Structural and functional analysis of CD81-Claudin-1, a hepatitis C virus receptor complex

Michelle Clare

Doctor of Philosophy

Aston University

July, 2014

Aston University

Structural and functional analysis of CD81-Claudin-1, a hepatitis C virus receptor complex

Michelle Clare

Doctor of Philosophy

2014

Thesis Summary

Many viruses initiate infection through a multistep process involving host cell membrane proteins. Hepatitis C virus (HCV) is an important human pathogen that infects more than 185 million people worldwide and results in progressive liver disease. Recent advances have identified an essential role for tetraspanin CD81 and tight junction protein Claudin-1 in HCV entry into hepatocytes in the liver. CD81 associates with Claudin-1 and this complex is necessary for virus internalisation; defining the full length interface of this membrane protein interaction is therefore important for the design of future anti-viral therapies. Structural information is lacking for CD81: indeed, there is no high resolution structure for any full-length tetraspanin. This thesis describes an analysis of the protein-protein interaction interface between CD81 and Claudin-1 full-length proteins using a split-ubiquitin yeast assay. Also, using recombinant protein production of CD81, this thesis describes work towards successful crystallisation trials of a full length tetraspanin.

CD81 homotypic and heterotypic interactions with Claudin-1 were analysed in a high-throughput format in yeast, showing that this interaction is specific and does not require other mammalian cell factors. This work demonstrates that the CD81 large extracellular loop and its first transmembrane domain are involved in the CD81-Claudin-1 interaction: a novel full length molecular model predicted interacting amino acid residues that were confirmed *in vivo* using yeast assays. Thermal stability assays used to investigate recombinant membrane protein found that both detergent and buffer components are vital for the stability of recombinant CD81, which shows increased thermostability in the presence of cholesteryl hemisuccinate. Using the improved protein solution environment found here, and the increased understanding of the tetraspanin interaction interface; this work paves the way for CD81 structural characterisation alone or in combination with Claudin-1.

Keywords: Tetraspanin, tight junction protein, site-directed mutagenesis, *P. pastoris*, recombinant protein production.

This thesis is dedicated to James and our unborn baby.

We can't wait to meet you.

Acknowledgements

Contents

Abbreviations

AMP	Ampicillin
APS	Ammonium persulphate
BCA	Bicinchoninic acid
BMGY	Buffered complex glycerol medium
BMMY	Buffered complex methanol medium
B-OG	n-octyl-β-D-glucopyranoside
BSA	Bovine serum albumin
BSM	Basal salts medium
CD81	Cluster of differentiation 81
CHS	Cholesteryl hemisuccinate
CMC	Critical micelle concentration
CPM	N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide
Cub	C-terminal half of ubiquitin
DDM	n-dodecyl-β-D-maltopyranoside
DLS	Dynamic light scattering
DM	n-decyl-β-D-maltopyranoside
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dO	Dissolved oxygen concentration
E. coli	*Escherichia coli*
ECL1	First extracellular loop
ECL2	Second extracellular loop
EGFR	Epidermal growth factor receptor
Fos-10	Foscholine-10
FRET	Fluorescence resonance energy transfer
GFP	Green fluorescent protein
GPCR	Guanine nucleotide-binding protein coupled receptor
HCV	Hepatitis C virus
HCVcc	Hepatitis C virus cell culture
HCVpp	Hepatitis C virus pseudoparticles
HEPES	(2-[4-(2-hydroxyethyl) piperazin-1-yl]ethanesulfonic acid)
KAN	Kanamycin
LB	Luria-Bertani
LDL	Low density lipoprotein
LDLR	Low density lipoprotein receptor
LEL	Large extracellular loop
MNG	Maltose neopentyl glycol
MPL	Membrane Protein Laboratory
NANB	Non-A, non-B hepatitis
Ni-NTA	Nickel nitrilotriacetic acid
Nub	N-terminal half of ubiquitin
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline Tween 20
PCR	Polymerase chain reaction
pH	Negative logarithm of the hydrogen ion concentration
p-null	Palmitoylation-null
P. pastoris	*Pichia pastoris*
RNA	Ribonucleic acid
S. cerevisiae	*Saccharomyces cerevisiae*

SD	Synthetic defined
SDS	Sodium dodecyl sulphate
sE2	Soluble glycoprotein E2
SEC-MALS	size exclusion chromatography coupled with a multi angle light scattering
SEL	Small extracellular loop
SMA	Poly (Styrene-co-maleic acid)
SR-B1	Scavenger receptor class B type I
TEM	Tetraspanin enriched microdomain
TEMED	N,N,N',N'-tetramethyl- ethane-1,2-diamine
TM	Transmembrane domain
Tris	Tris(hydroxymethyl)aminoethane
Tween 20	Polyoxyethylene sorbitan monolaurate
UBP	Ubiquitin specific protease
VLDL	Very low density lipoprotein
WT	Wild type
w/v	Weight/volume
YPAD	Yeast peptone adenine dextrose
YPD	Yeast peptone dextrose
YNB	Yeast nitrogen base

Units

Å	Angstrom
°C	Celsius
g	Gram
g	Gravitational force
Hz	Hertz
h	Hour
kDa	Kilo Dalton
L	Litre
µg	Microgram
mg	Milligram
min	Minute
ml	Millilitre
mM	Millimole
M	Mole
nm	Nanometer
pmol	Picomole
psi	Pounds per square inch
RCF	Reactive centrifugal force
RFU	Relative fluorescence units
rpm	Revolutions per minute
sec	Second
U	Units
V	Volts
µL	Microlitre
µg	Microgram
ng	Nanogram

Amino acids

Table 1: Standard amino acid abbreviations. Standard amino acid abbreviations used interchangeably throughout text.

Amino acid	3-letter	1-letter	Amino acid	3-letter	1-letter
Alanine	Ala	A	**Leucine**	Leu	L
Arginine	Arg	R	**Lysine**	Lys	K
Asparagine	Asn	N	**Methionine**	Met	M
Aspartic acid	Asp	D	**Phenylalanine**	Phe	F
Cysteine	Cys	C	**Proline**	Pro	P
Glutamic acid	Glu	E	**Serine**	Ser	S
Glutamine	Gln	Q	**Threonine**	Thr	T
Glycine	Gly	G	**Tryptophan**	Trp	W
Histidine	His	H	**Tyrosine**	Tyr	Y
Isoleucine	Ile	I	**Valine**	Val	V

List of figures

List of tables

1. Introduction

1.1 Hepatitis C virus (HCV)

Viruses can be defined as *"obligatory intracellular parasites, whose replication depends on pathways and functions of the host cell"* (Krausslich & Bartenschlager, 2009).

Viruses rely on a host eukaryotic or prokaryotic cell to replicate its genome, outside of the host cell, viruses are inert. Viruses contain either a RNA or DNA genome that is surrounded by a virus encoded protein coat. Classification of viruses depends on their size, shape, the structure of the genome and how they undergo replication (Gelderblom, 1996). For example, there are around 21 different families of viruses that can infect humans.

The hepatitis C virus (HCV) currently infects more than 185 million people worldwide (Hill and Cooke, 2014). HCV is transmitted through intravenous drug use or *via* blood transfusion, although post-transfusion acquired infection is now less common in developed parts of the world owing to donor screening for markers of HCV (Donahue et al., 1992). Typically, around 20% of those infected with HCV have acute hepatitis and clear infection. In contrast, approximately 80% of HCV infected individuals suffer chronic hepatitis. This involves progressive liver injury and an increased risk of developing fibrosis or hepatocellular carcinoma (Timpe and McKeating, 2008). Chronic hepatitis can arise due to the virus' ability of escaping the host immune response. Much of the biology and virology of HCV infection is understood (Timpe and McKeating, 2008), but the process by which the virus enters human liver cells is still incompletely described.

Throughout the 1970's, the responsible agent(s) for post-transfusion hepatitis was unknown. The disease, during this period, was referred to as non-A, non-B hepatitis (NANB). This name occurred as diagnostic tests for hepatitis A and B viruses were used to eliminate them as the cause for hepatitis (Zuckerman et al., 2001). It was not until 1985 when Bradley and others found that the causative agent of NANB was a small, enveloped virus (Bradley et al., 1985). Later, with the use of molecular

techniques, Choo *et al* (1990) characterised the genome of this virus, which was then referred to as HCV (Choo et al., 1990).

Chronic hepatitis caused by HCV is usually treated with a combination therapy of pegylated interferon (peginterferon) plus ribavirin. A clinical study suggested that out of the two commercially available forms of peginterferon (alfa-2a and alfa-2b); the combination of peginterferon alfa-2a with ribavirin demonstrated the most success in treating chronic HCV infection (Ascione et al., 2010). Unfortunately there is, at present, no successful vaccine against HCV and combination therapy has its disadvantages such as its high cost, severe side effects and limited efficacy (McHutchison et al., 2006). Hepatitis C is typically diagnosed at a late stage when signs of liver damage are apparent, therefore reducing the success rate of recovery. Clearly, new therapies against HCV are required that act early on, potentially targeting HCV entry into host cells (Timpe and McKeating, 2008).

1.1.1 Viral features important for host infectivity

HCV is a member of the *Hepacivirus* genus in the *Flaviviridae* family. The HCV genome is a positive sense, single-stranded RNA comprised of around 10,000 nucleotides (Choo et al. 1990). The genomic material is contained within an icosahedral nucleocapsid shell, which is enclosed within a host-derived membrane envelope (see figure 1.1). The envelope holds two virus-encoded glycoproteins E1 and E2 (Timpe and McKeating, 2008). The HCV genome encodes a single polyprotein, which produces ten viral proteins once cleaved by both host and viral proteases. Three of the proteins that are produced are structural proteins, including core, E1 and E2. The remaining seven proteins consist of a small ion channel protein (p7) and six non-structural proteins (Dubuisson, 2007).

HCV encoded glycoproteins, E1 and E2, which sit within the virus envelope play a critical role in viral entry (Timpe and McKeating, 2008) (figure 1.1). E1 and E2 bind specific host cell surface receptors, which allow the virus to enter the cell *via* clathrin mediated endocytosis. Following entry, the viral and endosomal membranes fuse in a pH dependent fashion, this leads to the release of the capsid and genome into the

cytoplasm. This viral genome is replicated and expressed within the host cell in order for virions to be produced. Subsequently virions are released from the cell to infect others, leading to the repetition of the HCV life cycle.

Aston University

Illustration removed for copyright restrictions

Figure 1.1: Hepatitis C virus structure. The core contains viral genetic material (RNA) surrounded by a protective protein shell. A host-derived lipid envelope surrounds the protein capsid and holds two viral glycoproteins E1 and E2 that are crucial for viral entry into a host cell. Picture taken from Sciencephoto.com.

Great progress has been made in understanding the virus interaction with the host cell surface, in order to gain entry. Following the characterisation of the HCV genome in 1989, studies that looked into HCV entry were complicated for over 15 years due to the deficiency of *in vitro* culture systems, that were not able to demonstrate efficient virus replication (Timpe and McKeating, 2008). This made research difficult as levels of infectious virus found in serum or liver tissue were insufficient for study (Fournier et al., 1998). Even so, few studies using primary cultures of human hepatocytes demonstrated *in vitro* HCV infection and HCV replication after being incubated with human serum samples from HCV carriers (Fournier et al., 1998), (Castet et al., 2002). Since then, the development of suitable *in vitro* models, have enabled major advancements into the study of HCV entry and its life cycle.

1.1.2 HCV entry and life cycle *in vitro* models

The first advance for an *in vitro* model of HCV entry used a recombinant soluble glycoprotein E2 (sE2). Virus-host cell surface interactions have been studied using sE2 as a model to discover host cell surface receptors responsible for viral entry (Flint et al., 1999). sE2 production, first shown by Selby & colleagues (1994) was achieved by removal of the C-terminus of E2, rendering it soluble (Selby et al., 1994). The tetraspanin CD81 was initially discovered as a HCV host receptor when sE2 was used as a probe (Pileri et al., 1998). Furthermore, neutralising antibody against HCV E2 is associated with the prevention of HCV infection (Rosa et al., 1996). Even though sE2 is a worthy model, in reality E2 does not exist alone on the HCV surface but in complex with E1, therefore it may not exhibit true interactions with the host cell surface.

Later research, in 2003, developed a further working model so that the E1E2 heterodimer was studied *in vitro*. This involved using retroviral pseudoparticles that carried HCV glycoproteins E1E2 at its surface, which were referred to as HCV pseudoparticles (HCVpp) (Bartosch et al., 2003), (Drummer et al., 2003), (Hsu et al., 2003). HCVpp were designed to carry a reporter gene such as green fluorescent protein (GFP) or luciferase so that entry was possible to measure. Using HCVpp, primary hepatocytes were shown to be a key target for infection and that both E1 and E2 were required (Bartosch et al. 2003). HCVpp is a valuable tool to study the initial events of HCV entry and infectivity, although limitations come into play once the particle is internalised. Retroviral proteins subsequently control and dictate remaining replication events and so do not model HCV *in vitro* replication (Timpe and McKeating, 2008).

In 2005, the revolution for HCV research came about when an exceptional HCV strain was discovered that had the ability to replicate and release infectious virions in culture conditions using a human hepatoma cell line (Huh7) (Wakita et al., 2005), (Zhong et al., 2005). This was an isolated strain from a Japanese patient with fulminant hepatitis C, so called JFH-1. The *in vitro* generated infectious virions were shown to be infectious to chimpanzees (Wakita et al. 2005). Consequently, a model, referred to as HCV cell culture (HCVcc), was now available to study the complete life cycle of HCV. Consequently, this model generated evidence to prove the involvement of multiple host

cell surface receptors and factors in HCV entry and provided a tool to focus on the generation of antiviral therapies.

1.2 Host proteins involved in the HCV entry process

HCV enters cells *via* a coordinated series of events that involves viral and host cell factors (Zeisel et al., 2011) (see figure 1.2). It is biologically crucial for the virus to gain entrance to its host cell, so mechanisms have evolved that allow passage through the host's plasma membrane that are highly specific and characterise the type of host and cell that is targeted by the virus. HCV circulates in the blood and directly makes contact with the basolateral membrane of hepatocytes, the cell type that it can enter and replicate inside. Agnello *et al* (1999) provided evidence to confirm that the low density lipoprotein receptor (LDLR) is an important factor in viral entry (Agnello et al., 1999). The group suggested HCV can indirectly interact with lipoprotein receptors due to its attachment to very low density lipoprotein (VLDL) or low density lipoprotein (LDL) prior to endocytosis. After initial attachment of HCV to the host hepatocyte cell, other cell surface molecules are required for efficient viral entry.

Significant cell surface receptors involved in HCV entry include CD81 human tetraspanin (Pileri *et al*., 1998) and scavenger receptor class B type I (SR-BI) (Scarselli et al., 2002), both proving to bind HCV envelope glycoprotein E2. Evans *et al* (2007) later discovered another membrane protein essential for HCV entry (see figure 1.2). This was Claudin-1, a tight junction protein, suggesting it is involved in a late stage of HCV entry (Evans *et al*., 2007). Another tight junction protein important in viral entry is Occludin, which like Claudin-1 acts at a post-attachment step. Further evidence is required to determine if Occludin directly binds HCV particles and its exact role in the entry process (Lindenbach and Rice, 2013). Cell signalling events, which are initiated by E2 binding with CD81, result in the lateral movement of CD81-HCV complexes to sites of cell-cell contact. CD81 interacts with Claudin-1 and these co-receptor complexes are required for HCV internalisation *via* clathrin-mediated endocytosis (Lindenbach and Rice, 2013) (figure 1.2).

Figure 1.2: The HCV entry process. Multiple cells surface molecules are involved in HCV attachment to the cell surface and subsequent HCV entry. The interaction of HCV with lipoprotein particles initiates the initial attachment of HCV to the low density lipoprotein receptor (LDLR) and SR-B1. The virus glycoprotein E2 then directly binds to CD81 and this activates signal transduction through the epidermal growth factor receptor (EGFR) and Ras. This promotes lateral movement of CD81 in the hepatocyte plasma membrane to cell-cell contacts, where CD81 bind Claudin-1. CD81-Claudin-1 complexes result in HCV internalisation *via* clathrin-mediated endocytosis. Schematic taken from Lindenbach and Rice, 2013.

1.2.1 The tetraspanin family

Tetraspanins are a family of membrane proteins found in organisms such as *C. elegans*, *Drosophila* and mammals (Levy et al., 1998). They are relatively small membrane proteins, typically around 200-300 amino acids in size. In humans, 33 tetraspanins are known to exist with some such as CD9, CD81 and CD151 being expressed in many different cell types, although not ubiquitous, whilst others are mostly found in one

particular cell type. Tetraspanins are involved in a variety of biological processes that include immune responses, sperm-egg fusion, host-parasite interactions and cell migration and metastasis (Rubinstein, 2011).

A unique feature, common to all tetraspanins, is the lateral associations they form in the plasma membrane with each other, as well as with non-tetraspanin proteins. Tetraspanin function correlates with their ability to interact with one another as well as with other proteins. This interacting network constitutes a so-called 'tetraspanin web' or a tetraspanin enriched microdomain (TEM), (Hemler, 2003), (Levy and Shoham, 2005). The tetraspanin web acts as a signalling platform that coordinates such cellular processes and it has been suggested that tetraspanin homodimers are the smallest unit that make up the tetraspanin web (Charrin et al., 2009). Imaging techniques have shown that interactions do not only occur within tetraspanin webs but also outside of these microdomains, with tetraspanins able to exchange between the two environments.

There is accumulating evidence that cholesterol plays a significant role in the organisation and functioning of the tetraspanin web. A physical interaction between tetraspanins and cholesterol is likely to exist. In one study, CD9, CD81 and CD82 were labelled with a cholesterol analogue, [3H]photocholesterol, suggesting the tetraspanin and lipid exist in close proximity (Charrin et al., 2003). Furthermore, digitonin, a cholesterol precipitating reagent, was found to disrupt tetraspanin-tetraspanin interactions, suggesting tetraspanins were precipitated by digitonin due to their close association with cholesterol. Since cholesterol has an effect on tetraspanin protein interactions it therefore can modulate processes that rely on such protein-protein interactions. For example, Harris *et al* (2010) reported that removal of cholesterol from the plasma membrane using MβCD treatment reduced Claudin-1 expression at the membrane and its interaction with partner protein, tetraspanin CD81. Since CD81-Claudin-1 interactions form a HCV entry receptor, the depletion of cholesterol caused a reduction in HCV entry into liver hepatocytes (Harris et al., 2010).

Figure 1.3: Tetraspanin structure. a) Schematic of the tetraspanin topology showing four transmembrane domains, intracellular N- and C- terminus and two extracellular loops (SEL and LEL). The LEL has a structurally variable region and a constant region throughout the tetraspanin family. b) The CD81 LEL crystal structure (Kitadokoro et al., 2001) shows a soluble homodimer, each monomer composed of 5 α-helices (A-E). The LEL variable region is made up of C and D helices and the constant region A, B and E helices.

Tetraspanins can be distinguished from other four transmembrane domain proteins by their unique structural characteristics. Tetraspanins are composed of four transmembrane helices and two extracellular loops (the first is a small extracellular loop (SEL) and the second is a large extracellular loop (LEL)) and intracellular N- and C-terminal tails (see figure 1.3). A major milestone in structural tetraspanin research was in 2001: Kitadokoro and others successfully solved the crystal structure of a homodimer of human CD81 LEL (Kitadokoro et al., 2001), (Kitadokoro et al., 2002) confirming by sequence analysis that the key structural features of the LEL are conserved throughout the tetraspanin family. The X-ray crystallography data illustrated an antiparallel homodimer with each monomer being made up of five α-helices (A-E) stabilised by two disulphide bridges (Kitadokoro et al., 2001) (as shown in figure 1.3). The LEL has two important regions defined as a variable region (also known as the head domain) comprised of C and D helices that has shown involvement in tetraspanin interactions with partner proteins and ligands such as CD81 binding HCV glycoprotein E2 and a highly conserved region (known as the stalk domain) comprising A, B and E helices

that was suggested to be responsible for tetraspanin homodimerisation (Stipp et al., 2003).

Common throughout the tetraspanin superfamily is the conserved cysteine-cysteine-glycine (CCG) motif that exists within the large extracellular loop (LEL) along with 2-4 disulphide bridges. Tetraspanins also have conserved cysteine residues in the cytoplasmic domain of the protein, acting as palmitoylation sites (Seigneuret et al. 2001).

Tetraspanin conserved transmembrane domains are involved in tetraspanin intra- and intermolecular interactions, with TM1 and TM2 suggested to be involved in tetraspanin dimeric interactions (Kovalenko et al., 2005). The SEL has been shown to influence tetraspanin cell surface expression, in particular for tetraspanin CD81 (Masciopinto et al., 2001).

Other tetraspanin family members, uroplakins Ia and Ib have their three-dimensional structure solved to 6 Å resolution. They were shown to form a hexameric structure of heterodimers and primary interactions with other membrane proteins form through tight interactions *via* their transmembrane and extracellular domains (Min et al. 2006), an aspect that may be common throughout the tetraspanin family.

1.2.1.1 CD81

CD81 was first discovered in the late 1980s, when it was initially named target of anti-proliferative antibody 1 (TAPA-1) and later renamed as CD81 (cluster of differentiation 81) (Levy et al., 1998) . CD81 is a 26kDa membrane protein consisting of four transmembrane domains, a short intracellular loop, and two extracellular loops referred to as small extracellular loop (SEL) and large extracellular loop (LEL) that support two disulphide bridges (see figure 1.4). CD81 is widely expressed in mammals including cells such as hepatocytes, epithelial cells, B lymphocytes, but is not expressed in erythrocytes or neutrophils. Multiple biological functions of CD81 exist in humans including cell adhesion, tissue differentiation, immune cell maturation and host-parasite interactions (Hemler 2003).

Tetraspanin functions are dependent on their association with tetraspanins or other partner proteins. CD81 interacts with many different host partner proteins such as integrin α4β1, immunoglobulin family members EWI-F and EWI-2. CD81 is also well documented to be involved in pathogen infection by viruses, bacteria and parasites (Zona et al., 2014). The work in this thesis focuses on CD81 as a HCV entry receptor, into its target host cell, the hepatocyte.

Pileri and colleagues in 1998 discovered the initial HCV cell surface receptor as tetraspanin CD81, using sE2 in their investigation. Interestingly, sE2 was observed to competitively inhibit the binding of anti-CD81 to EBV-B cell lines (Pileri et al. 1998). HCV entry into hepatocytes is dependent on CD81. This was shown experimentally using a HepG2 cell line, which does not normally express CD81, but becomes permissive to HCVpp infection when transfected to express CD81 (Flint et al., 2006). The opposite scheme was true, when small interfering RNAs were used to silence CD81 expression in Huh-7.5 cells: HCVpp infectivity was hampered (Zhang et al., 2004).
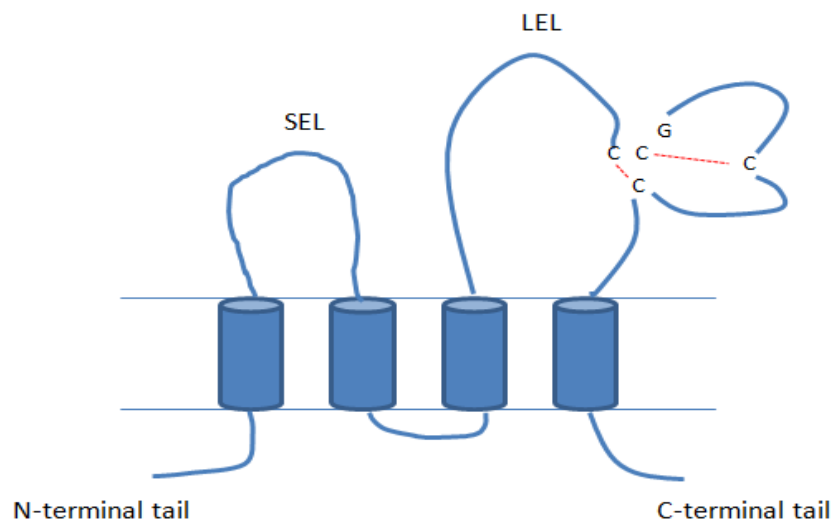
Figure 1.4: CD81 structure. CD81 is composed of four transmembrane domains; two extracellular loops referred to as the SEL and the LEL and intracellular N- and C- terminal domains. The LEL shows four cysteines and the CCG motif that is characteristic of tetraspanin family members.

The HCV entry process into hepatocytes is dependent on CD81 LEL (figure 1.2). A soluble recombinant form of human CD81 LEL binds HCV glycoprotein E2 as shown by protein immunoblot (Pileri et al. 1998). Also, full-length recombinant CD81 expressed in *Pichia pastoris* bound HCV E2 (Jamshad et al. 2008). Further studies

25

provide evidence for the necessity of CD81 LEL for viral entry, showing that *in vitro* HCVpp and HCVcc infection were inhibited by recombinant CD81 LEL and by monoclonal antibodies against E2 and CD81 (Zhong et al. 2005), (Zhang et al., 2004), (Lavillette et al., 2005).

Mutagenesis studies have revealed particular residues that are key for sE2 binding to CD81 LEL. Drummer and co-workers (2002) suggested that residues L162, I182, N184 and F186 in CD81 LEL are important for binding sE2 (Drummer et al., 2002). Furthermore, Bertaux & Dragic (2006) compared sequences of human and mouse CD81 to find that residues K171, I181, I182 and F186 were important for sE2 interactions with CD81 in addition to HCVpp entry.

The CD81 LEL was structurally solved as a soluble homodimer, which is the only solved structure available for any member of the tetraspanin family (Kitadokoro et al., 2001) (figure 1.3b). There is also a full length CD81 molecular model available, generated using information from the structure of soluble LEL and using homology modelling for the other regions in the full-length protein (Seigneuret et al., 2001). There is some debate as to how functionally relevant the LEL crystal structure is when used to predict the full length CD81 structure for a few reasons. Firstly, dimeric forms of CD81 are believed to be a significant basic unit as shown in the LEL crystal structure and the form involved in HCV E2 binding (Kitadokoro et al., 2002), (Drummer et al., 2005) although, more recent studies using recombinant full length CD81 show evidence that the CD81 monomer can bind E2 (Jamshad et al., 2008). Furthermore, when residues on the dimeric interface observed in the CD81 LEL crystal structure were mutated, soluble LEL dimerisation was disrupted but the same effect was not observed when tested on full-length CD81 oligomers in mammalian cells (Drummer et al., 2005). Therefore, Drummer & colleagues (2005) hypothesised that the LEL has a more robust structure in the full length tetraspanin compared to the soluble LEL counterpart. Regions other than in the LEL are likely to be vital for CD81 dimerisation and stability, such as in the transmembrane domains, since the CD9 dimeric interface was found in TM1 and TM2 (Kovalenko et al., 2005). Also, when viewing the crystal structure LEL dimer, the two loops are facing away from each other; incorporating the transmembrane domains into this orientation is impossible (Rajesh et al., 2012) (figure 1.3b). This leads to differences between what was found in the LEL soluble structure and what was later

modelled in the full length CD81 molecule (Seigneuret et al., 2001). Further research is still required to extend our knowledge on tetraspanins, and more specifically, CD81 full length structure and also the dimerisation interface.

Although an extensive collection of data is published to show that CD81 is a critical receptor for HCV entry, it was previously apparent to researchers that other molecules were required for this process. Human CD81 renders human hepatoma cells permissive to HCV infection, although CD81 from a range of other species that do not get infected by HCV also had a similar ability. Therefore, it was clear that other human specific factors / receptors must be involved for HCV to specifically infect humans and even more specifically, hepatocytes (Flint et al., 2006). Additionally, CD81 is ubiquitously expressed in humans and so other molecules, along with CD81, must define HCV tropism for the liver.

1.2.2 Claudin superfamily

Claudins were first described in the literature in 1998 (Tsukita and Furuse, 1998). There are approximately 27 family members found in mammals that are important components of tight junctions. Claudins have four transmembrane domains, separated by a first extracellular loop (ECL1) and a second smaller extracellular loop (ECL2) and small intracellular domains (see figure 1.5a).

a)

ECL1

ECL2

N-terminal tail   C-terminal tail

b)

Aston University

Illustration removed for copyright restrictions
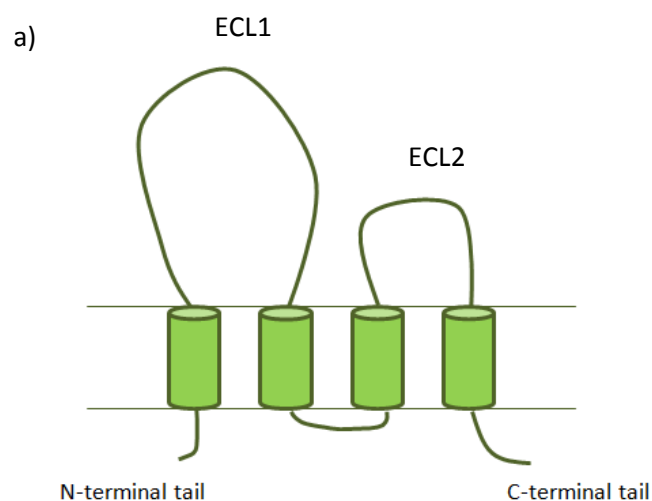
Figure 1.5: Claudin-1 structure and function. a) A schematic of the topology of Claudin-1 showing four transmembrane domains, intracellular N- and C- terminal domains and two extracellular loops (ECL1 and ECL2). b) Tight junction proteins Claudin and Occludin involved in the formation of junctional strands found at tight junctions. Tight junction strands from a seal to prevent uncontrolled leakage of water or solutes in epithelial or endothelial tissues. Claudins interact with other Claudins laterally and on neighbouring cells *via* their extracellular loops.

Claudin, from the Latin word *claudere*, means to close. Claudins constitute the paracellular seal in epithelial and endothelial tissues to allow for controlled movement of solutes and water *between* cells, as well as acting as a fence-like structure to separate the apical and basolateral sides of a membrane (Gunzel and Fromm, 2012) (see figure 1.5b). Claudin proteins can form oligomers laterally with themselves, *via* their extracellular loops, or with partner proteins in opposing cells to form a tight junction seal. It is possible to find Claudin homo- and hetero-oligomers in the epithelial membrane since multiple Claudins have been found in one tight junction strand. It is this combination that defines the sealing properties of a given tight junction (Van Itallie and Anderson, 2013).

Up until recently, a high resolution structure for Claudin family members did not exist. Suzuki and colleagues (2014) successfully solved the full length structure of mouse Claudin-15 at high resolution (2.4Å) (see figure 1.6). High level expression of Claudin-15 was achieved using Sf9 insect cells and the membrane protein was purified using maltose neopentyl glycol (MNG) detergent (see later for description) before crystallising in lipidic cubic phase (Suzuki et al., 2014). The structure shows that the TM domains form a typical left-handed four helix bundle and the extracellular loop

regions form β-sheet structure comprised of five β strands; four contributed by the first extracellular loop and one by the second extracellular loop. Interestingly, the crystal lattice demonstrates how Claudin-15 monomers form a linear homo-oligomer, with interactions between adjacent molecules dependent on extracellular domains involving the extracellular loop 1 of one monomer (residue Met68) and the transmembrane domain 3 and extracellular loop 2 (residues Phe146, Phe147 and Leu158) of the adjacent monomer since transmembrane domain 3 extends above the membrane surface (unlike the other three transmembrane domains). This may be physiologically relevant and represent the linear monomer interactions that form in tight junction strands (Suzuki et al., 2014).

Aston University

Illustration removed for copyright restrictions

Figure 1.6: Claudin-15 full length crystal structure. A ribbon representation of the crystal structure of monomeric mouse Claudin-15 solved at a resolution of 2.4 angstroms (PDB number: 4P79). Image taken from Suzuki *et al*., 2014.

1.2.2.1 Claudin-1

Claudin-1 is a 23 kDa membrane protein known to be highly expressed in the liver (Evans et al. 2007). Claudin-1 is functionally responsible for constituting tight junction strands and can associate *via* homotypic or heterotypic interactions in the cell membrane of a single cell and with proteins of an opposing cell through interactions between extracellular loops. In 2007, Evans and co-workers discovered Claudin-1 was

an entry receptor for HCV as a result of a repetitive expression cloning approach. This used the complementary DNA library of a cell line that is highly infected by HCV (Huh-7.5 cell line) to find specific genes that cause a cell line (CD81$^+$ SR-BI$^+$ 293 T cells) that is not typically infected by HCV to one that is susceptible to infection by HCVpp (Evans *et al.* 2007). Interestingly, Claudin-7 is one of the closest family members to Claudin-1, with only five residues differing between them in the LEL but is not a HCV entry receptor. Using a mutagenesis approach, it was found that residues I32 and E48 in the LEL of Claudin-1 are critical for HCV infection (Evans *et al.* 2007). Furthermore, other members of the Claudin family, Claudin-6 and Claudin-9 were also shown to support HCV entry (Zheng et al., 2007), (Meertens et al., 2008). Even so, Claudin-6 and -9 are not recognised to be critical factors defining HCV entry into hepatocytes as they are not known to be highly expressed in the liver (Timpe and McKeating, 2008).

Claudin-1 can locate in non-junctional areas of the basolateral surface of hepatocytes, co-localising with tetraspanin CD81 (Reynolds et al., 2008). Claudin-1 associates with CD81 and this receptor complex has been shown to be essential for HCV entry (Harris et al., 2010).

1.2.3 CD81-Claudin-1 as a HCV entry co-receptor complex

To date, there has not been any direct evidence for Claudin-1 to interact with HCV glycoproteins E1/E2, unlike HCV receptors CD81 and SR-BI. This may indicate the requirement of HCV to bind its cell surface receptors in a defined chain of events or potentially due to the low sensitivity of cell-based methods (Harris et al., 2008). Studies have provided evidence for Claudin-1 association with CD81 and the co-localisation of both proteins in the plasma membrane. Harris and co-workers (2008) demonstrated this using fluorescently tagged CD81 and Claudin-1 and used fluorescence resonance energy transfer (FRET) analysis to show that homotypic (CD81-CD81) (Claudin-1-Claudin-1) and heterotypic (CD81-Claudin-1) interactions occur in non-junctional areas of the plasma membrane (Harris et al. 2008). Claudin homotypic interactions are supported by previous work by Blasig and co-workers, reporting FRET between tagged

Claudin-5 molecules, indicating that homodimers exist in the plasma membrane of the same cell (Blasig et al., 2006).

The CD81-Claudin-1 complex has also been demonstrated using recombinant full length proteins, produced in *Pichia pastoris*, in which Claudin-1 oligomers associate with CD81 at a 1:2 molar ratio using dynamic light scattering (DLS) (Bonander et al., 2013). The complex was stabilised in the presence of cholesteryl hemisuccinate (CHS), suggesting that the complex is specific and does not require other cellular proteins or factors that might contribute to CD81-Claudin-1 complex formation.

Further evidence to demonstrate CD81-Claudin-1 is an essential complex in HCV entry was shown when residues I32 and E48 in Claudin-1 LEL disrupted association with CD81 and also disturbed HCV receptor function (Harris et al., 2010). In contrast, substituting the same residues on Claudin-7 (a non-HCV entry receptor) enabled the protein to interact with CD81 and allow HCV entry into 293T cells (Harris et al. 2010). The location of CD81-Claudin-1 complexes is important for HCV entry. CD81-Claudin-1 complexes located at basolateral areas of polarised hepatoma cells (Harris et al. 2010). This possibly explains the initial location of HCV infection, as the virus enters the liver *via* the sinusoidal blood; this would permit interactions between the virus and basal located forms of the co-receptor complex.

Other studies have focused on other Claudin-1 regions for its association with CD81, including the C-terminal tail and putative palmitoylation sites. Using FRET analysis, it was shown that Claudin-1 without its C-terminal tail associates with CD81 as does the wild-type full length Claudin-1 (Harris et al. 2008). This was also the case for intracellular palmitoylation sites, which did not prove a significant region for HCVpp entry (Evans *et al.*, 2007).

More recently, residues in CD81 LEL T149, E152 and T153 were shown to lie on the interaction interface with Claudin-1 ECL1 motif between residues 62-66 (Davis et al., 2012). In order to find the interaction interface in the extracellular loops, the study used a molecular modelling approach to predict the residues followed by site-directed mutagenesis and FRET to assess the interacting mutant proteins. It was found that the residues in the loop regions of both proteins are important for CD81-Claudin-1 complex formation and HCV entry. Importantly, mutant CD81 proteins (T149A, E152A and T153A) bound HCV E2 to a similar degree as WT CD81 (Davis et al., 2012) supporting

a previous study that showed the HCV E2 binding site is in a separate region (Drummer et al., 2002). This evidence may suggest the occurrence of a HCV-CD81-Claudin-1 complex prior to particle internalisation and infection.

It is clear in the literature that CD81-Claudin-1 loop regions are crucial for the protein interaction and consequently HCV entry. Little attention has focused on the transmembrane domains in this heterodimer but it is recognised that TM domains are important to drive tetraspanin dimerisation, more specifically TM1 and TM2 are reported to be on the dimerisation interface of CD9 (Kovalenko et al., 2005). Also, regarding HCV entry, Bertaux & Dragic (2006) found that mutating residues in CD81 TM1 (N18), TM2 (C80) and TM4 (E219), reduced HCVpp entry by approximately 40%. Since the mutants' ability to bind HCV E2 was not affected it seems apparent that another mechanism(s) is involved, such as a mutant knock down of a protein-protein interaction that is necessary for HCV infection to occur. Therefore, further investigation into transmembrane residues in their role in HCV entry and four TM protein interactions is required.

1.3 Recombinant protein production of membrane proteins

Integral membrane proteins are pharmaceutically important proteins as they are key drug targets. More than 50% of all drugs target them due to their many vital biological functions such as their involvement in cellular signal transduction processes. To a great extent, progress in the field of research concerning membrane proteins has been possible over the past few years due to developments in expression using a heterologous host and subsequent purification and crystallisation (Freigassner et al., 2009). High resolution structures are useful for drug development and to understand the function of a membrane protein but initial reliance is placed on the ability to produce high yields of functional and stable protein, since the majority are expressed in low levels naturally and their hydrophobic qualities means that they can be a challenge to produce and maintain in a functionally folded state (see Bill et al., 2011). There are a small number of membrane proteins that occur in large quantities in their native membrane such as rhodopsins and aquaporins, which were among the first membrane proteins to have their structures solved (Bill et al., 2011).

Hosts used to produce challenging proteins such as integral membrane proteins in the past include prokaryotes such as *Escherichia coli*, mammalian cells, insect cells and yeast such as *P. pastoris* (Fernandez and Vega, 2013). Even though prokaryotes have been hugely successful in the structural determination of many prokaryotic and eukaryotic membrane proteins, the production of a eukaryotic membrane protein may be favourable using a more evolutionary related recombinant host, such as the eukaryote yeast. The target membrane protein may benefit from eukaryotic features yeast can offer such as post-translational modifications and the eukaryotic secretory pathway for correct protein processing and functional expression (Ahmad et al., 2014).

1.3.1 *Pichia pastoris* **as a recombinant protein production host**

P. pastoris is a methanol-utilising yeast (methylotrophic) that can be readily grown to high volumes and cell densities (Daly and Hearn, 2005). The ability of yeast to use methanol as a sole carbon source was discovered over 40 years ago. In 1985 the alcohol oxidase I gene (*AOX1*), which is regulated by methanol, was isolated from *P. pastoris* (Ellis et al., 1985). The promoter of the *AOX1* gene is now known to be one of the most tightly regulated and effective promoters, which provides *P. pastoris* with an attractive feature to molecular biologists with an aim to control the expression of heterologous proteins (Cregg et al. 2000). Researchers are currently able to purchase a *Pichia* expression kit available from Invitrogen.

P. pastoris is widely used as a host to produce recombinant membrane proteins, in order to yield large quantities for structural and functional studies. Typically, it is the high cell densities and culture volumes that allow high yields of protein to be achieved, since even if a target protein is expressed at a low level on a cell basis, the volumetric yield is sufficient (Ahmad et al., 2014). Membrane protein production in *P. pastoris* has previously been successful in expression through to crystallisation. Shiroishi and co-workers had success using *P. pastoris* for the first time to determine the structure of a G protein-coupled receptor (GPCR). The human histamine H1 receptor was structurally determined at 3.1 Å resolution (Shiroishi et al., 2011). Solving the structures of GPCRs is a particular challenge in the membrane protein field because it is difficult to first of all express sufficient amounts of recombinant protein, then success must continue to

purification of stable proteins that are functional and properly folded and can pack together to form crystals (see Bill et al., 2011). Each stage of this process will likely need optimisation for a particular target protein, in order to reach the final goal of a solved structure.

P. pastoris was previously used at Aston University to generate milligram quantities of full length, correctly folded, functional CD81, for which all the palmitoylation sites were removed therefore referred to as CD81 p-null (Jamshad et al., 2008). The structural integrity of the full length CD81 p-null was determined through binding of specific monoclonal antibodies that were raised against native CD81. Further, the function of CD81 p-null was confirmed due to the specific binding with HCV E2 glycoprotein. Using the same recombinant host system, human Claudin-1 was also expressed to similar quantities and quality (Bonander et al., 2013). In both cases, the vector pPICZB was used to independently clone PCR products of the respective genes into *P. pastoris* wild-type X33 strains (Jamshad et al., 2008).

Recombinant Claudin-1 was previously characterised alone and in complex with CD81 p-null (Bonander et al., 2013). Yeast membrane bound and detergent solubilised Claudin-1 were shown to be antigenic using specific monoclonal antibodies, confirming the correct protein confirmation. Furthermore, Claudin-1 did not bind HCV sE2, which is consistent with current literature understanding. DLS analysis of the CD81-Claudin-1 complex found that Claudin-1 oligomers interacted with CD81 at a molar ratio of 1:2 (Bonander et al., 2013), also this complex was stabilised with the addition of CHS. This is consistent with findings that tetraspanins interact with membrane cholesterol (Charrin et al., 2003) and that cholesterol is important for the stability of tetraspanin protein interactions, such as in the CD81-Claudin-1 interaction for efficient HCV entry (Harris et al., 2010).

Tetraspanins have been produced recombinantly elsewhere in hosts other than yeast; these include CD81 in mammalian cells (Takayama et al., 2008) and CD151 and Uroplakin 1b in *E. coli* (Tarry et al., 2012). CD81 produced in a tetracycline-inducible stable mammalian cell line and purified using a single step immunoaffinity method (>95% purity) was shown to be folded and functional by binding HCV E2 (Takayama et al., 2008). Tarry and colleagues (2012) produced tetraspanins in *E.coli* and out of approximately 18 constructs screened for expression, only three were taken forward for

optimisation of expression and purification; CD81 was not taken forward. The study found cymal-5 as an optimal detergent for solubilisation. Although sufficient levels of recombinant full length tetraspanins were achieved, the protein was not homogeneous as seen by size exclusion chromatography, so unsuitable for structural studies, therefore purification optimisation using other detergents and conditions are still required in this area (Tarry et al., 2012). This would then enable further structural studies, such as protein X-ray crystallography or NMR spectroscopy, after obtaining sufficient concentrations of stable, functional and monodispersed CD81. To date, there has not been any high resolution full length structure for any member of the tetraspanin family.

1.3.2 Membrane protein solubilisation

Once a membrane protein of interest has been expressed to sufficient levels, the next step is to extract it from the host membrane. The most common approach is to solubilise the target protein using detergent (Prive, 2007). Detergents are used to mimic the lipid bilayer from where the membrane protein will be extracted due to their similarities in amphipathic properties. Detergent monomers are composed of a hydrophilic head group and a hydrophobic alkyl chain. Monomers in solution naturally form globular micelles at or above the critical micelle concentration (CMC) of a given detergent; so that the head group is solvent exposed and the hydrophobic alkyl chains are concealed from water molecules, forming a micelle core (Oliver et al., 2013). At detergent concentrations below the CMC, detergent exists as monomers, though at concentrations above the CMC, an equilibrium forms between monomers and micelles (Moraes et al., 2014). Figure 1.7 shows the complexity and various stages of membrane solubilisation by detergent, which is dependent on the detergent concentration. At low detergent concentration, monomers bind to the membrane although it is not disrupted. Membrane disruption occurs when the detergent concentration is increased (see figure 1.7) and a further concentration increase results in a heterogeneous mixture of micelles containing protein in complex with either detergent or detergent and lipid (Kalipatnapu and Chattopadhyay, 2005).
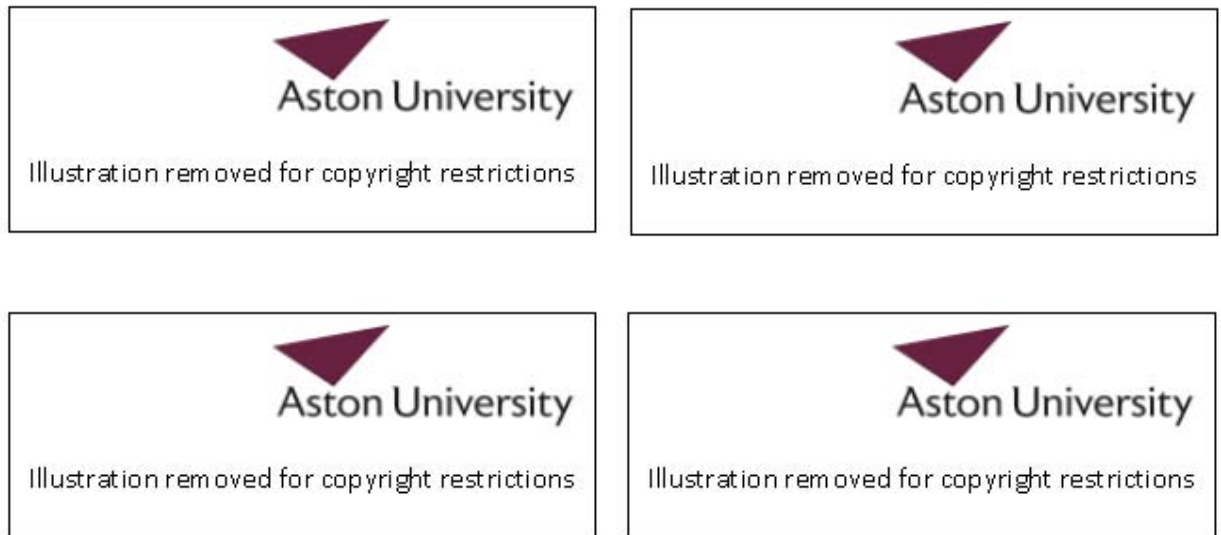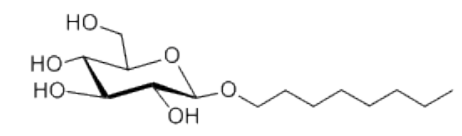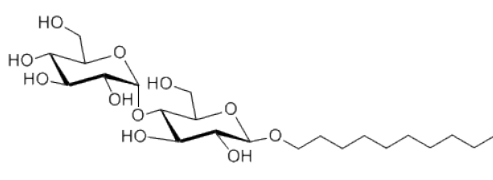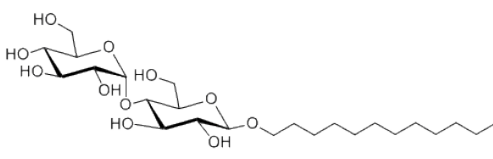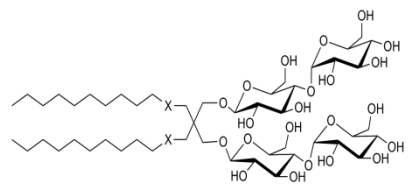
Figure 1.7: Detergent solubilisation of membrane proteins in a lipid bilayer. A) the lipid bilayer, B) Low detergent (grey tails) concentration bound to the membrane, C) Higher concentration of detergent causes disruption of the lipid bilayer and D) shows at higher detergent concentrations a heterogeneous solution exists of lipid-protein-detergent mixed micelles. Image taken from Kalipatnapu and Chattopadhyay (2005).

The main objective when extracting membrane proteins from the membrane is to achieve a soluble, functional and properly folded protein sample that can then be purified. Although, only too often, membrane proteins will denature and aggregate throughout this process since their environment is too distinct from the lipidic environment from which they have come from. Therefore, the addition of lipids to the protein detergent complex can sometimes benefit the stability of certain membrane proteins (Prive, 2007). Other innovative approaches to improve membrane protein detergent stability are now being used to obtain X-ray crystal structures, which include screening protein mutants, ligand addition and the use of novel detergents (Kang et al., 2013).

Table 1.1: Detergents used for membrane protein solubilisation and purification. Commonly used detergents to solubilise and maintain the folded structure and function of membrane proteins include β-OG, DM and DDM, also used in this thesis. MNG represent a more recently developed amphiphile used to solubilise membrane proteins. In contrast to conventional detergents, MNG compounds have two hydrophilic and two hydrophobic subunits.

Detergent	Formula Weight	CMC (%)	Structure
n-octyl-β-D-gluocopyranoside (β-OG)	292.4	0.53	
n-decyl-β-D-maltopyranoside (DM)	482.6	0.087	
n-dodecyl-β-D-maltopyranoside (DDM)	510.6	0.0087	
maltose neopentyl glycol (MNG)	-	0.0010	

The most common detergents used previously to obtain crystals have included the following three nonionic detergents: *n*-octyl-β-D-gluocopyranoside (β-OG), *n*-decyl-β-D-maltopyranoside (DM) and *n*-dodecyl-β-D-maltopyranoside (DDM) (Oliver et al., 2013). These detergents are regarded as 'mild' since they are known to break lipid-lipid and protein-lipid interactions rather than protein-protein interactions (Moraes et al., 2014). This is in contrast to 'harsh' detergents such as ionic detergents (for example, sodium dodecyl sulfate (SDS)) which are likely to denature the protein of interest. Table 1.1 shows how head group and tail length properties can differ between three

detergents. It has been found that the detergent head group is involved in the detergent-protein interaction, whereas the alkyl chain has an effect on the CMC value (Prive, 2007). Shorter chains such as in β-OG (8 carbons) have a much higher CMC than longer chains in DM and DDM (10 and 12 carbons, respectively) and so large amounts of β-OG are used to solubilise membrane proteins as compared to that used for DM and DDM. The detergent used to solubilise is typically optimised for each individual target protein by assessing the stability and monodispersity of the protein-detergent complex. Also, the detergent used for extraction may be different to the detergent used for later stages of purification and crystallisation, after performing a detergent buffer exchange (Moraes et al., 2014).

Since detergents may replace annular lipids surrounding the protein, which can result in the loss of protein structure and function, alternative solubilisation approaches have now been developed in order to improve membrane protein stability. One example includes the use of a novel set of amphiphiles composed of a quaternary carbon derived from neopentyl glycol and hydrophilic head groups derived from maltose, therefore the compounds are called maltose neopentyl glycol (MNG) (Chae et al., 2010). In contrast to current detergents (table 1.1), MNG compounds have two hydrophilic and two hydrophobic subunits. It was found that MNG could extract and maintain proper folding and function of various membrane proteins, with the ability to assist crystal formation for structural characterisation. Such findings established MNG amphiphiles as being more beneficial for protein stability (for those proteins that were tested) than standard detergents used such as DM (Chae et al., 2010). Recently, MNG was used for the crystallisation of membrane protein Claudin-15 (Suzuki et al., 2014).

Given that membrane lipids surrounding target membrane proteins are crucial for stability, other approaches have made successful efforts to purify the protein with the annular lipid environment intact (Jamshad et al., 2011). For example, the poly(styrene-co-maleic acid) (SMA) is a amphipathic polymer that is composed of alternating hydrophilic (maleic acid) and hydrophobic (styrene) units so that at neutral and basic pH the SMA forms a disc-like structure surrounding the protein plus membrane. Adding the SMA directly to lipid bilayers containing the protein of interest is sufficient to form the nano-discs and therefore detergent is not required. This enables purification of the target protein in particles with diameters between 9-11nm (Knowles et al., 2009),

which maintains the protein structure and function in order to study further using biophysical techniques such as analytical ultracentrifugation.

