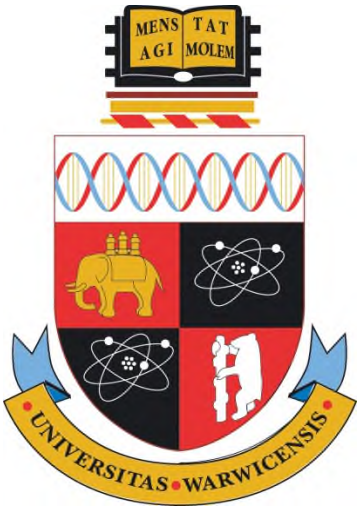


Investigation and Characterisation of Transmembrane Domain Interactions in Carnitine Palmitoyltransferase 1 Isoforms

Leo Bowsher

A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy.

Department of Chemistry

University of Warwick

September 2015

Table of Contents

List of Figures

List of Tables

Abbreviations

2D	Two dimensional
ANOVA	Analysis of variance
CAT	Carnitine acyltransferases
CD	Circular dichroism
CHI	CNS searching of helical interactions
CMC	Critical micelle concentration
CNS	Crystallography and NMR system
CoA	Coenzyme A
COT	Carnitine octanoyltransferase
CPT1	Carnitine palmitoyltransferase 1
CPT2	Carnitine palmitoyltransferase 2
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DPC	Dodecylphosphocholine
EDTA	Ethylenediaminetetraacetic acid
FID	Free induction decay
GpA	Glycophorin A
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Lysogeny Broth
MALDI-TOF	Matrix assisted linear desorption ionisation - time of flight
MBP	Maltose binding protein
MD	Molecular dynamics

MWCO	Molecular weight cut off
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
OD	Optical density
OMM	Outer mitochondrial membrane
ONP	Ortho-nitrophenol
ONPG	Ortho-nitrophenyl-β-galactoside
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PSI	Pounds per square inch
RF	Radio Frequency
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TAE	Tris acetate EDTA
TBST	Tris buffered saline tween
TCA	Trichloroacetic acid
TM	Transmembrane
TOCSY	Total correlation spectroscopy
YPD	Yeast peptone d-glucose

Acknowledgements

I would like to thank both of my supervisors, Dr. Ann Dixon, and Prof. Victor Zammit, for the opportunity to pursue this interesting project. Their guidance throughout the project has been vital to the successful completion of this research.

I thank Dr. Ivan Prokes for NMR training and providing me with spectrometer time, Dr. Nikola Chmel for help with HPLC, Dr. Lijiang Song and Philip Aston for help with mass spectrometry, and Dr. Claudia Blindauer and Dr. Manuela Tosin for their suggestions and input acting on my advisory panel.

My thanks go to all the members of the Dixon group both past and present. Their friendship, help and interesting discussions have made the lab and office a great place to work. Special thanks go to Gemma Warren for her initial training, Mike Chow for too many suggestions to count, and Fay Probert, Anthony Nash, and Christine Lockey for their friendship and support. I would also like to thank Dhadchi Jeyaharan and Muhammad Hasan for their camaraderie throughout our PhD projects together.

Lastly I would like to thank my parents who have supported me without hesitation throughout my time at Warwick, and my girlfriend Cordelia Rampley for her love and encouragement through all the ups and downs.

Declaration

The work presented in this thesis is original, and was conducted by the author unless otherwise stated. This work was performed under the supervision of Dr. Ann Dixon and Prof. Victor Zammit, and has not been presented previously for another degree at any institution.

The work was funded by the BBSRC.

All sources of information have been acknowledged by means of reference.

Summary

Carnitine Palmitoyltransferase 1 (CPT1) is an outer mitochondrial membrane protein that plays an important role in the β-oxidation of long chain fatty acids by regulating their entry into the mitochondrial matrix. This regulatory property is due to its inhibition by malonyl-CoA. Three isoforms of CPT1 have been identified: CPT1A (which is abundant in the liver); CPT1B (which occurs in heart and skeletal muscles and other highly oxidative tissues e.g. brown adipose) and a brain isoform (CPT1C). CPT1A and CPT1B have a high degree of similarity in primary sequence, however CPT1B has 30-100 fold greater sensitivity to malonyl-CoA. CPT1A and CPT1B are both predicted to have two transmembrane (TM) domains which are thought to interact both inter- and intramolecularly, and to sense the membrane environment in which they occur. These interactions are thought to be important for the structure and the function of CPT1. The potential for modulating the function of the three different CPT1 isoforms is of therapeutic interest for the treatment of many metabolic disorders, most notably diabetes.

The main focus of this research was to systematically study the TM domains of CPT1A and CPT1B to characterise the homotypic and heterotypic interactions that occur between them. Using an *in vivo* assay designed to measure interactions between TM domains in the *E. coli* inner membrane, it has been shown that the TM domains in CPT1A and CPT1B interact, in both a homotypic and a heterotypic fashion, to a relatively strong degree. Using mutagenesis, several important residues have been identified in the stabilisation of homotypic and heterotypic interactions in the TM domains of CPT1A and CPT1B in this manner. The residues identified illustrate some key differences and similarities between the interaction profiles of these two isoforms. The CPT1B TM domains were also expressed and purified as peptides and *in vitro* biophysical experiments were used to corroborate interaction data in the full TM domains of CPT1B.

1. Introduction

1.1 Membrane Proteins

Membranes in biological systems allow the separation of cellular processes and reactions from the rest of the cell or organism and are a key factor in the evolution of more complex eukaryotic organisms. For example, membranes allow cells to control the chemical conditions in different compartments to optimise processes and to keep proteolytic enzymes away from unintended targets (Fuerst, 2005). This separation is important, but the products of these reactions are often required in other organelles, cells or tissues within an organism. It is membrane proteins that regulate the transport of molecules across membranes allowing them to reach their intended targets. The selective transport of substances across membranes is vital to a great deal of cellular functions, such as secretion, cell signalling and many metabolic processes (Filmore, 2004). Due to their central role in many fundamental biological processes, membrane proteins are a prevalent target for therapeutics. It has been estimated that ~30% of human proteins are membrane proteins but greater than 50% of therapeutic drugs have a membrane protein as their target (Overington, et al., 2006).

There are two main types of membrane protein: peripheral membrane proteins which interact and bind to a component of the membrane; and integral membrane proteins which are permanently embedded in a membrane. The integral membrane proteins can be broken down into two further categories: monotopic, where they are embedded or attached to just one leaflet of the bilayer; and bitopic where part of the protein (known as the transmembrane (TM) domain) completely

spans the bilayer so there are parts of the protein exposed on both aspects of the membrane. It is the integral membrane proteins which pose the greatest challenges for purification and crystallisation. These bitopic transmembrane proteins occur with two main forms of membrane spanning domain: those with an α helical structure and those which employ a β barrel structure (**Figure 1.1**).

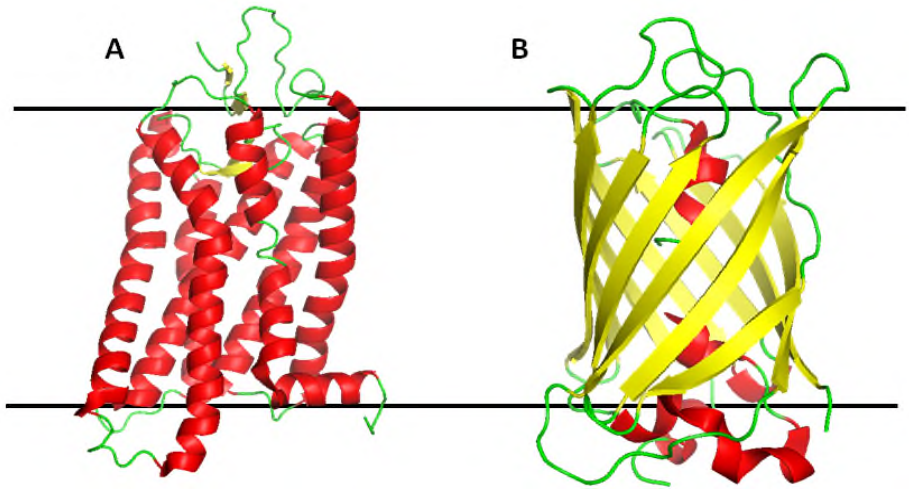


Figure 1.1 – Representative structures of the two main types of transmembrane protein. (**A)** The α helical (red) protein, bovine rhodopsin (PDB file: 1GZM), and (**B)** the beta barrel (yellow) of GFP (PDB file: 1RRX).

1.1.1 Structure Elucidation of Membrane Proteins

Membrane proteins are clearly of great importance, but they pose significant challenges to structural investigation that do not apply to soluble globular proteins. This is demonstrated by the number of membrane protein structures currently in the Protein Data Bank (PDB). The total number of structures held in the PDB in 2015 is 112387, of which only 2927 are for membrane proteins (Bernstein, et al., 1977, Raman, et al., 2006). This is only approximately 2.6% of all the structures held, however it has been estimated that around 27% of human proteins are alpha helical membrane proteins and greater than 50% of human proteins interact with the membrane in some way (Almen, et al., 2009). The main obstacle that must be

overcome in studying membrane proteins is their solubilisation without at least partial denaturation of the native structure. This makes conventional crystallography diffraction studies for structure determination difficult.

Nuclear Magnetic Resonance (NMR) does not require formation of crystals, and can be used for structural elucidation as well as investigation of dynamics in solution. Membrane proteins are most often solubilised in detergents in solution state NMR, or in lipid vesicles in solid state NMR experiments. In this way a more dynamic natural structure of proteins can be studied (Wuthrich, 1990). There are however advantages to crystallography when screening of multiple samples is required.

1.1.2 Motifs that Stabilise Transmembrane Helix-Helix Interactions

Helical TM domains exist largely within the hydrophobic core of the bilayer and consequently consist of a much higher proportion of hydrophobic amino acids than most other domains. However several TM helix-helix interaction motifs have been identified which appear frequently in TM domains, some of which include charged residues.

A common sequence motif found in α helical TM domains is the G-xxx-G motif, in which two glycine residues are separated by three amino acids, and was first identified in the study of Glycophorin A (GpA) (MacKenzie et al., 1997). GpA is a TM protein found in red blood cells which forms a coating that prevents the red blood cell from adhering to other cells (Russ and Engelman, 2000). In this motif, the three residue spacer means that the two small glycine residues fall on the same face of a helical TM domain (**Figure 1.2**) forming a groove (there are 3.6 residues per

turn of an α helix (Dunitz, 2001)). This allows favourable close packing with a similar groove on a neighbouring helical domain. Other small residues, such as serine and alanine, in addition to glycine have also been shown to form favourable packing in this manner (Senes, et al., 2004). The G-xxx-G motif in GpA has a vital role in the dimerisation of this protein, and is over represented across helical TM domains.

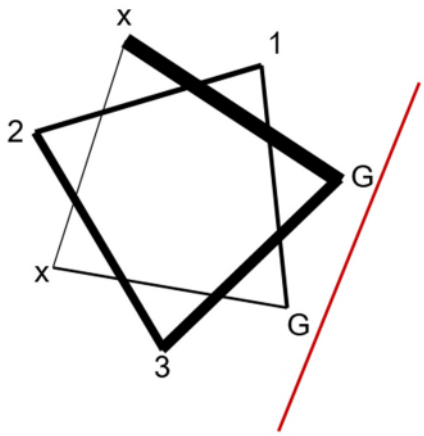


Figure 1.2 – A helical wheel representation showing a G-xxx-G motif. The two glycine residues, separated by three amino acids, fall on the same face of the helix forming an interface of small residues.

There are several other motifs that have been shown to affect TM domain interactions such as π-π stacking between aromatic amino acid side chains (phenylalanine, tyrosine and tryptophan) on separate helices (Dougherty, 1996, Johnson, et al., 2007). Hydrogen bond formation between polar/charged residues is also an effective stabilising motif. The formation of a hydrogen bond brings two polar residues together and helps to shield them from the hydrophobic membrane environment, thus minimising unfavourable (polar-hydrophobic) interactions (Zhou, et al., 2001). Heptad repeats, repeating patterns of seven amino acids, known to support leucine zipper type interactions in soluble proteins have also been shown to promote interactions in TM domains of membrane proteins such as bacteriorhodopsin (Langosch and Heringa, 1998).

Due to the constrained structure of the amino acid proline and the inability of its amide group to form the normal backbone-stabilising hydrogen bonding interactions in a helix, proline residues in α helical TM domains can affect packing of helices by either disrupting the helix or allowing the formation of a tightly coiled helix.

1.2 Carnitine Palmitoyltransferase 1 (CPT1)

1.2.1 Catalytic Activity of CPT1

Carnitine palmitoyltransferase 1 (CPT1) is a vital enzyme in fatty acid oxidation where it catalyses the reaction between long chain fatty acyl Coenzyme A (CoA) and carnitine (**Figure 1.3**). There are two other enzymes responsible for carrying out this trans-esterification reaction on other chain length fatty acyl CoA molecules: carnitine octanoyltransferase (COT) for medium chain length and carnitine acyltransferases for short chain length. Additionally carnitine palmitoyltransferase 2 (CPT2), which is associated with the matrix side of the inner mitochondrial membrane, can also perform this reaction. The actions of CPT1 and CPT2 allow for the transport of fatty acids into the mitochondrial matrix.

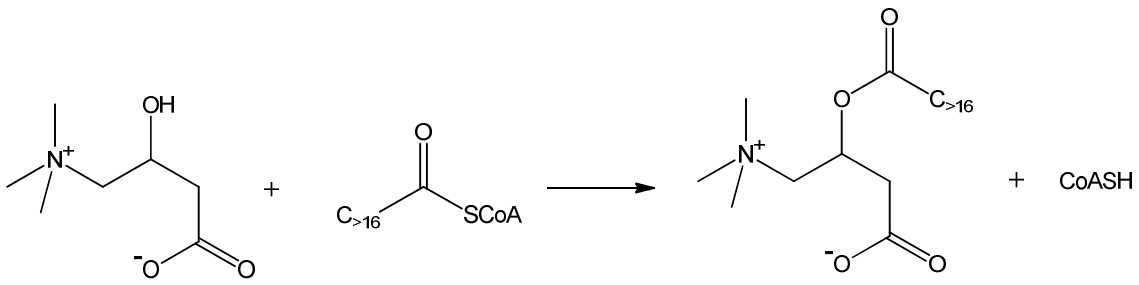


Figure 1.3 – The reaction catalysed by all carnitine acyltransferases (CATs) – the trans-esterification of carnitine with fatty acyl CoA. CPT1 specifically catalyses the reaction of long chain (greater than 16 carbon atoms) fatty acyl CoAs.

1.2.2 The Three Isoforms of CPT1

There are three known isoforms of CPT1 which occur primarily in three different tissue types: CPT1A in the liver; CPT1B in heart and skeletal muscle, as well as other highly oxidative tissues such as brown adipose; and CPT1C in the brain. All CPT1 isoforms are transmembrane proteins (Zammit, 2008) with two TM domains which span the outer mitochondrial membrane (OMM). CPT1C has however been shown to reside in the endoplasmic reticulum membrane in neurons (Sierra, et al., 2008). Comparatively little is known about CPT1C as it is the most recently discovered and is thought to have a significantly different role in metabolism (Price, et al., 2002). Its catalytic efficiency is extremely low, with an activity at least 2 orders of magnitude lower than either CPT1A or CPT1B (Price, et al., 2002, Wolfgang, et al., 2006).

1.2.2.1 CPT1A and CPT1B

CPT1A (773 amino acids) and CPT1B (772 amino acids) have the same catalytic specificity, but markedly different sensitivity to inhibition by malonyl-CoA,

the product of the first reaction in the pathway of fatty acid biosynthesis (Shi, et al., 2000). The inhibition of CPT1 by malonyl-CoA ensures that the synthesis and breakdown of fatty acids does not occur simultaneously and therefore unnecessarily. Because malonyl-CoA arises primarily from the metabolism of glucose (via pyruvate and citrate), this means that CPT1 has a major role in balancing the metabolism of glucose and fatty acids for energy production.

CPT1 allows long chain fatty acids to cross the OMM by linking them to carnitine. The inner membrane of mitochondria is impermeable to acyl-CoA esters, whereas a specific transporter for acylcarnitines (and carnitine) is expressed in this membrane. This function effectively makes it the rate limiting step in the oxidation of fatty acids because it regulates the entry of long chain fatty acids into the mitochondria where β oxidation can occur (**Figure 1.4**) (Ramsay and Zammit, 2004, Zammit, 2008). CPT1 utilises long-chain acyl-CoA esters, and so it has effects not just on the rate of fatty acid oxidation, but also on any process that requires a supply of long chain acyl-CoA, such as lipid synthesis. In addition, acyl-CoA esters are highly bioactive molecules that can affect gene transcription and the activity of ion channels. Thus modulation of CPT1 activity can have wide-ranging metabolic effects.

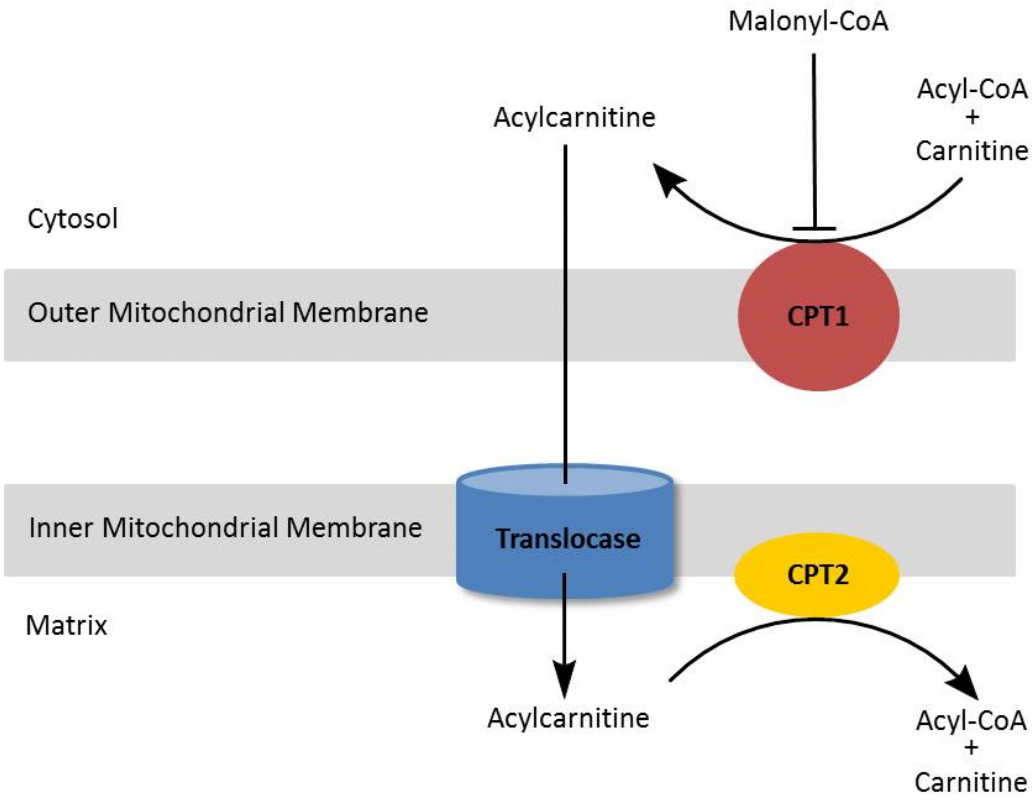


Figure 1.4 – Schematic of the enzymes responsible for the shuttling of long chain fatty acids into the mitochondrial matrix.

1.2.2.2 CPT1C

CPT1C is specific to the brain and occurs in areas responsible for appetite control and regulation of diurnal rhythm (Zammit, 2008). It binds malonyl-CoA with the same affinity as that of CPT1A but has very low catalytic activity. CPT1C could therefore act as a cellular 'sink' for malonyl-CoA in the areas in which it is expressed. It is involved in appetite control and regulation of bodyweight. Experiments in transgenic mice with a knockout CPT1C gene have shown that regulation of body weight is disrupted, but overexpression of CPT1C in the hypothalamus can protect against weight gain (Dai, et al., 2007, Sierra, et al., 2008).

1.2.3 The Structures of CPT1A and CPT1B

The topology of the CPT1 enzyme was determined using partial proteolysis, immobilised malonyl-CoA and substrates, anti-peptide antibody binding and immunogold electron microscopy experiments (Fraser, et al., 1997, van der Leij, et al., 1999). The results of these studies indicated that CPT1A and CPT1B both have a large C-terminal domain (~600 residues) that contains the catalytic domain, and a much smaller (~47 residue) N-terminal domain (**Figure 1.5**). These two domains are connected by the two TM domains (TM1 and TM2) and a short loop (Faye, et al., 2005, Shi, et al., 2000, Swanson, et al., 1998, Zammit, 2008).

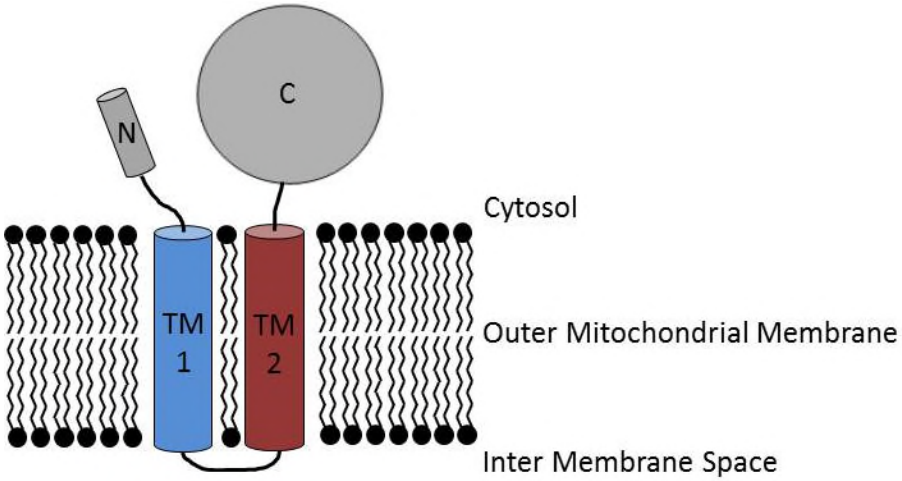


Figure 1.5 – Diagrammatic representation of the topology of the membrane proteins CPT1A and CPT1B showing the large C-terminus containing the catalytic site and small regulatory N-terminus (Zammit, 1999).

CPT 1 adopts a polytopic conformation with both the N- and C-termini on the cytosolic side of the OMM (Zammit, 1999). This means that almost the entire protein is exposed to the cytosol, except for the transmembrane domains and the short loop connecting them (**Figure 1.5**). Several experimental approaches have provided evidence that interactions between the N- and C-terminal domains have an effect on

the regulation of the catalytic activity by the inhibitor malonyl-CoA. The N- and C-termini can be chemically cross-linked by an agent with a spacer length of 15.8 Å. Furthermore the ease of crosslinking observed is dependent on the physical-chemical state of the OMM, as determined by the physiological state of the donor animal. Thus, there is less crosslinking when CPT1A is less sensitive to malonyl-CoA, owing to the greater fluidity of the membrane e.g. in membranes isolated from fasted or diabetic animals (Jackson, et al., 2000, Jackson, et al., 2001, Jackson, et al., 2000).

There has also been *in silico* modelling that suggests electrostatic interactions between opposing residues on the two domains which stabilise the docking of the N-terminus onto the C-terminus (Lopez-Vinas, et al., 2007). However for the N- and C-termini to interact there must also be interactions between the two transmembrane helices and it is into this area that my research is focussed.

1.2.4 Implications for the Role of Transmembrane Domains in the Function of CPT1

CPT1B has a 30-100 fold greater sensitivity to malonyl-CoA than CPT1A (Shi, et al., 2000). A series of deletions and truncations of the N-terminal domain were shown to retain the catalytic activity of CPT1 but to greatly alter the sensitivity to malonyl-CoA (Shi, et al., 1999). This provided evidence that the N-terminal domain is indeed the regulatory domain which modulates the binding of malonyl-CoA to the catalytic domain (Shi, et al., 2000).

In addition to this difference in sensitivity to malonyl-CoA between the two isoforms, it has been found that the sensitivity of CPT1A is modulated by the membrane environment in which it is situated (see above) whereas that of CPT1B is

not (McGarry, et al., 1983, Park and Cook, 1998, Rao, et al., 2011, Zammit, et al., 1997). The membrane environment is significantly affected by the physiological state of the organism, such as when fasting, as well as by some diseases, such as diabetes. Both the composition as well as the fluidity (which is affected by composition) of the membrane may be profoundly altered under these conditions. This sensing of the membrane environment provides a strong indication that intramolecular interactions in the transmembrane domains may have some effect on the sensitivity to malonyl-CoA, at least in CPT1A. These interactions could be both, between the two transmembrane helices, as well as between the helices and the surrounding lipid components of the membrane (Zammit, et al., 1998, Zammit, et al., 1989).

Another characteristic that can alter considerably is the curvature of the membrane. In mitochondria there are contact sites where the outer and the inner membranes come into close proximity. It has been observed that CPT1 is enriched at these contact sites between the two membranes (Fraser and Zammit, 1998). These contact sites are also the location of the electron transport chain and β-oxidation complex on the inner mitochondrial membrane (Kerner and Bieber, 1990). It is also known that the mitochondrial membranes have a very specific lipid composition at these contact sites which would likely also affect the kinetic properties of CPT1A. The importance of membrane curvature on the regulation of malonyl-CoA sensitivity of CPT1A was demonstrated recently (Rao, et al., 2011).

Due to the difficulties involved in obtaining structural data for transmembrane proteins, there is no complete structure of any of the CPT1 isoforms available, and because of this there is no definitive structural information for the active site at which the trans-esterification reaction occurs, or of malonyl-CoA

inhibition. Recently a structure of the N-terminal domain of CPT1A has been elucidated by solution NMR and a structure for the C-terminal domain of CPT1A has been proposed by homology modelling with rat carnitine acetyltransferase (Morillas, et al., 2004) (**Figure 1.6**).

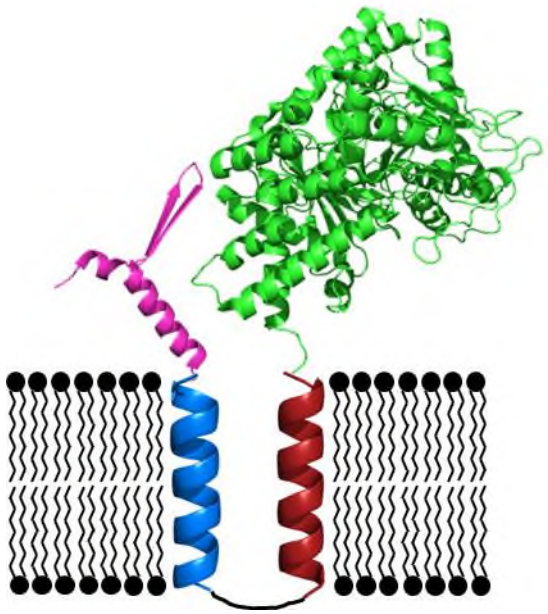


Figure 1.6 – The current structural information known about CPT1A. The N-terminal NMR structure (pink), C-terminal homology structure (green), and the unknown structure of the two TM domains, TM1 (blue) and TM2 (red) shown as generic helices.

One of the most interesting mechanisms proposed for the transport across the mitochondrial outer membrane involves the oligomerisation of CPT1 and the formation of a pore in the membrane. It has been shown that CPT1A exists in dimeric, trimeric and possibly hexameric states *in vivo* so this is a distinct possibility (Jenei, et al., 2009, Jenei, et al., 2011, Morillas, et al., 2004). If this is the case then there is even greater importance in investigating any transmembrane interactions as these would play an important role in the stability of oligomeric complexes. Both hetero (intramolecular) as well as homo (intermolecular) interactions could

contribute to any stabilising and destabilising forces (**Figure 1.7**) (Unterreitmeier, et al., 2007).

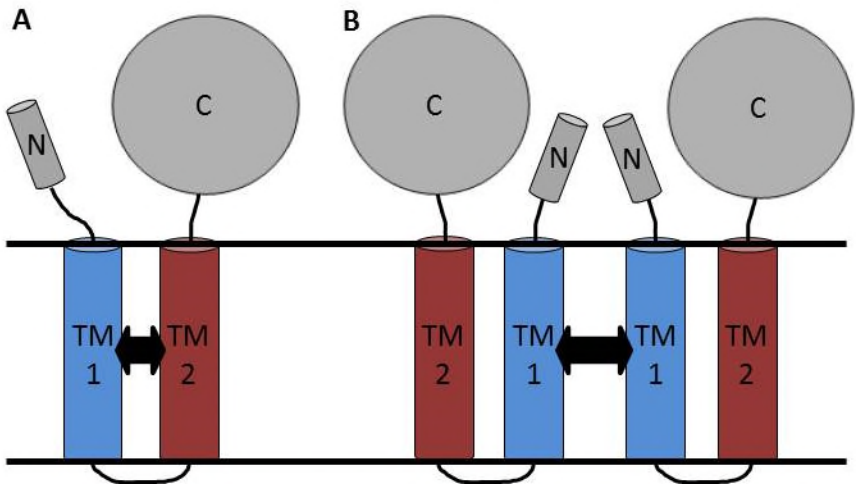


Figure 1.7 – Schematic of the types of interactions being investigated. (**A**) Heterotypic, intramolecular interactions between TM1 (blue) and TM2 (red), and (**B**) homotypic, intermolecular interactions between TM2 (red) and another TM2 (red) on a neighbouring molecule. These homotypic interactions could also occur between two TM1 domains (blue).

The study of CPT1 has important implications for pharmacological strategies aimed at the control of fatty acid oxidation and acyl-CoA control in tissues in conditions such as insulin resistance and diabetes. Hyperglycemia associated with type 2 diabetes causes increased levels of malonyl-CoA and consequently increased inhibition of CPT1 (McCormick, et al., 1988, Zammit, et al., 2009). This inhibition leads to a reduction in the transport of long chain fatty acids into mitochondria for oxidation, and subsequently to an increase in free fatty acid levels causing fat to accumulate in skeletal muscle (Dobbins, et al., 2001, Rasmussen, et al., 2002). An improved structural understanding of membrane bound CPT1 will greatly facilitate the design of therapeutics that could target this key enzyme.

1.3 Nuclear Magnetic Resonance (NMR)

Other than X-ray crystallography and, more recently, cryo-electron microscopy, NMR spectroscopy is one of the only methods able to produce high resolution structural information about proteins. It is capable of studying proteins in much more native like environments than X-ray crystallography; proteins can be studied in solution if soluble, or for insoluble membrane proteins a variety of membrane mimetic substances such as detergents and lipids can be used to solubilise them. NMR spectroscopy was the method of choice in this project and an introduction to the theory of NMR follows.

1.3.1 Spin

NMR spectroscopy is a technique that exploits the intrinsic atomic property known as 'spin'. Spin, or the intrinsic angular momentum, of an atom is characterised by its spin quantum number (I). The spin quantum number is defined by the nuclear composition of an atom: if the nucleus contains an even number of protons and neutrons I = 0; if the number of protons added to the number of neutrons in a nucleus is odd then the nucleus has half integer spin; if the number of neutrons and the number of protons are both odd then the nucleus has integer spin. The gyromagnetic ratio (γ), a fundamental unique constant for each type of nuclei, is related to the spin quantum number and the magnetic moment (μ) of a nucleus following **equation 1.1**.

$$\gamma = \frac{2\pi\mu}{hI} \qquad \textbf{1.1}$$

Atoms with spin not equal to 0 are NMR active and are capable of undergoing transitions between nuclear spin energy levels when placed in an external

magnetic field. While all nuclei with I > 1 are observable via NMR, nuclei with I = ½ are the easiest to observe experimentally. For protein NMR the most important nuclei that have I = ½ are ^{1}H, ^{13}C and ^{15}N. Quantum mechanics states that a nucleus of spin I has 2I + 1 spin states which are of equal energy in the absence of an external magnetic field, however NMR spectroscopy exploits the fact that in the presence of a magnetic field these energy levels split. This is known as the Zeeman Effect and can be seen in **Figure 1.8**.

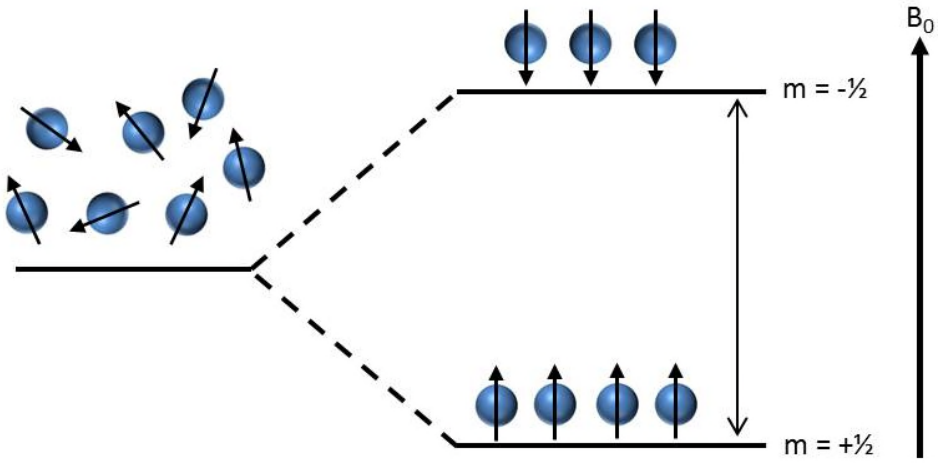


Figure 1.8 – Zeeman splitting on introduction of a spin I =½ nuclei to a magnetic field.

The initial populations of these energy levels is described by the Boltzmann distribution and is thermodynamically governed resulting in a slight population excess in lower energy states. Nuclei with I = ½ will have 2 possible spin states denoted as, m = +½ aligned with the external magnetic field, and m = -½ energetically unfavourably aligned against the external magnetic field. In NMR spectroscopy nuclei are excited into higher energy states by the use of electromagnetic radiation, and in the case of nuclei with I= ½ there is only one transition.

The frequency of electromagnetic radiation required to transition between the two energy states is proportional to the difference in energies of the states. This energy difference is governed by the strength of the external magnetic field (conventionally labelled as B_0) and an inherent property of each nucleus, the gyromagnetic ratio. The energy difference increases with an increase in either of these parameters as shown in **equation 1.2** and directly relates to the extent to which the lower energy state is favourable.

$$\Delta E = \frac{\gamma h B_0}{2\pi} \qquad \textbf{1.2}$$

The population of each state (N) is given by **equation 1.3** where ΔE is the difference in energies, k is the Boltzmann constant and T is the temperature. This shows that at thermal equilibrium at room temperatures the ratio between the two states is extremely small.

$$\frac{N_{upper}}{N_{lower}} = e^{-\frac{\Delta E}{kT}} \qquad \textbf{1.3}$$

This very small population difference is one of NMR spectroscopy's largest problems as this is the cause of low sensitivity in NMR experiments. As can be seen by **equations 2** and **3** the energy difference and hence the population difference can be increased by increasing the magnetic field strength (B_0) or by decreasing the temperature. Additionally increasing the sample concentration and averaging multiple experiments also increases the signal to noise. Despite this, there is a very small difference in population which results in a net magnetisation (M_0).

1.3.2 Net Magnetisation

When a nucleus with intrinsic angular momentum is placed in a magnetic field the nucleus' magnetic moment will precess around the axis of the magnetic field (conventionally this is the z axis) (**Figure 1.9**). The frequency of this precession (ω) is given by **equation 1.4** and is known as the Larmor frequency.

$$\omega = \gamma B \qquad \textbf{1.4}$$

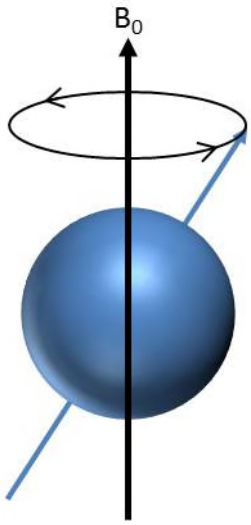


Figure 1.9 – A single nucleus has spin angular momentum and a specific frequency of precession that is dependent on its gyromagnetic ratio and the magnitude of the magnetic field.

In an NMR sample all nuclei of a particular type will all precess at those nuclei's Larmor frequency resulting in an ensemble average magnetic moment with a common angular frequency. This bulk magnetisation is shown as M_0 in **Figure 1.10** where the net magnetisation is aligned with B_0. M_0 is positive due to the slight population excess caused by the external magnetic field.

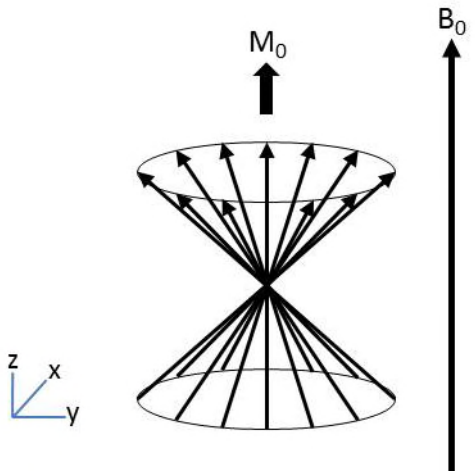


Figure 1.10 – Application of a magnetic field causes the ensemble average magnetisation (M_0) to precess around the z axis with a net magnetisation parallel to the magnetic field. This net magnetisation is caused by the excess of nuclear spins in the lower energy state.

In an NMR experiment, once a sample is placed in the magnetic field M_0 will align with B_0 until equilibrium is reached. Electromagnetic radiation in the form of a radio frequency (RF) pulse at the Larmor frequency of the nuclei to be studied is applied to the system which perturbs this equilibrium. This rotates M_0 into the x-y plane where it precesses and induces a voltage in the detector coil that is placed in this plane. The RF energy applied to the system equalises the probability of transitions between the higher and lower energy spin states, and because there is a greater population in the lower state there will be more transitions to the higher state than from higher to lower resulting in a perturbation from equilibrium. Once the pulse is complete M_0 will gradually relax back to its equilibrium position aligning with B_0. This is longitudinal, or spin-lattice, relaxation known as T_1. This together with spin-spin relaxation, known as T_2, caused by interactions that redistribute the energy across the nuclei within a spin system, allow this relaxation back to equilibrium and cause the NMR signal to decay with time. A simple 90° pulse is shown along with the relaxation afterwards in **Figure 1.11**. This decaying signal is

known as a free induction decay (FID) and it is this that is the detected signal in an NMR experiment. This FID has contributions from all the nuclei affected by the RF pulse and so to be analysed it is converted from a time domain signal to a frequency domain spectrum by Fourier transformation.

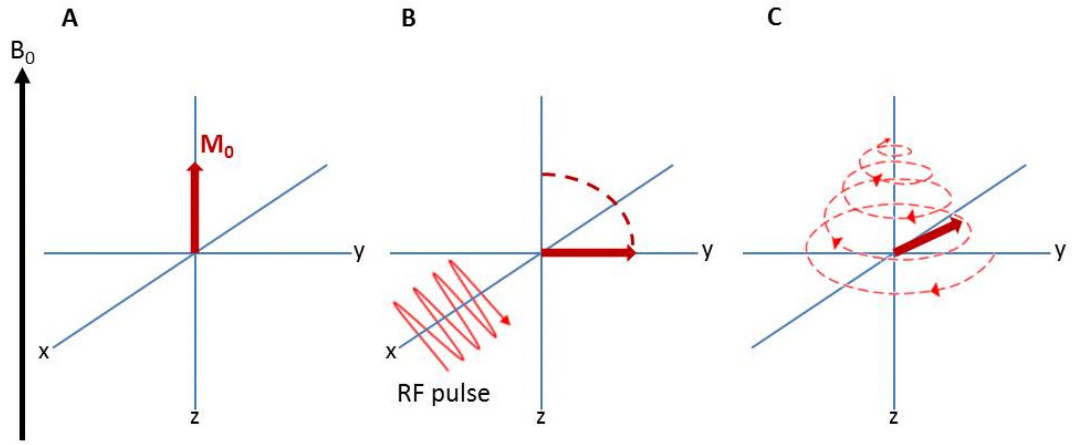


Figure 1.11 – A simple 90° RF pulse. (**A**) The position of M_0 at equilibrium. (**B**) An RF pulse is introduced which flips M_0 into the x-y plane. (**C**) After the RF pulse M_0 precesses around the z axis as it relaxes to the equilibrium position.

1.3.3 Chemical Shift

Different nuclei give signals at different frequencies due to their different gyromagnetic ratios and different Larmor frequencies. However the Larmor frequency of a particular nucleus is also affected by that particular nucleus' local electronic environment. The external magnetic field used in NMR experiments induces a local magnetic field in the electron clouds of each nucleus opposite to that of the external field. This induced local magnetic field has a small, but measureable, effect on each nucleus. This effect can either serve to cancel out a small amount of the external magnetic field which is known as shielding, or enhance the field, known as deshielding. These effects cause all the nuclei of the same type but in different

electronic environments to have slightly different Larmor frequencies. These differences in the Larmor frequency are referred to as chemical shifts. The chemical shift is conventionally expressed in parts per million (ppm) difference in frequency (**equation 1.5** where ω_{ref} is the frequency of a reference compound) rather than in Hz to remove the effect of the magnetic field strength.

$$\delta_{ppm} = \frac{\omega_0 - \omega_{ref}}{\omega_{ref}} \times 10^6 \qquad \textbf{1.5}$$

1.3.4 Two Dimensional (2D) NMR

The differences in chemical shifts are enough to distinguish all the different nuclei in small molecule samples, however when working with large macromolecular samples, such as proteins, there are a great deal of chemical shifts all very close to each other and often overlapping. To address this problem, two or higher dimensional NMR experiments can be performed to separate the observed chemical shifts across a second frequency axis. This second axis can either be for the same type of nucleus (homonuclear experiments) or for a different type of nucleus (heteronuclear experiments).

2D NMR experiments all have four main stages: preparation where the magnetisation is established; evolution where the spins are allowed to precess; mixing where the magnetisation is transferred between nuclei; and finally detection. A 2D experiment is in reality a series of 1D experiments with different evolution and detection times. Magnetisation can be transferred by one of two mechanisms: scalar coupling (through bond) or via dipolar interactions (through space).

1.3.4.1 Total Correlation Spectroscopy (TOCSY)

TOCSY is a ^{1}H detected homonuclear 2D experiment where magnetisation is allowed to disperse throughout each spin system via scalar coupling prior to detection. This experiment is particularly important in analysing small proteins as each amino acid is an isolated spin system. Magnetisation disperses throughout each amino acid giving a characteristic pattern depending on the side chain allowing identification of the amino acid.

1.3.4.2 Nuclear Overhauser Effect Spectroscopy (NOESY)

NOESY is a ^{1}H detected homonuclear 2D experiment which relies on the Nuclear Overhauser effect (NOE) to transfer magnetisation by dipolar coupling through space. This type of experiment is extremely important in protein structure determination experiments. The intensity of the NOE is approximately proportional to $1/r^{6}$ and consequently usually only protons within approximately 5 Å are observed. This allows certain constraints to be derived even in the case where amino acids which are distant in primary sequence but are close in space giving valuable information about the tertiary structure of a protein.

1.3.4.3 Heteronuclear Single Quantum Coherence (HSQC)

HSQC is a heteronuclear 2D experiment which is extremely important for protein studies. It is commonly used to correlate the nitrogen and proton of NH groups. As there is an NH group present in the peptide bond of each amino acid this type of experiment results in a signal for each amino acid in a protein, except for proline as there is no directly bonded proton in a proline peptide bond. Signals are

also observed for the NH_2 groups of asparagine and glutamine, and the NH of tryptophan and histidine. ^{14}N has a spin I = 1 and therefore is quadrupolar and difficult to observe using NMR however so in this experiment when correlating with nitrogen, ^{15}N must be used. The natural abundance of ^{15}N is however very low (0.368%) so isotopic enrichment must be used for this type of experiment to be successful except at high concentrations of protein.

1.4 Aims and Objectives

The overall aim of this project is to investigate the interactions present in the TM and juxtamembrane regions of CPT1A and CPT1B and how the TM domains may result in the markedly different kinetic characteristics of the two isoforms, which share > 65% sequence identity. This broad aim can be broken down into several objectives, as listed below:

- Investigation of homo- and hetero-association of the isolated TM domains of both CPT1A and CPT1B. This will be accomplished using both a genetic *in vivo* assay (GALLEX) and *in vitro* peptide-based assays using circular dichroism and NMR spectroscopy.
- To determine the residues and/or motifs (e.g. those discussed in **Section 1.1.2**) that are important for maintaining these interactions using mutagenesis strategies.
- The expression of isotopically labelled peptides of CPT1A and CPT1B TM domains for use in heteronuclear NMR spectroscopy for structure determination.

2. Materials and Methods

2.1 Reagents and Materials

All reagents and materials used were supplied by Sigma Aldrich (UK), Fisher Scientific (UK), Pierce (UK), BioRad (UK), and Avanti-polar lipids (USA) unless otherwise specified.

2.2 Bacterial Strains

Six strains of *E. coli* were used in these experiments (**Table 2.1**), all of which were available from lab stocks stored at -80°C. DH5α and TOP10 cells were used for all the cloning and for replication and maintenance of the plasmids after mutagenesis. NT326 cells, which lack a native Maltose Binding Protein (MBP), were used as a control for the GALLEX assay. *E. coli* SU101 and SU202 cells were used to perform the GALLEX assay: SU101 was used to measure homo interactions and SU202 for hetero interactions. Finally BL21 (DE3) pLysS cells were used to express the trpLE-TM constructs. SU101 and SU202 cells were kindly provided by Prof. Dirk Schneider.

Strain	Genotype	Source
DH5α	F´ proA+B+ lacI^{q} Δ lacZ M15/ fhuA2 Δ(lac-proAB) glnV gal R(zgb-210::Tn10)Tet^{S} endA1 thi-1 Δ(hsdS-mcrB)5	NEB (UK)
TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^{R}) endA1 λ^{-}	Invitrogen (UK)
BL21 (DE3) pLysS	F^{-} ompT gal dcm lon hsdS$_{B}$(r_{B}^{-} m_{B}^{-}) λ(DE3) pLysS(cm^{R})	NEB (UK)
SU101	lexA71::Tn5 (Def)sulA211 D(lacIPOZYA)169/F¢lacIqlacZDM15::Tn9 op+/op+	D. Schneider (University of Freiburg, Germany)
SU202	lexA71::Tn5 (Def)sulA211 D(lacIPOZYA)169/F¢lacIqlacZDM15::Tn9 op408/op+	D. Schneider (University of Freiburg, Germany)
NT326	F−(argF-lac)U169, rpsL150, relA1, rbsR, flbB5301, ptsF25, thi-1, deoC1, ΔmalE444, recA, srlA	D. Engelman (Yale University, USA)

Table 2.1 – *E. coli* strains used in this project (Cymer, et al., 2013).

2.3 Vectors

Table 2.2 provides a list of the plasmids that had been generated prior to this study. The pALM148 and pBLM 100 were kindly provided by Prof. Dirk Schneider.

Plasmid	**Description**
pALM148	Described in detail in reference (Schneider & Engelman, 2003)
pBLM100	Described in detail in reference (Schneider & Engelman, 2003)
pALM148 GpA	pALM148 as above with Glycophorin A TM domain insert
pALM148 $G_{83}I$	pALM148 as above with Glycophorin A Gly83 to Ile mutation TM domain insert
pBLM100 GpA	pBLM100 as above with Glycophorin A TM domain insert
pBLM100 $G_{83}I$	pBLM100 as above with Glycophorin A Gly83 to Ile mutation TM domain insert

Table 2.2 – Previously generated plasmids used in this project.

Table 2.3 lists the mutation carrying plasmids generated during this study to produce CPT1A TM1 fusion proteins for homo and hetero GALLEX experiments. All the listed mutations were generated in both pALM148 and pBLM100.

