc deprotection, the N-terminus was capped with an acetyl group using Method G. Lysine Dde was then deprotected using Method C. The peptide was immediately cleaved from the resin and purified using Method I.(Yield: 3.3 mg, 8.9%)

HRMS:

$C_{45}H_{73}ClN_{12}O_{15}$ requires $[M+H]^+$ 981.5364, found: $[M+H]^+$ 981.5400

5.3.17 HaloTag-PEG-AcCl-PEG-YQSRL (23)
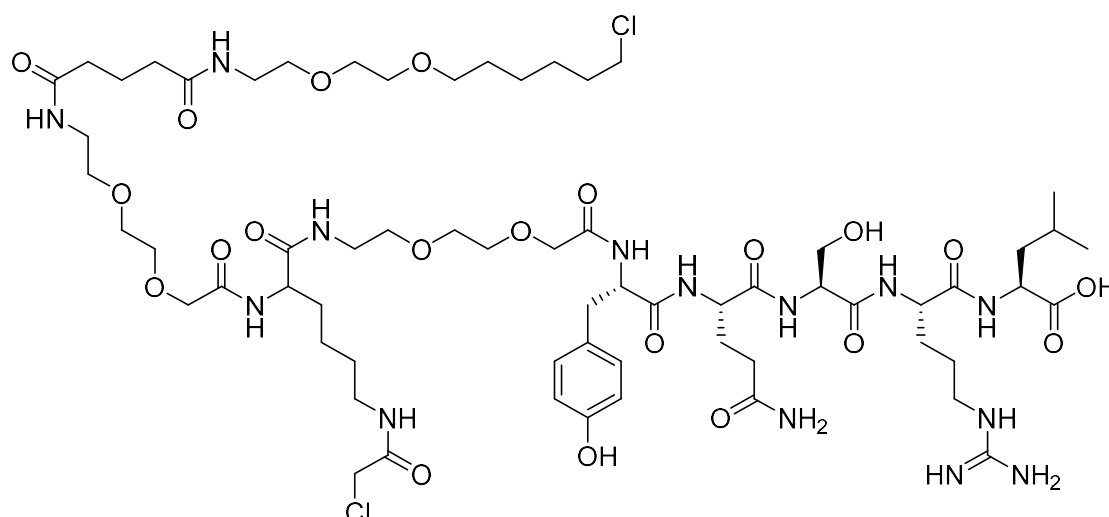
YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A.A second Fmoc-PEG-COOH was coupled using method B followed by coupling of the HaloTag substrate using Method H. The Lysine Dde was then deprotected using Method C and the chloroacetyl group coupled to the lysine side chain using Method F. The peptide was then immediately cleaved from the resin and purified using Method I. (Yield: 5.6mg, 10.0%)

HRMS:

$C_{64}H_{108}Cl_2N_{14}O_{21}$ requires $[M+H]^+$ 1479.7263, found: $[M+H]^+$ 1479.7270

5.3.18   HaloTag-PEG-AcCl-PEG-YQLRS (24)
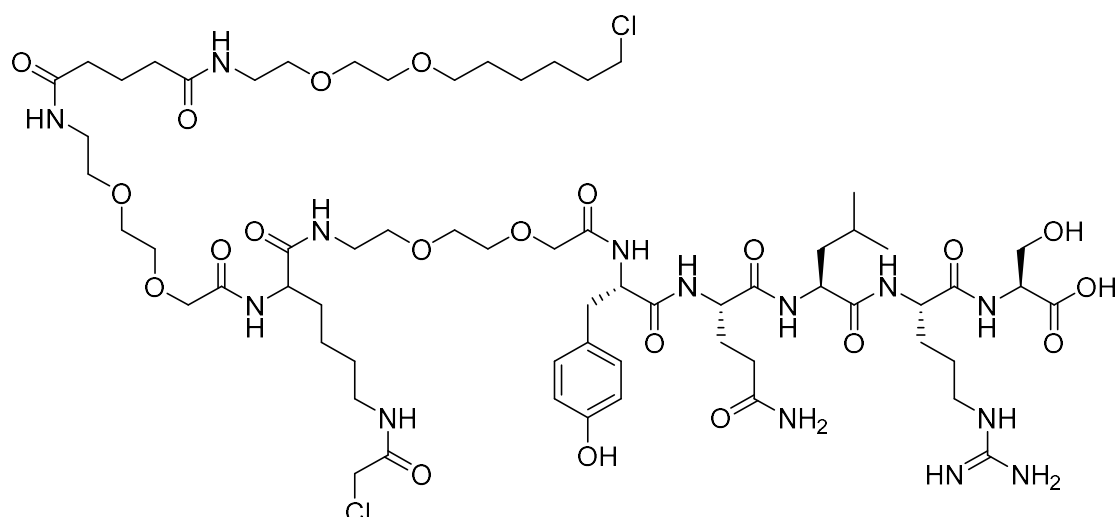
YQLRS peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A.A second Fmoc-PEG-COOH was coupled using method B followed by coupling of the HaloTag substrate using Method H. The Lysine Dde was then deprotected using Method C and the chloroacetyl group coupled to the lysine side chain using Method F. The peptide was then immediately cleaved from the resin and purified using Method I. (Yield: 3.3mg, 5.9%)

HRMS:

$C_{64}H_{108}Cl_2N_{14}O_{21}$ requires [M+H]$^+$ 1479.7263, found: [M+H]$^+$ 1479.7276

5.3.19 HaloTag-PEG-Ac-PEG-YQSRL (25)
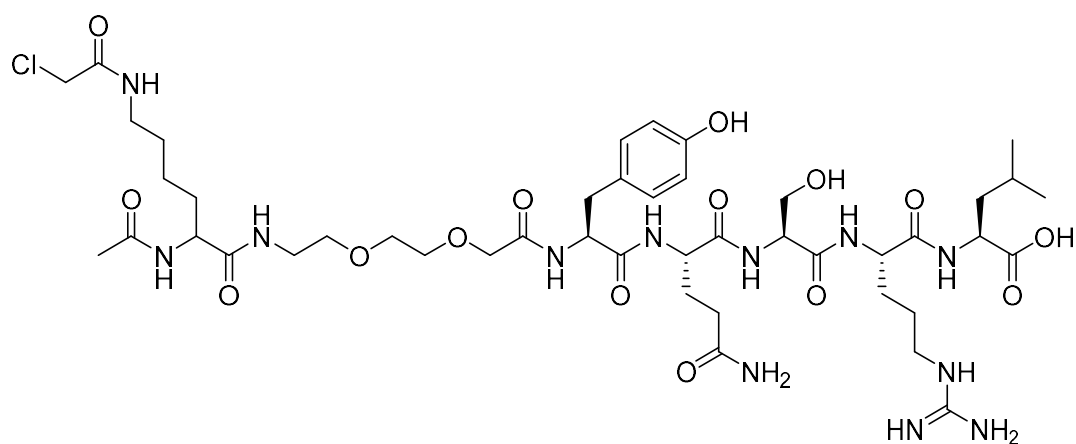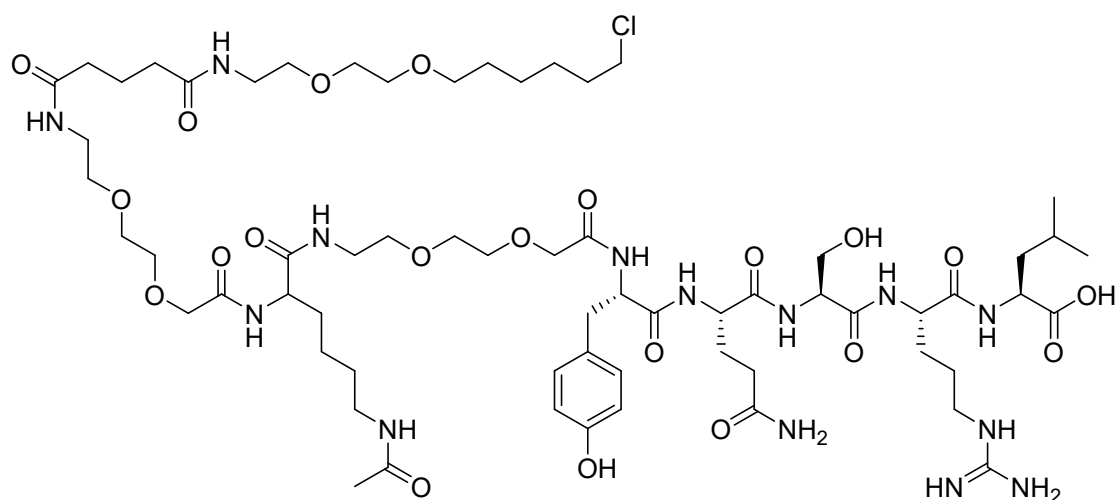

YQSRL peptide was synthesised on 50mg pre-loaded Leu resin (Novabiochem) using Method A. Fmoc-PEG-COOH was coupled using Method B. Fmoc-Lysine Dde was coupled using Method A. A second Fmoc-PEG-COOH was coupled using method B followed by coupling of the HaloTag substrate using Method H. The Lysine Dde was then deprotected using Method C and the side chain capped with an acetyl group using Method G. The peptide was then immediately cleaved from the resin and purified using Method I. (Yield: 4.9mg, 8.9%)

HRMS:

$C_{64}H_{109}ClN_{14}O_{21}$ requires [M+H]$^+$ 1445.7653, found: [M+H]$^+$ 1445.7687

5.3.20   Lissamine-YQSKL (26)
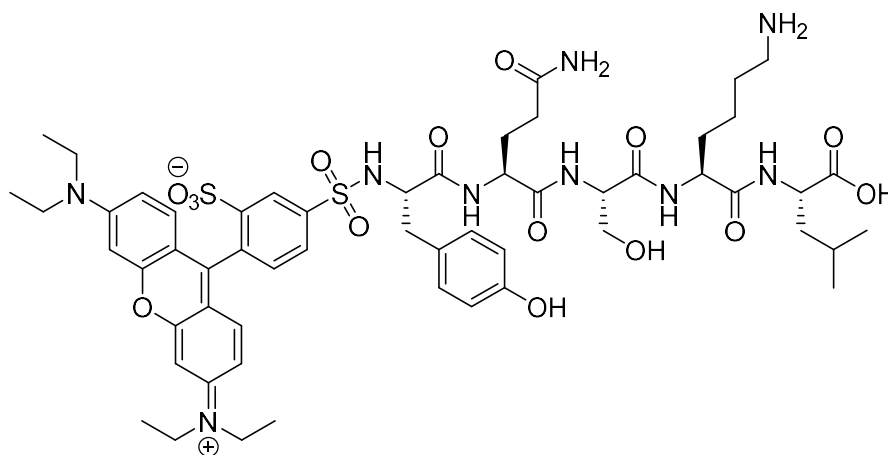
YQSKL peptide was synthesised on 30mg pre-loaded Leu resin (Novabiochem) using Method A. Lissamine was coupling using Method E. Peptide was cleaved and purified using Method I. (Yield: 12mg, 44%)

HRMS:

$C_{56}H_{75}N_9O_{15}S_2$ requires [M+H]$^+$ 1178.4897, found: [M+H]$^+$ 1178.4919

References

AGRAWAL, G. & SUBRAMANI, S. 2016. De novo peroxisome biogenesis: Evolving concepts and conundrums. *Biochimica et Biophysica Acta,* 1863**,** 892-901.

AGRIMI, G., RUSSO, A., SCARCIA, P. & PALMIERI, F. 2012. The human gene SLC25A17 encodes a peroxisomal transporter of coenzyme A, FAD and NAD+. *Biochemical Journal,* 443**,** 241-247.

ALBERTINI, M., REHLING, P., ERDMANN, R., GIRZALSKY, W., KIEL, J. A. K. W., VEENHUIS, M. & KUNAU, W.-H. 1997. Pex14p, a Peroxisomal Membrane Protein Binding Both Receptors of the Two PTS-Dependent Import Pathways. *Cell,* 89**,** 83-92.

ALSFORD, S. A. M., KELLY, J. M., BAKER, N. & HORN, D. 2013. Genetic dissection of drug resistance in trypanosomes. *Parasitology,* 140**,** 1478-1491.

ANDRADE, M. A., CHACÓN, P., MERELO, J. J. & MORÁN, F. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Engineering, Design and Selection,* 6**,** 383-390.

ANTONENKOV, V. D. & HILTUNEN, J. K. 2006. Peroxisomal membrane permeability and solute transfer. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1697-1706.

ANTONENKOV, V. D., SORMUNEN, R. T. & HILTUNEN, J. K. 2004. The rat liver peroxisomal membrane forms a permeability barrier for cofactors but not for small metabolites in vitro. *Journal of Cell Science,* 117**,** 5633-5642.

ARANOVICH, A., HUA, R., RUTENBERG, A. D. & KIM, P. K. 2014. PEX16 contributes to peroxisome maintenance by constantly trafficking PEX3 via the ER. *Journal of Cell Science,* 127**,** 3675-3686.

ARGYRIOU, C., D'AGOSTINO, M. D. & BRAVERMAN, N. 2016. Peroxisome biogenesis disorders. *Translational Science of Rare Diseases,* 1**,** 111-144.

AUBOURG, P., MOSSER, J., DOUAR, A. M., SARDE, C. O., LOPEZ, J. & MANDEL, J. L. 1993. Adrenoleukodystrophy gene: Unexpected homology to a protein involved in peroxisome biogenesis. *Biochimie,* 75**,** 293-302.

BAKER, A., LANYON-HOGG, T. & WARRINER, S. L. 2016. Peroxisome protein import: a complex journey. *Biochemical Society Transactions,* 44**,** 783-789.

BARGLOW, K. T. & CRAVATT, B. F. 2004. Discovering Disease-Associated Enzymes by Proteome Reactivity Profiling. *Chemistry & Biology,* 11**,** 1523-1531.

BARRETT, M. P., ZHANG, Z. Q., DENISE, H., GIROUD, C. & BALTZ, T. 1995. A diamidine-resistant Trypanosoma equiperdum clone contains a P2 purine transporter with reduced substrate affinity. *Molecular and Biochemical Parasitology,* 73**,** 223-229.

BARROS-BARBOSA, A., FERREIRA, M. J., RODRIGUES, T. A., PEDROSA, A. G., GROU, C. P., PINTO, M. P., FRANSEN, M., FRANCISCO, T. & AZEVEDO, J. E. 2019a. Membrane topologies of PEX13 and PEX14 provide new insights on the mechanism of protein import into peroxisomes. *The FEBS Journal,* 286**,** 205-222.

BARROS-BARBOSA, A., RODRIGUES, T. A., FERREIRA, M. J., PEDROSA, A. G., TEIXEIRA, N. R., FRANCISCO, T. & AZEVEDO, J. E. 2019b. The intrinsically disordered nature of the peroxisomal protein translocation machinery. *The FEBS Journal,* 286**,** 24-38.

BARØY, T., KOSTER, J., STRØMME, P., EBBERINK, M. S., MISCEO, D., FERDINANDUSSE, S., HOLMGREN, A., HUGHES, T., MERCKOLL, E., WESTVIK, J., WOLDSETH, B., WALTER, J., WOOD, N., TVEDT, B., STADSKLEIV, K., WANDERS, R. J. A., WATERHAM, H. R. & FRENGEN, E. 2015. A novel type of rhizomelic chondrodysplasia punctata, RCDP5, is caused by loss of the PEX5 long isoform. *Human Molecular Genetics,* 24**,** 5845-5854.

BELLU, A. R., SALOMONS, F. A., KIEL, J. A. K. W., VEENHUIS, M. & VAN DER KLEI, I. J. 2002. Removal of Pex3p Is an Important Initial Stage in Selective Peroxisome Degradation in Hansenula polymorpha. *Journal of Biological Chemistry,* 277**,** 42875-42880.

BENKE, P. J., REYES, P. F. & PARKER, J. C. 1981. New form of adrenoleukodystrophy. *Human Genetics,* 58**,** 204-208.

BHATAYA, A., SCHMIDT-DANNERT, C. & LEE, P. C. 2009. Metabolic engineering of Pichia pastoris X-33 for lycopene production. *Process Biochemistry,* 44**,** 1095-1102.

BHOGAL, M. S., LANYON-HOGG, T., JOHNSTON, K. A., WARRINER, S. L. & BAKER, A. 2016. Covalent Label Transfer between Peroxisomal Importomer Components Reveals Export-driven Import Interactions. *Journal of Biological Chemistry,* 291**,** 2460-2468.

BLOK, N. B., TAN, D., WANG, R. Y.-R., PENCZEK, P. A., BAKER, D., DIMAIO, F., RAPOPORT, T. A. & WALZ, T. 2015. Unique double-ring structure of the peroxisomal Pex1/Pex6 ATPase complex revealed by cryo-electron microscopy. *Proceedings of the National Academy of Sciences,* 112**,** E4017-E4025.

BOETTCHER, M. & MCMANUS, MICHAEL T. 2015. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular Cell,* 58**,** 575-585.

BOSCH, P. J., CORRÊA, I. R., JR., SONNTAG, M. H., IBACH, J., BRUNSVELD, L., KANGER, J. S. & SUBRAMANIAM, V. 2014. Evaluation of fluorophores to label SNAP-tag fused proteins for multicolor single-molecule tracking microscopy in live cells. *Biophysical Journal,* 107**,** 803-814.

BOTTGER, G., BARNETT, P., KLEIN, A. T., KRAGT, A., TABAK, H. F. & DISTEL, B. 2000. Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner. *Molecular Biology of the Cell,* 11**,** 3963-3976.

BRAVERMAN, N., DODT, G., GOULD, S. J. & VALLE, D. 1998. An Isoform of Pex5p, the Human PTS1 Receptor, is Required for the Import of PTS2 Proteins Into Peroxisomes. *Human Molecular Genetics,* 7**,** 1195-1205.

BRAVERMAN, N., STEEL, G., OBIE, C., MOSER, A., MOSER, H., GOULD, S. J. & VALLE, D. 1997. Human PEX7 encodes the peroxisomal PTS2 receptor and is responsible for rhizomelic chondrodysplasia punctata. *Nature Genetics,* 15**,** 369-376.

BROCARD, C. & HARTIG, A. 2006. Peroxisome targeting signal 1: Is it really a simple tripeptide? *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1565-1573.

BROCARD, C., KRAGLER, F., SIMON, M. M., SCHUSTER, T. & HARTIG, A. 1994. The Tetratricopeptide Repeat Domain of the PAS10 Protein of Saccharomyces cerevisiae Is Essential for Binding the Peroxisomal Targeting Signal -SKL. *Biochemical and Biophysical Research Communications,* 204**,** 1016-1022.

BROWN, L.-A. & BAKER, A. 2008. Shuttles and cycles: transport of proteins into the peroxisome matrix (Review). *Molecular Membrane Biology,* 25**,** 363-375.

BUCKLEY, D. L., RAINA, K., DARRICARRERE, N., HINES, J., GUSTAFSON, J. L., SMITH, I. E., MIAH, A. H., HARLING, J. D. & CREWS, C. M. 2015. HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chemical Biology,* 10**,** 1831-1837.

CAMPBELL, A. E. & BENNETT, D. 2016. Targeting protein function: the expanding toolkit for conditional disruption. *The Biochemical journal,* 473**,** 2573-2589.

CARVALHO, A. F., COSTA-RODRIGUES, J., CORREIA, I., COSTA PESSOA, J., FARIA, T. Q., MARTINS, C. L., FRANSEN, M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2006. The N-terminal Half of the Peroxisomal Cycling Receptor Pex5p is a Natively Unfolded Domain. *Journal of Molecular Biology,* 356**,** 864-875.

CARVALHO, A. F., GROU, C. P., PINTO, M. P., ALENCASTRE, I. S., COSTA-RODRIGUES, J., FRANSEN, M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2007a. Functional characterization of two missense mutations in Pex5p—C11S and N526K. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1773**,** 1141-1148.

CARVALHO, A. F., PINTO, M. P., GROU, C. P., ALENCASTRE, I. S., FRANSEN, M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2007b. Ubiquitination of Mammalian Pex5p, the Peroxisomal Import Receptor. *Journal of Biological Chemistry,* 282**,** 31267-31272.

CHAN, A., SCHUMMER, A., FISCHER, S., SCHRÖTER, T., CRUZ-ZARAGOZA, L. D., BENDER, J., DREPPER, F., OELJEKLAUS, S., KUNAU, W.-H., GIRZALSKY, W., WARSCHEID, B. & ERDMANN, R. 2016. Pex17p-dependent assembly of Pex14p/Dyn2p-subcomplexes of the peroxisomal protein import machinery. *European Journal of Cell Biology,* 95**,** 585-597.

CHEESEMAN, M. D., CHESSUM, N. E. A., RYE, C. S., PASQUA, A. E., TUCKER, M. J., WILDING, B., EVANS, L. E., LEPRI, S., RICHARDS, M., SHARP, S. Y., ALI, S., ROWLANDS, M., O'FEE, L., MIAH, A., HAYES, A., HENLEY, A. T., POWERS, M., TE POELE, R., DE BILLY, E., PELLEGRINO, L., RAYNAUD, F., BURKE, R., VAN MONTFORT, R. L. M., ECCLES, S. A., WORKMAN, P. & JONES, K. 2017. Discovery of a Chemical Probe Bisamide (CCT251236): An Orally Bioavailable Efficacious Pirin Ligand from a Heat Shock Transcription Factor 1 (HSF1) Phenotypic Screen. *Journal of Medicinal Chemistry,* 60**,** 180-201.

CHEN, R., RISHI, H. S., POTAPOV, V., YAMADA, M. R., YEH, V. J., CHOW, T., CHEUNG, C. L., JONES, A. T., JOHNSON, T. D., KEATING, A. E., DELOACHE, W. C. & DUEBER, J. E. 2015. A Barcoding Strategy Enabling Higher-Throughput Library Screening by Microscopy. *ACS Synthetic Biology,* 4**,** 1205-1216.

CHENG, Y. 2015. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell,* 161**,** 450-457.

CHESSUM, N. E. A., SHARP, S. Y., CALDWELL, J. J., PASQUA, A. E., WILDING, B., COLOMBANO, G., COLLINS, I., OZER, B., RICHARDS, M., ROWLANDS, M., STUBBS, M., BURKE, R., MCANDREW, P. C., CLARKE, P. A., WORKMAN, P., CHEESEMAN, M. D. & JONES, K. 2018. Demonstrating In-Cell Target Engagement Using a Pirin Protein Degradation Probe (CCT367766). *Journal of Medicinal Chemistry,* 61**,** 918-933.

CHOWDHARY, G., KATAYA, A. R., LINGNER, T. & REUMANN, S. 2012. Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. *BMC Plant Biology,* 12**,** 142-142.

CHU, T.-T., GAO, N., LI, Q.-Q., CHEN, P.-G., YANG, X.-F., CHEN, Y.-X., ZHAO, Y.-F. & LI, Y.-M. 2016. Specific Knockdown of Endogenous Tau Protein by Peptide-Directed Ubiquitin-Proteasome Degradation. *Cell Chemical Biology,* 23**,** 453-461.

CINIAWSKY, S., GRIMM, I., SAFFIAN, D., GIRZALSKY, W., ERDMANN, R. & WENDLER, P. 2015. Molecular snapshots of the Pex1/6 AAA+ complex in action. *Nature Communications,* 6**,** 7331.

COLE, N. B. 2013. Site-Specific Protein Labeling with SNAP-Tags. *Current Protocols in Protein Science,* 73**,** 30.1.1-30.1.16.

COLE, R. D., STEIN, W. H. & MOORE, S. 1958. On the Cysteine Content of Human Hemoglobin. *Journal of Biological Chemistry,* 233**,** 1359-1363.

COOPER, T. G. & BEEVERS, H. 1969. β Oxidation in Glyoxysomes from Castor Bean Endosperm. *Journal of Biological Chemistry,* 244**,** 3514-3520.

CORREA, I. R., BAKER, B., ZHANG, A., SUN, L., PROVOST, C. R., LUKINAVICIUS, G., REYMOND, L., JOHNSSON, K. & XU, M.-Q. 2013. Substrates for Improved Live-Cell Fluorescence Labeling of SNAP-tag. *Current Pharmaceutical Design,* 19**,** 5414-5420.

COSTA-RODRIGUES, J., CARVALHO, A. F., FRANSEN, M., HAMBRUCH, E., SCHLIEBS, W., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2005. Pex5p, the Peroxisomal Cycling Receptor, Is a Monomeric Non-globular Protein. *Journal of Biological Chemistry,* 280**,** 24404-24411.

CREWS, C. M. 2010. Targeting the Undruggable Proteome: The Small Molecules of My Dreams. *Chemistry & Biology,* 17**,** 551-555.

CROSS, L. L. 2016. *Re-Design of a Receptor-Targeting Signal Interaction to Create a New Peroxisomal Trafficking Pathway.* Doctor of Philosophy, University of Leeds.

CROSS, L. L., PAUDYAL, R., KAMISUGI, Y., BERRY, A., CUMING, A. C., BAKER, A. & WARRINER, S. L. 2017. Towards designer organelles by subverting the peroxisomal import pathway. *Nature Communications,* 8**,** 454.

D'ANDREA, L. D. & REGAN, L. 2003. TPR proteins: the versatile helix. *Trends in Biochemical Sciences,* 28**,** 655-662.

DANPURE, C. J. 1989. Recent advances in the understanding, diagnosis and treatment of primary hyperoxaluria type 1. *Journal of Inherited Metabolic Disease,* 12**,** 210-224.

DANPURE, C. J. 2005. Molecular Etiology of Primary Hyperoxaluria Type 1: New Directions for Treatment. *American Journal of Nephrology,* 25**,** 303-310.

DANPURE, C. J. & JENNINGS, P. R. 1986. Peroxisomal alanine:glyoxylate aminotransferase deficiency in primary hyperoxaluria type I. *FEBS Letters,* 201**,** 20-34.

DAWIDOWSKI, M., EMMANOUILIDIS, L., KALEL, V. C., TRIPSIANES, K., SCHORPP, K., HADIAN, K., KAISER, M., MÄSER, P., KOLONKO, M., TANGHE, S., RODRIGUEZ, A., SCHLIEBS, W., ERDMANN, R., SATTLER, M. & POPOWICZ, G. M. 2017. Inhibitors of PEX14 disrupt protein import into glycosomes and kill *Trypanosoma* parasites. *Science,* 355**,** 1416-1420.

DE DUVE, C. & BAUDHUIN, P. 1966. Peroxisomes (microbodies and related particles). *Physiological Reviews,* 46**,** 323-357.

DE MARCOS LOUSA, C., VAN ROERMUND, C. W. T., POSTIS, V. L. G., DIETRICH, D., KERR, I. D., WANDERS, R. J. A., BALDWIN, S. A., BAKER, A. & THEODOULOU, F. L. 2013. Intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding cassette transporter is required for transport and metabolism of fatty acids. *Proceedings of the National Academy of Sciences,* 110**,** 1279-1284.

DE VET, E. C. J. M., IJLST, L., OOSTHEIM, W., WANDERS, R. J. A. & VAN DEN BOSCH, H. 1998. Alkyl-Dihydroxyacetonephosphate Synthase: FATE IN PEROXISOME BIOGENESIS DISORDERS AND IDENTIFICATION OF THE POINT MUTATION UNDERLYING A SINGLE ENZYME DEFICIENCY. *Journal of Biological Chemistry,* 273**,** 10296-10301.

DEAN, J. M. & LODHI, I. J. 2018. Structural and functional roles of ether lipids. *Protein & Cell,* 9**,** 196-206.

DEL RÍO, L. A., SANDALIO, L. M., PALMA, J., BUENO, P. & CORPAS, F. J. 1992. Metabolism of oxygen radicals in peroxisomes and cellular implications. *Free Radical Biology and Medicine,* 13**,** 557-580.

DELOACHE, W. C., RUSS, Z. N. & DUEBER, J. E. 2016. Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. *Nature Communications,* 7**,** 11152.

DEOSARAN, E., LARSEN, K. B., HUA, R., SARGENT, G., WANG, Y., KIM, S., LAMARK, T., JAUREGUI, M., LAW, K., LIPPINCOTT-SCHWARTZ, J., BRECH, A., JOHANSEN, T. & KIM, P. K. 2013. NBR1 acts as an autophagy receptor for peroxisomes. *Journal of Cell Science,* 126**,** 939-952.

DIAS, A. F., RODRIGUES, T. A., PEDROSA, A. G., BARROS-BARBOSA, A., FRANCISCO, T. & AZEVEDO, J. E. 2017. The peroxisomal matrix protein translocon is a large cavity-forming protein assembly into which PEX5 protein enters to release its cargo. *Journal of Biological Chemistry,* 292**,** 15287-15300.

DISTEL, B., GOULD, S. J., VOORN-BROUWER, T., VAN DER BERG, M., TABAK, H. F. & SUBRAMANI, S. 1992. The carboxyl-terminal tripeptide serine-lysine-leucine of firefly luciferase is necessary but not sufficient for peroxisomal import in yeast. *The New Biologist,* 4**,** 157-65.

DODT, G., BRAVERMAN, N., WONG, C., MOSER, A., MOSER, H. W., WATKINS, P., VALLE, D. & GOULD, S. J. 1995. Mutations in the PTS1

receptor gene, PXR1, define complementation group 2 of the peroxisome biogenesis disorders. *Nature Genetics,* 9**,** 115-125.

DODT, G., WARREN, D., BECKER, E., REHLING, P. & GOULD, S. J. 2001. Domain Mapping of Human PEX5 Reveals Functional and Structural Similarities to Saccharomyces cerevisiae Pex18p and Pex21p. *Journal of Biological Chemistry,* 276**,** 41769-41781.

DOUANGAMATH, A., FILIPP, F. V., KLEIN, A. T. J., BARNETT, P., ZOU, P., VOORN-BROUWER, T., VEGA, M. C., MAYANS, O. M., SATTLER, M., DISTEL, B. & WILMANNS, M. 2002. Topography for Independent Binding of α-Helical and PPII-Helical Ligands to a Peroxisomal SH3 Domain. *Molecular Cell,* 10**,** 1007-1017.

DRYSDALE, G. R. & LARDY, H. A. 1953. Fatty Acid Oxidation by a Soluble Enzyme System from Mitochondria. *Journal of Biological Chemistry,* 202**,** 119-136.

EBBERINK, M. S., MOOIJER, P. A. W., GOOTJES, J., KOSTER, J., WANDERS, R. J. A. & WATERHAM, H. R. 2011. Genetic classification and mutational spectrum of more than 600 patients with a Zellweger syndrome spectrum disorder. *Human Mutation,* 32**,** 59-69.

EFFELSBERG, D., CRUZ-ZARAGOZA, L. D., TONILLO, J., SCHLIEBS, W. & ERDMANN, R. 2015. Role of Pex21p for Piggyback Import of Gpd1p and Pnc1p into Peroxisomes of Saccharomyces cerevisiae. *Journal of Biological Chemistry,* 290**,** 25333-25342.

ELGERSMA, Y., VOS, A., VAN DEN BERG, M., VAN ROERMUND, C. W., VAN DER SLUIJS, P., DISTEL, B. & TABAK, H. F. 1996. Analysis of the carboxyl-terminal peroxisomal targeting signal 1 in a homologous context in Saccharomyces cerevisiae. *Journal of Biological Chemistry: Cell Biology and Metabolsim,* 271**,** 26375-82.

EMANUELSSON, O., ELOFSSON, A., VON HEIJNE, G. & CRISTÓBAL, S. 2003. In Silico Prediction of the Peroxisomal Proteome in Fungi, Plants and Animals. *Journal of Molecular Biology,* 330**,** 443-456.

EMMANOUILIDIS, L., GOPALSWAMY, M., PASSON, D. M., WILMANNS, M. & SATTLER, M. 2016. Structural biology of the import pathways of peroxisomal matrix proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1863**,** 804-813.

ERDMANN, R. & SCHLIEBS, W. 2005. Peroxisomal matrix protein import: the transient pore model. *Nature Reviews Molecular Cell Biology,* 6**,** 738-742.

ERDMANN, R. S., BAGULEY, S. W., RICHENS, J. H., WISSNER, R. F., XI, Z., ALLGEYER, E. S., ZHONG, S., THOMPSON, A. D., LOWE, N., BUTLER, R., BEWERSDORF, J., ROTHMAN, J. E., ST JOHNSTON, D., SCHEPARTZ, A. & TOOMRE, D. 2019. Labeling Strategies Matter for Super-Resolution Microscopy: A Comparison between HaloTags and SNAP-tags. *Cell Chemical Biology,* 26**,** 584-592.e6.

FARMER, L. M., RINALDI, M. A., YOUNG, P. G., DANAN, C. H., BURKHART, S. E. & BARTEL, B. 2013. Disrupting Autophagy Restores Peroxisome Function to an *Arabidopsis* lon2 Mutant and Reveals a Role for the LON2 Protease in Peroxisomal Matrix Protein Degradation. *The Plant Cell,* 25**,** 4085-4100.

FARRÉ, J.-C., CAROLINO, K., STASYK, O. V., STASYK, O. G., HODZIC, Z., AGRAWAL, G., TILL, A., PROIETTO, M., CREGG, J., SIBIRNY, A. A. & SUBRAMANI, S. 2017. A New Yeast Peroxin, Pex36, a Functional

Homolog of Mammalian PEX16, Functions in the ER-to-Peroxisome Traffic of Peroxisomal Membrane Proteins. *Journal of Molecular Biology,* 429**,** 3743-3762.