1.3.3 Membrane protein stability studies

It is well established that achieving stable recombinant eukaryotic membrane protein preparations can be a difficult challenge, but the optimisation of stability as well as solubility and homogeneity have been shown to be essential for a successful crystallisation trial (Alexandrov et al., 2008). Therefore, methods used to assess and optimise the stability of purified samples are imperative prior to solving a high resolution protein structure. One relatively recent approach exploits a thiol-specific fluorochrome N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) used to assess the thermal stability of membrane proteins in a high throughput format, therefore requiring small amounts of purified material (Alexandrov et al., 2008) (see figure 1.8).

Aston University

Illustration removed for copyright restrictions

Figure 1.8: CPM structure. Thiol reactive CPM is non-fluorescent until reacted with thiols and becomes highly fluorescent. This can be used to measure temperature-induced membrane protein unfolding since most cysteine residues are typically buried in transmembrane domains (Alexandrov et al., 2008).

Thermal stability is assessed in this assay using a fluorescence read-out when a membrane protein (in a particular buffer) is incubated at a certain temperature for unfolding to occur. Cysteine residues that are buried within the hydrophobic interior of the target protein, typically within transmembrane domains, are exposed after temperature-induced protein unfolding. This causes the thiol-specific CPM to covalently bind to exposed cysteines and the fluorescence readout can be recorded over time using an appropriate spectrofluoremeter (Alexandrov et al., 2008). Various studies

have used this method to improve the stability of the target membrane protein(s). Sonoda *et al*., (2011) used the CPM assay to optimise detergent stability of 24 eukaryotic and prokaryotic membrane proteins and found a correlation between protein stability and the likeliness of obtaining well-ordered crystals for high resolution structure determination (Sonoda et al., 2011).

An alternative approach to improve protein stability in a different study used alanine scanning mutagenesis. For example, the GPCR β1-adrenergic receptor with 6 point mutations and bound antagonist shown an increase in thermal stability as compared to the WT receptor. Thermal stability was assessed by heating detergent-solubilised receptors to 32°C and then performing radioligand binding experiments to measure the remaining number of functional receptors (Serrano-Vega et al., 2008).

1.4 Determination of protein-protein interactions

As described in earlier sections, membrane proteins are typically organised in molecular complexes at the plasma membrane in order to perform their role in biological processes. Therefore, it is fundamental for biologists to unpick these complexes and determine specific protein-protein interactions that occur and for what reason. Typically, if a biological function is dependent on a membrane protein binding to its partner protein, it is also likely to be linked to disease and seen as a potential drug target. This is true for the CD81-Claudin-1 complex. Although it is not understood why tetraspanins interact with tight junction proteins normally outside of the tight junction, it is clear that this complex is required for HCV entry into hepatocytes (Harris et al., 2010).

Biochemical means to study protein-protein interactions exist such as coimmunoprecipitation and cross-linking, which require cell lysis using harsh detergents that may interfere with weak protein-protein interactions (Thaminy et al., 2003). Consequently, alternative techniques employed to map protein-protein interactions at the plasma membrane (*in vivo*) are extremely useful to determine novel interaction partners and also to define the binding interface between proteins. The hydrophobic nature of membrane proteins adds on an extra level of complexity, since

the lipid bilayer is the optimal environment for membrane proteins to maintain structure and function. Therefore, techniques that can measure protein-protein interactions *in situ* may be advantageous over those where membrane proteins need to be extracted prior to measurement with partner proteins.

1.4.1 The yeast split-ubiquitin Dualmembrane system

A technique used in this thesis to investigate protein-protein interactions is the yeast split-ubiquitin Dualmembrane system (Dualsystems Biotech Ltd). The method allows detection of interactions between integral membrane proteins or soluble proteins *in situ* in the yeast, *Saccharomyces cerevisiae*. It has been developed over time from the original yeast-two-hybrid system (Fields and Song, 1989), which was primarily suited to soluble proteins and was problematic for membrane proteins since the proteins of interest would be required to enter the nucleus for transcriptional activation of reporter genes.

Modifications of the original yeast-two-hybrid have meant that a wider range of protein targets can now be studied. One example uses the split-ubiquitin idea first described by Johnsson and Varshavsky (1994). The genetic assay uses ubiquitin, a protein that usually tags other proteins for degradation with chains of ubiquitin; this process flags a signal on the protein so that it is transported to the proteasome for degradation. Ubiquitin specific proteases (UBPs) recycle ubiquitin monomers by cleaving a site after the last residues of ubiquitin. In *S. cerevisiae* when ubiquitin is split into two halves and expressed independently, the two halves have strong affinity for each other and so re-assemble within the cell. UBPs can then recognise and cleave the reconstituted ubiquitin (Johnsson and Varshavsky, 1994). In order to make these phenomena useful when being applied to a protein interaction assay, the N-terminal half of ubiquitin was mutated at residue 3 from an isoleucine to a glycine, which results in the two halves of ubiquitin having negligible affinity for each other and so UBPs do not recognise either half within the cell. Therefore, reconstitution of ubiquitin and cleavage by UBPs can only occur as a result of two interacting proteins that are fused to either half of ubiquitin (Johnsson and Varshavsky, 1994) (see figure 1.9).

The yeast split-ubiquitin Dualmembrane system allows heterologously expressed membrane proteins to reside in the membrane, *in vivo*, where they can undergo protein-protein interactions, which can be measured using a reporter gene system (Stagljar et al., 1998). After selecting two proteins of interest, one is assigned as bait and is fused to the C-terminal half of ubiquitin (Cub) plus a transcription factor LexA-VP16 and the other is prey, which is fused to the N-terminal half of mutated ubiquitin (NubG) (see figure1.9). When co-expressed within the same cell, if the two proteins of interest have natural affinity for each other, they will interact, which forces NubG and Cub into close proximity. This results in re-assembly of the two halves of ubiquitin and so UBPs now recognise functional ubiquitin and cleave the polypeptide chain between Cub and LexA-VP16.

LexA-VP16 is composed of two domains; LexA is a DNA binding domain and VP16 is a transcriptional activating domain. When it is released upon cleavage by UBPs it translocates to the nucleus and activates transcription of reporter genes (Stagljar et al., 1998). The reporter genes used in this assay are two autotrophic growth markers; *HIS3* and *ADE2*. When a protein-protein interaction occurs between the proteins of interest, the yeast co-expressing them can grow on minimal media that lacks addition of histidine or adenine. The other reporter gene that can be used for measurements of protein-protein interactions is *lacZ*, which when activated, causes expression of β-galactosidase that can be measured using colorimetric and fluorometric assays.
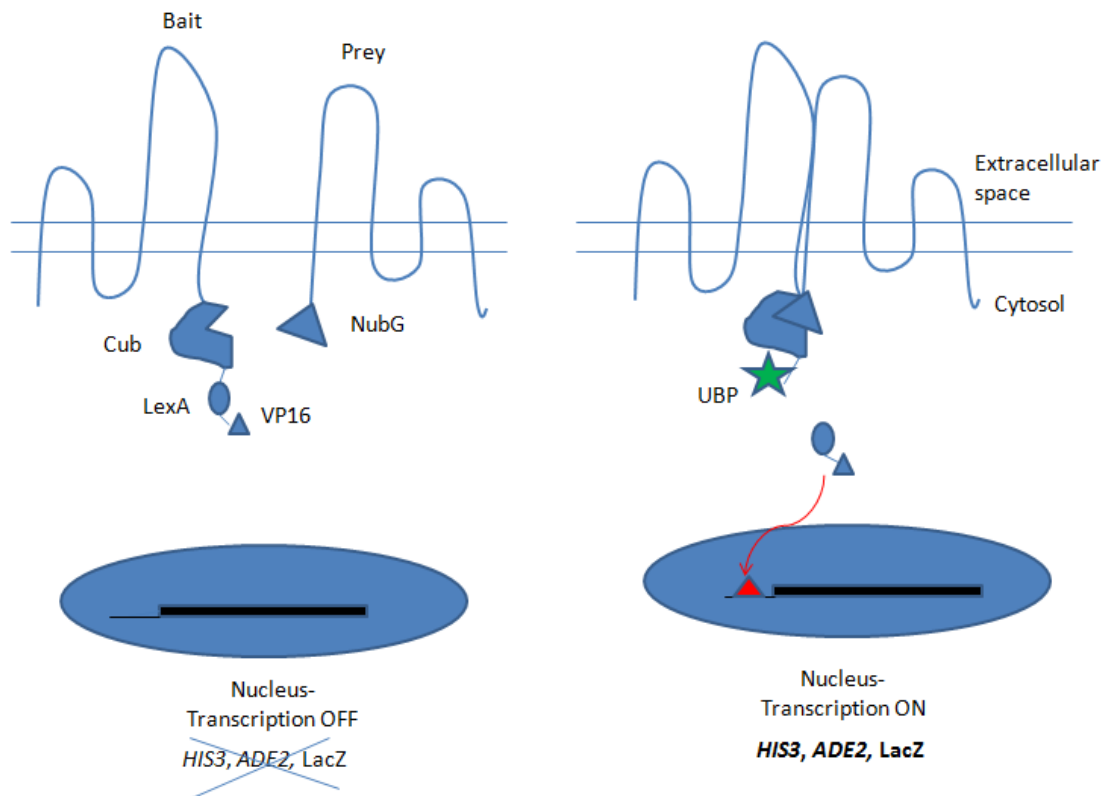
Figure 1.9: Yeast two hybrid split-ubiquitin Dualmembrane system. Bait and prey proteins of interest are fused to half of ubiquitin (Cub-LexA-VP16 and NubG, respectively). Upon a protein-protein interaction of the bait and prey proteins, Cub-LexA-VP16 and NubG are forced into close proximity and reconstitute 'whole' ubiquitin, which is recognised by UBPs that cleave LexA-VP16. The artificial transcription factor can then translocate to the nucleus where it initiates the expression of reporter genes *HIS3*, *ADE2* and *lacZ*. Therefore, growth on histidine and adenine deficient plates or activity in a β-galactosidase assay is an indirect measurement of a protein-protein interaction (Johnsson and Varshavsky, 1994).

1.5 Aims and Objectives

This thesis describes an investigation designed to increase current understanding of the structural and functional features of CD81 and Claudin-1. This was done using two yeast strains, *S. cerevisiae* and *P. pastoris* in order to determine the interaction interface of CD81-Claudin-1 and the structural stability of recombinant proteins, respectively, with the following objectives:

- To optimise a yeast split-ubiquitin Dualmembrane system assay in order to detect CD81 homodimers and CD81-Claudin-1 heterodimers in the yeast membrane;

- To determine the full length interaction interface of the CD81-Claudin-1 complex, specifically focusing on the role of transmembrane domains and to learn if there are any similarities between the binding interface of homo- and hetero-interactions;

- To select residues to test in a protein-protein interaction yeast assay guided by molecular modelling;

- To structurally characterise full length recombinant CD81 by means of improvements to solubilisation, purification and stabilisation procedures.

2. Materials and methods

2.1 Vector generation and Dualmembrane split ubiquitin yeast assay

The DUALmembrane pairwise interaction kit was purchased from Dualsystems Biotech Ltd and contained the following items that were used in this thesis:

Table 2.1: DUALmembrane pairwise interaction kit contents. Taken from the user manual (Dualsystems Biotech Ltd).

Aston University

Illustration removed for copyright restrictions

2.1.1 Yeast strain NMY51

S. cerevisiae strain NMY51 (see table 2.1) was grown on YPAD agar plates by incubating at 30°C for three days. Single colonies were picked into 10ml overnight YPAD cultures (shaking at 30°C, 220rpm). The following day, glycerol stocks were prepared using 1:1 YPAD NMY51 saturated culture and 50% sterile glycerol and vortexed. Using a cryovial, glycerol stocks were stored at -80°C until needed.

NMY51 on agar plates were kept at 4°C wrapped in parafilm and re-streaked every 2 weeks from the glycerol stock. Yeast plates older than 2 weeks were not used for transformation.

2.1.2 Yeast media and plates

YPAD

Liquid media was made using the Dualsystems Biotech Ltd media starter package. Each pouch required 1L water and YPAD was stirred until completely dissolved. Liquid YPAD was autoclaved. If making YPAD agar plates, 500ml dissolved YPAD was added to 10g agar and autoclaved. Subsequently, YPAD agar was poured into sterile petri dishes and allowed to solidify overnight. Agar plates were stored at 4°C, wrapped in parafilm, until use.

Synthetic Defined (SD) media

Liquid or agar plates were made using Dualsystems Biotech Media Package 1 (for DUALmembrane), following manual guidelines. Liquid media and agar plates were made excluding; SD-Leu, SD-Trp, SD-Trp-Leu, SD-Trp-Leu-His and SD-Trp-Leu-His-Ade. Amino acids (purchased from Dualsystems Biotech) were added at the following final concentration:

- Adenine (A) - 10 mg/L
- Histidine (H) - 20mg/L
- Tryptophan (W) - 20mg/L
- Leucine (L) - 100mg/L

Yeast SD plates were made using 2% agar, left to solidify overnight and stored at 4°C, wrapped in parafilm, until use.

2.1.3 *E.coli* competent cells

XL-1 Blue supercompetent cells and XL-10 Gold Ultracompetent cells were purchased from Stratagene (Agilent technologies) and stored at -80°C until use.

2.1.4 Bacterial media

Luria-Bertani (LB) broth- For liquid broth 20g/L LB was dissolved completely into water and autoclaved. For LB-agar, 20g/L LB and 15g/L agar was autoclaved and allowed to cool before pouring into sterile petri dishes. If an antibiotic was to be added, such as Ampicillin (LB-AMP) or Kanamycin (LB-KAN) the LB (+/- agar) was allowed to cool to around 40°C before antibiotic addition and mixing with a magnetic stirrer.

2.1.4.1 Antibiotics

Kanamycin

Stored at -20°C at 50mg/ml and used at a working concentration of 40µg/ml).

Ampicillin

Stored at -20°C at 100mg/ml and used at a working concentration of 100µg/ml).

2.1.5 *E. coli* transformation

In order to obtain further vector DNA to use for yeast transformation, *E.coli* was transformed and minipreps performed for isolation of the given vector.

For *E.coli* transformation, 100µl XL-1 Blue supercompetent cells were treated with 1.7µl β-mercaptoethanol (1.42 M) prior to transformation, mixing on ice every 2 min for 10 min in total. 20µl of cells were taken and mixed with 2µl vector DNA (~500ng/µl) and incubated on ice for 30 min. A heat block was set to 42°C and tubes containing bacteria and DNA were exposed to a heat shock for 45 seconds and then put immediately back on to ice for 2 min. 1ml LB was added to each tube of bacteria and DNA and incubated at 37°C, 200rpm for 1h. After, bacterial samples (30µl and 60µl of mix) were added to LB-agar plates containing an appropriate antibiotic (AMP or KAN) in order to select for cells that harbour the selected DNA vector. Plates were then incubated at 37°C for no more than ~18h and subsequently stored at 4°C until single bacterial colonies were required to be picked. Using 3ml LB-AMP or LB-KAN (depending upon vector being used), a single bacterial colony was added into media and incubated overnight at 37°C, 200rpm. The following morning, cultures were taken, centrifuged and vectors were isolated using a GenElute Plasmid miniprep kit (using the spin method) according to the manufacturers manual. Vector concentration would be determined using a NanoDrop and the quality determined using agarose gel electrophoresis. Successful DNA isolations were then sent for DNA sequencing (Eurofins).

2.1.6 Vector generation

Upon arrival, 50µl MilliQ water was added to lyophilized plasmids and incubated at 50°C for 5 min, vortexed then stored at -20°C until further use.

The Oligonucleotide primers to sub-clone *CLDN1* and *CD81* from pTrip vectors (kindly provided by Professor Jane McKeating, Birmingham University) were generated using Dualsystems Biotech guidelines. The primers synthesised (Invitrogen) are shown in table. The polymerase chain reaction (PCR) used to amplify the insert is outlined in table.

Table 2.2: CD81 and Claudin-1 oligonucleotide primers. The Oligonucleotide primers were designed using the manufacturer's guidelines containing *Sfi* I sites (DUALsystems Biotech Ltd). The oligonucleotide primers were synthesised by Invitrogen and on arrival diluted with MilliQ water so that the stock concentration was 100pmol/µl and stored at -20°C.

Vector	Primer sequence
pBT3-N- *CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCATGGGAGTGGAGGGCTGCACC '3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCTCAGTACACGGAGCTGTTCCG'3
pPR3-N-*CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCATGGGAGTGGAGGGCTGCACC '3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCTCAGTACACGGAGCTGTTCCG'3
pBT3-C- *CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCAAAAATGGGAGTGGAGGGCTGC '3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCGTACACGGAGCTGTTCCGGATGCC'3
pPR3-C- *CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCAAAAATGGGAGTGGAGGGCTGC '3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCGTACACGGAGCTGTTCCGGATGCC'3
pBT3-STE- *CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCGGAGTGGAGGGCTGCACCAAG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCGTACACGGAGCTGTTCCGGATGCC'3
pPR3- STE-*CD81*	Forward primer: 5'ATTAACAAGGCCATTACGGCCTTGGAGTGGAGGGCTGCACCAAG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCGTACACGGAGCTGTTCCGGATGCC'3
pBT3-N- *CLDN1*	Forward primer: 5' ATTAACAAGGCCATTACGGCCATGGCCAACGCGGGGCTGCAG'3 Reverse primer: 5' AACTGATTGCCGAGGCGGCCCTACACGTAGTCTTTCCCGCTGGA'3
pPR3-N- *CLDN1*	Forward primer: 5' ATTAACAAGGCCATTACGGCCATGGCCAACGCGGGGCTGCAG'3 Reverse primer: 5' AACTGATTGCCGAGGCGGCCCTACACGTAGTCTTTCCCGCTGGA'3

49

pBT3-C- *CLDN1*	Forward primer: 5' ATTAACAAGGCCATTACGGCCAAAAATGGCCAACGCGGGGCTGCAG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCCACGTAGTCTTTCCCGCTGGA'3
pPR3-C- *CLDN1*	Forward primer: 5' ATTAACAAGGCCATTACGGCCAAAAATGGCCAACGCGGGGCTGCAG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCCACGTAGTCTTTCCCGCTGGA'3
pBT3-STE- *CLDN1*	Forward primer: 5'ATTAACAAGGCCATTACGGCCGCCAACGCGGGGCTGCAGCTGTTG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCCACGTAGTCTTTCCCGCTGGA'3
pPR3-STE- *CLDN1*	Forward primer: 5'ATTAACAAGGCCATTACGGCCTTGCCAACGCGGGGCTGCAGCTGTTG'3 Reverse primer: 5'AACTGATTGGCCGAGGCGGCCCCCACGTAGTCTTTCCCGCTGGA'3

Using primers in table 2.2, the PCR reaction detailed in table 2.3 was performed using a GeneAmp PCR system 9700 (Applied Biosystems).

Table 2.3: CD81 and Claudin-1 PCR protocol. All PCR reactions were conducted in the GeneAmp PCR system 9700 (Applied Biosystems). The reactions were conducted in 0.25ml domed capped PCR tubes (Fisher Scientific Ltd).

PCR reagents	PCR program
5μl *pfu* buffer	1) 94˚C for 3 min
1μl Forward primer (10pmol/μl)	2) 94˚C for 1 min
1μl Reverse primer (10pmol/μl)	3) 55˚C for 1 min
1 μl dNTP (100 mM)	4) 68˚C for 2 min
1μl Template plasmid (50ng)	5) Repeat steps 2,3 and 4
1 μl *pfu* Turbo polymerase	30 times
40 μl ddH2O	6) 72˚C for 7 min
	7) 4˚C until collection
Total volume = 50μl	

The PCR products were then purified using a QIAquick PCR purification kit (Qiagen), following the manufacturers' guidelines. Subsequently, the purified PCR product and vectors were digested with a *Sfi I* enzyme (see table 2.4 for reaction) at 50°C for 2 h.

Table 2.4: Digestion reaction. The concentration of DNA was assessed using a NanoDrop 1000. All digestion reactions were conducted in a heat block at 50°C for 2h.

Digestion reaction mixture
34 μl MilliQ H_2O
5 μl NEB buffer 4
0.5 μl BSA (x100)
2 μg DNA (insert or vector)
20 units *Sfi* I enzyme (1 μl)
Total volume = 50μl

Subsequently, Antarctic phosphatase was added to the digested vectors but not to the inserts. Then a further PCR purification step was carried out (QIAquick PCR purification kit, Qiagen) on both the insert and digest in order to remove the *Sfi* I enzyme as it is not amenable to heat inactivation. Next, the insert was ligated into the appropriate vector using T4 ligase by adding a 1:3 molar ratio of vector: insert (using 100ng of vector). The reaction was conducted for 2 h at 22°C, followed by 10 min incubation at 65°C to heat inactivate the T4 ligase enzyme.

The ligated vector was then transformed into XL10-Gold cells (Stratagene), using a heat shock method (42°C for 45 seconds) then cells were incubated in LB broth at 37°C for approximately 1h for recovery. LB agar plates were made with the appropriate antibiotic (Kanomycin or Ampicillin) depending on the vector used. After incubation at 37°C, XL10-Gold cells were plated onto LB-agar plates (with the correct antibiotic selection) and incubated at 37°C overnight. Single colonies were subsequently selected and grown overnight in LB-antibiotic and then the vector of interest was extracted using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich), following the spin method. Subsequently, the vector was digested using the reaction in table 2.4 and products ran on an agarose gel (1%) to identify if the insert and the vector appeared as separate bands due to *Sfi I* digestion. Vector DNA sequences were confirmed by the Functional Genomics Lab (Birmingham University).

2.1.7 Site-directed mutagenesis of split-ubiquitin fusion proteins

Table 2.5: Site-directed mutagenesis primers. Primers were stored at 100pmol/µl in MilliQ water at -20°C.

Split ubiquitin vector	Mutagenesis primers	Mutation
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GCTGTGGTGAAGGCCTTCCACGAG 3' Reverse primer: 5' CTCGTGGAAGGCCTTCACCACAGC 3'	T149A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GAAGACCTTCCACGCCACGCTTGACTGCTG 3' Reverse primer: 5' CAGCAGTCAAGCGTGGCGTGGAAGGTCTTC 3'	E152A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GAAGACCTTCCACGAGGCCCTTGACTGCTGTGGC 3' Reverse primer: 5' GCCACAGCAGTCAAGGGCCTCGTGGAAGGTCTTC 3'	T153A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GTGCATCAAGTACCTGGCCTTCGTCTTCAATTTCG 3' Reverse primer: 5' CGAAATTGAAGACGAAGGCCAGGTACTTGATGCAC 3'	L14A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GTACCTGCTCTTCGTCGCCAATTTCGTCTTCTG 3' Reverse primer: 5' CAGAAGACGAAATTGGCGACGAAGAGCAGGTAC 3'	F17A

Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GTACCTGGCCTTCGTCGCCAATTTCGTCTTCTG 3' Reverse primer: 5' CAGAAGACGAAATTGGCGACGAAGGCCAGGTAC 3'	L14A-F17A N.B. used L14A mutant vector to make the double mutant
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' CAGGAATCCCAGGCCCTGCTGGGGAC 3' Reverse primer: 5' GTCCCCAGCAGGGCCTGGGATTCCTG 3'	C89A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' CTTCAATTTCGTCGCCTGGCTGGCTGGAG 3' Reverse primer: 5' CTCCAGCCAGCCAGGCGACGAAATTGAAG 3'	F21A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' CTGGCTGGCTGCCGGCGTGATCCTG 3' Reverse primer: 5' CAGGATCACGCCGGCAGCCAGCCAG 3'	G25A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' CTGGAGGCGTGGCCCTGGGTGTGG 3' Reverse primer: 5' CCACACCCAGGGCCACGCCTCCAG 3'	I28A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GGAGGCGTGATCGCCGGTGTGGCCCTG 3' Reverse primer: 5' CAGGGCCACACCGGCGATCACGCCTCC 3'	L29A
Bait-STE-*CD81* or Prey-N-*CD81*	Forward primer: 5' GGCCCTGTGGGCCCGCCATGACC 3' Reverse primer: 5' GGTCATGGCGGGCCCACAGGGCC 3'	L35A

Mutagenesis was performed using a QuikChange II site-directed mutagenesis kit (Agilent Technologies) according to the manufacturers' guidelines. Mutagenesis forward and reverse primers (see table 2.5) were used to mutate Bait-STE-*CD81* and Prey-N-*CD81* vectors using the reaction in table 2.6 (N.B. all work was performed on ice):

Table 2.6: Site-directed mutagenesis reaction.

Mutagenesis reaction	PCR program
40µl water (MilliQ) 5µl 10× Buffer 1µl Template plasmid (50ng) 1µl Forward primer (10pmol/µl) 1µl Reverse primer (10pmol/µl) 1µl DNTp 1µl PFU ULTRA Total volume= 50µl	1) 95°C for 2min 2) 95°C for 50 sec 3) 55°C for 50 sec 4) 68°C for 15 min 5) repeat steps 2,3 and 4, 18 times 6) 68°C for 7 min 7) 4°C until collection

After the mutagenesis reaction was complete (as shown in table 2.6) 1µl Dpn1 was added to each reaction and incubated at 37°C for 4h. This is to remove template DNA and leave DNA that has undergone mutagenesis. Subsequently, mutated vectors were transformed into XL-10 Gold Ultracompetent cells, which were selected on either LB-AMP or LB-KAN agar plates. Single colonies were then selected and grown in LB-antibiotic media overnight and minipreps performed the following day. Isolated vectors were then sent off to be sequenced (Eurofins) and results would verify the presence of a mutation and if the DNA sequence was in-frame.

2.1.8 Split-ubiquitin NMY51 yeast transformation

2.1.8.1 Yeast transformation buffers

Deoxyribonucleic acid sodium salt from salmon testes (salmon sperm DNA)

Salmon sperm DNA was dissolved in 10mM Tris, 1mM EDTA, pH 8.0, before autoclaving and aliquoted at 10mg/ml and stored at -20°C.

40% PEG solution

PEG (40% w/v) was dissolved using 10mM Tris pH 7.4, 1% Lithium acetate, 1mM EDTA pH 8.

Lithium acetate Tris EDTA (LiAc/TE) buffer

10mM Tris pH 7.4, 1% Lithium acetate, 1mM EDTA pH 8.

2.1.8.2 Yeast transformation method

Yeast 2% SD agar plates were prepared and allowed to solidify overnight. 4 single NMY51 colonies were picked into 10ml YPAD media and grown overnight at 30°C, shaking at 220rpm. The following day, yeast NMY51 pre-cultures (×4) were diluted to an OD_{546} 0.2 and grown until the OD_{546} reached between OD_{546} 0.6-0.8. Cells were centrifuged for 5 min at 2,000rpm to obtain pellets and subsequently washed in 5ml LiAc/TE buffer. Cells were pelleted again then resuspended in 1ml LiAc/Te buffer into a homogenised solution. In the meantime, salmon sperm DNA was heated to 95°C for 5 min using a heat block and subsequently put straight on to ice before use.

For each yeast transformation, 2μg vector DNA was added, if performing a single (one vector) transformation the total amount of DNA used was 2μg, if performing a double transformation (i.e. bait and prey vectors) 4μg DNA was used per reaction. Each transformation used 120μg salmon sperm DNA, 100μl homogenised yeast culture and 1ml 40% PEG solution. Tubes were inverted and mixed before incubating at 30°C for 30 min and transferred to a heat block for 20 min at 42°C for heat shock. Subsequently, tubes were centrifuged at 3,000rpm for 5min and PEG was carefully removed from the yeast pellet. For single (one vector) transformations, 100μl MilliQ water was used to resuspend yeast pellets and added to an appropriate SD agar plate and cells were spread using a glass spreader. For double (two vector) transformations, 150μl MilliQ water was added to resuspend yeast pellets and 50μl yeast solution was added to three SD agar plates (SD-W-L, SD-W-L-H and SD-W-L-H-A). All plates were incubated at 30°C for approximately 4-5 days before counting and recording colony counts for subsequent data analysis.

2.1.8.3 Split ubiquitin protein interactions measured using yeast growth curves

After counting NMY51 colonies selected on SD-Trp-Leu plates for specific protein-protein interaction conditions, colonies were then pooled into 5ml SD-W-L-H-A media. Samples were then vortexed for 5min and then OD_{546} measured using a spectrophometer. Initially cultures were diluted to an OD_{546} 1 in 1ml SD-W-L-H-A and then diluted further to an OD_{546} 0.02 in both SD-W-L and SD-W-L-H-A media. Cells were then added to sterile 96-well plates (Greiner) in four replicate wells each containing 200μl culture. The plate was subsequently covered using a sterile Breathe-Easy seal (Sigma) and put into a Multiskan GO (Thermo Scientific). The OD_{546} of all wells was measured every 1h and cells were incubated at 30°C whilst shaking, between readings. Measurements were taken for at least 19h and raw data was exported to excel. Using GraphPad prism (San Diego), growth curves were generated by plotting the OD_{546} against time (h) using a point-to-point connecting line. Representative growth curves are shown. In order to compare the growth of WT CD81 with mutant CD81 (in a particular protein interaction) the internal positive control (Claudin-1 + NubI) was included in all experiments. Using GraphPad prism, when the OD_{546} of the NubI control reached an OD_{546} 0.2, at that given time point, the OD_{546} of WT and mutant protein conditions were taken and used to plot the mean OD_{546} value of at least 3 repeat experiments (see figure 2.1). This approach took into account differences in lag time of yeast growth between independent experiments.
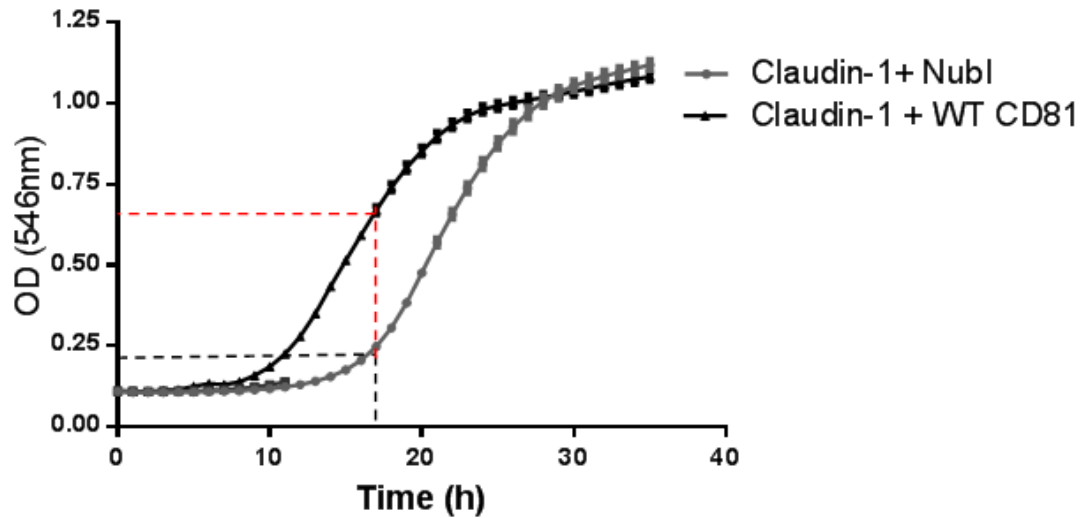
Figure 2.1: Growth curves in SD-W-L-H-A. *S. cerevisiae* strain NMY51 was co-transformed with two vectors (either encoding bait Claudin-1 and prey NubI or bait Claudin-1 with prey WT CD81) and allowed to grow for 5 days selected on SD-W-L plates. Colonies were pooled from each plate and diluted to an OD_{546} 0.02 before measuring the OD_{546} over time using a 96 well format in SD-W-L-H-A for 35 h at 30°C. Representative growth curves are shown as the mean and ±SEM of four growth curves measured in parallel. Both bait + NubI and bait + WT CD81 were included in all subsequent experiments for data analysis as the NubI reaction was used as an internal control to compare WT and mutant CD81 growth at a certain point of the curve (when NubI reaches OD_{546} 0.2, the WT and mutant reaction OD_{546} is taken and plotted as the mean of at least three experiments in a separate graph). The black dashed line represents the time when the NubI control has reached OD_{546} 0.2, the red dashed line shows how this time point is used to determine the OD_{546} of the WT reaction in GraphPad prism (San Diego).

2.1.8.4 Split ubiquitin protein interactions measured using the *lacZ* reporter gene

Two different approaches were used to quantify *lacZ* reporter gene expression using β-galactosidase activity.

The first method used the HTX β-galactosidase assay kit (Dualsystems Biotech Ltd). NMY51 cells transformed with bait and prey vectors were selected for on SD-W-L plates. Multiple colonies were selected and grown overnight in SD-W-L media, and diluted the next morning to OD_{546} 0.2 and grown until they reached log phase around 0.5-0.8. An OD_{546} 1 in 1ml was taken for each reaction and centrifuged (2,000×g, 5 min). The supernatant was removed and 100µl lysis solution was added and incubated

for 30 min. Samples were centrifuged to remove unlysed cells (2,000×g, 1 min) and 100µl supernatant was added to a 96 well plate and the colour development was monitored using a Multiskan GO (Thermo Scientific) for 10h. Measurements were taken every 30 min using absorbance at 615nm. The time (min) against absorbance (615nm) was plotted using GraphPad prism for each reaction.

The second method used Fluorescein-di-β-D-galactopyranoside (FDG, Invitrogen) as substrate for β-galactosidase. Stocks of 20 mM FDG in 100% DMSO were stored at -20°C. Similar to the first method cells transformed with two vectors were grown and diluted so as to reach log phase before analysis. 2µl FDG was added to 4 mLmedia (SD-W-L and SD-W-L-H-A) and cells were diluted to an OD_{546} 0.02 in FDG-medium. 100µl of culture was added in a 96-well plate (Black, clear bottom, Greiner) and a fluorescence plate reader (SpectraMAX Gemini EM, Molecular Devices) was used at 30°C over a 24 h period with measurement taken every 5 min (Excitation 485nm, Emission 535nm and cut-off 530nm). Cells were mixed for 20 sec between each reading. Using GraphPad prism the relative fluorescence units (RFUs) were plotted against time (min).

2.1.8.5 Split-ubiquitin fusion protein expression

NMY51 cells were transformed with a single vector, bait or prey vector, and selected on plates SD-Leu and SD-Trp, respectively. Single colonies were picked into 10ml SD-Leu or SD-Trp media and grown overnight at 30°C, 220rpm. The following day, the 10ml pre-cultures were added to 40ml media and grown in a baffled shake flask overnight. Cells were subsequently centrifuged and yeast pellets were either subjected to lysis to obtain the yeast membrane fraction or alternatively used for confocal microscopy analysis.

2.2 Culturing *P. pastoris*

2.2.1 Stock solutions, media and yeast strains:

10X YNB (13.4% Yeast Nitrogen Base with Ammonium Sulphate without amino acids)

134 g of yeast nitrogen base (YNB) with ammonium sulphate and without amino acids was dissolved in 1 litre of water. The solution was heated to dissolve YNB completely in water. Alternatively, 34 g of YNB without ammonium sulphate and amino acids and 100 g of ammonium sulphate was used. The solution was filter sterilised and stored at 4°C.

0.02% Biotin

20 mg Biotin was dissolved in 100ml of water, filter sterilised and subsequently stored at 4°C.

20% Dextrose

200 g of D-glucose was dissolved in 1 litre of water, autoclaved on a liquid cycle and stored at room temperature.

5% Methanol

5 mL of methanol mixed with 95ml of water. The solution was filter sterilised and stored at 4°C.

10% Glycerol

100ml of glycerol mixed with 900ml of water. The solution was sterilised either by filtering or autoclaving, then stored at room temperature

1 M potassium phosphate buffer, pH 6.0

132ml of 1 M K_2HPO_4 and 868ml of 1 M KH_2PO_4 were mixed together before confirming the pH as $6.0 \pm 0.1 \bullet$ (if the pH was adjusted, phosphoric acid or KOH were used). The solution was sterilised by autoclaving then stored at room temperature.

P. pastoris media

Yeast Extract Peptone Dextrose Medium (YPD) (+/- Zeocin)

1% yeast extract, 2% peptone, 2% dextrose, 2% agar (optional 100 mg/ml Zeocin)

For 1 litre solution, 10 g yeast extract, 20 g of peptone were dissolved in 900ml water and 20 g of agar was added when making YPD plates. The solution was autoclaved for 20 minutes on a liquid cycle. The solution was then cooled to ~60°C before adding 100 mL of a 20% Dextrose stock solution and 1ml of a 100 mg/ml stock of Zeocin (if required).

Buffered Glycerol-complex Medium (BMGY) and Buffered Methanol-complex Medium (BMMY)

1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, 1% glycerol or 0.5% methanol

To make a 1 L solution, 10 g yeast extract and 20 g peptone were dissolved in 700ml water. The solution was autoclaved for 20 minutes on a liquid cycle and then cooled to room temperature. The following was then added and subsequently mixed well; 100ml 1 M potassium phosphate buffer (pH 6.0), 100ml 10× YNB, 2ml of a 0.02% Biotin stock and 100ml 10% glycerol (for BMGY) or 100ml 5% methanol (for BMMY). The media was stored at +4°C and kept for no longer than two months.

***P. pastoris* strains**

This project employs *P. pastoris* strains that were designed and developed by M. Jamshad (Jamshad *et al.*, 2008) and N. Bonander (Bonander *et al*., 2011, Bonander et al., 2013), shown in table 2.7.

Table 2.7: ***P. pastoris* yeast strains used to produce recombinant proteins.**

P. pastoris recombinant strain	Additional details
X33 CD81 palmitoylation- null (p-null)	Zeocin resistant
X33 Claudin-1 Wild-type (WT)	Zeocin resistant
X33 Claudin-1 p-null	Zeocin resistant
X33 CD82 p-null	Zeocin resistant
X33 WT	N/A

2.2.1.1 Storage of *P. pastoris*

Short term working stocks of *P. pastoris* were prepared by streaking one colony of each strain on to YPD-Zeocin (100µg/ml) plates, then incubated at 30°C for three days or until single colonies were visible. Plates were subsequently stored at 4°C until required for further use. Fresh plates were renewed by re-streaking yeast glycerol freeze stocks every two weeks.

Long term glycerol stocks of *P.pastoris* were prepared by adding one colony of each strain from a YPD-Zeocin plate to YPD (2-5 ml) and grown overnight at 30°C, 220rpm. After the incubation, 1ml of the saturated sample was added to 1ml of a 50% sterile glycerol stock and vortexed. All glycerol stocks were stored at -80°C until further use.

2.2.2 Expression of recombinant membrane proteins

2.2.2.1 Shake flask culture

For small scale expression of CD81 or Claudin-1 from *P. pastoris*, single colonies were picked into 10ml BMGY and grown overnight at 30°C, 220rpm. The pre-culture (10ml) was then added to 40ml BMGY and grown overnight (30°C, 220rpm) in small baffled shake flasks (250ml total volume) so that a 1:5 working volume was achieved. The following day, cells were centrifuged and resuspended into 50ml BMMY to begin induction from an OD_{600} 1. Cells were supplemented with 1% total volume of 100% methanol after 24 h induction and again at 48h. Total induction was for a total of ~ 50h.

For large scale recombinant protein expression, large baffled shake flasks were used. For each production batch, a 50ml pre-culture originating from a single colony was used to inoculate 8× 200ml working volume (1L shake flasks), therefore 1600ml BMGY to build yeast biomass. Once grown to an OD_{600} ~7-10, cell pellets were obtained by centrifugation and BMGY was removed. Cells were resuspended in BMMY for the start of induction using 8 x 500ml working volume cultures (using 2 L baffled shake flasks). Cells were supplemented with methanol (1% v/v) at 24h post-induction and subsequently harvested at around 48-54 h post-induction. This approach typically yielded ~80g total wet cell weight. Yeast cell pellets were subsequently used to isolate total membranes and further analysis of recombinant proteins.

2.2.2.2 Bioreactor expression of CD81

The following method used a 2L Applikon bioreactor.

Inoculum seed preparation

Pre-cultures were set up by adding a single colony of a *P. pastoris* strain to a falcon tube containing 15ml BMGY and left to incubate overnight. The pre-culture was then added to a baffled flask containing 50ml BMGY. This culture was grown overnight and

used to inoculate the bioreactor the following day to an OD_{600} of 1. All cultures were incubated at 30°C and 220 rpm.

Glycerol batch phase

Media used in bioreactors:

Basal salts medium (BSM)

26.7ml 85% phosphoric acid, 0.93 g calcium sulphate, 18.2 g potassium sulphate, 14.9 g magnesium sulfate heptahydrate, 4.13 g potassium hydroxide and 40.0 g glycerol were dissolved in water to a total of 1 L. The solution was autoclaved in the bioreactor.

PTM1 trace salts

Dissolved 6.0 g cupric sulfate-5H2O, 0.08 g sodium iodide, 3.0 g manganese sulfate-H2O, 0.2 g sodium molybdate-2H2O, 0.02 g boric acid, 0.5 g cobalt chloride, 20.0 g zinc chloride, 65.0 g ferrous sulfate-7H2O, 0.2 g biotin and 5.0 ml sulphuric acid in water to a final volume of 1 L and filter sterilised. The solution was stored at room temperature.

Yeast extract, peptone and glycerol (YPG)

10 g yeast extract and 20 g peptone were dissolved in 900ml water, added to a bioreactor and autoclaved. Separately, 10 g or 40 g glycerol were dissolved in 90ml water, was autoclaved then added to an autoclaved bioreactor using a needle and syringe.

Method:

Bioreactors were either set up using BSM and PTM_1 trace salts or using complex medium YPG. Antifoam J673A (1ml) was added to each vessel using a needle and syringe. If a minimal medium was used then PTM_1 trace salts (4.35ml) were added in the same way. The dissolved oxygen (dO) probe was allowed to polarise for > 6 h. Next, the stirrer (700 rpm), the temperature (set to 30°C) and aeration were applied to

the system. The pH of the vessel (either set to pH 5 or in later experiments pH 7) was controlled using 50% phosphoric acid and 28% ammonium hydroxide. The bioreactor was then inoculated with the yeast culture previously prepared so that the final OD_{600} was 1. The dO was maintained at ~30% by the automatic addition of air from a compressor or oxygen enriched air from a cylinder when the culture required it for growth. The culture was grown until the glycerol content was completely consumed; this was when the dO trace increased towards 100%. The culture was then starved to ensure all glycerol was consumed before beginning a methanol induction phase.

Methanol fed batch phase

Methanol was added to the bioreactor vessel in one of two ways. Initially 10ml methanol (100% sterile filtered) was added directly to the vessel using a needle and syringe. The culture was allowed to grow for 4-6 h prior to the addition of a further 10ml methanol, and then the culture was left overnight. The following morning 10ml methanol was added to the vessel in an aseptic manner and left to grow for a further 4-6 h prior to harvesting the yeast. Alternatively, methanol was added to the bioreactor using a feed rate of 4.0 ml/h. The feed rate was increased after 4 h of growth to 8 ml/h for the remainder of the run. The induction period lasted for ~26 h before harvesting yeast.

2.3 Protein determination

2.3.1 Yeast membrane preparations

Buffers and reagents:

Breaking buffer-

      50mM Na_2HPO_4 (N.B. use $Na_2H_2PO_4$ also)
      2mM EDTA

100mM NaCl

5% Glycerol (w/v)

pH 7.4 (used NaOH), autoclaved

Resuspension buffer (Buffer A)-

20mM HEPES

50mM NaCl

10% glycerol

pH 7 (NaOH), autoclaved

Small scale membrane preparation:

Yeast cell pellets were resuspended in 1ml breaking buffer and transferred to a 2 ml breaking tube. Acid washed glass beads (Sigma) measured to 1ml in an eppendorf were added to the breaking tube and kept on ice. Protease inhibitors (1:500) were then added to each tube.

A Quiagen Tissue lyser was used to break yeast cells. Briefly, the metal tissue lyser sample holder was placed into a -80°C freezer for 30 minutes prior to use. Samples were then balanced in the sample holder and the tissue lyser was set at 50 Hz for 10 min. Yeast lysates were then centrifuged at 700×g for 3 min to separate cell lysate from glass beads. The sample was subsequently resuspended and transferred to 1.5ml tubes and centrifuged at 700×g for 20 min at 4°C. The supernatant was then centrifuged at 100,000×g for 1h at 4°C. The resulting pellet was the yeast total membrane and the supernatant the soluble/ cytoplasmic yeast fraction. To yeast pellets, 100µl buffer A was added and allowed to soak overnight at 4°C. The following day the membrane fraction was resuspend in the buffer and used for total protein quantification using a BCA assay or stored at -20°C for later use.

Large scale membrane preparation:

Large cell pellets (i.e. harvested from a bioreactor or from 2L shake flasks) were passed through an Emulsiflex-C3 cell disrupter in order to isolate yeast membranes. The cell pellet (20 g cells) was resuspended in ice-cold breaking buffer at a ratio of 2:1 (v/w) buffer to pellet. Yeast protease inhibitors were added (1:2000 dilution). Cells were passed through an Emulsiflex-C3 cell disrupter (Avestin) ~6 times or approximately 15 minutes, reaching around 25,000-30,000 psi. Sample was collected from the Emulsiflex-C3 and remaining unbroken cells and cell debris was removed by centrifugation (10,000× g, for 30 min, at 4°C). The supernatant was collected and ultracentrifuged (100,000× g, 1h, at 4°C). The supernatant was discarded and the membrane pellet was re-suspended in ice-cold resuspension buffer using a glass homogenizer (adding 10ml resuspension buffer per g of membrane). Homogenised membrane fractions were either kept on ice if being used immediately or aliquoted and stored at -80°C for future analysis.

2.3.2 Protein solubilisation and purification

2.3.2.1 Buffers:

Solubilisation buffer:

1×PBS (137mM NaCl, 2.7mM KCl, 10mM Na$_2$HPO$_4$, and 2mM KH$_2$PO$_4$, pH 7.4)
150mM NaCl
10% Glycerol
20mM Imidazole
Yeast membrane stock (20% aliquots)

Detergent (either 6% OG or 1% DM for CD81 and 3% Fos-10 for Claudin-1).

Purification buffers:

Wash 1- 1×PBS,150mM NaCl, 10% Glycerol, 30mM Imidazole, detergent used at 3× CMC

Wash 2- 1×PBS, 150mM NaCl, 10% Glycerol, 40mM Imidazole, detergent used at 3× CMC

Elution buffer- 1×PBS, 150mM NaCl, 10% Glycerol, 250mM Imidazole, detergent used at 3× CMC

2.3.2.2 Method

Yeast membranes (18ml of 20% stock added to make up 70ml solubilisation buffer) of CD81 or Claudin-1 were solubilised using a solubilisation buffer for 2h at 4°C prior to ultracentrifugation at 100,000 ×g for 1h at 4°C to obtain the solubilised material in the supernatant and non-solubilised in the pellet. The solubilised fraction was then incubated with Ni-NTA agarose (Qiagen) for 1h at 4°C. The resin was added to a Bio-Rad Econo-column Chromatography column and at 4°C the flow through was collected followed by wash 1 (5 column volumes) and wash 2 (5 column volumes) of the resin. The His-tagged protein was then eluted from the resin using 2 column volumes of elution buffer. Samples from each stage were collected so that the protein could be detected on an SDS-gel. The protein concentration was determined using a NanoDrop, using the elution buffer as a blank). Protein was subsequently concentrated using an ultrafiltration unit Vivaspin 20 (50,000 Da molecular weight cut off) (Sartorius) by centrifugation at 3,900×g at 4°C. Every 5 min centrifugation, the protein samples was mixed by pipetting up and down to ensure the PES membrane did not get saturated and also to prevent protein aggregation. NanoDrop readings confirmed the protein concentration and typically concentration of ~10mg/ml CD81 was reached.

Protein was analysed by various instruments at the Membrane Protein Laboratory (Diamond Light source). The size exclusion chromatography coupled with a multi angle light scattering (SEC-MALS) detector was performed using a SEC-MALS system (Malvern) by Matthew Jennions (MPL).

Crystallisation trials were performed in collaboration with James Birch (MPL) using a Mosquito lipidic cubic phase (TTP Labtech). The crystallisation screen used two 96 well plates, MemGold and MemGold 2 (Molecular Dimensions) which encompass a range of 96 conditions each with varying pH, PEGs and salt additives based on crystallisation conditions taken from the Protein Data Bank. The Mosquito lipidic cubic phase (TTP Labtech) was used to automate the delivery of 100nl of CD81 protein sample to 100nl of each buffer in both 96 well plates (either MemGold or MemGold2). The 2 plates were then incubated at 20°C for 2-3 weeks for crystal formation.

2.3.3 CPM (7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin) thermostability assay

CPM (Invitrogen) was dissolved in DMSO at 4mg/ml and aliquots were stored at -80°C. Before use, CPM dye was diluted 1:100 in buffer (20mM tris, pH7.5, 150mM NaCl and 3× CMC detergent). Protein (CD81 or Claudin-1) was previously semi-purified using a Ni-NTA approach and typically 6-8µg CD81 and 4µg Claudin-1 was used for each condition tested. Using a black, clear bottom 96 well plate (Greiner) 150µl buffer (standard conditions- 20mM Tris, pH 7.5, 150mM NaCl and 3× CMC detergent) was added in duplicate. The standard buffer changed depending on what component was being investigated. Protein was then added to wells and incubated for 10 min at room temperature prior to adding the CPM dye, in order to equilibrate buffer conditions. 3µl diluted CPM dye was then added to wells (kept in the dark as much as possible). Using a SpectraMAX Gemini EM (Molecular devices) the plate was incubated at 40°C for 3h. Fluorescence was measured from the bottom of the plate every 5 min (Excitation 387nm, Emission 463nm, Cut-off 455nm). The plate was mixed prior to each measurement. In excel, at each time point raw data was converted to the fraction of folded protein by division of the highest fluorescence value reached for that condition as 'maximal unfolded protein'. A single one-phase decay curve was subsequently plotted using GraphPad Prism software (San Diego, CA) and the mean unfolding half life (min) was obtained for each condition tested. Experiments were repeated with at least two different protein preparations.

2.3.4 Bicinchoninic acid (BCA) protein assay

Protein concentration was determined using 1mg/ml Bovine serum albumin (BSA) as standard in a sterile 96-well flat bottom plate. A standard curve was generated by plating BSA standard in triplicate wells at concentrations of 0, 0.2, 0.4, 0.6, 0.8, 1.0 mg/ml, making up a total 10µl volume with PBS. Samples were plated in triplicate to a total volume of 10µl (a range of dilutions can be attempted to fall within the standard range). 200µl BCA reagent (a 50:1 ratio (v/v) between BCA solution and 4% (w/v) copper II sulphate solution) was added to all triplicate wells of standards and samples. The plate was incubated at 37ºC for 30 min. Absorbance values were then measured on a Biotek EL800 microplate reader using a 570nm filter. A standard curve was generated to calculate the protein content of samples, expressed as mg/ml.

2.3.5 SDS page and Western blot

2.3.5.1 Buffers and reagents

Laemmli sample buffer-

A 4x Laemmli buffer was made using 2.4ml 1M Tris pH 6.8, 0.8 g SDS stock, 4ml 100% glycerol, 0.01% bromophenol blue, 1ml β-mercaptoethanol and 2.8ml H_2O. The solution was stored in aliquots at -20°C (N.B. β-mercaptoethanol was added after thawing and before use).

Running buffer (Tris/ Glycine/ SDS buffer)-

100mL ultra pure 10× SDS Tris buffer (0.25M tris, 1.92M glycine and 1% SDS) (National Diagnostics) was diluted in 900 mL water.

Transfer buffer (Tris/ Glycine buffer)-

100ml ultra pure 10× Tris/ Glycine buffer (0.25M tris and 1.92M glycine) (National Diagnostics) was diluted in 200ml methanol and 700ml of water.

Phosphate buffered saline (PBS)-

10 PBS tablets (Fisher) were added to 1L water

Blocking buffer-

5% milk (Marvel dried milk) was added to PBS buffer (5 g milk in 100 ml PBS)

Wash buffer-

Tween-20 (0.2%) was added to PBS

Antibodies-

Antibodies shown in table 2.8 were used for western blot analysis.