Plasmid	Sequence	Description
CPT1ATM1	VVVILWSSPNAPFVGTII	Wild type TM1 from CPT1A
CPT1ATM1 $V_{65}A$	VV**A**ILWSSPNAPFVGTII	Valine 65 to alanine
CPT1ATM1 $I_{64}A$	VVV**A**LWSSPNAPFVGTII	Isoleucine 64 to alanine
CPT1ATM1 $L_{63}A$	VVVI**A**WSSPNAPFVGTII	Leucine 63 to alanine
CPT1ATM1 $W_{62}A$	VVVIL**A**SSPNAPFVGTII	Tryptophan 62 to alanine
CPT1ATM1 $S_{61}I$	VVVILW**I**SPNAPFVGTII	Serine 61 to isoleucine
CPT1ATM1 $S_{60}I$	VVVILWS**I**PNAPFVGTII	Serine 60 to isoleucine
CPT1ATM1 $P_{59}A$	VVVILWSS**A**NAPFVGTII	Proline 59 to alanine
CPT1ATM1 $N_{58}A$	VVVILWSSP**A**APFVGTII	Asparagine 58 to alanine
CPT1ATM1 $A_{57}I$	VVVILWSSPN**I**PFVGTII	Alanine 57 to isoleucine
CPT1ATM1 $P_{56}A$	VVVILWSSPNA**A**FVGTII	Proline 56 to alanine
CPT1ATM1 $F_{55}A$	VVVILWSSPNAP**A**VGTII	Phenylalanine 55 to alanine
CPT1ATM1 $V_{54}A$	VVVILWSSPNAPF**A**GTII	Valine 54 to alanine
CPT1ATM1 $G_{53}I$	VVVILWSSPNAPFV**I**TII	Glycine 53 to isoleucine

Table 2.3 – CPT1A TM1 sequences constructed using site directed mutagenesis in both pALM148 and pBLM100.

Table 2.4 lists the mutation carrying plasmids generated during this study to produce CPT1A TM2 fusion proteins for homo and hetero GALLEX experiments. The listed mutations had previously been introduced into the pBLM100 plasmid, but in this study the mutations were also introduced to the pALM148 plasmid to allow hetero GALLEX experiments to be performed (GALLEX explained in detail in **Section 2.8**).

Plasmid	Sequence	Description
CPT1A TM2	IVSGVLFGTGLWVAVIMT	Wild type TM2 from CPT1A
CPT1A TM2 $G_{107}I$	IVS**I**VLFGTGLWVAVIMT	Glycine 107 to isoleucine
CPT1A TM2 $V_{108}A$	IVSG**A**LFGTGLWVAVIMT	Valine 108 to alanine
CPT1A TM2 $L_{109}A$	IVSGV**A**FGTGLWVAVIMT	Leucine 109 to alanine
CPT1A TM2 $F_{110}A$	IVSGVL**A**GTGLWVAVIMT	Phenylalanine 110 to alanine
CPT1A TM2 $G_{111}I$	IVSGVLF**I**TGLWVAVIMT	Glycine 111 to isoleucine
CPT1A TM2 $T_{112}A$	IVSGVLFG**A**GLWVAVIMT	Threonine 112 to alanine
CPT1A TM2 $G_{113}I$	IVSGVLFGT**I**LWVAVIMT	Glycine 113 to isoleucine
CPT1A TM2 $L_{114}A$	IVSGVLFGTG**A**WVAVIMT	Leucine 114 to alanine
CPT1A TM2 $W_{115}A$	IVSGVLFGTGL**A**VAVIMT	Tryptophan 115 to alanine
CPT1A TM2 $V_{116}A$	IVSGVLFGTGLW**A**AVIMT	Valine 116 to alanine
CPT1A TM2 $A_{117}I$	IVSGVLFGTGLWV**I**VIMT	Alanine 117 to isoleucine
CPT1A TM2 $V_{118}A$	IVSGVLFGTGLWVA**A**IMT	Valine 118 to alanine
CPT1A TM2 $G_{107,113}I$	IVS**I**VLFGT**I**LWVAVIMT	Glycine 107 and 113 to isoleucine

Table 2.4 – CPT1A TM2 sequences constructed using site directed mutagenesis in both pALM148 and pBLM100.

Table 2.5 lists the mutation carrying plasmids generated during this study to produce CPT1B TM1 fusion proteins for homo and hetero GALLEX experiments. All the listed mutations were generated in both pALM148 and pBLM100.

Plasmid	Sequence	Description
CPT1B TM1	MVVVLWSTPSGPYVGRLI	Wild type TM1 from CPT1B
CPT1B TM1 $V_{64}A$	MVV**A**LWSTPSGPYVGRLI	Valine 64 to alanine
CPT1B TM1 $L_{63}A$	MVVV**A**WSTPSGPYVGRLI	Leucine 63 to alanine
CPT1B TM1 $W_{62}A$	MVVVL**A**STPSGPYVGRLI	Tryptophan 62 to alanine
CPT1B TM1 $S_{61}I$	MVVVLW**I**TPSGPYVGRLI	Serine 61 to isoleucine
CPT1B TM1 $T_{60}A$	MVVVLWS**A**PSGPYVGRLI	Threonine 60 to alanine
CPT1B TM1 $P_{59}A$	MVVVLWST**A**SGPYVGRLI	Proline 59 to alanine
CPT1B TM1 $S_{58}I$	MVVVLWSTP**I**GPYVGRLI	Serine 58 to isoleucine
CPT1B TM1 $G_{57}I$	MVVVLWSTPS**I**PYVGRLI	Glycine 57 to isoleucine
CPT1B TM1 $P_{56}A$	MVVVLWSTPSG**A**YVGRLI	Proline 56 to alanine
CPT1B TM1 $Y_{55}A$	MVVVLWSTPSGP**A**VGRLI	Tyrosine 55 to alanine
CPT1B TM1 $V_{54}A$	MVVVLWSTPSGPY**A**GRLI	Valine 54 to alanine
CPT1B TM1 $G_{53}I$	MVVVLWSTPSGPYV**I**RLI	Glycine 53 to isoleucine

Table 2.5 – CPT1B TM1 sequences constructed using site directed mutagenesis in both pALM148 and pBLM100.

Table 2.6 lists the mutation carrying plasmids generated during this study to produce CPT1B TM2 fusion proteins for homo and hetero GALLEX experiments. All the listed mutations were generated in both pALM148 and pBLM100.

Plasmid	Sequence	Description
CPT1B TM2	LLSMVIFSTGVWATGIFL	Wild type TM2 from CPT1B
CPT1B TM2 $M_{109}A$	LLS**A**VIFSTGVWATGIFL	Methionine 109 to alanine
CPT1B TM2 $V_{110}A$	LLSM**A**IFSTGVWATGIFL	Valine 110 to alanine
CPT1B TM2 $I_{111}A$	LLSMV**A**FSTGVWATGIFL	Isoleucine 111 to alanine
CPT1B TM2 $F_{112}A$	LLSMVI**A**STGVWATGIFL	Phenylalanine 112 to alanine
CPT1B TM2 $S_{113}I$	LLSMVIF**I**TGVWATGIFL	Serine 113 to isoleucine
CPT1B TM2 $T_{114}A$	LLSMVIFS**A**GVWATGIFL	Threonine 114 to alanine
CPT1B TM2 $G_{115}I$	LLSMVIFST**I**VWATGIFL	Glycine 115 to isoleucine
CPT1B TM2 $V_{116}A$	LLSMVIFSTG**A**WATGIFL	Valine 116 to alanine
CPT1B TM2 $W_{117}A$	LLSMVIFSTGV**A**ATGIFL	Tryptophan 117 to alanine
CPT1B TM2 $A_{118}I$	LLSMVIFSTGVW**I**TGIFL	Alanine 118 to isoleucine
CPT1B TM2 $T_{119}A$	LLSMVIFSTGVWA**A**GIFL	Threonine 119 to alanine
CPT1B TM2 $G_{120}I$	LLSMVIFSTGVWAT**I**IFL	Glycine 120 to isoleucine

Table 2.6 – CPT1B TM2 sequences constructed using site directed mutagenesis in both pALM148 and pBLM100.

2.4 Antibiotics

Unless otherwise stated all antibiotics were used at the following working concentrations, and were typically taken from 1000 × stock solutions. Solutions of ampicillin and kanamycin were made in distilled water and then filter sterilised. Tetracycline solutions were made in 70% ethanol and then filter sterilised, and finally chloramphenicol solutions were made in 100% ethanol.

Antibiotic	Working concentration (µg/ml)
Ampicillin	100
Kanamycin	50
Chloramphenicol	35
Tetracycline	10

Table 2.7 – Antibiotics and their concentrations used in this project.

2.5 *E. coli* Growth

All *E. coli* strains were grown in Lysogeny Broth (LB) medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl), either in liquid cultures or on plates with added agar (15 g/l). For the liquid cultures, 5 ml of LB liquid medium was inoculated with a single colony taken from an LB agar plate and incubated for ~16 hours at 37 °C and shaken at 180 r.p.m. Standard outgrowths taken from these liquid cultures were 1 in 100 dilutions into fresh media e.g. 50 µl of culture into 5 ml of fresh LB.

2.6 Cloning Methods

2.6.1 Restriction Enzyme Digestion

All restriction enzyme digests were carried out according to the enzyme manufacturer's instructions. A typical reaction contained 2 µl 10x reaction buffer, 17 µl plasmid DNA, and 0.5 µl of each restriction enzyme required. This reaction mixture was incubated for 0.5 – 1 hour (enzyme dependent) at 37 °C in a water bath. The DNA was then run on an agarose gel and after excision was purified using a QIAprep gel extraction kit (Qiagen, UK) according to the manufacturer's instructions.

2.6.2 Preparation of Synthetic Oligonucleotide Inserts

Synthetic oligonucleotides were used to clone several required TM sequences into expression vectors. In order to ligate these inserts into the plasmid vectors of interest, they first must be phosphorylated and coding and template strands annealed.

The enzyme manufacturer's instructions were followed. A typical phosphorylation reaction contained:

Component	Volume (µl)
10 µM oligonucleotide	5
10x kinase buffer	2
10 mM ATP	1
dH_2O	10
T4 polynucleotide kinase	2

This was incubated at 37 °C in a water bath for 30 minutes and then at 65 °C for 10 minutes to denature the enzyme and stop the reaction. To anneal coding and template oligonucleotide inserts the following procedure was followed.

The solution described below was heated to 95 °C for 7 minutes and then allowed to cool to room temperature slowly over approximately 20 minutes in a thermal cycler block. The annealed primers contained the required sticky ends for ligation into the plasmid vector.

Component	Volume (µl)
Forward oligonucleotide (2.5 µM)	4
Reverse oligonucleotide (2.5 µM)	4
Annealing buffer	2

Annealing buffer (200 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 500 mM NaCl)

2.6.3 Ligation of Synthetic Oligonucleotide Inserts into Vectors

The T4 ligase enzyme manufacturer's instructions (NEB, UK) were followed. A typical ligation reaction contained the following:

Component	Volume (µl)
Digested plasmid vector	2
Phosphorylated and annealed oligonucleotides	6
Ligation buffer (manufacturer supplied)	2
10 mM ATP	1
dH_2O	8
T4 ligase	1

This reaction was incubated at 16 °C overnight and then transformed into *E. coli* by the addition of 10 µl of the ligation mixture to 100 µl of competent cells.

2.6.4 Preparation of Chemically Competent Cells

A small amount of LB (typically 5 ml) was inoculated with a single colony of *E. coli* and grown overnight. This culture was then used to inoculate a larger volume (typically 100-200 ml) of fresh LB which was grown to an OD_{600} of 0.6. The culture was chilled on ice for 30 minutes and the cells harvested by centrifugation (3000 x g, 4 °C, 10 minutes) and resuspended in chilled 0.1 M $CaCl_2$ solution. This solution was left on ice for 30 minutes and again harvested by centrifugation as before. The cells were resuspended in chilled 0.1 M $CaCl_2$ with 15% glycerol added as a cryoprotectant. This solution was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C until required.

2.6.5 *E. coli* Transformations

Transformations of plasmids were performed into the required *E. coli* strain using the following conditions. 1 µl of plasmid DNA was added to a 100 µl aliquot of suitable chemically competent cells and allowed to incubate for 30 minutes at 4

°C. The cells were then subjected to a heat-shock at 42 °C for 45 seconds, followed by further incubation at 4 °C for 2 minutes. 1 ml of room temperature LB was then added and the cells incubated at 37 °C and shaken at 180 r.p.m. for 1 hour. These cells were sedimented at 3,000 x g for 5 minutes, resuspended in ~100 µl of LB medium and then plated onto LB-agar plates which included the relevant antibiotic(s). The plates were incubated at 37 °C for ~16 hours.

15 µl of plasmid was used for transformations into DH5α cells after site directed mutagenesis. This was to ensure good transformation efficiency of the nicked plasmid produced by the Quikchange mutagenesis strategies employed.

2.6.6 Checking Transformants by Colony Polymerase Chain Reaction (PCR)

Any colonies that grew from the transformation were checked for the presence of the insert by colony PCR. Single colonies were selected and dispersed in 50 µl dH_2O and added to the following reaction:

Component	**Volume (µl)**
10x taq buffer (manufacturer supplied)	1
10 mM dNTP mixture	0.1
10µM forward primer	0.1
10 µM reverse primer	0.1
W1 detergent	0.5
dH_2O	7
Taq polymerase	0.1

This reaction mixture was then cycled in a PCR thermal cycler using the following conditions with steps 2-4 repeated 30 times:

Step	Temperature (°C)	Time (s)
1	94	120
2	94	15
3	50	30
4	68	30
5	68	300

Once complete, the reactions were analysed on a 2% agarose gel to verify presence of the desired insert by comparison with a commercial DNA marker (NEB, UK).

2.6.7 Agarose Gels

Agarose gels were prepared by adding agarose to TAE buffer (40 mM Tris, pH 8, 20 mM acetic acid, 1 mM EDTA) and heating to dissolve. All gels were prepared at either 1%, to analyse plasmid DNA, or 2%, to analyse smaller DNA fragments such as inserts or primers. The gel was allowed to cool before the addition of a 10000 × stock of GelRed ™ (Biotium, USA)

2.6.8 DNA Sequencing

All DNA sequencing was performed by GATC Biotech, UK. The sequencing primers used are shown below:

Primer	Sequence
T7 forward	TAATACGACTCACTATAGGG
pABLM forward	GGGATTCGTCTGTTGCAGGAAGAGGAAGAA
pABLM reverse	CGACTTCAGCGAGACCGTTATAG

2.7 *Pichia pastoris* Growth

Two strains of *Pichia Pastoris* yeast were used in experiments: X33 wildtype and X33:SUC2 which is protease deficient. Both yeast strains were grown in Yeast Peptone D-glucose (YPD) medium, either in liquid cultures or on plates with added agar. A variety of volumes of liquid cultures were used, all of which were inoculated by taking a loop of cells from an agar plate and then incubating at 30 °C with shaking at 220 r.p.m.. YPD-agar plates were incubated at 30 °C.

2.7.1 *Pichia pastoris* Transformations

Two different plasmid vectors were used for transformations into *P. pastoris*. The constitutively expressed pGAPZ plasmid and the methanol induced pPICZ plasmid, both of which are available from Invitrogen.

The strain to be transformed into was streaked out from stocks stored at -80 °C onto YPD-agar plates and incubated as described for ~2 days. Small scale overnight liquid cultures were set up by inoculating 5 ml of YPD with cells from the plate and incubated as described for ~22 hours. Three larger scale overnight liquid cultures were then set up by adding 50, 100 and 150 µl of the previous culture to 250 ml YPD medium and incubated for ~22 hours. Three different concentrations are used to ensure that at least one culture will be close to the required optical density (OD). One of these cultures was selected and grown to an OD_{600} of 1.3-1.5. These cells are harvested by centrifugation at 5,000 r.p.m. for 10 minutes and washed three times in decreasing volumes of dH_2O at 4 °C. The cells were washed a final time in 1 M sorbitol at 4 °C and harvested again before resuspension in 375 µl of 1 M sorbitol at 4 °C.

The DNA to be transformed was prepared by linearising the plasmid using restriction enzymes. For pGAPZ this was achieved by incubating plasmid DNA obtained from a miniprep extraction with AvrII obtained from New England Biolabs Inc. for 30 minutes at 37°C. The enzyme was then heat denatured by heating to 70 °C for 10 minutes and the resulting linear DNA cleaned up by using a Qiagen Qiaquik Gel Extraction Kit following the manufacturers protocol. This protocol was the same for linearising pPICZ plasmids except the restriction enzyme used was PmeI also obtained from New England Biolabs Inc.

30 µl of the linearised DNA was added to 60 µl of the prepared cells and incubated at 4 °C for 5 minutes. The cells were transferred to an electroporator cuvette and electroporated. 1 ml of ice cold 1 M sorbitol was immediately added after electroporation and incubated at 30 °C for 1 hour. 1 ml of 30 °C YPD was added and the cells incubated at 30 °C for 4 hours with shaking at 220 r.p.m.. 250 µl of these cells was then plated onto YPDS-agar plates (YPDS is YPD which includes sorbitol at a concentration of 1 M) which contained the antibiotic zeocin at a concentration of 0.1 mg/ml. a separate 250 µl aliquot of cells was also plated onto YPDS-agar plates with a zeocin concentration of 0.5 mg/ml as well in an attempt to obtain high copy number inserts of the vector. These plates were incubated at 30 °C for 3 days.

2.8 Protein Detection and Analysis

2.8.1 SDS-PAGE

Protein detection was performed by separation using sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) and then western blotting to detect specific proteins using antibodies. Separation by SDS-PAGE was achieved using a 12% polyacrylamide gel and a Bio-Rad gel tank system. Samples were prepared by normalising cultures to an OD_{600} of 0.3 in 1 ml, harvesting by centrifugation at 13,000 × g for 1 minute and resuspending in 1× SDS loading buffer from Invitrogen. These samples were then vortexed and heated to 90 °C for 10 minutes before being vortexed again and then loaded onto the gel. ColorPlus Prestained Protein Marker from New England Biolabs Inc. was used as a molecular weight marker on all gels. The gels were run in 1X Tris/Glycine running buffer under the following conditions: 125 V, 35 mA for 15 minutes then 200 V, 35 mA for 35 minutes.

2.8.2 Western Blotting

After SDS-PAGE was complete, the proteins were electro-blotted onto a nitrocellulose or PVDF membrane using a Bio-Rad Mini Trans-Blot® Cell with these conditions: 100 V, 350 mA for 1 hour. The membrane was blocked in a 5% milk Tris Buffered Saline with Tween-20 (TBS-T) overnight with gentle shaking. To detect the MBP fusion proteins which are used in the GALLEX assay the membranes were incubated with a mouse anti-MBP primary antibody (Sigma Aldrich, UK) at a concentration of 1 in 4000 in TBS-T. For detection of His-tagged trpLE-CPT1 proteins the membranes were incubated with a mouse anti His-tag primary antibody

(Roche Diagnostics, UK) at a concentration of 1 in 5000 in TBS-T. The membranes were then washed in TBS-T for 3 × 10 minutes with shaking. After the washes the membranes were incubated with the secondary antibody, an alkaline phosphatase conjugated antimouse IgG antibody (Sigma Aldrich, UK) at a concentration of 1 in 10000. Again 3 × 10 minute washes with shaking were performed and the membrane was then developed. Developing was achieved using the SigmaFast™ kit (Sigma Aldrich, UK) following the manufacturer's protocol.

2.8.3 Coomassie Staining

Protein bands were detected by staining with Coomassie Brilliant Blue stain (56ml dH_2O, 4 ml acetic acid, 0.01 g Coomassie Brilliant Blue) overnight. The gels were subsequently destained in Coomassie destain (40% methanol 10% acetic acid in dH_2O) for 2-3 hours until protein bands could be clearly identified.

2.8.4 Silver Staining

To detect small peptides and proteins that do not stain well with Coomassie, staining with silver nitrate was used. Gels were soaked in fixer solution (60 ml 50% acetone, 1.5 ml 50% TCA, 25 µl 37% formaldehyde) for 15 minutes. The gel was then washed 3 times in dH_2O and soaked in dH_2O for a further 5 minutes before being washed 3 times and soaked in 50% acetone for 5 minutes. The gel was soaked in 60 ml 1 mM $Na_2S_2O_3$ solution for 5 minutes before being washed 3 times in dH_2O. The gel was soaked in stain solution (160 mg $AgNO_3$, 600 µl 37% formaldehyde, 60 ml dH_2O) for 8 minutes. Following 2 more washes in dH_2O, the gel was soaked in developer solution (1.2 g Na_2CO_3, 25 µl 37% formaldehyde, 25 mg $Na_2S_2O_3$, 60 ml

dH_2O) for 10-20 seconds. Once the protein bands could be visualised easily the reaction was quenched with 1% acetic acid in dH_2O.

2.9 GALLEX

2.9.1 Preparing Constructs for GALLEX

The GALLEX assay requires that the TM sequences to be studied be cloned into specific plasmids necessary for the assay. To measure homo interactions in GALLEX the pBLM100 plasmid was used, and to measure hetero interactions both pBLM100 and pALM148 were required (Schneider and Engelman, 2003). These plasmids were cut using the SpeI and SacI restriction sites and the wild type TM sequences to be studied were ligated in using phosphorylated synthetic oligonucleotides of the TM sequences. The protocols used are described in **Sections 2.6.1, 2.6.2** and **2.6.3**.

2.9.2 Preparing Mutants for GALLEX Assay

The wild type constructs were then subjected to Quikchange site directed mutagenesis to create all the mutant constructs studied. The protocol from the Quikchange kit produced by Stratagene and using Phusion® High-Fidelity DNA Polymerase (New England Biolabs Inc.) was followed in the construction of all the mutants. Each mutation was attempted with three different concentrations of template plasmid DNA present. A typical reaction was performed using 10 µl 5X Phusion buffer, 1 µl of dNTPs (100 mM of each dNTP), 1 µl of forward primer (125 ng/ml), 1 µl of reverse primer (125 ng/ml), 1 µl of Phusion DNA polymerase and either 1, 4 or 10 µl of template plasmid. The volume of each reaction was 50 µl and

this was made up using autoclaved dH_2O. The mutagenesis was carried out using 25 cycles of the following three steps: denaturation at 98 °C for 10 seconds, annealing at 68 °C for 30 seconds, then extension at 72 °C for 1 minute. An initial denaturation at 98 °C for 30 seconds was included before cycling and a final extension step at 72 °C for 10 minutes was included at the end of the cycles.

Following the thermal cycling the reactions were incubated with DpnI (New England Biolabs Inc.) for 1 hour at 37 °C to remove the methylated template DNA and leaving the newly synthesised mutant plasmid. This was then directly transformed into competent *E. coli* DH5α cells (prepared as described in **Section 2.6.4**) for repair and replication.

These plasmids were then isolated using the miniGeneJET™ Plasmid Miniprep Kit (Fermentas) and following the protocols included with the kit. The isolated plasmids were sequenced (GATC **Section 2.6.8**) to ensure they contained the desired mutations.

2.9.3 GALLEX Assay

The GALLEX assay is an *in vivo* assay performed in *E. coli* which can be used to detect interactions between TM domains (Schneider & Engelman, 2003). The assay is able to detect and measure both homotypic interactions as well as heterotypic interactions. Chimeric protein constructs are expressed in *E. coli* with the TM domain to be studied in the centre flanked by an MBP domain C terminally and a LexA domain N terminally. The MBP domain acts as a periplasmic anchor to ensure that the whole construct is inserted into the *E. coli* cell membrane in the correct orientation with MBP in the periplasm and the LexA domain in the

cytoplasm. The LexA domain acts as the reporter system in the assay. When two of the constructs come into close proximity, facilitated by the interaction of the TMs, the two LexA domains dimerise allowing them to bind to DNA. This LexA dimer is an effective repressor of *lacZ* transcription which results in inhibition of the expression of β-galactosidase (**Figure 2.1**).

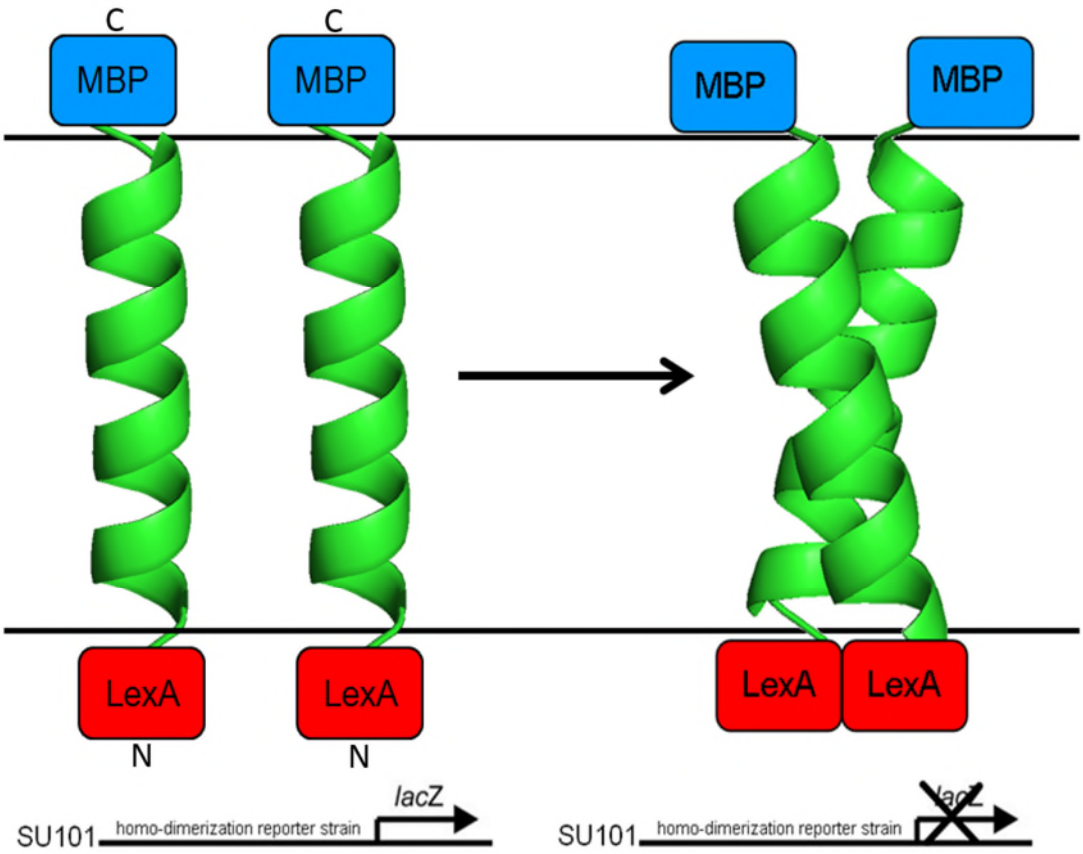


Figure 2.1 – Schematic of the homo GALLEX assay showing the N terminal LexA domain and the C terminal MBP domain flanking the TM domain sequence.

In the hetero assay the dimer formed is not between two wildtype LexA domains but is instead between one wildtype LexA which binds to the wildtype op+ DNA as well as a mutant LexA (LexA*) which binds to the mutant op+408 DNA. The *E. coli* SU202 strain has a mutated op+/op408 DNA region which allows only a LexA-LexA* dimer to bind and should exclude any LexA-LexA and LexA*-LexA* homo dimers from binding and inhibiting β-galactosidase expression. This exclusion

allows the study of hetero interactions even when homo interactions are also present (**Figure 2.2**).

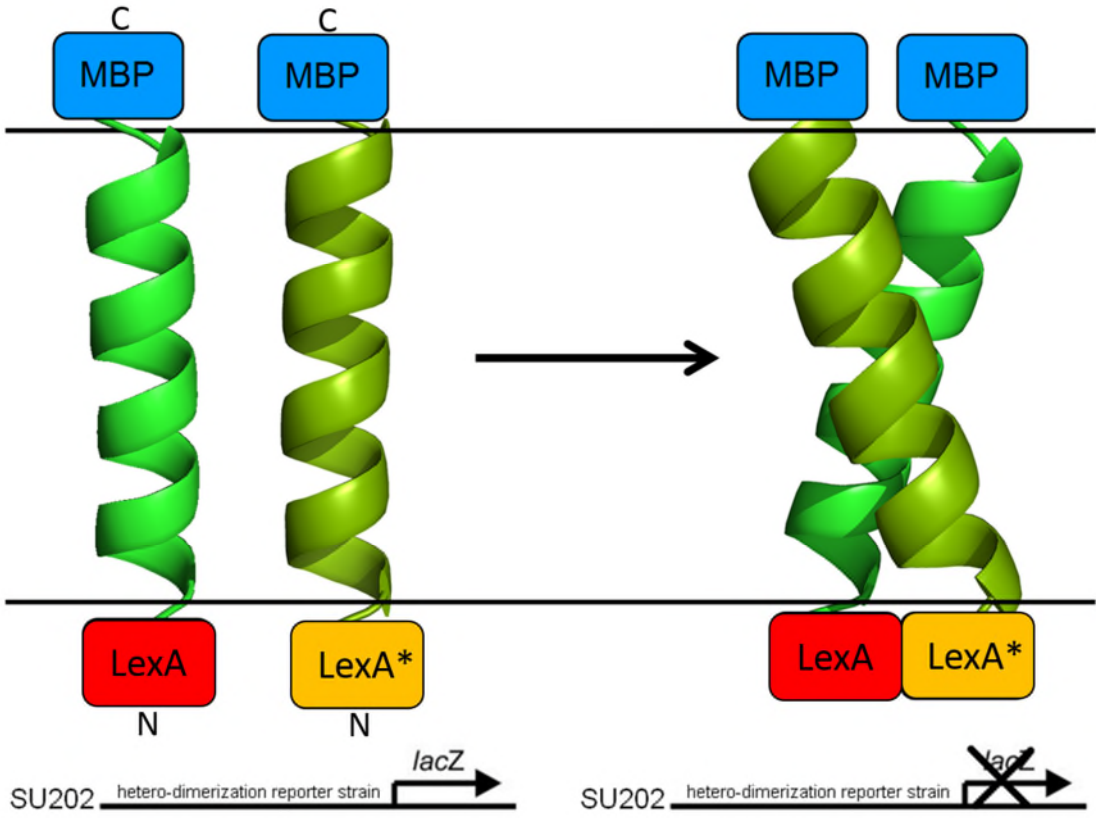


Figure 2.2 – Schematic of the hetero GALLEX assay showing the N terminal LexA/LexA* domain and the C terminal MBP domain flanking the TM domain sequence.

This inhibition can be detected by a simple colourimetric absorbance assay. Ortho-nitrophenyl-β-galactoside (ONPG) is added to lysed cells after growth with the chimeric protein constructs. ONPG is hydrolysed by β-galactosidase to form galactose and Ortho-Nitrophenol (ONP) which has a yellow colour and a maximum absorbance at 420 nm (**Figure 2.3**). The production of ONP is proportional to the level of β-galactosidase and so the concentration of ONP can be used to deduce the level of inhibition of lacZ expression.

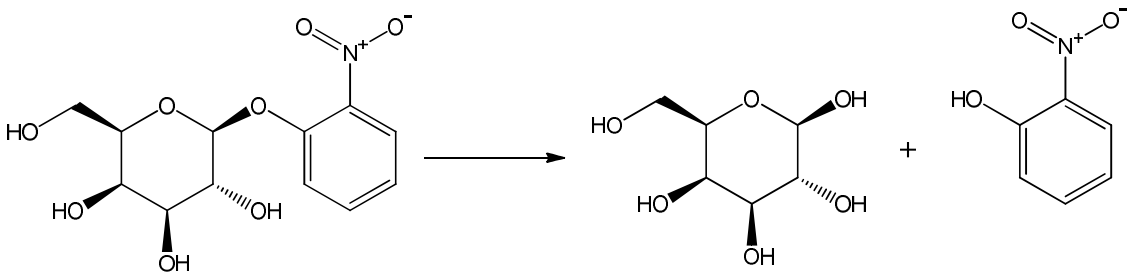


Figure 2.3 – Reaction showing the breakdown of ONPG (left) to galactose and ONP (right) by β-galactosidase. It is inhibition of this reaction and detection and quantification of the yellow colour of ONP that functions as the reporting mechanism in the GALLEX assay.

To measure homo interactions cultures were prepared from *E. coli* strain SU101 transformed with the desired TM in the pBLM100 plasmid. Outgrowths were grown up with the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 0.01 mM to induce expression of the chimeric protein via the lac operon until an OD_{600} of 0.6-0.8 is obtained. The antibiotics ampicillin, chloramphenicol and kanamycin are included at a final concentration of 100 µg/ml, 5 µg/ml and 5 µg/ml respectively. The SU101 strain is resistant to chloramphenicol and kanamycin and additional resistance to ampicillin is conferred by the pBLM plasmid if successfully transformed. A 50 µl aliquot of culture was taken and added to 900 µL 1× Z-buffer with 50 mM β-mercaptoethanol (1× Z-buffer = 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl and 1 mM $MgSO_4$). The cells were then lysed using 10 µl 0.1% SDS and 2 drops of chloroform and then thoroughly vortexed for 30 seconds. 200 µl of 4 mg/ml ONPG in 1× Z-buffer was added to start the reaction and after a time of 8 minutes 500 µl of 1 M Na_2CO_3 was added to stop the reaction. The solution was centrifuged at 13,400 x g for 15 minutes to sediment cell debris and then transferred to a cuvette. The absorbance at 420 and 550 nm was recorded. These absorbance values together with the absorbance at 600 nm, recorded earlier, were used to calculate the levels of β-galactosidase using **equation 2.1** where

t is the time in minutes and v is the volume of culture used in ml.

$$\beta\text{-}galactosidase\ units = \frac{1000(A_{420} - 1.75 \times A_{550})}{t \times v \times A_{600}} \qquad \textbf{2.1}$$

The expression level of the chimeric protein construct was detected using SDS-PAGE and Western blotting using anti-MBP antibodies described in **Section 2.7.2** and the β-galactosidase units calculated were normalised to the expression levels of each sample obtained by analysing the Western blots using ImageJ (Schneider, et al., 2012).

To measure hetero interactions using the GALLEX assay, cultures were prepared from *E. coli* strain SU202 transformed with one TM in the pALM148 plasmid and the other in the pBLM100 plasmid. This was done by sequentially transforming with the pALM148 plasmid, growing and making the cells competent and then transforming the pBLM100 plasmid. Outgrowths were grown with 0.01 mM IPTG to an OD_{600} of 0.6-0.8 as in the homo assay. The additional antibiotic, tetracycline, was added to a concentration of 3 µg/ml as resistance to this is conferred by the additional pALM plasmid. The rest of the hetero assay was performed in the same manner as the homo assay.

2.9.4 Controls

Controls were employed to check the correct insertion of the chimeric protein constructs into the *E. coli* membrane as well as to ascertain the expression levels of the proteins.

2.9.4.1 Spheroplast Assay

An outgrowth was taken from an overnight culture of *E. coli* NT326 transformed with the GALLEX plasmid to be tested and grown in the presence of 0.01 mM IPTG to an OD_{600} of ~0.6. The cells from a 1.5 ml volume of this outgrowth were harvested by centrifugation at 13,400 × g. for 2 minutes and resuspended in 0.5 ml buffer (100 mM tris-acetate pH 8.2, 0.5 M sucrose, 5 mM EDTA). 6 µl of 10 mg/ml lysozyme was added and incubated at 4 °C for 1 minute before adding 0.5 ml 4 °C dH_2O and incubated for a further 4 minutes. 20 µl of 1 M $MgSO_4$ was added before the resulting spheroplasts were pelleted at 13,400 × g and 4 °C for 2 minutes.

The spheroplasts were resuspended in 300 µl of buffer (10 mM HEPES pH7.6, 2 mM EDTA) and split into three equal sample fractions in 1.5 ml Eppendorf tubes. One fraction of the three, the whole spheroplast fraction, is then precipitated using trichloroacetic acid (TCA). 1 ml of 10% TCA is added and the sample incubated at 4 °C for 30 minutes before being pelleted at 13,400 × g for 15 minutes. All the supernatant is removed and 1 ml of acetone is added and incubated at 4 °C for 5 minutes before being pelleted again at 13,400 × g for 10 minutes. All the supernatant is removed and the pellet air dried.

One of the other fractions is treated with 2.68 µl of 19.6 mg/ml proteinase K and incubated at 4 °C for 30 minutes before being precipitated with TCA as described for the whole spheroplast fraction. The final fraction is freeze-thawed 5 times using a dry ice-ethanol bath and 37 °C water bath before added 2.68 µl of 19.6 mg/ml proteinase K and incubated at 4 °C for 30 minutes. This was then precipitated using TCA as described above.

All three fractions are then run on a standard 12% SDS-PAGE gel as described and the bands visualised by Western blotting using anti-MBP antibodies. MBP should be detected in all three fractions but due to the addition of proteinase K, which cleaves the MBP domain from the rest of the protein, in the last two, the band in this case will be seen at a lower molecular weight corresponding to just the MBP domain. This assay shows correct insertion because if the MBP domain is not exposed to the periplasm, proteinase K will not be able to cleave the domain in the second fraction in which whole spheroplasts are treated with proteinase K.

2.9.4.2 MalE Complementation Assay

The second control of insertion is growing *E. coli* NT326 cells, which lack native MBP, with the plasmids coding for the chimeric proteins on maltose minimal medium with no other carbon source. If the MBP domain is exposed to the periplasm, it should confer the ability to utilise the maltose as a carbon source on the cells and allow them to grow.

Maltose minimal medium is prepared by making sterilised 1 M $MgSO_4$, 1 M $CaCl_2$ and 20% maltose (weight/volume) solutions. Agar and dH_2O (final agar concentration 15 g/l) was autoclaved and to this, 20 ml of 5X M9 salts, 0.2 ml 1 M $MgSO_4$, 2 ml 20% maltose and 10 µl 1 M $CaCl_2$ was added. 5X M9 salts were made according to the recipe in Molecular Cloning vol. 3 (64 g $Na_2HPO_4 \cdot 7H_2O$, 15 g KH_2PO_4, 2.5 g NaCl, 5 g NH_4Cl in 1 litre of dH_2O).

2.9.4.3 Surface Expression Dot Blot

An outgrowth was taken from an overnight culture of *E. coli* and grown in the presence of 0.01 mM IPTG to an OD_{600} of ~0.6. The cells from a 2 ml volume of this outgrowth were harvested by centrifugation at 3,000 × g. for 2 minutes and resuspended in 0.5 ml buffer (100 mM tris-acetate pH 8.2, 0.5 M sucrose, 5 mM EDTA). 6 µl of 10 mg/ml lysozyme was added and incubated at 4 °C for 1 hour before adding 0.5 ml 4 °C dH_2O and incubating for a further 4 minutes. 20 µl of 1 M $MgSO_4$ was added before the resulting spheroplasts were pelleted at 3,000 × g. and 4 °C for 2 minutes. A 2 µl aliquot of this sample was pipetted directly onto a nitrocellulose membrane and allowed to dry. This membrane was then probed for MBP expression in the same way as the standard Western blot outlined in **Section 2.7.2**. Untransformed SU101 and NT326 cells were also included as controls to check for cell lysis.

2.9.4.4 NaOH Extraction

An overnight culture was inoculated 1:100 into fresh LB media with the required antibiotics and 10 µM IPTG. The culture was grown until an OD_{600} of 0.6 was reached and then the cells were collected by centrifugation at 13000 × g for 1 minute. The cell pellet was resuspended in 90 µl H_2O, 2.5 µl 1 M $MgCl_2$, 5 µl DNase (10 mg/ml), and 5 µl lysozyme (10 mg/ml). This mixture was incubated at room temperature for 1 hour and then cooled on ice. 150 µl of ice cold (~4 °C) H_2O was added and then 125 µl of this was taken as the whole cell fraction for analysis. 125 µl ice cold (~4 °C) 0.1 M NaOH was added to the remainder and vortexed for 1 minute before centrifugation at 13000 × g for 15 minutes. The supernatant was taken as the

soluble protein fraction, and the pellet as the membrane protein fraction. The whole cell and the soluble protein fractions were precipitated using TCA as described above in **Section 2.8.4.1**.

2.9.4.4. Expression Check of Chimeric GALLEX Proteins

When performing the homo GALLEX assay a sample of cell culture was taken prior to running the assay. These samples were used to normalise the assay's results to the protein expression level as well as to the cell density of the culture. The samples were normalised to OD_{600} for cell density and then analysed by SDS-PAGE and Western blotting using anti-MBP antibodies to specifically detect the chimeric GALLEX proteins as described in **Section 2.8.2**.

2.10 Construction of trpLE Expression Constructs

Synthetic oligonucleotides of the desired TM sequences were ligated into the pMMHb plasmid which had been cut using HindIII and BamHI restriction enzymes.

2.10.1 Expression

After the relevant plasmid was transformed into *E. coli* BL21 (DE3) pLysS as described above, a single colony was picked from the plate and incubated in liquid LB with antibiotics overnight at 37 °C and 180 r.p.m.. This starter culture was then used to inoculate 1 litre of LB in a 5 l flask and grown to an OD_{600} of 0.6 and then induced by the addition of IPTG. The final concentration of IPTG added was optimised for the individual constructs being expressed. After growth overnight at 37 °C these cells were harvested by centrifugation at 3000 r.p.m. and stored at -80 °C

after the supernatant was removed.

2.10.2 TrpLE Protein Extraction and Purification

The cell pellet was defrosted and resuspended in 30 ml lysis buffer (50 mM Tris-HCl, pH 8.0, 200 mM NaCl) with 3 ml BugBuster® HT Protein Extraction Reagent added. This solution was gently rocked at room temperature for 30 minutes and then lysed using a cell disruptor (Constant Systems Ltd., UK) at a pressure of 30 kPSI. The inclusion bodies and insoluble matter were pelleted and the supernatant removed; they were then resolubilised in guanidine buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 6 M guanidine HCl, 15 mM imidazole) by end over end rotation overnight.

The supernatant was applied to Ni^{2+} sepharose resin in a chromatography column which had previously been equilibrated in guanidine buffer and allowed to bind for ~1 hour. The remaining sample was allowed to drain from the column and the resin was washed in 10 bed volumes of guanidine buffer. The bound protein was eluted by the addition of 10 bed volumes of elution buffer (20 mM Tris-HCl, pH 7.0, 200 mM NaCl, 6 M guanidine HCl, 350 mM imidazole). The elution fraction was dialysed against dH_2O (several changes of dH_2O over ~ 1 day) using 3.5 kDa molecular weight cut off Snakeskin™ dialysis tubing and the resulting precipitated protein was lyophilised.

2.10.3 Protein Cleavage

The lyophilised protein was dissolved in formic acid and cleaved by the addition of ~1 g CNBr. This reaction was allowed to proceed for 3 hours and then

loaded into 2kDa molecular weight cut off dialysis tubing and dialysed against dH_2O overnight and lyophilised once complete.

2.10.4 High Performance Liquid Chromatography (HPLC) Purification

A Phenomenex 5 µm luna C5 100Å (250mm x 10 mm) reverse phase HPLC column with an Agilent 1100 series pump were used to separate the peptide from the cleaved trpLE tag. The mobile phase was composed of 95% H_2O, 5% isopropanol and 58% isopropanol, 37% acetonitrile, 5% H_2O. All solvents used were HPLC grade and were degassed prior to use. Elution of the proteins was monitored using the absorbance at 280 nm as all the proteins contained aromatic residues that absorb significantly at this wavelength. Fractions containing peptide were identified by mass spectrometry and SDS-PAGE and then dried and stored at -20 °C until required.

2.10.5 Mass Spectrometry

Bruker MicrOTOF and Autoflex MALDI-TOF instruments were used to analyse HPLC fractions and pure peptide samples. For analysis by MicrOTOF, HPLC fractions were analysed directly. Dry purified peptide was dissolved in 69% water, 20% isopropanol, 10% acetonitrile, 1% formic acid before analysis. Analysis was performed using positive ion mode between m/z of 200 – 3000. Spectra were recorded for 1 minute, averaged and deconvoluted. Matrix solutions were prepared for MALDI-TOF analysis by solubilising matrix (α-cyano-4-hydroxy cinnamic acid or sinapinic acid) in 150 µl acetonitrile for several hours and then adding 300 µl 0.1% TFA to this. 4 µl of matrix was mixed with 2 µl of sample and then 1 µl of this solution was spotted onto a MALDI plate and allowed to air dry for ~ 1 hour. Spectra

were acquired in the positive ion and linear mode. The mass range was calibrated using lysozyme.

2.11 Circular Dichroism (CD) Spectroscopy

Circular dichroism spectra were recorded using either a Jasco J815 or J1500 instrument (Jasco, Dunmow UK). Spectra were recorded between 190 and 280 nm using 2 nm bandwidth, 1 second response time, 0.2 nm data pitch and 100 nm/minute scanning speed. Peptide samples (50-200 µM) were prepared in 30 mM sodium phosphate buffer, pH 5.8, 20 mM NaCl, 100 mM DPC.

CD spectra were analysed using the online tool Dichroweb (Compton and Johnson, 1986, Manavalan and Johnson, 1987, Sreerama and Woody, 2000) to calculate secondary structure content using reference set 4 between 190 and 240 nm (Lees, et al., 2006).

2.12 NMR

2.12.1 NMR Sample Preparation

Isotopically labelled peptide was dissolved in NMR buffer (30 mM sodium phosphate buffer, pH 5.8, 20 mM NaCl) containing 100 mM deuterated DPC and 10% D_2O and mixed by vortexing. The sample was spun down and added to a 3 mm NMR tube ready for analysis. NMR samples were stored at 4 °C when not in use.

2.12.2 Protein Concentration Determination

All the CPT1 TM domain peptides used for NMR naturally contained a tryptophan residue. This allowed the protein concentration to measured using the absorbance of tryptophan at 280 nm.

2.12.3 NMR Experiments

A Bruker 700 MHz Avance spectrometer with a cryoprobe was used for all measurements. All spectra were recorded at 37 °C unless otherwise stated.

2.12.3.1 1D Experiments

1D proton spectra were acquired using a zgpr and zggpw5 pulse sequences which use pre-saturation pulses for water suppression. 64 scans of 32768 data points were recorded and processing was performed using an EM exponential window function.

2.12.3.2 2 and 3D Heteronuclear Experiments

HSQC spectra were recorded using the hsqcetfpf3gpsi2 pulse sequence with 2048 data points in the t_2 dimension and 128 planes in the t_1 dimension with 40 scans. HSQC spectra were processed using a GM Lorentz-to-Gauss window function and automatic polynomial baseline correction was applied. TOCSY (dipsi2esgpph) spectra were acquired with 2048 data points in the t_2 dimension and 128 planes in t_1 with ~300 scans with mixing times 50-100 ms. TOCSY spectra were processed with a QSINE window function and automatic polynomial baseline correction.

2.12.3.3 NMR Titration Experiments

HSQC spectra were used to monitor interactions between peptides during the titration of unlabelled peptide into an isotopically labelled NMR sample. These HSQC spectra were recorded as described above. A concentrated stock of the peptide in NMR buffer (30 mM sodium phosphate buffer, pH 5.8, 20 mM NaCl 100mM DPC) was prepared and this was used to add the unlabelled protein to the NMR sample. Once added the NMR sample was vortexed, spun down and allowed to stand for 1 hour to allow the peptides to equilibrate within micelles before the next spectrum was recorded.

2.12.3.4 Software for Data Processing and Analysis

All spectra were processed in Topspin 3.2 and analysed using either Topspin 3.2 or CCPN Analysis version 2.4.1.

2.13 Software

The CNS searching of Helix Interactions (CHI) program was used to search for possible structures that fit the experimental results obtained from using the GALLEX assay. Searches have been done for CPT1A TM1 and CPT1B TM2 homotypic interactions. Left and right crossing angles were searched from a starting angle of 25° with a starting distance of 10.4 Å between the two helices. The helices underwent a series of rotational steps of 45° between 0 and 360° with Molecular Dynamics (MD) and energy minimisation steps after each rotation. Predicted structures were generated and clustered if their backbone RMSD was ≤1 with a minimum of 10 structures required for a cluster. All resulting predicted structure

clusters were analysed and visualised using PyMOL, an open source pdb file viewer.