FERDINANDUSSE, S., DENIS, S., CLAYTON, P. T., GRAHAM, A., REES, J. E., ALLEN, J. T., MCLEAN, B. N., BROWN, A. Y., VREKEN, P., WATERHAM, H. R. & WANDERS, R. J. A. 2000a. Mutations in the gene encoding peroxisomal α-methylacyl-CoA racemase cause adult-onset sensory motor neuropathy. *Nature Genetics,* 24**,** 188-191.

FERDINANDUSSE, S., DENIS, S., FAUST, P. L. & WANDERS, R. J. A. 2009. Bile acids: the role of peroxisomes. *Journal of Lipid Research,* 50**,** 2139-2147.

FERDINANDUSSE, S., DENIS, S., IJLST, L., DACREMONT, G., WATERHAM, H. R. & WANDERS, R. J. A. 2000b. Subcellular localization and physiological role of α-methylacyl-CoA racemase. *Journal of Lipid Research,* 41**,** 1890-1896.

FERDINANDUSSE, S., DENIS, S., MOOIJER, P. A. W., ZHANG, Z., REDDY, J. K., SPECTOR, A. A. & WANDERS, R. J. A. 2001a. Identification of the peroxisomal beta-oxidation enzymes involved in the biosynthesis of docosahexaenoic acid. *Journal of Lipid Research,* 42**,** 1987-1995.

FERDINANDUSSE, S., DENIS, S., MOOYER, P. A. W., DEKKER, C., DURAN, M., SOORANI-LUNSING, R. J., BOLTSHAUSER, E., MACAYA, A., GÄRTNER, J., MAJOIE, C. B. L. M., BARTH, P. G., WANDERS, R. J. A. & POLL-THE, B. T. 2006a. Clinical and biochemical spectrum of D-bifunctional protein deficiency. *Annals of Neurology,* 59**,** 92-104.

FERDINANDUSSE, S., JIMENEZ-SANCHEZ, G., KOSTER, J., DENIS, S., VAN ROERMUND, C. W., SILVA-ZOLEZZI, I., MOSER, A. B., VISSER, W. F., GULLUOGLU, M., DURMAZ, O., DEMIRKOL, M., WATERHAM, H. R., GÖKCAY, G., WANDERS, R. J. A. & VALLE, D. 2014. A novel bile acid biosynthesis defect due to a deficiency of peroxisomal ABCD3. *Human Molecular Genetics,* 24**,** 361-370.

FERDINANDUSSE, S., KOSTOPOULOS, P., DENIS, S., RUSCH, H., OVERMARS, H., DILLMANN, U., REITH, W., HAAS, D., WANDERS, R. J. A., DURAN, M. & MARZINIAK, M. 2006b. Mutations in the gene encoding peroxisomal sterol carrier protein X (SCPx) cause leukencephalopathy with dystonia and motor neuropathy. *American Journal of Human Genetics,* 78**,** 1046-1052.

FERDINANDUSSE, S., OVERMARS, H., DENIS, S., WATERHAM, H. R., WANDERS, R. J. A. & VREKEN, P. 2001b. Plasma analysis of di- and trihydroxycholestanoic acid diastereoisomers in peroxisomal α-methylacyl-CoA racemase deficiency. *Journal of Lipid Research,* 42**,** 137-141.

FODOR, K., WOLF, J., ERDMANN, R., SCHLIEBS, W. & WILMANNS, M. 2012. Molecular Requirements for Peroxisomal Targeting of Alanine-Glyoxylate Aminotransferase as an Essential Determinant in Primary Hyperoxaluria Type 1. *PLOS Biology,* 10**,** e1001309.

FODOR, K., WOLF, J., REGLINSKI, K., PASSON, D. M., LOU, Y., SCHLIEBS, W., ERDMANN, R. & WILMANNS, M. 2015. Ligand-Induced Compaction of the PEX5 Receptor-Binding Cavity Impacts Protein Import Efficiency into Peroxisomes. *Traffic,* 16**,** 85-98.

FOULON, V., ANTONENKOV, V. D., CROES, K., WAELKENS, E., MANNAERTS, G. P., VAN VELDHOVEN, P. P. & CASTEELS, M. 1999.

Purification, molecular cloning, and expression of 2-hydroxyphytanoyl-CoA lyase, a peroxisomal thiamine pyrophosphate-dependent enzyme that catalyzes the carbon–carbon bond cleavage during α-oxidation of 3-methyl-branched fatty acids. *Proceedings of the National Academy of Sciences,* 96**,** 10039-10044.

FOURCADE, S., RUIZ, M., CAMPS, C., SCHLÜTER, A., HOUTEN, S. M., MOOYER, P. A. W., PÀMPOLS, T., DACREMONT, G., WANDERS, R. J. A., GIRÒS, M. & PUJOL, A. 2009. A key role for the peroxisomal ABCD2 transporter in fatty acid homeostasis. *American Journal of Physiology-Endocrinology and Metabolism,* 296**,** E211-E221.

FRAIN, K. M., ROBINSON, C. & VAN DIJL, J. M. 2019. Transport of Folded Proteins by the Tat System. *The Protein Journal,* 38**,** 377-388.

FRANCISCO, T., RODRIGUES, T. A., FREITAS, M. O., GROU, C. P., CARVALHO, A. F., SÁ-MIRANDA, C., PINTO, M. P. & AZEVEDO, J. E. 2013. A cargo-centered perspective on the PEX5 receptor-mediated peroxisomal protein import pathway. *The Journal of Biological Chemistry,* 288**,** 29151-29159.

FRANSEN, M. 2014. HaloTag as a Tool to Investigate Peroxisome Dynamics in Cultured Mammalian Cells. *In:* IVANOV, A. I. (ed.) *Exocytosis and Endocytosis.* New York, NY: Springer New York.

FRANSEN, M., BREES, C., BAUMGART, E., VANHOOREN, J. C. T., BAES, M., MANNAERTS, G. P. & VAN VELDHOVEN, P. P. 1995. Identification and Characterization of the Putative Human Peroxisomal C-terminal Targeting Signal Import Receptor. *Journal of Biological Chemistry,* 270**,** 7731-7736.

FRANSEN, M., NORDGREN, M., WANG, B. & APANASETS, O. 2012. Role of peroxisomes in ROS/RNS-metabolism: Implications for human disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1822**,** 1363-1373.

FRANSEN, M., TERLECKY, S. R. & SUBRAMANI, S. 1998. Identification of a human PTS1 receptor docking protein directly required for peroxisomal protein import. *Proceedings of the National Academy of Sciences,* 95**,** 8087-8092.

FREITAS, M. O., FRANCISCO, T., RODRIGUES, T. A., ALENCASTRE, I. S., PINTO, M. P., GROU, C. P., CARVALHO, A. F., FRANSEN, M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2011. PEX5 Protein Binds Monomeric Catalase Blocking Its Tetramerization and Releases It upon Binding the N-terminal Domain of PEX14. *Journal of Biological Chemistry,* 286**,** 40509-40519.

FUKUI, S., TANAKA, A., KAWAMOTO, S., YASUHARA, S., TERANISHI, Y. & OSUMI, M. 1975. Ultrastructure of methanol-utilizing yeast cells: appearance of microbodies in relation to high catalase activity. *Journal of Bacteriology,* 123**,** 317-328.

FURUYA, T., KESSLER, P., JARDIM, A., SCHNAUFER, A., CRUDDER, C. & PARSONS, M. 2002. Glucose is toxic to glycosome-deficient trypanosomes. *Proceedings of the National Academy of Sciences,* 99**,** 14177-14182.

GALDEANO, C. & BARRIL, X. 2021. CHAPTER 7 Targeted Protein Degradation Chemical Probes. *The Discovery and Utility of Chemical Probes in Target Discovery.* The Royal Society of Chemistry.

GARDNER, B. M., CASTANZO, D. T., CHOWDHURY, S., STJEPANOVIC, G., STEFELY, M. S., HURLEY, J. H., LANDER, G. C. & MARTIN, A. 2018. The peroxisomal AAA-ATPase Pex1/Pex6 unfolds substrates by processive threading. *Nature Communications,* 9**,** 135.

GARDNER, B. M., CHOWDHURY, S., LANDER, G. C. & MARTIN, A. 2015. The Pex1/Pex6 Complex Is a Heterohexameric AAA+ Motor with Alternating and Highly Coordinated Subunits. *Journal of Molecular Biology,* 427**,** 1375-1388.

GATTO, G. J., MAYNARD, E. L., GUERRERIO, A. L., GEISBRECHT, B. V., GOULD, S. J. & BERG, J. M. 2003. Correlating Structure and Affinity for PEX5:PTS1 Complexes. *Biochemistry,* 42**,** 1660-1666.

GATTO JR, G. J., GEISBRECHT, B. V., GOULD, S. J. & BERG, J. M. 2000. Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nature Structural Biology,* 7**,** 1091.

GEERTS, S., HOLMES, P. H., EISLER, M. C. & DIALL, O. 2001. African bovine trypanosomiasis: the problem of drug resistance. *Trends in Parasitology,* 17**,** 25-28.

GEUZE, H. J., MURK, J. L., STROOBANTS, A. K., GRIFFITH, J. M., KLEIJMEER, M. J., KOSTER, A. J., VERKLEIJ, A. J., DISTEL, B. & TABAK, H. F. 2003. Involvement of the Endoplasmic Reticulum in Peroxisome Formation. *Molecular Biology of the Cell,* 14**,** 2900-2907.

GEWIES, A. & GRIMM, S. 2003. UBP41 Is a Proapoptotic Ubiquitin-specific Protease. *Cancer Research,* 63**,** 682-688.

GIRZALSKY, W., REHLING, P., STEIN, K., KIPPER, J., BLANK, L., KUNAU, W.-H. & ERDMANN, R. 1999. Involvement of Pex13p in Pex14p Localization and Peroxisomal Targeting Signal 2–dependent Protein Import into Peroxisomes. *Journal of Cell Biology,* 144**,** 1151-1162.

GOULD, S. J., KELLER, G. A., HOSKEN, N., WILKINSON, J. & SUBRAMANI, S. 1989. A conserved tripeptide sorts proteins to peroxisomes. *The Journal of Cell Biology,* 108**,** 1657-1664.

GOUVEIA, A. M., GUIMARÃES, C. P., OLIVEIRA, M. E., REGUENGA, C., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2003a. Characterization of the Peroxisomal Cycling Receptor, Pex5p, Using a Cell-free in Vitro Import System. *Journal of Biological Chemistry,* 278**,** 226-232.

GOUVEIA, A. M., GUIMARÃES, C. P., OLIVEIRA, M. E., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2003b. Insertion of Pex5p into the Peroxisomal Membrane Is Cargo Protein-dependent. *Journal of Biological Chemistry,* 278**,** 4389-4392.

GOUVEIA, A. M. M., REGUENGA, C., OLIVEIRA, M. E. M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2000. Characterization of Peroxisomal Pex5p from Rat Liver: Pex5p in the Pex5p-Pex14p Membrane Complex is a Transmembrane Protein. *Journal of Biological Chemistry,* 275**,** 32444-32451.

GREWAL, P. S., SAMSON, J. A., BAKER, J. J., CHOI, B. & DUEBER, J. E. 2021. Peroxisome compartmentalization of a toxic enzyme improves alkaloid production. *Nature Chemical Biology,* 17**,** 96-103.

GROU, C. P., CARVALHO, A. F., PINTO, M. P., ALENCASTRE, I. S., RODRIGUES, T. A., FREITAS, M. O., FRANCISCO, T., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2009a. The peroxisomal protein import machinery – a case report of transient ubiquitination with a new flavor. *Cellular and Molecular Life Sciences,* 66**,** 254-262.

GROU, C. P., CARVALHO, A. F., PINTO, M. P., HUYBRECHTS, S. J., SÁ-MIRANDA, C., FRANSEN, M. & AZEVEDO, J. E. 2009b. Properties of the Ubiquitin-Pex5p Thiol Ester Conjugate. *Journal of Biological Chemistry,* 284**,** 10504-10513.

GROU, C. P., CARVALHO, A. F., PINTO, M. P., WIESE, S., PIECHURA, H., MEYER, H. E., WARSCHEID, B., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2008. Members of the E2D (UbcH5) Family Mediate the Ubiquitination of the Conserved Cysteine of Pex5p, the Peroxisomal Import Receptor. *Journal of Biological Chemistry,* 283**,** 14190-14197.

GROU, C. P., FRANCISCO, T., RODRIGUES, T. A., FREITAS, M. O., PINTO, M. P., CARVALHO, A. F., DOMINGUES, P., WOOD, S. A., RODRÍGUEZ-BORGES, J. E., SÁ-MIRANDA, C., FRANSEN, M. & AZEVEDO, J. E. 2012. Identification of Ubiquitin-specific Protease 9X (USP9X) as a Deubiquitinase Acting on Ubiquitin-Peroxin 5 (PEX5) Thioester Conjugate. *Journal of Biological Chemistry,* 287**,** 12815-12827.

GULICK, T. 2003. Transfection Using DEAE-Dextran. *Current Protocols in Cell Biology,* 19**,** 20.4.1-20.4.10.

HAANSTRA, J. R., VAN TUIJL, A., KESSLER, P., REIJNDERS, W., MICHELS, P. A. M., WESTERHOFF, H. V., PARSONS, M. & BAKKER, B. M. 2008. Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. *Proceedings of the National Academy of Sciences,* 105**,** 17718-17723.

HAGEN, S., DREPPER, F., FISCHER, S., FODOR, K., PASSON, D., PLATTA, H. W., ZENN, M., SCHLIEBS, W., GIRZALSKY, W., WILMANNS, M., WARSCHEID, B. & ERDMANN, R. 2015. Structural Insights into Cargo Recognition by the Yeast PTS1 Receptor. *Journal of Biological Chemistry,* 290**,** 26610-26626.

HAGMANN, V., SOMMER, S., FABIAN, P., BIERLMEIER, J., VAN TREEL, N., MOOTZ, H. D., SCHWARZER, D., AZEVEDO, J. E. & DODT, G. 2018. Chemically monoubiquitinated PEX5 binds to the components of the peroxisomal docking and export machinery. *Scientific Reports,* 8**,** 16014.

HAJRA, A. K. & BISHOP, J. E. 1982. GLYCEROLIPID BIOSYNTHESIS IN PEROXISOMES VIA THE ACYL DIHYDROXYACETONE PHOSPHATE PATHWAY*. *Annals of the New York Academy of Sciences,* 386**,** 170-182.

HAJRA, A. K. & DAS, A. K. 1996. Lipid Biosynthesis in Peroxisomes. *Annals of the New York Academy of Sciences,* 804**,** 129-141.

HARANO, T., NOSE, S., UEZU, R., SHIMIZU, N. & FUJIKI, Y. 2001. Hsp70 regulates the interaction between the peroxisome targeting signal type 1 (PTS1)-receptor Pex5p and PTS1. *The Biochemical J ournal,* 357**,** 157-165.

HARDER, B., SCHOMBURG, A., PFLANZ, R., KÜSTNER, K. M., GERLACH, N. & SCHUH, R. 2008. TEV protease-mediated cleavage in Drosophila as a tool to analyze protein functions in living organisms. *BioTechniques,* 44**,** 765-772.

HASAN, S., PLATTA, H. W. & ERDMANN, R. 2013. Import of proteins into the peroxisomal matrix. *Frontiers in Physiology,* 4**,** 261-261.

HATAKEYAMA, H., NAKAHATA, Y., YARIMIZU, H. & KANZAKI, M. 2017. Live-cell single-molecule labeling and analysis of myosin motors with quantum dots. *Molecular Biology of the Cell,* 28**,** 173-181.

HEINIG, M. & FRISHMAN, D. 2004. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Research,* 32**,** W500-W502.

HERSHKO, A. & CIECHANOVER, A. 1998. THE UBIQUITIN SYSTEM. *Annual Review of Biochemistry,* 67**,** 425-479.

HOCH, D. G., ABEGG, D. & ADIBEKIAN, A. 2018. Cysteine-reactive probes and their use in chemical proteomics. *Chemical Communications,* 54**,** 4501-4512.

HOEFLER, G., HOEFLER, S., WATKINS, P. A., CHEN, W. W., MOSER, A., BALDWIN, V., MCGILLIVARY, B., CHARROW, J., FRIEDMAN, J. M., RUTLEDGE, L., HASHIMOTO, T. & MOSER, H. W. 1988. Biochemical abnormalities in rhizomelic chondrodysplasia punctata. *The Journal of Pediatrics,* 112**,** 726-733.

HOFFER, L., VOITOVICH, Y. V., RAUX, B., CARRASCO, K., MULLER, C., FEDOROV, A. Y., DERVIAUX, C., AMOURIC, A. S., BETZI, S. P., HORVATH, D., VARNEK, A., COLLETTE, Y., COMBES, S. B., ROCHE, P. & MORELLI, X. 2018. Integrated Strategy for Lead Optimization Based on Fragment Growing: The Diversity-Oriented-Target-Focused-Synthesis Approach. *Journal of Medicinal Chemistry,* 61**,** 5719 - 5732.

HUHSE, B., REHLING, P., ALBERTINI, M., BLANK, L., MELLER, K. & KUNAU, W.-H. 1998. Pex17p of Saccharomyces cerevisiae Is a Novel Peroxin and Component of the Peroxisomal Protein Translocation Machinery. *Journal of Cell Biology,* 140**,** 49-60.

HUYBRECHTS, S. J., VAN VELDHOVEN, P. P., BREES, C., MANNAERTS, G. P., LOS, G. V. & FRANSEN, M. 2009. Peroxisome Dynamics in Cultured Mammalian Cells. *Traffic,* 10**,** 1722-1733.

IACOVACHE, I., VAN DER GOOT, F. G. & PERNOT, L. 2008. Pore formation: An ancient yet complex form of attack. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1778**,** 1611-1623.

INCARBONE, M., RITZENTHALER, C. & DUNOYER, P. 2018. Peroxisomal Targeting as a Sensitive Tool to Detect Protein-Small RNA Interactions through in Vivo Piggybacking. *Frontiers in Plant Science,* 9.

ISLINGER, M., LI, K. W., SEITZ, J., VÖLKL, A. & LÜERS, G. H. 2009. Hitchhiking of Cu/Zn Superoxide Dismutase to Peroxisomes – Evidence for a Natural Piggyback Import Mechanism in Mammals. *Traffic,* 10**,** 1711-1721.

JANSEN, G. A., OFTNAN, R., FERDINANDUSSE, S., IJLST, L., MUIJSERS, A. O., SKJELDAL, O. H., STOKKE, O., JAKOBS, C., BESLEY, G. T. N., WRAITH, J. E. & WANDERS, R. J. A. 1997. Refsum disease is caused by mutations in the phytanoyl–CoA hydroxylase gene. *Nature Genetics,* 17**,** 190-193.

JANSEN, G. A., VAN DEN BRINK, D. M., OFMAN, R., DRAGHICI, O., DACREMONT, G. & WANDERS, R. J. A. 2001. Identification of Pristanal Dehydrogenase Activity in Peroxisomes: Conclusive Evidence That the Complete Phytanic Acid α-Oxidation Pathway Is Localized in Peroxisomes. *Biochemical and Biophysical Research Communications,* 283**,** 674-679.

JANSSEN, D. B. 2004. Evolving haloalkane dehalogenases. *Current Opinion in Chemical Biology,* 8**,** 150-159.

JOHNSON, M. A., SNYDER, W. B., LIN CEREGHINO, J., VEENHUIS, M., SUBRAMANI, S. & CREGG, J. M. 2001. Pichia pastoris Pex14p, a

phosphorylated peroxisomal membrane protein, is part of a PTS–receptor docking complex and interacts with many peroxins. *Yeast,* 18**,** 621-641.

JONES, A. T. & SAYERS, E. J. 2012. Cell entry of cell penetrating peptides: tales of tails wagging dogs. *Journal of Controlled Release,* 161**,** 582-591.

JONES, J. M., MORRELL, J. C. & GOULD, S. J. 2004. PEX19 is a predominantly cytosolic chaperone and import receptor for class 1 peroxisomal membrane proteins. *The Journal of Cell Biology,* 164**,** 57-67.

JOSHI, A. S., HUANG, X., CHOUDHARY, V., LEVINE, T. P., HU, J. & PRINZ, W. A. 2016. A family of membrane-shaping proteins at ER subdomains regulates pre-peroxisomal vesicle biogenesis. *The Journal of Cell Biology,* 215**,** 515-529.

JUILLERAT, A., GRONEMEYER, T., KEPPLER, A., GENDREIZIG, S., PICK, H., VOGEL, H. & JOHNSSON, K. 2003. Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chemistry & Biology,* 10**,** 313-317.

KABEYA, Y., MIZUSHIMA, N., UENO, T., YAMAMOTO, A., KIRISAKO, T., NODA, T., KOMINAMI, E., OHSUMI, Y. & YOSHIMORI, T. 2000. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *The EMBO Journal,* 19**,** 5720-5728.

KAWAGUCHI, K., MUKAI, E., WATANABE, S., YAMASHITA, A., MORITA, M., SO, T. & IMANAKA, T. 2021. Acyl-CoA thioesterase activity of peroxisomal ABC protein ABCD1 is required for the transport of very long-chain acyl-CoA into peroxisomes. *Scientific Reports,* 11**,** 2192.

KELLEY, R. I., DATTA, N. S., DOBYNS, W. B., HAJRA, A. K., MOSER, A. B., NOETZEL, M. J., ZACKAI, E. H., MOSER, H. W., OPITZ, J. M. & REYNOLDS, J. F. 1986. Neonatal adrenoleukodystrophy: New cases, biochemical studies, and differentiation from Zellweger and related peroxisomal polydystrophy syndromes. *American Journal of Medical Genetics,* 23**,** 869-901.

KEMPIŃSKI, B., CHEŁSTOWSKA, A., POZNAŃSKI, J., KRÓL, K., RYMER, Ł., FRYDZIŃSKA, Z., GIRZALSKY, W., SKONECZNA, A., ERDMANN, R. & SKONECZNY, M. 2020. The Peroxisomal Targeting Signal 3 (PTS3) of the Budding Yeast Acyl-CoA Oxidase Is a Signal Patch. *Frontiers in Cell and Developmental Biology,* 8.

KEPPLER, A., GENDREIZIG, S., GRONEMEYER, T., PICK, H., VOGEL, H. & JOHNSSON, K. 2003. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology,* 21**,** 86-89.

KEPPLER, A., KINDERMANN, M., GENDREIZIG, S., PICK, H., VOGEL, H. & JOHNSSON, K. 2004. Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods,* 32**,** 437-444.

KERSSEN, D., HAMBRUCH, E., KLAAS, W., PLATTA, H. W., DE KRUIJFF, B., ERDMANN, R., KUNAU, W.-H. & SCHLIEBS, W. 2006. Membrane Association of the Cycling Peroxisome Import Receptor Pex5p. *Journal of Biological Chemistry,* 281**,** 27003-27015.

KIKUCHI, M., HATANO, N., YOKOTA, S., SHIMOZAWA, N., IMANAKA, T. & TANIGUCHI, H. 2004. Proteomic Analysis of Rat Liver Peroxisome: Presence of Peroxisome-Specific Isozyme of LON protease. *Journal of Biological Chemistry,* 279**,** 421-428.

KIM, P. K., HAILEY, D. W., MULLEN, R. T. & LIPPINCOTT-SCHWARTZ, J. 2008. Ubiquitin signals autophagic degradation of cytosolic proteins and peroxisomes. *Proceedings of the National Academy of Sciences,* 105**,** 20567-20574.

KIM , P. K., MULLEN , R. T., SCHUMANN , U. & LIPPINCOTT-SCHWARTZ , J. 2006. The origin and maintenance of mammalian peroxisomes involves a de novo PEX16-dependent pathway from the ER. *Journal of Cell Biology,* 173**,** 521-532.

KIM, Y.-I., NAM, I.-K., LEE, D.-K., BHANDARI, S., CHARTON, L., KWAK, S., LIM, J.-Y., HONG, K., KIM, S.-J., LEE, J. N., KWON, S. W., SO, H.-S., LINKA, N., PARK, R. & CHOE, S.-K. 2019. Slc25a17 acts as a peroxisomal coenzyme A transporter and regulates multiorgan development in zebrafish. *Journal of Cellular Physiology,* 235**,** 151-165.

KINDERMAN, M. & SCHWAB, M. 2006. *Pyrimidines Reacting with O6-Alykylguanine-DNA Alkyltransferase.*

KLOUWER, F. C. C., BERENDSE, K., FERDINANDUSSE, S., WANDERS, R. J. A., ENGELEN, M. & POLL-THE, B. T. 2015. Zellweger spectrum disorders: clinical overview and management approach. *Orphanet Journal of Rare Diseases,* 10**,** 151.

KNIGHT, J., JIANG, J., ASSIMOS, D. G. & HOLMES, R. P. 2006. Hydroxyproline ingestion and urinary oxalate and glycolate excretion. *Kidney International,* 70**,** 1929-1934.

KOEPKE, J. I., NAKRIEKO, K.-A., WOOD, C. S., BOUCHER, K. K., TERLECKY, L. J., WALTON, P. A. & TERLECKY, S. R. 2007. Restoration of Peroxisomal Catalase Import in a Model of Human Cellular Aging. *Traffic,* 8**,** 1590-1600.

KOIKE, S. & JAHN, R. 2017. Probing and manipulating intracellular membrane traffic by microinjection of artificial vesicles. *Proceedings of the National Academy of Sciences,* 114**,** E9883-E9892.

KRAGLER, F., LANGEDER, A., RAUPACHOVA, J., BINDER, M. & HARTIG, A. 1993. Two independent peroxisomal targeting signals in catalase A of Saccharomyces cerevisiae. *The Journal of Cell Biology,* 120**,** 665-673.

KRISTJANSON, P. M., SWALLOW, B. M., ROWLANDS, G. J., KRUSKA, R. L. & DE LEEUW, P. N. 1999. Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research. *Agricultural Systems,* 59**,** 79-98.

KUNZE, M., NEUBERGER, G., MAURER-STROH, S., MA, J., ECK, T., BRAVERMAN, N., SCHMID, J. A., EISENHABER, F. & BERGER, J. 2011. Structural requirements for interaction of peroxisomal targeting signal 2 and its receptor PEX7. *The Journal of Biological Chemistry,* 286**,** 45048-45062.

KUNZE, M., PRACHAROENWATTANA, I., SMITH, S. M. & HARTIG, A. 2006. A central role for the peroxisomal membrane in glyoxylate cycle function. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1441-1452.

LAMETSCHWANDTNER, G., BROCARD, C., FRANSEN, M., VAN VELDHOVEN, P., BERGER, J. & HARTIG, A. 1998. The Difference in Recognition of Terminal Tripeptides as Peroxisomal Targeting Signal 1 between Yeast and Human Is Due to Different Affinities of Their Receptor Pex5p to the Cognate Signal and to Residues Adjacent to It. *Journal of Biological Chemistry,* 273**,** 33635-33643.

LANYON-HOGG, T., HOOPER, J., GUNN, S., WARRINER, S. L. & BAKER, A. 2014. PEX14 binding to Arabidopsis PEX5 has differential effects on PTS1 and PTS2 cargo occupancy of the receptor. *FEBS letters,* 588**,** 2223-2229.

LANYON-HOGG, T., WARRINER, S. L. & BAKER, A. 2010. Getting a camel through the eye of a needle: the import of folded proteins by peroxisomes. *Biology of the Cell,* 102**,** 245-263.

LATRUFFE, N. & VAMECQ, J. 1997. Peroxisome proliferators and peroxisome proliferator activated receptors (PPARs) as regulators of lipid metabolism. *Biochimie,* 79**,** 81-94.

LAZAROW, P. B. & DE DUVE, C. 1976. A fatty acyl-CoA oxidizing system in rat liver peroxisomes; enhancement by clofibrate, a hypolipidemic drug. *Proceedings of the National Academy of Sciences,* 73**,** 2043-2046.

LAZAROW, P. B. & FUJIKI, Y. 1985. Biogenesis of Peroxisomes. *Annual Review of Cell Biology,* 1**,** 489-530.

LEGAKIS, J. E., KOEPKE, J. I., JEDESZKO, C., BARLASKAR, F., TERLECKY, L. J., EDWARDS, H. J., WALTON, P. A. & TERLECKY, S. R. 2002. Peroxisome senescence in human fibroblasts. *Molecular Biology of the Cell,* 13**,** 4243-4255.

LEGAKIS, J. E. & TERLECKY, S. R. 2001. PTS2 Protein Import into Mammalian Peroxisomes. *Traffic,* 2**,** 252-260.

LI, H., TSUI, T. Y. & MA, W. 2015. Intracellular Delivery of Molecular Cargo Using Cell-Penetrating Peptides and the Combination Strategies. *International Journal of Molecular Sciences,* 16**,** 19518-19536.

LILL, P., HANSEN, T., WENDSCHECK, D., KLINK, B. U., JEZIOREK, T., MIEHLING, J., BENDER, J., DREPPER, F., GIRZALSKY, W., WARSCHEID, B., ERDMANN, R. & GATSOGIANNIS, C. 2020. Towards the molecular architecture of the peroxisomal receptor docking complex. *bioRxiv***,** 854497.

LINGARD, M. J., MONROE-AUGUSTUS, M. & BARTEL, B. 2009. Peroxisome-associated matrix protein degradation in *Arabidopsis. Proceedings of the National Academy of Sciences,* 106**,** 4561-4566.

LOS, G. V., ENCELL, L. P., MCDOUGALL, M. G., HARTZELL, D. D., KARASSINA, N., ZIMPRICH, C., WOOD, M. G., LEARISH, R., OHANA, R. F., URH, M., SIMPSON, D., MENDEZ, J., ZIMMERMAN, K., OTTO, P., VIDUGIRIS, G., ZHU, J., DARZINS, A., KLAUBERT, D. H., BULLEIT, R. F. & WOOD, K. V. 2008. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology,* 3**,** 373-382.

LÉON, S., GOODMAN, J. M. & SUBRAMANI, S. 2006. Uniqueness of the mechanism of protein import into the peroxisome matrix: Transport of folded, co-factor-bound and oligomeric proteins by shuttling receptors. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1552-1564.

MA, C., HAGSTROM, D., POLLEY, S. G. & SUBRAMANI, S. 2013. Redox-regulated cargo binding and release by the peroxisomal targeting signal receptor, Pex5. *The Journal of Biological Chemistry,* 288**,** 27220-27231.

MANJITHAYA, R., NAZARKO, T. Y., FARRÉ, J.-C. & SUBRAMANI, S. 2010. Molecular mechanism and physiological role of pexophagy. *FEBS letters,* 584**,** 1367-1373.

MATSUMOTO, N., TAMURA, S. & FUJIKI, Y. 2003. The pathogenic peroxin Pex26p recruits the Pex1p–Pex6p AAA ATPase complexes to peroxisomes. *Nature Cell Biology,* 5**,** 454-460.

MATSUZAKI, T. & FUJIKI, Y. 2008. The peroxisomal membrane protein import receptor Pex3p is directly transported to peroxisomes by a novel Pex19p- and Pex16p-dependent pathway. *The Journal of Cell Biology,* 183**,** 1275-1286.

MAYERHOFER, P. U., BAÑÓ-POLO, M., MINGARRO, I. & JOHNSON, A. E. 2016. Human Peroxin PEX3 Is Co-translationally Integrated into the ER and Exits the ER in Budding Vesicles. *Traffic,* 17**,** 117-130.

MAYNARD, E. L., GATTO JR., G. J. & BERG, J. M. 2004. Pex5p binding affinities for canonical and noncanonical PTS1 peptides. *Proteins: Structure, Function, and Bioinformatics,* 55**,** 856-861.

MCDOWELL, G. S. & PHILPOTT, A. 2013. Non-canonical ubiquitylation: Mechanisms and consequences. *The International Journal of Biochemistry & Cell Biology,* 45**,** 1833-1842.

MCNEW, J. A. & GOODMAN, J. M. 1994. An oligomeric protein is imported into peroxisomes in vivo. *The Journal of Cell Biology,* 127**,** 1245-1257.

MICHELS, P. A. M., BRINGAUD, F., HERMAN, M. & HANNAERT, V. 2006. Metabolic functions of glycosomes in trypanosomatids. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1463-1477.

MIHALIK, S. J., RAINVILLE, A. M. & WATKINS, P. A. 1995. Phytanic Acid α-oxidation in Rat Liver Peroxisomes. *European Journal of Biochemistry,* 232**,** 545-551.

MIYATA, N., OKUMOTO, K., MUKAI, S., NOGUCHI, M. & FUJIKI, Y. 2012. AWP1/ZFAND6 Functions in Pex5 Export by Interacting with Cys-Monoubiquitinated Pex5 and Pex6 AAA ATPase. *Traffic,* 13**,** 168-183.

MIYAZAWA, S., HASHIMOTO, T. & YOKOTA, S. 1985. Identity of Long-Chain Acyl-Coenzyme A Synthetase of Microsomes, Mitochondria, and Peroxisomes in Rat Liver1. *The Journal of Biochemistry,* 98**,** 723-733.

MOFFAT, J. G., RUDOLPH, J. & BAILEY, D. 2014. Phenotypic screening in cancer drug discovery — past, present and future. *Nature Reviews Drug Discovery,* 13**,** 588-602.

MOLLWITZ, B., BRUNK, E., SCHMITT, S., POJER, F., BANNWARTH, M., SCHILTZ, M., ROTHLISBERGER, U. & JOHNSSON, K. 2012. Directed Evolution of the Suicide Protein O6-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability. *Biochemistry,* 51**,** 986-994.

