Investigating helix-helix interactions in the transmembrane domains of membrane proteins

Gavin W.King MSc BSc

A thesis submitted for the degree of Doctor of Philosophy

University of Warwick
Department of Chemistry
December 2009

Table of contents

List of figures

List of tables

Acknowledgements

Firstly, I would like to thank my supervisor Ann Dixon for giving me the opportunity to carry out the work presented in this thesis and for her guidance throughout my research and in the preparation of this manuscript.

I would also like to thank Joanne Oates for all her help initiating me into ways of the laboratory and the mysterious art of “cloning”, and for her collaborating on the NMR study of E5, Zsuzsanna Jenei for discussions regarding the GALLEX assay, Andrew Beevers and David Scott (University of Nottingham) for performing the AUC analyses, Oksana Leszczyszyn for help with NMR spectroscopy, Matthew Hicks for assistance with performing the FRET experiments, James Sturgis for guidance in analysing the data from FRET experiments, Dirk Schneider for providing plasmids and guidance on performing the GALLEX assay, Donald Engelman for providing plasmids for performing the TOXCAT assay, my academic panel Alison Rodger and Gegory Challis for their suggestions, and Janet Crawford for synthesising the peptides used in this study.

Additionally, I owe numerous debts of gratitude to all the members of the Dixon and Rodger research groups for their helpful contributions and friendship.

Finally, and most importantly, my appreciation and thanks go to my daughters Lizzie and Abi for the joy they bring me everyday, and to my wife Jo whose hard work and tireless support enabled me to carry out this work.

Declaration

The work in this thesis is original, and was conducted by the author, unless otherwise stated, under the supervision of Dr Ann M. Dixon. It has not previously been presented for another degree.

Funding was provided by an EPSRC studentship.

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

Collaborative work has been carried out contributing towards the material in this thesis. The extent of the collaborative work is indicated in the relevant sections. In summary collaborative work was carried out with the following:

Dr Joanne Oates, Department of Chemistry, University of Warwick, performed work that contributed to Chapter 6.

Dr Andrew Beevers, Department of Chemistry, University of Warwick, performed work that contributed to Chapter 3.

Abbreviations

A

A	absorbance
ACN	acetonitrile
AFM	ammonium persulfate
AU	absorbance units

B

bp	base pair
BS^3	bis[sulfosuccinimidyl]-suberate

C

C	concentration
CAT	chloramphenicol acetyltransferase
CD	circular dichroism
CHI	CNS searching of helix interactions
CNBr	cyanogen bromide
COSY	correlated spectroscopy
CNS	crystallography and NMR system

D

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
dH_2O	deionized water
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DNA	deoxyribonucleic acid
DPC	dodecylphosphocholine

E

EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionisation mass spectrometry

F

FRET	Förster resonance energy transfer
FA	formic acid
FT	Fourier transform

G

g	grams
GpA	glycophorin A
GST	glutathione s-transferase
x g	times gravity

H

hr	hour
HFIP	hexafluoroisopropanol
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence

I

Ii	invariant chain protein
IPTG	isopropyl-β-D-thiogalactopyranoside

K

kb	kilo bases or kilo basepairs
kDa	kilo Dalton
KSI	ketosteroid isomerase

L

LB	Luria-Bertani
LPPG	1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)]

M

MALDI-TOF	matrix-assisted laser desorption ionisation time of flight
MHC	class II majorhistocompatibility complex
MHCα	alpha chain of class II majorhistocompatibility complex protein
MHCβ	beta chain of class II majorhistocompatibility complex protein
mL	millilitre
MW	molecular weight

N

NMR	nuclear magnetic resonance spectroscopy
NOE	nuclear Overhauser enhancement
NOESY	nuclear Overhauser enhancement spectroscopy

O

OCD	oriented circular dichroism
OD	optical density
OG	octyl-glucoside

P

PCR	polymerase chain reaction
pH	p of hydrogen in solution

R

RMSD	root mean squared deviation
rpm	revolutions per minute

S

SDS	sodium dodecylsulphate

T

TCA	trichloroacetic acid
TFA	trifluoroacetic acid
TFE	triflouroethanol
TMD	transmembrane domain
TOCSY	total correlation spectroscopy

U

UV	ultraviolet light

W

w/v	weight per volume

Greek Symbols

μg	microgram
μL	microlitre
μM	micromolar

Summary

Helix-helix interactions between membrane-spanning transmembrane (TM) domains have been shown to drive the assembly of α-helical membrane proteins within biological membranes. However, the rules that determine these interactions are not yet fully understood, despite such interactions being found in an increasing number of proteins. Recent work has implicated TM domain interactions in the formation of the protein complex Ii-MHC, formed from the association of Major Histocompatibility Complex Class II (MHC) and the MHC-associated-Invariant Chain (Ii) proteins. Following biosynthesis, three MHC α/β heterodimers bind to the Ii homotrimer to form a nonameric Ii-MHC complex within the endoplasmic reticulum. This is a critical step in the export of MHC molecules to the antigen presentation system and hence the activation of an immune response to a pathogen. In this study we have explored the TM domain interactions within the Ii-MHC complex. Results from *in vivo* and *in vitro* experiments revealed the TM domains of the α- and β-chains of MHC have a propensity to self-associate into homo-dimers and to associate with one another to form hetero-dimers. Highly conserved GxxxG motifs (known to drive dimerization) were implicated in these interactions. The TM domain of Ii was confirmed to self-associate to form trimers by *in vivo* and *in vitro* methods, but surprisingly also displayed additional oligomeric states suggesting the interaction is not as specific as was previously thought. Furthermore, we show that *in vivo*, the TM domain of Ii can associate with those of the α- and β-chains of MHC, whilst *in vitro* methods suggested Ii preferentially binds to α-chains. Collectively, these findings strongly suggest that the TM domains of Ii and MHC have a role to play in the assembly of the Ii-MHC complex, and hence the very important process of antigen presentation. Additionally, in this study we have undertaken development of NMR spectroscopy methods that have the potential to increase our understanding of not only the Ii-MHC complex, but protein-protein interactions in general.

1 Introduction

1.1 Membrane Proteins

With around 30% of sequenced genomes encoding for membrane-associated proteins and around two thirds of all drugs targeting these proteins, the importance of membrane proteins is now well established (MacKenzie, 2006; Rath, Johnson et al., 2007). These proteins are involved in a myriad of functions critical to survival of individual cells and multicellular organisms, including the transport of small molecules and ions, and receptors for signalling molecules. Furthermore, the malfunction of membrane proteins has been implicated in several diseases, such as autoimmunity, diabetes and cancer (Sanders and Myers, 2004).

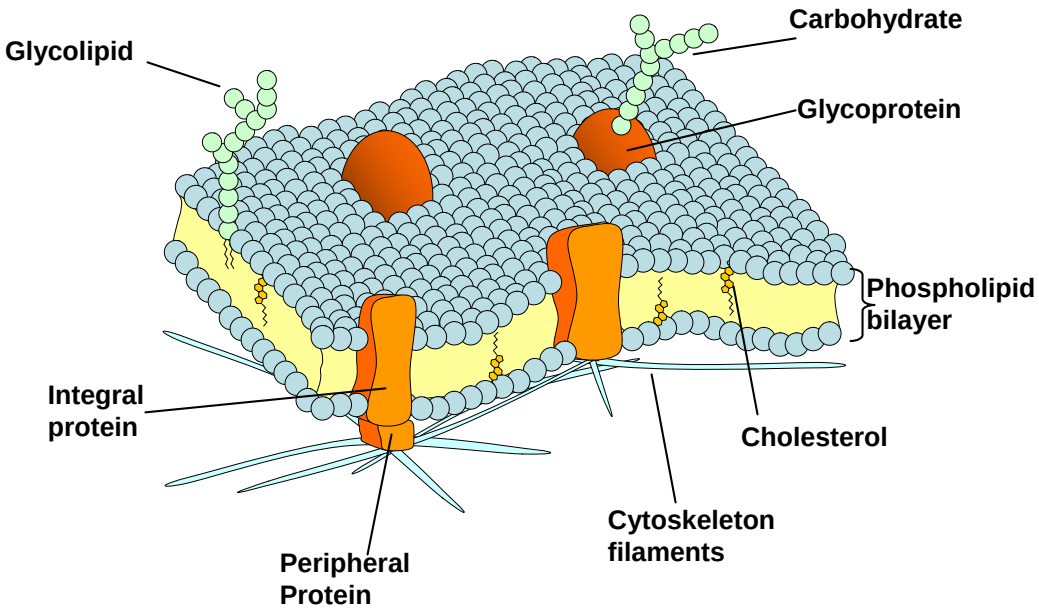


Figure 1.1. The fluid mosaic model of biological membranes

A cross-sectional schematic of the fluid mosaic model of biological membranes, which represents our current understanding of these highly heterogeneous mixtures of lipids of various types and proteins. The external face of membrane is often coated with carbohydrate molecules that are either embedded in the bilayer or are associated with proteins known as glycoproteins. Proteins can span the membrane (integral) or associate with the outside of the membrane (peripheral).

Cellular membranes define the boundary of individual cells and are important features of the internal structure of cells, where they define organelles and serve to compartmentalise cellular functions. The fluid mosaic model first proposed by Singer and Nicholson and represented in cartoon form in Figure 1.1, summarises our current understanding of the organisation of biological membranes (Singer and Nicolson, 1972). Membranes are highly complex heterogeneous environments mainly composed of a bilayer of amphiphilic phospholipids into which are embedded other components such as glycolipids, cholesterol and proteins (Engelman, 2005). Proteins that associate with the membrane can be embedded in the membrane via bilayer spanning regions (i.e. integral membrane proteins), bound reversibly or irreversibly to the surface of the membrane (peripheral membrane proteins), or be tethered to the membrane by a lipid anchor.

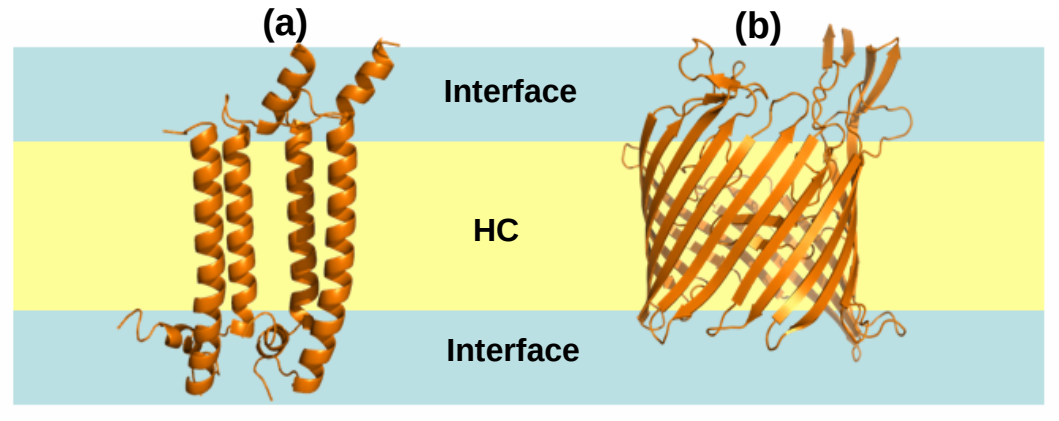


Figure 1.2. Observed folds of integral membrane proteins

To date the only two folds that have been observed for the membrane spanning domains of integral membrane proteins are (a) α-helical and (b) β-barrel. These proteins insert into and span heterogeneous cell membranes composed of amphiphilic lipids that are arranged as bilayers with polar interfacial regions (interface) and a hydrocarbon core (HC).

Across a single leaflet of the lipid bilayer the environment changes from an aqueous solvent, to an interfacial region of polar, zwitterionic, or charged lipid headgroups, to a central core primarily composed of hydrophobic hydrocarbon chains. This complex amphiphilic nature of membranes is a significant contributor to the stability of membrane proteins making it very difficult to study their structures and interactions in isolation from their native environment. As a result of these technical challenges there is much less biochemical and structural data on membrane proteins compared to soluble proteins, yet improved techniques and

advances in membrane mimetics are beginning to address this problem (King and Dixon, 2008).

Remarkably, the structures of membrane spanning regions (or transmembrane (TM) domains) of integral membrane proteins have so far been observed to take one of only two structural forms, namely alpha helices or beta barrels (Bowie, 2005). In the cases of the beta barrel, the polypeptide chain transverses the membrane several times whilst for alpha helical proteins the chain may cross the membrane several times, to form a bundle of alpha helices or only once for proteins that contain only one TM domain. Representative models for these two structural motifs are shown in Figure 1.2. It is likely that the conformational space available to membrane proteins is constrained by the fact that the burial of hydrophilic peptide bonds in the hydrophobic core of a biological membrane is energetically costly (White, 2003). These two structural motifs allow for the formation of the greatest number of internal hydrogen bonds, which is able to offset the high energetic cost of desolvating the protein upon insertion into the membrane (White, 2003).

Since the focus of this study is on discerning the rules that govern the folding of alpha helical membrane proteins we will now consider in more detail the properties of this important class of proteins.

1.2 Helix-helix interactions in α-helical membrane proteins

The most abundant class of integral membrane proteins are those that span the membrane with a domain comprised of stretches of residues with alpha helical secondary structure; often termed transmembrane (TM) helices. A typical TM helix is formed from around 20 predominantly hydrophobic amino acids that span a lipid bilayer with a 30 Å thick hydrocarbon core (Hessa, White et al., 2005).The hydrophobic effect from the burial of hydrophobic residues in the core of the bilayer has a free energy of ~40 kcal mol^{-1} whilst the energetic cost to desolvate the hydrophilic backbone is only ~30 kcal mol^{-1}. Thus, there is a favourable free energy of ~10 kcal mol^{-1} for stabilising the helix in the membrane (Hessa, White

et al., 2005). A single protein may be polytopic and possess several helices that span the membrane bilayer, as is the case for the well known family of G-protein coupled receptors or monotopic and possess a single TM helix, as exemplified by the protein Glycophorin A (GpA) from erythrocytes (Lemmon, Flanagan et al., 1992; Kobilka, 2007).

Over the last twenty years or so it has become increasingly apparent that these TM helices serve a much more important function than merely anchoring the proteins within the bilayer. Study into the factors that drove the association of GpA monomers into a dimer revealed the ability of helix-helix interactions between TM domains to govern the folding of membrane proteins (Lemmon, Flanagan et al., 1992; MacKenzie, Prestegard et al., 1997). Specifically, in the case of GpA it was found that the specific arrangement of two Gly residues within the TM domains was responsible for the formation of the GpA dimer. Much has been achieved in delineating the rules that govern helix-helix interactions leading to the identification of several motifs that drive assembly of alpha helical proteins within membranes (Harrington and Ben-Tal, 2009). However, there still remains much to do to completely understand these assemblies.

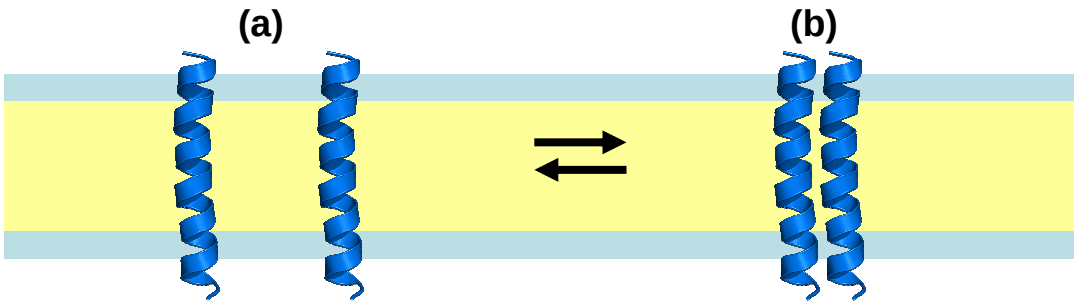


Figure 1.3. The two stage model of α-helical membrane protein folding

As proposed by Popot et al the two stage model describes how alpha helical membrane proteins can attain their native conformation. Membrane spanning stretches composed predominantly of hydrophobic amino acids that insert across membrane as (a) independently stable α-helices within the plane of the bilayer then (b) associate through lateral helix-helix interactions to form bundles of helices.

An important framework for understanding the association of α-helical transmembrane domains in membrane proteins is the “Two Stage Model” as proposed by Popot et al (Popot and Engelman, 1990). In this model (see Figure

1.3), stretches of predominantly hydrophobic amino acids insert into the membrane as independently stable α-helices (stage 1) and subsequently laterally associate in the plane of the bilayer (stage 2). This determines the folding and assembly of many integral membrane proteins. Perhaps, the most important contribution of the two stage model to the field of membrane protein folding is that individual transmembrane helices can be thought of as independently stable domains. This has guided the development of strategies for studying the assembly of alpha helical membrane proteins using protein fragments corresponding to the TM domains. A further key contribution of the two stage model is that it suggests the side-chain interactions are the determinants of specificity in helix-helix interactions. Other models for the assembly of membrane proteins have been proposed, namely the three stage model (Jacobs and White, 1989) and four stage model (White and Wimley, 1999), which expands upon the two stage model of Popot et al to incorporate the means by which the TM domain enters the membrane.

It is important to note that the two stage model says nothing about how the protein becomes inserted into the membrane but rather focuses attention on how the amino acid sequence of the helices might determine their interactions and proposes the question “What are the sequence determinants that drive association of transmembrane helices and hence determine the stability of membrane proteins?”

1.2.1 Helix-helix interaction motifs

The first indication that TM helix-helix interactions have sufficient specificity to drive tertiary structure formation came from early experiments performed on the α-helical membrane protein bacteriorhodopsin, in which the native protein fold was reconstituted from fragments (Popot, Trewhella et al., 1986; Popot, Gerchman et al., 1987). Similarly, this has also been shown for other proteins including lactose permease (Bibi and Kaback, 1990), rhodopsin (Ridge, Lee et al., 1995) and the red cell anion exchanger protein (Groves and Tanner, 1995).

For soluble proteins the hydrophobic effect is generally considered to be the major driving force for driving their folding. However, this cannot be the force driving association of TM α-helices because the hydrophobic effect arises solely from the increase in entropy upon dehydration of a non polar surface, and this is expended after the helices are inserted within the membrane (Von Heijne, 2003). Therefore, helix-helix association is more likely to be driven by van der Waals forces, such as the London dispersion force, or hydrogen bonding (White, Ladokhin et al., 2001; White, 2006). The presence of a polar amino acid which can form H bonds in the hydrophobic core of a membrane is energetically unfavourable but it can be compensated for by the insertion of the surrounding hydrophobic residues. The high energetic cost of breaking H-bonds in the hydrocarbon core of the bilayer should provide a strong stabilising force for helix association. However, such bonds are thought to have weak specificity relative to Van der Waals interactions and are believed to cause non-specific aggregation (White, 2006). It has been proposed that van der Waals interactions in close packing helices are the main determinants for TM helix association and that H-bonds serve to stabilize a preformed oligomer (Schneider, 2004).

Over the last two decades several amino acid sequence motifs have been identified that mediate helix-helix interactions within membranes (Senes, Engel et al., 2004; Harrington and Ben-Tal, 2009). One very significant motif is the small-xxx-small motif which frequently occurs in TM helices, and is found conserved amongst families of proteins whose functions include signal transduction, channels, transporters, toxins, and enzymes (Russ and Engelman, 2000). In this motif, two small residues such as alanine, serine or more commonly glycine are separated by three other residues. The GxxxG motif occurs most frequently and is over represented in statistical analysis of TM domains and is very often found flanked by β-branched amino acids (Senes, Gerstein et al., 2000). It was work carried out primarily in the Engelman lab on the protein Glycophorin A (Lemmon, Flanagan et al., 1992; Lemmon, Flanagan et al., 1992; Adams, Engelman et al., 1996) that identified the importance of the glycine variant of this motif, the GxxxG motif, for stabilising the GpA dimer and culminated in the determination of the solution NMR structure of the dimeric GpA TM domain (MacKenzie, Prestegard et al., 1997).

The separation of the two small residues by three other residues in the small-xxx-small sequence motif has the effect of placing the two small residues sequentially on the same face of the helix. The lack of bulky side groups at this position creates a pocket that allows the close approach of two TM helices, facilitating the formation of van der Waals interactions. The strength of the interaction is dependent on the sequence context, and can be enhanced by the presence of nearby β-branched residues (Russ and Engelman, 2000) and that GxxxG motifs located centrally in the TM helix mediate stronger helix-helix interactions than those at the ends (Johnson, Rath et al., 2006). It is possible that the close proximity of the protein backbones from the two helices establish networks of weak interhelical hydrogen bonds forming between alpha protons and carbonyl O atoms. Furthermore, the small-xxx-small motif may also act as a pivot point about which structural rearrangement can occur, as is observed when Ca^{2+} dissociates from Ca^{2+}-ATPase (Senes, Engel et al., 2004). It should be noted however, that although the presence of a GxxxG motif is highly indicative of a propensity for a helix to oligomerise, it has also been found in some cases to play no role in helix-helix interactions (Kobus and Fleming, 2005).

1.3 MHC Class II and Invariant Chain proteins

To further our understanding of the structural determinants for the assembly of TM helices in alpha helical membrane proteins this study has focused on two very important alpha helical membrane proteins, namely the Major Histocompatibility Class II protein (MHC) and the MHC-associated Invariant Chain. These proteins play a role in the immune system, and previous research has suggested they display helix-helix interactions within and between their TM domains. We will now consider what is currently known about these proteins.

1.3.1 MHC Class II proteins

Major Histocompatibility Class II proteins (MHC) are a diverse family of heterodimeric membrane proteins encoded by a large array of genes found in most vertebrates. They are composed of two polypeptide chains of α and β, that are 230 and 240 residues long with molecular weights of 33 kDa and 28 kDa, respectively. MHC proteins are part of the endosomal antigen presentation system

(Watts, 2004; Vyas, Van der Veen et al., 2008). They are responsible for the display of self or foreign peptides (termed antigens) to T-helper cells through the binding of the $CD4^+$ receptor on T-helper cells, which either ignore the self-peptides or recognise the foreign peptide and trigger an immune response, in a process called antigen presentation. MHC Class II proteins to be loaded with the antigenic peptide of 3 to 18 residues in length which occupies a groove formed by the association of MHC α and β chains, and transported to the cell surface. The initial stages of antigen presentation following the biosynthesis of MHC involves the chaperone protein MHC Class II associated invariant chain (Cresswell, 1994).

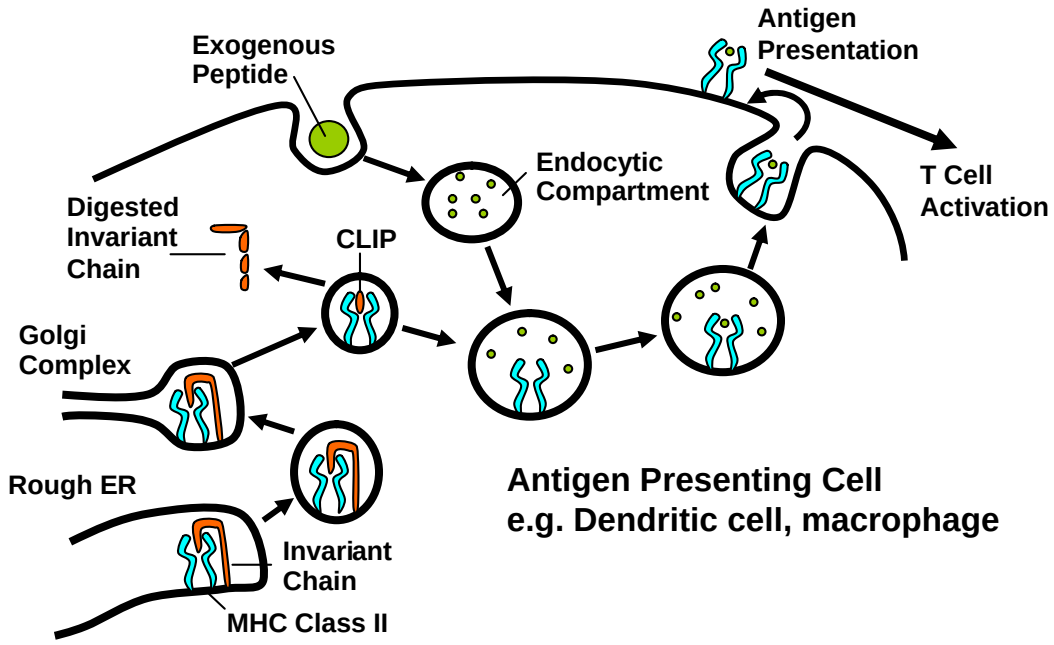


Figure 1.4. The mammalian antigen presentation pathway

Within the rough endoplasmic reticulum (ER) of antigen presenting cells (APCs) three MHC hetero-dimers bind to an invariant chain trimer (Ii) enabling export of the Ii-MHC complex to the Golgi complex for post translational modification of MHC. Ii is then digested by proteolysis leaving part of Ii known as the CLIP domain bound in the antigen binding domain of MHC. MHC is subsequently loaded with antigenic peptide derived from an invading pathogen. The loaded MHC is then transported to the plasma membrane and the complex presented to T cells at the APC surface.

The process of antigen presentation, outlined in Figure 1.4, takes place in specialised cells called antigen presenting cells (e.g. Dendritic cells, macrophages, and B-cells). Mutations in the MHC Class II proteins that disrupt antigen presentation are associated with diseases such as rheumatoid arthritis (Holmdahl,

2000), type I diabetes (Jones, Fugger et al., 2006), muscular sclerosis (Jones, Fugger et al., 2006), HIV infection (Schindler, Wurfl et al., 2003), asthma (Ye, Finn et al., 2003), and certain cancers (Ishigami, Natsugoe et al., 2001).

1.3.2 MHC Class II-associated Invariant Chain

As illustrated in Figure 1.4, a critical step in the endosomal pathway is the association of MHC Class II α and β-subunits with the MHC Class II-associated invariant chain (Ii)(Anderson and Miller, 1992). This has been shown to occur in the endoplasmic reticulum (ER) shortly after synthesis (Anderson and Miller, 1992; Peterson and Miller, 1992; Germain and Rinker, 1993; Romagnoli, Layet et al., 1993; Simonsen, Momburg et al., 1993). Ii is a 216 residue integral membrane protein that forms a homotrimer in the ER (Marks, Blum et al., 1990; Lamb and Cresswell, 1992). A single Ii trimer binds to three Class II α/β heterodimers to form a nonameric (nine chain) complex (Sung and Jones, 1981; Kvist, Wiman et al., 1982). It is only as part of this complex that MHC Class II proteins can be released from the ER, avoid rapid degradation, and be targeted to the endosomal pathway for use in antigen presentation (Ericson, Sundstrom et al., 1994; Thery, Brachet et al., 1998). Trimerisation of Ii is therefore a vital first step in antigen presentation and subsequently a vital step for activation of an immune response.

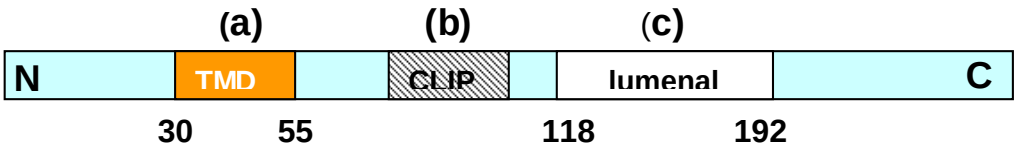


Figure 1.5. Schematic of the organisation of Invariant Chain

Schematic diagram of Human Invariant Chain showing the organisation running from the N- to the C-term of (a) the transmembrane domain (TMD) composed of residues 30-55 (b) the CLIP region which occupies the antigen binding site of MHC Class II proteins and (c) the luminal domain composed of residues 118-192 which is known to trimerise, Recent work has also implicated the TM domain in playing a role in the self-association of Ii.

Due to the significant role Ii plays in the immune system several studies have focused on investigating the structure of the Ii trimer in recent years. Trimerisation of Ii is thought to be mediated by the C-terminal lumenal domain

composed of residues 118-192 (Bijlmakers, Benaroch et al., 1994; Jasanoff, Wagner et al., 1998) (see Figure 1.5). It has been shown conclusively by structure determination that the lumenal domain forms a trimeric structure leading to the belief that this domain was solely responsible for trimerisation of Ii (Bijlmakers, Benaroch et al., 1994; Bertolino, Staschewski et al., 1995; Gedde-Dahl, Freisewinkel et al., 1997; Jasanoff, Wagner et al., 1998). However, there are indications that the transmembrane (TM) domain composed of residues 30-56 also has a role to play in the self association of Ii (Ashman and Miller, 1999).

In the last decade studies have shown that the TM domain of Ii is a site of important helix-helix interactions that impact upon both the structure and function of Ii. The sequence of the TM domain is highly conserved across species (Bremnes, Rode et al., 2000) and mutations in this region can prevent formation of the nonameric complex (Ashman and Miller, 1999) and therefore disrupt antigen presentation (Frauwirth and Shastri, 2001). It is unclear at present whether this is a result of the mutations destabilising the MHC Ii trimer or disrupting interactions between Ii and MHC Class II molecules. An 80-residue fragment of Ii derived from the N-terminal and TM domains has been shown to form trimers (Ashman and Miller, 1999). Mutational studies revealed that the mutation of the polar residues Gln49, Thr49 and Thr50 in the TM domain could prevent trimerisation of the full length Ii protein (Ashman 1999). Subsequently the secondary structure and tilt angle of the Ii TM domain in a bilayer was determined, leading to the prediction of a left-handed coiled coil trimeric model for the Ii TM (Kukol, Torres et al., 2002). This model suggested the presence of a stabilising hydrogen bonding network involving Gln47, Thr49 and Thr50 (Gratkowski, Lear et al., 2001; Zhou, Merianos et al., 2001).

Following on from this work, an investigation of a fragment of Ii corresponding to the TM domain (residues K26-R60), using the in vivo TOXCAT assay and biophysical methods, showed that in isolation the TM domain of Ii can form trimers and were able to determine a dissociation constant for the trimer in detergent micelles of DPC (Dixon, Stanley et al., 2006). Furthermore, the same study rationalised the deleterious effects of mutating Gln47 and Thr50 (Q47A, T50A) in terms of the large disruptions these could cause in Ii TM domain helix-

helix interactions by removing the potential for inter-helical hydrogen bonding. This study also showed that the role of the Ii TM domain in trimer formation can be conceptualised using the "two-stage model" of membrane protein folding described above.

As a result of this work on the TM domain of Ii a revised model for MHC Class II complex assembly was proposed (see Figure 1.6), that recognises the importance of Ii TM domain trimerisation in the formation of the nine-chain complex (Dixon, Stanley et al., 2006). However, the dissociation constant for the luminal domain has not been determined so the relative contributions of the TM and luminal domains to the trimerisation of Ii cannot be assessed.

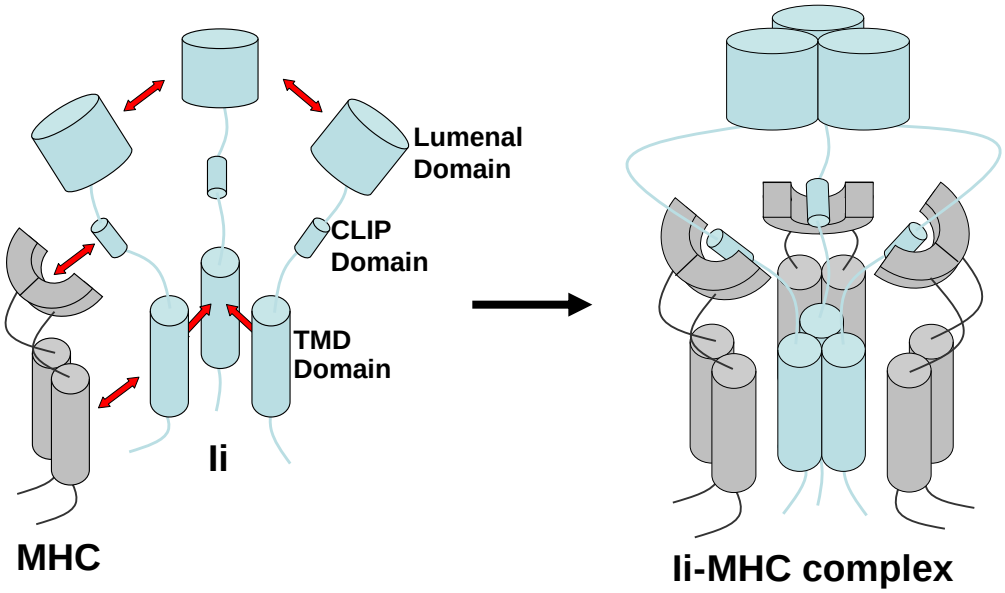


Figure 1.6. Proposed structure of Ii-MHC complex

Within the ER of mammalian cells, Invariant chain (blue) forms a trimer by association of the luminal and possibly also the transmembrane domains. Three MHC class II heterodimers of α and β chains (grey) associate with the Invariant Chain trimer through interactions between the CLIP domain of invariant chain and the antigen binding sites of the MHC molecule forming a nine chain (nonameric) complex that is subsequently exported from the ER to the antigen presentation pathway. The TM domains of MHC have also been implicated as playing a role in stabilising the complex. The formation of this complex is an absolute requirement for MHC to be exported from the ER.

The TM domains of Ii and MHC Class II proteins have been shown to be a possible site of important protein-protein interactions. The TM domains of MHC Class II proteins are thought to be important for intracellular trafficking and

antigen presentation (Barabanova, Kang et al., 2004). It has been shown that the Ii TM domain can interact directly and specifically with MHC class II proteins (Castellino, Han et al., 2001).

1.4 Development of NMR methods for studying helix-helix interactions

In addition to investigating the helix-helix interactions within the MHC-Ii complex we have also investigated the use of solution state NMR methods to further our understanding of helix-helix interactions in α-helical membrane proteins. The use of solution NMR to solve the atomic structure of a TM domain oligomer was investigated using E5 as a test subject for developing protocols for the optimisation of sample preparation and experimental parameters. We also investigated the use of paramagnetic NMR techniques and designed a novel method for determining helix-helix interactions. The development of NMR methods is discussed further in Chapter five.

1.5 Aims and objectives

The aim of this study was to further our understanding of the role of interactions between α-helical TM domains in driving the assembly of membrane protein complexes. Specifically, we have focussed on the complex formed between MHC and Ii, of key importance to the immune response. Using a wide range of methods we have investigated the formation of Ii trimers, and the role of helix-helix interactions in the formation of the MHC hetero-dimer and the Ii-MHC complex. Our results have allowed us to construct a model of the Ii-MHC complex stabilised by TM interactions.

2 Materials and methods

2.1 Reagents and materials

All laboratory reagents and materials used in this study were of the highest grade available and unless otherwise stated were supplied by Sigma Aldrich (UK), Fisher Scientific (UK), Avanti-polar lipids (USA), Pierce (UK) or Cole-Parmer (UK).

2.2 Bacterial Strains

Table 2.1 provides a list of the *Escherichia coli* used in this study. These were either commercially available or kindly provided by other research groups, as indicated.

Strain	Notes	Source	Reference
DH5α	*sup*E44 Δ*lac*U169 (Ø80lacZΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 thi-	Novagen (UK)	(Sambrook and Russell, 2001)
NT326	*F−(argF-lac)U169, rpsL150, relA1, rbsR, flbB5301, ptsF25, thi-1, deoC1, ΔmalE444, recA, srlA*	D.Engelman (Yale University, USA)	(Treptow and Shuman, 1985)
BL21(DE3)	Deficient in lon and ompT proteases, *ompT hsdS*B(rB- mB-) *gal dcm*	Novagen (UK)	
SU101	Possesses Wt LexA promotor	D. Schneider (University of Freiburg, Germany)	(Dmitrova, Younes-Cauet et al., 1998)
SU202	Possesses hybrid wt/mutant LexA promotor	D. Schneider (University of Freiburg, Germany)	(Dmitrova, Younes-Cauet et al., 1998)

Table 2.1. *E. coli* strains used in this study

2.3 Vectors

Table 2.2 provides a list of the plasmids used in this study. These were either commercially available or kindly provided by other research groups, as indicated.

Name	Notes	Source	Reference
pCC-Kan	New England Biolabs pMAL-c2 and –p2 vectors (pBR322 + *lacI* and Maltose binding protein)	D.Engelman (Yale University, USA)	(Russ and Engelman, 1999)
pCC-GpA	pCC-KAN with glycophorin A (GpA) transmembrane (TM) domain	D.Engelman (Yale University, USA)	(Russ and Engelman, 1999)
pCC-G83I	pCC-GpA with Gly83 to Ile substitution	D.Engelman (Yale University, USA)	(Russ and Engelman, 1999)
pALM100	Tetracycline resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pBLM100	Ampicillin resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pBLM-GpA	Ampicillin resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pBLM-G83I	Ampicillin resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pALM-GpA	Tetracycline resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pALM-G83I	Tetracycline resistant, IPTG inducible	D. Schneider (University of Freiburg, Germany)	(Schneider and Engelman, 2003)
pGEX-6p-3	GST expression vector, Ampicillin resistant	GE Healthcare (UK)	
pET31b(+)	IPTG inducible Ampicillin resistant	Invitrogen (UK)	

Table 2.2. DNA vectors used in this study

2.4 Affinity tags

A 6-amino acid His tag was present at the carboxy-terminus of KSI fusion proteins in peptides expressed using the pET31b(+) expression system. A glutathione S-transferase (GST) fusion tag was incorporated into fusion proteins in expression using the pGEX expression system.

2.5 Methods concerning growth of E. coli strains

E. coli cells were routinely cultured in Luria-Bertani (LB) liquid medium. Typically, growth of the culture was achieved by inoculating the LB medium with a single colony from an agar plate, followed by incubating for 16 hrs at 37°C under aerobic conditions (shaking at 250 rpm). For growth on agar plates (LB medium plus agar), cultures were incubated in a 37°C oven for 16 hrs. Stocks of each strain were maintained on agar plates stored at 4°C and were re-plated every two months.

2.5.1 Antibiotics

The following antibiotics were added where necessary and used at the following concentrations: ampicillin (100 μg/mL); chloramphenicol (34 μg/mL); tetracycline (3 μg/mL).

2.6 Cloning methods

2.6.1 Preparation of competent *E. coli* cells

Transformation competent cells were prepared using the calcium chloride method described in Sambrook and Russell (Sambrook and Russell, 2001). 10 mL of a mid-exponential phase culture (OD_{600} ~0.6) was pelleted and re-suspended in 10 mL of 100 mM $MgCl_2$ and incubated on ice for 5 mins. The cells were then pelleted and re-suspended in 1 mL of 100 mM $CaCl_2$. Cells were incubated for 2-24 hrs at 4°C before use.

2.6.2 Purification of vector DNA

Vector DNA was isolated from 2 - 5 mL of stationary phase overnight culture using a “QIAprep Mini-Prep” kit from Qiagen (Germany). The method was carried out as detailed in the manufacturer's instructions.

2.6.3 Digestion of vector DNA and purification

Restriction endonuclease digestions of vector DNA were carried out according to the enzyme manufacturer's guidelines. Where possible, a double digest was performed simultaneously using a suitable buffer that would maintain activity of both enzymes. A typical double digest reaction contained: 17 µL vector DNA (from a 50 µL plasmid mini-prep), 2 µL buffer (as supplied with the enzyme), and 0.5 µL of each restriction enzyme. Typically, reactions were incubated for 1 hr in a 37°C water bath. DNA was purified by gel extraction. DNA was excised from agarose gels and extracted using a QIAprep Gel Extraction kit (Qiagen, Germany) according to the manufacturer's instructions.

2.6.4 Preparation of phosphorylated oligonucleotide inserts

The TM domain sequences are cloned into the TOXCAT and GALLEX assays by using annealed long oligonucleotide primers that must first be phosphorylated for use in the ligation reaction. Complementary forward and reverse oligonucleotides encompassing the TM domain sequence of interest and restriction digest products at the 5' and 3' ends, were designed and purchased from Invitrogen (UK). 10 µM stock solutions were prepared of forward and reverse oligonucleotides. The oligonucleotides were phosphorylated in a reaction that typically contained 5 µL 10 µM oligonucleotide, 2 µL 10× kinase buffer (500 mM Tris-HCl (pH 7.6 at 25°C), 100 mM $MgCl_2$, 50 mM DTT, 1 mM spermidine and 1 mM EDTA), 1 µL 10 mM adenosine triphosphate (ATP), 2 µL T4 Pol Kinase (Fermentas), 10 µL sterile deionised H_2O (dH_2O). The reaction was incubated at 37°C for 30 mins followed by heating to 56°C for 10 mins to inactivate the kinase. Phosphorylated oligonucleotides were annealed in a reaction that typically contained 4 µL forward oligonucleotide, 4 µL reverse oligonucleotide, 2 µL annealing buffer (250 mM tris-HCl pH7.5, 20 mM $MgCl_2$, 600 mM NaCl), and 10 µL sterile dH_2O. The reaction was incubated at 95°C for 7 mins and then allowed to cool to room temperature.

2.6.5 Ligation of digested vector and phosphorylated insert

Ligation reactions were carried out according to the guidelines supplied with the T4 DNA ligase (Fermentas, UK). Digested vector and phosphorylated insert were

combined in a 1:3 ratio respectively, in a total reaction volume of 20 μL. Ligation reactions typically contained 6 μL 10 mM oligonucleotide insert, 2 μL digested vector, 2 μL 10× ligase buffer (400 mM Tris-HCl pH 7.8, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP), 1 μL 10 mM ATP, 1.5 μL T4 DNA Ligase (Fermentas), 7.5 μL H_2O. The reaction was incubated at room temperature for 1-4 hrs and then overnight on ice.

2.6.6 Transformation of competent cells

Typically, 2 μL of plasmid DNA or 10μL of a ligation mix was added to 0.1 mL of competent *E. coli* cells and incubated on ice for 30 mins. The cells were then heat shocked at 42°C in a water bath for 90 secs, after which 0.5 mL of LB was added and the cells incubated at 37°C for 30 mins. The cells were pelleted and re-suspended in 100 μL of LB before being spread on to agar plates containing appropriate antibiotic for plasmid resistance followed by incubation overnight at 37°C.

2.6.7 DNA sequencing

All DNA sequencing reactions were performed by the Molecular Biology Service, University of Warwick or GATC Biotech (Germany). The sequencing primers used in this study are given in Table 2.3.

Name	Sequence
pcckan_f	CCTTCATCAGCCACTGTAGTGAAC
pGEX_f	GGGCTGGCAAGCCACGTTTGGTG
pABLM_f	GGGATTCGTCTGTTGCAGGAAGAGGAAGAA

Table 2.3. Sequencing primers used in this study

2.7 Analytical gel electrophoresis (SDS-PAGE)

SDS sample loading buffer (60 mM Tris-HCl (pH 6.8), 2% SDS, 10% glycerol, and 0.01% bromophenol blue) was added to all samples. Electrophoresis was carried out at room temperature on 4-12% NuPAGE NOVEX Bis-Tris Mini Gels (Invitrogen) in MOPS-SDS running buffer.

Detection of protein bands by Coomassie blue staining: Protein gels were placed in fixer solution (50% methanol, 10% acetic acid) for 30 mins, then in stain solution (56 mL dH_20, 4 mL glacial acetic acid, 0.01 g Coomassie Brilliant Blue) for 1 hr or overnight with gentle shaking. Gels were then destained (20% methanol, and 7 % glacial acetic acid), until proteins bands were visible (1-2 hrs). Gels were rinsed well with deionised water before an image was taken.

Detection of protein bands by silver staining: Gels used in cross-linking analyses were stained using silver nitrate. Polyacrylamide gels were soaked in fixer solution (60 mL 50% acetone, 1.5 mL 50% TCA, 25 µL formaldehyde) for 15 mins with gentle shaking. The gel was then washed three times with dH_2O and soaked in dH_2O for a further 5 min. The gel was washed again 3 times with dH_2O before soaking in 50% acetone for 5 mins. The gel was then soaked in sodium thiosulphate solution (10 mg $Na_2S_2O_3$ in 60 mL dH_2O) for 1 min followed by 3 washes with dH_2O. The gel was then soaked in stain solution (160 mg silver nitrate, 600 µL formaldehyde, 60 mL dH_2O) for 8 min. Following two washes with dH_2O, the gel was soaked in developer (1.2 g Na_2CO_3, 25 µL formaldehyde, 25 mg $Na_2S_2O_3$ in 60 mL dH_2O) for 10-20 seconds. The development of protein bands was quenched by discarding the stain solution and soaking the gel in a 1% acetic acid solution in dH_2O for 1-2 mins. The gel was then rinsed and stored in dH_2O until an image was taken.

2.8 TCA precipitation of samples for SDS-PAGE

Trichloroacetic acid (TCA) dehydrates proteins and leading to aggregation of hydrophobic regions and the eventual precipitation of the protein. 40 µL samples were brought up to a total volume of 100 µL with H_2O. 100 µL of 10% TCA was added and the reaction left on ice for 20 mins. Samples were centrifuged for 15 mins at 15,000 × g. The resulting pellet was washed with ethanol and allowed to dry in air. The pellet was re-suspended in LDS loading buffer and loaded onto a gel immediately after boiling for 7 mins at 95°C.

2.9 Detection of proteins by immuno-blotting

The transfer of proteins from an acrylamide gel to a nitrocellulose membrane was achieved using the NuPAGE western blotting system according to the supplied instructions (Invitrogen, UK). After transfer, the membranes were blocked with 3% milk solution (1.2 g dried milk in 20 mL TTBS buffer (8 g NaCl, 2 g KCl, 3 g Tris base, 999 mL dH_2O, 1 mL Tween-20, pH 7.4)) for 1 hr or overnight. The membrane was then incubated under agitation with primary antibody, Anti-MBP (Sigma, UK), in 3% milk solution for 1 hr followed by washing in TTBS (typically, 3×5min washes). The membrane was then incubated with the secondary antibody, anti-mouse (alkaline phosphatise conjugate), in 3% milk solution for 1 hr with agitation. Immuno-reactive bands were detected using Sigma Fast BCIP/NBT Detection kit (Sigma, UK) according to the supplied instructions. BCIP (5-Bromo-4-chloro-3-indolyl phosphate dipotassium) is hydrolysed by alkaline phosphatise to form a blue intermediate which is then oxidised by NBT (nitrotetrazolium blue chloride) to produce an insoluble, dark purple dye. All antibodies used in this study were purchased from Sigma (UK) and included Anti His c-term, Anti MBP and Anti Mouse.

2.10 Synthesis and purification of synthetic peptides

2.10.1 Peptide synthesis

Peptides corresponding to the TM domains of interest, with amino acid sequences given in the relevant sections, were synthesised at the Keck Facility, Yale University, using standard Fmoc (*N*-(9-fluorenyl) methoxycarbonyl) chemistry (Fisher and Engelman, 2001).

2.10.2 RP-HPLC purification of peptides

Synthetic peptides were supplied as a crude reaction product and subsequently purified by optimised reversed-phase high performance liquid chromatography using strategies based on those reported to be effective for hydrophobic peptides (Lew and London, 1997; Kochendoerfer, Salom et al., 1999; Jones, Ball et al., 2000; Fisher and Engelman, 2001; Tiburu, Dave et al., 2003). Typically, 4 - 6 mg

of the crude product was dissolved in 400µL trifluoroacetic acid (TFA), 200µL trifluoroethanol (TFE), 400µL acetonitrile (ACN), 70/30 ACN/H_2O ACN and 200 µL of HPLC grade water or 400 µL of a mixture of formic acid-acetic acid-chloroform-trifluoroethanol (FACT) (Jones, Ball et al., 2000). or 1:4 formic acid:hexaflouroisopropanol (HFIP), 200 µL 70:30 HPLC grade isopropanol: ACN and 200 µL of HPLC grade water or 400 µL of a mixture of formic acid-acetic acid-chloroform-trifluoroethanol (FACT) (Jones, Ball et al., 2000). Samples were loaded onto either Phenomenex Jupiter 5 µm C4 (250 mm × 10 mm) or a Phenomenex Luna 5 µm CN (250 mm × 10 mm) reversed phase columns with a typical flow rate of 2 mL/min. The mobile phase was composed of water (Buffer A) and either 100% ACN or 70:30 isopropanol:ACN (Buffer B). Typically, initial experiments used gradients of 1% per min before being optimised to 0.3 – 0.5 % min following determination of the retention time of the desired peptide. The optimised gradients along with the HPLC spectra for each peptide will be shown in the following chapters when the results for purification are presented. Elution was monitored by absorbance of aromatic residues at 280 nm. All solvents were HPLC grade. Fractions containing the pure peptide were identified by mass spectrometry as described in Section 2.11. Multiple runs of RP-HPLC purification were performed, and fractions containing pure peptide were pooled and lyophilised. Following lyophilisation, the purity of the peptide was assessed by mass spectrometry as described in Section 2.11. Purified peptides were then stored at -20°C until required.

2.11 Mass spectrometry analysis of peptides

HPLC fractions and purified peptides were analysed by electrospray ionisation time-of-flight mass spectrometry (ESI-MS) on a Bruker MicroTOF or matrix-assisted laser desorption ionisation mass spectrometry (MALDI-MS) on a Bruker MALDI-TOF. For analysis by ESI-MS, 10 µL of 10% formic acid solution was added to 90 µL of samples prior to analysis. Spectra were acquired in positive ion mode and detection was between 50 and 3000 mass/charge (m/z). Spectra were typically recorded for 2 mins, averaged and deconvoluted to determine the mass of the main species. For analyses by MALDI-MS, samples were prepared by combining 5 µL of peptide with 5 µL of matrix solution (10 mg/mL α-cyano-4-

hydroxy cinnamic acid in 50% ACN, 50% H_2O and 0.1 % TFA). 1 µL of the sample/matrix solution was spotted onto a MALDI target plate and allowed to air dry for 30 mins. MALDI-MS spectra were acquired in the positive ion and linear mode. The mass range from m/z 2000 to 5000 was externally calibrated with polyethyleneglycol 2000.

2.12 Peptide concentration determination

Peptide concentration was determined by measuring the UV absorbance at 280 nm. For each peptide, the extinction coefficient at 280 nm was determined using ProtParam (Gasteiger, Hoogland et al., 2005) and used to calculate the peptide concentration using the Beer-Lambert law (1)

$$A = \varepsilon lc \quad (1)$$

where, A is the absorbance at 280 nm, ε is the extinction coefficient (M^{-1} cm^{-1}), l is the optical path length (cm) and c is concentration (mol L^{-1}).

2.13 Covalent cross-linking of peptides

Cross-linking reactions were carried out for 20 µM solutions of the peptide of interest in 20 mM sodium phosphate (pH 7), 0.15 mM NaCl and detergent in a reaction volume of 40 uL The concentrations of detergents were varied to give different micelle:peptide ratios using equation 2,

$$[\text{detergent}] = [\text{peptide}] \times R \times A + CMC \quad (2)$$

where [detergent] is the detergent concentration in mol L^{-1}, [peptide] is the peptide concentration, R is the peptide:micelle ratio, and A and CMC are the aggregation number and critical micelle concentration of the detergent, respectively. A 20 fold excess of the cross-linker Bis[sulfosuccinimidyl]-suberate (BS^3) (Pierce, UK) was used to cross-link the peptide in solution according to the manufacturer's protocol. BS^3 reacts specifically with the terminal NH_2 groups on lysine side chains provided that the reactive groups are within 11.4 Å of one another (Staros, 1982). The reaction was incubated at room temperature for 30 mins before being quenched by the addition of 1 M Tris-HCl (pH 7.5) to a final concentration of 20 µM and incubation at room temperature for 15 mins. Samples

uncross-linked and cross-linked in SDS were prepared as controls. 10 μL of 4X NUPAGE SDS loading buffer was added to all samples prior to analysis by SDS-PAGE and visualization of protein bands was achieved by staining with silver nitrate as described in Section 2.7.

2.14 Peptide reconstitution into lipid vesicles

Lipids and proteins were co-dissolved in TFE (0.25 mg/mL of peptide and 2.5 mg/mL lipid in 200 μL final volume). Solvent was then removed under a stream of nitrogen gas. The sample was then held under vacuum overnight to ensure total removal of solvent. The sample was redissolved in 50 mM $NaPO_4$ buffer, pH 7. Samples were then freeze-thawed six times using an ethanol-dry-ice bath and a water bath at 40°C followed by sonication for 5 mins at 40°C. This resulted in a loss of sample turbidity.

2.15 Heterologous expression of TM domain peptides

A GST protein expression system (GE Healthcare, UK) was employed in this study to generate a TM domain peptide fused with a lanthanide binding tag (LBT) sequence. The fusion protein was purified using a GST affinity column and the peptide released from the GST moiety by on-column cleavage with Prescission protease (GE Healthcare, UK). The protease is itself a GST fusion and remains bound to the column whilst the TM peptide is released and eluted. Expected yields from the system are up to 50 mg of pure peptide per 1 L of liquid culture (GE Healthcare, UK).

2.15.1 Cloning into the pGEX GST fusion expression system

An oligonucleotide insert was produced encoding for the TM of interest and the LBT sequence and possessing BamHI and XhoI restriction sites at the 3' and 5' ends, respectively (Genscript, USA). The synthetic gene was digested with BamHI and XhoI and purified by gel electrophoresis. DNA was excised from agarose gels and extracted using a QIAprep Gel Extraction kit (Qiagen, Germany) according to the manufactures instructions. The plasmid pGEX-6P-3 was digested

with BamHI and XhoI and purified as outlined in Section 2.6.3. The oligonucleotide insert was ligated downstream of a Glutathione-S Transferase (GST) gene to produce the construct pGEX-TM-LBT as outlined in section 2.6.5. The expression host *E. coli* BL21 was made competent as describe in section 2.6.1, and then transformed with pGEX-TM-LBT as outlined in Section 2.6.6. The success of ligation was confirmed by DNA sequencing of transformants using the primer pGEX_f as described in Section 2.6.7.

2.15.2 Induction checks of GST fusion protein constructs

Expression of the fusion protein from 10 mL cultures of BL21 cells transformed with the plasmid pGEX-TM-LBT was induced by the addition of 1 mM Isopropyl-β-D-1-thiogalactoside (IPTG) to the culture medium when an OD600 of ~0.3 was achieved. Aliquots normalised to an OD_{600} of 0.1 were taken prior to induction and every hour for ~6hrs post-induction for analysis of protein expression level by SDS-PAGE.

2.15.3 Small scale purification analysis of GST constructs

20 mL of cell culture was resuspended in 2 mL 1 × PBS (140 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3). 2 μL of 100 mg/mL DNAse and 100 mg/mL lysozyme were then added and the suspension incubated on ice for 15 mins. The detergent sarkosyl was then added to a final concentration of 1% w/v. The suspension was then sonicated using a probe sonicator for 3 × 30 secs on ice followed by centrifugation for 15 mins at 1700 × g to pellet insoluble material. A sample was taken prior to centrifugation for SDS-PAGE analysis of the protein content of the whole cell fraction. The supernatant was mixed with 100 μL of 75% Glutathione Sepharose™ 4B (GE Healthcare, UK), that had previously been washed four times with 800 μL of 1× PBS and 1% sarkosyl, to remove all traces of ethanol from the storage buffer. The sample-sepharose mixture was incubated at 4°C for 2 hrs with gentle mixing. Fusion protein bound to the sepharose was pelleted by a few seconds of centrifugation. A sample of the supernatant was taken for SDS-PAGE analysis in order to assess the level of fusion-protein binding to the sepharose. The sepharose was then washed four times with cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH

7) plus detergent. Samples of washes were taken for SDS-PAGE analysis. All residual buffer was removed from the sepharose. 4 μL of Prescission Protease (GE Healthcare, UK) was added to 96 μL of cleavage buffer and this was then added to the sepharose. The cleavage reaction was incubated for 16 hrs at 4 °C and then for a further 2 hrs at room temperature. The sepharose was pelleted and the supernatant removed. A sample of the supernatant was taken for analysis by SDS-PAGE. 20 μL of 2× NUPAGE LDS sample buffer was added to the sepharose and all other samples, and samples were then heated at 90°C for 7 mins before being analysed by SDS-PAGE. Protein bands were visualised by Coomassie staining.

2.16 Molecular modelling using CHI

Computational analysis of the helical interactions was performed using the CNS searching of helix interactions (CHI) program, the details of which have been described previously (Adams, Arkin et al., 1995; Adams, Engelman et al., 1996; Adams, Lee et al., 1998), on an 8-node dual 2.66-GHz Xenon processor Linux cluster (Streamline Computing, Warwick). The CHI program uses in vacuo computational modelling and molecular dynamics to generate an in silico representation of TM domains that can be used to identify oligomeric structures with energetically favourable interhelical interactions. Using CHI, canonical α-helices were built with sequences indicated in the Results sections. The starting structures incorporated both right-handed (-25°) and left-handed (25°) crossing angles and an axis-to-axis distance between the helices of 10.4 Å. During the search of interactions for TM domain dimers, the two helices were simultaneously rotated about their central axis in 30° increments from 0 to 360°. After each rotation, molecular dynamics (MD) simulations were performed using simulated annealing of atomic coordinates. Four different MD simulations were performed for each structure generated by helical rotation. Groups of structures with a backbone root mean squared deviation (rmsd) of ≤ 1 Å were placed in clusters of 10 or more members, followed by calculation of an average structure for each cluster and energy minimisation.