3. Homotypic Interactions of the CPT1A and CPT1B TM Domains

3.1 Introduction and Aims

This study focuses on the two catalytically active isoforms of CPT1: CPT1A found in the liver and other oxidative tissues; and CPT1B found in muscle. Although there has been a significant amount of work performed to investigate these enzymes' activity and topology (Zammit, et al., 1998, Zammit, et al., 1989, Zammit, et al., 1997), there has been relatively little work aimed at studying their inter and intramolecular interactions (e.g. oligomerisation) and complete three-dimensional structure. There have been structural studies which have yielded partial structures of both the N (Rao, et al., 2011) and C terminal (Morillas, et al., 2001, Morillas, et al., 2004) soluble domains, but no detailed structural work has been performed on the membrane spanning and juxtamembrane regions such as the loop.

The major differences between CPT1A and CPT1B are their ability to oligomerise and their sensitivity to inhibition by malonyl-CoA. It has also been shown that the sensitivity to inhibition by malonyl-CoA is modulated by the membrane environment in which CPT1A is located (Zammit et al., 2008). Specifically, there is evidence that the composition, fluidity and even curvature of the membrane can affect CPT1A's sensitivity to inhibition. There has been no modulation in inhibition sensitivity observed in CPT1B. There have also been reports demonstrating that the transmembrane regions of CPT1A interact strongly and could be a driving force for oligomerisation (Jenei et al. 2009 and 2011), and it is logical to

assume that the transmembrane and juxtamembrane regions would be the most affected by any changes to the membrane composition and environment. Finally, there is also significant evidence that interactions between the N terminal and catalytic C terminal domains regulate binding of malonyl-CoA and sensitivity to inhibition either within the enzyme, or possibly as part of a larger complex involving multiple CPT1 subunits. The loop and juxtamembrane region is two residues shorter in CPT1A than CPT1B, and indeed the differences in the loop region have been implicated to be important in explaining the different sensitivities to inhibition of these two isoforms (Borthwick, et al., 2006) (**Figure 3.1**).

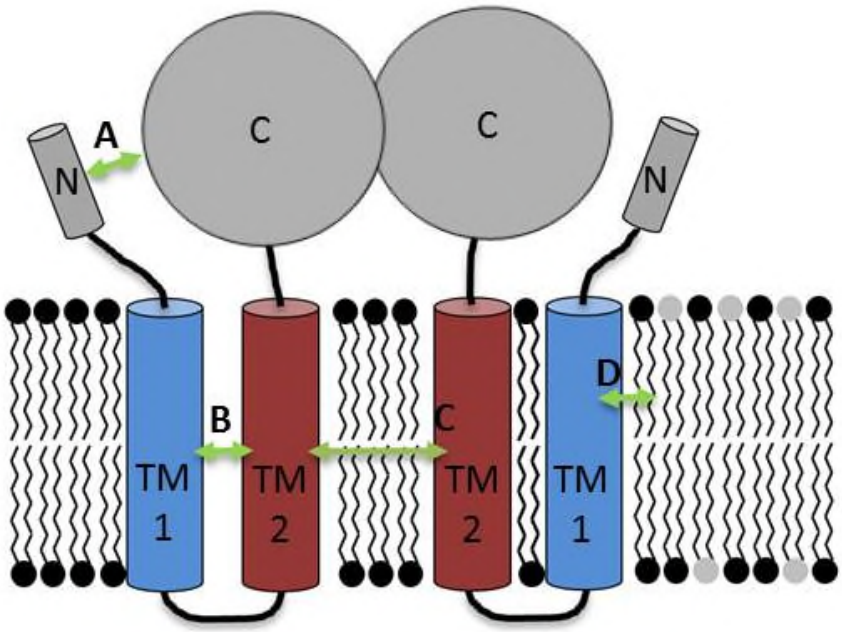


Figure 3.1 – The possible interactions that can affect the function of CPT1A and CPT1B including: (**A**) interactions between the N and C termini; (**B**) heterotypic intramolecular interactions; (**C**) homotypic intermolecular interactions; and (**D**) interactions between the TM domains of CPT1 and the lipids of the membrane.

This potential link between malonyl-CoA binding sensitivity and oligomerisation was the aim of the work detailed in this chapter. Specifically, we sought to identify and characterise any and all interactions between transmembrane

regions of CPT1A or CPT1B in a natural membrane bilayer (the *E. coli* outer membrane) using the GALLEX assay. The results were then used to establish whether there are differences in the oligomerisation propensities of CPT1A and CPT1B isoforms and, if so, whether this could help explain the observable differences in inhibition sensitivity. The results discussed in this chapter focus exclusively on homotypic self-association interactions of the TM domains of CPT1A and CPT1B. This type of interaction would be vital for the formation of any oligomers either initiated or supported by the TM domains.

3.2 Identification of the CPT1 TM Domains

The full length CPT1 sequences (UniProtKB database - CPT1A sequence IDs: human P50416, rat P32198, mouse P97742, and horse Q68Y62. CPT1B: human Q92523, rat Q63704, mouse Q924X2, cow Q58DK1, and pig Q8HY46) were analysed using both Topcons (Bernsel, et al., 2009, Tsirigos, et al., 2015) and TMHMM (Krogh, et al., 2001, Moller, et al., 2001, Sonnhammer, et al., 1998) in order to predict the location of the TM domains (**Figure 3.2**). These regions agreed well with previous topology studies (Zammit, et al., 1997) as well as with previous studies in the group (Jenei, et al., 2009, Jenei, et al., 2011). CPT1 sequences from several organisms were also investigated to assess any significant differences between species. Generally a high level of conservation was found between all the sequences investigated (**Figure 3.3**). CPT1B TM2 showed significantly lower sequence identity to the rest of the protein; however it still maintained a very high sequence similarity. The rat CPT1 enzymes were selected for study as these have been well-characterized in the past (Brown, 2003, Lewandowski, et al., 2013, Rao, et al., 2011), facilitating direct comparison of results obtained here to previous work.

The use of rat CPT1 could also allow animal samples to be used later in the project if required.

TM Domain	TMHMM	Topcons	(Zammit, et al., 1997)	GALLEX Experiments
CPT1A TM1	49-71	53-74	48-75	50-68
CPT1A TM2	104-122	104-125	103-122	104-120
CPT1B TM1	53-71	53-74	48-75	50-68
CPT1B TM2	105-124	105-126	103-122	106-122

```
          40          50          60          70
CPT1A TM1 KKKFIRFKNG IITGVFPANP SSWLIVVVGV ISSMHAKVDPS
CPT1B TM1 KKRLIRIKNG ILRGVYPGSP TSWLVVVMAT VGSNYCKVDIS

          90          100         110         120
CPT1A TM2 TLDTTGRMS- -SQTKNIVSGVL FGTGLWVAV IMTMRYSLKVLL
CPT1B TM2 CLPTRYGSYG TPQTETLLSMVI FSTGVWATG IFLFRQTLKL
```

Figure 3.2 – Results from using two TM domain prediction tools: Topcons and TMHMM, as well as the sequences used in the GALLEX experiments discussed in this chapter. The bold residues are those predicted by TMHMM, the highlighted are those predicted by Topcons, and the underlined residues are those which were used in the GALLEX assay.

The motifs discussed later were taken into account when picking the sequences to use as well as the reasonably strict length requirements for the GALLEX assay (Schneider and Engelman, 2003). To maintain alignment of conserved motifs within CPT1A and CPT1B TM domains the CPT1A sequence was shifted by two residues in TM2.

3.3 Interaction Motifs in CPT1 TM Domains

Once the TM domain sequences were identified (**Figure 3.2**), they were analysed for the presence of sequence motifs that are known to promote TM helix-helix interactions in membrane proteins (DeGrado, et al., 2003) (**Section 1.1.2**).

Three such interaction motifs were identified in the TM domains of CPT1A and B, namely the G-xxx-G motif (or other small-xxx-small motif, with serine or alanine in place of glycine), motifs of aromatic residues, and motifs of polar residues. Additionally, there are two conserved proline residues in both CPT1A and CPT1B TM1. Proline residues can affect helicity due to their rigid restricted peptide bond, which can serve to break the helix, introduce a kink, or in some cases cause a tighter helix to form. These motifs and residues are highlighted in **Figure 3.3**.

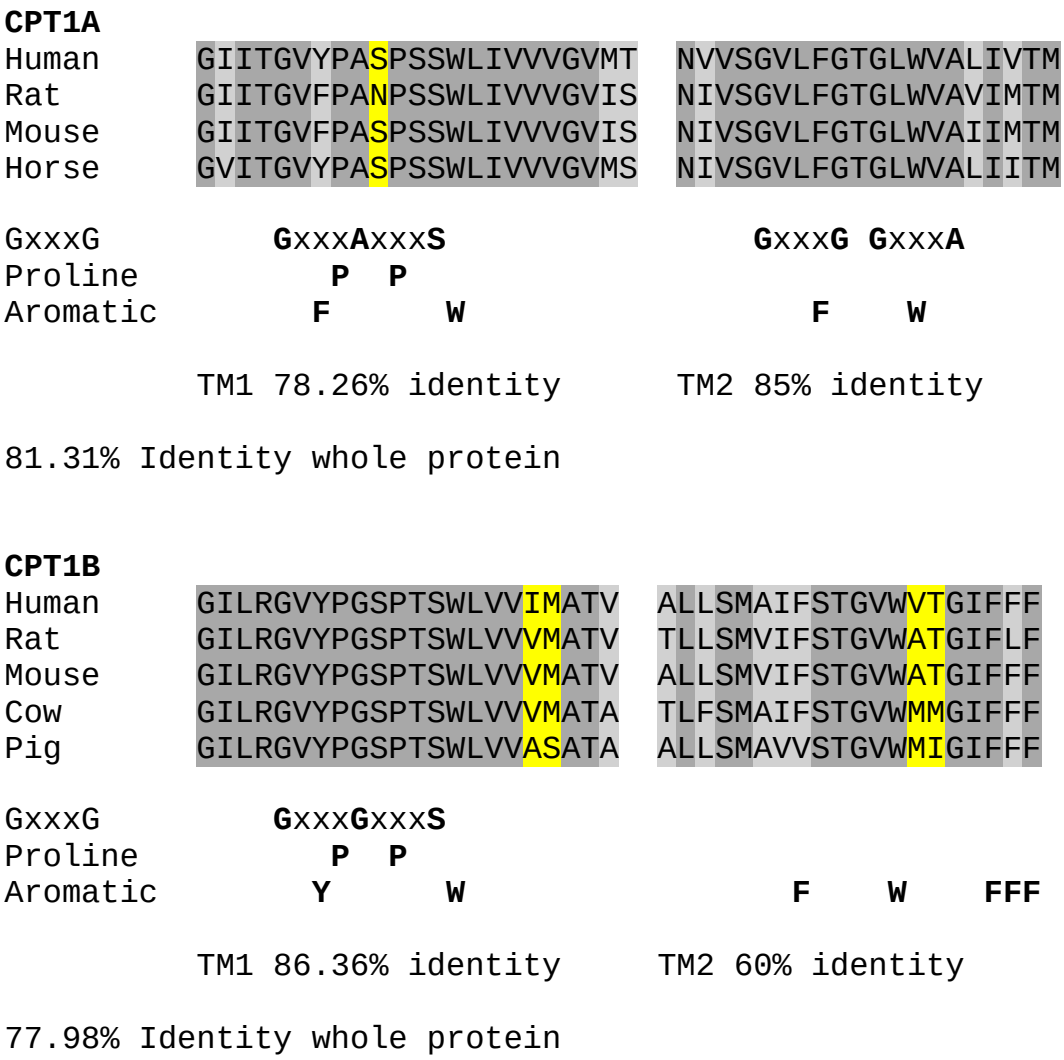


Figure 3.3 – Sequence similarity across species of CPT1. Identical residues shown in dark grey, similar residues in light grey, and differences highlighted in yellow. Sequences were taken from the UniProtKB database using only reviewed sequences. CPT1A sequence IDs: human P50416, rat P32198, mouse P97742, and horse Q68Y62. CPT1B: human Q92523, rat Q63704, mouse Q924X2, cow Q58DK1, and pig Q8HY46. Conservation calculations were performed using the online tool at: http://www.bioinformatics.org/sms2/color_align_cons.html. G-xxx-G, aromatic and proline interaction motifs and residues are shown below each sequence alignment.

3.4 The GALLEX Assay

The GALLEX assay is an *in vivo* genetic assay that measures the ability of TM domains to self-associate (see **Section 2.8.3** for more details). It offers several advantages over similar assays such as the TOXCAT assay (Russ and Engelman,

1999), the ToxR assay (Langosch, et al., 1996) and the POSSYCCAT assay (Lindner, et al., 2007) in that it can be used to measure heterotypic as well as homotypic interactions and utilises higher copy number plasmids for more rapid protein expression.

TOXCAT and GALLEX have been used to study several different proteins previously, from small single TM domain proteins such as the self-association of DAP12 (Wei, et al., 2013), to individual TM domains of GPCRs (Lock, et al., 2014) to novel synthetically designed TM domains (Nash, et al., 2015). Initially, both the TOXCAT and GALLEX assays were to be used to measure TM helix self-association, as these methods have shown very good agreement in past studies (Jenei, et al., 2011), however problems both with cloning and reproducibility meant that the TOXCAT assay was abandoned and GALLEX was selected as the sole method of choice here.

In order to perform any GALLEX assay the TM domain sequences to be tested must be inserted into the required plasmids for the assay. This was performed by ligating custom primers encoding for the TM domains of CPT1A and CPT1B into the pBLM100 plasmid after double digestion by SacI and SpeI restriction enzymes (see **Sections 2.6.1, 2.6.2, 2.6.3** and **2.8.1** for details). The same plasmids can be used to study heterotypic interactions using hetero GALLEX, and due to the topology of CPT1 the two TM domains are antiparallel to one another. To maintain this topology the sequences for TM1 of both CPT1A and CPT1B were inserted inverted to that of TM2. The inverted sequences are denoted by a lower case ‘i’ prefix (**Figure 3.4**).

The length of the CPT1 TM domains used had been optimised by a previous researcher in the group (Jenei et al., 2011). A selection of lengths between 16 and 22

residues of the predicted TM domains were tested in TOXCAT and GALLEX for acceptable signal. The 18 amino acid long TM domains for the controls had the greatest discrepancy between positive and negative controls and so this length was chosen for all the TM domains throughout.

CPT1A TM domains	**Sequence used for GALLEX experiments**
iTM1	VV**VILWSSPNAPFVG**TII
TM2	IVS**GVLFGTGLWVAV**IMT
CPT1B TM domains	
iTM1	MVV**VLWSTPSGPYVG**RLI
TM2	LLS**MVIFSTGVWATG**IFL
Controls	
GpA	ITLIIFGVMAGVIGTILL
$G_{83}I$	ITLIIFGVMA**I**VIGTILL

Figure 3.4 – Sequences of the TM domains used in all GALLEX experiments. This includes the four wild type CPT1A and CPT1B TM domains as well as the positive (GpA), and negative control ($G_{83}I$) sequences used. The residues which were mutated in the alanine scan experiments as well as the $G_{83}I$ mutation are shown in bold. All sequences are shown N terminus to C terminus.

The transmembrane domain of Glycophorin A (GpA), a cell surface TM protein found in erythrocytes, was used as a positive control throughout these assays. GpA is a well characterised protein and the TM domain from it is known to dimerise strongly via close packing of G-xxx-G motifs (Doura and Fleming, 2004, Russ and Engelman, 2000). A point mutation in the GpA TM domain ($G_{83}I$) strongly disrupts dimerisation and was used throughout as a negative control (sequences shown in **Figure 3.4**).

3.4.1 Controls for Homo GALLEX

Before GALLEX can be used to investigate interactions, the expression and correct orientation of the LexA-TM-MBP fusion proteins must be confirmed. Several methods were employed to achieve this. Firstly the protein expression level was measured using SDS-PAGE and Western blotting against MBP. Once expression was confirmed, NaOH extraction, which separates soluble from membrane-associating proteins (**Section 2.8.4.4**), was used to confirm that the GALLEX fusion protein was associating with the membrane. Finally, to check for correct orientation, the spheroplast and MalE complementation assays were used (see **Sections 2.8.4.1** and **2.8.4.2**).

The principle behind the spheroplast assay is to strip away the outer membrane of the *E.coli* cells, converting them into spheroplasts, and then using proteinase K to digest the exposed surface proteins. If MBP is exposed, as expected for correctly inserted GALLEX fusion proteins, then an anti-MBP Western blot will show full-length fusion protein and lower molecular weight cleaved MBP when proteinase K is added to intact spheroplasts as well as lysed cells. If the MBP domain is not exposed then proteolysis will only occur in lysed cells. In the MalE complementation assay the expression plasmids are transformed into *E. coli* NT326 cells which lack native MBP. These cells are then grown on M9 minimal media with maltose included as the sole carbon sauce. If MBP from the fusion proteins is present in the periplasm these cells should be able to utilise the maltose in the media and grow. Fusion proteins inserted in the incorrect orientation (i.e. MBP in the cytoplasm) will not grow.

In this study, these two assays presented significant problems with reproducibility and reliability (**Figure 3.5**). CPT1A TM2 grew well in the MalE complementation assay but none of the other CPT1 TM domains ever did, however they consistently gave good expression results and sensible results in the GALLEX assay itself. In the spheroplast assay incomplete or sometimes no proteolysis was observed on the addition of proteinase K, and these inconsistent results were not reflected in expression or other assay data. Due to these issues a simpler assay was developed to supplement and potentially replace them.

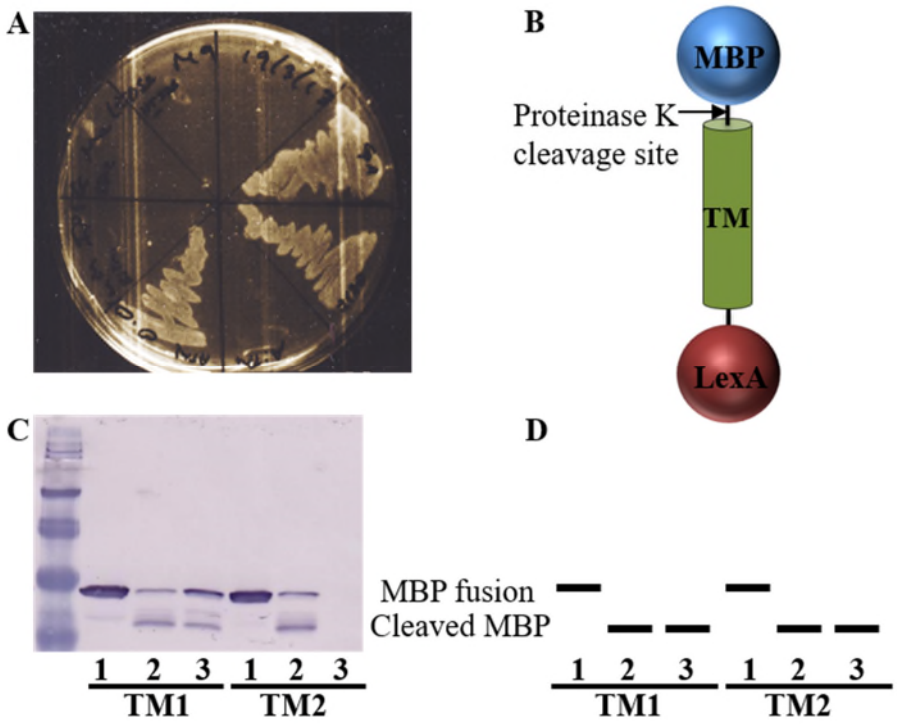


Figure 3.5 – Representative data obtained from old GALLEX controls (**A**) the MalE complementation assay showing only the controls and CPT1A TM2 growing. (**B**) Schematic to show the expected proteinase K cleavage site on the GALLEX fusion protein. (**C**) An anti-MBP Western blot of samples from the spheroplast assay where lane (**1**) shows whole spheroplasts before proteolysis, lane (**2**) shows whole spheroplasts after proteolysis, and lane (**3**) shows whole spheroplasts after proteolysis and cell lysis. (**D**) The expected pattern of the MBP protein bands from the spheroplast assay.

This assay was based on the same principle as the spheroplast assay, but omits the troublesome proteinase K digestion step and uses Western blotting to directly detect surface expression by using dot blots. This method was tested with cell lines both with and without native MBP to prove that only cell surface proteins were being detected **(Figure 3.6)**. The theory was that NT326 cells which do not have native MBP should not be visible unless transformed with a GALLEX construct and that SU101 cells, which do contain native MBP, with no transformed GALLEX construct would only be detected on the blot if the cells are being lysed (**Figure 3.6**). It was important to ensure that the spheroplasts were remaining intact and only the surface proteins were being detected, as without this conformation the assay does not provide information on the orientation and membrane insertion of the protein (**Section 2.8.4.3**). Once satisfactory expression and orientation were established, the GALLEX self-association measurements were performed.

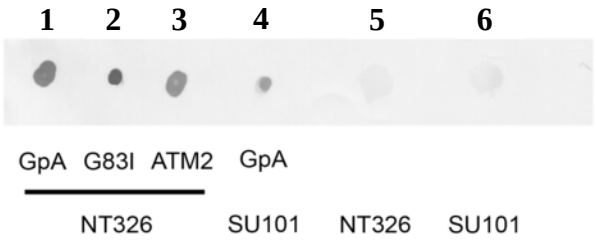


Figure 3.6 – Representative results from the spheroplast dot blot assay developed to increase reliability and reproducibility. Dots **1-3** have no native *E. coli* MBP but have been transformed with the GALLEX expression plasmids so any signal observed is due to the GALLEX fusion protein being expressed. Dot **4** has native MBP as well as a transformed GALLEX expression plasmid. Dot **5** has no native MBP and has not been transformed – no MBP should be present in this sample at all. Dot **6** has native cytosolic MBP only – this sample shows that the cells are not being lysed significantly and only surface protein expression is being detected as there is a similar signal to that of dot 5 where no MBP is present.

3.4.2 Optimising Homo GALLEX

There are some parameters which can be optimised in the GALLEX assay: length of the TM domain, IPTG induction concentration and the length of time the samples are left to react before measurement. The length of the TM domain has already been discussed above (**Section 3.4**) and a length of 18 residues was used here.

From previous experiments and literature an IPTG induction concentration of 10 µM was chosen here (Cymer, et al., 2013). Although protein expression level was increased at higher IPTG concentrations, up to 100 µM was tested, it was felt that subtle differences between interactions were being lost when more protein was present. These conditions were tested in preliminary experiments, where a range of IPTG concentrations between 2.5 and 10 µM were used, and results were found to agree with the previous findings that 10 µM IPTG was optimal for these experiments (results shown in **Figure 3.7**). A clear degradation in range between positive and negative results can be seen when moving from 10 µM through to 2.5 µM.

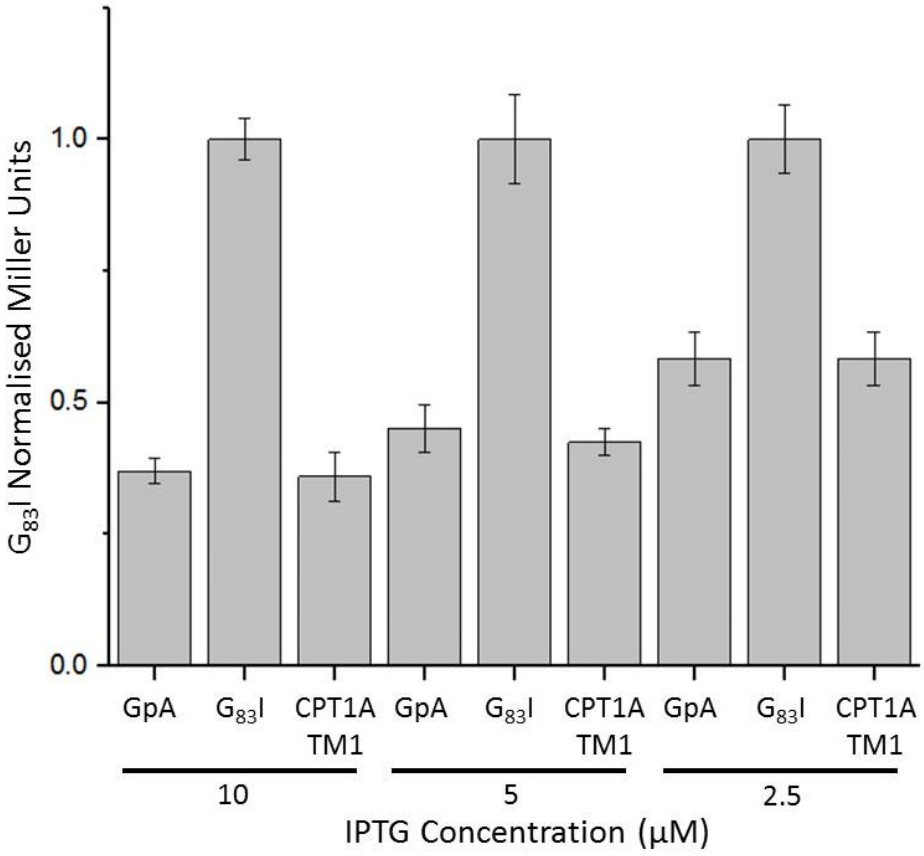


Figure 3.7 – The optimisation of IPTG concentration used in GALLEX experiments. Three concentrations were tested with both positive and negative controls as well as one of the TM domains from CPT1A to ensure this was representative for the samples to be tested as well as the controls. All results have been adjusted for expression levels and normalised to the results for $G_{83}I$ which have been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

In homo GALLEX assays the results can be normalised to the protein expression levels of each sample in order to eliminate higher protein concentration from being interpreted as a stronger signal. This was done by analysing each sample by SDS-PAGE and then Western blotting against MBP. This allowed the protein concentration in each sample to be visualised and analysed using ImageJ, where each band was isolated and analysed by measuring the area under the plot produced by ImageJ. However this initial method led to subjectivity as to where the baseline should be placed to measure the area of the Western blot band and not background.

This was especially problematic on blots with suboptimal staining resulting in high background values potentially obscuring small expression differences.

To improve on this, significant optimisations, regarding how the Western blotting data was analysed were incorporated as studies progressed. A rolling ball background subtraction algorithm in ImageJ was introduced to improve blots with high background staining. This allowed the background to be brought down to pure white in between the bands on clean blots hence removing the baseline ambiguity problem. In addition it was found that significantly more accurate and reproducible results could be obtained by using ImageJ to analyse the whole gel rather than isolating each band (**Figure 3.8**). These optimisations helped a great deal in minimising the variability between samples seen in the earliest experiments and were incorporated into all future experiments.

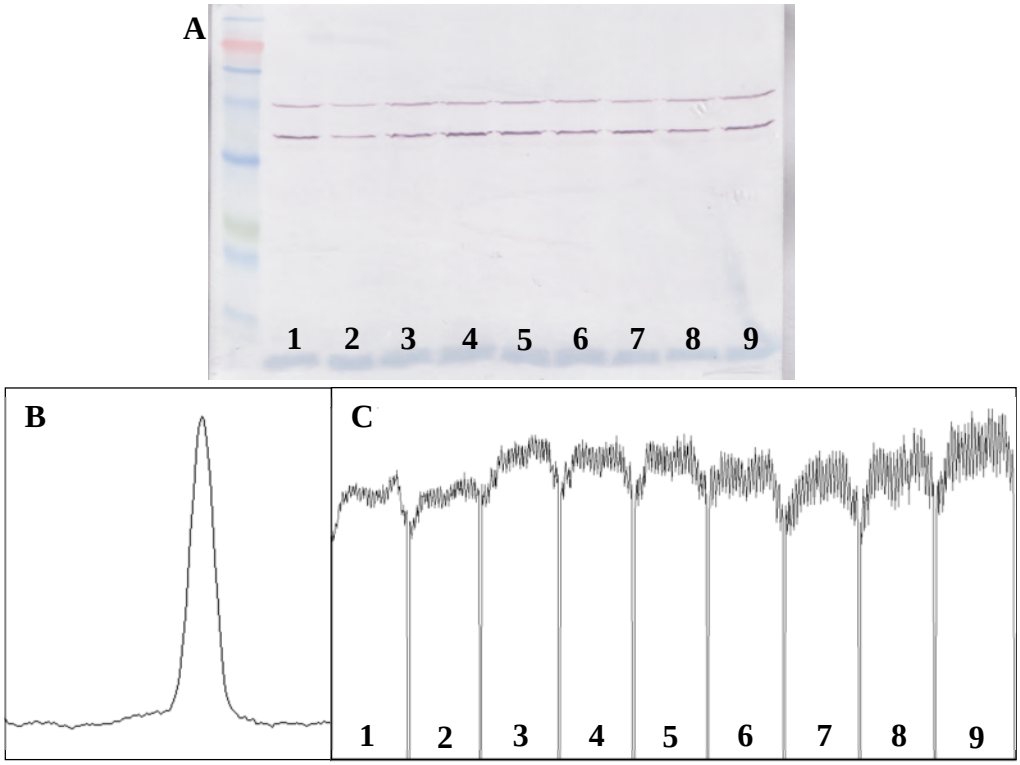


Figure 3.8 – Representative data produced during analysis with ImageJ. (**A**) An example Western blot obtained from the analysis of nine samples used in a GALLEX assay. The upper band corresponds to the GALLEX fusion protein, and the lower band to the native MBP present in these cells (SU101). (**B**) An example of the type of data plot obtained using the initial method of analysis in ImageJ. A plot like this was produced for each band on the gel. (**C**) Example of data obtained after the optimisations in the analysis described above. This data shows the whole gel analysis with each numbered section corresponding to each numbered lane on the blot. The background subtraction algorithm allows the background to be reduced all the way to zero in this plot.

The GALLEX measurements were always made in triplicate, using both technical and biological repeats. It was discovered that the errors were not significantly larger when only using technical repetition. This allowed significant streamlining of the assay. Finally, error in the GALLEX measurements was estimated using a one way ANOVA and Fisher least significant difference between the means with a confidence interval of 95%.

3.4.3 Transmembrane Domain Self-Association in Wild Type CPT1 Isoforms

Previous studies have shown that CPT1A TM2 self-associated (Jenei et al., 2009 and 2011) and there was some information to suggest that CPT1A TM1 did as well, but there was very little information available with regards to CPT1B. The wildtype sequences of the four TM domains from CPT1A and CPT1B were tested, and the results (as well as those of the positive and negative controls) are shown in **Figure 3.9**. All the homo GALLEX data shown were normalised to expression levels using the optimised technique for quantifying bands from Western blots (Section 3.4.2) as well as the value obtained for the negative control ($G_{83}I$). All four of the CPT1 TM domains interacted strongly, yielding β-gal activity similar to that for the positive control, GpA. Indeed, CPT1A TM1 and CPT1B TM2 appeared to interact more strongly than the TM domain of GpA. This contradicts previously reported results (Jenei et al., 2009) stating that CPT1A TM2 self-interacts but TM1 does not. This discrepancy is very likely due to the length of the TM domain constructs used in the previous study, as it has since been shown that the length of the TM domains used in both GALLEX and TOXCAT can significantly affect the signal observed in the assays. It does now appear that not only does CPT1A TM1 interact, but it does in fact have stronger homo interactions than TM2. These results also show a difference between CPT1A and B: in CPT1A, TM1 self-associates more strongly than TM2; in CPT1B this strength of association is reversed.

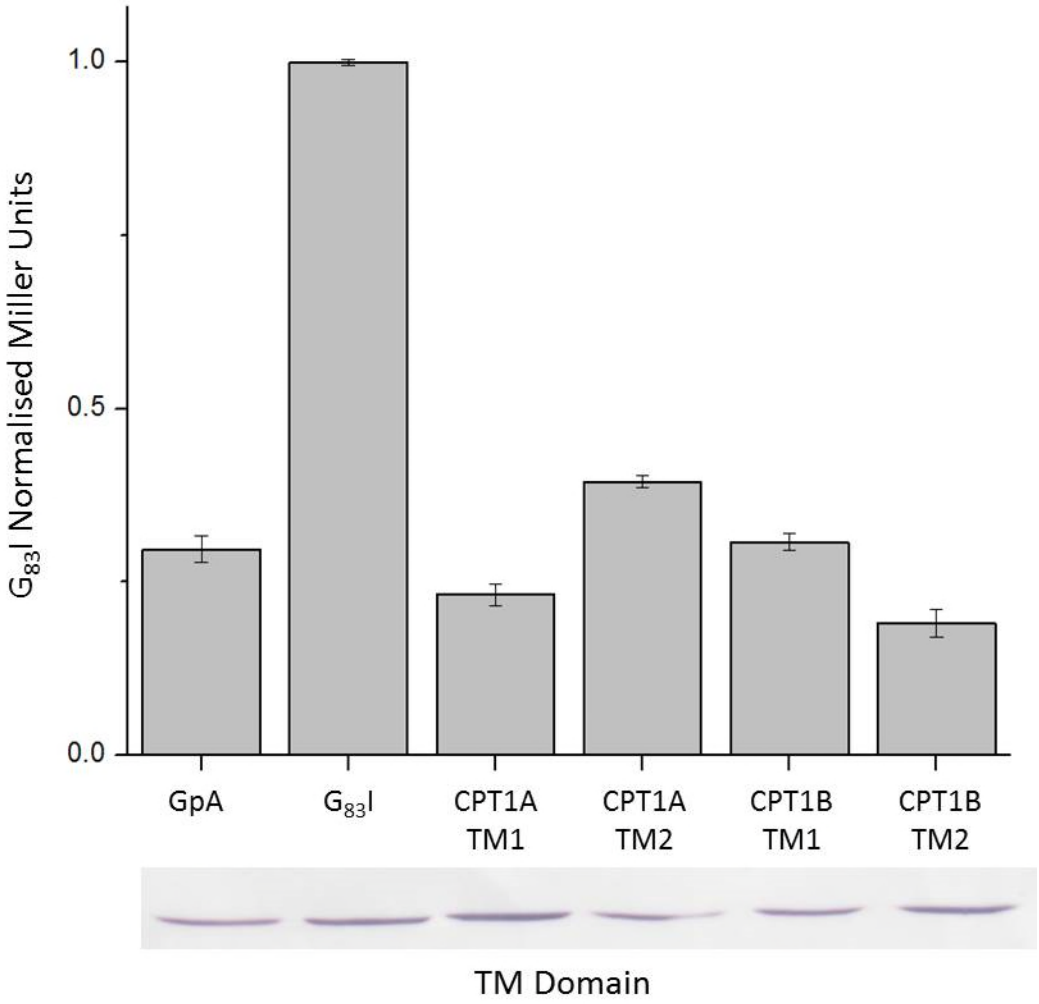


Figure 3.9 – Homo interactions of the wild type TM domains of CPT1A and B as measured in the homo GALLEX assay. All results have been adjusted for expression levels and normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. The relative expression levels measured by anti-MBP Western blot of each GALLEX fusion protein used to normalise the results are shown beneath each lane.

3.4.4 Sequence Dependence of CPT1A TM1 Self-Association

Once the relative strengths of homotypic interactions were confirmed for all four of the wild type CPT1A and CPT1B TM domains, a comprehensive mutagenesis strategy (Alanine-scanning mutagenesis) was employed to investigate the sequence dependence of these interactions. Each of the core twelve residues at the centre of each TM domain was mutated to Alanine or Isoleucine (depending on

the size of the residues) via Quikchange site directed mutagenesis. The rationale for this strategy was that a significant change in the steric bulk of the amino acid involved in helix-helix interaction would disrupt favourable packing and thus destabilise the interaction. In addition, changes in the polarity or aromaticity of an amino acid (via Alanine or Isoleucine substitution) could also destabilise helix-helix interactions.

The results for CPT1A TM1 are shown in **Figure 3.10**, again normalized to expression level and negative control and compared to wild-type. Interestingly, the three residues predicted to be part of a G-xxx-G like motif (S_{61}-xxx-A_{57}-xxx-G_{53}) did not show any disruption of the wildtype strength of interaction. Mutation of one or both proline residues (P_{56} and P_{59}) showed a small increase in interaction strength, but surprisingly this increase was less than other mutations. F_{55} showed the largest disruption from wildtype interaction levels, suggesting that this residue lies on the helix-helix interaction surface of CPT1A TM1. L_{63} stands out as weaker interacting than most of the other mutations tested and would lie on the same face of the helix as F_{55}; however this result is not statistically significantly different to the wildtype level of interaction, so is unlikely to be important.

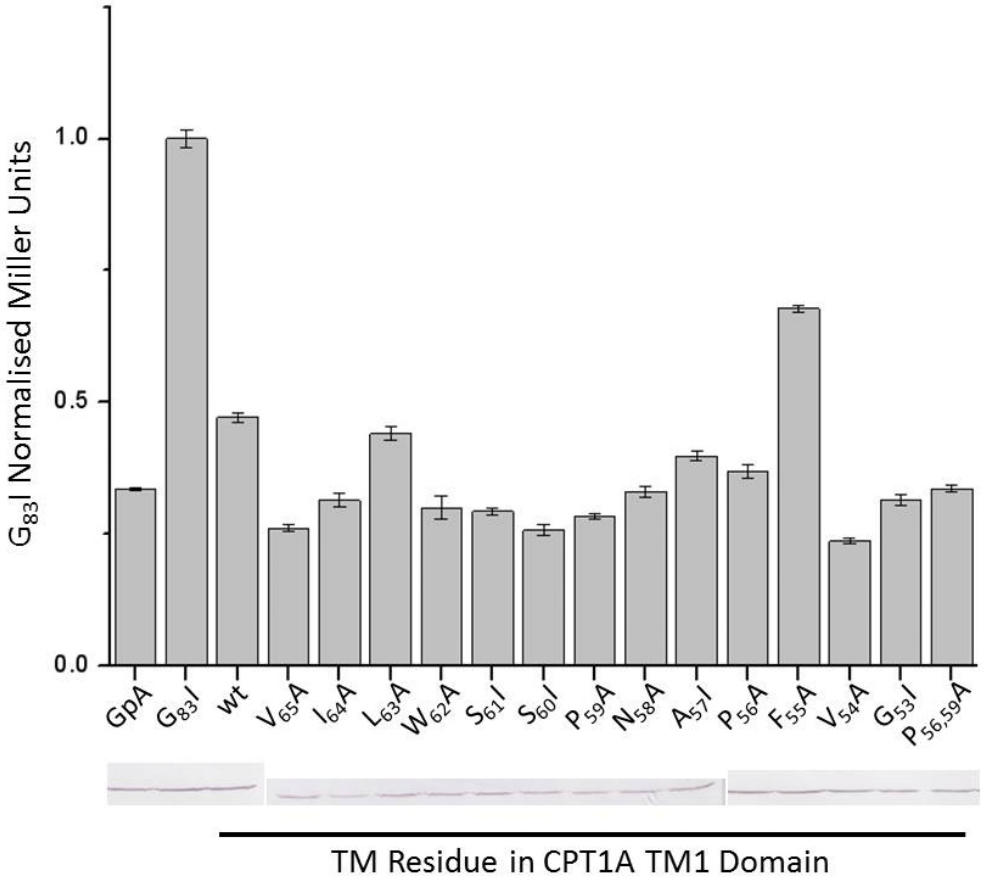


Figure 3.10 – Homotypic interactions of the mutant TM domains of CPT1A TM1 as measured in the homo GALLEX assay. All results have been adjusted for expression levels and normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. The relative expression levels measured by anti-MBP Western blot of each GALLEX fusion protein used to normalise the results are shown beneath each lane.

To complement the results obtained from homo GALLEX experiments, computational models were generated using the crystallography and NMR system (CNS) (Brunger, 2007, Brunger, et al., 1998) searching for helical interactions (CHI) software (Adams, et al., 1995, Adams and Brunger, 1997, Adams, et al., 1996). CHI is a set of commands using CNS that allow for global searching for protein helix – helix interactions. The techniques used were developed by modelling TM domains from glycophorin A and phospholamban (for more detail see **Section 2.12**). Structures predicted using CHI were grouped into clusters with greater than 10 members and a backbone RMSD ≤ 1. All clusters were analysed by collating the data

output from the CHI simulations in the form of a free energy of interaction for each residue as well as by eye using Pymol (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC). All the CHI searches were performed on the 18 amino acid sequences used in the GALLEX assay as well as on the complete predicted TM domains in case the length of the helices affected the interactions found. All of the relevant structures found that are discussed here were found to be present in both lengths tested.

Nine clusters were identified after running a CHI search for CPT1A TM1 homodimers, of which three were found to be symmetrical interactions (those between the same amino acid residues on each helix). One of these clusters showed the phenylalanine (F_{55}) identified in the homo GALLEX alanine scan at the dimerisation interface (**Figure 3.11**). The results for this cluster indicated that W_{62} and N_{58} would also lie along this interface, but neither of these residues showed significant changes on mutation, and so are presumably not important for interaction. The break in the helix and slight kink seen in the model is caused by the two proline residues.

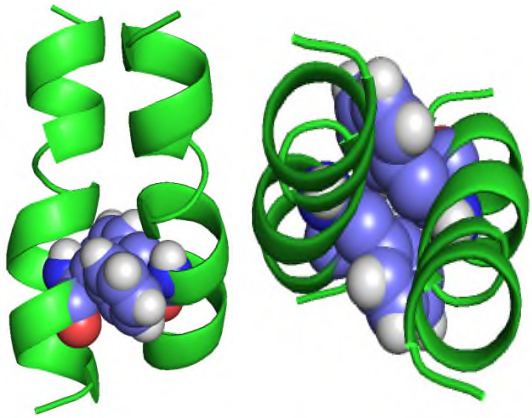


Figure 3.11 – Model of CPT1 TM1 homotypic interactions. The predicted structure from the cluster found in the CPT1A TM1 homo interactions CHI search, with the helix depicted as a cartoon representation (green) and the F_{55} residue shown as a space filling representation (coloured by atom). The model is shown from the side of the helices (left) as well as down the length of the helices (right) to show the proximity of the F_{55} residues.

3.4.5 Sequence Dependence of CPT1A TM2 Homotypic Interactions

The data shown in **Figure 3.12** was previously recorded by another researcher in the group, Zsuzsana Jenei (Jenei, et al., 2011). It is presented and discussed here as the outcome is relevant to further data collected in this project in the following chapter. The proposed G_{107}-xxx-G_{111} motif showed a significant disruption of interaction on mutation to isoleucine; however the second proposed G_{113}-xxx-A_{117} motif did not. While mutation of the first residue (G_{113}) showed a large disruption, mutation of the second (A_{117}) was not significantly different to wildtype levels. Instead V_{116} was shown to also be important for CPT1A TM2 homotypic interactions. As this data was collected by a previous researcher there may be subtle differences in how this data was collected and normalised. As discussed above there were some significant optimisations incorporated into the normalisation protocol during this project, but all results of importance found previously were corroborated in these studies including the optimisations.

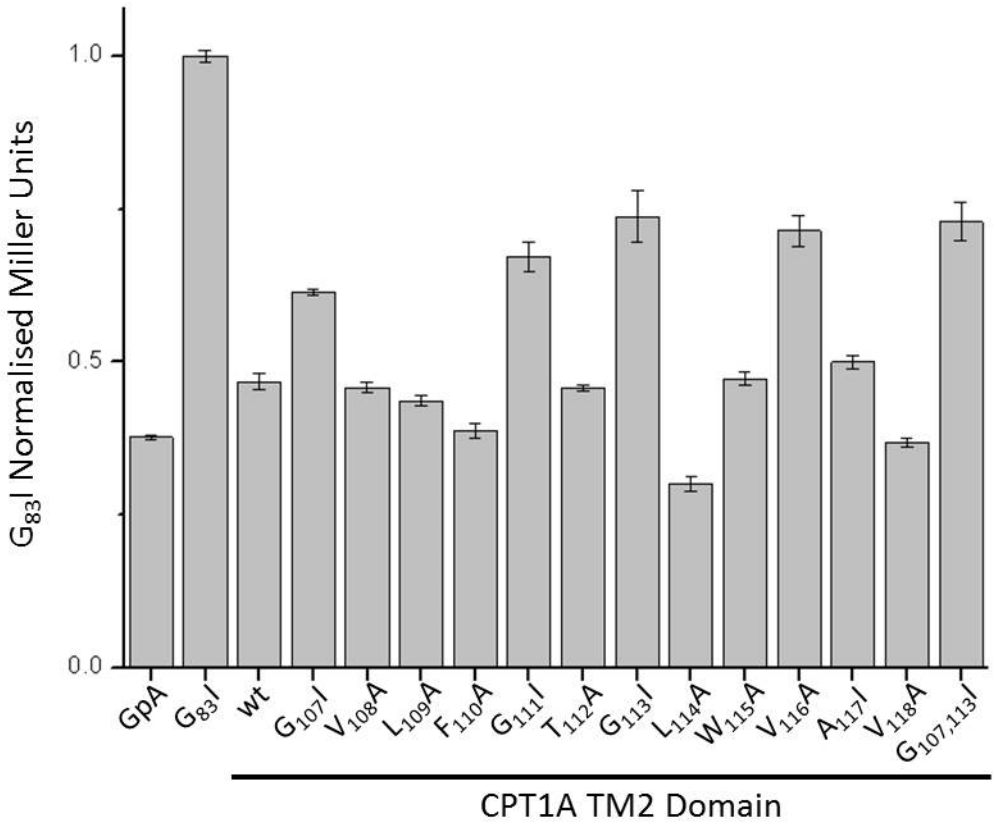


Figure 3.12 – Homotypic interactions of the mutant TM domains of CPT1A TM2 as measured in the homo GALLEX assay. All results have been adjusted for expression levels and normalised to the result for G_{83}I which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. Data collected by Zsuzsana Jenei (Jenei, et al., 2011).

CHI models of CPT1A TM2 were generated and analysed as part of this project however. Eight clusters were identified from this modelling. From the previous GALLEX experiments two distinct interfaces had been shown to be potential homodimer interfaces. These two interfaces could not both be active concurrently in a dimer as they lie on opposite faces of the TM domain helix. It also seemed unlikely that one could interact with the other in an asymmetric manner as they would not be present at the same depth in the membrane. Due to this symmetric models were expected, and indeed CHI models were found that represented both of these interface states (**Figure 3.13**).

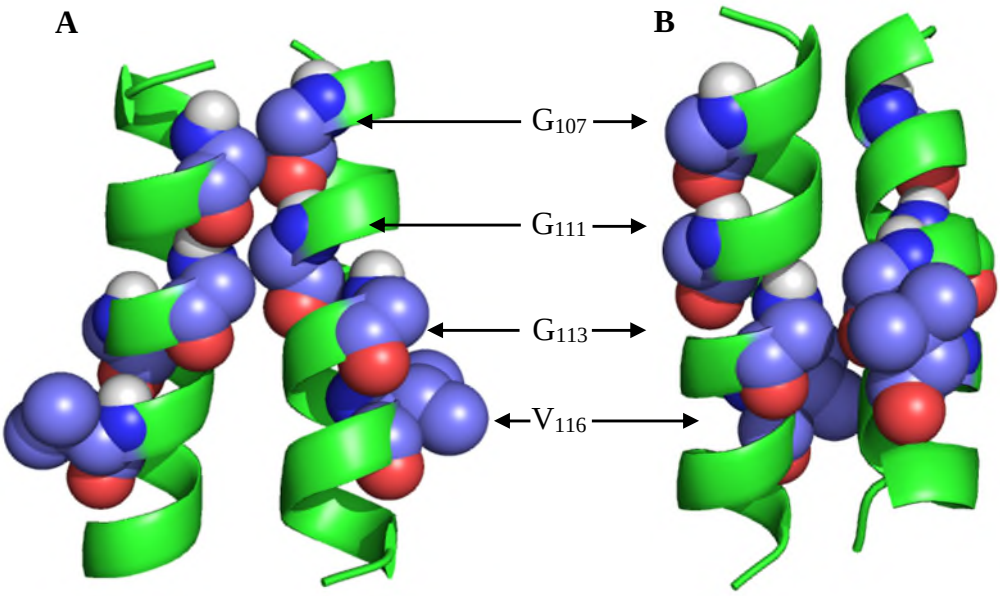


Figure 3.13 – Model of CPT1A TM2 homotypic interactions. The predicted structures from the clusters found in the CPT1A TM2 homo interactions CHI search, with the helix depicted as a cartoon representation (green) and the residues shown as a space filling representation (coloured by atom). (**A**) Structure illustrating the G_{107} G_{111} interface on one side of the helix, and (**B**) the structure illustrating the G_{113} V_{116} interface on the other side of the helix.