MORGAN, K. 2014. Plasmids 101: The Promoter Region – Let's Go! *blog.addgene.org* [Online]. [Accessed 04/03/2014 2020].

MOSER, H. W., LOES, D. J., MELHEM, E. R., RAYMOND, G. V., BEZMAN, L., COX, C. S. & LU, S.-E. 2000. X-Linked Adrenoleukodystrophy: Overview and Prognosis as a Function of Age and Brain Magnetic Resonance Imaging Abnormality. A Study Involving 372 Patients. *Neuropediatrics,* 31**,** 227-239.

MOTLEY, A. M., HETTEMA, E. H., KETTING, R., PLASTERK, R. & TABAK, H. F. 2000. Caenorhabditis elegans has a single pathway to target matrix proteins to peroxisomes. *EMBO reports,* 1**,** 40-46.

NEUBERGER, G., MAURER-STROH, S., EISENHABER, B., HARTIG, A. & EISENHABER, F. 2003. Motif Refinement of the Peroxisomal Targeting

Signal 1 and Evaluation of Taxon-specific Differences. *Journal of Molecular Biology,* 328**,** 567-579.

NEUHAUS, A., KOOSHAPUR, H., WOLF, J., MEYER, N. H., MADL, T., SAIDOWSKY, J., HAMBRUCH, E., LAZAM, A., JUNG, M., SATTLER, M., SCHLIEBS, W. & ERDMANN, R. 2014. A novel Pex14 protein-interacting site of human Pex5 is critical for matrix protein import into peroxisomes. *The Journal of Biological Chemistry,* 289**,** 437-448.

NIEDERHOFF, K., MEINDL-BEINKER, N. M., KERSSEN, D., PERBAND, U., SCHÄFER, A., SCHLIEBS, W. & KUNAU, W.-H. 2005. Yeast Pex14p Possesses Two Functionally Distinct Pex5p and One Pex7p Binding Sites. *Journal of Biological Chemistry,* 280**,** 35571-35578.

NITO, K., HAYASHI, M. & NISHIMURA, M. 2002. Direct Interaction and Determination of Binding Domains among Peroxisomal Import Factors in Arabidopsis thaliana. *Plant and Cell Physiology,* 43**,** 355-366.

NORDGREN, M., FRANCISCO, T., LISMONT, C., HENNEBEL, L., BREES, C., WANG, B., VAN VELDHOVEN, P. P., AZEVEDO, J. E. & FRANSEN, M. 2015. Export-deficient monoubiquitinated PEX5 triggers peroxisome removal in SV40 large T antigen-transformed mouse embryonic fibroblasts. *Autophagy,* 11**,** 1326-1340.

OELJEKLAUS, S., REINARTZ, B. S., WOLF, J., WIESE, S., TONILLO, J., PODWOJSKI, K., KUHLMANN, K., STEPHAN, C., MEYER, H. E., SCHLIEBS, W., BROCARD, C., ERDMANN, R. & WARSCHEID, B. 2012. Identification of Core Components and Transient Interactors of the Peroxisomal Importomer by Dual-Track Stable Isotope Labeling with Amino Acids in Cell Culture Analysis. *Journal of Proteome Research,* 11**,** 2567-2580.

OFMAN, R., HETTEMA, E. H., HOGENHOUT, E. M., CARUSO, U., MUIJSERS, A. O. & WANDERS, R. J. A. 1998. Acyl-CoA: Dihydroxyacetonephosphate Acyltransferase: Cloning of the Human cDNA and Resolution of the Molecular Basis in Rhizomelic Chondrodysplasia Punctata Type 2. *Human Molecular Genetics,* 7**,** 847-853.

OHANA, R. F., ENCELL, L. P., ZHAO, K., SIMPSON, D., SLATER, M. R., URH, M. & WOOD, K. V. 2009. HaloTag7: A genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification. *Protein Expression and Purification,* 68**,** 110-120.

OKAMOTO, T., KAWAGUCHI, K., WATANABE, S., AGUSTINA, R., IKEJIMA, T., IKEDA, K., NAKANO, M., MORITA, M. & IMANAKA, T. 2018. Characterization of human ATP-binding cassette protein subfamily D reconstituted into proteoliposomes. *Biochemical and Biophysical Research Communications,* 496**,** 1122-1127.

OKUMOTO, K., NODA, H. & FUJIKI, Y. 2014. Distinct Modes of Ubiquitination of Peroxisome-targeting Signal Type 1 (PTS1) Receptor Pex5p Regulate PTS1 Protein Import. *Journal of Biological Chemistry,* 289**,** 14089-14108.

OLAKU, V., MATZKE, A., MITCHELL, C., HASENAUER, S., SAKKARAVARTHI, A., PACE, G., PONTA, H. & ORIAN-ROUSSEAU, V. 2011. c-Met recruits ICAM-1 as a coreceptor to compensate for the loss of CD44 in Cd44 null mice. *Molecular Biology of the Cell,* 22**,** 2777-2786.

OPPERDOES, F. R. 1984. Localization of the initial steps in alkoxyphospholipid biosynthesis in glycosomes (microbodies) of Trypanosoma brucei. *FEBS Letters,* 169**,** 35-39.

OSMAN, G., RODRIGUEZ, J., CHAN, S. Y., CHISHOLM, J., DUNCAN, G., KIM, N., TATLER, A. L., SHAKESHEFF, K. M., HANES, J., SUK, J. S. & DIXON, J. E. 2018. PEGylated enhanced cell penetrating peptide nanoparticles for lung gene therapy. *Journal of Controlled Release,* 285**,** 35-45.

OTERA, H., OKUMOTO, K., TATEISHI, K., IKOMA, Y., MATSUDA, E., NISHIMURA, M., TSUKAMOTO, T., OSUMI, T., OHASHI, K., HIGUCHI, O. & FUJIKI, Y. 1998. Peroxisome targeting signal type 1 (PTS1) receptor is involved in import of both PTS1 and PTS2: studies with PEX5-defective CHO cell mutants. *Molecular and cellular biology,* 18**,** 388-399.

OTERA, H., SETOGUCHI, K., HAMASAKI, M., KUMASHIRO, T., SHIMIZU, N. & FUJIKI, Y. 2002. Peroxisomal targeting signal receptor Pex5p interacts with cargoes and import machinery components in a spatiotemporally differentiated manner: conserved Pex5p WXXXF/Y motifs are critical for matrix protein import. *Molecular and Cellular Biology,* 22**,** 1639-1655.

PAKER, A. M., SUNNESS, J. S., BRERETON, N. H., SPEEDIE, L. J., ALBANNA, L., DHARMARAJ, S., MOSER, A. B., JONES, R. O. & RAYMOND, G. V. 2010. Docosahexaenoic acid therapy in peroxisomal diseases: results of a double-blind, randomized trial. *Neurology,* 75**,** 826-830.

PALMIERI, L., ROTTENSTEINER, H., GIRZALSKY, W., SCARCIA, P., PALMIERI, F. & ERDMANN, R. 2001. Identification and functional reconstitution of the yeast peroxisomal adenine nucleotide transporter. *The EMBO Journal,* 20**,** 5049-5059.

PAN, D., NAKATSU, T. & KATO, H. 2013. Crystal structure of peroxisomal targeting signal-2 bound to its receptor complex Pex7p–Pex21p. *Nature Structural & Molecular Biology,* 20**,** 987.

PAP, E. H. W., DANSEN, T. B. & WIRTZ, K. W. A. 2001. Peptide-based targeting of fluorophores to peroxisomes in living cells. *Trends in Cell Biology,* 11**,** 10-12.

PARKES, J. A., LANGER, S., HARTIG, A. & BAKER, A. 2003. PTS1-independent targeting of isocitrate lyase to peroxisomes requires the PTS1 receptor Pex5p. *Molecular Membrane Biology,* 20**,** 61-69.

PATEL, S. G., SAYERS, E. J., HE, L., NARAYAN, R., WILLIAMS, T. L., MILLS, E. M., ALLEMANN, R. K., LUK, L. Y. P., JONES, A. T. & TSAI, Y.-H. 2019. Cell-penetrating peptide sequence and modification dependent uptake and subcellular distribution of green florescent protein in different cell lines. *Scientific Reports,* 9**,** 6298.

PEDROSA, A. G., FRANCISCO, T., BICHO, D., DIAS, A. F., BARROS-BARBOSA, A., HAGMANN, V., DODT, G., RODRIGUES, T. A. & AZEVEDO, J. E. 2018. Peroxisomal monoubiquitinated PEX5 interacts with the AAA ATPases PEX1 and PEX6 and is unfolded during its dislocation into the cytosol. *The Journal of Biological Chemistry,* 293**,** 11553-11563.

PEDROSA, A. G., FRANCISCO, T., FERREIRA, M. J., RODRIGUES, T. A., BARROS-BARBOSA, A. & AZEVEDO, J. E. 2019. A Mechanistic Perspective on PEX1 and PEX6, Two AAA+ Proteins of the Peroxisomal Protein Import Machinery. *International Journal of Molecular Sciences,* 20**,** 5246.

PLATTA, H. W., EL MAGRAOUI, F., BÄUMER, B. E., SCHLEE, D., GIRZALSKY, W. & ERDMANN, R. 2009. Pex2 and Pex12 Function as Protein-Ubiquitin Ligases in Peroxisomal Protein Import. *Molecular and Cellular Biology,* 29**,** 5505-5516.

PLATTA, H. W., GRUNAU, S., ROSENKRANZ, K., GIRZALSKY, W. & ERDMANN, R. 2005. Functional role of the AAA peroxins in dislocation of the cycling PTS1 receptor back to the cytosol. *Nature Cell Biology,* 7**,** 817-822.

POIRIER, Y., ERARD, N. & MACDONALD-COMBER PETÉTOT, J. 2002. Synthesis of polyhydroxyalkanoate in the peroxisome of Pichia pastoris. *FEMS Microbiology Letters,* 207**,** 97-102.

POLL-THE, B. T., ROELS, F., OGIER, H., SCOTTO, J., VAMECQ, J., SCHUTGENS, R. B., WANDERS, R. J., VAN ROERMUND, C. W., VAN WIJLAND, M. J., SCHRAM, A. W. & ET AL. 1988. A new peroxisomal disorder with enlarged peroxisomes and a specific deficiency of acyl-CoA oxidase (pseudo-neonatal adrenoleukodystrophy). *American Journal of Human Genetics,* 42**,** 422-434.

POOLE, B., LEIGHTON, F. & DE DUVE, C. 1969. The synthesis and turnover of rat liver peroxisomes. II. Turnover of peroxisome proteins. *The Journal of Cell Biology,* 41**,** 536-546.

PRACHAROENWATTANA, I. & SMITH, S. M. 2008. When is a peroxisome not a peroxisome? *Trends in Plant Science,* 13**,** 522-525.

PRIOLO, C., TANG, D., BRAHAMANDAN, M., BENASSI, B., SICINSKA, E., OGINO, S., FARSETTI, A., PORRELLO, A., FINN, S., ZIMMERMANN, J., FEBBO, P. & LODA, M. 2006. The isopeptidase USP2a protects human prostate cancer from apoptosis. *Cancer Research,* 66**,** 8625-32.

PURDUE, P. E. & LAZAROW, P. B. 1996. Targeting of human catalase to peroxisomes is dependent upon a novel COOH-terminal peroxisomal targeting sequence. *Journal of Cell Biology,* 134**,** 849-862.

PURDUE, P. E. & LAZAROW, P. B. 2001. Peroxisome Biogenesis. *Annual Review of Cell and Developmental Biology,* 17**,** 701-752.

RATBI, I., FALKENBERG, K. D., SOMMEN, M., AL-SHEQAIH, N., GUAOUA, S., VANDEWEYER, G., URQUHART, J. E., CHANDLER, K. E., WILLIAMS, S. G., ROBERTS, N. A., EL ALLOUSSI, M., BLACK, G. C., FERDINANDUSSE, S., RAMDI, H., HEIMLER, A., FRYER, A., LYNCH, S.-A., COOPER, N., ONG, K. R., SMITH, C. E. L., INGLEHEARN, C. F., MIGHELL, A. J., ELCOCK, C., POULTER, J. A., TISCHKOWITZ, M., DAVIES, S. J., SEFIANI, A., MIRONOV, A. A., NEWMAN, W. G., WATERHAM, H. R. & VAN CAMP, G. 2015. Heimler Syndrome Is Caused by Hypomorphic Mutations in the Peroxisome-Biogenesis Genes PEX1 and PEX6. *American Journal of Human Genetics,* 97**,** 535-545.

REGLINSKI, K., KEIL, M., ALTENDORF, S., WAITHE, D., EGGELING, C., SCHLIEBS, W. & ERDMANN, R. 2015. Peroxisomal Import Reduces the Proapoptotic Activity of Deubiquitinating Enzyme USP2. *PLOS ONE,* 10**,** e0140685.

REGUENGA, C., OLIVEIRA, M. E. M., GOUVEIA, A. M. M., SÁ-MIRANDA, C. & AZEVEDO, J. E. 2001. Characterization of the Mammalian Peroxisomal Import Machinery: Pex2p, Pex5p, Pex12p, AND Pex14p are Subunits of the Same Protein Assembly. *Journal of Biological Chemistry,* 276**,** 29935-29942.

REUMANN, S. 2004. Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiology,* 135**,** 783-800.

REUMANN, S., QUAN, S., AUNG, K., YANG, P., MANANDHAR-SHRESTHA, K., HOLBROOK, D., LINKA, N., SWITZENBERG, R., WILKERSON, C. G., WEBER, A. P. M., OLSEN, L. J. & HU, J. 2009. In-Depth Proteome Analysis of Arabidopsis Leaf Peroxisomes Combined with in Vivo Subcellular Targeting Verification Indicates Novel Metabolic and Regulatory Functions of Peroxisomes. *Plant Physiology,* 150**,** 125-143.

RODRIGUES, T. A., FRANCISCO, T., DIAS, A. F., PEDROSA, A. G., GROU, C. P. & AZEVEDO, J. E. 2016. A cell-free organelle-based in vitro system for studying the peroxisomal protein import machinery. *Nature Protocols,* 11**,** 2454-2469.

ROKKA, A., ANTONENKOV, V. D., SOININEN, R., IMMONEN, H. L., PIRILÄ, P. L., BERGMANN, U., SORMUNEN, R. T., WECKSTRÖM, M., BENZ, R. & HILTUNEN, J. K. 2009. Pxmp2 Is a Channel-Forming Protein in Mammalian Peroxisomal Membrane. *PLOS ONE,* 4**,** e5090.

ROSE, I. A. & WARMS, J. V. 1983. An enzyme with ubiquitin carboxy-terminal esterase activity from reticulocytes. *Biochemistry,* 22**,** 4234-7.

ROSENTHAL, M., METZL-RAZ, E., BÜRGI, J., YIFRACH, E., DRWESH, L., FADEL, A., PELEG, Y., RAPAPORT, D., WILMANNS, M., BARKAI, N., SCHULDINER, M. & ZALCKVAR, E. 2020. Uncovering targeting priority to yeast peroxisomes using an in-cell competition assay. *Proceedings of the National Academy of Sciences,* 117**,** 21432-21440.

RUSSELL, D. W. 2003. The enzymes, regulation, and genetics of bile acid synthesis. *Annual Review of Biochemistry,* 72**,** 137-74.

SADZUKA, Y., KISHI, K., HIROTA, S. & SONOBE, T. 2003. Effect of Polyethyleneglycol (PEG) Chain on Cell Uptake of PEG-Modified Liposomes. *Journal of Liposome Research,* 13**,** 157-172.

SAENGKHAE, C., LOETCHUTINAT, C. & GARNIER-SUILLEROT, A. 2003. Kinetic analysis of rhodamines efflux mediated by the multidrug resistance protein (MRP1). *Biophysical Journal,* 85**,** 2006-2014.

SAIDOWSKY, J., DODT, G., KIRCHBERG, K., WEGNER, A., NASTAINCZYK, W., KUNAU, W.-H. & SCHLIEBS, W. 2001. The Di-aromatic Pentapeptide Repeats of the Human Peroxisome Import Receptor PEX5 Are Separate High Affinity Binding Sites for the Peroxisomal Membrane Protein PEX14. *Journal of Biological Chemistry,* 276**,** 34524-34529.

SAKAMOTO, K. M., KIM, K. B., KUMAGAI, A., MERCURIO, F., CREWS, C. M. & DESHAIES, R. J. 2001. Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proceedings of the National Academy of Sciences,* 98**,** 8554-8559.

SAKUMA, S., SUITA, M., YAMAMOTO, T., MASAOKA, Y., KATAOKA, M., YAMASHITA, S., NAKAJIMA, N., SHINKAI, N., YAMAUCHI, H., HIWATARI, K.-I., HASHIZUME, A., TACHIKAWA, H., KIMURA, R., ISHIMARU, Y., KASAI, A. & MAEDA, S. 2012. Performance of cell-penetrating peptide-linked polymers physically mixed with poorly membrane-permeable molecules on cell membranes. *European Journal of Pharmaceutics and Biopharmaceutics,* 81**,** 64-73.

SALOMONS, F. A., KIEL, J. A. K. W., FABER, K. N., VEENHUIS, M. & VAN DER KLEI, I. J. 2000. Overproduction of Pex5p Stimulates Import of

Alcohol Oxidase and Dihydroxyacetone Synthase in a Hansenula polymorpha pex14Null Mutant. *Journal of Biological Chemistry,* 275**,** 12603-12611.

SANTOS, M., IMANAKA, T., SHIO, H., SMALL, G. & LAZAROW, P. 1988. Peroxisomal membrane ghosts in Zellweger syndrome--aberrant organelle assembly. *Science,* 239**,** 1536-1538.

SCHELL-STEVEN, A., STEIN, K., AMOROS, M., LANDGRAF, C., VOLKMER-ENGERT, R., ROTTENSTEINER, H. & ERDMANN, R. 2005. Identification of a Novel, Intraperoxisomal Pex14-Binding Site in Pex13: Association of Pex13 with the Docking Complex Is Essential for Peroxisomal Matrix Protein Import. *Molecular and Cellular Biology,* 25**,** 3007-3018.

SCHENBORN, E. T. & GOIFFON, V. 2000. DEAE-Dextran Transfection of Mammalian Cultured Cells. *In:* TYMMS, M. J. (ed.) *Transcription Factor Protocols.* Totowa, NJ: Humana Press.

SCHLIEBS, W., SAIDOWSKY, J., AGIANIAN, B., DODT, G., HERBERG, F. W. & KUNAU, W.-H. 1999. Recombinant Human Peroxisomal Targeting Signal Receptor PEX5: Structural Basis for Interaction of PEX5 with PEX14. *Journal of Biological Chemistry,* 274**,** 5666-5673.

SCHMIDT, F., DIETRICH, D., EYLENSTEIN, R., GROEMPING, Y., STEHLE, T. & DODT, G. 2012. The Role of Conserved PEX3 Regions in PEX19-Binding and Peroxisome Biogenesis. *Traffic,* 13**,** 1244-1260.

SCHMIDT, T. G. M., BATZ, L., BONET, L., CARL, U., HOLZAPFEL, G., KIEM, K., MATULEWICZ, K., NIERMEIER, D., SCHUCHARDT, I. & STANAR, K. 2013. Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expression and Purification,* 92**,** 54-61.

SCHMIDT, T. G. M. & SKERRA, A. 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protocols,* 2**,** 1528-1535.

SCHRADER, M. 2001. Tubulo-Reticular Clusters of Peroxisomes in Living COS-7 Cells:Dynamic Behavior and Association with Lipid Droplets1. *Journal of Histochemistry & Cytochemistry,* 49**,** 1421-1429.

SCHRADER, M., GRILLE, S., FAHIMI, H. D. & ISLINGER, M. 2013. Peroxisome Interactions and Cross-Talk with Other Subcellular Compartments in Animal Cells. *In:* DEL RÍO, L. A. (ed.) *Peroxisomes and their Key Role in Cellular Signaling and Metabolism.* Dordrecht: Springer Netherlands.

SCHREIBER, S. L., KOTZ, J. D., LI, M., AUBÉ, J., AUSTIN, C. P., REED, J. C., ROSEN, H., WHITE, E. L., SKLAR, L. A., LINDSLEY, C. W., ALEXANDER, B. R., BITTKER, J. A., CLEMONS, P. A., DE SOUZA, A., FOLEY, M. A., PALMER, M., SHAMJI, A. F., WAWER, M. J., MCMANUS, O., WU, M., ZOU, B., YU, H., GOLDEN, J. E., SCHOENEN, F. J., SIMEONOV, A., JADHAV, A., JACKSON, M. R., PINKERTON, A. B., CHUNG, T. D. Y., GRIFFIN, P. R., CRAVATT, B. F., HODDER, P. S., ROUSH, W. R., ROBERTS, E., CHUNG, D.-H., JONSSON, C. B., NOAH, J. W., SEVERSON, W. E., ANANTHAN, S., EDWARDS, B., OPREA, T. I., CONN, P. J., HOPKINS, C. R., WOOD, M. R., STAUFFER, S. R., EMMITTE, K. A. & TEAM, N. I. H. M. L. P. 2015. Advancing Biological Understanding and Therapeutics Discovery with Small-Molecule Probes. *Cell,* 161**,** 1252-1265.

SELMER, T. & PINKENBURG, O. 2008. *Method of cloning at least one nucleic acid molecule of interest using type IIS restriction endonucleases, and corresponding cloning vectors, kits and system using type IIS restriction endonucleases*. Germany patent application PCT/EP2008/051396.

SHIMOZAWA, N., TSUKAMOTO, T., NAGASE, T., TAKEMOTO, Y., KOYAMA, N., SUZUKI, Y., KOMORI, M., OSUMI, T., JEANNETTE, G., WANDERS, R. J. A. & KONDO, N. 2004. Identification of a new complementation group of the peroxisome biogenesis disorders and PEX14 as the mutated gene. *Human Mutation,* 23**,** 552-558.

SHIOZAWA, K., KONAREV, P. V., NEUFELD, C., WILMANNS, M. & SVERGUN, D. I. 2009. Solution structure of human Pex5.Pex14.PTS1 protein complexes obtained by small angle X-ray scattering. *The Journal of Biological Chemistry,* 284**,** 25334-25342.

SINGH, V., WANG, S. & KOOL, E. T. 2013. Genetically Encoded Multispectral Labeling of Proteins with Polyfluorophores on a DNA Backbone. *Journal of the American Chemical Society,* 135**,** 6184-6191.

SKOULDING, N. S., CHOWDHARY, G., DEUS, M. J., BAKER, A., REUMANN, S. & WARRINER, S. L. 2015. Experimental Validation of Plant Peroxisomal Targeting Prediction Algorithms by Systematic Comparison of In Vivo Import Efficiency and In Vitro PTS1 Binding Affinity. *Journal of Molecular Biology,* 427**,** 1085-1101.

SONG, H., LIU, Y., XIONG, L., LI, Y., YANG, N. & WANG, Q. 2013. Design, Synthesis, and Insecticidal Evaluation of New Pyrazole Derivatives Containing Imine, Oxime Ether, Oxime Ester, and Dihydroisoxazoline Groups Based on the Inhibitor Binding Pocket of Respiratory Complex I. *Journal of Agricultural and Food Chemistry,* 61**,** 8730-8736.

SOUTH, S. T., BAUMGART, E. & GOULD, S. J. 2001. Inactivation of the endoplasmic reticulum protein translocation factor, Sec61p, or its homolog, Ssh1p, does not affect peroxisome biogenesis. *Proceedings of the National Academy of Sciences,* 98**,** 12027-12031.

SRIKUN, D., ALBERS, A. E., NAM, C. I., IAVARONE, A. T. & CHANG, C. J. 2010. Organelle-Targetable Fluorescent Probes for Imaging Hydrogen Peroxide in Living Cells via SNAP-Tag Protein Labeling. *Journal of the American Chemical Society,* 132**,** 4455-4465.

STANLEY, W. A., FILIPP, F. V., KURSULA, P., SCHÜLLER, N., ERDMANN, R., SCHLIEBS, W., SATTLER, M. & WILMANNS, M. 2006. Recognition of a Functional Peroxisome Type 1 Target by the Dynamic Import Receptor Pex5p. *Molecular Cell,* 24**,** 653-663.

STEINBERG, D., HERNDON, J. H., UHLENDORF, B. W., MIZE, C. E., AVIGAN, J. & MILNE, G. W. A. 1967. Refsum's Disease: Nature of the Enzyme Defect. *Science,* 156**,** 1740-1742.

STUDIER, F. W. & MOFFATT, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology,* 189**,** 113-130.

SUGIURA, A., MATTIE, S., PRUDENT, J. & MCBRIDE, H. M. 2017. Newly born peroxisomes are a hybrid of mitochondrial and ER-derived pre-peroxisomes. *Nature,* 542**,** 251-254.

SUZUKI, Y., IAI, M., KAMEI, A., TANABE, Y., CHIDA, S., YAMAGUCHI, S., ZHANG, Z., TAKEMOTO, Y., SHIMOZAWA, N. & KONDO, N. 2002. Peroxisomal acyl CoA oxidase deficiency. *The Journal of Pediatrics,* 140**,** 128-130.

SWINKELS, B. W., GOULD, S. J. & SUBRAMANI, S. 1992. Targeting efficiencies of various permutations of the consensus C-terminal tripeptide peroxisomal targeting signal. *FEBS Letters,* 305**,** 133-136.

TITORENKO, V. I. & RACHUBINSKI, R. A. 1998. Mutants of the yeast Yarrowia lipolytica defective in protein exit from the endoplasmic reticulum are also defective in peroxisome biogenesis. *Molecular and Cellular Biology,* 18**,** 2789-2803.

TITORENKO, V. I. & RACHUBINSKI, R. A. 2001. Dynamics of peroxisome assembly and function. *Trends in Cell Biology,* 11**,** 22-29.

TITUS, D. E. & BECKER, W. M. 1985. Investigation of the glyoxysome-peroxisome transition in germinating cucumber cotyledons using double-label immunoelectron microscopy. *The Journal of Cell Biology,* 101**,** 1288-1299.

TORO, A. A., ARAYA, C. A., CÓRDOVA, G. J., ARREDONDO, C. A., CÁRDENAS, H. G., MORENO, R. E., VENEGAS, A., KOENIG, C. S., CANCINO, J., GONZALEZ, A. & SANTOS, M. J. 2009. Pex3p-dependent peroxisomal biogenesis initiates in the endoplasmic reticulum of human fibroblasts. *Journal of Cellular Biochemistry,* 107**,** 1083-1096.

TOURE, M. & CREWS, C. M. 2016. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angewandte Chemie International Edition,* 55**,** 1966-1973.

VALM, A. M., COHEN, S., LEGANT, W. R., MELUNIS, J., HERSHBERG, U., WAIT, E., COHEN, A. R., DAVIDSON, M. W., BETZIG, E. & LIPPINCOTT-SCHWARTZ, J. 2017. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature,* 546**,** 162-167.

VAN DEN BRINK, D. M. & WANDERS, R. J. A. 2006. Phytanic acid: production from phytol, its breakdown and role in human disease. *Cellular and Molecular Life Sciences,* 63**,** 1752.

VAN DER KLEI, I. J., HILBRANDS, R. E., KIEL, J. A., RASMUSSEN, S. W., CREGG, J. M. & VEENHUIS, M. 1998. The ubiquitin-conjugating enzyme Pex4p of Hansenula polymorpha is required for efficient functioning of the PTS1 import machinery. *The EMBO journal,* 17**,** 3608-3618.

VAN ROERMUND, C. W. T., DRISSEN, R., VAN DEN BERG, M., IJLST, L., HETTEMA, E. H., TABAK, H. F., WATERHAM, H. R. & WANDERS, R. J. A. 2001. Identification of a Peroxisomal ATP Carrier Required for Medium-Chain Fatty Acid β-Oxidation and Normal Peroxisome Proliferation in *Saccharomyces cerevisiae. Molecular and Cellular Biology,* 21**,** 4321-4329.

VAN ROERMUND, C. W. T., IJLST, L., WAGEMANS, T., WANDERS, R. J. A. & WATERHAM, H. R. 2014. A role for the human peroxisomal half-transporter ABCD3 in the oxidation of dicarboxylic acids. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1841**,** 563-568.

VAN ROERMUND, C. W. T., VISSER, W. F., IJLST, L., VAN CRUCHTEN, A., BOEK, M., KULIK, W., WATERHAM, H. R. & WANDERS, R. J. A. 2008. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl–CoA esters. *The FASEB Journal,* 22**,** 4201-4208.

VAN ROERMUND, C. W. T., VISSER, W. F., IJLST, L., WATERHAM, H. R. & WANDERS, R. J. A. 2011. Differential substrate specificities of human

ABCD1 and ABCD2 in peroxisomal fatty acid β-oxidation. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1811**,** 148-152.

VAN VELDHOVEN, P. P., DE SCHRYVER, E., YOUNG, S. G., ZWIJSEN, A., FRANSEN, M., ESPEEL, M., BAES, M. & VAN AEL, E. 2020. Slc25a17 Gene Trapped Mice: PMP34 Plays a Role in the Peroxisomal Degradation of Phytanic and Pristanic Acid. *Frontiers in Cell and Developmental Biology,* 8.

VAN VELDHOVEN, P. P., JUST, W. W. & MANNAERTS, G. P. 1987. Permeability of the peroxisomal membrane to cofactors of beta-oxidation. Evidence for the presence of a pore-forming protein. *Journal of Biological Chemistry,* 262**,** 4310-4318.

VAN DER ZAND, A., GENT, J., BRAAKMAN, I. & TABAK, HENK F. 2012. Biochemically Distinct Vesicles from the Endoplasmic Reticulum Fuse to Form Peroxisomes. *Cell,* 149**,** 397-409.

VEENHUIS, M., MATEBLOWSKI, M., KUNAU, W. H. & HARDER, W. 1987. Proliferation of microbodies in Saccharomyces cerevisiae. *Yeast,* 3**,** 77-84.

VERHOEVEN, N. M., SCHOR, D. S. M., TEN BRINK, H. J., WANDERS, R. J. A. & JAKOBS, C. 1997. Resolution of the Phytanic Acid α-Oxidation Pathway: Identification of Pristanal as Product of the Decarboxylation of 2-Hydroxyphytanoyl-CoA. *Biochemical and Biophysical Research Communications,* 237**,** 33-36.

VISSER, W. F., VAN ROERMUND, C. W. T., WATERHAM, H. R. & WANDERS, R. J. A. 2002. Identification of human PMP34 as a peroxisomal ATP transporter. *Biochemical and Biophysical Research Communications,* 299**,** 494-497.

VIVÈS, E., BRODIN, P. & LEBLEU, B. 1997. A Truncated HIV-1 Tat Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus. *Journal of Biological Chemistry,* 272**,** 16010-16017.

VOSS, S. & SKERRA, A. 1997. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein Engineering, Design and Selection,* 10**,** 975-982.

WALTON, P. A., BREES, C., LISMONT, C., APANASETS, O. & FRANSEN, M. 2017. The peroxisomal import receptor PEX5 functions as a stress sensor, retaining catalase in the cytosol in times of oxidative stress. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1864**,** 1833-1843.

WALTON, P. A., HILL, P. E. & SUBRAMANI, S. 1995. Import of stably folded proteins into peroxisomes. *Molecular Biology of the Cell,* 6**,** 675-683.

WANDERS, R. J. & BRITES, P. 2010. Biosynthesis of ether-phospholipids including plasmalogens, peroxisomes and human disease: new insights into an old problem. *Clinical Lipidology,* 5**,** 379-386.

WANDERS, R. J. A., DEKKER, C., HOVARTH, V. A. P., SCHUTGENS, R. B. H., TAGER, J. M., VAN LAER, P. & LECOUTERE, D. 1994. Human alkyldihydroxyacetonephosphate synthase deficiency: A new peroxisomal disorder. *Journal of Inherited Metabolic Disease,* 17**,** 315-318.

WANDERS, R. J. A., FERDINANDUSSE, S., BRITES, P. & KEMP, S. 2010. Peroxisomes, lipid metabolism and lipotoxicity. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1801**,** 272-280.

WANDERS, R. J. A., SCHUMACHER, H., HEIKOOP, J., SCHUTGENS, R. B. H. & TAGER, J. M. 1992a. Human dihydroxyacetonephosphate acyltransferase deficiency: A new peroxisomal disorder. *Journal of Inherited Metabolic Disease,* 15**,** 389-391.

WANDERS, R. J. A., SCHUTGENS, R. B. H., SCHRAKAMP, G., VAN DEN BOSCH, H., TAGER, J. M., SCHRAM, A. W., HASHIMOTO, T., POLL-THÉ, B. T. & SAUDUBRAU, J. M. 1986. Infantile Refsum disease: deficiency of catalase-containing particles (peroxisomes), alkyldihydroxyacetone phosphate synthase and peroxisomal β-oxidation enzyme proteins. *European Journal of Pediatrics,* 145**,** 172-175.

WANDERS, R. J. A., VAN ROERMUND, C. W. T., BRUL, S., SCHUTGENS, R. B. H. & TAGER, J. M. 1992b. Bifunctional enzyme deficiency: Identification of a new type of peroxisomal disorder in a patient with an impairment in peroxisomal β-oxidation of unknown aetiology by means of complementation analysis. *Journal of Inherited Metabolic Disease,* 15**,** 385-388.

WANDERS, R. J. A., VAN ROERMUND, C. W. T., VAN WIJLAND, M. J. A., SCHUTGENS, R. B. H., VAN DEN BOSCH, H., SCHRAM, A. W. & TAGER, J. M. 1988. Direct demonstration that the deficient oxidation of very long chain fatty acids in X-linked adrenoleukodystrophy is due to an impaired ability of peroxisomes to activate very long chain fatty acids. *Biochemical and Biophysical Research Communications,* 153**,** 618-624.