2.17 TOXCAT assay for homo-association of TM domains

2.17.1 Principle of the TOXCAT assay

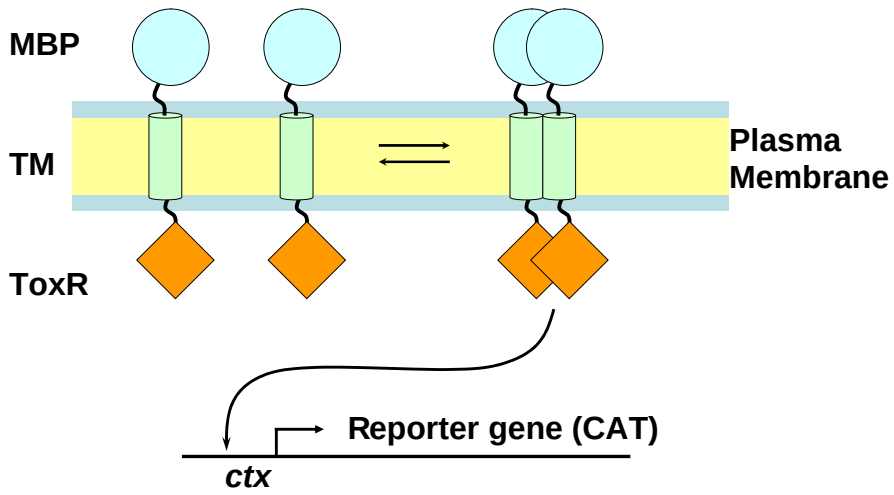


Figure 2.1. Principle of the TOXCAT assay

The TOXCAT assay is used to monitor self-association of α-helical TM domains. A chimeric protein is constructed of Maltose Binding Protein (MalE), the transmembrane domain (TM) of interest and the ToxR promoter. ToxR is a functional dimer that is incapable of dimerizing alone. Interactions between TMDs drive dimerisation of ToxR, which binds the ctx promoter activating transcription of the reporter gene Chloramphenicol Acyltransferase (CAT).

The propensity for TM domains to form homo-oligomers in *E. coli* membranes was investigated using the TOXCAT assay (Russ and Engelman, 1999). An overview of the assay is presented in Figure 2.1. This assay uses cloning techniques to insert the DNA sequence encoding the TM domain of interest into a gene encoding a fusion protein with the structure ToxR-TM-MBP, where ToxR is a dimerisation-dependent transcriptional activator from *vibrio cholera*, and MBP (maltose binding protein) is a monomeric periplasmic anchor protein that correctly orients the construct in the membrane. The TM domain being assayed replaces the native TM domain of the ToxR protein. Oligomerisation of the TM domain in the periplasmic membrane of *E. coli* allows the ToxR domains to dimerise and bind the ctx promoter, resulting in expression of the reporter gene product, chloramphenicol acetyltransferase (CAT). The amount of CAT expressed is directly proportional to the extent of oligomerisation. The TM domain of

glycophorin A (GpA), which is known to strongly dimerise, and its dimerisation defective mutant (G83I), were used as a positive and negative controls respectively.

2.17.2 Cloning TM domains into the TOXCAT assay

Using the restriction enzymes NheI and BamHI, oligonucleotide inserts corresponding to the DNA sequence of the TMD of interest (see Table 2.4) were generated from long primers and ligated into the vector pccKan, and the resulting plasmid was transformed into the host strain *E. coli* NT326, using the methods described in Section 2.6. Oligonucleotides were purchased as primers from Invitrogen, UK.

Name	**Sequence**
MHCα_F	CTAGCACTGTGGTCTGTGCCCTGGGGTTGTCTGTGGGC CTCGTGGGCATCGTGGTGGGCACCATCTTCATCATTCA AGGCCTGGG
MHCα_R	GATCCCCAGGCCTTGAATGATGAAGATGGTGCCCACC ACGATGCCCACGAGGCCCACAGACAACCCCAGGGCAC AGACCACAGTG
MHCα$G_{225}LG_{229}$L_F	CTAGCACTGTGGTCTGTGCCCTGCTGTTGTCTGTGCTG CTCGTGGGCATCGTGGTGGGCACCATCTTCATCATTCA AGGCCTGGG
MHCα$G_{225}LG_{229}$L_R	GATCCCCAGGCCTTGAATGATGAAGATGGTGCCCACC ACGATGCCCACGAGCAGCACAGACAACAGCAGGGCAC AGACCACAGTG
MHCα$G_{232}LG_{236}$L_F	CTAGCACTGTGGTCTGTGCCCTGGGGTTGTCTGTGGGC CTCGTGCTGATCGTGGTGCTGACCATCTTCATCATTCA AGGCCTGGG
MHCα$G_{232}LG_{236}$L_R	GATCCCCAGGCCTTGAATGATGAAGATGGTCAGCACCAC TCAGCACGAGGCCCACAGACAACCCCAGGGCACAGACC AGTG
MHCβ_F	CTAGCATGCTGAGCGGCATTGGCGGCTGCGTGCTGGGCG ATTTTTCTGGGCCTGGGCCTGTTTATTGG
MHCβ_R	GATCCCAATAAACAGGCCCAGGCCCAGAAAAATCACG CCCAGCACGCAGCCGCCAATGCCGCTCAGCATG
MHCβ$G_{233}LG_{237}$L_F	CTAGCATGCTGAGCGGCATTGGCCTGTGCGTGCTGCTG GTGATTTTTCTGGGCCTGGGCCTGTTTATTGG
MHCβ$G_{233}LG_{237}$L_R	GATCCCAATAAACAGGCCCAGGCCCAGAAAAATCACC AGCAGCACGCACAGGCCAATGCCGCTCAGCATG

Table 2.4. Oligonucleotide sequences used in the TOXCAT assay

Forward (_F) and their complementary reverse (_R) oligonucleotides corresponding to the TM domain of interest were ordered from Invitrogen and used as inserts for cloning into the TOXCAT assay.

2.17.3 Disc diffusion assay for CAT activity

Cells to be assayed were incubated at 37°C until they reached mid-exponential phase (OD_{600} = 0.6-0.8). An aliquot of the cells was then normalised to give an OD_{600} of 0.1, and 100 μL of cells were plated out on LB-agar media containing 100 μg/mL ampicillin. 42 μL of 90 mg/mL chloramphenicol (CAM) in ethanol was dried onto a Whatman grade 1 filter paper disk (diameter 30 mm) and placed in the centre of the plate. The plate was then incubated overnight at 37°C. The zone of cell growth inhibition surrounding the disk was then measured.

2.17.4 Chloramphenicol acetyltransferase activity assay

Cells to be assayed were incubated at 37°C until an OD_{600} of 0.6 was reached. 200 μL of cells were centrifuged at 14,000 rpm for 10 mins and resuspended in 50 μL of 100 mM tris-HCl, pH 8.0. 20μL of lysis solution (100 mM EDTA, 100 mM DTT, 50mM tris-HCl, pH 8.0) was then added. A drop of toluene was added to the top of the solution and the solution incubated at 30°C for 30 mins. The levels of CAT expression were assayed using the FAST CAT® Green (deoxy) Chloramphenicol Acetyltransferase assay kit (Invitrogen) according to the manufacturer's instructions. The reaction was terminated by adding 300 μL of xylene. The reaction was then mixed and centrifuged for 3 mins at 12,000 rpm and the upper phase collected. The fluorescence emission of the upper phase at 525 nm (excitation of 495 nm) was then measured using a Perkin Elmer LS50B fluorimeter. Emission from samples was normalised to that of the positive control, GpA.

2.17.5 Analysis of expression levels for TOXCAT chimera

Cells were grown until mid-exponential phase (OD_{600} ~0.6). 1 mL samples normalised to OD_{600} of 0.1 were centrifuged and the cell pellet resuspended in 80 μL of SDS loading buffer and analysed by SDS-PAGE. MBP bands were visualised by immunoblotting with anti-MBP and anti-mouse as described in Section 2.9.

2.17.6 Maltose plate assay for determining insertion and orientation

Overnight cell cultures were streaked out on maltose minimal media agar plates, containing M9 salts (48 mM Na_2HPO_4; 22 mM KH_2PO_4; 8.6 mM NaCl; 18.7 mM NH_4Cl), 2 mM $MgSO_4$, 100 μM CaCl, 0.4% maltose, and 15% (w/v) agar, and incubated at 37 °C for 2-3 days. If the construct is correctly oriented in the membrane (i.e. MBP on the periplasmic side), then the bacteria were able to utilise maltose as a carbon source as evidenced by growth of the colonies.

2.18 GALLEX assay for determining homo- and hetero-association of TM domains

2.18.1 Principle of the GALLEX assay

Whilst the TOXCAT assay is a valuable tool for the study of TM domain interactions, it is limited to studying the self-association of TM domains only. The GALLEX assay was designed by Schneider et al to allow the *in vivo* monitoring of the hetero-association of α-helical TM domains (Schneider and Engelman, 2003; Finger, Volkmer et al., 2006). Conveniently, the GALLEX assay can also be use to monitor homo-association providing a means of corroborating the findings from the TOXCAT assay.

An overview of the principle of the GALLEX assay is shown in Figure 2.2. Similar to the TOXCAT assay it involves the generation of a fusion protein containing the TM domain, MBP, and a DNA binding promoter. However, the promoter is the N-terminal part (residues 1-87) of LexA protein from *E.coli.* Two plasmids are used in the assay depending on whether homo- or hetero- association is being studied. The plasmid pBLM100 and the *E. coli* strain SU101 are employed to determine homo-association, whilst both pBLM100 and pALM100 and the *E .coli* strain SU202 are used to study hetero-association. A homo-associating fusion protein from the wt-LexA plasmid pBLM will bind to the wt-LexA promoter/operator and repress expression of lacZ (encodes β-Galactosidase) in the genome of the reporter strain SU101.

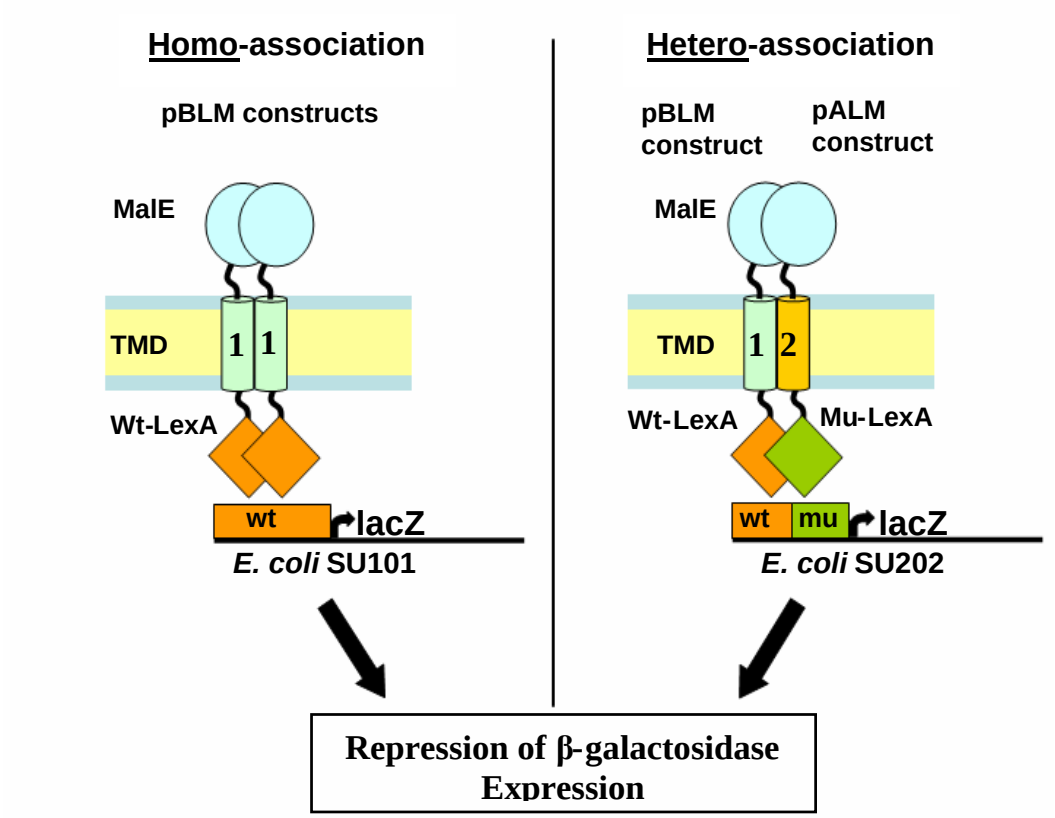


Figure 2.2. Principle of the GALLEX assay

The GALLEX assay can be used to monitor homo- and hetero-association of TMDs. In the homo-association assay a fusion protein composed of maltose binding protein (MalE), TMD of interest, and wild type LexA repressor (wt-LexA) is expressed from pBLM plasmid. Oligomerisation of this construct driven by the association of the TMDs enables LexA to dimerise and bind to the wild type (wt) lacZ promoter in *E. coli* SU101, repressing the expression of the enzyme β-Galactosidase. In the hetero-association assay two fusion proteins are expressed from pALM and pBLM plasmids. pBLM constructs are identical to those of the homo- assay whilst pALM constructs possess a mutant LexA domain (Mu-LexA). Oligomerisation of the different TMDs enables dimerisation of the wild type and mutant LexA domains, which bind a wild-type/mutant hybrid lacZ promoter (wt/mu) in *E.coli* SU202, repressing expression of β-galactosidase. The activity of β-galactosidase can be measured using a standard Miller assay.

For monitoring hetero-association fusion proteins are simultaneously expressed from a wt-LexA plasmid (pBLM) and from a mutated-LexA plasmid (pALM). A hetero-associated fusion protein will bind to the hybrid LexA promoter/operator and repress the expression of lacZ in the genome of the reporter strain SU202. In the *E. coli* strains SU101 and SU202, the lacZ reporter genes are under the control of the wild-type LexA recognition sequence (op+), and op408/op+ hybrid recognition sequence, respectively, which have been integrated into the genome of the host. The op408/op+ hybrid operator is composed of half of the wild-type promoter plus an altered half that recognises the mutant LexA domain (LexA408). The homodimers do not recognise the hydrid operator so do not interfere with the

measurement of hetero-association. Association of fusion constructs is then monitored by the activity of β-galactosidase (Daines and Silver, 2000).

2.18.2 Cloning TM domains into GALLEX plasmids

Name	Sequence
MHCα_F	CGACTGTGGTCTGTGCCCTGGGGTTGTCTGTGGGCCTCGTGGGCATCGT GGTGGGCACCATCTTCATCATTCAAGGCCTGA
MHCα_R	CTAGTCAGGCCTTGAATGATGAAGATGGTGCCCACCACGATGCCCACGA GGCCCACAGACAACCCCAGGGCACAGACCACAGTCGAGCT
MHCβ_F	CGATGCTGAGCGGCATTGGCGGCTGCGTGCTGGGCGTGATTTTTCTGGGCC GGCCTGTTTATTA
MHCβ_R	CTAGTAATAAACAGGCCCAGGCCCAGAAAAATCACGCCCAGCACGCAGCCG CAATGCCGCTCAGCATCGAGCT
MHCα-16_F	CGGCTCTTGGTCTTTCTGTTGGTCTTGTTGGTATTGTTGTTGGTACTGT TA
MHCα-16_R	CTAGTAACAGTACCAACAACAATACCAACAAGACCAACAGAAAGACCAA GAGCCGAGCT
MHCα-17_F	CGGCTCTTGGTCTTTCTGTTGGTCTTGTTGGTATTGTTGTTGGTACTGT TTTTA
MHCα-17_R	CTAGTAAAAACAGTACCAACAACAATACCAACAAGACCAACAGAAAGAC CAAGAGCCGAGCT
MHCα-18_F	CGGCTCTTGGTCTTTCTGTTGGTCTTGTTGGTATTGTTGTTGGTACTGT TTTTATTA
MHCα-18_R	CTAGTAATAAAAACAGTACCAACAACAATACCAACAAGACCAACAGAAA GACCAAGAGCCGAGCT
MHCβ-19_F	CGGCTCTTGGTCTTTCTGTTGGTCTTGTTGGTATTGTTGTTGGTACTGT TTTTATTATTA
MHCβ-19_R	CTAGTAATAATAAAAACAGTACCAACAACAATACCAACAAGACCAACAG AAAGACCAAGAGCCGAGCT
MHCα-16_F	CGATGCTTTCTGGTGTTGGTGGTTTTGTTCTTGGTGTTATTTTTCTTGG TA
MHCα-16_R	CTAGTACCAAGAAAAATAACACCAAGAACAAAACCACCAACACCAGAAA GCATCGAGCT
MHCα-17_F	CGATGCTTTCTGGTGTTGGTGGTTTTGTTCTTGGTGTTATTTTTCTTGG TGCTA
MHCα-17_R	CTAGTAGCACCAAGAAAAATAACACCAAGAACAAAACCACCAACACCAG AAAGCATCGAGCT
MHCα-18_F	CGATGCTTTCTGGTGTTGGTGGTTTTGTTCTTGGTGTTATTTTTCTTGG TGCTGGTA
MHCα-18_R	CTAGTACCAGCACCAAGAAAAATAACACCAAGAACAAAACCACCAACAC CAGAAAGCATCGAGCT
MHCα-19_F	CGATGCTTTCTGGTGTTGGTGGTTTTGTTCTTGGTGTTATTTTTCTTGG TGCTGGTCTTA
MHCα-19_R	CTAGTAAGACCAGCACCAAGAAAAATAACACCAAGAACAAAACCACCAA CACCAGAAAGCATCGAGCT
Ii-23_F	CGTATACTGGTTTCTCAATTTTAGTTACTTTATTATTAGCTGGTCAAGC TACTACTGCTTATTTCTTATATA
Ii-23_R	CTAGTATATAAGAAATAAGCAGTAGTAGCTTGACCAGCTAATAATAAAG TAACTAAAATTGAGAAACCAGTATACGAGCT

Table 2.5. Oligonucleotide sequences used in the GALLEX assay

Forward (_F) and their complementary reverse (_R) oligonucleotides corresponding to the TM domain of interest were ordered from Invitrogen and used as inserts for cloning into the GALLEX assay.

All plasmids used in this study were constructed by ligating synthetic oligonucleotide inserts encoding the TMD sequence of interest (see Table 2.5) into SpeI/SacI restriction digested pALM100 and pBLM100 vectors (see Table 2.2 and Section 2.6). Insertion of TMD sequence into pALM and pBLM plasmids was confirmed by DNA sequencing using the primer pABLM_f (see Section2.6).

For monitoring homo-association competent *E. coli* SU101 cells were transformed with pBLM100 constructs and plated out on LB-agar media containing ampicillin at 100 μg/ml (see Section 2.6). For monitoring hetero-association, competent *E. coli* SU202 cells were first transformed with the pALM100 plasmids and plated out on LB-agar media containing tetracycline at 6 μg/ml. The transformants were then made competent again, re-transformed with pBLM100 plasmids, and plated out on LB-agar media containing tetracycline and ampicillin at 6 and 100 μg/ml, respectively (see Section 2.6).

2.18.3 Monitoring β-galactosidase activity

In bacteria the enzyme β-galactosidase performs the function of hydrolysing the disaccharide lactose to yield galactose and glucose. The compound o-nitro-phenyl-galactopyranose (ONPG) is a homologue of lactose and can be hydrolysed by β-galactosidase to yield galactose and the yellow coloured compound o-nitrophenol (ONP). This reaction is used to monitor the repression of β-galactosidase caused by association of the TM domains bringing together the LexA dimer.

Clones of each transformant were grown overnight in LB (37°C) in the presence of IPTG (1 mM) and the relevant antibiotics for the plasmid construct. The next day, cells were diluted to an OD_{600} of 0.1 in LB medium containing the appropriate antibiotics and 1 mM IPTG and grown at 37°C to an OD_{600} of ~0.6, at which point cells were harvested. 900 μL of 1×Z buffer (300 mM $Na_2HPO_4{\cdot}7H_2O$, 200 mM $NaH_2PO_4{\cdot}H2O$, 50 mM KCl, 5 mM $MgSO_4{\cdot}7H_2O$, 50 mM 2-mercaptoethanol) was added to 50 μL of cell culture. 2-Mercaptoethanol is known to activate β-galactosidase. Cells were lysed by adding 10 μL of 0.1% SDS and 2 drops of chloroform with vortexing for 10 seconds until the solution became turbid. Samples were allowed to equilibrate to room temperature before adding

200 µL of o-nitro-phenyl-galactopyranose (ONPG, 4 mg/mL in 1×Z buffer) and mixing thoroughly. ONPG is colourless but is hydrolysed by β-galactosidase to produce the coloured compound ONP. The length of time taken for the colour to appear was recorded and the reaction stopped by addition of 0.5 mL of 1 M Na_2CO_3 solution. β-galactosidase operates optimally at pH 7.0, adding sodium carbonate increases the pH to around 9, thus inactivating the enzyme and quenching the reaction. Cell debris was pelleted by centrifugation. The absorbance of the supernatant was monitored at 420 and 550 nm, which corresponds to the absorbance from the ONP and the scattering from the cell debris, respectively. The Miller units of β-galactosidase were calculated using equation 3:

$$\text{Miller units} = \frac{1000 \times A_{420} - (1.75 \times A_{550})}{t\,(\text{min}) \times v\,(\text{mL}) \times OD_{600}} \quad (3)$$

where, A_{420} and A_{550} are the absorbance at 420 and 550 nm, respectively, t is the time taken for the colour to appear, v is the reaction volume, and OD_{600} is the optical density of the starting cell culture at 600 nm.

2.18.4 Controls for GALLEX assay

The *E. coli* strains SU101 and SU202 used in the GALLEX assay are not MBP deficient, so the MBP-deficient *E. coli* strain NT326 is used to confirm expression and membrane insertion of the chimera as described for the TOXCAT assay. To test for membrane insertion and orientation (spheroplast assay), competent NT326 cells were prepared as described in Section 2.6.6. Cells were grown overnight in LB plus ampicillin (100 µg/ml) and streptomycin (50 µg /mL). The next day the culture was diluted to an OD_{600} of 0.1 and incubated at 37 °C for ~1.5 hrs until the OD_{600} reached between 0.6 and 0.8. 1 mL of cells was normalised to an OD_{600} of 0.6 and the cells pelleted by centrifugation at 1700 × g. Cells were separated into periplasmic, cytoplasmic and membrane fractions using a procedure based on the EDTA/lysozyme/cold osmotic shock method (Randall and Hardy, 1986). Typically cells were harvested and resuspended in 1 mL chilled buffer I (100 mM tris-acetate pH 8.2; 0.5 M sucrose; 5 mM EDTA). 40 µL lysozyme (2 mg/mL) and 500 µL dH_2O was added before incubation on ice for 5 min followed by the

addition of 20 µL $MgSO_4$. The spheroplasts were pelleted by centrifugation at 10,000×g and the supernatant was collected as the periplasmic fraction. Spheroplasts were washed in 1 mL chilled buffer II (50 mM tris-acetate pH 8.2; 0.25 mM sucrose; 10 mM $MgSO_4$) and pelleted by centrifugation. The supernatant was discarded and the spheroplasts resuspended in proteolysis buffer (10 mM HEPES pH 7.6, 2 mM EDTA). The spheroplasts were then either treated with Proteinase K (to a final concentration of 0.25 mg/mL) for 30 min on ice, or broken open by 5× freeze-thaw cycles before treatment with Proteinase K. The whole cell and soluble fractions were TCA precipitated as described in Section 2.8.

2.19 Förster Resonance Energy Transfer (FRET)

2.19.1 Principles of FRET

The phenomenon of Förster resonance energy transfer (FRET) can be observed between pairs of molecules where the emission wavelength of one molecule (i.e. the donor) overlaps or resonates with the excitation wavelength of the other (i.e. the acceptor) (Johnson, 2005). This resonance between these molecules, termed a FRET pair, allows the transfer of energy through nonradiative dipole–dipole coupling from the donor molecule to the acceptor molecule. The energy transfer can be monitored using spectroscopic techniques, and in those cases where the molecules have fluorescent properties, the transfer can be measured using a fluorimeter, so often this technique is termed fluorescence resonance energy transfer. FRET can be exploited to study protein-protein interactions because the energy transfer displays a strong dependency on the distance between the FRET pair. The distance at which the energy transfer efficiency is 50% is known as the Förster distance. The combining of different FRET pairs enables Förster distances to be fine-tuned, and is thus an important consideration when choosing a FRET pair. Thus the FRET pair should be chosen to ensure the structural details of association do not impede observation of FRET. In the case of studying helix-helix interactions, consideration must therefore be made of possible antiparallel association or staggered alignment of the helices.

2.19.2 Selection of fluorophores and peptide synthesis

Peptides corresponding to the TM domains of Ii, MHCα and MHCβ were synthesised and labelled at the N-term with either fluorescein or rhodamine fluorophores (see Table 2.6), as described in Section 2.10.1. The fluorescein/rhodamine FRET pair was chosen on the basis that it has a favourable Förster distance for the study of TM helix-helix interaction at ~50 Å (Cardullo, Agrawal et al., 1988), and they have been used previously in the study of TM helix interactions (You, Li et al., 2005; Merzlyakov and Hristova, 2008).

Name	Amino acid sequence	Mass (Da)
Ii-Fl	Fl-KASRGALYTGFSILVTLLLAGQATTAYFLYQQQKK	4210.53
Ii-Rh	Rh-KASRGALYTGFSILVTLLLAGQATTAYFLYQQQKK	4264.52
MHCα-Fl	Fl-KELTETVVCALGLSVGLVGIVVGTVFIIRGLRSWK	4072.2
MHCβ-Rh	Rh-KSESAQSKMLSGVGGFVLGVIFLGAGLFIYFRNQK	4165.64

Table 2.6. Amino acid sequences of FRET peptides used in this study

Peptides were synthesised and labelled with fluorophores at the Keck Facility (Yale University, USA). The peptides were labelled at the N-term with the fluorophores fluorescein (Fl) and rhodamine (Rh), as indicated. Labelling was performed at the time of synthesis. Underlined residues indicate the extent of the predicted transmembrane domain. K residues were added to the sequences to aid purification and for cross-linking purposes. Peptides were supplied as crude product from the synthesis and purified by RP-HPLC. Expected masses were used to identify fractions of pure peptides during RP-HPLC purification. Peptides were end capped.

2.19.3 RP-HPLC purification of fluorophore labelled peptides

The peptides listed in Table 2.6 were purified using RP-HPLC according to protocols developed for the unlabelled counterparts and the purity was determined by mass spectrometry using the methods outlined in Section 2.11.

2.19.4 FRET sample preparation

Aliquots of lyophilised peptides were reconstituted in TFE and the protein concentration determined as described in Section 2.12. The peptides were reconstituted into detergent solution by mixing TFE-solubilised peptides with detergent solubilised in TFE. The mixtures of peptide and detergent in TFE were

dried under vacuum to a film, which was then solubilised in 20 mM phosphate buffer, 200 mM NaCl, pH 7.0 (Fisher, Engelman et al., 1999).

2.19.5 Fluorescence spectroscopy

Fluorescence spectra were recorded on a JASCO FP-6500 fluorimeter. Typically, the spectral resolution of the excitation and emission monochromators were both set to 5 nm and the scan speed at 200 nm/min. Emission spectra between 450 and 650 nm were recorded with an excitation wavelength of 439 nm (excitation wavelength of the fluorescein donor fluorophore) using a quartz cuvette (Starna Optiglass Ltd, Hainault, UK).

2.19.6 FRET calculations

Determining energy transfer: The percentage energy transfer, E, was calculated from measurements of the donor intensity at 520 nm in the absence and presence of the acceptor. E is described by equation 4:

$$E(\%) = (D - DA)/(D) \times 100 \quad (4)$$

where D and DA are the emission at 520 nm of samples containing only donor-labelled peptides and mixed samples of both donor- and acceptor-labelled peptides, respectively. The contribution to the emission at 520 nm from the direct excitation of the acceptor was removed by subtracting the spectra of sample containing only acceptor-labelled peptide from that of both donor- and acceptor-labelled peptides.

Determining oligomeric state: As shown by Veatch et al (Veatch and Stryer, 1977) the relative fluorescence yield of the donor as a function of the mole fraction of the acceptor can be used to distinguish amongst various oligomeric models. FRET data was fitted to the equation 5:

$$Q / Q0 = 1 - E (X_d - X_d^N) / X_d \quad (5)$$

where, Q is the absorbance at 520 nm, Q0 is the donor-only absorbance at 520 nm normalised to X_d, E is the transfer efficiency, X_d is the mole fraction donor and N is the oligomeric state.

2.20 NMR Spectroscopy

Sample preparation for NMR analyses: Samples for use in solution-state NMR analyses were prepared by dissolving each of the selectively labelled TM peptides shown in Figure 6.1, in deuterated sodium dodecylsulfate (SDS) (Avanti Polar Lipids, Alabaster, USA) detergent micelles. A series of samples were prepared in which peptide concentration was kept approximately constant (0.5–0.7 mM) while the concentration of SDS detergent was steadily increased from ~ 25 mM up to 700 mM to achieve a range of peptide: micelle molar ratios as indicated in Chapter 6. Briefly, the desired amount of peptide dissolved in trifluoroethanol (TFE) (Sigma-Aldrich, UK) was added to an aqueous SDS solution of the appropriate concentration. The final volume of TFE was kept to a minimum to prevent precipitation of SDS. The peptide–detergent solution was then lyophilized and resuspended in 20 mM sodium phosphate buffer (pH 5.8) containing 10% D_2O. The sample was subsequently mixed using a vortex mixer followed by sonication at 40°C for 15–30 mins. Samples were allowed to equilibrate for 24 hours before acquiring measurements. The peptide: micelle molar ratio for each sample was calculated using a value for the SDS critical micelle concentration (CMC) of 8 mM and an average aggregation number of 62 according to published values (le Maire, Champeil et al., 2000). This resulted in samples with peptide: micelle ratios ranging from 2:1 (or 2) to 1:22 (or 0.05).

NMR spectroscopy: Sensitivity enhanced 2D ^{15}N-^{1}H heteronuclear single quantum correlation (HSQC) spectra were recorded at 40°C on either a 500 or 700 MHz Bruker Avance spectrometer fitted with a cryoprobe. All spectra were processed using Topspin 2.0 (Bruker, UK) and analyzed with CcpNMR software (Vranken, Boucher et al., 2005). Backbone amide proton and nitrogen chemical shift assignments of the $E5_{TM}$ peptides were obtained from ^{15}N-edited total correlation spectroscopy (TOCSY) experiments with a spin-lock time of 60 ms, and ^{15}N-edited nuclear Overhauser spectroscopy (NOESY) spectra with mixing times of 60 to 100 ms. Average backbone amide chemical shift differences ($\Delta\delta$) were calculated according to equation 6:

$$\Delta\delta = ((\Delta\delta_{HN})^2 + (\Delta\delta_N / 5)^2)^{0.5} \quad (6)$$

where $\Delta\delta_{HN}$ and $\Delta\delta_{N}$ are the chemical shift differences between monomeric and dimeric species for the amide proton and nitrogen atoms, respectively according to the method reported by Wu et al (Wu, Shih et al., 2007).

2.21 Circular Dichroism

CD spectra were measured using a Jasco J715 spectropolarimeter (Jasco UK, Great Dunmow, UK) and 1.0 mm path-length quartz cuvettes (Starna Optiglass Ltd, Hainault, UK). All spectra were recorded from 190 to 260 nm (data below 200 nm are not shown due to high noise of light scattering) using 2.0 nm spectral bandwidth, 0.2 nm step resolution, 100 nm min $^{-1}$ scanning speed, and 1 s response time. 40 μM peptide samples were prepare in 50mM sodium phosphate buffer (pH 7), 100 mM NaCl containing various concentrations of detergent as indicated in Chapter 6. CD spectra of the buffer and detergent alone were subtracted to obtain the final spectra.

2.22 Oriented Circular Dichroism

Oriented circular dichroism is a useful tool for determining the insertion of TM domains into lipid bilayers, and involves a slight modification to the technique of circular dichroism. Differential absorbance of right and left-handed circularly polarised light by proteins of different secondary structure give characteristic spectra for alpha helix, beta sheet, and random coil (Merzlyakov, You et al., 2006). Peptides and lipids (DMPC) were codissolved in organic solvent (0.25 mg/mL peptide and 2.5 mg/mL lipid in 50 μL TFE). The solution was deposited dropwise on to a quartz slide, and the solvent removed under a stream of nitrogen forming multilamellar vesicles containing lipid bilayers and peptide. To ensure total solvent removal, the sample was held under vacuum overnight. The sample was hydrated by placing in a vessel with a drop of water into the chamber containing the sample slide, and equilibrated overnight. The slide was then mounted into a custom holder with the multilamellar layers oriented perpendicular to the light path. A CD spectrum was acquired using the same parameters as described in Section 2.21. The slide was then rotated in increments of 45°, acquiring a CD spectrum each time for a total of 8 measurements through a total rotation of 360°. The spectra were then averaged to minimise artefacts arising

from linear dichroism to generate the OCD spectra. An OCD spectrum of lipid alone was then subtracted from this spectrum to give the final OCD spectra.

2.23 Analytical Ultracentrifugation

Sedimentation velocity data were collected for peptide concentrations of 67, 134, 268 μM solubilised in 15 mM DPC (Avanti Polar Lipids) and 100mM NaCl using a double-channel centrepiece, a speed of 60,000 rpm, and a temperature of 25°C using a Beckman XL-1/A analytical ultracentrifuge. Buffer was prepared in 52.5% D_2O (Cambridge Isotope Laboratories, Andover MA) to match the buoyant density of the detergent. When the solvent matches the buoyant density of the detergent micelles, the only contribution to the buoyant molecular weight is from the peptide (Kochendoerfer, Salom et al., 1999). A total of 400 scans were recorded in each case, with 50 s between each scan. The moving boundary was monitored by repetitive radial scanning at a constant step size of 0.003 cm at 280 nm using a UV absorption optical system. Fitting of the resulting profiles to various oligomeric state models was achieved using SEDFIT (Schuck, 2000) to generate a continuous sedimentation coefficient distribution, which was subsequently converted to a molecular mass distribution using a peptide monomeric molecular mass of 3924 Da, a buffer density of 1.05971 g ml^{-1}, a buffer viscosity of 1.0267 centipoise, and a partial specific volume of 0.7792 ml g^{-1} (calculated using SEDNTERP (Hayes, D, B., Lane, T., Philo, J., University of New Hampshire, USA)).

3 TM domain interactions of Invariant Chain

3.1 Introduction and objectives

As discussed in Section 1.3.2, MHC Class II-Associated Invariant Chain (Ii) is known to be a trimeric protein that binds to three MHC Class II α/β hetero-dimers facilitating the release of the nonomeric complex from the ER and its subsequent involvement in the antigen presentation pathway of mammalian immunity (Frauwirth, Sanderson et al., 1995). The soluble domain of Ii is known to trimerise and the MHC hetero-dimers associate with Ii via the extra-membranous CLIP domain of Ii which occupies the antigen binding domain of MHC (Jasanoff, Wagner et al., 1998).

The TM domain of Ii has previously been studied in isolation using *in vivo* assays and synthetic peptides and was found to form a specific trimer that was potentially stabilised by inter-helical hydrogen bonding involving residues Q49 and T50 (Ashman and Miller, 1999; Kukol, Torres et al., 2002; Barabanova, Kang et al., 2004; Dixon, Stanley et al., 2006). Of particular note is the recent study by Dixon, Stanley et al where the association of the TM domain of Ii was studied by a combination of the in vivo TOXCAT assay and analysis of model peptides using cross-linking and AUC. This was the first study to analyse the oligomeric state of the of Ii TM domain in isolation from the rest of the molecule and suggested it could strongly and very specifically self associate to form a trimeric oligomeric state. Interestingly, the results from their cross-linking analysis of a peptide corresponding to residues K26-R60 of Ii identified oligomeric states of only trimer and hexamer in DPC whilst the AUC data was found to fit best to a monomer-trimer model for association. Notably, Ii TM domain was not found to associate to form intermediate oligomeric states e.g. dimer, tetramer, or pentamer. In order to progress with our investigation of the association of the TM domain of Ii with the α- and β-chains of MHC, the self-association of the Ii TM domain and its oligomeric state were substantiated by further *in vivo* and *in vitro* experiments, as described in this section. In addition, the effect of detergent concentration on

the oligomeric state of Ii has not previously been considered but is addressed in the studies presented here. The Ii TM domain also provided the test sequence for developing in-house protocols.

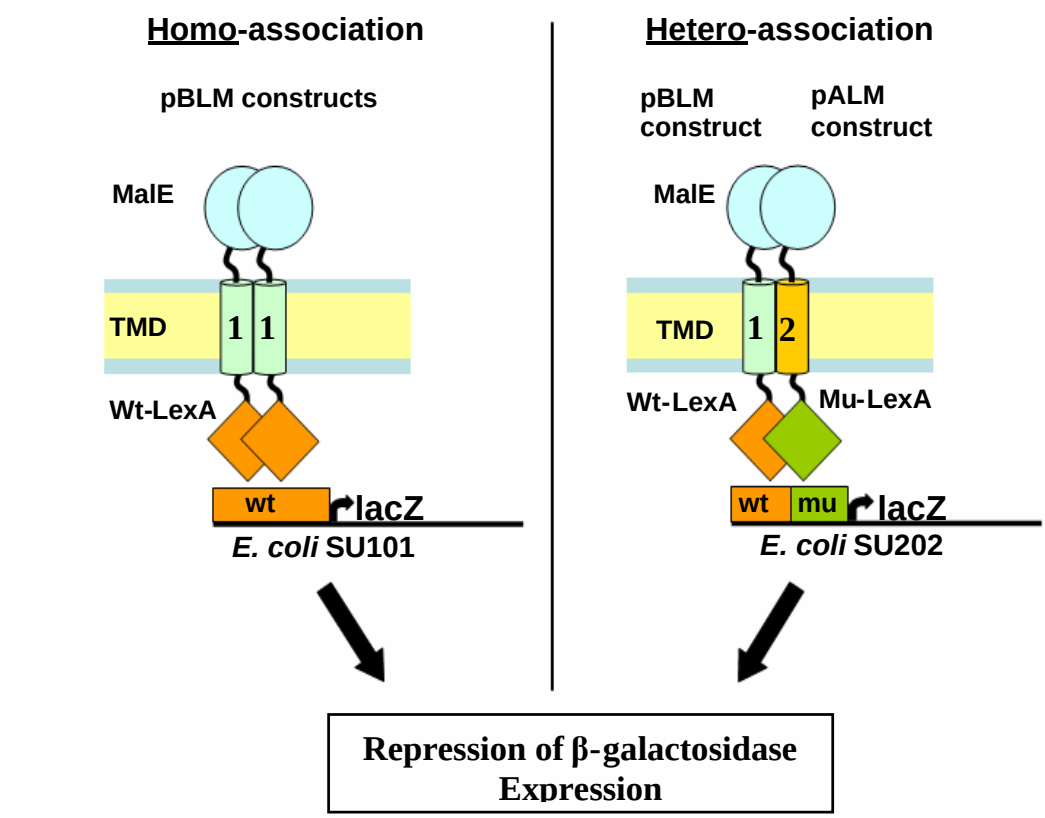


Figure 3.1. Principle of the GALLEX assay

The GALLEX assay can be used to monitor homo- and hetero-association of TMDs. In the homo-association assay a fusion protein composed of maltose binding protein (MalE), TMD of interest, and wild type LexA repressor (wt-LexA) is expressed from pBLM plasmid. Oligomerisation of this construct driven by the association of the TMDs enables LexA to dimerise and bind to the wild type (wt) lacZ promoter in *E. coli* SU101, repressing the expression of the enzyme β-Galactosidase. In the hetero-association assay two fusion proteins are expressed from pALM and pBLM plasmids. pBLM constructs are identical to those of the homo- assay whilst pALM constructs possess a mutant LexA domain (Mu-LexA). Oligomerisation of the different TMDs enables dimerisation of the wild type and mutant LexA domains, which bind a wild-type/mutant hybrid lacZ promoter (wt/mu) in *E.coli* SU202, repressing expression of β-galactosidase. The activity of β-galactosidase can be measured using a standard Miller assay.

The objectives of this part of the project are as follows: to confirm that the Ii TM domain can weakly self-associate and that its oligomeric state is trimeric; to confirm the GALLEX assay (discussed in Section 2.18.1, and presented again in Figure 3.1) can be used to monitor self-association of TM domains *in vivo* using Ii and optimise the methodology for use in further studies involving MHC TM

domains; to explore the possible effects of detergent concentration on self-association of Ii TM domain peptide; to explore the use of FRET experiments in studying TM domain association using Ii and optimise methodology for use in subsequent studies with MHC TM domains

3.2 Monitoring self-association of Ii TM domain with the GALLEX assay

Using the *in vivo* TOXCAT assay, previous studies revealed the propensity for the Ii TM domain to oligomerise in *E.coli* membranes (Dixon, Stanley et al., 2006). An alternative method for determining TM domain oligomerisation is the GALLEX assay as proposed by Schneider et al which has the additional benefit of enabling association of different TM domain sequences to be monitored (i.e. hetero-association) (Schneider and Engelman, 2003).

The GALLEX assay has now been implemented at Warwick, optimized in house and used to confirm the self-association of the Ii TM domain. The principle of the GALLEX assay for measuring TM helix-helix association was described in detail, and the methods provided in Section 2.18. In brief, the system enables TM domain association in the inner membrane of *E. coli* to be detected by the repression of the reporter gene β-galactosidase. In the GALLEX assay self-association can be assayed by inserting the TM of interest into the plasmid pBLM100 followed by transformation of the resulting plasmid into the host strain SU101. The sequence for the wild type TM domain of Ii (see Figure 3.2a) was cloned into the plasmid pBLM100 and assayed using the GALLEX assay as described in Section 2.18. Positive and negative controls for self-association were provided by the dimeric TM domain of GpA and its oligomerization-deficient mutant GpA_{G83I}, respectively.

The results from the GALLEX assay are presented in Figure 3.2b. The Ii signal was intermediate between that of the positive and negative control, confirming that Ii can self-associate in *E.coli* membranes. This is consistent with the self-association observed with the alternative assay TOXCAT (Dixon, Stanley et al., 2006).

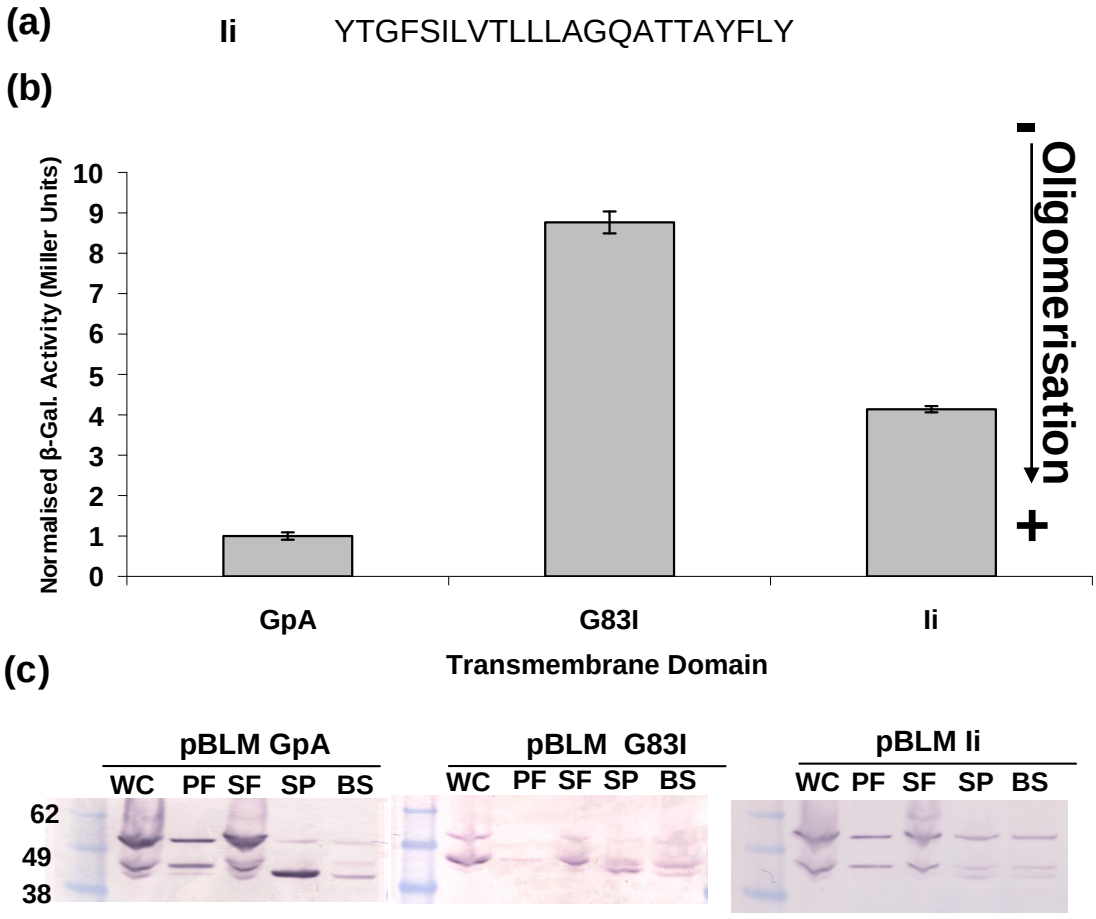


Figure 3.2. Monitoring self association of Ii TM domain using the GALLEX assay

Self-association was monitored using the GALLEX system as described in the text (a) Amino acid sequence of the Ii TM domain insert cloned into the plasmid pBLM100 and subsequently transformed into *E.coli* strain SU101 as described in Section 2.18.2 (b) β-galactosidase activity mediated by the oligomerisation propensity of the expressed constructs in *E.coli* SU101. Repression of activity is indicative of association of the TM domains. Data is an average from three independent measurements. Expression of the chimeric proteins was induced by the addition of 1 mM IPTG. Details of the β-galactosidase assay and the calculation of Miller Units are described in the Section 2.18.3. All plasmids and *E. coli* strains were kindly proved by Dr Schneider. GpA and the dimerisation deficient mutant of GpA, G83I, act as positive and negative controls respectively. (c) Test for insertion and orientation of the expressed chimera. Western blot analysis of E.coli extracts: WC, whole cell; PF, periplasmic fraction; SF, spheroplast fraction; SP, spheroplast proteolysis; BS, broken spheroplast proteolysis. The expressed chimeric proteins with a molecular mass of 54kDa are found predominantly in the inner membrane fraction and correctly oriented in the membrane.

3.3 Synthesis and purification of Ii TMD peptides for in vitro studies

The reductionist approach of using model synthetic peptides in conjunction with in vitro techniques is proving to be a productive strategy for the study of TM domain interactions. However, the synthesis and purification of these highly hydrophobic transmembrane α-helices are notoriously difficult. A commonly used strategy that has met with success is the use of Fmoc synthesis coupled with RP-HPLC that is optimised to each peptide(Fisher and Engelman, 2001). Even with advances in this approach it still can be difficult to separate the desired peptide from the crude product and very often broad overlapping peaks will be observed that may correspond to truncation products, conformational differences or oligomers along with reactants from the synthesis. Nevertheless, this approach has been utilised in this study to explore the self-association of the Ii TM domain. Peptides that correspond to the TM domain of Ii were synthesised at the Keck Facility (Yale University, USA) in addition peptides labelled with the fluorophores fluorescein and rhodamine were produced for use in FRET experiments. The amino acid sequences of the peptides synthesized for this study are shown in Table 3.1.

Name	Amino acid sequence	Mass (Da)
Ii	KASRGALYTGFSILVTLLLAGQATTAYFLYQQQGR	3808.4
Ii_K	KASRGALYTGFSILVTLLLAGQATTAYFLYQQQKK	3892.55
Ii-Fl	Fl-KASRGALYTGFSILVTLLLAGQATTAYFLYQQQKK	4210.53
Ii-Rh	Rh-KASRGALYTGFSILVTLLLAGQATTAYFLYQQQKK	4264.52

Table 3.1. Amino acid sequences of Ii_{TM} peptides used in the in vitro studies of Ii TM association

Peptides were synthesised at the Keck Facility (Yale University, USA). Ii-Fl and Ii-Rh were synthetic peptides produced for use in FRET studies and were labelled at the N-term with the fluorophores fluorescein (Fl) and rhodamine (Rh), respectively. The wild type sequence of Ii contains a C residues at the second position identified in bold. This was mutated to A in the synthetic peptides to remove the complication of cysteine bond formation. Labelling was performed at the time of synthesis. Underlined residues indicate the extent of the predicted transmembrane domain. K residues were added to the sequence in Ii_K at N-term for cross-linking purposes. Peptides were supplied as crude product from the synthesis and purified by RP-HPLC. Expected masses were used to identify fractions of pure Ii peptides during RP-HPLC purification. Peptides were end capped.

The general details of peptide synthesis, purification and analysis by mass spectrometry were as described in Section 2.10. The peptides were supplied as a crude product of the synthesis and thus contained undesirable contaminants such as fmoc protecting groups from the synthesis and truncated peptides, requiring the peptide to be purified. Reversed phase high pressure liquid chromatography (RP-HPLC) was employed for this task as it is a widely used technique for the purification of hydrophobic peptides.

Typical RP-HPLC chromatograms and mass spectra of pooled pure fractions for the peptides Ii, Ii_K, Ii-Fl and Ii-Rh are shown in Figures 3.2 to 3.5 respectively. In each case, the peak corresponding to the peptide in the RP-HPLC chromatogram is indicated in the figure and the major component of the pooled fractions are the desired peptide.

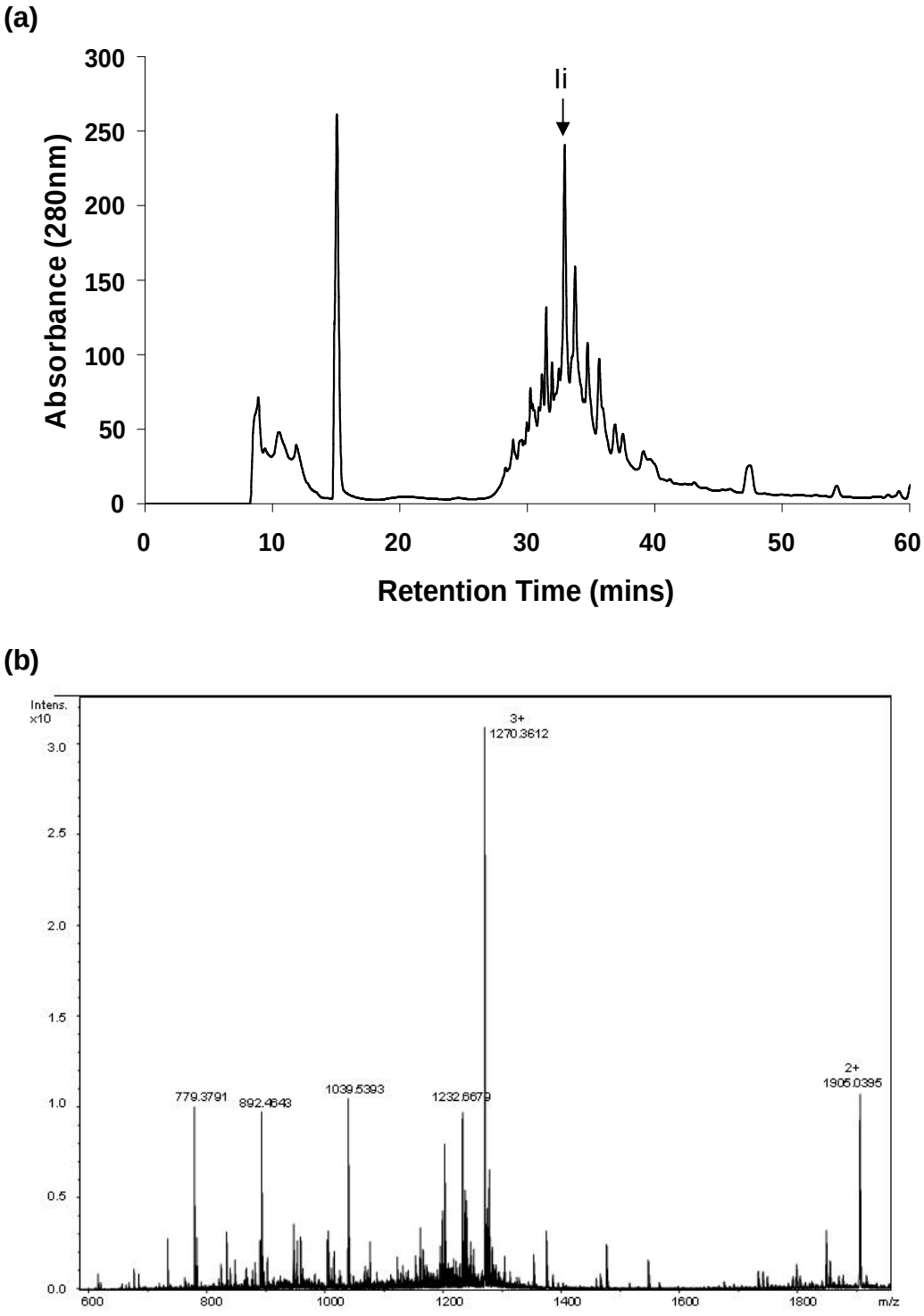


Figure 3.3. RP-HPLC purification of Ii peptide

(a) The Ii peptide was purified by reverse phase HPLC (solid line) using a linear 1% per min Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent , on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the Ii peptide is indicated. (b) Purity of pooled fractions was analysed using ESI mass spectrometry. The major peak with a mass of 1270 Da corresponds to the expected mass for the 3^+ charge state of the Ii peptide.

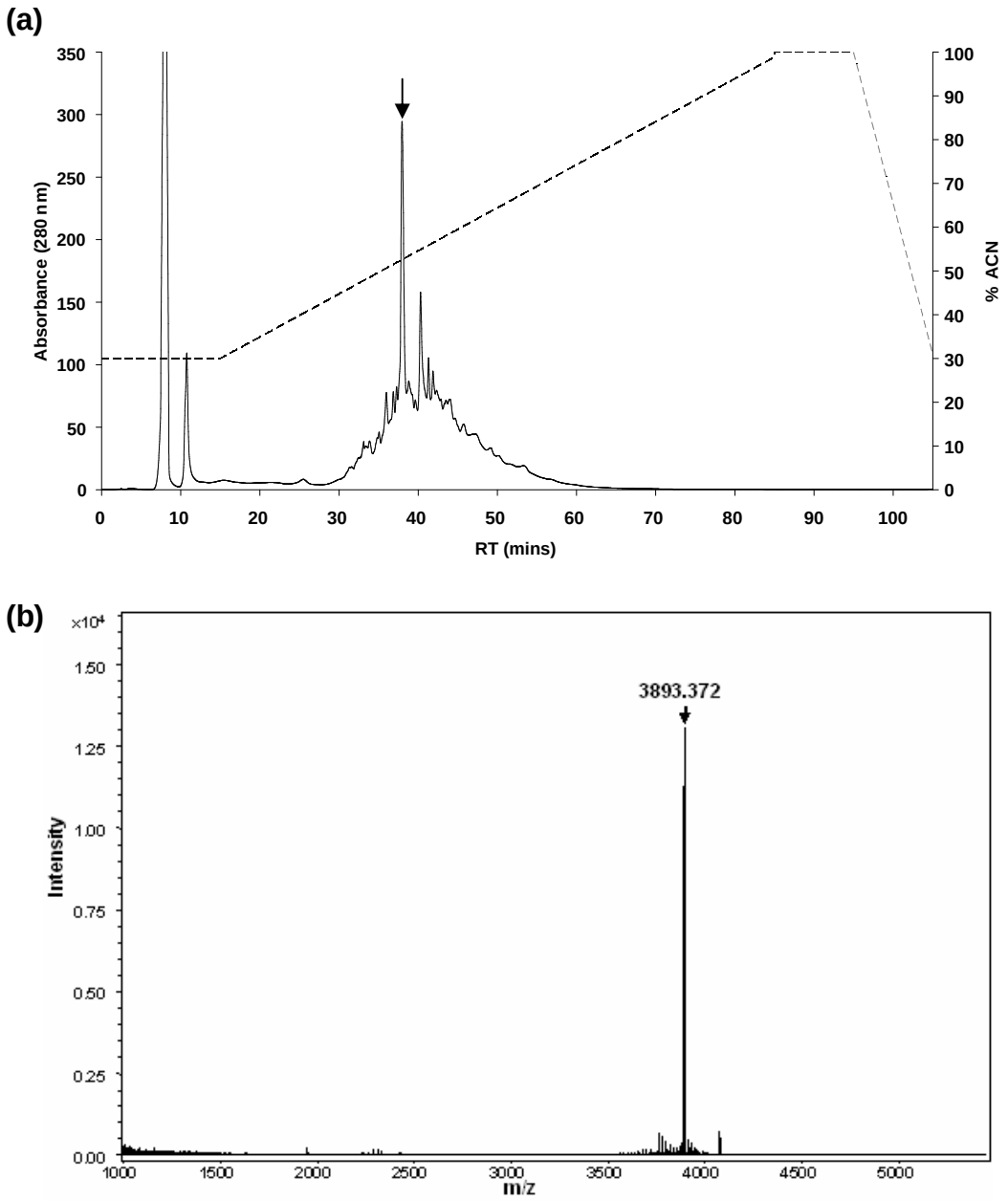


Figure 3.4. RP-HPLC purification of Ii_K peptide

(a) The Ii_K peptide was purified by reverse phase HPLC (solid line) using a linear 1% per min Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent , on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the Ii_K peptide is indicated. (b) Purity of pooled fractions was analysed using MALDI mass spectrometry. The major peak with a mass of 3893 Da corresponds to the expected mass for Ii_K peptide.