Table 2.8. Primary and secondary antibodies. Primary antibodies were used to detect CD81, Claudin-1, Pma1p and a His_6 tag with an appropriate secondary antibody.

Primary Antibody	Dilution
Anti-CD81 (2.131) (mouse) (provided by Jane McKeating, Birmingham University)	1:100
Anti-Claudin-1 (rabbit) (Invitrogen)	2µg/ml
Anti-Pma1p (mouse) (Abcam)	1:200
Anti-His (mouse) (Clontech)	1:5000
Secondary Antibody	
Goat anti-mouse (HRP conjugated) (Sigma)	1:2000
Goat-anti-rabbit (HRP conjugated) (Abcam)	1:2000

Sodium dodecyl sulphate polyacrylamide gel electrophoresis gel preparation-

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) gels were prepared as detailed in table 2.9 and 2.10. Note that the separating gel and the stacking gel contained 12% and 4% polyacrylamide respectively.

Table 2.9: SDS-PAGE 12% separating gel components.

12 % separating gel	Volume to prepare gels (×2) (In order of addition)
Polyacrylamide (30%)	4.5 ml
Water	3.6 ml
Tris-HCl 1.5M, pH 8.8	3 ml
SDS (10%)	120 μl
Tetramethylethylnediamine (TEMED)	9μl
Ammonium persulphate (APS) 20%	40 μl

Table 2.10: SDS-PAGE 4% stacking gel components.

4% stacking gel	Volume to prepare gels (×2) (In order of addition)
Polyacrylamide (30%)	0.7 ml
Water	3.1 ml
Tris-HCl 0.5M, pH 6.8	1.3 ml
SDS (10%)	50 μl
TEMED	5 μl
APS 20%	20 μl

APS was added last to each gel solution. Propan-2-ol was added to the top of the separating gel resulting in a gel that set straight at the top. Propan-2-ol was removed from the gel after leaving to set, by rinsing with water and subsequently drying using Whatman filter paper. The stacking gel was then added to the top of the separating gel and a comb was inserted and removed when the gel solidified.

2.3.5.2 SDS PAGE

5µl of a 4x Laemmli sample buffer was added to a given amount of total protein sample (e.g. yeast membrane preparations) as determined by a BCA assay. Samples were made up to 25µl total volume with buffer A (20mM HEPES, 50mM NaCl, 10% glycerol, pH 7). The samples were vortexed and loaded on to an SDS PAGE gel in the presence of a Tris/ Glycine/ SDS running buffer. A Protometrics national diagnostics ladder (Fisher) was added to each gel (5µl) along with a transfer ladder (PageRuler^{TM} Plus prestained protein ladder from Fermentas). Using the Bio-Rad PROTEAN 3 cell the SDS PAGE was run at 150 volts for ~1 h or until the dye front reached the bottom of the gel.

2.3.5.3 Western Blotting

Following SDS PAGE, the separating gel was removed from the stacking gel. Fibre pads, filter paper (Whatman 3mm chromatography paper) and nitrocellulose (Whatman PROTRAN nitrocellulose transfer membrane) were soaked in transfer buffer to allow for equilibration. After, using a Bio-Rad colour coded easy lock cassette a transfer insert was arranged in the following order with the cassette placed black side-down.

- 1 fibre pad
- 3 filter papers
- Separating gel
- Nitrocellulose membrane
- 3 filter papers
- 1 fibre pad

The cassette, after being locked, was then placed into a colour-coded electrophoretic blotting cell. The blotting cell and a Bio-ice cooling unit were added to a Bio-Rad PROTEAN 3 cell. Using transfer buffer, the cell was then filled to the top and the transfer was carried out at 100V for 1h.

The nitrocellulose membrane was then removed and transferred to a container with blocking buffer (50 ml blocking buffer per membrane). The membrane was then

blocked overnight, covered at 4°C. The following day, the primary antibody was added directly to the blocking buffer (dilution dependent on antibody used, see table 2.8). The membrane was placed on to a rocker and incubated for 1 h (30 rpm). Then the membrane was washed three times in PBS-Tween (0.2%) for ~5-10 mins. The secondary antibody (see table 2.8) was then added to each membrane in blocking buffer and incubated for 1 h (30 rpm). A further 3 washes were performed on each membrane with PBS-Tween (0.2%). An EZ-ECL chemiluminescence detection kit (Geneflow) was used to detect protein bands, as suggested by the manufacturers' instruction guidelines.

2.3.6 Confocal Microscopy

Split-ubiquitin vectors were transformed into NMY51 cells and selected on either SD-W or SD-L plates, for example Prey CD81 expression was selected using SD-W. Single colonies were picked into 10ml SD media and grown overnight at 30°C, 220rpm. 500µl formaldehyde (3.7% in PBS) was added to the culture and incubated for ~15 min at 30°C, 220rpm. An OD_{600} 1 of cells were taken and washed twice in 1 ml PBS/0.5% Tween-20. Cells were then resuspended in 0.5ml 50µg/mL Zymolyase diluted in PBS and incubated for 20 min at 30°C. Washes were performed three times in 1ml PBS using centrifugation at 4,000rpm for 3min.

To stain cells with appropriate antibodies, cells were resuspended in 100µl primary antibody (anti-CD81 1:100 dilution in PBS/1% BSA) and incubated for 1h at room temperature, 30rpm. Cells were subsequently washed twice in 500µl PBS/BSA. Following washes, cells were resuspended in secondary antibody, Alexa Fluor-488 goat anti-mouse IgG (H+L) (Life Technologies) using 40µg/mL in PBS/1% BSA. Samples were incubated for 1h at room temperature in the dark and two washes performed after in PBS/BSA before centrifugation and resuspension of cells in 20µl PBS. Microscope slides were prepared by adding 10µl sample on to slide with 3µl of mounting solution. Cover slips were then attached and slides viewed under a confocal microscope by Charlotte Bland (Aston University).

2.3.7 Anti-CD81 ELISA

NMY51 cells expressing either a bait or prey protein fusion were grown and the cell pellet lysed for total membrane preparation. A BCA assay was performed on crude membrane preparations to determine the total protein concentration. 50µl total yeast membranes were added to a 96-well plate (Immulon II ELISA plates Nunc) in duplicate at 0.2µg/µl and left overnight to coat the plate overnight at 4°C. Unbound protein was removed by washing three times with PBS. Plates were blocked with 2% BSA/PBS adding 100µl per well at RT for 20min. After three washes, primary antibodies (anti-CD81 2.131 and 2.66, kindly provided by Professor Jane McKeating) diluted 1:2 in 0.05% Tween20/PBS were added to wells (50µl per well) and left to bind at RT for 1h. After three PBS washes, secondary antibody (ECL anti-mouse IgG, HRP linked whole antibody (from sheep), Amersham) was added at a dilution of 1:1000 in 0.05% Tween20/PBS (50µl per well) and incubated at RT for 1h covered in foil. Three washes were subsequently performed before adding TMB supersensitive one component HRP microwell substrate (50µl/well) and incubated at RT for 20min. Then, 450nm stop reagent for TMB microwell substrate was used and the absorbance measured on a Fusion plate reader (Perkin-Elmer) using a 450nm wavelength.

2.4 Molecular modelling

Molecular models were produced by Dr Jonathan Mullins (Swansea University) and were used in this thesis to guide amino acid mutagenesis targets.

2.4.1 Claudin-1 structural modelling by protein threading

Claudin-1 was modelled by protein threading using the I-Tasser program (Roy et al., 2010). Overall percentage sequence identity of the whole template chains with the query sequence ranged from 11-22%, which are typical values obtained for modelling with low and remote homologues. The top PDB hit in the list of templates used by I-Tasser is 3A0O (chain A), which is alginate lyase from Agrobacterium tumefaciens C58 (Ochiai et al., 2010).

2.4.2 Structural modelling by homology modelling

The CD81 and Claudin-1 structures were modelled using a homology pipeline assembled in the Biskit structural bioinformatics platform (Grunberg et al., 2007), which scans the entire PDB database for candidate homologs (Mullins, 2012). The pipeline incorporates elements of the NCBI tools platform (Wheeler et al., 2008) including BLAST for similarity searching of sequences (Altschul et al., 1990). T-COFFEE was used to align candidate sequences with the template (Notredame et al., 2000) and the MODELLER program was used for model assembly, iterated to final structure (Eswar et al., 2003). This software pipeline has been widely employed in the structural modelling of several membrane protein families (Mullins et al., 2011), (Davis et al., 2012).

CD81 was modelled using PDB: 2AVZ (Seigneuret, 2006) and includes the known coordinates derived from the CD81 crystal structure (PDB: 1G8Q) (Kitadokoro et al., 2001). Claudin-1 was ultimately modelled by 34% homology with the recently determined structure for mouse Claudin-15 (PDB: 4P79) (Suzuki et al., 2014). Interactions between these *in silico* models of Claudin-1 and CD81 were simulated using Hex 8.0 (Ritchie and Venkatraman, 2010), fitting for shape and electrostatic interactions.

2.4.3 *In silico* mutagenesis using FoldX analysis

Using the structural model of CD81-Claudin-1 produced by Dr Jonathan Mullins as template (as described in section 2.4.2), the computer program FoldX was used in a graphical user interface plugin for the YASARA molecular graphics suite (Van Durme et al., 2011) to calculate the free energy ($\Delta\Delta G$) difference for CD81 mutants. CD81 TM residues investigated in this thesis were substituted to alanine and the protein stability and the stability of the CD81-Claudin-1 complex were determined. A positive $\Delta\Delta G$ suggested the disruption of protein/complex stability and negative values implied increased stability. The default FoldX settings were used; temperature 298K, pH 7.0, Ionic strength 50mM, van der Waals design 2 and number of runs 3.

3. Optimisation of a DUALmembrane split ubiquitin yeast assay to explore CD81-Claudin-1 interactions

3.1 Chapter 3 objectives

A yeast hybrid based system designed for membrane proteins referred to as the DUALmembrane system (Dualsystems Biotech Ltd) was used to identify and explore protein- protein interactions using CD81 and Claudin-1. Chapter 4 describes site-directed mutagenesis studies on the interaction interface between CD81 homodimers, Claudin-1 homodimers and CD81-Claudin-1 heterodimers. Since these proteins had not been previously characterised in this experimental system, this chapter describes the optimisation steps taken to establish a suitably robust assay, such as:

- Optimisation of vector selection;
- Characterisation of protein specificity in assay;
- Reporter gene selection and in what format it is used;
- Protein localisation and expression analysis.

3.2 Overview of the split-ubiquitin method

The assay uses the split-ubiquitin system as devised by Nils Johnsson and Alexander Varshavsky (1994). Yeast ubiquitin can be split into two halves which are referred to as Nub (amino- terminal ubiquitin) and Cub (carboxy- terminal ubiquitin). When in close proximity the two halves bind to each other and adopt the native folded conformation. Ubiquitin proteases (UBPs) are able to recognise and cleave the 'whole' ubiquitin. Wild type Nub has an isoleucine at position 3 which is required for high affinity binding with Cub. When isoleucine is mutated to a glycine, the two halves do not associate readily and therefore UBPs do not cleave the protein. This feature is exploited in the DUALmembrane system because the two halves must be brought into close proximity by a protein-protein interaction as shown in figure 3.1.

The first protein of interest (Bait) can be fused to Cub (which is also fused to the synthetic transcription factor LexA-VP16) and the second protein of interest (Prey) can

be fused to mutant Nub (NubG) (see figure 3.1a). If the bait and prey protein have natural affinity for each other it will in turn force NubG and Cub to re-assemble. Therefore, the 'whole' ubiquitin will be recognised and cleaved by UBPs. Furthermore, the UBPs cleave the polypeptide chain between Cub and LexA-VP16. The artificial transcription factor, LexA-VP16, which is a fusion of the LexA DNA binding domain and the transcriptional activator VP16 (from the Herpes simplex virus) translocates to the nucleus after UBP cleavage. In the nucleus it binds LexA operators positioned upstream of a reporter gene. Transcriptional activation of the reporter gene takes place after VP16 recruitment of RNA polymerase II. The reporter genes used in the yeast hybrid system are two growth markers (*HIS3* and *ADE2*), when activated, allow the yeast to grow on minimal medium lacking histidine or adenine. Also *lacZ* is used as a reporter gene, which encodes the enzyme β-galactosidase, therefore when two membrane proteins interact this can be measured in a quantitative β-galactosidase assay (figure 3.1b)

b)

a)

Aston University

Illustration removed for copyright restrictions

Aston University

Illustration removed for copyright restrictions
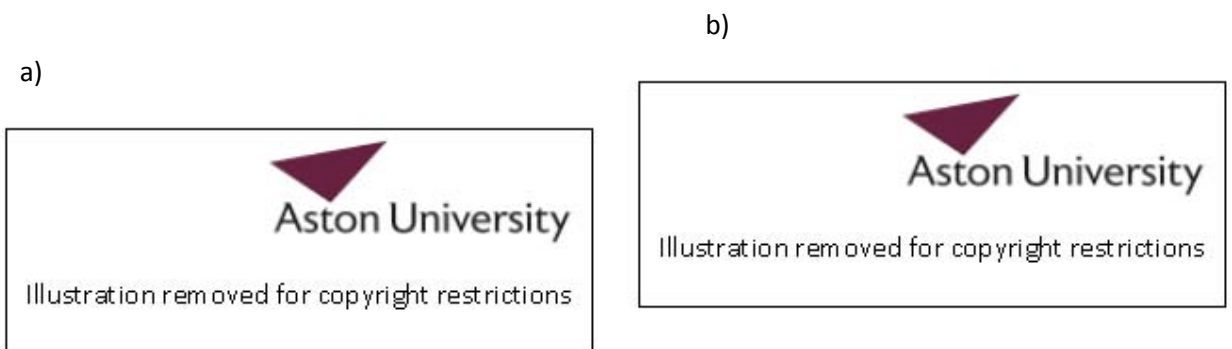
Figure 3.1: Bait and Prey reconstitution of split ubiquitin. a) Protein of interest chosen as Bait is fused to Cub (carboxy- terminal ubiquitin) and artificial transcription factor LexA-VP16. The second protein of interest is fused to mutant NubG (amino- terminal ubiquitin). If bait and prey proteins have a natural affinity and can bind to each other, this causes reconstitution of split ubiquitin. Ubiquitin proteases (UBPs) recognise reconstituted ubiquitin and therefore cleave a site between Cub and LexA-VP16. b) In the nucleus transcription is switched on as a result of LexA-VP16 translocating to the nucleus and activating reporter genes *ADE2*, *HIS3* and *lacZ* since there are LexA binding sites in the promoter region. Schematic taken from the DUALmembrane pairwise interaction kit manual (Dualsystems Biotech Ltd).

The initial aim was to clone CD81 and Claudin-1 into vectors that are compatible with the DUALmembrane system (which are provided by Dualsystems Biotech Ltd). The

choice of vector is dependent on the topology of the protein of interest. The cloning strategy employed here intended to increase the likelihood of success of this assay.

The *S. cerevisiae* yeast reporter strain used here, NMY51, is an auxotroph for tryptophan, leucine, adenine and histidine. The *ADE2* phenotype results in characteristic pink/red yeast colonies due to a red coloured intermediate that builds up in the adenine metabolic pathway. If the *ADE2* gene is activated here due to the presence of a protein-protein interaction then the cells display a white or pink colour depending on the interaction strength (white colonies representing the strongest interactions).

3.3 Generation of split-ubiquitin expression vectors

The human genes, *CD81* and *CLDN1*, were cloned from mammalian pTrip vectors (kindly provided by Professor Jane McKeating, Birmingham University) into various vectors that were compatible with the DUALmembrane system. The choice of vector was dependent on the topology of the protein of interest; in this case both membrane proteins contain their C- and N- terminus within the cytosol. In order for the system to work, Nub and Cub are required to be present on the cytosolic side of the membrane, therefore Nub and Cub could be fused to either the C- or N- terminus of both CD81 and Claudin-1 (see figure 3.2). According to the DUALmembrane user manual, three different vectors could have been suitable for each protein of interest (the reasoning is explained in figure 3.2). Therefore the genes encoding *CD81* and *CLDN1* were sub-cloned into six different vectors (three bait vectors and three prey vectors). Consequently, twelve novel vectors were produced and later tested in the DUALmembrane split ubiquitin assay.

a)

Aston University

Illustration removed for copyright restrictions

b)

Aston University

Illustration removed for copyright restrictions

c)

Aston University

Illustration removed for copyright restrictions

Figure 3.2: Rationale for generating different vectors encoding bait and prey with four transmembrane domains. a) Vectors that encode both prey and bait proteins. Carboxyl- and amino-terminal NubG and Cub LexA-VP16 proteins were generated to assess the potential effect of the relative orientations of each half of ubiquitin. Vectors that encode the *STE2* sequence were selected based on the manufacturer's guidelines as it has been found in certain targets to enhance translation. b) Schematic showing one combination of protein fusion pairing that were tested, C-terminal CD81 Cub fusion and N-terminal NubG fusion on Claudin-1, whilst c) shows N- and C- fusions, respectively.

Both *CD81* and *CLDN1* cDNA were cloned into bait and prey vectors (see materials and methods) in order to optimise the orientation of the Cub- NubG interaction. All

vectors were sequenced verified (example shown in figure 3.3) and used for subsequent interaction studies. A summary of vectors made, correctly sequenced and in frame can be found in table 3.1.
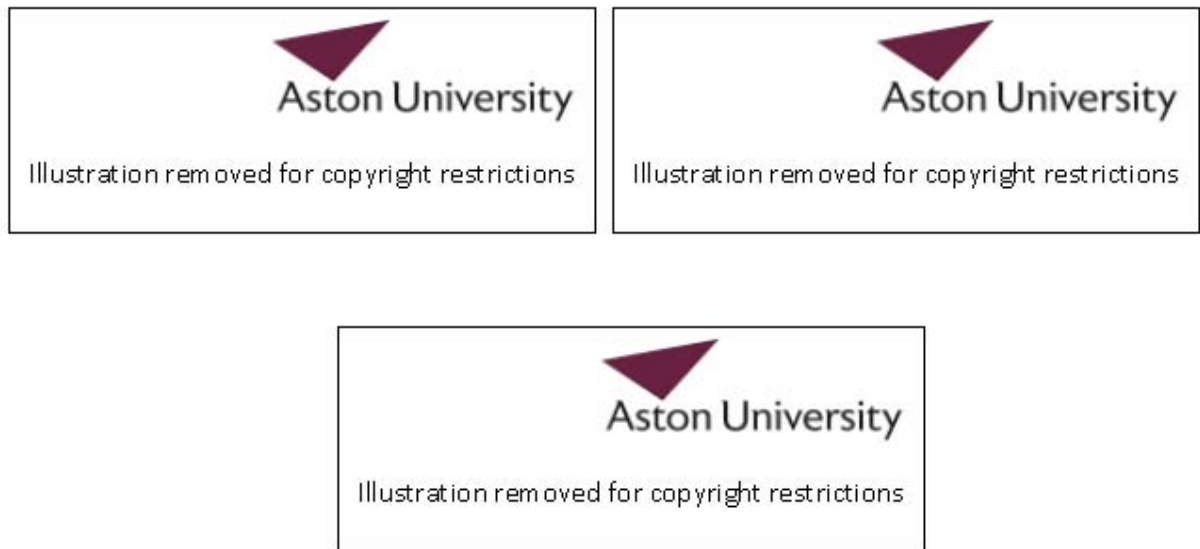
Figure 3.3: Sequence confirmation of pBT3-C-*CLDN1*. a) pBT3-C-*CLDN1* was sequenced with the pBT3-C forward primer. *CLDN1* gene sequence begins with start codon (ATG) which is preceded by 'AAAA' to ensure efficient initiation of translation. The 'AAAA' sequence is underlined and the start codon (ATG) can be found immediately after. b) pBT3-C-*CLDN1* was sequenced with the pBT3-C reverse primer. The blue triangle represents the end of *CLDN1*coding sequence. Note, that the stop codon of *CLDN1* gene has been removed. c) Represents the multiple cloning site of the pBT3-C vector and the correct reading frame of Cub coding sequence. *CLDN1* gene was sub-cloned using the two differing *Sfi I* sites. The figure represents that the *CLDN1* gene has been successfully sub-cloned into the pBT3-C vector in the correct reading frame. All sequencing primers can be found in DUALmembrane system manual (Dualsystems Biotech Ltd).

Table 3.1: Summary of final vectors generated. Bait and prey vector names show which gene was sub-cloned (*CD81* or *CLDN1* and if a *STE2* leader sequence (referred to as STE) is present at the 5´ end of the gene. Information is also given as to what protein is synthesised from a given vector and specifically whether the Cub (bait) or NubG (prey) fusion is on its C- or N-terminus.

Bait vector name	Protein expressed	Prey vector name	Protein expressed
pBT3- N- *CD81*	VP16-LexA- Cub-CD81	pPR3- C- *CD81*	CD81-NubG
pBT3-STE-*CD81*	CD81-Cub-LexA-VP16	pPR3- N- *CD81*	NubG-CD81
pBT3- C- *CD81*	CD81-Cub-LexA-VP16	pPR3- STE- *CD81*	CD81-NubG
pBT3- STE- *CLDN1*	Claudin-1-Cub-LexA-VP16	pPR3- C- *CLDN1*	Claudin-1-NubG
pBT3- C- *CLDN1*	Claudin-1-Cub-LexA-VP16	pPR3- STE- *CLDN1*	Claudin-1-NubG

3.4 Selection of optimal vectors for use in CD81-Claudin-1 interaction assay

To establish a robust read-out for CD81 homodimer interactions or CD81 heterodimer interactions with Claudin-1, a screen encompassing five bait vectors and five prey vectors was performed (see table 3.1). Since this was a screen intended to be executed quickly, there were no controls incorporated into these initial experiments. Any conditions that appeared interesting from the screen were repeated with all necessary controls.

Theoretically, each bait fusion protein has the potential to interact with each prey fusion protein, which equated to 25 conditions to be included in a screen. The NMY51 yeast strain was transformed with paired vectors (all shown in table 3.1), in different combinations, and then grown on three different plates lacking various components SD-W-L, SD-W-L-H and SD-W-L-H-A. The first plate selects for vector transformation as bait vectors include a *LEU2* marker and prey vectors a *TRP1* marker, therefore if both vectors are transformed yeast will grow in the absence of leucine (L) and tryptophan (W), respectively. The further two plates lacking histidine (H) and adenine (A) are selecting for yeast with activated reporter genes (*HIS3* and *ADE2*, which are activated due to a protein-protein interaction). All plates were left to incubate at 30°C and colonies were counted on day five of incubation as was suggested in the manufacturers'

manual. The numbers of red and white colonies on SD-W-L plates were recorded since in all conditions colonies will grow on this plate if both vectors are transformed, but they will turn red if cells express a non-interacting protein pair and white if they express a protein-interacting pair. The red phenotype is due to a block in the adenine synthesis pathway, therefore this colour accumulates over time if yeast cannot synthesise adenine due to not expressing a pair of interacting proteins (see table 3.2).

Table 3.2: Yeast colony counts from initial protein interaction screen. Each bait vector was co-transformed in NMY51 with each prey vector. After yeast transformation (see materials and methods) cells were incubated on selective plates for 4-5 days at 30°C before counting and recording the number and colour of individual colonies on each plate. The growth on SD-W-L-H and SD-W-L-H-A plates is the % of the total counted on the SD-W-L control plate for each reaction. Reactions in bold show where successful protein interactions were found.

Reaction	Bait and prey vectors involved in reaction	SD- W-L		SD-W-L-H		Growth on SD-W-L-H (%)	SD-W-L-H-A		Growth on SD-W-L-H-A (%)
		Colony colour		Colony colour			Colony colour		
		RED	WHITE	RED	WHITE		RED	WHITE	
1	pBT3-STE-*CD81* + pPR3-C-*CD81*	360	0	0	0	0	0	0	0
2	pBT3-STE-*CD81* + pPR3-C-*CLDN1*	432	0	0	0	0	0	0	0
3	**pBT3-STE-*CD81* + pPR3-N-*CD81***	**11**	**329**	**2**	**44**	**13.5**	**0**	**22**	**6.5**
4	pBT3-STE-*CD81* + pPR3-STE-*CD81*	412	1	0	0	0	0	2	0.5
5	pBT3-STE-*CD81* + pPR3-STE-*CLDN1*	407	1	0	0	0	0	0	0
6	pBT3-STE-*CLDN1* + pPR3-C-*CD81*	314	9	0	5	1.5	0	2	0.6
7	pBT3-STE-*CLDN1* + pPR3-C-*CLDN1*	500	0	0	10	2	0	6	1.2
8	**pBT3-STE- *CLDN1* + pPR3-N-*CD81***	**10**	**308**	**0**	**16**	**5**	**0**	**14**	**4.4**
9	pBT3-STE-*CLDN1* + pPR3-STE-*CD81*	338	8	0	3	0.8	0	3	0.8
10	pBT3-STE-*CLDN1* + pPR3-STE-*CLDN1*	304	12	0	9	2.8	0	6	1.9
11	pBT3-C-*CLDN1* + pPR3-C-*CD81*	221	13	0	2	0.9	0	1	0.4
12	pBT3-C-*CLDN1* + pPR3-C-*CLDN1*	136	5	0	0	0	0	0	0
13	pBT3-C-*CLDN1* + pPR3-N-*CD81*	266	21	0	3	1.0	0	1	0.3
14	pBT3-C-*CLDN1* + pPR3-STE-*CD81*	259	6	0	0	0	0	0	0
15	pBT3-C- *CLDN1* + pPR3-STE-*CLDN1*	234	6	0	0	0	0	0	0

Reaction	Bait and prey vectors involved in reaction	SD- W-L		SD-W-L-H		Growth on SD-W-L-H (%)	SD-W-L-H-A		Growth on SD-W-L-H-A (%)
		Colony colour		Colony colour			Colony colour		
		RED	WHITE	RED	WHITE		RED	WHITE	
16	pBT3-N-*CD81* + pPR3-C-*CD81*	201	7	0	0	0	0	0	0
17	pBT3-N-*CD81* + pPR3-C- *CLDN1*	251	11	0	0	0	0	0	0
18	pBT3-N-*CD81* + pPR3-N-*CD81*	129	6	0	0	0	0	0	0
19	pBT3-N-*CD81* + pPR3-STE-*CD81*	155	4	0	0	0	0	0	0
20	pBT3-N-*CD81* + pPR3-STE-*CLDN1*	206	15	0	0	0	0	0	0
21	pBT3-C-*CD81* + pPR3-C-*CD81*	246	0	0	0	0	0	0	0
22	pBT3-C-*CD81* + pPR3-C-*CLDN1*	252	0	0	0	0	0	0	0
23	pBT3-C-*CD81* + pPR3-N-*CD81*	136	4	0	3	2.1	0	3	2.1
24	pBT3-C-*CD81* + pPR3-STE-*CD81*	150	0	0	0	0	0	0	0
25	pBT3-C-*CD81* + pPR3-STE-*CLDN1*	212	3	0	0	0	0	0	0

Table 3.3: Overview of screen results. 5 bait vectors were used in combination with 5 different prey vectors to co-transform yeast strain NMY51, the expressed protein fusions are shown in the table in bold. After an initial screen of 25 vector combinations, colonies were counted on various selective plates. Symbols represent (-) no protein-protein interaction found, (?) low colony numbers but no protein-protein interaction and (y) shows positive protein-protein interactions.

PREYS BAITS	pPR3-C-*CD81* **CD81-NubG**	pPR3-C-*CLDN1* **Claudin-1-NubG**	pPR3-N-*CD81* **NubG-CD81**	pPR3-STE-*CD81* **CD81-NubG**	pPR3-STE-*CLDN1* **Claudin-1-NubG**
pBT3-STE-*CD81* **CD81-Cub-LexA-VP16**	-	-	y	-	-
pBT3-STE-*CLDN1* **Claudin-1-Cub-LexA-VP16**	?	?	y	?	?
pBT3-C-*CLDN1* **Claudin-1-Cub-LexA-VP16**	?	-	?	-	-
pBT3-N-*CD81* **VP16-LexA-Cub-CD81**	-	-	-	-	-
pBT3-C-*CD81* **CD81-Cub-LexA-VP16**	-	-	?	-	-

Results from the protein interaction screen indicated which reactions provided the optimal orientation for NubG and Cub fusions to detect membrane protein-protein interactions. Firstly, the colour of colonies on SD-W-L plates indicated if protein interactions had occurred (white) and if they had not (red). Reaction 3 (pBT3-STE-*CD81* + pPR3-N-*CD81*) and 8 (pBT3-STE-*CLDN1* + pPR3-N-*CD81*) gave the highest numbers of white colonies overall, being 329 and 308 respectively (see table 3.2). In contrast a negative result such as reaction 1 (pBT3-STE-*CD81* + pPR3-C-*CD81*) resulted in 360 red colonies and 0 white colonies on a SD-W-L plate.

Furthermore, a positive protein-protein interaction should result in the activation of two growth reporter genes *HIS3* and *ADE2*. Table 3.2 shows Reaction 3 resulted in 13.5% growth on SD-W-L-H and 6.5% growth on SD-W-L-H-A, whilst reaction 8 showed 5% growth on SD-W-L-H and 4.4% growth on SD-W-L-H-A. Both reactions demonstrated a result positive for protein-protein interactions compared to other reactions in the screen which did not show >1% growth on SD-W-L-H-A; therefore were considered negative because protein interaction signals were not detected.

3.5 Characterisation of Bait proteins CD81 and Claudin-1

Once having selected bait proteins, Bait-CD81 and Bait-Claudin-1, expressed from vectors pBT3-STE-*CD81* and pBT3-STE-*CLDN1*, respectively. It was then necessary to functionally characterise them by initial genetic experiments using control prey vectors (see figure 3.4). The positive prey control vector is pOst1-NubI that expresses a fusion of the yeast resident ER protein Ost1 to wild-type (WT) Nub portion of yeast ubiquitin (referred to as NubI). This should readily associate with the Cub fusion of the bait protein of interest if the bait protein is correctly expressed and the Cub-LexA-VP16 fusion is accessible to the NubI fusion. The negative control vector construct pPR$_3$N expresses NubG (mutated Nub) fused to a nonsense peptide fusion. Since NubG has no affinity for Cub there should not be activation of reporter genes as a result of the expression with the bait (protein of interest) with NubG. Using NubG expression with the bait protein would determine if the bait protein causes any 'leakiness' of reporter genes.

a)

NubI = **Activation** of reporter genes *HIS3*, *ADE2* and LacZ

b)

NubG = **No activation** of reporter genes *HIS3*, *ADE2* and LacZ
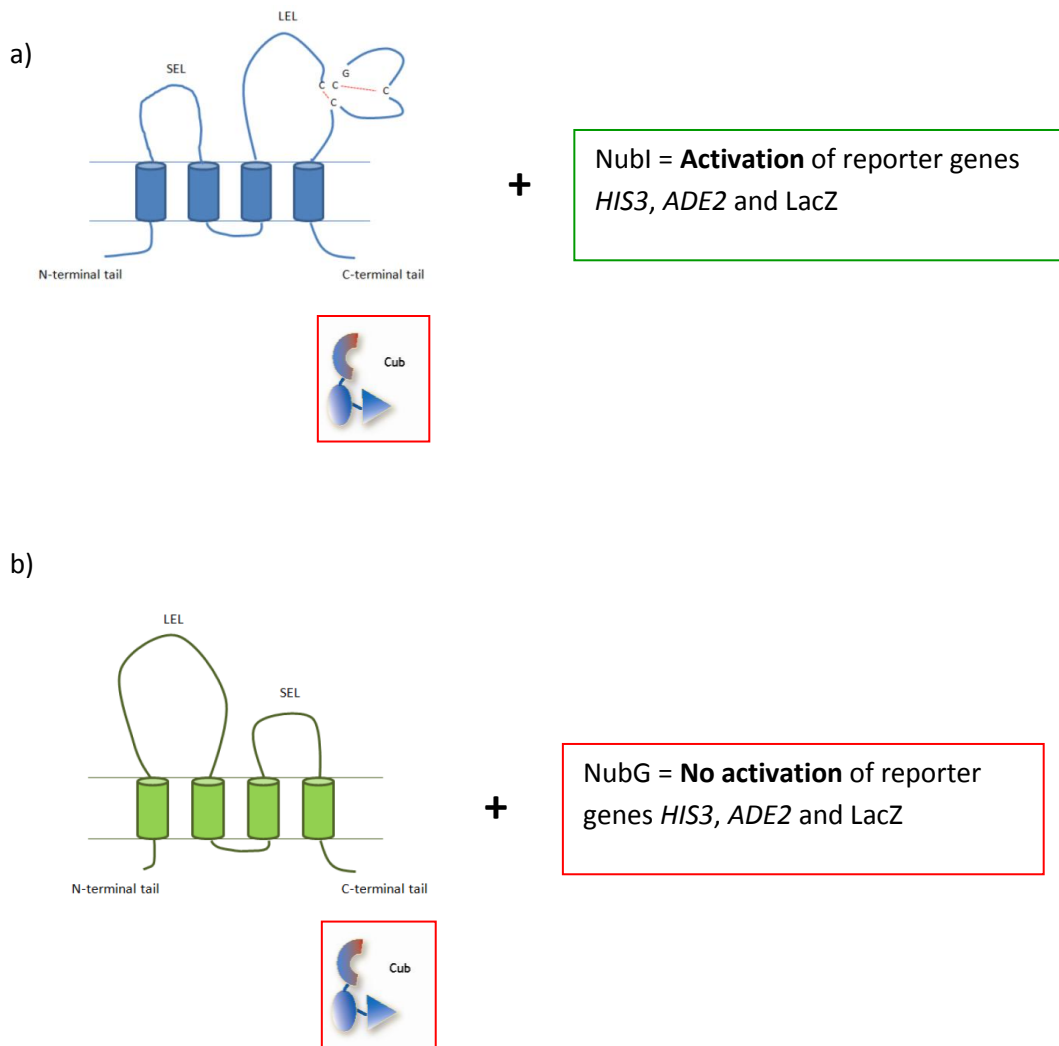
Figure 3.4: Schematic of prey NubI and NubG controls. a) Prey control vector pOst1-NubI expresses a fusion of the yeast resident ER protein Ost1 to WT Nub (referred to as NubI). NubI should associate with the Cub-LexA-VP16 fusion if the bait protein (this could be CD81 or Claudin-1) is properly expressed. Activation of reporter genes results as a consequence of ubiquitin proteases cleaving LexA-VP16 from the re-assembled ubiquitin molecule. b) Prey control vector pPR$_3$N expresses NubG (mutated Nub) fused to a nonsense peptide and has no natural affinity for Cub fused to the bait protein of interest. When co-expressed in yeast this should not lead to the activation of reporter genes.

Bait proteins, Bait-CD81 and Bait-Claudin-1 were tested for their expression and potential to activate growth reporter genes when co-expressed with one of the two control prey proteins, NubI or NubG (positive and negative control, respectively).
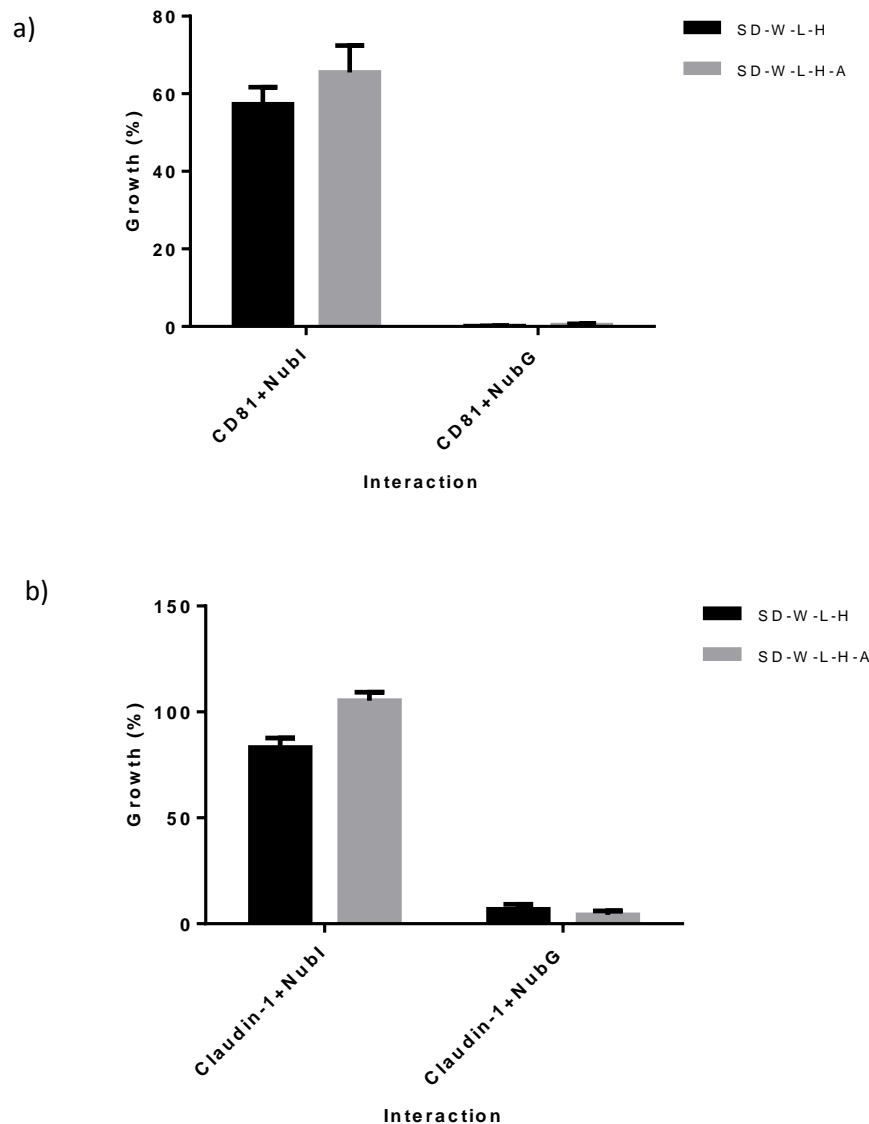
Figure 3.5: Bait CD81 and Bait Claudin-1 genetic experiment to determine expression and function using the split-ubiquitin system. Bait vectors pBT3-STE-*CD81* and pBT3-STE-*CLDN1* (expressing bait CD81 (a) or bait Claudin-1 (b) respectively) were co-expressed with either pOst1-NubI or pPR3N control prey vectors, expressing NubI or NubG portions of ubiquitin, respectively. After yeast (NMY51) co-transformations (see materials and methods) yeast were incubated for 5 days at 30°C before counting the number of colonies on three plates; SD-W-L, SD-W-L-H and SD-W-L-H-A. The percentage growth on SD-W-L-H and SD-W-L-H-A plates was calculated by using the number of colonies on SD-W-L as 100% growth. Experiments were repeated (n=3).

Figure 3.5 shows that the Bait-CD81 Cub fusion demonstrates high binding with NubI since 57.3% and 65.5% growth was observed on SD-W-L-H and SD-W-L-H-A plates respectively. This suggests successful CD81 expression and insertion into the membrane so that Cub and NubI can bind to each other and subsequently activate growth reporter genes. The same trend was seen for Bait-Claudin-1 with NubI (see figure 3.5) giving higher 83.2% and 105.3% growth on SD-W-L-H and SD-W-L-H-A plates respectively.

In contrast, minimal growth was observed for Bait-CD81 with NubG, 0.11% and 0.39% on SD-W-L-H and SD-W-L-H-A plates respectively (figure 3.5). This demonstrates CD81 has negligible levels of self auto-activation of reporter genes. Slightly higher signals of auto-activation were shown for Bait-Claudin-1 with NubG, 6.7% and 4.19% on SD-W-L-H and SD-W-L-H-A plates respectively, but this was lower than the signal observed for specific protein interaction signals later. (See figure 3.6).

3.5.2 Specificity of CD81 and Claudin-1 interactions

To demonstrate that the split-ubiquitin system in yeast can detect specific protein interactions, a positive control protein pair was used. APP (amyloid A4 precursor protein) APP, a type I integral membrane protein, was used as bait along with the cytosolic protein Fe65 (amyloid beta A4 precursor protein-binding family B member 1) as prey. Figure 3.6 demonstrates that APP and Fe65 interact in this assay as shown by 56.2% growth on SD-W-L-H and 58.3% on SD-W-L-H-A plates. The APP bait protein does not show any signal of auto-activation as shown using prey protein NubG as a negative control.

Since CD81 is used as a prey protein in this assay it was expressed with APP to assess its non-specific binding. Figure 3.6 show that APP + CD81 produce a negligible signal. Therefore, Prey-CD81 was then co-expressed with either Bait-CD81 or Bait-Claudin-1. The homodimer CD81-CD81 condition showed a positive interaction and gave 27.8% growth on SD-W-L-H and 24.8% growth on SD-W-L-H-A. A reaction demonstrating heterotypic interactions by Claudin-1 and CD81 also showed a positive interaction by resulting in 24.9% growth on SD-W-L-H and 17.2% on SD-W-L-H-A.
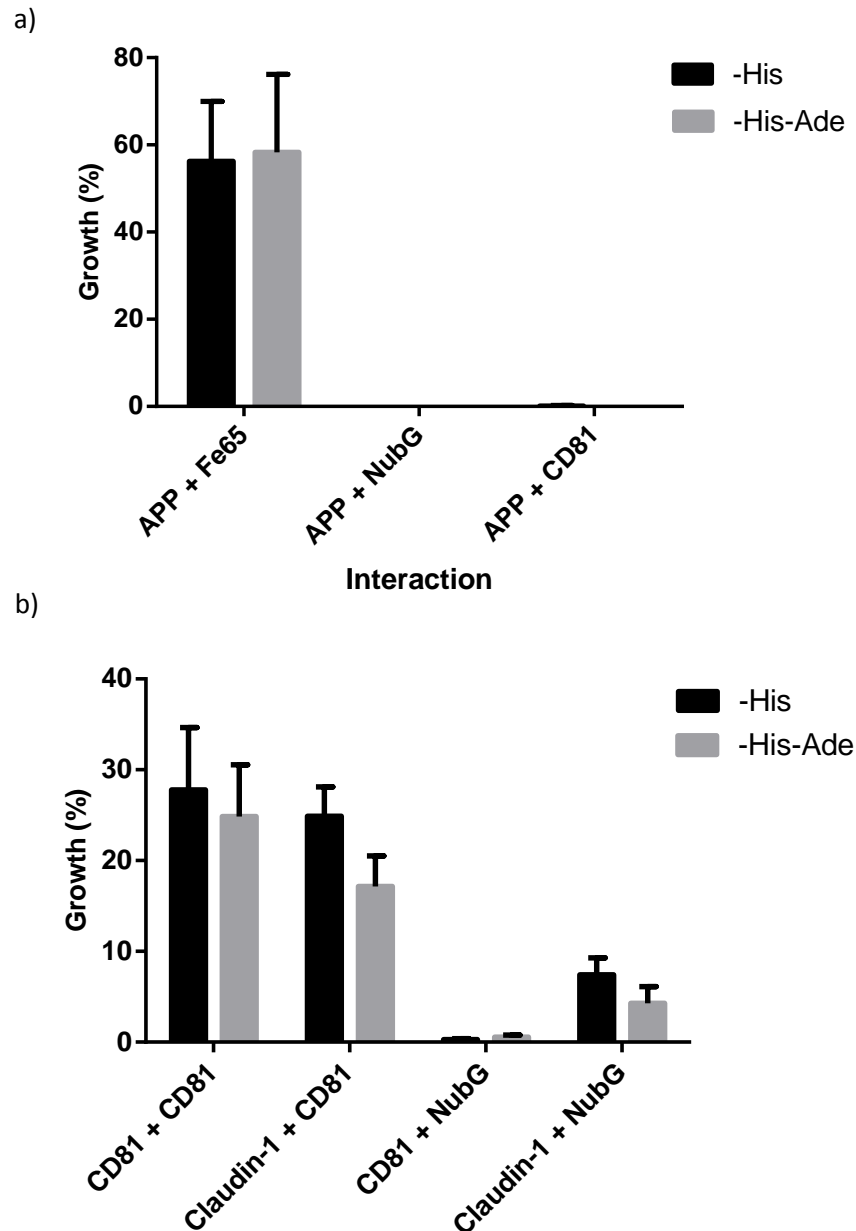
Figure 3.6: WT CD81 homo- and heterotypic interactions with Claudin-1 using the DUALmembrane split-ubiquitin assay. a) APP and Fe65 were used as a positive control protein interaction pair in the yeast split-ubiquitin assay. Prey negative control protein NubG was used to show bait auto-activation of reporter genes. b) Prey-CD81 was co-transformed with Bait-CD81 or Bait-Claudin-1. Following co-transformations (see materials and methods) yeast were incubated for 5 days at 30°C before counting the number of colonies on the three following plates; SD-W-L, SD-W-L-H and SD-W-L-H-A. The percentage growth on SD-W-L-H and SD-W-L-H-A plates was calculated by using the number of colonies on SD-W-L as 100% growth. Experiments were repeated (n=3).

3.5.3 Frequency of CD81 and Claudin-1 interaction

The frequency of a protein-protein interaction can be measured using a yeast split-ubiquitin assay as an indirect measurement of the rate of reporter gene expression, which results in the growth of yeast on selective agar plates. To demonstrate the frequency of WT interactions of CD81-CD81 and CD81-Claudin-1, a yeast spot assay was used on selective agar plates. Colonies selected to co-express two transformed plasmids were grown on SD-W-L media and then diluted to an OD_{546} 0.2. Diluted cultures were used to perform 10-fold serial dilutions on to selective agar plates (see figure 3.7).

1) SD-W-L   2) SD-W-L-H-A

a) Bait-CD81: NubI

b) Bait-CD81: NubG

c) Bait-CD81: Prey-CD81

d) Bait-CD81: Prey-Claudin-1

3) SD-W-L-H-A + 5 mM 3-AT

a) Bait-CD81: NubI

b) Bait-CD81: NubG

c) Bait-CD81: Prey-CD81

d) Bait-CD81: Prey-Claudin-1
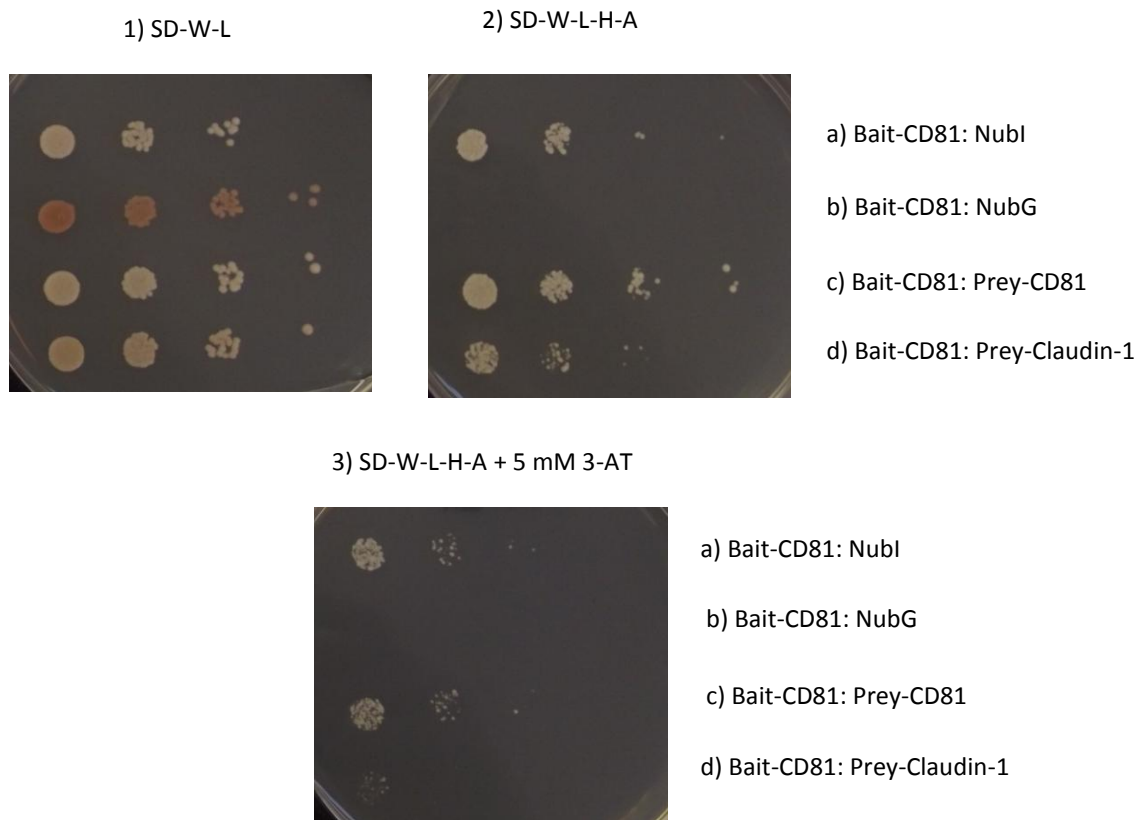
Figure 3.7: Yeast spotting assay to demonstrate the frequency of membrane protein interactions. NMY51 cells were selected in SD-W-L media to maintain the transformation of two plasmids (bait and prey vectors). Cultures were subsequently diluted to an OD_{546} 0.2 and used to perform 10-fold serial dilutions on to agar plates; SD-W-L, SD-W-L-H-A and SD-W-L-H-A + 5 mM 3AT. Serial dilutions were carried out from left to right (high to low cell concentrations, respectively) for co-expression of Bait-CD81 with a) positive prey control b) negative prey control c) CD81 homodimerisation and d) CD81-Claudin-1 heterodimerisation.

Figure 3.7 shows yeast growth spots on three different plates for four different protein reactions. The first plate, SD-W-L, is a control to show cells are transformed with two plasmids (bait and prey) that are selected for by their ability to grow. All four reactions labelled a-d can grow over the four dilutions, except the positive control, which may be due to reduced levels of yeast transformation using the vector that expresses NubI. Since growth on plates 2) and 3) is relative to growth on 1), this result does not suggest NubI is not a good positive control.

The second plate, SD-W-L-H-A, shows the level of activation of growth reporter genes *ADE2* and *HIS3* as a result of bait and prey protein interaction. Figure 3.7 demonstrates that the positive (a) and CD81 homodimerisation (c) reaction can grow successfully over four dilutions on SD-W-L-H-A whereas the negative control (b) does not grow at all. Finally the CD81-Claudin-1 reaction (d) can grow over three dilutions in contrast to four dilutions for CD81-CD81, suggesting that the CD81 homotypic interaction is more frequent than the heterotypic interaction as a result of increased levels of reporter gene activation. This observation is also shown in the third plate, SD-W-L-H-A + 5mM 3AT. Since 3-AT is a competitive inhibitor of the product of the *HIS3* gene, colonies cannot grow to the same extent as what they did on SD-W-L-H-A. Presence of 3-AT emphasises that CD81 homodimerisation (colonies shown over three dilutions) is more frequent than the CD81-Claudin-1 heterodimerisation interaction (colonies only shown in first dilution).

3.5.4 Localisation of bait proteins

After finding that WT CD81 and Claudin-1 interactions could be monitored/ detected in the yeast strain NMY51, it was imperative to investigate the expression of these WT fusion proteins and if they were present in the yeast membrane.