3.4.6 Sequence Dependence of CPT1B TM1 Homotypic Interactions

The same strategy of alanine scanning mutagenesis was applied to the two TM domains of CPT1B and results are shown in **Figure 3.14**. Similar to CPT1A TM1, the predicted G-xxx-G like motif (including residues S_{61}, G_{57}, and G_{53}) in CPT1B TM1 did not demonstrate large disruption to interaction on mutation of any of these residues. The only residue that did show significant disruption on mutation in CPT1B TM1 was Y_{55}. There was also a small but statistically significant weakening effect on mutation of G_{57} and G_{53} but this was not the greatest effect observed. The tyrosine residue (Y_{55}) identified is in the same location in the protein (position 55) as the phenylalanine (F_{55}) shown to affect TM1 – TM1 interactions in CPT1A TM1. As both of these residues have aromatic side chains and mutation to alanine strongly weakens these homo interactions, it seems logical to suggest that an aromatic residue at position 55 is critical for homodimers formation, and that it is the aromatic side chain, and not the backbone that is involved. The residue L_{63} would also be on the same face of the helix as Y_{55}, however here unlike CPT1A, the $L_{63}A$ mutation does show some weakening of the interaction compared to the wildtype. Data is not shown for the $W_{62}A$ mutation because it proved extremely difficult to work with. On induction of expression of this construct with IPTG in preparation for the GALLEX assay, a sudden drop in cell density was observed and consequently sufficiently dense samples could not be obtained to run the assay. This effect was not seen with any other construct used throughout all the GALLEX experiments.

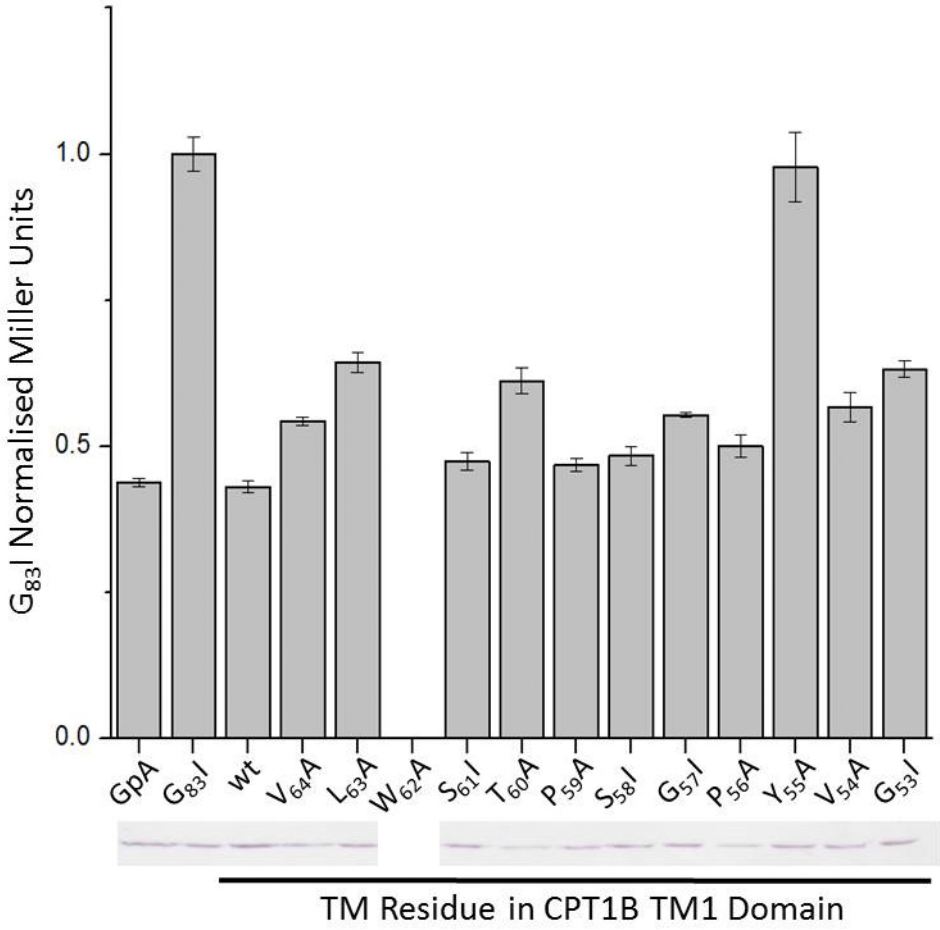


Figure 3.14 – Homotypic interactions of the mutant TM domains of CPT1B TM1 as measured in the homo GALLEX assay. All results have been adjusted for expression levels and normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. The relative expression levels measured by anti-MBP Western blot of each GALLEX fusion protein used to normalise the results are shown beneath each lane.

CHI models were once again generated to check for agreement with the GALLEX experimental results. Four clusters were predicted from the CHI search for homo interactions in CPT1B TM1. In this case, one cluster was found that showed close proximity of Y_{55} residues, possibly forming a hydrogen bond, shown to be important in the GALLEX results (**Figure 3.15 A**). However another cluster was also observed that showed packing of the predicted S_{61}-xxx-G_{57}-xxx-G_{53} (**Figure 3.15 B**) which was not seen at all experimentally.

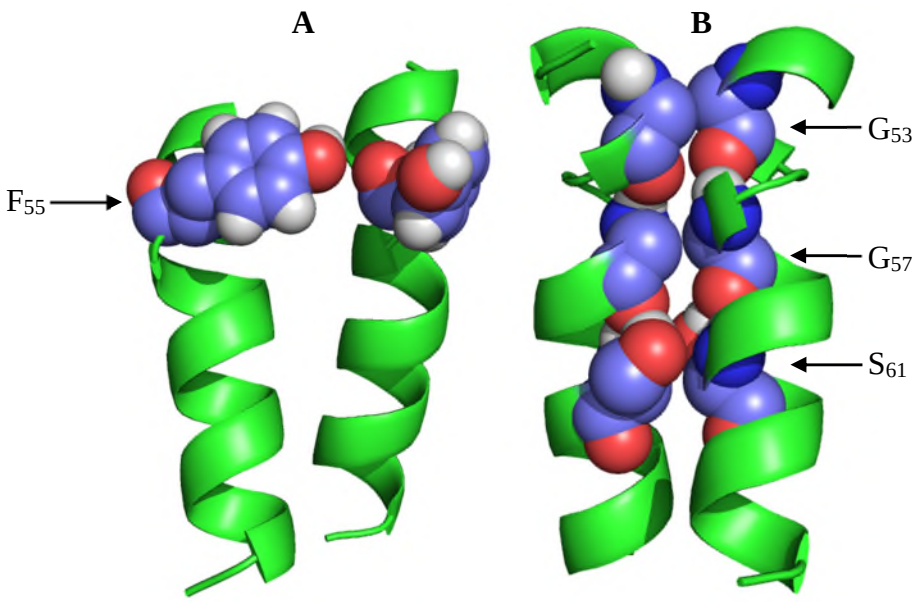


Figure 3.15 – Model of CPT1B TM1 interactions. The predicted structures from the clusters found in the CPT1B TM1 homo interactions CHI search, with the helix depicted as a cartoon representation (green) and the residues shown as a space filling representation (coloured by atom). (**A**) The cluster found which agrees with the GALLEX data showing the Y_{55} residues in close proximity, and (**B**) the cluster showing the original predicted G_{53}, G_{57}, and S_{61} small residues engaging in close packing.

3.4.7 Sequence Dependence of CPT1B TM2 Homotypic Interactions

Unlike the other TM domains of CPT1A and B, CPT1B TM2 did not have any obvious G-xxx-G or other well-known motifs to predict sites of interaction. The phenylalanine residue was selected after discovering the importance of the phenylalanine in the TM1 – TM1 interactions of CPT1A. The phenylalanine in CPT1B TM2 was not found to have any great affect (**Figure 3.16**). Instead, two small residues (S_{113} and G_{120}) were found to greatly disrupt homo interactions, and one additional residue (G_{115}) was disruptive to a slightly lesser degree.

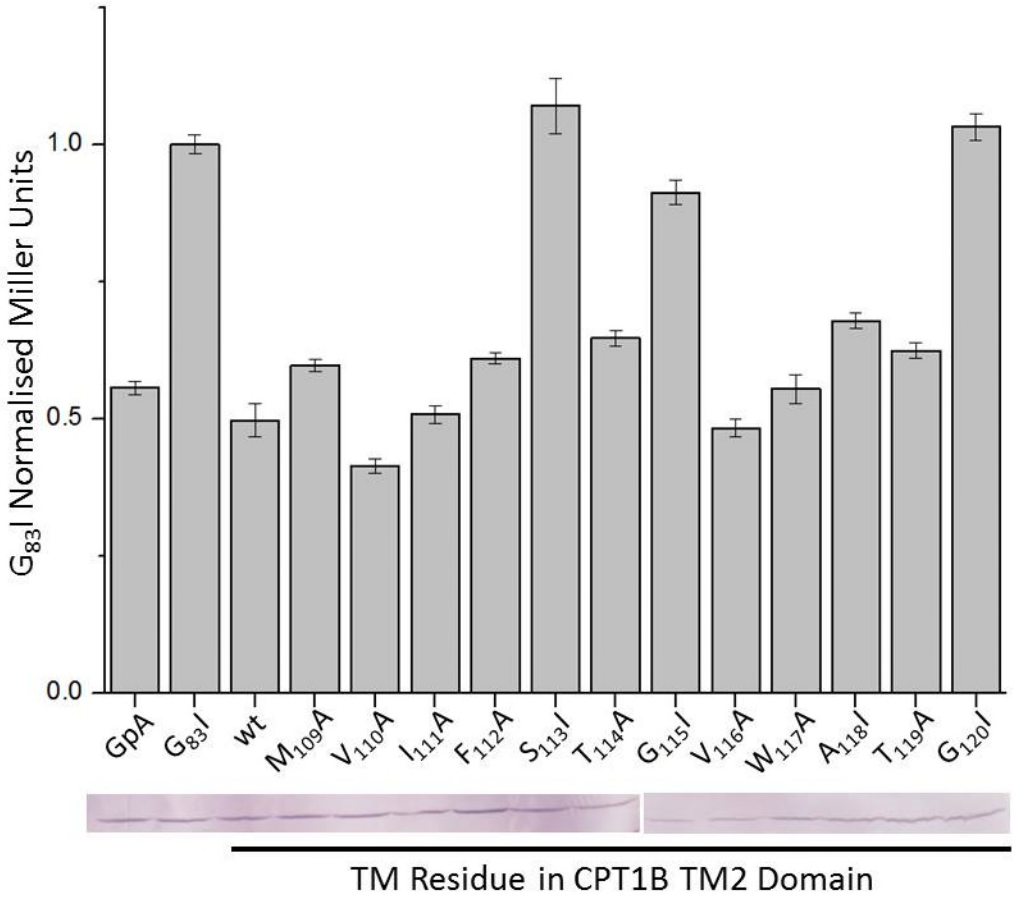


Figure 3.16 – Homotypic interactions of the mutant TM domains of CPT1B TM2 as measured in the homo GALLEX assay. All results have been adjusted for expression levels and normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. The relative expression levels measured by anti-MBP Western blot of each GALLEX fusion protein used to normalise the results are shown beneath each lane.

Twelve clusters were identified after running a CHI search for CPT1B TM2 homotypic interactions, of which four were found to be symmetrical interactions (those between the same amino acid residues on each helix). One of these clusters showed S_{113} and G_{120} in very close proximity at the interface, but this structure placed G_{115} on the opposite face, away from the interface (**Figure 3.17**). G_{115} did not show as much of a disruption in interaction as the other two residues when it was mutated in the homo GALLEX assay, so perhaps there are two possible TM2 – TM2 interacting faces, one of which has more favourable interactions over the other. Since G_{115} showed less of a disruption in the GALLEX assay and no structure with this

residue at the interaction interface was found in the CHI search, it would suggest that this is the secondary, less favourable, interaction and that the G_{113} and S_{120} face is preferred if this is indeed the case.

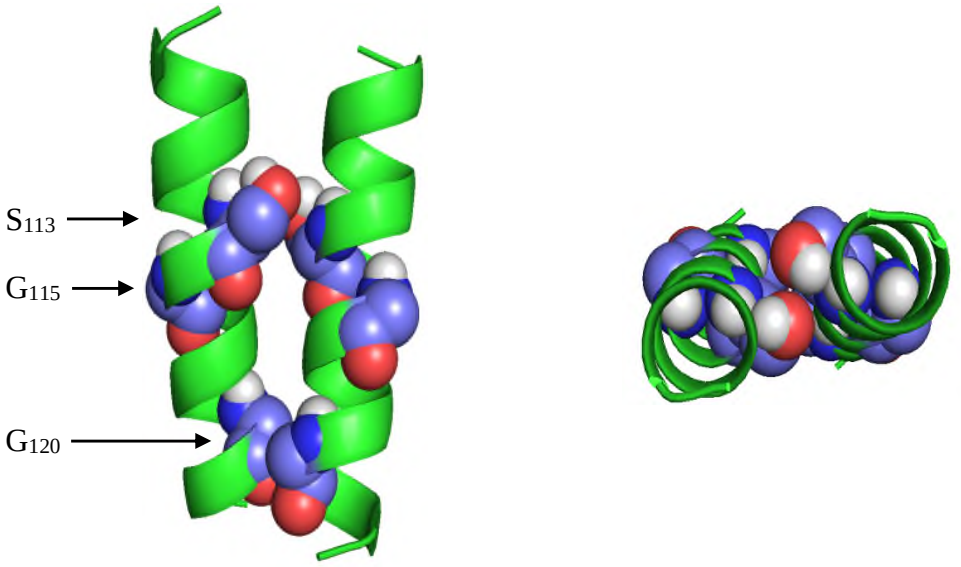


Figure 3.17 – Model of CPT1B TM2 interactions. The predicted structure found from the cluster identified in the CHI search of CPT1B TM2 homo interactions with the helix depicted as a cartoon representation (green) and the S_{113} and G_{120} residues in close interaction and G_{115} shown as a space filling representation (coloured by atom). The model is shown from the side of the two helices (left) to show the two points of contact, as well as along the length of the two helices (right) to show the close proximity of the S_{113} and G_{120} residues.

3.4.8 Discussion and Working Model of CPT1A and CPT1B Oligomerisation

The most important residue for TM1 – TM1 homotypic interactions in CPT1A appears to be phenylalanine 55 in which the aromatic ring is involved in π interactions, and not the S_{61}-xxx-A_{57}-xxx-G_{53} motif that was predicted (**Figure 3.18 A**). For CPT1A TM2 – TM2 homotypic interactions the critical residues are in the G_{105}-xxx-G_{111} motif, as well as another distinct interface formed primarily by G_{113} and V_{116} (**Figure 3.18 B**)

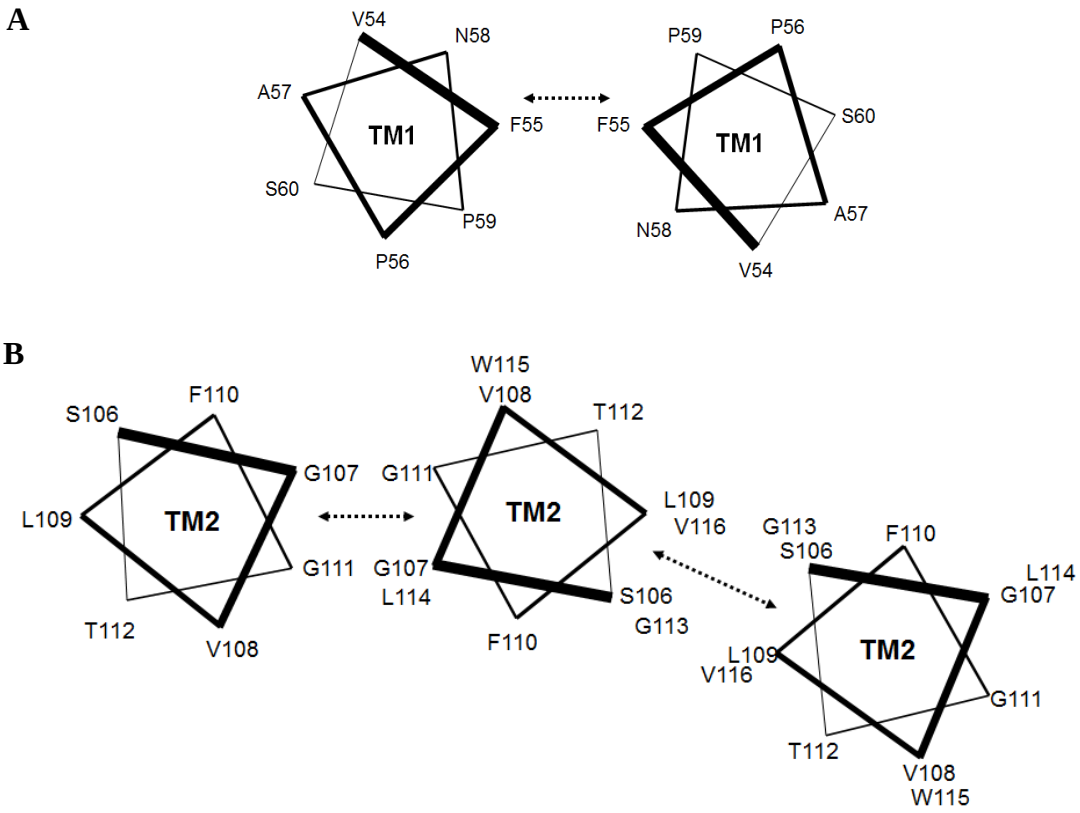


Figure 3.18 – Schematic of CPT1A homotypic TM interactions. Helical wheel representations of the primary homotypic interactions discovered in CPT1A (**A**) TM1 and (**B**) TM2.

As in CPT1A, the most important residue for TM1 – TM1 homotypic interactions in CPT1B is an aromatic residue at position 55 (**Figure 3.19 A**). In this case it is a tyrosine rather than a phenylalanine as found in CPT1A. The L_{63} residue, which is conserved in both CPT1A and CPT1B, exists on the same face of the helix as the F_{55}/Y_{55} residue, but it is only in CPT1B where the $L_{63}A$ mutation demonstrated weakening of TM1 – TM1 homotypic interactions. For TM2 – TM2 homotypic interactions the critical residues are S_{113} and G_{120}, and G_{115} to a slightly lesser extent (**Figure 3.19 B**). The S_{113} and G_{120} residues lie on the same face of the helix at either end of the TM domain with the G_{115} residue on the opposite face in the centre of the TM helix.

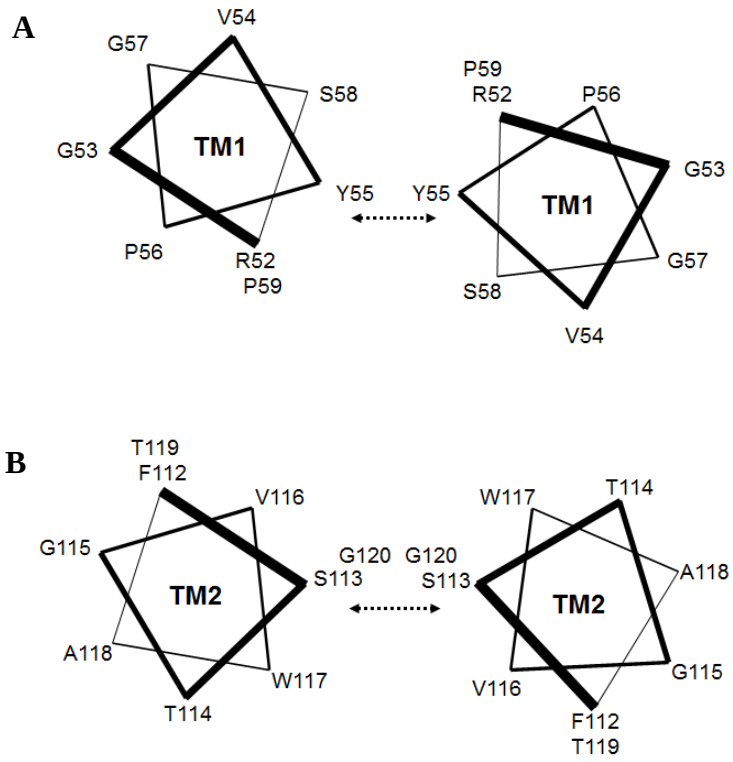


Figure 3.19 – Schematic of CPT1B homotypic TM interactions. Helical wheel representations of the primary homotypic interactions discovered in CPT1B (**A**) TM1 and (**B**) TM2.

In both CPT1A and CPT1B TM2 self-association there appear to be two distinct interfaces that can promote homotypic interactions, and in both cases these lie on separate faces of the TM domain helix. In CPT1A, the two distinct interfaces (G_{107},G_{111} and G_{113},V_{116}) showed equal, and relatively strong, disruption upon mutation in the GALLEX assay and both potential structures were observed in the CHI search. However in CPT1B it appears that one of these interfaces is favoured over the other; the G_{115} residue showed less disruption on mutation in the GALLEX assay and no structure was observed in the CHI search. It was only the G_{113}, S_{120} interface that was found here. As self-association, either between TM1 or TM2, must necessarily promote oligomerisation to some degree, these results suggest that the TM domains of both CPT1A and CPT1B are capable of self-association and thus promotion of oligomerisation. Whereas CPT1B TM2 shows one strongly interacting

motif and a secondary significantly weaker interacting motif, CPT1A TM2 has two strongly interacting motifs. These two motifs could allow for significantly more self-association in CPT1A than CPT1B. This difference in the interactions that are favoured could directly relate to the degree and propensity of oligomerisation observed in full length CPT1.

4. Heterotypic Interactions in the CPT1A and CPT1B TM Domains

4.1 Introduction and Aims

In the previous chapter it was found that all the TM domains from both CPT1A and CPT1B are capable of forming homotypic, intermolecular interactions, and it has been shown previously that all of these TM domains can form higher order oligomers (Faye, et al., 2007). However there are several potential interaction motifs that were identified in the CPT1A and CPT1B TM sequences that were not responsible for these homotypic interactions. This chapter describes experiments conducted to study whether any of these 'unused' motifs were responsible for alternative, heterotypic interactions instead. There is good prior evidence to suggest that the sensitivity to inhibition by malonyl CoA is regulated by intramolecular interactions between the N-terminal and C-terminal domains of CPT1A (Shi, et al., 2000) which are in some (currently uncharacterised) way regulated by the TM domains and the loop region (Zammit, et al., 1997). These studies demonstrated that the N-terminal domain is essential for high efficiency malonyl CoA binding and inhibition, but also that loss of sensitivity to inhibition is enhanced when the loop region is destabilised.

4.2 Hetero GALLEX Assay

As discussed previously (**Sections 2.8.3** and **3.4**), the GALLEX assay can be used to study heterotypic interactions as well as homotypic. This heterotypic version of the assay uses two chimeric proteins expressed separately from two different

plasmids. One of these chimeras is the same as that used for the homo GALLEX assay, resulting in the LexA-TM-MBP protein expressed from the pBLM100 plasmid. The other chimera is similar but contains a mutant LexA protein (denoted here and throughout as LexA*) to yield the LexA*-TM-MBP fusion protein expressed from the pALM148 plasmid. These two proteins are expressed simultaneously in the *E. coli* SU202 reporter strain which contains a mutant LexA binding region that only recognises LexA – LexA* heterodimers (**Figure 4.1**). This allows the hetero GALLEX assay to detect only heterotypic interactions. For the GpA and $G_{83}I$ controls used in hetero experiments, the same TM domain is used in both plasmids.

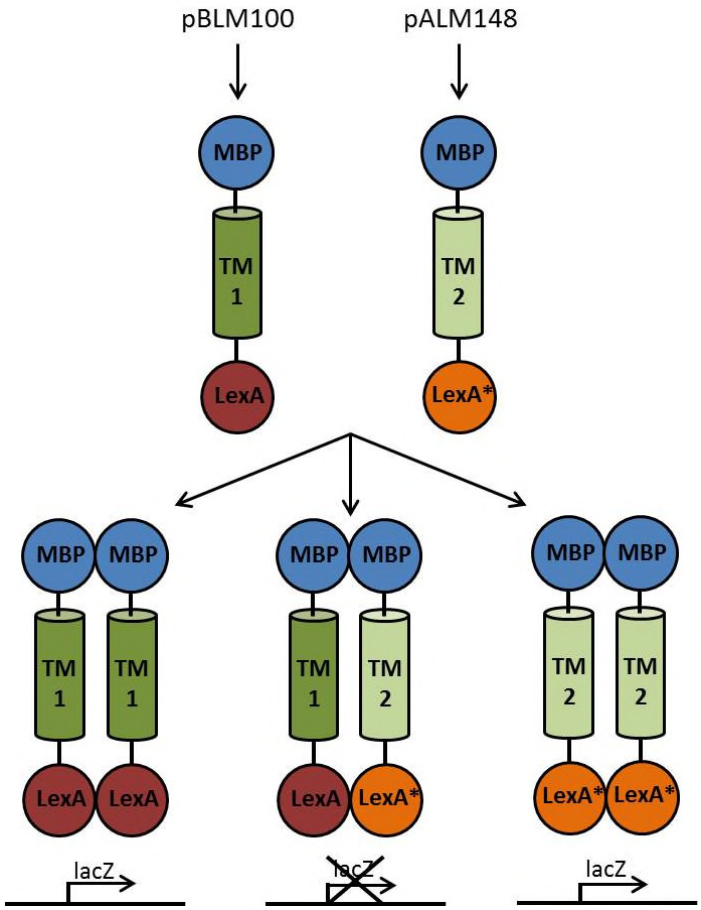


Figure 4.1 – The different possible interactions expected in a hetero GALLEX experiment. Even if TM1 – TM1 and TM2 – TM2 interactions occur, it is only the LexA – LexA* (red and orange) dimer that can bind to lacZ and inhibit expression.

The length optimised constructs (sequences shown in **Figure 3.4**) were used from the previous homo GALLEX studies and the same controls for expression and orientation were carried out. As before, the NaOH assay gave consistent results showing all the constructs were associating with the membrane, however the MalE complementation assay and the spheroplast assay proved unreliable and consequently the spheroplast dot blot assay was used to gather this information instead. Several IPTG concentrations were tested (10, 5 and 2.5 µM) just as in the homo GALLEX optimisation and again 10 µM was found to be the optimal concentration, as there was significant loss of signal at lower IPTG concentrations.

4.2.1 Wildtype TM Domain Heterotypic Interactions in CPT1A and CPT1B

The wild type TM domains from CPT1A and CPT1B were tested to investigate the extent of heterotypic interactions present. As the hetero GALLEX assay requires the use of both pALM148 and pBLM100 plasmids, TM1 was cloned into pALM148 and TM2 was cloned into pBLM100. In order to maintain the correct antiparallel topology of the two TM domains found in the full length protein, the TM1 domain constructs were designed with the TM sequence inverted. A lower case i prefix is used throughout to denote these inverted sequences. To ensure that the inversion of the helices and the subsequent inversion of the dipole moment of the backbone hydrogen bonding did not alter the interactions detected, the sequences were studied in both orientations using CHI modelling (**Figures 4.5** and **4.10**). The two TM domains were also cloned into the partner plasmid (i.e. TM1 cloned into pBLM100 and TM2 into pALM148) to rule out any differences in expression from each plasmid affecting the results. As shown in **Figure 4.2**, all of the wild type TM domains from both CPT1A and CPT1B showed evidence of relatively strong

heterotypic interactions, although not as strong as the positive control (GpA). These TM1 – TM2 interactions appeared to be of a similar strength in both CPT1 isoforms. The plasmid that the GALLEX chimera was expressed from did not appear to have a large effect on the GALLEX signal observed.

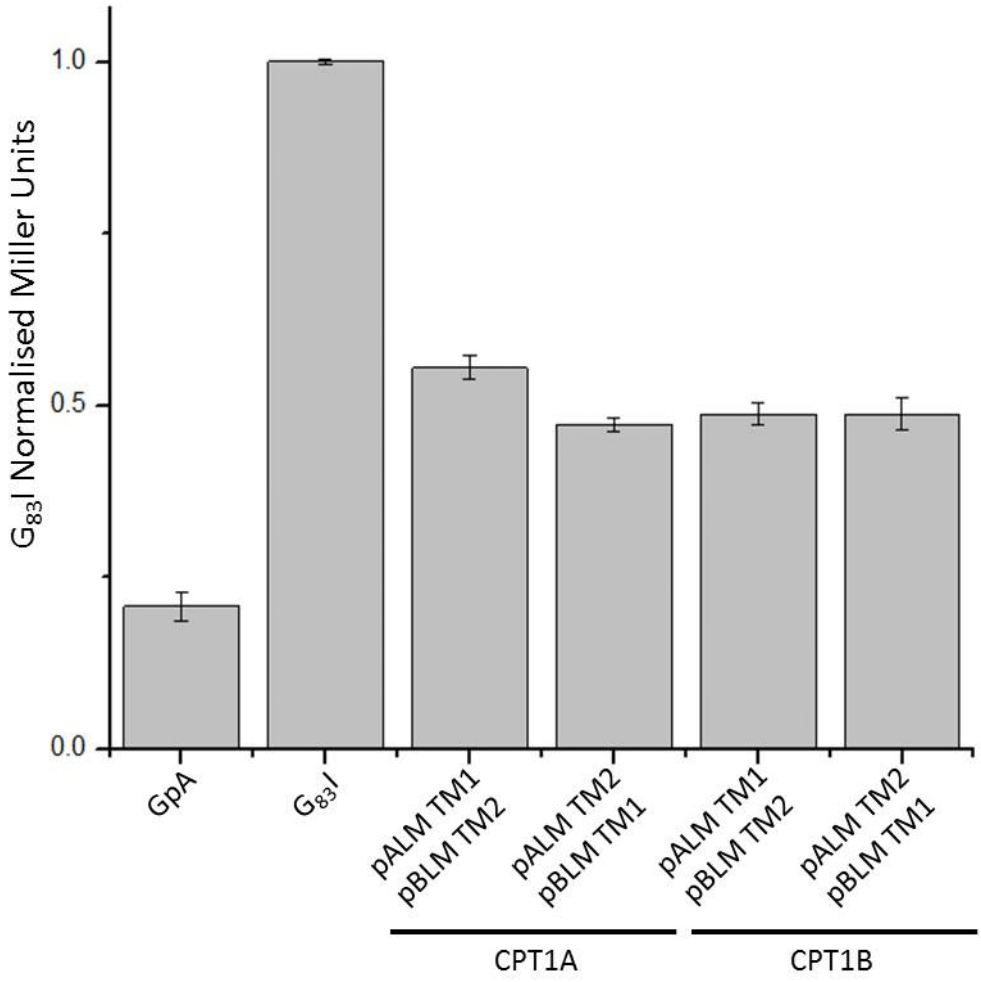


Figure 4.2 – Heterotypic interactions of the wild type TM domains of CPT1A and CPT1B as measured using the hetero GALLEX assay. All results have been normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

4.2.2 Sequence Dependence of Heterotypic Interactions in CPT1A

Wild type CPT1A and CPT1B TM domains in isolation showed significant TM1 – TM2 heterotypic interaction, but there was no difference in the strength of these interactions between the two isoforms. The effect of specific residues on these

interactions was then studied to elucidate any differences between the isoforms. Our hypothesis was that if differences exist, they must lie in the mechanism of the interactions since the overall propensities were virtually identical. To this end, alanine scanning mutagenesis was carried out. All of the mutations already generated in the pBLM100 CPT1 constructs (listed in **Tables 2.3, 2.4, 2.5,** and **2.6**) were transferred into pALM148. This was performed either by directly mutagenising the wildtype CPT1 sequences in pALM148, or by amplifying the mutant sequences from the pBLM100 construct by PCR and then ligating this sequence into pALM148.

Figure 4.3 shows the results of a hetero GALLEX experiment in which wild type CPT1A TM2 was measured with mutants of CPT1A TM1. Only one residue (F_{55}) was found to significantly disrupt TM1 – TM2 heterotypic interactions, and this was also previously identified as important for CPT1A TM1 self-association. This phenylalanine residue, and possibly the whole helical face of TM1 on which it lies, is important in stabilising both homo and hetero interactions in CPT1A. The small magnitude of the disruption to this interaction by the $F_{55}A$ mutation suggests that TM2 may have more of an impact in stabilizing these interactions.

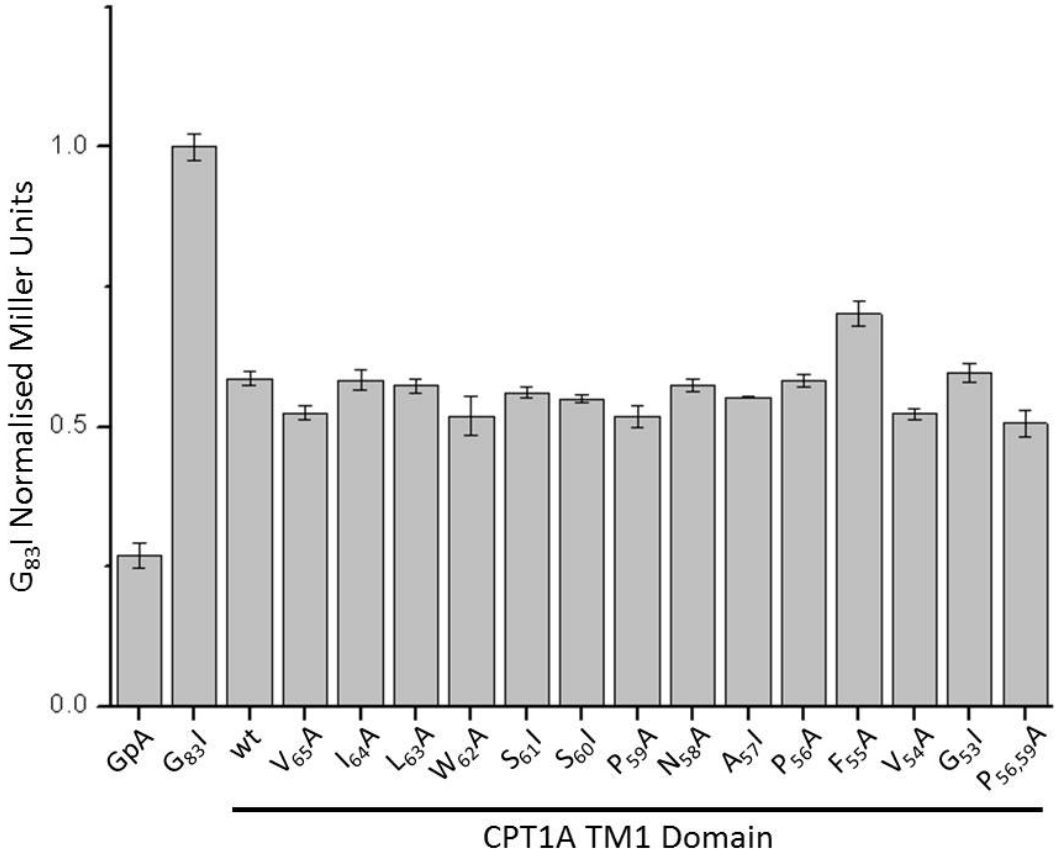


Figure 4.3 – Heterotypic interactions of the mutant TM domains of CPT1A TM1 with the wildtype TM2 as measured by the hetero GALLEX assay. All results have been normalised to the result for G_{83}I which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

As shown in **Figure 4.4**, all of the residues identified as important in the self-association of CPT1A TM2 (**Figure 3.12**) were also found to be important for CPT1A TM1-TM2 hetero interactions; however the relative strengths of these interactions are different. For TM2 self-association, the G_{107}-xxx-G_{111} motif had the strongest disruption on mutation in homo GALLEX experiments (**Section 3.4.5**), whereas this motif appears less important in heterotypic interactions and is overshadowed by the greater disruption in the G_{113}I and V_{116}A mutants (**Figure 4.4**). In addition the $G_{107,113}$I double mutant showed only a slight further disruption over just the G_{113}I single mutation alone, implying most of the disruption to the interaction is due to the G_{113}I mutation. As the G_{107}-xxx-G_{111} motif and G_{113} and V_{116}

are on different faces of the TM2 helix (shown in inset in **Figure 4.4**), it follows that the G_{113}, V_{116} face is more important for heterotypic interactions, while the G_{107}-xxx-G_{111} face is more important for TM2 self-association.

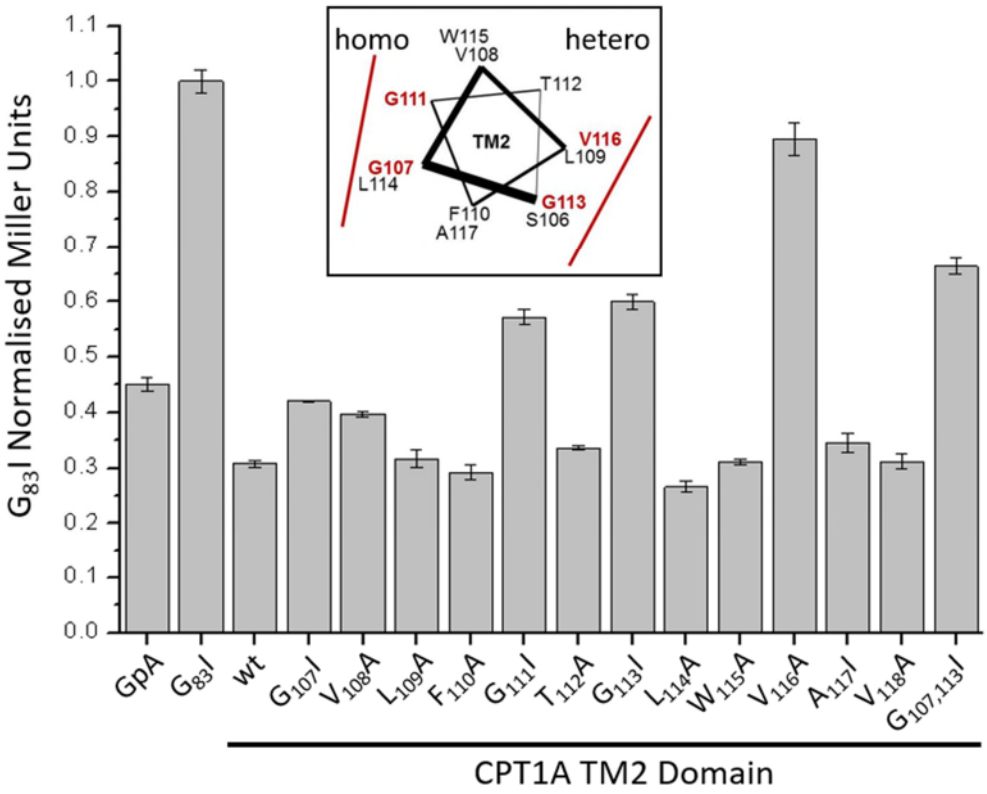


Figure 4.4 – Heterotypic interactions of the mutant TM domains of CPT1A TM2 with the wildtype TM1 as measured by the hetero GALLEX assay. All results have been normalised to the result for G_{83}I which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results. Inset shows the two faces of interaction highlighted in red on a helical wheel representation of CPT1A TM2.

CHI modelling was then used to identify possible structures that agree with the data obtained from the hetero GALLEX assay. Fourteen clusters of structures were returned from the CHI search for heterotypic interactions in CPT1A, four of which contained the F_{55} residue in TM1 located at the heterodimer interface. There was only one cluster however that contained this feature as well as the G_{113} and V_{116} residues (highlighted in the alanine scanning mutagenesis hetero GALLEX

experiment) also at the heterodimer interface (**Figure 4.5**). The structure suggests that the F_{55} and G_{113}, V_{116} residues are close enough to experience backbone hydrogen bonding between them, and the aromatic ring of the phenylalanine is projected away from the interface, unlike it's positioning in the TM1 homodimer structure **(Figure 3.11**).

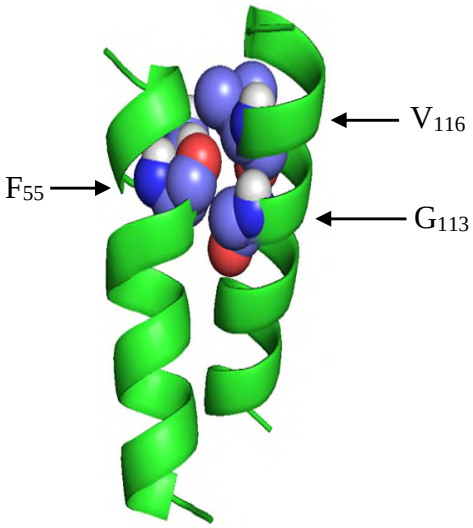


Figure 4.5 – Model of CPT1A heterotypic interactions. The predicted structure from the cluster found in the CHI search of CPT1A TM1 heterotypic interactions, with the helices depicted in cartoon representation (green) and the F_{55} residue on TM1 and the G_{113} and V_{116} residues from TM2 shown in a space-filling representation (coloured by atom).

4.2.2.1 Double Mutation Hetero GALLEX Experiments

As some of the effects observed in these hetero GALLEX experiments were quite small, and to confirm the contributions of the amino acids discovered to promote heterotypic interactions in CPT1A, a further hetero GALLEX experiment was performed in which a mutation on each TM domain was used to investigate if a further reduction in interaction was observed when both helices were mutated. This experiment will also allow the relative importance of each residue to be studied; this could be valuable here to determine, for example, whether either TM domain triggers the heterotypic interactions observed in CPT1A. These experiments were performed in the same way as the alanine scanning hetero GALLEX assays, but with each mutation tested individually and then along with a mutation on the other TM domain. For CPT1A, the $F_{55}A$ mutation on TM1 was screened against the three TM2 mutations ($G_{111}I$, $G_{113}I$ and $V_{116}A$) found from the alanine scanning hetero GALLEX experiments. All the mutations were once again found to significantly disrupt, to a similar degree, the wild type hetero interactions detected individually (**Figure 4.6**). There was a significant enhancement of disruption when the $F_{55}A$ and $V_{116}A$ mutations were combined with the effect of both of these even surpassing the negative control in this experiment. This result showed significantly greater disruption than either of these single mutations alone.

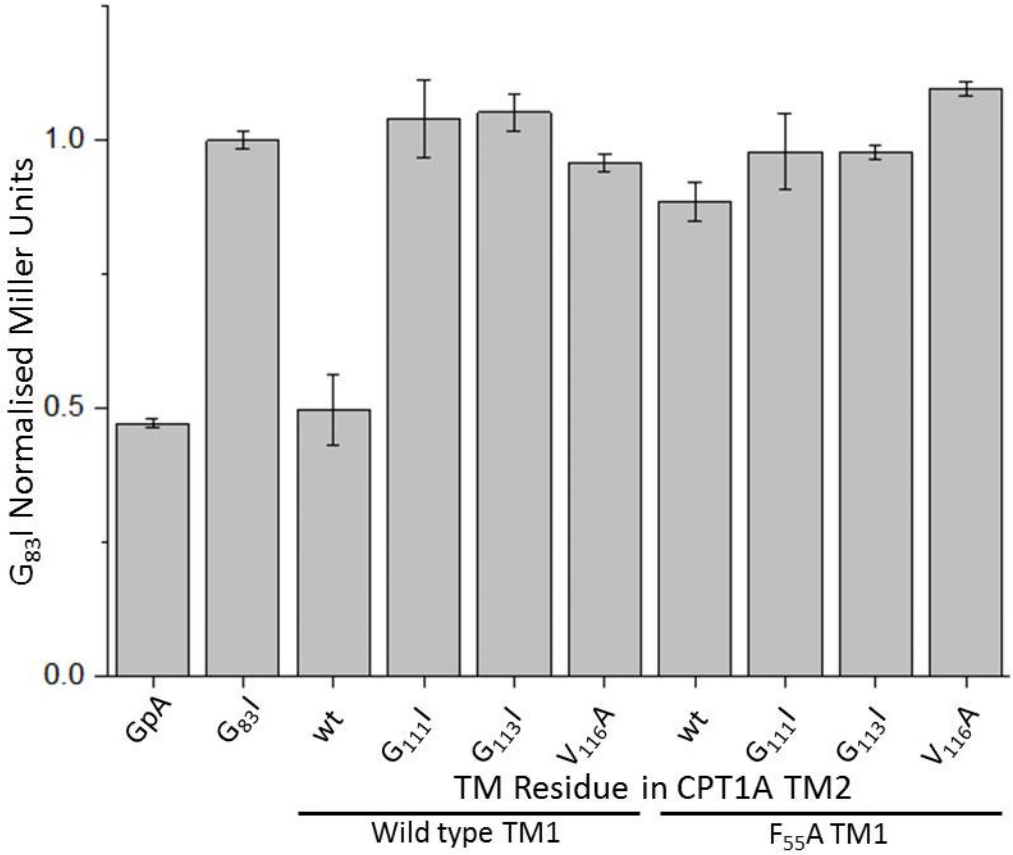


Figure 4.6 – Results from the double mutant hetero GALLEX assay for CPT1A. The $G_{111}I$, $G_{113}I$ and $V_{116}A$ mutations on CPT1A TM2 were screened against wild type and the $F_{55}A$ mutation on TM1. All results have been normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

4.2.2.2 Summary of Heterotypic Interactions in CPT1A

The TM1-TM2 heterotypic interactions in CPT1A appear to be mediated by the F_{55} residue on TM1 and primarily the G_{113} and V_{116} residues on TM2. Mutation of the G_{111} residue in the G_{107}-xxx-G_{111} motif on TM2 also showed some disruptive effect, so there may be some contribution from this residue as well. These two separate motifs do not lie on the same face of the TM2 helix so, just as in TM2 self-association, this indicates that both of these interactions could coexist. This leads to the possibility that both intermolecular and intramolecular interactions occur

simultaneously; this is naturally a requirement for oligomerisation as has been observed in CPT1A. The F_{55} residue is required for both TM1 self-association as well as TM1-TM2 heterotypic associations. This dual purpose nature could indicate a change in state for the enzyme as it switches between a monomeric state, in which intramolecular TM1-TM2 interactions are preferred, to an oligomeric state where intermolecular TM1-TM1 associations are dominant.

These interactions are summarised in **Figure 4.7** with the TM helices depicted as helical wheels showing the two turns of the helix that are relevant to the interactions discussed.

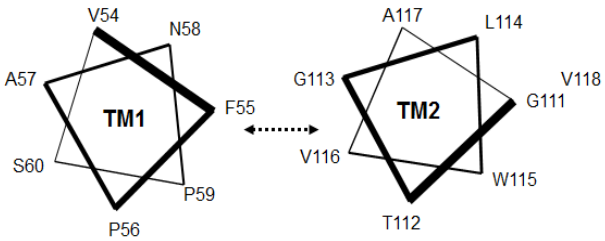


Figure 4.7 – Schematic of CPT1A heterotypic interactions. Helical wheel diagram showing the F_{55} residue in TM1, and the G_{113} and V_{116} residues in TM2 which were found experimentally from the GALLEX experiments, and computationally from the CHI searches are shown forming an interface of interaction.

4.2.3 Sequence Dependence of Heterotypic Interactions in CPT1B

The effect of specific residues was explored as in CPT1B heterotypic interactions in order to compare this to the findings already made for CPT1A. The only significant weakening of TM1 – TM2 interactions shown in **Figure 4.8** is due to mutation of G_{57} to isoleucine. This was not seen in the TM1 – TM1 homo GALLEX experiment. Also, unlike CPT1A TM1, it appears that the aromatic residue at position 55 (Y_{55} in this case) is not relevant for hetero interactions in CPT1B.