WANDERS, R. J. A. & WATERHAM, H. R. 2006a. Biochemistry of Mammalian Peroxisomes Revisited. *Annual Review of Biochemistry,* 75**,** 295-332.

WANDERS, R. J. A. & WATERHAM, H. R. 2006b. Peroxisomal disorders: The single peroxisomal enzyme deficiencies. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1763**,** 1707-1720.

WANDERS, R. J. A., WATERHAM, H. R. & FERDINANDUSSE, S. 2016. Metabolic Interplay between Peroxisomes and Other Subcellular Organelles Including Mitochondria and the Endoplasmic Reticulum. *Frontiers in Cell and Developmental Biology,* 3.

WATERHAM, H. R., FERDINANDUSSE, S. & WANDERS, R. J. A. 2016. Human disorders of peroxisome metabolism and biogenesis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1863**,** 922-933.

WHITMORE, L. & WALLACE, B. A. 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research,* 32**,** W668-W673.

WIERZBICKI, A. S., LLOYD, M. D., SCHOFIELD, C. J., FEHER, M. D. & GIBBERD, F. B. 2002. Refsum's disease: a peroxisomal disorder affecting phytanic acid α-oxidation. *Journal of Neurochemistry,* 80**,** 727-735.

WILLIAMS, C. 2014. Going against the flow: A case for peroxisomal protein export. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1843**,** 1386-1392.

WILLIAMS, C., BENER AKSAM, E., GUNKEL, K., VEENHUIS, M. & VAN DER KLEI, I. J. 2012. The relevance of the non-canonical PTS1 of peroxisomal catalase. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1823**,** 1133-1141.

WILLIAMS, C., VAN DEN BERG, M. & DISTEL, B. 2005. Saccharomyces cerevisiae Pex14p contains two independent Pex5p binding sites, which are both essential for PTS1 protein import. *FEBS Letters,* 579**,** 3416-3420.

WILLIAMS, C., VAN DEN BERG, M., SPRENGER, R. R. & DISTEL, B. 2007. A Conserved Cysteine Is Essential for Pex4p-dependent Ubiquitination of the Peroxisomal Import Receptor Pex5p. *Journal of Biological Chemistry,* 282**,** 22534-22543.

WILLIAMS, C. & VAN DER KLEI, I. J. 2013. Pexophagy-linked degradation of the peroxisomal membrane protein Pex3p involves the ubiquitin–proteasome system. *Biochemical and Biophysical Research Communications,* 438**,** 395-401.

WILLIAMS, E. L., ACQUAVIVA, C., AMOROSO, A., CHEVALIER, F., COULTER-MACKIE, M., MONICO, C. G., GIACHINO, D., OWEN, T., ROBBIANO, A., SALIDO, E., WATERHAM, H. & RUMSBY, G. 2009. Primary hyperoxaluria type 1: update and additional mutation analysis of the AGXT gene. *Human Mutation,* 30**,** 910-917.

WOLF, J., SCHLIEBS, W. & ERDMANN, R. 2010. Peroxisomes as dynamic organelles: peroxisomal matrix protein import. *The FEBS Journal,* 277**,** 3268-3278.

WOODWARD, A. W. & BARTEL, B. 2005. The Arabidopsis peroxisomal targeting signal type 2 receptor PEX7 is necessary for peroxisome function and dependent on PEX5. *Molecular Biology of the Cell,* 16**,** 573-583.

WRIESSNEGGER, T., AUGUSTIN, P., ENGLEDER, M., LEITNER, E., MÜLLER, M., KALUZNA, I., SCHÜRMANN, M., MINK, D., ZELLNIG, G., SCHWAB, H. & PICHLER, H. 2014. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of Pichia pastoris. *Metabolic Engineering,* 24**,** 18-29.

YANG, Z. & KLIONSKY, D. J. 2010. Eaten alive: a history of macroautophagy. *Nature Cell Biology,* 12**,** 814-822.

YOKOTA, S., ODA, T. & FAHIMI, H. D. 2001. The Role of 15-lipoxygenase in Disruption of the Peroxisomal Membrane and in Programmed Degradation of Peroxisomes in Normal Rat Liver. *Journal of Histochemistry & Cytochemistry,* 49**,** 613-621.

ZHANG, J., TRIPATHI, D. N., JING, J., ALEXANDER, A., KIM, J., POWELL, R. T., DERE, R., TAIT-MULDER, J., LEE, J.-H., PAULL, T. T., PANDITA, R. K., CHARAKA, V. K., PANDITA, T. K., KASTAN, M. B. & WALKER, C. L. 2015. ATM functions at the peroxisome to induce pexophagy in response to ROS. *Nature Cell biology,* 17**,** 1259-1269.

ZHOU, Y. J., BUIJS, N. A., ZHU, Z., GÓMEZ, D. O., BOONSOMBUTI, A., SIEWERS, V. & NIELSEN, J. 2016. Harnessing Yeast Peroxisomes for Biosynthesis of Fatty-Acid-Derived Biofuels and Chemicals with Relieved Side-Pathway Competition. *Journal of the American Chemical Society,* 138**,** 15368-15377.

ZIENTARA-RYTTER, K. & SUBRAMANI, S. 2016. Autophagic degradation of peroxisomes in mammals. *Biochemical Society Transactions,* 44**,** 431-440.

ZOLMAN, B. K., MONROE-AUGUSTUS, M., SILVA, I. D. & BARTEL, B. 2005. Identification and Functional Characterization of *Arabidopsis* PEROXIN4 and the Interacting Protein PEROXIN22. *The Plant Cell,* 17**,** 3422-3435.

Appendix A

A.1 *Hs*-His$_6$-PEX5C

A.1.1 *Hs*-His$_6$-PEX5C plasmid map

A.1.1 *Hs*-His$_6$-PEX5C DNA Sequence

ATGGGCAGCAGCCATCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGGAGTTTGAACGAGCCAAGTCAGCTATAGAGTCTGATGTCGATTTCT
GGGACAAGTTGCAGGCAGAGTTGGAGGAGATGGCAAAACGGGATGCTGAGGCC
CACCCCTGGCTTTCTGACTATGATGACCTTACGTCAGCTACCTATGATAAGGGGTA
CCAGTTTGAGGAGGAGAACCCCTTGCGTGATCACCCTCAGCCTTTTGAAGAAGGG
CTGCGGCGCCTTCAGGAGGGGGACCTGCCAAATGCTGTGCTGCTTTTTGAGGCA
GCTGTGCAGCAGGATCCTAAGCACATGGAAGCTTGGCAGTATCTGGGTACCACC
CAGGCAGAGAATGAACAAGAACTATTAGCCATCAGTGCATTGCGGAGGTGTCTGG
AGCTAAAGCCAGATAACCAGACAGCACTGATGGCGCTGGCTGTGAGCTTCACCAA
CGAGTCCCTGCAGCGACAGGCCTGTGAAATCCTACGAGACTGGCTGCGGTACAC
ACCAGCCTATGCCCATCTGGTGACACCTGCTGAAGAAGGGGCTGGTGGGGCAGG
ACTGGGCCCCAGCAAGCGTATCCTGGGATCTCTCTTGTCTGACTCCCTGTTTCTT
GAAGTGAAAGAGCTCTTCCTGGCAGCTGTGCGGCTGGACCCTACCTCCATTGACC

CTGATGTGCAGTGTGGCTTGGGAGTCCTTTTCAACCTGAGTGGGGAGTATGACAA
GGCCGTGGACTGCTTCACAGCTGCCCTCAGCGTTCGTCCCAATGACTATTTGCTG
TGGAATAAGCTAGGCGCCACCCTGGCCAATGGAAACCAGAGTGAAGAAGCAGTA
GCTGCGTACCGCCGGGCCCTCGAGCTCCAGCCTGGCTATATCCGGTCCCGCTAT
AACCTGGGCATCAGCTGCATCAACCTCGGGGCTCACCGGGAGGCTGTGGAGCAC
TTTCTGGAGGCCCTGAACATGCAGAGGAAAAGCCGGGGCCCCCGGGGTGAAGG
AGGTGCCATGTCGGAGAACATCTGGAGCACCCTGCGTTTGGCATTGTCTATGTTA
GGCCAGAGCGATGCCTATGGGGCAGCCGACGCGCGGGATCTGTCCACCCTCCTA
ACTATGTTTGGCCTGCCCCAG<mark style="background-color:red">TGA</mark>

Green: Start codons

Red: Stop codon

Cyan: His₆ tag

Magenta: Thrombin recognition sequence

A.1.2 *Hs*-His₆-PEX5C protein sequence

MGSS<mark style="background-color:cyan">HHHHHH</mark>SSG<mark style="background-color:magenta">LVPRGS</mark>HMEFERAKSAIESDVDFWDKLQAELEEMAKRDAEAHP
WLSDYDDLTSATYDKGYQFEEENPLRDHPQPFEEGLRRLQEGDLPNAVLLFEAAVQ
QDPKHMEAWQYLGTTQAENEQELLAISALRRCLELKPDNQTALMALAVSFTNESLQR
QACEILRDWLRYTPAYAHLVTPAEEGAGGAGLGPSKRILGSLLSDSLFLEVKELFLAAV
RLDPTSIDPDVQCGLGVLFNLSGEYDKAVDCFTAALSVRPNDYLLWNKLGATLANGN
QSEEAVAAYRRALELQPGYIRSRYNLGISCINLGAHREAVEHFLEALNMQRKSRGPRG
EGGAMSENIWSTLRLALSMLGQSDAYGAADARDLSTLLTMFGLPQ

Cyan: His₆ tag

Magenta: Thrombin recognition sequence

Arrow: site of thrombin cleavage

Number of amino acids: 387

Molecular weight: 42982.17 Da

Formula: $C_{1889}H_{2938}N_{534}O_{589}S_{13}$

Extinction coefficient: 51130 $M^{-1}cm^{-1}$

A.1.3 *Hs*-His₆-PEX5C protein mass spectrum
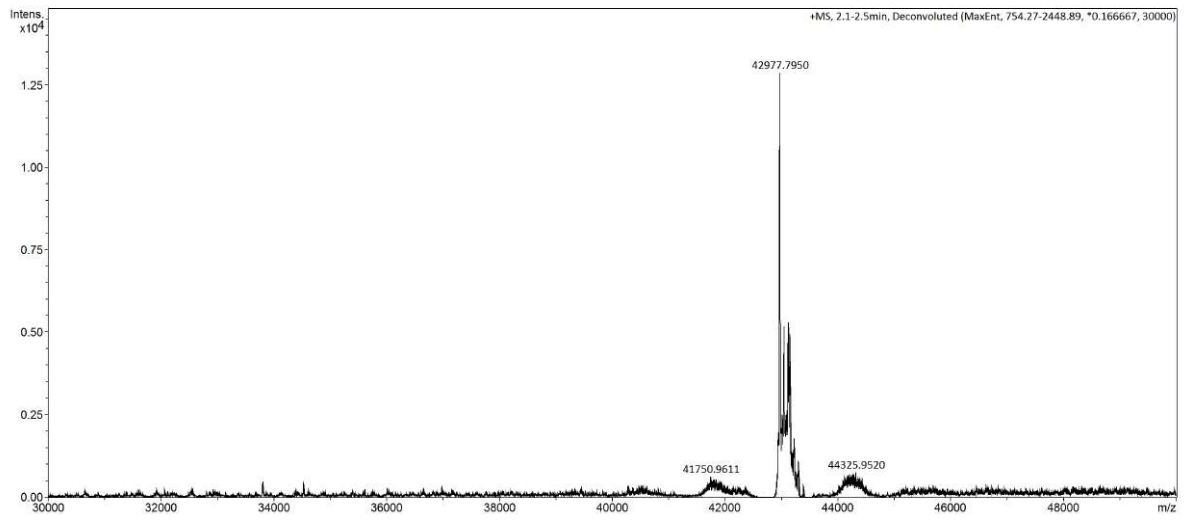
Expected Mass: 42982.17 Da, Observed Mass: 42977.8 Da

A.2 *At*-His₆-PEX5C

A.2.1 *At*-His₆-PEX5C plasmid map
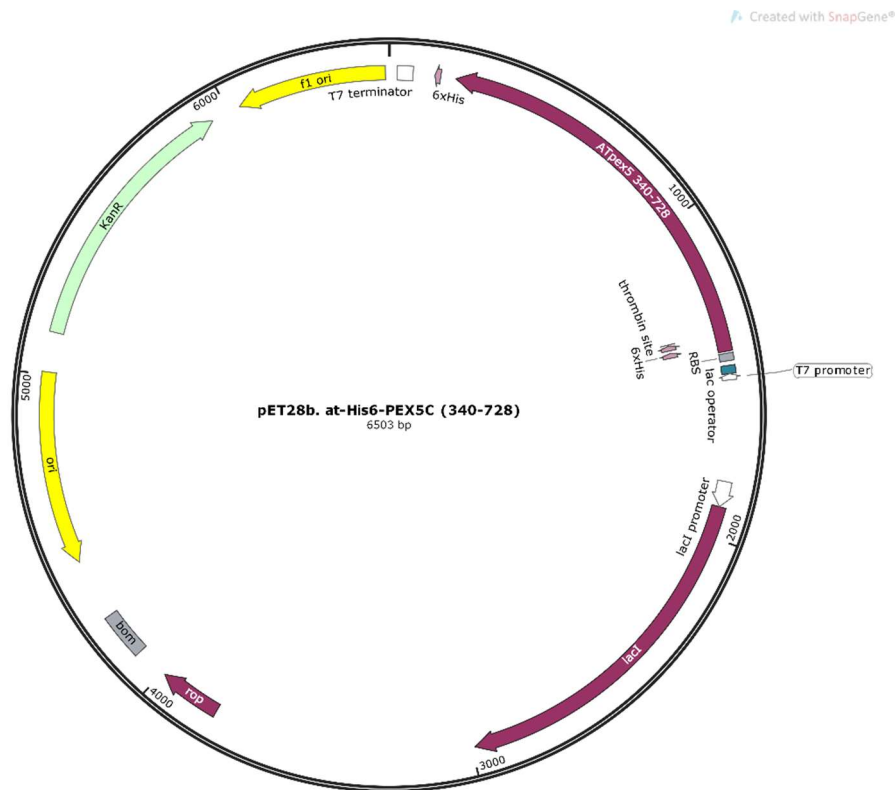
A.2.2 *At*-His₆-PEX5C DNA Sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGCAAGCTTCAGCCCCCGGGGAATGGGCTACTGAATATGAACAGCAGT
ATCTGGGGCCACCAAGTTGGGCTGATCAATTTGCAAATGAGAAACTTTCACATGG
ACCAGAACAGTGGGCTGATGAGTTTGCTTCCGGGAGAGGACAGCAAGAAACAGC
TGAGGACCAATGGGTTAATGAGTTTTCAAAGTTGAATGTTGATGACTGGATAGATG
AATTTGCTGAAGGTCCCGTGGGTGATAGTTCAGCTGATGCATGGGCAAATGCTTA
CGATGAGTTTCTGAATGAGAAAAATGCTGGAAAACAAACCAGTGGTGTCTACGTCT
TCTCTGACATGAATCCTTATGTGGGTCACCCTGAACCTATGAAAGAAGGGCAAGA
ATTGTTTCGAAAAGGACTTCTGAGTGAAGCAGCGCTTGCTCTAGAAGCTGAGGTT
ATGAAAAACCCTGAGAATGCTGAAGGTTGGAGATTACTTGGGGTCACACACGCAG
AGAACGATGATGATCAACAGGCAATAGCTGCAATGATGCGTGCACAGGAGGCTGA
TCCCACAAATCTAGAGGTGCTTCTTGCGCTTGGTGTGAGTCATACCAACGAGTTA
GAGCAAGCAACTGCTTTGAAATATCTATATGGATGGCTGCGAAATCACCCAAAGTA
TGGAGCAATTGCGCCTCCGGAGCTAGCGGATTCTTTGTACCATGCTGATATTGCT
AGATTATTCAATGAAGCTTCTCAGTTGAATCCTGAGGACGCCGATGTGCATATAGT
GTTGGGCGTGCTCTACAATCTGTCGAGAGAGTTCGATAGAGCAATCACATCCTTC

CAAACAGCATTACAACTAAAACCAAACGATTATTCTCTGTGGAATAAGCTAGGTGC
AACGCAAGCCAACAGTGTCCAGAGTGCTGATGCCATATCTGCTTATCAACAGGCT
CTAGATTTAAAACCAAATTATGTTCGTGCTTGGGCAAACATGGGAATCAGTTACGC
AAACCAGGGGATGTACAAGAATCAATCCCGTATTATGTCCGTGCCCTTGCGATG
AATCCTAAAGCTGATAACGCATGGCAATACTTGAGACTCTCGTTAAGTTGTGCATC
AAGGCAAGACATGATAGAAGCTTGTGAGTCAAGGAATCTCGATCTCTTGCAGAAA
GAATTCCCGCTG<mark style="background-color:red">TGA</mark>

Green: Start codons

Red: Stop codon

Cyan: His$_6$ tag

Magenta: Thrombin recognition sequence

A.2.3 *At*-His$_6$-PEX5C protein sequence

MGSS<mark style="background-color:cyan">HHHHHH</mark>SSG<mark style="background-color:magenta">LVPRGS</mark>HMQASAPGEWATEYEQQYLGPPSWADQFANEKLSH
GPEQWADEFASGRGQQETAEDQWVNEFSKLNVDDWIDEFAEGPVGDSSADAWAN
AYDEFLNEKNAGKQTSGVYVFSDMNPYVGHPEPMKEGQELFRKGLLSEAALALEAEV
MKNPENAEGWRLLGVTHAENDDDQQAIAAMMRAQEADPTNLEVLLALGVSHTNELE
QATALKYLYGWLRNHPKYGAIAPPELADSLYHADIARLFNEASQLNPEDADVHIVLGVL
YNLSREFDRAITSFQTALQLKPNDYSLWNKLGATQANSVQSADAISAYQQALDLKPNY
VRAWANMGISYANQGMYKESIPYYVRALAMNPKADNAWQYLRLSLSCASRQDMIEA
CESRNLDLLQKEFPL

Cyan: His$_6$ tag

Magenta: Thrombin recognition sequence

Arrow: site of thrombin cleavage

Number of amino acids: 408

Molecular weight: 45580.35 Da

Formula: $C_{2016}H_{3038}N_{554}O_{635}S_{12}$

Extinction coefficient: 87445 $M^{-1}cm^{-1}$

A.2.4 *At*-His₆-PEX5C protein mass spectrum
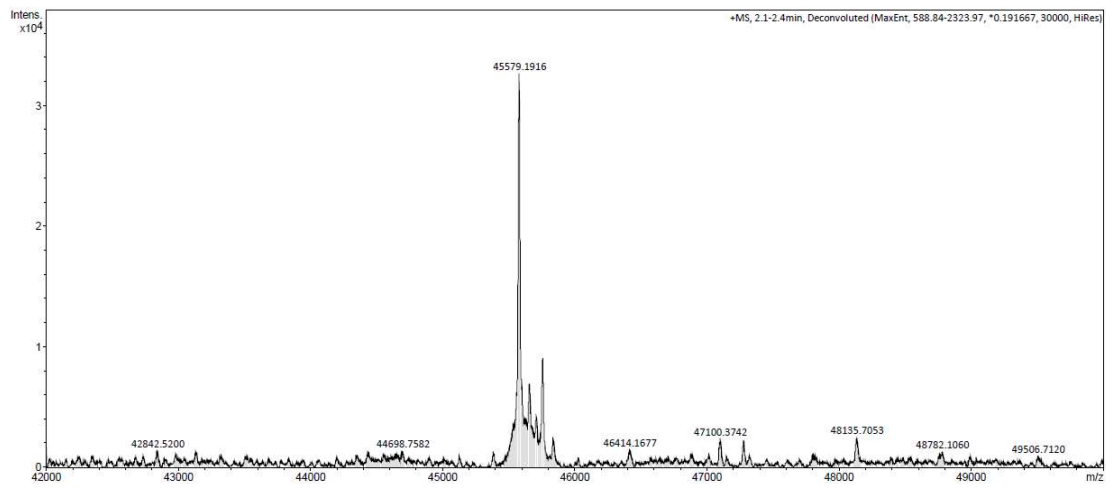
Expected mass: 45,580.3 Da. Observed mass: 45,579.2 Da

A.3 *At*-His₆-PEX5C F606C

A.3.1 *At*-His₆-PEX5C F606C plasmid map

pET28b. at-His6-PEX5C (340-728)
6503 bp
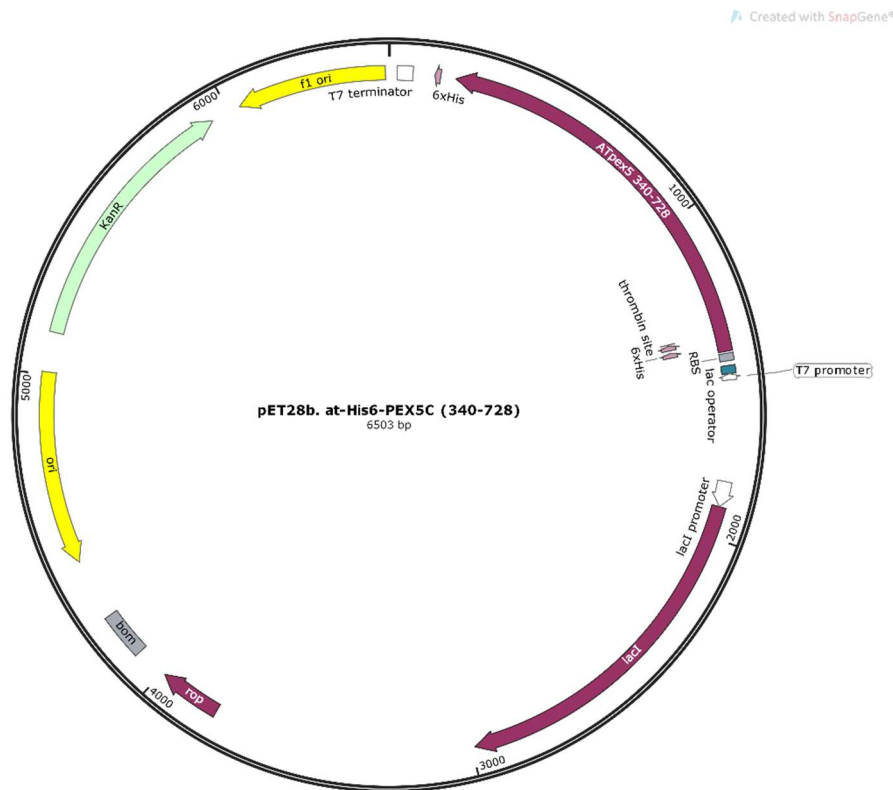
A.3.2 *At*-His₆-PEX5C F606C DNA Sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGCAAGCTTCAGCCCCCGGGGAATGGGCTACTGAATATGAACAGCAGT
ATCTGGGGCCACCAAGTTGGGCTGATCAATTTGCAAATGAGAAACTTTCACATGG
ACCAGAACAGTGGGCTGATGAGTTTGCTTCCGGGAGAGGACAGCAAGAAACAGC
TGAGGACCAATGGGTTAATGAGTTTTCAAAGTTGAATGTTGATGACTGGATAGATG
AATTTGCTGAAGGTCCCGTGGGTGATAGTTCAGCTGATGCATGGGCAAATGCTTA
CGATGAGTTTCTGAATGAGAAAAATGCTGGAAAACAAACCAGTGGTGTCTACGTCT
TCTCTGACATGAATCCTTATGTGGGTCACCCTGAACCTATGAAAGAAGGGCAAGA
ATTGTTTCGAAAGGACTTCTGAGTGAAGCAGCGCTTGCTCTAGAAGCTGAGGTT
ATGAAAAACCCTGAGAATGCTGAAGGTTGGAGATTACTTGGGGTCACACACGCAG
AGAACGATGATGATCAACAGGCAATAGCTGCAATGATGCGTGCACAGGAGGCTGA
TCCCACAAATCTAGAGGTGCTTCTTGCGCTTGGTGTGAGTCATACCAACGAGTTA
GAGCAAGCAACTGCTTTGAAATATCTATATGGATGGCTGCGAAATCACCCAAAGTA
TGGAGCAATTGCGCCTCCGGAGCTAGCGGATTCTTTGTACCATGCTGATATTGCT
AGATTATTCAATGAAGCTTCTCAGTTGAATCCTGAGGACGCCGATGTGCATATAGT
GTTGGGCGTGCTCTACAATCTGTCGAGAGAGTGCGATAGAGCAATCACATCCTTC

CAAACAGCATTACAACTAAAACCAAACGATTATTCTCTGTGGAATAAGCTAGGTGC
AACGCAAGCCAACAGTGTCCAGAGTGCTGATGCCATATCTGCTTATCAACAGGCT
CTAGATTTAAAACCAAATTATGTTCGTGCTTGGGCAAACATGGGAATCAGTTACGC
AAACCAGGGGATGTACAAAGAATCAATCCCGTATTATGTCCGTGCCCTTGCGATG
AATCCTAAAGCTGATAACGCATGGCAATACTTGAGACTCTCGTTAAGTTGTGCATC
AAGGCAAGACATGATAGAAGCTTGTGAGTCAAGGAATCTCGATCTCTTGCAGAAA
GAATTCCCGCTG**TGA**

Green: Start codons

Red: Stop codon

Cyan: His6 tag

Magenta: Thrombin recognition sequence

A.3.3 *At*-His6-PEX5C F606C protein sequence

MGSSHHHHHHSSGLVPRGSHMQASAPGEWATEYEQQYLGPPSWADQFANEKLSH
GPEQWADEFASGRGQQETAEDQWVNEFSKLNVDDWIDEFAEGPVGDSSADAWAN
AYDEFLNEKNAGKQTSGVYVFSDMNPYVGHPEPMKEGQELFRKGLLSEAALALEAEV
MKNPENAEGWRLLGVTHAENDDDQQAIAAMMRAQEADPTNLEVLLALGVSHTNELE
QATALKYLYGWLRNHPKYGAIAPPELADSLYHADIARLFNEASQLNPEDADVHIVLGVL
YNLSRECDRAITSFQTALQLKPNDYSLWNKLGATQANSVQSADAISAYQQALDLKPNY
VRAWANMGISYANQGMYKESIPYYVRALAMNPKADNAWQYLRLSLSCASRQDMIEA
CESRNLDLLQKEFPL

Cyan: His6 tag

Magenta: Thrombin recognition sequence

Arrow: site of thrombin cleavage

Number of amino acids: 408

Molecular weight: 45536.31Da

Formula: $C_{2010}H_{3034}N_{554}O_{635}S_{13}$

Extinction coefficient: 87445 $M^{-1}cm^{-1}$

A.3.4 *At*-His₆-PEX5C F606C protein mass spectrum
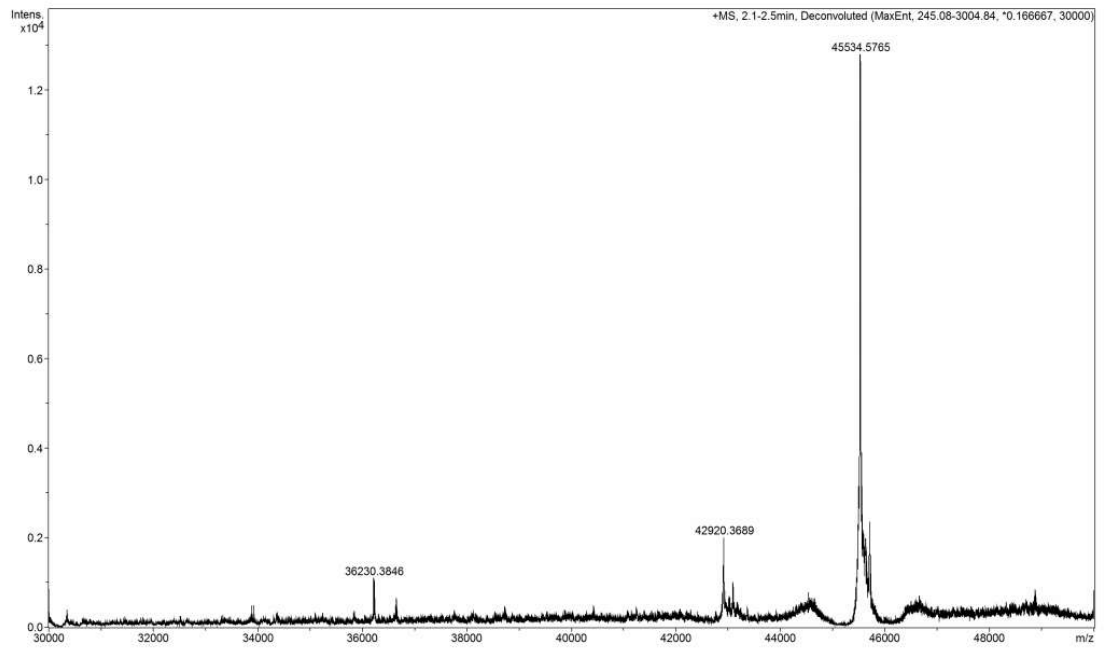
Expected mass: 45,536.31 Da. Observed mass: 45,534.5765 Da

A.4 SNAP-Strep

A.4.1 SNAP-strep Plasmid map

pET12b SNAP-twinstrep
6028 bp

A.4.2 SNAP-strep DNA sequence

ATGAAGCTTGGTCCCGGTTCCGACAAGGACTGCGAGATGAAGCGTACCACCCTT
GACTCCCCCCTTGGTAAGTTAGAGCTTTCCGGTTGCGAGCAAGGTCTTCACGAGA
TTATTTTCCTTGGTAAGGGTACATCCGCTGCTGACGCTGTCGAGGTCCCCGCTCC
CGCTGCTGTCCTTGGTGGTCCCGAGCCCCTTATGCAAGCTACCGCTTGGCTTAAC
GCTTACTTCCACCAACCCGAGGCTATTGAGGAGTTCCCCGTCCCCGCTCTTCACC
ACCCCGTCTTCCAACAAGAGTCCTTCACCCGTCAAGTCCTTTGGAAGTTACTTAAG
GTCGTCAAGTTCGGTGAGGTCATTTCCTACTCCCACCTTGCTGCTCTTGCTGGTAA
CCCCGCTGCTACCGCTGCTGTCAAGACCGCTCTTTCCGGTAACCCCGTCCCCATT
CTTATTCCCTGCCACCGTGTCGTCCAAGGTGACCTTGACGTCGGTGGTTACGAGG
GTGGTCTTGCTGTCAAGGAGTGGCTTCTTGCTCACGAGGGTCACCGTCTTGGTAA
GCGTTCCGCTTGGTCCCACCCCCAATTCGAGAAGGGTGGTGGTTCCGGTGGTGG
TTCCGGTGGTTCCGCTTGGTCCCACCCCCAATTCGAGAAGTAA

Cyan: Strep Tag

Green: Start codon

Red: Stop Codon

A.4.1 SNAP-Strep protein sequence

MKLGPGSDKDCEMKRTTLDSPLGKLELSGCEQGLHEIIFLGKGTSAADAVEVPAPAAV
LGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFG
EVISYSHLAALAGNPAATAAVKTALSGNPVPILIP<mark>C</mark>HRVVQGDLDVGGYEGGLAVKEW
LLAHEGHRLGKRSA<mark>WSHPQFEK</mark>GGGSGGGSGGSA<mark>WSHPQFEK</mark>

Cyan: Strep tag

Yellow: Reactive Cysteine of SNAP-Tag protein

Number of amino acids: 204

Molecular weight: 22814.02 Da

Formula: $C_{1024}H_{1583}N_{279}O_{295}S_5$

Extinction coefficient: 32095 $M^{-1}cm^{-1}$

A.4.2 SNAP-Strep protein mass spectrum
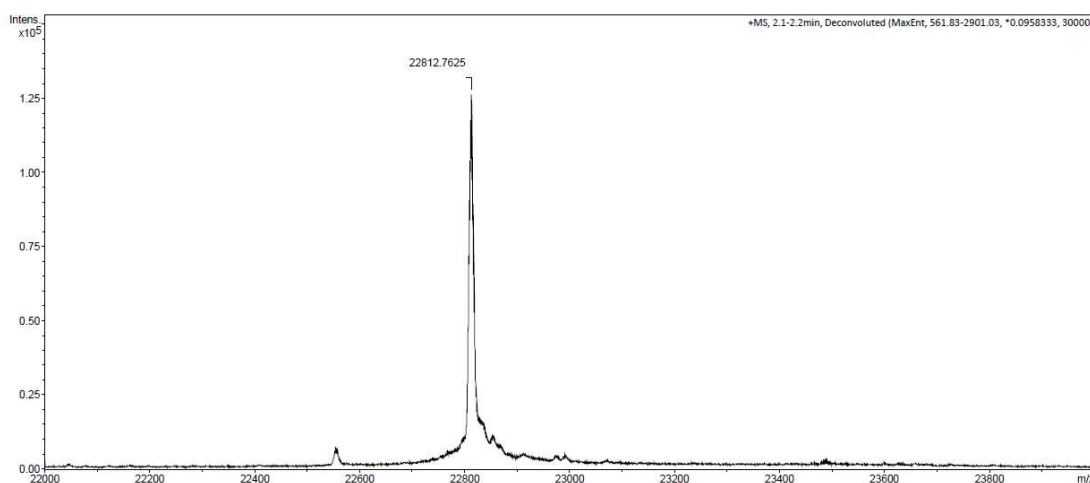
Expected mass: 22814.02 Da, Observed mass: 22.812.7625 Da