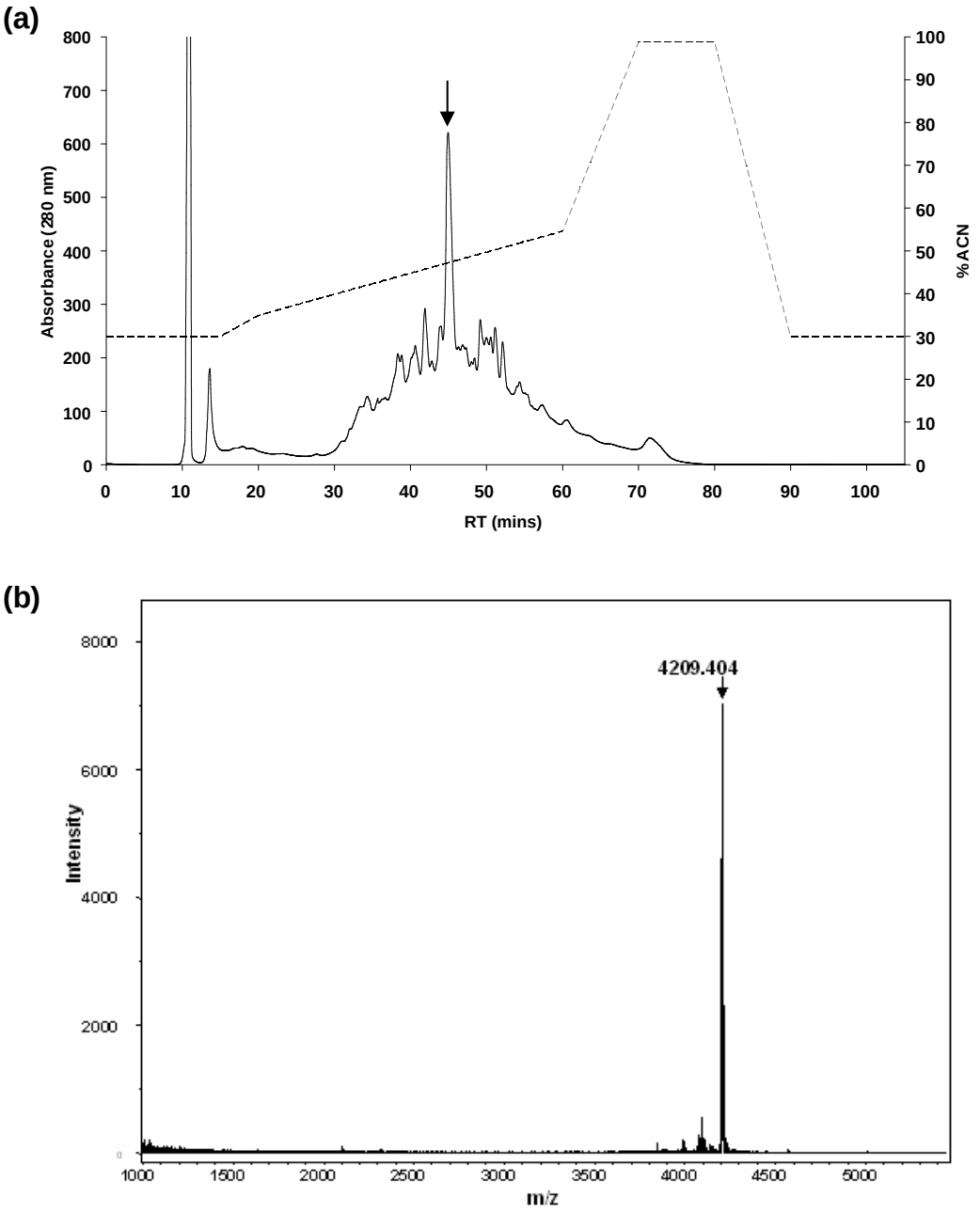


Figure 3.5. RP-HPLC Purification of Ii-Fl peptide

(a) The Ii-Fl peptide was purified by reverse phase HPLC (solid line) using a Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent, on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the Ii-Fl peptide is indicated. (b) Purity of pooled fractions was analysed using MALDI mass spectrometry. The major peak with a mass of 4209 Da corresponds to the expected mass for Ii-Fl peptide.

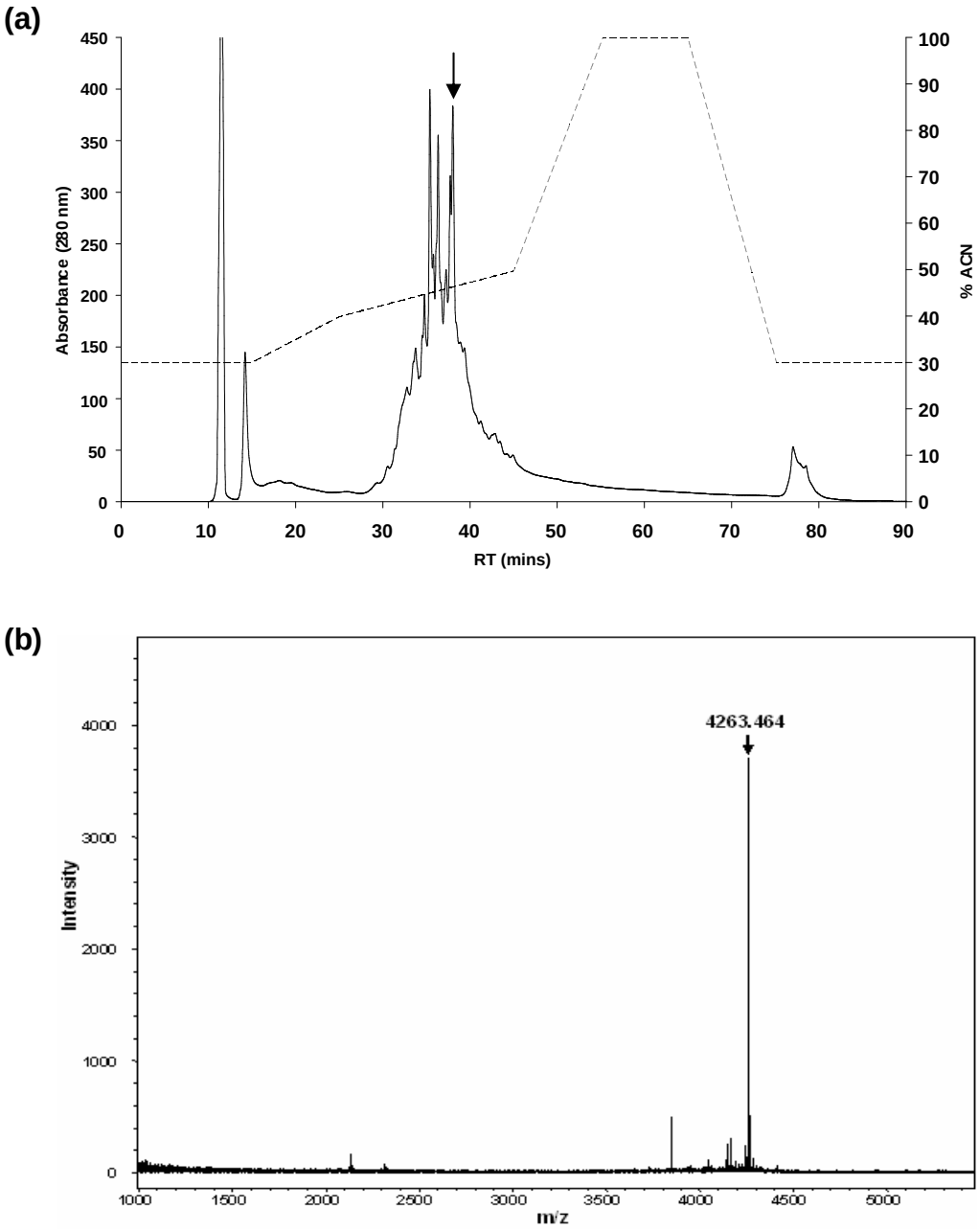


Figure 3.6. RP-HPLC purification of Ii-Rh peptide

The Ii-Rh peptide was purified by reverse phase HPLC (solid line) using a linear 1% per min Acetonitrile (ACN) gradient (broken line) with H_2O as the second solvent , on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the Ii-Rh peptide is indicated. (b) Purity of pooled fractions was analysed using MALDI mass spectrometry. The major peak with a mass of 4209 Da corresponds to the expected mass for Ii-Rh peptide.

3.4 Analysis of the secondary structure of Ii TM domain using CD spectroscopy

In this study we sought to determine the effect of the peptide:micelle molar ratio on the oligomeric state of the Ii TM peptide. Since it is well known that detergents can denature proteins, it was not known what effect (if any) this would have on the secondary structure of the peptide. To answer this question, circular dichroism (CD) spectroscopy was used to assess the secondary structure of the Ii TM peptide in the detergent DPC at varying peptide:micelle molar ratios. As shown in Figure 3.7a, the peptide was soluble at all concentrations of DPC tested, enabling CD spectra to be acquired for all peptide:micelle ratios, and in all cases data were truncated at 200 nm below which the absorbance was too high to give reliable data (high tension values above 600 volts (Figure 3.7b).

As shown in Figure 3.7a, negative maxima were observed at 208 and 222nm which is characteristic of the presence of α-helical secondary structure. Interestingly, greater signal was observed with increasing detergent concentration which is likely due to the increasing solubilisation of the peptide since the peptide concentration was constant for each measurement. As shown in Figure 3.7c, analysis of the CD spectra using the program CDSSTR (Johnson, 1999) revealed that the percentage of α-helical content increases as the peptide:micelle ratio is decreased (i.e. detergent concentration is increased), rising to a maximum of ~80% α-helix with a ratio of 1:1. This increase may reflect increasing solubilisation of the protein. Interestingly, the percentage then begins to decrease slightly as the ratio increases further. This possibly indicates that above 1:1 the DPC detergent is beginning to denature the peptide, although the decrease is slight.

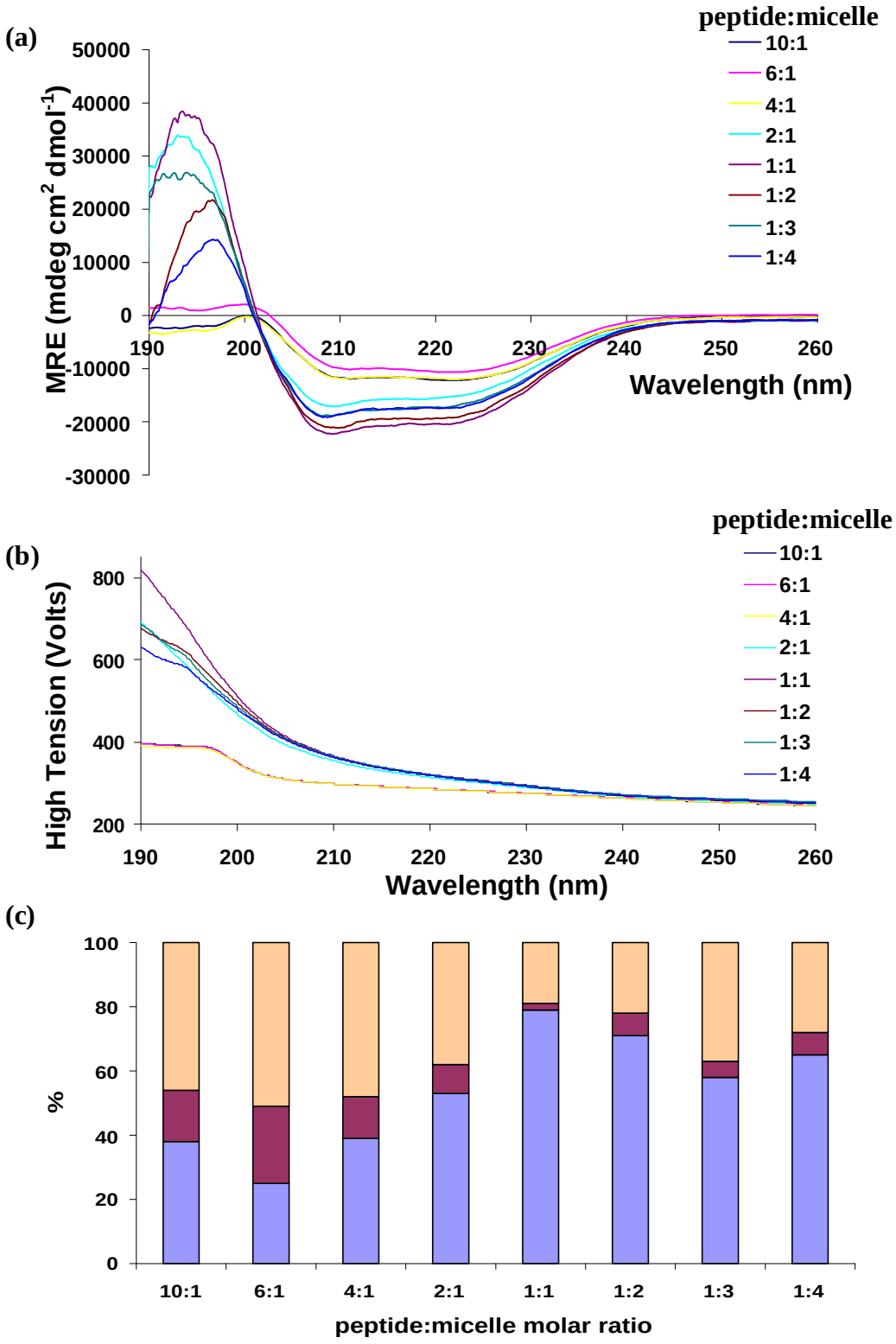


Figure 3.7. Analysis of the secondary structure of Ii TM peptide

(a) Circular Dichroism spectra of Ii TM peptide reconstituted into the detergent DPC at varying peptide: DPC micelle ratios as indicated. Mean residue ellipticity (MRE) was calculated from the measured ellipticity as described in Materials and Methods. (b) High tension for CD spectra, typically the CD data is taken to be reliable whilst this remains below 600. (c) Percentage secondary structure content (α-helix (blue), β-sheet (red), Random coil (yellow) at varying peptide: DPC micelle ratios as calculated from the CD spectra using CDSSTR (Johnson, 1999).

3.5 Cross-linking analyses of Ii TM domain self-association

The use of covalent cross-linking of TM domains in mild detergents enables the visualisation of oligomeric states using SDS-PAGE, which would otherwise be denatured in SDS detergent. Such analyses were performed in this study to confirm the oligomeric state of the Ii peptide. It is known that the concentration of detergent can affect the dissociation constant (k_d) and thus the oligomeric states of transmembrane interactions (Fisher, Engelman et al., 1999). Therefore the effect of varying the micelle concentration (i.e. micelle:peptide molar ratio) upon the oligomeric state of Ii peptide was explored here as it has not previously been considered with respect to the association of the Ii TM domain.

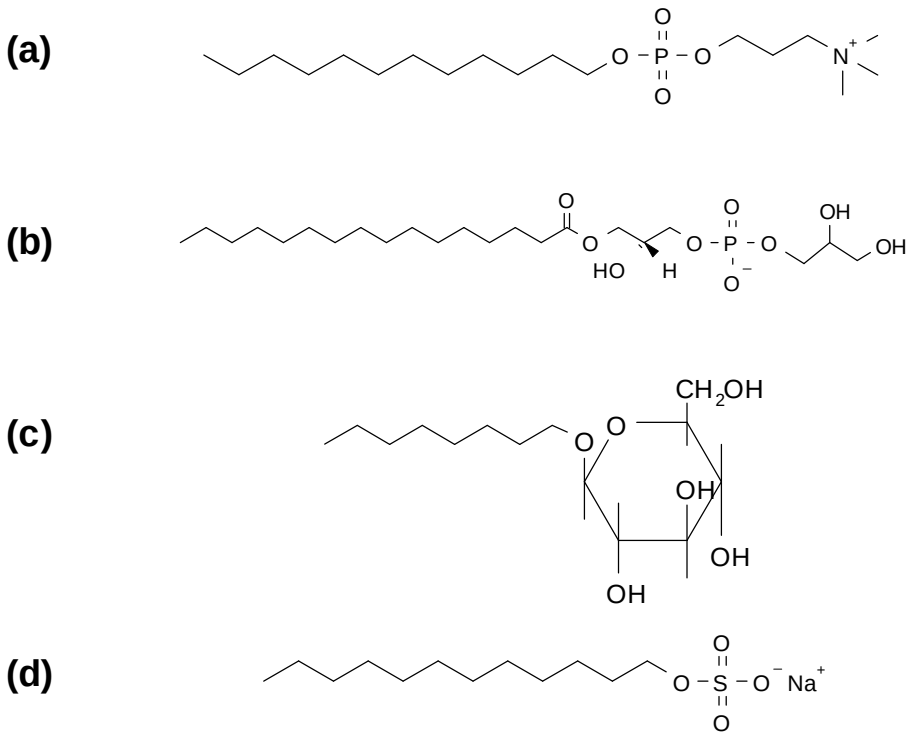


Figure 3.8. Structure of detergents used in cross-linking studies

Structure of the detergents (a) DPC, (b) LPPG, (c) OG (d) SDS, used in this study to solubilise the hydrophobic model TM peptides.

The Ii peptide was dissolved in varying micelle concentrations of the detergents octylglucoside (OG), 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)] (LPPG), dodecylphosphocholine (DPC), and sodium dodecylsulfate

(SDS) (Figure 3.8), cross-linked with Bis(sulfosuccinimidyl)suberate (BS^3) (Figure 3.9), and analysed using SDS-PAGE as described in Section 2.13. As shown in Figure 3.10, Ii self assembles in the detergents OG and LPPG and achieves higher order oligomeric states as the micelle concentration is decreased. The Ii peptide self assembles in DPC, but the trend of increasing oligomeric state with decreasing micelle concentration was not observed. It is possible that this could be due to the higher micelle concentrations used in this case, which are disrupting the higher order oligomers. Notably, the major oligomeric state observed using cross-linking is dimer, which contradicts earlier work which indicated the TM domain of Ii to be predominantly if not entirely a trimer (Dixon, Stanley et al., 2006).

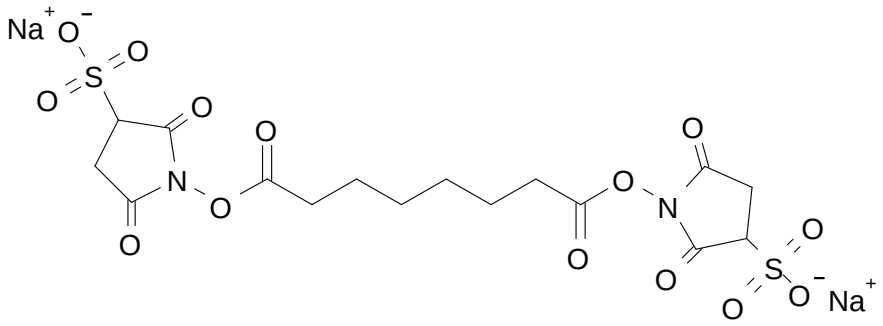


Figure 3.9. Structure of cross-linking agent BS^3

Structure of the cross-linking agent Bis(sulfosuccinimidyl)suberate (BS^3) (Pierce, UK) used in this study to monitor oligomeric states of model peptides in mild detergents by SDS-PAGE analyses.

Since the predominant oligomeric state observed was dimer it was hypothesised that the presence of only a single K residue in the sequence of the Ii peptide may be limiting the cross-linking of a trimer, since this residue possesses the only free amino group available to react with the cross-linking agent BS^3. In order to test this, a second peptide, Ii_K, was synthesised that possessed a greater number of terminal K residues (see Table 3.1) and purified as described in Section 3.3. It should be noted that the presence of additional K residues at the termini of TM domain peptides has been shown to have no affect on the association of strongly-associating TM domains (Melnyk, Partridge et al., 2003).

As shown in Figure 3.11, cross-linking of Ii_K was performed on the peptide solubilised in the detergents (a) LPPG and (b) DPC at varying peptide:micelle ratio. In LPPG (Figure 3.11a), Ii_K peptide seems to be forming oligomeric states from monomer to trimer in a detergent-concentration dependent manner similar to the Ii peptide. For the samples of Ii_K cross-linked in DPC, higher peptide:micelle ratios were explored to those used with Ii and as shown in Figure 3.11b, multiple oligomeric states are observed. With the higher peptide:micelle ratio the same pattern of decreasing oligomeric size as seen in the other detergents is observed. This result also shows that DPC is a better solubilising agent for the Ii peptide than the other detergents and on this basis was chosen as the detergent for use in other techniques.

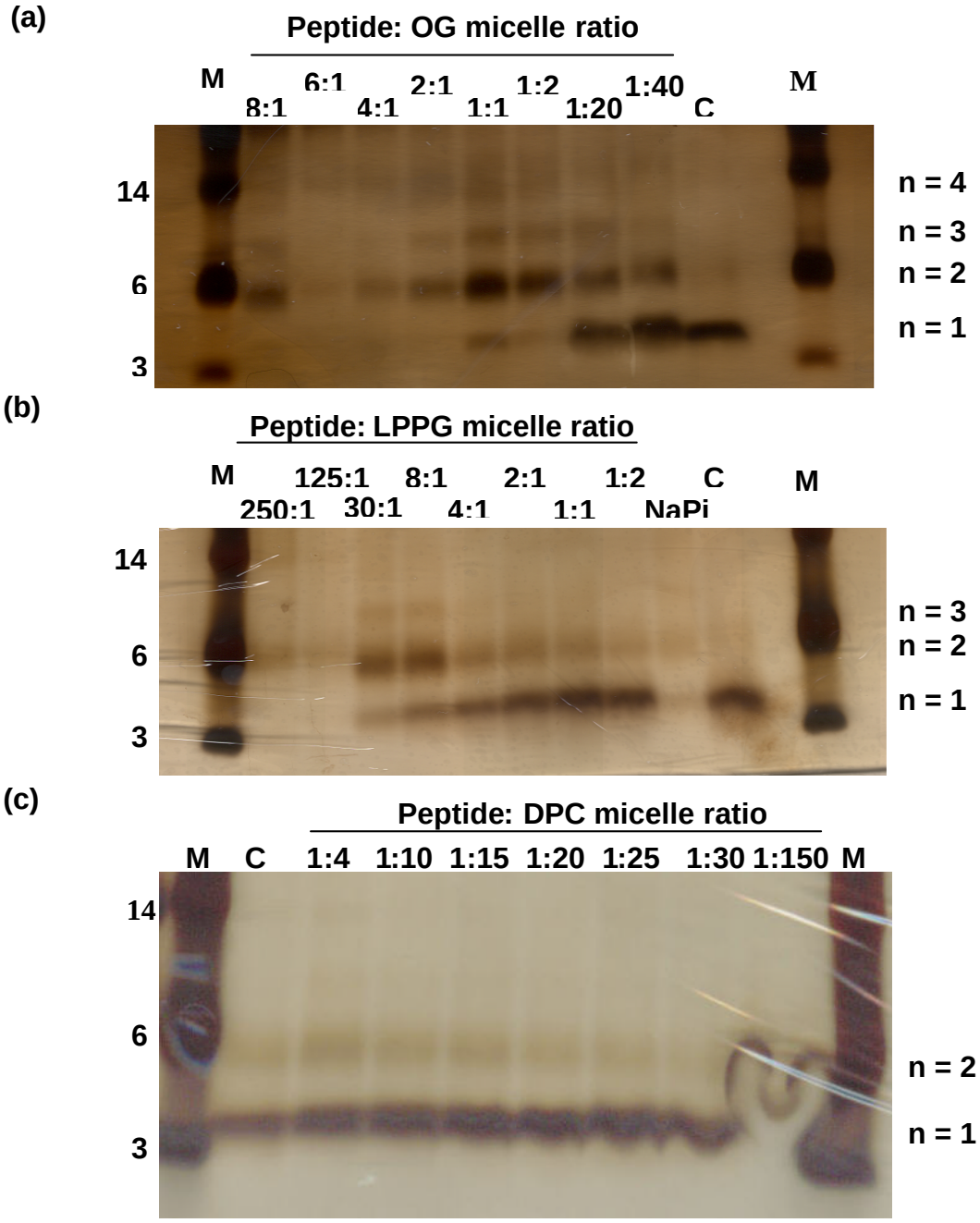


Figure 3.10. SDS PAGE analysis of cross-linked Ii peptides in DPC

Crosslinking of Ii at varying peptide:micelle ratios of the detergents: (a) Octylglucoside (OG). The cmc of OG was taken as 18 mM and the aggregation number as 90 for micelle:peptide calculations. (b) LPPG. The cmc of LPPG was taken as 18 μM and the aggregation number as 125 for micelle:peptide calculations. (c) DPC. The cmc of DPC was taken as 1000 μM and the aggregation number as 56 for micelle:peptide calculations. Peptide:micelle ratios were calculated as described in Section 2.13. Crosslinking agent was 1 mM BS3 supplied by Pierce. Lanes marked C are cross-linked Ii in 150 mM SDS used as a control for non-specific oligomerisation. Lanes marked with M are molecular weight standards. In all cases the concentration of Ii was 20 μM. Ii has a mass of 3.8 kDa. Possible oligomeric states are indicated where n is the stoichiometry.

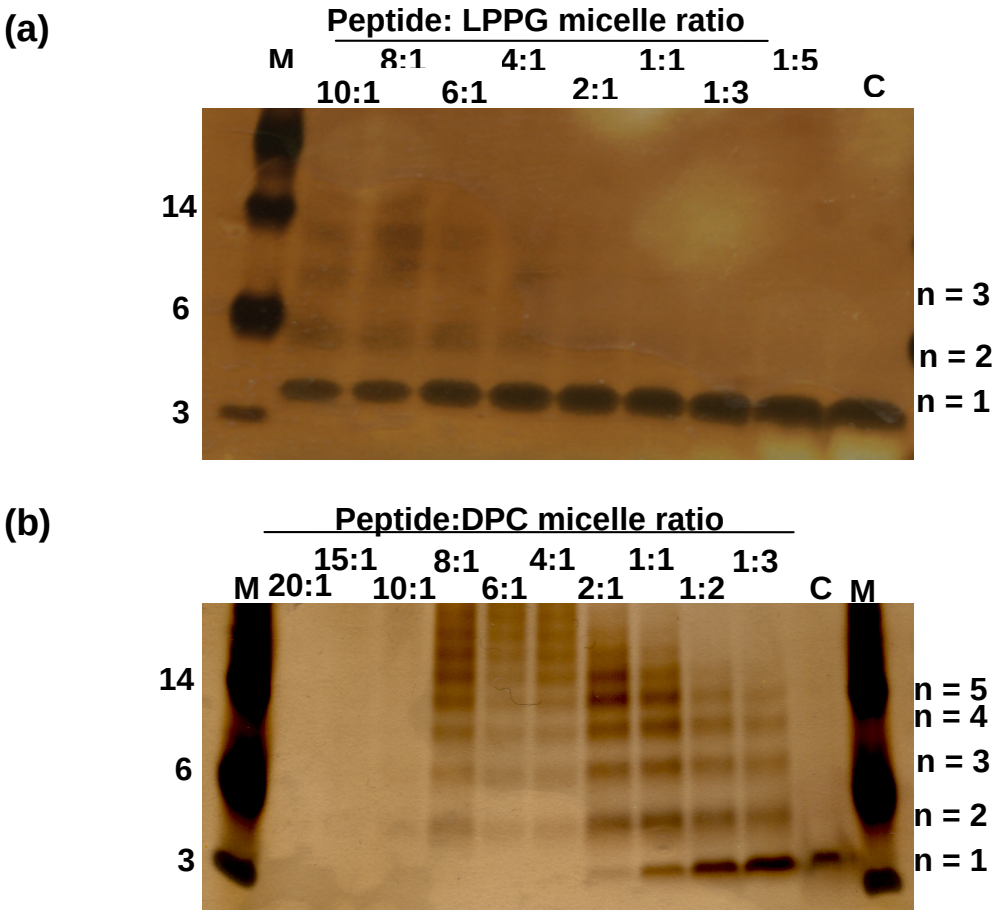


Figure 3.11. SDS PAGE analysis of cross-linked Ii_K peptides in DPC

Crosslinking of Ii_K at varying peptide:micelle ratios of the detergents: (a) LPPG. The cmc of LPPG was taken as 18 μM and the aggregation number as 125 for micelle:peptide calculations. (b) DPC. The cmc of DPC was taken as 1000 μM and the aggregation number as 56 for micelle:peptide calculations. Crosslinking agent was 1 mM BS3 supplied by Pierce. Peptide:micelle ratios were calculated as described in Section 2.13. Lanes marked C are cross-linked Ii_K in 150 mM SDS used as a control for non-specific oligomerisation. Lanes marked with M are molecular weight standards. In all cases the concentration of Ii was 20 μM. Ii_K has a mass of 3.89 kDa. Possible oligomeric states are indicated where n is the stoichiometry.

3.6 Assessing the oligomeric state of the Ii TM domain peptide using AUC

The lack of agreement between the cross-linking studies of the Ii TM domain in this study and those in the literature cast doubt on the oligomeric state of this peptide. In the study by Dixon et al, the Ii peptide was also analysed by sedimentation equilibrium analytical ultracentrifugation (AUC) where it was found that the data fit best to a monomer trimer model (Dixon, Stanley et al., 2006). (Dixon, Stanley et al., 2006). Therefore, in this work, sedimentation velocity experiments were performed in collaboration with Dr Andrew Beevers (University of Warwick, UK) in order to corroborate this finding as described in Section 2.23. Sedimentation velocity can be difficult with peptides of low molecular mass due to the high speeds necessary to minimize the rate of back diffusion. A speed of 60,000 rpm was required to provide data with the required mass resolution and this is in agreement with the previous study of Ii (Dixon, Stanley et al., 2006).

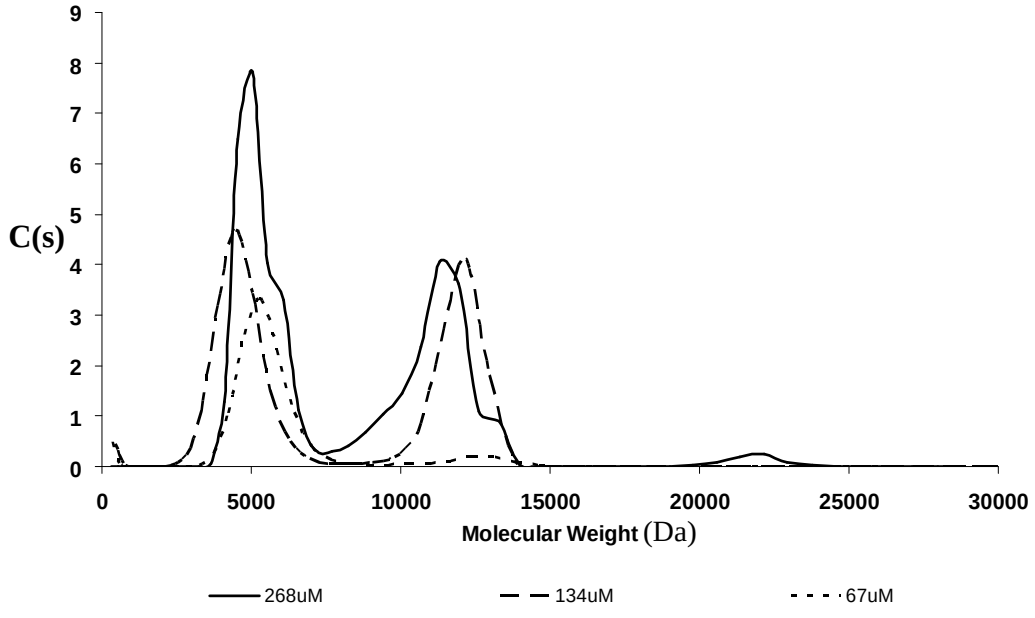


Figure 3.12. Sedimentation velocity analysis of Ii oligomeric state

Sedimentation velocity data obtained for Ii TM peptide in buffer containing 15 mM DPC and 52.5% D_2O. Sedimentation coefficient distribution profile was calculated using SEDFIT and converted to molecular mass. Monomer mass for Ii is 3808 Da. Data were collected and analyzed by Dr Andrew Beevers.

The molecular weights and distribution of the species observed for the Ii TM peptide dissolved in the detergent DPC are shown in Figure 3.12. The data show three distinct species in solution with molecular weights at approximately around 5000, 12000 and 22000 Da. Given that the monomer mass of Ii_{TM} peptide is 3808 Da, it is likely that the 5000 Da distribution corresponds to a mean value for the monomer-dimer species, whilst the distributions centred at 12,000 and 22,000 Da indicate the presence of trimer and hexamer, respectively. Again, like in the cross-linking data a shift in oligomeric state from higher to lower order is observed as the ratio of detergent to the peptide increases. This data agrees with that from cross-linking and GALLEX in this work, since they show the Ii peptide is self-associating and is forming oligomers from dimer to hexamer that are sensitive to the concentration of detergent.

3.7 FRET analyses of Ii TM domain self-association

As discussed in Section 2.19, Förster resonance energy transfer, or FRET analyses rely upon the phenomenon that energy can be transferred between molecules possessing overlapping emission and excitation wavelengths (termed FRET pairs) that are in close spatial proximity. The distance dependence of the energy transfer, known as the Förster distance, provides a tool for determining interactions between molecules. The FRET pairs that are often used in these studies are fluorophores (e.g. fluorescein and rhodamine) allowing the energy transfer to be monitored by fluorescence spectroscopy. This is done by monitoring either the decrease in donor emission or the increase in acceptor emission with the former approach being more commonly used.

There have been numerous examples of FRET studies in recent years that make use of this *in vitro* technique to measure TM domain interactions for model synthetic peptides solubilised in membrane mimetics (Duneau, Vegh et al., 2007). An advantage of this technique is that it enables hetero-association to be studied through the selective labelling of peptides with FRET pairs (Duneau, Vegh et al., 2007), it therefore providing a means for monitoring association between the TM domains in the Ii-MHC complex. Before proceeding to explore these hetero-associations the technique was first applied to the study of the self-association of Ii TM peptides in order to optimise the methodology.

The transfer of energy between the donor and acceptor can be measured via monitoring either the quenching of the donor emission or the enhancement of the acceptor emission. Typically, the donor quenching is monitored using fluorescence spectroscopy with spectra acquired at fixed excitation wavelength, which in the case of fluorescein labelled peptides was 439 nm, whilst the emission is monitored over the wavelength range 450 to 650 nm.

3.7.1 Design and synthesis of fluorophore labelled peptides

FRET pairs are known to vary in their Förster distances (i.e. the distance over which energy transfer can occur). Therefore, a donor and acceptor FRET pair was chosen that produce a FRET signal that would be independent of the manner in which the peptides associate (i.e. parallel versus antiparallel association and crossing angle) and would therefore report on the amount of oligomer present. Fluorescein and rhodamine are a commonly used FRET pair which can report over a distance of 40-90 Å and have been employed successfully in the study of other TM domain interactions (Li, You et al., 2005; You, Li et al., 2005) making them an appropriate choice for use in this study.

Peptides corresponding to the TM domain of Ii were synthesised, labelled with fluorescein (Ii-Fl) and rhodamine (Ii-Rh) at their N-termini and purified as described in Section 3.3. As shown in Figure 3.13, the excitation and emission spectra of the labelled peptides display properties of the fluorophores that are useful for FRET measurements. Specifically, the wavelengths over which fluorescein labelled Ii emits display considerable overlap with the excitation wavelengths of rhodamine labelled Ii. Advantageously, the excitation wavelengths of Ii-Fl display minimal overlap with that of Ii-Rh below around 450 nm. An excitation wavelength of 439 nm was chosen for use in FRET experiments, such that fluorescein could be selectively excited with minimal direct excitation of rhodamine.

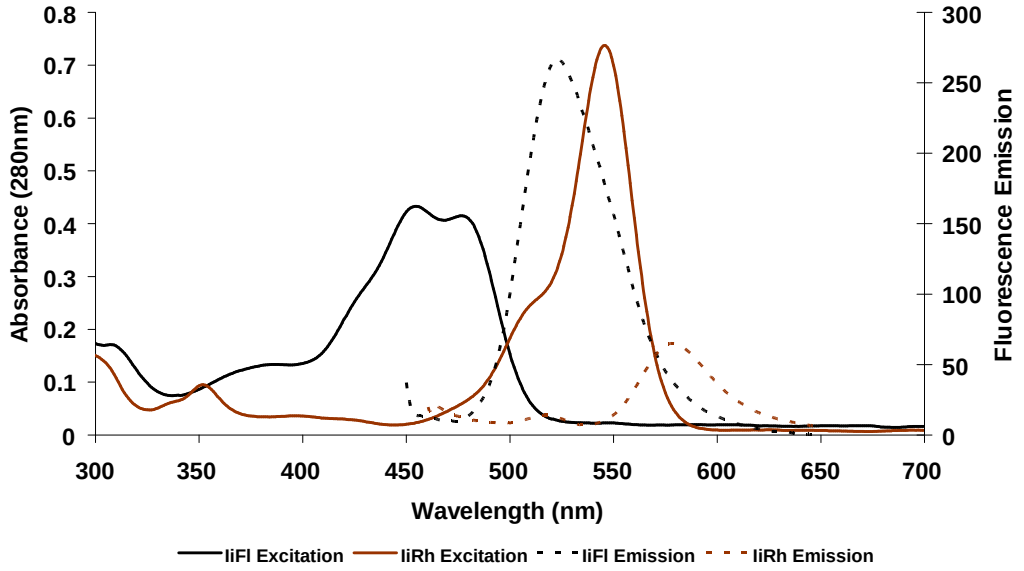


Figure 3.13. Absorbance and emission spectra for fluorescein and rhodamine labelled Ii peptide

Emission and absorbance spectra for Ii peptide labelled with fluorescein (Ii-Fl) and rhodamine (Ii-Rh) were collected at peptide concentrations of 4 μM. The emission spectra for Ii-Fl and Ii-Rh were collected using excitation wavelengths of 439 nm and 540 nm, respectively. All peptides were solubilised in TFE.

3.7.2 Optimising sample preparation for FRET experiments

Detergents have been successfully employed in FRET experiments to determine interactions between TM domain peptides (Fisher, Engelman et al., 1999; Fisher, Engelman et al., 2003; Li, You et al., 2005; Duneau, Vegh et al., 2007). Since previous experiments on the Ii TM domain described in the preceding chapters have revealed the propensity of this sequence to oligomerise in mild detergents such as DPC, it seemed logical to conduct the FRET experiments using this detergent. Since it is believed that TM peptides are in fast exchange between detergent micelles in solution it was assumed that dissolving the donor and acceptor peptides in buffered DPC solutions prior to mixing to produce the 'FRET' sample would yield a FRET signal if the peptides could indeed freely exchange between micelles enabling energy transfer to occur (assuming that the peptides interact). As shown in Figure 3.14, mixtures of Ii-Fl and Ii-Rh peptides separately solubilised in the detergent DPC display a low energy transfer that fluctuates considerably around a mean value of 7 % as the peptide:micelle molar ratio is changed. This data indicates that no significant energy transfer is

occurring. The samples were left for a further three days at room temperature before taking repeat measurements but no further improvement in signal was observed. Since other experiments present here suggest Ii can self-associate we would have expected to observe a FRET signal.

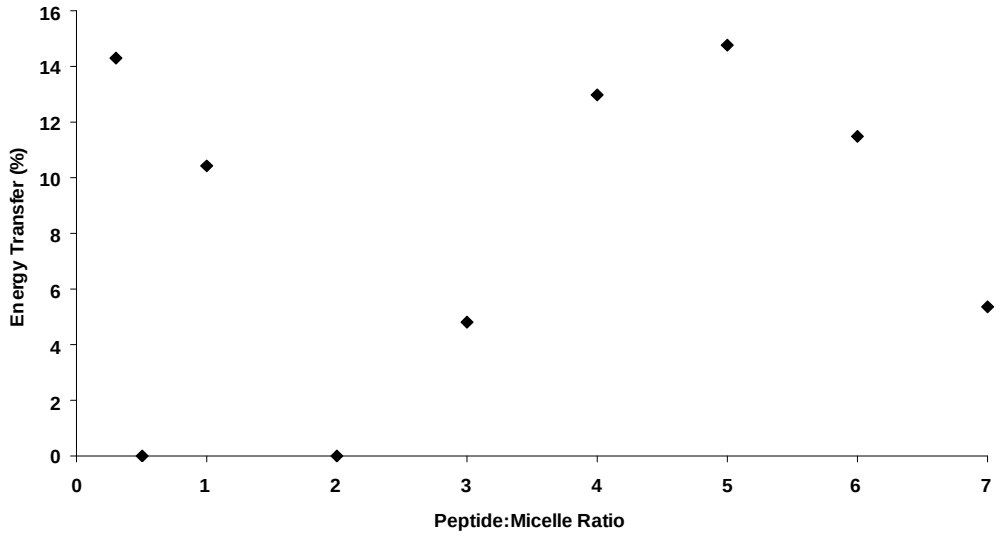


Figure 3.14. Exploring preparation of FRET samples in detergent

Energy transfer for mixtures of Ii-Fl and Ii-RH peptides pre-dissolved in DPC buffer solution and subsequently mixed. The DPC concentration was adjusted to provide varying peptide:micelle ratios whilst keeping the individual peptide concentration at 2 μM and total concentration at 4 μM. Energy transfer was calculated as described in Section 2.19.6.

It was considered that the sample preparation method may be restricting the observation of FRET signal and that the peptides were not rapidly exchanging between the micelles. To test this hypothesis a sample preparation method was employed where the peptides and detergent were first co-dissolved in TFE, followed by lyophilisation to remove the organic solvent and subsequent re-dissolving of the peptide-detergent film in aqueous buffer. This method had been used previously in the cross-linking studies shown above. It was believed that this would allow for the random mixing of the donor and acceptor peptides upon formation of the micelles thus enabling a FRET signal to be observed.

This method was carried out for Ii-Fl and Ii-Rh peptides. The peptides and DPC were dissolved in TFE separately and then mixed before removing the solvent and

reconstituting in aqueous buffer. As shown in Figure 3.15 this co-dissolving method resulted in observation of a significant FRET signal, as indicated by the decrease in donor emission at 520 nm and increase in acceptor emission at 570 nm. The experiments were initially carried out at a peptide:micelle molar ratio of 3:1, which was chosen on the basis that oligomers of the Ii TM domain were observed in cross-linking experiments at this ratio. This data provides further evidence of the ability of the Ii TM domain to self-associate and indicates that the co-dissolving method of sample preparation is suitable for FRET analysis.

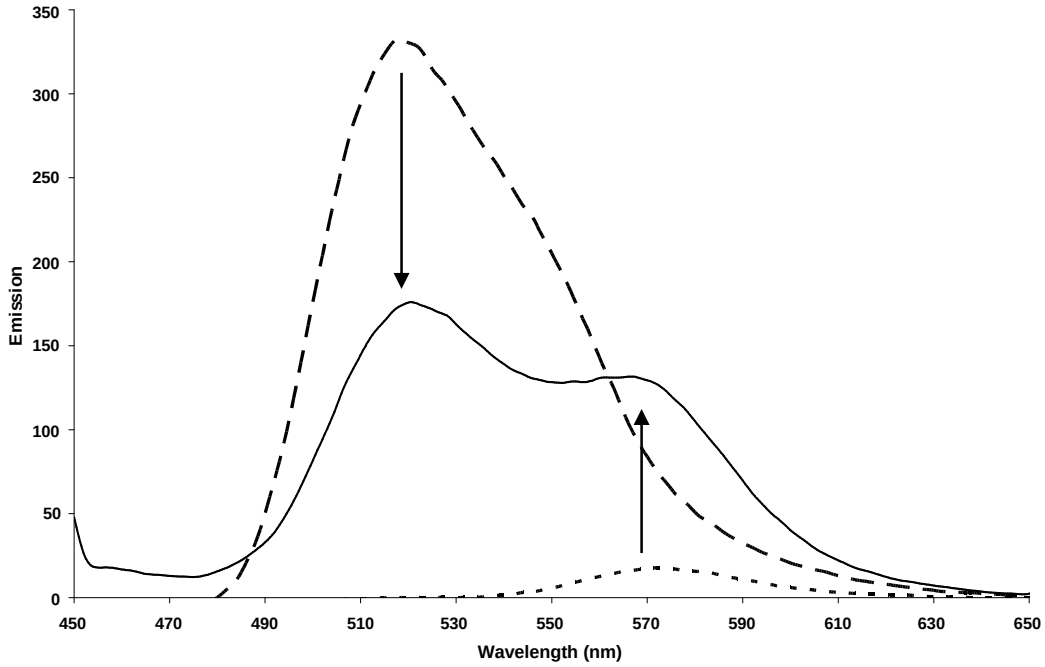


Figure 3.15. FRET spectra for Ii samples prepared using co-dissolving method

FRET analyses of Ii-Fl and Ii-Rh peptides in DPC micelles at a peptide:micelle molar ratio of 3:1. All samples were prepared using the co-dissolving method where peptides and DPC were first mixed in TFE prior to lyophilisation and re-suspension in buffer solution. The broken line is the spectrum given by Ii-Fl, the dotted line is the spectrum observed for Ii-Rh whilst the solid line is the spectrum for the mixture of Ii-Fl and Ii-Rh. FRET signal is evident from the decrease in the donor emission at 520 nm and an increase in acceptor emission at 570 nm as indicated by the arrows.

3.7.3 Effect of the peptide:micelle molar ratio on the association of the Ii TM domain

Since it has been shown in this study that the oligomeric state of the Ii TM domain peptide can be modulated by detergent concentration, the dependency of the FRET signal on the peptide:micelle molar ratio was explored. FRET samples of Ii-Fl and Ii-Rh were prepared using the co-dissolving method, and the peptide:micelle ratio was varied between 20:1 and 1:3 whilst keeping the total peptide concentration constant. For each ratio, the energy transfer was calculated as described in Section 2.19.6 and plotted versus the peptide:micelle molar ratio.

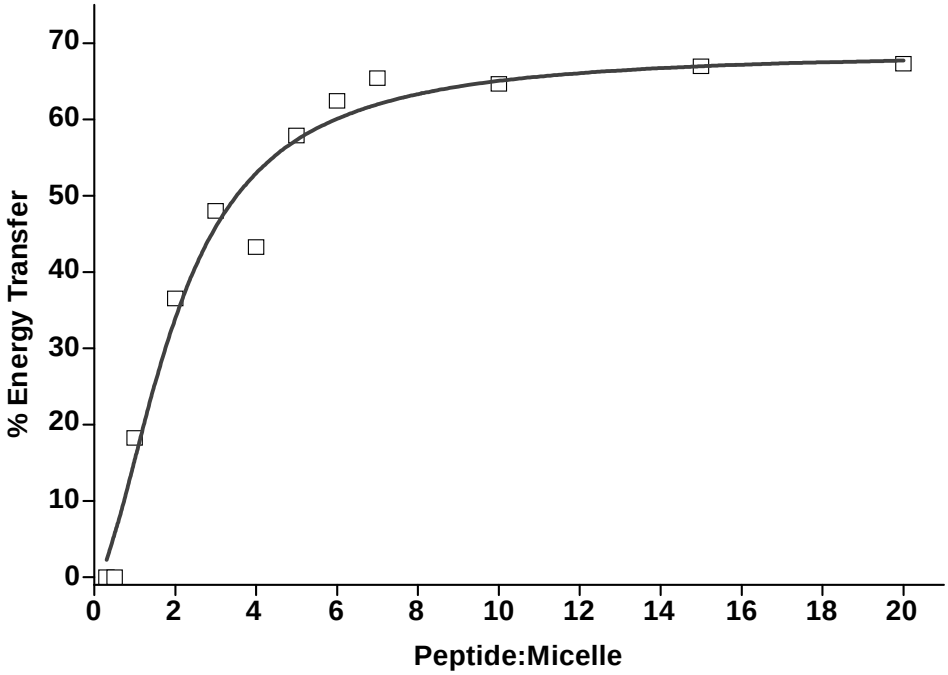


Figure 3.16. Dependency of energy transfer on the peptide:micelle ratio

Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM Ii-Fl, 2 μM Ii-Rh) while the detergent concentration was varied. Samples were prepared using the co-dissolving method by mixing peptide and DPC pre-solubilised in TFE. A CMC of 1 mM was used in calculations of the peptide:micelle molar ratio. Excitation spectra were collected and energy transfer was calculated as described in Materials and Methods.

As shown in Figure 3.16, the energy transfer displayed a strong dependency on the peptide:micelle molar ratio, with the highest energy transfer occurring at higher ratios and reaching a maximum value of 65%. Since the energy transfer is

directly related to the formation of oligomers, this indicates a shift in oligomeric state from monomer to higher order oligomers as the ratio is increased. This is in keeping with the known effect of detergents upon oligomeric state reported in the literature and also with the cross-linking data on for Ii TM peptides in this study. This data has implications for further FRET experiments which can be performed to identify oligomeric state and thermodynamic properties since these experiments must be performed at a constant detergent concentration.

3.7.4 Determining specificity of the Ii TM domain interaction

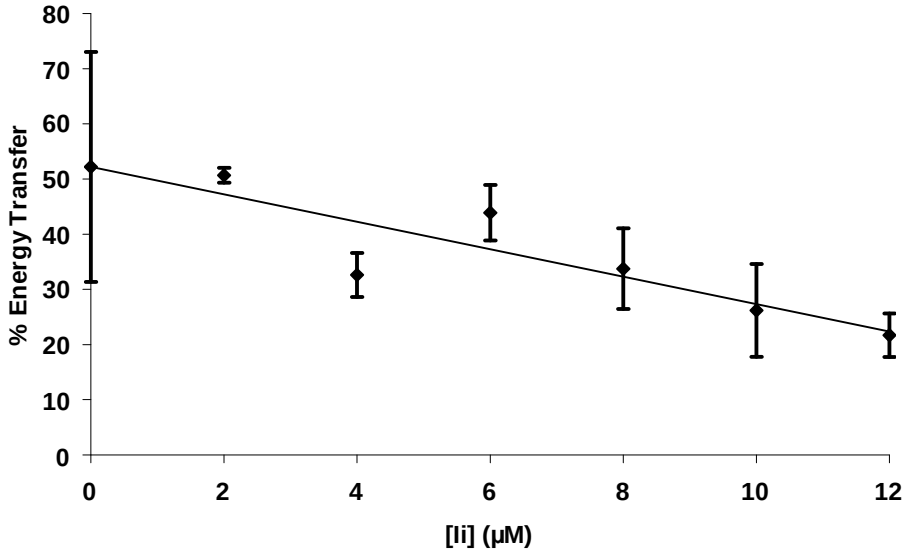


Figure 3.17. Effect of unlabelled Ii peptide on FRET of Ii-Fl to Ii-Rh

Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM Ii-Fl, 2 μM Ii-Rh) while the concentration of unlabelled Ii was varied. Samples were prepared by co-dissolving all peptides and DPC dissolved in TFE. Experiments were performed at a peptide:micelle ratio of 10:1. The reduced FRET efficiency suggests that sequence-specific oligomerisation contributes to the measured FRET efficiency.

It has been reported that in order to ascertain if the measured FRET signal for Ii-Fl and Ii-Rh arises from a specific interaction, and is not due to the peptides merely occupying the same micelle and thus being in close proximity, it is necessary to conduct a titration with unlabelled peptide (Fisher, Engelman et al., 1999). Since this experiment must be carried out at a fixed detergent concentration the peptide:micelle ratio of 10:1 was chosen on the basis of the results described in the preceding section, where significant oligomerisation was

observed at this ratio. The concentrations of Ii-Fl and Ii-Rh peptides were kept constant whilst varying the concentration of unlabelled Ii peptide. As shown in Figure 3.17, the energy transfer decreases from ~60% to ~20% with increasing concentration of unlabelled peptide. This is indicative of the unlabelled peptide disrupting the formation of donor and acceptor partners.

3.7.5 Determining the oligomeric state of the Ii TM domain

In FRET experiments, the stoichiometry of donors and acceptors generating the observed FRET signal (and thus the oligomeric state of the complex) can be determined by measuring the energy transfer as a function of the donor acceptor ratio whilst keeping the total peptide and detergent concentrations constant (Veatch and Stryer, 1977). Since it has been shown by cross-linking that the oligomeric state of Ii TM peptide is modulated by the peptide:micelle molar ratio, this experiment was performed for Ii TM domain at a range of detergent concentrations whilst keeping the peptide concentrations constant.

For Ii TM domain, the mole fraction of the acceptor peptide Ii-Rh to donor peptide Ii-Fl was varied between 0 and 1 whilst keeping the total peptide concentration constant at 4 μM. The energy transfer was measured and the ratios of emission at 520 nm in the donor only sample (Q0) to that in the FRET sample (Q) were calculated. The value of Q0 was normalised to the mole fraction of donor present in the FRET samples.

The value of Q/Q0 for varying peptide:micelle molar ratios is plotted in Figure 3.18, as are the predicted lines for various oligomeric states calculated according to Equation 5 in Section 2.19.6. At a ratio of 1:3 (Figure 3.18a) the data does not fit well to any of the proposed models, and is therefore difficult to interpret but could suggest Ii is primarily monomeric since there is no significant effect on Q/Q0 with increasing acceptor. At a ratio of 2:1 (Figure 3.18b) the data fits best to trimer and tetramer with reduced CHI^2 values of 2.41 and 2.27, respectively, and FRET efficiency values of 53 % and 48 %, respectively. At a ratio of 4:1 (Figure 3.18c) the data fits best to trimer with a reduced CHI^2 value of 1.25 and a FRET

efficiency value of 50 %. At a ratio of 10:1 (Figure 3.18c) the data fits best to tetramer with CHI^2 value of 8.75 and a FRET efficiency of 66 %.

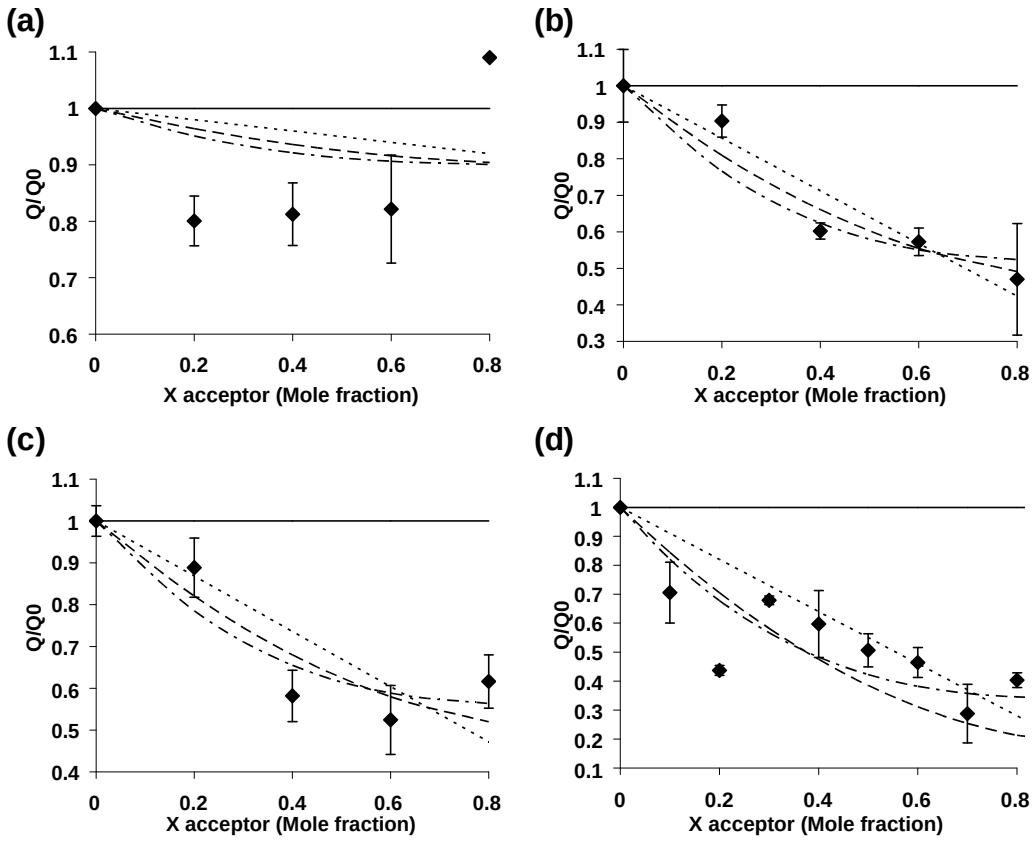


Figure 3.18. Determining oligomeric state of Ii TM domain association

Stoichiometry of Ii-Fl and Ii-Rh association in DPC detergent at peptide:micelle ratios of (a) 1:3 (b) 2:1 (c) 4:1 (d) 10:1. The energy transfer was measured and the ratios of emission at 520 nm in the donor only sample (Q0) to that in the FRET sample (Q) were calculated. The value of Q0 was normalised to the mole fraction of donor present in the FRET samples. The ratio of Ii-Fl to Ii-Rh was varied between 0.2 and 1.0 whilst keeping the total peptide concentration constant at 4 μM and the peptide: Calculated curves for monomer (solid), dimer (dotted), trimer (broken), and tetramer (broken dotted) are shown and were calculated using Equation 5 as described in Section 2.19.6. The goodness-of-fit for the experimental data to the calculated curves was determined using a standard reduced CHI^2 curve fitting procedure.

These data strongly suggest that the Ii TM domain has a propensity to oligomerise in DPC micelles and to adopt an oligomeric state that is detergent concentration dependent. Interestingly, the FRET data suggests Ii TM domain adopts a trimeric oligomeric state along with a tetrameric state. Notably, there are issues with the reproducibility of the methodology leading to significant errors and outliers in the data sets which confound the issue of conclusively assigning an oligomeric state,

although this problem can be negated to some extent through the use of statistical analysis such as CHI^2 curve fitting as used in this study.

3.7.6 FRET analysis of Ii TM domain self-association in lipid bilayers

Although detergent micelles are widely used and are convenient for analysing TM domain interactions, they can be considered to be poor membrane mimetics due to their high degree of curvature relative to planar bilayers. The forces applied to the peptides by this curvature may disrupt the native folding. Lipid vesicles or liposomes possessing a bilayer similar to a native membrane can be formed from lipid molecules. It has been reported that FRET analysis can be performed on peptides in such systems (Merzlyakov, You et al., 2006). The use of lipids is therefore a natural progression from our studies performed in detergent micelles.

