THE EXPRESSION AND ANALYSIS OF TETRASPANIN LARGE EXTRACELLULAR DOMAINS

FRANCINE MARTIN. BSc.

Department of Molecular Biology and Biotechnology

University of Sheffield

February 2006

ABSTRACT

The Expression and Analysis of Tetraspanin Large Extracellular Domains

There are 33 tetraspanin proteins in mammals and these are widely expressed on the majority of cells and tissues. Tetraspanin proteins are differentiated from other tetraspan proteins by a number of conserved motifs, in particular an absolutely conserved Cys-Cys-Gly motif in the large extracellular loop (EC2). This EC2 domain is the most variable region between tetraspanin proteins in length and amino acid sequence, and consequently is the most intriguing domain. Soluble recombinant EC2 domains have been used to study tetraspanin function, and so one of the main aims of this work was to optimise the production of GST-EC2 proteins from CD9, CD81, CD63, CD151. Methods were developed to produce these domains in large quantities, and they were shown to have biological activity in the inhibition of giant cell formation in human monocytes. The soluble EC2 domains from CD63, CD9, CD81 and CD151 were all shown to potently inhibit HIV-1 infection of macrophages. Infection of the two major HIV-1 strains R5 and X4 was prevented by all the tetraspanins tested, suggesting an involvement of a tetraspanin enriched microdomain. The inhibition was macrophage specific since infection of PBMC was not prevented. Another aim was to clone and express EC2 domains from CD82 and CD231. CD82 and CD231 EC2 domains are larger and consequently proved more difficult to express. Both were cloned successfully into the pGEX-2T expression vector and CD82 was expressed in Origami cells (Novagen). CD82 EC2 was considered to be a valuable tool to make because of the involvement of CD82 in tumor cell metastasis. The original aim of the work was to characterise the binding between tetraspanin EC2 domains and primary ligands. However, no appropriate ligands were found. We did show by ELISA that mouse, but not human, CD9 bound to murine PGS17 but we could not detect the reported binding between CD9 and fibronectin. Additionally, efforts were made to clone and express two immunoglobulin-like proteins that had recently been identified as CD9 and CD81 primary binding partners. The cloning of full length EWI DNA as well and double Ig domains was carried out successfully into four different expression vectors, however, no protein expression was detected in eukaryotic and a range of prokaryotic expression systems. Structural studies were performed using CD63 EC2. Strategies to cleave GST and to purify the EC2 domain were successfully developed. Instability problems of CD63 EC2 in the absence of GST did not allow for concentration of the cleaved product to levels required for crystal trials. CD63 EC2-His6 recombinant protein was used in X-ray crystallography studies, and over 1000 crystallisation conditions were tested in preliminary structural studies.

Acknowledgements

I would like to thank my supervisors, Peter Monk and Lynda Partridge and co-workers at the University of Sheffield, in particular Adrian Higginbottom. I also thank Pat Baker and the X-ray crystallography lab for their help and advice in the crystallography work described here.

I would like to thank the numerous collaborators who have contributed to the work described in this thesis, in particular Cecilia Cheng Mayer, Siu-Hong Ho, Peter Lopez, Gregory Moseley and Varadarajan Parthasarathy for the HIV-1 work and Christopher Liu for supplying the CD63 EC2-His6.

I thank the Humane Research Trust for the funding for this work.

This PhD would not have been completed without the continual support of my friends and family, in particular Andrew Parker.

Abbreviations

BCR	.....	B-cell receptor
BSA	.....	Bovine serum albumin
BSS	.....	Balanced salt solution
CD	.....	Cluster of differentiation
CEA	.....	Carcinoembryonic antigen
CDV	.....	Canine distemper virus
CHO	.....	Chinese hamster ovary
Da	.....	Daltons
DMEM	.....	Dulbecco's Modified Eagle's Medium
DMSO	.....	Dimethyl sulphoxide
EC2	.....	Tetraspanin large extracellular domain
ECM	.....	Extracellular matrix
E. coli	.....	*Escherichia Coli*
EDTA	.....	Ethylenediaminine tetra-acetic acid
ELISA	.....	Enzyme-linked immunosorbent assay
Fc	.....	Fragment crystalline
FIV	.....	Feline immunodeficiency virus
FN	.....	Fibronectin
FPRP	.....	Prostaglandin F2α receptor regulatory protein
GFP	.....	Green fluorescent protein
GST	.....	Glutathione-S-transferase
HB-EGF	.....	Heparin binding – epidermal growth factor
HBSS	.....	Hanks' balanced salt solution
HCV	.....	Hepatitis C Virus
His6	.....	6 x histidine tag
HIV	.....	Human immunodeficiency virus
HRP	.....	Horseradish peroxidase
HTLV-1	.....	Human T-leukaemia virus-1
Ig	.....	Immunoglobulin
IgSF	.....	Immunoglobulin superfamily
IFN	.....	Interfoeron
kDal	.....	kilo Dalton
LB	.....	Luria-Bertani
mAb	.....	Monoclonal antibody
MDM	.....	Monocyte derived macrophages
MHC	.....	Major histocompatibility complex
MNC	.....	Mononuclear cell
MVB	.....	Multivesicular bodies
NK	.....	Natural Killer cells
PBMC	.....	Peripheral blood mononuclear cells
PBS	.....	Phosphate buffered saline
PCR	.....	Polymerase chain reaction
PGRL	.....	Prostaglandin regulatory-like protein

PI4 kinase	.....	Phosphatidyl inositol 4 kinase
PSG	.....	Pregnancy specific glycoprotein
RBL	.....	Rat basophilic leukaemia
RT	.....	Room temperature
SDS-PAGE	.....	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	.....	Standard error of the mean
TCR	.....	T-cell receptor
Tween-20	.....	Polyoxyethylene sorbitan monolaurate
UV	.....	Ultra violet
WT	.....	Wild type

Amino Acid Abbreviations

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

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 THE TETRASPANIN SUPERFAMILY

Over the past decade there has been increasing interest in a recently identified family of membrane proteins, the tetraspanins (transmembrane 4 (TM4) superfamily, TM4SF, 4TM or tetraspan family). This large and abundantly expressed family of proteins that span the membrane four times is highly intriguing. At least one member of the family appears to be expressed on all mammalian cells and their unique structure gives no obvious clues as to their function. The first tetraspanin to be discovered was the melanoma-associated protein, ME491, now more commonly known as CD63 (Hotta 1988). CD9 was originally described as a surface antigen on T-cells and developing B lymphocytes and was also known to be highly expressed on platelets. In 1991, the CD9 antigen was cloned and shown to be similar to CD63 (Boucheix et al. 1991): the novel family of proteins was named the tetraspanin family (Horejsi et al. 1991). Murine CD81 was identified due to its role as the target of antiproliferative antibody-1 (TAPA-1), and this mouse antigen was later found to be similar to human CD81 and CD9 (Andria et al. 1991; Levy et al. 1991; Andria et al. 1992). The advances in gene cloning and sequence database technologies have allowed the rapid identification of more tetraspanin family members, and in 1998 alone, six new members were discovered (Todd et al. 1998). Now there are 32 or 33 members of the tetraspanin family in humans (Table 1.1), 37 in *Drosophila*, and 20 in *Caenorhabditis elegans*. The conserved nature of tetraspanins over such diverse species suggests an ancient evolutionary origin of these membrane spanning proteins,

and it is likely that homologous members of this family will be discovered in other species as more genomes are researched.

Indeed, in the invertebrates *Schistosoma mansoni* and *Schistosoma japonicum* cell surface proteins have been identified and assigned to the tetraspanin family. PLSI, a tetraspanin encoding gene, has been identified in a plant pathogenic fungus, *Magriaporthe grisea* as well as other fungi including *B. cinerea, N. crassa* and *C. lindemuthianum* (Gourgues et al. 2002).

Tetraspanins are expressed on cell surface membranes and are also found in intracellular compartments, such as the endosome and lysosome. Some tetraspanins, such as CD9, are ubiquitously expressed whereas others are restricted to particular cell lineages (Table 1.2). Recent studies have shown that tetraspanins are implicated in a wide range of diverse biological functions. Here, the properties of mammalian tetraspanin proteins will be discussed and an overview of the main tetraspanin functions relevant to this work will be given.