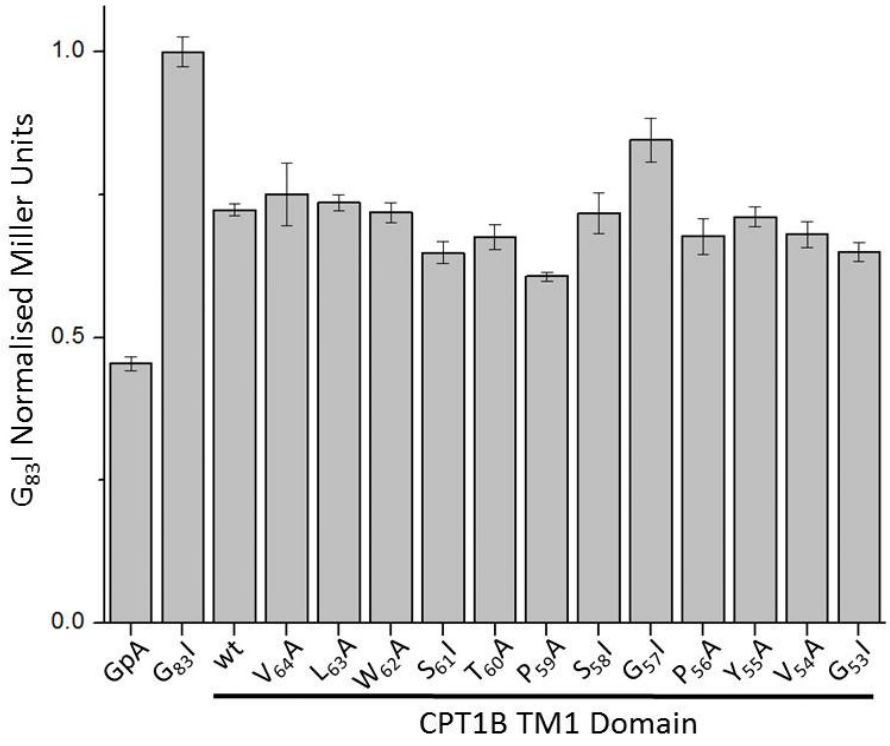


Figure 4.8 – Heterotypic interactions of the mutant TM domains of CPT1B TM1 with the wildtype TM2. All results have been normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

The single largest disruption of interactions from wildtype in **Figure 4.9** is residue G_{115}. While this residue was identified in TM2 – TM2 homo interactions in CPT1B it seemed to play a more minor role than the other two residues. Here, in

TM1 – TM2 interactions, that trend is reversed with the G_{115} residue playing a larger role than the S_{113} and G_{120} residues, which also show some disruption. The S_{113} and G_{120} are important for TM2 – TM2 homo interactions, whereas the G_{115} residue is more important for hetero interactions. This correlates well in terms of location in the membrane with the single residue (G_{57}) found in **Figure 4.8**. These two glycine residues could perhaps pack tightly together.

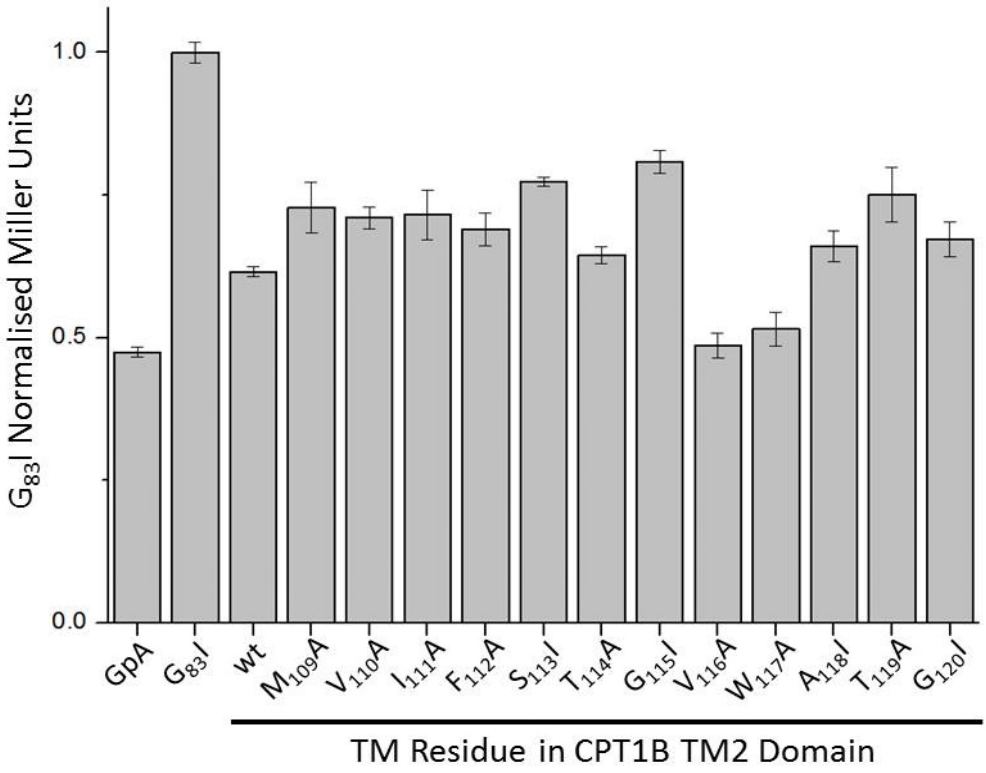


Figure 4.9 – **Heterotypic interactions of the mutant TM domains of CPT1B TM2 with the wildtype TM1** as measured by the hetero GALLEX assay. All results have been normalised to the result for G83I which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

Six clusters were identified after running a CHI search for CPT1B hetero interactions, two of which had the G_{57} and G_{115} residues near to the interface. However neither of them showed these residues to be particularly close in space, so it would appear that it is not close packing that is the stabilising interaction here (**Figure 4.10**).

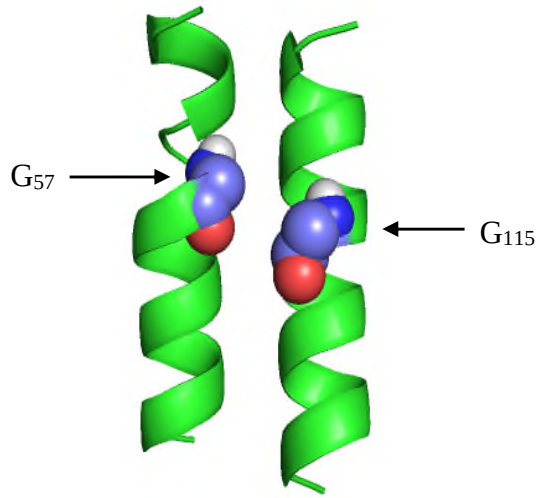


Figure 4.10 – Model of CPT1B heterotypic interactions. The predicted structure from the cluster found in the CPT1B TM1 heterotypic interactions CHI search, with the helices depicted as a cartoon (green) and the G_{57} residue on TM1, and the G_{115} residue on TM2 shown with an electron cloud for each atom (coloured by atom).

4.2.3.1 Double Mutation Hetero GALLEX CPT1B Experiments

Similarly to CPT1A, a further hetero GALLEX experiment was undertaken to check the previous results obtained from the alanine scanning mutation experiments. The G_{57}I mutation from TM1 and the G_{115}I from TM2 found to be relevant in heterotypic interactions in CPT1B were tested again individually and together. In this case there did not appear to be any synergistic enhancement of heterotypic interactions when both mutations were introduced, instead suggesting that it is the G_{115} residue on TM2 that is almost solely playing a role in these interactions (**Figure 4.11**). This agrees with the CHI models generated which failed to show any obvious interactions between these two residues. Following this result the CHI models were inspected again for any likely models showing G_{115} at the interface of interaction.

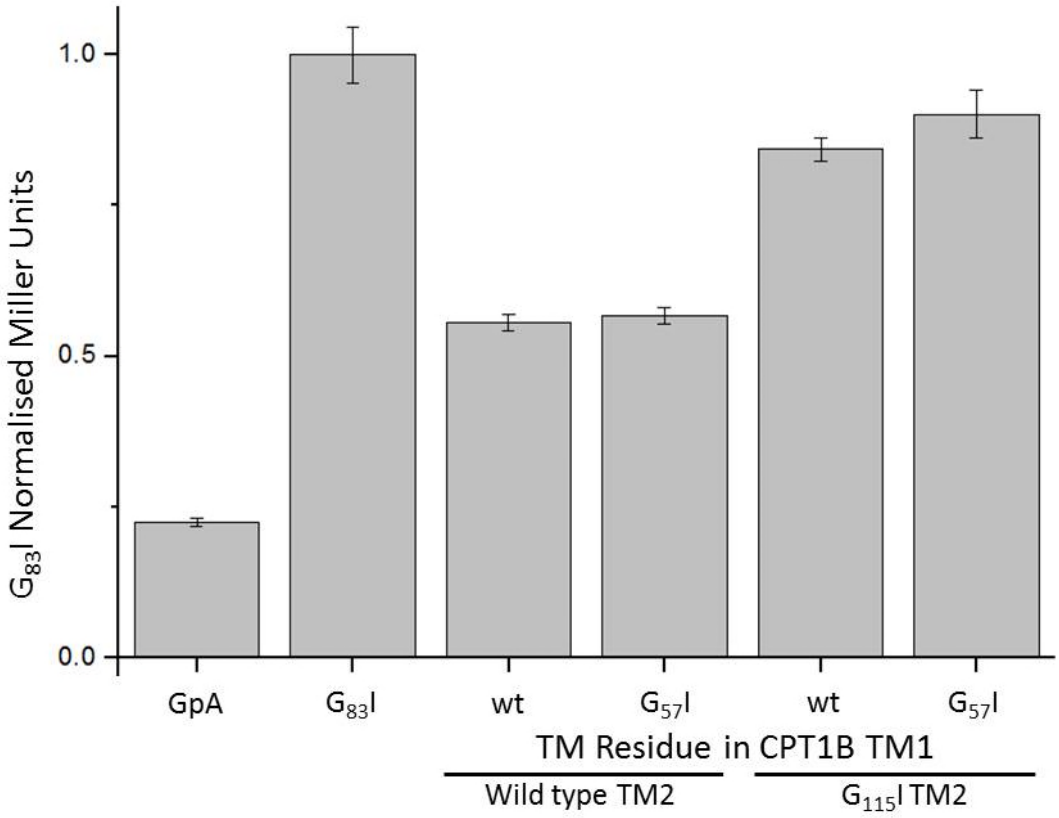


Figure 4.11 – Results from the double mutant hetero GALLEX assay for CPT1B. The G_{57}I mutation on TM1 was screened against the G_{115}I mutation on TM2 to check for their respective involvement in forming an heterotypic interaction. All results have been normalised to the result for G_{83}I which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

4.2.3.2 Summary of Heterotypic Interactions in CPT1B

Unlike CPT1A, only a single residue on each TM domain of CPT1B was observed to affect heterotypic interactions on mutagenesis: the G_{57} residue on TM1 and the G_{115} residue on TM2. As these are both small glycine residues, this could imply that it is a steric effect, which allows either the glycine residues themselves or residues near them to pack together more efficiently than when a larger residue is present, thus promoting interaction. In the double mutation experiment however, it was found that G_{115} on TM2 had a much greater effect on this interaction than G_{57} on TM1. These two residues would lie at the same depth in the membrane so interactions would be possible; however a convincing structure from CHI searches was not found.

This interaction is shown in **Figure 4.12,** with the TM helices depicted as helical wheels showing the two turns of the helix that are relevant to the interaction. The dashed arrow represents the interaction found from the hetero GALLEX experiments. Unlike CPT1A, this heterotypic interaction found in CPT1B could exist concurrently with the homotypic interactions found in the previous chapter (**Section 3.4.6** and **3.4.7**) as the Y_{55} residue on TM1 shows no role in heterotypic interactions. The only exception would be if G_{115} on TM2 was involved in both homotypic and heterotypic interactions; it was the only residue to be identified in both assays.

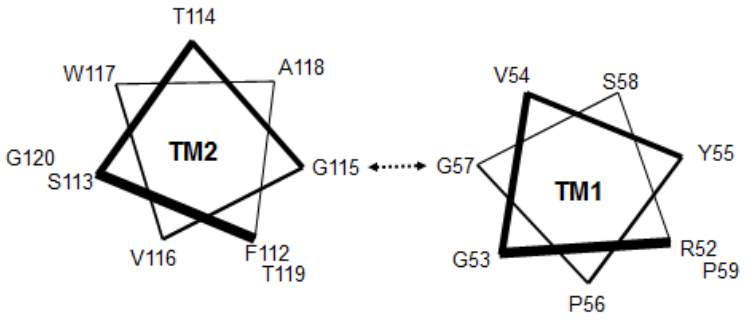


Figure 4.12 – Schematic of heterotypic interactions in CPT1B. Helical wheel diagram showing the G_{57} residue in TM1 and the G_{115} residue in TM2 which were found experimentally from the GALLEX experiments are shown.

4.2.4 Differences in Heterotypic Interactions of CPT1A and CPT1B

The two most important physiological differences are that CPT1A is 30-100 fold less sensitive to inhibition by malonyl-CoA than CPT1B and that CPT1A shows the ability to modulate its sensitivity to inhibition. Given that it is known that the composition, curvature and thickness of the membrane can affect this modulation in sensitivity it is logical to assume that the transmembrane domains, the part of the protein most closely in contact with the membrane, are at least partially responsible for this.

The differences in interactions identified between the two isoforms of CPT1 studied suggest that CPT1A has several distinct forms in which either homotypic interactions (TM1 – TM1 and/or TM2 – TM2) or TM1 – TM2 hetero interactions dominate. Perhaps a switch between these two possible states is responsible for changes in the full length protein which lead to a change in the inhibition of the protein. The interactions identified in CPT1B, on the other hand, can all exist concurrently with separate faces of the helices mediating homotypic and heterotypic interactions. If a balance between homotypic and heterotypic TM domain

interactions is involved in modulating the sensitivity of CPT1 to inhibition, then this could explain why CPT1B does not show the same variation in sensitivity to inhibition as CPT1A, despite their conserved structures and sequences.

4.2.5 Relating the relative strengths of homo and hetero GALLEX results - Competition GALLEX

In all of the GALLEX experiments described here, the data obtained is always qualitative within each type of assay, either homo or hetero. This is useful to compare homotypic interactions with other homotypic interactions, and heterotypic with other heterotypic interactions, but presents difficulty when comparing homotypic with heterotypic interactions. Partly this is due to the extra normalisation that the homo assay allows, using protein expression levels rather than just normalising to cell density, however there are also other areas of uncertainty in comparing the two assays. It was unknown as to whether the expression levels from the two plasmids were the same at a specific IPTG concentration, and whether the mutated DNA binding site in the hetero reporting strain SU202 was completely specific to the mutated LexA*. The mutated DNA binding site in SU202 and the corresponding mutations in the LexA* domain are relatively minor and result in only a few amino acid changes from wildtype LexA so there was potential here for a loss in specificity (Dmitrova, et al., 1998).

To address this, control experiments were developed here to assess the plasmid expression levels and the specificity of the homo and hetero assays. These controls are discussed in the following section.

4.2.5.1 Expression Levels of pALM148 and pBLM100

In order to gain some understanding of the relative populations of each of the chimeric GALLEX proteins in a hetero assay, the relative expression levels from each of the plasmids (pALM148 and pBLM100) containing the same TM sequences were tested. If one of the proteins was being significantly overexpressed from one plasmid compared to the other, this could bias the results to indicate that an interaction was stronger than it actually was simply because there is a larger population of proteins capable of interaction. The results in **Figure 4.13** show that not only are the expression levels of the pALM148 and pBLM100 plasmids very similar for the same TM sequence, but that even across multiple TM sequences the expression levels did not vary to a great extent. This was important in validating the standard hetero GALLEX assay as well as providing confidence in the possibility of extending GALLEX methodologies to include competition measurements. Initially there were significant discrepancies in the expression levels between pALM148 and pBLM100 but, as in the standard homo GALLEX assay, the optimised method using ImageJ to quantify bands from Western blots improved the data dramatically (**Section 3.4.2**). There were some difficulties in growing the pALM148 $G_{83}I$ samples to the same OD_{600} as the other samples in this assay. Despite normalisation to cell density, this poor growth is likely to account for the lower expression levels observed in this construct as this was significantly lower than any other tested. The $G_{83}I$ TM domain did not suffer difficulties in growth in the pBLM100 plasmid so it does not appear to be anything to do with this specific TM sequence.

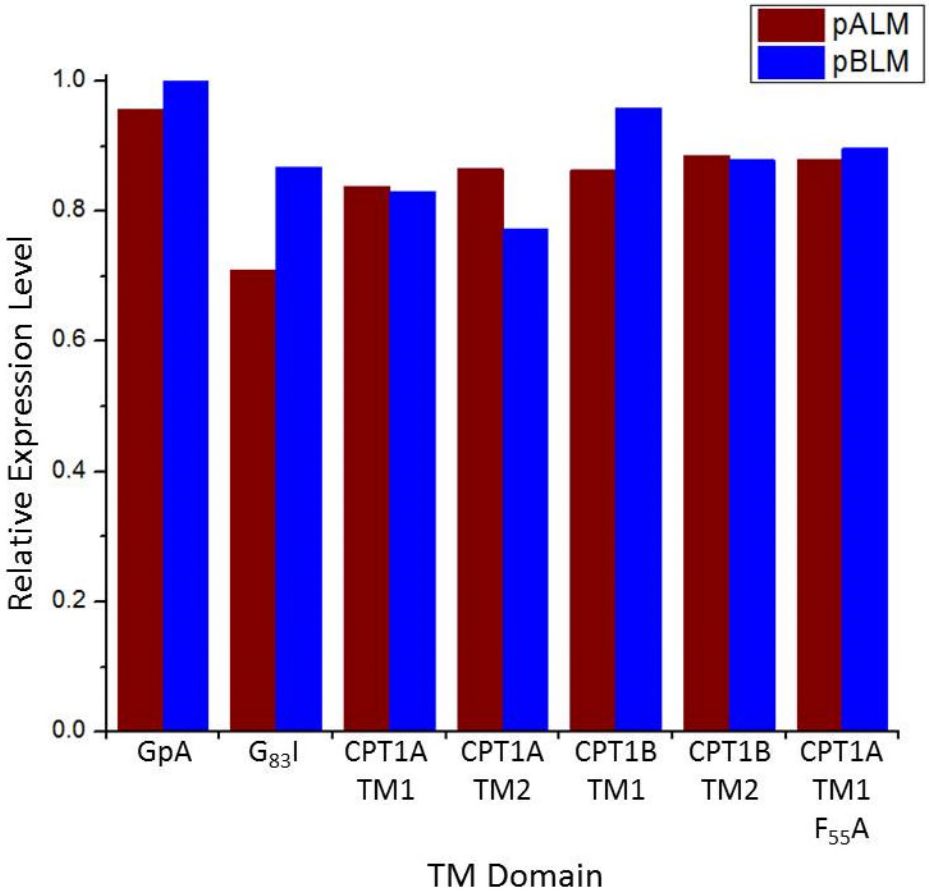


Figure 4.13 – The relative expression levels of the pALM148 and pBLM100 plasmids with the controls (GpA and $G_{83}I$) and the wildtype TM domains of CPT1A and CPT1B. The largest expression level (GpA in pBLM100) was set to a value of 1, and other levels are shown relative to this. Significantly poorer growth to OD_{600} was observed in preparing the $G_{83}I$ samples than any of the others; this is reflected in the lowest expression levels observed.

4.2.5.2 'Homo in Hetero' Control for Hetero GALLEX Assay

For the 'homo in hetero' control experiment, as the name suggests, the hetero reporting stain SU202 was transformed with only pBLM100 as used in the homo assay. This should mean that only wildtype LexA chimeric proteins are present and if interactions occur this can only result in a LexA-LexA dimer which should not be capable of binding to the mutant DNA binding domain in SU202. If this is the case then all the samples should report as if no interactions are occurring and the signal observed should be of a similar level as the negative control $G_{83}I$. The results from this control assay rule out any 'crosstalk' from the homo assay appearing in the hetero assay (**Figure 4.14**). This was an initial concern as several of the same residues in both CPT1A and CPT1B were identified in both the homo and hetero GALLEX assay. As in the standard homo GALLEX assay and the expression level tests in this chapter (**Figure 4.13**) the optimised method using ImageJ to quantify bands from Western blots improved this data dramatically (**Section 3.4.2**).

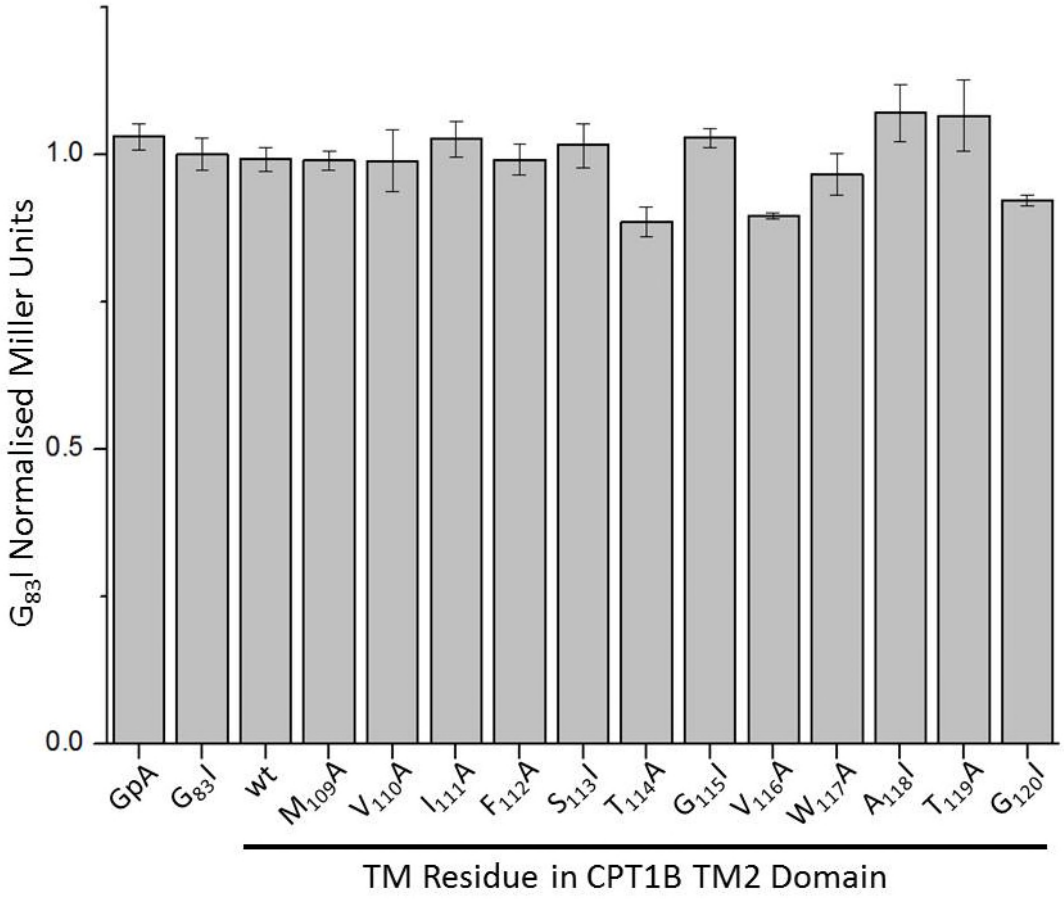


Figure 4.14 – 'Homo in hetero' control for hetero GALLEX. Results showing the attempted measurement of homotypic interactions measured in the hetero reporting *E.coli* GALLEX strain (SU202). The alanine scan mutant TM domains of CPT1B TM2 were used to demonstrate the lack of signal observed. All results have been adjusted for expression levels and normalised to the result for $G_{83}I$ which has been set at 1. Error bars represent +/- the standard error of the mean of 3 separate results.

Now that the specificity of homo and hetero detection in the GALLEX assays had been confirmed, a method for comparing homo and heterotypic interactions was developed. An issue with systems such as the CPT1 TM domains, where each component is capable of both homotypic and heterotypic interactions, is that if homotypic interactions occur to a much greater extent than heterotypic, the signal observed during the hetero assay could be obscured because all the protein is involved in homotypic interactions and there is none free to participate in heterotypic interactions. For the CPT1 TM domains, because both components are indeed

capable of homotypic interactions, it would be possible for both TM domains to exclusively self-associate and yield no observable signal in the hetero assay. A method to compare the relative strengths of the homotypic and heterotypic interactions was needed to investigate whether there was a reduction in heterotypic association signal due to homotypic interactions.

In homo GALLEX the only observable population is LexA-LexA dimers, and in hetero GALLEX, the observable population is LexA-LexA* dimers. It was hoped that by measuring both of these, the potential population of LexA*-LexA*, which cannot be directly measured in either assay, could be inferred. This would confer information about the level of homotypic interactions within the hetero assay. To achieve this, the hetero assay was performed in the homo reporting strain (SU101). This allowed measurement of the LexA-LexA population while LexA* protein was present. This result, when compared to the same sample in the absence of LexA* protein, could be used to infer the relative strengths of homo and hetero interactions in the system. As both the TM domains were present, it was assumed that whatever interactions were preferred natively would likewise be preferred here. The competition assay with both TM domains was run in parallel with a standard homo GALLEX experiment on each TM domain individually. The results from each assay for each TM domain were then compared. It was expected that if heterotypic interactions were similarly or more favourable than homotypic interactions, then a disruption to the homotypic interactions measured in the homo GALLEX assay would be observed.

This assay was performed on the four CPT1A and CPT1B TM domains with and without their heterotypic interacting TM domain present e.g. CPT1A TM1 with

and without CPT1A TM2. It was found that further optimisation was necessary here over and above what had already been carried out for the standard homo and hetero GALLEX assays. The concentration of IPTG used for inducing protein expression was the main parameter that required adjustment. Using the same concentration as in the standard GALLEX assays (10 µM) there was no observable difference between the two results. As the disrupting effect could be quite small if the heterotypic interactions were not significantly stronger than the homotypic interactions, it was reasoned that perhaps these subtle differences were being obscured with high protein concentrations. To investigate this, several lower IPTG concentrations were screened. An IPTG concentration of 5 µM was found to provide almost as good definition between positive and negative controls as 10 µM but could allow for more subtle effects to be discerned. 5 µM was used for all the competition GALLEX experiments conducted. At concentrations below this, the range between positive and negative results was severely degraded as can be seen in **Figure 4.15**.

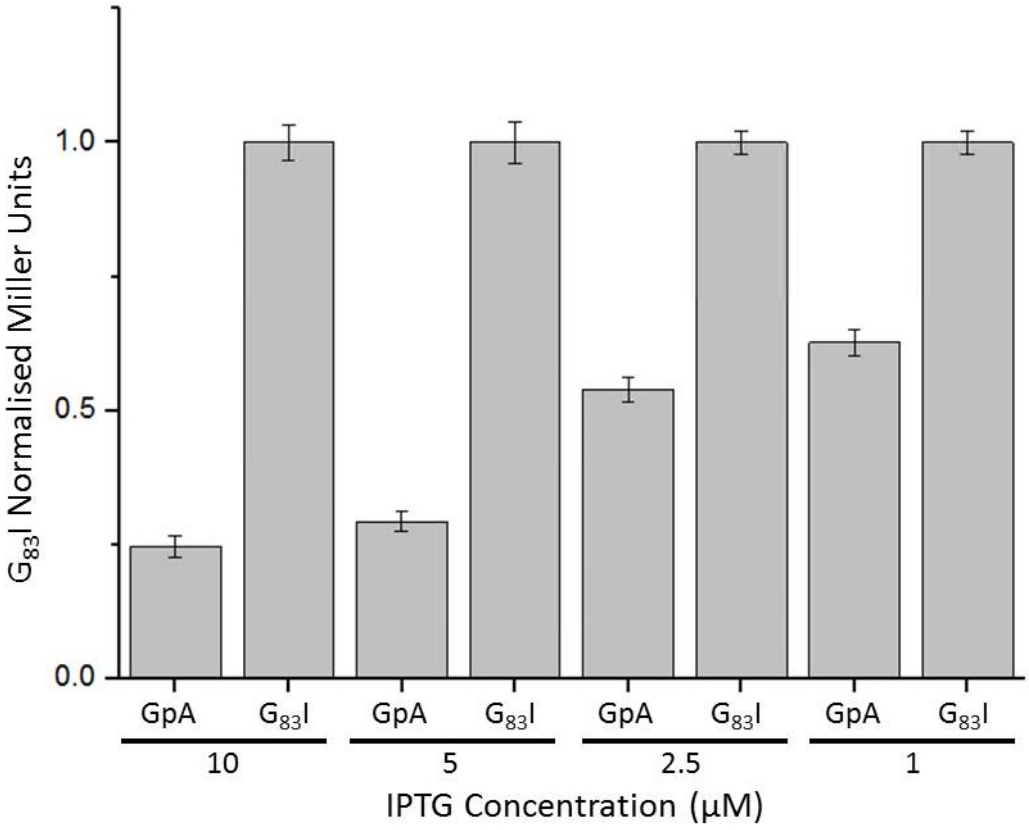


Figure 4.15 – The effect of IPTG concentration in homo GALLEX. As IPTG concentration is increased, less protein is produced and the signal of positive interactions decreases significantly. This degrades the range between strong and weak interactions.

After the expression levels had been optimised and the competition GALLEX assay performed on the CPT1 TM domains using this IPTG concentration, differences between the CPT1 TM domains became clearer (**Figure 4.16**). Both TM1 and TM2 from CPT1A showed similar levels of disruption to their self-association on the addition of the other 'competing' TM domain. In CPT1B it appears that TM1 self-association is not as strongly affected by the presence of TM2 as either of the CPT1A TM domains. In contrast, TM2 self-association was the most affected by the addition of TM1 out of all the CPT1 TM domains.

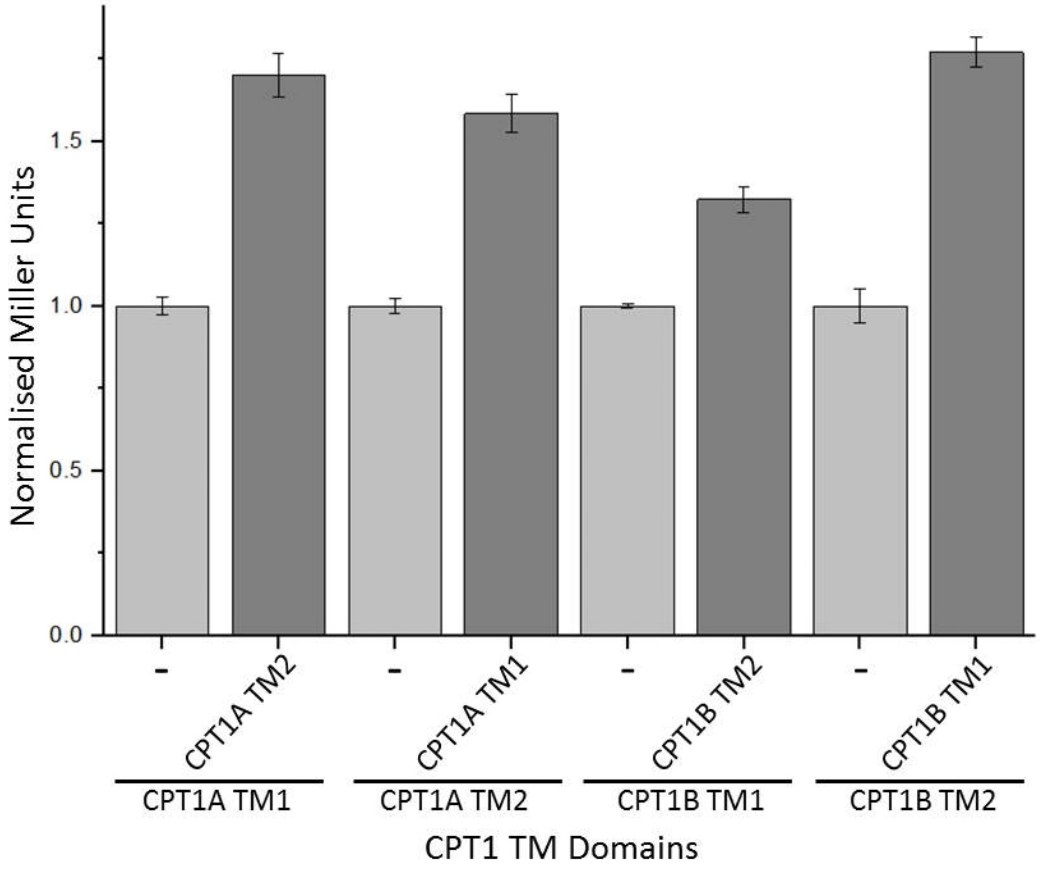


Figure 4.16 – Competition GALLEX assay for the wildtype CPT1A and CPT1B TM domains. These results are shown as relative differences where the wild type homo GALLEX results (shown in light grey) are set at a value of 1 and each competition result (shown in dark grey) is scaled relative to the appropriate homo GALLEX result. The difference in signal observed between these two results indicates the level that the homotypic interaction is affected when heterotypic interactions are also present and available.

The differences in these competition results indicate that self-association of both TM domains of CPT1A are equally disrupted by the presence of heterotypic interactions, whereas CPT1B TM2 self-association is significantly more disrupted than that of TM1 when heterotypic interactions are available. This could indicate that CPT1B favours heterotypic intramolecular interaction over intermolecular TM2 self-association. This could suggest that when CPT1B oligomerises, it is TM1 that drives this process and not TM2. It is thought that CPT1A can form higher order oligomers than CPT1B; these results suggest that this could be because both TM domains in

CPT1A self-associate favourably versus only TM1 in CPT1B. This, coupled with the two faces of interaction found for CPT1A TM2, suggest a greater range of oligomerising interactions are possible, certainly within the TM domains. Higher order oligomers can be propagated through only CPT1A TM2 whereas for complexes higher than a dimer in CPT1B this would require intermolecular TM1-TM2 interactions as well as TM2 self-association. These levels of oligomerisation could be very important in the modulation of sensitivity to inhibition by malonyl CoA as it is believed that interactions between the N- and C-termini form the malonyl CoA binding site. These interactions could occur within a monomer through intramolecular TM1 - TM2 interactions, however it is also possible that the malonyl CoA binding site could be formed by intermolecular TM1 - TM2 interactions in a dimer or higher order complex. Perhaps the level of oligomerisation is related to the sensitivity to inhibition and CPT1A shows this modulation due to the greater amount of available oligomeric configurations made possible by the greater amount of TM domain interaction configurations.

5. Expression and Purification of CPT1A and CPT1B TM Domain Peptides

5.1 Introduction

In order to study CPT1A and CPT1B using biophysical methods, a reliable method of protein production must be achieved first. Heterologous protein expression in *E. coli* is one of the simplest and most cost efficient methods to produce proteins in large enough yields for experimentation. This is especially true when isotopically labelled protein is required for heteronuclear NMR experiments. Due to the cost and concentrations required it is not reasonable to prepare protein or peptides synthetically for these types of experiment, and so an heterologous protein expression system that is capable of producing hydrophobic peptides at a reasonable yield was required.

The following chapter describes the development and optimisation of a protocol for the expression of CPT1 TM domain peptides using the vector pMMHb. This protocol was adapted and optimised using the protocol described by Claridge and Schnell (Claridge and Schnell, 2012) and the vector was also kindly provided by Jason Schnell, Oxford University, UK. This vector encodes for a fusion protein with an N-terminal His_9 tag followed by the TrpΔLE1413 fusion tag (trpLE) (Staley and Kim, 1994). A cyanogen bromide (CNBr) cleavage site between these two tags and the desired protein sequence allows for cleavage of both the His-tag and the trpLE fusion tag in one step after expression and purification is complete (schematic of the fusion protein in **Figure 5.1**).

The trpLE tag is comprised of the leader sequence of the trp operon of *E. coli* which is fused to a 97 residue C terminal sequence of the anthranilate synthase gene (Bertrand, et al., 1976, Miozzari and Yanofsky, 1978). This sequence allows for high expression levels of trpLE fusion proteins in inclusion bodies. Expressing hydrophobic proteins as inclusion bodies is advantageous as this can increase the yield over strategies where the protein is inserted into the cell membrane. Expressing proteins as inclusion bodies can also significantly lower their toxicity to the cell, as well as protecting them from cytosolic proteases in *E. coli* (Kleid, et al., 1981). Several small hydrophobic proteins containing either one or two TM domains have been expressed using a trpLE fusion protein including: the viral protein Vpu from HIV (Ma, et al., 2002); individual TM domains from GPCRs (Zheng, et al., 2005); caveolins (Diefenderfer, et al., 2009); the zetazeta TM dimer (Call, et al., 2006); TM sections of two viral glycoproteins, gp45 (Chong, et al., 1991), and avian leukosis (Smith, et al., 2004).

Figure 5.1 – Schematic of the protein to be expressed. This shows the N terminal His_9 tag (blue), followed by the trpLE sequence (red), the single methionine residue to enable CNBr cleavage (green), and the CPT1 TM sequence (orange).

A good protein expression yield is important for all biophysical experiments but is especially important for isotopically labelled NMR experiments due to the expense of the isotopic labels and the inherent insensitivity of NMR experiments. Although initial expression yields in LB media may be promising, there is often a significant reduction in yield when expressed in a minimal M9 medium which is required for isotopic labelling. NMR experiments require a concentration in the

range of 0.1-1 mM for sufficient signal to noise which in this case meant up to 1 mg of protein per 180 µl NMR sample. It is obviously desirable to achieve significantly higher expression levels per litre than this to minimise total culture volumes and hence minimise the amounts of isotopically labelled metabolites required.

Before this work to express the CPT1A and CPT1B TM domains as peptides in *E. coli*, the expression of full length CPT1 was attempted in the yeast *P. pastoris*. Previous attempts to express full length CPT1 in *E. coli* had resulted in very high levels of cell death on induction of protein expression, and it was hoped that by expressing CPT1 in a eukaryotic system it would allow for the native insertion of CPT1 into the outer mitochondrial membrane. Two different expression plasmids were tested: pGAPZ which results in constitutive expression; and pPICZ in which expression is induced in the presence of methanol. Low level constitutive expression was observed using the pGAPZ system, however this could not be achieved reliably, and the pPICZ system never resulted in higher expression levels than pGAPZ and often resulted in no detectable expression at all by anti-CPT1 Western blots. Unfortunately, despite many attempts to optimise growth and induction conditions, this could not be improved upon. Following these disappointing results, efforts were redirected to the expression of the CPT1 TM domain peptides in *E. coli* instead.

This chapter describes experiments to optimise expression by adjusting the following conditions: temperature, IPTG concentration, induction time, and cell harvest time. Optimisations made to the purification methods used post expression are also discussed.

5.2 Construction of the Expression Vectors

To make the expression constructs in the pMMHb plasmid, custom primers encoding for the four CPT1A and CPT1B TM domains were ligated into the plasmid between the HindIII and BamHI sites (see **Section 2.6** for details). In the design of the custom primers any cysteines were mutated to alanines to prevent disulphide bond formation, and methionines were mutated to leucines (final sequences shown in **Figure 5.2**). These mutations were to prevent further cleavage when CNBr was added to remove the trpLE and His-tags during purification. A single serine residue was omitted from the beginning of the CPT1A TM2 sequence as residues with hydroxyl groups, and particularly serine, which immediately follow the methionine residue can react with an intermediate formed during the cleavage reaction and reform the peptide bond (Kaiser and Metzka, 1999).

CPT1A TM domains	
TM1	FKNGIITGVFPANPSSWLIVVVGVISS**L**HAK
TM2	QTKNIVSGVLFGTGLWVAVI**L**T**L**RYSL
CPT1B TM domains	
TM1	KNGILRGVYPGSPTSWLVVVLATVGSNY**A**KV
TM2	PQTETLLS**L**VIFSTGVWATGIFLFRQTL

Figure 5.2 – Sequences of the CPT1 TM domain peptides. Primers encoding for these sequences were ligated into the pMMHb plasmid for expression. Residues shown in bold are those that had to be mutated: methionine was mutated to leucine, and cysteine to alanine.

5.3 Cell Growth and Expression

Initial expression trials in LB medium showed good expression levels (**Figure 5.3**), however on switching to M9 minimal medium a significant reduction was observed. This reduction was exacerbated when ^{12}C glucose and ^{14}N ammonium chloride were substituted for ^{13}C and ^{15}N isotopically labelled metabolites (Cambridge Isotope Laboratories Inc., USA and Sigma Aldrich, UK). Due to the size of the final peptides to be expressed (~3 kDa) it was decided that triple resonance NMR experiments would not be required and hence all future expression in labelled M9 media was performed using only ^{15}N ammonium chloride and standard ^{12}C glucose. This provided an immediate increase in expression yields in M9 minimal medium (**Figure 5.4**).

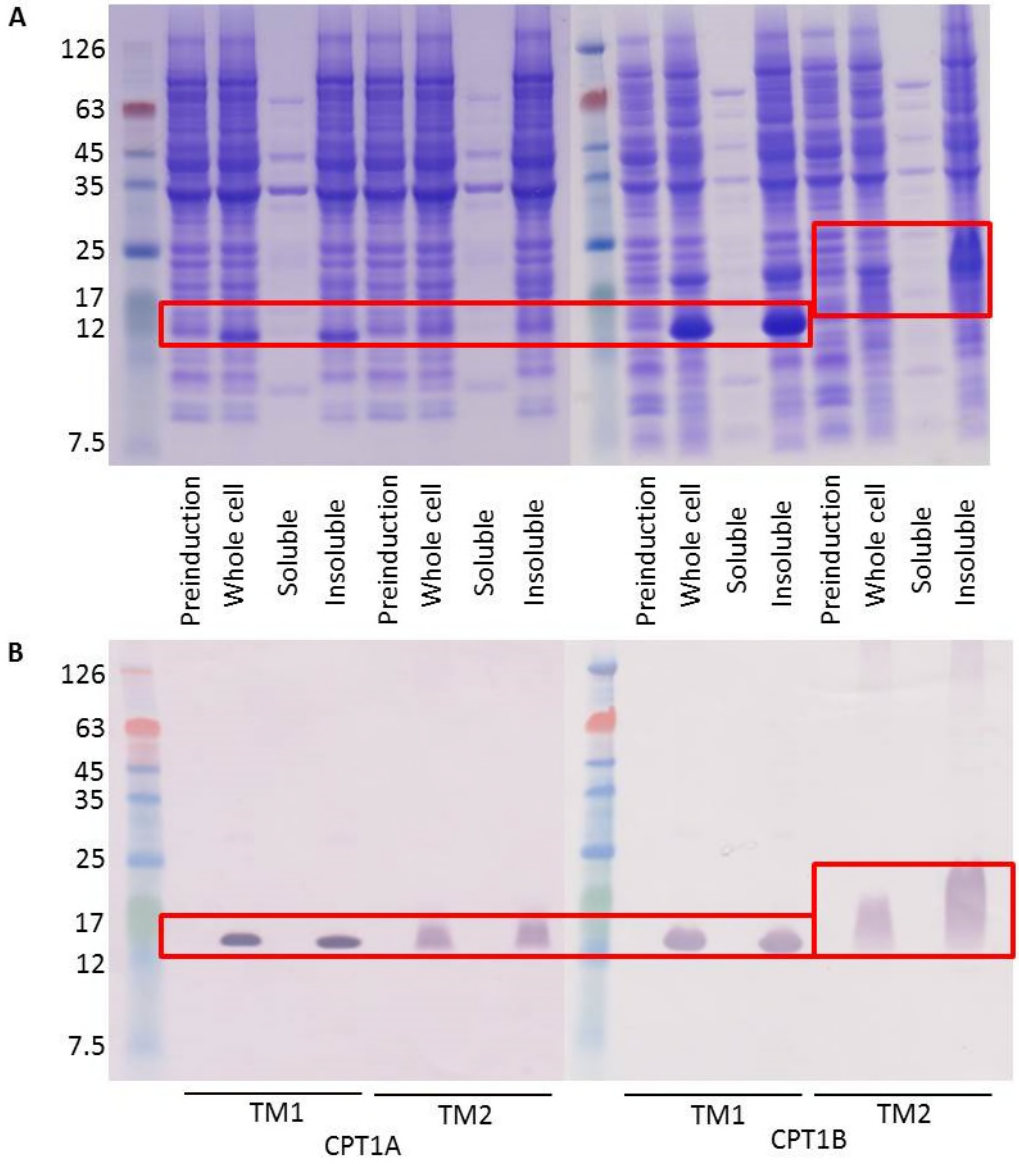


Figure 5.3 – Initial expression in LB medium. (**A**) SDS-PAGE and (**B**) anti His-tag Western blots showing the successful expression of the CPT1A and CPT1B TM domain trpLE fusion proteins in LB medium. Protein bands have been highlighted with a red box. The bands for CPT1B TM2 were particularly diffuse. All of the observed expressed protein is found in the insoluble fraction showing the trpLE domain is successfully directing protein expression to inclusion bodies. No protein expression is observed preinduction. All cell growth was performed at 37 °C with an IPTG concentration of 1 mM, and the cells were harvested after overnight growth of ~ 16 hours.

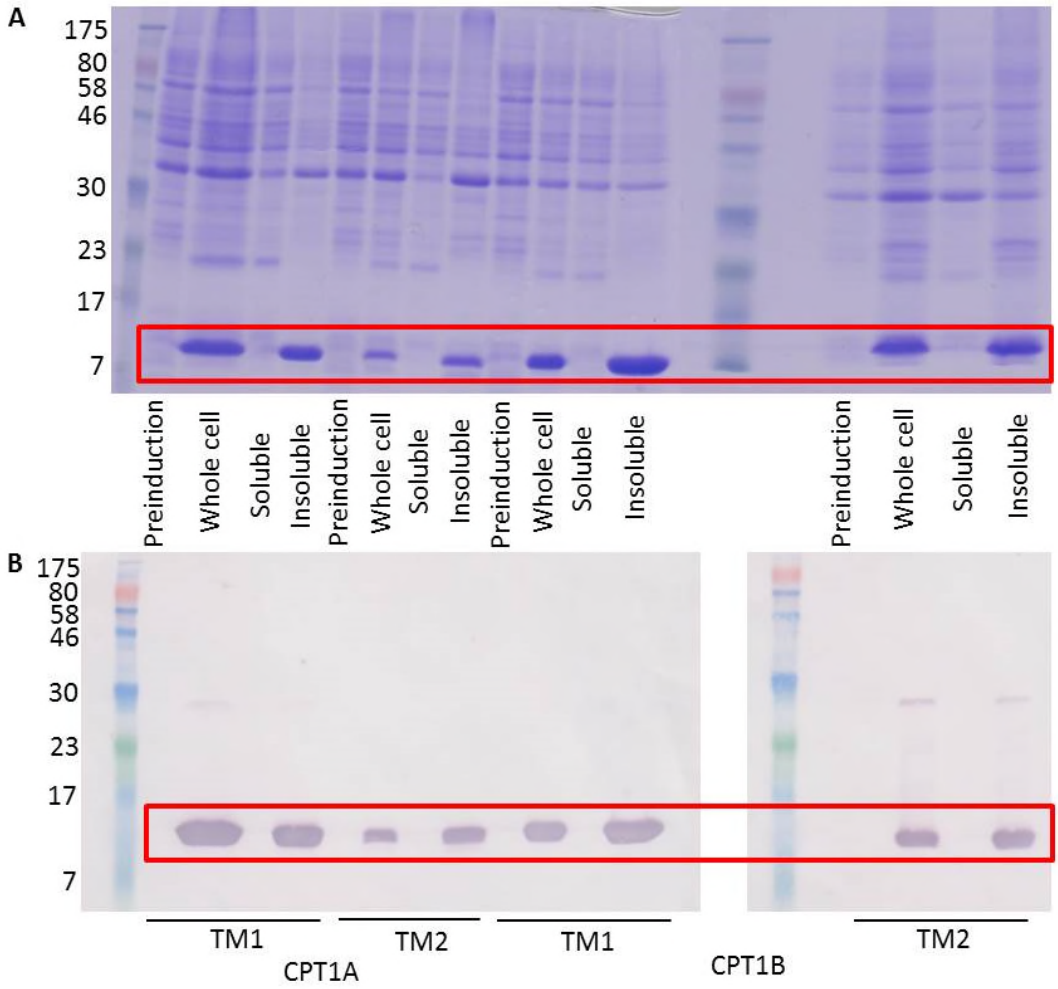


Figure 5.4 – Initial expression in M9 minimal medium. (**A**) SDS-PAGE and (**B**) anti His tag Western blots showing the successful expression of the CPT1A and CPT1B TM domain trpLE fusion proteins in M9 minimal media. Protein bands have been highlighted with a red box. Again, all of the observed expressed protein is found in the insoluble fraction, and no protein expression is observed preinduction. All cell growth was performed at 37 °C with an IPTG concentration of 1 mM, and the cells were harvested after overnight growth of ~ 16 hours.