3.5.4.1 Immunoblot analysis

A few methods were used to assess expression of CD81 and Claudin-1. The first approach used western blot analysis which allowed independent analysis of protein

expression either in yeast membranes or cytosolic fractions. A method used by Gisler *et al.*, 2008 was modified slightly to isolate the yeast total membrane fraction from the cytosolic fraction using a glass bead cell lysis method followed by an ultracentrifugation spin to obtain the membrane (see materials and methods, section 2.3.1) (Gisler et al., 2008). In order to assess if the method was valid for membrane isolation a yeast plasma membrane ATPase, PMA-1p, was chosen as a housekeeping control protein that should be detected in the membrane but not in the cytosolic fraction.
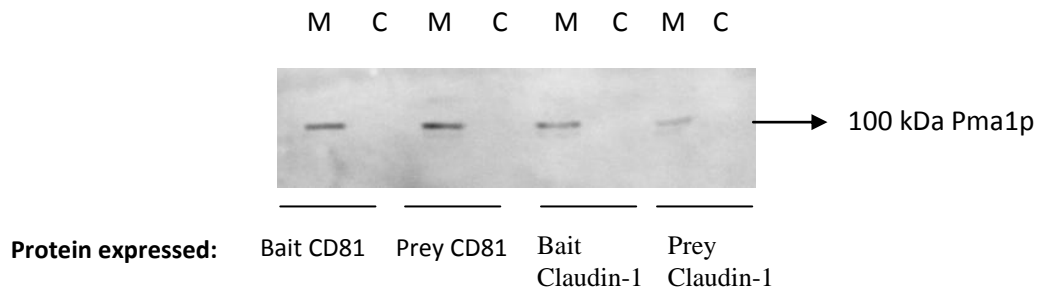
Figure 3.8: Method validation to isolate yeast membrane fractions. NMY51 cells were transformed with a single vector and grown for 3 days at 30°C. Single colonies were picked and grown overnight before harvesting. The cells were subsequently broken using a glass bead method followed by a high speed ultracentrifugation spin to separate yeast total membrane found in the pellet and the cytoplasmic fraction in the supernatant. Samples were quantified for total protein using a BCA assay and 40µg was loaded per sample on to an SDS gel (see materials and methods). The membrane (M) and cytosolic (C) fractions were run on a 12% separating gel and Pma1p was detected using an anti-Pma-1p antibody.

Figure 3.8 shows that the plasma membrane ATPase can be detected only in yeast membrane fractions and not present in any of the cytosolic fractions for cells expressing both CD81 and Claudin-1 bait and prey plasmids. This implies that this method for membrane isolation is valid and that it can be used to identify the correct expression of CD81 and Claudin-1 in the yeast membrane.

Subsequently, WT bait and prey CD81 and Claudin-1 proteins were transformed into the yeast strain NMY51 in order to investigate membrane protein expression using western blot analysis. As the previous validation method used 40µg total protein in order to observe Pma1p expression, the same total amount was loaded again. Results

revealed that only Prey CD81 and Prey Claudin-1 could be observed using this amount of total protein (see figure 3.9); in contrast neither bait proteins were detectable. This is likely to be due to the difference in yeast copy number depending on which plasmid the protein is expressed. Prey proteins are present in a high copy yeast plasmid whereas Bait proteins are in low copy yeast plasmids. Therefore, for western blot analysis, prey proteins were subsequently loaded and observed using only 15 µg of total protein (figure 3.9). Bait proteins were observed using confocal microscopy since the sensitivity of confocal microscopy is higher than that of a western blot (see figure 3.10).
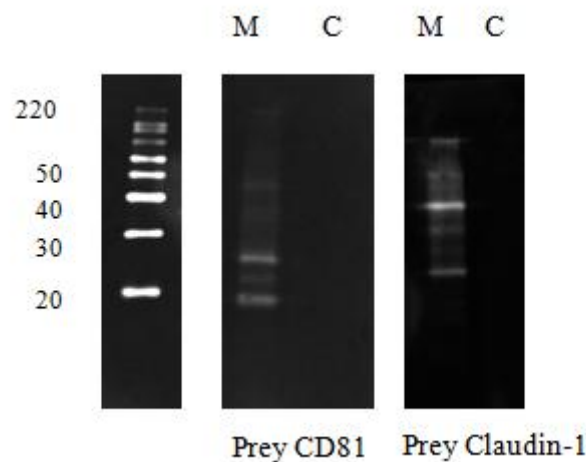
Figure 3.9: Expression analysis of Prey CD81 and Claudin-1. NMY51 cells were transformed with a single prey (CD81 or Claudin-1) vector and grown for 3 days at 30°C. Single colonies were picked and grown overnight before harvesting. The cells were subsequently broken using a glass bead method followed by a high speed ultracentrifugation spin to separate yeast total membrane found in the pellet and the cytoplasmic fraction in the supernatant. Samples were quantified for total protein using a BCA assay and 15µg was loaded per sample on to an SDS gel (see materials and methods). The membrane (M) and cytosolic (C) fractions were run on a 12% separating gel and proteins were detected using anti-CD81 or anti-Claudin-1 antibodies.

Western blot analysis shown Prey CD81 and Prey Claudin-1 were both present in membrane fractions of yeast and absent from cytosolic fractions (see figure 3.9). Non-transformed NMY51 cells did not show any protein bands using either antibody (data not shown). Prey CD81 is estimated to be 33 kDa and Claudin-1 30 kDa. Both ran approximately at the correct molecular weight but slightly lower which is characteristic of a membrane protein. Also both appeared as higher oligomeric bands which are expected in a non-reducing SDS-PAGE.

3.5.4.2 Confocal microscopy analysis

To gain an insight into the localisation of WT CD81 expression in the NMY51 yeast strain, protoplasts were prepared (see materials and methods) followed by staining with a specific anti-CD81 antibody and then viewed using confocal microscopy.

a)

b)
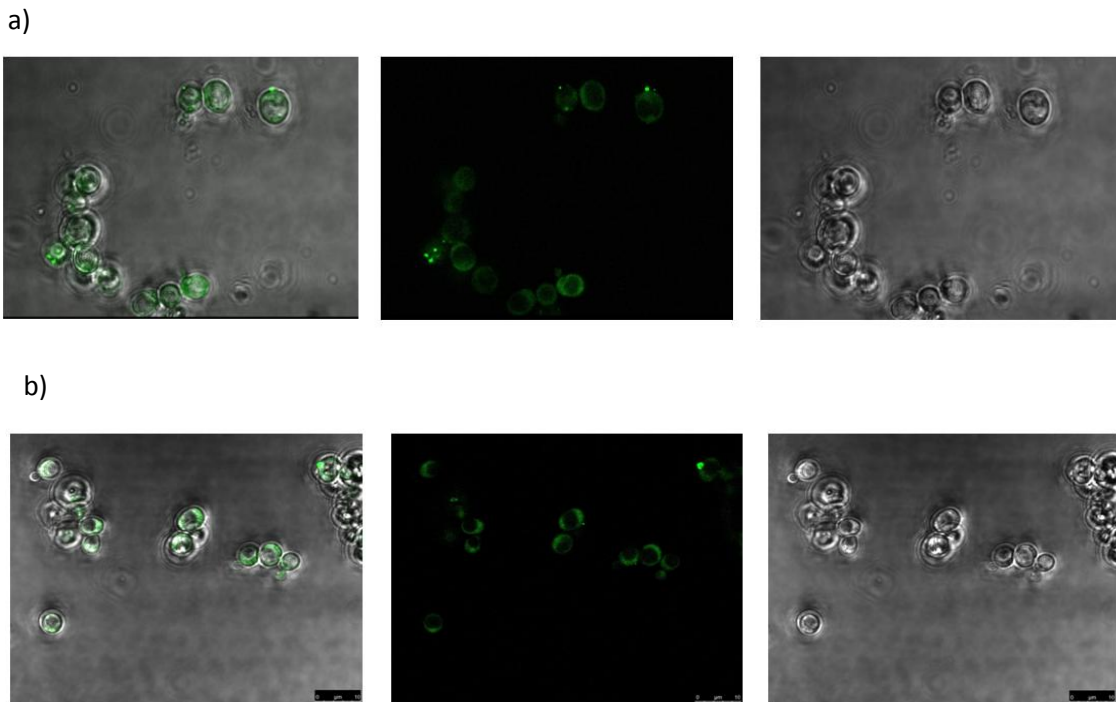
Figure 3.10: Visualisation of CD81 in *S. cerevisiae* NMY51 strain. NMY51 cells transformed with Bait-CD81 or Prey-CD81 were partially digested with Zymolyase to produce yeast protoplasts that were then stained with anti-CD81 2.131 followed by secondary antibody (Alexa Fluor-488 goat anti-mouse IgG (H+L), Life Technologies) (see materials and methods, section 2.3.6). Slides were prepared and viewed on a confocal microscope with the help of Charlotte Bland (Aston University). a) Bait CD81 expression and b) Prey CD81 expression in NMY51.

Figure 3.10 show that anti-CD81 2.131 can probe for Bait and Prey CD81 in yeast protoplasts. Gisler *et al*., 2008 also used a similar approach to show the expression of fusion proteins in the yeast membrane and although this method does not provide 100% evidence that the membrane protein resides in the plasma membrane specifically, the ring structures shown in figure 3.10 suggest that this may be the case. Also, since bait proteins are more difficult to capture using a western blot since expression levels are low, confocal microscopy provided an alternative method to visualise the expression of Bait-CD81 using specific antibodies against the protein.

3.6 Analysis of reporter genes

Now that the split-ubiquitin system was initially explored to look at WT protein interactions using the *HIS3* and *ADE2* growth reporter genes, it was then interesting to explore the third reporter gene, *lacZ*. The aim here was to first of all identify if the experiment could be performed in a timelier manner using any of the reporter genes and in a higher-throughput format i.e. using a 96-well plate. Secondly, if a quantitative assay was possible, that could support data generated, up to now, from manual counting of yeast colonies on agar plates.

3.6.1 *lacZ* reporter gene

Two different approaches were used to monitor *lacZ* reporter gene activation. The first used a HTX β-galactosidase assay kit (Dualsystems Biotech) and the second used Fluorescein-di-β-D-galactopyranoside (FDG, Invitrogen). The *lacZ* gene encodes β-galactosidase and active enzyme can be detected using substrate X-gal, which is an analog of lactose. X-gal is hydrolysed by β-galactosidase and produces galactose and 5-bromo-4-chloro-3-hydroxyindole (the latter product contributes to a blue product that is measured as the activity of the *lacZ* gene). The HTX β-galactosidase assay kit is an efficient approach for lysing yeast cells and subsequent detection of β-galactosidase activity using X-gal as substrate to provide a colorimetric read-out. The second approach uses FDG as substrate for β-galactosidase, which hydrolyses non-fluorescent FDG initially to fluorescein monogalactoside and then further to highly fluorescent fluorescein. lacZ gene activation can therefore be quantitatively measured by a highly sensitive fluorescent read-out. The similarities and differences between the two approaches are summarised in table 3.4.

Table 3.4: A comparison of two methods that quantitate *lacZ* activation. The HTX β-galactosidase assay kit allows efficient lysis and detection of β-galactosidase activity using an X-gal substrate and a colorimetric read-out. FDG is a sensitive substrate for β-galactosidase and is hydrolysed in its presence to produce highly fluorescent fluorescein.

Method	HTX β-galactosidase assay kit (Dualsystems Biotech)	Fluorescein di-β-D-galactopyranoside (FDG)
Format	96 well	96 well
Material used	Lysed yeast cells	Viable yeast cells
Type of measurement	Colorimetric	Fluorescence
Data generated	Quantitative	Quantitative
Substrate used	X-gal (blue colour development)	FDG (hydrolysed by β-galactosidase)

Two quantitative assays to measure β-galactosidase activity were performed on lysed or unlysed yeast cells (depending on the substrate used, see table 3.4). Cells were co-transformed with both Bait-CD81 and Prey-CD81 to detect homo-oligomers or Bait-Claudin-1 and Prey-CD81 to detect hetero-interactions.
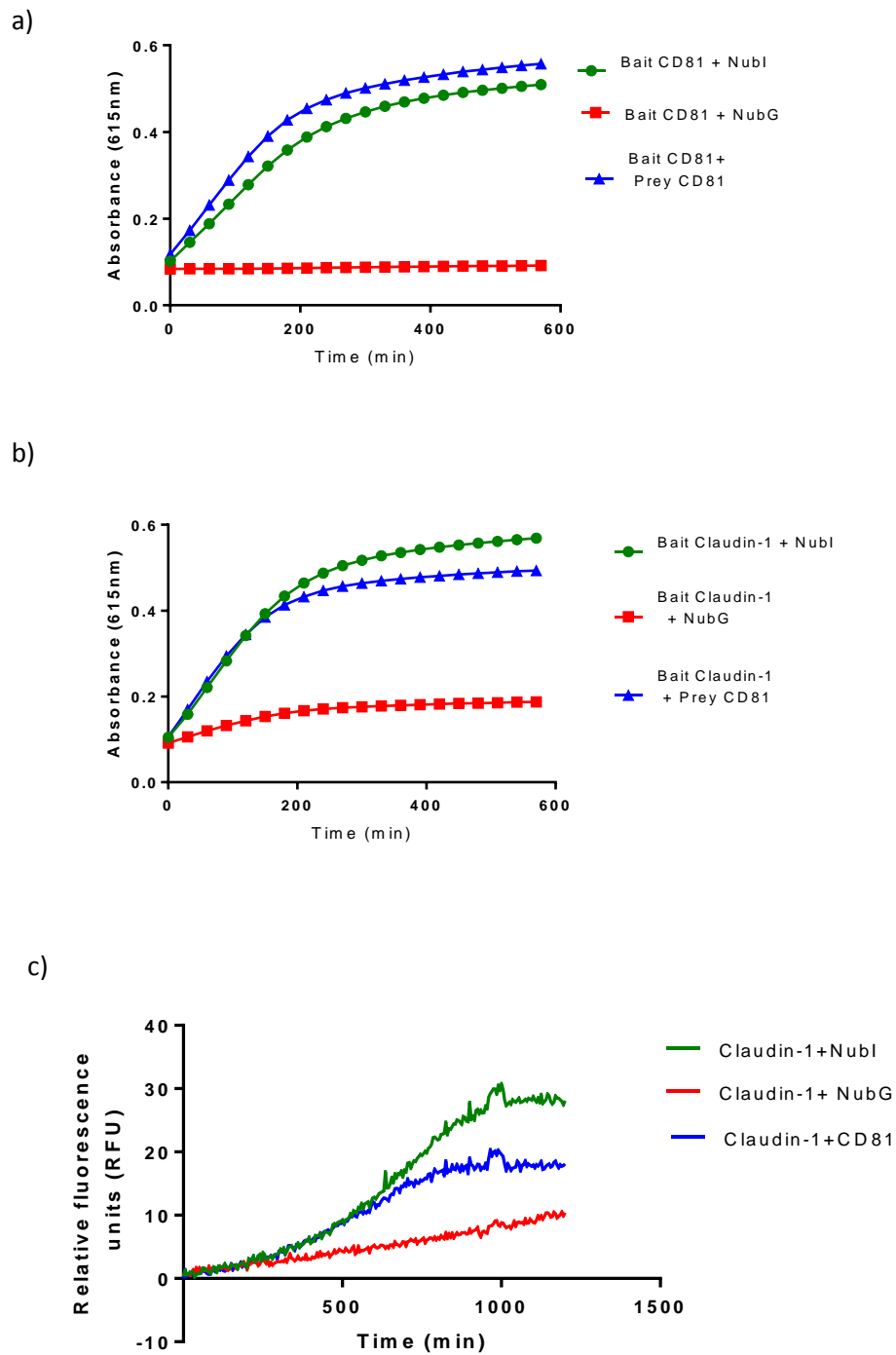
Figure 3.11: *lacZ* quantification using two different substrates. NMY51 cells co-transformed with bait and prey vectors were grown on agar plates for 4-5 days at 30°C. Pre-cultures were set up in SD-W-L media and grown overnight. a) CD81 homodimerisation and b) CD81-Claudin-1 heterodimerisation show *lacZ* expression using a HTX β-galactosidase assay kit (Dualsystems Biotech) of lysed cells at an OD_{546} 0.5-0.8. c) CD81-Claudin-1 heterodimerisation show β-galactosidase quantification over time using Fluorescein-di-β-D-galactopyranoside. Viable yeast cells grown from a starting OD_{546} 0.02 were selected for in SD-W-L, and then grown in SD-W-L-H-A during data point measurements.

Figure 3.11 indicates that tetraspanin interactions can be detected using this approach by measuring colour development of β-galactosidase over 10 hours. The positive control, Bait CD81 with NubI gave an increasing positive signal over time whilst the negative control, Bait CD81 with NubG gave a negligible signal over time. Also, figure 3.11b shows that Claudin-1- CD81 interactions can be detected using this method. The positive control (Bait Claudin-1 with NubI) showed a higher response than for the WT Claudin-1–CD81 interaction. In contrast the negative control (Claudin-1 + NubG) is low, although being slightly higher than the negative control for Bait CD81 plus NubG in figure 3.11a, which is consistent with what was observed using growth reporter genes, as shown in figure 3.6.

Dowell & Brown (2009) used yeast reporter genes to look at GPCR agonism using FDG (Dowell and Brown, 2009). Here, FDG has been used to quantitate *lacZ* reporter gene expression over time using viable yeast cells co-transformed with various vectors pairs. Figure 3.11c indicates FDG can be used to show CD81-Claudin-1 interactions and the results run parallel with that shown in 3.11b using a different substrate. Although, using this method, cells were grown in SD-W-L-H-A to provide these growth curves and so essentially three reporter genes; *HIS3*, *ADE2* and *lacZ* were used to generate this data.

Overall, both approaches allow determination of a protein-protein interaction using the *lacZ* reporter gene.

3.6.2 *HIS3* and *ADE2*

Since growth reporter genes *HIS3* and *ADE2* worked well to observe protein-protein interactions using an agar plate format and colony counting (seen earlier in this chapter), we then investigated if *HIS3* and *ADE2* could be used in a 96-well plate growth curve format using a spectrophotometer (OD_{546} vs. time). More specifically, alongside counting colonies on agar plates, transformed cells would then be picked and grown in liquid culture whilst the OD_{546} was monitored over time.

3.6.3 Optimisation of a growth curve read-out in 96-well plates

To begin with, when cultures were grown in 96-well plates there was an evaporation problem. Since yeast were grown at 30°C over 24 h, at the end of the incubation period there would be outer columns and rows with a major loss in culture volume, therefore data would be hugely variable and could not be relied upon. In order to prevent this issue, Breathe-Easy sealing membrane (Sigma) were used thereafter, which prevented excessive vapour loss throughout the course of the experiment incubation period.

NMY51 single colonies co-transformed with WT proteins (Claudin-1 and CD81) were pooled into dropout media, vortexed for 5min and diluted to an OD_{546} 0.02 in both SD-W-L and SD-W-L-H-A media. 200μl cells were added to wells across a 96-well plate and then incubated at 30°C in a shaking spectrophotometer and the OD_{546} was read every 1h to generate yeast growth curves.

a)

SD-W-L

b)

SD-W-L-H-A
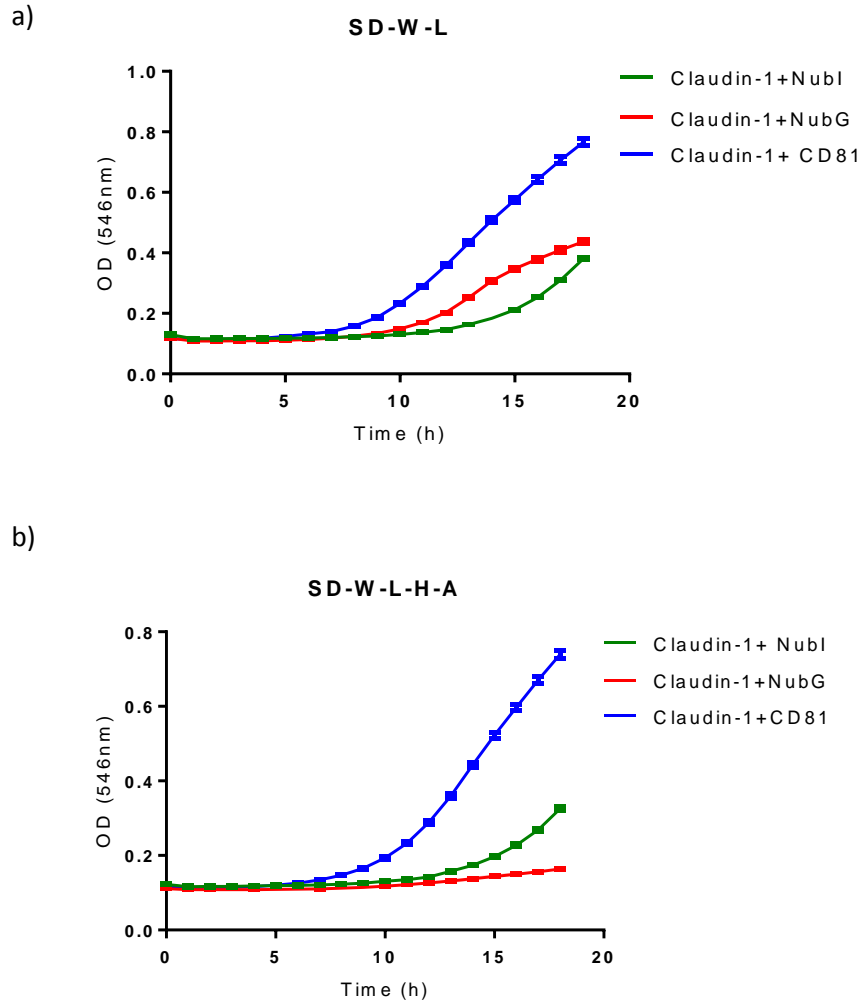
Figure 3.12: Liquid culture growth curves monitored over 18h in a 96-well plate. NMY51cells were co-transformed with appropriate vectors and grown on agar plates for 4-5 days at 30°C. Colonies selected on SD-W-L plates were pooled into dropout media, vortexed and diluted to an OD_{546} 0.02 in SD-W-L and SD-W-L-H-A media. 200µl of cultures were added to quadruplicate wells, plate covered with a Breathe Easy seal (Sigma) and then incubated at 30°C in a spectrophotometer. The OD_{546} was measured every 1h between shaking. a) growth in SD-W-L media selecting for double vector transformation, b) growth in SD-W-L-H-A selecting for protein-protein interactions using *HIS3* and *ADE2* reporter genes. Representative curves are shown of a least three independent experiments.

Figure 3.12 suggests that measuring OD_{546} over time is sufficient to monitor growth reporter gene activation (*HIS3* and *ADE2*), therefore protein-protein interactions. Figure 3.12a demonstrates growth over 18h in SD-W-L and b) growth in SD-W-L-H-A. It is clear that growth differs amongst the various reactions. This is apparent using SD-W-L for which the presence of two vectors is selected, showing the Claudin-1-CD81 interaction reaches the highest OD_{546} (~0.75), whilst the positive and negative reactions fall much lower (~0.4).

In more stringent selection media, SD-W-L-H-A, cells expressing Claudin-1-CD81 reached a higher OD_{546} than the positive control over 18 h initial growth. The negative control does not increase in comparison, reaching an approximate OD_{546} 0.16 at 18h growth (see figure 3.12b).

Since reactions differed regarding growth rate in SD-W-L (figure 3.12a), it was interesting to investigate why this was the case, as it might dictate how data were later analysed. One possibility was that adenine and histidine were not in excess in SD-W-L; therefore this would affect the growth ability of cells expressing a positively interacting protein pair and cells that were not. Those that expressed a positive protein-protein interaction would produce more adenine and histidine and in turn show a higher growth rate as a consequence of *HIS3* and *ADE2* being activated.

This idea was subsequently tested by using a negative control for a protein interaction, Claudin-1+ NubG. In theory a negative control should serve as a good condition to test as negligible levels of adenine and histidine should be produced by such transformed cells. The question would then be; can you increase growth by supplementing with higher levels of adenine or histidine than what is usually found in SD-W-L standard media? The experiment was set up using standard SD-W-L, as well as media supplemented with increasing amounts of adenine, histidine or both (see figure 3.13).
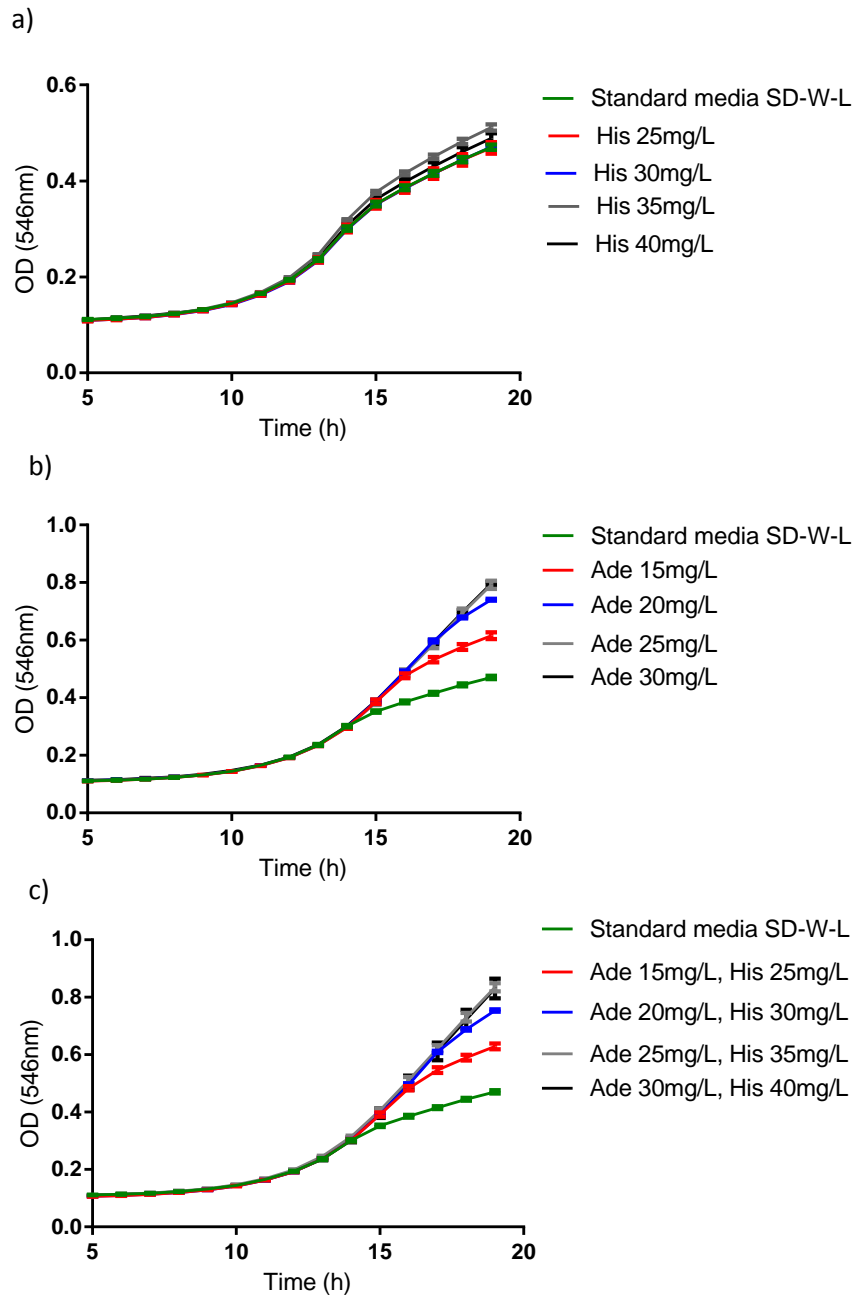
Figure 3.13: Growth curves to monitor supplementation with adenine or histidine and both. NMY51 cells co-transformed with vectors for Claudin-1 and NubG expression were selected on SD-W-L plates. Colonies were pooled into dropout media, vortexed and diluted to an OD_{546} 0.02 in appropriate media. 200µl cultures were added to a 96 well plate in triplicate wells and incubated at 30°C in a spectrophotometer whilst shaking. OD_{546} measurements were taken every 1h in between shaking for 20h. Representative curves are shown for triplicate mean values ±SEM. Experiment was repeated at least twice.

Figure 3.13a shows that increasing the concentration of histidine up to 40mg/L did not make a substantial difference in yeast growth over 20h compared to the levels of histidine already in standard SD-W-L. In contrast, figure 3.11b demonstrates increasing the concentration of adenine results in higher OD_{546} values measured over 20h. OD_{546} values approximately doubled from ~0.4 OD_{546} in standard SD-W-L to ~0.8 when supplementing with at least 25mg/L adenine, 30mg/L did not produce a further increase in OD_{546}.

When both adenine and histidine are supplemented together (figure 3.13c) the growth curve profiles are very similar to that seen with adenine alone, suggesting further that histidine does not provide additional yeast cell growth as compared to levels already in standard SD-W-L, whereas adenine levels maybe are lacking in standard SD-W-L for these specific cells. These data would suggest that cells expressing a positive protein-protein interaction may grow to higher OD_{546} values over time in SD-W-L compared to cells expressing a negative protein interacting pair. This could affect the way in which data is interpreted using yeast growth assays since data generated using SD-W-L-H-A may or may not be normalised to data generated using SD-W-L (in which vectors are selected for), since positive interacting protein pairs would likely grow to higher cell densities in SD-W-L than negative interacting protein pairs. Hereafter (in chapter 4), data from growth curves will not be normalised to SD-W-L, but will be treated independently when SD-W-L-H-A is used.

In summary, growth reporter genes can demonstrate the ability to monitor the occurrence of protein-protein interactions using a split-ubiquitin yeast assay. It is a relatively inexpensive approach since a substrate does not need to be used in contrast to using the *lacZ* reporter gene. Also, relatively little time is taken using transformed yeast to set up a 96-well plate prior to measuring OD_{546} over time. A limiting factor using this approach is that OD_{546} values appear to be accurate between 0.2-0.8. The spectrophotometer cannot read accurately above OD_{546} 1 and so data will plateau even if cells are still growing above this in a given well.

4. Exploring the CD81-Claudin-1 binding interface

4.1 Chapter 4 objectives

Claudin-1 and CD81 oligomerise with themselves (homo-oligomers) and with other interaction partners (hetero-oligomers) at the plasma membrane. Claudin-1 direct associations with tetraspanins, such as CD9 and CD81, were observed previously by Kovalenko and colleagues (2007) using a covalent chemical cross-linking approach. In a different study, CD81-Claudin-1 interactions were detected using fluorescence resonance energy transfer (FRET) (Harris *et al.*, 2010) and it was found that the complexes act as essential co-receptors in HCV entry. Interestingly, two residues in Claudin-1 extracellular loop 1 (ECL1) (see figure 4.1), I32 and E48, when mutated to alanine cause a knock-down in CD81-Claudin-1 association which in turn inhibits HCV entry, demonstrating that the receptor complex is essential in the infection process.
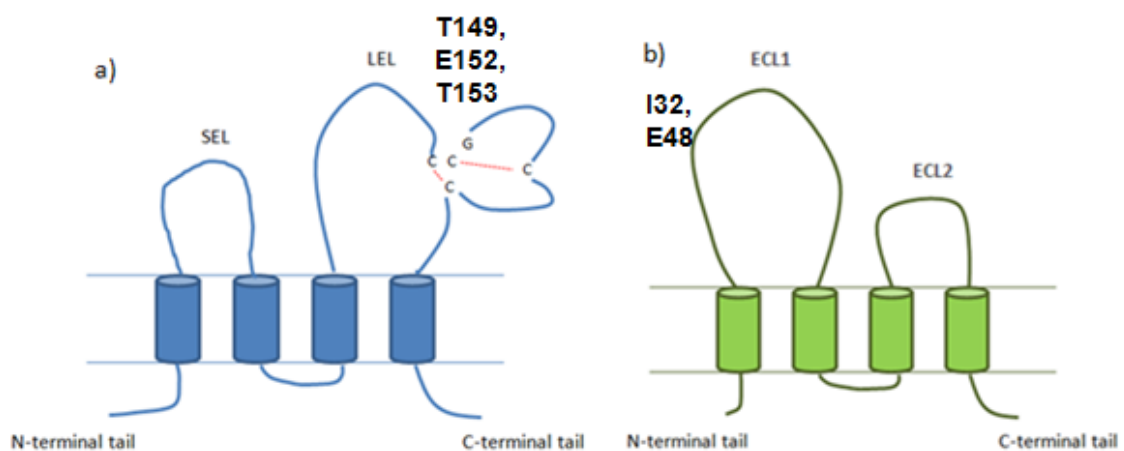
Figure 4.1: CD81 and Claudin-1 schematic. The two membrane proteins explored in this study. a) CD81 as shown with four transmembrane domains, a small extracellular loop (SEL) and a large extracellular loop (LEL). b) Claudin-1 also crosses the plasma membrane four times and has a larger extracellular loop 1 (ECL1) and a smaller extracellular loop 2 (ECL2). Residues found important for CD81-Claudin-1 interactions in CD81 LEL (T149, E152 and T153) (Davis et al., 2012) and Claudin-1 ECL1 (I32 and E48) (Evans et al., 2007), (Davis et al., 2012) are in bold.

Since HCV entry into hepatocytes is a potential target for antiviral therapy (Xiao et al., 2014) research into the binding interface between CD81-Claudin-1 is of great importance. Davis *et al.,* (2012) used a molecular modelling approach to predict that the CD81 large extracellular loop (LEL) residues T149, E152 and T153, are important for binding to Claudin-1 ECL1 (predicted motif between residues 62-66) (see figure 4.1). Mutagenesis and FRET were performed to test the prediction and results showed that CD81 LEL residues (T149, E152 and T153) are imperative in CD81-Claudin-1 interactions and consequently in HCV infection.

It has been described in the literature that the LEL also contributes to CD81 dimerisation (Drummer et al., 2005), although residues throughout the full length molecule, including extracellular and transmembrane (TM) regions, participate in CD81-CD81 interactions (Kovalenko *et al*., 2004). Therefore, investigation into CD81 transmembrane regions involved in the CD81-Claudin-1interaction would extend our knowledge of the full length protein binding interface.

This chapter describes an investigation of both LEL and transmembrane regions of CD81 involved in CD81-CD81 and CD81-Claudin-1 associations, using the yeast split-ubiquitin method (explored in Chapter 3), site-directed mutagenesis and molecular modelling. Specific objectives include:

- Assess CD81 LEL residues T149, E152 and T153 in the CD81-Claudin-1 interaction and the CD81-CD81 interaction;
- Explore CD81 TM1 residues; L14, F17 and C89 in CD81 homo-oligomer and hetero-oligomer interactions;
- Assess the expression of WT CD81 and mutant CD81 using western blot analysis and ELISA;
- Test CD81 TM residues as predicted by a molecular model in the yeast split-ubiquitin method.

4.2 Assessment of protein-protein interactions

Details on optimisation of the yeast split-ubiquitin method used in this chapter to monitor protein-protein interactions were presented in chapter 3. This chapter explores

specific residues in CD81 that are predicted to be involved in the CD81-Claudin-1 interaction. Data were collected in two different ways in order to determine the most sensitive method for comparing WT and mutant protein interactions:

1) The first approach used yeast colony counts as a quantitative measure of the output of a protein interaction. Yeast strain NMY51 was co-transformed with two vectors (one encoding the bait and the other encoding the prey protein) then grown on plates (SD-W-L, SD-W-L-H-A) for 5 days prior to enumerating colonies. Data was represented in two ways; the number of colonies grown on SD-W-L-H-A (selecting for a protein-protein interaction normalised to NubG and NubI controls to account for day to day variability) and secondly, the percentage growth on SD-W-L-H-A relative to colonies grown on SD-W-L to account for vector transformation efficiency (as reported by the ability of cells to grow without tryptophan and leucine) (see figure 4.2).

a)

SD-W-L-H-A →

WT Bait and Prey reaction
Claudin-1-CD81

e.g. 50 colonies, therefore 24.2% when
normalised to positive and negative controls

SD-W-L-H-A →

WT Bait and Prey reaction
Claudin-1 + NubI (positive control)

e.g. 200 colonies (100% growth)

SD-W-L-H-A →

WT Bait and Prey reaction
Claudin-1 + NubG (negative control)

e.g. 2 colonies (0% growth)

b)

SD-W-L

SD-W-L-H-A →

WT Bait and Prey reaction

Claudin-1-CD81

E.g. 10% Growth

Vector transformation
Efficiency
e.g. 500 colonies

Selection of protein-
protein interaction using
HIS3 and *ADE2* reporter
genes
e.g. 50 colonies

Figure 4.2: Yeast agar colony count data analysis. Colony count data analysis was performed in two different ways and both are shown in subsequent figures to analyse WT and mutant protein interactions. a) Data is represented as the % of colonies on SD-W-L-H-A plates. The colony number counted on the WT protein-protein test condition is normalised to that grown for the positive (NubI) and negative (NubG) internal controls (co-transformed with the same bait protein as the test condition). Example colony numbers and associated normalised data are shown. b) Data is represented by using the results from two agar plate colony counts, that on SD-W-L (selecting for vector transformation) and SD-W-L-H-A (selecting for a protein-protein interaction). The final % growth for a reaction is the number on SD-W-L-H-A as a percentage (%) of the number grown on the SD-W-L plate. Example colony and growth (%) is shown as an example.

2) The second approach used to quantify protein-protein interactions used an automated measure of cell growth over time in SD-W-L-H-A liquid media in a 96-well format. Comparing WT and mutant CD81 binding to Claudin-1 was achieved as an indirect measurement of reporter gene (histidine and adenine) activation. Growth curves were plotted as OD_{546} versus time (h). Growth curves were used to determine the OD_{546} of the WT interaction compared to a mutant interaction at a specific point when the internal positive control (Claudin-1 + NubI) reaches an OD_{546} of 0.2. This point was chosen to analyse data between independent experiments because OD_{546} 0.2 is approximately the beginning of the exponential phase of growth for the NubI with Claudin-1. Note that this interaction shows a longer lag growth phase in comparison to WT CD81-Claudin-1 growth profiles (see figure 4.3), maybe due to a lower occurrence of NubI interactions with Claudin-1 than that of CD81 with Claudin-1. Therefore, as an internal positive control, it provides a point at which to compare the OD_{546} reached of WT and mutant CD81 interactions. The frequency of a protein-protein interaction is relative to the activation of reporter genes (histidine and adenine) and the ability of yeast to grow in SD-W-L-H-A lacking histidine and adenine demonstrates the occurrence of such interactions.
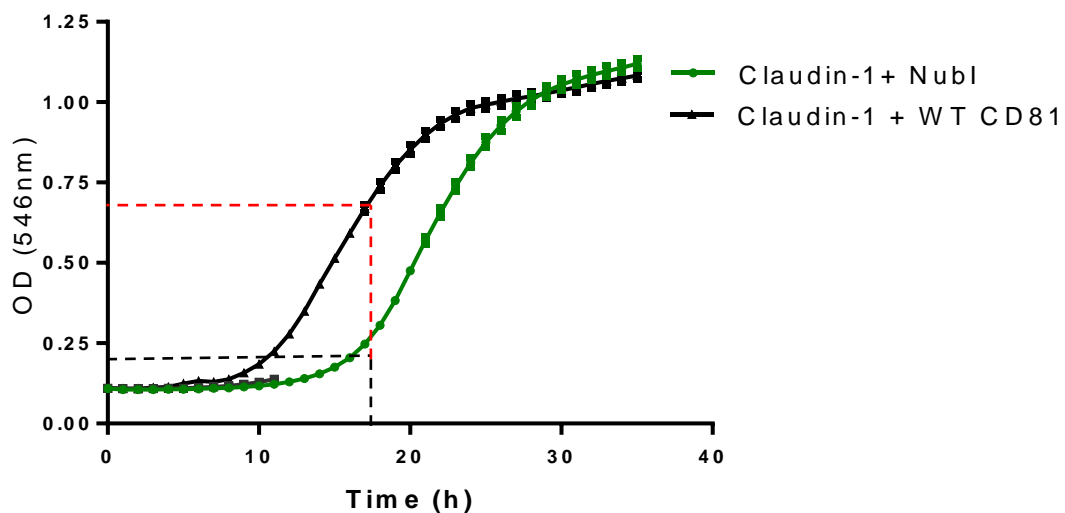
Figure 4.3: Growth curves in SD-W-L-H-A showing differences in lag phase before exponential growth of two protein reactions. *S. cerevisiae* strain NMY51 was co-transformed with two vectors (either encoding bait Claudin-1 and prey NubI or bait Claudin-1 with prey WT CD81) and allowed to grow for 5 days selected on SD-W-L plates (see materials and methods). Colonies were pooled from each plate and diluted to an OD_{546} 0.02 before measuring the OD_{546} over time using a 96 well format in SD-W-L-H-A for 35 h at 30°C. Representative growth curves are shown as the mean and SEM of four growth curves measured in parallel. Both bait and NubI and bait and WT CD81 were included in all subsequent experiments for data analysis as the NubI reaction was used as an internal control to compare WT and mutant CD81 growth

at a certain point of the curve (when NubI reaches OD_{546} 0.2, the WT and mutant reaction OD_{546} is taken and plotted in a separate graph at this point).

4.3 CD81 LEL residues involved in Claudin-1 interactions

To investigate previously reported residues in CD81 LEL that, when mutated, disrupt CD81-Claudin-1 interactions, three residues T149, E152 and T153 were mutated using site-directed mutagenesis to alanine. *S. cerevisiae* strain NMY51 was then co-transformed with bait and prey vectors (WT or mutant CD81) and grown on selective plates before counting colonies formed. This would allow for validation of the yeast-based assay.

Aston University

Illustration removed for copyright restrictions

Figure 4.4: CD81 LEL interaction with Claudin-1 ECL1. Molecular modelling of the loop domains, CD81 LEL and Claudin-1 ECL1 show the residues predicted and demonstrated to be involved in CD81-Claudin-1 interaction. Residues labelled in CD81 LEL (green) tested here using a split ubiquitin assay are T149, E152 and T153 (taken from Davis *et al*., 2012).

Figure 4.4 shows the position of the mutated residues in CD81 LEL along with residues in Claudin-1 ECL1 (between residues 62-66) that are involved in the interaction as described previously by Davis et al., (2012). Figure 4.5a and c show the results of Claudin-1 interaction with CD81 WT, T149A, E152A or T153A in the yeast split-ubiquitin method. A colony count approach showed that there is a decrease in the interaction of Claudin-1 with CD81 T149A, E152A and T153A (2.48%, 4.70% and 4.53%) as compared to WT CD81 (6.58%) when normalising to SD-W-L to account for transformation efficiencies (figure 4.5a). Only CD81 T149A showed a statistically significant decrease compared to WT CD81 (figure 4.5a and c). When the transformation efficiency is not considered, the same trend is observed for the number of colonies grown on SD-W-L-H-A alone (but normalised to internal NubI and NubG controls), with CD81 T149A showing further reduced binding to Claudin-1 than CD81 E152A or T153A (figure 4.5c).

Expression analysis performed on total yeast membranes showed that WT CD81 and mutant prey proteins (E152A and T153A) are expressed to similar levels to WT CD81, whilst CD81 T149A expression is slightly higher than WT (figure 4.5b). Reduced interactions shown between mutated CD81 and Claudin-1 (figure 4.5 a and c) are likely a result of the residues being important in the WT binding interface rather than due to the lack of expression of mutant CD81, as compared to WT CD81.
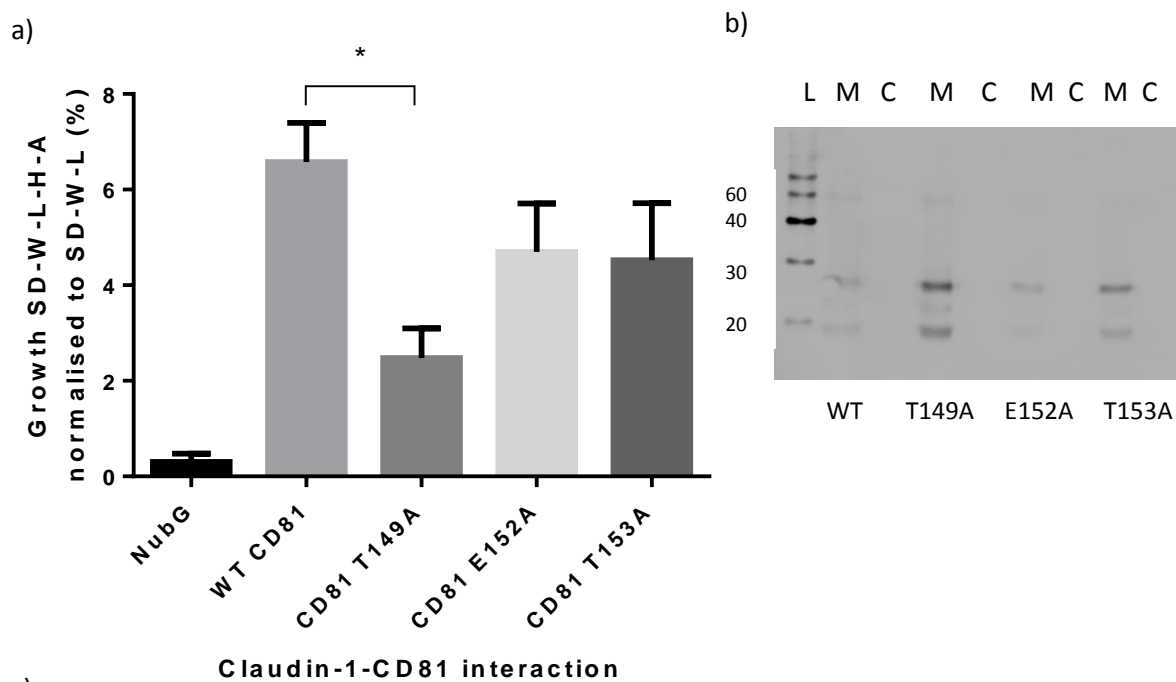
a)

b)

c)

Reaction Bait Claudin-1 +	SD-W-L-H-A colony count normalised to NubI and NubG controls Mean ± SEM (%)	Significance	Growth on SD–W-L-H-A normalised to SD–W-L Mean ± SEM (%)	Significance
NubI	100	-	56.13 ± 10.02	-
NubG	0	ns	0.30 ± 0.18	***
WT CD81	17.39 ± 11.11	-	6.58 ± 0.82	-
CD81T149A	6.25 ± 2.26	ns	2.48 ± 0.62	*
CD81 E152A	10.7 ± 4.39	ns	4.70 ± 1.01	ns
CD81 T153A	10.97 ± 4.85	ns	4.53 ± 1.19	ns

Figure 4.5: CD81 LEL residues involved in Claudin-1 interactions. NMY51 cells were transformed, using lithium acetate (see materials and methods), with appropriate bait and prey vector pairs (WT or mutant). Cells were incubated at 30°C and selected on plates; SD-W-L or SD-W-L-H-A, for 5 days before colonies were counted. a) WT Claudin-1 and CD81 interactions (WT or mutant) represented as growth on SD-W-L-H-A relative to the SD-W-L control. b) Western blot analysis of yeast total membranes expressing prey WT CD81 and mutants (T149A, E152A and T153A). c) Table summarising data normalised in two ways; SD-W-L-H-A colony counts normalised to internal NubI and NubG controls and growth on SD-W-L-H-A normalised to SD-W-L to take into account yeast transformation efficiency. Experiment was repeated four times to compare mean ± SEM of WT to mutant CD81 using a one-way ANOVA followed by a Dunnett's multiple comparison test.

Colonies grown on SD-W-L for each protein-protein interaction were further used to assess if protein interactions can be monitored using growth in liquid culture (SD-W-L-H-A) in a 96 well plate. Others have attempted to grow yeast in this format to generate useful data (Toussaint et al., 2006) but it is not typically used in this context to measure membrane protein-protein interactions. Figure 4.6 shows that yeast growth curves are useful to analyse mutants that may knock-down a protein-protein interaction, since the initial onset of growth (as a consequence of reporter gene activation) is delayed compared to a WT interaction. For mutants CD81 T149A, E152A and T153A the growth curve shifts to the right as compared to WT CD81 (figure 4.6 a, b and c, respectively) due to a decrease in protein-protein interaction frequency at the membrane which reduces activation in reporter gene expression and consequently yeast growth.

Yeast growth curves generated were used to quantitatively compare WT CD81-Claudin-1 interaction and CD81 mutant interactions. At the time point when the OD_{546} of the internal NubI control reached 0.2 (not shown), the OD_{546} reached at this point by WT and mutant CD81 was taken and the mean of three experiments was plotted (figure 4.6d). Again this approach showed the same trend as in figure 4.5. All three CD81 LEL mutants knock-down CD81-Claudin-1 interactions. Even though CD81 T149A, E152A and T153A resulted in similar mean OD_{546} values to each other (0.47, 0.45 and 0.43 respectively) only E152A and T153A were significantly different to WT (mean OD_{546} 0.55). CD81 T149A was not significantly different from WT although a reduction in the mean value is visually apparent (figure 4.6). These data, taken together, confirm that T149, E152 and T153 in CD81 LEL are important for binding to Claudin-1, presumably to Claudin-1 ECL1. Furthermore, it is possible to assess mutant protein interactions using the yeast split-ubiquitin method as it was consistent with results found elsewhere in the literature using FRET with mammalian cells and the same CD81 mutations in the LEL.

Both approaches used here; yeast colony counts and growth in liquid culture, provide useful data as they assess growth at distinct stages (after 5 days for colony counts and 0-23h for liquid growth cultures) and should be used in conjunction with one another to get the most out of the split-ubiquitin method. This work started by using yeast colony counts alone to assess protein interactions but later progressed to use both colony counts and yeast growth curves.
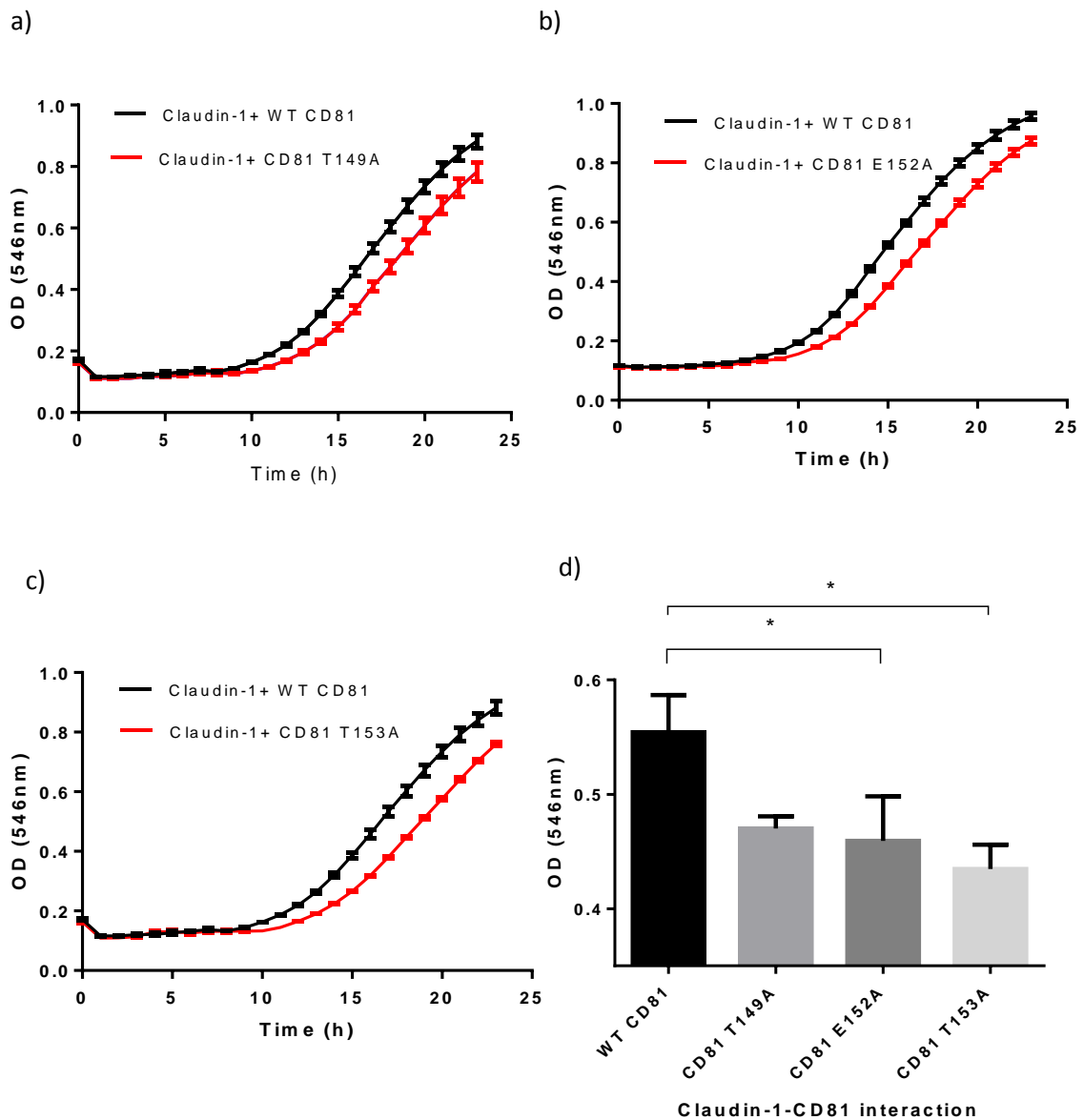
Figure 4.6: Liquid growth curves show CD81 LEL residues T149, E152 and T153 regulate interaction with Claudin-1. *S. cerevisiae* strain NMY51 was transformed with bait and prey vector pairs and were grown on SD-W-L plates. Colonies from a single plate were pooled, vortexed and diluted to an OD_{546} 0.02 before adding to a 96-well plate for growth analysis. OD_{546} measurements were taken every 1h and cells were shaken between readings and incubated at 30°C. a-c) OD_{546} was plotted against time to generate growth curves for WT and mutant CD81 interactions with Claudin-1. d) The mean OD_{546} value of the WT and mutant interactions was compared at the time point where the internal NubI control reached an OD_{546} of 0.2. Experiments repeated 3 times, each experiment performed 4 growth curves in parallel. Data analysed using a one way ANOVA followed by a Fisher's LSD test, n=3, p 0.05.