Tetraspanin	Alternative names	New Name#1	Protein #	Chr
MGC50844	Penumbra	TM4SF24?	NP_848657	7q32.3
NET1	TM4-C	TM4SF17	NP_005718	1p34.1
Tspan2	FLJ12082	TM4SF14	NP_005716	1p13.1
Tspan3 *	TM4-A, TM4SF8.1,2	TM4SF8	NP_005715, NP_944492	15q24.3
NAG2	TM4SF7	TM4SF7	NP_003262	11p15.5
NET4		TM4SF9	NP_005714	4q23
TM4SF6	T245	TM4SF6	NP_003261	Xq22
CD231	A15,MXS1,CCG-B7,TALLA-1, TM4SF2b,DXS1692E	TM4SF2	NP_004606	Xp11.4
CO-029		TM4SF3	NP_004607	12q14.1q21.1
NET5		TM4SF19	NP_006666	12p13.33-32
Oculospanin		TM4SF21	NP_114151	17q25.3
TSPAN11	DN930492		ENST00000261177	12p11.21
NET2		TMFSF12	NP_036470	7q31.31
NET6	TM4SF13, FLJ22934	TM4SF13	NP_055214	7p21.1
DC-TM4F2	DC-TM4F2	TM4SF14	NP_112189	10q22.3
NET7	2700063A19Rik	TM4SF15	NP_036471	10q22.1
TM4B		TM4SF16	NP_036598	19p13.2
CAD35489	F-Box23, SB134	TM4SF17b	NP_569732	5q31
BAB55318		TM4SF16	NP_570139	11p11.2
XP_084868		TM4SF20?	XP_084868	12q21.31
Uroplakin Ib	UPIB, UPK1	TM4SF30	NP_008883	3q13.3-q21
Uroplakin Ia	UP1A, UPIA, UPKA, MGC14388	TM4SF31	NP_008931	19q13.13
Peripherin	RP7, AVMD, PRPH, AOFMD, PRPH2	TM4SF29	NP_000313	6p21.2-p12.3
Rom1	ROSP1	TM4SF28	NP_000318	11q13
CD151 (2 variants, UTR)	PETA3, SFA1, gp27	TM4SF32	NP_004348, NP_6205999	11p15.5
CD53	MOX44	TM4SF5	NP_000551	1p13
CD37	GP52-40	TM4SF4	NP_001765	19p13-q13.4
CD82	Kangai1,R2, 4F9,C33,IA4,ST6GR15,KAI1, SAR2	TM4SF11	NP_002222	11p11.2
CD81	TAPA-1, S5.7	TM4SF10	NP_004347	11p15.5
CD9	BA2, P24, GIG2, MIC3, MRP-1, BTCC-1, DRAP-27	TM4SF2	NP_001760	12p13.3
CD63	MEL1,MLA1,ME491, granulophysin, LAMP3, OMA81H,	TM4SF1	NP_001771	12q12-q13
SAS		TM4SF26	NP_005972	12q13.3
TSSC6 (3 variants; UTR)	PHEMX, MGC22455	TM4SF15	NP_620593, NP_620591, NP_005696	11p15.5

Table 1.1. The human tetraspanin proteins. The table lists the 33 human tetraspanin proteins. Alternative names, TM4SF names and the suggested new nomenclature are also shown. *isoform 2 is 30 residues shorter than i1.

Tetraspanin	Tissue distribution	Refs
CD9	Platelets, early B cells, activated and differentiating B cells, activated T cells, eosinophils, basophils, endothelial cells, megakaryocytes, epithelium, dendritic cells, brain and peripheral nerves, fibroblasts, lung, kidney, liver, vascular smooth muscle, skeletal muscle, keratinocytes	(Boucheix 2001); (Wright 1994); (Sincock 1997)
CD63	Wide lymphoid and non-lymphoid distribution including platelets, neutrophils, monocytes, macrophages, dendritic cells,endothelium, megakaryocytes, epithelium, fibroblasts, lung, kidney, liver, smooth muscle, skeletal muscle, peripheral nerves, pancreas, cardiac muscle	(Boucheix 2001); (Wright 1994); (Sincock 1997)
CD81	Broad expression on non-lymphoid tissues and on lymphocytes, thymocytes, follicular dendritic cells, eosinophils and epithelium	(Levy 1998); (Jones 1997); (Silvie 2003)
CD82	B and T cells, NK cells, monocytes, granulocytes, platelets. Various non-haemopoietic cells.	(Boucheix 2001); (Lagaudriere-Gesbert 1997)
CD151	Endothelial cells, platelets, dendritic cells, megakaryocytes, epithelium, lung, kidney, liver, smooth and skeletal muscle and peripheral nerves, keratinocytes, pancreas, cardiac muscle	(Ashman 1997; Sincock 1997)
CD231	Heart, brain, lung, liver, skeletal muscle, kidney, pancreas, immature T-cells	(Zemni R 2000); (Wright 1994)
Rom1	Photoreceptor cells	(Goldberg 1996)
Peripherin	Photoreceptor cells	(Goldberg 1996)

Table 1.2. Tissue distribution of a selection of the better characterised tetraspanin proteins.

1.2 TETRASPANIN STRUCTURE AND DOMAIN PROPERTIES

Tetraspanins are type III membrane glycoproteins which traverse the membrane four times producing short intracellular N- and C-termini and two extracellular loops (the small extracellular loop (SEL or EC1) and the large extracellular loop (LEL or EC2)), as well as a small intracellular loop located between TM2 and TM3 (Figure 1.1). Members of the tetraspanin family are differentiated from other tetraspan proteins by a number of structural and sequence motifs that are discussed in this chapter.

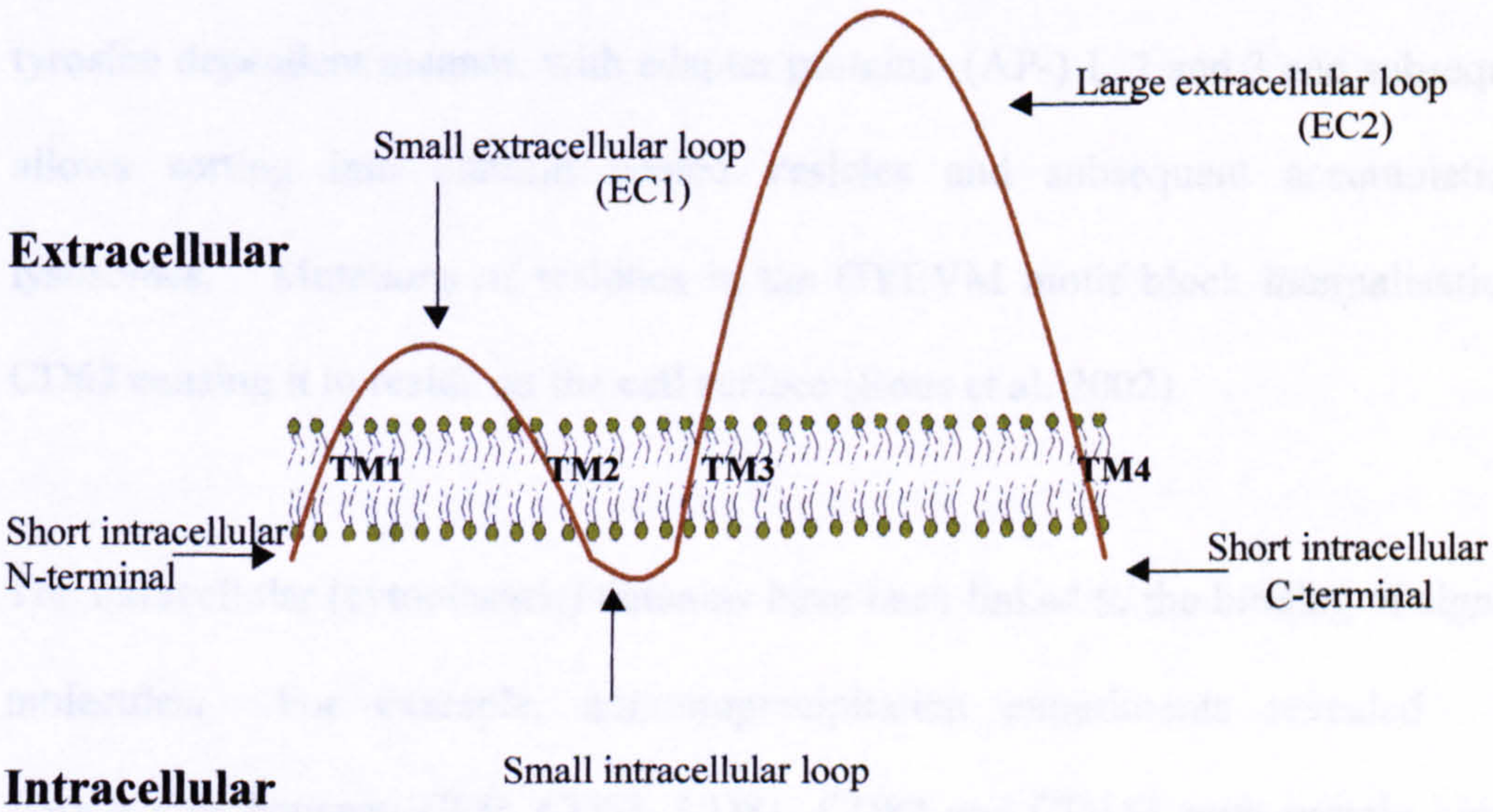


Figure 1.1. Simplified schematic sketch of a tetraspanin. This simplified diagram shows the basic outline of a tetraspanin protein. The small, intracellular termini and the intracellular loop between transmembrane regions 2 and 3 are noted, as are the extracellular loops; EC1 between TM1 and TM2 and the larger EC2 between TM3 and TM4.