FRET analyses in lipid vesicles was performed after reconstituting the Ii-Fl and Ii-Rh peptides into multilamellar vesicles (MLVs) composed of DMPC lipids as described in Section 2.14. As shown in Figure 3.19, a possible FRET signal was observed between Ii-Fl and Ii-Rh peptides as indicated by the decrease in emission of the donor signal. However, the acceptor showed only a minimal increase in emission. This suggests that the Ii TM peptides are self-associating in these artificial bilayers, but is not conclusive due to the lack of a concomitant increase in the emission of the acceptor. Further work will need to be performed to discern if this is a real FRET signal. An improvement to the method may be the formation of unilamellar vesicles by extrusion of the MLVs, which will place all peptides in the same membrane and allow greater self association, since in the MLVs the peptides may be distributed unevenly between the many bilayers forming the MLV.

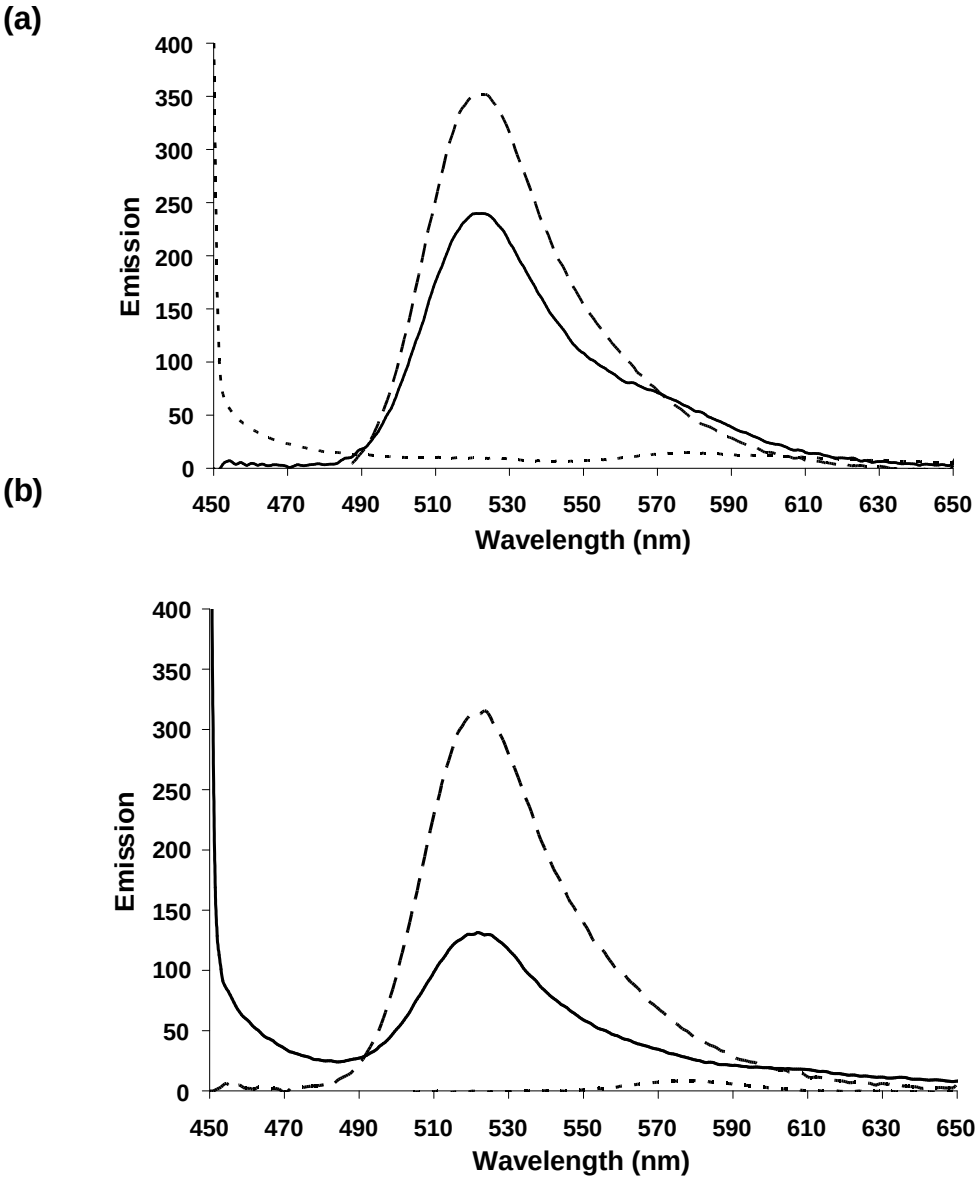


Figure 3.19. FRET analysis of Ii TM domain association in lipid bilayers.

Fluorescence emission spectra of fluorescein (Fl) and rhodamine (Rh) labelled Ii TM peptides in DMPC vesicles. Continuous line 0.1 mol% Ii-Fl and 0.1 .mol% Ii-Rh. Broken line 0.1 mol% Ii-Fl. Dotted line 0.1 mol% Ii-Rh. Labelling yield was f_d= 1.0 and f_a = 0.9 for the donor and acceptor, respectively. The total peptide concentration was 0.25 mol%. In the experiments, labelled peptides were co-dissolved with DMPC lipids in TFE. The solvent was evaporated, and the samples were hydrated and freeze-thawed four times to achieve equilibrium, as described in methods. The excitation wavelength was fixed at 439 nm, such that only the Fl was directly excited whilst the emission was scanned from 450 to 650 nm. The FRET efficiency was calculated from the decrease in Fl fluorescence at 520 nm (equation 1). (a) 1 mg/mL (b) 0.25 mg/mL

3.8 Observing insertion of Ii TM peptide into lipid bilayers using Oriented CD

For peptides solubilised in detergent micelles we expect that the detergent monomers aggregate in a 'donut-like' manner shielding the hydrophobic region of the peptides from the aqueous environment, so the orientation of the TM in the micelle is of little consequence (le Maire, Champeil et al., 2000). However, for studies carried out in liposomes, such as the FRET experiment just described, the orientation is of considerable importance so it must be established that the peptides are inserted perpendicular to the membrane normal i.e. that the peptides are membrane spanning and not simply associated with the membrane surface.

The use of oriented CD (OCD) has been reported in the literature as an appropriate technique for confirming the insertion of TM peptides in lipid bilayers (You, Li et al., 2005). It has been shown that the OCD spectra of helices that are parallel and perpendicular to the plane of the bilayer are significantly different (Wu, Huang et al., 1990). For helices oriented parallel to the plane of the bilayer the OCD spectra have two minima at 205 and 225 nm and a maximum at around 192 nm. However, helices that are perpendicular to the plane (i.e. span the membrane) display a minima and maxima at around 230 and 200 nm, respectively. This results from the $\pi \rightarrow \pi^*$ transition component at ~208 nm being polarised parallel to the helical axis. Thus, light propagating at an angle perpendicular to the lipid bilayer but parallel to the axis of the helix is not absorbed eliminating the 205 nm minima.

This technique was employed to analyse the insertion of Ii TM peptide in DMPC bilayers. A solution of lipid and peptide in TFE was deposited on a quartz slide generating a multilamellar film consisting of aligned lipid bilayers. The slide was placed into the spectropolarimeter such that the incident light propagates perpendicular to the film. A lack of sample homogeneity can be averaged out by rotating the sample through 360°, and acquiring CD spectra at 45° increments. Spectra were then averaged and baseline corrected to give the OCD spectrum shown in Figure 3.20a. The CD spectra of Ii-Fl and Ii-Rh in DMPC bilayers are

shown in Figure 3.20b, and display the two minima at 208 and 222 nm characteristic of an α-helical peptide. Comparison of the OCD and CD spectra reveal the lack of a 208 nm minima in the OCD spectra which indicates that the peptide is inserted in the lipid bilayer.

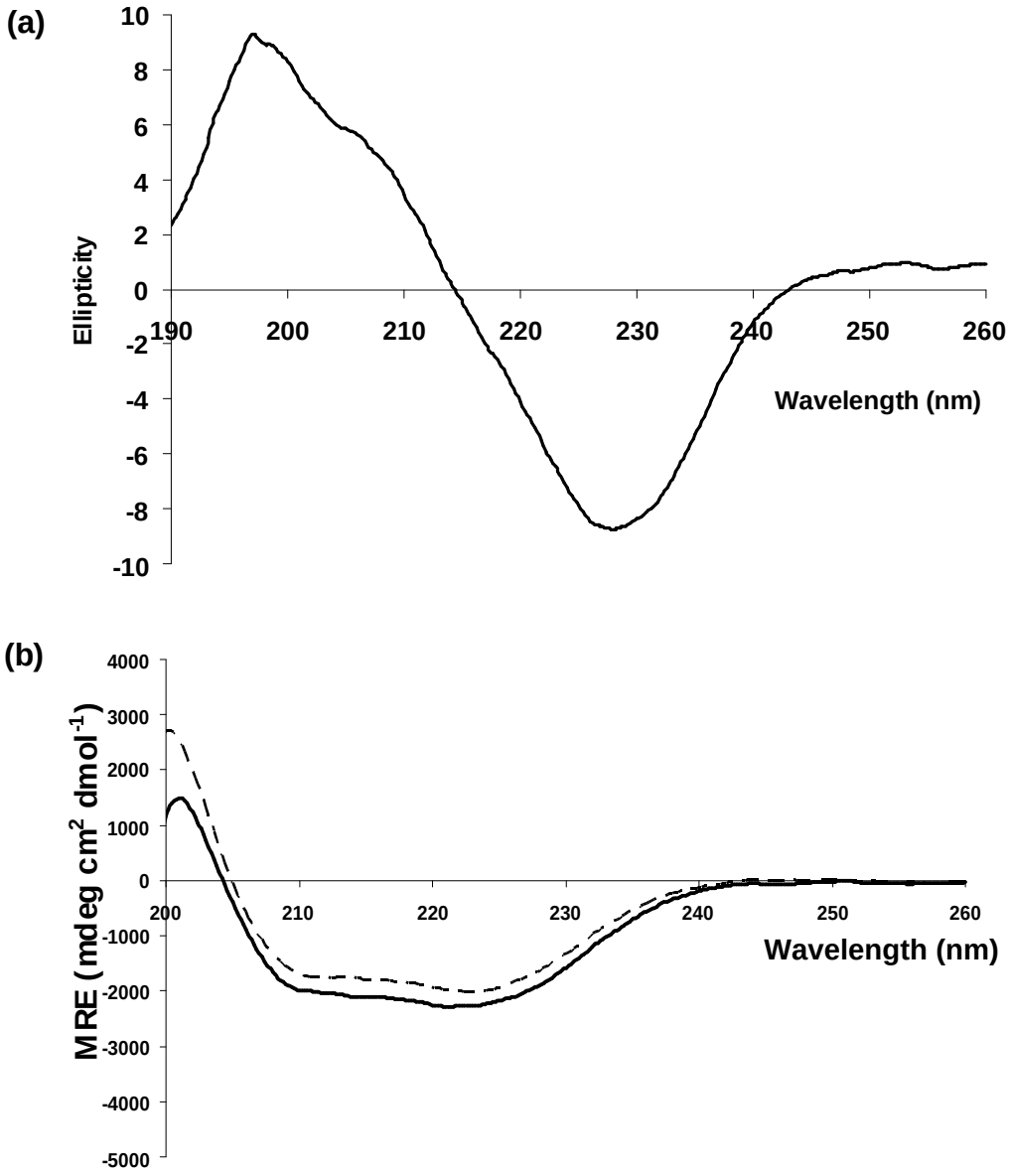


Figure 3.20. Oriented CD and CD spectra of Ii in DMPC liposomes

(a) Oriented Circular Dichroism spectrum of Ii TM domain in oriented DMPC bilayers. The samples were oriented multilayers on a quartz slide, deposited from the organic solvent, TFE. The multilayers were placed in the spectropolarimeter perpendicular to the optical path. The sample was rotated in increments of 45°, and spectra were collected and averaged. This spectrum is indicative of a helix that is spanning a membrane. (b) CD spectrum of Ii-Fl (solid line) and Ii-Rh (broken line) in DMPC lipid bilayers.

3.9 Conclusions on self-association of Ii TM domain

The full length Ii protein is known to be trimeric and previous studies on the TM domain of Ii in isolation suggest it too can oligomerise to form a strongly interacting trimer (Dixon, Stanley et al., 2006). In order to progress to studying the interactions of the Ii TM domain with the TM domains of MHC, we sought to confirm the self-association of Ii and its oligomeric state in addition to using this protein as a test-bed for developing methodologies.

Use of the in vivo assay GALLEX confirmed that Ii self-associates, corroborating reported results with the alternative assay TOXCAT. An advantage of the GALLEX assay over the TOXCAT assay is that it allows hetero-association of TM domains to be explored. Therefore, the plasmid constructs created for the studies in this section can be further utilised for exploring hetero-association between the TM domain of Ii and those of MHCα and MHCβ.

A peptide corresponding to TM domain of Ii was synthesised and purified for *in vitro* studies. Ii TM domain does not form SDS stable oligomers, therefore it was necessary to use cross-linking in conjunction with SDS PAGE to analyse the oligomeric state. The cross-linking results show that this peptide is capable of assembling to form a range higher order oligomeric states with dimer being the most abundant. The lack of a specific oligomeric state for Ii TM peptide is in conflict with the literature where a specific trimeric oligomer was observed (Dixon, Stanley et al., 2006). Interestingly, in the previous study the sequence contained only one lysine residue providing only one cross-linkable group.

This lead to a question of how a trimer may be cross-linked if this is the case? A further Ii peptide was designed with additional K residues at the N- and C-termini. Cross-linking of this peptide reveals the same pattern with increased dimer formation. Therefore from cross-linking results we conclude that it was not possible to establish a dominant oligomeric state for Ii TM domain only that it can self-assemble. Furthermore, this study shows using cross-linking that the detergent concentration or more specifically the peptide:micelle ratio is a strong determinant of the observed oligomeric state. This is the first time that this has been shown for Ii. The ratios presented here represent relatively small changes in

detergent concentration highlighting the importance for careful sample preparation.

Using CD spectroscopy this study reveals that the peptide:micelle ratio as well as affecting the oligomeric state also affects the secondary structure of Ii TM peptide. This is not something that has been widely reported in the literature and indeed for the well characterised TM domain of GpA the detergent concentration was found to modulate only its oligomeric state and not its helicity.

The use of fluorophore labelled Ii TM peptides provided a test case for developing the FRET assays that will be used later in the study of hetero-association. From a technical aspect there were concerns about adding a large hydrophobic group in the form of the fluorophore to an already very hydrophobic peptide sequence and the impact this would have upon purification. Using RP-HPLC and standard methods a 100% level of purity for Ii-Fl was obtained whilst Ii-Rh labelled peptide contained around 10% unlabelled peptide. The FRET analyses showed that the Ii TM domain is capable of homo-oligomerisation and that the energy transfer and oligomeric state were dependent on the peptide:micelle ratio further testifying to the importance of this parameter in the study of these domains. The oligomeric state of the peptide was explored using FRET analyses, and suggested that the Ii TM peptide could self-assemble to form trimeric and/or tetrameric states depending on the detergent concentration in the form of the peptide:micelle molar ratio. However, it is important to note that the results from the FRET experiments do display significant variation which hinders the assignment of a specific oligomeric state to the TM domain of Ii.

In summary, there is some evidence from this study that in isolation the TM domain of Ii indeed self assembles to form a homo-trimer, as has been reported in the literature (Dixon, Stanley et al., 2006). However, FRET and cross-linking analyses suggest it also assembles into additional oligomeric states (i.e. tetramer and higher), with the relative proportions of which are highly dependent on the detergent concentrations in the form of the peptide:micelle ratio. One explanation for the problem of assigning a definitive oligomeric state to the TM domain of Ii may be because that the helix-helix interactions are mediated by H-bonds. H-bonding between TM helices is thought to have weak specificity relative to Van

der Waals interactions and are believed to trigger non-specific aggregation (White, 2006). It has been proposed that van der Waals interactions in close packing helices are the main determinants for TM helix association and that H-bonds serve to stabilize a preformed oligomer (Schneider, 2004). If this is the case then it suggests that although Ii TM domain can self-associate it is not the driving force behind assembly of the full-length Ii trimer and instead plays a secondary role to that of the luminal trimerisation domain.

4 TM domain interactions of MHC Class II α and β chains

4.1 Introduction and objectives

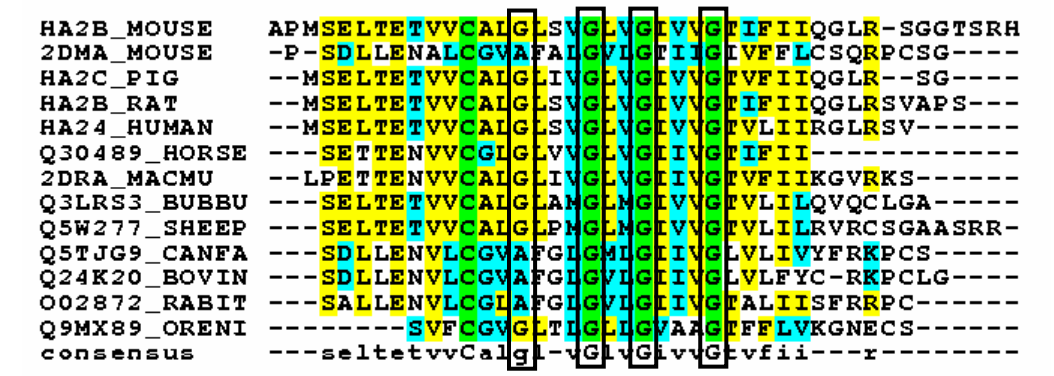
As discussed in Section 1.3.1, Major Histocompatibility Complex Class II proteins (MHC) are heterodimeric membrane proteins composed of non-covalently linked α and β chains. MHC present peptides derived from an invading pathogen to T cells, triggering an immune response to that pathogen. The crystal structure of the soluble extracellular domain of MHC (HLA-DR1) has been solved confirming it can oligomerise to form dimers which can also associate to form tetramers (Stern, Brown et al., 1994; Schafer, Malapati et al., 1998). MHC α- and β-chains are known to bind to the MHC Class II associated Invariant Chain (Ii) through an interaction mediated by their respective soluble domains. However, the TM domains of MHC and Ii have also been implicated in this association (Ashman and Miller, 1999; Castellino, Han et al., 2001; Barabanova, Kang et al., 2004). This section describes the results from studies into the self- and hetero-associations of the TM domains of the α- and β-chains of MHC. The objectives of the work presented in this chapter are to determine if the TM domains of α (MHCα) and β (MHCβ) can self-associate to form homo-oligomers, if they can associate with one another to form hetero-oligomers; and should they display helix-helix interactions, identify the residues involved.

4.2 MHC TM domains display conserved dimerisation motifs

As discussed in Section 1.2.1, TM domains have been found to display sequence motifs that are indicative of helix-helix interactions. Therefore, multiple sequence alignments of the putative TM domains of MHCα and MHCβ were performed using ClustalW2 (Larkin, Blackshields et al., 2007), to identify conserved residues, which would suggest these residues are functionally important. As shown in Figure 4.1, MHCα and MHCβ were found to contain highly conserved small-xxx-small transmembrane motifs. In the case of MHCα the small residue is

either Gly or Ala whilst for MHCβ it is predominantly Gly. As discussed in Section 1.2.1, this motif is an important structural feature that can stabilise helix-helix interactions in TM domains. It is therefore possible that this conserved motif may play a role in mediating TM helix-helix interactions in the MHC Class II-Ii complex.

(a)



(b)

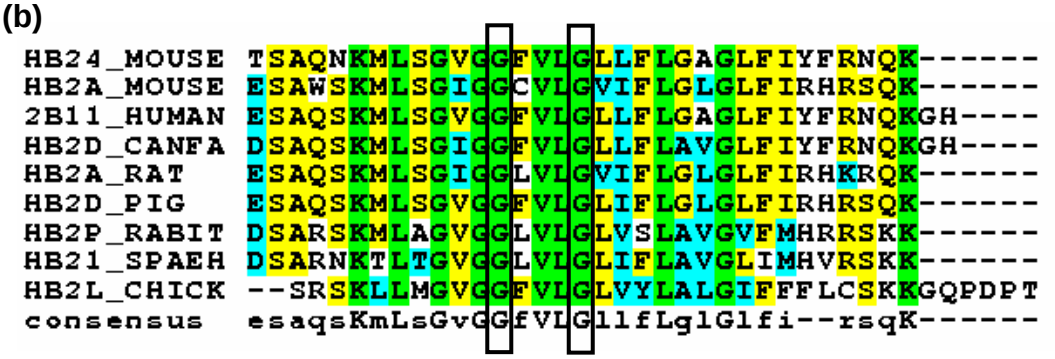


Figure 4.1. Conservation of GxxxG motifs in the TM domains of MHCα and MHCβ

Sequence alignments of the predicted TM domains of (a) 13 MHC Class II α-chain transmembrane domains and (b) nine MHC Class II β-chain transmembrane domains. The TM domains are both highly conserved as are two small-xxx-small motifs (boxed) in MHCα and a single GxxxG motif (boxed) in MHCβ. Alignments were generated using ClustalW2 (Larkin, Blackshields et al., 2007).

The propensity for human MHCα and MHCβ to self-associate to form homo-dimers was explored using molecular dynamics simulations with the program CHI (Adams, Arkin et al., 1995; Adams, Engelman et al., 1996). As shown in Figure 4.2a, this resulted in an energy minimised model for an MHCα homo-dimer where the two GxxxG motifs pack at the interface between the dimer. Similarly, as shown Figure 4.2b, in this analysis resulted in a model for MHCβ homo-dimers where the single GxxxG motif was also packing at the interface of the dimer.

These data are the first suggestion that the TM domains of MHCα and MHCβ are capable of self-association to form homo-dimers and implicate the GxxxG motifs in stabilising that interaction. The biological relevance of this observation is not known since homo-oligomers have not as yet been observed in the full length proteins of either MHCα or MHCβ. A tentative explanation is that they could mediate the formation of inactive oligomers or serve to streamline the process of multiple chains associating with Ii.

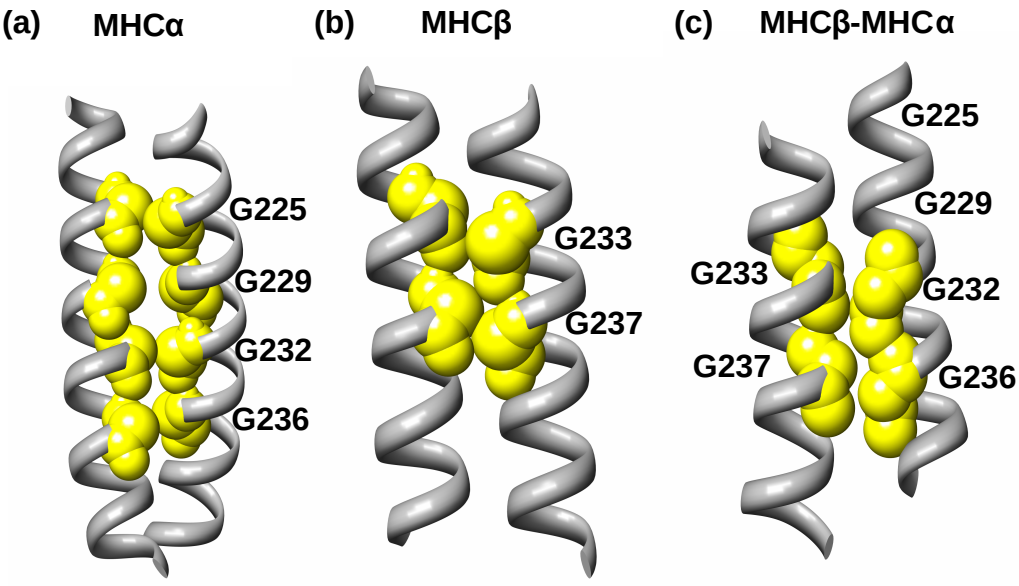


Figure 4.2. Models of homo and hetero-dimers of MHC α/β TM domains

Molecular simulations for homo-dimers of (a) MHCα, (b) MHCβ and a hetero-dimer of (c) MHCα and MHCβ TM domains generated using CHI software (See Section 2.16). Simulated annealing predicted that the dimers are stabilised by packing of the GxxxG motifs (shown in yellow).

Since the full length α and β chains of MHC are known to associate through non-covalent interactions between their soluble domains, the propensity for the TM domains of these two proteins to self-associate and further stabilise α/β hetero-dimers was explored using molecular dynamics studies. As shown in Figure 4.2c, this analysis revealed a model for the MHCα and MHCβ hetero-dimer where the second GxxxG motif of MHCα is packing with the GxxxG motif of MHCβ at the dimer interface. These data, therefore predict that the TM domains of α and β may be able to associate with one another to form hetero-dimers, and that the GxxxG motifs may also be important for stabilising that interaction. These data are of

course in vacuo simulations that will require corroboration from experimental data, but they do provide an important rationale for experimental mutagenesis studies.

4.3 In vivo assays reveal self-association of TM domains of α- and β- chains of MHC

In order to test the hypothesis that MHCα and MHCβ are able to self-associate, the TOXCAT assay was used. This *in vivo* assay enables monitoring of the association of TM domains within the inner membrane of *E. coli* via expression of the reporter gene CAT, as described in more detail in Section 2.17, and presented again in Figure 4.3.

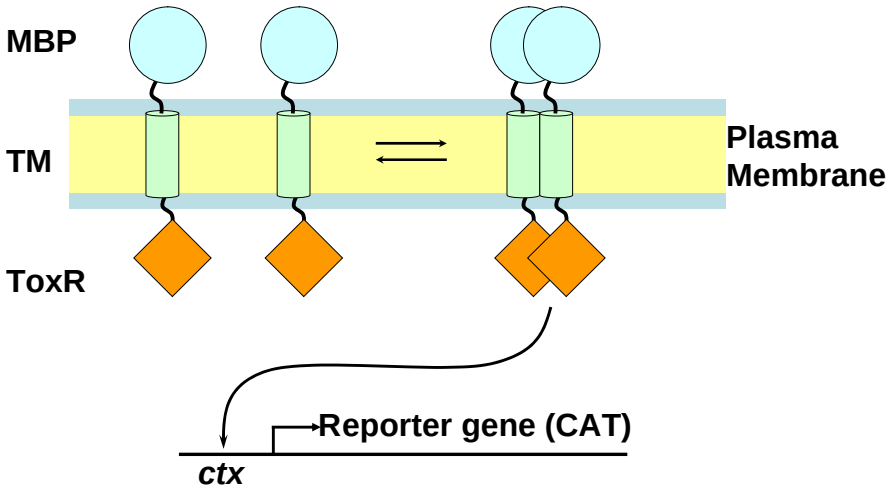


Figure 4.3. Principle of the TOXCAT assay

The TOXCAT assay is used to monitor self-association of α-helical TM domains. A chimeric protein is constructed of Maltose Binding Protein (MalE), the transmembrane domain (TM) of interest and the ToxR promoter. ToxR is a functional dimer that is incapable of dimerizing alone. Interactions between TMDs drive dimerisation of ToxR, which binds the ctx promoter activating transcription of the reporter gene Chloramphenicol Acyltransferase (CAT).

The amino acid sequences for the putative human MHCα and MHCβ proteins were cloned into the vector pccKan, as described in Section 2.17.2. The dimeric TM domain of GpA is used as a positive control for association whilst the G83I mutant of GpA which impairs association is used as a negative control. As shown in Figure 4.4, the level of CAT activity is comparable to the positive control for both MHCα and MHCβ suggesting these TM domains are strongly self-

associating within the *E. coli* inner membrane. This is the first time that the self-association of the MHC TM domains has been observed and substantiates the predictions from both the presence of a GxxxG motif and the molecular dynamics studies.

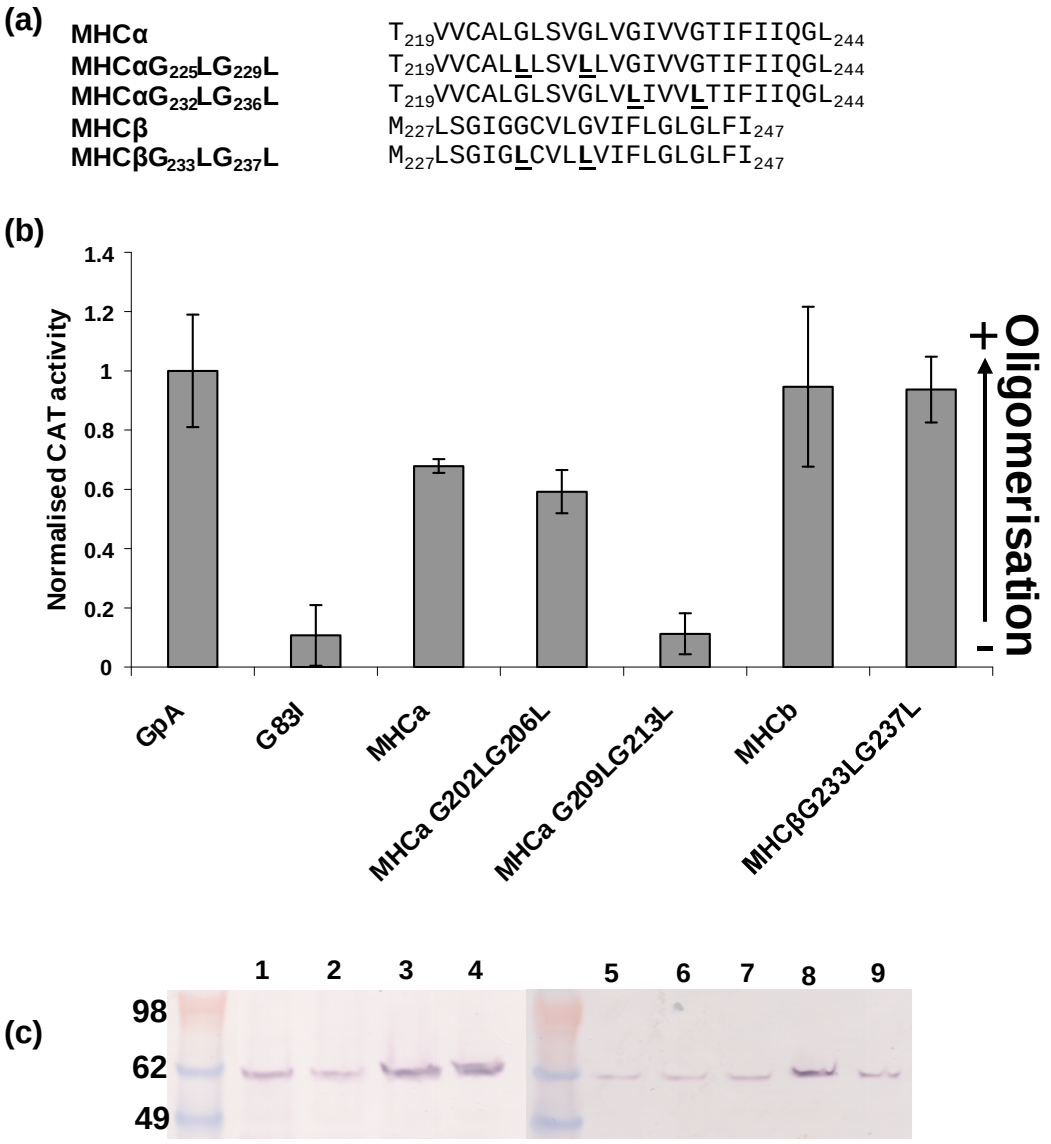


Figure 4.4. Monitoring the self-association of MHCα and MHCβ

Self-association of TM domains was monitored using the TOXCAT assay. (a) Amino acid sequences for the predicted TM domains of Human MHC α (MHCα) and β (MHCβ) chains, and the GxxxG double mutants, cloned into the plasmid pccKan. (b) Self-association was monitored via activity of the CAT reported gene with higher values indicating association. GpA and the dimerisation deficient mutant of GpA, G83I, act as positive and negative controls respectively. Error bars represent the standard error from three separate measurements. (c) Western blot from showing expression levels of the constructs. Molecular weight markers in kDa are shown to the left of each blot. (1 and 5) GpA, (2 and 6) G83I, (3) MHCα (4) MHCβ (7) MHCαG225LG229L (8) MHCαG232LG236L (9) MHCβG233LG237L.

Molecular dynamics simulations predicted that the GxxxG motifs of MHCα and MHCβ may play a role in stabilising the self association. To determine if this is the case, mutagenesis of the wild type sequences was performed. Double mutants were produced where the two Gly residues of the GxxxG motifs were replaced with Leu residues. It was hypothesised that the increased steric hindrance from the presence of a bulkier side chain at these positions would disrupt the self association. The TOXCAT assay was used to monitor the association of these double mutants. As shown in Figure 4.4, the mutation $MHC\alpha_{G225LG229L}$ had little effect on the self-association of MHCα, whilst the mutation $MHC\alpha_{G232LG236L}$ significantly reduced the self association implicating these residues as being situated at the interface of the oligomer. In contrast, the mutation of the GxxxG contained within MHCβ did not result in a loss of oligomerisation, suggesting that this motif is not involved in the self assembly of this TM domain. Therefore, these data suggest a role for GxxxG motifs in the self-association of MHCα but not that of MHCβ, and lend further evidence to the significance of this motif in stabilising the self-association of TM domains.

4.4 In vivo evidence for the hetero-association of α and β TM domains

After establishing the propensity for MHCα and MHCβ to self-associate using the TOXCAT assay we sought to determine if these two TM domains could interact with each other (i.e. can they form hetero-oligomers). This measurement can be achieved *in vivo* by the use of the GALLEX assay, the principle of which is described in Section 2.18.1. In order to progress to monitoring hetero-association with this assay, it was first necessary to determine if self association could be monitored by this method, and thus in the process confirm the efficacy of this assay.

Oligonucleotide inserts encoding the amino acid sequences for MHCα and MHCβ were designed and cloned into the pBLM100 plasmid and transformed into the host strain *E. coli* SU101, as described in the Section 2.18.2. Initial experiments were performed with the same amino acid sequence length as was used in the TOXCAT assay in Section 4.3. As shown in Figure 4.5, the MHCα sequence with

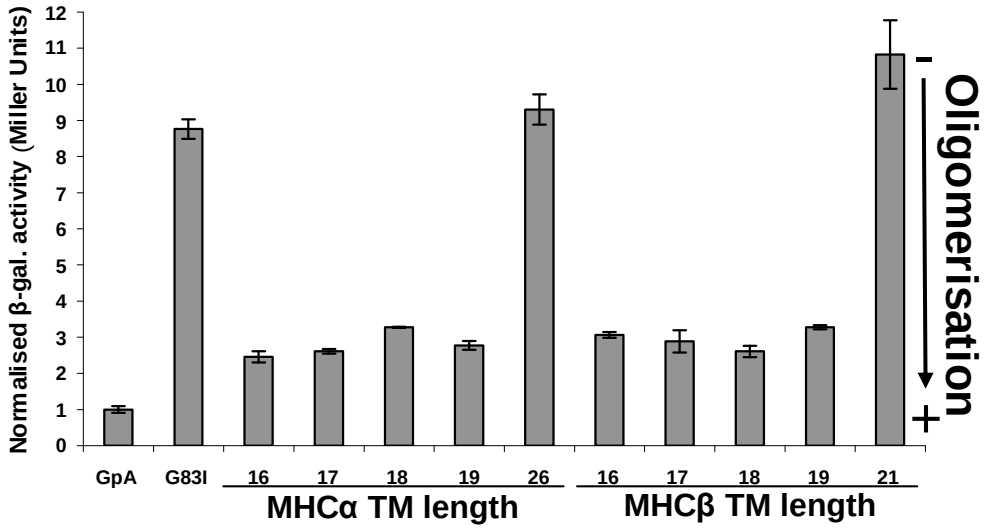
length 26 (MHCα26) and the MHCβ sequence with length 21 (MHCβ21) did not repress β-galactosidase activity, giving results greater than the negative control. This indicates that these sequences of MHCα and MHCβ are not self-associating and contradicts the findings from the TOXCAT assay (see Figure 4.4).

It was hypothesised that the discrepancy between the results from the GALLEX and TOXCAT assays may be due to the length of the α-helices incorporated into the chimera. The crossing angle of the helices and their length may be such that the LexA domains are held apart by the interacting TM domains thus preventing LexA dimerisation and leading to a false negative result. To determine if this was the case for MHCα and MHCβ, the length of the TM domain sequence was varied. In each case the length was varied in order to maintain a centralised location for the GxxxG motifs in the sequence. This was because it has been noted that the strength of the interaction is dependent on the position of the GxxxG motif (Johnson, Rath et al., 2006). As shown in Figure 4.5b, the repression of β-galactosidase activity displays a strong dependency on the length of the sequence studied with TM lengths of between 16 and 19 residues for both MHCα and MHCβ repressing β-galactosidase expression. The results therefore indicate MHCα and MHCβ TM domains are interacting in *E.coli* membranes. This result would seem to indicate that there is a critical sequence length for the TM insert, above which any TM interactions can not be observed. Therefore after, sequence length optimisation therefore, the results for the self-association of MHCα and MHCβ TM domains were found to be consistent between the TOXCAT and GALLEX assays. The propensity for the self-association of MHCα and MHCβ TM domains within a native membrane has therefore been substantiated by two *in vivo* assays.

(a)

MHCα26	T_{219}VVCAL**G**LSV**G**LV**G**IVV**G**TVFIIQGL$_{244}$
MHCα19	A_{223}L**G**LSV**G**LV**G**IVV**G**TVFII$_{241}$
MHCα18	A_{223}L**G**LSV**G**LV**G**IVV**G**TVFI$_{240}$
MHCα17	A_{223}L**G**LSV**G**LV**G**IVV**G**TVF$_{239}$
MHCα16	A_{223}L**G**LSV**G**LV**G**IVV**G**TV$_{238}$
MHCβ21	M_{227}LSGVG**G**FVL**G**VIFLGAGLFI$_{247}$
MHCβ19	M_{227}LSGVG**G**FVL**G**VIFLGAGL$_{245}$
MHCβ18	M_{227}LSGVG**G**FVL**G**VIFLGAG$_{244}$
MHCβ17	M_{227}LSGVG**G**FVL**G**VIFLGA$_{243}$
MHCβ16	M_{227}LSGVG**G**FVL**G**VIFLG$_{242}$

(b)



(c)

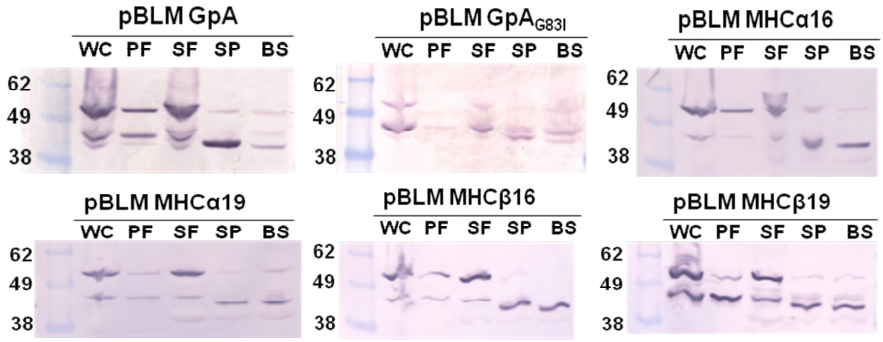


Figure 4.5. Self-association of MHCα and MHCβ TM domains in GALLEX assay

Self-association of MHCα and MHCβ TM domains was monitored using the GALLEX assay, as described in the text (a) Amino acid sequence of the human MHCα and MHCβ TM domains cloned into pBLM100 and subsequently transformed into *E.coli* strain SU101 as described in Section 2.18.2 (b) β-galactosidase activity mediated by the oligomerisation propensity of the expressed constructs in *E.coli* SU101. Repression of activity is indicative of association of the TM domains. Error bars represent the standard error from three independent measurements. Expression of the chimeric proteins was induced by the addition of 1 mM IPTG. Details of the β-galactosidase assay and the calculation of Miller Units are described in the Section 2.18.3. All plasmids and *E. coli* strains were kindly proved by Dr Schneider. GpA and the dimerisation deficient mutant of GpA, G83I, act as positive and negative controls respectively. Data were normalised to GpA (c) Test for insertion and orientation of the expressed chimera. Western blot analysis of E.coli extracts: WC, whole cell; PF, periplasmic fraction; SF, spheroplast fraction; SP, spheroplast proteolysis; BS, broken spheroplast proteolysis. The expressed chimeric proteins with a molecular mass of 54 kDa are found solely in the inner membrane fraction and correctly oriented in the membrane.

As discussed in Section 4.3, it was found using the TOXCAT assay that the C-terminal GxxxG motif of the MHCα TM domain played a role in its self-association and that the same motif in MHCβ did not. In order to confirm that this effect could still be observed with the optimised sequences used in the GALLEX assay, this mutagenesis study was repeated. As shown in Figure 4.6b, the mutation $MHC\alpha_{G225LG229L}$ had no effect on the self-association of this TM domain, whilst $MHC\alpha_{G232LG236L}$ reduced its self-association to the level of the negative control GpA_{G83I}. These results corroborate those from the TOXCAT assay and lend further support to the significance of the GxxxG motif in this interaction.

However, contrary to what was observed using the TOXCAT assay (see Figure 4.4), the GALLEX assay indicates that the $G_{233}LG_{237}L$ mutation of MHCβ does cause a moderate but significant decrease in self-association of MHCβ (Figure 4.6b) The decrease is not as large as that observed for the $MHC\alpha_{G232LG236L}$ mutant, but would suggest that the GxxxG motif in MHCβ may play a role in self-association. The cause for the discrepancy between these two assays is unknown, but may be related to the sensitivity of the GALLEX assay to the length of the TM domain, which required a shorter sequence for GALLEX measurements.

As discussed in Section 2.18.1, the GALLEX assay was predominantly developed to enable determination of hetero-association between α-helical TM domains. Since the full length MHC is known to be a hetero-dimer we therefore used the GALLEX assay to determine the propensity for the MHCα and MHCβ sequences to associate with each other and form hetero-oligomers. As shown in Figure 4.7, a β-galactosidase activity was observed that is comparable to that of the positive control GpA for MHCα and MHCβ, suggesting they associate within the inner membrane of *E.coli*. This is the first data to demonstrate association between the TM domains of α- and β-chains of MHC.

Mutagenesis of the GxxxG motifs was performed and the effect on the association of MHCα and MHCβ monitored by the GALLEX assay. As shown in Figure 4.7, the mutation of the second GxxxG motif in MHCα ($MHC\alpha_{G232LG236L}$) significantly reduced the association of MHCα and MHCβ. This effect was also apparent when the GxxxG in MHCβ was mutated ($MHC\beta_{G233LG237L}$) although to a less extent.

The mutation of the first GxxxG in MHCα (MHCα$_{G225LG229L}$) did not affect the hetero-association of MHCα and MHCβ. These data are in agreement the findings from molecular modelling where the G232G236 motif in MHCα was shown to pack with the G233G237 motif in MHCβ. These findings therefore strongly implicate the GxxxG motif as stabilising heterodimer formation of in the TM domain of the α- and β-chains of MHC.

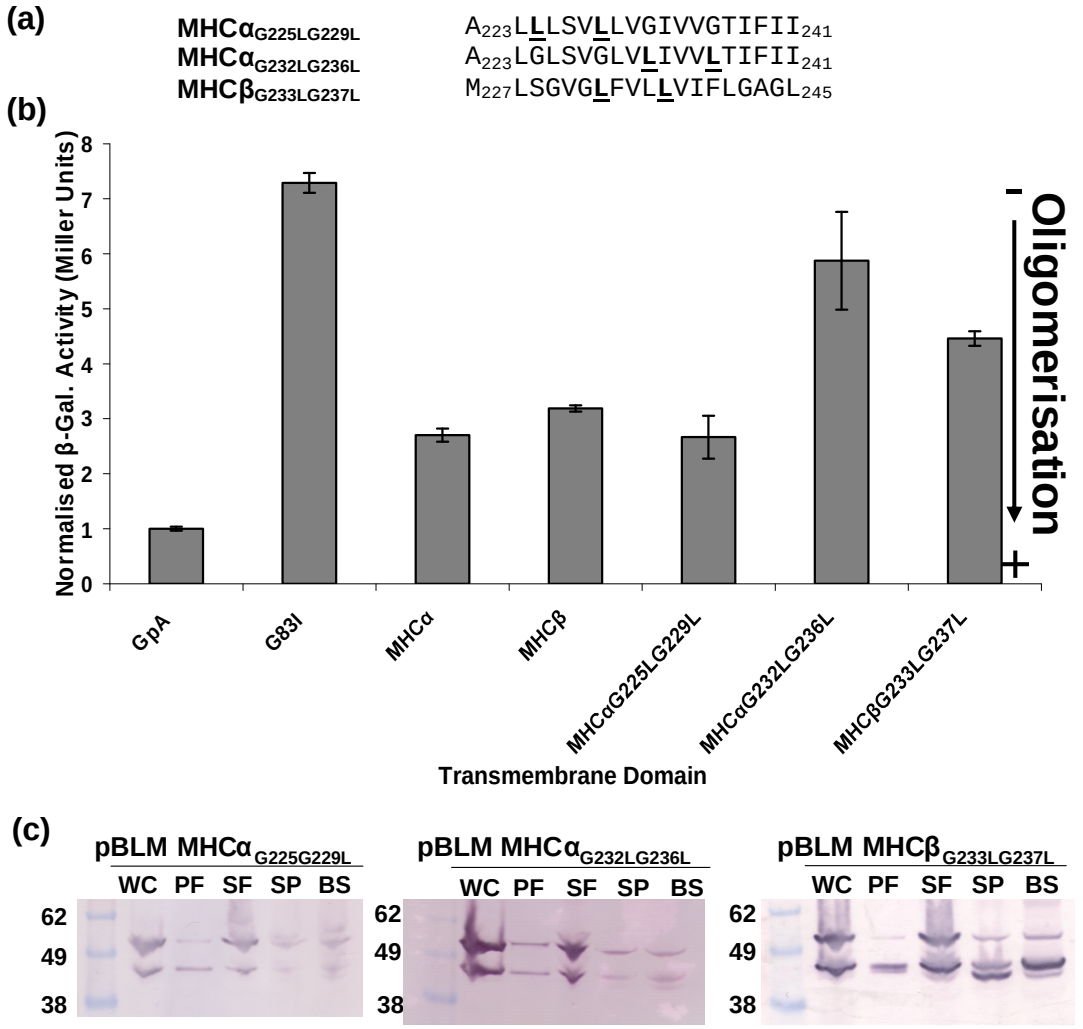


Figure 4.6. Self-association of mutated MHCα and MHCβ TM domains in GALLEX assay

Self-association of MHCα and MHCβ TM domains. (a) Sequences of the human MHCα and MHCβ TM domains cloned into pBLM100 as described in methods. Positions of the GxxxG motifs are indicated. (b) The propensity for the TM domains to self-associate was measured with the GALLEX assay (Schneider and Engelman, 2003). Internal standards of human GpA (very strong interaction producing minimal β-galactosidase activity) and G83I (mutant of GpA that shows minimal interaction producing a high β-galactosidase activity). Data were normalised to GpA (c) Test for insertion and orientation of the expressed chimera. Western blot analysis of E.coli extracts: WC, whole cell; PF, periplasmic fraction; SF, spheroplast fraction; SP, spheroplast proteolysis; BS, broken spheroplast proteolysis. The error bars represent the standard error from three separate measurements.

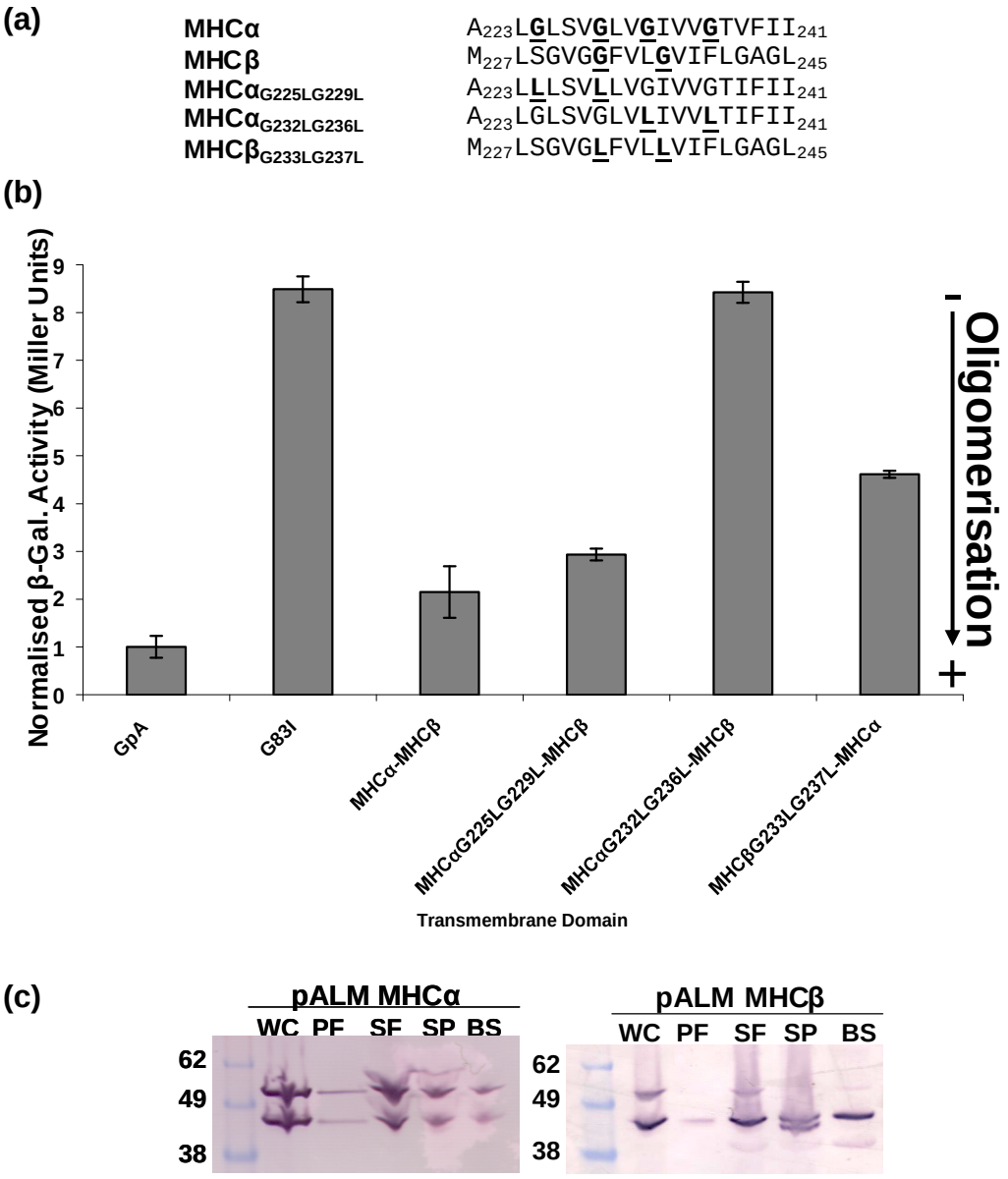


Figure 4.7. Hetero-association of MHCα and MHCβ TM domains

Monitoring hetero-association of MHCα and MHCβ TM domains using the GALLEX assay. (a) Sequences of the human MHCα and MHCβ TM domains cloned into pALM100 and pBLM100 respectively and transformed together into *E.coli* SU202, as described in Section 2.18.2. (b) Measurement of β-galactosidase activity. Internal standards of human GpA (very strong interaction producing minimal β-galactosidase activity) and G83I (mutant of GpA that shows minimal interaction producing a high β-galactosidase activity). Data were normalised to GpA. (c) Test for insertion and orientation of the expressed chimera. Western blot analysis of E.coli extracts: WC, whole cell; PF, periplasmic fraction; SF, spheroplast fraction; SP, spheroplast proteolysis; BS, broken spheroplast proteolysis. The error bars represent the standard error from three individual measurements.

4.5 Synthesis and purification of MHC TM domain analogues

In Section 4.4, the results from in vivo assays suggested that the α-helical TM domains MHCα and MHCβ have a propensity to self-associate to form homo-oligomers and to associate with one another to form hetero-oligomers, via well known GxxxG interaction. In order to confirm that these motifs oligomers represent helix-helix interactions, and determine the possible oligomeric states these domains could adopt in those oligomers, a strategy of studying model peptides that are derived from the TM domains of the α- and β-chains of MHC was employed.

As was discussed for studies of Ii in Chapter 3, the approach of studying peptide analogs has in recent years been useful in discerning the helix-helix interactions of TM domains. Model peptides for MHCα and MHCβ were synthesised at the Keck Facility (Yale University, USA), using the amino acid sequences given in Table 4.1 and purified by RP-HPLC as described in Section 2.10. Additionally, two further peptides, MHCα-Fl and MHCβ-Rh, were synthesised that were differentially labelled at the N-termini with the fluorophores fluorescein and rhodamine, respectively, for use in subsequent FRET experiments to study the hetero-association of these peptides.

Name	Sequence	Mass (Da)
MHCα	KELTETVVCALGLSVGLVGIVVGTVFIIRGLRSWK	3757.55
MHCβ	KSESAQSKMLSGVGGFVLGVIFLGAGLFIYFRNQK	3791.72
MHCα-Fl	Fl-KELTETVVCALGLSVGLVGIVVGTVFIIRGLRSWK	4072.2
MHCβ-Rh	Rh-KSESAQSKMLSGVGGFVLGVIFLGAGLFIYFRNQK	4165.64

Table 4.1. TM domain sequences and labels used in the MHCα-MHCβ FRET assay

Peptides were synthesised at the Keck Facility (Yale University, USA). MHCα-Fl and MHCβ-Rh were synthetic peptides produced for use in FRET studies and were labelled at the N-term with the fluorophores fluorescein (Fl) and rhodamine (Rh), respectively. Labelling was performed at the time of synthesis. Underlined residues indicate the extent of the predicted transmembrane domain. K residues were added to the sequences to aid purification and for cross-linking purposes. Peptides were supplied as crude product from the synthesis and purified by RP-HPLC. Expected masses were used to identify fractions of pure Ii peptides during RP-HPLC purification.

In brief, purification of MHCα was achieved by solubilisation of the crude peptide in formic acid and HFIP, followed by purification on a reverse phase C4 column. The peptide was eluted using a solvent gradient that consisted of water and a mixture of isopropanol and ACN as the mobile phase. A typical RP-HPLC chromatogram for MHCα is shown in Figure 4.8a, and the fraction containing the MHCα peptide is indicated. It is likely that the separation could be improved and the peptide yield increased by the use of a more hydrophobic column, however this equipment was not available for testing. Typically, eight runs of RP-HPLC purification were performed and the purity of the fractions confirmed by MALDI mass spectrometry, as described in Section 2.11. Pure fractions were pooled and lyophilised. As shown in Figure 4.8a, the MALDI mass spectrum of the pooled fractions reveals the major component is the MHCα peptide with a mass of 3756 Da which corresponds to the H^+ ionisation state.

The purification of MHCβ by RP-HPLC was optimised as detailed in Section 2.10.2. Typically, 8 runs of RP-HPLC purification were performed and fractions checked by mass spectrometry. A typical chromatogram MHCβ is shown in Figure 4.9a, with the fraction containing the MHCβ peptide indicated. As shown in Figure 4.9b, the mass spectrum of the pooled fractions reveals the major component is the MHCβ peptide with a mass of 4209 Da which corresponds to the H^+ ionisation state.

The purification of the fluorophore labelled peptides MHCα-Fl and MHCβ-Rh were carried out according to the protocols developed for the purification of their unlabelled counterparts. For MHCα-Fl, a typical RP-HPLC chromatogram and mass spectrum of the purification product are shown in Figure 4.10. The presence of the unlabelled peptide can be detected in the mass spectrum indicating that the purity of MHCα-Fl is not 100%. This is due to the fluorophore not significantly altering the hydrophobicity of this peptide relative the unlabelled peptide, thus 100% separation is difficult to achieve.