4.4 Exploring CD81 TM1 residues

4.4.1 Rationale for exploring CD81 TM1

A previous study (Bertaux & Dragic, 2006) explored CD81 transmembrane residues in the context of HCV entry. CD81 residues C80, N18 and E219 when mutated to alanine, reduced HCVpp entry by approximately 40% as compared to control (Bertaux & Dragic, 2006). Binding ability of the CD81 mutants to soluble E2 (sE2) was not affected, suggesting that another mechanism had caused reduced infection. One hypothesis that was not tested at the time was that these transmembrane residues are involved in CD81 binding to a partner protein such as its co-receptor Claudin-1 or another CD81 molecule, which in turn is important for the HCV entry process.

When mapping residues using protein viewer software (Swiss-Pdb Viewer v 4.0.4) on to a CD81 full length homology model (Seigneuret, 2006), it was observed that C80, N18 and E219 combined with LEL residues (T149, E152 and T153) that are known to reduce CD81-Claudin-1 interactions (Davis et al., 2012), together pointed towards CD81 TM1 and TM2 as a potential binding interface for Claudin-1 or another CD81 molecule (see figure 4.7 a-c). Looking from a bird's-eye view, the LEL residues T149, E152 and T153 appear to line up on the same side face as TM2 residue C80 between the space of TM1 and TM2 (figure 4.7b).

Furthermore, a computed hydrogen bond is predicted to exist between N18 (TM1) and E219 (TM4) (figure 4.7c) by Swiss-Pdb Viewer version 4.0.4. Consequently, the double mutant (N18A-E219A) used in Bertaux & Dragic (2006) possibly caused a reduction in HCV entry as the residues may be involved in stabilising the TM domains, particularly CD81 TM1 in this hypothetical scenario. This evidence taken together highlights the possibility that CD81 TM1 is involved in partner protein interactions, which may have a functional effect for HCV entry.

In a separate study that also focused on CD81 TM1 and TM2, Kovalenko *et al*., (2005) provided insight into the dimeric interface of tetraspanin CD9 using a disulphide crosslinking approach of single cysteine mutants. Results suggested TM1 and TM2 are involved in the dimeric interface of CD9 including residues L14, F17 and G80. Bringing together this information, here we performed a homology alignment between

CD9 and CD81 (93% similarity) and selected the corresponding residues in CD81 (L14, F17 and C89). Such residues were then mutated to alanine and tested in CD81 homo- and heterotypic interactions in comparison to WT CD81. Therefore, these residues were selected and tested without using a molecular model to predict the interaction interface of CD81-Claudin-1.

a)

b)

c)

Figure 4.7: CD81 predicted full length model. Full length CD81 homology model accessed using protein data bank (PDB) accession 2AVZ, Seigneuret (2006). a) CD81 predicted full length model mapped with LEL residues T149 (yellow), E152 (purple) and T153 (pink). TM2 residue also highlighted C80 (cyan). The zoomed-in area shows residues in the large extracellular loop (LEL) b) Birds-eye view of residues indicated in part a. c) TM1 and TM4 showing residues N18 (yellow) and E219 (orange), hydrogen bond shown with white dotted line. Pictures produced using Accelrys, Discovery studio 4.0.

4.4.2 CD81 TM1 involvement in Claudin-1-CD81 interactions

a)

b)

Reaction Bait Claudin-1 +	SD-W-L-H-A colony count normalised to NubI and NubG controls Mean ± SEM (%)	Significance	Growth on SD–W-L-H-A normalised to SD–W-L Mean ± SEM (%)	Significance
NubI	100.0	-	72.27 ± 9.91	-
NubG	0.0	**	0.557 ± 0.557	***
WT CD81	17.81 ± 5.33	-	8.41 ± 1.41	-
CD81 L14A	6.39 ± 0.65	ns	3.62 ± 0.53	*
CD81 F17A	7.15 ± 3.53	ns	5.56 ± 1.38	ns
CD81 C89A	2.59 ± 2.42	*	1.82 ± 0.42	**
CD81 L14A-F17A	4.33 ± 1.65	*	2.64 ± 0.72	**
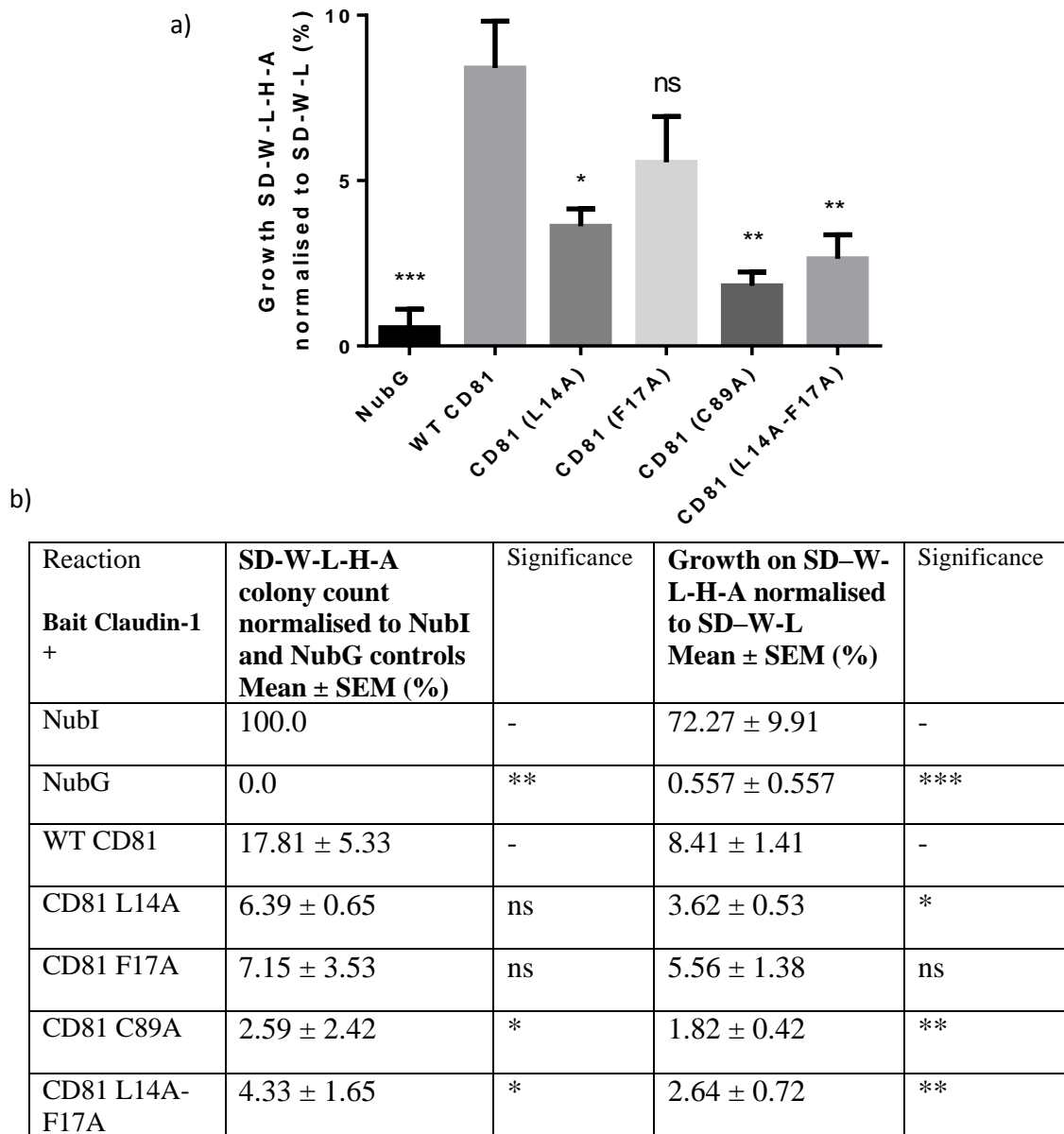
Figure 4.8: CD81 TM1 mutants knock-down Claudin-1 interaction. NMY51 cells were transformed, using lithium acetate (see materials and methods), with appropriate bait and prey vector pairs (WT or mutant). Cells were incubated at 30°C and selected on plates; SD-W-L or SD-W-L-H-A, for 5 days before colonies were counted. a) WT Claudin-1 and CD81 interactions (WT or mutant) represented as growth on SD-W-L-H-A relative to the SD-W-L control. b) Table summarising data normalised in two ways; SD-W-L-H-A colony counts normalised to internal NubI and NubG controls and growth on SD-W-L-H-A normalised to SD-W-L to take into account yeast transformation efficiency. Experiment was repeated three times to compare mean ± SEM of WT to mutant CD81 using a one-way ANOVA followed by a Dunnett's multiple comparison test.

Figure 4.8 demonstrates residues on CD81 TM1 are important in its association with Claudin-1. Residues L14, F17 and C89, when mutated to alanine show a reduction in the level of CD81-Claudin-1 interactions as compared to WT (see figure 4.8a and b). Normalisation of data to the SD-W-L control plate (accounting for transformation efficiency) showed higher levels of significance than when normalising colony counts on SD-W-L-H-A to internal controls (figure 4.8b). When normalising growth on SD-W-L-H-A to SD-W-L, residues L14A, C89A and a double mutant L14A-F17A show a significant reduction in the CD81-Claudin-1 interaction (3.62%, 1.82% and 2.64% respectively) as compared to WT (8.41%). Residue F17A on CD81 TM1 also showed a reduction in yeast growth (5.56%), although this was not statistically different from WT (figure 4.8a and b). Even so, introducing a double mutation in CD81 (L14A-F17A) demonstrated a further knock-down than L14A alone. Therefore, both residues are likely to play a part in the binding interface of CD81-Claudin-1, but L14A is shown to have a larger knock-down effect than F17A when mutated and involved in the interaction.

These data suggest that CD81 residues at the bottom of TM1 are likely to be involved or on the interface of the Claudin-1 interaction binding site. This was explored further using a molecular model of the complex and mutating residues that continue up CD81 TM1 from L14 and F17.

4.4.3 Expression of CD81 TM1 mutants

We next investigated if CD81 TM1 mutants were being expressed in order to determine whether the mutation had decreased the expression thereby decreasing potential CD81-Claudin-1 interactions or whether the interaction was a direct result of the mutation. Western blots were performed on yeast membrane fractions to determine expression and furthermore membrane fractions were tested in an ELISA to detect binding of anti-CD81 antibodies that bind specifically to the folded LEL of CD81.
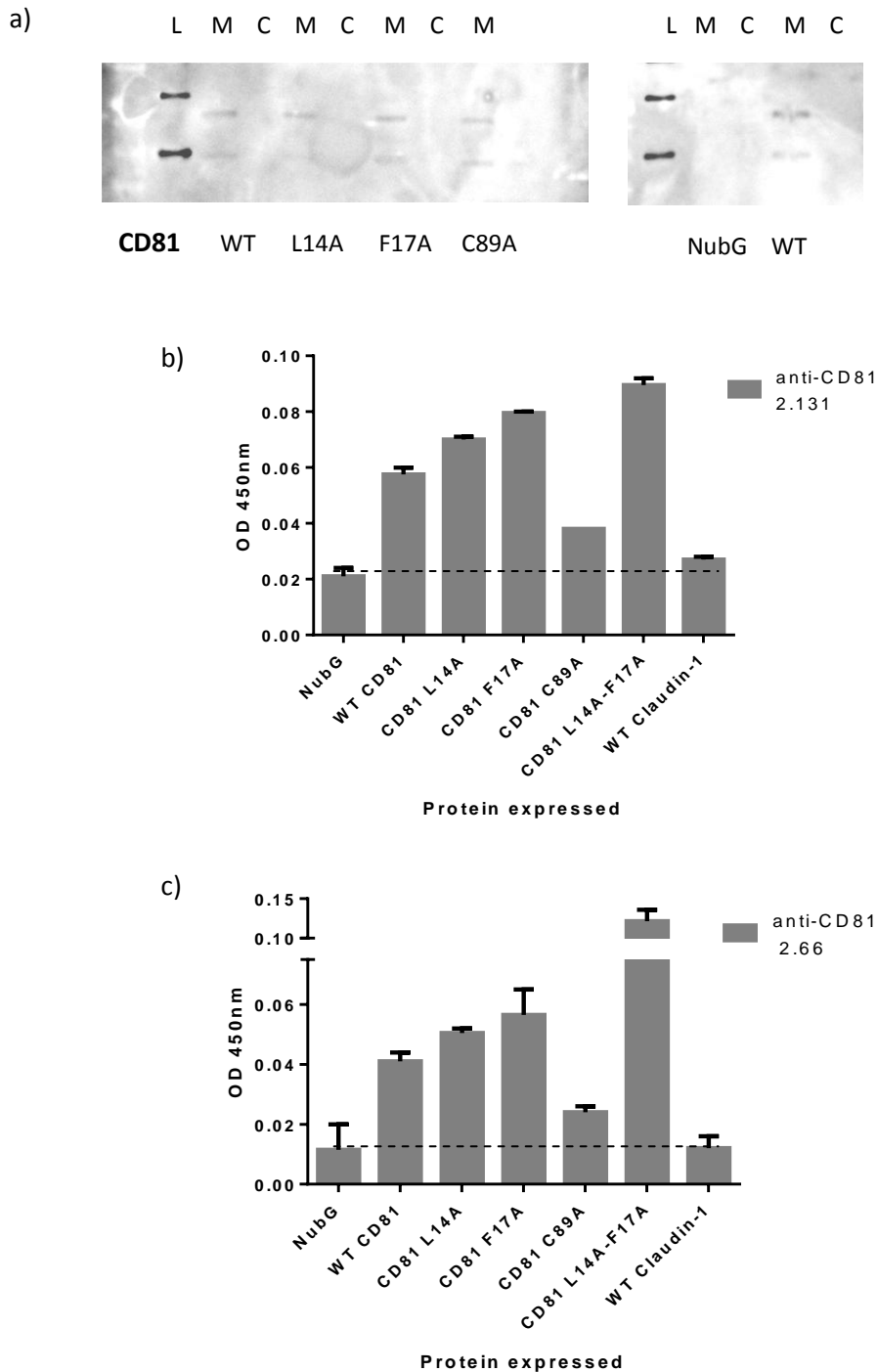
Figure 4.9: WT and mutant CD81 expression. NMY51 cells were transformed (see materials and methods) to express WT or mutant (L14A, F17A, C89A or L14A-F17A) Prey CD81. Cells grown on SD-W plates were selected and grown in SD-W liquid culture overnight. Cells were harvested and then lysed in order to isolate the yeast total membrane fraction using a tissue lyser method. Membranes were resuspended in buffer and quantified for total protein. a) Western blot of yeast membrane (M) and cytosolic fractions (C) expressing WT or mutant Prey CD81 or NubG (negative control). b) ELISA on yeast membranes expressing NubG, WT or mutant CD81 and Claudin-1 using anti-CD81 2.131 and c) anti-CD81 2.66. Dashed line indicates negative control NubG background signal. Data represents the mean ± SEM of two independent experiments.

Expression data for CD81 TM1 mutants (L14A, F17A, C89A and L14A-F17A) suggest that expression is similar to that found for WT CD81 (see figure 4.9). The same amount of total protein was loaded for western blot analysis of CD81 yeast membranes; figure 4.9a shows that the protein band intensity is similar between WT and mutant CD81, suggesting comparable expression. Also, both WT and mutant CD81 show both monomer and dimer protein bands. The negative control, cells expressing NubG, does not show any equivalent banding pattern (figure 4.9a).

To further assess the expression of CD81, equivalent levels of total protein of WT and mutant CD81 membranes (as determined by a protein BCA assay) were coated on to ELISA plates before probing with anti-CD81 antibodies 2.131 or 2.66. Both antibodies target the CD81 LEL and do not bind if the structure is unfolded. Negative control samples (NubG and Claudin-1) provided the background signal of the ELISA (the average background indicated with a dashed line) (figure 4.9b). Anything above the dashed line indicates higher than background for the ELISA and reports positive binding to both anti-CD81 antibodies tested.

Figures 4.9 b and c show ELISA data for different CD81 samples and suggest both antibodies bind slightly increased levels of mutant CD81 (L14A, F17A and L14A-F17A) in yeast membranes as compared to WT CD81. Furthermore, double mutant CD81 L14A-F17A shows the highest increase in binding antibodies 2.131 and 2.66 as compared to WT CD81 (figure 4.9b and c, respectively). The double mutation may have an effect on the accessibility of both antibodies to the site in which they bind to the CD81 folded loop. One theory may be that the double mutant decreases tetraspanin oligomerisation and therefore may reveal more antibody binding sites in the LEL.

CD81 C89A in contrast shows a decrease as compared to WT CD81, suggesting that the structure of CD81 is compromised when residue C89 is mutated to alanine. The mutation may result in an unstructured LEL or more specifically a change in loop conformation in the region where anti-CD81 2.131 and 2.66 bind to WT.

Overall, the expression data show that L14A and F17A CD81 mutants are expressed and can bind similar levels of antibody as WT CD81 and so their knock-down of CD81-Claudin-1 interactions (figure 4.8) suggest that the residues are involved in the protein binding interface. L14A-F17A supports this idea and the increased level of ELISA antibody binding may suggest a decrease in CD81-CD81 interactions. Mutation of

residue C89 shows a decrease in CD81-Claudin-1 interactions (figure 4.8) but also a reduction in binding anti-CD81 antibodies as compared to WT CD81. Therefore, the reduction in the Claudin-1- CD81 C89A interaction (figure 4.8) may be a result of an unfolded CD81 LEL that cannot bind anti-CD81 antibodies to WT level (figure 4.9 b and c) or Claudin-1 (figure 4.8).

Plasma membrane expression data of WT and mutant CD81 would have been useful evidence to further interpret mutant knock down results seen in figure 4.8. Even so, evidence shown in this chapter to demonstrate that the CD81-Claudin-1 interaction can be monitored using the DUALmembrane system shows that the Nub and Cub moieties are facing the cytosolic side of the plasma membrane in order to be cleaved by UBPs, in contrast to residing in the ER or Golgi membrane (hence Nub and Cub would not be in the cytoplasm) where the Cub fusion would not be cleaved by UBPs, translocate to the nucleus and activate reporter genes.

4.5 CD81 LEL and TM1 residues involved in homo-oligomerisation

Residues in CD81 LEL T149, E152 and T153 and in TM1 L14, F17 and C89 were tested in the CD81 homotypic interaction to determine if there was a common binding interface with CD81-Claudin-1 heterotypic interactions.

Figure 4.10 shows how CD81 mutations affect homo-oligomerisation. The data suggests that CD81 loop residue, T153, when mutated to alanine significantly reduces CD81-CD81 interactions as shown by reduced yeast growth and reporter gene activation (figure 4.10a). This result is apparent when growth data is normalised to NubI and NubG internal controls (figure 4.10a) where WT CD81 shows 30.5% growth as compared to CD81 T153A shows 5.9% growth in comparison. CD81 T149A and E152A also show reduced growth compared to WT (24.2% and 18.5%, respectively) but are not statistically different. The same trend in the data is shown when growth is normalised to the SD-W-L control (figure 4.10a). This data suggests that T153A is involved in CD81-CD81 interaction and T149A and E152A may have less of an involvement. T149A and T153A are known to be involved in CD81 dimerisation (Martino and Guido, 2002) and so the knock-down seen with the same mutants in the

CD81-Claudin-1 interaction may result from indirect effects on CD81-CD81 interactions.

When CD81 TM1 mutants L14A, F17A, C89A and L14A-F17A were tested in CD81 homo-oligomerisation, L14A and L14-F17A mutants exhibited significantly lower growth than WT CD81 when data was normalised to internal controls (15.9%, 12.6% and 30.6% respectively) (figure 4.10b). This may suggest that the residues are also involved in CD81-CD81 interactions as well as CD81-Claudin-1 interactions seen earlier in this chapter. Again, an indirect effect of reducing CD81-CD81 interactions may be the case on the frequency of CD81-Claudin-1 interactions. Mutants F17A and C89A do not have a significant effect on yeast growth in CD81-CD81 interactions, if only showing a small knock-down as compared to WT, but they did not reach statistical significance (figure 4.10b).

This data suggests that CD81-Claudin-1 oligomerisation may be directly linked to the status of CD81-CD81 oligomerisation. When CD81-CD81 interactions are reduced this may cause a reduction in the occurrence of CD81-Claudin-1 interactions. Previous studies suggest that two CD81 molecules bind one Claudin-1 molecule (Bonander et al., 2013) so potentially a CD81 dimer is required to interact with Claudin-1. Alternatively CD81 may be a monomer, dimer or higher order oligomer and binds Claudin-1 on the same side 'face' as it would another CD81 molecule, suggesting a common binding interface of CD81 and Claudin-1.
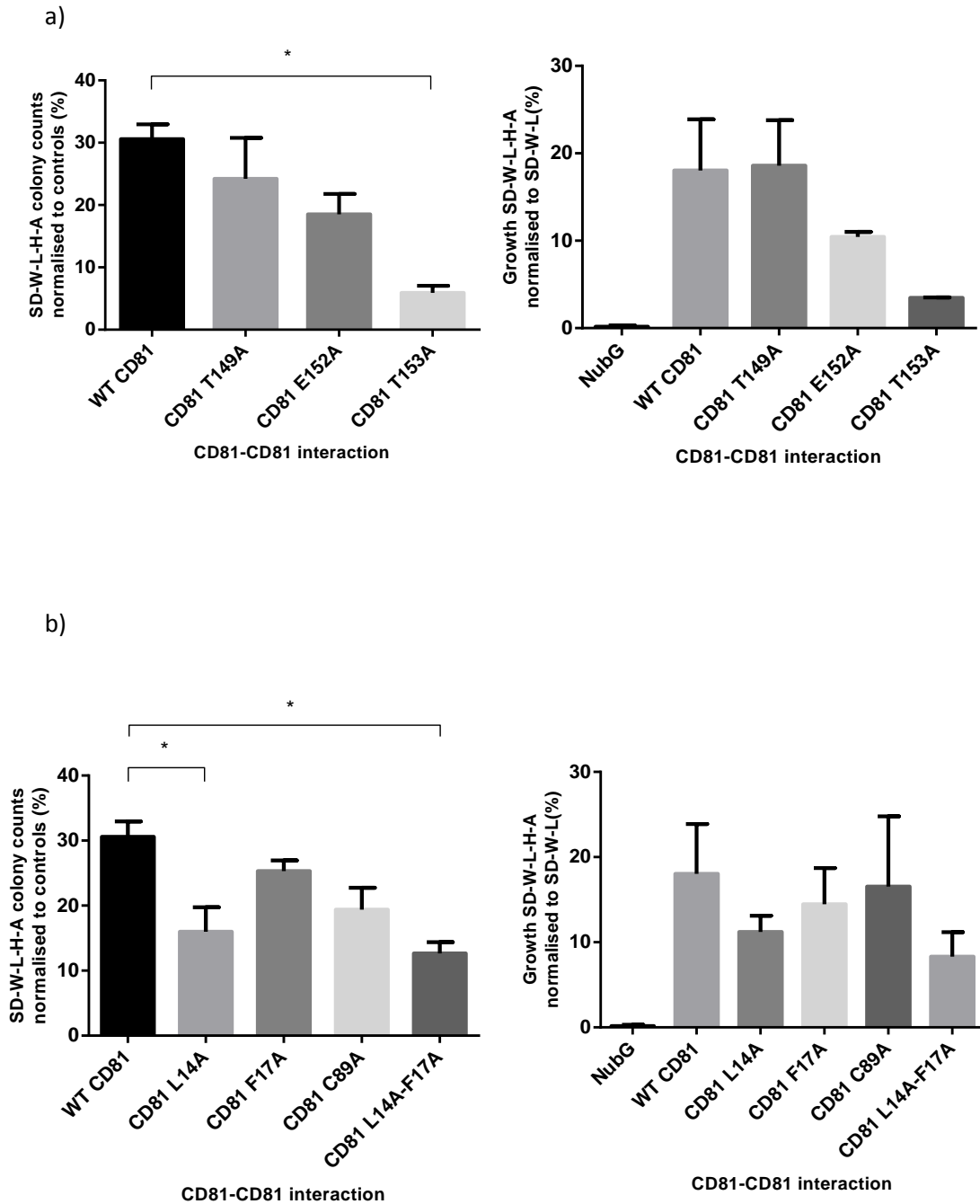
Figure 4.10: CD81-CD81 interactions. NMY51 cells were transformed with appropriate CD81 bait and prey vector pairs (WT or mutant). Cells were incubated at 30°C and selected on plates; SD-W-L or SD-W-L-H-A, for 5 days before colonies were counted. a) CD81-CD81 interactions (WT or loop residue mutant) represented as growth on SD-W-L-H-A normalised to internal controls NubI and NubG and CD81-CD81 interactions normalised using SD-W-L control. b) CD81-CD81 interactions (WT or TM1 residue mutant) represented as growth on SD-W-L-H-A normalised to internal controls NubI and NubG and CD81-CD81 interactions normalised using SD-W-L control. Experiment was repeated twice to compare mean ± SEM of WT to mutant CD81 using a one-way ANOVA followed by a Dunnett's multiple comparison test.

4.6 Using a molecular modelling approach to predict further CD81 TM1 residues involved in the CD81-Claudin-1 binding interface

In collaboration with Jonathan Mullins (Swansea University), we produced a model of the full length CD81-Claudin-1 complex using the Seigneuret (2006) CD81 homology model (shown in figure 4.11) docked with a threading based model of Claudin-1, generated by Dr Mullins, using Hex 8.0 (a protein docking program). We explored CD81 TM1 further using the model complex which allowed us to choose appropriate residues along the TM1 interface to perform site-directed mutagenesis and evaluate in the yeast split-ubiquitin system.

4.6.1 Molecular model complex of CD81-Claudin-1 using a Claudin-1 threading based model

Figure 4.11 shows the CD81-Claudin-1 structural model complex, using the Claudin-1 threading-based model, which was used as a guide to direct mutagenesis experiments. Residues L14 and F17 on CD81 TM1 appeared to form part of a hydrophobic ridge that could possibly fit into a groove present on Claudin-1 TM2 (residues G87 and G91). Using this information, we next selected residues on CD81 TM1 from L14 and F17 that faced the same direction to probe this face for its interaction with Claudin-1. Residues were F21, G25, I28, L29 and L35. On Claudin-1 TM2 G87 and G91 also appeared to be good mutagenesis targets. The initial plan was to mutate all chosen residues to alanine and assess the mutants in protein-protein interaction experiments using the yeast split-ubiquitin assay.

Aston University

Illustration removed for copyright restrictions

Figure 4.11: CD81-Claudin-1 complex model. A CD81-Claudin-1 homology based model using CD81 full length homology model by Seigneuret (2006) with a Claudin-1 threading-based model generated and docked using Hex 8.0 by Dr Jonathan Mullins (Swansea University). CD81 TM1 (yellow) shows residues L14, F17, F21, G25, I28 and L35 pointing outwards as a possible interaction interface with Claudin-1 TM2 (residues predicted to be involved are G87 and G91 in orange).

4.6.2 Testing CD81 TM residues predicted to be involved in the Claudin-1 interaction

Residues on CD81 TM1 (F21, G25, I28, L29 and L35) predicted to interact with Claudin-1 using a full length model were tested in the yeast split-ubiquitin assay.

Figure 4.12 shows data generated using yeast colony counts on plates selecting for a protein-protein interaction (SD-W-L-H-A). A similar data trend is shown when growth on SD-W-L-H-A is normalised to internal controls or to the SD-W-L control plate (see figure 4.12a-c). Using data that was normalised to SD-W-L, CD81 I28A and L35A resulted in a reduction in yeast growth (3.01% and 3.82% respectively) in comparison to WT CD81 (7.88%) (figure 4.12b and c), even though the apparent reduction was not statistically significant. This would suggest that introducing mutations at I28 and L35 have an effect on the CD81-Claudin-1 interaction, possibly due to disrupting the protein binding interface on CD81 TM1. CD81 mutants F21A, G25A and L29A showed similar levels of growth on SD-W-L-H-A as WT CD81 (figure 4.12) suggesting that these CD81 mutants can bind equally well to Claudin-1 as WT CD81.
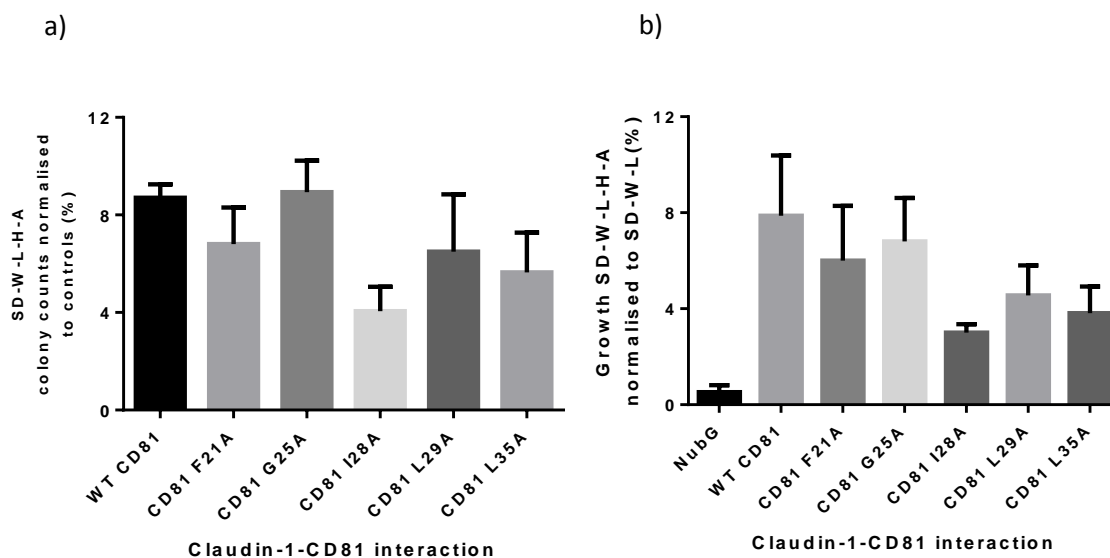
a)   b)

Reaction Bait Claudin-1 +	SD-W-L-H-A colony count normalised to NubI and NubG controls Mean ± SEM (%)	Significance	Growth on SD–W-L-H-A normalised to SD–W-L Mean ± SEM (%)	Significance
NubI	100	-	104.8 ± 26.17	-
NubG	0	**	0.53 ± 0.29	*
WT CD81	8.71 ± 0.54	-	7.88 ± 2.50	-
CD81 F21A	6.81 ± 1.50	ns	6.00 ± 2.29	ns
CD81 G25A	8.94 ± 1.29	ns	6.80 ± 1.81	ns
CD81 I28A	4.05 ± 1.00	ns	3.01 ± 0.35	ns
CD81 L29A	6.49 ± 2.35	ns	4.56 ± 1.25	ns
CD81 L35A	5.65 ± 1.63	ns	3.82 ± 1.11	ns

Figure 4.12: Predicted CD81 TM1 residue mutants in Claudin-1-CD81 interactions. NMY51 cells were transformed with appropriate bait and prey vector pairs (WT or mutant). Cells were incubated at 30°C and selected on plates; SD-W-L or SD-W-L-H-A, for 5 days before colonies were counted. a) WT Claudin-1 and CD81 interactions (WT or mutant) represented as growth on SD-W-L-H-A internal controls NubI and NubG. b) WT Claudin-1-CD81 interactions normalised using SD-W-L control. c) Table summarising data normalised in two ways. Experiment was repeated five times to compare mean ± SEM of WT to mutant CD81 using a one-way ANOVA followed by a Dunnett's multiple comparison test.

In order to determine the effects of predicted CD81 TM1 residues further in the CD81-Claudin-1 interaction, yeast liquid growth curve analysis was performed. Colonies grown on SD-W-L plates selected for double vector transformations were pooled from each test condition. Yeast was diluted to an OD_{546} 0.02 and incubated for approximately 20h whilst measuring the OD_{546} reached by the sample every hour.

Data in figure 4.13 show that the slowest growth, hence a reduction in CD81-Claudin-1 interaction was due to L14A-F17A, I28A and L35A CD81 mutants. This was demonstrated in representative growth curves by a shift to the right of the mutant growth profiles as compared to WT (figure 4.13b-g). Other mutants tested, F21A, G25A and L29A did not result in different growth profiles as compared to WT CD81. Furthermore, figure 4.13a compares WT and mutant CD81 average OD_{546} at a specific point in the curve for the CD81-Claudin-1 interaction. CD81 L14A-F17A and L35A had significantly different OD_{546} values in comparison to WT (0.47, 0.50 and 0.62 respectively) and I28A demonstrated a decrease (OD_{546} 0.53) although not significantly different from WT CD81. F21A, G25A and L29 caused a slight growth reduction as compared to WT but this was not significantly different.

This suggests that mutants that reduce CD81-Claudin-1 interactions as determined by yeast growth are L14A-F17A, I28A and L35A. Possibly due to being involved in the CD81-Claudin-1 binding interface. Although predicted to be facing the same direction and on the interaction interface, residues F21A, G25A and L29A did not have such pronounced effects on yeast growth, hence are likely to not disrupt the CD81-Claudin-1 interaction when mutated to alanine. Substitution of another amino acid other than alanine may have shown a different result.

This approach suggests that subtle differences in the interaction of CD81 TM mutants with Claudin-1 can be shown in liquid growth cultures, even on a small scale in a 96 well format.

e)

f)

g)
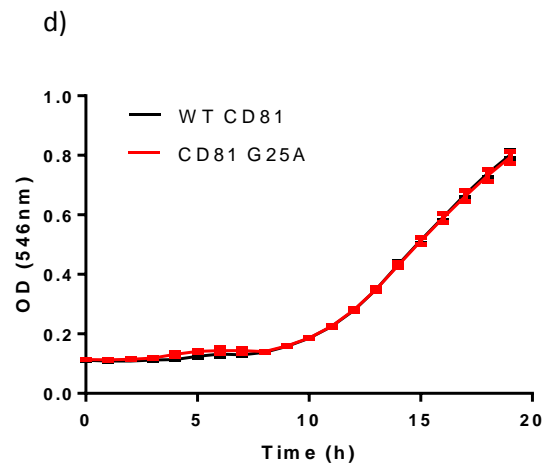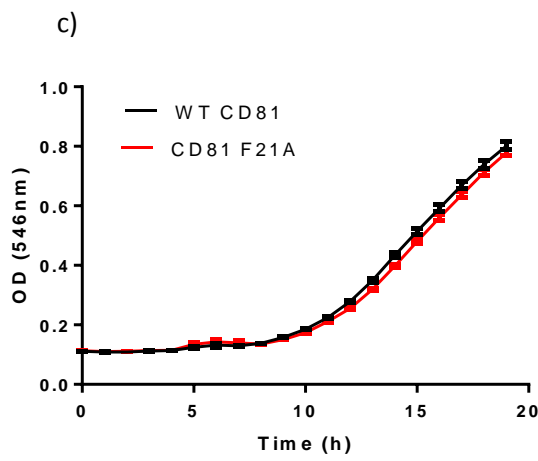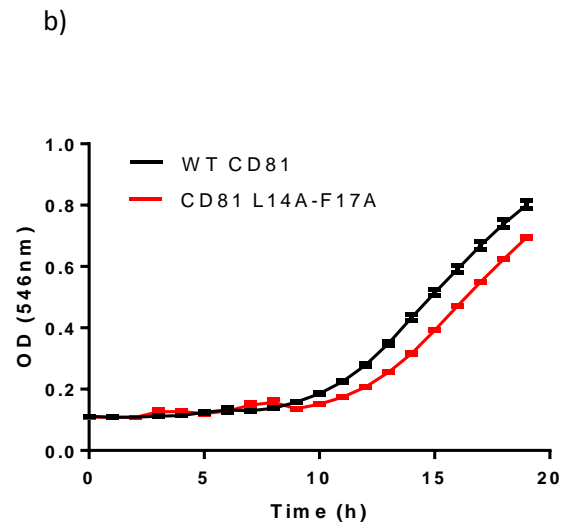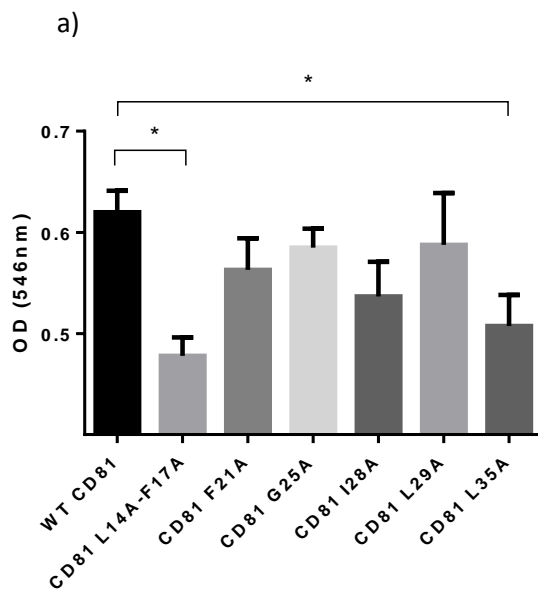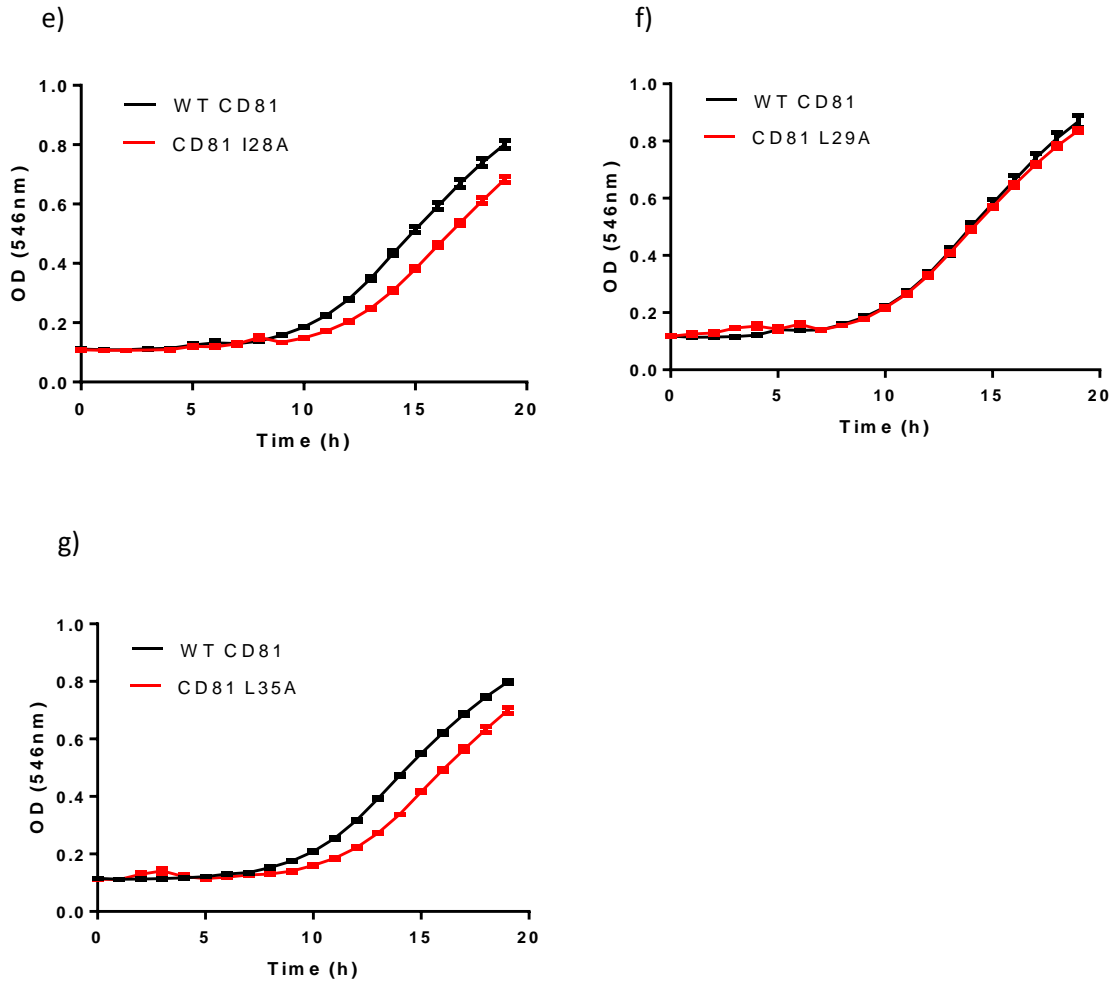
Figure 4.13: CD81 TM1 residues involved in the CD81-Claudin-1 interaction. NMY51 was transformed with bait and prey vector pairs and were grown on SD-W-L plates. Colonies from a single plate were pooled, vortexed and diluted to an OD_{546} 0.02 before adding to a 96-well plate for growth analysis. OD_{546} measurements were taken every 1h and cells were shaken and incubated at 30°C. a) The mean OD of the WT and mutant interactions were compared at the time point where internal NubI control reached OD 0.2. b-g) OD_{546} was plotted against time(h) for WT and mutant CD81 interactions with Claudin-1. Experiments repeated 4 times, each experiment performed 4 growth curves in parallel. Data analysed using a one way ANOVA followed by a Fisher's LSD test. n=4, p 0.05.

4.6.3 CD81 TM1 mutant expression

In order to determine expression of the predicted CD81 TM1 mutants in yeast as compared to WT CD81, western blot analysis and ELISA were performed using specific anti-CD81 antibodies. Western blot analysis on CD81 WT or mutant membrane and cytosolic fractions (figure 4.14a) shown that mutant protein bands were of the same density as WT as seen by monomer and dimer molecular weight sizes. This was except CD81 I28A, which showed evidence of reduced protein band intensity which would point towards reduced expression as compared to WT CD81.

Further experiments were performed using an ELISA with anti-CD81 antibodies. The negative controls (NubG and PBS) were used to determine background signal with anything above the dotted line showing positive binding of antibodies to CD81 samples (figure 4.14b and c). Anti-CD81 2.131 and 2.66 both bound relatively equal levels of WT CD81 as mutant CD81 in yeast membranes (figure 14b and c). This suggests that firstly the expression level in yeast membranes is similar between all proteins and secondly, the ability to bind CD81 folded LEL specific antibodies is consistent between WT and mutants. These data provide evidence that the reduced interaction between CD81-Claudin-1 by mutants I28A and L35A (figure 4.13) is possibly due to their involvement in the protein binding interface and not due to reduced expression or protein mis-folding as compared to WT CD81. Evidence shown here to suggest folded mutant CD81 demonstrates that the mutants are likely not mis-folded and retained in the ER or Golgi, hence similar plasma membrane expression as WT CD81.
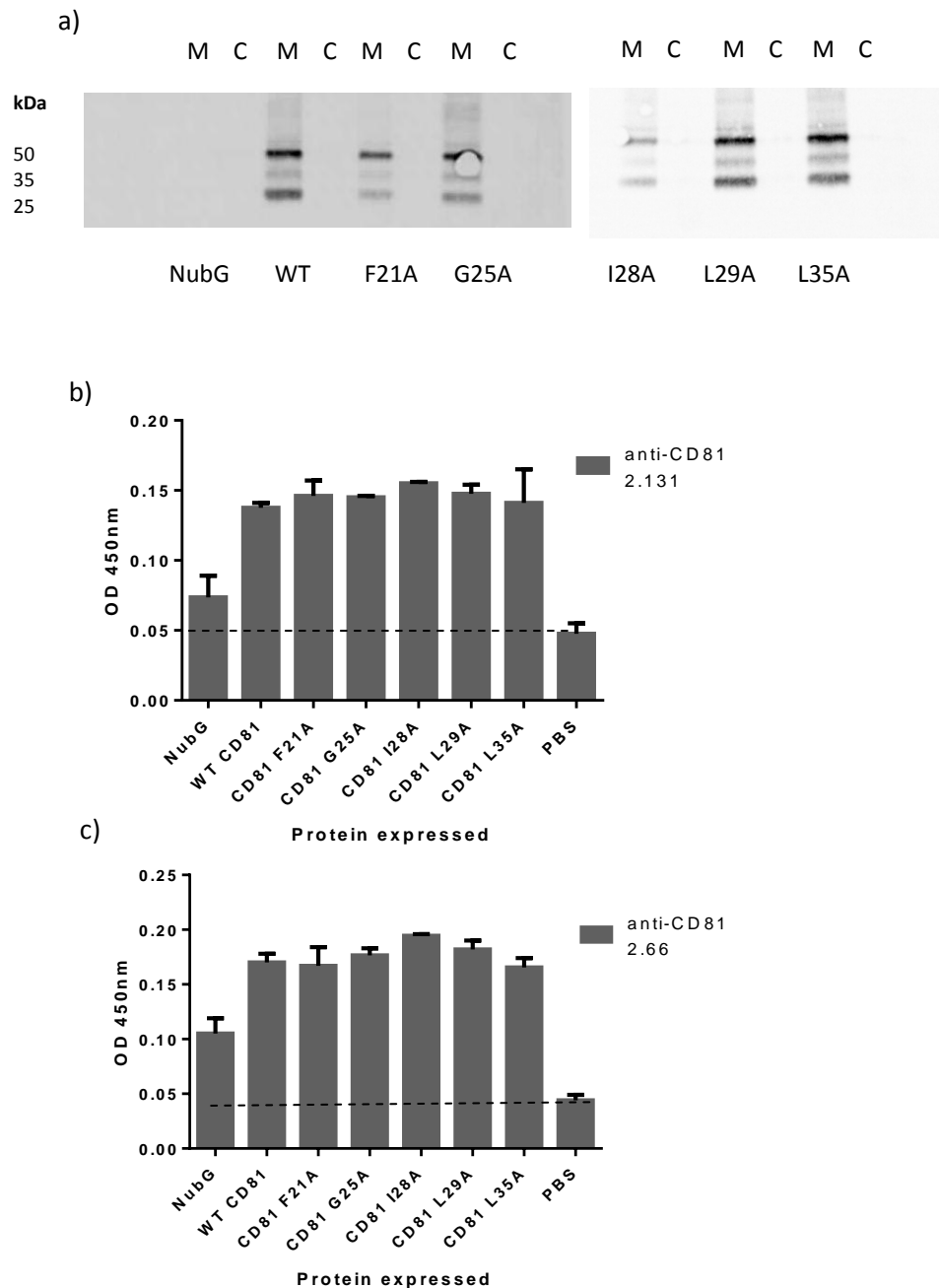
Figure 4.14: CD81 TM1 mutant expression. NMY51 cells were transformed (see materials and methods) to express WT Prey CD81 or mutant (F21A, G25A, I28A, L29A or L35A). Cells grown on SD-Trp plates were selected and grown in SD-Trp liquid culture overnight. Cells were harvested and then lysed in order to isolate the yeast membrane fraction using a tissue lyser method. Membranes were resuspended in buffer and quantified for total protein. a) Western blot of yeast membrane (M) and cytosolic fractions (C) expressing WT or mutant Prey CD81 or pPR3N (negative control). b) ELISA on yeast membranes expressing pPR3N, WT or mutant CD81 using anti-CD81 2.131 and c) anti-CD81 2.66 monoclonal antibodies. Dashed line indicates negative control NubG background signal. Data represents the mean ± SEM of two independent experiments.

4.6.4 Molecular model complex of CD81-Claudin-1 using a Claudin-15 based homology model

A major milestone for the Claudin family, the structure of mammalian Claudin-15 solved to 2.4 Angstroms (Suzuki et al., 2014) allowed us to improve our CD81-Claudin-1 complex model. Since the threading based model of Claudin-1 was based on alginate lyase from Agrobacterium tumefaciens C58 as template, the revised model based on Claudin-15 (a close homologue of Claudin-1 and therefore a good template) would likely be more reliable.

a)   b)

Aston University

Illustration removed for copyright restrictions

Aston University

Illustration removed for copyright restrictions
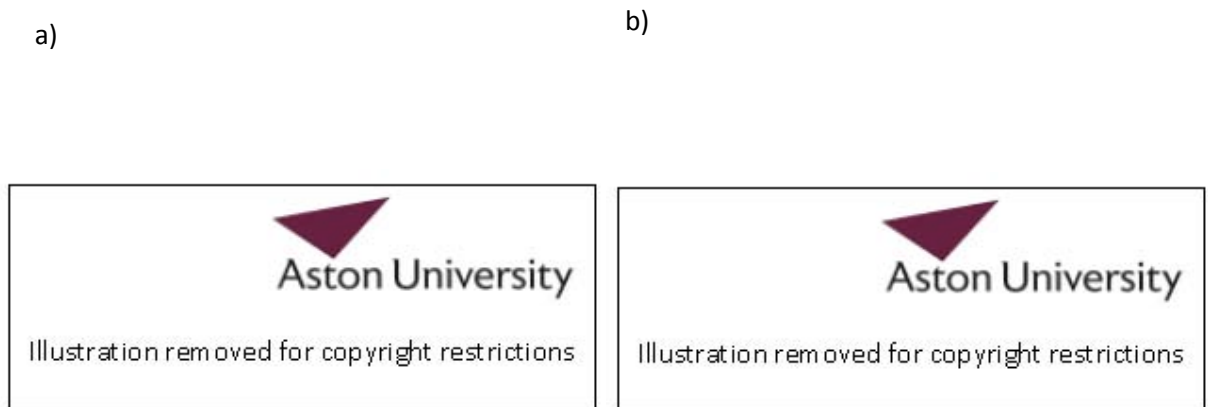
Figure 4.15: Two Claudin-1 full length models used in this study. a) Cytoplasmic view of full length Claudin-1 threading based model showing TM domains (TM1-4) in red going in a clockwise direction. b) Cytoplasmic view of full length Claudin-1 homology based model. TM1-TM4 (red) is shown to be going in an anti-clockwise direction. Models produced by Dr Jonathan Mullins. Pictures using PDB files were generated using Accelrys Discovery Studio 4.0 Visualizer.

Figure 4.15 shows how the two Claudin-1 full length models differ regarding TM bundle direction as viewed from the cytoplasm. The reasons for the difference depend on the method and the template that was used to generate each model i.e. threading template alginate lyase from Agrobacterium tumefaciens C58 and the homology model used the Claudin-15 structure as template. The latter is a much more reliable template

than the former due to increased homology with Claudin-1, but was not available until 2014. This shows the development of molecular modelling over time.

Figure 4.16: Claudin-15 homology based model used to generate a CD81-Claudin-1 complex model. CD81-Claudin-1 complex docked together using Hex 8.0. The Claudin-1 homology based model (blue) was used with CD81 full length homology model by Seigneuret (2006) (red). Residues on CD81 TM1 L14, F17, I28 and L35 are highlighted in green and predicted to have an effect on the CD81-Claudin-1 interface via Claudin-1 TM2 (V85 and V92) and the intracellular loop (M105). Produced by Dr Jonathan Mullins.