1.2.1 Intracellular domains

The intracellular N- and C- termini and the short intracellular loop between TM2 and TM3 are 72 – 78 % identical between mammal and zebra fish homologues, but share only 21 – 38 % identity among different family members (Stipp 2003), suggesting important conserved functions that differ between each tetraspanin. The C- terminal tail is the most divergent of the intracellular domains and so may have functional significance. For instance, this region of CD63 possesses the lysosomal targeting motif YXXØ, where X is any amino acid and Ø is a bulky hydrophobic amino acid (Bonifacino et al. 1999). CD63 possess the motif GYEVM which interacts, in a tyrosine dependent manner, with adapter proteins- (AP-) 1, 2 and 3 and subsequently allows sorting into clathrin coated vesicles and subsequent accumulation in lysosomes. Mutations of residues in the GYEVM motif block internalisation of CD63 causing it to reside on the cell surface (Rous et al. 2002).

The intracellular (cytoplasmic) domains have been linked to the binding of signalling molecules. For example, immunoprecipitation experiments revealed direct associations between CD9, CD53, CD81, CD82 and CD151 with protein kinase C (PKC) (Zhang et al. 2001) and mutagenesis studies demonstrated that TM1, TM2 and the intracellular domain are likely to be involved in these interactions. In addition, phosphatidylinositol 4-kinase has been found in complexes with CD151, CD63 and CD81 (Berditchevski et al. 1997; Berditchevski et al. 2002).

1.2.2 Transmembrane regions

The hydrophobic transmembrane domains are the regions of greatest homology between humans and zebra fish (86 %) underlying the crucial function of these regions (Stipp et al. 2003). Transmembrane regions are thought to form the interface of both homo- and hetero tetraspanin interactions in tetraspanin rich microdomains (TEMs) (see 1.3.). Highly conserved polar residues in the TM regions are postulated to promote tight packing and stabilisation of the membrane spanning regions. Juxtamembrane cysteine residues in all 4 TM regions have been identified by site-directed mutagenesis to be palmitoylation sites in CD9 and CD151 (discussed in 1.3). The importance of the TM regions was emphasised by the use of truncated forms of CD82 lacking one or more transmembrane regions (Cannon et al. 2001); these truncated forms accumulate in the ER, have no Golgi processing and are not transported to the cell surface. Absence of TM3 prevents the folding of the EC2 suggesting that the TM3 and TM4 play crucial roles in bringing together membrane proximal regions of the EC2 to allow proper folding to occur (Cannon et al. 2001). Recent structural studies using molecular modelling computer programs and the crystal structure of CD81 EC2 predict that the transmembrane helices in the majority of tetraspanins are positioned in a square pattern in the membrane and are left-handed coiled-coils (Seigneuret 2006) (Figure 1.2).

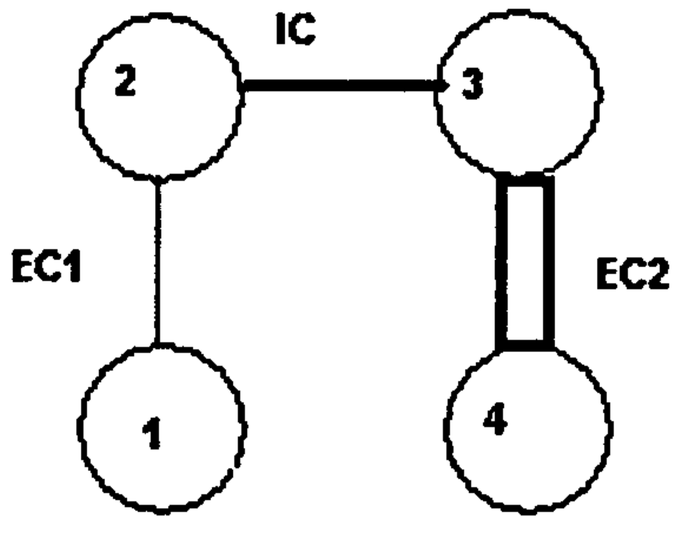


Figure 1.2. Arrangement of helices in the TM bundle. The helices are arranged in a square in the membrane. The EC1, EC2 and intracellular loop (IC) are shown.

1.2.3 The small extracellular domain (EC1)

The small extracellular region, EC1 (small extracellular loop SEL), is located between TM1 and TM2 and is composed of 13 – 30 amino acids. It has been reported that the EC1 is required for optimal surface expression of the larger extracellular loop (Masciopinto et al. 2001), where is was noted that in the absence of the EC1, EC2 is expressed at lower levels because it is retained intracellularly. Molecular modelling of EC1 domains has predicted small β-strands enriched in hydrophobic residues in the EC1 of most tetraspanins (Seigneuret 2006).

1.2.4 The large extracellular loop (EC2)

The EC2 domains show more variation than the other tetraspanin domains both in length and amino acid sequence (Figure 1.2). These large extracellular domains are typically composed of between 70 and 140 amino acids and are topologically located between TM3 and TM4. In 2001, the crystal structure of EC2 from CD81 was determined at 1.6 Å resolution (Kitadokoro et al. 2001). As previously predicted by modelling, it was confirmed that the CD81 EC2 is composed of 5 helical regions (A-E) spanning residues:- Asn 115-Asp 137, Ala 140–Asp 155, Leu 165–Asn 172, Asn

180–Phe 186 and Asp 189–Gly 200. Helices A, B and E are collectively known as the constant region and form a hydrophobic surface, that was initially thought to be involved in CD81 homodimerization as the crystal studies revealed that this region was the point of contact between adjacent CD81 EC2 proteins (Kitadokoro et al. 2001). However, molecular modelling of full tetraspanin proteins predicts that this hydrophobic patch or "groove" interacts with the hydrophobic β-strand from the EC1 domain (Seigneuret 2006). This modelling proposes that the EC1 and EC2 are packed together to form a structurally conserved extracellular domain. A so-called, "stalk" region is formed by EC1 and EC2 helices A and E and is held together by two salt bridges in the EC2 (Lys 124–Asp195 and Asp 128–His 191) as well as a number of hydrophobic contacts. This stalk region forms the base for the "mushroom-like" head which is composed of the remaining helices. The C and D helix region forms a low polarity patch that is predicted to be the site of specific protein–protein interactions and, consistent with this theory, this is the region of most variability among family members and species homologues (Kitadokoro et al. 2001; Seigneuret et al. 2001). Several studies support the evidence that this low polarity patch is crucial for specific protein – protein interactions. This subloop, or tetraspanin fold, is formed by disulphide bonding between four absolutely conserved cysteines that are present in three variously contained motifs (Cys-Cys-Gly, Pro-X-Ser-Cys [X = any amino acid], and Glu-Gly-Cys). In addition to the disulphide bridges, the head subdomain is also stabilised by Gly158 and Pro176 residues whose unique backbone conformations allow the structural constraints imparted by the disulphides to be accommodated (Kitadokoro et al. 2001). A number of salt bridges in the head domain further act to stabilise the structure (Kitadokoro et al. 2001). The EC2 fold is

unique: it does not have significant homology to any other tertiary structure (Figure 1.3.).