Common reasons for low expression yields were explored and mitigated where required. Leaky expression prior to induction can cause lower than optimum yields, especially if the expressed protein is significantly toxic to the cells. To prevent this, the BL21 (DE3) pLysS cell line was used for expression, where the pLysS plasmid is present to repress leaky expression. In these cells, protein overexpression is governed by the bacteriophage T7 RNA polymerase which is

expressed via the IPTG inducible *lac*UV5 promoter. Due to the very high activity of the T7 RNA polymerase, which can transcribe approximately 5 times faster than *E. coli* RNA polymerase, there is very likely to be some low level expression even with no induction. If the expressed protein is toxic to the cell, as membrane proteins often are, this causes cells which produce low levels of the protein to be selected for which consequently reduces yields (Schlegel, et al., 2012). Through the expression of T7 lysozyme, from the pLysS plasmid the activity of the T7 RNA polymerase can be inhibited until desired. Using the BL21 (DE2) pLysS cells no detectable expression was observed prior to induction in any cultures, either in LB or M9 media, when analysed by SDS-PAGE and subsequent Western blots against the His-tag (**Figures 5.3** and **5.4**).

Another potential cause for low expression levels is poor aeration of the media during growth. Aeration is important for consistent *E. coli* growth because, without sufficient oxygen levels in the medium, *E. coli* produce high levels of acetic acid. If enough acetic acid is produced to significantly acidify the media this can seriously affect the growth rate of the bacteria. Anaerobic growth also leads to lower energy levels in the cells and therefore a subsequent lowering of protein expression levels. All large scale expression was performed with 1 L of culture in 5 L flasks, keeping to the recommended 1:5 culture volume to flask ratio and indeed lower expression was observed when 2 L cultures were grown in the same flasks.

After transformation into BL21 (DE3) pLysS cells, 10 mL starter cultures were grown overnight in the presence of kanamycin and chloramphenicol. These larger 1 L cultures for expression were then inoculated 1:100 (10 mL starter culture in 1 L of media) and allowed to grow to an OD_{600} ~0.6, which took approximately 4

hours, before induction with IPTG. While these growth optimisations helped to increase protein yields, further conditions could still be optimised.

To further decrease the cost of isotope labelling, a cell condensation method was used (Sivashanmugam, et al., 2009) where cells were initially grown in LB medium, then when the culture had reached a sufficient OD_{600} the cells were gently harvested by centrifugation, washed in M9 medium and then resuspended in half the volume of M9 media containing isotopically labelled metabolites. In this way a higher OD_{600} could be reached in a faster time than by growth in M9 medium from the point of inoculation. Additionally higher cell numbers per litre of culture were achievable without having to resort to larger volume cultures which would have required extra labelled metabolites.

5.3.1 Induction Time and Cell Density

Variations in the point of induction as well as the length of time of induction were explored to characterise their effect on culture growth and consequently protein expression yields. Cultures were induced at a range of OD_{600} from 0.4 – 1.0 and then incubated for 24 hours with regular OD_{600} readings taken. It was found that cultures induced at lower OD_{600}=0.6 or lower actually grew to higher final cell densities when allowed to for sufficient lengths of time than those induced at OD_{600}=0.8 (**Figure 5.5**).

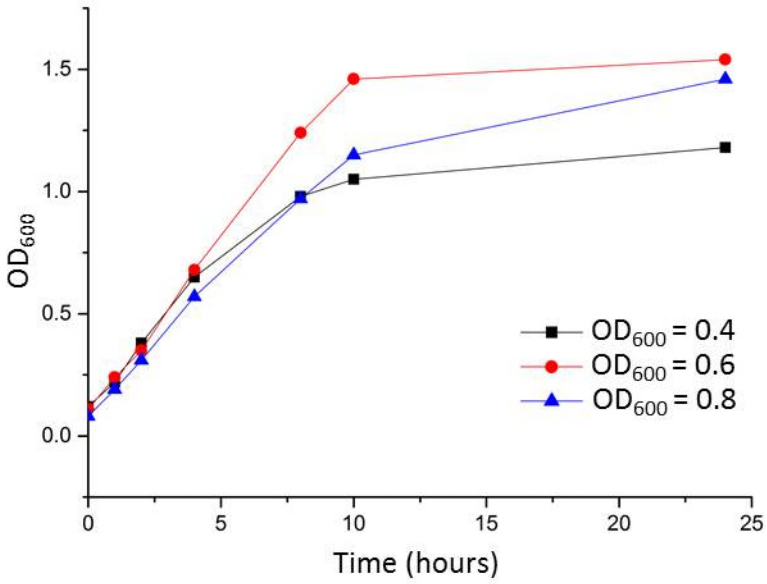


Figure 5.5 – Cell growth on induction with IPTG. OD_{600} measurements of a CPT1B TM1-trpLE fusion protein expression in LB medium at 37 °C and an IPTG induction concentration of 1 mM induced at three different OD_{600} readings (0.4 (black), 0.6 (red), and 0.8 (blue)).

The concentration of IPTG used to induce protein expression was also investigated to find optimal conditions. Three IPTG concentrations (0.3, 0.5 and 1 mM) were used to induce cultures and samples were harvested after 4, 16 and 24 hours. Samples were analysed using SDS-PAGE and Western blotting. All samples were normalised using the OD_{600} reading recorded when the sample was taken. This ensures that the same number of cells is present in each sample so any variation observed is due to the relative expression levels of those cells. Surprisingly, given that the expression system was the same for all the trpLE-CPT1 fusion proteins expressed, the expression levels of CPT1B TM1 did not vary significantly across the three IPTG concentrations tested, but CPT1B TM2 expression was affected strongly (**Figure 5.6**). The largest increase in protein expression was observed at 16 hours of induction time with the lowest IPTG concentration (0.3 mM). This protein expression level was maintained up to the 24 hour time point but did not offer a

significant increase over 16 hours. All subsequent expression was performed with an IPTG concentration of 1 mM for CPT1B TM1 and 0.3 mM for CPT1B TM2 for 16 hours before harvest. These concentrations were used for CPT1A TM1 and TM2 respectively as well.

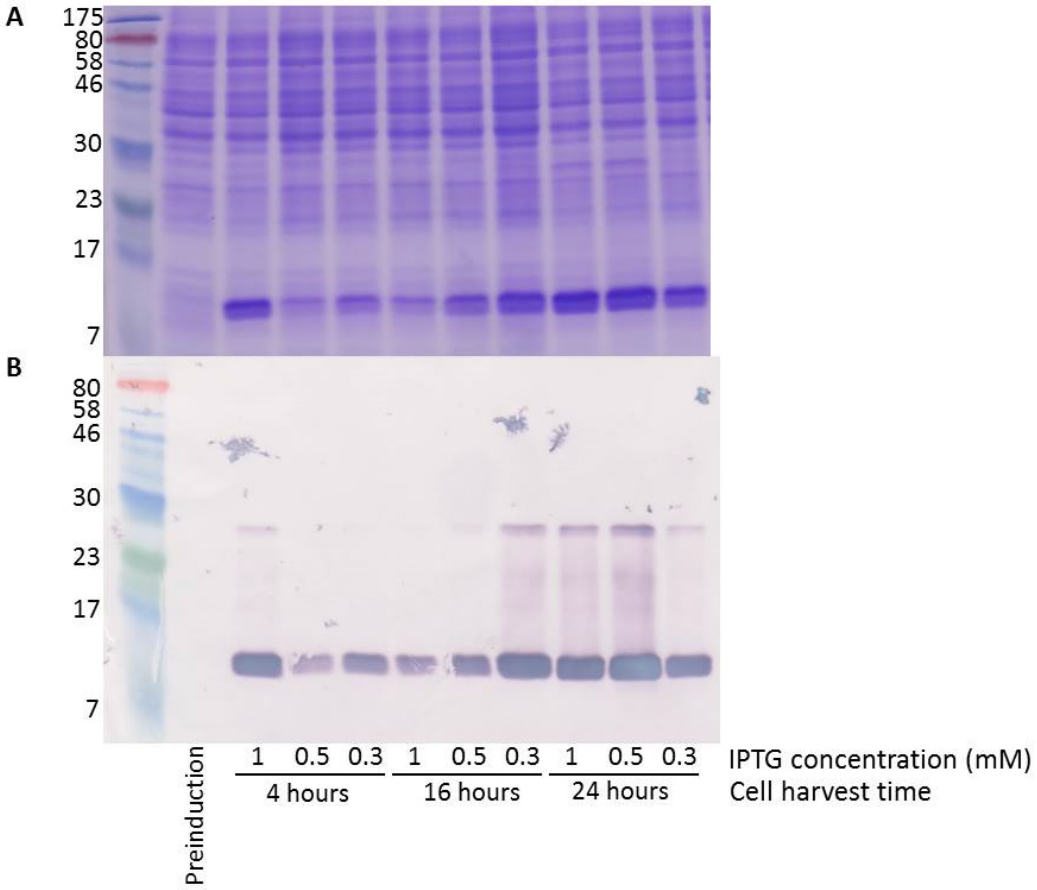


Figure 5.6 – Optimising IPTG concentration and cell harvest time. (**A**) SDS-PAGE and (**B**) anti His tag Western blot showing the optimisation of IPTG induction concentration and time of cell harvest for CPT1B TM2. All cell growth was performed at 37 °C in LB medium.

5.3.2 Temperature

The starting protocol (Claridge and Schnell, 2012) performed all expression at 37 °C , and this seemed to give good expression levels. A lower temperature of 25 °C was tested but this did not seem to affect the expression levels per cell by a great deal, however the growth rate was understandably slower at lower temperatures. 37 °C was therefore used for all subsequent expression as it gave good yields and fast growth.

5.3.3 Condensation Method for Labelled Protein Production

As mentioned above (**Section 5.2**) the cell condensation method (Sivashanmugam, et al., 2009) was used to improve the yield and reduce the cost of isotopically labelled cell cultures. Several different condensation factors from a 2 to an 8 fold concentration were tested for efficacy. In all tests the 2 fold condensation showed an increase in yield over no condensation but there appeared to be no further gain on moving to a higher condensation factor. All subsequent expressions were carried out using a 2 fold condensation when resuspending the cells in M9 minimal media.

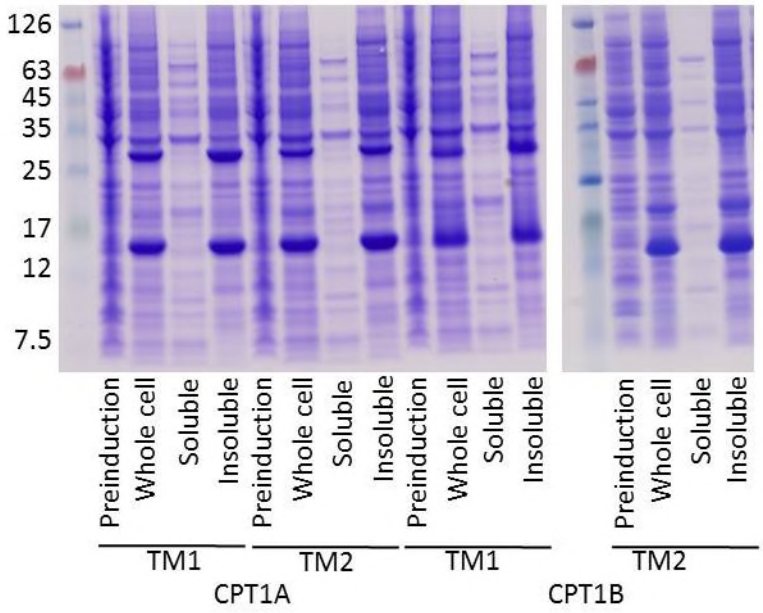


Figure 5.7 – Optimised expression in LB medium. SDS-PAGE showing the expression of the four TM domains from CPT1A and CPT1B as trpLE fusion proteins using the optimised conditions discussed: 37 °C induced with an IPTG concentration of 1 mM for TM1 and 0.3 mM for TM2 induced at an OD_{600} of 0.6, and cell harvest after 16 hours.

5.4 Improvements in Protein Extraction and Purification

Once the protein expression levels were optimised sufficiently, protein extraction and purification methods were optimised. SDS-PAGE indicated that 100% of the protein was being expressed to insoluble inclusion bodies as desired (**Figure 5.7**), hence the next step was to efficiently extract and solubilise the protein in guanidine ready for purification by Ni^{2+} metal affinity chromatography.

Initially, cells were lysed by resuspension in phosphate buffered saline (PBS) and then using a Vibra-cell tip sonicator (Sonics & Materials Inc., USA) at a power level of 10 Watts with an 80% duty cycle for 2 minutes. This lysis strategy worked well for cells grown in LB but proved insufficient for those grown in M9 minimal medium. This was an unexpected effect; however, it has been observed in the laboratory before. Longer sonication times could be used to successfully lyse cells

grown in M9 minimal medium, but this was inefficient and allowed significant heating of the samples. To improve on this, lysis was performed using a cell disruptor (Constant Systems Ltd., UK) at 30 kPSI. This proved much quicker than using the tip sonicator and so this method was adopted for all large scale cell lysis. Cells grown in M9 minimal medium often still required two passes through the cell disruptor to achieve good cell lysis levels, and so to further aid lysis BugBuster™ HT protein extraction reagent (Merck Millipore, UK) was added as a supplement to the resuspension PBS buffer.

The insoluble matter after cell lysis and centrifugation was initially solubilised by vortexing in 6 M guanidine buffer, and then allowed to mix overnight by end over end rotation. A Dounce homogeniser was tested to try and increase the amount of protein solubilised in this manner, however it was found that a significant amount of insoluble matter, and therefore protein, was not adequately solubilised and was lost in transfer to and from the homogeniser. To avoid these losses, and to replace the homogeniser, a sonication step was introduced instead prior to end over end rotation. Each sample was sonicated using a tip sonicator for 1 minute at a power level of 10 Watts and an 80% duty cycle. This served to initially break up the insoluble pellet and allow a greater amount to be resolubilised by end over end rotation overnight.

Immobilised metal affinity chromatography IMAC is a commonly used technique to purify His-tagged proteins after expression, as it relies on the affinity of the His-tag to bind to metal ions immobilised in an agarose resin. Once the His-tagged protein is bound to the Ni^{2+} resin it can be washed to remove impurities and then eluted from the resin by adding a competitive Ni^{2+} binding agent such as

imidazole. It is common to include low concentrations of imidazole throughout the binding and washing phases to keep nonspecific binding to a minimum.

After solubilisation in 6 M guanidine buffer Ni^{2+}, IMAC was used to purify the His-tagged trpLE-peptide fusion protein. The imidazole concentrations used for binding and elution were optimised (as visualised through SDS-PAGE and Western blots) to ensure the protein remained bound to the column resin during wash steps and eluted efficiently. A range of imidazole concentrations between 0 and 25 mM were tested in the wash steps, and 200 and 350 mM in the elution steps. The optimum concentrations were found to be 15 mM in the wash buffer, and 350 mM in the elution buffer (sample results for CPT1B TM domain peptides shown in **Figure 5.8**).

The number of column volumes of buffer used to wash and elute was also optimised, and ten column volumes (20 ml) were used for both washing and elution. The quality of the SDS-PAGE was compromised due to the samples containing residual guanidine and the Coomassie staining was poor. Western blotting (**Figure 5.8 B** and **D**) was performed to supplement the SDS-PAGE and alleviate the issues with poor staining.

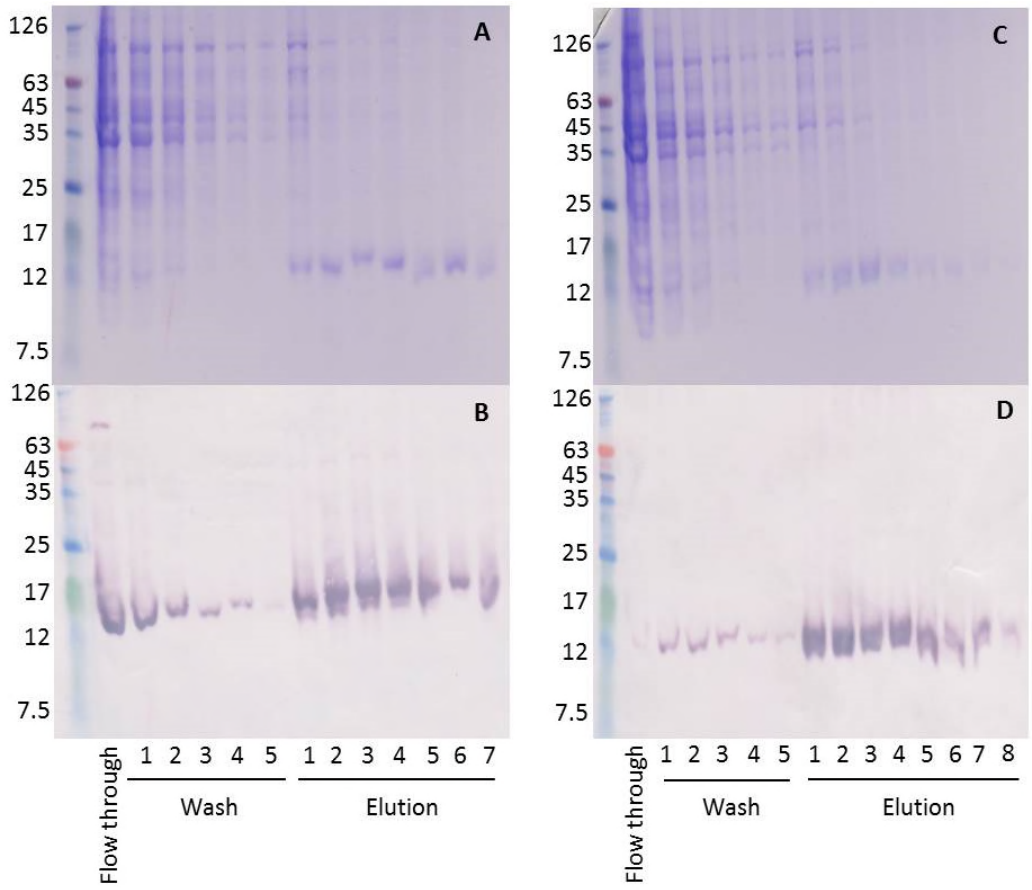


Figure 5.8 – IMAC purification of CPT1B TM1 and TM2. (**A**) SDS-PAGE and (**B**) anti-His tag Western blot of the IMAC purification of CPT1B TM1. (**C**) SDS-PAGE and (**D**) anti-His tag Western blot of the purification of CPT1B TM2. The wash and elution fractions were each 1 column volume (2 ml).

5.5 Optimisation of Protein Cleavage

CNBr was used to cleave the trpLE expression tag and the His purification tag from the desired CPT1 peptides. There were many variations in cleavage protocols found in the literature (Andreev, et al., 2010, Gupta A, 2012) which use different acidic conditions as well as different reaction times. These often vary significantly with reaction times ranging between a few hours up to several days. As a starting point a protocol from literature (Claridge and Schnell, 2012) was used, as this used the same expression system. Initially the reaction was allowed to proceed for 2-3 hours at room temperature in 70% formic acid. After further analysis of the

literature this cleavage time was increased to a full 3 hours and the conditions were altered to 100% formic acid and the reaction was protected from light by covering the reaction vessel in aluminium foil. These alterations significantly improved the quantity of cleaved protein observed in subsequent HPLC purification steps (as estimated by peak area). Although the reaction did not go to completion, the risk of generating side reaction products when using reaction times longer than 3 hours would have meant that more rigorous purification methods would have to be used, and so these conditions were not altered further.

After cleavage from the trpLE domain, the small size of the CPT1 TM peptides meant that they did not stain well using Coomassie brilliant blue unless present at high concentrations. Where high concentrations could not be achieved silver staining was used in concert with Coomassie staining to detect the peptides (**Figure 5.9**).

After cleavage with CNBr the reaction mixture was dialysed against water and then lyophilised. During this dialysis step no precipitation was observed, and all of the lyophilised material could be redissolved in pure water. This was an unexpected result as the expressed peptides were derived from TM domains and were expected to be hydrophobic. The inclusion of residues beyond the hydrophobic core of the CPT1 TM domains apparently provided enough hydrophilic residues to allow for reasonable water solubility. Unfortunately, this precluded the purification of the peptides by precipitation in water.

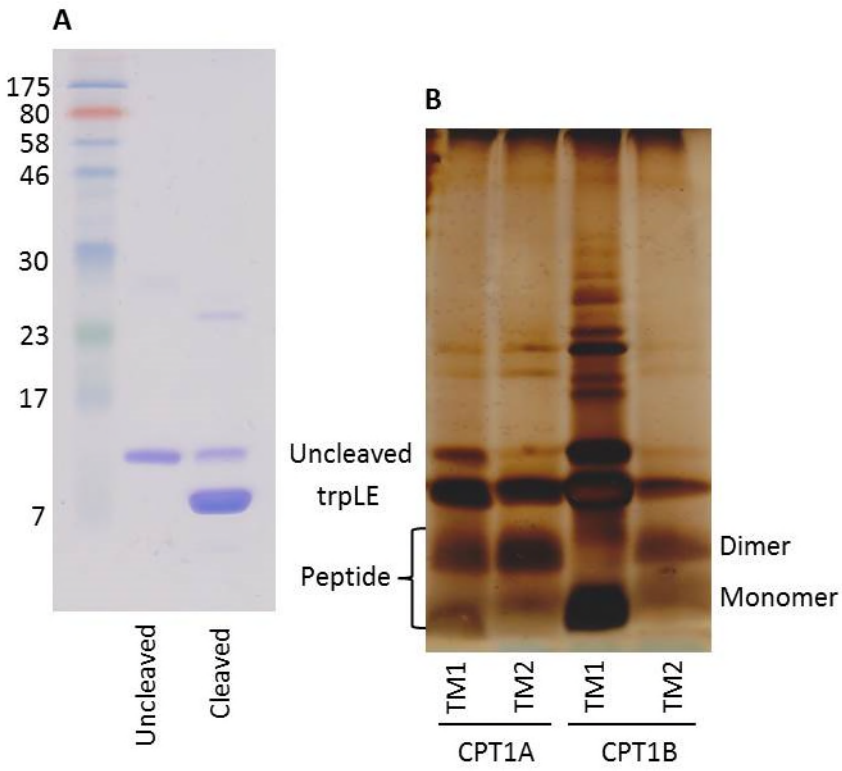


Figure 5.9 – Cleavage with CNBr. (**A**) SDS-PAGE stained using Coomassie brilliant blue showing samples of CPT1B TM1 before and after cleavage using CNBr. The uncleaved trpLE-CPT1B fusion protein and the cleaved trpLE domain can be clearly seen, however the band for the cleaved CPT1B TM1 peptide is extremely faint. (**B**) SDS-PAGE stained using silver showing samples of CPT1A and CPT1B TM peptides after cleavage with CNBr. In addition to those bands seen when staining with Coomassie, there are prominent bands at lower molecular weights corresponding to the CPT1 peptides. In CPT1A TM1, TM2 and CPT1B TM2 samples, bands can be seen for both monomer and dimer.

5.6 Purification After Protein Cleavage

What initially seemed like a relatively simple purification, the separation of the cleaved peptide from the His-tagged trpLE domain and any remaining uncleaved protein, required multiple techniques and significant optimisation to achieve efficiency and reliability. Reversed phase high performance liquid chromatography (RP-HPLC) was the initial technique tested and was eventually the one selected, however due to early difficulties several other techniques were also tested. The first of these utilised a dialysis size exclusion strategy. As dialysis tubing had been used

effectively in previous purification steps, after the IMAC column to purify after expression and again after the CNBr cleavage, this seemed like a sensible approach. Both of these previous steps resulted in excellent retention of the desired proteins when appropriate MWCO dialysis tubing was used. Protein was loaded into dialysis tubing with MWCO 10 kDa, which was surrounded by larger diameter tubing with MWCO 2 kDa, just as is used in the post CNBr cleavage purification step. This should allow the peptides to migrate through the inner tube but remain trapped between the inner and outer tubing. In practice, sufficiently large volumes could not be maintained in this space and consequently a majority of the peptide was in fact retained within the inner tubing even over several days and multiple swaps of water. To try and improve on this strategy, centrifugal concentrators were then tested. A carefully chosen MWCO centrifugal concentrator should allow the peptide to elute but retain the larger unwanted proteins. Initially 10 kDa MWCO concentrators were tested but when these showed complete retention of both peptides and larger proteins, a range of MWCOs (5-30 kDa) were tested. Unfortunately none of these were successful and in almost all cases complete retention of all peptides and proteins was observed even at 30 kDa MWCO (**Figure 5.10**). This was unexpected as even the largest protein (uncleaved trpLE-peptide fusion) with a molecular weight of ~17 kDa should have eluted using these concentrators. Nonspecific binding of the proteins to the membrane in the concentrators or protein aggregation could explain this observation.

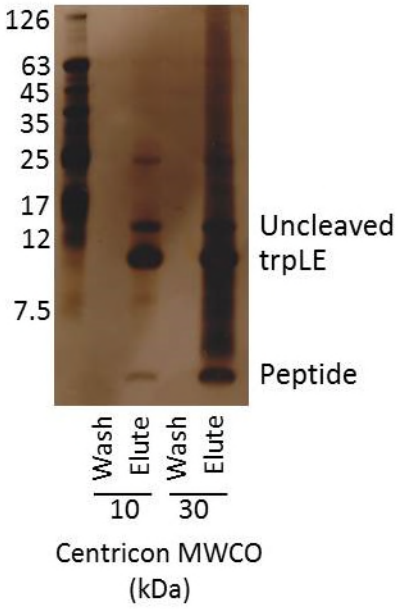


Figure 5.10 – Purification using centrifugal concentrators. Representative data showing the failure of any MWCO centrifugal concentrators to separate CPT1 TM peptide from trpLE and uncleaved fusion protein. The peptide was expected to appear in the wash lanes with the trpLE and uncleaved protein retained until elution.

After these failures, other strategies were considered carefully and two likely techniques were tested further. The first, gel filtration, has been successfully used to purify other proteins within the group. To this end a HiLoad 16/600 Superdex 75 pg (GE Healthcare Life Sciences, UK) was used with an AKTA-FPLC (GE Healthcare Life Sciences, UK). If successful this system would have been a relatively fast method to obtain purified peptide in pure water. Results from this attempt are shown in **Figure 5.11**. Unfortunately it was discovered that the uncleaved protein, trpLE, and the CPT1B TM1 peptide tested all co eluted in a single peak. This method was not pursued further as the sizes of the proteins to be separated were at the very low end of the manufacturers specifications for this column.

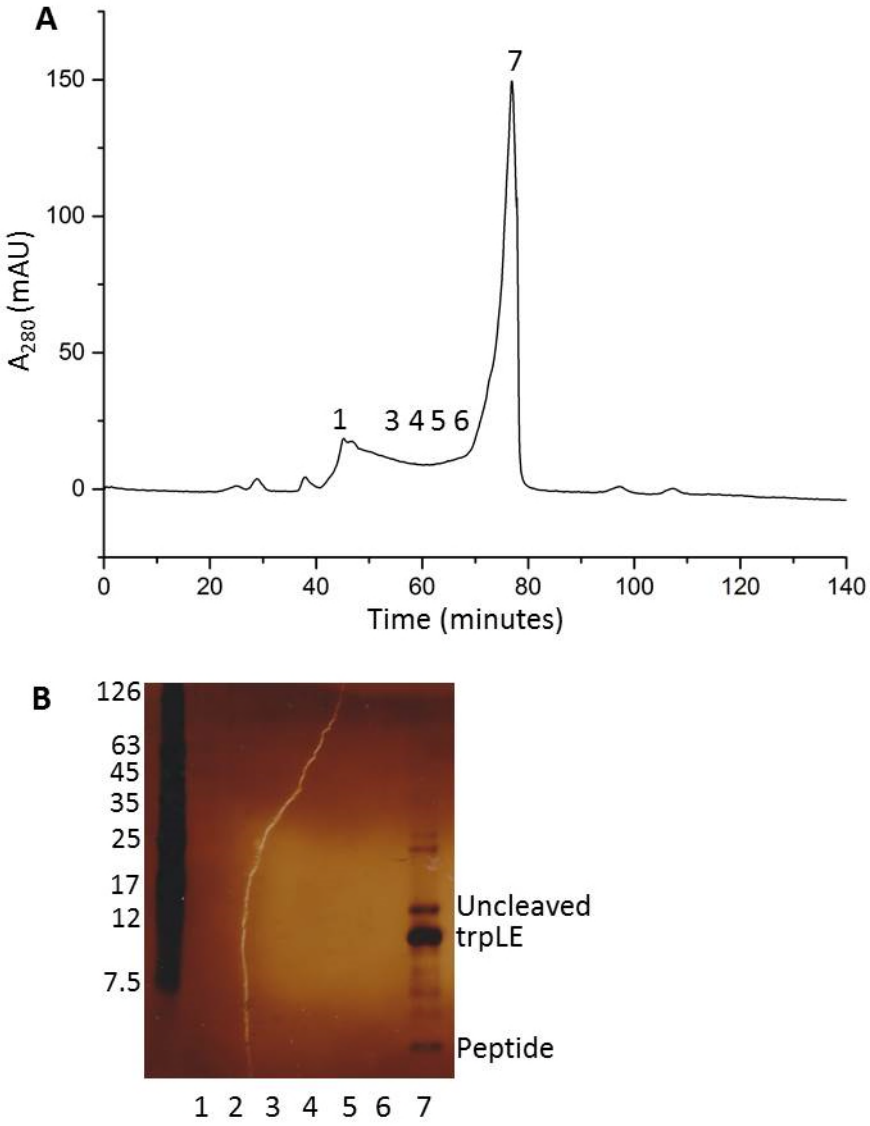


Figure 5.11 – Purification by FPLC. (**A**) FPLC chromatogram using a HiLoad 16/600 Superdex 75 pg column. (**B**) Several fractions (1-7) were analysed by SDS-PAGE and then silver staining to visualise elution of the proteins. Uncleaved protein, trpLE and CPT1B TM1 peptide are all co eluted.

Due to the effectiveness of using IMAC columns earlier to purify the fusion protein after expression, the use of another IMAC column to separate the His tagged fusion protein and trpLE tag from the cleaved peptide was investigated. In this case the pure peptide should not bind to the Ni resin and would be found in the eluate. This seemed like a promising strategy, however once again experiments showed that the majority of the peptide was being maintained on the column and only very low amounts were eluting in the column flow-through and the wash steps. In an attempt

to overcome this, the Ni^{2+} metal ions on the resin were replaced with Co^{2+}. Co^{2+} has been shown to have significantly better specificity of binding, so should reduce non-specific protein binding, if this indeed the cause of the peptide retention observed. The disadvantage to using Co^{2+} over Ni^{2+} is a reduction in the overall binding affinity, however if pure peptide could be isolated this would be a minor inconvenience in comparison. The use of Co^{2+} ions however did not appear to make a difference to any of the protein binding and once again all three protein/peptide components were retained on the column until elution with imidazole.

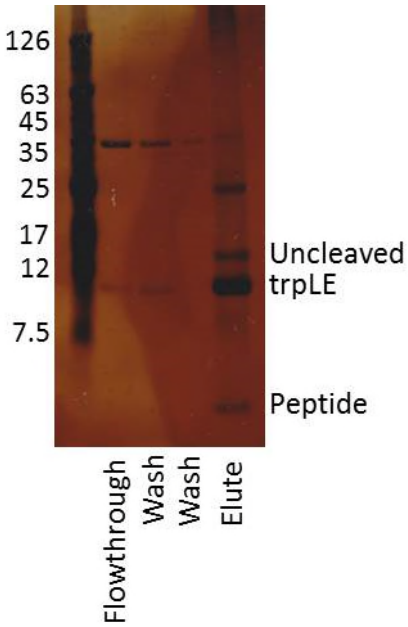


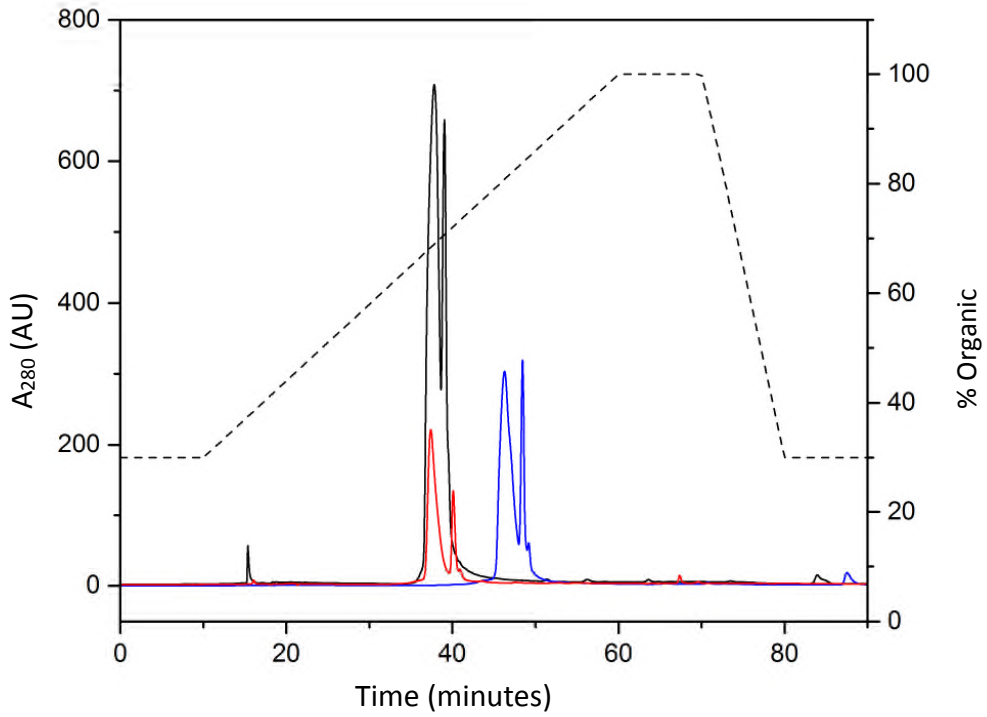
Figure 5.12 – Purification by IMAC. Results of the Co^{2+} IMAC purification of CPT1B TM1 peptide showing almost complete retention of uncleaved protein, trpLE and peptide on the column until elution with imidazole.

These further failures using alternate techniques required a return to reverse phase HPLC and a large amount of optimisations in solvent conditions, flow rate, column substrate and solvent gradient for each peptide to be purified. These individual optimisations will be discussed in greater depth in the following sections.

5.6.1 Optimisations to the HPLC Purification Conditions

Due to the difficulties in obtaining pure peptide described so far, the decision was made to focus efforts on the CPT1B TM peptides first to allow for biophysical experiments to begin as soon as possible. Several different reverse phase HPLC columns were used in initial attempts to optimise purification including a phenyl hexyl, a C4 and a C5 column (Phenomenex, UK) (results from C4 shown in **Figure 5.13**). Once the C5 column had been selected as providing the greatest reliability and resolution of the available columns further parameters could be investigated. Firstly several organic phases were tested: two common organic phases (acetonitrile and isopropanol); as well as one found from literature (Claridge and Schnell, 2012) 57% isopropanol, 38% acetonitrile, and 5% water. For the acetonitrile and isopropanol phases 100% water was used as the aqueous phase, however for the combination organic phase a modified aqueous phase was used as well (95% water, 5% isopropanol) (**Figure 5.13 B**). All organic and aqueous phases tested contained 0.1% trifluoroacetic acid (TFA) to ensure an optimal pH range. These different solvent systems were tested with a standard 30-100% organic gradient over 90 minutes.

A



B

	Organic	**Aqueous**
Black	100% acetonitrile	100% water
Blue	100% isopropanol	100% water
Red	57% isopropanol 38% acetonitrile 5% water	95% water 5% isopropanol

Figure 5.13 – Initial purification by HPLC. (**A**) Differences in HPLC retention between the (**B**) 3 sets of phases tested using a C4 column.

Although the mixed isopropanol/acetonitrile/water and water/isopropanol solvents provided an improvement in the resolution, it was still not sufficient for easy separation of the peptide from other components without further improvements to the running conditions. The peaks for all three expected components (uncleaved protein, peptide, and His-tagged trpLE) were eluting very close together on the C4 and phenyl hexyl columns; a switch to a C5 column gave an immediate improvement in resolution. Retention times were however significantly shorter than the C4 with elution of some components immediately after the void volume of the column. To correct for this the starting organic content was adjusted and tested for elution under

isocratic solvent conditions of 10, 20 and 30% organic phase. No elution was observed under either 10 or 20% isocratic conditions after 120 minutes and so 20% organic was selected as the new starting condition. Furthermore, as a significant amount was eluting at 30%, but not at 20%, organic, the gradient was significantly reduced at low organic percentages and then increased at higher organic concentrations. This coupled with over all longer run times of up to 110 minutes greatly improved the separation of the observe peaks in the chromatogram, and allowed isolation of pure CPT1 peptides from the cleaved trpLE domain, and any remaining uncleaved protein. The flow rate was also increased from 1 ml/min which helped to improve and narrow the peak shapes when using these longer run times with shallow gradients. A final flow rate of 1.5 ml/min was used to purify CPT1B TM1 and 2 ml/min for CPT1B TM2.

Fractions collected during HPLC purification were analysed by SDS-PAGE to verify purity of the CPT1 peptides. SDS-PAGE analysis was useful in addition to mass spectrometry, used later, as it was discovered that the trpLE domain or residual uncleaved protein was not always detectable using just mass spectrometry. Data is shown in **Figure 5.14** for the purification of CPT1B TM1 and **Figure 5.15** for CPT1B TM2. Fractions collected from the peaks were also analysed with anti His-tag Western blotting to confirm removal of the His-tag from the CPT1 peptides after cleavage.

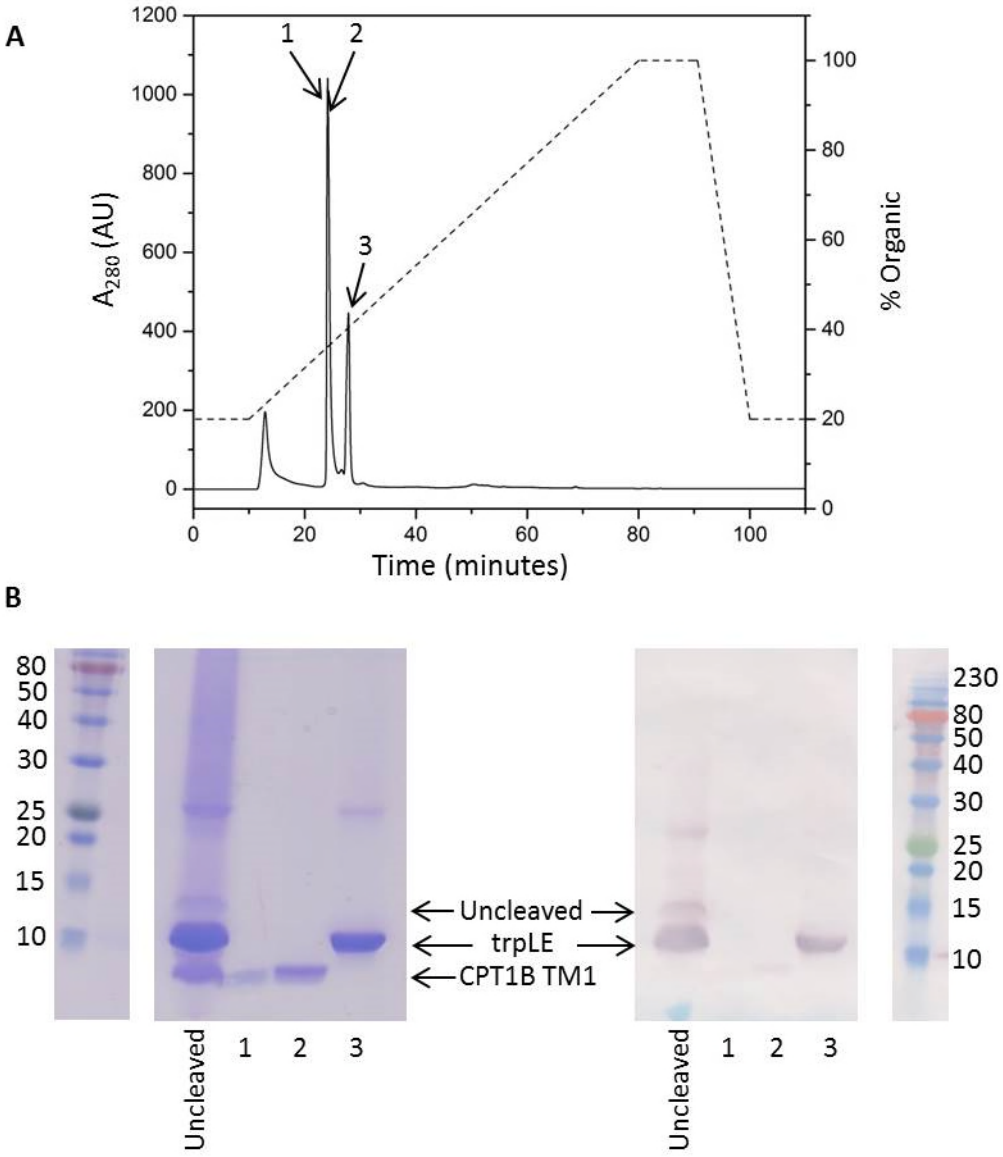


Figure 5.14 – Optimised purification of CPT1B TM1 by HPLC. (**A**) HPLC chromatogram of the purification of CPT1B TM1 and (**B**) SDS-PAGE and anti His-tag Western blot of three fractions indicated by the numbered lanes showing uncleaved protein, trpLE and CPT1B TM1 peptide when stained with Coomassie, and only the uncleaved protein and trpLE on the Western blot as the His-tag has been removed from the CPT1B TM1 peptide. The dashed line on the HPLC chromatogram shows the solvent gradient used.

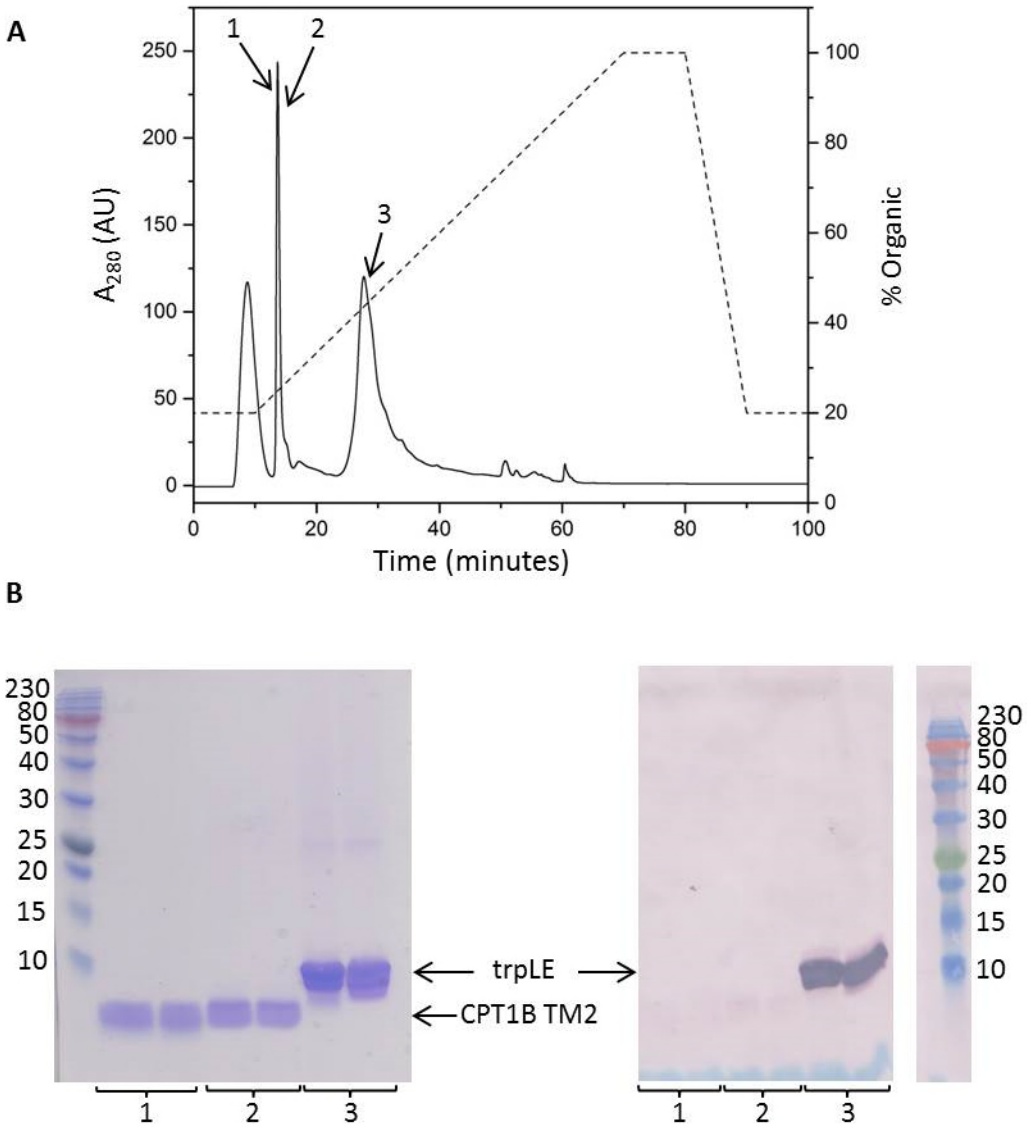


Figure 5.15 – Optimised purification of CPT1B TM2 by HPLC. (**A**) HPLC chromatogram of the purification of CPT1B TM2 and (**B**) SDS-PAGE and anti His-tag Western blot of three fractions indicated by the numbered lanes showing uncleaved protein, trpLE and CPT1B TM2 peptide when stained with Coomassie, and only the uncleaved protein and trpLE on the Western blot as the His-tag has been removed from the CPT1B TM1 peptide. The dashed line on the HPLC chromatogram shows the solvent gradient used.

5.7 Analysis of CPT1B Peptides by Mass Spectrometry

After confirming the purity of the CPT1 peptides, they were then analysed by mass spectrometry using a MicrOTOF electrospray ionisation time of flight instrument (Bruker, UK). The correct masses, calculated from the peptide sequences,

were found for CPT1B TM1 (3245 Da) (**Figure 5.16**) and TM2 (3137 Da) (**Figure 5.17**). ^{15}N isotopically labelled CPT1B TM1 was also tested (**Figure 5.18**) to check the increase in mass over the unlabelled peptide. This shows the level of ^{15}N incorporated into the peptide. The expected increase in mass with 100% ^{15}N incorporation for CPT1B TM1 was 38 (+1 for each backbone N atom, +1 for each tryptophan, +1 for each lysine, +1 for each asparagine, and +2 for each arginine). The difference in mass between labelled and unlabelled CPT1B TM1 experimentally was ~35.5 giving an ^{15}N incorporation of >93%.