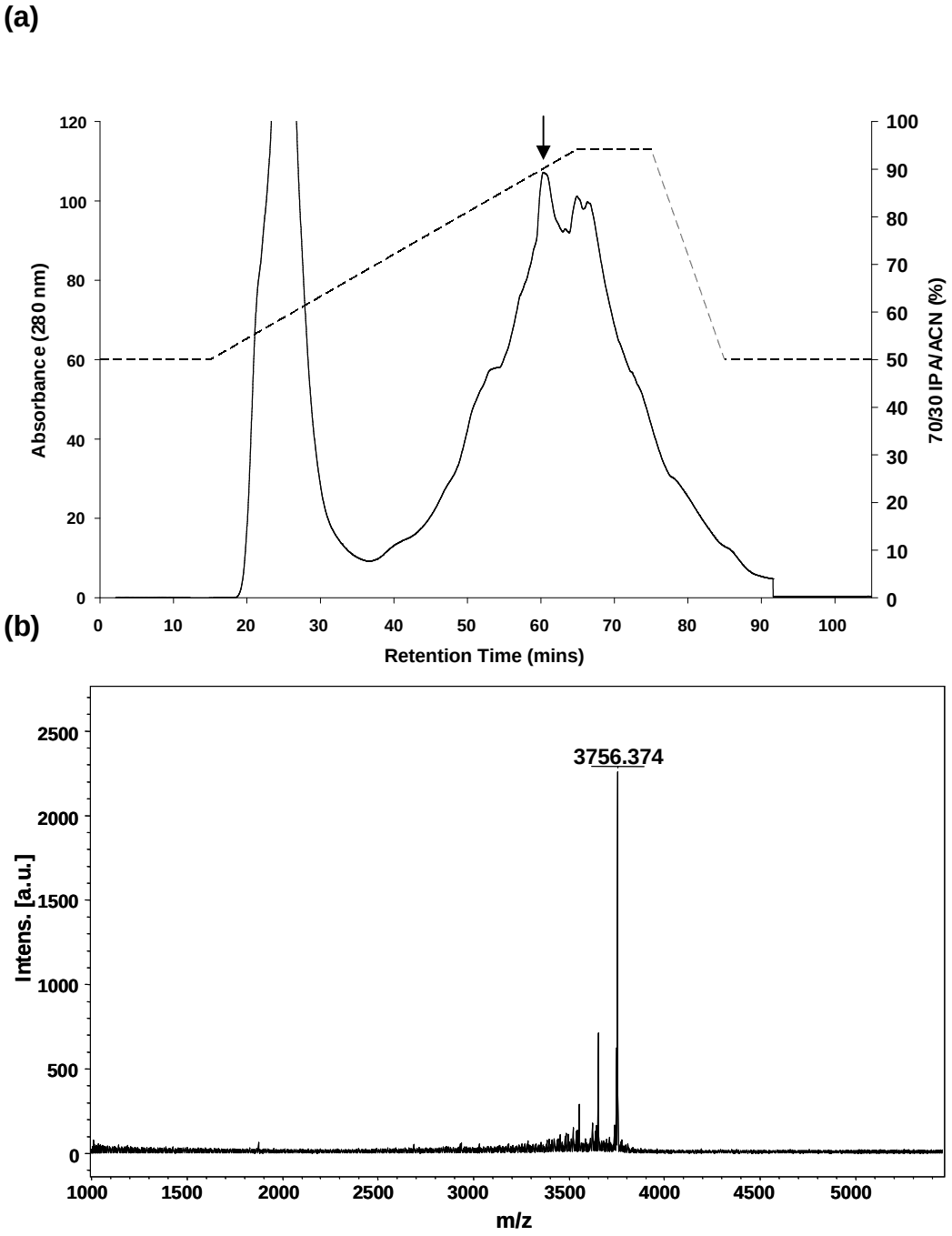


Figure 4.8. RP-HPLC purification of MHCα peptide

(a) The MHCα peptide was purified by reverse phase HPLC (solid line) using a Isopropanol (IPA)/Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent, on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the MHCα peptide is indicated. (b) Purity of pooled fractions from reverse-phase HPLC purification of MHCα peptide was analysed using MALDI mass spectrometry. The major peak with a mass of 3756 Da corresponds to the expected mass for MHCα peptide.

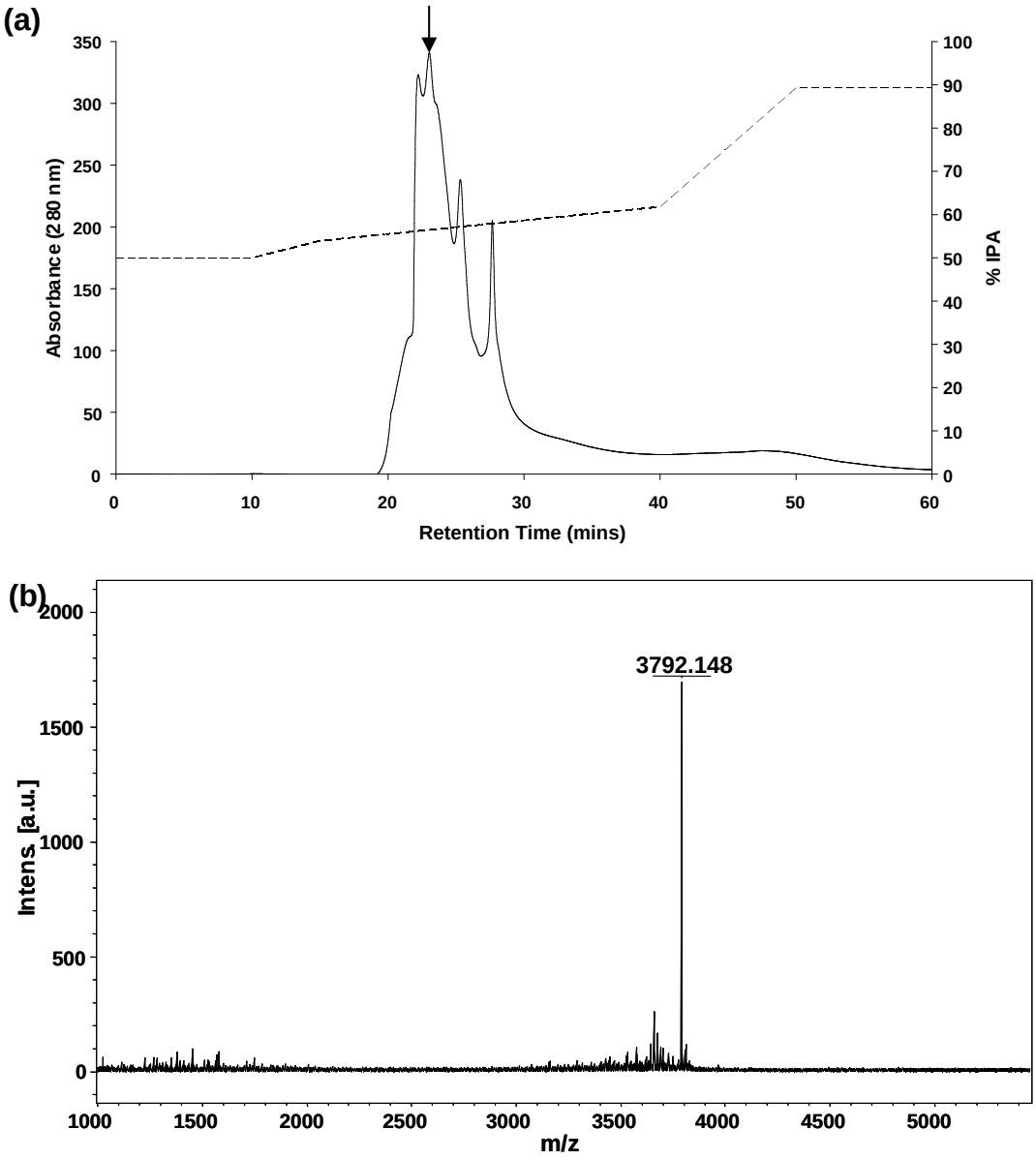


Figure 4.9. RP-HPLC purification of MHCβ peptide

(a) The MHCβ peptide was purified by reverse phase HPLC (solid line) using a Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent, on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the MHCβ peptide is indicated. (b) Purity of pooled fractions from reverse-phase HPLC purification of MHCβ peptide was analysed using MALDI mass spectrometry. The major peak with a mass of 3792 Da corresponds to the expected mass for MHCβ peptide.

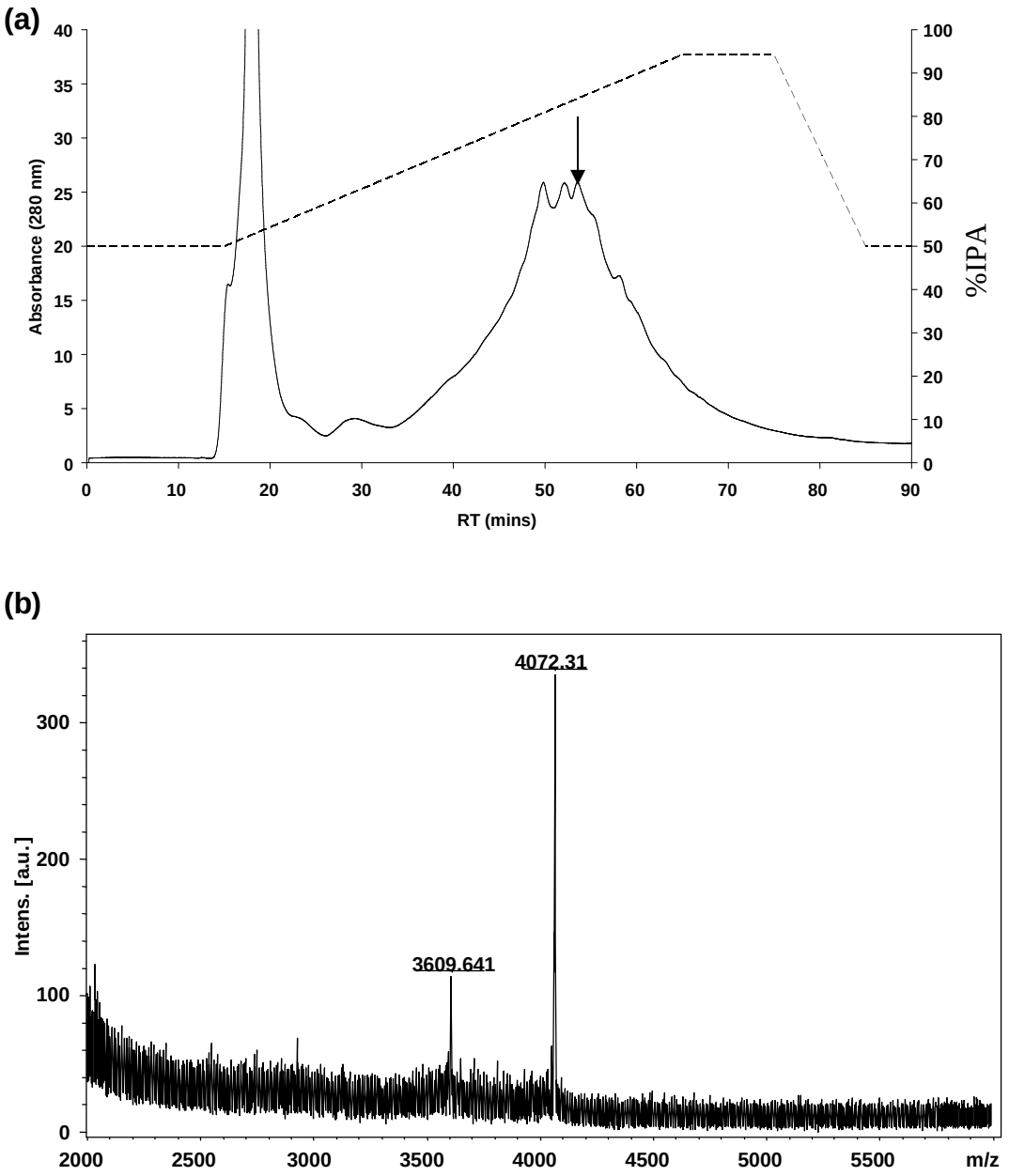


Figure 4.10. RP-HPLC purification of MHCα-Fl peptide

(a) The MHCα-Fl peptide was purified by reverse phase HPLC (solid line) using a isopropanol (ACN) gradient (broken line) and H_2O as the second solvent, on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the MHCα-Fl peptide is indicated. (b) Purity of pooled fractions from reverse-phase HPLC purification of MHCα-Fl peptide was analysed using MALDI mass spectrometry. The major peak with a mass of 4072 Da corresponds to the expected mass for MHCα-Fl peptide.

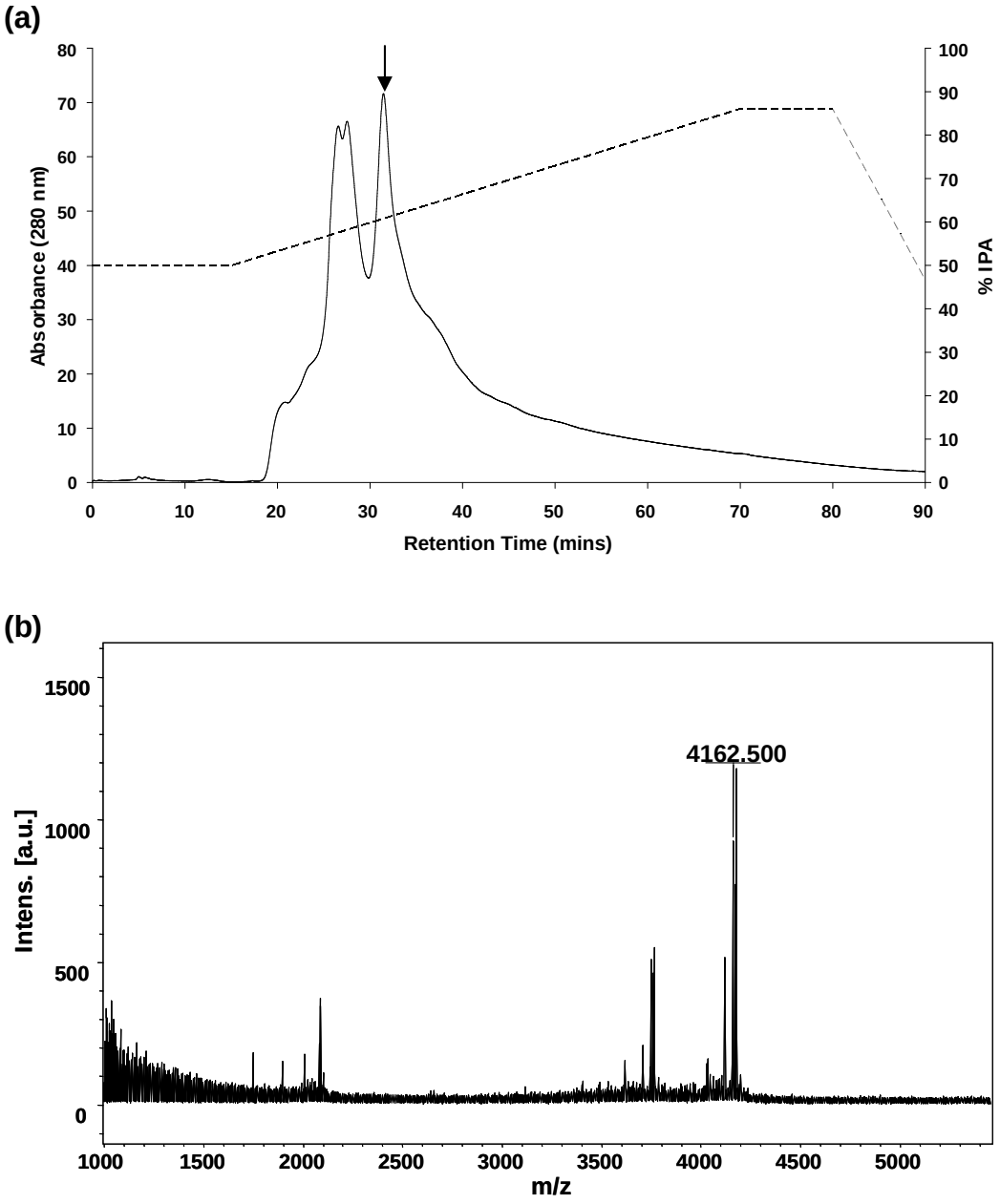


Figure 4.11. RP-HPLC purification of MHCβ-Rh peptide

(a) The MHCβ-Rh peptide was purified by reverse phase HPLC (solid line) using a Acetonitrile (ACN) gradient (broken line) and H_2O as the second solvent, on a Phenomenex Jupiter C4 column (Phenomenex, UK). 0.1% TFA was present in both solvents. Elution of fractions was monitored by the absorbance at 280 nm. The peak generated by the elution of the MHCβ-Rh peptide is indicated. (b) Purity of pooled fractions from reverse-phase HPLC purification of MHCβ-Rh peptide was analysed using MALDI mass spectrometry. The major peak with a mass of 4162 Da corresponds to the expected mass for MHCβ-Rh peptide.

4.6 Secondary structure of peptide analogues of α and β TM domains

Since the amino acid sequences of the model peptides MHCα and MHCβ correspond to the predicted α-helical TM domains of the respective full length α and β proteins, it was necessary to characterise their secondary structure to confirm they were α-helical, which is most readily achieved by using circular dichroism (CD). CD spectra were acquired for both the MHCα and MHCβ peptides at a range of peptide:micelle molar ratios since it had been shown previously for the Ii peptide (see Section 3.4) that this ratio could impact upon the secondary structure of TM peptide analogues.

As shown in Figure 4.12a, for MHCα, minima were observed at 208 and 222 nm for all ratios, indicating the presence of α-helical content in the secondary structure. The CD spectra showed a dependency on the peptide:micelle ratio. As shown in Figure 4.12c, analysis of the CD spectra using the program CDSSTR (Johnson, 1999) revealed that the percentage of α-helix increases as the peptide:micelle ratio is decreased (i.e. detergent concentration is increased), and rises to a maximum of ~40 % α-helix at a ratio of 1:5. The greater signal and increased α-helical content observed at the higher detergent concentrations possibly reflect the improved solubilisation of the peptide.

As shown in Figure 4.13a, for the MHCβ peptide, as the ratio approaches 1:1, an increasing proportion of α-helical content is apparent which rises to a maximum of ~60 % with a concomitant decrease in the percentage of β-sheet and random coil, as shown in Figure 4.13c. It is likely this shift from β-sheet to α-helical secondary structure reflects the increasing solubility of the peptides, and that the b-sheet observed at high peptide:micelle molar ratio represents an aggregated state.

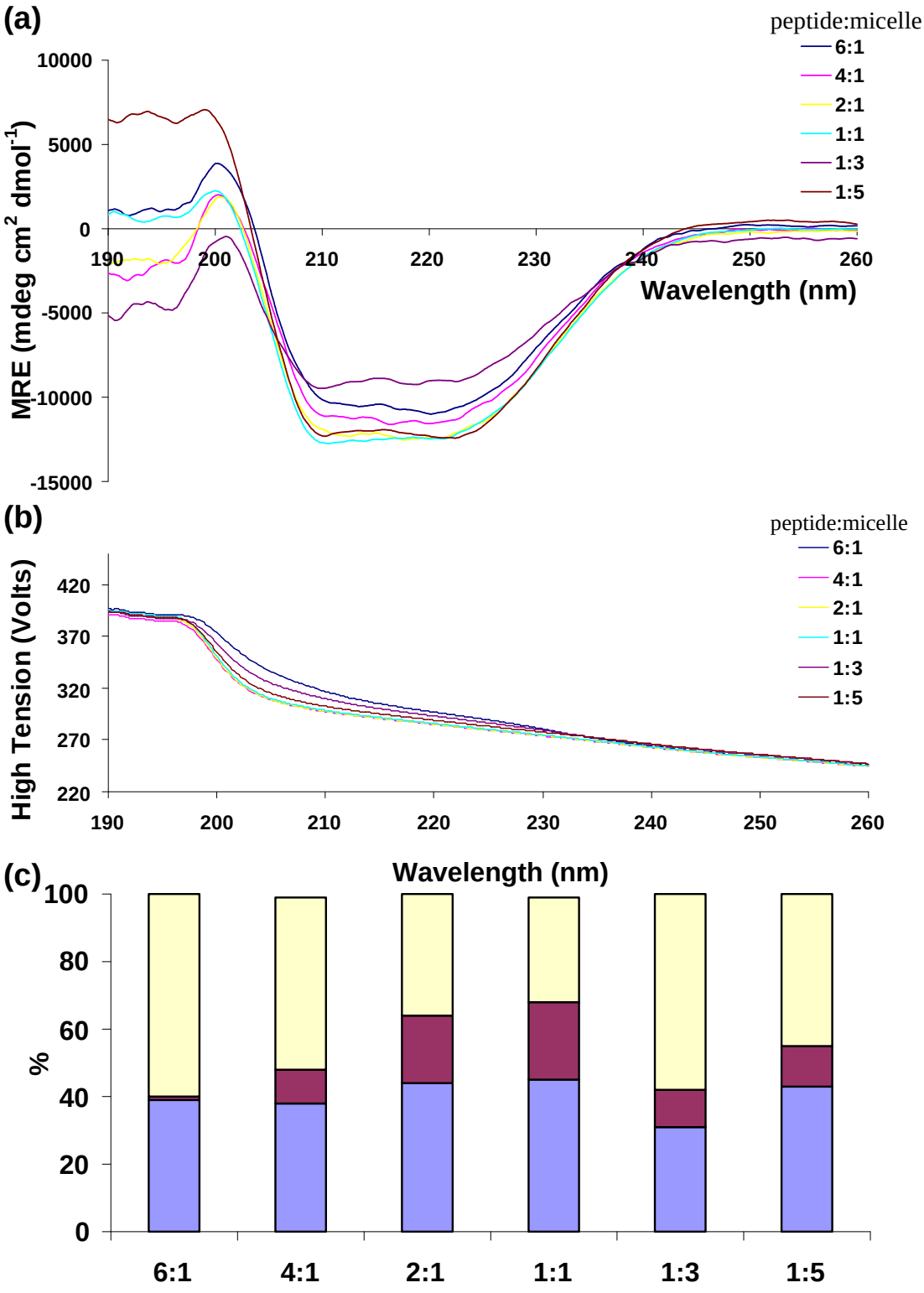


Figure 4.12. CD spectra for MHCα in varying DPC concentrations

(a) Circular Dichroism spectra of MHCα TM peptide reconstituted into the detergent DPC at varying peptide: DPC micelle ratios. Mean residue ellipticity (MRE) was calculated from the measured ellipticity as described in Materials and Methods. (b) High tension for CD spectra, typically the CD data is taken to be reliable whilst this remains below 600. (c) Percentage secondary structure content (α-helix (blue), β-sheet (red), Random coil (yellow) at varying peptide: DPC micelle ratios as calculated from the CD spectra using CDSSTR (Johnson, 1999).

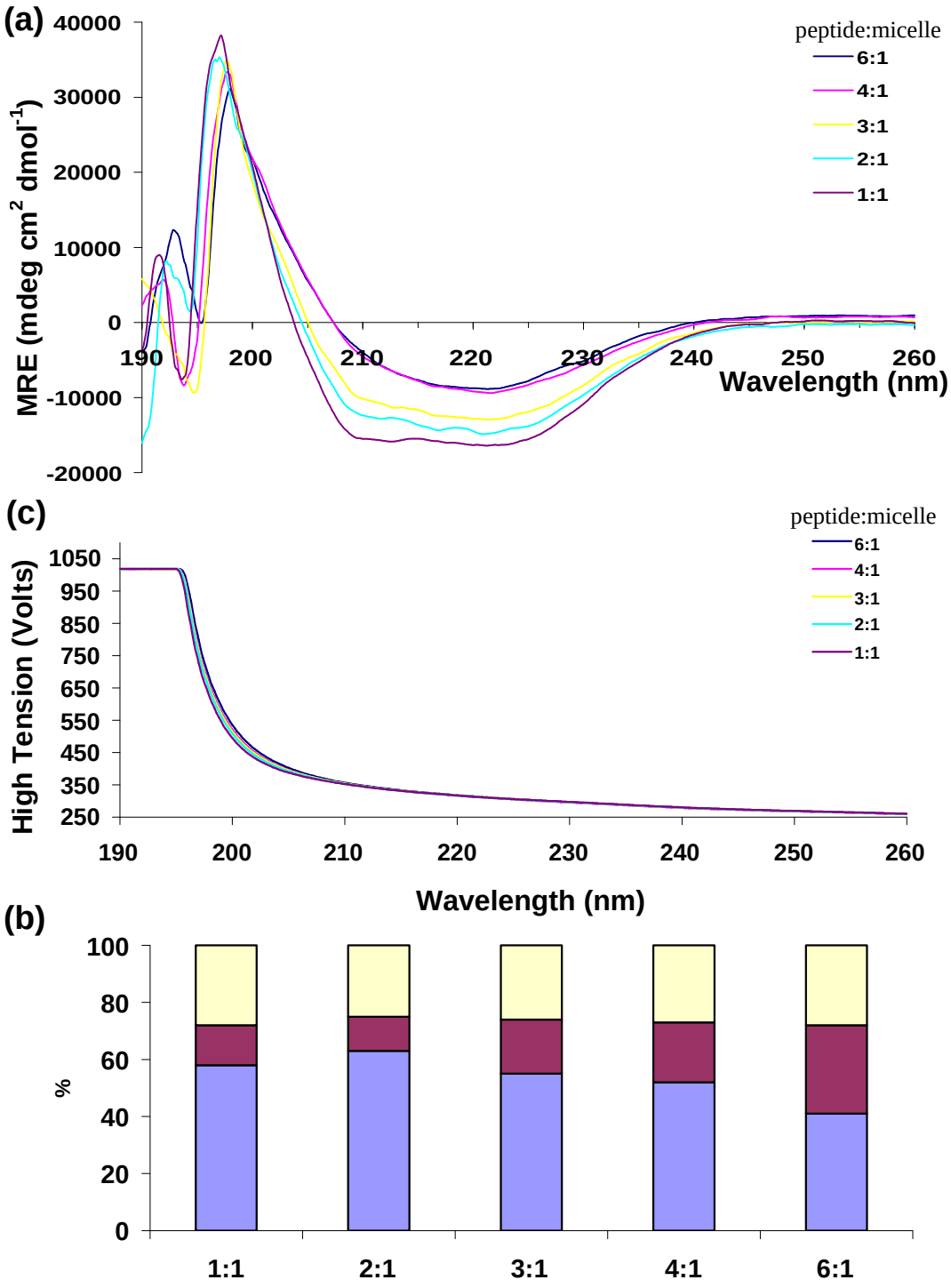


Figure 4.13. CD spectra for MHCβ in varying DPC concentrations

(a) Circular Dichroism spectra of MHCβ TM peptide reconstituted into the detergent DPC at varying peptide: DPC micelle ratios. Mean residue ellipticity (MRE) was calculated from the measured ellipticity as described in Materials and Methods. (b) High tension for CD spectra, typically the CD data is taken to be reliable whilst this remains below 600. (c) Percentage secondary structure content (α-helix (blue), β-sheet (red), Random coil (yellow) at varying peptide: DPC micelle ratios as calculated from the CD spectra using CDSSTR (Johnson, 1999).

It is likely, that these data are not wholly representative of the actual percentage of α-helical content in the MHCα and MHCβ peptides, since this is a global average and the fitting programs typically place a heavy weighting on the data between 190 and 200. This is particularly problematic for peptides solubilised in detergent micelles since the presence of the micelles introduces a significant level of noise in this region due to light scattering. MHCα and MHCβ peptides are expected to be analogous to the TM domains of the α- and β-chains of MHC. The data presented in this section suggests that they are indeed forming structures with significant α-helical content in membrane mimetics and therefore were considered representative models for these α-helical TM domains.

4.7 SDS-PAGE analysis of MHCα and MHCβ TM domain peptides

The preceding results from this study implicated the conserved small-xxx-small motifs in the TM domains of the α- and β-chains of MHC in the self-association of these TM domains. It has been shown for TM domains possessing small-xxx-small motifs, that their peptide analogues have the potential to form highly stable homo-oligomers that are observable by SDS-PAGE (Lemmon, Flanagan et al., 1992). In order to determine if this behaviour could be observed for MHCα and MHCβ peptides and also for their fluorophore labelled counterparts, SDS-PAGE analyses were performed.

The results for the MHCα peptide and the fluorescein labelled MHCα peptide are shown in Figure 4.14, at peptide concentrations over a range of 25 to 125 μM. For MHCα two distinct bands are observed that possibly correspond to monomer (3.76 kDa) and dimer (7.52 kDa) oligomeric states, and indicate that the MHCα peptide is self-associating. These results corroborate those from the *in vivo* TOXCAT and GALLEX self-association studies and indicate it is a strong interaction is since it occurs even in the denaturing detergent SDS. The fluorophore-labelled variant of MHCα peptide, MHCα-Fl was also assessed by SDS-PAGE and gave similar results to the unlabelled peptide, although the lower

band is stronger and appears to be at slightly higher mass in the lanes with higher concentrations of peptide.

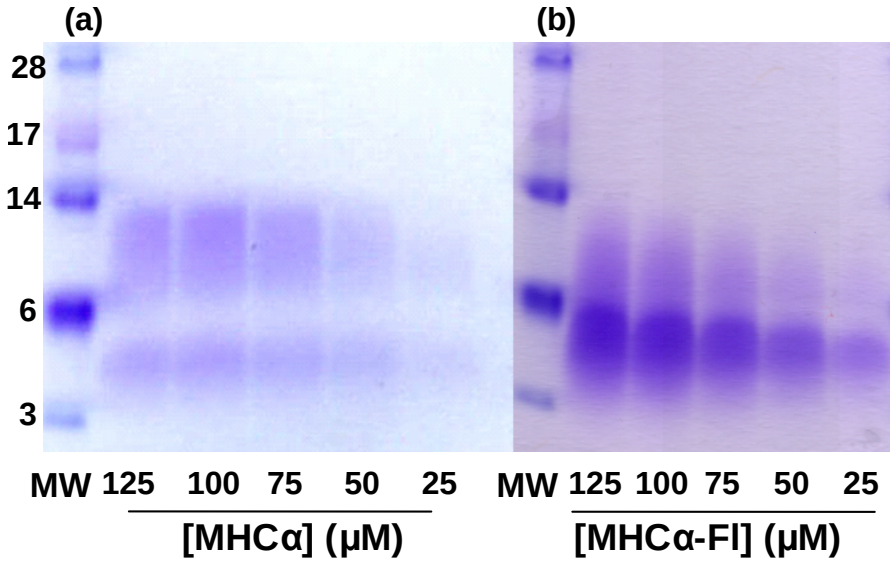


Figure 4.14. SDS PAGE analysis of MHCα and MHCα-Fl peptides

Analysis of (a) MHCα (MW 3.76 kDa) and (b) MHCα-Fl (MW 4.07 kDa) peptides carried out over a range of concentrations as indicated below each lane were dissolved in SDS sample loading buffer, analysed by SDS-PAGE and visualized using coomassie-G250. Molecular mass standards (MW) with masses in kDa are shown in the far left-hand lane of each gel.

The results of SDS-PAGE analysis for the MHCβ peptide and its rhodamine-labelled counterpart, MHCβ-Rh, are shown in Figure 4.15a. For MHCβ peptide a single band is observed for both peptides at a mass that is possibly intermediate between monomer (3.79 kDa) and dimer (7.58 kDa). This is also the case for the rhodamine-labelled MHCβ peptide. There is possibly a second band in the unlabelled peptide but the origin of this is not known. Cross-linking analyses were performed before an oligomeric state was assigned.

Recent data shows transmembrane peptides can run anomalously on SDS-PAGE gels which could lead to a possible mistaken assignment of oligomeric states for the MHCα and MHCβ peptide bands (Rath, Glibowicka et al., 2009; Walkenhorst, Merzlyakov et al., 2009). Therefore, cross-linking analysis was performed before bands were definitively assigned oligomeric states.

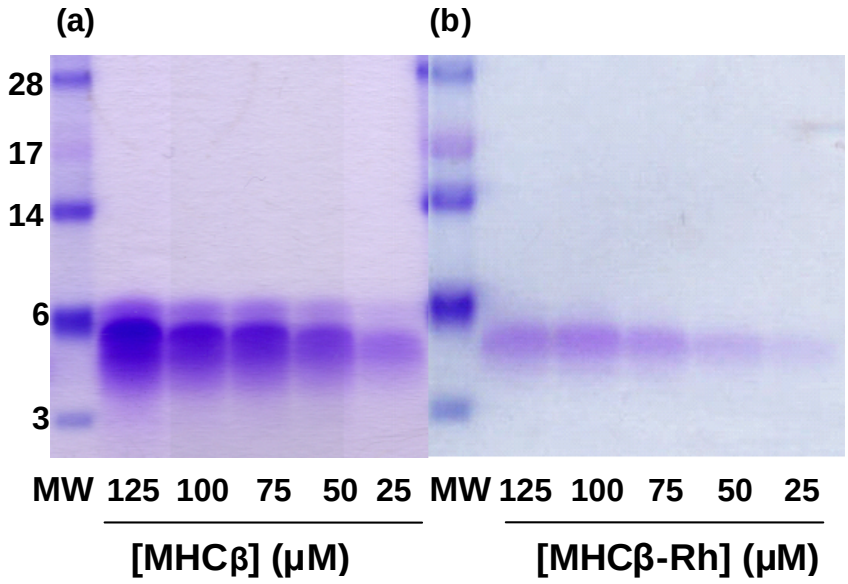


Figure 4.15. SDS PAGE analysis of MHCβ and MHCβ-Rh peptides

Analysis of (a) MHCβ (MW 3.79 kDa) and (b) MHCβ-Rh (MW 4.17 kDa) TM peptides carried out over a range of concentrations as indicated below each lane were dissolved in SDS sample loading buffer, analysed by SDS-PAGE and visualized using coomassie-G250. Molecular mass standards (MW) with masses in kDa are shown in the far left-hand lane of each gel.

4.8 In vitro analysis of self-association of α and β TM domains

As noted in Section 3.5, the solubilisation of TM peptides in the detergent SDS can disrupt some of the weaker helix-helix interactions of TM domains preventing the assembly of the peptides into their native oligomeric states. In order to investigate the formation of oligomers by MHCα and MHCβ peptides in a milder detergent, the peptides were subjected to cross-linking prior to analysis by SDS-PAGE, as described in Section 2.13. MHCα and MHCβ peptides were dissolved in DPC detergent micelles and then treated with the water soluble cross-linker BS^3. Cross-linking reactions were carried out at increasing peptide:micelle molar ratio to investigate the effect of detergent concentrations on the oligomeric state. Cross-linked species were analyzed using SDS-PAGE and visualised by staining with silver nitrate for its increased sensitivity over coomassie staining.

As shown in Figure 4.16a, for MHCα several bands are observed at peptide:micelle ratios between 10:1 and 4:1 that correspond to oligomeric states

from monomer (n=1) to tetramer (n=4) and higher, suggesting the peptides are not soluble at these detergent concentrations leading to aggregation. As the peptide:micelle ratio is varied from 2:1 to 1:3 the number of oligomeric states is reduced with monomer (3.76 kDa), dimer (7.52 kDa), trimer (11.28 kDa) and tetramer (15.04 kDa) bands being most clearly delineated. Furthermore, the abundance of the oligomeric states clearly shifts to lower order states as the peptide:micelle molar ratio is reduced. This data shows that the bands observed in SDS-PAGE correspond to dimer and monomeric states for MHCα peptide. Therefore, in the milder detergent DPC, MHCα is capable of forming higher order oligomers above dimer but these states are not stable since they can be modulated by the detergent concentration. As a negative control, cross linking was carried out with the peptide dissolved in SDS and two bands were observed as observed in the results of the previous section.

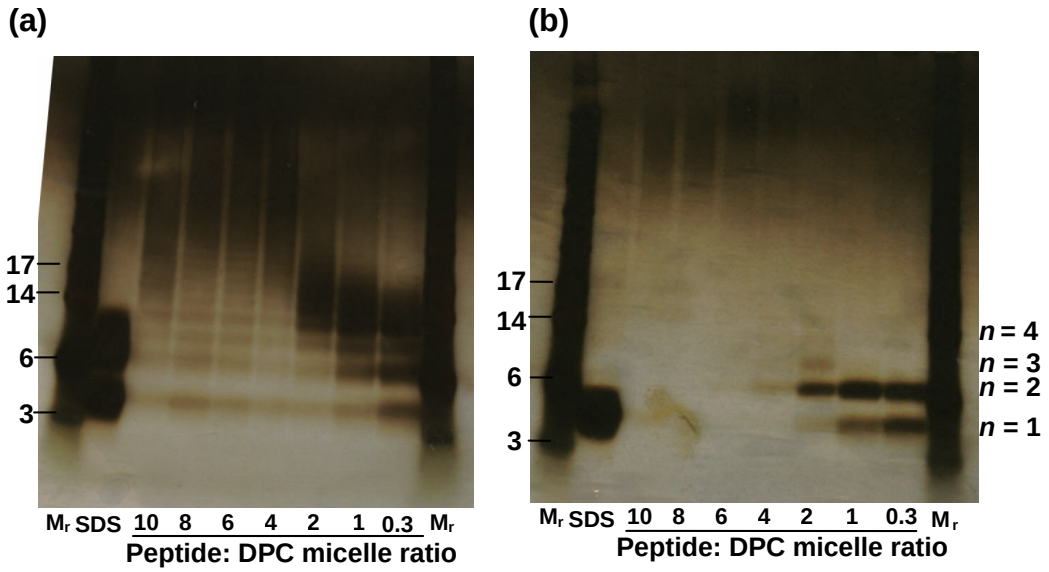


Figure 4.16. Analysis of self-association propensity of MHCα and MHCβ TM derived peptides in DPC detergent

SDS-PAGE analysis of BS_3-mediated cross-linking of (a) MHCα (MW 3.76 kDa) and (b) MHCβ (MW 3.79 kDa) TM peptides dissolved in DPC at varying peptide:micelle ratios as indicated. Molecular weight markers are shown in the far left- and right hand lanes (M_r). Protein bands were visualized by staining with silver nitrate. Oligomeric states (e.g. dimer indicated by n = 2) are indicated at the far right of the gels. A negative control reaction in which cross-linking was carried out for the peptide dissolved in SDS buffer is shown in the first lane.

The results from the cross-linking of MHCβ peptide are shown in Figure 4.16b. No distinct bands are observed at peptide:micelle ratios below 4, but peptide is evident at the top of the gel near the wells suggesting that the peptide is insoluble in DPC at these ratios. At peptide:micelle ratios of 4 and 2 bands corresponding to monomer (3.79kDa), dimer (7.58 kDa) and trimer (11.37 kDa) are observed. At ratios of 1 and 0.3 (i.e. 1:3) we see bands corresponding to monomer and dimer only, with monomer becoming more prevalent as the ratio decreases. Compared to MHCα, MHCβ predominantly self-assembles into dimers, this is interesting considering that MHCα has two GxxxG motifs whilst MHCβ has only one, indicating that MHCα may have different modes of interaction. In a similar manner to MHCα and Ii, the oligomeric state of MHCβ can be modulated by the detergent concentration.

4.9 FRET analyses of TM domain associations of MHC

Results from the *in vivo* GALLEX assay described in Section 4.4 suggested that the TM domains of the α- and β-chains of MHC can associate to form hetero-oligomers. We would therefore expect to observe this same behaviour in the model peptides. However, to date relatively few techniques are available to study the hetero-association of hydrophobic peptides *in vitro.* As discussed in Section 2.19, and as was shown in the preceding studies of Ii (see Section 3.7), the phenomenon of FRET can be used to monitor the self-association of synthetic peptides in membrane mimetics such as detergents or lipids. Fortunately, it is possible to extend this approach to the study of hetero-association through the differential labeling of peptides with fluorophores that constitute a FRET pair, which for the purposes of this study were fluorescein and rhodamine. FRET analyses were therefore employed in this study to confirm the hetero-association of the MHCα and MHCβ peptides predicted from our in vivo studies and to determine the oligomeric state of that interaction.

4.9.1 Association of the α- and β-chain TM domains

The FRET between MHCα-Fl and MHCβ-Rh in the detergent DPC was monitored using fluorescence as described in Section 2.19.5. Since it has been shown that the association of the model peptides is dependent on the detergent concentration or more specifically the peptide:micelle molar ratio, the dependence of the FRET with this parameter was explored.

Model peptides of MHCα and MHCβ were synthesised and labelled with the fluorophores fluorescein and rhodamine to produce the peptides MHCα-Fl and MHCβ-Rh, respectively, and purified using RP-HPLC as described in Section 4.5. For the purposes of FRET measurements MHCα-Fl provides the donor and MHCβ-Rh, the acceptor. All FRET samples used in the studies in this section were prepared using the co-dissolving method as described in Section 2.19.4. The FRET between MHCα-Fl and MHCβ-Rh was monitored at peptide:micelle molar ratios of between 1:3 and 4:1 in the detergent DPC whilst keeping the total peptide concentration constant. For each ratio the energy transfer was calculated as described in Section 2.19.6, and plotted versus the peptide:micelle molar ratio.

As shown in Figure 4.17, energy transfer between MHCα and MHCβ peptide was observed strongly suggesting these peptides are interacting. Furthermore, the FRET was found to be dependent on the peptide:micelle molar ratio with the greatest energy transfer being observed at ratios above 1:1 whilst below this ratio no or minimal energy transfer is observed. A maximum energy transfer of around 50% is observed. At a ratio above 4:1, the decreased solubility of the peptides prevents the interpretation of FRET measurements.

Since both MHCα and MHCβ have been shown in this study to self-associate it is possible that the observed energy transfer is being modulated by the formation of homo-oligomers. Furthermore, the presence of unlabelled MHCβ could be modulating the energy transfer leading to a reduced FRET signal. To see if this is the case it would be necessary to analyse the homo interaction by FRET in future work.

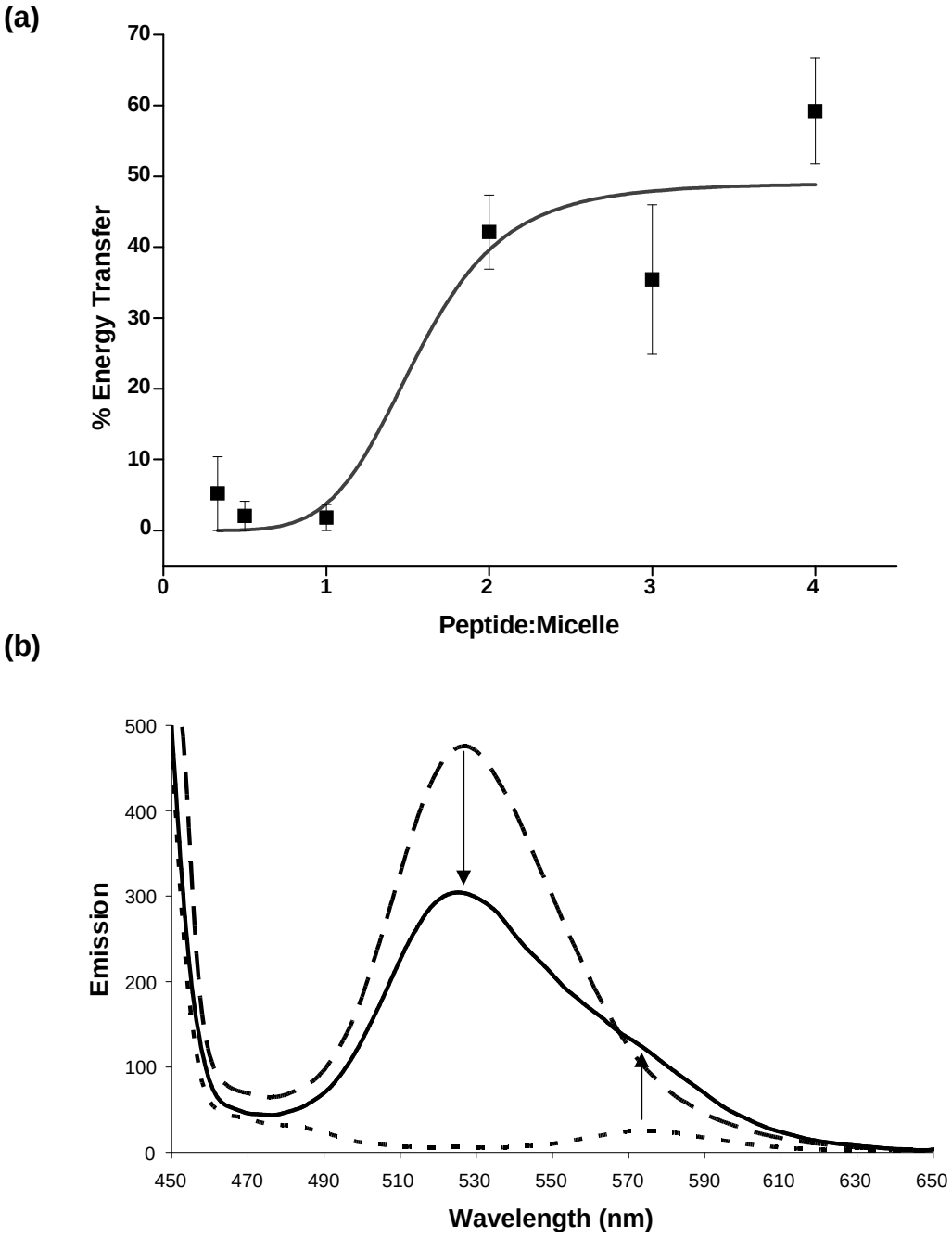


Figure 4.17. Change in energy transfer between MHCα-Fl and MHCβ-Rh with detergent concentration

(a) Plot of percentage energy transfer versus peptide:micelle ratio. The total donor and acceptor peptide concentrations were kept constant at 4 μM (2 μM MHCα-Fl, 2 μM MHCβ-Rh) while the detergent concentration was varied. Samples were prepared by mixing peptide and DPC pre-solubilised in TFE. A CMC of 1 mM and aggregation number of 56 were used in calculations of the peptide:micelle molar ratio. Emission spectra were collected and energy transfer was calculated as described in Materials and Methods. (b) FRET spectra for MHCα-Fl and MHCβ-Rh in DPC micelles at a peptide:micelle ratio of 4:1. The broken line is MHCα-Fl only spectra, the dotted line is the spectra of MHCβ-Rh on its own whilst solid line is the spectra for a mixture of MHCα-Fl and MHCβ-Rh. A FRET signal is evident from the decrease in the donor emission at 520 nm and an increase in acceptor emission at 570 nm as indicated by the arrows.

4.9.2 Determining specificity of MHC TM domain association

In order to determine if the FRET signal measured between MHCα and MHCβ arises from a specific interaction, a competition assay was performed by the titration with increasing concentrations of unlabelled MHCβ peptide. Since this experiment must be carried out at a fixed detergent concentration, the peptide:micelle molar ratio of 3:1 was chosen since this yielded a high energy transfer and . The concentrations of MHCα-Fl and MHCβ-Rh peptides were kept constant whilst varying the concentration of unlabelled peptide. As shown in Figure 4.18, the energy transfer decreases with increasing concentration of unlabelled MHCβ peptide. This is indicative of unlabelled MHCβ peptide disrupting the formation of donor and acceptor partners by competing with MHCβ-Rh for binding to MHCα-Fl.

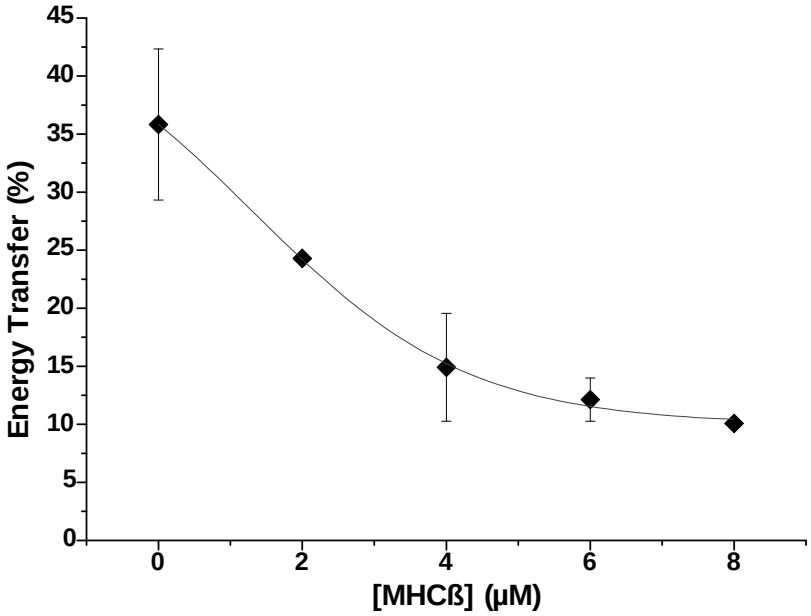


Figure 4.18. Effect of unlabelled MHCβ peptide on FRET from MHCα-Fl to MHCβ-Rh

Plot of energy transfer between MHCα-Fl and MHCβ-Rh versus the concentration of unlabelled MHCβ. Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM MHCα-Fl, 2 μM MHCβ-Rh) while the concentration of unlabelled MHCβ was varied. Samples were prepared by co-dissolving all peptides and DPC dissolved in TFE. Experiment was performed at a peptide:micelle molar ratio of 3:1. The reduced FRET efficiency suggests that sequence-specific oligomerisation contributes to the measured FRET efficiency.

4.9.3 Determining the oligomeric state of MHC TM domain

It has been noted in the preceding chapters that the oligomeric state of interacting TM domain peptides can be monitored using FRET analyses. This is achieved by the measurement of energy transfer as a function of the mole fraction of acceptor. Since it was shown in Section 4.9.1 that the energy transfer between MHCα-Fl and MHCβ-Rh was dependent on the peptide:micelle molar ratio, the oligomeric state as a function of this ratio was explored. As shown in Figure 4.19, at the peptide:micelle ratio of 1:1 the data fits best to the calculated line for monomer with a reduced CHI^2 value of 2.26, indicating that there is no association of the peptides at this ratio. This corroborates the findings of Section 4.9.1 where no FRET signal was observed at this ratio. At a ratio of 2:1 the data fits to the calculated curve for a dimer oligomeric state with a reduced CHI^2 value of 3.86 and FRET efficiency of 60%, indicating MHCα-Fl and MHCβ-Rh are assembling into dimers in the DPC micelles. At a ratio of 3:1 the data fits to the calculated curve for a tetramer oligomeric state with a reduced CHI^2 value of 3.17 and FRET efficiency of 30%. Therefore MHCα and MHCβ peptides seem to be associating to form dimers and with decreasing detergent concentration they are assembling into higher order oligomeric states, which is in keeping with the findings for the other TM peptides in this study. This is the first data to show that the MHCα and MHCβ TM domains can associate to form hetero-dimers and hints at a possible role for the TM domain in stabilising the assembly of the full length MHC hetero-dimer.

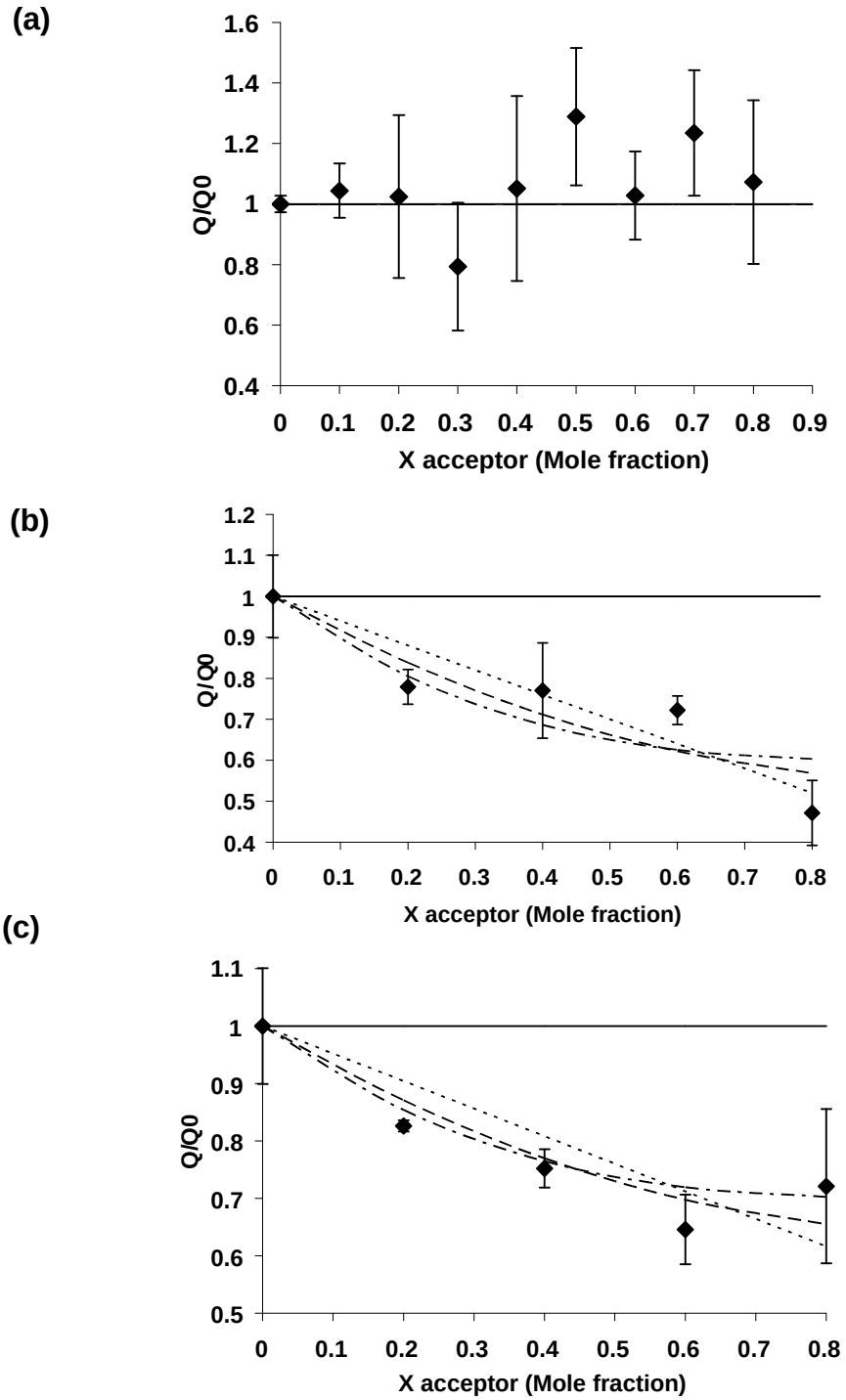


Figure 4.19. Determining oligomeric state of the MHCα and MHCβ association

Stoichiometry of MHCα-Fl and MHCβ-Rh association in DPC detergent at peptide:micelle ratios of (a) 1:1 (b) 2:1 (c) 3:1. The energy transfer was measured and the ratios of emission at 520 nm in the donor only sample (Q0) to that in the FRET sample (Q) were calculated. The value of Q0 was normalised to the mole fraction of donor present in the FRET samples. The ratio of MHCα-Fl to MHCβ-Rh was varied between 0.2 and 0.8 whilst keeping the total peptide concentration constant at 4 μM. Calculated curves for monomer (solid), dimer (dotted), trimer (broken), and tetramer (broken dotted) are shown and were calculated using Equation 5 as described in Section 2.19.6. The goodness-of-fit for the experimental data to the calculated curves was determined using a standard reduced CHI^2 curve fitting procedure.

4.10 Conclusions on the association of the MHC TM domains

MHC Class II proteins are hetero-dimeric α-helical membrane proteins composed of α- and β-chains. The work presented here represents the first studies of the TM domains of the MHC Class II hetero-dimer, here termed MHCα and MHCβ respectively, in isolation from the soluble domains.

MHCα and MHCβ TM domains possess highly conserved dimerisation motifs

Sequence analysis revealed the presence of highly conserved small-xxx-small motifs in MHCα and MHCβ, where for both proteins the small residues are predominantly Gly. MHCα contains two such motifs whilst MHCβ contains just one. The GxxxG and similar small-xxx-small motifs are well known to stabilise dimer formation in TM domains including that of the extensively characterised TM domain of GpA (ref). The presence of these motifs in MHCα and MHCβ TM domains suggests they may be capable of self-association to form dimers. Molecular models of the TM domains of MHCα and MHCβ generated using CHI suggested the GxxxG motifs could mediate helix-helix interactions by packing of the residues at the interface of homo-dimers.

MHCα and MHCβ TM domain sequences can self associate

Using the in vivo assay TOXCAT it was shown that MHCα and MHCβ TM domains are capable of self-associating in *E.coli* membranes. This observation was further corroborated through the use of the GALLEX in vivo assay which also showed these domains are capable of self-association. Intriguingly, the GALLEX assay showed a strong dependence on the length of TM domain sequence used, requiring this parameter to be optimised for both MHCα and MHCβ. Initial attempts with the sequences used in TOXCAT generated negative results with GALLEX but as the sequence was shortened the result resembled that observed with TOXCAT. This length dependence suggests caution should be applied when interpreting data from these in vivo assays and moreover shows why it is necessary to study the association of TM domains with a number of

complimentary techniques in order to reach a consistent result before a conclusion on the association of TM domains can be made. Mutation of the two GxxxG motifs in MHCα monitored by the GALLEX assay confirmed the TOXCAT result that the mutation $G_{232}LG_{236}L$ could disrupt the self association of this TM domain whilst mutation of the other motif did not. The GALLEX results for the GxxxG mutation in MHCβ were a little more inconclusive and seemed to hint that there had been a disruption which was not apparent from the results from the TOXCAT assay.

Purification of α-helical peptide analogues of MHCα and MHCβ

The TOXCAT and GALLEX assays identified that MHCα and MHCβ can self-associate but they are incapable of reporting on the oligomeric state of that interaction e.g. dimer, trimer, tetramer..etc. To explore the oligomeric state of MHCα and MHCβ TM domains required the use of in vitro methods in conjunction with model synthetic peptides. Synthetic peptides corresponding to the TM domains of MHCα and MHCβ were synthesised and purified using standard fmoc chemistry and RP-HPLC, respectively. The MHCα and MHCβ peptides contained a high proportion of hydrophobic residues making them difficult to purify. It was found that formic acid was a better solubilising agent than TFA and that a combination of formic acid and HFIP could be used successfully to purify these highly hydrophobic peptides. Since these peptides were predicted to be TM domains, CD analysis was performed on MHCα and MHCβ peptides which showed that when solubilised in detergents they possessed significant α-helical content and were therefore likely to represent the TM domains of the α and β-chains of MHC.

Detergent sensitive self-association of MHCα and MHCβ revealed by SDS PAGE

SDS PAGE analysis of MHCα and MHCβ TM peptides was performed to identify possible self association since the presence of GxxxG motifs in other TM domains has been known to mediate the formation of SDS stable dimers. The result for MHCα TM domain peptide suggested that this TM domain can form oligomers at a range of peptide concentrations, even in the strongly denaturing detergent SDS,

suggesting strong helix-helix interactions. SDS stable dimers have been observed for other TM domains that have been subsequently been shown to have very specific and strong helix-helix interactions e.g. GpA (Lemmon, Flanagan et al., 1992), E5 (Oates, Hicks et al., 2008). MHCβ TM domain displayed a single band that could not be assigned an oligomeric state until cross-linking analysis was performed.