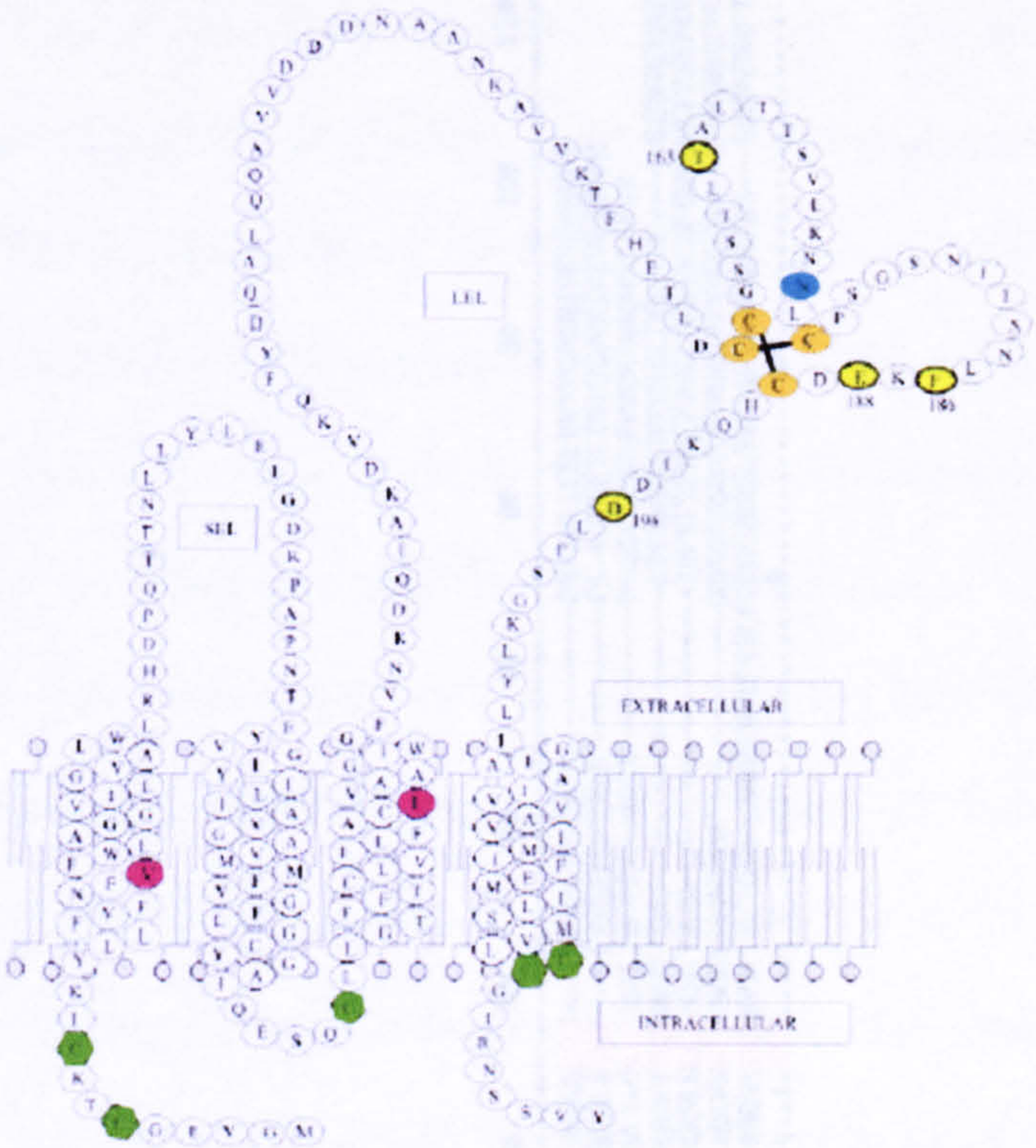


Figure 1.3 The tetraspanin CD81(Figure taken from (Martin et al. 2005). The primary sequence of CD81 (single letter amino acid code) is shown schematically in the context of its transmembrane architecture. Key conserved residues in the tetraspanin superfamily include charged transmembrane residues and canonical cysteines which are disulphide bonded to form a functional subloop in the EC2. Sites of potential modification by palmitoylation and N glycans are shown, as are approximate locations of sites of molecular association.

Figure 1.4. Multiple sequence alignment of tetraspanin EC2 domains. The absolutely conserved motif CCG is aligned and highlighted in red. Residues that are > 50 % conserved are shown in blue.

Due to the varying nature of the EC2 domains of tetraspanins, it is commonly predicted that this large extracellular domain is responsible for the unique properties of individual members of the tetraspanin family. The EC2 domains have been suggested to mediate the primary interactions of tetraspanin proteins (Stipp 2003), and in the few cases where primary binding partners have been identified, this hypothesis has proved to be correct. For example: The interaction between CD9 and CD81 and two members of the EWI Immunoglobulin sub-family(IgSF) utilises the EC2 domain, (Charrin et al. 2001; Stipp et al. 2001) (see 1.3.1.1.1), murine PSG17 interacts with murine CD9 via the EC2 region (Waterhouse et al. 2002) (see 1.3.1.1.2), and the high affinity interaction between CD151 and integrin alpha-3 beta-1 maps within the EC2 region of CD151, (Berditchevski et al. 2001). As with all transmembrane proteins, the study of this family of membrane proteins is limited because the full length proteins cannot be solubly expressed and used to study tetraspanin functions *in vitro*. The fact that the EC2 domains are likely to be essential for function coupled with the need for new tools to investigate tetraspanin function gave rise to the production of recombinant tetraspanin EC2 proteins. The first recombinant EC2 fusion proteins were produced to study the epitope mapping of various anti-CD53 mAbs (Tomlinson et al. 1993; Tomlinson et al. 1995). Recombinant tetraspanin EC2 domains, made here in Sheffield and elsewhere, have since been used to study tetraspanin functions (Higginbottom et al. 2000; Zhu, G. Z. et al. 2002; Higginbottom et al. 2003). Critical residues involved in the interaction of HCV glycoprotein E2 with CD81 were determined by constructing mutant CD81-EC2 domains. Human CD81 EC2 binds to the HCV glycoprotein E2 but the EC2 from African Green Monkey CD81 does not, although the EC2 regions from these

two species differ by only 4 amino acids (highlighted in Figure 1.3). Each of the 4 residues was sequentially mutated in human CD81 to match that of AGM and the mutants were expressed as GST-EC2 fusion proteins. Binding studies revealed that Phe186 in human EC2 was the critical residue for binding to HCV E2, since when this residue alone was mutated to match the AGM homologue, binding was abolished (Higginbottom et al. 2000) (see 1.4 for more detail). In another example, the recombinant fusion proteins were used in sperm-egg binding and fusion assays. Previously, antibodies against CD9 were shown to inhibit oocyte fusion and recombinant-CD9 EC2 was demonstrated to have the same effect. In our laboratory, the critical cysteine residues that are predicted to form the disulphides were mutated to alanines (Figure 1.5) and it was shown that these constructs were not recognised by the conformation-dependent anti-CD9 antibody, ALMA-1. The mutant EC2 domains, lacking at least part of their native structure, no longer possessed the inhibitory property in the sperm-egg fusion assays, thus suggesting a structural requirement for CD9 in this role (Higginbottom et al. 2003).

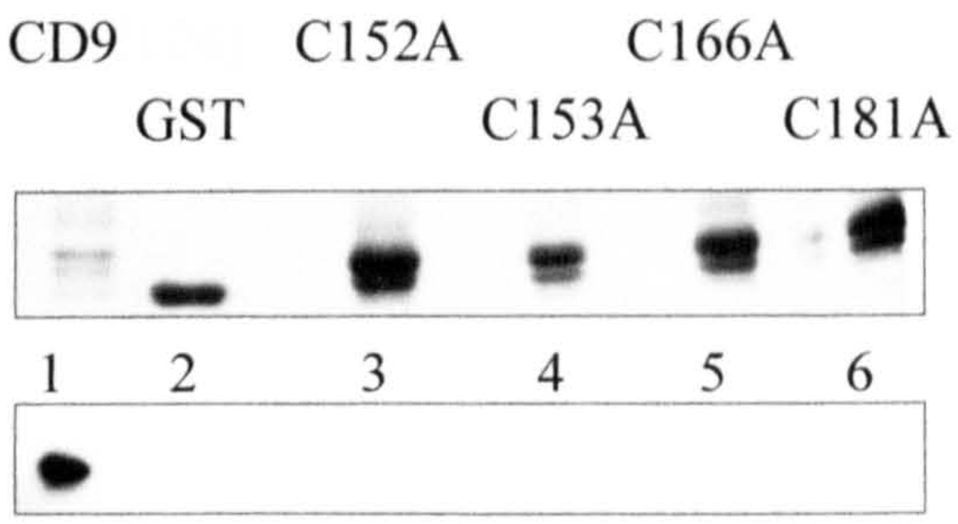


Figure 1.5. Coomassie stained gel (top) and Western blot showing WT GST-CD9 EC2 and the four Cys-Ala mutants (Higginbottom et al. 2003). The Western blot shows that CD9 EC2 (lane 1) is recognised by the structurally sensitive antibody, ALMA-1 but the CD9 cysteine mutants are not.

1.3 TETRASPANIN ENRICHED MICRODOMAINS

It is predicted that the predominant function of tetraspanin proteins is to act as facilitators and organisers of large, multimolecular complexes on the cell surface, know as tetraspanin enriched microdomains (TEMs) or tetraspanin webs (Boucheix et al. 2001; Charrin et al. 2001; Charrin et al. 2003; Stipp et al. 2003). To participate in these microenvironments on the cell surface, tetraspanins exploit their ability to form homo and heterodimers and multimers, which, coupled with their ability to bind a diverse range of other molecules, helps to form the complex networks of organised proteins on the cell surface that are involved in complex cellular events such as cell fusion (Figure 1.6). Analogies have been drawn between TEMs and lipid rafts. In the latter, protein-protein interactions are facilitated by association with membrane regions enriched for cholesterol and glycosphingolipids (Simons et al. 1997). In either lipid rafts or TEMs the associated proteins are believed to form large integrated signalling platforms. Although TEMs and lipid rafts exist as separate entities, they have been shown to interact physically and functionally (Cherukuri et al. 2004; Delaguillaumie et al. 2004).