Figure 4.16 shows the evolved CD81-Claudin-1 complex model using a homology approach. Using this model we were able to plot residues in CD81 TM1 that were mutated and tested earlier in the yeast split-ubiquitin assay and subsequently predict residues in Claudin-1 that may be on the interaction interface. The predictions on Claudin-1 included V85, V92 on Claudin-1 TM2 and M105 on Claudin-1 intracellular

loop (see figure 4.16). Further experiments are required to test if the Claudin-1 predictions are correct.

Using the revised CD81-Claudin-1 homology model, FoldX, which is a computer algorithm that can predict the effect of an amino acid mutation on the stability of proteins and protein complexes, was used to predict the effect of CD81 TM1 mutations on protein stability and on the stability of CD81-Claudin-1 interactions. Each CD81 TM1 residue that was mutated to alanine and explored in yeast split-ubiquitin assays were analysed (see table 4.1).

Table 4.1: FoldX analysis of CD81 TM1 mutants on protein stability and stability of CD81-Claudin-1 interactions. The CD81-Claudin-1 homology based model generated by Dr Jonathan Mullins was used in FoldX. Each CD81 TM residue was mutated to alanine and scored (Gibbs free energy (ΔG)) for protein stability and protein interaction stability. If a mutation increases protein stability, $\Delta G <0$, if it is less stable $\Delta G >0$.

CD81 mutant	Average protein stability $\Delta\Delta G$ (Kcal/mol)	Average interaction energy $\Delta\Delta G$ (Kcal/mol)
L14A	2.02	0.15
F17A	0.79	0.69
C89A	0.29	0.00
F21A	2.01	0.80
G25A	1.82	0.00
I28A	3.22	0.99
L29A	2.58	0.01
L35A	1.56	-0.28

FoldX analysis was useful to analyse the CD81-Claudin-1 homology model (figure 4.16) and help make a judgement of the analysis in context with experimental data of CD81 mutants in the split-ubiquitin yeast assay. Mutants that caused instability to the CD81-Claudin-1 complex as predicted using the model were L14A, F17A, F21A and I28A. Results from the split-ubiquitin assay are consistent with L14A, F17A and I28A, but not F21A, which is in the middle of CD81 TM1 and so may be involved in non-specific hydrophobic interactions. L35A resulted in a reduction in CD81-Claudin-1 interactions in yeast but was not shown to in FoldX analysis (table 4.1), even though the molecular model does show a close association between L35 on CD81 and V85 on Claudin-1 (figure 4.16). Taken together, results from yeast assays and FoldX analysis

using a molecular model suggest that residues at the bottom (L14 and F17) and at the top (I28A and L35A) of CD81 may be involved in the interaction interface with Claudin-1 TM2 (see figure 4.16).

Most of the mutations, according to FoldX analyses caused protein instability since ΔG scores >0 were regarded as less stable. The mutation which was suggested to cause the highest level of protein instability was I28A and so this may suggest that protein instability is the cause of any reduction seen in CD81-Claudin-1 interactions. Although, ELISA experimental data did not find a difference in expression of CD81 I28A compared to WT CD81 (figure 4.14b) but western blot analysis did find reduced protein expression as compared to WT. Further work is required to clarify the effect of CD81 I28A on protein stability (figure 4.14a).

5. Improving the thermal stability of recombinant, full length CD81 for structural and functional characterisation

5.1 Chapter 5 overview and objectives

Recombinant protein production is a major route investigators take to produce protein samples in order to characterise the structural and functional elements of prokaryotic and eukaryotic membrane proteins. In contrast to naturally abundant membrane proteins such as the G-protein coupled receptor, rhodopsin (Stenkamp et al., 2002), the majority are found in relatively low quantities naturally and so prokaryotic or eukaryotic host systems are chosen to specifically overproduce a protein of interest for further study, which is likely to require high protein concentrations.

This chapter focuses on the yeast recombinant host, *P. pastoris*, as it has many advantages for over-expression of human proteins. Firstly, *P. pastoris* are unicellular organisms able to grow to high cell densities at fast rates, they are relatively inexpensive to culture, and the yeast also offers its eukaryote features for protein production such as post-translational modifications as well as a secretory pathway (Mattanovich et al., 2012). Moreover, strong inducible promoters are available such as the alcohol oxidase I (*AOX1*) promoter, utilised in this chapter, which enables protein expression induced in the presence of methanol (see Ahmad *et al*., 2014 for review).

P. pastoris has previously shown success in producing eukaryotic membrane proteins that were subsequently structurally solved (Carpenter et al., 2008). Once overexpressed, membrane proteins must be solubilised from the yeast membrane, most commonly using detergents but this in itself introduces problems due to the hydrophobic nature of membrane proteins. Detergents are often screened for their ability to yield soluble, stable and homogenous membrane protein from the host membrane. It has been shown using various approaches that optimising conditions such as buffer, detergent or the addition of lipids, which may render a protein more stable, can result in improved crystallisation efforts (Carpenter et al., 2008).

One relatively recent approach, used to optimise and guide membrane protein purification conditions, measures thermal stability using the thiol-specific fluorochrome N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) (Alexandrov et

al., 2008). The method relies on native cysteines embedded within a folded protein which bind CPM if exposed due to temperature-induced protein unfolding, which results in a measureable fluorescent signal.

This chapter describes the expression, solubilisation, purification and optimisation of the thermal stability of human CD81 and Claudin-1. Our aim was to focus on CD81 in particular, in order to aspire towards the first full length tetraspanin solved structure. Specific objectives included:

- Expression of full length human CD81 and Claudin-1 in *P. pastoris*;
- Solubilisation and purification of CD81 and Claudin-1;
- Thermal stability assessment of CD81 to optimise solution conditions such as buffer composition, pH, ionic strength, glycerol content and lipid addition;
- CD81 crystallisation trials.

5.2 CD81 and Claudin-1 expression, solubilisation and antigenicity

Recombinant, full length human CD81, mutated at palmitoylation sites (from Cys to Ala) was previously produced in mg quantities using *P. pastoris* as the eukaryotic host (Jamshad et al., 2008). Similar quantity and quality of human wild-type Claudin-1 was also produced using the same system (Bonander et al., 2013) (see figure 5.1).

a)

b)
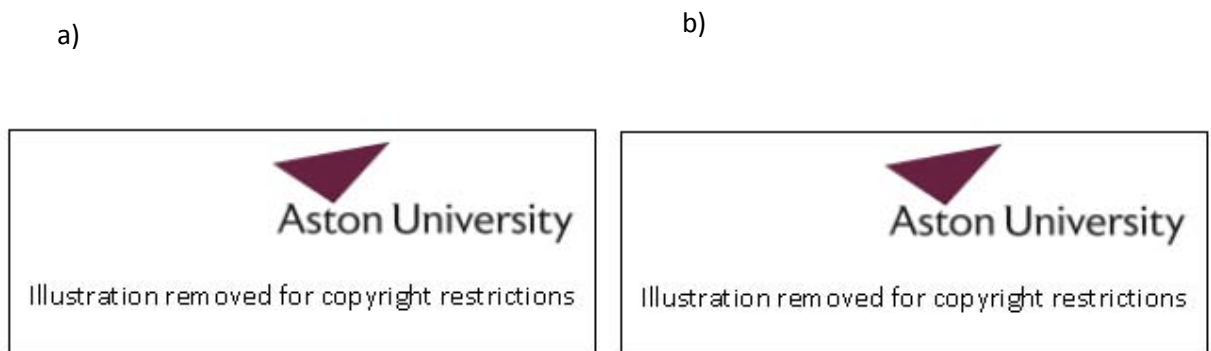
Figure 5.1: Membrane protein constructs of CD81 and Claudin-1 used in this study. Recombinant full length human CD81 (Jamshad et al., 2008) and Claudin-1 (Bonander et al., 2013) were previously produced in mg quantities. *P. pastoris* X33 cells were used to produce four transmembrane proteins under tight regulation using the *AOX1* promoter in the pPICZB vector (Invitrogen). a) Full length human CD81 was mutated at palmitoylation sites (Cys to Ala) and a His_6 tag was incorporated on the C-terminal tail for detection and purification purposes. b) Full length human Claudin-1 used was wild-type that has a His_6 tag at the C-terminal tail.

P. pastoris colonies, selected on YPD-zeocin plates were initially tested for their expression profile for CD81 and Claudin-1 in shake flasks over time using western blot analysis and two different approaches. The first approach determined if the *AOX1* promoter was tightly regulated by methanol induction using CD81 expression as an example. Yeast cells were grown for approximately 22h in BMGY to accumulate biomass prior to induction. Cells were then diluted to an OD_{600} 5 and grown for 47h in BMMY in order to induce CD81 expression. Samples were collected pre- and post-induction (see figure 5.2a) and the total membrane was isolated. The same amount of

total protein (25µg) for each sample was then loaded on to an SDS-gel for western blot analysis and CD81 detection.

The second approach focused on post-induction expression of both CD81 and Claudin-1 over a time course. After cell growth in BMGY, cells were induced in BMMY and 3ml samples were collected between 0-50 h induction. Samples were analysed for protein expression by isolating the total membrane, the total protein concentration was this time not quantified, but equal volumes of membrane preparation (20µl) was loaded per sample on to an SDS gel to determine the volumetric yield of membrane protein. An anti-His antibody was used to detect CD81 and Claudin-1 using a western blot approach (see figure 5.2b).
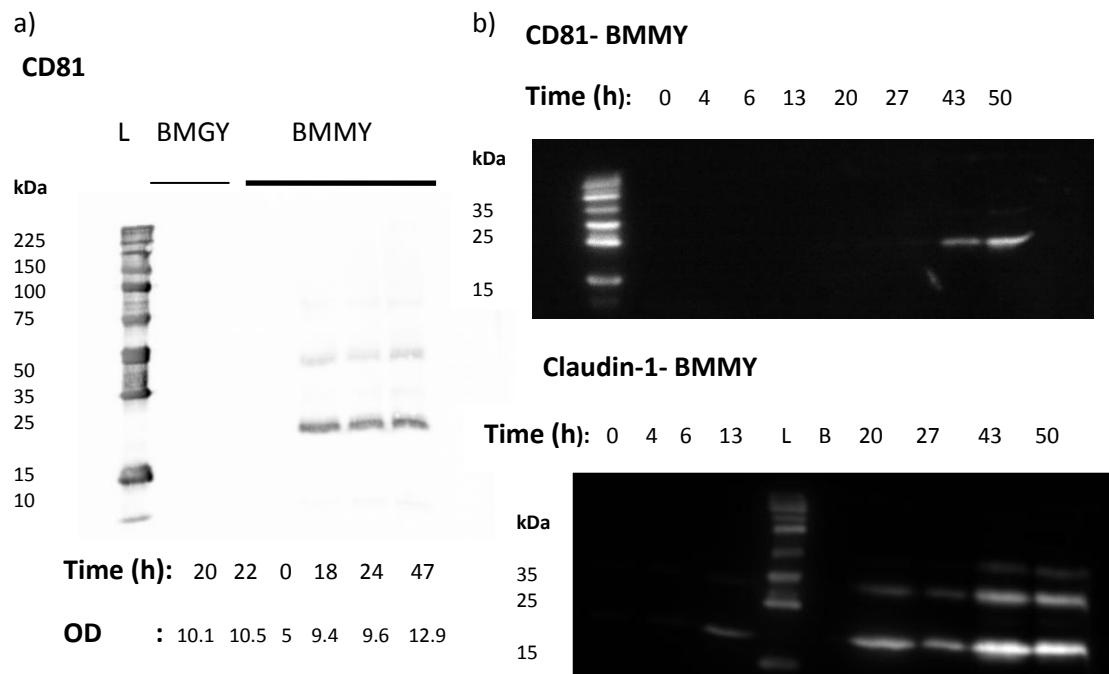
Figure 5.2: CD81 and Claudin-1 full length expression in *P. pastoris*. a) CD81 expression pre- and post induction. Yeast were grown in BMGY for ~22 hours to an OD of 7-10 and then induced in BMMY for ~47h prior to harvesting the cell pellet. Samples equating to the same OD were collected both pre-induction (BMGY) and post-induction (BMMY) and the total membrane fraction isolated and total protein quantified using a BCA assay. 25µg total protein was loaded on to an SDS gel for each sample and then CD81 detected using an anti-His antibody. b) CD81 and Claudin-1 expression over time (0-50 h) during induction in BMMY. 3ml samples collected at each time point were broken and the total membrane was isolated. Equal volumes (20µl) of membrane were loaded on to an SDS-gel and CD81 and Claudin-1 were detected using an anti-His antibody. Molecular size markers are indicated in part a) using a ProtoMetrics ladder (national diagnostics) (also shown in part b). L=ladder, B= blank lane.

Figure 5.2a demonstrates tight regulation of the *AOX1* promoter by methanol as induction of CD81 expression is only apparent at 18h post-induction and not before. This continues to at least 47 h post-induction as the cellular biomass continues to increase up to an OD_{600} of 12.9 in BMMY. Increasing biomass in BMMY over time is more apparent in figure 5.2b. The same volume of yeast membranes were loaded for each sample collected between 0-50 h induction (and therefore not corrected for total protein concentration). It was observed that as the time increased, the total yield of recombinant CD81 and Claudin-1 increased in the collected sample (figure 5.2b). This is likely due to an increase in cellular biomass over time as yeast use methanol as an alternative carbon source. In this instance, CD81 detection was apparent at 43h induction, whilst Claudin-1 at 13h induction. This may account for slower growth of CD81 expressing cells in comparison to Claudin-1 expressing cells since we know CD81 is expressed much earlier than this (18h post-induction) as seen in figure 5.2a. Based on these findings and a quote from Shi *et al*., (2002), '*production is roughly proportional to cell density*', all subsequent expression of recombinant CD81 and Claudin-1 was performed for 50h induction in BMMY prior to harvesting cells.

Recombinant CD81 is predicted to have a molecular weight of 26.7 kDa and Claudin-1 23.7 kDa, since both proteins are expressed with a C-terminal His_6 tag. CD81 shows the ability to form monomers and dimers (figure 5.2a) on a non-reducing SDS-PAGE and Claudin-1 monomers, dimers and trimers (figure 5.2b).

After producing CD81 and Claudin-1 independently in *P. pastoris* membranes, it was then necessary to extract the proteins from the membrane, rendering them soluble. This was previously achieved using detergents (Jamshad et al., 2008), (Bonander et al., 2013). CD81 was solubilised using β-OG and Claudin-1 using foscholine-10 (Fos-10); the detergent for each protein was previously selected in our laboratory.

Figure 5.3a demonstrates detergent CD81 and Claudin-1 extractions from the yeast membrane after an ultracentrifugation spin of the solubilised material. The supernatant was run on a SDS gel in order to detect solubilised protein. Subsequently, Claudin-1, CD81 and a non-HCV receptor CD82 solubilised in detergent were further analysed for HCV sE2 functional binding in an ELISA. Figure 5.3b demonstrates that CD81, but not Claudin-1 or CD82 bind to sE2, which is consistent with current literature (Bonander et

al., 2013) where Claudin-1 has not shown sufficient evidence of binding directly to HCV, or more specifically sE2, in contrast to CD81

a)

b)

Figure 5.3: HCV sE2 functional binding of detergent solubilised CD81 but not Claudin-1. a) CD81 and Claudin-1 were extracted from yeast membranes using detergents, β-octyl-glucopyranoside (β-OG) and foscholine-10 (Fos-10), respectively and detected with an anti-His antibody using western blot analysis. Solubilisation also achieved with polymers (referred to as polymer 1 and polymer 2) produced in Professor Brian Tighe's laboratory, Aston University, 1) non-solubilised protein and 2) solubilised protein. A ProtoMetrics ladder is shown for molecular weight markers (National Diagnostics).b) Claudin-1, CD81 and a non-HCV receptor CD82 were solubilised in detergent and subsequently analysed for HCV sE2 functional binding in an ELISA (performed by Ke Hu at Birmingham University).

In addition to extracting CD81 and Claudin-1 from yeast membranes using detergents, a more recent novel approach for solubilisation was performed using polymers. Amphipathic polymers known as styrene maleic acid lipid particles (SMALPs) directly

extract membrane proteins from the native membrane and form a disc-like structure that surrounds both protein and maintains surrounding lipids. This approach has been previously shown to enhance protein stability and function (Rajesh et al., 2011), (Jamshad et al., 2011). Figure 5.3 shows CD81 and Claudin-1 solubilisation achieved with polymers (referred to as polymer 1 and polymer 2) produced in Prof Brian Tighe's laboratory, Aston University. Polymer 1 was able to extract monomeric CD81, but was not 100% efficient as CD81 was found in the non-soluble fraction (figure 5.3a, lanes 2 and 1, respectively). Claudin-1 was extracted using a different polymer (referred to as polymer 2) and was shown to exist as higher oligomeric forms in the solubilised fraction, although a large level of protein remained unsolubilised (figure 5.3a, lanes 2 and 1, respectively). Future work using polymers would include optimisation of solubilisation conditions to improve extraction efficiency. Furthermore, protein functional assays would be useful to determine structural stability of CD81 and Claudin-1 in polymers. If structure and function were maintained, polymers may be useful for downstream crystallisation trials.

5.2.1 Scaling up recombinant protein production

Having confirmed functional expression could be achieved in shake flasks, *P. pastoris* recombinant protein expression was scaled up from small shake-flasks to either a 2 L bioreactor or multiple 2L baffled shake flasks.

Scaling up from a shake-flask to a bioreactor should increase the total yield of protein produced in the system because the cells grow in a more controlled environment on account of improved aeration. This was investigated by preparing a bioreactor containing basal salts medium (BSM) as a standard medium to use in a controlled environment, in order to grow and induce a culture of *P. pastoris* strain expressing CD81.

Figure 5.4 is a bioreactor trace, which shows variables such as pH, temperature, dissolved oxygen (dO) and evolved carbon dioxide (CO_2). The level of dO used by the bioreactor culture decreases very slowly and does not fall with a simultaneous rise in CO_2 until ~80 h from the beginning of the run (see figure 5.4). This suggested that the

culture did not undergo high levels of respiration and growth early on in the run and therefore would result in low levels of biomass, which is not optimal prior to induction of recombinant protein expression.

dO decreased at ~80 h
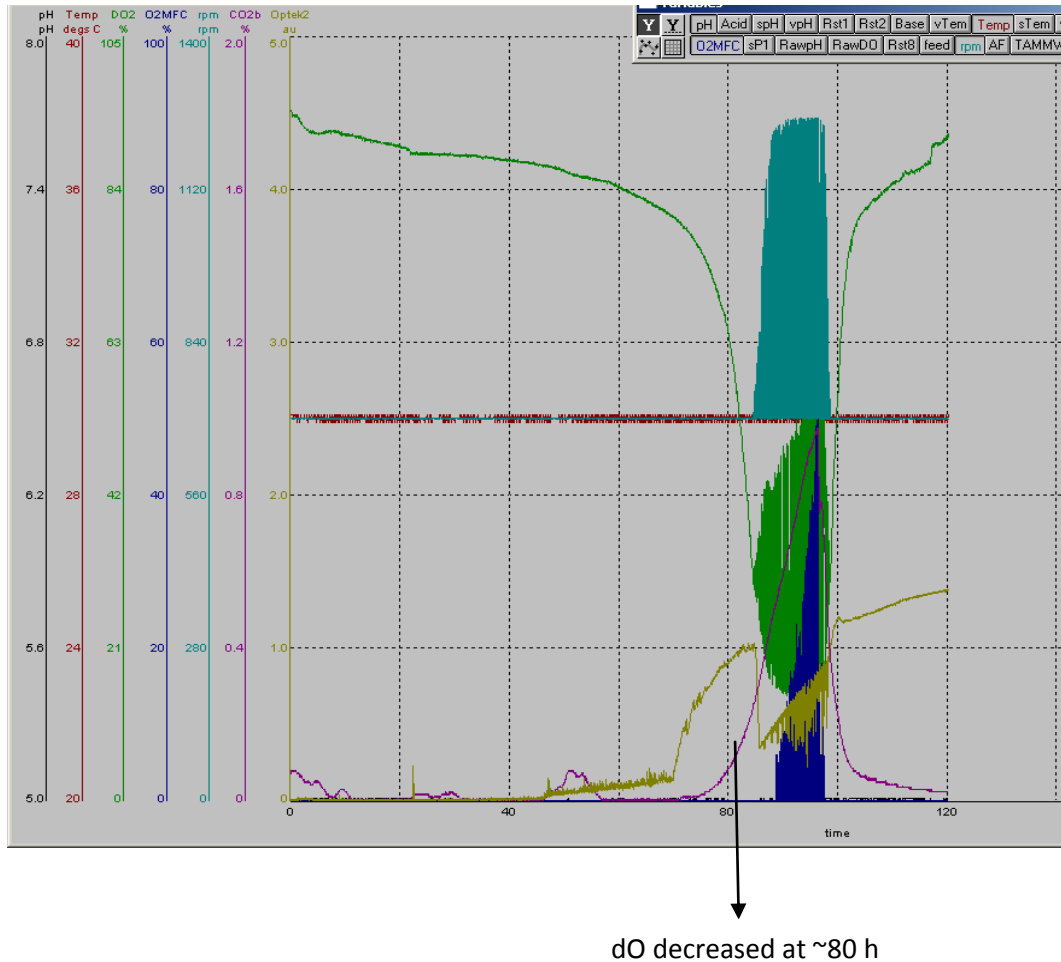
Figure 5.4: ***P. pastoris* CD81 recombinant strain grows at a slow rate in minimal medium.** *P. pastoris* expressing recombinant human CD81 was inoculated into a bioreactor consisting of a minimal medium, BSM. Culture conditions were set at 30°C, pH 5 and 30% dissolved oxygen (dO). This bioreactor run did not include an induction phase with methanol as the yeast cells did not grow efficiently, as shown by the lack of respiration up until ~80h by the dO and CO_2 traces.

Since *P. pastoris* CD81 production had earlier been achieved in shake flasks using a complex medium (BMGY), we investigated the effect of *P. pastoris* growth in a bioreactor by changing the medium from a minimal (BSM) to a complex (YPG) medium, which had been used successfully in a laboratory protocol to produce recombinant CD81 and Claudin-1 (Bonander et al., 2011).

143

Results indicated that the yeast strain grew successfully using a complex medium and we were able to assess the growth profile of the culture using online bioreactor measurements (see figure 5.5). The growth profile demonstrates a drop in the dO and rise in CO_2 within 12h of bioreactor inoculation, in contrast to that seen using a minimal medium (see figure 5.5 and 5.4 respectively). The culture OD_{600} was measured throughout the run and reached 16.5 just prior to the start of induction. Over time, there was a reduced capability for the culture to respire as the dO gradually increased. This immediately changed at the beginning of induction, when methanol was added to the vessel, the dO dramatically declined and CO_2 increased (see figure 5.5). This also occurred as a result of the third methanol supplementation to the vessel. Furthermore, the OD_{600} of the culture increased throughout the induction period and reached an OD_{600} of 31.8 at 52 h of the total bioreactor run.

a)

b)

CD81 26.7 kDa →

Methanol addition to vessel
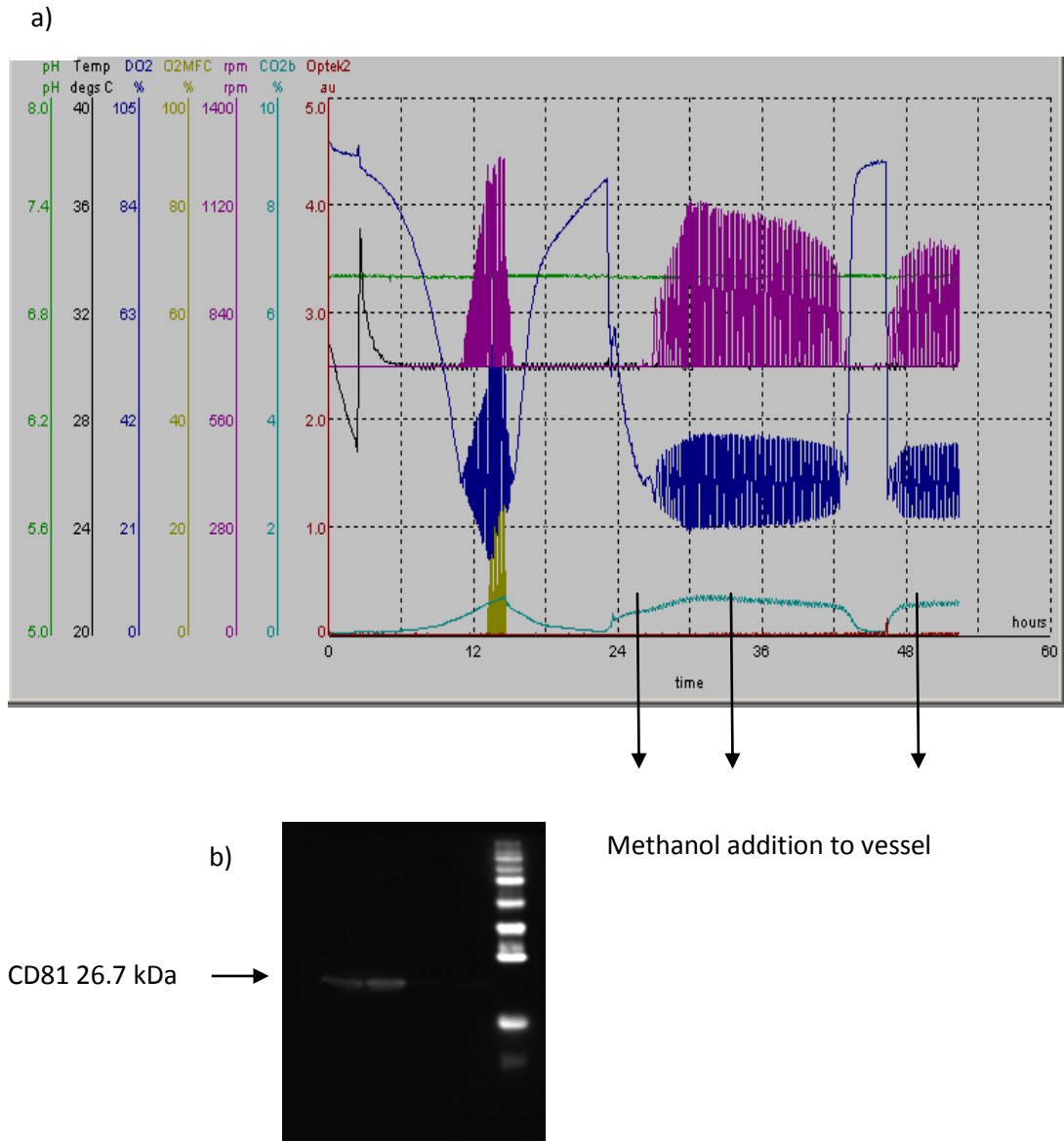
Figure 5.5: *P. pastoris* **CD81 expressing strain is suited for growth in a bioreactor using a complex medium.** a) *P. pastoris* strain expressing CD81 was inoculated into a bioreactor consisting of YPG (10g/L of glycerol in YPG, 1 L working volume) and methanol was added as indicated on the trace. Culture conditions were set at 30°C, pH 7 and 30% dO. The level of dissolved oxygen (dO) in the bioreactor and the amount of CO_2 evolved was recorded over time. The trace shows a bioprocess response. As *P. pastoris* demands a high level of O_2 for growth, the system attempts to maintain a specific set-point (i.e. the dO will be maintained at ~30%). When this is not achieved the system responds by increasing the stirrer speed; following this more air (40% O_2; 60% N_2) is pumped into the vessel in order to reach set-point. This can be seen here by the dO trace fluctuating around 30% and also the stirrer trace (rpm) responding accordingly. b) Western blot analysis confirming CD81 expression.

CD81 expression was successful using a bioreactor containing complex medium such as YPG (figure 5.5b). Even so, in order to produce high yields of membrane protein easily and efficiently from *P. pastoris*, large baffled shake flasks were tested in comparison to a bioreactor and were deemed just as successful regarding cell biomass reached at the end of the batch run and regarding set-up time was a more efficient approach. For each production batch, 1600ml BMGY was used to grow the yeast strain using 8× 200ml working volume (1L shake flasks). Once grown to an OD_{600} ~7-10 the cells were used for BMMY induction in 8× 500ml working volume cultures (using 2L flasks). Cells were supplemented with methanol (1% v/v) at 24h post-induction and subsequently harvested at around 48-54 h post-induction. This approach typically yielded ~80g total wet cell weight. In comparison a bioreactor with 1L working volume would yield approximately 60-70g wet cell weight. Since there was not a great difference between the resulting cell weight yields of the two production approaches, ultimately large 2L flasks were used routinely for CD81 production hereafter due to ease of use and sufficient material obtained for solubilisation and purification.
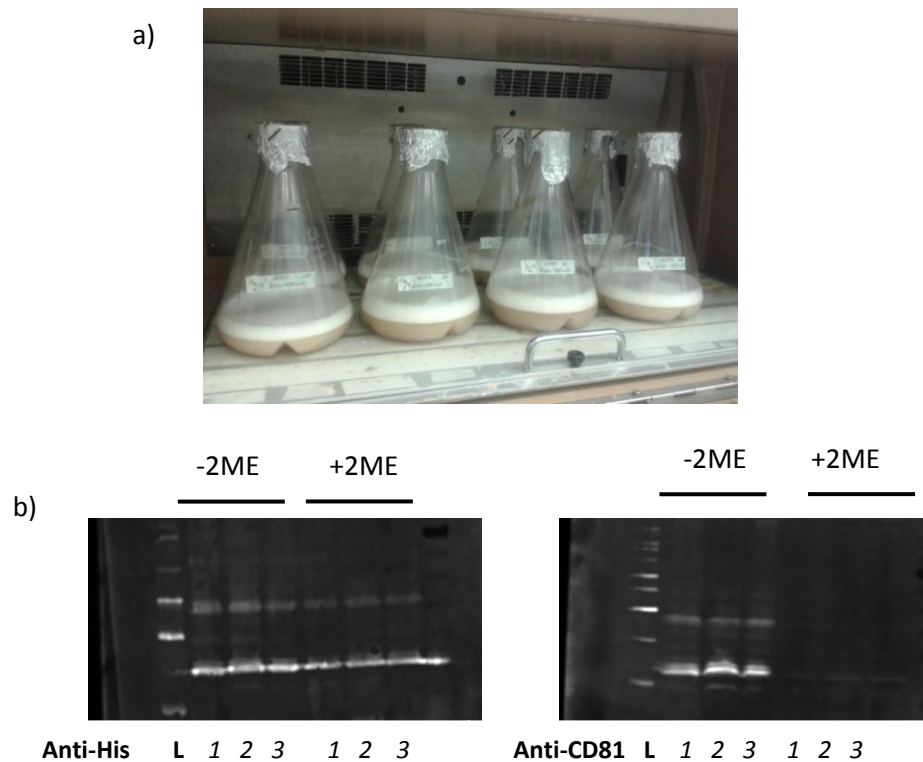
Figure 5.6: CD81 expression in large 2L baffled shake flasks. a) One batch culture would consist of 8 x 2L shake flasks (500ml working volume) and cells were induced in BMMY for ~50h. b) 3 independent batches of cells were broken using a glass bead method and the total membrane isolated. 5μl per sample was loaded on to a gel with/without 2-mercaptoethanol (2ME) included in the western loading sample buffer. A western blot was performed and CD81

was either detected using an anti-His (Clontech) antibody or anti-CD81 (2.131). +/- 2ME in sample buffer is indicated in figure. L= ProtoMetrics ladder (national diagnostics) and *1, 2* and *3* indicates if the protein sample was from batch culture *1, 2* or *3* produced independently.

Figure 5.6b indicates CD81 expression in three independent large scale batch cultures using 2L baffled shake flasks. The protein band intensity appears consistent between batches suggesting reproducibility in the method used. Recombinant CD81 was detected using both anti-His and also a specific anti-CD81 antibody, showing both monomers and dimers are present in the yeast membrane. The anti-His antibody bound both +/- 2-mercaptoethanol (2ME) treated samples in contrast to the anti-CD81 monoclonal antibody used, which could only bind -2ME samples. The anti-CD81 antibody is conformationally sensitive; therefore this experiment suggests CD81 is correctly folded in yeast membranes since the anti-CD81 antibody cannot bind to reduced CD81 samples. Since anti-CD81 2.131 specifically binds the folded LEL it seems likely that the absence of disulphide bonds in this structure, as a result of 2ME in the sample buffer, prevents binding of the antibody.

5.3 Large scale CD81 solubilisation and purification

Now that there was sufficient evidence for CD81 production, antigenicity and insertion into *P. pastoris* membranes, the next challenge was to extract high yields of CD81 from yeast membranes using detergent. In order to do this, time was spent at the Membrane Protein Laboratory (MPL) at the Diamond Light Source (Oxfordshire) in order to gain expert advice on dealing with membrane proteins from Isabel Moraes, James Birch and other group members with an aim to put CD81 through crystallisation trials. Initially, CD81 was solubilised and purified in DDM as this is a standard detergent used in the membrane protein field for crystallisation attempts (Gutmann et al., 2007) and seemed like a good place to begin. Standard buffer conditions tried initially were 1× PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, and 2mM KH_2PO_4, pH 7.4); 150mM NaCl, 10% glycerol and DDM (1% for solubilisation and 0.1% in purification buffers) (figure 5.7).

After solubilisation in DDM for 2h at 4°C, the sample was spun in an ultracentrifuge to remove unsolubilised material in the pellet. Solubilised material in the supernatant was

then bound to Ni-NTA resin in order to purify CD81. Subsequently two imidazole washes (30mM and 40mM) and a final elution at 250mM imidazole rendered semi-pure CD81 observed as a monomer, dimer and possibly higher order forms (as shown in figure 5.7). The sample was then concentrated using a 50 kDa cut off centrifugal concentrator (Sartorius), although soon into the concentration procedure CD81 at around 0.593 mg/ml appeared to have aggregated under these conditions, therefore either the buffer conditions, detergent or both were not suitable to maintain CD81 stability.

Sample	Solubilisation/purification fraction
1	Prior to ultracentrifugation spin
2	After ultracentrifugation spin
3	Column flow through (unbound material)
4	30mM imidazole wash
5	40mM imidazole wash
6	250mM imidazole elution
7	Concentrated CD81 sample
8	Aggregated pellet
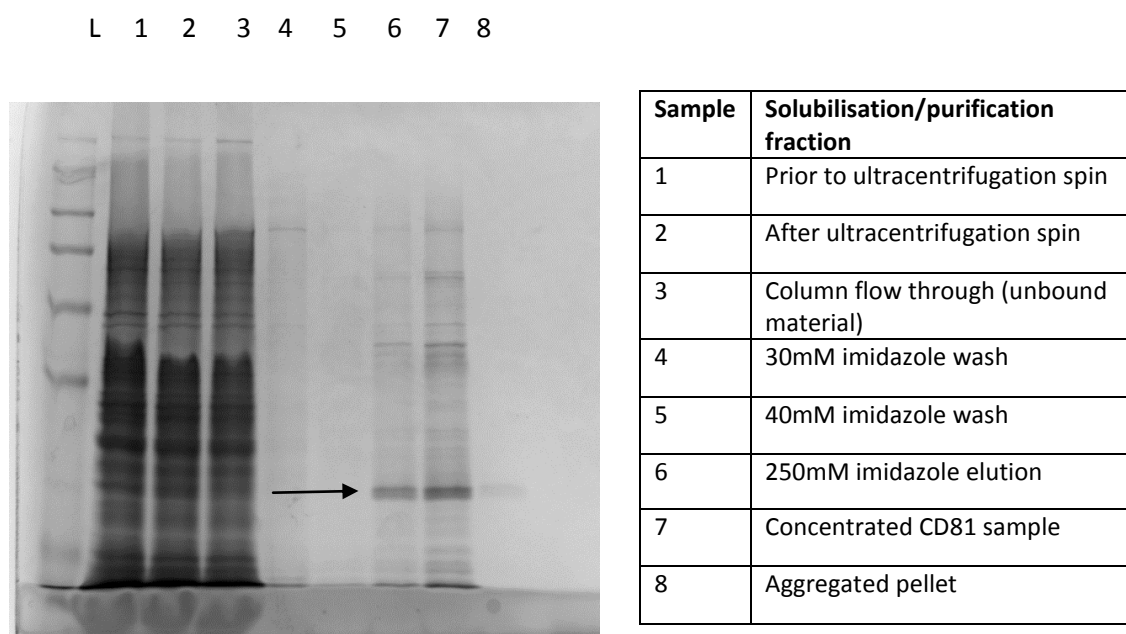
Figure 5.7: Solubilisation and purification of CD81 using DDM. CD81 was solubilised using 1% DDM in buffer (1× PBS, 150 mM NaCl, 10% glycerol, 20mM imidazole) for 2h at 4°C. After an ultracentrifuge spin, the solubilised material was taken as the supernatant and bound to nickel resin for 1h at 4°C. After performing 30mM and 40mM imidazole washes, CD81 was eluted from the column containing resin using 250mM imidazole. The protein was concentrated using a 50 kDa cut off centrifuge spin tube with PES membrane (Sartorius). Samples collected at each point of the solubilisation and purification procedure were loaded on to an SDS-gel and later developed using Instant blue (Expedeon). Arrow represents size of monomeric CD81.

5.4 Optimisation of the thermal stability of CD81 for crystallisation trials using a CPM assay

CD81 must be extracted from the host membrane in order for purification to occur; crucial annular lipids that maintain protein structure are lost and replaced by detergent molecules forming a protein-detergent complex. Naturally, detergents cannot mimic the lipid bilayer exactly and so membrane proteins typically undergo instability problems, which may lead to protein aggregation. Since ideally 10mg/ml was needed for crystallisation trials, conditions that yield soluble and stable CD81 were investigated using a CPM thermal stability assay (Alexandrov *et al*., 2008), since this is one factor that may contribute to successful crystallisation trials. Figure 5.8 shows cysteine residues in CD81 and Claudin-1 used in the assay that are involved in binding CPM during protein unfolding, which is measured by a fluorescent read-out over time at 40°C.

a)

Protein sequence for CD81 palmitoylation null mutant:

```
M S G V E G A T K A I K Y L L F V F N F V F W L A G G V I L G V A L W
L R H D P Q T T N L L Y L E L G D K P A P N T F Y V G I Y I L I A V G
A V M M F V G F L G A Y G A I Q E S Q A L L G T F F T C L V I L F A C
E V A A G I W G F V N K D Q I A K D V K Q F Y D Q A L Q Q A V V D D D
A N N A K A V V K T F H E T L D C C G S S T L T A L T T S V L K N N L
C P S G S N I I S N L F K E D C H Q K I D D L F S G K L Y L I G I A A
I V V A V I M I F E M I L S M V L A A G I R N S S V Y G G G H H H H H
H Stop
```
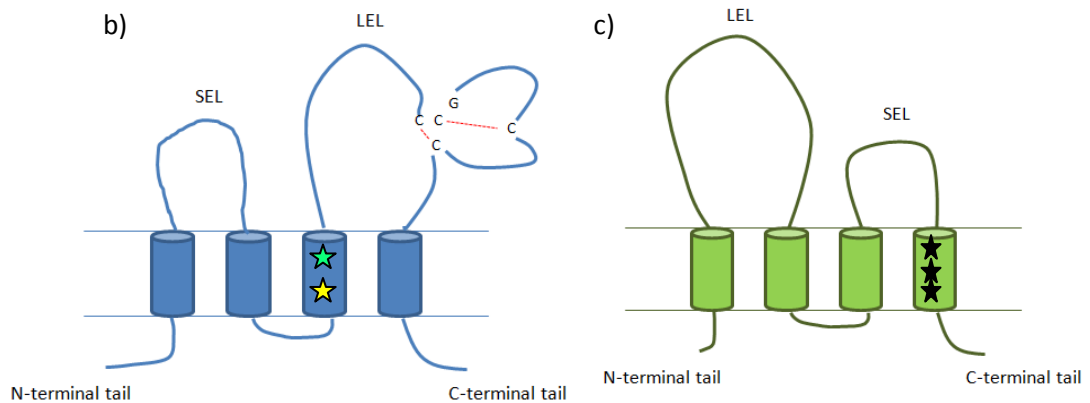

Figure 5.8: Cysteines available in CD81 and Claudin-1 for CPM reactivity. Temperature induced protein unfolding of CD81 or Claudin-1 causes embedded cysteine residues to be exposed to CPM reactivity, which is measured as fluorescence emitted over time. a) Mutant CD81 protein sequence highlighting putative palmitoylation residues mutated from Cys to Ala (alanine residues are blue and bold) (Jamshad et al., 2008). Cysteines remaining for CPM reactivity are in TM 3 and highlighted in yellow (C97) and green (C104). Cysteines involved in disulphide bonds are highlighted in red and found in the LEL. b) CD81 showing TM3 cysteines (yellow C97) and green (C104) and those involved in the two disulphide bonds in the LEL. C) WT Claudin-1 showing cysteines embedded in TM4 (black stars), C175, C183 and C184.

5.4.1 CD81 thermal stability in detergent solutions

Optimising the detergent used to maintain solubility and stability is imperative for a recombinant membrane protein. The CPM assay was used to assess the stability of CD81 in different detergents as aggregation issues were apparent for CD81 in DDM. All detergents were tested at 3× critical micelle concentration (CMC) in buffer; 20mM Tris, pH7.5, 150mM NaCl, and using 8µg total CD81 for each condition. Buffer and protein were allowed to equilibrate at room temperature for ten minutes before addition of CPM and subsequent incubation at 40°C for 180 minutes, measuring the fluorescence read-out every 5 minutes in a 96-well spectrofluoremeter. Using the raw data from each time point, the fraction of folded protein was calculated using the highest fluorescent score for each test detergent condition as the 'maximal unfolded state'. An exponential one phase decay curve was fitted, in order to deduce the unfolding half life in minutes.

The results shown that all detergents tested were more stabilising than DDM (see figure 5.9), which correlated well with observations that CD81 aggregated in this detergent. CD81 shows improved stability in smaller detergents such as β-OG (292 Da) resulting in a half life of 19.5 min in comparison to larger detergents such as DDM (511 Da) with a half life of only 9.6 min (figure 5.9b). Furthermore, the addition of cholesteryl hemisuccinate (CHS) with β-OG results in further stability of CD81 (mean half life 25 min). β-OG + CHS was found to be significantly more stabilising as compared to DDM, UDM and DM alone and also when compared to β-OG and CYMAL6 alone, but the latter did not reach statistical significance (figure 5.9b). Cholesterol has previously shown importance in its contribution to stability and function of tetraspanins in the mammalian plasma membrane (Harris et al., 2010) and also stability to recombinant CD81 when in complex with Claudin-1 (Bonander et al., 2013).

a)
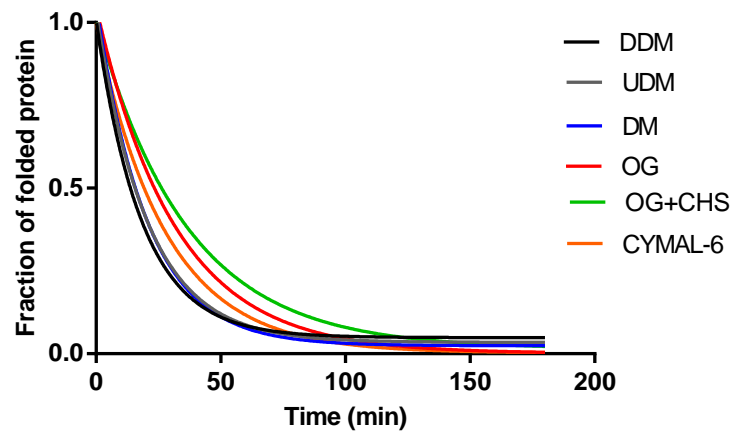
b)
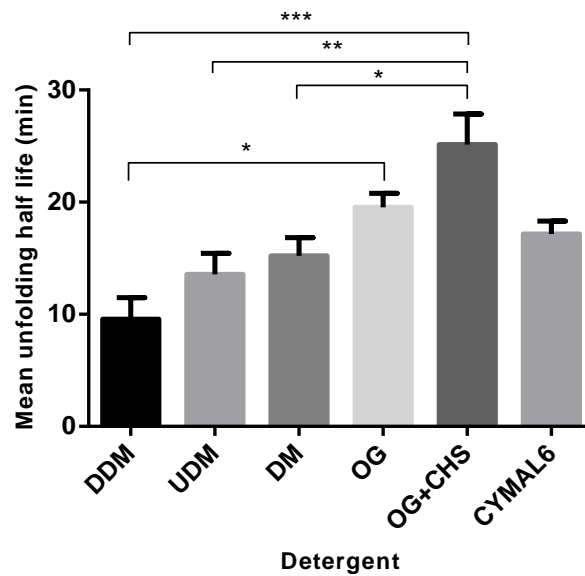
Figure 5.9: CD81 thermal stability in detergent solutions. 8µg of purified CD81 was incubated with different detergents at 3× critical micelle concentration (CMC) in buffer (20mM Tris, pH 7.5, 150mM NaCl) for 10 minutes at room temperature prior to adding CPM and incubating at 40°C for 180 minutes. Using a spectrofluoremeter, measurements were recorded every 5 minutes and the raw data for each condition were converted to a fraction of folded protein, using the highest fluorescence value reached for that condition as 'maximal unfolded protein'. a) Representative exponential one phase decay curve. b) The mean unfolding half life in minutes for n=3. ANOVA was performed followed by a Tukey's multiple comparison test.

DM was used for CD81 solubilisations since it had shown evidence of extracting higher yields of CD81 in comparison to β-OG (0.64 mg/ml and 0.29 mg/ml in the column eluate, respectively). Also, a larger pellet was observed after the β-OG solubilisation ultracentrifugation step in comparison to DM solubilised material suggesting that DM is able to extract CD81 more efficiently than β-OG.

During the purification procedure CD81 was exchanged into β-OG with CHS or left in DM. Samples collected from solubilisation through to purification are shown on a gel in figure 5.10. It can be seen that CD81 exists predominantly in a monomer and dimer in both detergents (lane 7 and 11 in figure 5.10). Furthermore, in contrast to CD81 in DDM, the ability to concentrate CD81 in β-OG or DM was achieved without protein aggregation to concentrations of 9 mg/ml and 13 mg/ml, respectively. The ability to concentrate a membrane protein without aggregation issues suggests a more stable protein preparation, therefore an improvement in comparison to using DDM (seen earlier). Since we had CD81 at reasonably high concentrations, we next used the samples for crystallisation trials and analysis using size-exclusion chromatography coupled to a multi-angle light scattering detector system (SEC-MALS).
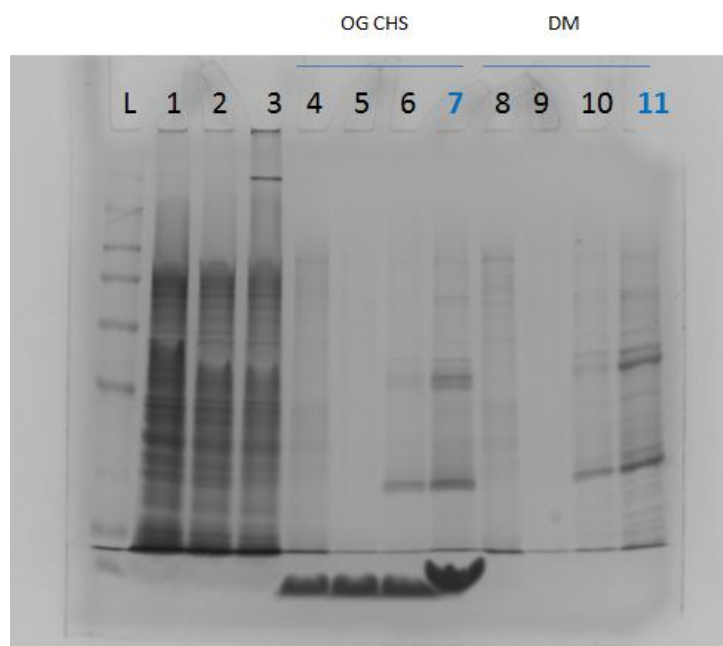
Sample	Solubilisation/purification fraction run on gel
1	Pre-ultracentrifugation spin (unsolubilised material)
2	Post ultracentrifugation spin (solubilised material)
3	Flow through (unbound to column resin)
4	30mM imidazole wash (OG+CHS)
5	40mM imidazole wash (OG+CHS)
6	Elution (OG+CHS)
7	Concentrated CD81 in OG+CHS (9 mg/ml)
8	30mM imidazole wash (DM)
9	40mM imidazole wash (DM)
10	Elution (DM)
11	Concentrated CD81 in DM (13 mg/ml)

Figure 5.10: CD81 solubilisation and purification in β-OG+CHS and DM. CD81 was solubilised from yeast membranes using 1× PBS, 150mM NaCl, 10% glycerol, 20mM imidazole and 1% DM for 2h at 4°C. After binding solubilised material to nickel resin for 1h at 4°C the column resin was washed in 30mM and 40mM imidazole followed by elution in 250mM imidazole using either a DM buffer or β-OG+CHS buffer (3X CMC). CD81 was then concentrated using 50kDa cut off concentrating device (Sartorius) and reached concentrations of 9mg/ml in OG+CHS (lane 7) and 13mg/ml in DM (lane 11). Samples from each stage of solubilisation and purification were run on an SDS-gel and developed using Instant blue (Expedeon).

5.4.1.1 Crystallisation trials for CD81

Using CD81 purified in DM (13 mg/ml) and in β-OG+CHS (9 mg/ml) crystallisation plates were set up, guided by James Birch (MPL). CD81 samples were spun down to remove any particulate matter. The crystallisation screen used two 96 well plates, MemGold and MemGold 2 (Molecular Dimensions) which encompass a range of 96 conditions each with varying pH, PEGs and salt additives based on crystallisation conditions taken from the Protein Data Bank. This aims to find a condition for crystallisation that is suited to the protein target. A Mosquito lipidic cubic phase (TTP Labtech) was used to automate the delivery of 100nl of CD81 protein sample to 100nl of each buffer in both 96 well plates (either MemGold or MemGold2). The 2 plates were then incubated at 20°C and after 2-3 weeks crystals had formed, but were later found not be CD81 crystals as confirmed by Isabel Moraes (MPL). Therefore, at this point it was concluded that further optimisation regarding preparation of purified CD81 was required before repeating this process again.

5.4.1.2 SEC-MALS analysis of CD81

SEC-MALS is a useful method to analyse a protein sample regarding the size of a protein-detergent complex, the ratio between how much is protein and how much is detergent and also the level of free detergent in a sample and its molecular weight. The MALS part of the system is composed of a low-angled light scattering and right angled light scattering detectors. The system includes a refractive index (RI) detector and an ultraviolet light detector (information from MPL). CD81 were processed using SEC-MALS by Matthew Jennions (MPL).

CD81 purified in DM at 1mg/ml was taken and analysed using SEC-MALS in order to assess the effect of two different buffers on CD81 homogeneity, the size of protein-detergent complexes and the level of excess detergent in the sample. Large levels of free detergent in comparison to the level seen in protein-detergent complexes are not typically optimal for crystallisation trials.

The first buffer was identical to that used during nickel resin purification, 1× PBS, 150mM NaCl, 10% glycerol, 250mM imidazole and 3× CMC DM, whilst the second buffer was composed of 20mM Tris, pH 7.5, 150mM NaCl and 3× CMC DM.

The SEC-MALS trace given for CD81 in the PBS buffer (see figure 5.11) shows that the protein-detergent complex is predominantly a dimer with a molecular weight of ~ 140 kDa (see purple line, peak 3, figure 5.11). Potentially the glycerol and high salt content of the buffer may be pushing the equilibrium to a CD81 dimer. There is also a much larger oligomer present that is ~570 kDa (peak 2 purple trace, figure 5.11). The refractive index trace (red trace) represents the level of detergent present. The trace in figure 5.11 suggests that there is no detergent present in the oligomer complex but there is in the dimer protein-detergent complex. The CD81 oligomer may have formed due to the lack of detergent surrounding the CD81 molecules. It is also clear from the refractive index (red trace) that there is excess detergent present in the protein sample since the level of excess detergent is higher (indicated by * in figure 5.11) than that present in the dimeric detergent-protein complexes and so this particular sample would be predicted as non-optimal for crystallisation trials because of the crystallising propensity of detergent themselves, which is likely to have occurred in the previous crystallisation trial.