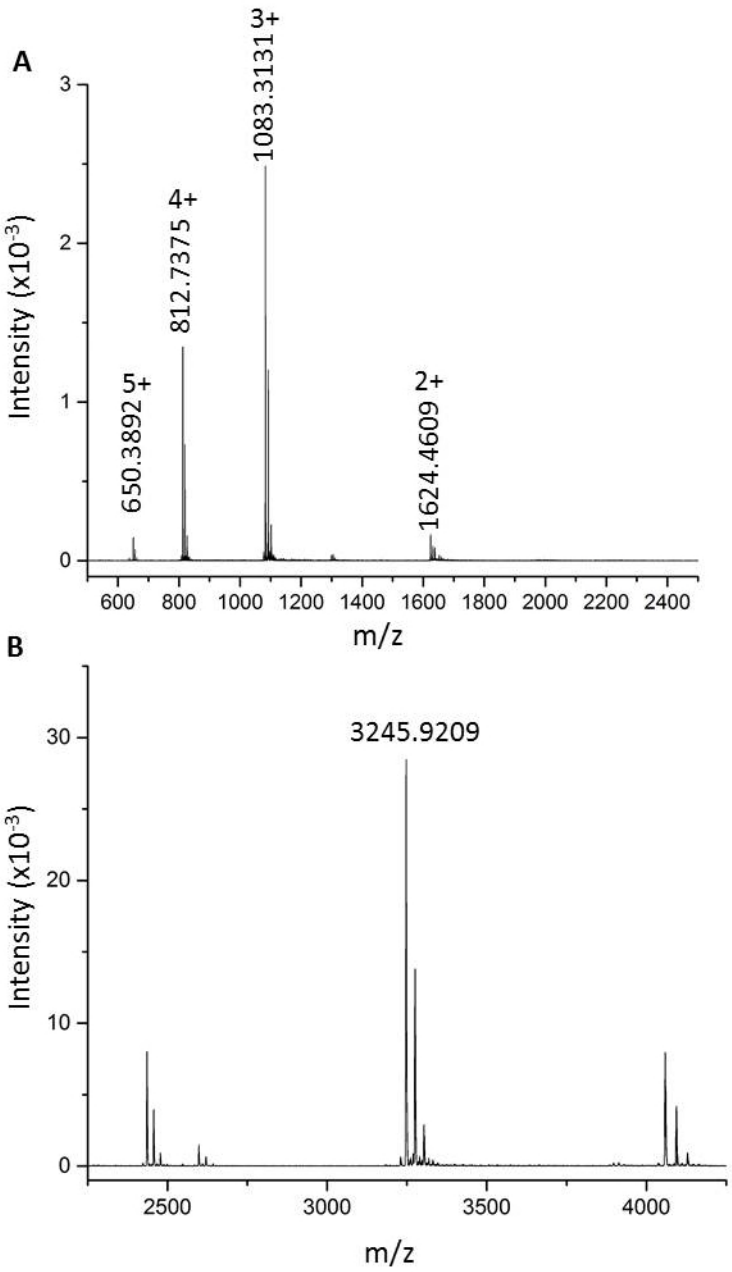


Figure 5.16 – Mass spectrum of purified CPT1B TM1. (**A**) Four distinct charge states that deconvolute to spectrum (**B**) showing the major peak at the expected mass for this peptide (3245 Da).

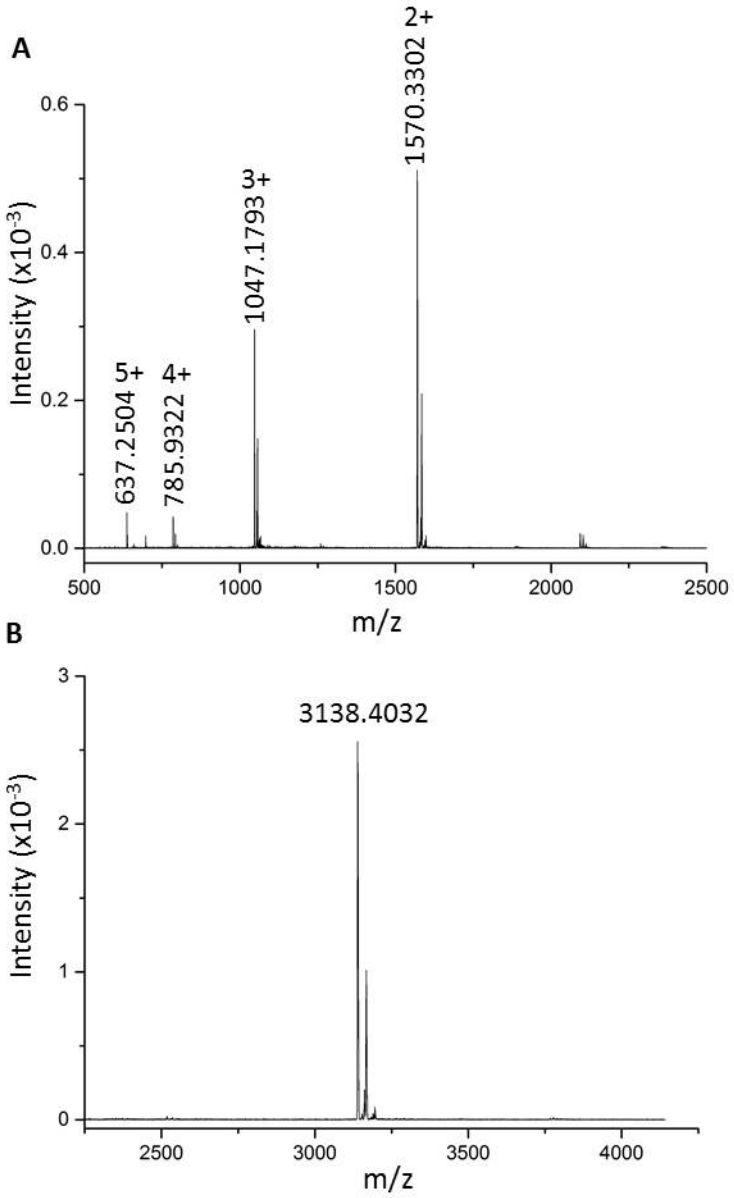


Figure 5.17 – Mass spectrum of purified CPT1B TM2 peptide. (**A**) Four distinct charge states that deconvolute to spectrum (**B**) showing the major peak at the expected mass for this peptide (3137 Da).

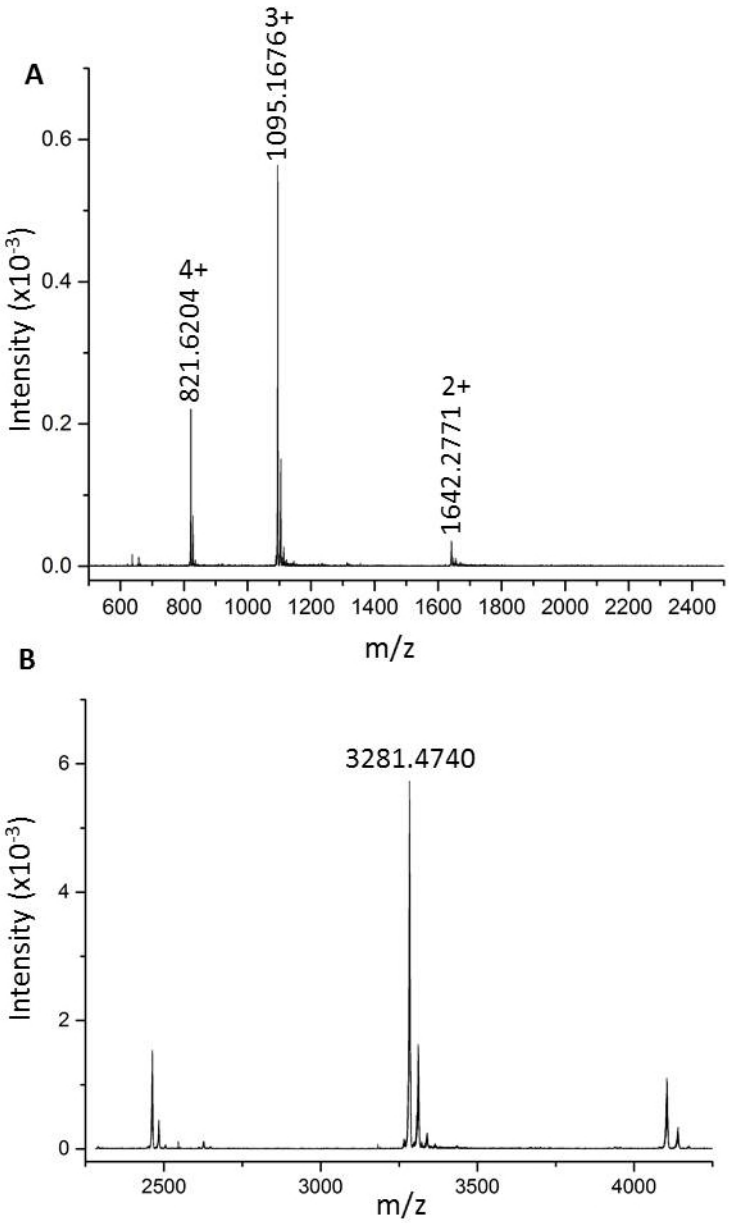


Figure 5.18 – Mass spectrum of purified ^{15}N labelled CPT1B TM1. (**A**) Three distinct charge states that deconvolute to spectrum (**B**) showing the major peak at a mass of 3281.4740 Da. This is an increase of 35.5531 Da over the unlabelled CPT1B TM1 peptide giving >93% ^{15}N incorporation on labelling.

5.8 Summary

The expression conditions using *E. coli* BL21 (DE3) pLysS were optimised for the expression of His-tagged CPT1-trpLE fusion proteins. These optimum conditions were found to include cell growth at 37 °C, IPTG induction at an OD_{600} of 0.6 and a concentration of 1 mM for CPT1 TM1 peptides, and 0.3 mM for CPT1 TM2 peptides with a harvest time of 16 hours. Now that the purity and mass of the CPT1B TM1 and TM2 peptides had been established, further biophysical experiments were performed, and due to the good ^{15}N labelling efficiency demonstrated, heteronuclear NMR experiments were also implemented. These experiments will be discussed in the following chapter. Unfortunately, despite repeated attempts to purify the TM domain peptides of CPT1A using the same methods as for CPT1B, on analysis by SDS-PAGE and NMR the purity of these peptides was compromised. Data is shown to illustrate this in the following chapter in **Figure 6.11**.

6. Measuring CPT1B TM Domain Interactions Using *In Vitro* Experiments

6.1 Introduction

In **Chapters 3** and **4** the interactions between CPT1A and CPT1B TM domains were studied *in vivo* in the outer membrane of *E. coli* using the GALLEX assay. The TM domains from both CPT1A and CPT1B were found to be capable of homotypic and heterotypic interactions and the sequence specificity of these interactions was studied. Motifs/residues that promoted these interactions were identified for all the TM domains, however the residue implicated by hetero GALLEX experiments in CPT1B TM1 (G_{57}) (**Figure 4.8**) did not show the same effect in the double mutant hetero GALLEX experiment (**Figure 4.11**). Due to difficulties in purification explained in **Chapter 5** the decision was made to focus on the TM domains of CPT1B in order to attempt to refine the model proposed with regard to the sequence specificity of CPT1B heterotypic interactions.

The optimal length for TM domain sequences in the GALLEX experiments was 18 amino acids in order to provide sufficient difference in signal between positive and negative controls in the assay. The average length a TM domain needs to be to span a membrane is ~17.3 amino acids (Hildebrand, et al., 2004) so this restriction does not pose a serious issue for many TM domains, however the native TM domains of CPT1A and CPT1B are predicted to be 22 amino acids in length (predicted sequences in **Section 3.2**). This additional length could lead to tilting of the CPT1 TM domains within the membrane, facilitating interactions that are not allowed with the length restrictions in GALLEX. We wished to study the structure and interactions of the entire TM domains, as well as obtain a measure of the

strength of association and order of oligomerisation.

The expressed CPT1B TM domain peptides were studied *in vitro* using two biophysical techniques: circular dichroism (CD) and NMR. This study aimed to corroborate the information obtained through the GALLEX experiments described, as well as to investigate whether the length restrictions for the TM domains in the GALLEX assay are preventing native interactions from being observed.

6.2 Circular Dichroism (CD) Measurement of Secondary Structure for CPT1 TM Domains

CD spectroscopy is a form of absorbance spectroscopy that measures the difference in the absorbance of right-handed and left-handed circularly polarised light. These measurements are extremely sensitive to the secondary structure of peptides and proteins. CD spectra of proteins can be used to estimate the secondary structure composition in a protein (i.e. alpha helices, parallel and antiparallel beta sheets, turns and disordered regions). The key peaks are all located in the ultraviolet portion of the spectrum between 190 and 260 nm and are shown in **Figure 6.1**.

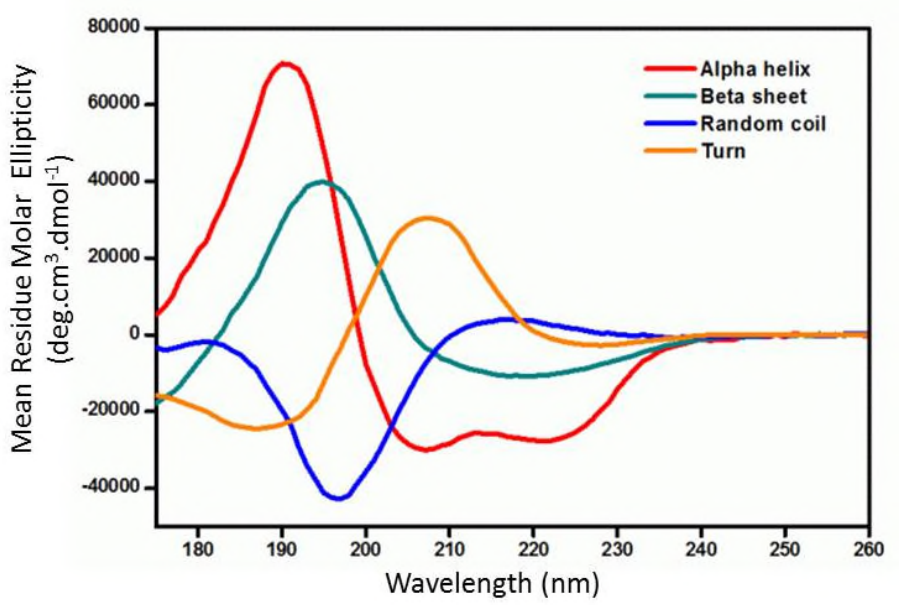


Figure 6.1 – The characteristic maxima and minima of the different secondary structures in proteins: alpha helical (red); beta sheet (green); random coil (blue); and turns (yellow). Figure taken from online fitting tool CD_Fit5 here: http://srcd_fit5.nsrrc.org.tw/

The CPT1B TM domains are thought to be helical in the full length protein and so a strong alpha helical signal was expected. The characteristic CD spectrum for alpha helical proteins contains a strong positive maximum at 195 nm and two negative maxima at 208 and 222 nm. To allow easy comparison between CD spectra, it is common to convert the machine units of millidegrees to mean residue molar ellipticity (MRE) which normalizes the signal for the numbers of residues and the protein concentration. This was performed using **equation 6.1** where n is the number of peptide bonds:

$$MRE = \frac{Ellipticity\ (mdeg) \times 10^6}{Pathlength\ (mm) \times Protein\ Concentration\ (\mu M) \times n} \qquad \textbf{(6.1)}$$

6.2.1 CD Conditions and Optimisation

The conditions used for CD experiments were designed to be consistent with the NMR experiments discussed later in this chapter. To this end, a sodium phosphate buffer containing dodecylphosphocholine (DPC) detergent was used. A range of peptide and DPC concentrations were tested.

When changing the DPC concentration, it was important to estimate the concentration of micelles at a specific DPC monomer concentration. **Equation 6.2** shows the calculation used to convert monomer concentration to micelle concentration.

$$Micelle\ Concentration = \frac{Monomer\ Concentration - CMC}{Aggregation\ Number} \qquad \textbf{(6.2)}$$

Representative values for the critical micelle concentration (CMC) and aggregation number of DPC were taken from the literature; the CMC of DPC is 1.5 mM and the aggregation number is ~75 at 25 °C (Arora and Tamm, 2001). To allow for 100% incorporation of the peptides into DPC micelles, a ratio of at least one micelle per peptide was maintained. For a peptide concentration of 50 µM this equated to a DPC monomer concentration of at least 5.25 mM. To minimise peptide-peptide interactions, a significantly higher micelle:peptide ratio was also used: ~20:1 or 100 mM DPC. This increased detergent concentration increases the likelihood of micelle incorporation as well as preventing the crowding of multiple peptides into the same micelle (potentially promoting interactions that may not occur without this pressure).

Since the CPT1B TM peptides were discovered to be soluble in water, an important question to be answered was whether they were forming their native alpha helical state in buffer and/or water, or whether a membrane mimetic was required to stimulate helix formation. To this end, increasing concentrations of DPC were added to a sample of CPT1B TM1 to measure any enhancement in the alpha helical signal with increasing DPC concentration (**Figure 6.2 A**). The CD spectra were also analysed using the online tool Dichroweb (**Section 2.11**) to obtain percentage helical content values. These are shown plotted against DPC concentration in **Figure 6.2 B**.

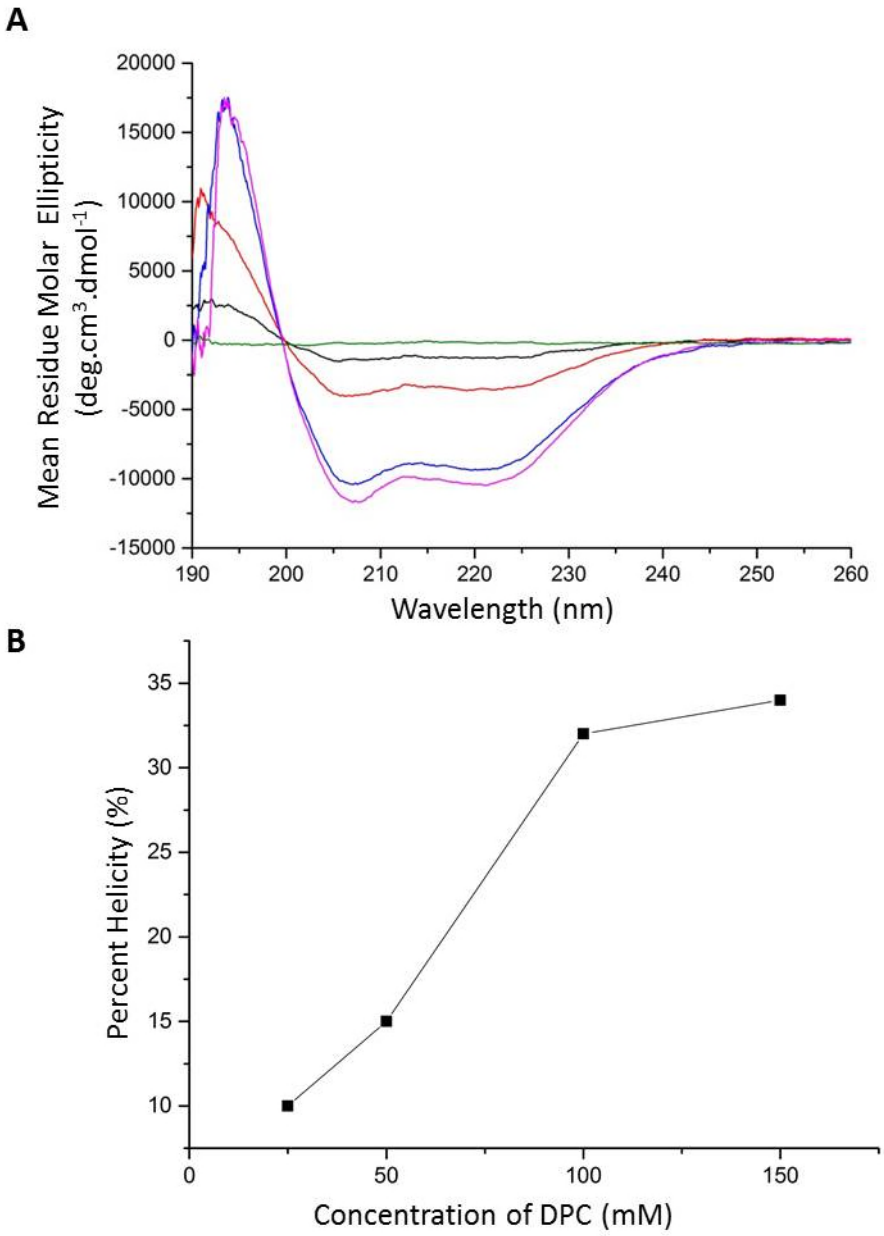


Figure 6.2 – Detergent is required for helical structure in CPT1B TM1. (**A**) CD spectra of 50 µM CPT1B TM1 peptide solubilised in sodium phosphate buffer (30 mM NaPi, 20 mM NaCl, pH 5.8) containing increasing DPC concentrations of 25 mM (black), 50 mM (red), 100 mM (blue), and 150 mM (magenta). Blank buffer shown in green.

These spectra clearly show that the alpha helical content of CPT1B TM1 was greatly enhanced in the presence of DPC and increases with DPC concentration up to and including [DPC]=150 mM. Since the helical content does not vary significantly in samples containing 100 mM vs. 150 mM DPC, and to reduce any noise introduced

by light scattering at higher micelle concentrations, a concentration of 100 mM DPC was used in all future experiments. Individual CD spectra were then collected for CPT1B TM1 (**Figure 6.3**) and TM2 (**Figure 6.4**).

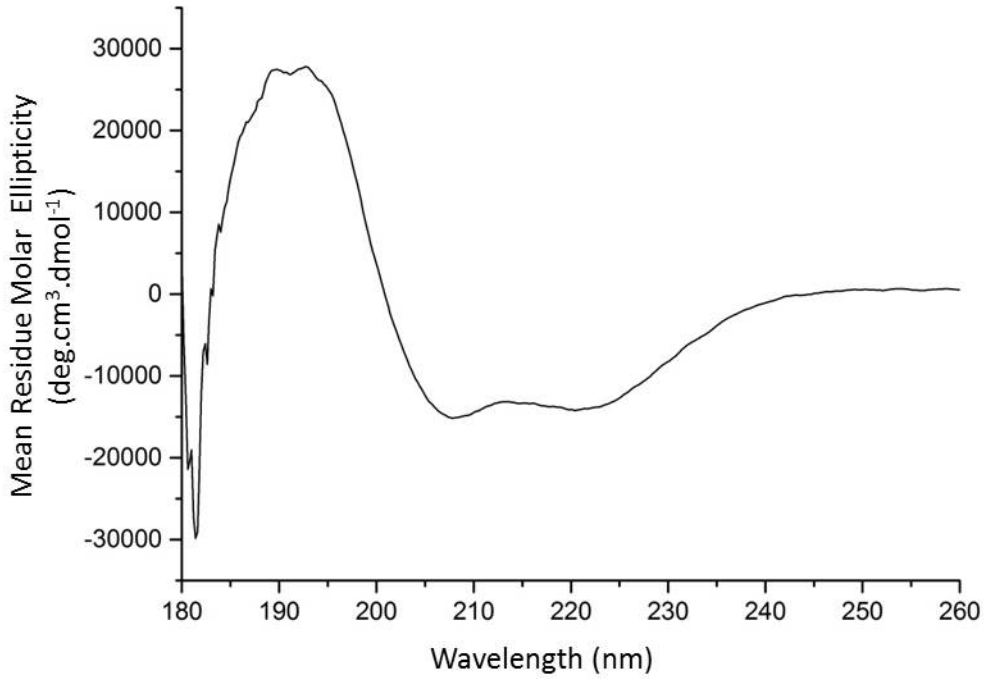


Figure 6.3 – CD spectrum of CPT1B TM1. 50 µM CPT1B TM1 in 100 mM DPC in buffer (30 mM NaPi, 20 mM NaCl, pH 5.8) showing characteristic alpha helical peaks at 195, 208 and 222 nm.

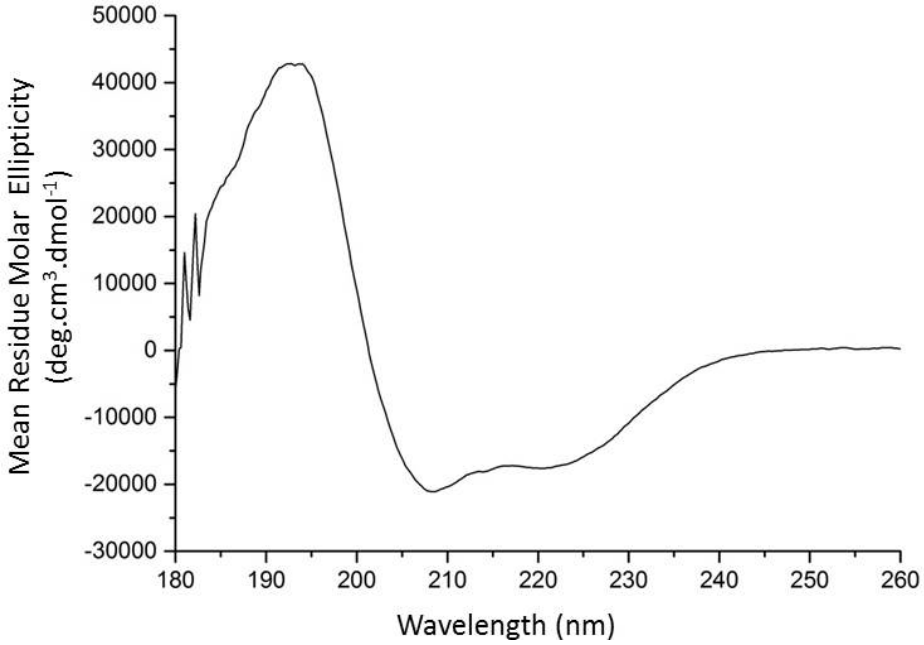


Figure 6.4 – CD spectrum of CPT1B TM2. 50 µM CPT1B TM2 in 100 mM DPC in buffer (30 mM NaPi, 20 mM NaCl, pH 5.8) showing characteristic alpha helical peaks at 195, 208 and 222 nm.

6.2.2 Measuring CPT1B TM1 – TM2 Interactions Using CD Spectroscopy

Once the alpha helicity of both CPT1B TM1 and TM2 had been confirmed in DPC micelles, CD spectroscopy was used as a method for detecting TM1 – TM2 interactions. This could be achieved, either through an enhancement in their helicity on interaction, or if a coiled coil structure is formed this can be observed. This approach has been used in the past to study synthetic peptides of the kinesin neck region (Tripet, et al., 1997). Sections of the protein were analysed for secondary structure and if the two helical peptides interact through a coiled coil interaction, the ratio of the two negative alpha helical signals is altered. A 222:208 nm ratio >1 indicates coiled coil formation, and <1 indicates a non-coiled coil alpha helical structure or a monomer (Tripet, et al., 1997).

A sample was prepared containing 25 µM each of CPT1B TM1 and TM2 (to maintain a total peptide concentration of 50 µM) in 100 mM DPC. A CD spectrum was measured and then compared against one arithmetically produced by the addition of the individual CPT1B TM1 and TM2 spectra (results shown in **Figure 6.5**). If no interactions between the two peptides occurred then the experimental spectrum should match closely with the arithmetically produced spectrum. If interactions between the two peptides were occurring then a significant difference between the two spectra was expected.

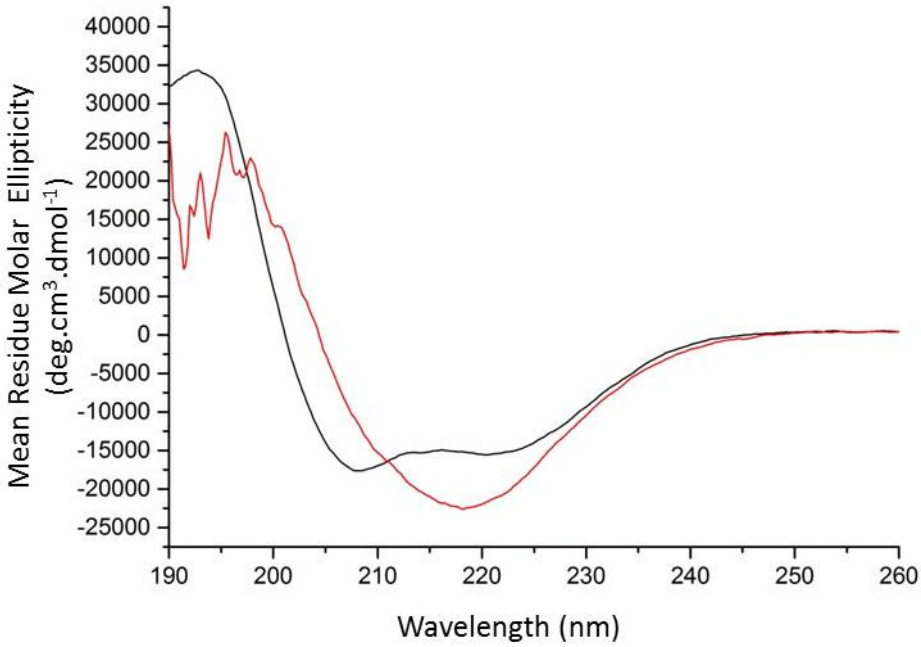


Figure 6.5 – Measuring CPT1B TM interactions using CD. Calculated arithmetic addition of the CPT1B TM1 and TM2 spectra recorded before (black) and the measured experimental result on mixing the two peptides (red).

A difference was observed between the arithmetically added CPT1B peptide spectra, and that resulting from the mixture of the two. The change in the signal however was unexpected. There was a loss in the characteristic alpha helical minimum at 208 nm entirely and a shift of the minimum at 222 nm. Despite this surprising result, it does show that there is a change upon addition of CPT1B TM2 to

TM1 indicating that the two peptides are interacting, and so this suggests that CD spectroscopy can be used to detect CPT1B TM domain interactions in DPC micelles. More experiments, including titration experiments in which one of the TM domain peptides would be slowly added to the other, were planned in order to measure this effect more closely, however a shortage of materials prevented these experiments.

6.3 NMR

Due to the difficulties in crystallising membrane proteins, NMR spectroscopy is a valuable tool in the study of their structure. Many soluble protein structures have been solved using NMR spectroscopy, however due to the hydrophobicity of membrane proteins a membrane mimetic is required to stabilise the native fold of these proteins. This membrane mimetic is often a detergent micelle used to solubilise the hydrophobic parts of the protein. This is advantageous in presenting a native membrane-like environment to the membrane protein but also significantly increases the size of the protein-micelle complex to be studied. This increase in size makes the study of even relatively small membrane proteins more difficult by NMR.

NMR spectroscopy has been used to calculate structures and measure oligomerisation in isolated TM domain sections in proteins previously (Bocharov, et al., 2012). These studies have relied on being able to assign NMR signals to specific amino acids in the protein sequence. Due to a reduction in expression yield when including both isotopic labels, the peptides expressed and purified in **Chapter 5** were only ^{15}N labelled so a three dimensional sequential assignment using HNCA and HN(CO)CA type experiments was not possible as these rely on ^{13}C labelling as well as ^{15}N (Bocharov, et al., 2010). ^{1}H homonuclear experiments like total correlation

spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY) can however be used to assign small proteins and peptides by assigning the amino acid side chains, and this was the strategy to be employed here.

In addition to ^{1}H homonuclear experiments, ^{15}N labelling allows ^{1}H-^{15}N heteronuclear experiments such as HSQC to be performed. HSQC experiments are useful in providing a fingerprint of a protein. An HSQC spectrum records all NH correlations in a protein which means a peak is observed for each amide group in the protein backbone as well as any NH groups in the amino acid side chains. As a signal is observed for each non-proline amino acid in an HSQC spectrum, this makes HSQC experiments an excellent method for detecting changes in the environment of a protein. If the HSQC can be assigned, then any perturbations in the chemical shift of these peaks can be directly mapped back to a specific amino acid. This method was used to investigate the interactions between the CPT1B TM1 and TM2 peptides; increasing amounts of CPT1B TM2 was added to a sample of CPT1B TM1 and chemical shift perturbations in the HSQC spectrum were recorded. Here experiments are presented to assess purity of the expressed peptides (**Chapter 5**) as well as to assign individual residues and study TM domain interactions in membrane mimicking detergent micelles.

6.3.1 Preparation of CPT1 Peptides for NMR Spectroscopy

All NMR samples were prepared by dissolving lyophilised peptide (~0.4 mM) in a buffer (30 mM NaP, pH 5.8, 20 mM NaCl, 100 mM deuterated DPC, and 10% D_2O). Sodium phosphate was selected as the buffering agent and produced good initial spectral quality, so no other buffers were tested. Deuterated DPC (98%

deuteration) was used to eliminate signals from the DPC, present at much higher concentration than the peptides to ensure solubilisation in micelles, and allow proton detected experiments to be performed. 10% D_2O was added to this buffer as the spectrometer requires this for the lock signal. A pH of 5.8 was used as lower pH ensures that amino acid side chains are fully protonated and can help to improve spectra quality due to slower NH exchange at lower pH (Matthew and Richards, 1983) .

As the CPT1B TM1 peptide had been successfully expressed and purified with ^{15}N isotopic labelling (**Figure 5.18**), the first experiment performed was a ^{1}H-^{15}N HSQC. HSQC experiments are useful to check the integrity and fold of a protein; a well folded protein should produce a spectrum with well dispersed and resolved peaks, whereas an unfolded or aggregated protein will have broader and less dispersed peaks. This experiment should produce a spectrum with a single peak for each NH group in the peptide; this includes one NH for each peptide bond (except that of proline) and NH groups in the side chains of tryptophan, asparagine, glutamine, histidine, lysine, and arginine. HSQC data was acquired at two temperatures (25 °C and 37 °C), and it was found that the peaks were slightly better resolved and two additional peaks were visible at 37 °C (**Figure 6.6**). The dispersion and peak shape were good at both temperatures. The additional peak observed at 25 °C in the bottom right of the spectrum could be seen at lower contour levels at 37 °C.

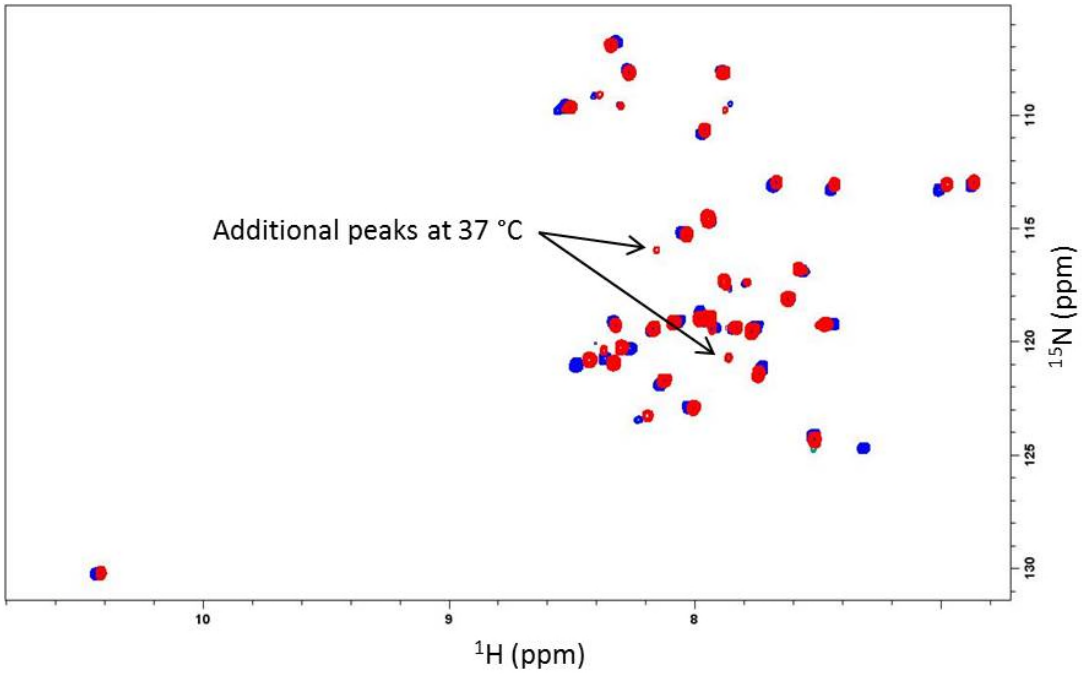


Figure 6.6 – Temperature optimisation of NMR. Two ^{1}H-^{15}N HSQC spectra of CPT1B TM1 peptide at 25 °C (blue) and 37 °C (red). The two additional peaks observed at 37 °C have been highlighted with arrows. The spectrum was recorded with a peptide concentration of 0.38 mM, and detergent and buffer concentrations of 30 mM NaP, pH 5.8, 20 mM NaCl, 100 mM deuterated DPC.

The HSQC spectrum (**Figure 6.7**) was also used to assess the purity of the sample by comparing the number of observed peaks against the theoretical number of peaks calculated from the sequence. In the case of CPT1B TM1, the theoretical number of peaks in a ^{1}H-^{15}N HSQC spectrum is 36 (29 for each backbone NH, and 7 additional peaks for 1 tryptophan (1), 2 asparagines (2), 1 arginine (2), and 2 lysines (2)). However 39 peaks were observed in the HSQC spectrum and an additional peak in the tryptophan side chain region was found at low contour levels despite only one tryptophan being present in the sequence. One possibility to explain these additional peaks is that the peptide exists in multiple states either within the micelles, or as this peptide was soluble in water, possibly a low concentration of the peptide exists partially outside of the micelles.

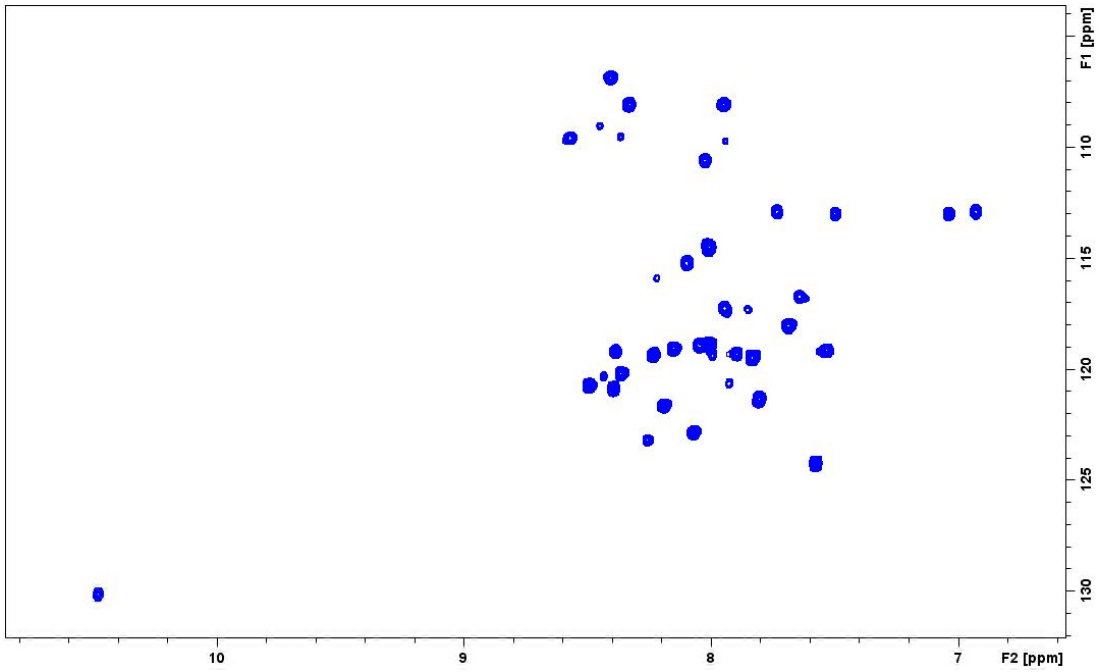


Figure 6.7 – ^{1}H-^{15}N HSQC spectrum of CPT1B TM1 peptide. 0.38 mM CPT1B TM1 peptide solubilised in 100 mM deuterated DPC, collected at 37 °C.

6.3.2 Assignment of CPT1B TM1 Spectrum

As the CPT1B TM peptides are reasonably small, and without ^{13}C labelling, the assignment strategy involved collecting TOCSY and NOESY ^{1}H homonuclear experiments, as well as ^{15}N-filtered versions of these. By ^{15}N-filtering, any peak overlaps in the two dimensional TOCSY and NOESY could be separated by the ^{15}N chemical shift as well. Unfortunately due to lack of material and extremely low signal in the two dimensional NOESY experiments tested, only TOCSY spectra could be recorded.

Due to issues with the ^{15}N-filtered experiments and 2D NOESY that could have been used for assignments, we then turned to homonuclear 2D TOCSY NMR experiments for as much assignment data as possible. For small peptides most, if not all, residues can often be assigned by residue type using 2D TOCSY experiments. In TOCSY, magnetisation is transferred via spin-spin coupling throughout a complete

spin system during a spin-lock phase. Each amino acid in the protein sequence is a separate spin system, and each residue type gives rise to a unique chemical shift pattern in a TOCSY spectrum. An example of this is shown in **Figure 6.8,** where the spin system of isoleucine has been highlighted with a red line. Isoleucine has a relatively distinctive chemical shift pattern with three peaks below 2 ppm. In some cases, either due to peak overlap, missing peaks, or similar connectivity patterns, assignments can be difficult or impossible without further experimental data.

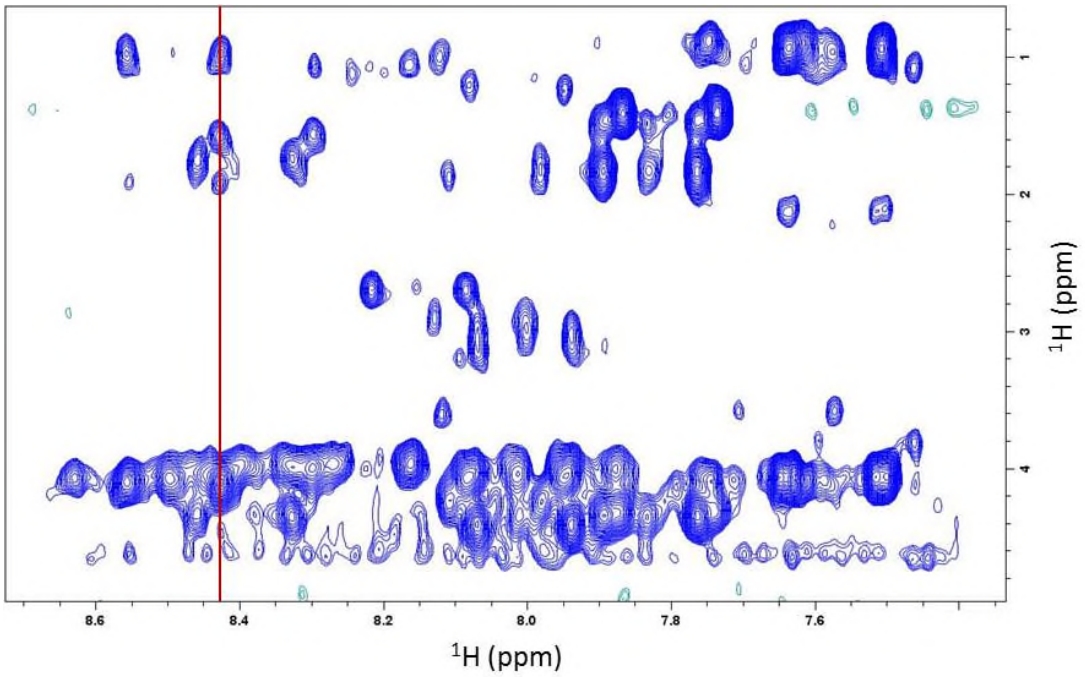


Figure 6.8 – Fingerprint region of a TOCSY spectrum of CPT1B TM1 peptide. The spin system of an isoleucine residue has been highlighted with a red line showing the αH peak above 4 ppm, and the βH, γH and γCH_3 peaks below 2 ppm. This spectrum was recorded at 37 °C.

All of the residues that were (non-sequentially) assigned are listed in **Table 6.1**, along with the chemical shifts of each of their ^{1}H peaks. Out of 31 residues in the CPT1B TM1 peptide, 22 were assigned by residue type. Where possible the assignments from the 2D TOCSY were mapped onto the HSQC spectrum using the NH amide chemical shifts, and 19 of the 22 assigned peaks were able to be mapped

in this manner (shown in **Figure 6.9**). Further experiments, such as 2D NOESY, will yield a more complete sequential assignment but could not be achieved in this project.

	Residue Type	**HN**	**Hα**	**Hβ**	**Others**
1	Val	7.506	4.086	2.085	0.965
2	Val	7.575	3.590	2.240	0.932
3	Val	7.635	4.062	2.127	0.932
4	Val	7.746	4.097	2.045	0.886
5	Leu	7.763	4.302	1.821	1.457
6	Leu	7.831	4.348	1.828	1.483
7	Leu	7.895	4.358	1.821	1.466
8	Ser	7.939	4.418	3.024	
9	Gly	8.000	2.972		
10	Ser	8.013	4.335	4.037	
11	Ser	8.069	4.418	3.069	
12	Val	8.083	4.063	2.696	1.192
13	Ala	8.120	3.599		1.000
14	Gly	8.128	2.896		
15	Ala	8.164	3.963		1.055
16	Gly	8.270	3.963		
17	Lys	8.298	4.007		1.547, 1.055
18	Thr	8.325	4.345	3.981	1.739
19	Ile	8.428	4.090	1.903	1.575, 0.973
20	Thr	8.458	4.315		1.739
21	Val	8.555	4.097	1.914	0.969
22	Gly	8.629	4.064		

Table 6.1 – Non-sequential peak assignment for the CPT1B TM1 peptide. All assignments were obtained from TOCSY data (mixing times = 80 ms) (spectrum shown in **Figure 6.8**) and are listed with residue type, chemical shifts of each ^{1}H, and sorted by ascending HN chemical shift.

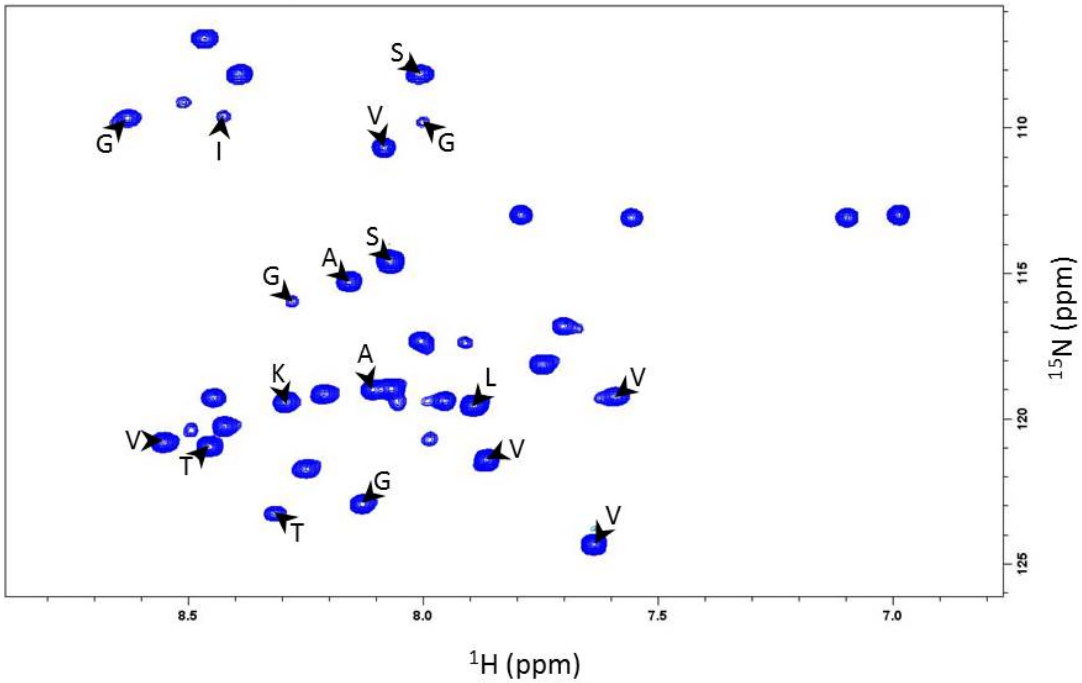


Figure 6.9 – HSQC residue assignments of CPT1B TM1. All the residue type assignments that could be mapped onto the HSQC spectrum using the NH amide chemical shifts found from the TOCSY assignment.

6.3.3 Chemical Shift Perturbations on Addition of CPT1B TM2

To investigate whether the two TM domains in CPT1 interact, a series of HSQC spectra were acquired upon stepwise titration of the unlabelled CPT1B TM2 peptide into the ^{15}N labelled CPT1B TM1 sample. An initial ^{1}H-^{15}N HSQC spectrum of CPT1B TM1 peptide (0.38 mM) in deuterated DPC (100 mM) was acquired. Then CPT1B TM2 peptide was added to this sample and allowed to equilibrate for 1 hour before further data acquisition. CPT1B TM2:TM1 concentration ratios of 0:1, 0.5:1, 1:1, 2:1, 4:1, 8:1, and 16:1 were tested in this manner (**Figure 6.10**). It was expected that if the CPT1B TM2 peptide interacted with the CPT1B TM1 peptide, then any residues involved in the interaction would experience a change of environment, and the peaks associated with them would undergo a corresponding chemical shift perturbation.