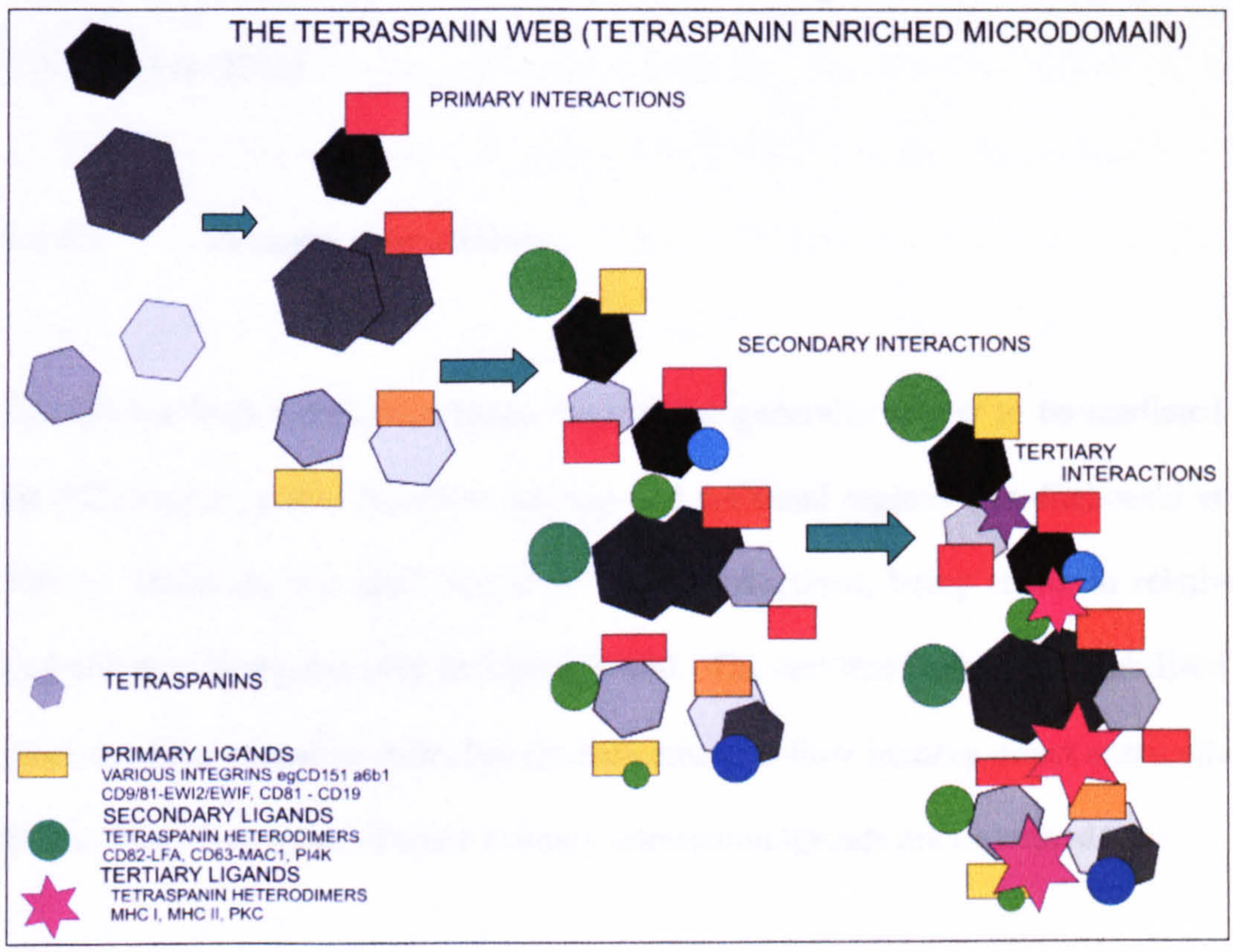


Figure 1.6. Predicted mechanism for the formation of tetraspanin enriched microdomains. Tetraspanin molecules (hexagons) bind to their individual primary ligands (rectangles). Tetraspanin homo- and hetero-dimerisation and binding of shared, secondary ligands causes clustering of complexes. Tertiary interactions form (stars) by the ligation of neighbouring complexes to create a multimolecular complex, a tetraspanin enriched microdomain.

1.3.1 Hierarchy within tetraspanin enriched microdomains

Recent studies assessing the strength/proximity of tetraspanin interactions have indicated that TEMs are constructed from a hierarchy of interactions (Figure 1.6) (Charrin et al. 2003)

1.3.1.1 Primary interactions

Primary (or high order) tetraspanin interactions generally appear to be mediated by the EC2 region, particularly the subloop and proximal regions (Berditchevski et al. 2001). These are the most robust of TEM interactions, being stable in relatively hydrophobic detergents such as Triton X-100. The fact that they can be stabilised by short covalent chemical cross linkers indicates that they involve direct extracellular interactions. Examples of some primary tetraspanin ligands are discussed here.

1.3.1.1.1 EWI proteins

In 2001, three independent groups identified two novel primary binding partners for CD9 and CD81 (Charrin et al. 2001; Clark et al. 2001; Stipp et al. 2001). The first major molecular partner for CD9 was identified by using digitonin solubilisation to reveal high affinity complexes in tetraspanin webs in carcinoma cells (Serru et al. 1999). This 135 kDal protein, initially named CD9P-1 and now more commonly called EWI-F, is a member of the IgSF and is the human homologue of the rat prostaglandin Fα2 regulatory protein (FPRP). EWI-F was isolated by affinity purification using anti-CD9 antibodies and, using CD9/CD82 chimeric molecules, the

binding region on CD9 was defined as the EC2 and fourth TM domain. The same study also revealed that CD81 could form similar complexes with CD9P-1 and other tetraspanins were identified in complexes with CD9P-1 under milder conditions, highlighting this protein's importance in TEMs (Charrin et al. 2001). A second protein, prostaglandin regulatory-like protein (PGRL) was identified around the same time as being a major molecular partner for CD81 in T-cells. Again, this 75 kDal protein was identified as a member of the IgSF with most similarity to FPRP, suggesting the existence of a new protein sub-family (Clark et al. 2001). A third study, again searching for CD9 and CD81 molecular partners, also identified a 70 kDal protein that was later realised to be PGRL. In this study the protein was named EWI-2 and was shown to be widely expressed and 91% identical to the murine homologue (Stipp et al. 2001). As in the case of EWI-F, the other groups also pinpointed the tetraspanin EC2 domains as the region that associates with the primary ligand. Further database searches identified two additional members of this novel sub-family of the IgSF, CD101 and IgSF3. Since all four members of this subfamily contain a Glu-Trp-Ile (EWI) motif in their second Ig-like domain, the nomenclature EWI-2 and EWI-F was given to PGRL and CD9P-1/hFPRP respectively, CD101 was renamed EWI101 and IgSF3 became EWI-3 (Stipp et al. 2001). EWI-2 (or IgSF8) is the smallest protein in the family, being composed of four Ig-like domains. EWI-F has six (Fig. 1.7), EWI101 and EWI-3 are much larger with 14 and 16 Ig-like domains, respectively. Neither EWI101 nor EWI-3 bind to CD9 and CD81 (Stipp et al. 2001).