The SEC-MALS trace (see figure 5.12), which used a Tris buffer shows that there are two CD81 species present, a dimer (~110 kDa, peak 2) and a monomer (~53 kDa, peak 3) involved in protein-detergent complexes. Again, it is shown by the refractive index that the excess free detergent present is much higher than detergent involved in protein-detergent complexes (see *, figure 5.12). Optimally, free detergent should have a lower level than that seen in protein-detergent complexes and may explain why crystallisation trials were unsuccessful in this attempt. DM was used in this sample at 3× CMC, therefore future work could try using CMC values lower than this, although not reducing lower than 1× CMC (advice from Isabel Moraes, MPL). Also, CD81 could be concentrated using a 100 kDa cut off device rather than 50kDa cut off in order to aim towards a more homogenous CD81 preparation. Since CD81 buffers used thus far had not been optimised for tetraspanin stability, investigations next explored all solution components using the CPM thermal stability assay.

Aston University

Illustration removed for copyright restrictions

Figure 5.11: SEC-MALS analysis of CD81 in a PBS buffer. Purified CD81 (1mg/ml) was injected into a SEC-MALS system equilibrated with 1×
PBS, 150mM NaCl, 10% glycerol, 250mM imidazole and 3× CMC DM. The purple trace represents ultra violet (UV) (280nm) detecting the absorption
of protein present in the sample, detergent show negligible UV signal. The red trace represents the refractive index (RI) which detects both protein and
detergent in solution. The difference between RI and UV allows estimation of the molecular weight of the protein involved in protein-detergent
complexes. The run was executed by Matthew Jennions (MPL). UV peaks 1, 2 and 3 are shown in the figure. Excess detergent present in the sample
(*).

Aston University

Illustration removed for copyright restrictions

Figure 5.12: SEC-MALS analysis of CD81 in a tris buffer. Purified CD81 (1mg/ml) was injected into a SEC-MALS system equilibrated with 20mM Tris, pH 7.5, 150mM NaCl and 3× CMC DM. The purple trace represents ultra violet (UV) (280nm) detecting the absorption of protein present in the sample, detergent show negligible UV signal. The red trace represents the refractive index (RI) which detects both protein and detergent in solution. The difference between RI and UV allows estimation of the molecular weight of the protein involved in protein-detergent complexes. The run was executed by Matthew Jennions (MPL). UV peaks 1, 2 and 3 are shown in the figure. Excess detergent present in the sample (*).

5.4.2 Effect of buffer ionic strength on CD81 thermal stability

In order to determine if ionic strength has an effect on CD81 stability, CD81 was purified in DM and resuspended in buffer containing 20mM Tris, pH 7.4, 0.26% DM and increasing concentrations of NaCl from 100mM to 1M. Buffer and 8µg total protein were allowed to equilibrate at room temperature for ten minutes before addition of CPM and subsequent incubation at 40°C for 180 min, measuring the fluorescence read-out every 5 min.

As shown in figure 5.13, the NaCl concentration did not have a significant effect on CD81 stability. 100mM NaCl gave the highest mean unfolding half life (20.24 ± 1.80 min). When the NaCl concentration was increased up to 1000mM, the mean unfolding half life for CD81 decreased (17.55 ± 2.97 min), although it was not significantly different to 100mM NaCl. These data suggest that the stability of CD81 does not differ in solutions containing increasing NaCl concentrations from 100-1000Mm.
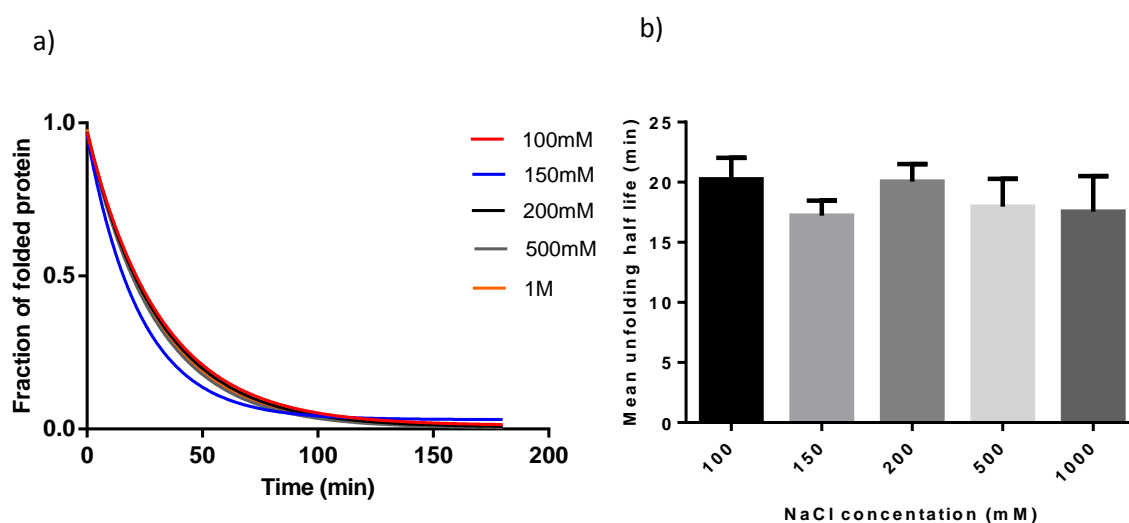
Figure 5.13: Effect of buffer ionic strength on CD81 thermal stability. 8µg of purified CD81 was incubated with increasing concentrations of NaCl, 100mM-1M in buffer (20mM Tris, pH 7.5, 0.26% DM) for 10 min at room temperature prior to adding CPM and incubating at 40°C for 180 min. Using a spectrofluoremeter, measurements were recorded every 5 minutes and the raw data for each condition were converted to a fraction of folded protein, using the highest fluorescence value reached for that condition as 'maximal unfolded protein'. a) Representative exponential one phase decay curve. b) The mean unfolding half life in minutes for n=2 (using 2 different protein preparations). ANOVA was performed followed by a Tukey's multiple comparison test.

5.4.3 Effect of glycerol on CD81 thermal stability

Glycerol included in solubilisation and purification buffers can be used to stabilise protein structure (Alexandrov et al., 2008). Tetraspanin stability with or without glycerol has not been shown before using a thermal stability assay, therefore glycerol was investigated using 0, 10, 20 and 30% glycerol in buffer (20mM Tris, pH 7.5, 150mM NaCl, 0.26% DM). 8µg total protein was added in each buffer condition and left to incubate for ten minutes before exposing to CPM and recording fluorescence measurements every 5 minutes at 40°C.

CD81 in the presence of glycerol has increased stability in comparison to without (see figure 5.14). CD81 in buffer without glycerol resulted in a mean unfolding half life of 19.29 $\pm$ 0.39 minutes in comparison to 24.71 $\pm$ 1.66 minutes in the presence of 10% glycerol. This indicates that glycerol may increase the thermal stability of purified CD81, although the result did not reach statistical significance. Increasing the glycerol concentration to 20 and 30% did not make a significant difference when compared to 10%, and so it can be concluded that 10% glycerol provides CD81 with sufficient stability.

Although this experiment does not provide evidence as to why glycerol improves the stability of CD81, it has been previously found that glycerol reduces a protein's ability to undergo conformational changes and therefore stabilises the native protein structure (Tsai et al., 2000). Therefore, glycerol may improve CD81 thermal stability as shown here by reducing conformational changes in the extracellular loop domains for example.
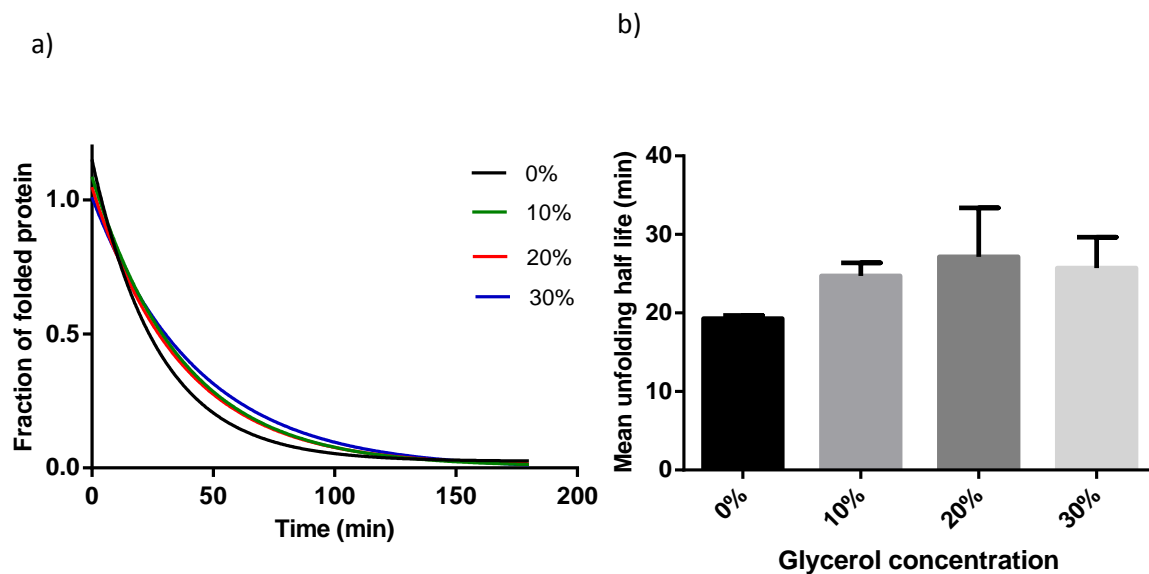
Figure 5.14: Effect of glycerol on CD81 thermal stability. 8μg of purified CD81 was incubated in buffer (20mM Tris, pH 7.5, 150mM NaCl, 0.26% DM) with or without glycerol (0, 10, 20 and 30%) for 10 min at room temperature prior to adding CPM and incubating at 40°C for 180 min. Using a spectrofluoremeter, measurements were recorded every 5 min and the raw data for each condition were converted to a fraction of folded protein, using the highest fluorescence value reached for that condition as 'maximal unfolded protein'. a) Representative exponential one phase decay curve. b) The mean unfolding half life in minutes for n=3 (using 2 different protein preparations). ANOVA was performed followed by a Tukey's multiple comparison test.

5.4.4 The influence of buffer composition and pH on CD81 thermal stability

The buffer composition and pH can have a significant effect on protein stability as shown previously for the human apelin receptor, a GPCR family member (Alexandrov et al., 2008). We next investigated the stability of CD81 in different buffers in a range of pH values between 6.5- 8.3. Buffers at different pH values were incubated with 8µg total CD81 for ten minutes before exposure to CPM and fluorescence recorded over time at 40°C.

The results demonstrated that CD81 has increased stability at lower pH (~6.5- 7) (see figure 5.15). This trend appeared in all three different buffers tested, which were Tris, HEPES and MOPS. In each buffer, when the pH was increased, the stability of CD81 decreased as shown by the fraction of folded protein (figure 5.15 a-e). The least favourable combination for CD81 stability was MOPS at pH 8.3 which gave a mean unfolding half life of 13.86 ± 3.04 minutes (figure 5.15 e). In comparison, MOPS at pH 6.5 provided the highest level of stability out of the conditions tested, since the mean unfolding half life for CD81 was 22.51 ± 4.41 minutes, although this result was not shown to be statistically significant. These data may suggest that the pH rather than the buffer composition has more influence on CD81 stability.
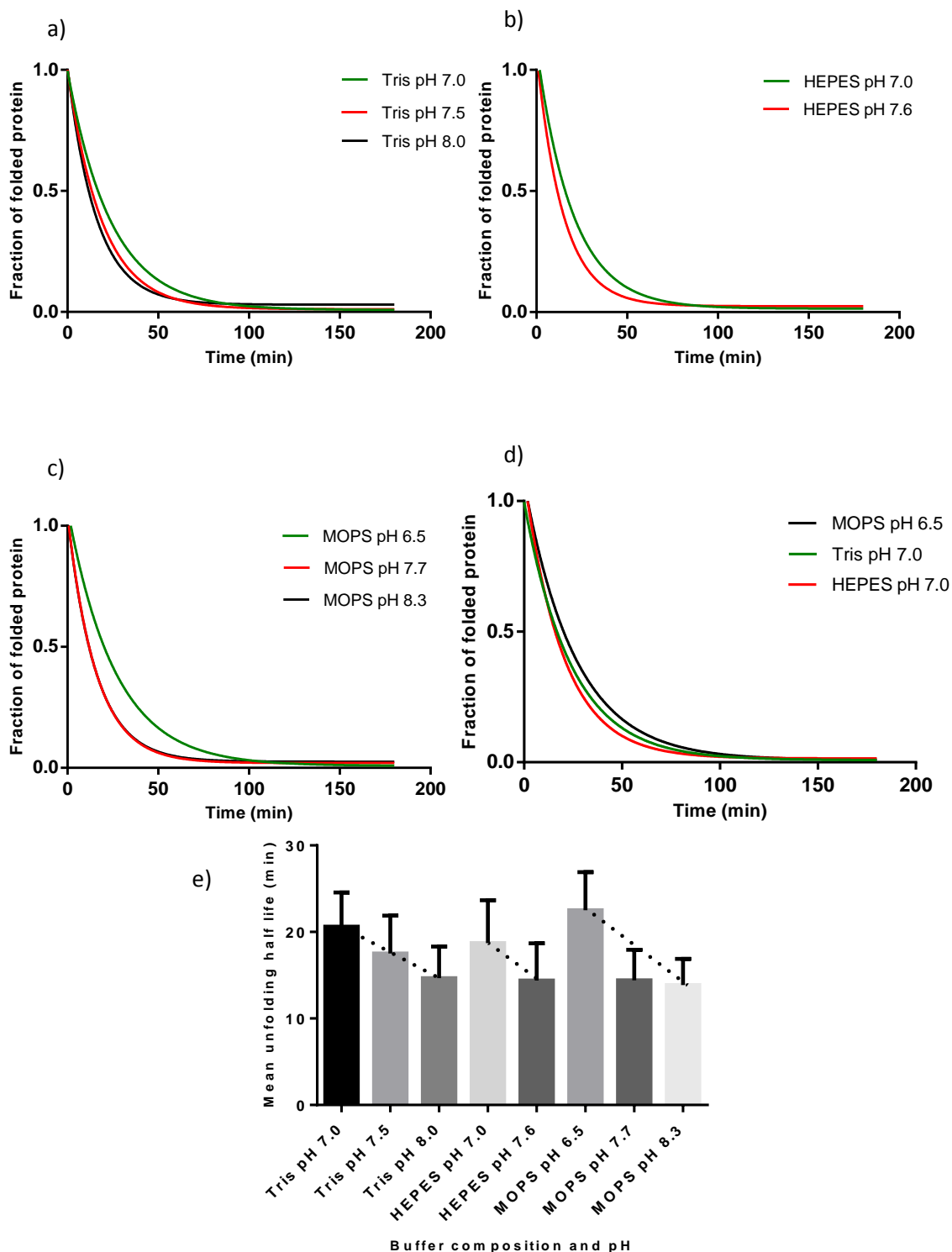
Figure 5.15: The influence of buffer composition and pH on CD81 thermal stability. 8μg of purified CD81 was incubated in buffer (20mM buffer, pH 6.5-8.3, 150mM NaCl, 0.26% DM) for 10 min at room temperature prior to adding CPM and incubating at 40°C for 180 min. Using a spectrofluoremeter, measurements were recorded every 5 min and the raw data for each condition were converted to a fraction of folded protein, using the highest fluorescence value reached for that condition as 'maximal unfolded protein'. a-d) Representative exponential one phase decay curves for a) Tris pH 7-8, b) HEPES pH 7-7.6, c) MOPS pH 6.5-8.3 and d) all three buffers tested at the lowest pH. e) The mean unfolding half life in minutes for n=2 (using 2 different protein preparations). Dotted line indicates the decrease in unfolding half life of CD81 as the pH is increased. ANOVA was performed followed by a Tukey's multiple comparison test.

5.4.5 Effect of cholesteryl hemisuccinate (CHS) on CD81 and Claudin-1 thermal stability

DM was used to extract CD81 from *P. pastoris* membranes, since it was found to extract more CD81 than β-OG. Since initial findings, using the CPM assay, suggested CHS increased the thermal stability of CD81 in β-OG (see figure5.9), we then investigated further to determine if the same applied to DM and CHS. Also, we investigated if the amount of CHS present (0.1% and 0.2% (v/v)) with DM or OG had any effect on protein stability using the CPM assay. CHS was previously found to stabilise the CD81-Claudin-1 complex (Bonander et al., 2013) as shown by dynamic light scattering. Therefore, we investigated using the CPM assay whether this previous finding was due to a CD81-specific CHS stabilising effect or whether CHS has a stabilising influence on Claudin-1 also. To test this, Claudin-1 was extracted and purified using Fos-10 and incubated with two different concentrations of CHS (0.1% and 0.2%).

Results shown that CD81 stability is significantly increased when DM is used in combination with 0.2% CHS, as the mean unfolding half life for CD81 in DM alone is 16.33± 1.13 minutes in comparison to 24.51 ± 1.25 minutes when DM and 0.2% CHS is added (see figure 5.16a). The same trend applied to the stability of CD81 when in β-OG with or without 0.2% CHS. The mean unfolding half life of CD81 in β-OG alone was 18.68 ± 1.19 minutes, in contrast to a significant increase for β-OG with 0.2% CHS at 24.26 ± 1.62 minutes (see figure 5.16a). The stability CHS provides CD81 appears to be concentration dependent since unfolding half life mean values slightly increase from 0.1% to 0.2% CHS in both DM and β-OG. These data suggest that the stability improvement of CD81 is CHS dependent, but also with some contribution to which detergent is used, with β-OG providing more stability than DM (see figures 5.16).

Results also demonstrated that CHS has a stabilising effect on Claudin-1 (figure 5.16b). When Claudin-1 is in Fos-10 alone the mean unfolding half life is 16.64 ± 0.23 minutes. Claudin-1 is further stabilised by the presence of CHS, with both concentrations 0.1% (21.16 ± 1.09 minutes) and 0.2% (22.2 ± 0.10 minutes) being a significant increase as compared to Fos-10 alone (see figure 5.16b). Although in this case, for Claudin-1, CHS does not appear to be providing protein stability in a concentration dependent manner,

as there is not a great difference between mean unfolding half life for 0.1% and 0.2% CHS (21.16 and 22.2 minutes, respectively).

It seems apparent that both CD81 and Claudin-1 independently benefit, regarding protein structure stability, in the presence of CHS. These data may explain why cholesterol in mammalian membranes is crucial for the CD81-Claudin-1 complex to carry out HCV entry, since depletion of cholesterol inhibits HCV infection (Harris et al., 2010). Furthermore, it provides more evidence as to why CHS prolongs the stability of recombinant forms of the CD81-Claudin-1 complex seen previously (Bonander et al., 2013) and that both membrane proteins benefit from CHS addition to reduce protein unfolding and increase stability.
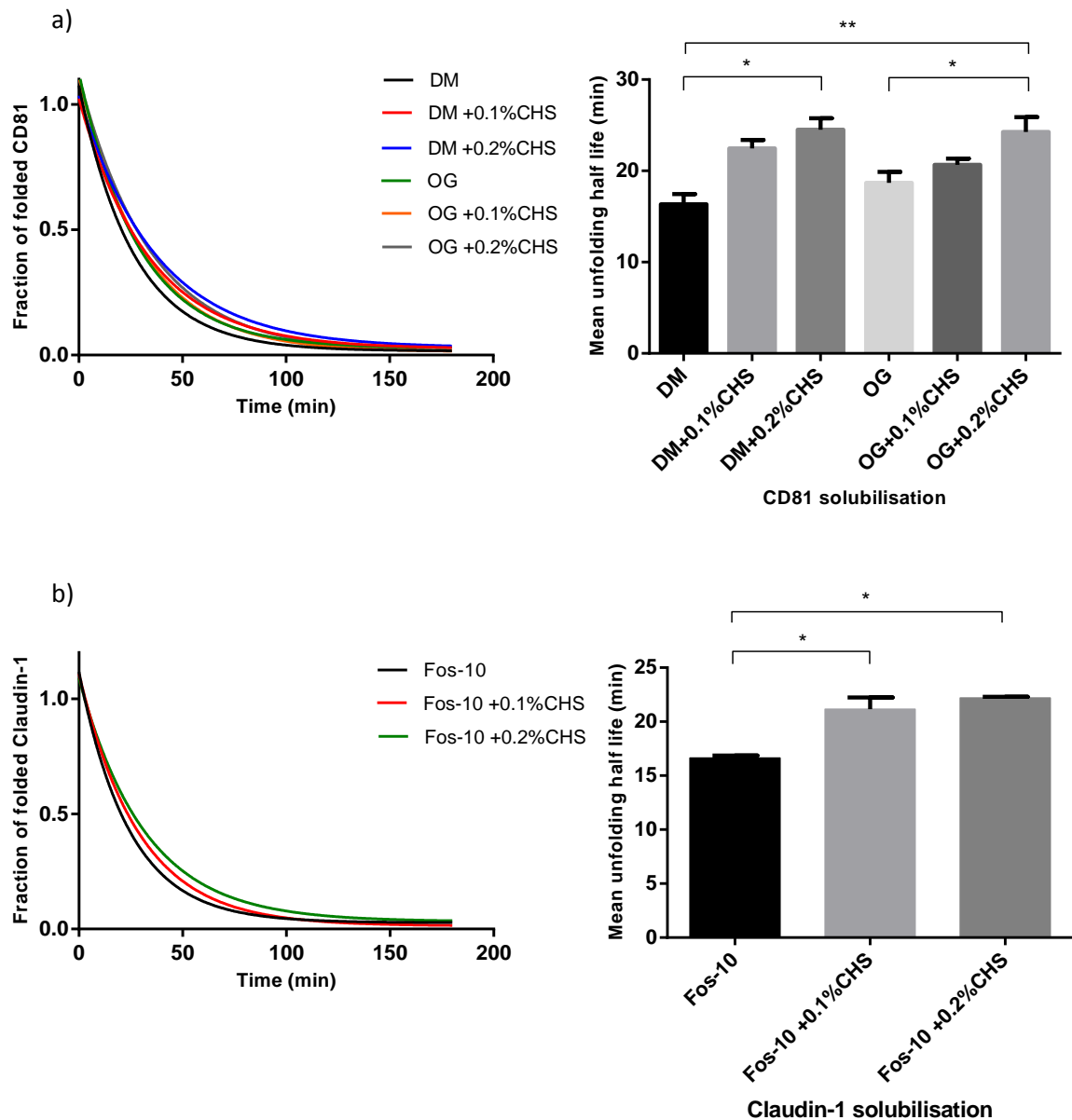
Figure 5.16: Effect of CHS on CD81 and Claudin-1 thermal stability. a) CD81 representative exponential one phase decay curve and the mean unfolding half life in minutes. 6-8µg purified CD81 was incubated with DM or β-OG with/without 0.1% or 0.2% CHS in buffer (20mM Tris, pH 7.5, 150mM NaCl). b) 4 µg purified Claudin-1 incubated in Fos-10 with/without 0.1 or 0.2% CHS in buffer. Solution conditions were incubated in a 96-well plate for 10 min at room temperature prior to adding CPM and incubating at 40°C for 180 min. Using a spectrofluoremeter, measurements were recorded every 5 min and raw data converted to a fraction of folded protein, using the highest fluorescence value reached for that condition as 'maximal unfolded protein'. ANOVA was performed followed by a Dunnett's multiple comparison test, n=5 CD81, n=2 Claudin-1. One batch of CD81 was produced and purified for CPM analysis in collaboration with Dr Justin MacDonald (University of Calgary).

5.4.6 Optimised conditions for CD81 stability

The CPM thermal stability assay was able to predict which solution conditions may be suitable for CD81 and possibly other tetraspanin family members to remain stable and folded in solution. The optimised conditions that came out of the thermal stability assessments are summarised in table 5.1, along with those conditions that were initially used but were unsuccessful due to CD81 aggregation.

The optimised conditions will now be used in purification buffers and tested in subsequent SEC-MALS analysis and crystallisation trials of CD81. The detergent DM could be used for the initial extraction of CD81 from yeast membranes (since it extracts a higher level than β-OG), but then CD81 would be exchanged into β-OG during purification procedures. Also, we have previously noted difficulty in concentrating CD81 in β-OG plus CHS using a 50 kDa cut off concentration device (Sartorius), therefore in future preparations the protein may be concentrated then subsequently mixed with CHS for stability purposes.

Table 5.1: Optimised thermal stability conditions for CD81. The table shows buffer components that were tested in a CPM thermal stability assay. The initial conditions caused CD81 to aggregate in solution. The optimised conditions show solution components that result in the most thermally stable form of CD81 according to a CPM assay and will be tested in protein sample analytical methods such as SEC-MALS and used for CD81 crystallisation trials.

Sample component	Initial conditions	Optimised condition
Detergent	DDM (3x CMC)	OG (3x CMC)
NaCl	150mM	100mM
Glycerol	10%	10%
Buffer and pH	1X PBS, pH 7.4	MOPS, pH 6.5- 7
Lipid content	None	0.2% CHS

6. Discussion

The CD81-Claudin-1 receptor complex is necessary for HCV entry and infection of hepatocytes, but the normal biological function of such complexes remains to be determined. The hetero- and homo- interactions involving all four HCV receptors; CD81, Claudin-1, SR-B1 and Occludin and a high resolution structure of any full length HCV receptor component remain elusive. The focus of this thesis was to extend knowledge on how two of the HCV receptor components, CD81 and Claudin-1, interact *via* protein domains such as extracellular loops and TM domains using full-length membrane proteins. Yeast was used as the experimental host to characterise the role of specific amino acids in the interaction of hetero- and homo-oligomers and also to produce recombinant forms of CD81 and Claudin-1 to move towards understanding the 3D structure of the HCV receptor complex, which may in turn help to direct future structure-aided anti-viral drug design.

A yeast split-ubiquitin assay was optimised to quantify membrane protein-protein interactions of CD81 and Claudin-1 in the absence of other HCV receptor components present in higher eukaryotic cell models. This enabled analysis of the specific role of amino acid residues previously known and those predicted here in CD81 homo- and hetero-oligomerisation. CD81 LEL residues and TM1 residues have been shown in this thesis to play a part in both CD81-Claudin-1 interactions and also in CD81-CD81 interactions, which contribute to an understanding of the complex association of the 4TM membrane proteins with each other. Furthermore, this thesis explored CD81 thermal stability under different solution conditions that may be used in crystallisation attempts, the optimised thermal stability conditions found here may also apply to other tetraspanin family members for which no full length member has had a high resolution structure solved to date. Thermal stability studies showed that cholesterol plays an important role in the structural integrity of both CD81 and Claudin-1 and may contribute in future to obtain the 3D structure of CD81 alone or in complex with Claudin-1. Table 6.1 presents the four main objectives of this thesis, what was achieved and in which direction future work may go from these findings. This chapter will discuss each of the four objectives in turn and draw general conclusions about the main findings of the thesis.

Thesis objectives	Achievements	Main outcomes	Future work
1) Optimisation of a yeast split-ubiquitin Dualmembrane system assay to detect CD81-Claudin-1 heterodimers and CD81 homodimers in yeast membranes.	• Optimal vector selection (Table 3.2/3.3) to determine CD81 homo- and hetero- interactions with Claudin-1 (Figure 3.6), expressed in yeast membranes (Figures 3.9, 3.10). • High throughput measurement of protein interactions using growth reporter genes in a 96-well plate format (Figure 3.12)	• Capability to detect specific CD81-CD81 and CD81-Claudin-1 interactions in the yeast membrane. • Potential to measure protein interactions in a high-throughput manner using yeast growth curve analysis.	• Using a 96-well plate format to measure yeast growth reporter gene activation, the split-ubiquitin assay can be used in a high-throughput manner, including the use of cDNA libraries to screen for novel interaction partners.
2) Analysis of CD81 LEL and TM1 amino acid residues in the full length interaction interface of the CD81-Claudin-1 complex and a comparison of the binding interface of CD81 homo- and hetero-interactions.	• CD81 LEL mutants T149A, E152A and T153A reduce Claudin-1-CD81 interactions as shown by yeast colony counts (figure 4.5) and growth curves (figure 4.6). • CD81 TM1 is involved in Claudin-1- CD81 interactions as residues L14A, F17A and C89A knock-down interactions (figure 4.8). • CD81 homo-oligomerisation is reduced by E152A and T153A in the LEL. In TM1 residue L14A and L14A-F17A reduce CD81-CD81 interactions (Figure 4.10)	• CD81 LEL mutants validate the yeast split-ubiquitin assay, since shown to reduce CD81-Claudin-1 interaction and HCV entry previously (Davis et al., 2012). • CD81 TM1 is involved in CD81-Claudin-1 interactions. • Common residues reduce protein interactions for CD81-Claudin-1 and CD81-CD81 interactions; in CD81 LEL; E152 and T153 and in TM1; L14A (possibly F17A).	• Investigation into CD81 TM1 residues, L14, F17 and C89 regarding their functional involvement in HCV entry into hepatocytes. • Further repeated experiments for CD81-CD81 interactions to determine if F17 and C89 are important in homo-oligomerisation.

3) Prediction of CD81 and Claudin-1 TM residues involved in the interaction as guided by molecular modelling and tested in a split-ubiquitin yeast assay.	• Residues in CD81 TM1 predicted to be on the CD81-Claudin-1 interface are; F21, G25, I28 and L35 (figure 4.11) • CD81 mutants tested in the CD81-Claudin-1 interaction that result in reduced protein interactions are I28A and L35A (figure 4.12 and 4.13).	• Experimental data demonstrate that the full length CD81-Claudin-1 molecular model can be used to predict the TM binding interface.	• The predicted CD81 TM1 residues that knock-down CD81-Claudin-1 could be investigated further regarding function in HCV entry. • Investigation of predicted TM1 residues, found here using a molecular model, in CD81-CD81 heterotypic interactions.
4) Structural characterisation of full length recombinant CD81 by means of improvements to solubilisation, purification and stabilisation procedures.	• Expression, solubilisation and purification of full length CD81 using *P. pastoris* as a recombinant host (figures 5.2-5.6). Thermal stability investigation of CD81 for optimisation of crystallisation trials (figure 5.9 and 5.13- 5.16).	• CD81 was produced to mg quantities in a functional and correctly folded state. • CD81 can be solubilised in many detergents but optimal detergent and buffer conditions were optimised using a thermal stability assay, which shows cholesterol stabilises the structure of CD81. Optimisation steps should lead to future crystallisation success.	• Assessment of optimised CD81 thermal stability conditions will be analysed using SEC-MALS and other techniques to determine the homogeneity and the level of excess detergent compared to in protein-detergent complexes, this will lead to crystallisation trials of CD81.

Table 6.1: Thesis summary. An overview of thesis objectives; how each one was achieved by referring to specific figures, the main outcomes from the achievements and what future work will be done as a result of the findings.

6.1 The yeast split-ubiquitin Dualmembrane system assay can be optimised to detect CD81-Claudin-1 heterodimers and CD81 homodimers in the yeast membrane.

Previous studies that monitored CD81-Claudin-1 interactions used mammalian cells (293-T cells) and FRET analysis (Davis et al., 2012). Yeast systems to study 4TM membrane protein interactions have been less commonly used with no other study assessing CD81-Claudin-1 interactions *in vivo* using yeast. Therefore, optimisation was required to ascertain if specific interactions could be measured and provide reliable data. The orientation of the split ubiquitin fusion partner with CD81 or Claudin-1 N- or C- terminus was important for protein interaction detection. For bait proteins a C-terminal fusion of Cub in combination with a prey protein N-terminal NubG fusion was optimal for both Claudin-1-CD81 interactions and CD81-CD81 interactions (see figure 6.1 for a possible combination). This result may be due to the presence of a *STE2* leader sequence at the 5' end of bait vector *CD81* and *CLDN1* genes, which may have enhanced protein translation. In the absence of a *STE2* leader sequence, other bait C-terminal fusions were unsuccessful in showing a protein-protein interaction (refer to table 3.2/3.3). This was also the case for the prey protein fusion, it was required on the N-terminal end of CD81, if fused to the C-terminus, a protein-protein interaction was not detected (see table 3.2). This work was crucial for a successful assay, since once the correct protein fusions were chosen, the expressed proteins were characterised for specific expression in yeast membranes and showed evidence of specific CD81 homo- and hetero- interactions (see chapter 3).
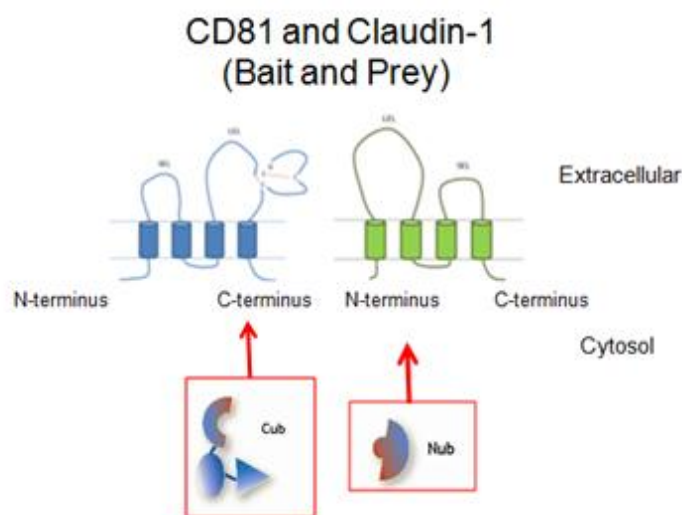
Figure 6.1: Split ubiquitin fusion orientation using 4TM membrane proteins. Multiple combinations of Cub and NubG fusions were tested to determine the correct orientation to detect protein-protein interactions using the split-ubiquitin reporter gene system. Optimal vectors used expressed bait proteins with a 5' end *STE2* leader sequence and the Cub fusion on the C-terminus. Prey proteins (either CD81 or Claudin-1) were optimal when an N-terminal fusion was used for NubG.

An informative feature of the split-ubiquitin assay was to show higher frequency of CD81 homo-oligomers than that of CD81 hetero-oligomers in *vivo* (see chapter 3, figure 3.7). This was demonstrated using a yeast spotting assay, where 10µl yeast culture is grown on an agar plate in serial dilutions. This approach showed that cells co-transformed with CD81-CD81 would grow at higher dilutions than CD81-Claudin-1 transformed cells on plates lacking histidine or adenine, suggesting higher frequency of homo-oligomer protein-protein interactions activated increased expression of *HIS3* and *ADE2* growth reporter genes. Previous studies have shown CD81 to associate with Claudin-1 at a ratio of 2:1 (Bonander et al., 2013) and so it may be that CD81 oligomers are predominant over CD81-Claudin-1 complexes as a minimum of a CD81 dimer is thought to be required for complex formation with Claudin-1 (Harris et al., 2010). This result may also demonstrate the function of tetraspanins as regulators at the membrane, binding to other interaction partners due to breaking away from higher order CD81 oligomers (Hemler, 2005).

Progress was made in chapter 3 regarding the way in which protein-protein interactions were quantitatively measured using reporter genes. Expression of the *lacZ* reporter gene was measured using a β-galactosidase substrate such as X-gal or FDG and both gave

signals for a specific protein-protein interaction by using positive and negative controls to interpret data (figure 3.8). Numerous studies in the literature have used the *lacZ* reporter for previous studies that monitor protein-protein interactions using similar assays as the split-ubiquitin approach shown here (Stagljar et al., 1998), (Hou et al., 2010). Although the *lacZ* reporter is useful, it is advantageous to collect data from intact yeast cells in order to cut down on sample preparation time and on artefacts that may be introduced during cell lysis. The X-gal approach requires lysis of yeast cells for β-galactosidase quantification but the FDG approach uses viable cells, although the substrate (FDG) is costly if used in large amounts. Due to these reasons, a growth curve assay approach was explored to measure OD_{546} yeast over time in a 96-well plate without using any substrates and using the activation of growth reporter genes *HIS3* and *ADE2* for the measurement of protein-protein interactions.

This approach provided real-time protein interaction data over a 24-48 h period, which showed distinct differences between the positive control (using NubI), negative control (using NubG) and the CD81-Claudin-1 test condition (figure 3.10). It was appropriate to use this format to monitor homo- and hetero- interactions using WT and mutant proteins (chapter 4) as well as using the original approach of colony counts on selective agar plates. Previous studies including Toussaint et al., (2006) utilised yeast growth curves on a micro-scale (100µl volume) by assessing WT and DNA repair-deficient yeast strains to the treatment of DNA damaging agents such as ultraviolet light. The study found that generating data from yeast growth curves, in contrast to cell dilution spot test colony formation, was much more sensitive to phenotypic differences found with certain cytotoxic compounds and was more efficient for high-throughput screening (Toussaint et al., 2006). Other studies have used growth curves for real-time monitoring of peptide-protein interactions. For example, Rid et al., (2013) used the yeast two-hybrid system with a peptide library to find novel interaction partners for a vitamin-D receptor.

WT and mutant membrane protein interactions monitored in real-time using growth curves (as demonstrated in chapter 3 and 4 of this thesis) has not been performed in the same way previously in the literature, as far as we are aware. Therefore, this approach may be useful for defining the binding interface or finding novel interaction partners of other membrane proteins. It is important to note that using the DUALmembrane system

to define a protein binding interface may be possible, but a high number of biological replicates may be required to show a statistically significant result, as found in this study. This may be easier to achieve using yeast growth curves on a 96-well plate in contrast to counting yeast colonies on agar plates.

Even though the yeast split ubiquitin method was used successfully here to study the CD81-Claudin-1 interaction, there are limitations to this method which must be considered. A well known limitation of this method includes the high frequency of false positive read-outs for a protein-protein interaction (Sjohamn and Hedfalk, 2014). For this reason, multiple controls must be included to ensure that a genuine interaction pair is being monitored. Alternative techniques that could be used to support data generated using this method include FRET in a mammalian cell system and DLS using recombinant CD81 and Claudin-1.

6.2 CD81 LEL and TM1 amino acid residues are involved in the interaction with Claudin-1 and share common residues with CD81 homotypic interactions.

Previous work by Davis and colleagues (2012) found that CD81 LEL residues T149, E152 and T153 when mutated to alanine reduced CD81 FRET with Claudin-1 and also impaired virus entry. The CD81 mutants had no effect on CD81 interaction with HCV E2 which suggested that the residues play a distinct part in Claudin-1 association and viral entry. The study predicted residues in CD81 LEL would associate with Claudin-1 ECL1 in a region between 63-66 using a molecular model of the loop regions (Davis et al., 2012). Additionally it was apparent that the LEL of CD81 is required to bind HCV E2 and Claudin-1 to result in HCV entry, forming an HCV E2-CD81-Claudin-1 ternary complex. Residues outside T149, E152 and T153 are responsible in the binding of HCV E2; therefore there are distinct functional regions in CD81 LEL (Drummer et al., 2002).

CD81 LEL loop residues (T149, E152 and T153) were investigated in this thesis (chapter 4) to determine CD81-Claudin-1 associations in the yeast split-ubiquitin assay. This was necessary to indicate if differences exist when CD81-Claudin-1 are expressed and associate in a yeast plasma membrane rather than a mammalian host. Furthermore, yeast is a lower eukaryote and does not express other HCV receptor components or

CD81/Claudin-1 homologues so it was interesting to see if this may have an effect on CD81-Claudin-1 associations.

Growth on agar plates and subsequent colony counts revealed that CD81 mutant T149A showed a significant decrease in association with Claudin-1 as compared to WT (figure 4.4), whilst residues E152A and T153A only showed a significant decrease when observed using yeast growth curves (figure 4.5). In both approaches (colony counts and growth curves) it was apparent by mean values that all three CD81 mutants show a knock-down in their interaction with Claudin-1 in comparison to WT CD81, which is in agreement with current literature that used FRET and mammalian cells to measure the heterotypic protein interaction (Davis et al., 2012). These data would suggest that the binding interface of CD81-Claudin-1 can be investigated reliably using a yeast split-ubiquitin assay and site-directed mutagenesis. Furthermore, such findings suggest CD81-Claudin-1 is a highly specific and direct interaction likely to not require additional HCV receptor components or other proteins that are expressed in mammalian cells but are absent in yeast.

Residues in the LEL E152A and T153A also reduced CD81-CD81 interactions, as they did for CD81-Claudin-1 interactions, with T153A showing a significant reduction as compared to WT (figure 4.14). Interestingly, residue T149A did not show the same effect, but was more similar to WT CD81, although the mean value of T149A when data were normalised to internal controls was lower than WT (24.2% and 30.6%, respectively) and so increasing the number of experimental repeats may have shown this mutant to also knock-down CD81-CD81 interactions. Previous studies have suggested that CD81 LEL residues T149 and T153 may be involved in interactions with CD81 SEL, therefore are possibly involved in both homo- and hetero-oligomerisation of CD81 (Yalaoui et al., 2008). It is known that the SEL stabilises the conformation of the LEL, and CD81 expressed without the SEL shows reduced expression levels at the cell surface (Masciopinto et al., 2001). Previous findings may suggest reasons as to why mutant T153A caused a significant reduction in CD81-CD81 interactions, such as a reduction in LEL stability or that it is directly involved in putative homotypic interactions. Previous studies that focused on determination of the CD81 dimerisation binding interface used LEL recombinant soluble forms and found that mutants that affected LEL dimers did not show the same effects in the full length protein, therefore the orientation of the loops is likely to be different in the full-length protein (Drummer

et al., 2005).

Expression analysis of CD81 mutants was performed on yeast total membranes by performing western blot analysis and anti-CD81 ELISA. This data provided evidence of 1) total expression of mutant compared to WT and 2) the folded nature of CD81 and if it could bind specific anti-CD81 antibodies. If a CD81 mutant could bind CD81 antibodies it was likely to be folded and not retained in the ER or Golgi, although further evidence of specific plasma membrane expression would guide the interpretation of mutant CD81-Claudin-1 interaction data as compared to WT CD81. Although, for the CD81-Claudin-1 interaction to be measured successfully using the DUALmembrane system the Cub and Nub moieties must be located on the cytosolic side of the plasma membrane in order to be cleaved by UBPs and for LexA-VP16 to translocate to the nucleus, hence, this signal indicates the location of a given interacting protein pair at the plasma membrane.

After investigating CD81 LEL residues, we then set out to explore the role of TM domains of CD81 and their involvement in the CD81-Claudin-1 and CD81-CD81 interaction. Our initial focus was to look at TM1 and TM2 since various studies had shown specific residues in tetraspanins to either be involved in HCV entry (residues N18 TM1 and C80 TM2, (Bertaux and Dragic, 2006) or form the interface of tetraspanin dimerisation (on CD9, residues L14, F17 and G80, (Kovalenko et al., 2005)). After performing an alignment between CD9 and CD81, residues L14, F17 and C89 were chosen to mutate to alanine and test in both homotypic and heterotypic CD81-Claudin-1 interactions (see section 4.4.1 for a more detailed rationale for picking CD81 TM1 residues).

Results showed that residues L14, F17, C89 have a role in CD81-Claudin-1 complex association, as when mutated to alanine, all significantly reduced the interaction. F17A did not result in a significant reduction, but data showed that the mean was reduced compared to WT and the double mutant L14A-F17A caused a larger knock-down than L14A alone, suggesting that F17A contributed to the reduction (see figure 4.7). These data revealed that there may be a link between tetraspanin dimerisation and tetraspanin-Claudin-1 dimerisation binding interfaces. Kovalenko and colleagues (2005) predicted L14, F17 and G80 to be on the CD9 dimerisation interface and results from this current study suggest the corresponding residues in CD81 are involved in CD81-Claudin-1

binding. This may not only apply to CD81, but also for other tetraspanins that bind Claudin-1, since Claudin-1 has also shown evidence of associating with CD9 (Kovalenko et al., 2007). Further work would be required to show if CD81 and CD9 have a common binding interface with Claudin-1.

To look further into the residues picked from the CD9 dimerisation study, we then tested them in CD81-CD81 interactions. CD81 TM1 mutants L14A and L14A-F17A demonstrated a significant reduction in CD81-oligomerisation as compared to WT CD81. Such findings are therefore in agreement with what was suggested to be involved in a tetraspanin dimerisation interface (Kovalenko et al., 2005). Furthermore, common residues that reduce CD81-Claudin-1 found in this study are responsible for reducing CD81 homotypic interactions. Residues F17A and C89A did not show significantly different results to WT for the CD81-CD81 interaction, which for CD81-Claudin-1 had a more pronounced effect on reducing the interaction. Further experiments would be required to determine the role of F17 and C89 in CD81 dimerisation.

Kovalenko and colleagues (2005) suggest that the tetraspanin dimerisation interface involves conserved TM1 and TM2 residues and evidence points to TM3 and TM4 being involved in a second dimeric interface, contributing to the oligomerisation of tetraspanin dimers. Also predicted from the study was that the TM3/TM4 interface may be involved in heterologous protein-protein interactions as this is where residues are more variable and less conserved throughout the family. Data found here for CD81-Claudin-1 suggests otherwise, since TM1 has a role in the heterotypic interaction; although TM1 mutants either have a direct effect on CD81-Claudin-1 or indirect due to the disruption of CD81-CD81 interactions. The distinction between which mechanism is correct is difficult to conclude from the results shown in chapter 4. Potentially combining specific mutations that are suspected to be on the CD81-Claudin-1 interface then producing recombinant proteins to analyse in a biophysical manner may provide further information on the oligomeric status of the 4TM proteins. In this thesis, we set out to learn more about the mechanism of how the mutants are causing a reduction in CD81-Claudin-1 interactions by using predictions from a molecular model (figure 6.2).

Figure 6.2: CD81-Claudin-1 complex model showing CD81 TM1 residues involved in Claudin-1 interactions. A CD81-Claudin-1 homology based model using CD81 full length homology model (PDB: 2AVZ) by Seigneuret (2006) with a Claudin-1 threading-based model generated and docked using Hex 8.0 by Dr Jonathan Mullins (Swansea University). CD81 TM1 (yellow) shows residues L14, F17, F21, G25, I28 and L35 pointing outwards as a possible interaction interface with Claudin-1 TM2 (residues predicted to be involved are G87 and G91 in orange).

Working in collaboration with Dr Mullins (Swansea University) we mapped the residues L14 and F17 on to a full length CD81 molecular model (Seigneuret, 2006) docked with a Claudin-1 threading-based model (see figure 6.2). The model showed that residues L14 and F17 face the same side of the helix pointing away from the CD81 core. Since L14 is a branched chain bulky hydrophobic residue and F17 an aromatic bulky residue, they form a hydrophobic ridge. This hydrophobic region appears to fit appropriately into a groove on Claudin-1 TM2 (residues G87 and G91) (figure 6.2). The model also suggests that a branched chain residue, L83, on Claudin-1 TM2 may associate with residue G25 on CD81 TM1. The CD81-Claudin-1 complex model (figure 6.2) suggests that the experimental reduction observed using L14A, F17A and L14A-F17A mutants in the split-ubiquitin yeast assay is due to the mutation of CD81 TM1 residues that directly interact with Claudin-1 TM2 residues. Future work may involve testing CD81 TM1 mutants in a HCV entry assay to determine if they have a role in viral entry due to their involvement in CD81-Claudin-1 interactions.

6.2.1 Position of key residues that interact within the extracellular domains in relation to the predicted packing of TM domains of CD81 and Claudin-1.

The recent mouse Claudin-15 crystal structure (Suzuki *et al.,* 2014) revealed the first extracellular loop contained an antiparallel beta sheet comprising of β1-β4 that was able to further interact with β5 located on the second extracellular loop. The beta sheet is further stabilised by a disulphide bridge found between C52 and C62, located on β3 and β4. This disulphide bridge is believed to be conserved across the family as the cysteines in these positions are highly conserved. The crystal structure also reveals that a short extracellular helix (ECH) found on the first extracellular loop prior to TM2 is essential for Claudin-15 homodimerisation (see figure 6.3). The packing of Claudin-15 monomers within the crystal lattice may resemble tight junction strands. M68 found in the ECH of mouse Claudin-15 fits into a hydrophobic pocket in the adjacent Claudin-15 monomer associating with residues found in TM3 and the second extracellular loop (residues F146, F147 and L158). Freeze fracture electron microscopy results shown that upon mutation of M68, F146 and F147 caused a disruption in tight junction strand assembly when mutated to charged or small residues (figure 6.3) (Suzuki *et al.,* 2014).
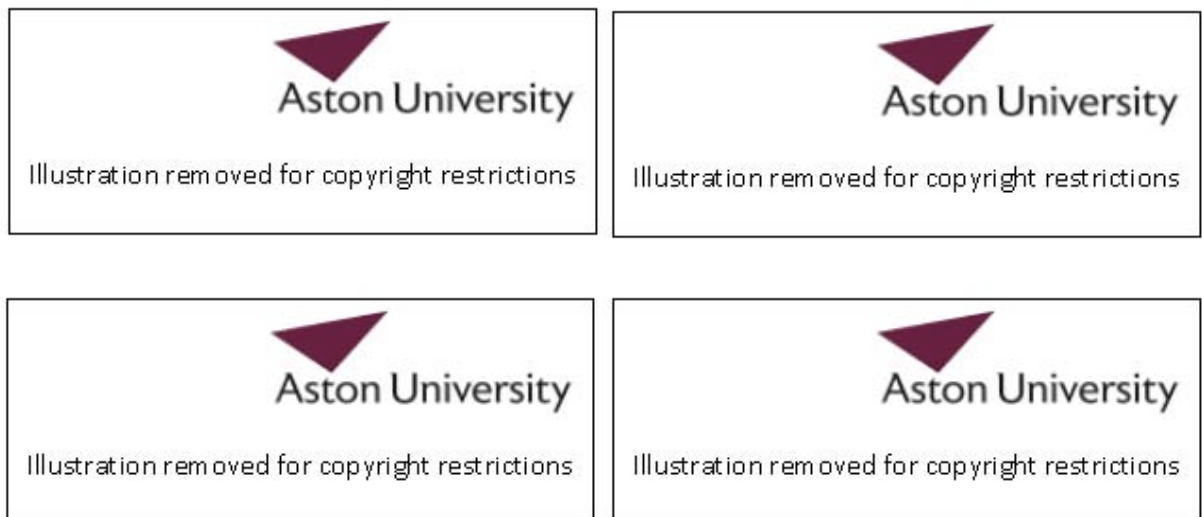
Figure 6.3 Identification of key residues involved in Claudin-15 oligomerisation *via* extracellular domains. A and B) A ribbon representation of a single row of Claudin-15 monomers. C) High resolution view of the extracellular domain interaction site. D and E) Electrostatic potential surface of the interface on each monomer. Taken from Suzuki *et al.,* (2014).

Mouse Claudin-15 and human Claudin-1 have been aligned. There is high conservation between the two family members (95% homology). The alignment shows that M68 in mouse Claudin-15 analogous residue in Claudin-1 is L70.

```
humanClaudin-1      1  MANAGLQLLGFILAFLGWIGAIVSTALPQWRIY   33
mouseClaudin-15     1  MSVA-VETFGFFMSALGLLMLGLTLSNSYWRVS   32


humanClaudin-1     34  SYAGDNIVTAQAMYEGLWMSCVSQSTGQIQCKV   66
mouseClaudin-15    33  TVHG-NVITTNTIFENLWYSCATDSLGVSNCWD   64


humanClaudin-1     67  FDSLLNLSSTLQATRALMVVGILLGVIAIFVAT   99
mouseClaudin-15    65  FPSMLALSGYVQGCRALMITAILLGFLGLFLGM   97


humanClaudin-1    100  VGMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLA  132
mouseClaudin-15    98  VGLRCTNVGNMDLSKKAKLLAIAGTLHILAGAC  130


humanClaudin-1    133  ILVATAWYGNRIVQEFYDPMTPVNARYEFGQAL  165
mouseClaudin-15   131  GMVAISWYAVNITTDFFNPLY-AGTKYELGPAL  162


humanClaudin-1    166  FTGWAAASLCLLGGALLCCSCPRKTTSYPTPRP  198
mouseClaudin-15   163  YLGWSASLLSILGGICVFSTCCCSSKEEPATRA  195


humanClaudin-1    199  YPKPAPS------------------SG-KDYV  211
mouseClaudin-15   196  GLPYKPSTVVIPRATSDESDISFGKYGKNAYV  227
```

Figure 6.4: Alignment between mouse Claudin-15 and human Claudin-1. Red residues show high conservation within alignment based on physicochemical properties of the amino acid, whereas orange residues show less conservation as assigned by T-Coffee. M68 and L70 have been highlighted in white font on the appropriate protein.