A.5 Halo-Strep

A.5.1 Halo-Strep Plasmid Map

pPSG-IBA103 HaloTag7-Strep
3766 bp
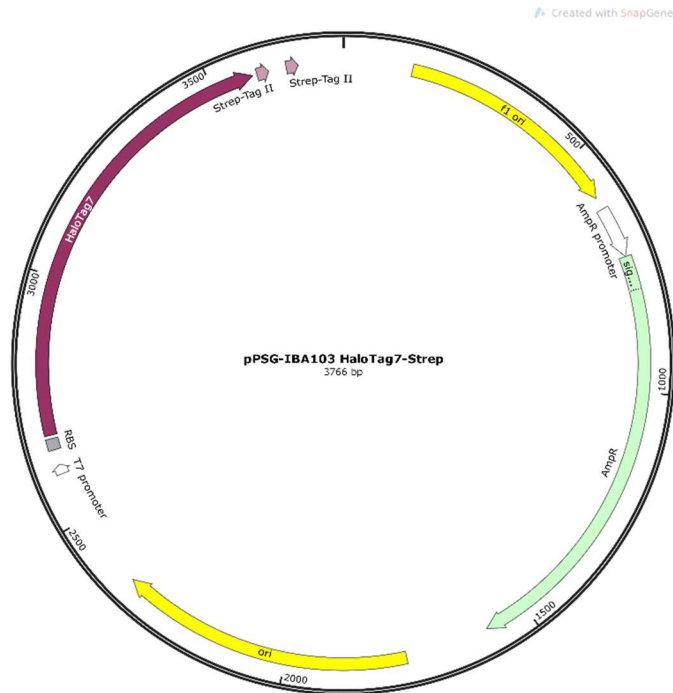
A.5.2 Halo-Strep DNA Sequence

ATGCCAGATCTCGGCAGCCATATGGCGGAAATTGGCACCGGCTTTCCGTTTGATC
CGCATTATGTGGAAGTGCTGGGCGAACGCATGCATTATGTGGATGTGGGCCCGC
GCGATGGCACCCCGGTGCTGTTTCTGCATGGCAACCCGACCAGCAGCTATGTGT
GGCGCAACATTATTCCGCATGTGGCGCCGACCCATCGCTGCATTGCGCCGGATC
TGATTGGCATGGGCAAAAGCGATAAACCGGATCTGGGCTATTTTTTTGATGATCAT
GTGCGCTTTATGGATGCGTTTATTGAAGCGCTGGGCCTGGAAGAAGTGGTGCTG
GTGATTCATGATTGGGGCAGCGCGCTGGGCTTTCATTGGGCGAAACGCAACCCG
GAACGCATTAAAGGCATTGCGTTTATGGAATTTATTCGCCCGATTCCGACCTGGG
ATGAATGGCCGGAATTTGCGCGCGAAACCTTTCAGGCGTTTCGCACCACCGATGT
GGGCCGCAAACTGATTATTGATCAGAACGTGTTTATTGAAGGCACCCTGCCGATG
GGCGTGGTGCGCCCGCTGACCGAAGTGGAAATGGATCATTATCGCGAACCGTTT
CTGAACCCGGTGGATCGCGAACCGCTGTGGCGCTTTCCGAACGAACTGCCGATT
GCGGGCGAACCGGCGAACATTGTGGCGCTGGTGGAAGAATATATGGATTGGCTG
CATCAGAGCCCGGTGCCGAAACTGCTGTTTTGGGGCACCCCGGGCGTGCTGATT
CCGCCGGCGGAAGCGGCGCGCCTGGCGAAAAGCCTGCCGAACTGCAAAGCGGT
GGATATTGGCCCGGGCCTGAACCTGCTGCAGGAAGATAACCCGGATCTGATTGG
CAGCGAAATTGCGCGCTGGCTGAGCACCCTGGAAATTAGCGGGAGCGCTTGGAG

CCACCCGCAGTTCGAAAAAGGTGGAGGTTCTGGCGGTGGATCGGGAGGTTCAGCGTGGAGCCACCCGCAGTTCGAGAAATAA

Cyan: Strep Tag

Green: Start Codon

Red: Stop Codon

A.4.1 Halo-Strep Protein Sequence

MPDLGSHMAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWR
NIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWG
SALGFHWAKRNPERIKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNV
FIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEY
MDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIG
SEIARWLSTLEISGSAWSHPQFEKGGGSGGGSGGSAWSHPQFEK*

Cyan: Strep tag

Yellow: Reactive Aspartic Acid Residue

Number of amino acids: 333

Molecular weight: 37257.40 Da

Formula: $C_{1705}H_{2561}N_{447}O_{476}S_{10}$

Extinction coefficient: 69440 $M^{-1}cm^{-1}$

A.4.2 Halo-Strep Protein Mass Spectrum
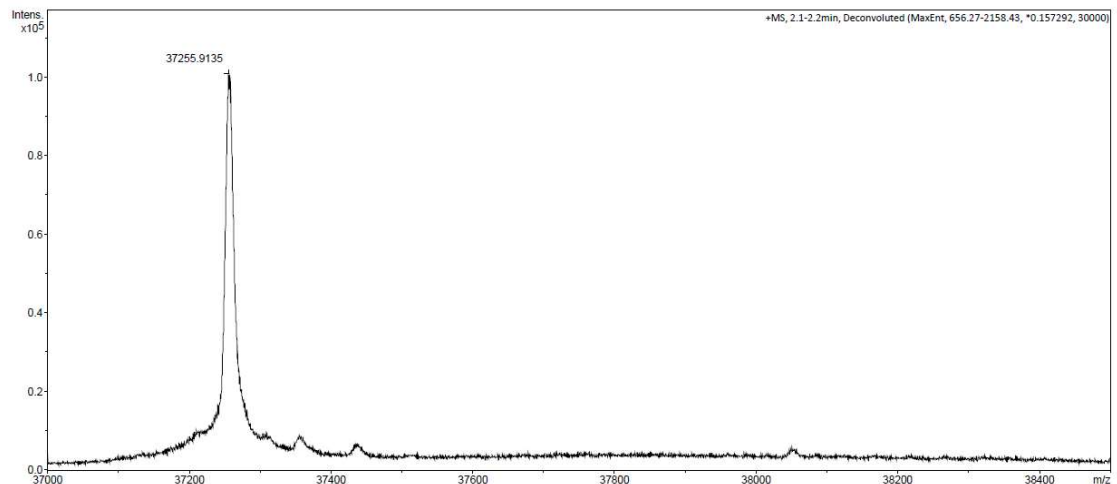
Expected mass: 37257.40 Da, Observed mass: 37255.9135 Da.

A.4.3 Stargate Cloning to generate plasmid

The process of Stargate cloning used to generate the HaloTag protein encoding plasmid is summarised in Figure A.1

pPSGIBA103 also encodes the LacZ α gene under a lac promotor (and so its expression can be induced by the addition of IPTG) in the segment between the two Esp3I sites. LacZ α encodes a peptide that complements the LacZ ω deletion mutation encoded in some strains of *E. coli,* such as the XL-1 Blue strain used here. When expressed together, LacZ α and ω form β-galactosidase. β-galactosidase can cleave the lactose-related compound 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) to 5-bromo-4-chloro-indoxyl, which dimerises to a blue precipitate, 5,5'-dibromo-4,4'-dichloro-indigo. This means, that colonies that receive the unchanged pPSGIBA103 plasmid will appear blue on an agar plate containing X-Gal and IPTG (Figure A.2A).

When the gBlock® containing the HaloTag gene is successfully cloned into the plasmid, the LacZ α gene is lost, and so these colonies will appear white as they will not be able to express β-galactosidase.
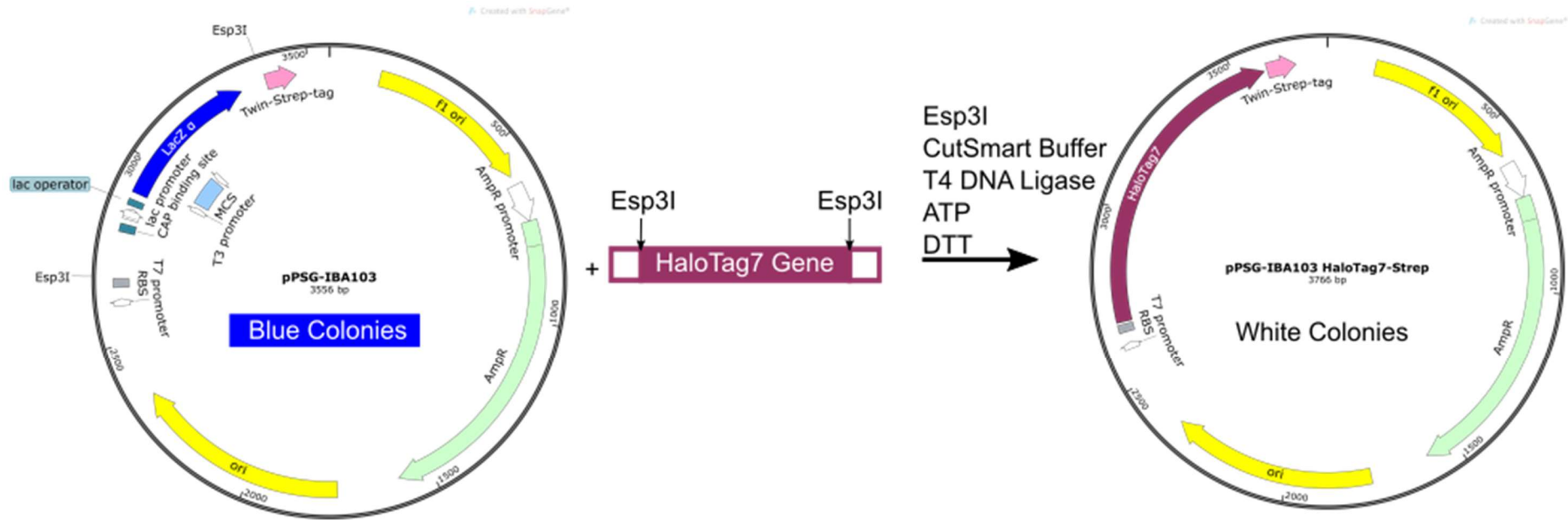
Figure A.1 Generation of the HaloTag Protein Expression Vector The generation of the pPSG-UBA103 HaloTag7-Twin-Strep-Tag plasmid using the Stargate cloning method is summarised. The plasmid used to express the HaloTag protein in E. coli was generated using StarGate Cloning. The reaction mixture was transformed into XL-1 Blue cells, which contain the plasmid to encode the LacZ ω peptide, which complements LacZ α to together allow the generation of blue colonies. Transformed cells with the HaloTag vector plasmid will give white colonies.
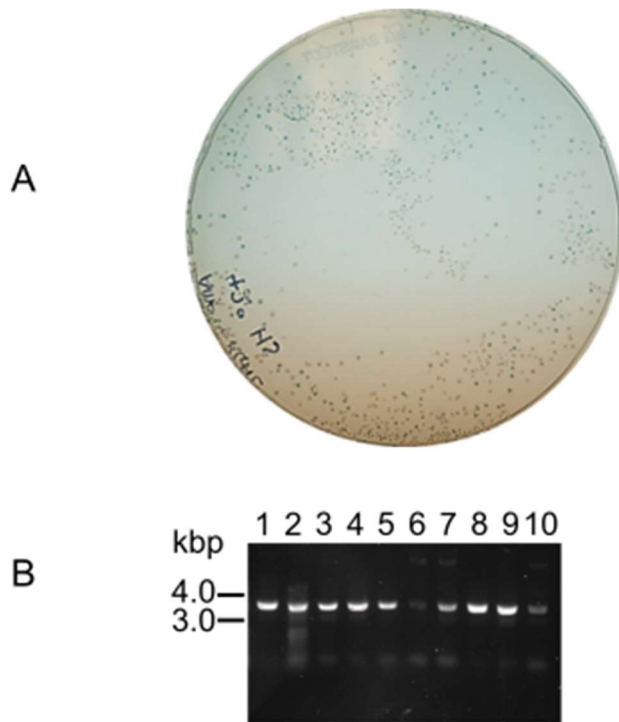
Figure A.2 Screening for successfully cloned HaloTag expression plasmids The reaction mixture from StarGate cloning was transformed into XL-1 Blue cells. **A.** The transformation gave a mixture of white and blue colonies. **B.** Ten white colonies were picked and subjected to colony PCR with primers that hybridised to the HaloTag7 gene. Eight of the ten colonies showed a strong positive result for containing a plasmid of the correct size (3766bp) with the HaloTag7 gene. Three of these colonies were sent for sequencing and all were found to encode the desired plasmid.

A.6 EGFP-SNAP-Strep

A.6.1 Primers to amplify insert from pET12b-SNAP-Strep

Forward (5'-3') GATGATGATGCTCAAGCTTTGGGTCCCGGTTCCGACAAG

Reverse (5'-3') ATCATCATCCGGTACCTTACTTCTCGAATTG

A.6.2 EGFP-SNAP-Strep Plasmid Map
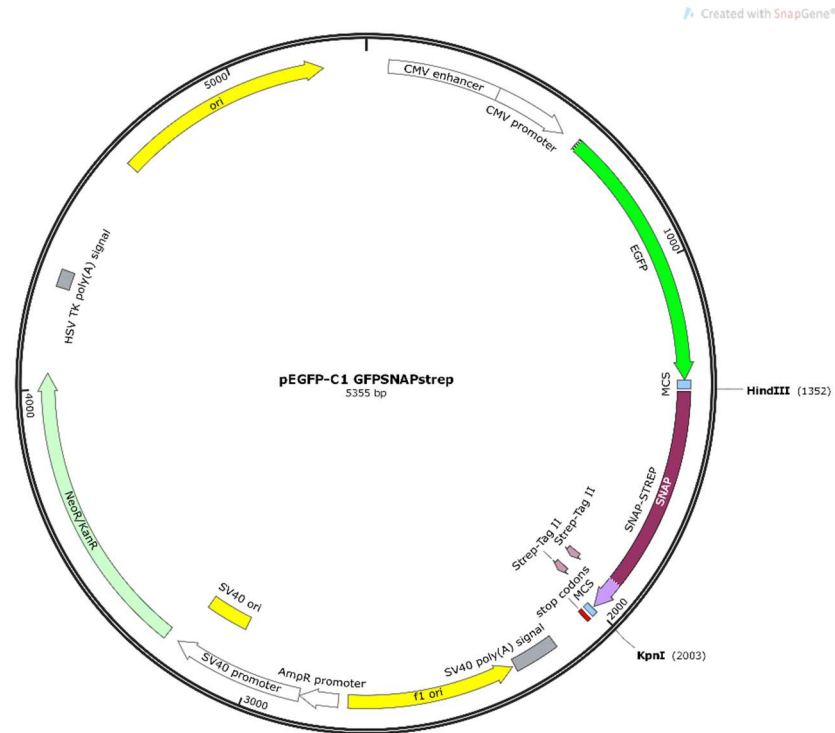
A.6.3 EGFP-SNAP-Strep DNA Sequence

```
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGA
GCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG
GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC
TGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCT
TCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
CGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAA
GACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCAT
CAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC
CGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
CAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG
CGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTCCGGACTCAGATCTCGAGCTCAAGCTTTGGGTCCCGG
```

TTCCGACAAGGACTGCGAGATGAAGCGTACCACCCTTGACTCCCCCCTTGGTAAG
TTAGAGCTTTCCGGTTGCGAGCAAGGTCTTCACGAGATTATTTTCCTTGGTAAGGG
TACATCCGCTGCTGACGCTGTCGAGGTCCCCGCTCCCGCTGCTGTCCTTGGTGG
TCCCGAGCCCCTTATGCAAGCTACCGCTTGGCTTAACGCTTACTTCCACCAACCC
GAGGCTATTGAGGAGTTCCCCGTCCCCGCTCTTCACCACCCCGTCTTCCAACAAG
AGTCCTTCACCCGTCAAGTCCTTTGGAAGTTACTTAAGGTCGTCAAGTTCGGTGA
GGTCATTTCCTACTCCCACCTTGCTGCTCTTGCTGGTAACCCCGCTGCTACCGCT
GCTGTCAAGACCGCTCTTTCCGGTAACCCCGTCCCCATTCTTATTCCCTGCCACC
GTGTCGTCCAAGGTGACCTTGACGTCGGTGGTTACGAGGGTGGTCTTGCTGTCAA
GGAGTGGCTTCTTGCTCACGAGGGTCACCGTCTTGGTAAGCGTTCCGCTTGGTCC
CACCCCCAATTCGAGAAGGGTGGTGGTTCCGGTGGTGGTTCCGGTGGTTCCGCT
TGGTCCCACCCCCAATTCGAGAAGTAA

Yellow: Start Codon

Red: Stop Codon

Green: GFP

Magenta: SNAP-Tag

Cyan: Strep-Tag

A.6.4 EGFP-SNAP-Strep Protein Sequence

VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWP
TLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYKSGLRSRAQALGPGSDKDCEMKRTTLDSPLGKLELSGCEQGLHEIIFLGKGTSA
ADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQ
VLWKLLKVVKFGEVISYSHLAALAGNPAATAAVKTALSGNPVPILIPCHRVVQGDLDVG
GYEGGLAVKEWLLAHEGHRLGKRSAWSHPQFEKGGGSGGGSGGSAWSHPQFEK

Green: GFP

Magenta: SNAP-Tag Protein

Yellow: Reactive Cysteine of SNAP-Tag protein

Cyan: Strep-Tag

A.7 EGFP-Halo-Strep

A.7.3 Primers to Amplify Insert from pPSG-IBA103-Halo-Strep

Forward (5'-3') TAGTAGTGTACAGCCTCGGCAGCCATATGGCG

Reverse (5'-3') AGTAGAAGCTTGGGGTGGCTCCCTTATTTCTCG

A.7.4 EGFP-Halo-Strep plasmid Map
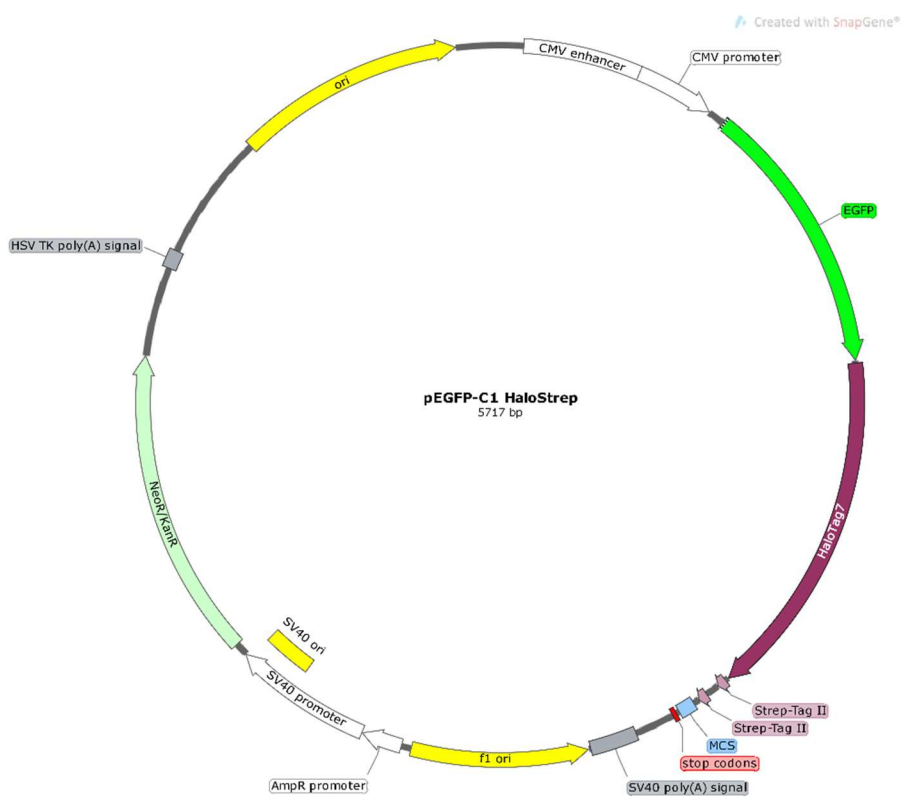
A.7.5 EGFP-Halo-Strep DNA Sequence

```
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGA
GCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG
GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC
TGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCT
TCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
CGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAA
GACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCAT
CAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC
CGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
```

CAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG
CGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAGCCTCGGCAGCCATATGGCGGAAATTGGCACCGGCTTTCC
GTTTGATCCGCATTATGTGGAAGTGCTGGGCGAACGCATGCATTATGTGGATGTG
GGCCCGCGCGATGGCACCCCGGTGCTGTTTCTGCATGGCAACCCGACCAGCAGC
TATGTGTGGCGCAACATTATTCCGCATGTGGCGCCGACCCATCGCTGCATTGCGC
CGGATCTGATTGGCATGGGCAAAAGCGATAAACCGGATCTGGGCTATTTTTTTGA
TGATCATGTGCGCTTTATGGATGCGTTTATTGAAGCGCTGGGCCTGGAAGAAGTG
GTGCTGGTGATTCATGATTGGGGCAGCGCGCTGGGCTTTCATTGGGCGAAACGC
AACCCGGAACGCATTAAAGGCATTGCGTTTATGGAATTTATTCGCCCGATTCCGAC
CTGGGATGAATGGCCGGAATTTGCGCGCGAAACCTTTCAGGCGTTTCGCACCAC
CGATGTGGGCCGCAAACTGATTATTGATCAGAACGTGTTTATTGAAGGCACCCTG
CCGATGGGCGTGGTGCGCCCGCTGACCGAAGTGGAAATGGATCATTATCGCGAA
CCGTTTCTGAACCCGGTGGATCGCGAACCGCTGTGGCGCTTTCCGAACGAACTG
CCGATTGCGGGCGAACCGGCGAACATTGTGGCGCTGGTGGAAGAATATATGGAT
TGGCTGCATCAGAGCCCGGTGCCGAAACTGCTGTTTTGGGGCACCCCGGGCGTG
CTGATTCCGCCGGCGGAAGCGGCGCGCCTGGCGAAAAGCCTGCCGAACTGCAA
AGCGGTGGATATTGGCCCGGGCCTGAACCTGCTGCAGGAAGATAACCCGGATCT
GATTGGCAGCGAAATTGCGCGCTGGCTGAGCACCCTGGAAATTAGCGGGAGCGC
TTGGAGCCACCCGCAGTTCGAAAAAGGTGGAGGTTCTGGCGGTGGATCGGGAGG
TTCAGCGTGGAGCCACCCGCAGTTCGAGAAATAA

Green: GFP

Magenta: Halo-Tag Protein

Cyan: Strep-tag

Yellow: Start Codon

Red: Stop Codon

A.7.6 EGFP-Halo-Strep Protein Sequence

VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWP
TLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE
GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM
DELYSLGSHMAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYV
WRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHD
WGSALGFHWAKRNPERIKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIID
QNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALV
EEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNP
DLIGSEIARWLSTLEISGSAWSHPQFEKGGGSGGGSGGGSAWSHPQFEK

Green: GFP

Magenta: HaloTag Protein

Yellow: Reactive Aspartic Acid of HaloTag protein

Cyan: Strep-Tag

Appendix B

B.1 Sequence Alignment of PEX5 Proteins

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                      10         20         30         40         50
H. sapiens    ---------- --MAMREL-- ---------- ---------- ----------
A. thalian    ---------- --MAMRDLVN GGAACAVPGS SSSSNPLGAL TNALLGSSSK
S. cerevis    MDV----GSC SV-------- ---------- ---------- ----------
B. taurus     ---------- --MAMREL-- ---------- ---------- ----------
M. musculu    ---------- --MYQ----- ---------- ---------- ----------
D. melanga    ------MVQS GIWRSKPILW RIRFCAAIYF IIIRALCNTV TKPIASITAD
N. tabacum    ---------- --MAMRDLVT GAPSCGEP-- SSSSNPLGAL ANALIGSSSK
K. pastori    MSLIGGGSDC AA-------- ---------- ---------- ----------
T. brucei     ---------- -----MDCGA GFALGQQLAK DALHMQGGVR PGTTGNVEQD


              ....|....| ....|....| ....|....| ....|....| ....|....|
                      60         70         80         90        100
H. sapiens    ---------- ---------- VEAECGGANP LMKLAGHFTQ DKALR-----
A. thalian    TQERLKEIPN ANRSGPRPQF YSEDD---Q- IRSLPGSELD QPLLQPGAQG
S. cerevis    ---------- ---------- ------GNNP LAQLHKHTQQ NKSLQFNQKN
B. taurus     ---------- ---------- VEAECGGANP LMKLAGHFTQ DKALR-----
M. musculu    ---------- ---------- ---------- -----GHMQL VNEQQ-----
D. melanga    KQLRSISTSN TMSFRP---L VEGDCGGVNP LMQLGGQFTR DVAHK-----
N. tabacum    TQERLKEIPT SVSTSSDGNF LAGFQ---EP LVSLPGSEFE QPL-HPNIQG
K. pastori    ---------- ---------- ------GSNP LAQFTKHTQH DTSLQQSMRN
T. brucei     ALMTGMMVPP TGPMEDWAQH FAAHQHHHQQ HQQMMMQRQH NDALMIQQQH


              ....|....| ....|....| ....|....| ....|....| ....|....|
                     110        120        130        140        150
H. sapiens    ---------- ---------Q EGLRPGPWPP GAPASEAAS- ------KPLG
A. thalian    SEFFRGFRSV DQNGLGAAWD EVQQGGPMPP MGPM------ --FEPVQPTF
S. cerevis    NGRL------ ---------N ESPLQGTNKP GI--SEAFI- ------SNV-
B. taurus     ---------- ---------Q EGLRPGPWPP GAPASEAVS- ------KPLG
M. musculu    ---------- ---------E ---------- -------S- ------RPLL
D. melanga    ---------- ---------D EGYVQ----- --RHFERAAR ----------
N. tabacum    SQFLQGFRSA DQNRLSDAWD EIQR--PQLP FGSQNMTNIP LEHARVQPDL
K. pastori    -GEF------ ---------Q ----QGNQRM MR--NE---- ------STM-
T. brucei     RDMEEAFRAS ARAG---APQ QANAGPLMMP PGPMMMAGGM ----------


              ....|....| ....|....| ....|....| ....|....| ....|....|
                     160        170        180        190        200
H. sapiens    VASEDELVAE FL-------- ---------- ---------- ----------
A. thalian    EGPPQRVLSN FLHSFVESSR GGIPFRPAPV PVLGLSQSDK QCIRDRSSIM
S. cerevis    NAISQENMAN MQ-------- ---------- ---------- ----------
B. taurus     VASEDELVAE FL-------- ---------- ---------- ----------
M. musculu    SPSIDDFLCE TK-------- ---------- ---------- ----------
D. melanga    --PEDQLINE FL-------- ---------- ---------- ----------
N. tabacum    NGPPQQVLSS FLHSFVNSSQ GGMQFRPTSL PLLGLSEGDK QCIRDRSTIM
K. pastori    SPMERQQMDQ FM-------- ---------- ---------- ----------
T. brucei     --APMMHAGG FM-------M GGM---PQMM PCAPMG---- ----------


              ....|....| ....|....| ....|....| ....|....| ....|....|
                     210        220        230        240        250
H. sapiens    ---------- ----QDQNA- --PLVS--RA PQTFKMDDLL AEMQQI-EQS
A. thalian    ARHFFADRGE EFINSQVNAL LSSLDIDDGI QARGHVPGRF RELDDY---W
S. cerevis    ---------- ----RFING- -EPLIDDKRR MEIGPSSGRL PPFSNV-HSL
```

```
B. taurus   ---------- ----QDQNA- --PLVS--RA PQTFKMDDLL AEMQEI-EQS
M. musculu  ---------- ----SEAIA- -KPVTS--NT AVLTTGLDLL DLSEPV-SQP
D. melanga  ---------- ----GQVTA- ---------P PQSFQMDTLL QEMRDI----
N. tabacum  ARHFFADKGE DFINGQVNAL LSSLEIDNDA RARGPVPGRY PELEEY---W
K. pastori  ---------- ----QQQNN- --PAFN---- ---------- --FQPMQHEL
T. brucei   ---------- ------MNMG MAPVAT-MSP ATTNTVSGAR EGATAV----


            ....|....| ....|....| ....|....| ....|....| ....|....|
               260        270        280        290        300
H. sapiens  NFRQAPQRAP GVADLALSEN WAQEFLAAGD AV---DVTQ- ----DYNETD
A. thalian  NESQAVVKPN LHPADNWAAE FNQHGMDHGG PDSWVQSFEQ QHG----VNG
S. cerevis  ---QTSANPT QIKGVNDISH WSQEFQGSNS I----QNRNA DTG--NSEKA
B. taurus   NFRQAPQRAP GVADLALSEN WAQEFLAAGD AV---DVTQ- ----EYNETD
M. musculu  QTK-AKKSEP SSKSSSLKKK ADGSDLISAD AE---QRAQA LRGPETSSLD
D. melanga  ---------- ---------- -----NIHGN PQ------QQ MHS--QQAEQ
N. tabacum  NESLAM-RPA PHVADGWINE FAQNRVEHAD PNAWAQSFEQ QHG----ANG
K. pastori  NVMQQNMNAP QQVANNS--- WNQEFRMKDP MVANAPSAQV QTP--VQSTN
T. brucei   -----SSAAP GVVDLGGDSA WAEK------ ---------- -----LHQAE


H. sapiens  WSQEFISEVT DPLSVSPARW AEEYLEQSEE KLWLGEPEGT AT-DRWYDEY
A. thalian  WATEFE---- ---------Q GQSQL-MS-- --------SQ MRSMDMQNIA
S. cerevis  WQRGST-TA- ------SSRF QYPNTMMNNY A---YASMNS LSGSRLQSPA
B. taurus   WSQEFISEVT DPLSVSPARW AEEYLEQSEE KLWLGEPEGT AAADRWYDEY
M. musculu  LDIQTQLEKW DDVKFHGDRT SKGHL-MAER KSCSSRTGSK EL--LWSAEH
D. melanga  WGQDFA---- ---------- ---------- -----RGLAP ALPNKMIHMQ
N. tabacum  WASEFE---- ---------H EQSQLGMI-- --------GQ MRGANIPNLA
K. pastori  WAQDFQ-QAG PEVQHHAQQH QHPILSVPGV RAGIYGGGRL MGGSMMNRAA
T. brucei   WGQDYK---- ---------- ---------- ---------- ----------


            ....|....| ....|....| ....|....| ....|....| ....|....|
               360        370        380        390        400
H. sapiens  HPEED--LQH TASDFVAKVD DPKLANSEFL KFVRQIGEGQ VSLESGAG--
A. thalian  AMEQTRKLAH TLSQDG---- NPKFQNSRFL QFVSKMSRGE LIIDENQVKQ
S. cerevis  FMNQQ---QS GRSKEGVNEQ EQQPWTDQFE KLEKEVSE-N LDINDEIEKE
B. taurus   QPEED--LQH TASDFVAKVD DPKLANSEFL KFVRQIGEGQ VSLESGAG--
M. musculu  RSQPE--LST GKSAL----- -----NSESA SELELVAPAQ ARL-------
D. melanga  AQQQD--LQH AQEFF----- DEPLISSQNF RSLPPLRQPL MPIAAGQQ--
N. tabacum  AMEQTRMLAH TLAQNN---- DPKFQNSKFL QFVSKMSRGE ITIEENQFKP
K. pastori  QMQQQNPAQA QTS-----EQ SQT------- QWEDQFKDIE SMLNSKTQ-E
T. brucei   -DVEVHTVEG STAQTV---- EEHAKTSKFY EFMDKIRKKE LLVDEDSG-E


            ....|....| ....|....| ....|....| ....|....| ....|....|
               410        420        430        440        450
H. sapiens  SGRAQAEQWA AEFI------ ---------- QQQGTSDAWV DQFTRPVNTS
A. thalian  AS-APG-EWA TEYEQQYLGP PSWADQFANE KLSHGPEQWA DEFASGRGQQ
S. cerevis  ENVSEVEQNK PETVEKEE-- ---------- ------GVYG DQYQSDFQ--
B. taurus   SGRAQAEQWA AEFI------ ---------- QQQGTSEAWV DQFTRPVNTS
M. musculu  ---TKEHRWG SALL------ ---------- SRN------- ---------H
D. melanga  --QDPFFDSA METIITDHLP QA-------- PQGESLDDWI SDYQRSTEQK
N. tabacum  ATVAPG-DWA AEYGQQYNGG QSWADQFAHE ELSRGPQGWV NEFSAEREQH
K. pastori  PKTKQQEQNT FEQV------ ---------- -------WD DIQVS-YA--
T. brucei   VVQGPGPDPD VEADTEYLAR LA-----AME GINVPPSVMD HMQGQDGVQR


            ....|....| ....|....| ....|....| ....|....| ....|....|
               460        470        480        490        500
H. sapiens  ALDMEFERAK SAIESDVDFW DKLQAELE-- ---EMAKRDA EAHPWLSDYD
A. thalian  ETAEDQWVNE FSKLNVDDWI DEF----A-- ---EGPVGDS SADAWANAYD
```