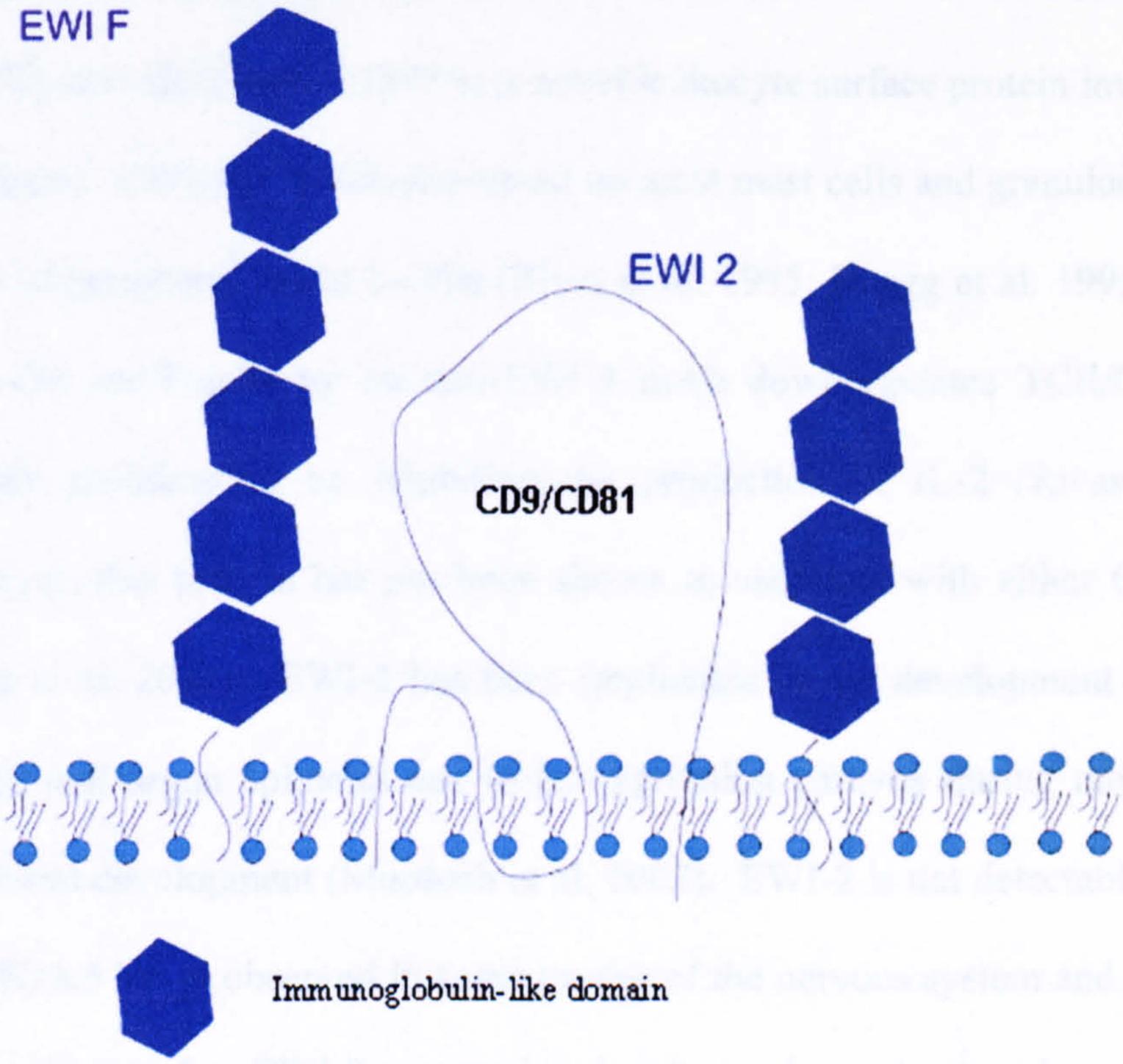


Figure 1.7. EWI-2 and EWI-F. A schematic diagram showing EWI-2 and EWI-F surrounding their primary ligand, CD9 (or CD81). The blue hexagonal shapes represent Ig-like domains, of which there are 4 and 6 in EWI-2 and EWI-F, respectively (not to scale).

There is no described function for human EWI-2, EWI-F or EWI-3. EWI-101 (V7, CD101) was identified in 1995 as a novel leukocyte surface protein involved in T cell activation. EWI-101 is also expressed on most mast cells and granulocytes as well as 30 % of peripheral blood T-cells (Rivas et al. 1995; Ruegg et al. 1995). Ligation of EWI-101 on T-cells by an anti-EWI-3 mAb downregulates TCR/CD3-dependent cellular proliferation by inhibiting the production of IL-2 (Rivas et al. 1995). However, this protein has not been shown to associate with either CD9 nor CD81 (Stipp et al. 2001). EWI-2 has been implicated in the development of the nervous system and organ epithelia due to its expression patterns during mouse embryonic and foetal development (Murdoch et al. 2003). EWI-2 is not detectable at embryonic day (E) 8.5 but is observed in some tissues of the nervous system and organ epithelia E9.5 and E10.5. EWI-2 expression levels in these tissues change continuously throughout embryonic development (Murdoch et al. 2003). Consistent with the described association between EWI-2 and CD9 and CD81, CD9 is present in the developing nervous system (Tole et al. 1993) and CD81 is expressed in pre-implantation mouse embryos (Andria et al. 1992) and in neural progenitor cells from human foetuses (Klassen et al. 2001). In the absence of signalling factor *Shh,* EWI-2 is not detected in the neural tube, suggesting that EWI-2 expression in the nervous system is regulated by *Shh* (Murdoch et al. 2003).

Lymphoid cells were screened for expression of EWI-2 and EWI-F (Charrin et al. 2003). EWI-F was not detected on peripheral blood cells, however, EWI-2 was expressed on these cells in a pattern that is identical to that of CD81 i.e. present on B-cells, T-cells and natural killer cells but not detected on monocytes, polynuclear cells or platelets (Charrin et al. 2003). Interestingly, this CD81 ligand is also expressed in

the other main type of cells susceptible to HCV, hepatocytes, and is co-expressed with CD81 in these cells. It is therefore possible that EWI-2 has some pathological significance in HCV infection.

This same study used chimeric CD9/CD82 molecules to define the region of CD9 involved in the binding to EWI-2 (Charrin et al. 2003). As expected, a 40 amino acid region within CD9 EC2 was involved in the binding but a second region was also required for optimum binding. This second region included a small part of TM2, the cytoplasmic loop, TM3 and a small part of EC1 (Charrin et al. 2003). This is the first example of a tetraspanin ligand interacting with 2 separate regions of a tetraspanin protein. Over-expression of EWI-2 in A431 cells impaired $\alpha 3\beta 1$ integrin-dependent cell adhesion on laminin-5. Cell re-aggregation and motility functions were also impaired but not cell adhesion or spreading. Interestingly, EWI-2 over-expression did not have any effect on cells plated on collagen 1, an $\alpha 2\beta 3$ integrin ligand (Stipp et al. 2003). The increased expression of EWI-2 caused CD81 to localise into filipodia at the periphery of the cell, an ideal localisation to affect cell motility. Also, the presence of EWI-2 caused an increase in association between CD81 and $\alpha 3\beta 1$. It is proposed that one prerequisite for cell motility is that CD81 and/or CD9 link EWI-2 to $\alpha 3\beta 1$ integrin in a TEM (Stipp et al. 2003).

EWI-2 has been implicated in cell motility on laminin-5 in a second report that also describes an association between EWI-2 and CD82 although it is still somewhat unclear as to whether this interaction takes place independently of CD81 or whether CD81 mediates the interaction between CD82 and EWI-2. (Zhang et al. 2003). This cell surface protein, initially named KASP (KAI1/CD82 associated protein), was

identified in Jurkat cells because of its ability to bind to CD82 and was shown by mass spectrometry to be to EWI-2.

1.3.1.1.2 Pregnancy Specific Glycoprotein 17: A CD9 primary binding partner?

Pregnancy-specific glycoproteins (PSGs) are a sub-group of the carcinoembryonic antigen (CEA) gene family that belongs to the Immunoglobulin gene superfamily (IgSF). PSGs are secreted proteins whereas the majority of other proteins in the CEA subgroup are cell surface bound. There are eleven human PSGs and they are composed of an Ig variable-like (IgV) domain at the N-terminus and a varying number of constant like domains. PSGs are produced by the placenta and the concentrations of PSGs in the blood of pregnant women increase during the course of pregnancy but are not detected in males or non-pregnant females (Lin et al. 1974). Low concentrations of PSGs can result in complications in pregnancy. It is predicted that the PSGs play a vital role in controlling the maternal immune system to prevent rejection of the foetus. Foetal rejection has been associated with increased production of inflammatory cytokines, such as interferon γ (IFN-γ) and tumor necrosis factor-α (TNF-α). However, secretion of the anti-inflammatory cytokines interleukin (IL) – 10 and transforming growth factor - β (TGF-β) have been associated with successful pregnancies. In 2001, Dveksler's group reported that human PSGs, PSG1, PSG6 and PSG11, induced the secretion of IL-10, IL-6 and TGF-β by human monocytes *in vitro* (Snyder et al. 2001). The N-terminal domain from PSG6 was enough to allow the induction, suggesting that it is this domain that binds to PSG receptors.

The first report describing the involvement of CD9 and PSGs came about when a cDNA expression library from a murine macrophage cell line, RAW 264.7, was screened by panning against a mouse PSG, PSG17, in an attempt to identify a cellular receptor for this glycoprotein (Waterhouse et al. 2002). Sequencing of positive clones revealed a perfect match to murine CD9. To confirm binding, HEK 293 cells transfected with CD9 were shown, by ELISA, to bind to PSG17-Myc-His6 in a dose dependent manner (Waterhouse et al. 2002). Flow cytometry studies demonstrated the binding of PSG17 to CD9 transfected HEK 293 cells where 66-87% of CD9 transfected cells showed PSG17 binding with less than 1% binding to cells transfected with the empty vector (Waterhouse et al. 2002). Furthermore, it was shown that murine PSG17 could not bind to other murine tetraspanins tested including CD53, CD63, CD81, CD82 and CD151. Pre-incubation of RAW 264.7 cells with anti-CD9 antibody resulted in a reduction of binding to PSG17, and in addition, binding of PSG17 to macrophages extracted from CD9 deficient mice was reduced to almost background levels (Waterhouse et al. 2002). This was the first report of a cellular receptor for a PSG and it was predicted that the variable EC2 region of CD9 would be responsible for the interaction.