The alignment in figure 6.4 showed L70 in Claudin-1 to be homologous to M68 in Claudin-15. Given their similar hydrophobic nature the residues may perform the same functional role in Claudin homodimerisation.

The role of L70 in Claudin-1 with CD81 interactions has not yet been investigated. However, highlighting this residue in our CD81-Claudin-1 full length model poses novel insight in how Claudin-1 and CD81 extracellular domains may interact. L70 on the ECH of Claudin-1 is predicted to be in close proximity to the helix domain found in the LEL of CD81. In particular L70 is interacting with T153 and close to T149 and

E152 (see Fig.6.5 for relative positions). These residues have been highlighted as important for heterodimerisation by Davis *et al.,* (2012) and in this thesis. This mechanism of interaction between the extracellular domains requires further investigation. Initial experiments may focus on L70 mutations to knock out the interaction. It is clear that this interface between CD81-Claudin-1 would competitively compete with Claudin-1-Claudin-1 interactions.

Figure 6.5: Relative position of L70 of Claudin-1 to LEL of CD81. L70 on ECH of Claudin-1 is highlighted as light blue. E152 (magenta), T149 (yellow) and T153 (green) located on the helix of LEL of CD81.

Data presented in this thesis and the information provided by the Claudin-15 structure may suggest that CD81 has two available binding interfaces for CD81 monomers involving TM1 and TM2 in one interface and TM3 and TM4 in the second, as also suggested by Kovalenko and colleagues (2005). Potentially a Claudin-1 monomer may bind an available TM1/TM2 site or displace a bound CD81 monomer for its interaction. This would leave dimeric CD81 bound via TM3/TM4. Evidence that supports this theory was found when recombinant forms of the proteins existed in a 2:1 ratio (CD81: Claudin-1) (Bonander et al., 2013). Further experiments are required to provide evidence of the involvement of two CD81 interaction interfaces in the complex.

6.3 Molecular modelling can predict TM residues involved in CD81-Claudin-1 interaction.

Using the CD81-Claudin-1 full length model (figure 6.2), residues F21, G25, I28 and L35 were picked based on facing in the same direction as L14 and F17 on CD81 TM1 and also based on α-helix periodicity. The selected residues were mutated to alanine and investigated using the yeast split-ubiquitin assay in CD81-Claudin-1 interactions. The results suggested I28A and L35A caused a reduction in CD81-Claudin-1 interactions as compared to WT CD81 with L35A but not I28A showing a significant decrease from WT, although mean growth was reduced for I28A compared to WT CD81. Yeast growth curve analysis in comparison to yeast agar colony counts were able to distinguish knock-down mutants in a more sensitive manner, since the growth curve would shift to the right compared to WT if a mutant had a diminishing effect on the protein interaction. Previous studies that also use yeast growth curves but to monitor the effects of cytotoxic compounds suggested that phenotypes were discovered using growth curves that may not have been using yeast spot tests using colony growth (Toussaint et al., 2006), which is in agreement to that found in this study.

Since the CD81-Claudin-1 model (figure 6.2) predicted residues I28 and L35 to be on the TM interface of the complex, the results described here would suggest that the model can predict TM residues that are important in structural heterotypic interactions (figure 6.2). Taking together, these data suggest that CD81 TM1 has cap regions either end of the helix, L14-F17 at one end and I28 and L35 at the other. Typically, TM cap regions are where specific structural interactions are found. Residues found in the middle of TM1 (F21 and G25) which did not display a knock-down in CD81-Claudin-1 interactions, may be due to non-specific hydrophobic interactions.

CD81 TM1 appears to play a role in CD81-Claudin-1 interactions, but there is also evidence to suggest it plays a role in CD81-CD81 interactions. Future work may investigate the role of TM1 in CD81-CD81 interactions by extending the number of mutant residues tested in the interaction, since only L14A-F17A were investigated here but found to reduce homotypic interactions. Results from analysing CD81 TM1 mutants imply that CD81-Claudin-1 and CD81-CD81 either share common features of their full length interaction interface, including LEL and TM residues or one interaction has a

direct result of the other. According to structural modelling of full length CD81-Claudin-1, using Claudin-15 as template for Claudin-1 (in collaboration with Dr Jonathan Mullins), it was observed that stable arrangements of CD81-Claudin-1 heterodimers were observed when Claudin-1 was superimposed upon either of the CD81 monomers in the CD81 homodimer, suggesting that a common face in CD81 (possibly including some of the residues mutated in this work) may be implicated in homo- and hetero interactions.

6.3.1 CD81-Claudin-1 homology based modelling

Recent availability of a Claudin-15 structure in the Protein Data Bank meant that it was used to produce a more reliable full length homology model for Claudin-1 (rather than using a threading based approach seen in figure 6.2) and subsequent modelling of CD81-Claudin-1 (produced by Dr Jonathan Mullins).Using the evolved homology model, FoldX (a computer algorithm that can predict the effect of an amino acid mutation on the stability of proteins and protein complexes) was used to predict the effect of CD81 TM1 mutations on protein stability and on CD81-Claudin-1 complexes. Each of the CD81 TM1 residues mutated to alanine in this thesis was analysed:

L14A- On the CD81-Claudin-1 homology model L14A was found on the first turn of TM1 facing into the CD81 TM bundle (see figure 6.6). FoldX showed that the mutation has an effect on CD81 monomeric stability because it is involved in a hydrophobic cluster that includes Val 225 (TM4) and L222 (TM4). The mutation of L14 has a minor indirect effect on CD81-Claudin-1 interaction energy because it supports F17, which is capable of forming a direct interaction with Claudin-1. Mutation of L14 may cause a rotamer shift in F17 that could impair the CD81-Claudin-1 interaction interface, which supports findings in chapter 4.

F17A- F17 is found on the second turn of CD81 TM1 facing towards Claudin-1 intracellular loop (figure 6.6). Mutation to alanine has a minor effect on CD81 stability because it supports Val 20 (TM1) and Phe 21 (TM1). F17 has a major effect in CD81-Claudin-1 interaction as its aromatic ring produces a hydrophobic interaction with Claudin-1 intracellular loop M105. Predictions of the role of L14 and F17 support

results found in this study and suggest why a double mutant of L14A-F17A causes a significant effect on the CD81-Claudin-1 interaction.

Figure 6.6: CD81 TM1 L14 and F17 in the CD81-Claudin-1 interaction. A homology based CD81-Claudin-1 model labelled with CD81 TM1 residues L14 (green) and F17 (light blue) and Claudin-1 intracellular loop residue M105 (yellow). Image created using Accelrys, Discovery studio 4.0. Docked homology model produced by Dr Jonathan Mullins.

C89A- Located at the cytosolic region just before CD81 TM3. It is found on the opposite side of CD81 away from the Claudin-1 predicted interface. This could be a second protein binding face for tetraspanins or other partner proteins as suggested previously by Kovalenko et al (2005). This could also be a potential palmitoylation site.

F21A- F21 is located on the third helix turn of CD81 TM1 and faces towards the Claudin-1 TM2 interface (figure 6.7). This residue is predicted to interact directly with F96 on Claudin-1 TM2. However, a large effect was not found experimentally in this study. This may be due to the fact that directly behind F21 is F218 (CD81 TM4) which may compensate for the hydrophobic interaction (e.g. redundancy in the receptor mechanism), as discussed earlier regarding non-specific mid-TM hydrophobic interactions. A future experiment could analyse a double mutation of F21A and F218A to evaluate if there was an effect on the CD81-Claudin-1 interaction.

Figure 6.7: CD81 TM1 F21 and I28 in the CD81-Claudin-1 interaction. A homology based CD81-Claudin-1 model labelled with a) CD81 TM1 residue F21 (green) and Claudin-1 TM2 residue F96 (pink). b) CD81 TM1 I28 and Claudin-1 TM2 V92. Image created using Accelrys, Discovery studio 4.0. Homology model made by Dr Jonathan Mullins.

G25A- This residue is located on CD81 TM1 fourth helix turn and faces towards the CD81 bundle (predicted originally using the threading based Claudin-1 –CD81 model). It has a predicted effect on protein stability because ILE215 (TM4 fourth turn) makes a hydrophobic interaction with G25 backbone. Also M73 located on CD81 TM2 packs against G25 side-chain. This mutation had no effect on CD81-Claudin-1 interactions,

which show the importance of Claudin-1 facing cap residues of CD81 TM1 in our experiments.

I28A- I28 is located on CD81 helix turn five on TM1 and is found in the CD81-Claudin-1 interface (figure 6.7). It is predicted to interact directly to Claudin-1 residue V92. I28A is predicted to impair CD81 stability by the deletion of essential hydrophobic interactions with V31 located on CD81.

L29A- L29 faces into CD81 TM bundle and is essential for stability as it predominately packs against M73. This residue was investigated as it was unclear whether the double glycine motif G25 and G26 could alter the periodicity of the alpha helix TM domain.

L35A- L35 was predicted by FoldX software to have an effect on protein stability. The residue is predicted to make a direct hydrophobic interaction with V85 on Claudin-1 TM2 (see figure 6.8), which supports experimental findings shown in this thesis (chapter 4).

Figure 6.8: CD81 TM1 L35 in the CD81-Claudin-1 interaction. A homology based CD81-Claudin-1 model labelled with CD81 TM1 L35 and Claudin-1 TM2 V85. Image created using Accelrys, Discovery studio 4.0. Homology model made by Dr Jonathan Mullins.

Analysis performed using FoldX has given a certain degree of insight. The CD81-Claudin-1 homology model used for analysis shows that the orientation of the loop domains in regard to TM bundles may differ in reality and when complexed as CD81-Claudin-1. The homology model of Claudin-1 based on the crystal structure of Claudin -15 suggests the loop sits on top of the bundle. CD81 LEL, in the full length homology model, is based on the crystal structure of a soluble dimer which has been critised for

not showing feasibly where the TM domains would be incorporated in reality. However, given that the loops (Claudin-1 ECL1 and CD81 LEL) have previously shown to interact with (Davis et al., 2012) and the results of which have been validated in this thesis, the ECL1 of Claudin-1 must extend for this interaction to take place or CD81 LEL would be in a different orientation in relation to its TM bundle.

The working full length molecular model (figure 6.5) has provided a useful tool to predict the CD81 TM1-Claudin-1 TM2 interface between the two proteins *via* their TM bundles. The empirical data collected shows excellent agreement between residues that face towards the CD81-Claudin-1 heterodimer interface compared to residues that face towards their own TM bundle. For example, the model and FoldX analysis predicted residues L14, F17, F21 and I28 to have an effect (once mutated to alanine) on the CD81-Claudin-1 interaction. Experimental findings agree with L14A, F17A and I28A and show a reduction on the heterotypic interaction. G25 and L29, according to the CD81-Claudin-1 homology model face towards the CD81 TM bundle, this is predicted in FoldX analysis not to have any involvement in the interaction and also experiments in yeast agree. L35A was found experimentally to be important for the heterotypic interaction, but FoldX analysis of the CD81-Claudin-1 homology model did not suggest the same, although when observing the model, L35 looks in very close proximity to V85 on Claudin-1 suggesting limitations with the structural model as it currently stands.

In summary, throughout the duration of this study, much progress has been made. At the beginning, CD81 LEL crystallised dimer provided more structural information than what was available for any Claudin family member. In contrast, at the end of this study, Claudin-1 has a very strong template for molecular modelling since full length Claudin-15 (a close homologue) was determined. Future work in this field will hopefully move towards the structural determination of a full length tetraspanin, be it CD81 or another family member, which is required for more accurate CD81-Claudin-1 modelling and knowledge of how these molecules interact.

6.4 Optimisation of full length recombinant CD81 for future structural studies.

As one of the major HCV receptor components for virus internalisation, CD81 is a key drug target and so knowledge of the full length structure would contribute to the design of antiviral therapy. Furthermore, tetraspanins are involved in a wide range of cellular

processes including their role in human cancer such as metastasis (Hemler, 2014) and so the advantages of solving the structure of a full length tetraspanin to aid its use as a therapeutic target are continually growing in number. To date, there is high resolution structural information available for a CD81 LEL soluble dimer (Kitadokoro et al., 2001) and also low resolution data for the tetraspanin Uroplakin complexes (Min et al., 2006), but thus far the hydrophobic nature of membrane proteins has limited the success of tetraspanin crystallisation to provide high resolution information. Previous studies have shown the production of full length functional and correctly folded CD81 using recombinant host *P. pastoris*, which was solubilised using detergent (Jamshad et al., 2008) and shown to complex with Claudin-1 (Bonander et al., 2013).

This thesis has taken steps used previously to produce, solubilise and purify CD81 but then focused on determination of optimal thermal stability conditions, an approach that has shown to be a pre-requisite for successful crystallisation efforts previously (Alexandrov et al., 2008). Results show that CD81 is amenable to thermal stability assessment using a CPM assay and various conditions such as detergent choice, buffer conditions and lipid addition contribute to the maintenance of a stable purified CD81 sample.

P. pastoris expression of CD81 was performed under tight regulation of the *AOX1* promoter (as shown in figure 5.2), which is consistent with what is known about the promoter controlling the expression of other recombinant protein targets (Cregg et al., 2000). Reproducible solubilisation of functional CD81 (as shown by HCV E2 binding) was then performed using β-OG (figure 5.3) consistent with that shown by a previous study (Jamshad et al., 2008). Scaling up production of CD81 was performed using 2L baffled shake flasks as growth in bioreactors using a minimal medium (BSM) was unsuccessful for unknown reasons. Typically, *P. pastoris* can be grown to high cell densities using a minimal medium (Macauley-Patrick et al., 2005). Even so, the *P. pastoris* X33 strain producing CD81 would only demonstrate growth in a complex medium such as YPG in a bioreactor and so the advantages of using a bioreactor to control elements such as medium composition, pH, dissolved oxygen etc were unsuitable for this particular *P. pastoris* clone, therefore further investigation would be required to determine the reasons for particularly slow growth shown here in minimal media.

Subsequent to large scale production of CD81 in yeast membranes, the material was taken to the MPL at the Diamond Light Source to gain expert advice in attempts to try and crystallise the full length membrane protein. Since using the detergent DDM to extract membrane proteins is relatively common and successful with other membrane proteins, we applied this standard approach to CD81, but CD81 in DDM detergent micelles aggregated upon a concentration procedure, therefore a CPM thermal stability assay was used to determine the optimal detergent for CD81 stability. The results suggested that CD81 had increased stability in shorter chain detergents such as β-OG (8 carbon chain length) in contrast to instability in longer chain detergents (figure 5.9), which explains why DDM (12 carbon chain length) caused CD81 aggregation. DM (10 carbon chain length) was found to provide intermediate CD81 stability using the CPM assay compared to DDM and OG, which suggests a correlation between acyl chain length and protein stability. Interactions of detergents with membrane proteins is protein dependent and factors such as detergent concentration, detergent properties, such as headgroup, length of acyl chain and micelle structure, all contribute to the stability of a given protein used with a particular detergent (Stangl et al., 2012). It is suggested in a previous study that detergents with increasing acyl-chain lengths are usually considered mild, but this is not always the case, as some transmembrane helix oligomers (such as Human glycophorin A) are destabilised in longer chain detergents (Stangl et al., 2012), as was true here regarding CD81. Therefore, this suggests optimisation of detergent conditions is crucial to obtain functional and correctly folded membrane protein for downstream analysis.

Determination of optimal CD81 detergents resulted in purification and concentration of CD81 in DM and OG +CHS at 13 mg/ml and 9 mg/ml, respectively. The ability to concentrate to such high concentrations (for a membrane protein) was further evidence to suggest that both detergents provided suitable stability for CD81 in protein-detergent complexes. Purification of CD81 in DM or OG+CHS showed that monomers and dimers were the predominant species in the samples (figure 5.10). Despite a lack of homogeneity, CD81 was set up for crystallisation trials using a lipid cubic phase approach, but was unsuccessful due to the presence of detergent crystals. Further analysis using SEC-MALS showed that in two different samples analysed (in two different buffers), there was excess detergent in samples, much higher than the level of detergent present in protein-detergent complexes (figure 5.11 and 5.12). Excess

detergent is unsuitable for crystallisation attempts as we had demonstrated with the trial we performed.

Two SEC-MALS analyses were performed on the same protein sample purified in DM. The only difference between the two attempts was the buffer in which CD81 was exchanged to in the analytical column. In a PBS buffer, the predominant peak was a CD81 dimer, which may be due to the high salt concentration pushing the equilibrium to a dimer. Furthermore, a higher order CD81 oligomer was also observed minus detergent, suggesting that the protein was moving towards aggregation on the column. Using a Tris buffer, predominantly monomer and dimers were observed with negligible protein aggregates. This suggests that the buffer conditions are fundamental to the protein state and stability in solution, this prompted further investigation using the CPM assay. Additionally, future work would focus on reducing the level of excess detergent in a CD81 sample but sufficient to maintain the formation of protein-detergent complexes to improve future crystallisation attempts.

6.4.1 CD81 thermal stability studies

A thermal stability CPM assay measures the fraction of folded protein over time at a given temperature. The method has been used previously to optimise conditions to improve membrane protein stability to guide purification and crystallisation efforts (Alexandrov et al., 2008). The assay was used in this thesis to assess the unfolding nature of CD81 over time at 40°C in given buffer conditions. Results suggested that CD81 is most stable (in the conditions tested in this work) in β-OG, 100mM NaCl, 10% glycerol, MOPS pH 6.5-7 and 0.2% CHS. CD81 did not show differences in stability when exposed to different levels of ionic strength (using NaCl); therefore the lowest concentration (100mM) used in the experiments was taken for future experiments. This data is in contrast to what has been found for other membrane proteins such as the human apelin receptor (a GPCR), which was found to be influenced with increasing ionic strength by an increase in protein stability (Alexandrov et al., 2008), suggested to be due to increased surface tension and exclusion of water from the protein interior. Even so, lower NaCl concentrations may be beneficial for downstream CD81 applications as, for example, high salt concentrations may interfere with crystallisation by forming salt crystals.

Glycerol has previously been shown to stabilise protein conformation by shifting the native structure to a more compact state and reduce conformational changes (Tsai et al., 2000), (Vagenende et al., 2009). CD81 was shown here to display increased stability in the presence of 10% glycerol in contrast to 0% glycerol (figure 5.14). Previous studies have explored the mechanism behind glycerol-dependent protein stability and suggest that it interacts with large protein hydrophobic surface regions, thus preventing protein unfolding and aggregation. This mechanism may apply to the effect glycerol has on CD81 and indeed other tetraspanin family members.

We had observed, using SEC-MALS, the buffer used for CD81 can affect the oligomeric state (figures 5.11 and 5.12), but also pH and the buffer composition can have great effects on protein stability. Using the thermal stability assay, we found for CD81, a MOPS buffer at pH 6.5-7 results in the greatest level of protein stability out of those conditions tested (figure 5.15). The theoretical pI value for CD81 is 5.94 and so using buffers between 6.5 -7 would seem appropriate. Careful interpretation of data must be done when pH values lower than 6.0 are tested, since the CPM dye shows decreased reactivity at pH >6 (Alexandrov et al., 2008).

Enhanced stability was also observed in this study when recombinant CD81 was combined with 0.2% CHS (see figure 5.16). This was also the case when Claudin-1 and 0.2% CHS was tested in a CPM thermal stability assay. Cholesterol has shown both functional and structural dependent effects for CD81 and Claudin-1 elsewhere in the literature; therefore data presented in this thesis provides supporting evidence that cholesterol may exert its effect by stabilising the structural integrity of the 4TM proteins. It was previously found that cholesterol is important for HCV entry since depletion of cholesterol at the plasma membrane caused a reduction in CD81-Claudin-1 complexes (Harris et al., 2010). In a different study, on a molecular level, recombinant CD81 and Claudin-1 formed specific complexes that were stabilised in cholesteryl supplemented environments (Bonander et al., 2013). A full length model of CD81 (Seigneuret, 2006) reveals that CD81 may have one or several cholesterol binding sites, due to aromatic residues that that are clustered on one side of the TM domain. Charrin and colleagues (2006) support this further by demonstrating a physical interaction exists between cholesterol and tetraspanins. Taken together, the data presented in this thesis contribute to what is already known by suggesting cholesterol supports the folded

nature of both CD81 and Claudin-1 and this will guide future structural and functional studies of the HCV receptor complex.

This thesis shows the first thermal stability assessment of recombinant full length CD81 and the results suggested may be applicable to other tetraspanin family members. Future work will determine if optimal thermal stability conditions shown here (β-OG (3×CMC), 100mM NaCl, 10% Glycerol, MOPS, pH6.5-7 and 0.2% CHS) will benefit other aspects deemed suitable for crystallisation trials such as homogeneity and the level of excess detergent in the protein-detergent complex sample. Currently, the optimised sample conditions shown here are being analysed for CD81 using SEC-MALS at the MPL. Predicted improvements to follow may include producing a more homogeneous protein sample using SEC, adjustments to the concentration of detergent used to maintain protein solubility or alternatively using a different solubilising agent such as a polymer.

6.5 Overall conclusions

This thesis has demonstrated both technical and molecular advancements in the tetraspanin and Claudin field. Technically, we have the ability to study CD81 homotypic and heterotypic interactions with Claudin-1 in yeast both *in vivo* using *S.cerevisiae* using a split-ubiquitin assay and *in vitro* using recombinant protein produced in *P. pastoris*. The ability to measure membrane protein-protein interactions in the split-ubiquitin assay in real-time to produce yeast growth curves provided a useful tool, which could be used in a high-throughput manner, to explore the role of specific amino acids in soluble and membrane associated domains of CD81. We now have a better understanding of the full length binding interface of CD81-Claudin-1 (involving CD81 LEL and TM1 and predicted Claudin-1 TM2) and CD81-CD81 (involving residues in LEL and TM1). These results have revealed that CD81-Claudin-1 interactions are likely dependent on the CD81 oligomerisation status in the plasma membrane and so using CD81-Claudin-1 as an anti-HCV target has associated complexities since inhibiting CD81-Claudin-1 may have a phenotypic effect due to reducing CD81-CD81 associations.

Since working on full length membrane proteins in this study, through collaboration we have developed a full length molecular model of the interaction to guide mutagenesis and define the protein binding interface *via* TM domains. Towards the end of completing this thesis, the Claudin-15 crystal structure was determined and so enabled us to generate a more reliable model that will guide future structural analysis of the CD81-Claudin-1 co-receptor complex.

To date, a full length tetraspanin structure solved to high resolution does not exist but we are now further towards a novel CD81 structure. Tetraspanin thermal stability has not been demonstrated previously in the literature and so conditions found here that provide increased CD81 stability will contribute towards CD81 crystallisation trials in the future. As a HCV receptor, solving the structure of CD81 will not only benefit anti-viral therapy design but will also advance treatment of other pathologies where CD81 has a role, such as in cancer progression.

7. References

AGNELLO, V., ABEL, G., ELFAHAL, M., KNIGHT, G. B. & ZHANG, Q. X. 1999. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 12766-12771.

AHMAD, M., HIRZ, M., PICHLER, H. & SCHWAB, H. 2014. Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol,* 98**,** 5301-17.

ALEXANDROV, A. I., MILENI, M., CHIEN, E. Y., HANSON, M. A. & STEVENS, R. C. 2008. Microscale fluorescent thermal stability assay for membrane proteins. *Structure,* 16**,** 351-9.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

ASCIONE, A., DE LUCA, M., TARTAGLIONE, M. T., LAMPASI, F., DI COSTANZO, G. G., LANZA, A. G., PICCIOTTO, F. P., MARINO-MARSILIA, G., FONTANELLA, L. & LEANDRO, G. 2010. Peginterferon Alfa-2a Plus Ribavirin Is More Effective Than Peginterferon Alfa-2b Plus Ribavirin for Treating Chronic Hepatitis C Virus Infection. *Gastroenterology,* 138**,** 116-122.

BARTOSCH, B., DUBUISSON, J. & COSSET, F.-L. 2003. Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes. *The Journal of Experimental Medicine,* 197**,** 633-642.

BERTAUX, C. & DRAGIC, T. 2006. Different domains of CD81 mediate distinct stages of hepatitis C virus pseudoparticle entry. *Journal of Virology,* 80**,** 4940-4948.

BILL, R. M., HENDERSON, P. J. F., IWATA, S., KUNJI, E. R. S., MICHEL, H., NEUTZE, R., NEWSTEAD, S., POOLMAN, B., TATE, C. G. & VOGEL, H. 2011. Overcoming barriers to membrane protein structure determination. *Nat Biotech,* 29**,** 335-340.

BLASIG, I. E., WINKLER, L., LASSOWSKI, B., MUELLER, S. L., ZULEGER, N., KRAUSE, E., KRAUSE, G., GAST, K., KOLBE, M. & PIONTEK, J. 2006. On the self-association potential of transmembrane tight junction proteins. *Cell Mol Life Sci,* 63**,** 505-14.

BONANDER, N., JAMSHAD, M., HU, K., FARQUHAR, M. J., STAMATAKI, Z., BALFE, P., MCKEATING, J. A. & BILL, R. M. 2011. Structural characterization of CD81-Claudin-1 hepatitis C virus receptor complexes. *Biochemical Society Transactions,* 39**,** 537-540.

BONANDER, N., JAMSHAD, M., OBERTHUR, D., CLARE, M., BARWELL, J., HU, K., FARQUHAR, M. J., STAMATAKI, Z., HARRIS, H. J., DIERKS, K., DAFFORN, T. R., BETZEL, C., MCKEATING, J. A. & BILL, R. M. 2013. Production, purification and characterization of recombinant, full-length human claudin-1. *PLoS One,* 8**,** e64517.

BRADLEY, D. W., MCCAUSTLAND, K. A., COOK, E. H., SCHABLE, C. A., EBERT, J. W. & MAYNARD, J. E. 1985. Posttransfusion non-A, non-B hepatitis in chimpanzees. Physicochemical evidence that the tubule-forming agent is a small, enveloped virus. *Gastroenterology,* 88**,** 773-779.

CARPENTER, E. P., BEIS, K., CAMERON, A. D. & IWATA, S. 2008. Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol,* 18**,** 581-6.

CASTET, V., FOURNIER, C., SOULIER, A., BRILLET, R., COSTE, J., LARREY, D., DHUMEAUX, D., MAUREL, P. & PAWLOTSKY, J.-M. 2002. Alpha interferon inhibits hepatitis C virus replication in primary human hepatocytes infected in vitro. *Journal of Virology,* 76**,** 8189-8199.

CHAE, P. S., RASMUSSEN, S. G., RANA, R. R., GOTFRYD, K., CHANDRA, R., GOREN, M. A., KRUSE, A. C., NURVA, S., LOLAND, C. J., PIERRE, Y., DREW, D., POPOT, J. L., PICOT, D., FOX, B. G., GUAN, L., GETHER, U., BYRNE, B.,

KOBILKA, B. & GELLMAN, S. H. 2010. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods,* 7**,** 1003-8.

CHARRIN, S., LE NAOUR, F., SILVIE, O., MILHIET, P. E., BOUCHEIX, C. & RUBINSTEIN, E. 2009. Lateral organization of membrane proteins: tetraspanins spin their web. *Biochem J,* 420**,** 133-54.

CHARRIN, S., MANIÉ, S., THIELE, C., BILLARD, M., GERLIER, D., BOUCHEIX, C. & RUBINSTEIN, E. 2003. A physical and functional link between cholesterol and tetraspanins. *European Journal of Immunology,* 33**,** 2479-2489.

CHOO, Q. L., WEINER, A. J., OVERBY, L. R., KUO, G., HOUGHTON, M. & BRADLEY, D. W. 1990. Hepatitis C virus: the major causative agent of viral non-A, non-B hepatitis. *British Medical Bulletin,* 46**,** 423-441.

CREGG, J. M., CEREGHINO, J. L., SHI, J. & HIGGINS, D. R. 2000. Recombinant Protein Expression in Pichia pastoris. *Molecular Biotechnology,* 16**,** 23-52.

DALY, R. & HEARN, M. T. 2005. Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. *J Mol Recognit,* 18**,** 119-38.

DAVIS, C., HARRIS, H. J., HU, K., DRUMMER, H. E., MCKEATING, J. A., MULLINS, J. G. & BALFE, P. 2012. In silico directed mutagenesis identifies the CD81/claudin-1 hepatitis C virus receptor interface. *Cell Microbiol,* 14**,** 1892-903.

DONAHUE, J. G., MUNOZ, A., NESS, P. M., BROWN, D. E., JR., YAWN, D. H., MCALLISTER, H. A., JR., REITZ, B. A. & NELSON, K. E. 1992. The declining risk of post-transfusion hepatitis C virus infection. *N Engl J Med,* 327**,** 369-73.

DOWELL, S. J. & BROWN, A. J. 2009. Yeast assays for G protein-coupled receptors. *Methods Mol Biol,* 552**,** 213-29.

DRUMMER, H. E., MAERZ, A. & POUMBOURIOS, P. 2003. Cell surface expression of functional hepatitis C virus E1 and E2 glycoproteins. *FEBS Letters,* 546**,** 385-390.

DRUMMER, H. E., WILSON, K. A. & POUMBOURIOS, P. 2002. Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. *Journal of Virology,* 76**,** 11143-11147.

DRUMMER, H. E., WILSON, K. A. & POUMBOURIOS, P. 2005. Determinants of CD81 dimerization and interaction with hepatitis C virus glycoprotein E2. *Biochemical and Biophysical Research Communications,* 328**,** 251-257.

DUBUISSON, J. 2007. Hepatitis C virus proteins. *World Journal of Gastroenterology: WJG,* 13**,** 2406-2415.

ELLIS, S. B., BRUST, P. F., KOUTZ, P. J., WATERS, A. F., HARPOLD, M. M. & GINGERAS, T. R. 1985. Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast Pichia pastoris. *Molecular and Cellular Biology,* 5**,** 1111-21.

ESWAR, N., JOHN, B., MIRKOVIC, N., FISER, A., ILYIN, V. A., PIEPER, U., STUART, A. C., MARTI-RENOM, M. A., MADHUSUDHAN, M. S., YERKOVICH, B. & SALI, A. 2003. Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res,* 31**,** 3375-80.

EVANS, M. J., VON HAHN, T., TSCHERNE, D. M., SYDER, A. J., PANIS, M., WÖLK, B., HATZIIOANNOU, T., MCKEATING, J. A., BIENIASZ, P. D. & RICE, C. M. 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature,* 446**,** 801-805.

FERNANDEZ, F. J. & VEGA, M. C. 2013. Technologies to keep an eye on: alternative hosts for protein production in structural biology. *Curr Opin Struct Biol,* 23**,** 365-73.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

FLINT, M., MAIDENS, C., LOOMIS-PRICE, L. D., SHOTTON, C., DUBUISSON, J., MONK, P., HIGGINBOTTOM, A., LEVY, S. & MCKEATING, J. A. 1999. Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. *Journal of Virology,* 73**,** 6235-6244.

FLINT, M., VON HAHN, T., ZHANG, J., FARQUHAR, M., JONES, C. T., BALFE, P., RICE, C. M. & MCKEATING, J. A. 2006. Diverse CD81 proteins support hepatitis C virus infection. *Journal of Virology,* 80**,** 11331-11342.

FOURNIER, C., SUREAU, C., COSTE, J., DUCOS, J., PAGEAUX, G., LARREY, D., DOMERGUE, J. & MAUREL, P. 1998. In vitro infection of adult normal human hepatocytes in primary culture by hepatitis C virus. *The Journal of General Virology,* 79 (Pt 10)**,** 2367-2374.

FREIGASSNER, M., PICHLER, H. & GLIEDER, A. 2009. Tuning microbial hosts for membrane protein production. *Microb Cell Fact,* 8**,** 69.

GISLER, S. M., KITTANAKOM, S., FUSTER, D., WONG, V., BERTIC, M., RADANOVIC, T., HALL, R. A., MURER, H., BIBER, J., MARKOVICH, D., MOE, O. W. & STAGLJAR, I. 2008. Monitoring protein-protein interactions between the mammalian integral membrane transporters and PDZ-interacting partners using a modified split-ubiquitin membrane yeast two-hybrid system. *Mol Cell Proteomics,* 7**,** 1362-77.

GRUNBERG, R., NILGES, M. & LECKNER, J. 2007. Biskit--a software platform for structural bioinformatics. *Bioinformatics,* 23**,** 769-70.

GUNZEL, D. & FROMM, M. 2012. Claudins and other tight junction proteins. *Compr Physiol,* 2**,** 1819-52.

GUTMANN, D. A., MIZOHATA, E., NEWSTEAD, S., FERRANDON, S., POSTIS, V., XIA, X., HENDERSON, P. J., VAN VEEN, H. W. & BYRNE, B. 2007. A high-throughput method for membrane protein solubility screening: the ultracentrifugation dispersity sedimentation assay. *Protein Science : A Publication of the Protein Society,* 16**,** 1422-8.

HARRIS, H. J., DAVIS, C., MULLINS, J. G. L., HU, K., GOODALL, M., FARQUHAR, M. J., MEE, C. J., MCCAFFREY, K., YOUNG, S., DRUMMER, H., BALFE, P. & MCKEATING, J. A. 2010. Claudin Association with CD81 Defines Hepatitis C Virus Entry. 285**,** 21092-21102.

HARRIS, H. J., FARQUHAR, M. J., MEE, C. J., DAVIS, C., REYNOLDS, G. M., JENNINGS, A., HU, K., YUAN, F., DENG, H., HUBSCHER, S. G., HAN, J. H., BALFE, P. & MCKEATING, J. A. 2008. CD81 and Claudin 1 Coreceptor Association: Role in Hepatitis C Virus Entry. *Journal of Virology,* 82**,** 5007-5020.

HEMLER, M. E. 2003. Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. *Annual Review of Cell and Developmental Biology,* 19**,** 397-422.

HEMLER, M. E. 2005. Tetraspanin functions and associated microdomains. *Nat Rev Mol Cell Biol,* 6**,** 801-811.

HEMLER, M. E. 2014. Tetraspanin proteins promote multiple cancer stages. *Nat Rev Cancer,* 14**,** 49-60.

HILL, A. & COOKE, G. 2014. Medicine. Hepatitis C can be cured globally, but at what cost? *Science (New York, N.Y.),* 345**,** 141-2.

HOU, J., RENIGUNTA, A., YANG, J. & WALDEGGER, S. 2010. Claudin-4 forms paracellular chloride channel in the kidney and requires claudin-8 for tight junction localization. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 18010-5.

HSU, M., ZHANG, J., FLINT, M., LOGVINOFF, C., CHENG-MAYER, C., RICE, C. M. & MCKEATING, J. A. 2003. Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 7271-6.

JAMSHAD, M., LIN, Y. P., KNOWLES, T. J., PARSLOW, R. A., HARRIS, C., WHEATLEY, M., POYNER, D. R., BILL, R. M., THOMAS, O. R., OVERDUIN, M. & DAFFORN, T. R. 2011. Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochem Soc Trans,* 39**,** 813-8.

JAMSHAD, M., RAJESH, S., STAMATAKI, Z., MCKEATING, J. A., DAFFORN, T., OVERDUIN, M. & BILL, R. M. 2008. Structural characterization of recombinant human CD81 produced in Pichia pastoris. *Protein Expression and Purification,* 57**,** 206-216.

JOHNSSON, N. & VARSHAVSKY, A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 10340-4.

KALIPATNAPU, S. & CHATTOPADHYAY, A. 2005. Membrane protein solubilization: recent advances and challenges in solubilization of serotonin1A receptors. *IUBMB Life,* 57**,** 505-12.

KANG, H. J., LEE, C. & DREW, D. 2013. Breaking the barriers in membrane protein crystallography. *Int J Biochem Cell Biol,* 45**,** 636-44.

KITADOKORO, K., BORDO, D., GALLI, G., PETRACCA, R., FALUGI, F., ABRIGNANI, S., GRANDI, G. & BOLOGNESI, M. 2001. CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *The EMBO Journal,* 20**,** 12-18.

KITADOKORO, K., PONASSI, M., GALLI, G., PETRACCA, R., FALUGI, F., GRANDI, G. & BOLOGNESI, M. 2002. Subunit association and conformational flexibility in the head subdomain of human CD81 large extracellular loop. *Biological Chemistry,* 383**,** 1447-1452.

KNOWLES, T. J., FINKA, R., SMITH, C., LIN, Y. P., DAFFORN, T. & OVERDUIN, M. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc,* 131**,** 7484-5.

KOVALENKO, O. V., METCALF, D. G., DEGRADO, W. F. & HEMLER, M. E. 2005. Structural organization and interactions of transmembrane domains in tetraspanin proteins. *BMC Struct Biol,* 5**,** 11.

KOVALENKO, O. V., YANG, X. H. & HEMLER, M. E. 2007. A novel cysteine cross-linking method reveals a direct association between claudin-1 and tetraspanin CD9. *Mol Cell Proteomics,* 6**,** 1855-67.

LAVILLETTE, D., TARR, A. W., VOISSET, C., DONOT, P., BARTOSCH, B., BAIN, C., PATEL, A. H., DUBUISSON, J., BALL, J. K. & COSSET, F.-L. 2005. Characterization of host-range and cell entry properties of the major genotypes and subtypes of hepatitis C virus. *Hepatology (Baltimore, Md.),* 41**,** 265-274.

LEVY, S. & SHOHAM, T. 2005. The tetraspanin web modulates immune-signalling complexes. *Nature Reviews. Immunology,* 5**,** 136-148.

LEVY, S., TODD, S. C. & MAECKER, H. T. 1998. CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. *Annual Review of Immunology,* 16**,** 89-109.

LINDENBACH, B. D. & RICE, C. M. 2013. The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol,* 11**,** 688-700.

MACAULEY-PATRICK, S., FAZENDA, M. L., MCNEIL, B. & HARVEY, L. M. 2005. Heterologous protein production using the Pichia pastoris expression system. *Yeast,* 22**,** 249-270.

MARTINO, B. & GUIDO, G. 2002. Modulation of tetraspanin function. Google Patents.

MASCIOPINTO, F., CAMPAGNOLI, S., ABRIGNANI, S., UEMATSU, Y. & PILERI, P. 2001. The small extracellular loop of CD81 is necessary for optimal surface expression of the large loop, a putative HCV receptor. *Virus Res,* 80**,** 1-10.

MATTANOVICH, D., BRANDUARDI, P., DATO, L., GASSER, B., SAUER, M. & PORRO, D. 2012. Recombinant protein production in yeasts. *Methods Mol Biol,* 824**,** 329-58.

MCHUTCHISON, J. G., BARTENSCHLAGER, R., PATEL, K. & PAWLOTSKY, J.-M. 2006. The face of future hepatitis C antiviral drug development: recent biological and virologic advances and their translation to drug development and clinical practice. *Journal of Hepatology,* 44**,** 411-421.

MEERTENS, L., BERTAUX, C., CUKIERMAN, L., CORMIER, E., LAVILLETTE, D., COSSET, F.-L. & DRAGIC, T. 2008. The Tight Junction Proteins Claudin-1, -6, and -9 Are Entry Cofactors for Hepatitis C Virus. *Journal of Virology,* 82**,** 3555-3560.

MIN, G., WANG, H., SUN, T.-T. & KONG, X.-P. 2006. Structural basis for tetraspanin functions as revealed by the cryo-EM structure of uroplakin complexes at 6-A resolution. *The Journal of Cell Biology,* 173**,** 975-983.

MORAES, I., EVANS, G., SANCHEZ-WEATHERBY, J., NEWSTEAD, S. & STEWART, P. D. 2014. Membrane protein structure determination - the next generation. *Biochim Biophys Acta,* 1838**,** 78-87.

MULLINS, J. G. 2012. Structural modelling pipelines in next generation sequencing projects. *Adv Protein Chem Struct Biol,* 89**,** 117-67.

MULLINS, J. G., PARKER, J. E., COOLS, H. J., TOGAWA, R. C., LUCAS, J. A., FRAAIJE, B. A., KELLY, D. E. & KELLY, S. L. 2011. Molecular modelling of the emergence of azole resistance in Mycosphaerella graminicola. *PLoS One,* 6**,** e20973.

NOTREDAME, C., HIGGINS, D. G. & HERINGA, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol,* 302**,** 205-17.

OCHIAI, A., YAMASAKI, M., MIKAMI, B., HASHIMOTO, W. & MURATA, K. 2010. Crystal structure of exotype alginate lyase Atu3025 from Agrobacterium tumefaciens. *J Biol Chem,* 285**,** 24519-28.

OLIVER, R. C., LIPFERT, J., FOX, D. A., LO, R. H., DONIACH, S. & COLUMBUS, L. 2013. Dependence of micelle size and shape on detergent alkyl chain length and head group. *PLoS One,* 8**,** e62488.

PILERI, P., UEMATSU, Y., CAMPAGNOLI, S., GALLI, G., FALUGI, F., PETRACCA, R., WEINER, A. J., HOUGHTON, M., ROSA, D., GRANDI, G. & ABRIGNANI, S. 1998. Binding of hepatitis C virus to CD81. *Science (New York, N.Y.),* 282**,** 938-941.

PRIVE, G. G. 2007. Detergents for the stabilization and crystallization of membrane proteins. *Methods,* 41**,** 388-97.

RAJESH, S., KNOWLES, T. & OVERDUIN, M. 2011. Production of membrane proteins without cells or detergents. *N Biotechnol,* 28**,** 250-4.

RAJESH, S., SRIDHAR, P., TEWS, B. A., FENEANT, L., COCQUEREL, L., WARD, D. G., BERDITCHEVSKI, F. & OVERDUIN, M. 2012. Structural basis of ligand interactions of the large extracellular domain of tetraspanin CD81. *Journal of Virology,* 86**,** 9606-16.

REYNOLDS, G. M., HARRIS, H. J., JENNINGS, A., HU, K., GROVE, J., LALOR, P. F., ADAMS, D. H., BALFE, P., HÜBSCHER, S. G. & MCKEATING, J. A. 2008. Hepatitis C virus receptor expression in normal and diseased liver tissue. *Hepatology,* 47**,** 418-427.

RITCHIE, D. W. & VENKATRAMAN, V. 2010. Ultra-fast FFT protein docking on graphics processors. *Bioinformatics,* 26**,** 2398-405.

ROSA, D., CAMPAGNOLI, S., MORETTO, C., GUENZI, E., COUSENS, L., CHIN, M., DONG, C., WEINER, A. J., LAU, J. Y., CHOO, Q. L., CHIEN, D., PILERI, P., HOUGHTON, M. & ABRIGNANI, S. 1996. A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 1759-1763.

ROY, A., KUCUKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc,* 5**,** 725-38.

RUBINSTEIN, E. 2011. The complexity of tetraspanins. *Biochem Soc Trans,* 39**,** 501-5.

SCARSELLI, E., ANSUINI, H., CERINO, R., ROCCASECCA, R. M., ACALI, S., FILOCAMO, G., TRABONI, C., NICOSIA, A., CORTESE, R. & VITELLI, A. 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *The EMBO Journal,* 21**,** 5017-5025.

SEIGNEURET, M. 2006. Complete predicted three-dimensional structure of the facilitator transmembrane protein and hepatitis C virus receptor CD81: conserved and variable structural domains in the tetraspanin superfamily. *Biophys J,* 90**,** 212-27.

SEIGNEURET, M., DELAGUILLAUMIE, A., LAGAUDRIÈRE-GESBERT, C. & CONJEAUD, H. 2001. Structure of the tetraspanin main extracellular domain. A partially conserved fold with a structurally variable domain insertion. *The Journal of Biological Chemistry,* 276**,** 40055-40064.

SELBY, M. J., GLAZER, E., MASIARZ, F. & HOUGHTON, M. 1994. Complex processing and protein:protein interactions in the E2:NS2 region of HCV. *Virology,* 204**,** 114-122.

SERRANO-VEGA, M. J., MAGNANI, F., SHIBATA, Y. & TATE, C. G. 2008. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 877-82.

SHIROISHI, M., KOBAYASHI, T., OGASAWARA, S., TSUJIMOTO, H., IKEDA-SUNO, C., IWATA, S. & SHIMAMURA, T. 2011. Production of the stable human histamine H(1) receptor in Pichia pastoris for structural determination. *Methods (San Diego, Calif.).*

SJOHAMN, J. & HEDFALK, K. 2014. Unraveling aquaporin interaction partners. *Biochim Biophys Acta,* 1840**,** 1614-23.

SONODA, Y., NEWSTEAD, S., HU, N. J., ALGUEL, Y., NJI, E., BEIS, K., YASHIRO, S., LEE, C., LEUNG, J., CAMERON, A. D., BYRNE, B., IWATA, S. & DREW, D. 2011. Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure,* 19**,** 17-25.

STAGLJAR, I., KOROSTENSKY, C., JOHNSSON, N. & TE HEESEN, S. 1998. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 5187-92.

STANGL, M., VEERAPPAN, A., KROEGER, A., VOGEL, P. & SCHNEIDER, D. 2012. Detergent properties influence the stability of the glycophorin A transmembrane helix dimer in lysophosphatidylcholine micelles. *Biophys J,* 103**,** 2455-64.

STENKAMP, R. E., TELLER, D. C. & PALCZEWSKI, K. 2002. Crystal structure of rhodopsin: a G-protein-coupled receptor. *Chembiochem,* 3**,** 963-7.

STIPP, C. S., KOLESNIKOVA, T. V. & HEMLER, M. E. 2003. Functional domains in tetraspanin proteins. *Trends Biochem Sci,* 28**,** 106-12.

SUZUKI, H., NISHIZAWA, T., TANI, K., YAMAZAKI, Y., TAMURA, A., ISHITANI, R., DOHMAE, N., TSUKITA, S., NUREKI, O. & FUJIYOSHI, Y. 2014. Crystal structure of a claudin provides insight into the architecture of tight junctions. *Science (New York, N.Y.),* 344**,** 304-7.

TAKAYAMA, H., CHELIKANI, P., REEVES, P. J., ZHANG, S. & KHORANA, H. G. 2008. High-level expression, single-step immunoaffinity purification and characterization of human tetraspanin membrane protein CD81. *PLoS One,* 3**,** e2314.

TARRY, M., SKAAR, K., HEIJNE, G., DRAHEIM, R. R. & HOGBOM, M. 2012. Production of human tetraspanin proteins in Escherichia coli. *Protein Expr Purif,* 82**,** 373-9.

THAMINY, S., AUERBACH, D., ARNOLDO, A. & STAGLJAR, I. 2003. Identification of novel ErbB3-interacting factors using the split-ubiquitin membrane yeast two-hybrid system. *Genome Res,* 13**,** 1744-53.

TIMPE, J. M. & MCKEATING, J. A. 2008. Hepatitis C virus entry: possible targets for therapy. *Gut,* 57**,** 1728-1737.

TOUSSAINT, M., LEVASSEUR, G., GERVAIS-BIRD, J., WELLINGER, R. J., ELELA, S. A. & CONCONI, A. 2006. A high-throughput method to measure the sensitivity of yeast cells to genotoxic agents in liquid cultures. *Mutat Res,* 606**,** 92-105.

TSAI, A. M., NEUMANN, D. A. & BELL, L. N. 2000. Molecular dynamics of solid-state lysozyme as affected by glycerol and water: a neutron scattering study. *Biophys J,* 79**,** 2728-32.

TSUKITA, S. & FURUSE, M. 1998. Overcoming barriers in the study of tight junction functions: from occludin to claudin. *Genes Cells,* 3**,** 569-73.

VAGENENDE, V., YAP, M. G. & TROUT, B. L. 2009. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry,* 48**,** 11084-96.

VAN DURME, J., DELGADO, J., STRICHER, F., SERRANO, L., SCHYMKOWITZ, J. & ROUSSEAU, F. 2011. A graphical interface for the FoldX forcefield. *Bioinformatics,* 27**,** 1711-2.

VAN ITALLIE, C. M. & ANDERSON, J. M. 2013. Claudin interactions in and out of the tight junction. *Tissue Barriers,* 1**,** e25247.

WAKITA, T., PIETSCHMANN, T., KATO, T., DATE, T., MIYAMOTO, M., ZHAO, Z., MURTHY, K., HABERMANN, A., KRÄUSSLICH, H.-G., MIZOKAMI, M.,

BARTENSCHLAGER, R. & LIANG, T. J. 2005. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature Medicine,* 11**,** 791-796.

WHEELER, D. L., BARRETT, T., BENSON, D. A., BRYANT, S. H., CANESE, K., CHETVERNIN, V., CHURCH, D. M., DICUCCIO, M., EDGAR, R., FEDERHEN, S., FEOLO, M., GEER, L. Y., HELMBERG, W., KAPUSTIN, Y., KHOVAYKO, O., LANDSMAN, D., LIPMAN, D. J., MADDEN, T. L., MAGLOTT, D. R., MILLER, V., OSTELL, J., PRUITT, K. D., SCHULER, G. D., SHUMWAY, M., SEQUEIRA, E., SHERRY, S. T., SIROTKIN, K., SOUVOROV, A., STARCHENKO, G., TATUSOV, R. L., TATUSOVA, T. A., WAGNER, L. & YASCHENKO, E. 2008. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 36**,** D13-21.

XIAO, F., FOFANA, I., THUMANN, C., MAILLY, L., ALLES, R., ROBINET, E., MEYER, N., SCHAEFFER, M., HABERSETZER, F., DOFFOEL, M., LEYSSEN, P., NEYTS, J., ZEISEL, M. B. & BAUMERT, T. F. 2014. Synergy of entry inhibitors with direct-acting antivirals uncovers novel combinations for prevention and treatment of hepatitis C. *Gut.*

YALAOUI, S., ZOUGBEDE, S., CHARRIN, S., SILVIE, O., ARDUISE, C., FARHATI, K., BOUCHEIX, C., MAZIER, D., RUBINSTEIN, E. & FROISSARD, P. 2008. Hepatocyte permissiveness to Plasmodium infection is conveyed by a short and structurally conserved region of the CD81 large extracellular domain. *PLoS Pathog,* 4**,** e1000010.

ZEISEL, M. B., FOFANA, I., FAFI-KREMER, S. & BAUMERT, T. F. 2011. Hepatitis C virus entry into hepatocytes: molecular mechanisms and targets for antiviral therapies. *Journal of Hepatology,* 54**,** 566-576.

ZHANG, J., RANDALL, G., HIGGINBOTTOM, A., MONK, P., RICE, C. M. & MCKEATING, J. A. 2004. CD81 is required for hepatitis C virus glycoprotein-mediated viral infection. *Journal of Virology,* 78**,** 1448-1455.

ZHENG, A., YUAN, F., LI, Y., ZHU, F., HOU, P., LI, J., SONG, X., DING, M. & DENG, H. 2007. Claudin-6 and Claudin-9 Function as Additional Coreceptors for Hepatitis C Virus. *Journal of Virology,* 81**,** 12465-12471.

ZHONG, J., GASTAMINZA, P., CHENG, G., KAPADIA, S., KATO, T., BURTON, D. R., WIELAND, S. F., UPRICHARD, S. L., WAKITA, T. & CHISARI, F. V. 2005. Robust hepatitis C virus infection in vitro. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 9294-9299.

ZONA, L., TAWAR, R. G., ZEISEL, M. B., XIAO, F., SCHUSTER, C., LUPBERGER, J. & BAUMERT, T. F. 2014. CD81-receptor associations--impact for hepatitis C virus entry and antiviral therapies. *Viruses,* 6**,** 875-92.

ZUCKERMAN, E., YESHURUN, D. & ROSNER, I. 2001. Management of hepatitis C virus-related arthritis. *BioDrugs: Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy,* 15**,** 573-584.