```
S. cerevis   ---------- -------EVW DSIHKDAEEV LPSELVNDDL NL------GE
B. taurus    ALDMEFERAK SAIESDVDFW DKLQAELE-- ---EMAKRDA EAHPWLSDHD
M. musculu   SLEEEFERAK AAVESDTEFW DKMQAEWE-- ---EMARRN- ----WISENQ
D. melanga   E--------Q TAANFNEKFW ERLQDEWQKL A-------DE NEHPWLSEYN
N. tabacum   GSVNDEWVNE FSKLNVNDWA DEFGNQVA-- ---EGAFGET SADSWAEAYD
K. pastori   ---------- -------DV- ---------- ---ELTNDQF QAQ-WEKDFA
T. brucei    GTDEDMEG-- ---MMGDDVY DPSADVEQ-- ---------- ----WAQEYA


             ....|....| ....|....| ....|....| ....|....| ....|....|
                  510        520        530        540        550
H. sapiens   D----L-TSA TYDKG----- YQFEEENPLR DHPQPFEEGL RRLQEG-DLP
A. thalian   EFLNEK-NAG KQTS----GV YVFSDMNPYV GHPEPMKEGQ ELFRKG-LLS
S. cerevis   DYLKYLGGRV NGNIE----- YAFQSNNEYF NNPNAYKIGC LLMENGAKLS
B. taurus    D----L-TSA SYDKG----- YHFEEENPLR DHPQPFEEGL RRLQEG-DLP
M. musculu   EAQNQV-TVS ASEKG----- YYFHTENPFK DWPGAFEEGL KRLKEG-DLP
D. melanga   DNMDAY---- ---KE----- YEFAEGNPMS DVENPFEKGK EYLSKG-DIP
N. tabacum   EYMNEQ-AAL KQQSDASRGV YVFSDLNPYV GHPNPLKEGQ ELFRKG-LLS
K. pastori   QYAE---GRL NYG-E----- YKYEEKNQFR NDPDAYEIGM RLMESGAKLS
T. brucei    QMQAMQ-ERL QNNTD----- YPFEANNPYM YHENPMEEGL SMLKLA-NLA


             ....|....| ....|....| ....|....| ....|....| ....|....|
                  560        570        580        590        600
H. sapiens   NAVLLFEAAV QQDPKHMEAW QYLGTTQAEN EQELLAISAL RRCLELKPDN
A. thalian   EAALALEAEV MKNPENAEGW RLLGVTHAEN DDDQQAIAAM MRAQEADPTN
S. cerevis   EAALAFEAAV KEKPDHVDAW LRLGLVQTQN EKELNGISAL EECLKLDPKN
B. taurus    NAVLLFEAAV QQDPKHMEAW QYLGTTQAEN EQELLAISAL RKCLELKPDN
M. musculu   VTILFMEAAI LQDPGDAEAW QFLGITQAEN ENEQAAIVAL QRCLELQPNN
D. melanga   SAVLCFEVAA KKQPERAEVW QLLGTSQTEN EMDPQAIAAL KRAYDLQPDN
N. tabacum   EAVLALEAEV LKNPENAEGW RLLGIAHAEN DDDQQAIAAM MRAQEADPTN
K. pastori   EAGLAFEAAV QQDPKHVDAW LKLGEVQTQN EKESDGIAAL EKCLELDPTN
T. brucei    EAALAFEAVC QKEPEREEAW RSLGLTQAEN EKDGLAIIAL NHARMLDPKD


             ....|....| ....|....| ....|....| ....|....| ....|....|
                  610        620        630        640        650
H. sapiens   QTALMALAVS FTNESLQRQA CETLRDWLRY TPA--YAHLV TPAEE-GA--
A. thalian   LEVLLALGVS HTNELEQATA LKYLYGWLRN HPK--YGAIA PPE-------
S. cerevis   LEAMKTLAIS YINEGYDMSA FTMLDKWAET KYPEIWSRIK QQDDKFQKEK
B. taurus    RTALMALAVS FTNESLQRQA CETLRDWLRY TPA--YAHLV APGEE-GA--
M. musculu   LKALMALAVS YTNTSHQQDA CEALKNWIKQ NPK--YKYLV ---KN-KK--
D. melanga   QQVLMALAAC YTNEGLQNNA VRMLCNWLTV HPK--YQHLV AAHPELQAE-
N. tabacum   LEVLLSLGVS HTNELEQQAA LKYLYSWLRH HPK--YGSIA PQD-------
K. pastori   LAALMTLAIS YINDGYDNAA YATLERWIET KYPDIASRAR SSNPDLD---
T. brucei    IAVHAALAVS HTNEHNANAA LASLRAWLLS QPQ--YEQLG SVNLQADVD-


             ....|....| ....|....| ....|....| ....|....| ....|....|
                  660        670        680        690        700
H. sapiens   GGAGLGPSKR ILGSLLSDS- -LFLEVKELF LAAVRLDPTS IDPDVQCGLG
A. thalian   ---------- LADSLYHA-- ----DIARLF NEA--SQLNP EDADVHIVLG
S. cerevis   GFTHIDMNAH ITKQFL---- -QLAN----- ------NLST IDPEIQLCLG
B. taurus    GGVGLGSSKR ILGSLLSDS- -LFLEVKELF LAAVRLDPTS IDPDVQCGLG
M. musculu   GSPGL--TRR MSKSPVDSS- -VLEGVKELY LEAAHQNGDM IDPDLQTGLG
D. melanga   -------GTS LASSLIGPS- -KLRDLQQIY LEAVRQHPSE VDAEVQDALG
N. tabacum   ---------- QPVSFYHA-- ----DVSRLF TDA--AQMSP DDADVHIVLG
K. pastori   GGDRIEQNKR VTELFMKAA- -QLSP----- ------DVAS MDADVQTGLG
T. brucei    ----IDDLNV QSEDFFFAAP NEYRECRTLL HAA--LEMNP NDAQLHASLG
```

```
                    ....|....| ....|....| ....|....| ....|....| ....|....|
                          710        720        730        740        750
H. sapiens   VLFNLSGEYD KAVDCFTAAL SVRPNDYLLW NKLGATLANG NQSEEAVAAY
A. thalian   VLYNLSREFD RAITSFQTAL QLKPNDYSLW NKLGATQANS VQSADAISAY
S. cerevis   LLFYTKDDFD KTIDCFESAL RVNPNDELMW NRLGASLANS NRSEEAIQAY
B. taurus    VLFNLSGEYD KAVDCFTAAL SVRPDDYLLW NKLGATLANG NQSEEAVAAY
M. musculu   VLFHLSGEFN RAIDAFNAAL TVRPEDYSLW NRLGATLANG DRSEEAVEAY
D. melanga   VLYNLSGEFD KAVDCYQSAL QVDPQNAKTW NRLGASLANG SRSVEAVEAY
N. tabacum   VLYNLSREYD KAIESFKTAL KLKPRDYSLW NKLGATQANS VQSADAILAY
K. pastori   VLFYSMEEFD KTIDCFKAAI EVEPDKALNW NRLGAALANY NKPEEAVEAY
T. brucei    VLYNLSNNYD SAAANLRRAV ELRPDDAQLW NKLGATLANG NRPQEALDAY


                    ....|....| ....|....| ....|....| ....|....| ....|....|
                          760        770        780        790        800
H. sapiens   RRALELQPGY IRSRYNLGIS CINLGAHREA VEHFLEALNM QRKSRGPRG-
A. thalian   QQALDLKPNY VRAWANMGIS YANQGMYKES IPYYVRALAM N---------
S. cerevis   HRALQLKPSF VRARYNLAVS SMNIGCFKEA AGYLLSVLSM HEVNTNNKKG
B. taurus    RRALELQPGY IRSRYNLGIS CINLGAHREA VEHFLEALNM QRKSRGPRG-
M. musculu   TRALEIQPGF IRSRYNLGIS CINLGAYREA VSNFLTALSL QRKSRNQQQV
D. melanga   QQALQLQPGF IRVRYNVGVC CMNLKAYKEA VEHLLTALTM QAHTNAAREL
N. tabacum   QQALDLKPNY VRAWANMGIS YANQGMYEDS IRYYVRALAM N---------
K. pastori   SRALQLNPNF VRARYNLGVS FINMGRYKEA VEHLLTGISL HEVEGVDAS-
T. brucei    NRALDINPGY VRVMYNMAVS YSNMSQYDLA AKQLVRAIYM QVGGTTPTG-


                    ....|....| ....|....| ....|....| ....|....| ....|....|
                          810        820        830        840        850
H. sapiens   EGGA------ ----MSENIW STLRLALSML GQSDAYGAAD AR-DLSTLLT
A. thalian   ---PK----- -----ADNAW QYLRLSLSCA SRQDMIEACE SR-NLDLLQK
S. cerevis   DVGSL----- -LNTYNDTVI ETLKRVFIAM NRDDLLQEVK PGMDLKRFKG
B. taurus    EGGA------ ----MSENIW STLRLALSML GQSDAYGAAD AR-DLPTLLA
M. musculu   PHPA------ ----ISGNIW AALRIALSLM DQPELFQAAN LG-DLDVLLR
D. melanga   PNAAMAATFR GQNQMSESIW STLKMVISLM GRSDLQSYVS DR-NLAALNE
N. tabacum   ---PK----- -----ADNAW QYLRISLSCA SRNDMLEACD AR-NLDVLQK
K. pastori   EMSSN----- -QGLQNNALV ETLKRAFLGM NRRDLVDKVY PGMGLAQFRK
T. brucei    EASREA---- -----TRSMW DFFRMLLNVM NRPDLVELTY AQ-NVEPFAK


                    ....|....
H. sapiens   MFGLPQ---
A. thalian   EFPL-----
S. cerevis   EFSF-----
B. taurus    MFGLPQ---
M. musculu   AFNLDP---
D. melanga   AFKD-----
N. tabacum   EFPL-----
K. pastori   MFDF-----
T. brucei    EFGLQSMLL
```

Peach: Conserved N-terminal cysteine   Green: TPR4

Pink: WXXXY Motifs   Blue: TPR5

Cyan: PEX7 binding domain   Indigo: TPR6

Red: TPR1   Violet: TPR7

Orange: TPR2   Yellow: TPR3

Residues in **bold** are included recombinant PEX5C proteins

B.2 STRIDE Analysis of *Hs*-PEX5C

REM		---Residue---				--Structure--			-Phi-			-Psi-			-Area-	
ASG	SER	A	280	1	T	Turn	360.00	142.41	84.3
ASG	ALA	A	281	2	T	Turn	-77.15	8.03	80.3
ASG	THR	A	282	3	T	Turn	-77.37	-13.60	105.0
ASG	TYR	A	283	4	T	Turn	-139.31	-5.84	22.3
ASG	ASP	A	284	5	T	Turn	-71.00	-24.55	67.1
ASG	LYS	A	285	6	T	Turn	-39.68	149.93	60.8
ASG	GLY	A	286	7	C	Coil	-80.14	166.31	64.1
ASG	TYR	A	287	8	C	Coil	-65.68	124.11	24.9
ASG	GLN	A	288	9	C	Coil	-94.81	102.57	71.0
ASG	PHE	A	289	10	C	Coil	-68.03	152.35	69.2
ASG	GLU	A	290	11	T	Turn	-71.57	118.12	83.6
ASG	GLU	A	291	12	T	Turn	-66.46	151.20	155.8
ASG	GLU	A	292	13	T	Turn	49.62	56.82	181.9
ASG	ASN	A	293	14	T	Turn	-73.05	112.36	13.0
ASG	PRO	A	294	15	T	Turn	-51.89	-33.79	110.0
ASG	LEU	A	295	16	T	Turn	-105.35	18.93	51.8
ASG	ARG	A	296	17	T	Turn	-50.73	-25.71	75.5
ASG	ASP	A	297	18	T	Turn	-121.48	37.79	133.4
ASG	HIS	A	298	19	T	Turn	-71.58	134.11	54.0
ASG	PRO	A	299	20	T	Turn	-47.07	-50.31	121.3
ASG	GLN	A	300	21	T	Turn	-126.71	74.64	32.1
ASG	PRO	A	301	22	H	AlphaHelix	-47.35	-61.24	8.2
ASG	PHE	A	302	23	H	AlphaHelix	-52.08	-56.12	11.4
ASG	GLU	A	303	24	H	AlphaHelix	-56.05	-51.28	88.6
ASG	GLU	A	304	25	H	AlphaHelix	-63.51	-31.38	48.7
ASG	GLY	A	305	26	H	AlphaHelix	-63.05	-49.17	0.0
ASG	LEU	A	306	27	H	AlphaHelix	-62.45	-33.98	51.8
ASG	ARG	A	307	28	H	AlphaHelix	-66.85	-43.18	126.0
ASG	ARG	A	308	29	H	AlphaHelix	-68.19	-43.37	94.8
ASG	LEU	A	309	30	H	AlphaHelix	-57.95	-45.86	22.3
ASG	GLN	A	310	31	H	AlphaHelix	-68.59	-16.81	170.9
ASG	GLU	A	311	32	H	AlphaHelix	-91.08	-1.50	144.3
ASG	GLY	A	312	33	C	Coil	85.61	4.85	10.7
ASG	ASP	A	313	34	C	Coil	-93.66	85.83	23.9
ASG	LEU	A	314	35	H	AlphaHelix	-63.83	-51.85	0.4
ASG	PRO	A	315	36	H	AlphaHelix	-50.94	-52.58	13.7
ASG	ASN	A	316	37	H	AlphaHelix	-65.58	-37.10	22.6
ASG	ALA	A	317	38	H	AlphaHelix	-56.85	-45.22	0.0
ASG	VAL	A	318	39	H	AlphaHelix	-59.10	-48.53	7.0
ASG	LEU	A	319	40	H	AlphaHelix	-55.06	-44.42	24.3
ASG	LEU	A	320	41	H	AlphaHelix	-66.33	-45.06	2.6
ASG	PHE	A	321	42	H	AlphaHelix	-64.28	-32.81	0.0
ASG	GLU	A	322	43	H	AlphaHelix	-70.49	-43.95	6.0
ASG	ALA	A	323	44	H	AlphaHelix	-56.43	-43.13	0.0
ASG	ALA	A	324	45	H	AlphaHelix	-60.44	-52.87	3.2
ASG	VAL	A	325	46	H	AlphaHelix	-69.72	-15.80	1.7
ASG	GLN	A	326	47	H	AlphaHelix	-82.35	-39.34	52.4
ASG	GLN	A	327	48	H	AlphaHelix	-89.88	4.05	87.3
ASG	ASP	A	328	49	T	Turn	-163.98	74.52	78.4
ASG	PRO	A	329	50	T	Turn	-50.99	-23.69	71.6
ASG	LYS	A	330	51	T	Turn	-99.05	16.56	150.6
ASG	HIS	A	331	52	T	Turn	-81.56	85.45	46.1
ASG	MET	A	332	53	H	AlphaHelix	-51.01	-51.92	47.3
ASG	GLU	A	333	54	H	AlphaHelix	-61.80	-31.62	84.8
ASG	ALA	A	334	55	H	AlphaHelix	-66.60	-39.24	0.0
ASG	TRP	A	335	56	H	AlphaHelix	-64.14	-42.30	11.8
ASG	GLN	A	336	57	H	AlphaHelix	-58.00	-48.33	19.4
ASG	TYR	A	337	58	H	AlphaHelix	-71.25	-30.97	60.3
ASG	LEU	A	338	59	H	AlphaHelix	-65.31	-37.72	0.0

```
ASG   GLY A  339    60     H    AlphaHelix   -70.26   -51.64      0.0
ASG   THR A  340    61     H    AlphaHelix   -66.28   -36.53     24.2
ASG   THR A  341    62     H    AlphaHelix   -67.55   -38.60      0.0
ASG   GLN A  342    63     H    AlphaHelix   -65.62   -36.74      3.6
ASG   ALA A  343    64     H    AlphaHelix   -61.36   -45.89     14.2
ASG   GLU A  344    65     H    AlphaHelix   -64.27   -29.59     14.4
ASG   ASN A  345    66     H    AlphaHelix   -88.47    13.23      4.4
ASG   GLU A  346    67     C         Coil    56.02    44.75      0.8
ASG   GLN A  347    68     C         Coil  -126.94    65.97     37.3
ASG   GLU A  348    69     H    AlphaHelix   -50.79   -52.04      9.4
ASG   LEU A  349    70     H    AlphaHelix   -62.25   -39.66     85.7
ASG   LEU A  350    71     H    AlphaHelix   -68.20   -35.80     38.0
ASG   ALA A  351    72     H    AlphaHelix   -62.09   -46.60      0.0
ASG   ILE A  352    73     H    AlphaHelix   -59.50   -41.48      0.0
ASG   SER A  353    74     H    AlphaHelix   -62.37   -48.51     12.9
ASG   ALA A  354    75     H    AlphaHelix   -62.62   -43.19      0.0
ASG   LEU A  355    76     H    AlphaHelix   -63.60   -46.26      0.0
ASG   ARG A  356    77     H    AlphaHelix   -64.53   -32.25     86.0
ASG   ARG A  357    78     H    AlphaHelix   -69.62   -40.19     52.1
ASG   CYS A  358    79     H    AlphaHelix   -62.12   -45.79      0.0
ASG   LEU A  359    80     H    AlphaHelix   -64.81   -22.47     22.1
ASG   GLU A  360    81     H    AlphaHelix   -67.66   -39.90    130.7
ASG   LEU A  361    82     H    AlphaHelix   -77.83   -41.56     52.4
ASG   LYS A  362    83     T         Turn  -134.98    88.60     73.2
ASG   PRO A  363    84     T         Turn   -55.61   -25.06     71.3
ASG   ASP A  364    85     T         Turn  -103.62    22.64     68.0
ASG   ASN A  365    86     T         Turn   -72.19   101.30      8.7
ASG   GLN A  366    87     H    AlphaHelix   -60.69   -37.20     77.1
ASG   THR A  367    88     H    AlphaHelix   -64.24   -44.21     76.9
ASG   ALA A  368    89     H    AlphaHelix   -66.28   -42.36      0.0
ASG   LEU A  369    90     H    AlphaHelix   -59.93   -50.88      3.8
ASG   MET A  370    91     H    AlphaHelix   -61.29   -40.84     20.2
ASG   ALA A  371    92     H    AlphaHelix   -67.44   -38.70     21.9
ASG   LEU A  372    93     H    AlphaHelix   -67.15   -34.24      3.8
ASG   ALA A  373    94     H    AlphaHelix   -65.81   -36.54      0.0
ASG   VAL A  374    95     H    AlphaHelix   -67.44   -43.91     10.0
ASG   SER A  375    96     H    AlphaHelix   -65.32   -42.54      0.6
ASG   PHE A  376    97     H    AlphaHelix   -60.00   -41.99      7.4
ASG   THR A  377    98     H    AlphaHelix   -63.16   -42.40      3.9
ASG   ASN A  378    99     H    AlphaHelix   -64.39   -27.68      3.0
ASG   GLU A  379   100     H    AlphaHelix   -89.96     7.69     71.0
ASG   SER A  380   101     C         Coil    62.28    31.95     86.2
ASG   LEU A  381   102     C         Coil  -105.58    77.20     36.0
ASG   GLN A  382   103     H    AlphaHelix   -56.67   -54.91     49.4
ASG   ARG A  383   104     H    AlphaHelix   -54.36   -48.63     30.5
ASG   GLN A  384   105     H    AlphaHelix   -62.26   -40.55     45.5
ASG   ALA A  385   106     H    AlphaHelix   -63.53   -42.30      0.0
ASG   CYS A  386   107     H    AlphaHelix   -65.48   -42.34      0.4
ASG   GLU A  387   108     H    AlphaHelix   -61.31   -43.45     92.5
ASG   ILE A  388   109     H    AlphaHelix   -64.97   -43.30     22.0
ASG   LEU A  389   110     H    AlphaHelix   -63.18   -34.93      0.2
ASG   ARG A  390   111     H    AlphaHelix   -66.55   -39.55     41.9
ASG   ASP A  391   112     H    AlphaHelix   -67.16   -33.93     46.5
ASG   TRP A  392   113     H    AlphaHelix   -59.72   -43.85      3.8
ASG   LEU A  393   114     H    AlphaHelix   -69.20   -39.64      0.0
ASG   ARG A  394   115     H    AlphaHelix   -70.99   -25.73     91.8
ASG   TYR A  395   116     H    AlphaHelix   -94.48     9.22     87.8
ASG   THR A  396   117     C         Coil  -111.49   105.60     14.2
ASG   PRO A  397   118     G     310Helix   -29.81   -58.91    108.5
ASG   ALA A  398   119     G     310Helix   -60.62   -22.76     63.3
ASG   TYR A  399   120     G     310Helix  -114.80    -8.99     23.1
ASG   ALA A  400   121     G     310Helix   -47.61   -31.27     37.5
ASG   HIS A  401   122     G     310Helix   -65.45   -28.52    131.7
```

```
ASG   LEU A  402  123    G     310Helix    -66.42    -25.30     50.3
ASG   VAL A  403  124    C       Coil      -93.39    122.77     17.1
ASG   THR A  404  125    C       Coil      -56.55    360.00    110.9
ASG   ARG A  420  126    T       Turn      360.00     -2.04    137.9
ASG   ILE A  421  127    T       Turn      -76.88    -13.63    115.6
ASG   LEU A  422  128    T       Turn     -128.32     22.98     50.2
ASG   GLY A  423  129    H    AlphaHelix   -54.60    -42.45     26.4
ASG   SER A  424  130    H    AlphaHelix   -59.26    -43.07     91.4
ASG   LEU A  425  131    H    AlphaHelix   -62.94    -49.86     37.6
ASG   LEU A  426  132    H    AlphaHelix   -63.42    -40.13     87.9
ASG   SER A  427  133    H    AlphaHelix   -65.81    -49.77     31.7
ASG   ASP A  428  134    H    AlphaHelix   -63.47    -43.49     74.9
ASG   SER A  429  135    H    AlphaHelix   -64.98    -39.03      4.7
ASG   LEU A  430  136    H    AlphaHelix   -66.35    -42.91     19.5
ASG   PHE A  431  137    H    AlphaHelix   -61.04    -44.86     23.5
ASG   LEU A  432  138    H    AlphaHelix   -57.25    -46.63     24.9
ASG   GLU A  433  139    H    AlphaHelix   -59.23    -57.48     31.3
ASG   VAL A  434  140    H    AlphaHelix   -62.43    -34.23      0.0
ASG   LYS A  435  141    H    AlphaHelix   -65.49    -48.15     17.6
ASG   GLU A  436  142    H    AlphaHelix   -61.91    -35.84     41.9
ASG   LEU A  437  143    H    AlphaHelix   -59.22    -48.46     20.6
ASG   PHE A  438  144    H    AlphaHelix   -68.40    -36.34      0.0
ASG   LEU A  439  145    H    AlphaHelix   -64.51    -35.95      4.3
ASG   ALA A  440  146    H    AlphaHelix   -66.80    -40.80     20.7
ASG   ALA A  441  147    H    AlphaHelix   -59.07    -37.62      3.1
ASG   VAL A  442  148    H    AlphaHelix   -69.79    -44.79     33.0
ASG   ARG A  443  149    H    AlphaHelix   -70.13    -17.06     87.0
ASG   LEU A  444  150    H    AlphaHelix   -50.33    -49.20     57.2
ASG   ASP A  445  151    T       Turn     -134.23     85.45     50.4
ASG   PRO A  446  152    T       Turn      -60.94    -12.05     47.1
ASG   THR A  447  153    T       Turn      -84.94    -23.92    100.7
ASG   SER A  448  154    T       Turn     -132.65    149.69     77.9
ASG   ILE A  449  155    C       Coil      -94.51    128.61     40.1
ASG   ASP A  450  156    C       Coil      -99.66    107.95     45.8
ASG   PRO A  451  157    H    AlphaHelix   -48.99    -45.51     17.4
ASG   ASP A  452  158    H    AlphaHelix   -66.34    -42.20     55.7
ASG   VAL A  453  159    H    AlphaHelix   -63.37    -38.12      1.6
ASG   GLN A  454  160    H    AlphaHelix   -65.19    -41.99     22.8
ASG   CYS A  455  161    H    AlphaHelix   -58.04    -49.44      6.8
ASG   GLY A  456  162    H    AlphaHelix   -58.99    -45.77      0.0
ASG   LEU A  457  163    H    AlphaHelix   -63.89    -36.87      7.2
ASG   GLY A  458  164    H    AlphaHelix   -65.06    -35.84      0.0
ASG   VAL A  459  165    H    AlphaHelix   -66.63    -44.41      0.0
ASG   LEU A  460  166    H    AlphaHelix   -60.10    -42.88      0.0
ASG   PHE A  461  167    H    AlphaHelix   -67.45    -22.67      2.0
ASG   ASN A  462  168    H    AlphaHelix   -74.12    -40.08     10.2
ASG   LEU A  463  169    H    AlphaHelix   -61.47    -37.06     32.2
ASG   SER A  464  170    H    AlphaHelix   -89.63     -3.87     46.4
ASG   GLY A  465  171    C       Coil       79.27     13.78     37.5
ASG   GLU A  466  172    C       Coil      -93.65     63.58     50.4
ASG   TYR A  467  173    H    AlphaHelix   -57.98    -35.62     62.8
ASG   ASP A  468  174    H    AlphaHelix   -59.13    -46.43     45.1
ASG   LYS A  469  175    H    AlphaHelix   -70.56    -33.06     15.8
ASG   ALA A  470  176    H    AlphaHelix   -61.81    -38.96      0.0
ASG   VAL A  471  177    H    AlphaHelix   -59.60    -46.19      1.0
ASG   ASP A  472  178    H    AlphaHelix   -60.06    -37.24     20.3
ASG   CYS A  473  179    H    AlphaHelix   -64.54    -43.25      5.2
ASG   PHE A  474  180    H    AlphaHelix   -76.03    -31.16      0.0
ASG   THR A  475  181    H    AlphaHelix   -73.01    -33.97      9.8
ASG   ALA A  476  182    H    AlphaHelix   -61.98    -46.19     11.2
ASG   ALA A  477  183    H    AlphaHelix   -66.10    -33.97      0.0
ASG   LEU A  478  184    H    AlphaHelix   -69.33    -28.47     13.4
ASG   SER A  479  185    H    AlphaHelix   -65.89    -20.74     56.8
```

```
ASG   VAL A  480   186     H    AlphaHelix    -97.46    -29.09    68.4
ASG   ARG A  481   187     T       Turn      -135.30     71.81    79.6
ASG   PRO A  482   188     T       Turn       -55.19    -26.68    39.1
ASG   ASN A  483   189     T       Turn      -107.98     22.17   106.3
ASG   ASP A  484   190     T       Turn       -86.40    101.39    35.1
ASG   TYR A  485   191     H    AlphaHelix    -61.25    -27.88    48.5
ASG   LEU A  486   192     H    AlphaHelix    -64.86    -44.25    52.5
ASG   LEU A  487   193     H    AlphaHelix    -64.70    -35.31     2.2
ASG   TRP A  488   194     H    AlphaHelix    -66.77    -38.25     7.2
ASG   ASN A  489   195     H    AlphaHelix    -63.49    -43.86     0.0
ASG   LYS A  490   196     H    AlphaHelix    -62.62    -36.31    19.0
ASG   LEU A  491   197     H    AlphaHelix    -65.82    -45.88     1.0
ASG   GLY A  492   198     H    AlphaHelix    -60.85    -51.77     0.0
ASG   ALA A  493   199     H    AlphaHelix    -55.98    -41.37     0.0
ASG   THR A  494   200     H    AlphaHelix    -64.62    -45.80     0.0
ASG   LEU A  495   201     H    AlphaHelix    -60.53    -44.06     1.7
ASG   ALA A  496   202     H    AlphaHelix    -62.82    -47.28     4.0
ASG   ASN A  497   203     H    AlphaHelix    -60.54    -27.29    25.7
ASG   GLY A  498   204     H    AlphaHelix    -84.94     13.98    22.1
ASG   ASN A  499   205     C       Coil        65.69     37.41   135.8
ASG   GLN A  500   206     C       Coil      -129.69     54.70    44.8
ASG   SER A  501   207     H    AlphaHelix    -61.12    -35.57     4.5
ASG   GLU A  502   208     H    AlphaHelix    -58.20    -44.59    64.2
ASG   GLU A  503   209     H    AlphaHelix    -68.97    -38.77    22.9
ASG   ALA A  504   210     H    AlphaHelix    -56.58    -31.78     0.0
ASG   VAL A  505   211     H    AlphaHelix    -57.94    -41.92    32.1
ASG   ALA A  506   212     H    AlphaHelix    -59.82    -37.42    26.3
ASG   ALA A  507   213     H    AlphaHelix    -69.15    -44.96     0.6
ASG   TYR A  508   214     H    AlphaHelix    -62.01    -41.12     0.0
ASG   ARG A  509   215     H    AlphaHelix    -55.35    -44.27   127.0
ASG   ARG A  510   216     H    AlphaHelix    -63.56    -49.77    52.9
ASG   ALA A  511   217     H    AlphaHelix    -55.35    -43.37     0.0
ASG   LEU A  512   218     H    AlphaHelix    -69.75    -32.07    13.2
ASG   GLU A  513   219     H    AlphaHelix    -64.39    -40.32   133.2
ASG   LEU A  514   220     H    AlphaHelix    -75.07    -35.62    54.4
ASG   GLN A  515   221     T       Turn      -145.13     83.11    19.4
ASG   PRO A  516   222     T       Turn       -52.49    -28.81    73.2
ASG   GLY A  517   223     T       Turn       -85.11     11.15    12.2
ASG   TYR A  518   224     T       Turn       -78.81     86.61     4.6
ASG   ILE A  519   225     H    AlphaHelix    -50.32    -52.64     1.4
ASG   ARG A  520   226     H    AlphaHelix    -48.34    -44.63     1.8
ASG   SER A  521   227     H    AlphaHelix    -67.16    -41.27     0.0
ASG   ARG A  522   228     H    AlphaHelix    -64.92    -41.91    37.7
ASG   TYR A  523   229     H    AlphaHelix    -58.68    -46.21     8.2
ASG   ASN A  524   230     H    AlphaHelix    -61.94    -35.45     0.0
ASG   LEU A  525   231     H    AlphaHelix    -64.85    -38.69     0.8
ASG   GLY A  526   232     H    AlphaHelix    -64.13    -39.27     0.0
ASG   ILE A  527   233     H    AlphaHelix    -62.91    -42.91     3.6
ASG   SER A  528   234     H    AlphaHelix    -62.33    -38.98     0.0
ASG   CYS A  529   235     H    AlphaHelix    -60.74    -41.90     5.6
ASG   ILE A  530   236     H    AlphaHelix    -58.13    -49.70     9.4
ASG   ASN A  531   237     H    AlphaHelix    -60.44    -26.15    54.5
ASG   LEU A  532   238     H    AlphaHelix    -89.90     -7.35    92.9
ASG   GLY A  533   239     C       Coil        86.93      2.38    41.6
ASG   ALA A  534   240     C       Coil       -88.50     72.96    28.2
ASG   HIS A  535   241     H    AlphaHelix    -64.93    -37.92    46.9
ASG   ARG A  536   242     H    AlphaHelix    -64.89    -43.11   114.0
ASG   GLU A  537   243     H    AlphaHelix    -64.21    -36.96   100.4
ASG   ALA A  538   244     H    AlphaHelix    -60.57    -48.19     0.0
ASG   VAL A  539   245     H    AlphaHelix    -56.66    -47.16     0.2
ASG   GLU A  540   246     H    AlphaHelix    -55.56    -47.77    64.5
ASG   HIS A  541   247     H    AlphaHelix    -59.29    -46.95    19.2
ASG   PHE A  542   248     H    AlphaHelix    -60.94    -43.77     0.0
```