In order to determine if the denaturing nature of SDS was disrupting the association of these peptides, covalent cross-linking was performed which involves covalently linking the peptides in the milder detergent DPC prior to analysis by SDS-PAGE. Cross-linking studies revealed MHCα TM domain peptide can self associate in detergent micelles to form higher order oligomers above those observed in the absence of cross-linking. The oligomeric state can be seen to be modulated by the concentration of DPC. At low peptide:micelle ratios the observation of laddering is likely due to low solubility of this very hydrophobic sequence. Notably as the peptide:micelle ratio approaches 1:3, MHCα forms dimer, trimer and tetrameric oligomers. Dimer and tetramer are the most prevalent oligomers suggesting that the tetramer is possibly a dimer of dimers, whilst the trimer is possibly a result of incomplete cross-linking of the tetramer. Interestingly, the cross-linking of MHCβ TM domain in DPC reveals it can form higher order oligomers, with dimers and trimer being observed between peptide:micelle ratios of 2:1 and 1:3. The fact that these are not observed in the uncross-linked SDS-PAGE suggests these interactions are weaker than those of MHCα. Bands are not observed below a peptide:micelle ratio of 4:1 suggesting that the peptide is insoluble beyond this ratio. At 2:1 a possible trimer band is observed. The cross-linking data helps to identify peptide:micelle ratios to be used in further experiments. This cross-linking data for MHCα and MHCβ corroborates the observation in vivo of self-association of MHCα and MHCβ TM domain peptides.

GxxxG motifs implicated in self-association of MHCα and MHCβ

CHI models suggested a possible role for the GxxxG motifs in the self-association of MHCα and MHCβ TM domains. To investigate their possible role mutation of these motifs in both TM domains was performed and their effect upon the self-

association monitored with TOXCAT and GALLEX. Mutation of both the GxxxG motifs in MHCα was performed with the Gly residues being changed to the bulkier residue Leu. The TOXCAT assay showed mutation of the G_{225} and G_{229} residues in the motif to the more bulky Leu residues could significantly reduce the self-association of MHCα TM domain. This result was further confirmed by making the same mutation in the homo- GALLEX assay. Mutation of the G232 and G236 residues in the second GxxxG motif in MHCα TM domain did not have an effect upon its oligomerisation as observed in both the TOXCAT and GALLEX assays. This was an interesting result since it has been noted for the TM domain of ErB1, which also possesses two GxxxG motifs that one motifs seemed to play a role in homo-dimerisation whilst the other played a role a in hetero0-dimerisation with ErB2 (Gerber, Sal-Man et al., 2004).

Intriguingly, a similar mutation of the GxxxG motif in MHCβ TM domain did not attenuate the TOXCAT or GALLEX signal suggesting this motif does not play a role in the self association of this TM domain. This raises the important questions of what the interacting residues are and what is the role of the highly conserved GxxxG motif in MHCβ?

The high propensity for self-association exhibited by MHCα and MHCβ is an unexpected finding since this interaction has not been observed in the full length proteins. The fact that MHCα and MHCβ can self-associate has important implications for further studies on the hetero-association and may lead to complications in data interpretation from such studies. The self-association may be due to the TM domain being studied in the absence of the soluble domain which may be the main driving force behind the control of oligomerisation. This work on the self-association of GxxxG containing TM domain sequences adds further evidence to the importance of such motifs in mediating TM domain oligomerisation.

MHCα and MHCβ TM domain analogues can associate to form hetero-oligomers

The preceding discussion focused on the self association of MHCα and MHCβ TM domains and the obvious next step is to determine if they can interact with one another. There are limited techniques available to study hetero-association

and little literature available where hetero-systems have been studied. The options for studying hetero-oligomerisation in vivo are particularly limited. The GALLEX assay was designed for use for studying hetero-interactions but surprisingly has been little used since its conception. FRET has been reported in the literature as also being suitable for studying hetero-association. Both of these techniques were applied in this study to determine if MHCα and MHCβ TM domains were sites of important interactions that would contribute to the stability of the MHC heterodimer.

Using the in vivo assay GALLEX it was shown that the sequences corresponding to the predicted TM domains of MHCα and MHCβ can interact in *E. coli* membranes. Furthermore, it was shown that the N-term GxxxG motif ($G_{225}xxxG_{229)}$ from MHCα and the single motif from MHCβ may play a role in this association since the signal could be attenuated by the double mutation of the Gly residues in this motif to a sterically bulker Leu residue. Use of the molecular modelling software CHI produced a structure where these two motifs were packing at the interface of the hetero-dimer. The purpose of the second GxxxG motif in the TM domain of MHCα is not clear.

MHCα and MHCβ TM domain analogues can associate to form a dimer

Monitoring hetero-association in biomolecules and determining the oligomeric state of that interaction is technically challenging for membrane proteins and few techniques have been developed for making such measurements. FRET has proven to be a useful tool for studying the self-association of model TM domain peptides as shown in the preceding chapter for Ii, and can be easily extended to the problem of monitoring the association of TM domains with differing sequence, by labelling the peptides accordingly with the FRET pair. FRET has been employed in this study in order to determine if peptides corresponding to the TM domains of MHCα and MHCβ can associate in vitro to form hetero-oligomers and to identify the oligomeric state for that interaction.

For fluorophore-labelled peptide analogues of MHCα and MHCβ an energy transfer was observed indicating they are associating in micelles of the detergent DPC. Furthermore, the energy transfer and therefore the association could be

disrupted by decreasing the peptide:micelle molar ratio i.e. increasing the detergent concentration. Since both MHCα and MHCβ have been shown in this study to self-associate it is possible that the observed energy transfer is being modulated by the formation of homo-oligomers. Furthermore the presence of unlabelled MHCβ as a contaminant could be modulating the energy transfer leading to a reduced FRET signal. To see if this is the case would need to analyse the homo interaction by FRET in future work. Using FRET it was also shown that the oligomeric state of the association is a dimer at peptide:micelle ratios of 2:1 and 3:1 indicating that these domains are forming hetero-dimers.

This data is the first indication that the interactions between the α- and β-chains of MHC may be important for the formation and hence function of MHC proteins. This is therefore in keeping with the literature regarding the assembly of the full length MHC α and β-chains prior to associating with Ii. It has been shown in vivo that Ii is essential for optimal presentation of MHC at the cell surface and that α and β-chains can form dimers in the absence of Ii (Elliott, Drake et al., 1994). It has additionally been shown in vitro that in the presence of microsomes (i.e. membrane vesicles formed from the ER by the disruption of eukaryotic cells), α and β-chains can form dimers (Bijlmakers, Benaroch et al., 1994). Therefore it was suggested that a preformed α/β heterodimer binds to Ii (Lamb and Cresswell, 1992; Bijlmakers, Benaroch et al., 1994). However, there is some controversy in the literature regarding the pre-assembly of the MHC hetero-dimer prior to association with Ii, since it has also been shown that Ii can associate with individual MHC subunits (Kvist, Wiman et al., 1982; Lotteau, Teyton et al., 1990; Teyton, Osullivan et al., 1990). This lead to the development of a further model for assembly of the Ii-MHC complex where Ii sequentially binds first to the α-chain then the Ii-α complex selects for matched β-chain to form the export-competent Ii-MHC complex (Koch, McLellan et al., 2007). The data presented in this chapter would seem to support the former model for Ii-MHC assembly.

5 TM domain interactions between Ii and MHC

5.1 Introduction and objectives

As described in Section 1.3.2, an essential first step in the process of antigen presentation is the association of the Ii homo-trimer to three MHC α/β hetero-dimers to form a nonomeric Ii_3(MHC α/β)$_3$ complex within the endoplasmic reticulum of antigen presenting cells. In the preceding chapters it was shown using multiple techniques that the TM domains of Ii can self associate to form oligomeric states including trimer and that the TM domains of MHCα and MHCβ can self-associate and with each other to form hetero-dimers. This implicates these domains as sites of important protein-protein interactions in the formation of the Ii-MHC complex. An association between the TM domain of Ii and full length MHC Class II proteins has been reported by Castellino et al but the exact details of this interaction are unknown (Castellino, Han et al., 2001). Furthermore, a sequential mechanism for the assembly of the Ii-MHC complex has been proposed in which Ii initially binds MHC α subunits, then β-subunits bind to the α-Ii complex (Koch, McLellan et al., 2007). Using the techniques outlined in the preceding chapters we sought to explore the TM domain interactions between Ii and MHC further and also determine if the proposed sequential assembly could be observed for the TM domains.

5.2 Monitoring hetero-association of Ii and MHC TM domains in natural membranes

As discussed in Section 2.18.1**,** the GALLEX assay can be used to monitor the hetero-association of TM domain sequences, making it an obvious choice for identifying possible TM domain interactions between Ii and the α- and β-chains of MHC. The sequence for the Ii TM domain was cloned into the plasmid pALM100 as described in Section 2.18.2, whilst the pBLM constructs of MHCα and MHCβ TM domains with a length of 19 residues, constructed for the work in Section 4.3, were used again. For the purposes of performing the assay pALM- and pBLM-

constructs were combined in the host strain *E.coli* SU202, the fusion protein from the two plasmids were simultaneously expressed, and the activity of the reporter gene β-Galactose (β-gal) assayed as described in Section 2.18.

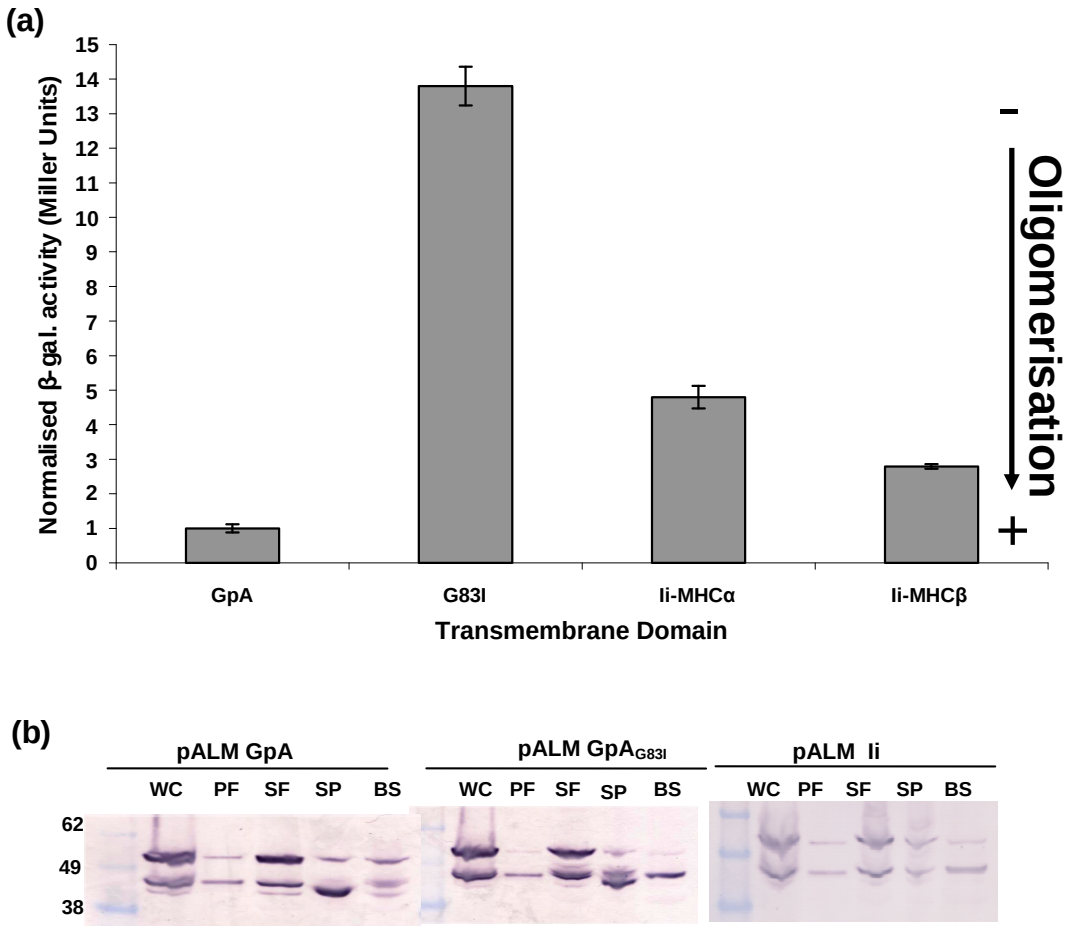


Figure 5.1. Hetero-association between TM domains of the Ii-MHC α/β complex

Hetero-association was monitored using the GALLEX assay as described in the text. Amino acid sequence of the Ii TM domains of Ii, GpA and GpA$_{G83I}$ were cloned into the plasmid pALM100 and subsequently transformed into *E.coli* strain SU202 along with the pBLM constructs of GpA, GpA$_{G83I}$, MHCα and MHCβ as described in Section 2.18.2. (a) β-galactosidase activity mediated by the oligomerisation propensity of the expressed constructs in *E.coli* SU202. Repression of activity is indicative of association of the TM domains. Data is an average from three independent measurements. Expression of the chimeric proteins was induced by the addition of 1 mM IPTG. Details of the β-galactosidase assay and the calculation of Miller Units are described in the Section 2.18.3. All plasmids and *E. coli* strains were kindly proved by Dirk Schneider. GpA and the dimerisation deficient mutant of GpA, G83I, act as positive and negative controls respectively. Error bars represent the standard error from three separate measurements (b) Test for insertion and orientation of the expressed chimera from pALM plasmid. The similar tests for the pBLM constructs were reported in the preceding chapters. Western blot analysis of E.coli extracts: WC, whole cell; PF, periplasmic fraction; SF, spheroplast fraction; SP, spheroplast proteolysis; BS, broken spheroplast proteolysis. The expressed chimeric proteins with a molecular mass of ~54kDa are found solely in the inner membrane fraction and correctly oriented in the membrane.

As shown in Figure 5.1a, the combination of Ii and MHCα TM domains leads to repression of β-galactosidase activity suggesting these two TM domains are associating within the inner membrane of *E .coli*. Similarly, the combination of Ii with MHCβ results in the repression of β-galactosidase activity also suggesting an association between these TM domains. These data indicate that, in isolation, the TM domains of the Ii and α and β-chains of proteins are sites of significant protein-protein interactions, which has important implications for the role of the TM domains of these proteins in the formation of the Ii-MHC complex. These findings support those of Castellino et al (Castellino, Han et al., 2001). Through the use of mutagenesis studies, it may be possible in future work to identify the residues that are mediating this interaction.

5.3 Analysis of self-association of Ii and MHC TM domains by cross-linking

In the preceding chapters, the method of covalently cross-linking peptides derived from the TM domains of Ii, α and β-chains of MHC in mild detergents was used to monitor self-association using SDS-PAGE. In an attempt to determine if cross-linking could be used to monitor hetero-association, mixtures of Ii, MHCα and MHCβ peptides were dissolved in DPC micelles and cross-linked with BS^3, as described in Section 2.13. Given the similar size of the peptides and the fact that they each exhibit self-association it was expected that this form of analysis would be difficult to interpret. When performing this analysis, we initially cross-linked each peptide independently, and then cross-linked the mixture. A result of no interaction was assigned if the banding pattern in the mixture was simply a combination of the bands from each component. Any differences between the banding patterns were attributed to association of the different peptides.

The results from cross-linking peptides separately and mixed and are shown in Figure 5.2b lanes 1-7 and uncross-linked samples are shown in Figure 5.2b lanes 1a-7a. For cross-linking of MHCα and MHCβ, comparison of lane 3 (MHCα plus MHCβ) to lanes 1 (MHCα) and 2 (MHCβ), shows the loss of a band at 14 KDa and possibly a more intense band at ~6 KDa (see highlighted regions of lane 3),

although the latter band could be the two dimers of each peptide coinciding in the gel. These differences possibly indicate the MHCα and MHCβ peptides have assembled into hetero-dimers. For the cross-linking of Ii and MHCα, comparison of lane 5 (Ii plus MHCα) to lanes 1 (MHCα) and 4 (Ii) does not reveal any significant differences due to the large number of bands, it is not possible to resolve any changes. For the cross-linking of Ii and MHCβ, comparison of lane 6 (Ii plus MHCβ) to lanes 2 (MHCβ) and 4(Ii) reveals the loss of the highest order band observed for Ii peptide alone. It is still possible to see the MHCβ dimer band and possibly the monomer band, indicating that any association is only weak.

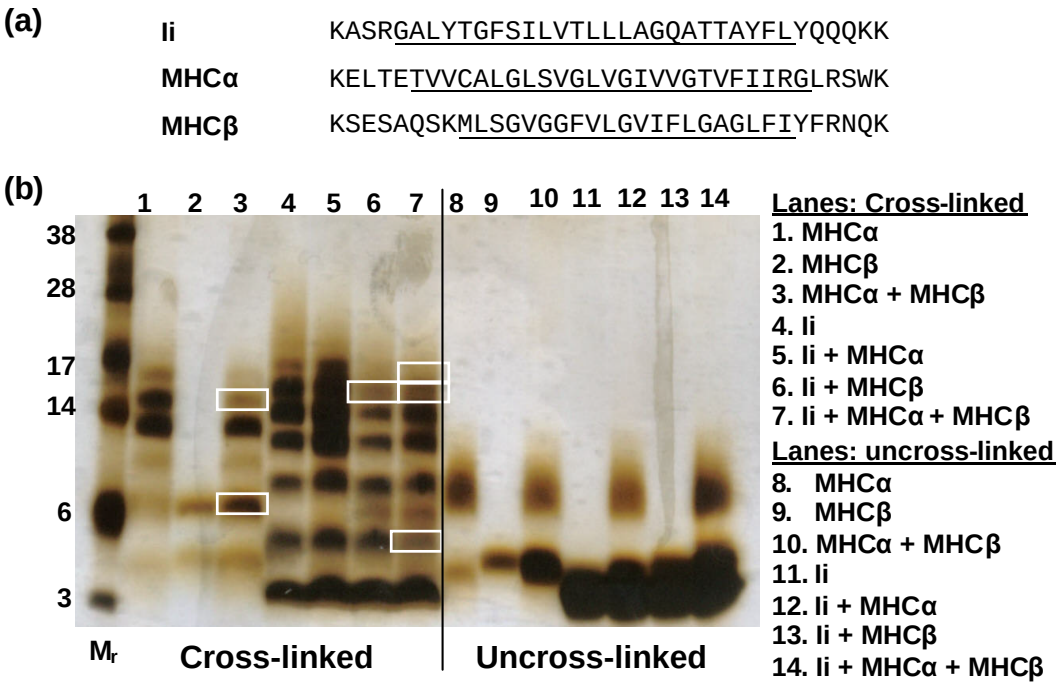


Figure 5.2. Analysis of hetero-association of Ii, MHCα and MHCβ peptides

(a) Amino acid sequence of the synthetic peptides used in cross-linking studies, corresponding to the TM of Invariant chain (Ii) and α and β chains of the Major Histocompatibility Complex Class II proteins (MHCα and MHCβ, respectively). Predicted TM domains are underlined. Additional Lys residues are added at the N and C term to aid solubility and avoid non-specific aggregation. (b) Lanes 1-7 show BS^3-mediated cross-linking of TM peptides dissolved in DPC at 3:1 peptide:micelle ratio also shown are peptides uncrosslinked, lanes 8-14. Molecular weight markers are shown in the far left-hand lane in kDa. Protein bands were visualized by staining with silver nitrate. Differences between lanes of hetero- and homo- cross-linked peptides are indicated by white boxes.

Finally, the interaction between all of the three TM domains was studied. For the cross-linking of Ii, MHCα and MHCβ, comparison of lane 7 (Ii plus MHCα and MHCβ) to lanes 1, 2 and 3 shows the loss of the higher order bands observed for

MHCα and Ii. Interestingly, there is a loss of the band corresponding to dimer for all three peptides suggesting the peptides have associated into higher order oligomers.

Despite the obvious difficulties inherent in using the approach of cross-linking to probe the association of TM domain peptides of similar mass, we can clearly see that it is possible to discern differences between the peptides when cross-linked separately and mixed. Although these results are not conclusive and are challenging to interpret they do suggest that Ii may be associating with MHCα and MHCβ.

5.4 FRET analyses to measure the interactions between the TM domains of Ii and MHC

In order to explore the interactions of the Ii and MHC TM domain peptides further, FRET analyses were performed. FRET analyses have been used so far in this study to confirm the self-association of Ii (Section 3.7), and reveal that MHCα and MHCβ TM domains can associate to form hetero-dimers (Section 4.9). This section describes the use of FRET analyses to determine if there are interactions between the TM domain of Ii, and those of MHCα and MHCβ.

5.4.1 FRET sample preparation

The synthesis and purification of fluorophore labelled peptides of Ii-Fl, Ii-Rh, MHCα-Fl and MHCβ-Rh, has been described previously in Sections 3.7 and 4.5 respectively. For the purposes of FRET measurements these peptides were combined as appropriate to generate a FRET pair of fluorescein- and rhodamine-labelled peptides (e.g. Ii-Fl and MHCβ-Rh). All FRET samples were prepared using the co-dissolving method as described in Section 3.7.2.

5.4.2 Monitoring FRET between MHCα and Ii peptides and its dependency on the peptide:micelle ratio

FRET samples of MHCα-Fl and Ii-Rh were prepared at peptide:micelle molar ratios of between 1:3 and 4:1 in the detergent DPC, whilst keeping the total peptide concentration constant at 8 μM. For each ratio the percentage energy

transfer was calculated as described in Section 2.19.4 and plotted versus the peptide:micelle molar ratio.

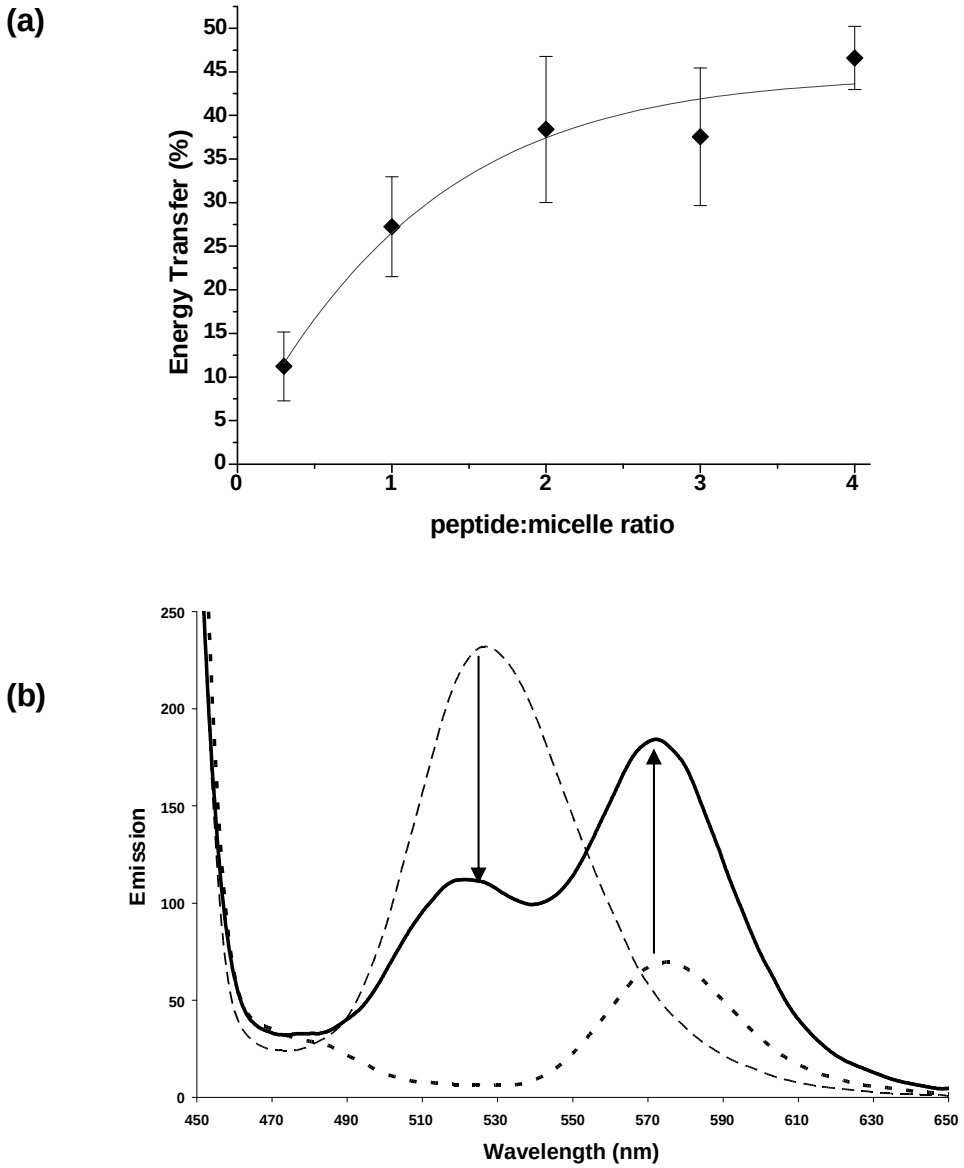


Figure 5.3. Change in energy transfer between MHC*α* and Ii with detergent concentration

Plot of Energy transfer (%) versus peptide:micelle ratio (a) Total donor and acceptor peptide concentrations were kept constant at 8 μM (4 μM MHCα-Fl, 4 μM Ii-Rh) while the detergent concentration was varied. Samples were prepared using the co-dissolving methods by mixing peptide and DPC pre-solubilised in TFE. A CMC of 1 mM and aggregation number of 56 were used in calculations of the peptide:micelle ratio. Emission spectra were collected and energy transfer was calculated as described in Section 2.19.6. (b) FRET spectra for MHCα-Fl and Ii-Rh in DPC micelles at a peptide:micelle ratio of 4:1. The broken line is spectrum of MHCα-Fl on its own, the dotted line is spectrum of MHCβ-Rh on its own, whilst the solid line is the spectrum for a mixture of MHCα-Fl and Ii-Rh. The FRET signal is evident from the decrease in the donor emission at 520 nm and an increase in acceptor emission at 570 nm as indicated by the arrows.

As shown in Figure 5.3, energy transfer was observed between the MHCα-Fl and Ii-Rh which indicates that these peptides are associating within DPC micelles. The greatest energy transfer is observed at peptide:micelle ratios of greater than 1 whilst below this ratio no or minimal energy transfer is observed. A maximum energy transfer of around 50% is observed between ratios of 1 and 4. These data agree well with GALLEX data and suggests the TM domains of Ii and the α-chain of MHC are associating.

5.4.3 Determining specificity of MHCα-Fl and Ii-Rh FRET signal by competition with unlabelled peptide

In order to ascertain if the FRET signal measured between MHCα-Fl and Ii-Rh arises from a specific interaction, and was not due to random co-localisation of the peptides merely occupying the same micelle and thus being in close proximity, it is necessary to do a competition assay with unlabelled peptide. Since this experiment must be carried out at a fixed detergent concentration the peptide:micelle molar ratio of 4:1 was chosen on the basis that it produced the highest energy transfer between Ii and MHCα. The concentrations of MHCα-Fl and Ii-Rh peptides were kept constant whilst varying the concentration of unlabelled Ii peptide. As shown in Figure 5.4, the energy transfer decreases with increasing concentration of unlabelled peptide. This is indicative of the unlabelled Ii peptide competing for binding sites on MHCα-Fl and disrupting the formation of donor and acceptor partners.

In order to confirm the specificity of the interaction between MHCα and Ii, the effect of adding unlabelled MHCβ peptide on the energy transfer between MHCα-Fl and Ii-Rh was explored. As shown in Figure 5.5, the addition of MHCβ had little or no impact upon the energy transfer between MHCα-Fl and Ii-Rh, requiring the addition of 10 μM unlabelled MHCβ to reduce the energy transfer by ~15%. This indicates that the peptides are not simply being forced together by co-localisation in the micelles. Furthermore since it has been shown that MHCα can interact with MHCβ this result suggests that MHCβ is not competing for the interaction interface between MHCα and Ii.

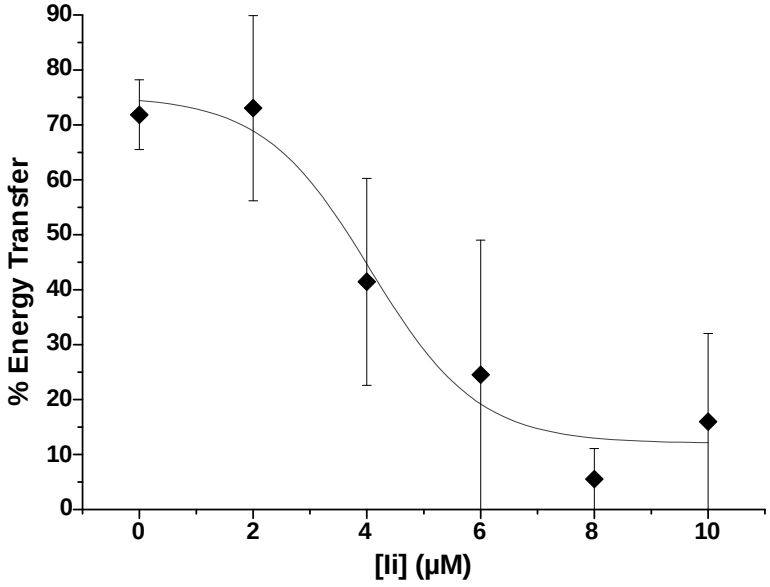


Figure 5.4. Effect of unlabelled Ii on FRET between MHCa-Fl and Ii-Rh

Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM MHCα-Fl, 2 μM Ii-Rh) while the concentration of unlabelled Ii was varied. Samples were prepared by co-dissolving all peptides and DPC dissolved in TFE. Experiments were performed at a peptide:micelle ratio of 4:1. The reduced FRET efficiency suggests that sequence-specific oligomerisation contributes to the measured FRET efficiency.

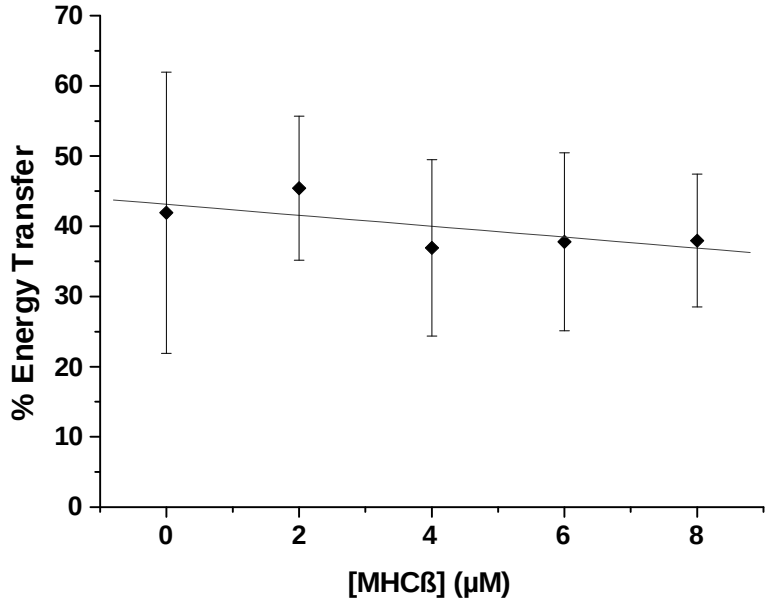


Figure 5.5. Effect of unlabelled MHCβ on FRET between MHCa-Fl and Ii-Rh

Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM MHCα-Fl, 2 μM Ii-Rh) while the concentration of unlabelled MHCβ was varied. Samples were prepared by co-dissolving all peptides and DPC dissolved in TFE. Experiments were performed at a peptide:micelle ratio of 3:1.

5.4.4 Determining the oligomeric state of MHCα and Ii TM domain association by FRET analysis

The measurement of energy transfer as a function of the mole fraction of acceptor was used as described in Section 2.19.5, to determine the oligomeric state of the MHCα and Ii hetero-oligomers at varying peptide:micelle molar ratios. As shown in Figure 5.6, at a peptide:micelle ratio of 1:1 the data fits to the calculated curves for tetramer oligomeric state with a reduced CHI^2 value of 0.66 and FRET efficiency of 52%, indicating MHCα-Fl and Ii-Rh are assembling into tetramers in the DPC micelles. Similarly, at a ratio of 2:1 the data also fits best to tetramer with a CHI^2 value of 0.68 and a FRET efficiency of 55%. As the peptide:micelle molar ratio is increased to 3:1 the data fits to the higher order oligomeric state of pentamer with a CHI2 value of 1.68 and a FRET efficiency of 54%, showing that the detergent can modulate the oligomeric state of the Ii-MHCα hetero-oligomer in a manner

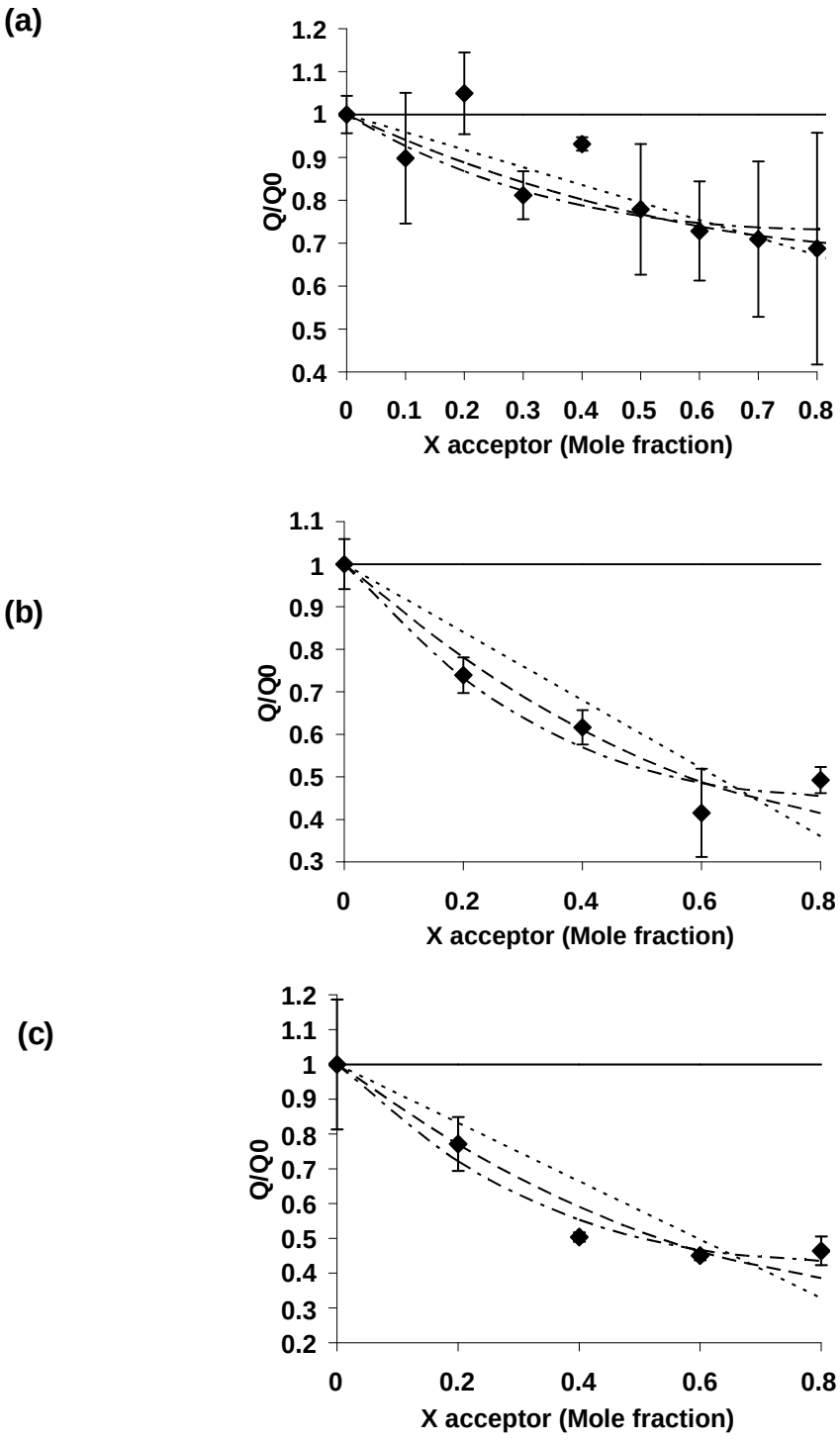


Figure 5.6. Oligomeric state of MHCα and Ii interaction

Stoichiometry of MHCα-Fl and Ii-Rh association in DPC detergent at peptide:micelle ratios of (a) 1:1 (b) 2:1 (c) 3:1. The energy transfer was measured and the ratios of emission at 520 nm in the donor only sample (Q0) to that in the FRET sample (Q) were calculated. The value of Q0 was normalised to the mole fraction of donor present in the FRET samples. The ratio of MHCα-Fl to Ii-Rh was varied between 0.2 and 1.0 whilst keeping the total peptide concentration constant at 4 μM. Calculated curves for monomer (solid), dimer (dotted), trimer (broken), and tetramer (broken dotted) are shown and were calculated using Equation 5 as described in Section 2.19.6. The goodness-of-fit for the experimental data to the calculated curves was determined using a standard reduced CHI^2 curve fitting procedure.

5.4.5 Monitoring FRET between MHCβ and Ii peptides and its dependency on the peptide:micelle molar ratio

In order to determine if the TM domain of Ii could associate with that of MHCβ, FRET analyses were performed using the fluorophore labelled peptides Ii-Fl and MHCβ-Rh. FRET was monitored at varying peptide:micelle molar ratios since it was expected this would impact upon any FRET signal observed. As shown in Figure 5.7, a FRET signal was observed with a maximum efficiency of ~35% for peptide:micelle molar ratios above 4:1. This is slightly lower than the values observed in our previous FRET studies, which were around 50-60%. Notably, the spectra for Ii-MHCβ at a ratio of 3:1, shown in Figure 5.7b, displays a very small increase in acceptor emission at 570 nm compared to those observed in previous FRET experiments. Furthermore, it is interesting that a FRET signal was observed at peptide:micelle ratios above 4:1 since in cross-linking experiments MHCβ does not appear to be soluble at these concentrations. These observations make it difficult to conclude that the energy transfer observed between Ii-Fl and MHC-Rh constitutes a real FRET signal. It is possible that the FRET signal is being reduced by the presence of unlabelled MHCβ peptide which could not be completely removed during the purification of MHCβ-Rh (Section 4.5). We therefore sought to validate the energy transfer observed between Ii and MHCβ using competition assays.

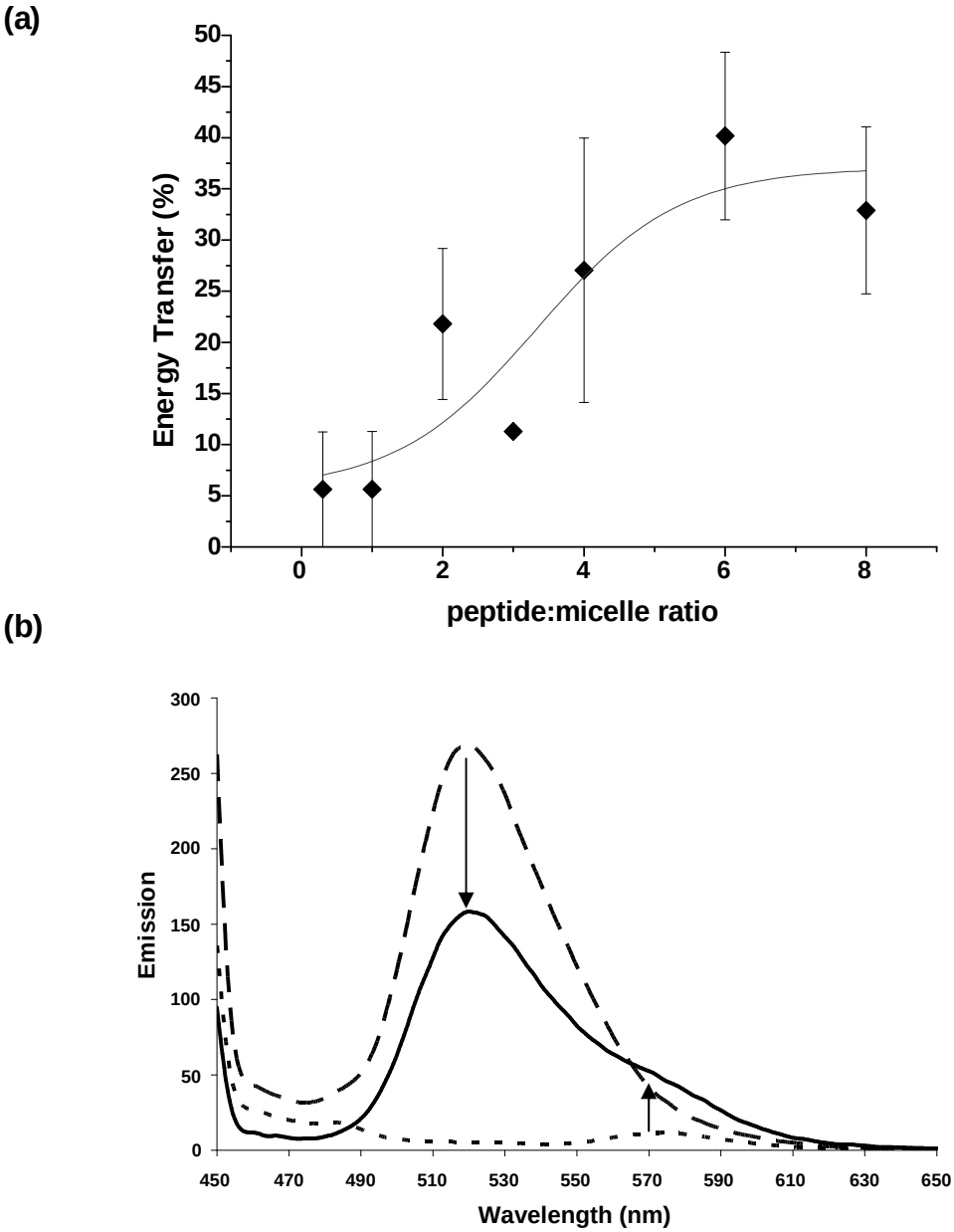


Figure 5.7. Change in FRET between Ii-MHCβ with detergent concentration

(a) Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM Ii-Fl, 2 μM MHCβ-Rh) while the detergent concentration was varied. Samples were prepared using the co-dissolving methods by mixing peptide and DPC pre-solubilised in TFE. A CMC of 1mM was used in calculations of the peptide:micelle molar ratio. Emission spectra were collected and energy transfer was calculated as described in Section 2.19.6 (b) FRET spectra for Ii-Fl and MHCβ-Rh in DPC micelles at a peptide:micelle molar ratio of 3:1. The broken line is a spectrum of Ii-Fl on its own, the dotted line is a spectrum of MHCβ-Rh on its own, whilst the solid line is the spectrum for a mixture of Ii-Fl and MHCβ-Rh. A possible FRET signal is evident from the decrease in the donor emission at 520 nm and an increase in acceptor emission at 570 nm as indicated by the arrows.

5.4.6 Determining specificity of Ii-MHCβ FRET signal by competition with unlabelled MHCβ peptide

In order to determine if the FRET observed between Ii-Fl and MHCβ-Rh arises from a specific interaction, a titration with unlabelled MHCβ peptide was performed. Since this experiment must be carried out at a fixed detergent concentration the peptide:micelle molar ratio of 4:1 was chosen, as measurement at this ratio resulted in the largest energy transfer, as described in the preceding section. The concentrations of Ii-Fl and MHCβ-Rh peptides were kept constant whilst varying the concentration of unlabelled MHCβ peptide. As shown in Figure 5.8, although the error bars for this experiment are quite large, it appears that the energy transfer is not affected by the presence of any concentration of unlabelled MHCβ peptide. This strongly suggests the unlabelled MHCβ peptide is not competing with the MHCβ-Rh peptide for binding to the Ii-Fl peptide and that the observed FRET is merely an experimental artefact and not indicative of an interaction between Ii and MHCβ.

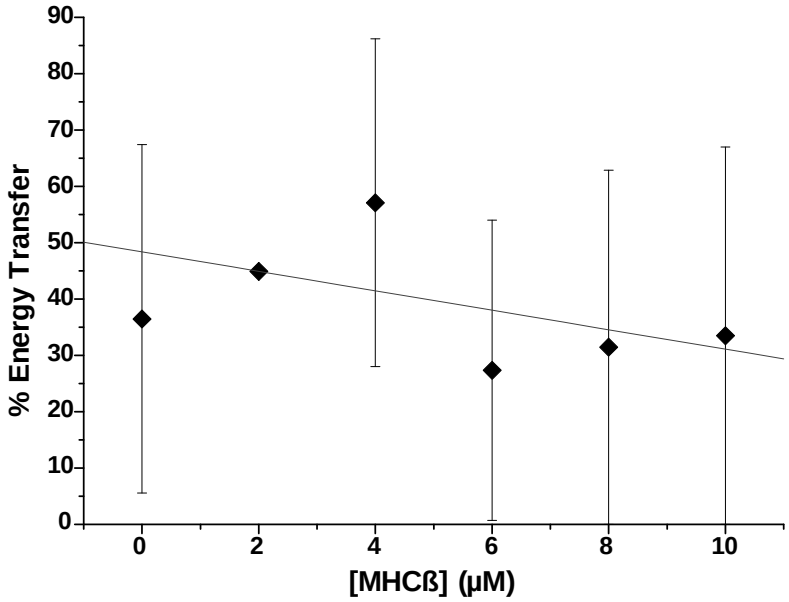


Figure 5.8. Effect of unlabelled MHCβ on FRET between Ii and MHCβ

Energy transfer between Ii-Fl and MHCβ-Rh peptide as a function of concentration of unlabelled MHCβ peptide. Total donor and acceptor peptide concentration was kept constant at 4 µM (2 µM Ii-Fl, 2 µM MHCβ-Rh) while the concentration of unlabelled MHCβ was varied. Samples were prepared by co-dissolving all peptides and DPC dissolved in TFE. Experiments were performed at a peptide:micelle molar ratio of 3:1.

5.4.7 Determining the oligomeric state of MHCβ and Ii association by FRET analysis

To further investigate the possibility of an interaction between Ii-Fl and MHCβ-Rh the measurement of energy transfer as a function of the mole fraction of acceptor was monitored at varying peptide:micelle molar ratios, as described in Section 2.19.5. As shown in Figure 5.9, at a peptide:micelle ratio of 1:1 the data fits best to a calculated curve for dimer with a CHI^2 value of 0.69 but with a FRET efficiency of just 20 %, suggesting a possible interaction between Ii and MHCβ peptide. At higher ratios of 2:1 and 3:1 however, the data fits best to monomer oligomeric state indicating no interaction is occurring. From this data and the preceding FRET data it is difficult therefore to confirm that Ii and MHCβ are specifically interacting since the general trend observed for the other TM peptides in this study of increasing oligomeric state with increasing peptide:micelle ratio was not observed. The observation of an interaction at a peptide:micelle ratio of 1:1 and not at higher ratios may result from poor solubility of the MHCβ peptide in DPC since it was found previously in the cross-linking analyses (see Figure 4.6b) that this peptide has limited solubility above a peptide:micelle ratio of 2:1 in this detergent. However, the observation of an interaction however weak it may be does fit with the findings from the GALLEX assay in this study. Further study will be required before conclusions can be drawn on whether the peptide models of the Ii and MHCβ TM domains also display this propensity for hetero-association.

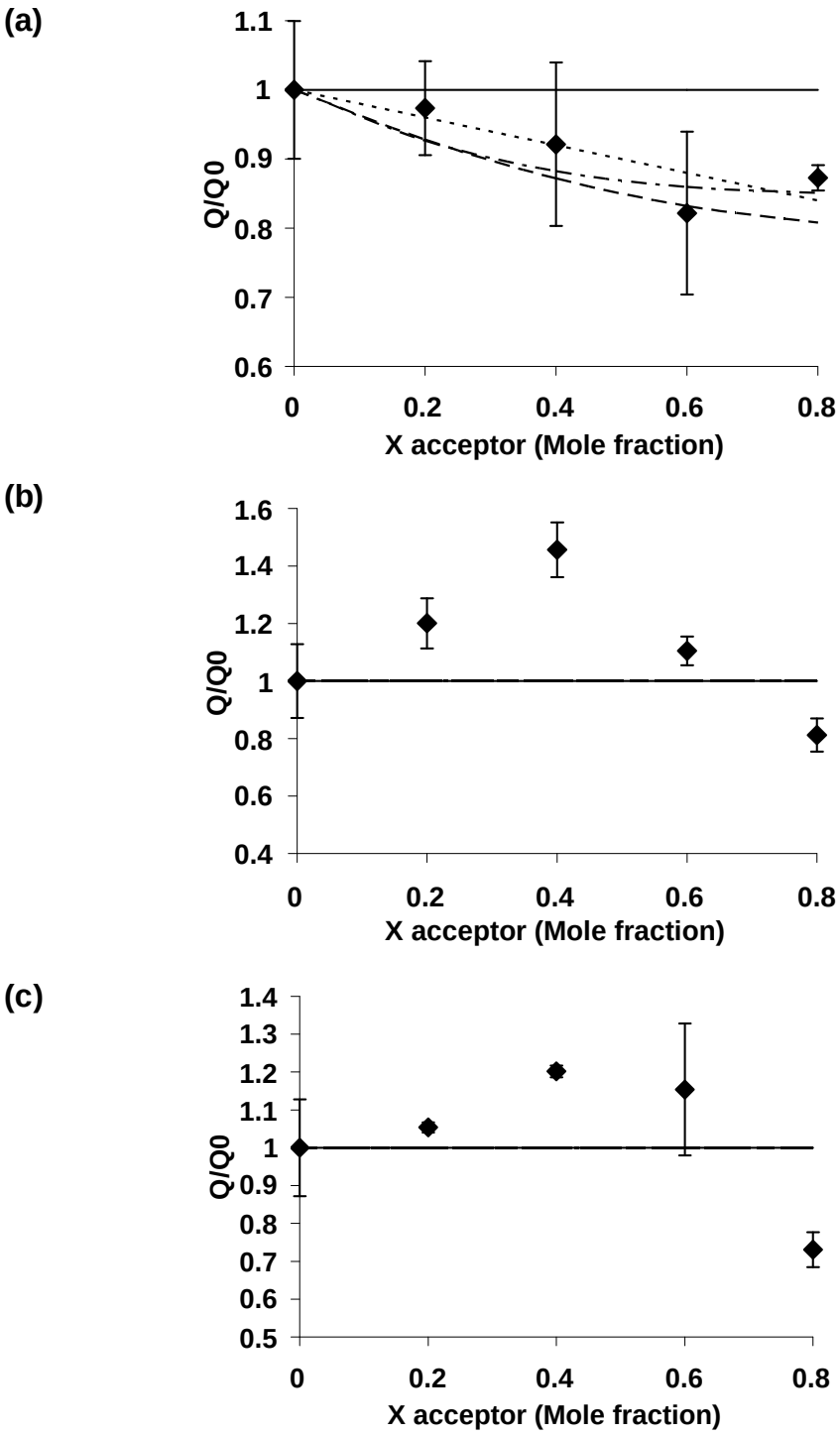


Figure 5.9. Determining oligomeric state of Ii MHCβ TM domain association

Stoichiometry of Ii-Fl and MHCβ-Rh in DPC detergent at peptide:micelle ratios of (a) 1:1 (b) 2:1 (c) 3:1. The energy transfer was measured and the ratios of emission at 520 nm in the donor only sample (Q0) to that in the FRET sample (Q) were calculated. The value of Q0 was normalised to the mole fraction of donor present in the FRET samples. The ratio of Ii-Fl to MHCβ-Rh was varied between 0.2 and 1.0 whilst keeping the total peptide concentration constant at 4 μM and the peptide: Calculated curves for monomer (solid), dimer (dotted), trimer (broken), and tetramer (broken dotted) are shown and were calculated using Equation 5 as described in Section 2.19.6. The goodness-of-fit for the experimental data to the calculated curves was determined using a standard reduced CHI^2 curve fitting procedure.

5.4.8 Determining effect of adding unlabelled peptide MHCα on the FRET between Ii and MHCβ

It has been suggested that in the assembly of the full-length Ii-MHC complex, the initial step is the binding of the Ii trimer to individual MHCα chains followed by the Ii-MHCα complex binding MHCβ chains (Koch, McLellan et al., 2007). Since it has been shown in this study that the TM domains of these proteins are sites of important protein-protein contacts, it was hypothesised that this order of interaction may be observable in the TM domains of these proteins using FRET analyses. If this is the case, then it would be expected that the observed energy transfer between Ii and MHCβ will increase upon the addition of MHCα.

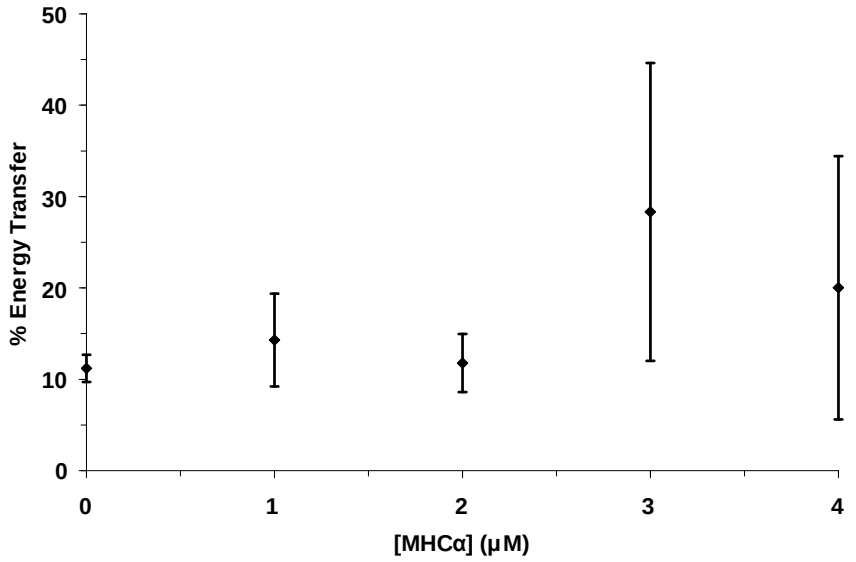


Figure 5.10. Effect of MHCα on the association of Ii and MHCβ

Energy transfer between Ii-Fl and MHCβ-Rh as a function of MHCα concentration. Total donor and acceptor peptide concentration was kept constant at 4 μM (2 μM Ii-Fl, 2 μM MHCβ-Rh) whilst unlabelled MHCα was added at increasing concentrations. Samples were prepared using the co-dissolving methods by mixing peptide and DPC pre-solubilised in TFE to give a peptide:micelle ratio of 2:1. A CMC of 1mM was used in calculations of the peptide:micelle ratio. Excitation spectra were collected and energy transfer was calculated as described in materials and methods.

In order to test this hypothesis, increasing concentrations of unlabelled MHCα peptide were added to a mixture of Ii-Fl and MHCβ-Rh peptides. The experiment was carried out in DPC at a peptide:micelle molar ratio of 2:1. This ratio was chosen on the basis that it had yielded a low energy transfer between Ii and MHCβ previously and was able to solubilise all the peptides. As shown in Figure

5.10, the addition of unlabelled MHCα did not enhance the energy transfer between Ii and MHCβ, thus suggesting that the TM domain of MHCβ does not play a role in the proposed sequential assembly of the Ii-MHC complex.

5.5 Conclusions on TM domain association of Ii and the α- and β-chains of MHC

The full length Ii trimer is known to associate with three MHC Class II hetero-dimers containing α (MHCα) and β (MHCβ) chains. This interaction is known to be mediated by the association of the CLIP domain of Ii occupying the antigen binding site of the MHC α/β hetero-dimer. Castellino et al mapped a further site of interaction between Ii and MHC to the transmembrane segment of Ii using mutagenesis (Castellino, Han et al., 2001).

The work presented in Chapters 3 and 4 of this study reveals that the TM domains of Ii and MHCα and MHCβ proteins are sites of potentially important helix-helix interactions that may further stabilise the Ii-MHC complex. Using the limited number of methods available for studying hetero-interactions between hydrophobic proteins we sought to determine if in isolation the TM domain of Ii could associate with those of MHCα and MHCβ and determine the stoichiometry of these interactions. This will verify the findings of Castellino et al. and provide the basis for further study into identifying the residues involved in these interactions.

A method for studying the hetero-association of isolated TM domains *in vivo* is the GALLEX assay. Use of this assay suggested that the TM domain of Ii can interact with those of MHCα and MHCβ in the *E.coli* inner membrane. This is the first time that such an interaction has been observed between these proteins. It is important to remember that GALLEX can only report on association and does not provide details on the oligomeric state of the interaction so that information has to be obtained using other techniques.

In order to confirm the findings from the GALLEX assay, and determine the stoichiometry of these interactions, model peptides derived from the TM domains of Ii, MHCα and MHCβ were produced and their association was monitored using

biophysical techniques. Chemical cross-linking of the peptides in a mild detergent hinted at the possible association of these domains since differences could be observed between the peptides cross-linked separately and when mixed. The analysis of such an experiment is complicated by to the similar size of the peptides under study and the complications involved when those peptides can self-associate as is the case for Ii, MHCα and MHCβ. Cross-linking is not frequently employed to study hetero-interactions of TM domains for this reason.

FRET analyses of fluorophore labelled model peptides suggested that the TM domains of Ii and MHCα could associate. Furthermore, the oligomeric states of that association could be determined and were found to from dimers or trimers depending on the peptide:micelle molar ratio. FRET analyses also suggested that the TM domain of Ii may interact with that of MHCβ but the findings were difficult to interpret conclusively. It is possible the solubility of the MHCβ peptide in the chosen detergent DPC, is an issue for this type of analyses. Intriguingly, it was found that the FRET between Ii with MHCβ could not be enhanced by the presence of MHCα. This data, is consistent with the finding in the full length proteins that individual α-chains coisolate with Ii, whereas β-chains exhibit only a low-affinity interaction with Ii (Neumann and Koch, 2005). Therefore, the TM domain of MHCβ would not appear to play a role in the assembly of the Ii-MHC complex. It should be noted that given the findings from the *in vivo* analysis of the association of Ii and MHCβ performed using the GALLEX assay (see Section 5.2), which suggested they were interacting, we would expect to see significant interaction in the FRET analyses also. The reasons for this discrepancy are uncertain at present, but it is possible that they may reflect the differing environment in which the analyses were performed, since the GALLEX assay is performed on proteins inserted into the inner membrane of *E.coli* compared to peptides solubilised in detergent for FRET measurements. Further study will be required to conclusively determine if the model peptides of Ii and MHCβ TM domains are interacting.