1.3.1.1.3 Fibronectin: A CD9 ligand?

There have been some reports describing the binding of CD9 to fibronectin, a component of the ECM, suggesting that rather than acting through integrins, tetraspanins might have a direct role in binding to the ECM and controlling cell motility (Cook et al. 1999); (Longhurst et al. 2002); (Cook et al. 2002). It is well

established that tetraspanins, including CD9, associate with β1, β2, β3 and β7 integrins in TEMs (Hadjiargyrou et al. 1996; Rubinstein et al. 1996; Berditchevski et al. 1999). Fibronectin is a widely expressed molecule that mediates a range of cellular interactions with the ECM such as cell adhesion, cell growth and motility (reviewed in (Pankov et al. 2004). The 250 kDal protein often exists as a dimer and contains multiple protein binding sites including the integrin binding site Arg-Gly-Asp (RGD). Fibronectin is a well established ligand for up to a dozen members of the integrin subfamily (Dedhar et al. 1990; Watt et al. 1994; Yokosaki et al. 1994; Chi-Rosso et al. 1997). The consensus is that fibronectin associates with integrin receptors on the plasma membrane during cellular processes such as adhesion and motility. However, recent reports by Jennings and co-workers describe the direct association between CD9 and fibronectin (Cook et al. 1999; Cook et al. 2002; Longhurst et al. 2002). Immobilised platelet CD9 and full length recombinant CD9 were shown to bind fibronectin and CHO cells transfected with CD9 showed decreased adhesion to fibronectin coated plates and an altered spreading morphology when compared to mock transfected CHO cells (Cook et al. 1999). When CHO cells were transfected with truncated CD9 mutants that lack the EC2 domain, cell behaviour was reversed back to that of mock-transfected cells (Cook et al. 1999). Short peptides derived from CD9 EC2 (especially "peptide 6", amino acids 168-192) were reported to bind to fibronectin in ELISA and to disrupt the binding of fibronectin to immobilised CD9 in competition ELISAs (Longhurst et al. 2002). These data suggested that fibronectin is a potential ligand for CD9 EC2.

1.3.1.2 Secondary interactions

Secondary interactions involve the transmembrane and/or intracellular regions (Berditchevski et al. 2001; Charrin et al. 2003). These interactions are less stable in high stringency detergents but are maintained in milder conditions (Brij-96, Brij-97) and involve tetraspanin-tetraspanin associations that show some dependency on palmitoylation (Charrin et al. 2002). The interactions are thought to indirectly link primary complexes and thereby facilitate their cross-talk.

1.3.1.3 Tertiary interactions

Tertiary interactions are stable only in low stringency detergents (CHAPS, Brij-99). While primary and secondary interactions are isolated as discrete entities, tertiary interactions are large complexes containing multiple tetraspanins and associated molecules. Thus, tertiary interactions are hypothesised to result from the large scale ligation of higher order complexes. The study of tertiary interactions is complicated by their poor stability, meaning that the composition of the complexes is less well defined than for higher order associations. The use of low stringency detergents to isolate these complexes suggests the possibility of non-specific capture of molecules in insoluble micelles. However, several tertiary interactions appear to be functionally relevant (Wright et al. 2004). Importantly, Claas et al., (Claas et al. 2001) have demonstrated that tertiary interactions can exist independently of lipid rafts/large detergent insoluble fractions. The localisation of tetraspanin complexes, intracellularly and in the plane of the membrane, is known to change in response to

certain stimuli, as does the molecular composition of isolated complexes, suggesting that TEMs are not static moieties (Penas et al. 2000; Israels et al. 2001).

1.3.2 Palmitoylation of tetraspanin proteins

There are a number of post-translational modifications that proteins undergo in order to enhance their stability, binding properties, enzymatic activity and localisation. Acylation is a common covalent modification of plasma membrane proteins and involves the covalent attachment of fatty acids to eukaryotic proteins. Protein acylation includes prenylation, myristoylation and palmitoylation, the latter of which is a shared feature amongst the tetraspanin proteins that have been researched in detail, namely CD9, CD63 CD81, CD82, and CD151 (Berditchevski et al. 2002; Charrin et al. 2002; Yang et al. 2002; Charrin et al. 2003; Linder et al. 2003; Stipp et al. 2003; Clark et al. 2004; Kovalenko et al. 2004). Diverse proteins can undergo palmitoylation and the purpose of this modification depends on the requirement of the protein. As palmitoylation increases the hydrophobicity of proteins, it contributes to their membrane association and since palmitoylation is a reversible process, this suggests a regulated modification to allow control of protein function, much like protein phosphorylation (Linder et al. 2003).

Palmitoylation of CD9, CD151 and CD81 has been extensively studied. One example of the functional role of palmitoylation is CD81 and B cell signalling. Upon B cell activation, the B cell receptor (BCR) associates with the BCR complex (CD19/CD21/CD81) to enhance the signalling through the BCR and lower the threshold for B cell activation. It is thought that the BCR complex participates in this

process by prolonging the residency of the BCR in lipid rafts (Cherukuri et al. 2004). Cherukuri et al., also showed that CD81 was palmitoylated upon colligation of the BCR and the CD19/CD21/CD81 complex. The palmitoylation was rapid and reversible and necessary for the prolonged residency of the BCR complex in lipid rafts (Cherukuri et al. 2004).

Juxtamembrane cysteine residues are prime candidates for palmitoylation. CD151 has 8 such cysteines, which, when mutated to alanines, reduce the association of CD151 with its primary partner, α3 integrin (Berditchevski et al. 2002). This mutant form of CD151 also loses the ability to associate with endogenously expressed CD63 and CD81. In addition, CD151 plays a crucial role in linking α3β1 integrin to other tetraspanins but the palmitoylation negative mutant fails to permit CD63 and CD81 association with the integrin (Berditchevski et al. 2002). Although there is significant difference in the binding properties of palmitoylated CD9, CD82 and CD151, the partitioning of tetraspanins into TEM appears not to depend on the presence of palmitate (Berditchevski et al. 2002; Zhou et al. 2004); (Charrin et al. 2002).

1.4 TETRASPANINS AND MALIGNANCY

Tetraspanins have frequently been identified in searches for differentially expressed genes in cancerous cells. Some tetraspanins are thought to work as metastasis suppressors whereas, in contrast, other tetraspanin proteins are over-expressed in certain cancers. For instance, high levels of CD9 in melanoma cells reduce their metastatic potential (Ikeyama et al. 1993) and CD82 is a metastasis suppressor of

numerous cancers including: lung cancer (Adachi et al. 1996), breast cancer (Huang et al. 1998) and prostate cancer (Bienstock et al. 2001). However, overexpression of tetraspanin CO-029 is seen in hepatocellular carcinoma (Kanetaka et al. 2001). Here, some of the possible mechanisms of action of tetraspanins in malignant cells are discussed.

1.4.1 Effects on motility

Tetraspanins are often found on the cell surface to be associated with integrins (Berditchevski et al. 1996; Hadjiargyrou et al. 1996; Hemler et al. 1996; Berditchevski et al. 1997; Berditchevski 2001; Zhang et al. 2001; Sterk et al. 2002). Integrins are cell surface receptors for extracellular matrix proteins (ECM) and transmit information both into and out of the cell. They are heterodimeric molecules consisting of one alpha and one beta subunit; there are many different types of alpha and beta subunit and the composition of the integrin determines its specificity. During integrin-dependent T cell migration, integrins adhere to fibronectin and laminin-5 (two components of the ECM), the cell develops filipodia and lamilipodia and the cell is "pulled" along the ECM tracks by integrins. Tetraspanins, by some unknown mechanism, can increase or decrease the amount of cell motility. This has been demonstrated by using anti-tetraspanin antibodies to inhibit T-cell motility or over-expression of tetraspanins to increase motility (Berditchevski 2001).

1.4.2 Tetraspanins and adhesion

KAI-1/CD82 and MRP-1/CD9 have been implicated in the impairment of adhesion and therefore the invasive and metastatic properties of tumor cells in ovarian cancer (Houle et al. 2002). This work also describes the altered tissue expression of these tetraspanins throughout the progression of the disease. In normal surface epithelium, CD82 and CD9 are expressed at very low levels or are absent altogether. In low grade tumours, or when the surface epithelium is modified, there is much higher expression that is completely lost in malignant tumours. This suggests that a loss of tetraspanins is consistent with a loss of adhesive properties during malignant transformation. However, in prostate cancer there are conflicting reports that the expression of CD82 is increased, not decreased, during the progression of the disease (Lijovic et al. 2002) and CO-029 is overexpressed in hepatocellular carcinoma (Kanetaka et al. 2001). One theory for the discrepancies in expression of tetraspanins is that it is their direct or indirect effect on integrin function that modifies motility and therefore metastatic potential. (Rubinstein et al. 1996). It may be that they interfere with integrins to prevent integrin function when over-expressed in a malignant tumor (tumor suppressor) or, when they are under-expressed in a malignant tumor, there are insufficient amounts to hold the integrins in place correctly and so adhesion is inadequate and metastasis takes place. This would fit in with the view that cancer cells have lost their tetraspanin regulatory function.