```
ASG   LEU A  543  249    H    AlphaHelix    -63.06   -41.54     0.0
ASG   GLU A  544  250    H    AlphaHelix    -61.87   -45.41    61.6
ASG   ALA A  545  251    H    AlphaHelix    -60.25   -52.04     0.0
ASG   LEU A  546  252    H    AlphaHelix    -58.65   -39.20     0.0
ASG   ASN A  547  253    H    AlphaHelix    -71.62   -30.63     8.4
ASG   MET A  548  254    H    AlphaHelix    -66.89   -36.39    62.4
ASG   GLN A  549  255    H    AlphaHelix    -71.38   -47.54     0.0
ASG   ARG A  550  256    H    AlphaHelix    -57.27   -42.20    49.9
ASG   LYS A  551  257    H    AlphaHelix    -66.82    30.58   137.8
ASG   SER A  552  258    T        Turn    -124.92   360.00    50.2
ASG   GLY A  559  259    T        Turn     360.00   117.27   111.1
ASG   GLY A  560  260    T        Turn     -73.32   159.17    13.7
ASG   ALA A  561  261    C        Coil     -73.68   128.92    42.7
ASG   MET A  562  262    C        Coil    -101.62   150.89    47.0
ASG   SER A  563  263    C        Coil     -77.75   120.17     0.0
ASG   GLU A  564  264    H    AlphaHelix    -62.36   -28.60   104.9
ASG   ASN A  565  265    H    AlphaHelix    -62.05   -37.61    58.5
ASG   ILE A  566  266    H    AlphaHelix    -73.75   -40.85     0.0
ASG   TRP A  567  267    H    AlphaHelix    -55.02   -49.03     1.8
ASG   SER A  568  268    H    AlphaHelix    -59.89   -45.20    72.9
ASG   THR A  569  269    H    AlphaHelix    -68.26   -34.67    20.9
ASG   LEU A  570  270    H    AlphaHelix    -63.40   -45.96     0.2
ASG   ARG A  571  271    H    AlphaHelix    -57.48   -42.18    54.7
ASG   LEU A  572  272    H    AlphaHelix    -62.73   -43.89    59.1
ASG   ALA A  573  273    H    AlphaHelix    -63.41   -37.63     0.0
ASG   LEU A  574  274    H    AlphaHelix    -61.54   -36.39     0.2
ASG   SER A  575  275    H    AlphaHelix    -68.16   -41.36    53.9
ASG   MET A  576  276    H    AlphaHelix    -65.07   -31.31    61.9
ASG   LEU A  577  277    H    AlphaHelix    -77.61   -31.35    47.5
ASG   GLY A  578  278    T        Turn      85.97    20.74    45.2
ASG   GLN A  579  279    T        Turn     -83.45    66.29    50.8
ASG   SER A  580  280    G     310Helix    -51.33   -24.01    77.4
ASG   ASP A  581  281    G     310Helix    -68.86   -18.10   128.3
ASG   ALA A  582  282    G     310Helix    -98.64    -7.56     0.0
ASG   TYR A  583  283    H    AlphaHelix    -57.35   -50.75    55.5
ASG   GLY A  584  284    H    AlphaHelix    -55.31   -49.65    49.0
ASG   ALA A  585  285    H    AlphaHelix    -62.34   -36.69    14.3
ASG   ALA A  586  286    H    AlphaHelix    -66.61   -49.98     0.4
ASG   ASP A  587  287    H    AlphaHelix    -55.57   -38.98    95.1
ASG   ALA A  588  288    H    AlphaHelix    -84.10     2.95    73.7
ASG   ARG A  589  289    C        Coil      58.51    34.69    83.4
ASG   ASP A  590  290    C        Coil     -91.48    84.87    62.8
ASG   LEU A  591  291    H    AlphaHelix    -64.10   -39.37     3.9
ASG   SER A  592  292    H    AlphaHelix    -54.61   -46.30    43.1
ASG   THR A  593  293    H    AlphaHelix    -61.35   -49.47    54.2
ASG   LEU A  594  294    H    AlphaHelix    -61.05   -41.88     0.6
ASG   LEU A  595  295    H    AlphaHelix    -55.25   -49.02     0.0
ASG   THR A  596  296    H    AlphaHelix    -63.75   -43.07    93.2
ASG   MET A  597  297    H    AlphaHelix    -57.16   -34.90    73.4
ASG   PHE A  598  298    H    AlphaHelix   -101.05    25.05    15.4
ASG   GLY A  599  299    C        Coil      63.08    29.68    66.9
ASG   LEU A  600  300    C        Coil    -108.31   143.84    23.4
ASG   PRO A  601  301    C        Coil     -73.07   152.20    98.4
ASG   GLN A  602  302    C        Coil     -67.72   360.00   107.4
```

B.3 Mutagenesis Primers

B.3.1 *At*-PEX5C A438C

Forward (5'-3'):   CGATGAGTTTCTGAATGAGAAAAATTGTGGAAAACAAACCAGTGG

Reverse (5'-3'):   CCACTGGTTTGTTTTCCACAATTTTTCTCATTCAGAAACTCATCG

B.3.2 *At*-PEX5C G439C

Forward (5'-3'): CTGAATGAGAAAAATGCTTGCAAACAAACCAGTGGTGTCTACGTC

Reverse (5'-3'): GACGTAGACACCACTGGTTTGTTTGCAAGCATTTTTCTCATTCAG

B.3.3 *At*-PEX5C Q508C

Forward (5'-3'): ACACACGAGAACGATGATGA**TTG**TCAGCGAATAGCTGCAATGATG

Reverse (5'-3'): CATCATTGCAGCTATTCGCTGACAATCATCATCGTTCTCGTGTGT

B.3.4 *At*-PEX5C L539C

Forward (5'-3'): CAAAGCAGTTGCTTGCTCGCACTCGTTGGTATGACTCAC

Reverse (5'-3'): GTGAGTCATACCAACGAGTGCGAGCAAGCAACTGCTTTG

B.3.5 *At*-PEX5C L566C