6 Developing NMR Methods for investigating protein interactions

Identifying helix-helix interactions and the structural determinants that drive α-helical membrane protein folding is a technically challenging problem due to their hydrophobicity. Of particular interest is accessing detailed structural information regarding the interacting side chains and hence enabling identification of the non-covalent bonds stabilising the association of the helices. Using current methods for studying TM interactions (e.g. TOXCAT, GALLEX, Cross-linking, FRET, AUC) it is possible to identify homo- and hetero- association between TM domains and to also determine their oligomeric state, as has been shown in the previous chapters of this work. Current studies on the hetero-association of TM domains have used the GALLEX assay, FRET experiments and immune-precipitation. Complementing these techniques, molecular modelling and mutagenesis studies can suggest residues that are important for the interactions. However, the results from these studies in our experience are often difficult to interpret conclusively. Furthermore, ultimately these methods, though powerful, do not provide information on the arrangements of atoms in these oligomers and ultimately the definitive test for any predictions made from such studies is to solve the atomic structure of the domains.

Currently the two most successful methods for solving the atomic structures of membrane proteins are X-ray crystallography and NMR spectroscopy. In particular, solution-state NMR spectroscopy has been applied to the study of transmembrane domains in isolation, which due to their small size are most amenable to analysis by this technique. We therefore sought to develop protocols for using NMR to solve the structure of TM domain oligomers and additionally designed a novel assay for determining the association of these domains that also has the potential to provide important structural information.

This section describes work carried out in this study on developing these NMR – based methodologies for identifying interacting α-helical TM domains and for determining the atomic structure of those oligomers. Preliminary work has been

performed on solving the structure of a TM domain from the well characterised E5 protein and on developing paramagnetic NMR methods for the rapid determination of TM domain structure. It is hoped that the results from these studies will facilitate future investigation of the Ii, and MHC proteins.

6.1 Towards solving the structure of the TM domain of E5

The E5 protein from Bovine Papillomavirus is the product of the smallest known oncogene at only 44 amino acid residues in length. E5 triggers tumour formation through activation of the platelet derived growth factor β receptor (PDGFβr) within the plasma membrane of host cells. Recently a 26-residue segment of this membrane protein, encompassing the α-helical TM domain, has been shown to form strongly interacting homo-dimers even in SDS that are stabilized by non-covalent helix-helix interactions (Oates, Hicks et al., 2008). AUC analysis of the E5 TM domain dimer revealed a ΔGapp for association of 7.4 Kcal mol^{-1} at 25 °C in DPC (Oates, Hicks et al., 2008). Interfacial residues that play a role in stabilising the dimer were predicted from in-vivo mutagenesis studies and computational models (Oates, Hicks et al., 2008). To date there are no atomic level structural data available on the E5 dimer identifying residues that stabilise its formation. This is due mainly to in vitro studies of these hydrophobic systems presenting significant technical challenges in terms of synthesis and purification and the requirement of NMR compatible membrane mimetics and sample optimisation.

Amphipathic detergents such as SDS are commonly used as membrane mimetics for in vitro membrane protein studies, provided it can be shown that the SDS will not disrupt oligomer formation for the protein of interest (which is common for SDS). However, when considering the use of detergents as solubilising agents for membrane proteins the micelle concentration (i.e. the concentration above which detergent monomers aggregate to form micellar structures) is often overlooked in favour of the bulk detergent concentration or specific physicochemical properties of the particular detergent. Strong detergents such as SDS are known to destabilise membrane protein structure and have been shown to modulate the oligomerisation of TM domains (Fisher, Engelman et al., 1999). Detergent

concentration has been shown to modulate the oligomerisation of TM domains but not helicity as shown for the dimerisation of the protein Glycophorin A (Fisher, Engelman et al., 1999). However, this work considered only the total detergent concentration and did not explicitly consider the detergent micelle concentration. We consider the micelle concentration here and show that it is crucial for maintaining non-covalent interactions in oligomeric species.

In this study we have performed a systematic investigation on the effect of SDS detergent micelle concentration on the NMR spectra of the transmembrane domain of E5 with a view to atomic level structural information for this domain using solution state NMR techniques. We also provide a rationale for the effect of micelle concentration on these oligomeric systems.

6.1.1 Synthesis of E5 Peptide and its purification

Previous studies revealed that a synthetic peptide analogous to the TM domain of E5 is able to self-associate to form a homo-dimer in detergent solutions (Oates, Hicks et al., 2008). This strategy of using synthetic peptides is advantageous for analysis by NMR spectroscopy since it enables NMR-active isotopes to be incorporated at specific positions. Two selectively ^{15}N-labelled peptides corresponding to residues F9 to H34 and encompassing the TM domain of E5 were synthesised with the sequences presented in Figure 6.1**,** as described in Section 2.10.

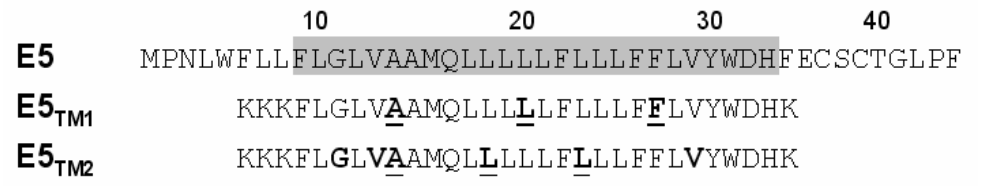


Figure 6.1. Primary sequence of full length E5 from Bovine Papillomavirus and synthetic peptides used in this study

Amino acid sequence of full length E5 with the predicted transmembrane domain highlighted in grey. The synthetic peptides $E5_{TM1}$ and $E5_{TM2}$ analogous to the TM domain of E5 were produced at the Keck facility (Yale University, USA). Underlined residues in the peptides are those predicted to be at the interface (Mattoon, Gupta et al., 2001). Backbone amide nitrogen atoms of residues highlighted in bold were selectively ^{15}N labelled during synthesis. In both peptides Ala12 was ^{15}N labelled as a control. Lysine residues were incorporated at the termini of the peptides to increase solubility and aid purification.

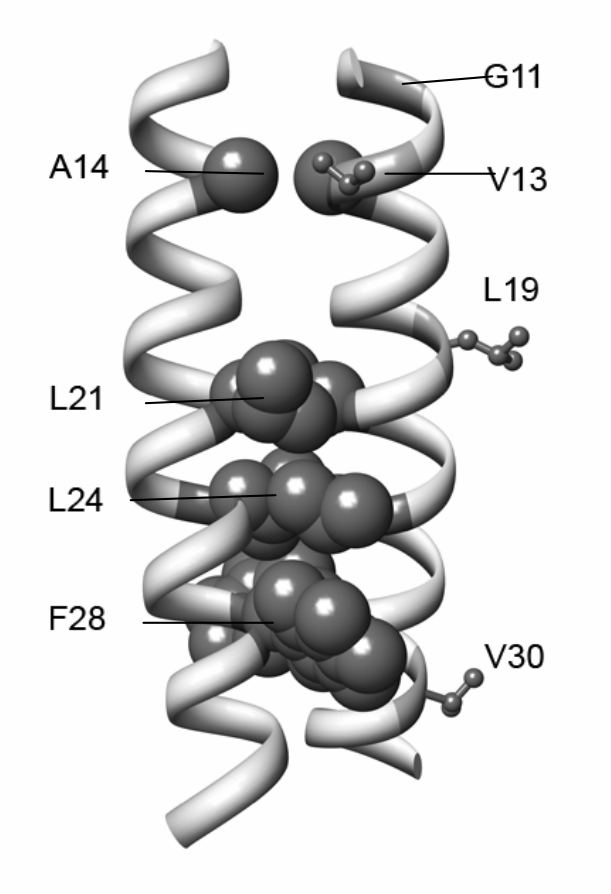


Figure 6.2. Molecular simulation of the E5 TM domain dimer

Predicted molecular model for the TM domain of E5. A homo-dimer was generated using the CHI program as described in Oates et al (Oates, Hicks et al., 2008) that predicted residues that are positioned at the interface and are therefore implicated in stabilising the helix-helix interactions. ^{15}N-labelled residues predicted to be at the interface are represented in space filling whilst those residues not expected to be interfacial are shown in ball and stick.

The amino acid sequences for $E5_{TM1}$ and $E5_{TM2}$ peptides are truncated compared to the full length E5 sequence so as to encompass just the TM domain. Lysine residues were incorporated at the N- and C- termini to aid solubility of this highly hydrophobic peptide and reduce non-specific aggregation, and this approach has been shown not to disrupt the oligomerisation of the E5 TM domain (Oates, Hicks et al., 2008). The peptide $E5_{TM1}$ was synthesised with ^{15}N-labelled residues at positions Ala14, Leu21 and Phe28 which are predicted to reside at the dimer interface, whilst $E5_{TM2}$ had ^{15}N-labelled residues at positions Ala14 and Leu24 that are predicted to be interfacial and at positions Gly11, Val13, Leu19, and Val30 which are expected to be outside the interface (Mattoon, Gupta et al., 2001).

A molecular model for the E5 TM domain generated using the CHI program, as described in Oates et al, shows the relative positions of the interfacial and non-interfacial ^{15}N labelled residues (Figure 6.2**)**(Oates, Hicks et al., 2008). The $E5_{TM1}$ and $E5_{TM2}$ peptides were purified by Dr Joanne Oates using RP-HPLC as described in Oates et al (Oates, Hicks et al., 2008).

6.1.2 NMR analyses of E5 TM peptides in trifluoroethanol

In order to characterise the $E5_{TM1}$ and $E5_{TM2}$ peptides, initial NMR experiments were performed on samples of the peptides solubilised in 80% deuterated trifluoroethanol (dTFE)/ 20% H_2O. The correct labelling of the peptides was confirmed by acquiring a 2D heteronuclear single quantum coherence spectrum (^{15}N-^{1}H HSQC), which enables resonances from ^{15}N-labelled amide groups to be observed. As shown in Figure 6.3a, the expected three resonances were observed in the spectra for the $E5_{TM1}$ peptide. This indicates that the labelling was successful and that the peptide is adopting one conformation in TFE. Presumably, this conformation is the monomeric helix since TFE is known to promote α-helix formation but disrupt helix-helix interactions (Luo and Baldwin, 1997). Similarly, ^{15}N -^{1}H HSQC spectrum of the $E5_{TM2}$ peptide revealed the expected six resonances, as shown in Figure 6.3b.

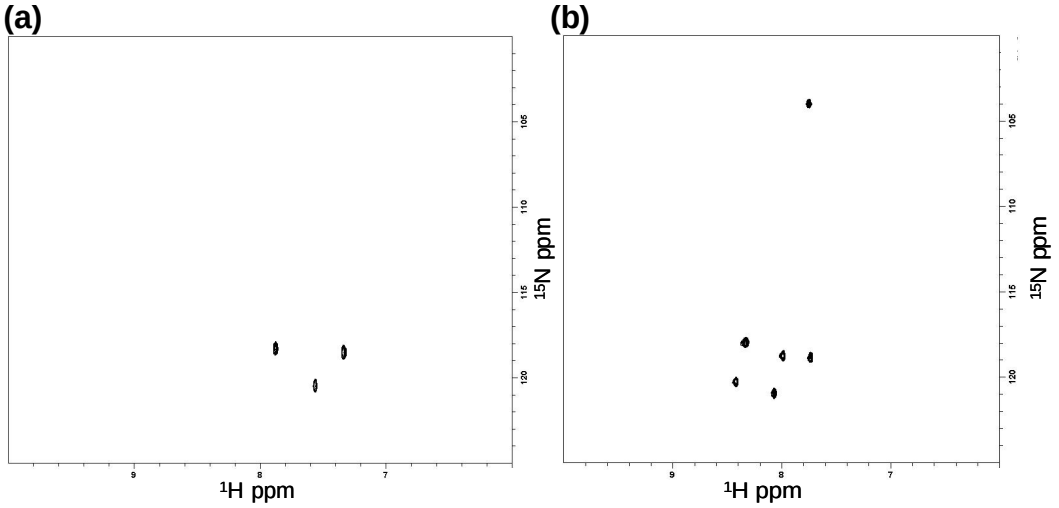


Figure 6.3. ^{15}N-^{1}H HSQC of $E5_{TM1}$ and $E5_{TM2}$ peptides in TFE

^{15}N-^{1}H HSQC spectra acquired on a Bruker 700 MHz instrument for (a) $E5_{TM1}$ and (b) $E5_{TM2}$ peptides solubilised in 80% dTFE/20% H_2O. The number of resonances observed corresponds to the number of ^{15}N-labelled residues in the peptides.

In order to assign the resonances observed for the E5 TM peptides in the ^{15}N-^{1}H HSQC spectra, 3D ^{15}N-edited-^{1}H-^{1}H HSQC TOCSY spectra were acquired which enable only those resonances from ^{15}N-labelled residues to be observed. By reference to published chemical shifts, the amide backbone and side chain resonances for the labelled residues in $E5_{TM1}$ and $E5_{TM2}$ were assigned as shown in Figure 6.4 and Figure 6.5, respectively using the proton labelling scheme shown in Figure 6.6. It was hoped that these would help guide assignments in subsequent NMR experiments on E5 peptides solubilised in detergent.

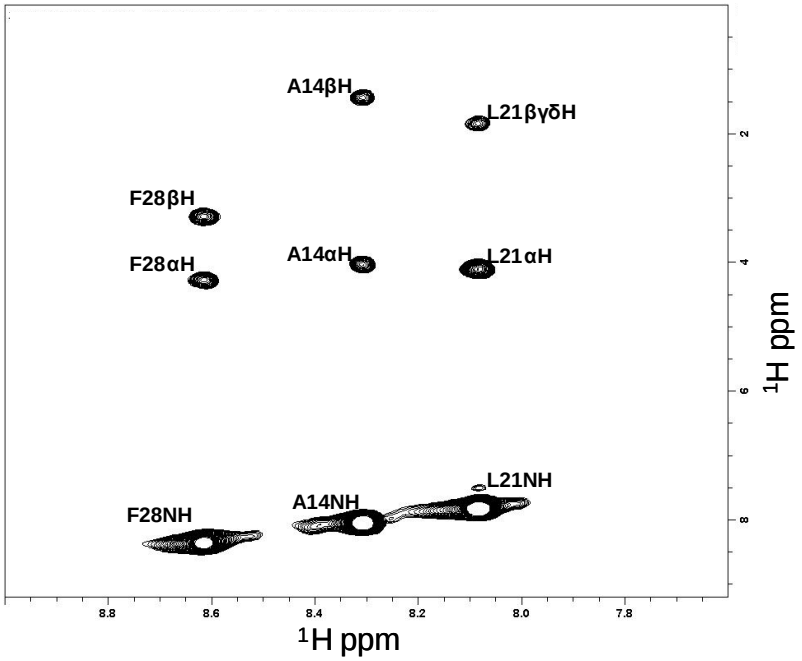


Figure 6.4. ^{15}N-edited ^{1}H-^{1}H TOCSY of $E5_{TM1}$ peptide in TFE

^{15}N-edited ^{1}H-^{1}H TOCSY spectra acquired on a Bruker 700 MHz instrument for $E5_{TM1}$ peptide solubilised in 80% dTFE/20% H_2O. Assignments made with reference to known resonance patterns and tables of chemical shifts.

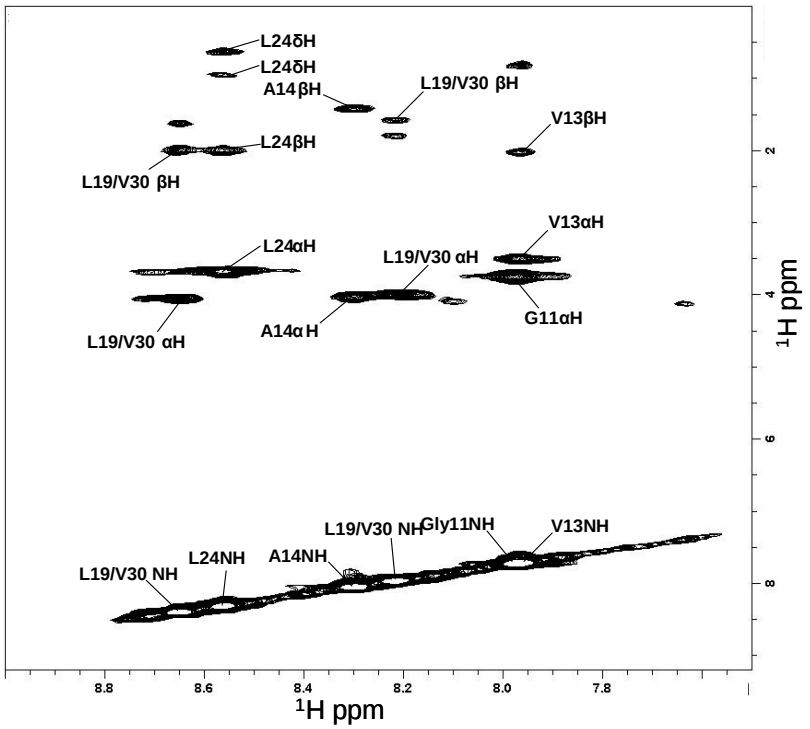


Figure 6.5. ^{15}N-edited ^{1}H-^{1}H TOCSY of E5$_{TM2}$ peptide in TFE

^{15}N-edited ^{1}H-^{1}H TOCSY spectra acquired on a Bruker 700 MHz instrument for E5$_{TM2}$ peptide solubilised in 80% dTFE/20% H_2O. Assignments made with reference to known resonance patterns and tables of chemical shifts.

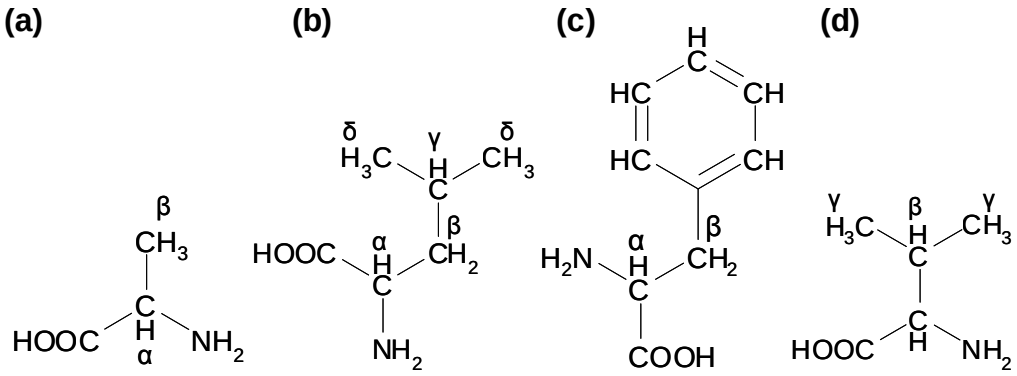


Figure 6.6. Labelling of amino acid side chains used in resonance assignment

Labelling of amino acid protons used in resonance assignments in ^{15}N-edited ^{1}H-^{1}H TOCSY spectra of E5 peptides, for (a) alanine, (b) leucine, (c) phenylalanine, (d) valine.

6.1.3 E5 adopts detergent-dependent conformations

TFE however is not representative of a native membrane environment, and the study then progressed to exploring the assembly of E5 homo-dimers in SDS detergent micelles which provide a more membrane-like environment and have been shown previously to maintain the formation of the E5 TM domain

homodimer (Oates, Hicks et al., 2008). Since it has been shown for other TM domains, including for Ii and MHC in this study, that the oligomeric state can be modulated by the detergent concentration, the affect of SDS concentration on the NMR spectra of E5 was explored and ^{15}N-^{1}H HSQC spectra of $E5_{TM1}$ were acquired at varying SDS concentrations.

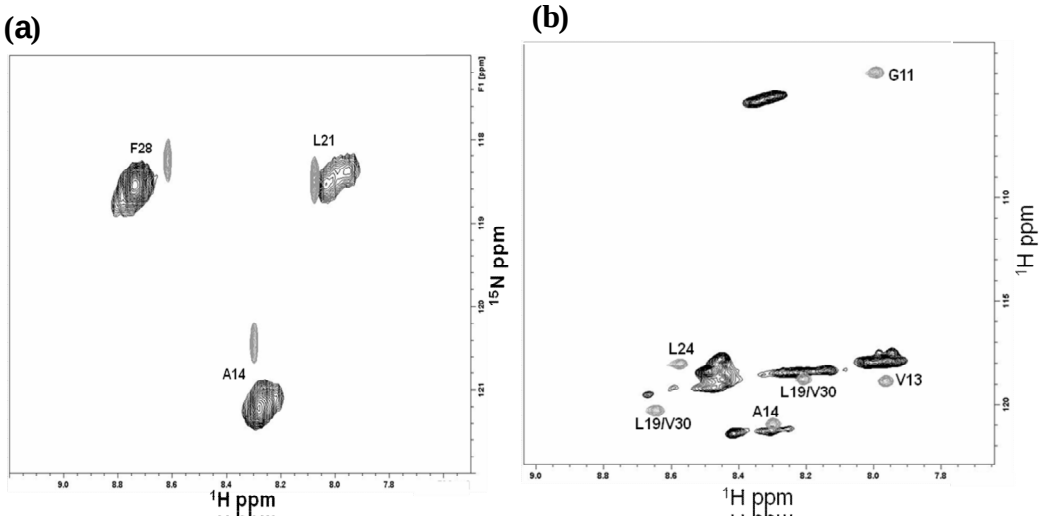


Figure 6.7. Overlay of ^{15}N-^{1}H HSQC spectra for E5 TM peptides in TFE and SDS

Comparison of ^{15}N-^{1}H HSQC spectra of (a) $E5_{TM1}$ and (b) $E5_{TM2}$ peptides acquired in TFE (grey) and SDS (black) showing the same pattern of distribution of the resonances and thus facilitating the assignment of the resonance observed for spectra acquired of E5 peptides in SDS.

Assignment of resonances for E5 peptides acquired in SDS is difficult due to the increased signal broadening introduced by the slower tumbling of the larger peptide-micelle complex. To aid assignment, the spectra acquired in SDS were compared to those acquired in TFE. Comparison of Figure 6.7 to Figure 6.3, the relative position of the resonances does not significantly change upon solubilization of the E5 TM peptides in SDS micelles, thus the assignments made for resonances acquired in TFE can aid those in SDS. Notably, as expected the peak widths are greatly increased for resonances from peptides solubilised in SDS micelles.

For spectra acquired in SDS, as shown in Figure 6.8a, 2D contour plots revealed the presence of two sets of three resonances with each doublet separated by

fractions of a ppm. Comparison of 1D projections through the HSQC spectrum (Figure 6.8b) revealed the relative intensities of these two sets were dependent upon the concentration of SDS with one diminishing as the concentration was increased with concomitant increase in the second set, as shown in Figure 6.8b. This is indicative of the presence of two species that are in slow exchange on the NMR timescale. It has been shown for other TM peptides that the oligomeric state of the protein can be modulated by the concentration of detergent with shifts to lower oligomeric states as the detergent concentration is increased, so it is likely therefore that these two sets of resonances represent dimeric and monomeric species of E5. Therefore, the set of resonances observed at low detergent concentration were assigned to the dimeric state of $E5_{TM1}$. This type of splitting pattern has been observed previously in other NMR studies of TM helix-helix interactions and also attributed to the assembly of monomers into oligomers (Gratkowski, Dai et al., 2002; Wu, Shih et al., 2007).

Interestingly, Leu21 displays a multiplet signal at low micelle concentration in contrast to the broad singlet observed for the other two resonances. It is possible that this is due to the close packing of the helices in the dimer interface restricting the motion of the Leu side chain causing it to adopt multiple rotameric forms. Rotamers of Leu side chains at closely packed dimer interfaces have been observed previously for closely packed dimers of transmembrane domains (MacKenzie, Prestegard et al., 1996). It is conceivable that this could lead to small changes in the backbone conformation and hence small changes in the amide ^{1}H chemical shift. Consistent with this interpretation is the observation that the multiplet collapses to just a single peak as the detergent is increased and E5 becomes more monomeric, which presumably releases the Leu side chain from its restricted motion at the interface.

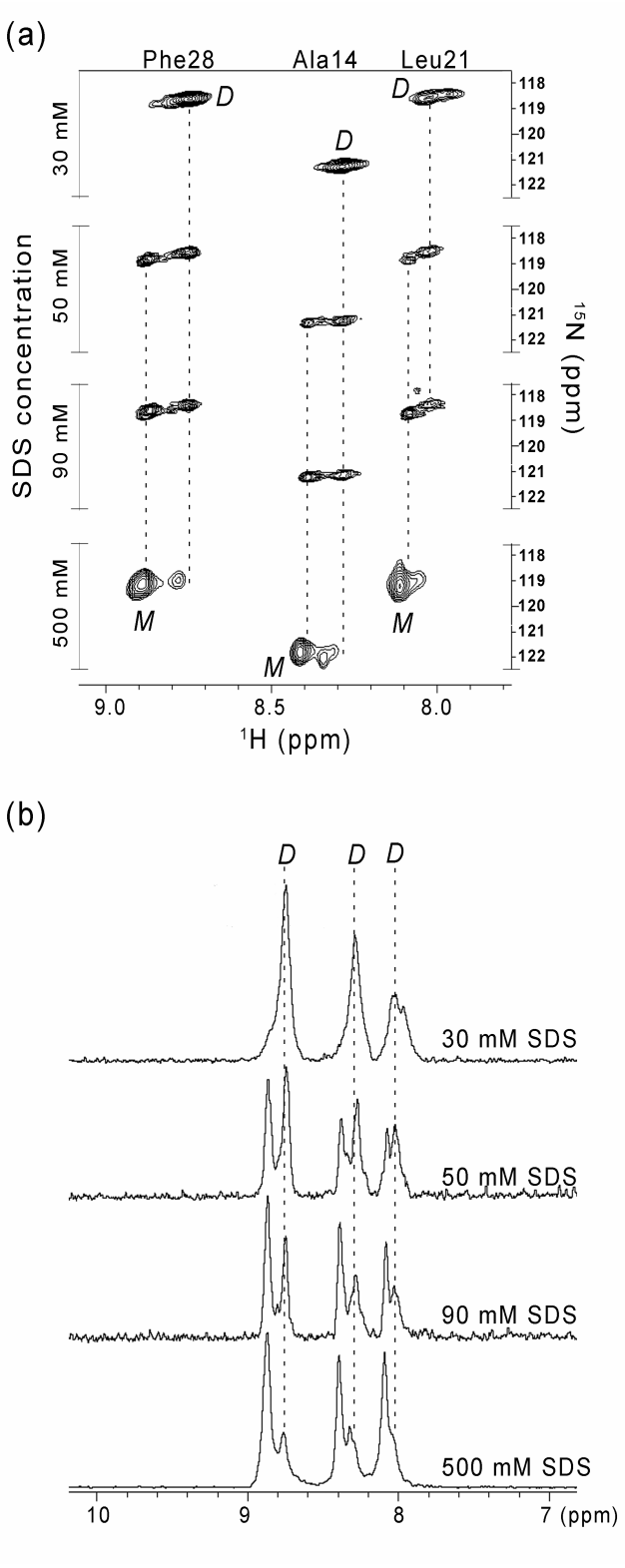


Figure 6.8. ^{15}N-^{1}H HSQC spectra of E5$_{TM1}$ in SDS detergent

(a) Stacked 2D contour plots from ^{15}N-^{1}H HSQC spectra of E5$_{TM1}$ as a function of SDS detergent concentration. (b) 1D projection of plane from 2D spectra showing the resonance doubling and change in relative intensity with detergent concentration. Residue assignments are indicated at the top and the set of peaks attributed to monomer (M) and dimer (D) are indicated.

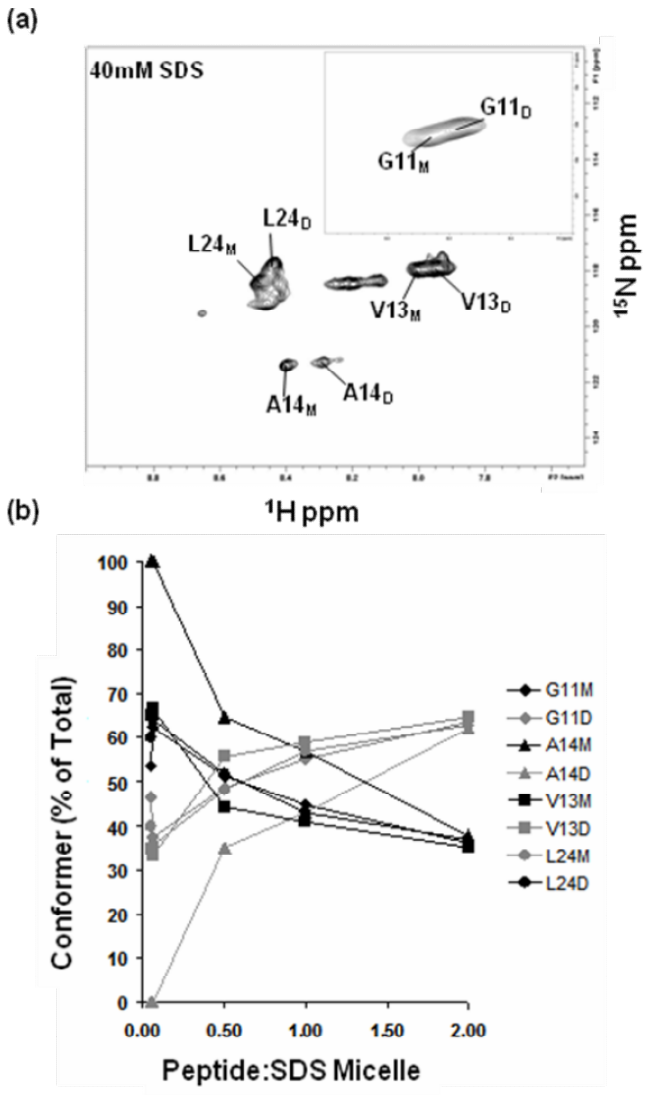


Figure 6.9. ^{15}N-^{1}H HSQC spectra of E5$_{TM2}$ in SDS detergent

Contour plot of two-dimensional ^{15}N-^{1}H HSQC spectra of E5$_{TM2}$ solubilised in 40 mM SDS detergent. (b) 1D projections of plane from 2D spectra showing the resonance doubling and change in relative intensity with the peptide:SDS micelle ratio.

It has been suggested that such a change could be due to non-specific interactions (Wu, Shih et al., 2007). To determine if this was a specific interaction or merely an artefact of low micelle concentration we designed a further peptide, E5$_{TM2}$ which possessed ^{15}N-labelled residues at three positions in the expected interface and at three positions distal to the interface. For residues residing at the interface of the E5 dimer we would expect to observe chemical shift changes upon association of the peptides due to their altered chemical environment. However, as shown in Figure 6.9a, all the resonances observed in the ^{15}N-^{1}H HSQC spectra of E5$_{TM2}$ also exhibit the same resonance splitting. Furthermore, as shown in Figure

6.9b, the relative intensities of the two sets of resonances are modulated by increasing detergent concentration in a similar fashion to $E5_{TM1}$, and were therefore assigned to monomer and dimer forms as indicated.

It has been reported that the average difference in backbone amide chemical shifts in a ^{15}N-^{1}H HSQC can be used to identify interfacial residues in oligomers since theoretically resonances from interfacial residues should undergo a more significant shift than those of other residues in the helix due to the significantly altered chemical environment (Wu, Shih et al., 2007). The average 15N 1H backbone chemical shift difference for the assigned resonances from E5TM1 and E5TM2 were calculated as described in Section 2.20, according to the method in Wu et al (Wu, Shih et al., 2007). As shown in Figure 6.10, the greatest average differences were observed for those residues in the helix expected to be at the interface namely, Ala14, Leu21, Leu24, and Phe28.

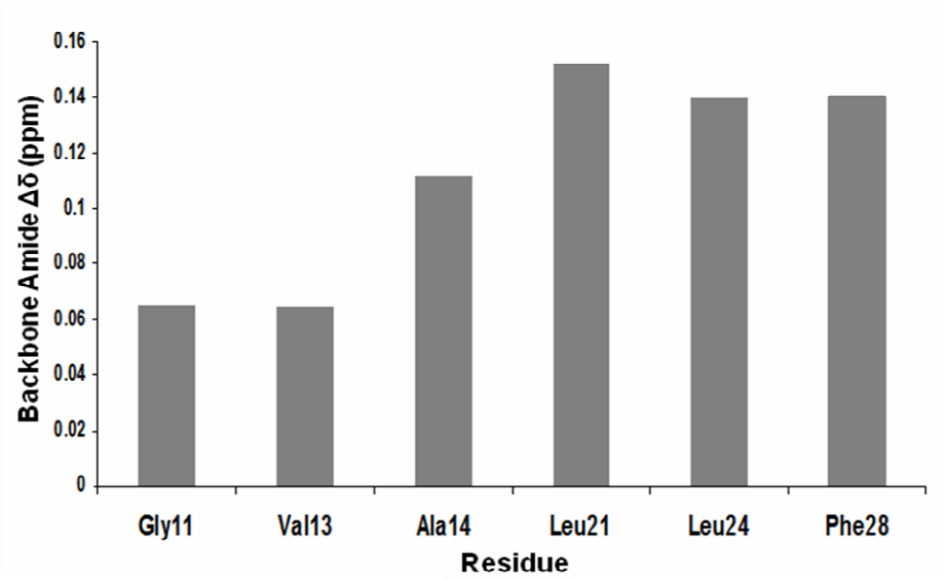


Figure 6.10. Average backbone ^{1}H and ^{15}N amide chemical shift differences

Average backbone ^{1}H and ^{15}N amide chemical shift differences for all assigned resonances from ^{15}N-^{1}H HSQC spectra of $E5_{TM}$ peptides in SDS. Δδ was calculated as described in Section 2.20.

These data show that the oligomeric state of TM domains in detergent micelles can be monitored by changes in chemical shift using ^{15}N-^{1}H HSQC spectra and that the shift in equilibrium between monomeric and oligomeric species is highly dependent on the detergent micelle concentration. Our data suggest that the E5

TM domain forms dimers in SDS micelles and moreover provides atomic level information about the structure of the E5 dimer consistent with previous published results.

6.1.4 Helical content of E5 is unaffected by peptide: micelle molar ratio.

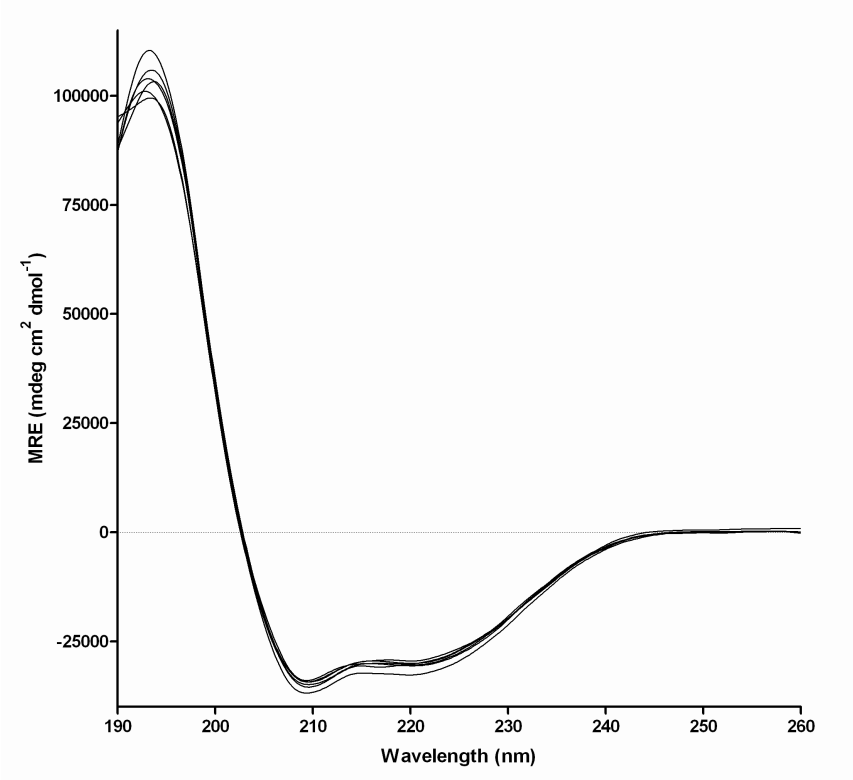


Figure 6.11. Secondary structure of $E5_{TM}$ peptide at varying detergent concentrations

Circular dichroism spectra of 40 μM $E5_{TM1}$ peptide reconstituted into the detergent SDS at peptide:micelle ratios of between 0.1 and 2.0 . Mean residue ellipticity (MRE) was calculated from the measured ellipticity as described in Section 2.21.

Since it is possible that the doubling of resonances observed in the ^{15}N- ^{1}H HSQC was a result of a conformational change such as an interchange between β-sheet and α-helical conformations of the E5 peptides. CD spectra were acquired to determine the effect of varying the concentration of SDS detergent micelles on the secondary structure of the $E5_{TM}$ peptides. As shown in Figure 6.11, all CD spectra show a characteristic α-helical profile, with negative absorption maxima at 208nm and 222nm. This demonstrates that the E5 peptide forms a very stable α-helix in SDS micelles with no significant differences in helicity being observed upon

varying the micelle concentration. This confirms that the two sets of resonances do not result from changes in conformation but instead represent monomeric and oligomeric species.

In summary the data the data presented here shows that E5 is forming a dimer in detergent micelles which is in agreement with all previous works, that the dimer is relatively stable to SDS being present at even high detergent concentrations, and that the interfacial residues may be identified by an increased backbone amide chemical shift relative to those residues not at the interface.

6.1.5 Conclusions on the study of E5 TM domain by NMR spectroscopy

The E5 protein is the smallest known oncogenic protein, little more than a TM domain that can activate cellular receptors in a completely unique way. It is therefore desirable to have an enhanced understanding of its structure and the implications this has for its function. Previous studies have shown E5 is dimeric and that in isolation its TM domain can self assemble to form homo-dimers (Oates, Hicks et al., 2008). *In vivo* and *in vitro* mutagenesis studies identified residues that potentially play a role in mediating the helix-helix interaction (Oates, Hicks et al., 2008). This study attempted to confirm those predictions by obtaining atomic level structural data for the E5 TM domain. The optimisation of sample preparation and the acquisition of NMR spectra were also explored.

^{15}N-^{1}H HSQC spectra of two selectively ^{15}N-labelled peptide analogues of the E5 TM domain revealed twice as many resonances as expected if the peptide were adopting a single conformation, which was expected to be that of a dimer. The relative intensities of these two sets of resonances were dependent on the detergent concentration, with one set of peaks increasing and the other decreasing as the detergent concentration was increased. The two sets were attributed to monomeric and dimeric forms of the E5 peptides with increasing monomer at higher detergent concentrations. The secondary structure of the E5 peptide was predominantly α-helical at all SDS detergent concentrations confirming that the splitting was not due to alterations in the secondary structure of the peptide. These results therefore correlate with other group's observations that detergent

concentration can modulate the oligomeric state but not the helicity of TM domain peptides (Fisher, Engelman et al., 1999). We suggest there are two possible explanations for this behaviour. One possibility is dynamic exchange of dissociated peptides with an ever increasing number of empty micelles reducing the chance of a peptide finding a partner as the micelle concentration is increased. Another scenario is one where, as the detergent concentration is increased, the SDS molecules compete for the interfacial region forcing the dimer apart as the monomer concentration is increased.

It has been reported that the E5 TM dimer has a low dissociation constant and was found to be predominantly dimeric. It is therefore surprising that such a significant proportion of monomer is observed in these NMR studies. However previous studies were carried out in DPC detergent which is a relatively mild detergent in comparison to SDS, so the results shown here are consistent with SDS being a more denaturing detergent. These findings highlight the need to consider the choice of detergent and the detergent concentration (particularly the peptide:micelle ratio) when studying these hydrophobic systems.

In this study we have provided the first atomic level structural information on the association of the E5 transmembrane domain dimer which has recently been found to be a functional subunit of the smallest known oncoprotein (Talbert-Slagle, Marlatt et al., 2009). We have also demonstrated the applicability of solution state NMR methods using detergent micelles as membrane mimetics to the study of the E5 oncoprotein which will provide a good foundation for a full structural characterisation of the protein.

6.2 Novel assay for determining protein-protein interactions

The phenomenon of paramagnetic alignment in NMR studies is a rapidly developing tool that has been applied to the problem of determining protein structure. Here we have exploited this phenomenon to develop a novel method for studying the helix-helix interactions of α-helical TM domains using NMR spectroscopy. It is hoped that this method will facilitate the identification of

homo- and hetero-interactions, as well as provide valuable molecular modelling restraints enabling the rapid structure determination of TM domain oligomers.

The method utilises the ability of paramagnetic lanthanide ions to drive the weak magnetic alignment of proteins with the magnetic field of an NMR spectrometer (Contreras, Ubach et al., 1999; Ikegami, Verdier et al., 2004). Lanthanides possess paramagnetic properties due to the presence of unpaired electrons in the f-orbitals of their trivalent ions. The presence of a paramagnetic ion results in a small anisotropic orientation of the protein with respect to the magnetic field. This weak alignment with the magnetic field facilitates the observation of the through-space interactions between the magnetic fields of bonded atoms, which are termed residual dipolar couplings (RDCs)(Fowler, Tian et al., 2000; Bax, 2003). Such interactions would otherwise be averaged to zero by the rapid tumbling of the molecules in the absence of the weak alignment. Weak alignment is preferential to complete alignment since extensive coupling would result in extensive resonance splitting in the NMR spectrum that would be too complicated to interpret. The magnitude of a RDC between two bonded atoms is dependent on the angle between the vector of the bond (e.g. N-H) and the vertical axis of the magnetic field (B_0). The determination of this angle provides protein backbone restraints for use in molecular modelling, enabling the global fold of a protein to be determined rapidly (Bax, 2003; Bax and Grishaev, 2005).

In this study we have developed a novel means for exploiting RDCs to identify interactions between α-helical TM domains, called HELICS (helix-LBT interactions via RDCs), but which could be extended to exploring any kind of protein-protein interactions. The scheme developed in this study for performing this kind of analysis is shown in Figure 6.12, and involves the differential incorporation of a lanthanide-binding tag (LBT) or isotopic labels into peptide analogues of TM domains. As an example, in the case of two peptides that can interact, one peptide would be isotopically labelled with ^{15}N whilst its partner would possess a LBT encompassing a paramagnetic lanthanide ion e.g. dysprosium (Dy^{3+}). The association of the two peptides would result in the weak alignment of the ^{15}N labelled peptide with the magnetic field of the NMR spectrometer, resulting in the observation of ^{15}N-^{1}H RDCs for this peptide. A key

point is that should no interaction between the two peptides occur then no RDCs would be observed from the isotopically labelled peptide. The measurement of RDCs can be achieved using standard ^{15}N-^{1}H HSQC NMR experiments on the peptides in the presence of a paramagnetic lanthanide or the control which would be the diamagnetic ion lanthanum. Thus, this scheme enables the interaction between two peptides to be determined rapidly by the presence or absence of a RDC.

In addition to molecular alignment in the magnetic field the presence of a paramagnetic metal ion results in distance-dependent paramagnetic shifts (Pintacuda, Park et al., 2006), and enhanced nuclear relaxation of neighbouring resonances (Prudencio, Rohovec et al., 2004), which can be used as molecular modelling restraints for structure refinement. Paramagnetic shifts are of particular interest as they can provide useful long-range structural information and have for some time been employed in structure determinations by NMR spectroscopy (Pintacuda, Park et al., 2006).

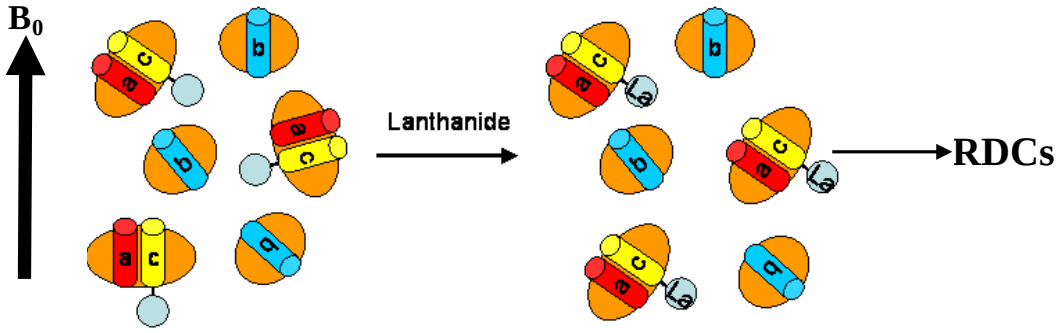


Figure 6.12. Overview of the HELICS assay for hetero-association

NMR-based assay for determining TM domain association. Molecular tumbling of the peptide-micelle complexes averages RDCs to zero. Labelling with paramagnetic lanthanide ions aligns peptides with the magnetic field (B_0). Interaction between isotopically labelled peptide a, and lanthanide labelled peptide c, results in alignment of the hetero-oligomer ac with B_0 and measurable RDCs for the isotopically labelled peptide a. There is no interaction with peptide b so no alignment occurs and hence no measurable RDCs are observed for this peptide.

In addition to molecular alignment in the magnetic field the presence of a paramagnetic metal ion results in distance-dependent paramagnetic shifts (Pintacuda, Park et al., 2006), and enhanced nuclear relaxation of neighbouring

resonances (Prudencio, Rohovec et al., 2004), which can be used as molecular modelling restraints for structure refinement. Paramagnetic shifts are of particular interest as they can provide useful long-range structural information and have for some time been employed in structure determinations by NMR spectroscopy (Pintacuda, Park et al., 2006).

A benefit to using lanthanide metals for generating alignment is that their ions vary in paramagnetic strength whilst retaining similar chemical properties, which enables them to be interchanged (Pintacuda, Keniry et al., 2004). This allows different members of the series to be used with the same lanthanide binding group. This is important for optimising signal loss due to paramagnetic relaxation, eliminating the degeneracy inherent in angle measurements from RDCs by varying the alignment. Moreover, it allows for use of the diamagnetic La^{3+} or Lu^{3+} ions as references for measuring RDCs and PCSs by taking the difference in chemical shift measured in the presence of a paramagnetic or diamagnetic lanthanide (Pintacuda, Keniry et al., 2004).

To instigate alignment of the proteins a paramagnetic lanthanide ion must be incorporated into the peptide of interest. For proteins that possess native metal binding sites this is a simple matter of exchanging the metals, however TM domains do not possess such sites, therefore other methods are required. Two methods have been reported for the incorporation of lanthanide ions into proteins that lack a native metal binding domain. One methods involves the covalent linking of a metal chelating group, usually one based on the metal chelating agent EDTA (Rodriguez-Castaneda, Haberz et al., 2006), to a Cys residue within the peptide, whilst another involves fusing the domain of interest to a protein metal-binding domain that has been optimised for the specific binding of lanthanide ions (Ma and Opella, 2000; Wohnert, Franz et al., 2003).

In this study we chose to produce a fusion protein consisting of the TM domain of interest coupled with a LBT sequence. This LBT sequence been reported to bind terbium (Tb^{3+}) ions with 50 nm binding affinity (Nitz, Franz et al., 2003), and its crystal structure has been determined (Nitz, Sherawat et al., 2004). Furthermore, the LBT sequence has the additional benefit of providing a means of performing FRET-like luminescence resonance energy transfer (LRET) experiments, and is

therefore a very versatile probe for protein studies (Sculimbrene and Imperiali, 2006). The LBT sequence YI**D**T**N**N**D**G**W**Y**E**GD**E**LLA includes 6 lanthanide coordinating residues (shown in bold) including a tryptophan residue that in addition to providing a coordinating carbonyl oxygen acts as a sensitizer for LRET experiments (Sculimbrene and Imperiali, 2006). To test the HELICS assay will require the use of two peptides; a ^{15}N-labelled peptide and a fusion protein consisting of the TM domain of interest and the lanthanide-binding tag (TM-LBT). For the purposes of performing the assay, the peptides will be reconstituted into detergent micelles. Synthetic and heterologous expression systems were explored in order to produce a TM-LBT fusion protein. This section details the efforts to-date in developing this novel method.

6.2.1 Choosing a model peptide to test efficacy of the novel method

It was first necessary to choose a TM domain with which to test the assay. The protein Glycophorin A (GpA) from erythrocytes represents perhaps the best understood example of a dimeric α-helical membrane protein. It has been shown to be a strongly associating homo-dimer that is stabilised by a GxxxG motif for which a solution state NMR structure has been determined (MacKenzie, Prestegard et al., 1997). Therefore, this TM domain was considered a suitable candidate for testing the efficacy of the HELICS assay presented above.

GpA KKITLIIFGVMAGVIGTILLISYGI
GpA-LBT KKITLIIFGVMAGVIGTILLISYGI**YIDTNNDGWYEGDELLA**

Figure 6.13. Sequences for ^{15}N-labelled GpA peptide and GpA-LBT Fusion protein

Sequence of the TM domain peptide analogues to be produced for testing the novel method developed in this study. The TM domain of GpA has been extensively characterised over many years and found to be a strongly associated dimer. ^{15}N-labelled residues were incorporated into GpA during synthesis at the underlined positions. The sequence for the peptide GpA-LBT is a fusion of the TM domain of GpA with the optimised LBT of Franz et al (Franz, Nitz et al., 2003). The sequence of the LBT is presented in bold at the C-terminus of the GpA TM domain sequence. Peptides were synthesised at the Keck Facility (Yale University, USA).

The sequences of peptides to be produced for this study are shown in Figure 6.13. GpA represents the TM domain of the full length GpA protein and was selectively ^{15}N labelled at the positions indicated, whilst GpA-LBT is a designed sequence

which fuses an optimised LBT developed by Franz et al to the C term of GpA TM domain sequence (Franz, Nitz et al., 2003). The placement of the LBT at the C- or N- termini was considered to be arbitrary since it was not known how or if this would impact on the results.

6.2.2 Attempts to synthesise and purify GpA-LBT and ^{15}N-labelled GpA

GpA and GpA-LBT peptides with the amino acid sequences shown in Figure 6.13, were synthesised at the Keck Facility (Yale University, USA). The peptides were supplied as a crude product containing undesirable contaminants such as fmoc protecting groups and truncated peptides, and thus required the desired peptide to be purified. Reverse Phase High Pressure Liquid Chromatography (RP-HPLC) was employed for this task as it is a widely used technique for the purification of hydrophobic peptides. The details of peptide purification were described in Section 2.19.

For the peptide GpA, solubilisation of the crude peptide was optimised and the solution loaded onto a C4 column equilibrated at 30% ACN and eluted against an optimised gradient of ACN and H_2O containing 0.1% TFA with a flow rate of 2 mL /min. Elution was monitored by the absorbance of the amide backbone at 222 nm. Multiple runs of RP-HPLC purification were performed and fractions containing pure peptide were pooled and lyophilised. Following lyophilisation the purity of the peptide was assessed by MALDI mass spectrometry as described in Section 2.11. The peptide was successfully purified (data not shown).

Attempts to purify the synthetic GpA-LBT peptide however were unsuccessful. Analysis of the crude synthesis product by MALDI mass spectrometry and SDS-PAGE revealed that the peptides synthesis had not been successful, since the GpA-LBT peptide was present at very low yield (data not shown). It is believed that this is due to the coupling of the hydrophobic sequence of the TM domain with the hydrophilic sequence of the LBT making synthesis of such a sequence very technically challenging.

6.2.3 Expression of lanthanide-binding tag fusion peptide

Due to the technical challenges of synthesising the GpA-LBT peptide, the use of heterologous expression systems was explored. Multiple systems for the production of hydrophobic peptides are commercially available. Two commonly used systems, the pET and pGEX systems, were identified and trialled in this study for the production of the GpA-LBT peptide. Both of these expression systems involve the fusing of the peptide to a larger domain that facilitates purification but they differ in the nature of that domain and the method used to remove the domain to release the peptide of interest.

Expression of hydrophobic peptides using the pET expression system

Substantial work was performed to optimise the pET system for expression of TM domain peptides. Attempts were made to produce Ii and MHCα peptides but without success. This expression system involves cloning of the TM domain of interest into the commercially available plasmid pET31 (Novagen, UK), to produce the fusion protein KSI-TM-His_6 where KSI is ketosteroid isomerise, a large hydrophobic protein that results in the fusion protein being directed to inclusion bodies during overexpression, TM is the TM domain of interest and His_6 is a 6-residue histidine tag for use in affinity chromatography. Purification of the fusion protein was successfully achieved by solubilisation of the inclusion bodies in 6M guanidine hydrochloride and Ni^{2+} chromatography. The cleavage of the KSI from the TM is reportedly achieved by treatment of the purified fusion protein with CNBr, which cleaves at junctional Met residues present between the TM and the KSI and His_6 moieties. The CNBr cleavage reaction was found to be unsuccessful at the standard reaction conditions reported in the literature. Attempts were made to optimise this reaction involving trialling several solvent conditions, but despite many attempts extensive cleavage of the product was apparent by SDS-PAGE analysis, and the low mass products which were expected to be the desired peptide could not be identified using amino acid sequence analysis. This approach was therefore abandoned in favour of an alternative expression system.

Production of GpA-LBT peptide using the pGEX expression system

An alternative expression system that has been successfully employed in the production of hydrophobic peptides is the pGEX system (Antolini, Lo Bello et al., 2003; Luo, Mo et al., 2007). This system involves the cloning of the TM of interest in the commercially available plasmid pGEX6T (GE Healthcare, UK) to produce the fusion protein GST-TM, as described in Section 2.15. GST, or Glutathione S-transferase, is a ~26 kDa protein that facilitates purification of the fusion protein using glutathione affinity chromatography. GpA-LBT was cloned into the pGEX vector to generate the vector pGEX-GpA-LBT with which BL21 cells were then transformed.

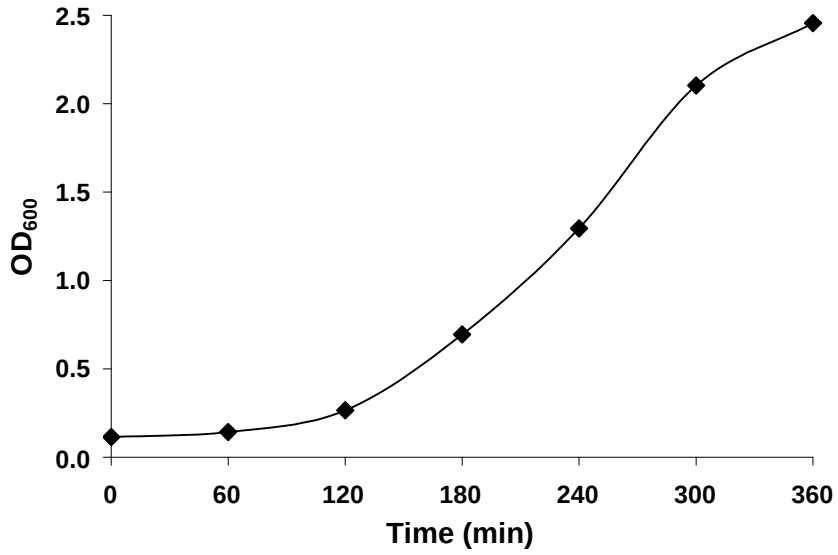


Figure 6.14. Growth curve for BL21 cells expressing GST-GpA-LBT fusion protein

Monitoring the growth of BL21 cells transformed with the plasmid pGEX-GpA-LBT. 10 mL of fresh LB media was inoculated with 1/40 dilution from overnight culture. When the OD_{600} reached 0.3, expression of the fusion protein was induced by addition of 1 mM IPTG. The OD_{600} was subsequently measured at hourly intervals until it began to plateau indicating the optimum point to harvest cells.

In order to test the effect of fusion protein expression on the ability of the BL21 cells to propagate, the cell growth following induction of fusion protein expression by the addition of 1 mM isopropyl β-thiogalactoside (IPTG) was monitored over 6 hours by measuring the optical density at 600 nm (OD_{600}). As shown in Figure 6.14, after 6 hours the OD_{600} reached over 2.0 showing that a

high cell density could be achieved before harvesting of the cells was performed. This indicates that the expression of the fusion protein is not detrimental to cell growth, and that a potentially high yield of fusion protein could be obtained, depending on the expression level of the fusion protein per cell.

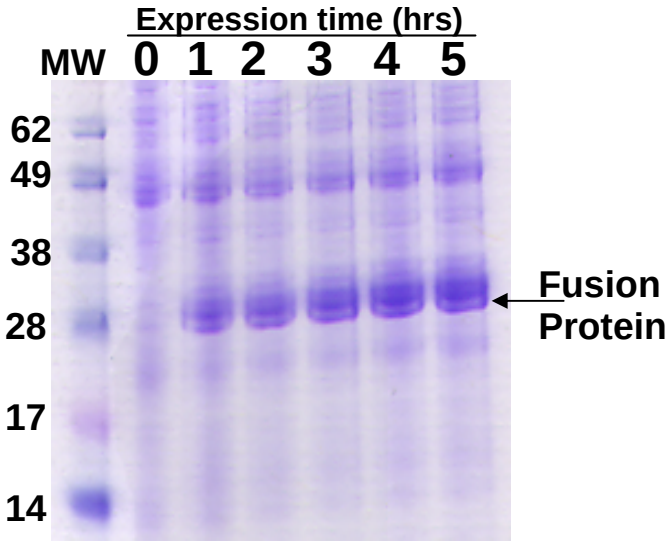


Figure 6.15. SDS-PAGE of induced expression of GST-GpA-LBT fusion protein

SDS-PAGE analysis of the induced expression of GST-GpA-LBT. BL21 cells were transformed with the plasmid pGEX-GpA-LBT. When an OD_{600} of 0.3 was reached the expression of the fusion protein was induced by the addition of 1 mM IPTG. At hourly intervals cell aliquots normalised to 0.6 were removed for analysis. The sample at time zero represents the cell expression immediately before induction. The band corresponding to the fusion protein is indicated. Molecular weight markers (MW) are shown in the left most lane, in kDa. Bands were visualised using Coomassie.