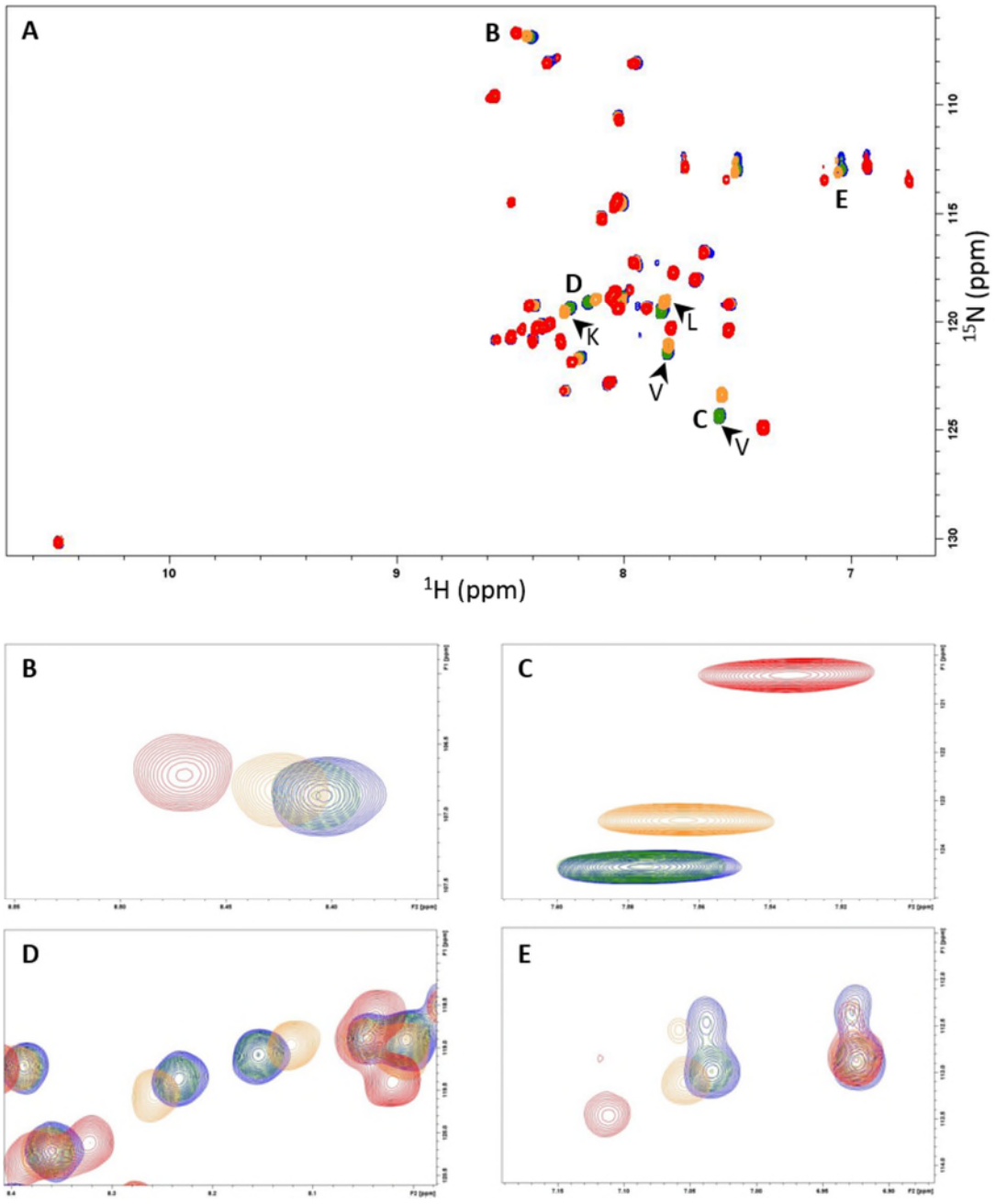


Figure 6.10 – The addition of unlabelled CPT1B TM2 to CPT1B TM1. (**A**) Four HSQC spectra overlaid showing the titration of CPT1B TM2 peptide into a sample of ^{15}N CPT1B TM1 peptide in DPC micelles. All spectra were recorded at 37 °C. TM2:TM1 concentration ratios of 0:1 (blue), 2:1 (green), 8:1 (orange), and 16:1 (red) are shown. (**B** – **E**) show the individual peak movements observed in more detail. The source of the enlargements has been marked on spectrum (**A**). Where a peak with a significant chemical shift perturbation already had an amino acid type assignment this has been labelled.

As seen in **Figure 6.10 A**, chemical shift perturbations were observed for ten of the peaks, however in all these cases, the effect was only observed at TM2:TM1 concentration ratios above 4:1. This could indicate that this is a weak interaction, and

an excess of TM2 peptide is needed in order for the interaction to be observable. Alternatively, further equilibration time may be required between the addition of the TM2 peptide and data acquisition to allow for TM2 peptide incorporation into the micelles. Unfortunately few of the peaks that underwent shifts were able to be assigned, and so without further experiments to assign more of the peptide little can be deduced about the sequence dependence of these interactions.

6.4 Purity of CPT1A TM1 NMR Samples

Some ^{15}N labelled CPT1A TM1 peptide was produced and purified following the same methods as that for CPT1B TM1. When analysed using mass spectrometry, this peptide had the correct mass (**Figure 6.11 B**), and there were no peaks characteristic of the trpLE fusion tag (**Figure 6.11 C**) which had been observed in other HPLC fractions. Despite this, when an HSQC spectrum was collected, significantly more peaks than expected were observed; especially in the regions of the spectrum normally associated with NH side chain peaks (highlighted in **Figure 6.11 A**). This suggested that either uncleaved fusion protein or trpLE had been purified with the peptide but had remained undetected (the trpLE sequence contains many more side chain NH's than the CPT1 TM peptides). Following this surprising result, SDS-PAGE was used alongside mass spectrometry to test for purity after HPLC.

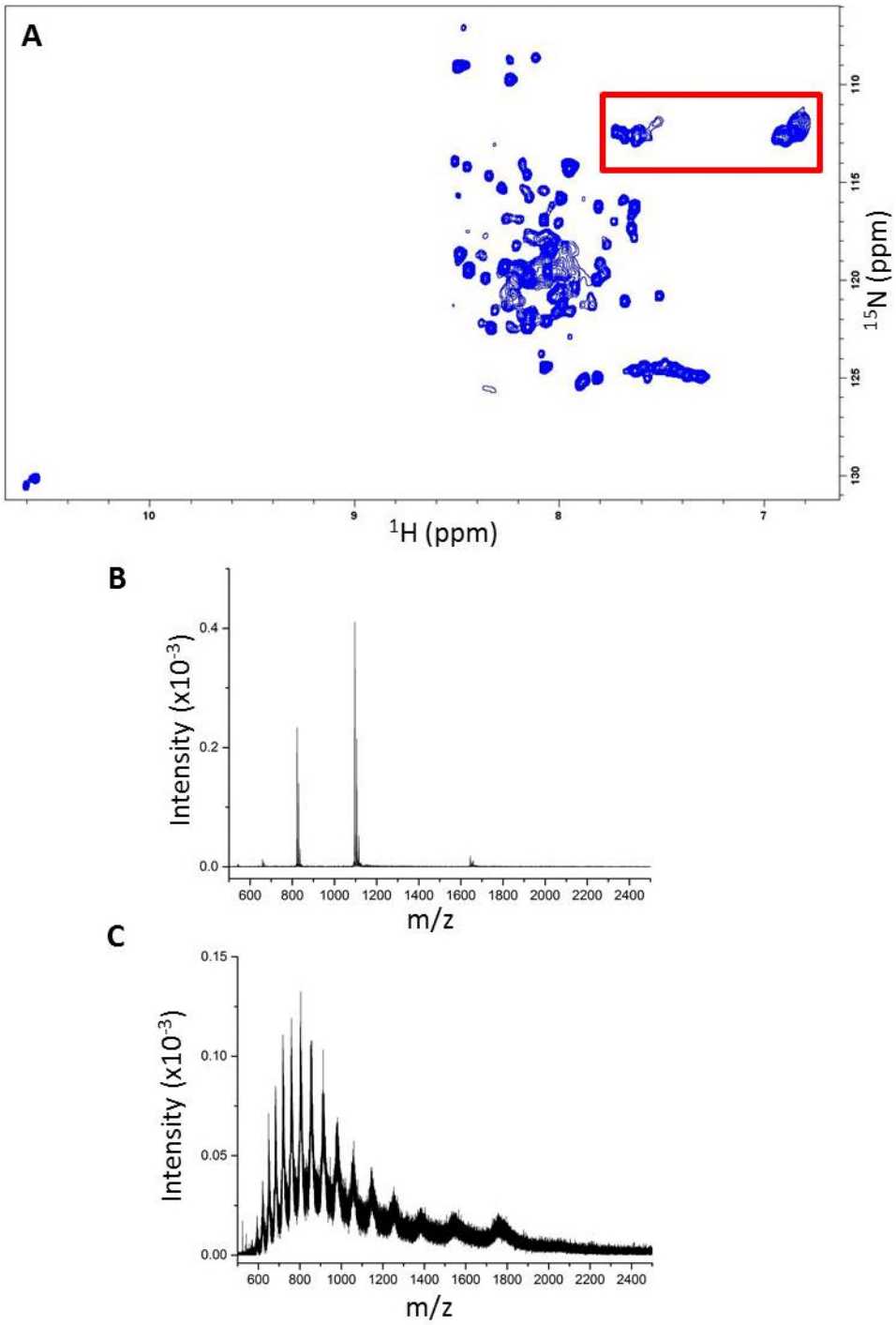


Figure 6.11 – Purity of CPT1A TM1 NMR sample. (**A**) HSQC spectrum of ^{15}N labelled CPT1A TM1 after HPLC purification showing significant amounts of trpLE fusion tag or uncleaved protein remaining in the sample. NH side chain peaks have been highlighted to show they are present in much greater numbers than expected for the CPT1A TM1 peptide alone. (**B**) mass spectrum of the CPT1A TM1 sample showing peaks at 822.7277, 1096.8890 and 1644.8310 which deconvolute to the correct mass of 3285.6490 for the CPT1A TM1 peptide. (**C**) mass spectrum showing the characteristic peaks when samples with trpLE are tested.

6.5 Summary

The results presented here have shown that the CPT1B TM domains do interact in the membrane mimicking environment of DPC micelles in both CD and NMR experiments. The CD experiments provided support for these interactions, however did not show the expected signals indicative of a helical or coiled coil structure. They did however show a significant change on introduction of CPT1B TM2 to the sample. The variable DPC concentration experiment on CPT1B TM1 demonstrated the requirement for detergent or another membrane mimetic to be present for folding of the CPT1B TM1 domain despite the solubility of the peptides in water. These experiments also allowed a larger section of the CPT1B TM regions to be investigated than in any previous experiments.

The NMR experiments demonstrated that CPT1 can be studied in this manner, providing good quality spectra. The 2D TOCSY experiment allowed for >70% assignment but unfortunately only by residue type. To assign sequentially and completely either 2D NOESY or ^{15}N-filtered experiments are needed in addition. Complete assignment would then allow for the characterisation of the specific residues affected by interactions.

7. Conclusions and Future Work

CPT1A and CPT1B are central to the functioning of mammalian cells as they have a vital role in energy homeostasis. Structural characterisation of full length CPT1 is difficult because it is associated with the outer mitochondrial membrane, however structural data is available for both soluble domains in the form of an NMR structure for the N-terminal domain (Rao, et al., 2011) and an homology model using rat acetyltransferase of the C-terminal domain (Morillas, et al., 2004). Intermolecular homotypic interactions between the TM domains of CPT1 have been implicated in oligomerisation of the enzyme (Faye, et al., 2007) (**Figure 7.1 B**). Intramolecular heterotypic interactions have been suggested to affect interactions between the N- and C-terminal domains, which are thought to form the malonyl-CoA binding site (Faye, et al., 2005) (**Figure 7.1 A**). It is known that CPT1A has variable sensitivity to inhibition by malonyl-CoA, and this sensitivity is modulated by changes in the membrane environment (McGarry and Brown, 2000). No variation in sensitivity to inhibition has been observed in CPT1B. If changes in the interactions between TM domains of CPT1A are responsible for this modulation, it is expected that significant differences should exist in these interactions between CPT1A and CPT1B.

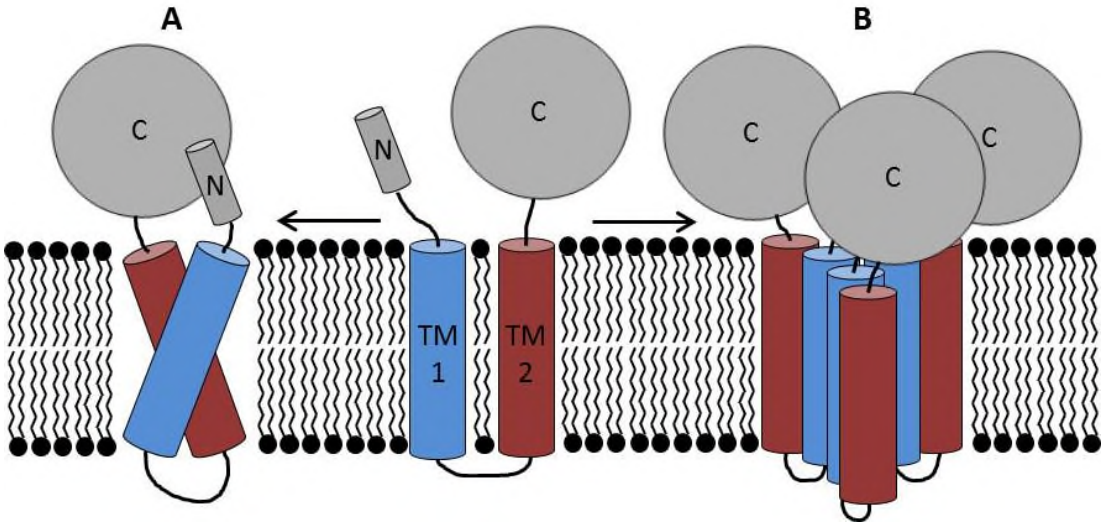


Figure 7.1 – Differences in TM interactions of CPT1. (**A**) shows intramolecular heterotypic interactions allowing the N- and C- termini to interact. (**B**) shows an example of oligomerisation.

7.1 Self-association of the CPT1A and CPT1B TM Domains

Previous studies had suggested that the TM domains of CPT1A and CPT1B are capable of self-association (Jenei, et al., 2011) but more detailed study, particularly of CPT1B, has not been pursued. The strength of TM domain interactions for CPT1A and CPT1B were studied here using the *in vivo* GALLEX assay, which confirmed that all the TM domains did indeed self-associate. It was found that TM1 of CPT1A had stronger self-association than TM2, while the opposite was true for CPT1B (i.e. TM2 had stronger self-association than TM1). This was the first difference we revealed between the two isoforms.

The sequence dependence of these interactions was then investigated using mutagenesis. The G_{103}-xxx-G_{107} and G_{113}-xxx-A_{117} motifs in CPT1A TM2 have previously been shown to reduce self-association on mutagenesis, however no detailed residue by residue data existed for any of the other TM domains. Through systematic mutagenesis it was discovered that three different residues in CPT1B

TM1, which do not form a small-xxx-small type motif, significantly reduced self-association on mutation. These residues were S_{113}, G_{115} and G_{120}, with mutation of the S_{113} and G_{120} residues reducing measured self-association to levels below that of the negative control in the GALLEX assay. This is another difference in the interactions between CPT1A and CPT1B.

In GALLEX experiments on TM1 of CPT1A and CPT1B, more similarities than differences were discovered. In both CPT1A and CPT1B TM1, a single residue was found to be the only mutation that significantly disrupted TM1 self-association. In both cases this residue was aromatic (phenylalanine in CPT1A, and tyrosine in CPT1B) and found at position 55 in the protein.

7.2 Heterotypic Interactions of the CPT1A and CPT1B TM Domains

Previous evidence suggested that heterotypic interactions occurred between TM1 and TM2 in both CPT1A and CPT1B (Jenei, 2010), however the strengths and the sequence dependence of these interactions had not been studied. Using hetero GALLEX experiments, it was observed that both CPT1A and CPT1B contain TM domains capable of heterotypic interactions. It was also discovered that the strength of these heterotypic interactions was the same for the two isoforms. However, differences in the sequence dependence were found. In CPT1A, the F_{55} residue in TM1 and the G_{113} and V_{116} residues in TM2 were found to cause significant disruption to heterotypic interactions on mutation; whereas in CPT1B, only a single glycine residue was found on each TM (G_{57} in TM1 and G_{115} in TM2) to disrupt interactions on mutation.

7.3 Overlap Between Self-Association and Heterotypic Interactions

As some residues in both CPT1A and CPT1B have been implicated in both homo- and heterotypic interactions, a novel competition version of the GALLEX assay was developed. This assay was used to gain some understanding into the preference of each isoform for either self-association or heterotypic interactions. The results from this assay indicated that TM2 of CPT1B had a stronger preference for heterotypic interactions than any of the other TM domains. In CPT1A, both TM domains showed equal disruption to self-association when heterotypic interactions were presented.

7.4 A Model for CPT1A and CPT1B TM Domain Interactions

To summarise all of these interactions, both homotypic and heterotypic, and present the differences identified between CPT1A and CPT1B, all the relevant residues were mapped onto helical wheel models (**Figures 7.2** and **7.3**).

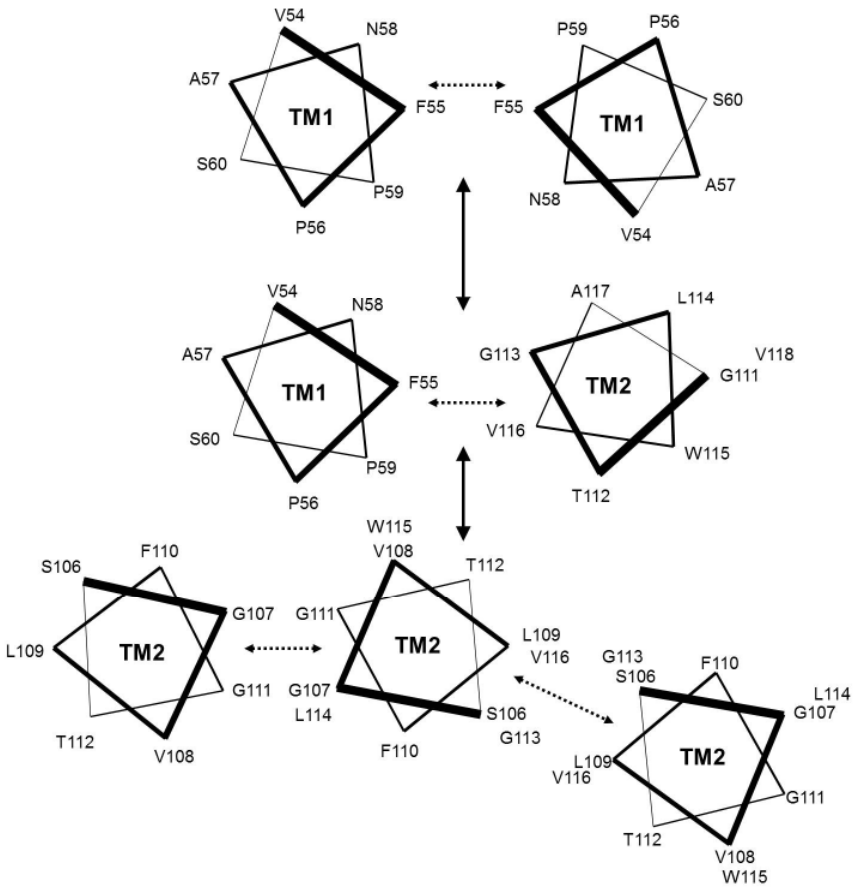


Figure 7.2 – Helical wheel schematic of all identified TM helix interactions in CPT1A. Interactions are shown with a dotted arrow and changes between homotypic and heterotypic interacting states are shown with a solid arrow.

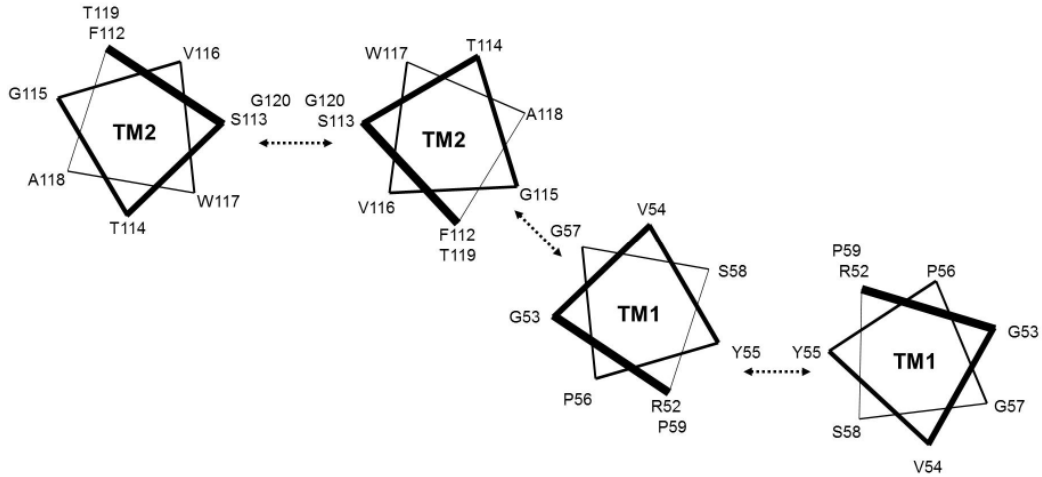


Figure 7.3 – Helical wheel schematic of all identified TM helix interactions in CPT1B. Interactions are shown with a dotted arrow.

It appears that CPT1A can adopt distinct forms in which either homo (TM1 - TM1 or TM2 - TM2) interactions or hetero (TM1 - TM2) interactions dominate, with some method of switching between these forms. This is possible because some of the

interfaces of interaction identified are capable of multiple types of interaction. On the other hand, CPT1B appears to have discrete interactions sites for homotypic vs heterotypic TM helix interactions, suggesting that these interactions could exist simultaneously. Intramolecular heterotypic interactions between TM1 and TM2 are thought to affect the formation of the malonyl-CoA binding site formed from the N- and C-termini. In CPT1B, these interactions are not disturbed by homotypic interactions between either TM1 or TM2. However in CPT1A, TM1 – TM1 homotypic interactions preclude heterotypic interactions from forming. This suggests less modulation in activity and inhibition of CPT1B compared to CPT1A.

7.5 Insights from *in vitro* Study of Interactions in CPT1

Although the information obtained from GALLEX experiments had demonstrated some clear differences between CPT1A and CPT1B, it is restricted to the study of 18 amino acid TM sections. However the full TM domains of CPT1 are predicted to be 22 amino acids in length, and so in order to study the full TM domains of CPT1 using *in vitro* biophysical techniques they were expressed as peptides, which had not been achieved before. **Chapter 5** describes the expression of these TM peptides and the eventual success in expressing and purifying the CPT1B TM domains. However, this success could not be replicated in a timely manner for the expression the TM domains of CPT1A. Due to great difficulties in the purification of all of the TM peptides, resulting in a very limited amount of pure material, only a limited amount of experimental data was collected.

Good quality NMR spectra were obtained for expressed CPT1B TM1 peptide, showing good peak dispersion and resolution, which resulted in partial

assignment of these spectra. Results from NMR and CD spectroscopy also corroborated the results obtained earlier, that the CPT1B TM domains do indeed heterotypically associate. This was important to confirm in experiments that could utilise the complete TM domains of CPT1B, and not the shortened 18 amino acid length constructs required for the GALLEX experiments.

7.6 Future Work

Further work is needed to successfully express the CPT1A TM domain peptides using the expression system described here, which would allow for comparisons with the results presented for the CPT1B TM domains. With the successful expression and purification of CPT1A and CPT1B TM domain peptides, complete assignment could be achieved using the techniques discussed. Further work is also needed to produce a reliable eukaryotic expression system for full length CPT1. This has been achieved in the past using *P. pastoris* (Brown, 2003), but attempts to replicate this did not prove successful here. Reliable expression of full length CPT1 would allow for the study of the key mutations revealed by GALLEX experiments in the context of the full length protein. Experiments such as Native-PAGE could be used to investigate full length oligomerisation, and assays to measure enzyme activity and sensitivity to inhibition (Zammit, et al., 1989) would allow the relationship between TM domain interactions and enzyme kinetics to be explored.

References

Adams, P. D., Arkin, I. T., Engelman, D. M. and Brunger, A. T. (1995) Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat Struct Biol*, **2**, 2, 154-62

Adams, P. D., Engelman, D. M. and Brunger, A. T. (1996) Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins*, **26**, 3, 257-61

Almen, M. S., Nordstrom, K. J., Fredriksson, R. and Schioth, H. B. (2009) Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol*, **7**, 50

Andreev, Y. A., Kozlov, S. A., Vassilevski, A. A. and Grishin, E. V. (2010) Cyanogen bromide cleavage of proteins in salt and buffer solutions. *Anal Biochem*, **407**, 1, 144-6

Arduini, A. and Bonomini, M. (2011) Carnitine palmitoyltransferase as a potential target for treating diabetes: is inhibition or activation preferable? *Hepatology*, **54**, 3, 1110-1

Arora, A. and Tamm, L. K. (2001) Biophysical approaches to membrane protein structure determination. *Curr Opin Struct Biol*, **11**, 5, 540-7

Bernsel, A., Viklund, H., Hennerdal, A. and Elofsson, A. (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res*, **37**, Web Server issue, W465-8

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Jr., E. E. M., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) The Protein Data Bank: A Computer-based Archival File For Macromolecular Structures. *J. of. Mol. Biol.*, **112**, 535

Bertrand, K., Squires, C. and Yanofsky, C. (1976) Transcription termination in vivo in the leader region of the tryptophan operon of Escherichia coli. *J Mol Biol*, **103**, 2, 319-37

Bocharov, E. V., Mayzel, M. L., Volynsky, P. E., Mineev, K. S., Tkach, E. N., Ermolyuk, Y. S., Schulga, A. A., Efremov, R. G. and Arseniev, A. S. (2010) Left-Handed Dimer of EphA2 Transmembrane Domain: Helix Packing Diversity among Receptor Tyrosine Kinases. *Biophysical Journal*, **98**, 5, 881-889

Bocharov, E. V., Mineev, K. S., Goncharuk, M. V. and Arseniev, A. S. (2012) Structural and thermodynamic insight into the process of “weak” dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1818**, 9, 2158-2170

Borthwick, K., Jackson, V. N., Price, N. T. and Zammit, V. A. (2006) The mitochondrial intermembrane loop region of rat carnitine palmitoyltransferase 1A is a major determinant of its malonyl-CoA sensitivity. *J Biol Chem*, **281**, 44, 32946-52

Brindle, N. P., Zammit, V. A. and Pogson, C. I. (1985) Regulation of carnitine palmitoyltransferase activity by malonyl-CoA in mitochondria from sheep liver, a tissue with a low capacity for fatty acid synthesis. *Biochem J*, **232**, 1, 177-82

Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protocols*, **2**, 11, 2728-2733

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54**, Pt 5, 905-21

Call, M. E., Schnell, J. R., Xu, C., Lutz, R. A., Chou, J. J. and Wucherpfennig, K. W. (2006) The structure of the zetazeta transmembrane dimer reveals features essential for its assembly with the T cell receptor. *Cell*, **127**, 2, 355-68

Chong, Y. H., Ball, J. M., Issel, C. J., Montelaro, R. C. and Rushlow, K. E. (1991) Analysis of equine humoral immune responses to the transmembrane envelope glycoprotein (gp45) of equine infectious anemia virus. *J Virol*, **65**, 2, 1013-8

Claridge, J. K. and Schnell, J. R. (2012) Bacterial production and solution NMR studies of a viral membrane ion channel. *Methods Mol Biol*, **831**, 165-79

Compton, L. A. and Johnson, W. C., Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem*, **155**, 1, 155-67

Cymer, F., Sanders, C. R. and Schneider, D. (2013) Analyzing oligomerization of individual transmembrane helices and of entire membrane proteins in E. coli: A hitchhiker's guide to GALLEX. *Methods Mol Biol*, **932**, 259-76

Dai, Y., Wolfgang, M. J., Cha, S. H. and Lane, M. D. (2007) Localization and effect of ectopic expression of CPT1c in CNS feeding centers. *Biochem Biophys Res Commun*, **359**, 3, 469-74

DeGrado, W. F., Gratkowski, H. and Lear, J. D. (2003) How do helix–helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Sci*, **12**, 4, 647-65

De Vivo, D. C., Bohan, T. P., Coulter, D. L., Dreifuss, F. E., Greenwood, R. S., Nordli, D. R., Jr., Shields, W. D., Stafstrom, C. E. and Tein, I. (1998) L-carnitine supplementation in childhood epilepsy: current perspectives. *Epilepsia*, **39**, 11, 1216-25

Diefenderfer, C., Lee, J., Mlyanarski, S., Guo, Y. and Glover, K. J. (2009) Reliable expression and purification of highly insoluble transmembrane domains. *Anal Biochem*, **384**, 2, 274-8

Dmitrova, M., Younes-Cauet, G., Oertel-Buchheit, P., Porte, D., Schnarr, M. and Granger-Schnarr, M. (1998) A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli. *Mol Gen Genet*, **257**, 2, 205-12

Dobbins, R. L., Szczepaniak, L. S., Bentley, B., Esser, V., Myhill, J. and McGarry, J. D. (2001) Prolonged inhibition of muscle carnitine palmitoyltransferase-1 promotes intramyocellular lipid accumulation and insulin resistance in rats. *Diabetes*, **50**, 1, 123-30

Dougherty, D. A. (1996) Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science*, **271**, 5246, 163-8

Doura, A. K. and Fleming, K. G. (2004) Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer. *J Mol Biol*, **343**, 5, 1487-97

Dunitz, J. D. (2001) Pauling's Left-Handed α-Helix. *Angewandte Chemie International Edition*, **40**, 22, 4167-4173

Faye, A., Borthwick, K., Esnous, C., Price, N. T., Gobin, S., Jackson, V. N., Zammit, V. A., Girard, J. and Prip-Buus, C. (2005) Demonstration of N- and C-terminal domain intramolecular interactions in rat liver carnitine palmitoyltransferase 1 that determine its degree of malonyl-CoA sensitivity. *Biochem J*, **387**, Pt 1, 67-76

Faye, A., Esnous, C., Price, N. T., Onfray, M. A., Girard, J. and Prip-Buus, C. (2007) Rat liver carnitine palmitoyltransferase 1 forms an oligomeric complex within the outer mitochondrial membrane. *J Biol Chem*, **282**, 37, 26908-16

Filmore, D. (2004) It's a GPCR world: Cell-based screening assays and structural studies are fuelling G-protein coupled receptors as one of the most popular classes of investigational drug targets. *Mod. Drug Discovery*, **7**, 11, 24-28

Fraser, F., Corstorphine, C. G. and Zammit, V. A. (1997) Topology of carnitine palmitoyltransferase I in the mitochondrial outer membrane. *Biochem J*, **323 (Pt 3)**, 711-8

Fraser, F. and Zammit, V. A. (1998) Enrichment of carnitine palmitoyltransferases I and II in the contact sites of rat liver mitochondria. *Biochem J*, **329**, Pt 2, 225-9

Fuerst, J. A. (2005) Intracellular compartmentation in planctomycetes. *Annu Rev Microbiol*, **59**, 299-328

Grantham, B. D. and Zammit, V. A. (1988) Role of carnitine palmitoyltransferase I in the regulation of hepatic ketogenesis during the onset and reversal of chronic diabetes. *Biochem J*, **249**, 2, 409-14

Gupta A, C. D., Mahalakshmi R (2012) Modified CNBr cleavage protocol for efficient separation of Met-Ser containing OmpX-Om14 membrane protein fusion. *International Review of Biophysical Chemistry*, **3**, 147-156
Hildebrand, P. W., Preissner, R. and Frommel, C. (2004) Structural features of transmembrane helices. *FEBS Lett*, **559**, 1-3, 145-51

Hong, H. and Bowie, J. U. (2011) Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *J Am Chem Soc*, **133**, 29, 11389-98

Jackson, V. N., Cameron, J. M., Fraser, F., Zammit, V. A. and Price, N. T. (2000) Use of six chimeric proteins to investigate the role of intramolecular interactions in determining the kinetics of carnitine palmitoyltransferase I isoforms. *J Biol Chem*, **275**, 26, 19560-6

Jackson, V. N., Price, N. T. and Zammit, V. A. (2001) Specificity of the interactions between Glu-3, Ser-24, and Gln-30 within the N-terminal segment of rat liver mitochondrial overt carnitine palmitoyltransferase (L-CPT I) in determining the malonyl-CoA sensitivity of the enzyme. *Biochemistry*, **40**, 48, 14629-34

Jackson, V. N., Zammit, V. A. and Price, N. T. (2000) Identification of positive and negative determinants of malonyl-CoA sensitivity and carnitine affinity within the amino termini of rat liver- and muscle-type carnitine palmitoyltransferase I. *J Biol Chem*, **275**, 49, 38410-6

Jenei, Z. (2010) PhD Thesis: Transmembrane domain interactions in carnitine palmitoyltransferase 1

Jenei, Z. A., Borthwick, K., Zammit, V. A. and Dixon, A. M. (2009) Self-association of transmembrane domain 2 (TM2), but not TM1, in carnitine palmitoyltransferase 1A: role of GXXXG(A) motifs. *J Biol Chem*, **284**, 11, 6988-97

Jenei, Z. A., Warren, G. Z., Hasan, M., Zammit, V. A. and Dixon, A. M. (2011) Packing of transmembrane domain 2 of carnitine palmitoyltransferase-1A affects oligomerization and malonyl-CoA sensitivity of the mitochondrial outer membrane protein. *Faseb j*, **25**, 12, 4522-30

Johnson, R. M., Hecht, K. and Deber, C. M. (2007) Aromatic and cation-pi interactions enhance helix-helix association in a membrane environment. *Biochemistry*, **46**, 32, 9208-14

Kaiser, R. and Metzka, L. (1999) Enhancement of Cyanogen Bromide Cleavage Yields for Methionyl-Serine and Methionyl-Threonine Peptide Bonds. *Analytical Biochemistry*, **266**, 1, 1-8

Kerner, J. and Bieber, L. (1990) Isolation of a malonyl-CoA-sensitive CPT/beta-oxidation enzyme complex from heart mitochondria. *Biochemistry*, **29**, 18, 4326-34

Kim, J. Y., Koves, T. R., Yu, G. S., Gulick, T., Cortright, R. N., Dohm, G. L. and Muoio, D. M. (2002) Evidence of a malonyl-CoA-insensitive carnitine palmitoyltransferase I activity in red skeletal muscle. *Am J Physiol Endocrinol Metab*, **282**, 5, E1014-22

Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H. and Bachrach, H. L. (1981) Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine. *Science*, **214**, 4525, 1125-9

Kolodziej, M. P. and Zammit, V. A. (1993) Mature carnitine palmitoyltransferase I retains the N-terminus of the nascent protein in rat liver. *FEBS Letters*, **327**, 3, 294-296

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, **305**, 3, 567-80

Langosch, D., Brosig, B., Kolmar, H. and Fritz, H.-J. (1996) Dimerisation of the Glycophorin A Transmembrane Segment in Membranes Probed with the ToxR Transcription Activator. *Journal of Molecular Biology*, **263**, 4, 525-530

Langosch, D. and Heringa, J. (1998) Interaction of transmembrane helices by a knobs-into-holes packing characteristic of soluble coiled coils. *Proteins*, **31**, 2, 150-9

Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics*, **22**, 16, 1955-62

Lewandowski, E. D., Fischer, S. K., Fasano, M., Banke, N. H., Walker, L. A., Huqi, A., Wang, X., Lopaschuk, G. D. and O'Donnell, J. M. (2013) Acute liver carnitine palmitoyltransferase I overexpression recapitulates reduced palmitate oxidation of cardiac hypertrophy. *Circ Res*, **112**, 1, 57-65

Lindner, E., Unterreitmeier, S., Ridder, A. N. and Langosch, D. (2007) An extended ToxR POSSYCCAT system for positive and negative selection of self-interacting transmembrane domains. *J Microbiol Methods*, **69**, 2, 298-305

Lock, A., Forfar, R., Weston, C., Bowsher, L., Upton, G. J., Reynolds, C. A., Ladds, G. and Dixon, A. M. (2014) One motif to bind them: A small-XXX-small motif affects transmembrane domain 1 oligomerization, function, localization, and cross-talk between two yeast GPCRs. *Biochim Biophys Acta*, **1838**, 12, 3036-51

Lopez-Vinas, E., Bentebibel, A., Gurunathan, C., Morillas, M., de Arriaga, D., Serra, D., Asins, G., Hegardt, F. G. and Gomez-Puertas, P. (2007) Definition by functional and structural analysis of two malonyl-CoA sites in carnitine palmitoyltransferase 1A. *J Biol Chem*, **282**, 25, 18212-24

Ma, C., Marassi, F. M., Jones, D. H., Straus, S. K., Bour, S., Strebel, K., Schubert, U., Oblatt-Montal, M., Montal, M. and Opella, S. J. (2002) Expression, purification, and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Sci*, **11**, 3, 546-57
MacKenzie, K. R., Prestegard, J. H. and Engelman, D. M. (1997) A transmembrane helix dimer: structure and implications. *Science*, **276**, 5309, 131-133

Manavalan, P. and Johnson, W. C., Jr. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem*, **167**, 1, 76-85

Matthew, J. B. and Richards, F. M. (1983) The pH Dependence oh Hydrogen Exchange in Proteins. *The Journal of Biological Chemistry*, **258**, 3039-3044,
McCormick, K., Mick, G. J., Mattson, V., Saile, D. and Starr, D. (1988) Carnitine palmitoyltransferase: effects of diabetes, fasting, and pH on the reaction that generates acyl CoA. *Metabolism*, **37**, 11, 1073-7

McCormick, K., Mick, G. J., Mattson, V., Saile, D. and Starr, D. (1988) Carnitine palmitoyltransferase: effects of diabetes, fasting, and pH on the reaction that generates acyl CoA. *Metabolism*, **37**, 11, 1073-7

McGarry, J. D. and Brown, N. F. (2000) Reconstitution of purified, active and malonyl-CoA-sensitive rat liver carnitine palmitoyltransferase I: relationship between membrane environment and malonyl-CoA sensitivity. *Biochem J*, **349**, Pt 1, 179-87

McGarry, J. D., Mills, S. E., Long, C. S. and Foster, D. W. (1983) Observations on the affinity for carnitine, and malonyl-CoA sensitivity, of carnitine palmitoyltransferase I in animal and human tissues. Demonstration of the presence of malonyl-CoA in non-hepatic tissues of the rat. *Biochem J*, **214**, 1, 21-8

Miozzari, G. F. and Yanofsky, C. (1978) Translation of the leader region of the Escherichia coli tryptophan operon. *J Bacteriol*, **133**, 3, 1457-66

Moller, S., Croning, M. D. and Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*, **17**, 7, 646-53

Morillas, M., Gomez-Puertas, P., Roca, R., Serra, D., Asins, G., Valencia, A. and Hegardt, F. G. (2001) Structural model of the catalytic core of carnitine palmitoyltransferase I and carnitine octanoyltransferase (COT): mutation of CPT I histidine 473 and alanine 381 and COT alanine 238 impairs the catalytic activity. *J Biol Chem*, **276**, 48, 45001-8

Morillas, M., Lopez-Vinas, E., Valencia, A., Serra, D., Gomez-Puertas, P., Hegardt, F. G. and Asins, G. (2004) Structural model of carnitine palmitoyltransferase I based on the carnitine acetyltransferase crystal. *Biochem J*, **379**, Pt 3, 777-84

Nash, A., Notman, R. and Dixon, A. M. (2015) De novo design of transmembrane helix-helix interactions and measurement of stability in a biological membrane. *Biochim Biophys Acta*, **1848**, 5, 1248-57

Overington, J. P., Al-Lazikani, B. and Hopkins, A. L. (2006) How many drug targets are there? *Nat Rev Drug Discov*, **5**, 12, 993-6

Park, E. A. and Cook, G. A. (1998) Differential regulation in the heart of mitochondrial carnitine palmitoyltransferase-I muscle and liver isoforms. *Mol Cell Biochem*, **180**, 1-2, 27-32

Price, N., van der Leij, F., Jackson, V., Corstorphine, C., Thomson, R., Sorensen, A. and Zammit, V. (2002) A novel brain-expressed protein related to carnitine palmitoyltransferase I. *Genomics*, **80**, 4, 433-42

Raman, P., Cherezov, V. and Caffrey, M. (2006) The Membrane Protein Data Bank. *Cell Mol Life Sci*, **63**, 1, 36-51

Ramsay, R. R. and Zammit, V. A. (2004) Carnitine acyltransferases and their influence on CoA pools in health and disease. *Mol Aspects Med*, **25**, 5-6, 475-93

Rao, J. N., Warren, G. Z., Estolt-Povedano, S., Zammit, V. A. and Ulmer, T. S. (2011) An environment-dependent structural switch underlies the regulation of carnitine palmitoyltransferase 1A. *J Biol Chem*, **286**, 49, 42545-54

Rasmussen, B. B., Holmback, U. C., Volpi, E., Morio-Liondore, B., Paddon-Jones, D. and Wolfe, R. R. (2002) Malonyl coenzyme A and the regulation of functional carnitine palmitoyltransferase-1 activity and fat oxidation in human skeletal muscle. *J Clin Invest*, **110**, 11, 1687-93

Ridder, A., Skupjen, P., Unterreitmeier, S. and Langosch, D. (2005) Tryptophan supports interaction of transmembrane helices. *J Mol Biol*, **354**, 4, 894-902

Russ, W. P. and Engelman, D. M. (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci U S A*, **96**, 3, 863-8

Russ, W. P. and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol*, **296**, 3, 911-9

Saggerson, D., Ghadiminejad, I. and Awan, M. (1992) Regulation of mitochondrial carnitine palmitoyl transferases from liver and extrahepatic tissues. *Adv Enzyme Regul*, **32**, 285-306

Schlegel, S., Lofblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D. J. and de Gier, J. W. (2012) Optimizing membrane protein overexpression in the Escherichia coli strain Lemo21(DE3). *J Mol Biol*, **423**, 4, 648-59

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Meth*, **9**, 7, 671-675

Schneider, D. and Engelman, D. M. (2003) GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. *J Biol Chem*, **278**, 5, 3105-11

Senes, A., Engel, D. E. and DeGrado, W. F. (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol*, **14**, 4, 465-79

Shi, J., Zhu, H., Arvidson, D. N. and Woldegiorgis, G. (1999) A single amino acid change (substitution of glutamate 3 with alanine) in the N-terminal region of rat liver carnitine palmitoyltransferase I abolishes malonyl-CoA inhibition and high affinity binding. *J Biol Chem*, **274**, 14, 9421-6

Shi, J., Zhu, H., Arvidson, D. N. and Woldegiorgis, G. (2000) The first 28 N-terminal amino acid residues of human heart muscle carnitine palmitoyltransferase I are essential for malonyl CoA sensitivity and high-affinity binding. *Biochemistry*, **39**, 4, 712-7

Sierra, A. Y., Gratacos, E., Carrasco, P., Clotet, J., Urena, J., Serra, D., Asins, G., Hegardt, F. G. and Casals, N. (2008) CPT1c is localized in endoplasmic reticulum of neurons and has carnitine palmitoyltransferase activity. *J Biol Chem*, **283**, 11, 6878-85

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Sci*, **18**, 5, 936-48

Smith, J. G., Mothes, W., Blacklow, S. C. and Cunningham, J. M. (2004) The mature avian leukosis virus subgroup A envelope glycoprotein is metastable, and refolding induced by the synergistic effects of receptor binding and low pH is coupled to infection. *J Virol*, **78**, 3, 1403-10

Sonnhammer, E. L., von Heijne, G. and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol*, **6**, 175-82

Sreerama, N. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem*, **287**, 2, 252-60

Staley, J. P. and Kim, P. S. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci*, **3**, 10, 1822-32

Swanson, S. T., Foster, D. W., McGarry, J. D. and Brown, N. F. (1998) Roles of the N- and C-terminal domains of carnitine palmitoyltransferase I isoforms in malonyl-CoA sensitivity of the enzymes: insights from expression of chimaeric proteins and mutation of conserved histidine residues. *Biochem J*, **335 (Pt 3)**, 513-9

Tripet, B., Vale, R. D. and Hodges, R. S. (1997) Demonstration of coiled-coil interactions within the kinesin neck region using synthetic peptides. Implications for motor activity. *J Biol Chem*, **272**, 14, 8946-56

Tsirigos, K. D., Peters, C., Shu, N., Käll, L. and Elofsson, A. (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Research*, **43**, 401-407

Unterreitmeier, S., Fuchs, A., Schaffler, T., Heym, R. G., Frishman, D. and Langosch, D. (2007) Phenylalanine promotes interaction of transmembrane domains via GxxxG motifs. *J Mol Biol*, **374**, 3, 705-18

van der Leij, F. R., Kram, A. M., Bartelds, B., Roelofsen, H., Smid, G. B., Takens, J., Zammit, V. A. and Kuipers, J. R. (1999) Cytological evidence that the C-terminus of carnitine palmitoyltransferase I is on the cytosolic face of the mitochondrial outer membrane. *Biochem J*, **341 (Pt 3)**, 777-84

Wei, P., Zheng, B. K., Guo, P. R., Kawakami, T. and Luo, S. Z. (2013) The association of polar residues in the DAP12 homodimer: TOXCAT and molecular dynamics simulation studies. *Biophys J*, **104**, 7, 1435-44

Wolfgang, M. J., Kurama, T., Dai, Y., Suwa, A., Asaumi, M., Matsumoto, S.-i., Cha, S. H., Shimokawa, T. and Lane, M. D. (2006) The brain-specific carnitine palmitoyltransferase-1c regulates energy homeostasis. *Proceedings of the National Academy of Sciences*, **103**, 19, 7282-7287

Wuthrich, K. (1990) Protein structure determination in solution by NMR spectroscopy. *J Biol Chem*, **265**, 36, 22059-62

Yun, R. H., Anderson, A. and Hermans, J. (1991) Proline in alpha-helix: stability and conformation studied by dynamics simulation. *Proteins*, **10**, 3, 219-28

Zammit, V., A, Corstorphine, C., G, Kolodziej, M., P and Fraser, F. (1998) Lipid molecular order in liver mitochondrial outer membranes, and sensitivity of carnitine palmitoyltransferase I to malonyl-CoA. *Lipids*, **33**, 4, 371-376

Zammit, V. A. (1999) Carnitine acyltransferases: functional significance of subcellular distribution and membrane topology. *Progress in Lipid Research*, **38**, 3, 199-224

Zammit, V. A. (2008) Carnitine palmitoyltransferase 1: Central to cell function. *IUBMB Life*, **60**, 5, 347-354

Zammit, V. A., Corstorphine, C. G. and Kolodziej, M. P. (1989) Target size analysis by radiation inactivation of carnitine palmitoyltransferase activity and malonyl-CoA binding in outer membranes from rat liver mitochondria. *Biochem J*, **263**, 1, 89-95

Zammit, V. A., Fraser, F. and Orstorphine, C. G. (1997) Regulation of mitochondrial outer-membrane carnitine palmitoyltransferase (CPT I): role of membrane-topology. *Adv Enzyme Regul*, **37**, 295-317

Zammit, V. A., Ramsay, R. R., Bonomini, M. and Arduini, A. (2009) Carnitine, mitochondrial function and therapy. *Adv Drug Deliv Rev*, **61**, 14, 1353-62

Zheng, H., Zhao, J., Wang, S., Lin, C. M., Chen, T., Jones, D. H., Ma, C., Opella, S. and Xie, X. Q. (2005) Biosynthesis and purification of a hydrophobic peptide from transmembrane domains of G-protein-coupled CB2 receptor. *J Pept Res*, **65**, 4, 450-8

Zhou, F. X., Merianos, H. J., Brunger, A. T. and Engelman, D. M. (2001) Polar residues drive association of polyleucine. **98**, 5, 2250-5