Forward (5'-3'): GGTACAAAGAATCCGCACACTCCGGAGGCGCAATTGC

Reverse (5'-3'): GCAATTGCGCCTCCGGAGTGTGCGGATTCTTTGTACC

B.3.6 *At*-PEX5C S603C

Forward (5'-3'): GGCGTGCTCTACAATCTGTGCAGAGAGTTCGATAGAGC

Reverse (5'-3'): GCTCTATCGAACTCTCTGCACAGATTGTAGAGCACGCC

B.3.7 *At*-PEX5C F606C

Forward (5'-3'): GCTCTACAATCTGTCGAGAGAGTGCGATAGAGCAATC

Reverse (5'-3'): GATTGCTCTATCGCACTCTCTCGACAGATTGTAGAGC

B.3.8 *At*-PEX5C A641C

Forward (5'-3'): GATAAGCAGATATGGCATCACAACTCTGGACACTGTTGGC

Reverse (5'-3'): GCCAACAGTGTCCAGAGTTGTGATGCCATATCTGCTTATC

B.3.9 *At*-PEX5C Q671C

Forward (5'-3'): GGGCAAACATGGGAATTGCTTACGCAAACCAGGGGATGTAC

Reverse (5'-3'): GTACATCCCCTGGTTTGCGTAAGCAATTCCCATGTTTGCCC

B.3.10   *At*-PEX5C L700C

Forward (5'-3'): CGCATGGCAATACTTGAGACGTTCGTTAAGTTGTGCATCAAGG

Reverse (5'-3'): CCTTGATGCACAACTTAACGAACGTCTCAAGTATTGCCATGCG

B.3.11   *Hs*-PEX5C Y467C

Forward (5'-3'): TTC AAC CTG AGT GGG GAG TGT GAC AAG GCC GTG GAC TGC

Reverse (5'-3'): GCA GTC CAC GGC CTT GTC ACA CTC CCC ACT CAG GTT GAA

B.4 Sequencing Primers

B.4.1 pET vector with T7 Promoter Sequencing Primers

Forward (5'-3'): TAATACGACTCACTATAGG

Reverse (5'-3'): CTAGTTATTGCTCAGCGG

Appendix C

HPLC Analysis of Final Peptides

C.1 SNAP-YQSKL (10)

C.1.1 SNAP-YQSKL Structure
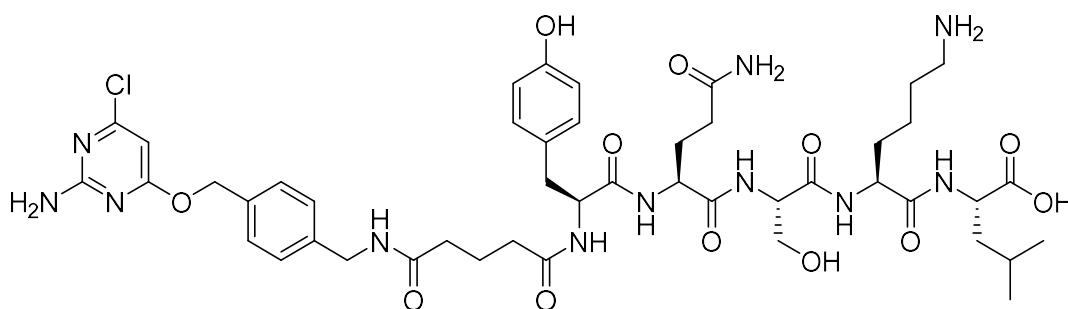
C.1.2 SNAP-YQSKL HPLC
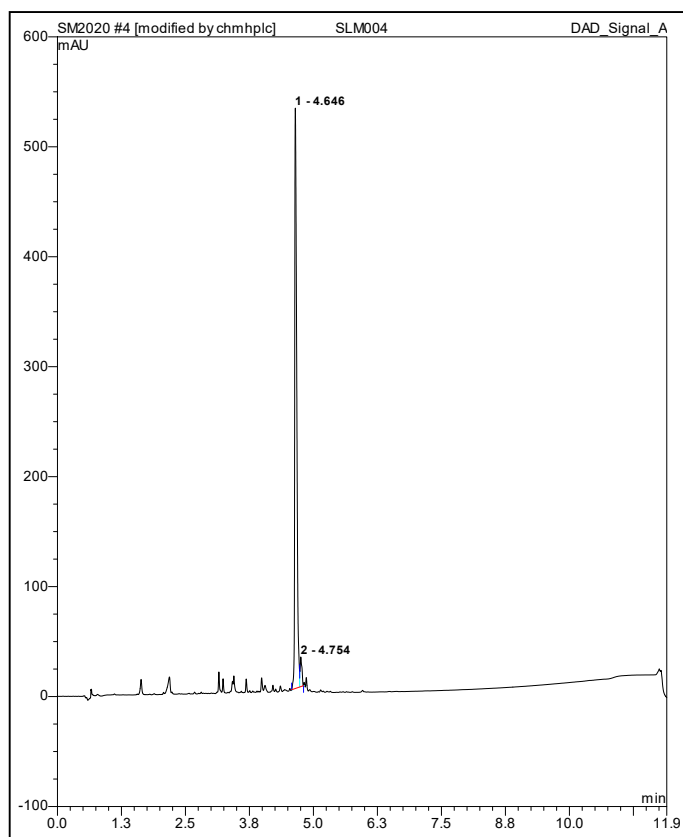
Purity= 95.71%

C.2 SNAP-YQLKS (11)

C.2.1 SNAP-YQLKS Structure
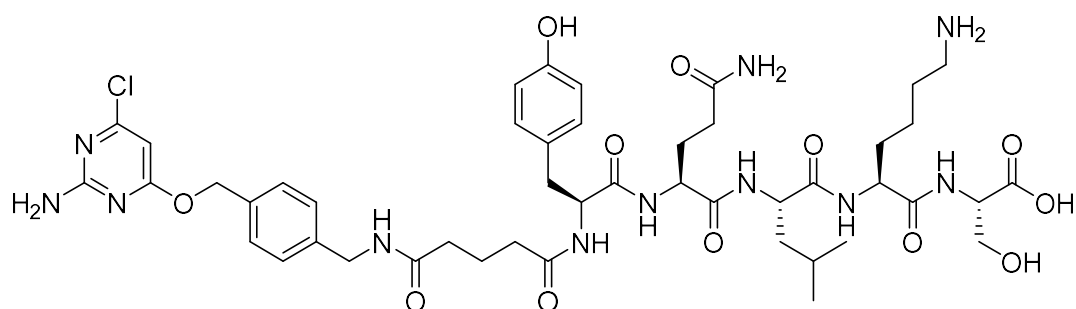
C.2.2 SNAP-YQLKS HPLC
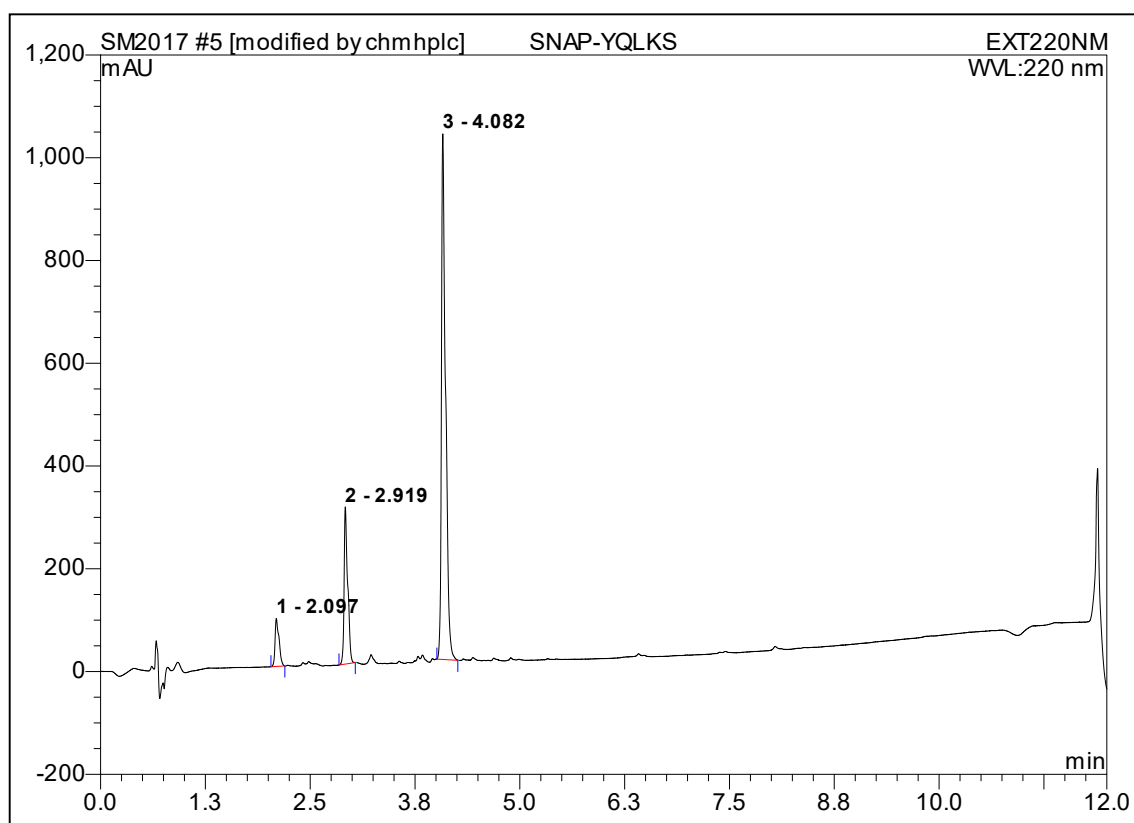
Purity= 74.41%

C.3 SNAP-PacificBlue-PEG-YQSKL (12)

C.3.1 SNAP-PacificBlue-PEG-YQSKL Structure
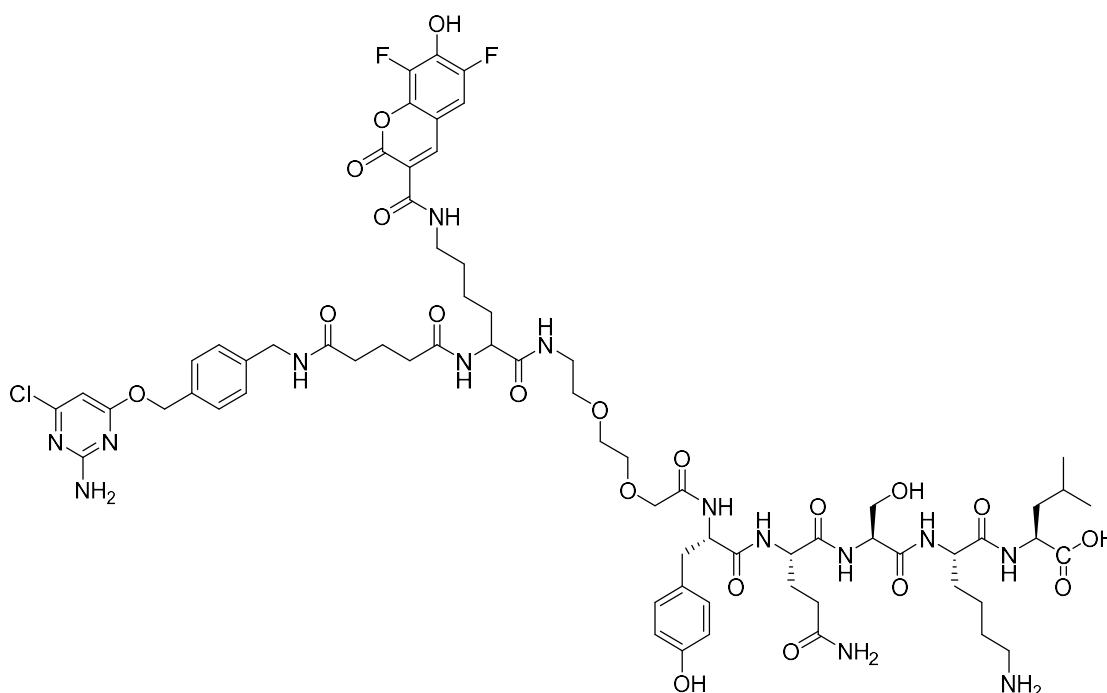
C.3.2 SNAP-PacificBlue-PEG-YQSKL HPLC
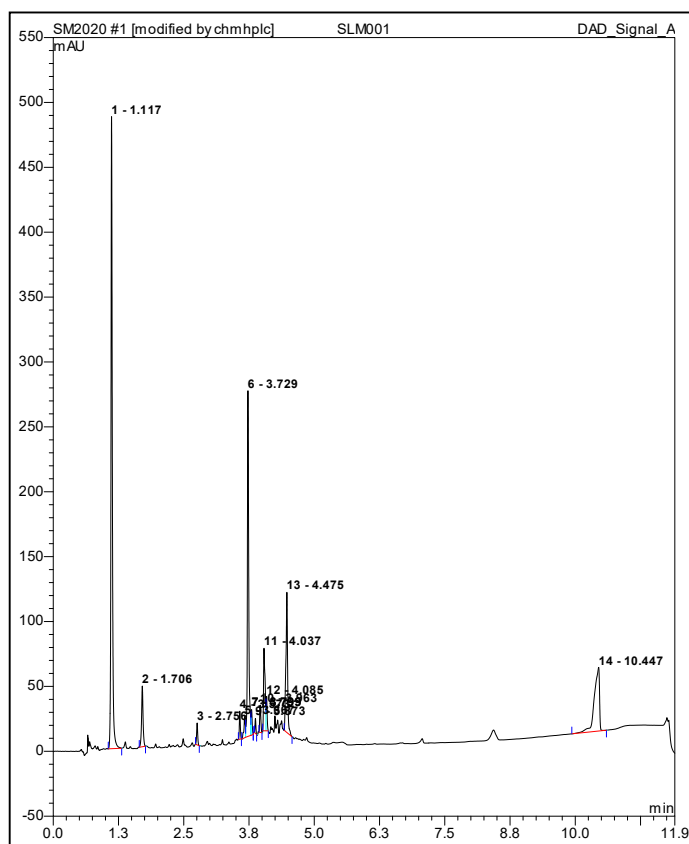
Purity= 37.72%

C.4 SNAP-PacificBlue-PEG-YQLKS (13)

C.4.1 SNAP-PacificBlue-PEG-YQLKS Structure
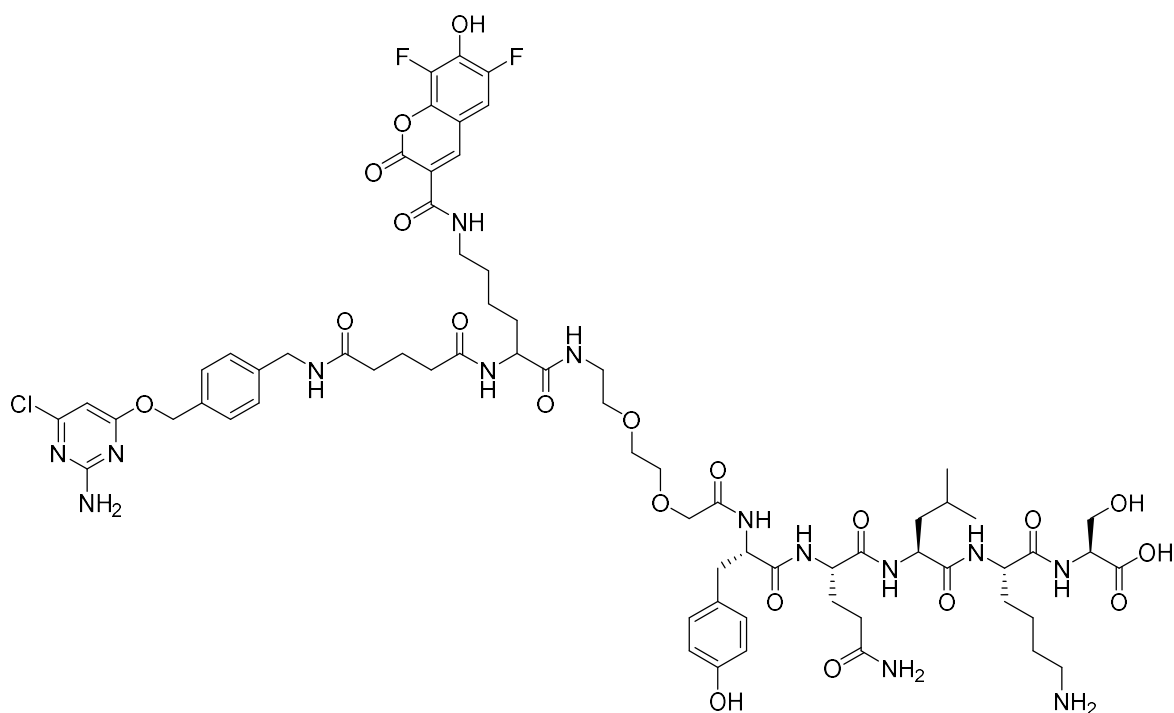
C.4.2 SNAP-PacificBlue-PEG-YQLKS HPLC
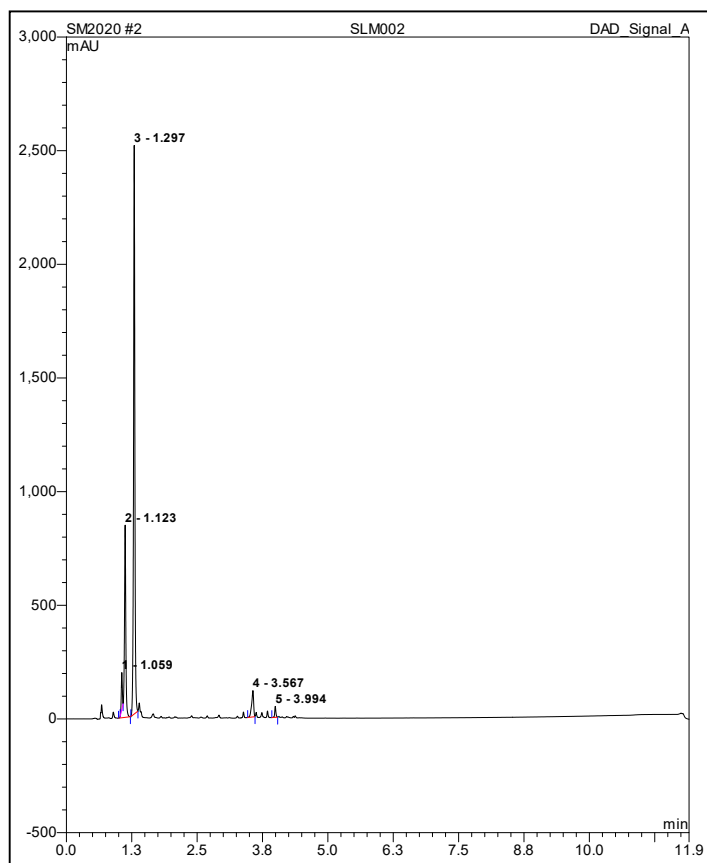
Purity= 66.00%

C.5 HaloTag-YQSKL (14)

C.5.1 HaloTag-YQSKL Structure
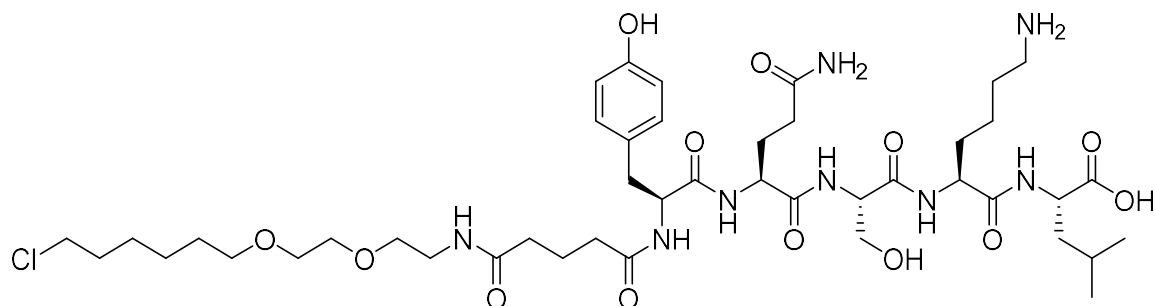
C.5.2 HaloTag-YQSKL HPLC
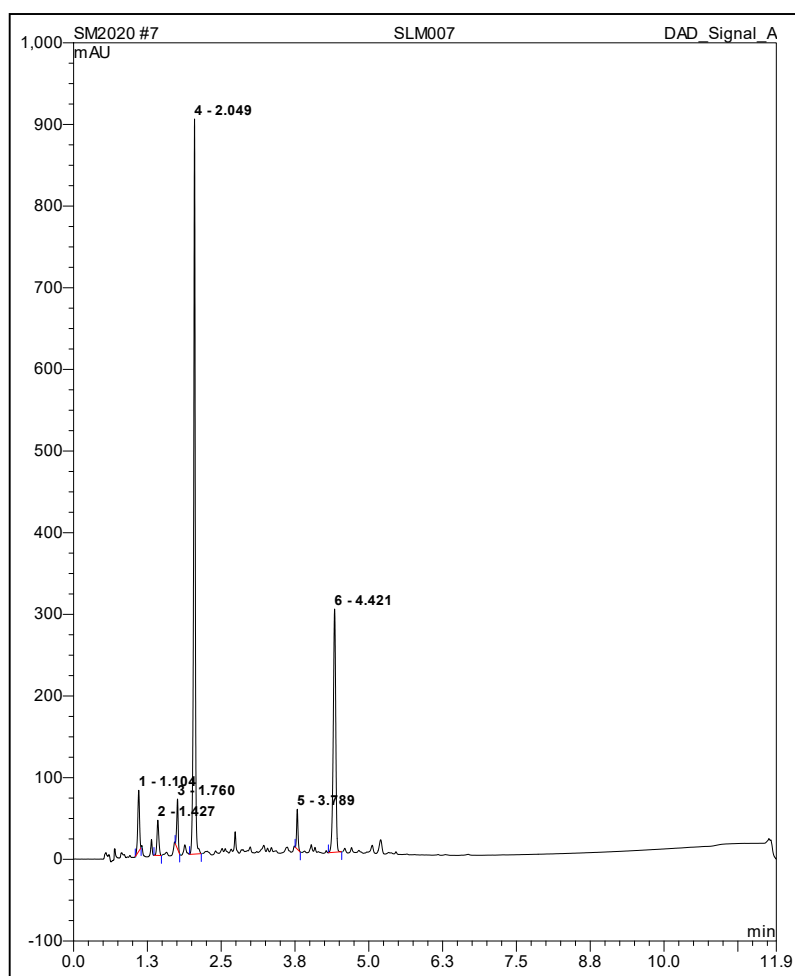
Purity= 56.39%

C.6 HaloTag-YQLKS (15)

C.6.1 HaloTag-YQLKS Structure
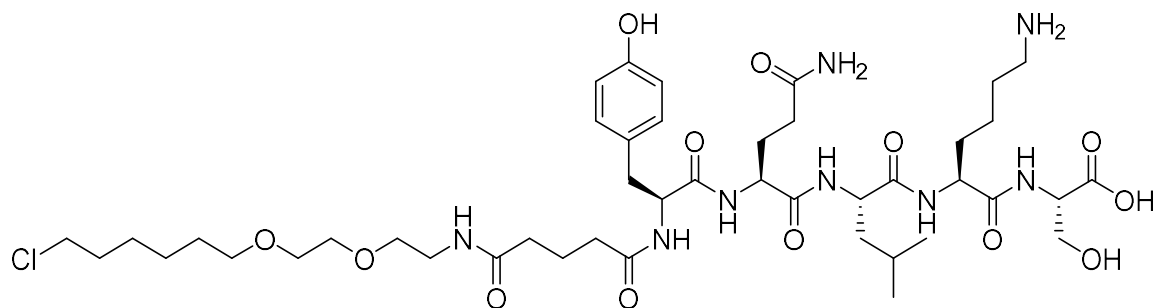
C.6.2 HaloTag-YQLKS HPLC
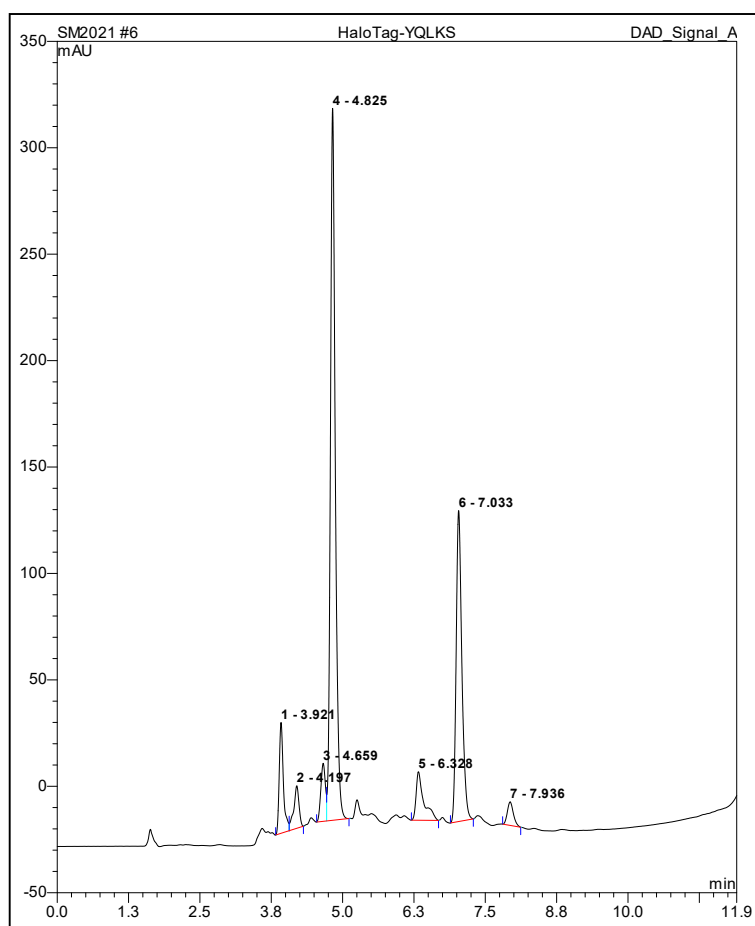
Purity= 51.87 %

C.7 HaloTag-PacificBlue-PEG-YQSKL (16)

C.7.1 HaloTag-PacificBlue-PEG-YQSKL Structure
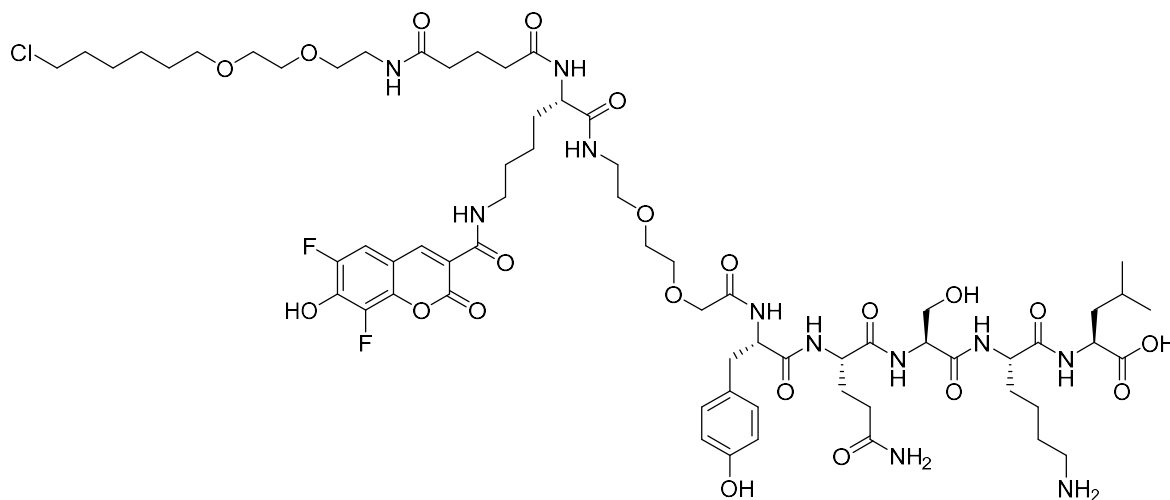
C.7.2 HaloTag-PacificBlue-PEG-YQSKL HPLC
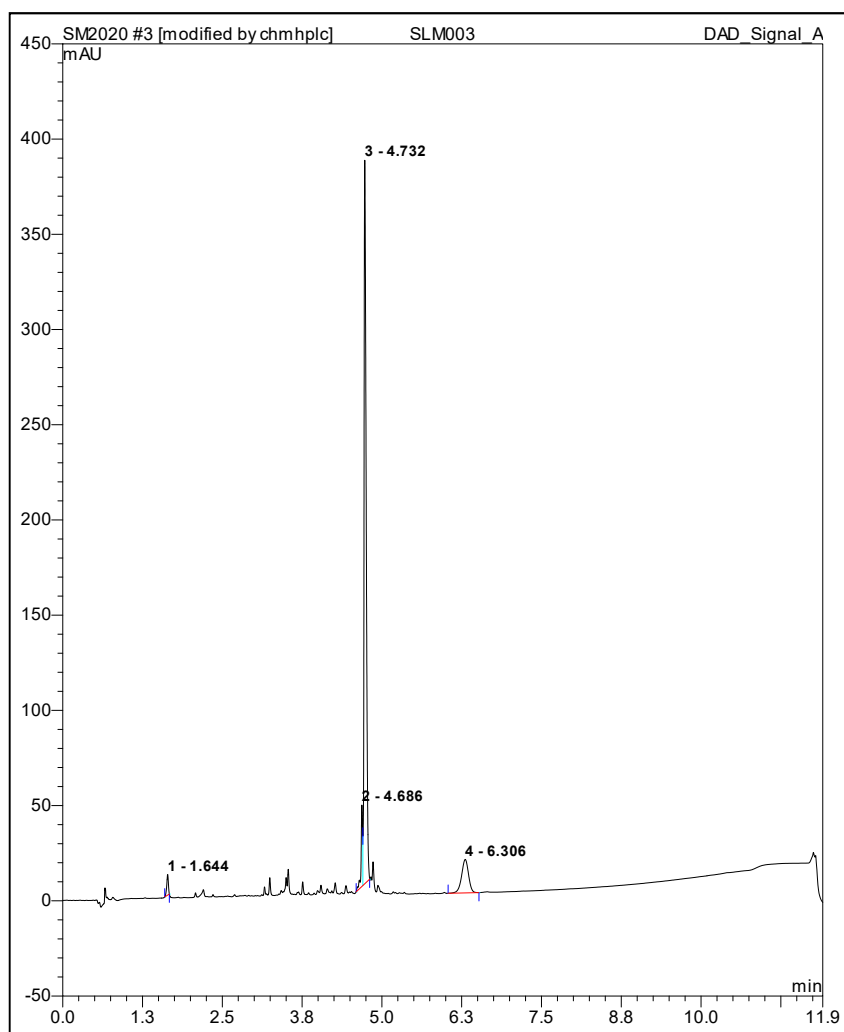
Purity= 79.96%

C.8 HaloTag-PacificBlue-PEG-YQLKS (17)

C.8.1 HaloTag-PacificBlue-PEG-YQLKS Structure
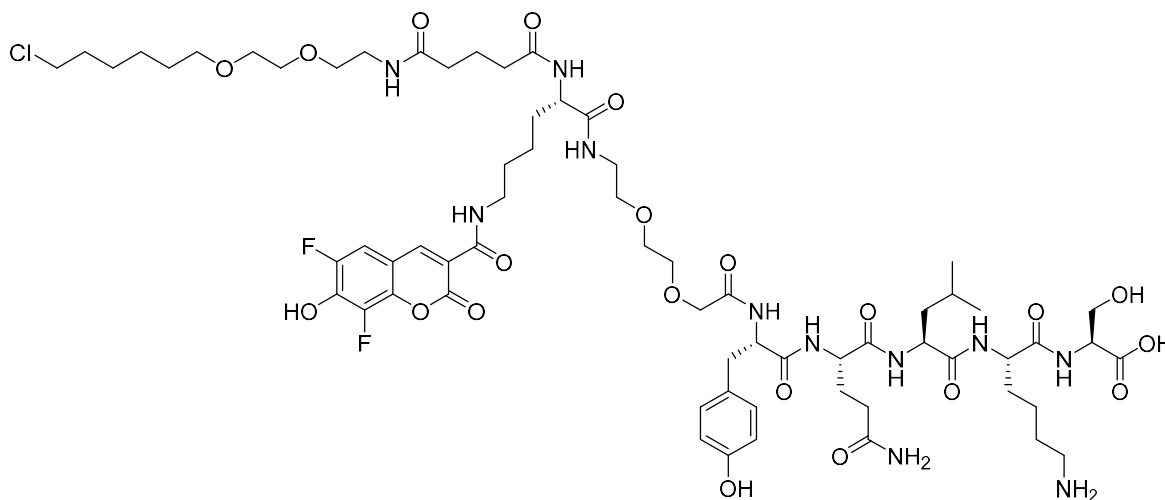

C.8.2 HaloTag-PacificBlue-PEG-YQLKS HPLC
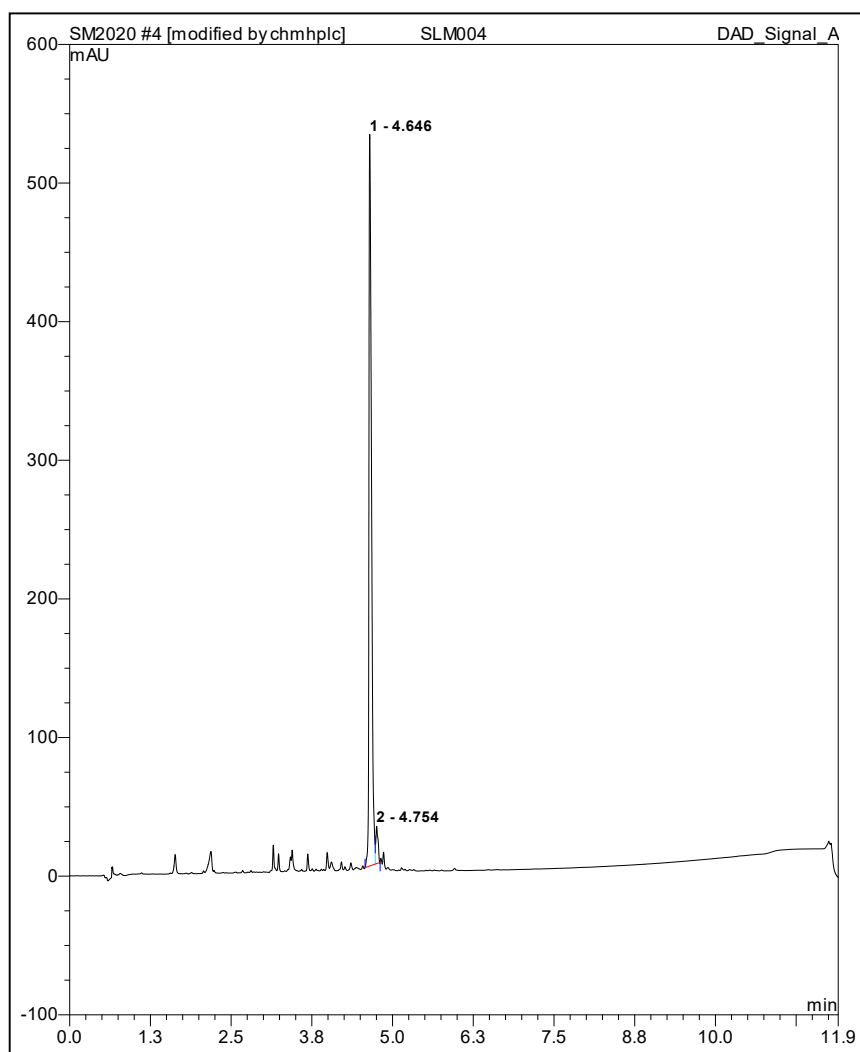
Purity= 95.71%

C.9 AcCl-No linker-YQSRL (18)

C.9.1 AcCl- No linker-YQSRL Structure
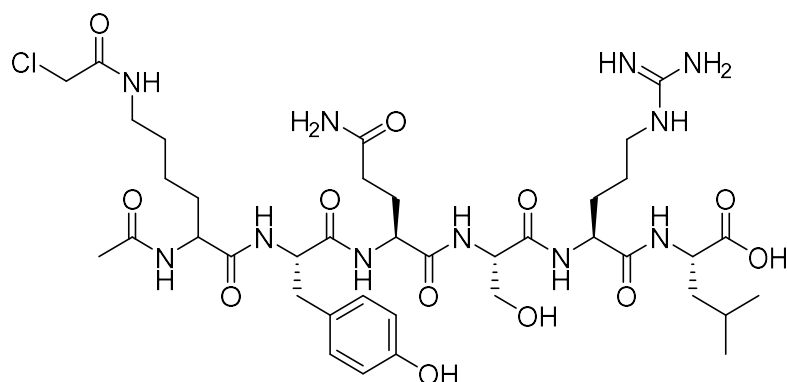
C.9.2 AcCl- No linker-YQSRL HPLC
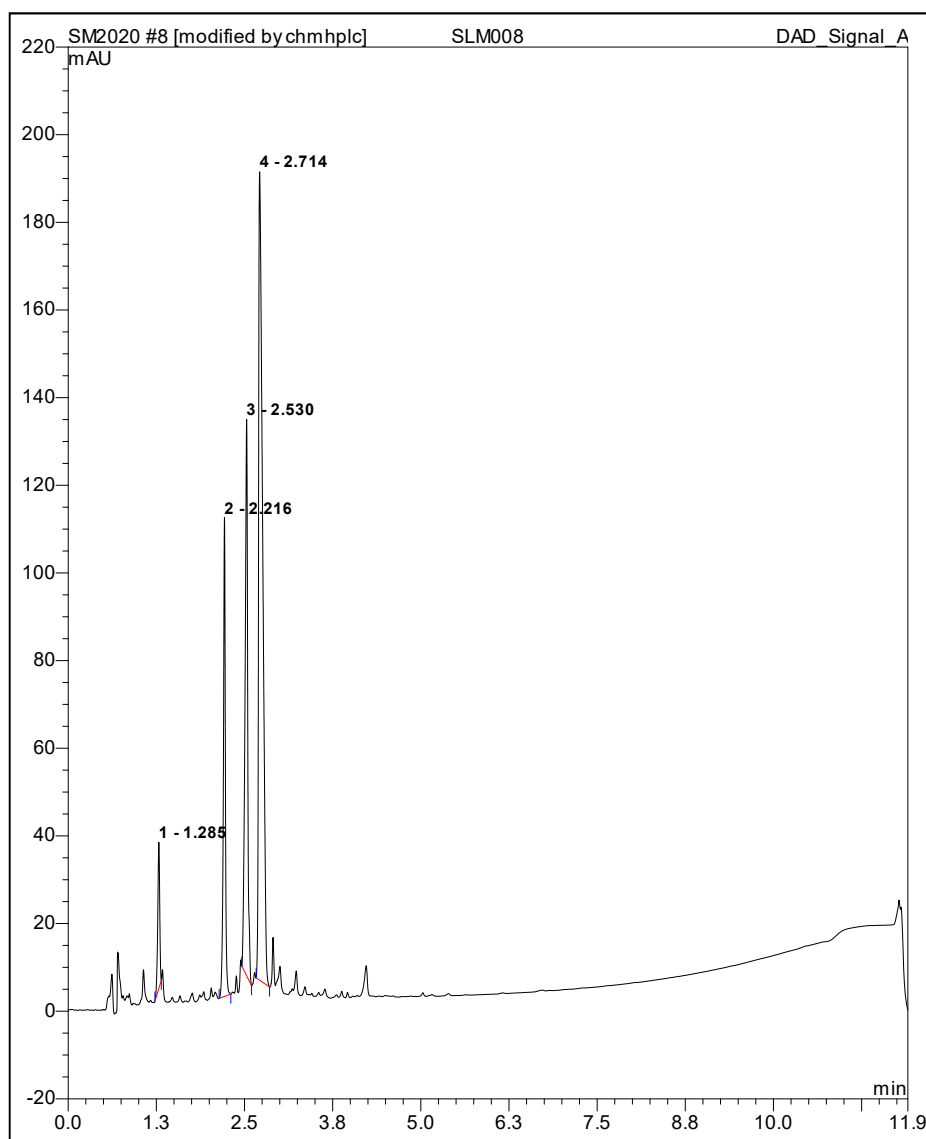
Purity= 54.53%

C.10 AcCl-PEG-YQSRL (19)

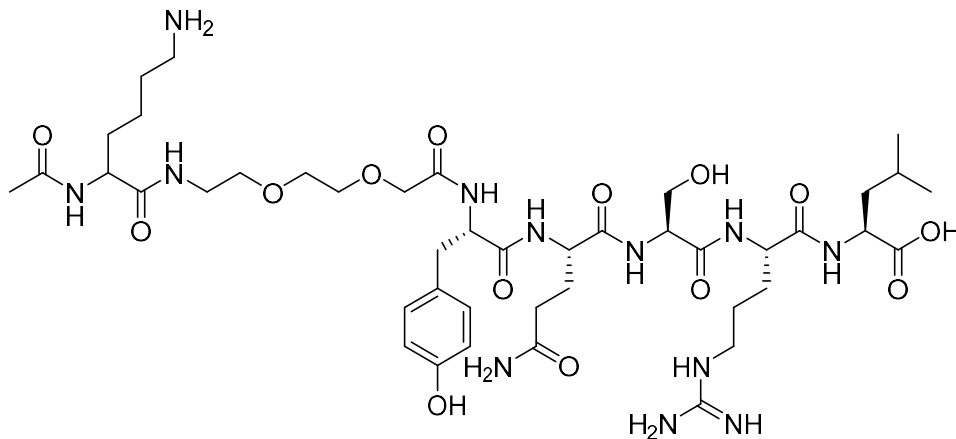### C.10.1 AcCl-PEG-YQSRL Structure

C.10.2 AcCl-PEG-YQSRL HPLC

Purity= 67.38%

C.11 AcCl-PEG₂-YQSRL (20)

C.11.1　　AcCl-PEG₂-YQSRL Structure
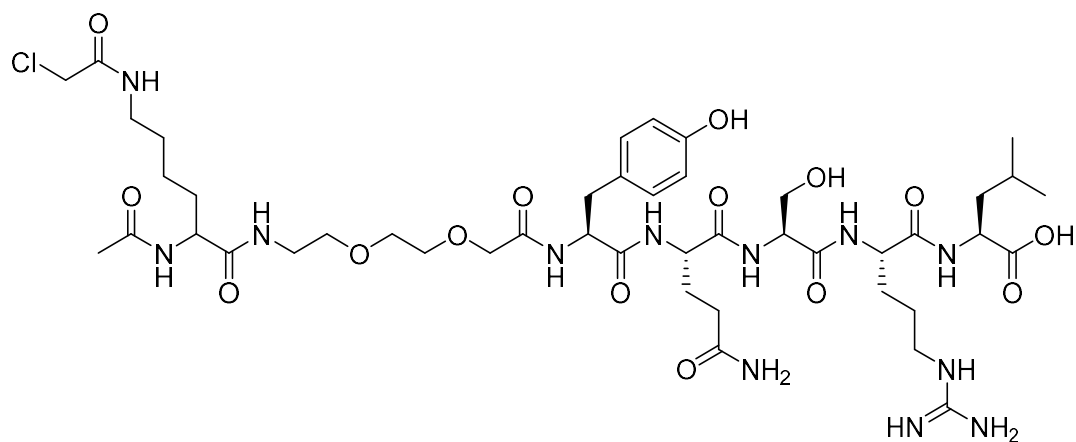
C.11.2　　AcCl-PEG₂-YQSRL HPLC
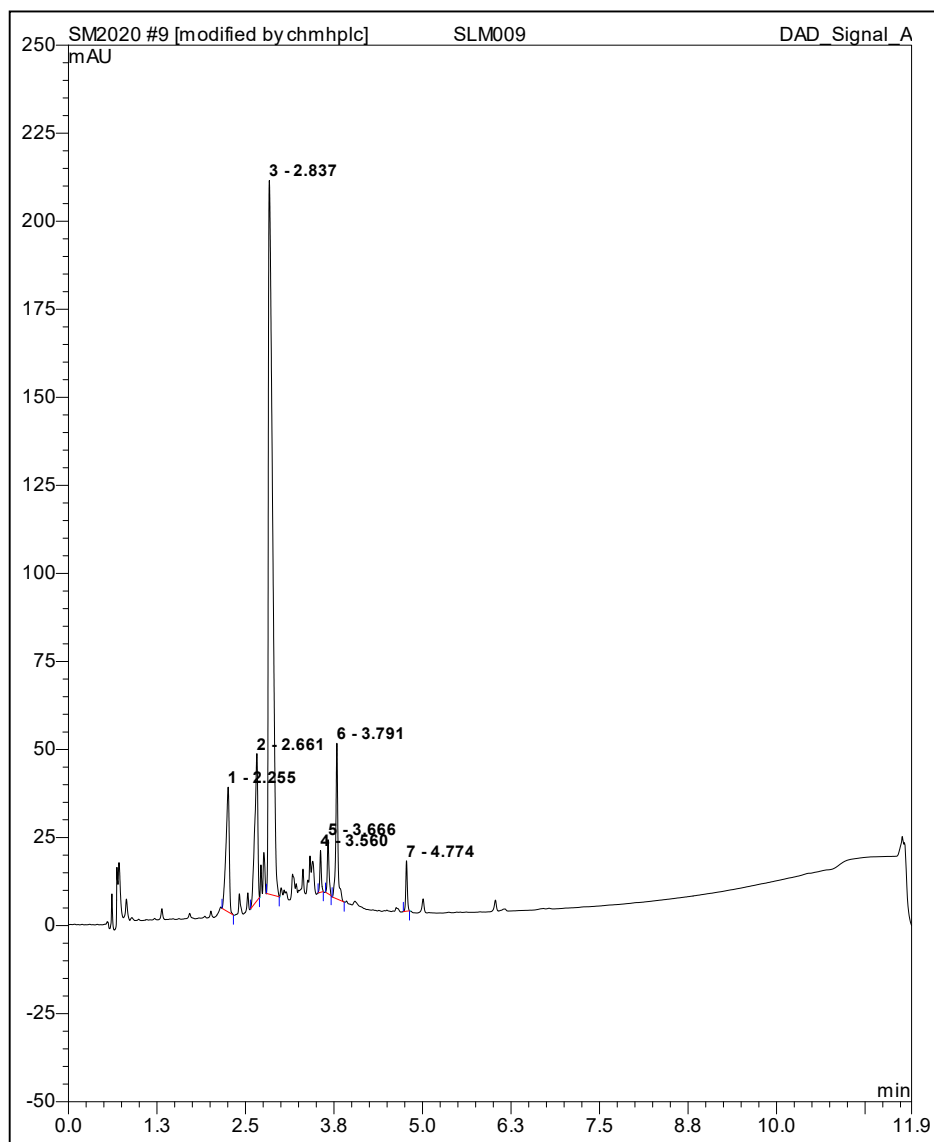
Purity= 69.94%

C.12 AcCl-PEG-YQLRS (21)

C.12.1   AcCl-PEG-YQLRS Structure
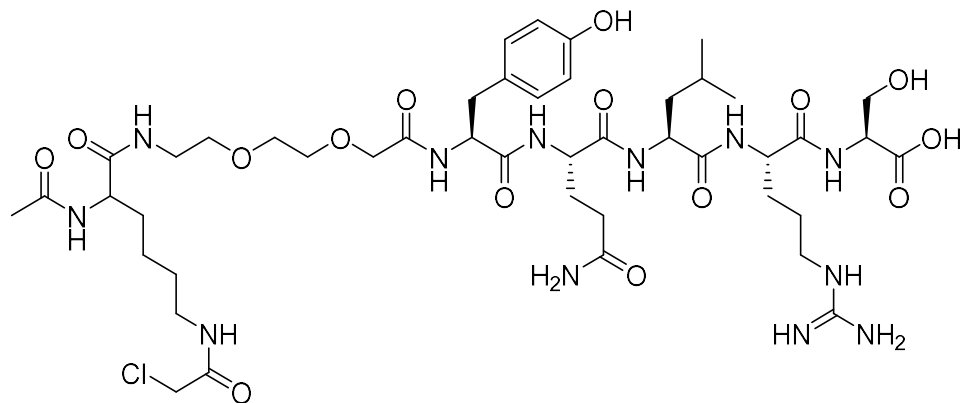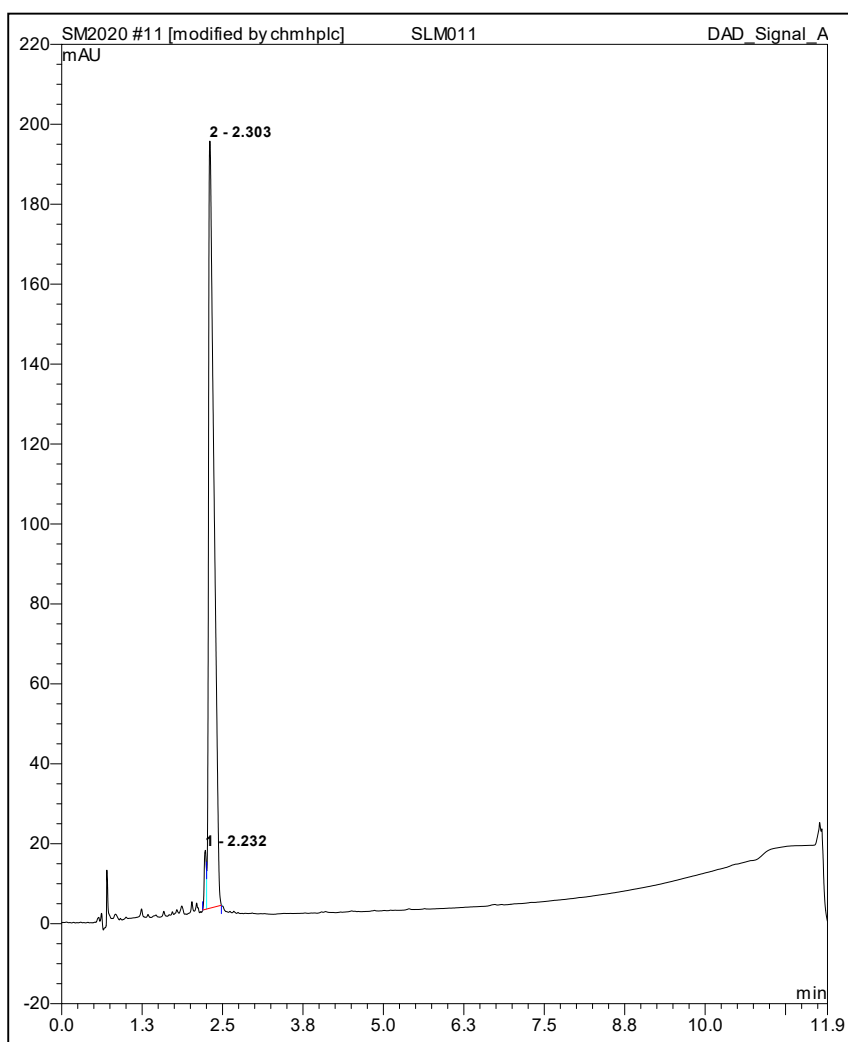

### C.12.2   AcCl-PEG-YQLRS HPLC
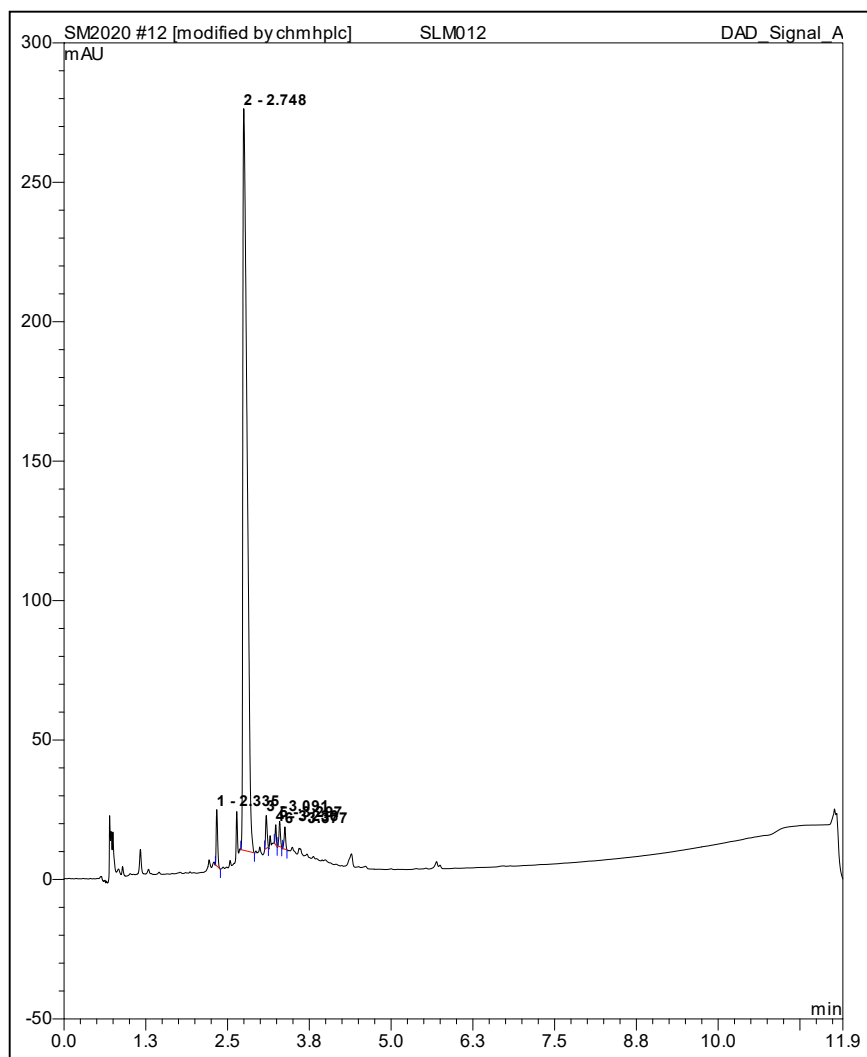
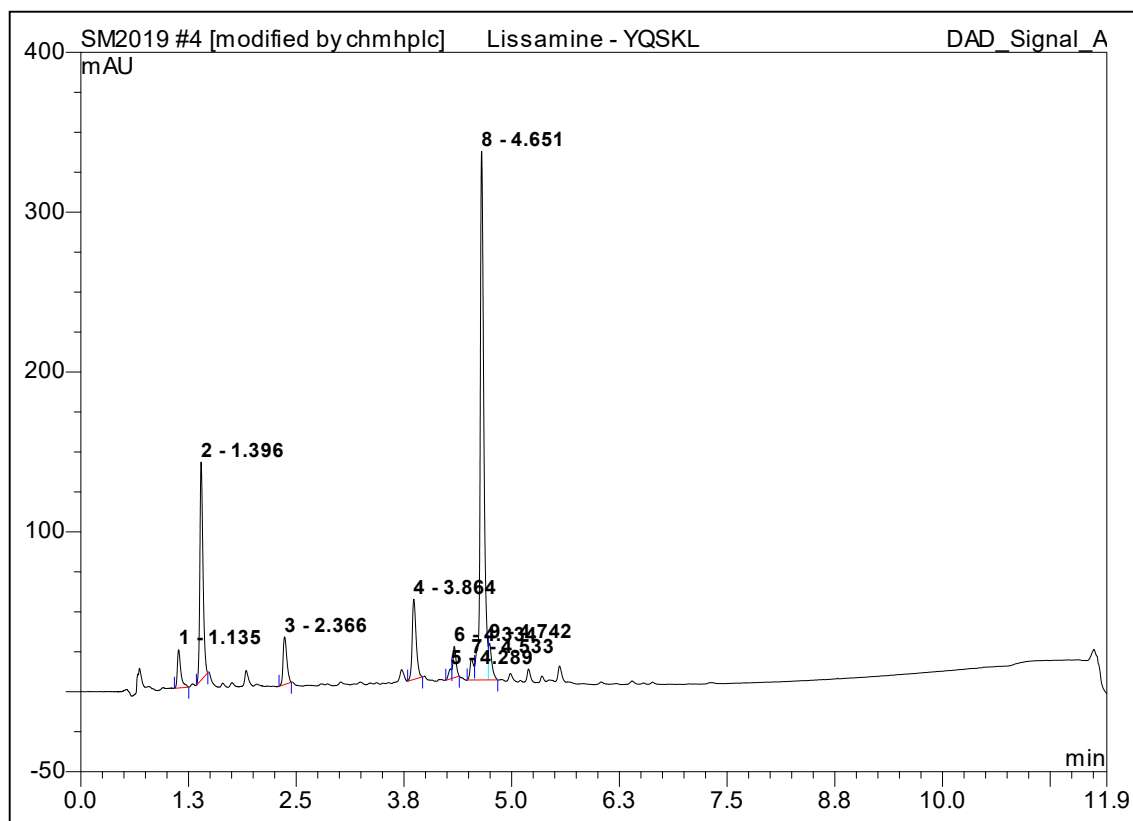Purity= 93.35%

C.13 Ac-PEG-YQSRL (22)

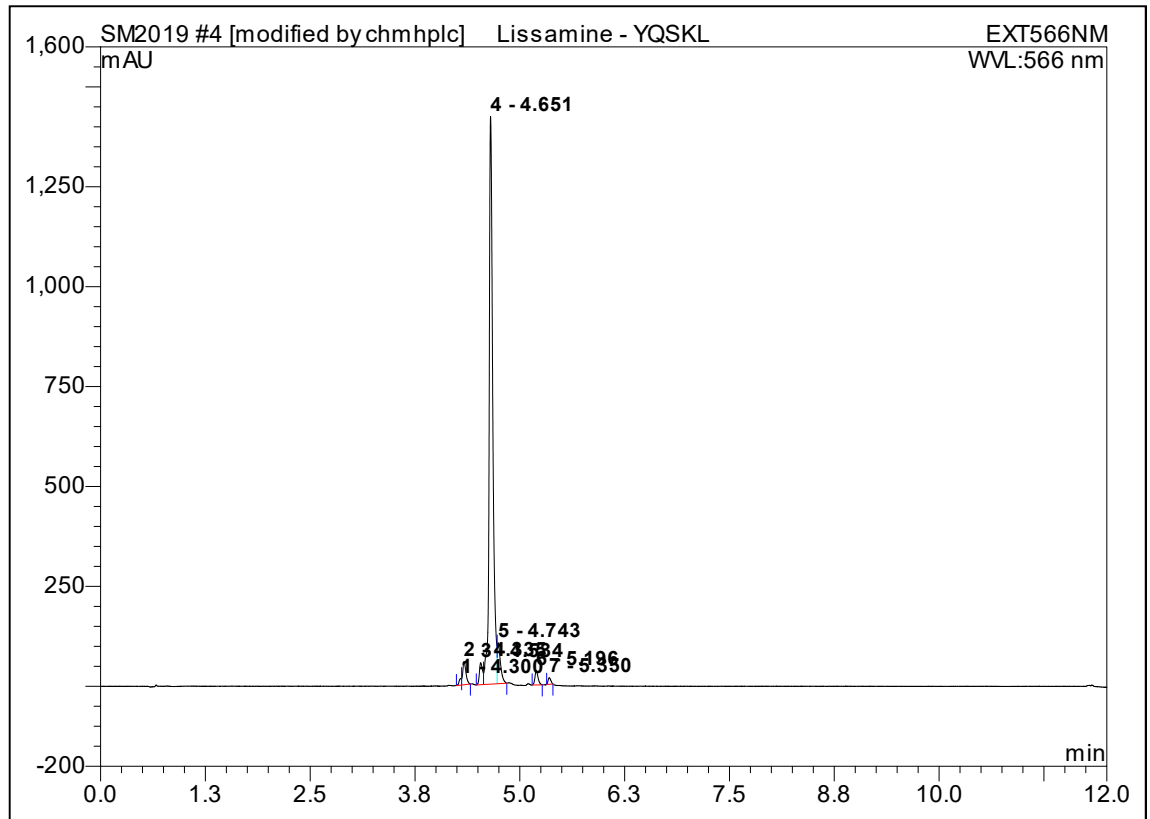### C.13.1   Unreactive-PEG-YQSRL Structure

C.13.2   Unreactive-PEG-YQSRL HPLC

Purity= 97.23%

C.14 HaloTag-PEG-AcCl-PEG-YQSRL (23)

C.14.1   HaloTag-PEG-AcCl-PEG-YQSRL

C.14.2   HaloTag-PEG-AcCl-PEG-YQSRL HPLC

Purity= 50.23%

C.15 HaloTag-PEG-AcCl-PEG-YQLRS (24)

C.15.1 HaloTag-PEG-AcCl-PEG-YQLRS Structure

C.15.2 HaloTag-PEG-AcCl-PEG-YQLRS HPLC

Purity= 44.61%

C.16 HaloTag-PEG-Ac-PEG-YQSRL (25)

C.16.1 HaloTag-PEG-Ac-PEG-YQSRL Structure

C.16.2 HaloTag-PEG-Ac-PEG-YQSRL HPLC

Purity= 72.78%

C.17 Lissamine-YQSKL (26)

C.17.1   Lissamine-YQSKL Structure
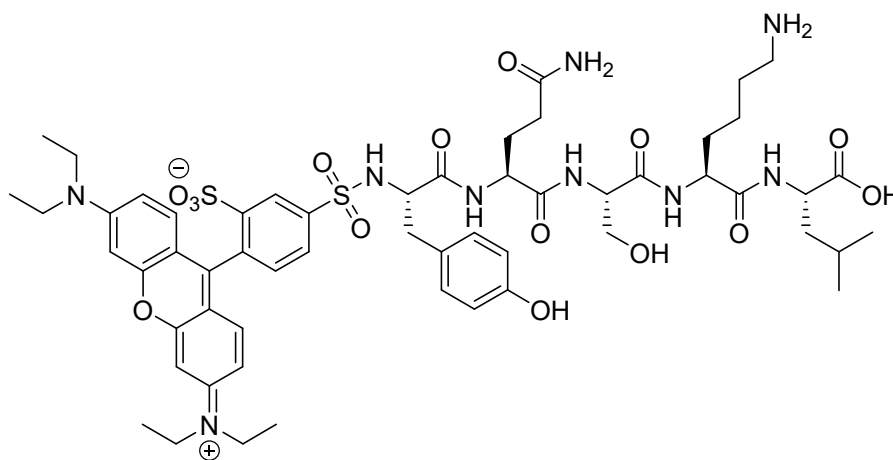
C.17.2   Lissamine-YQSKL HPLC

Purity = 54.59%

Purity= 85.48%