In order to confirm the expression of the fusion protein was being induced by the addition of 1 mM of IPTG and to measure the yield, the induction of expression was monitored by SDS-PAGE. Cells were grown to an OD_{600} of 0.3 and then expression of the fusion protein was induced by the addition of 1 mM IPTG. Aliquots of cells were taken immediately prior to induction and subsequently at hourly intervals over 5 hrs. Cell aliquots were normalised to an OD_{600} of 0.6 in order to monitor the level of expression per cell. As shown in Figure 6.15, at zero hours there is no expression of the fusion protein which indicates the expression of this protein is under tight control prior to induction. Following induction, an intense band at the expected mass for the fusion protein (~30 kDa) is observed which reaches a maximum after ~3 hours, and thus indicates that induction of

fusion protein was successful. This result indicates that a high yield can be expected from the expression of the fusion protein with the BL21 cell line.

In order to progress with purifying the fusion protein it was necessary to identify the cellular fraction (i.e. soluble or insoluble) containing the fusion protein. It is expected that the coupling of the hydrophobic GpA sequence to the larger soluble domain GST would result in the fusion protein being localised to the soluble fraction of the cell lysate. In order to identify whether the fusion protein was in the soluble or insoluble fraction after cell lysis, SDS-PAGE analysis of the cell lysate was performed, as described in Section 2.15.2. Lysis of the cells was performed in phosphate buffer and the lysate was then centrifuged to separate soluble and insoluble material. As shown in Figure 6.16, the fusion protein is present in whole cell samples, and very little is present in the soluble fraction. Instead, the fusion protein is found predominantly in the insoluble fraction. This is most likely in the form of inclusion bodies, which are commonly formed by the overexpression of hydrophobic proteins. This result indicates that to proceed with purifying the fusion protein will require the use of solubilising agents to solubilise this fraction.

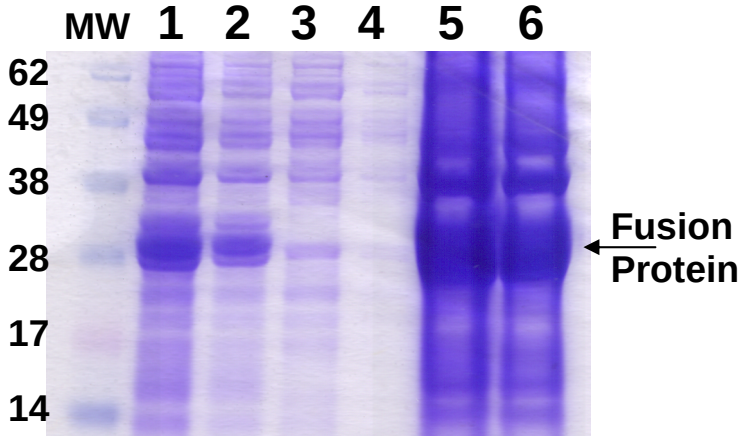


Figure 6.16. Isolating the fusion protein in cellular fractions

SDS-PAGE analysis of GST-GpA-LBT fusion protein overexpressed in BL21 cells. Cells were lysed in phosphate buffer followed by centrifugation of the lysate. The supernatant representing the soluble proteins and the pellet representing the insoluble proteins were analysed. Lanes 1 and 2 show the whole cell fraction with 2 being a 1 in 2 dilution. Lanes 3 and 4 show the soluble fraction with 4 being a 1 in 2 dilution. Lanes 5 and 6 show the insoluble fraction with 6 being a 1 in2 dilution. The GST-GpA-LBT fusion protein as indicated. Molecular weight markers (MW) are shown in the first lane. Bands were visualised using Coomassie.

The solubilising agent must possess certain properties to be suitable for use in this purification protocol. It must not only dissolve the insoluble fraction releasing the fusion protein, but must also not disrupt the native fold of the GST domain. The latter is necessary since it is a requirement for the binding of GST to the affinity chromatography column, which will be the next step in the purification. A review of the literature revealed that a 1% solution of the detergent sarkosyl can solubilise the fusion protein from inclusion bodies whilst maintaining the native GST fold (Frangioni and Neel, 1993). The protocol for production of the cell lysate was therefore modified by the incorporation of a detergent solubilisation step, where 1% sarkosyl is added to the cell lysate.

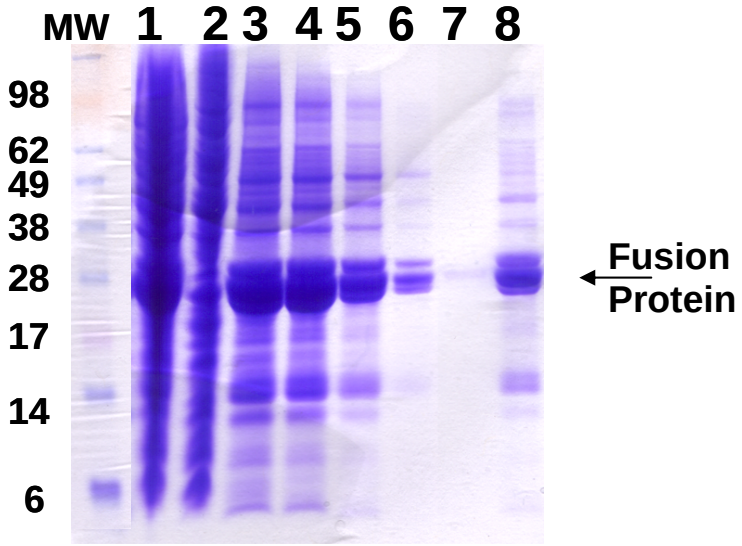


Figure 6.17. Purification of GST-GpA-LBT fusion protein

SDS-PAGE analysis of small scale purification of GST-GpA-LBT fusion protein to confirm binding to the affinity column. Cell lysate was solubilised in 1% sarkosyl and added to GST-binding sepharose matrix. Aliquots were taken at each stage of the purification. Lane 1 represents the whole cell fraction, Lane 2 the soluble fraction, Lane 3 the insoluble fraction, Lane 4 the supernatant after affinity chromatography, Lane 5 Wash 1, Lane 6 wash 2, Lane 7 wash 3, lane 8 elution of the Fusion protein. Molecular weight markers (MW) are shown in the first lane. Bands were visualised using Coomassie.

In order to confirm that the use of 1% sarkosyl detergent did not denature the GST domain, a small scale batch purification of the fusion protein was performed using GST-binding affinity chromatography, as described in Section 2.15.3. As shown in Figure 6.17, the fusion protein is observed in the whole cell (lane 1) and insoluble fractions (lane 3) prior to batch binding to GST sepharose. Following

binding not all the protein is bound due to the small scale nature of the purification (lane 4), and washing the matrix removed any unbound protein from the matrix (lane 5-7). Following the treatment of the matrix with glutathione elution buffer the fusion protein is observed to elute (lane 8). These data indicates solubilisation of the cell lysate with 1% sarkosyl does not impede binding of the fusion protein to the affinity matrix and furthermore suggests that the fusion protein may be purified in high yield.

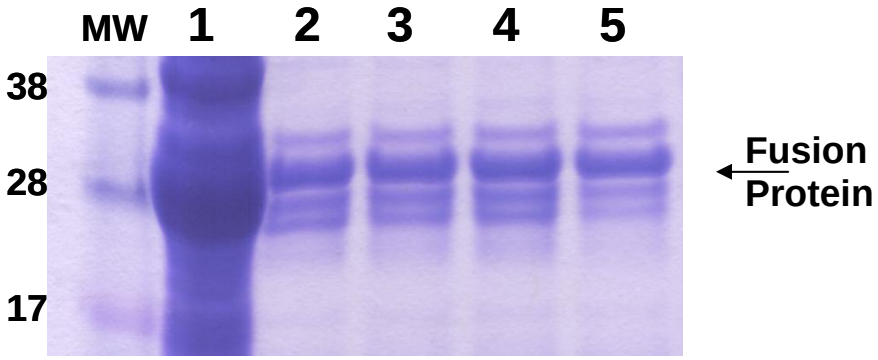


Figure 6.18. Purification of GST-GpA-LBT fusion protein

SDS-PAGE analysis of eluted GST-GpA-LBT fusion proteins measuring the effect of Triton X-100 on column binding. Aliquots were taken at each stage of the purification. Lane 1 represents the whole cell fraction. Lane 2 1% sarkosyl, Lane 3 plus 0.5% Triton, Lane 4 plus 1.0% Triton X100, Lane 5 1.5% Triton X100. Molecular weight markers (MW) are shown in the first lane. Bands were visualised using Coomassie.

Since the GST domain must remain in its native fold during the purification to bind the affinity column we attempted to optimise binding of the fusion protein to the matrix. It has been suggested that the addition of the detergent Triton-X100 can stabilise the fold of GST (Frangioni and Neel, 1993). To test the efficacy of this approach, the purification of fusion protein was performed at varying concentrations of Triton X-100 and aliquots of the eluted fusion protein analysed by SDS-PAGE. As shown in Figure 6.18, the intensity of the bands for samples that did (lanes 3-5) and did not (lane 2) contain Triton X-100 are identical. This suggests that the addition of triton does not enhance the binding of the fusion protein to the matrix.

With it established that the fusion protein can be purified from the cell lysate, the next stage in the purification is optimising the cleavage of the fusion protein to

release the GST domain which is followed by the purification of the GpA-LBT peptide. In the fusion protein a protease cleavage site is present between the GST domain and the GpA-LBT peptide, enabling the use of proteolysis to cleave the two domains. This can be performed using on-column cleavage by the addition of the protease Prescission **(**GE Healthcare, UK). In this method, the GST protein remains bound to the matrix and the desired peptide is released. Cleavage of the fusion protein was monitored by SDS-PAGE. However when this was performed the cleavage of the fusion protein could not be observed, since no band at the expected mass of the GpA-LBT peptide nor a decrease in the weight of the fusion protein band were observed.

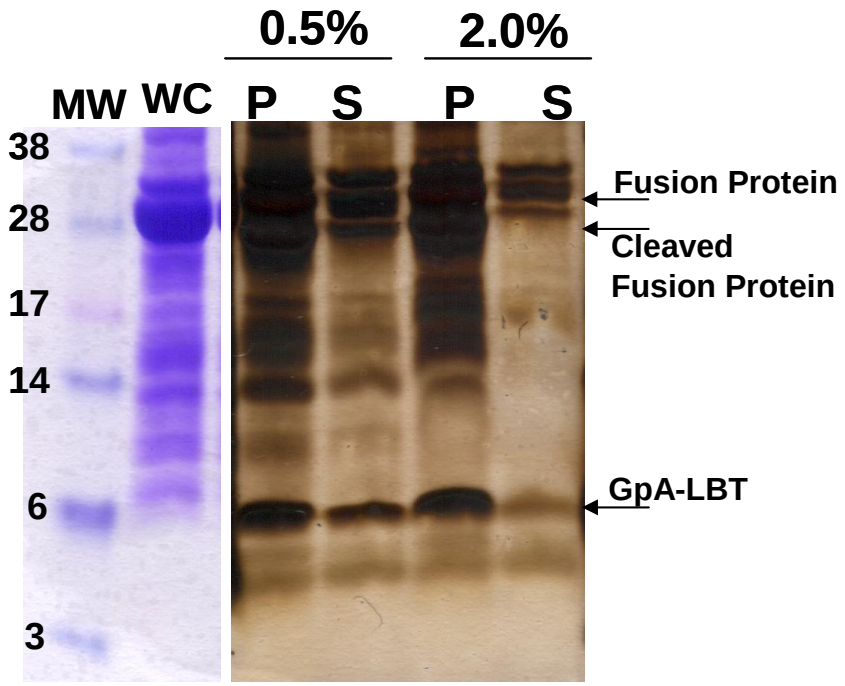


Figure 6.19. Cleavage of GST-GpA-LBT fusion protein in OG

SDS-PAGE analysis of cleaved GST-GpA-LBT fusion proteins. Molecular weight markers (MW) are shown in the first lane. Cleavage of fusion protein in OG at 0.5% and 2.0% OG. Following cleavage the sample was centrifuged. P is pellet, S is supernatant after centrifugation. Bands in left gel were visualised using Coomassie whilst those in the right gel were visualised using silver staining.

Attempts were subsequently made to optimise the number of units of protease used in the cleavage reaction but this had no effect. It was then hypothesised that the protease may be inactivated by the detergent. Attempts were then made to minimise the detergent concentration used in the purification, but without success.

A literature survey revealed the activity of the protease could be maintained in Triton X100. The use of this detergent in the cleavage reaction resulted in the fusion protein being cleaved. However, it was subsequently found that this detergent cannot be removed by dialysis due to its very low critical micellar concentration (CMC). It is necessary to able to remove or exchange the detergent since the large size of the Triton X-100 micelle will not be amenable to subsequent NMR analysis.

The detergent octylglucoside (OG) has a high CMC of ~19 mM making it a more suitable detergent for use in purification procedures requiring dialysis. Trials of fusion protein cleavage were performed in varying amounts of this detergent to determine if activity of the protease would be retained in the presence of OG. This was performed in batch mode enabling the sepharose matrix to be pelleted by centrifugation. The supernatant and the pellet were subsequently analysed by SDS-PAGE. It was expected that the GpA-LBT would be found in the supernatant. As shown in Figure 6.19, cleavage of the fusion protein was observed indicating that the protease does indeed retain its proteolytic activity in OG. However, the peptide was found predominantly in the pelleted fraction suggesting that either the peptide is remaining bound to the sepharose matrix or more likely that the peptide is not soluble at the concentrations of OG that are needed to maintain activity of the protease. Attempts were made to increase the OG concentration following the cleavage reaction in order to solubilise the peptide however these proved unsuccessful. Due to time constraints the project has not progressed past this stage. Future work on the purification of the GpA-LBT peptide would involve attempting to solubilise the products from cleavage in organic solvent such as TFE, filtering the sample to remove the sepharose matrix, and then performing RP-HPLC.

7 Conclusions

Determining the rules for helix-helix interactions that govern the assembly of α-helical membrane proteins represents an important and challenging area of research. Studies on membrane proteins such as Glycophorin A (GpA) have shown that the transmembrane domains of these proteins serve a greater function than merely anchoring them within the membrane and are actually centres of significant protein-protein interactions that play a role in the assembly of membrane proteins.

In this study we have explored the interactions between the TM domains of two α-helical membrane proteins from the immune system, using a range of *in vivo* and *in vitro* techniques. Major Histocompatibility Complex Class II (MHC) α- and β-subunits and Invariant Chain (Ii) associate in the ER to form a nonameric (nine chains) complex as part of the initial steps in the process of antigen presentation. This is necessary for optimal export of the MHC proteins from the ER. Although the association between the soluble domains of MHC and Ii has been well-characterised over many years and their respective structures are known, recent work has implicated interactions between their TM domains in the formation of this complex. The TM domain of Ii was shown to self assemble into a trimer whilst it was shown MHC α/β heterodimers could associate with Ii in a manner independent of the soluble domains. By conducting the work presented in this thesis we sought to gain further understanding of the role played by the α-helical TM domains of Ii and the α- and β-chains of MHC in the assembly of the Ii-MHC complex. This chapter summarises how the results of this study have furthered our knowledge of the assembly of Ii, MHC, and the Ii-MHC complex.

7.1 Studies on the helix-helix interactions in the Ii-MHC complex

As presented in Chapter 3, we sought to confirm the findings reported by Dixon et al (Dixon, Stanley et al., 2006), that in isolation, the TM domain of Ii self-assembled to form a trimer in the mild detergent DPC. In the process of

performing this work, the TM domains of Ii also served as the test subject for developing protocols for performing the GALLEX assay and FRET analyses which would be used in studying the association of Ii with MHC. Results from the *in vivo* GALLEX assay and *in vitro* cross-linking and FRET analyses of a model peptide homologous to that used in the study by Dixon et al were unable to confirm the TM domain of Ii forms a specific trimer. However, but they did confirm that Ii TM domain can self-assemble into a range of oligomeric states from dimer to pentamer, with dimer being the most prevalent. Furthermore, results from both cross-linking and FRET showed the oligomeric state of the Ii TM domain was highly dependent on the detergent concentration or more specifically the peptide:micelle ratio. This was shown in a range of detergents. So though the self-association of Ii TM domain is now well established by this study and others (Ashman and Miller, 1999; Kukol, Torres et al., 2002; Dixon, Stanley et al., 2006), it is uncertain that a single trimeric state for the Ii TM domain can be assigned. However, since the TM domain can self-associate and one of the observed oligomeric states was trimer and the full length protein is a trimer, it is possible that the native membrane environment may be contributing significantly to specifying the oligomeric state of the Ii TM domain. This could be tested by further studies on the model peptides in lipid bilayers that more closely resemble the native environment of these domains.

The difficulty in assigning a definitive oligomeric state to Ii TM domain may reflect the challenge of studying weakly associating TM domains in which the association is stabilised by H-bonds. To-date relatively few studies have been made of TM domains that are not stabilised by very strong interactions, such as in the case of the GpA TM dimer which is stabilised by the GxxxG motif to such an extent that its oligomeric state is visible by SDS-PAGE. As noted in Section 1.2.1, H-bonds bonds are thought to have weak specificity relative to Van der Waals interactions and are believed to cause non-specific aggregation (White, 2006). It has been proposed that van der Waals interactions in close packing helices are the main determinants for TM helix association and that H-bonds serve to stabilize a preformed oligomer (Schneider, 2004). Despite the difficulties in studying the TM domain of Ii, this section of the study facilitated the development of protocols for performing the GALLEX and FRET measurements which then enabled the study

to progress to studying hetero interactions between the TM domains of α- and β-chains of MHC and also the interactions between those and Ii.

In this study we have shown using the *in vivo* assays TOXCAT and GALLEX that the TM domains of the MHC α and β chains are capable of self-associating, which is the first time this behaviour has been observed. Furthermore, highly conserved small-xxx-small motifs that are known to infer a propensity for self-association were identified in the TM domains of the α- and β-chains of MHC. Modelling of the homo-dimers of MHCα and MHCβ using molecular dynamics suggested that the motifs could pack at the interface and stabilise homo-dimer formation. Subsequent mutagenesis studies using TOXCAT implicated one of the two motifs in the assembly of MHCα homo-oligomers, which suggests they are stabilised by to the formation of interhelical induced dipoles resulting from the close packing of the helices. Mutagenesis of the motif in MHCβ did not affect the homo-oligomerisation as measured with the TOXCAT assay, however the same mutation in an optimised sequence measured with the GALLEX assay did succeed in disrupting homo-oligomerisation. The discrepancy between the two assays is not fully understood at present, however the latter result is more in keeping with the results from molecular modelling and with what is known about GxxxG motifs. The sequence in the GALLEX assay was different from that used in the TOXCAT assay so future experiments could be designed to test this sequence in TOXCAT or perhaps a range of sequence lengths, and mutagenesis studies with model peptides could be performed in conjunction with cross-linking or FRET. The self-association of MHCα and MHCβ was also observed in studies on model peptide analogues using SDS-PAGE, with both forming SDS-stable dimers, which is in agreement with the known strength of GxxxG interactions in some proteins. Cross-linking studies showed MHCα could self-assemble into a range of oligomeric states whilst MHCβ formed predominantly dimers in mild detergent. The self-association of the MHC TM domains has not been observed before nor has this behaviour been observed in the full length MHC proteins, hence the biological significance of this finding is unknown. We could speculate that this interaction may drive the formation of inactive MHC homo-dimers, which has been observed for other membrane proteins such as receptor tyrosine kinases (Yu, Sharma et al., 2002; Seubert, Royer et al., 2003). Alternatively, self

association may facilitate the localisation of the MHC α- and β-chains for rapid assembly with their partner and Ii.

Since it is known that the full-length α- and β-chains of MHC associate via their soluble domains to form hetero-dimers we investigated the propensity for the MHCα and MHCβ TM domains to do the same. This was performed using the *in vivo* GALLEX assay and *in vitro* FRET analyses, which are two of the few methods currently available for performing analysis of hetero-oligomerisation. The results from the GALLEX assay suggested that the TM domains of MHCα and MHCβ could associate with one another to form hetero-oligomers, which is in keeping with what is known about the full length proteins. A subsequent mutagenesis study implicated the small-xxx-small motifs in this association. This result was corroborated by the findings from molecular modelling, which implicated these residues as being at the heterodimer interface. Subsequent analyses of model peptides with FRET showed that the peptides were also associating with one another, lending further support to our findings from GALLEX. FRET enabled the oligomeric state of the interaction to be monitored and the results suggest that MHCα and MHCβ are associating to form predominantly hetero-dimers, which is also in agreement with our understanding of the full length proteins.

It has been reported that MHC can bind to Ii in a manner that is independent of the soluble domains, and thus implicated the TM domain in mediating these interactions. We therefore explored the interactions between the TM domains of the α- and β-chains of MHC and Ii. Use of the *in vivo* GALLEX assay suggested that Ii could associate with MHCα and MHCβ in the inner membrane of *E. coli*, thus corroborating the reported findings. However, FRET analyses of peptide analogues suggested that Ii only associates strongly with MHCα and not MHCβ, since no energy transfer was observed between Ii and the MHCβ TM peptide. This discrepancy could be due to the environments the different measurement are carried out in, with GALLEX measurements being conducted for TM domains inserted into a biological membrane and FRET measurements being conducted for TM peptides solubilised in mild detergent. The observation in this work that Ii preferentially binds MHCα is interesting since it has been proposed previously

that, rather than MHCα and MHCβ associating to form a heterodimer prior to associating with Ii, MHCα and Ii associate first, followed by binding of MHCβ to the Ii-MHCα complex (Koch, McLellan et al., 2007).

7.1.1 Implications for the assembly of TM domains in the Ii-MHC complex

In summary, the results from this study suggest the following: (a) the TM domain of Ii self-associates into a range of specific oligomeric states from dimer to tetramer; (b) the TM domains of the α- and β-chains of MHC can self-associate and form α/β hetero-dimers; (c) the TM domains of Ii can associate with those of the α- and β-chains of MHC, but *in vitro* measurements indicate there is a preferential interaction between the TM domains of Ii and α. Collectively, these data show the TM domains of Ii and MHC are sites of potentially important protein-protein interactions that may play a role in the assembly of the MHC-Ii complex and ultimately in antigen presentation. We therefore consider how the findings from this study impact upon our current understanding of the assembly of the Ii-MHC complex.

Whilst this study was unable to show the TM domain of Ii self-associates into a specific trimer, as was observed previously, we did show that it can self-associate and that one of the oligomeric states it adopts is a trimer. Given that the full length protein is known to be trimeric it could be reasoned that the TM domain in its native environment is adopting this stoichiometry also, and that in detergents the forces (e.g. H-bonds) driving the self-assembly are too weak to specify that state. For the purposes of the following discussion we therefore treat the TM domain of Ii as if it were trimeric.

The results from this study suggest that the TM domains of MHC α and β-chains can associate to form a hetero-dimer, and that *in vivo* the TM domain of Ii can bind to those of the MHC α- and β-chains. Consistent with this finding, full length MHC α- and β-chains have been observed to associate in the absence of Ii (Bijlmakers, Benaroch et al., 1994; Elliott, Drake et al., 1994), leading to the suggestion that α- and β-chains chains of the same isotype first assemble to form a heterodimer prior to associating with Ii (Lamb and Cresswell, 1992; Bijlmakers,

Benaroch et al., 1994). However, in the presence of Ii, MHC α-β dimer intermediates have not been observed (Cresswell, 1994). This study also showed that *in vitro* Ii preferentially binds to MHC α-chains. This behaviour has also been observed in a number of studies on the full length proteins (Lamb and Cresswell, 1992; Bijlmakers, Benaroch et al., 1994), leading to an alternative model for the assembly of MHC α- and β-chains being proposed where MHC α-chains first bind to Ii, and then the MHCα-Ii complex selects for a isotypically matched MHC β-subunit, with Ii sandwiched in the middle (Neumann and Koch, 2005; Neumann and Koch, 2006). It is believed that this model has the advantage that a mismatched MHC β-subunit can dissociate from the MHCα-Ii matrix and be replaced by a matched MHC β-subunit in a more efficient manner. Our findings on the TM domain interactions in the Ii-MHC complex seem to support both models for the assembly of the full length proteins.

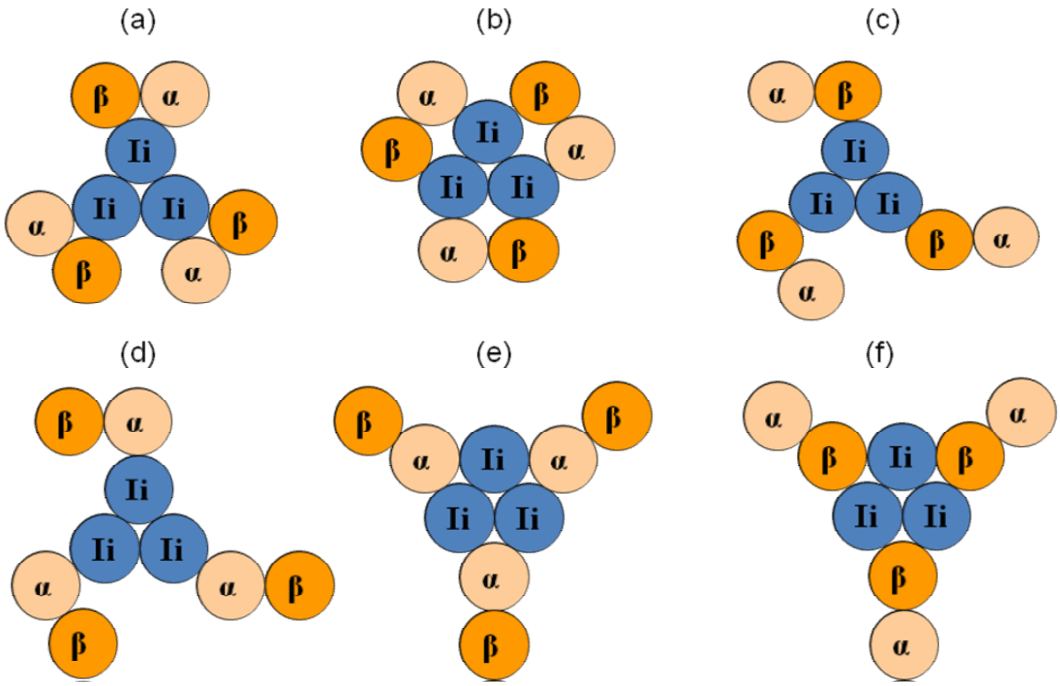


Figure 7.1. Models for association of TM domains in the Ii-MHC complex

Top down view of models for the association of the TM domains of Ii, and those of the α- and β-chains of MHC in the Ii-MHC complex.

In previous models of the assembly of the Ii-MHC complex, the role of the TM domains has not been considered; rather the emphasis is placed on the role of the soluble domains. However, the findings of this study and others (Castellino, Han

et al., 2001; Kukol, Torres et al., 2002; Dixon, Stanley et al., 2006) strongly implicate the TM domains in the formation of the Ii-MHC complex and therefore any future model for the assembly of the Ii-MHC complex should address the role played by the TM domains.

Using the results from this study and those reported in the literature we can begin to build a model for the association of the TM domains in the Ii-MHC complex. All of the possible arrangements that the TM domains could adopt are shown in Figure 7.1. MHCα and MHCβ were observed to associate to form hetero-dimers, which would support all the models in Figure 7.1. Ii was found to associate preferentially with MHCα, which is consistent with the models shown in Figure 7.1a, b, d, and e. There are difficulties in the data regarding the association of Ii with MHCβ, since an interaction was observed *in vivo* but this was more difficult to conclusively determine *in vitro*, as such it is not possible at the present time to discount any of the models shown in Figure 7.1a, b, d, or e until this has been resolved. Interestingly, the results from FRET experiments suggested that the interactions between Ii and MHCβ could not be enhanced by the presence of MHCα which we would not expect if the models shown in Figure 7.1a and b were correct. It may be possible in future work to discount the model shown in Figure 7.1a by exploring the interactions of MHC α and β TM domains with the TM domain mutants of Ii that were shown to disrupt its oligomer formation (Dixon, Stanley et al., 2006).

7.2 Development of NMR methods to study helix-helix interactions

Solution state NMR spectroscopy has the potential to conclusively determine TM domain interactions by investigation of model peptides solubilised in detergent micelles. In this study we therefore sought to develop NMR methodologies to facilitate the determination of the atomic structure of TM domain oligomers. Furthermore we sought to develop an NMR-based method that exploits paramagnetic effects to identify helix-helix interactions in TM domains and which also has the potential to provide valuable structural restraints for structure determination.

Recently, the TM domain of the protein E5 was shown to self-associate via non-covalent interactions to form a strongly interacting dimer in the detergent SDS using a range of techniques including TOXCAT, SDS-PAGE and AUC (Oates, Hicks et al., 2008). NMR spectra of selectively-labelled peptides corresponding to the TM domain of E5 showed that when solubilised in TFE, the expected numbers of cross-peaks are observed in the ^{15}N-^{1}H HSQC spectrum, which display resonances for every ^{15}N labelled residue in the peptide. When the peptides were solubilised in the detergent SDS multiple resonances were observed that indicated the presence of several conformational states. Subsequent HSQC spectra acquired at varying detergent concentrations revealed the resonances could be assigned to two groups whose relative heights could be modulated by the detergent concentration. CD spectra indicated that the change in detergent concentration was not modulating the secondary structure since the CD spectra were identical, indicating that the two populations were not folded and unfolded peptide. Furthermore, since work carried out in this thesis had shown that the oligomeric state of TM domains can be modulated by the peptide:micelle ratio, the two sets of resonances were assigned to monomer and dimer states of the E5 TM domain, with the dimer state dominating at lower detergent concentrations. This work provides a foundation for progressing with determination of the structure of the E5 TM domain dimer and for subsequent analysis of the TM domains of the Ii-MHC complex by NMR spectroscopy.

Methods to study self-association are well established, however of particular interest is determining the association between TM domains of differing proteins, since these kinds of interactions are important in a range of biological processes in particular the assembly of polytopic membrane proteins, such as GPCRs. However, there are relatively few methods available to make such measurements, two being the GALLEX assay and FRET measurements. We sought to develop a method for achieving this using NMR spectroscopy. A method was developed that could be used to determine helix-helix interactions between TM domains, which would also be applicable to identifying any protein-protein interactions. The method uses the phenomenon of magnetic alignment of proteins with the magnetic field of the NMR spectrometer to enable the measurement of RDCs

from ^{15}N-labelled peptides. The principle behind the methods was discussed in Section 6.2. The method calls for the production of a peptide, TM-LBT, which is a fusion protein containing the TM domain of interest and an optimised amino acid sequence that binds lanthanide ions, or lanthanide binding tag (LBT). In order to test the method the TM domain of GpA was used since this is perhaps the most well characterised example of a dimeric TM domain.

Attempts to produce the GpA-LBT peptide using the approach of fmoc synthesis and RP-HPLC purification used for other peptides in this study was not successful. Analysis of the product from synthesis using mass spectrometry showed that synthesis had not been successful and that the peptide was present at very low yield. We believe the poor yield is due to the technical difficulties in producing a peptide with the contrasting hydrophobic and hydrophilic segments that the GpA-LBT possesses. We therefore decided to attempt to heterologously express the GpA-LBT peptide. If this approach was successful it would present a much more cost effective method for generating such peptides in the future.

During the course of this study the pET expression system for the expression of hydrophobic peptides had been evaluated for the production of isotopically labelled peptides for use in structure determination by NMR spectroscopy. The pET system involves expression of a fusion protein composed of the TM of interest and ketosteroid isomerase (KSI). The successful use of this method has been reported by several researchers. In this system cleavage of the fusion protein to release the desired peptide is reportedly achieved by reaction with CNBr which it is alleged, specifically reacts with junctional methionine residues to cleave the protein and yield the peptide with a C-terminal homoserinelactone residue. However, despite attempts with the TM sequences of Ii, MHCα and MHCβ, the cleavage products from the treatment of the fusion protein with CNBr could not be identified by mass spectrometry or peptide sequencing. Numerous attempts at optimising the conditions of the CNBr cleavage reaction were made but none yielded the desired results. Therefore, this method was abandoned and an alternative approach was sought for producing the TM-LBT peptide.

The use of the pGEX expression system has been reported to be successful in the production of hydrophobic peptides by heterologous expression. In this system the

TM is expressed as part of a fusion protein with a GST moiety. This approach enables the fusion protein to be purified using affinity chromatography. The expression of the GST-GpA-LBT fusion protein was optimised, as was the purification of the fusion protein. Cleavage of the fusion protein with protease was optimised, however it was subsequently found that the GpA-LBT peptide was not soluble in the chosen detergent. Therefore it is expected that a further step will be required to separate the sepharose and GST tag from the GpA-LBT peptide following cleavage of the GST. This is likely to involve solubilisation in organic solvent, e.g. TFE, followed by filtration to remove the sepharose and then separation of the cleavage products by RP-HPLC.

7.3 Future Directions

Future work on the TM domain interactions of Ii-MHC could explore the necessity for the TM domain of Ii to assemble into an oligomer before associating with the TM domains of the α- and β-chains of MHC. This would allow us to distinguish between models a and b in Figure 7.1. Dixon et al showed that mutations of Thr and Gln residues in Ii disrupted oligomerisation. The affect of the mutants on the oligomeric state of the Ii is still to be determined, and could be studied by cross-linking or FRET of peptide analogues of the mutants. The deleterious effect of the mutation of the Ii TM domain should be corroborated with the GALLEX assay prior to determining the association of the Ii mutants with the TM domains of α and β-chains of MHC using the same assay. The oligomeric state of the Ii mutants could also be assessed using FRET analyses in a similar manner to that used in this study. Consideration should also be made as to the discrepancy between the results from the TOXCAT and GALLEX assays for the GxxxG mutant of MHCβ. Future work on the role played by the TM domains of Ii and MHC in the formation of the Ii-MHC complex should also involve mutational studies conducted on the full-length proteins *in vivo.* Regarding the development of NMR methods, the work should progress to attempting to solve the solution state structure of the E5 TM domain and continue to develop the methods utilising paramagnetic lanthanide ions for determining the interactions and structures of transmembrane domains.

8 References

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

Adams, P. D., A. S. Lee, et al. (1998). "Models for the transmembrane region of the phospholamban pentamer: which is correct?" Ann N Y Acad Sci **853**: 178-85.

Anderson, M. S. and J. Miller (1992). "Invariant Chain Can Function as a Chaperone Protein for Class-Ii Major Histocompatibility Complex-Molecules." Proceedings of the National Academy of Sciences of the United States of America **89**(6): 2282-2286.

Antolini, F., M. Lo Bello, et al. (2003). "Purified promyelocytic leukemia coiled-coil aggregates as a tetramer displaying low alpha-helical content." Protein Expression and Purification **29**(1): 94-102.

Ashman, J. B. and J. Miller (1999). "A role for the transmembrane domain in the trimerization of the MHC class II-associated invariant chain." Journal of Immunology **163**(5): 2704-2712.

Barabanova, Y. A., H. K. Kang, et al. (2004). "Role of the major histocompatibility complex class II transmembrane region in antigen presentation and intracellular trafficking." Immunology **111**(2): 165-172.

Bax, A. (2003). "Weak alignment offers new NMR opportunities to study protein structure and dynamics." Protein Sci **12**(1): 1-16.

Bax, A. and A. Grishaev (2005). "Weak alignment NMR: a hawk-eyed view of biomolecular structure." Curr Opin Struct Biol **15**(5): 563-70.

Bertolino, P., M. Staschewski, et al. (1995). "Deletion of a C-terminal sequence of the class II-associated invariant chain abrogates invariant chains oligomer formation and class II antigen presentation." J Immunol **154**(11): 5620-9.

Bibi, E. and H. R. Kaback (1990). "In vivo expression of the lacY gene in two segments leads to functional lac permease." Proc Natl Acad Sci U S A **87**(11): 4325-9.

Bijlmakers, M. J., P. Benaroch, et al. (1994). "Mapping functional regions in the lumenal domain of the class II-associated invariant chain." J Exp Med **180**(2): 623-9.

Bijlmakers, M. J. J. E., P. Benaroch, et al. (1994). "Assembly of Hla Dr1 Molecules Translated in-Vitro - Binding of Peptide in the Endoplasmic-Reticulum Precludes Association with Invariant Chain." Embo Journal **13**(11): 2699-2707.

Bowie, J. U. (2005). "Solving the membrane protein folding problem." Nature **438**(7068): 581-9.

Bremnes, B., M. Rode, et al. (2000). "The MHC class II-associated chicken invariant chain shares functional properties with its mammalian homologs." Exp Cell Res **259**(2): 360-9.

Cardullo, R. A., S. Agrawal, et al. (1988). "Detection of Nucleic-Acid Hybridization by Nonradiative Fluorescence Resonance Energy-Transfer." Proceedings of the National Academy of Sciences of the United States of America **85**(23): 8790-8794.

Castellino, F., R. Han, et al. (2001). "The transmembrane segment of invariant chain mediates binding to MHC class II molecules in a CLIP-independent manner." Eur J Immunol **31**(3): 841-50.

Contreras, M. A., J. Ubach, et al. (1999). "Measurement of one bond dipolar couplings through lanthanide-induced orientation of a calcium-binding protein." Journal of the American Chemical Society **121**(38): 8947-8948.

Cresswell, P. (1994). "Assembly, Transport, and Function of Mhc Class-Ii Molecules." Annual Review of Immunology **12**: 259-293.

Daines, D. A. and R. P. Silver (2000). "Evidence for multimerization of neu proteins involved in polysialic acid synthesis in Escherichia coli K1 using improved LexA-based vectors." Journal of Bacteriology **182**(18): 5267-5270.

Dixon, A. M., B. J. Stanley, et al. (2006). "Invariant chain transmembrane domain trimerization: A step in MHC class II assembly." Biochemistry **45**(16): 5228-5234.

Dmitrova, M., G. Younes-Cauet, et al. (1998). "A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli." Molecular and General Genetics **257**(2): 205-212.

Duneau, J. P., A. P. Vegh, et al. (2007). "A dimerization hierarchy in the transmembrane domains of the HER receptor family." Biochemistry **46**(7): 2010-9.

Elliott, E. A., J. R. Drake, et al. (1994). "The Invariant Chain Is Required for Intracellular-Transport and Function of Major Histocompatibility Complex Class-Ii Molecules." Journal of Experimental Medicine **179**(2): 681-694.

Engelman, D. M. (2005). "Membranes are more mosaic than fluid." Nature **438**(7068): 578-580.

Ericson, M. L., M. Sundstrom, et al. (1994). "Mutually Exclusive Binding of Peptide and Invariant Chain to Major Histocompatibility Complex Class-Ii Antigens." Journal of Biological Chemistry **269**(42): 26531-26538.

Finger, C., T. Volkmer, et al. (2006). "The stability of transmembrane helix interactions measured in a biological membrane." J Mol Biol **358**(5): 1221-8.

Fisher, L. E. and D. M. Engelman (2001). "High-yield synthesis and purification of an alpha-helical transmembrane domain." Analytical Biochemistry **293**(1): 102-108.

Fisher, L. E., D. M. Engelman, et al. (1999). "Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain." J Mol Biol **293**(3): 639-51.

Fisher, L. E., D. M. Engelman, et al. (2003). "Effect of detergents on the association of the glycophorin a transmembrane helix." Biophys J **85**(5): 3097-105.

Fowler, C. A., F. Tian, et al. (2000). "Rapid determination of protein folds using residual dipolar couplings." J Mol Biol **304**(3): 447-60.

Frangioni, J. V. and B. G. Neel (1993). "Solubilization and Purification of Enzymatically Active Glutathione-S-Transferase (Pgex) Fusion Proteins." Analytical Biochemistry **210**(1): 179-187.

Franz, K. J., M. Nitz, et al. (2003). "Lanthanide-binding tags as versatile protein coexpression probes." Chembiochem **4**(4): 265-71.

Frauwirth, K., S. Sanderson, et al. (1995). "Presentation of Endogenous Antigens on Class-Ii Mhc - Regulation by the Invariant Chain." Journal of Cellular Biochemistry: 95-95.

Frauwirth, K. and N. Shastri (2001). "Mutation of the invariant chain transmembrane region inhibits Ii degradation, prolongs association with MHC class II, and selectively disrupts antigen presentation." Cellular Immunology **209**(2): 97-108.

Gasteiger, E., C. Hoogland, et al. (2005). Protein Identification and Analysis Tools on the EXPASY Server. The Proteomics Protocols Handbook. J. M. Walker, Humana Press**:** 571-607.

Gedde-Dahl, M., I. Freisewinkel, et al. (1997). "Exon 6 is essential for invariant chain trimerization and induction of large endosomal structures." J Biol Chem **272**(13): 8281-7.

Gerber, D., N. Sal-Man, et al. (2004). "Two motifs within a transmembrane domain, one for homodimerization and the other for heterodimerization." J Biol Chem **279**(20): 21177-82.

Germain, R. N. and A. G. Rinker, Jr. (1993). "Peptide binding inhibits protein aggregation of invariant-chain free class II dimers and promotes surface expression of occupied molecules." Nature **363**(6431): 725-8.

Gratkowski, H., Q. H. Dai, et al. (2002). "Cooperativity and specificity of association of a designed transmembrane peptide." Biophysical Journal **83**(3): 1613-1619.

Gratkowski, H., J. D. Lear, et al. (2001). "Polar side chains drive the association of model transmembrane peptides." Proc Natl Acad Sci U S A **98**(3): 880-5.

Groves, J. D. and M. J. Tanner (1995). "Co-expressed complementary fragments of the human red cell anion exchanger (band 3, AE1) generate stilbene disulfonate-sensitive anion transport." J Biol Chem **270**(16): 9097-105.

Harrington, S. E. and N. Ben-Tal (2009). "Structural determinants of transmembrane helical proteins." Structure **17**(8): 1092-103.

Hessa, T., S. H. White, et al. (2005). "Membrane insertion of a potassium-channel voltage sensor." Science **307**(5714): 1427.

Holmdahl, R. (2000). "Association of MHC and rheumatoid arthritis - Why is rheumatoid arthritis associated with the MHC genetic region? An introduction." Arthritis Research **2**(3): 203-204.

Ikegami, T., L. Verdier, et al. (2004). "Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions." J Biomol NMR **29**(3): 339-49.

Ishigami, S., S. Natsugoe, et al. (2001). "Invariant chain expression in gastric cancer." Cancer Lett **168**(1): 87-91.

Jacobs, R. E. and S. H. White (1989). "The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices." Biochemistry **28**(8): 3421-37.

Jasanoff, A., G. Wagner, et al. (1998). "Structure of a trimeric domain of the MHC class II-associated chaperonin and targeting protein Ii." Embo J **17**(23): 6812-8.

Johnson, A. E. (2005). "Fluorescence approaches for determining protein conformations, interactions and mechanisms at membranes." Traffic **6**(12): 1078-92.

Johnson, R. M., A. Rath, et al. (2006). "The position of the Gly-xxx-Gly motif in transmembrane segments modulates dimer affinity." Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire **84**(6): 1006-1012.

Johnson, W. C. (1999). "Analyzing protein circular dichroism spectra for accurate secondary structures." Proteins **35**(3): 307-12.

Jones, D. H., E. H. Ball, et al. (2000). "Expression and membrane assembly of a transmembrane region from Neu." Biochemistry **39**(7): 1870-1878.

Jones, E. Y., L. Fugger, et al. (2006). "MHC class II proteins and disease: a structural perspective." Nat Rev Immunol **6**(4): 271-82.

King, G. and A. Dixon (2008). Techniques to study membrane protein structure. Wiley Encyclopedia of Chemical Biology, Wiley Blackwell Publishing.

Kobilka, B. K. (2007). "G protein coupled receptor structure and activation." Biochim Biophys Acta **1768**(4): 794-807.

Kobus, F. J. and K. G. Fleming (2005). "The GxxxG-containing transmembrane domain of the CCK4 oncogene does not encode preferential self-interactions." Biochemistry **44**(5): 1464-1470.

Koch, N., A. D. McLellan, et al. (2007). "A revised model for invariant chain-mediated assembly of MHC class II peptide receptors." Trends Biochem Sci **32**(12): 532-7.

Kochendoerfer, G. G., D. Salom, et al. (1999). "Total chemical synthesis of the integral membrane protein influenza A virus M2: Role of its C-terminal domain in tetramer assembly." Biochemistry **38**(37): 11905-11913.

Kukol, A., J. Torres, et al. (2002). "A structure for the trimeric MHC class II-associated invariant chain transmembrane domain." Journal of Molecular Biology **320**(5): 1109-1117.

Kvist, S., K. Wiman, et al. (1982). "Membrane insertion and oligomeric assembly of HLA-DR histocompatibility antigens." Cell **29**(1): 61-9.

Lamb, C. A. and P. Cresswell (1992). "Assembly and transport properties of invariant chain trimers and HLA-DR-invariant chain complexes." J Immunol **148**(11): 3478-82.

Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics **23**(21): 2947-8.

le Maire, M., P. Champeil, et al. (2000). "Interaction of membrane proteins and lipids with solubilizing detergents." Biochimica Et Biophysica Acta-Biomembranes **1508**(1-2): 86-111.

Lemmon, M. A., J. M. Flanagan, et al. (1992). "Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices." J Biol Chem **267**(11): 7683-9.

Lemmon, M. A., J. M. Flanagan, et al. (1992). "Sequence specificity in the dimerization of transmembrane alpha-helices." Biochemistry **31**(51): 12719-25.

Lew, S. and E. London (1997). "Simple procedure for reversed-phase high-performance liquid chromatographic purification of long hydrophobic peptides that form transmembrane helices." Anal Biochem **251**(1): 113-6.

Li, E., M. You, et al. (2005). "Sodium dodecyl sulfate-polyacrylamide gel electrophoresis and forster resonance energy transfer suggest weak interactions between fibroblast growth factor receptor 3 (FGFR3) transmembrane domains in the absence of extracellular domains and ligands." Biochemistry **44**(1): 352-60.

Lotteau, V., L. Teyton, et al. (1990). "Intracellular-Transport of Class-Ii Mhc Molecules Directed by Invariant Chain." Nature **348**(6302): 600-605.

Luo, P. Z. and R. L. Baldwin (1997). "Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water." Biochemistry **36**(27): 8413-8421.

Luo, S. Z., X. Mo, et al. (2007). "Glycoprotein Ib alpha forms disulfide bonds with 2 glycoprotein Ib beta subunits in the resting platelet." Blood **109**(2): 603-609.

Ma, C. and S. J. Opella (2000). "Lanthanide ions bind specifically to an added "EF-hand" and orient a membrane protein in micelles for solution NMR spectroscopy." J Magn Reson **146**(2): 381-4.

MacKenzie, K. R. (2006). "Folding and stability of alpha-helical integral membrane proteins." Chemical Reviews **106**(5): 1931-1977.

MacKenzie, K. R., J. H. Prestegard, et al. (1996). "Leucine side-chain rotamers in a glycophorin A transmembrane peptide as revealed by three-bond carbon-carbon couplings and C-13 chemical shifts." Journal of Biomolecular Nmr **7**(3): 256-260.

MacKenzie, K. R., J. H. Prestegard, et al. (1997). "A transmembrane helix dimer: Structure and implications." Science **276**(5309): 131-133.

Marks, M. S., J. S. Blum, et al. (1990). "Invariant chain trimers are sequestered in the rough endoplasmic reticulum in the absence of association with HLA class II antigens." J Cell Biol **111**(3): 839-55.

Mattoon, D., K. Gupta, et al. (2001). "Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein." Oncogene **20**(29): 3824-3834.

Melnyk, R. A., A. W. Partridge, et al. (2003). "Polar residue tagging of transmembrane peptides." Biopolymers **71**(6): 675-85.

Merzlyakov, M. and K. Hristova (2008). "Forster Resonance Energy Transfer Measurements of Transmembrane Helix Dimerization Energetics." Fluorescence Spectroscopy **450**: 107-127.

Merzlyakov, M., M. You, et al. (2006). "Transmembrane helix heterodimerization in lipid bilayers: probing the energetics behind autosomal dominant growth disorders." J Mol Biol **358**(1): 1-7.

Neumann, J. and N. Koch (2005). "Assembly of major histocompatibility complex class II subunits with invariant chain." Febs Letters **579**(27): 6055-6059.

Neumann, J. and N. Koch (2006). "A novel domain on HLA-DRbeta chain regulates the chaperone role of the invariant chain." J Cell Sci **119**(Pt 20): 4207-14.

Nitz, M., K. J. Franz, et al. (2003). "A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides." Chembiochem **4**(4): 272-276.

Nitz, M., M. Sherawat, et al. (2004). "Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide." Angewandte Chemie-International Edition **43**(28): 3682-3685.

Oates, J., M. Hicks, et al. (2008). "In Vitro Dimerization of the Bovine Papillomavirus E5 Protein." Biochemistry **47**: 8985-8992.

Peterson, M. and J. Miller (1992). "Antigen presentation enhanced by the alternatively spliced invariant chain gene product p41." Nature **357**(6379): 596-8.

Pintacuda, G., M. A. Keniry, et al. (2004). "Fast structure-based assignment of 15N HSQC spectra of selectively 15N-labeled paramagnetic proteins." J Am Chem Soc **126**(9): 2963-70.

Pintacuda, G., A. Y. Park, et al. (2006). "Lanthanide labeling offers fast NMR approach to 3D structure determinations of protein-protein complexes." J Am Chem Soc **128**(11): 3696-702.

Popot, J. L. and D. M. Engelman (1990). "Membrane protein folding and oligomerization: the two-stage model." Biochemistry **29**(17): 4031-7.

Popot, J. L., S. E. Gerchman, et al. (1987). "Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process." J Mol Biol **198**(4): 655-76.

Popot, J. L., J. Trewhella, et al. (1986). "Reformation of crystalline purple membrane from purified bacteriorhodopsin fragments." Embo J **5**(11): 3039-44.

Prudencio, M., J. Rohovec, et al. (2004). "A caged lanthanide complex as a paramagnetic shift agent for protein NMR." Chemistry **10**(13): 3252-60.

Randall, L. L. and S. J. Hardy (1986). "Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in E. coli." Cell **46**(6): 921-8.

Rath, A., M. Glibowicka, et al. (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins." Proc Natl Acad Sci U S A **106**(6): 1760-5.

Rath, A., R. M. Johnson, et al. (2007). "Peptides as transmembrane segments: decrypting the determinants for helix-helix interactions in membrane proteins." Biopolymers **88**(2): 217-32.

Ridge, K. D., S. S. Lee, et al. (1995). "In vivo assembly of rhodopsin from expressed polypeptide fragments." Proc Natl Acad Sci U S A **92**(8): 3204-8.

Rodriguez-Castaneda, F., P. Haberz, et al. (2006). "Paramagnetic tagging of diamagnetic proteins for solution NMR." Magn Reson Chem **44 Spec No**: S10-6.

Romagnoli, P., C. Layet, et al. (1993). "Relationship between invariant chain expression and major histocompatibility complex class II transport into early and late endocytic compartments." J Exp Med **177**(3): 583-96.

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

Sambrook, J. and D. W. Russell (2001). Molecular cloning: a laboratory manual. , Cold spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sanders, C. R. and J. K. Myers (2004). "Disease-related misassembly of membrane proteins." Annu Rev Biophys Biomol Struct **33**: 25-51.

Schafer, P. H., S. Malapati, et al. (1998). "The assembly and stability of MHC class II-(alpha beta)(2) superdimers." Journal of Immunology **161**(5): 2307-2316.

Schindler, M., S. Wurfl, et al. (2003). "Down-modulation of mature major histocompatibility complex class II and up-regulation of invariant chain cell surface expression are well-conserved functions of human and simian immunodeficiency virus nef alleles." J Virol **77**(19): 10548-56.

Schneider, D. (2004). "Rendezvous in a membrane: close packing, hydrogen bonding, and the formation of transmembrane helix oligomers." FEBS Lett **577**(1-2): 5-8.

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

Schuck, P. (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling." Biophysical Journal **78**(3): 1606-1619.

Sculimbrene, B. R. and B. Imperiali (2006). "Lanthanide-binding tags as luminescent probes for studying protein interactions." J Am Chem Soc **128**(22): 7346-52.

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

Senes, A., M. Gerstein, et al. (2000). "Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions." J Mol Biol **296**(3): 921-36.

Seubert, N., Y. Royer, et al. (2003). "Active and inactive orientations of the transmembrane and cytosolic domains of the erythropoietin receptor dimer." Molecular Cell **12**(5): 1239-1250.

Simonsen, A., F. Momburg, et al. (1993). "Intracellular distribution of the MHC class II molecules and the associated invariant chain (Ii) in different cell lines." Int Immunol **5**(8): 903-17.

Singer, S. J. and G. L. Nicolson (1972). "The fluid mosaic model of the structure of cell membranes." Science **175**(23): 720-31.

Staros, J. V. (1982). "N-hydroxysulfosuccinimide active esters: bis(N-hydroxysulfosuccinimide) esters of two dicarboxylic acids are hydrophilic, membrane-impermeant, protein cross-linkers." Biochemistry **21**(17): 3950-5.

Stern, L. J., J. H. Brown, et al. (1994). "Crystal-Structure of the Human Class-Ii Mhc Protein Hla-Dr1 Complexed with an Influenza-Virus Peptide." Nature **368**(6468): 215-221.

Sung, E. and P. P. Jones (1981). "The invariant chain of murine Ia antigens: its glycosylation, abundance and subcellular localization." Mol Immunol **18**(10): 899-913.

Talbert-Slagle, K., S. Marlatt, et al. (2009). "Artificial transmembrane oncoproteins smaller than the bovine papillomavirus E5 protein redefine sequence requirements for activation of the platelet-derived growth factor {beta} receptor." J Virol.

Teyton, L., D. Osullivan, et al. (1990). "Invariant Chain Distinguishes between the Exogenous and Endogenous Antigen Presentation Pathways." Nature **348**(6296): 39-44.

Thery, C., V. Brachet, et al. (1998). "MHC class II transport from lysosomal compartments to the cell surface is determined by stable peptide binding, but not by the cytosolic domains of the alpha- and beta-chains." Journal of Immunology **161**(5): 2106-2113.

Tiburu, E. K., P. C. Dave, et al. (2003). "An improved synthetic and purification procedure for the hydrophobic segment of the transmembrane peptide phospholamban." Analytical Biochemistry **318**(1): 146-151.

Treptow, N. A. and H. A. Shuman (1985). "Genetic-Evidence for Substrate and Periplasmic-Binding-Protein Recognition by the Malf and Malg Proteins, Cytoplasmic Membrane-Components of the Escherichia-Coli Maltose Transport-System." Journal of Bacteriology **163**(2): 654-660.

Veatch, W. and L. Stryer (1977). "Dimeric Nature of Gramicidin-a Transmembrane Channel - Conductance and Fluorescence Energy-Transfer Studies of Hybrid Channels." Journal of Molecular Biology **113**(1): 89-102.

Von Heijne, G. (2003). "Membrane protein assembly in vivo." Membrane Proteins **63**: 1-18.

Vranken, W. F., W. Boucher, et al. (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." Proteins **59**(4): 687-96.

Vyas, J. M., A. G. Van der Veen, et al. (2008). "The known unknowns of antigen processing and presentation." Nat Rev Immunol **8**(8): 607-18.

Walkenhorst, W. F., M. Merzlyakov, et al. (2009). "Polar residues in transmembrane helices can decrease electrophoretic mobility in polyacrylamide gels without causing helix dimerization." Biochim Biophys Acta.

Watts, C. (2004). "The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules." Nat Immunol **5**(7): 685-92.

White, S. H. (2003). "Translocons, thermodynamics, and the folding of membrane proteins." FEBS Lett **555**(1): 116-21.

White, S. H. (2006). "How hydrogen bonds shape membrane protein structure." Peptide Solvation and H-Bonds **72**: 157-+.

White, S. H., A. S. Ladokhin, et al. (2001). "How membranes shape protein structure." Journal of Biological Chemistry **276**(35): 32395-32398.

White, S. H. and W. C. Wimley (1999). "Membrane protein folding and stability: physical principles." Annu Rev Biophys Biomol Struct **28**: 319-65.

Wohnert, J., K. J. Franz, et al. (2003). "Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings." J Am Chem Soc **125**(44): 13338-9.

Wu, Y., H. W. Huang, et al. (1990). "Method of oriented circular dichroism." Biophys J **57**(4): 797-806.

Wu, Y., S. C. Shih, et al. (2007). "Probing the structure of the Ff bacteriophage major coat protein transmembrane helix dimer by solution NMR." Biochim Biophys Acta **1768**(12): 3206-15.

Ye, Q., P. W. Finn, et al. (2003). "MHC class II-associated invariant chain isoforms regulate pulmonary immune responses." J Immunol **170**(3): 1473-80.

You, M., E. Li, et al. (2005). "Forster resonance energy transfer in liposomes: measurements of transmembrane helix dimerization in the native bilayer environment." Anal Biochem **340**(1): 154-64.

Yu, X. C., K. D. Sharma, et al. (2002). "Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling." Molecular Biology of the Cell **13**(7): 2547-2557.

Zhou, F. X., H. J. Merianos, et al. (2001). "Polar residues drive association of polyleucine transmembrane helices." Proc Natl Acad Sci U S A **98**(5): 2250-5.