1.4.3 GM3 ganglioside

There has been evidence to suggest that CD9 and its involvement in the formation of protruding lamellipodia and filopodia during haptotactic cell motility is controlled or modulated by GM3 ganglioside (Ono et al. 2001). Three lines of evidence were tested to support this theory: When GM3 was added to colonic carcinoma cell lines expressing high amounts of CD9, motility was inhibited. Cell lines that express low levels of CD9 needed large amounts of CD9 gene transfection before becoming susceptible to GM3-dependent motility inhibition. Finally, GM-3 does not have any affect on motility except when it co-exists with CD9. Although dose-dependent inhibition by GM3 could not be demonstrated, they did show that GM3 preferentially binds CD9 and not integrins. Further work in this area showed that it was α3-dependent mobility on laminin 5 that was being inhibited. It is thought that the interaction between α3 and CD9 is activated by GM3. The stable complex thus formed blocks signalling to laminin-5 rendering the cells immobile (Kawakami et al. 2002).

1.5 TETRASPANIN INVOLVEMENT IN SPERM EGG FUSION

Gamete fusion is the first event that takes place in the making of a new life but surprisingly this process remains poorly understood. It is a strong possibility that one particular member of the tetraspanin superfamily, CD9, plays a vital role in the fusion of gamete membranes and so revealing its function could help our understanding of such a remarkable process.

1.5.1 Tetraspanin – Integrin Association in Oocytes

As discussed previously, it is a known fact that tetraspanins associate with integrins, especially the β1 integrin sub-family heterodimers such as α3β1, α4β1, α6β1 (Banerjee et al. 1997; Domanico et al. 1997; Berditchevski et al. 2001). This information has prompted individuals to look into the role that tetraspanins play in the specific binding of integrins and their known molecular partners. Integrins are known to be involved in sperm-egg binding, cell fusion and the passage of the fertilised egg from the oviduct to the uterus (Bowen et al. 2000). α6β1 integrin expressed on the surface of murine eggs was believed to bind a sperm surface glycoprotein, Fertilin β also known as ADAM 2 (A disintegrin and a metalloprotease domain), Fertilin α (ADAM 1) and Cyritestin or ADAM 3. Members of this family associate with integrins via their disintegrin domains. It is believed that these proteins play an important role in mammalian fertilization. However, the role of integrin α6β1 came into question when approximately 75% of eggs collected from α6 null mice fertilised normally, indicating that the role of α6β1 integrins in fertilisation is not as important as initially thought (Miller et al. 2000). In addition, a range of anti-integrin antibodies, pre-incubated with murine eggs before addition of sperm, were shown to only partially inhibit the binding of sperm (< 55 % inhibition) and had no effect on the fusion process (Sengoku et al. 2004).

An involvement of CD9 in sperm-egg fusion came about when an anti-CD9 mAb, JF9, was used to inhibit sperm – egg binding and fusion in a dose dependent manner. JF9 inhibited 99% of binding and 100 % of fusion (Chen et al. 1999). Anti-CD81 mAbs had very little affect on these processes (Takahashi et al. 2001).

At this early stage, it was generally believed that the anti-CD9 mAb either sterically hindered the binding of fertilin β to α6β1 integrin or somehow signalled for α6β1 to suppress binding. Similar results were obtained when another anti-CD9 mAb (KMC 8.8) was used to inhibit the binding of fertilin α to its binding partner on the egg surface (Wong et al. 2001). This same monoclonal antibody was used by Zhu and Evans who found inhibition of fertilin β binding to eggs only occurred when fertilin β was immobilised on small beads and not when it is in a recombinant soluble form (Zhu, X. et al. 2002). This could be due to other multi-valent interactions taking place with CD9 when fertilin is in the beaded form where there are multiple molecules on the beads. This data supports the evidence of a multimolecular tetraspanin web that organises the surface distribution and expression of membrane proteins (Zhu, G. Z. et al. 2002).

1.5.2 Gene Knock-Out Studies Reveal CD9 Is Critical For Sperm-Egg binding

Three independent groups performed CD9 gene knockout studies in mice that reinforced these findings (Kaji et al. 2000);(Miyado et al. 2000); (Le Naour et al. 2000). Heterozygotes for CD9 were apparently normal whereas female homozygotes could ovulate normally and sperm-egg binding took place but the eggs were unable to be fertilised because there was no sperm-egg fusion. In fact, fertilisation in these mice was less than 2 % of WT fertilisation levels but micro-injection of sperm into these eggs allowed normal development to resume, thus limiting the role of CD9 to the fusion event (Miyado et al. 2000). In WT eggs undergoing fertilisation, CD9 was shown to accumulate around the acrosomal region of the sperm head but in the

absence of fertilisation, CD9 was seen in patches all over the surface of the egg. Closer inspection revealed that, in fact, the CD9 was isolated at the tips of the microvilli that cover the egg. A possible theory is that CD9 is involved in the initial fusion of the two membranes, yet the actual ligand for CD9 on the sperm membrane remains unidentified (Kaji et al. 2000). The mice lacking CD9 were apparently normal in every other way apart from being severely infertile, suggesting that in tissues that also express high amounts of CD9, other tetraspanins are able to take over. This idea of redundancy in tetraspanins was highlighted in the work done in Drosophila (Fradkin et al. 2002). When 25% of the Drosophila tetraspanins were knocked out there were delays in embryonic synapse formation, but the organism still managed to have an apparently normal adult lifecycle (Fradkin et al. 2002).

The infertility seen in CD9 $^{-/-}$mice was reversed by CD9 poly (A)-RNA injection into the egg cytoplasm. It was also found that CD9 $^{+/-}$ mice have decreased expression of CD9 on the egg surface and hence a lower fertilisation rate compared to CD9 $^{+/+}$ mice and that injection of CD9 mRNA increased fertilisation rates (Kaji et al. 2002). Human (h) CD9 and murine (m) CD81 poly(A) mRNA was also injected into the cytoplasm of oocytes from CD9 deficient mice, restoring fertilisation rates to 78 and 75% respectively. This suggests that both hCD9 and mCD81 can compensate for lack of mCD9 in mouse egg fertilisation, (Kaji et al. 2002). Later studies confirmed an involvement of CD81 in mouse fertilisation when it was shown that CD81 $^{-/-}$ had reduced fertility and CD9 and CD81 double knock-out mice were completely infertile (Rubinstein et al. 2005). These data suggest complementary roles for CD9 and CD81, probably in a TEM.

1.5.3 The Mechanism of Action of CD9 in sperm/egg fusion

To investigate the binding mechanism of CD9, two question were asked: Does CD9 bind directly to a molecule on the sperm membrane, or does CD9 have more of a regulatory role, interacting with other molecules on the membrane of the egg thereby enabling these molecules to be the direct binding partners of the ADAM proteins? (Zhu, G. Z. et al. 2002). GST-EC2 CD9 fusion proteins were recognised by antibody KMC8 (see previously) suggesting that CD9 EC2 is responsible for the events leading up to sperm-egg fusion. Fertilisation was inhibited when GST-EC2 CD9 was incubated with eggs prior to insemination but when incubated with sperm, no significant inhibition was observed. This suggests that sperm do not have a receptor for CD9 and that CD9 probably interacts with another protein on the egg membrane. Mutant CD9 EC2 constructs were injected into null CD9 mice eggs to determine the specific residues required for fertilization to occur. Ser-Phe-Gln (173 – 175) sequence in the EC2 was pinpointed as the functional site (Zhu, G. Z. et al. 2002). Human GST-CD9 EC2 and GST-CD81 EC2 can also inhibit fertilisation of CD9 positive mouse oocytes but not the EC2 from CD63 (Higginbottom et al. 2003). In this work, the binding of sperm to the oocyte was not significantly inhibited but the fusion of the sperm with the oocyte was significantly inhibited, emphasizing the fact that CD9 is affecting the fusion event. Furthermore, a structural requirement of CD9 EC2 for the observed inhibition of sperm-egg binding and fusion was identified (Higginbottom et al. 2003). CD9-EC2 mutants lacking one or more cysteine residues that are predicted to form the disulphide bridges in the EC2 required for proper folding, failed to inhibit sperm-egg binding and fusion events *in vitro* when pre-

incubated with CD9 positive oocytes before addition of sperm (Higginbottom et al. 2003).

1.6 THE INVOLVEMENT OF TETRASPANINS IN VIRAL INFECTIONS

There are emerging roles for tetraspanins in the processes of viral infections. Interestingly, the functions of tetraspanins implicated in viral infection are diverse, as discussed here.

1.6.1 *Flaviviruses*

The most exhaustively documented involvement of a tetraspanin in viral infection is the role of CD81 in Hepititis C Virus (HCV) infection. CD81 was identified as the ligand for the HCV glycoprotein E2 using a cDNA expression library created from a sub-clone of human T-cell lymphoma MOLT-4, A2R, which showed high E2 binding capacity. Monoclonal antibodies recognising epitopes in CD81-EC2 can competitively inhibit the binding of recombinant E2 to EBV-transformed B cells (Pileri et al. 1998). Recombinant tetraspanin EC2 proteins were used to characterise the interaction between the EC2 of CD81 and HCV. African Green Monkey (AGM) CD81 differs from human CD81 by only 4 residues yet AGM-CD81 is unable to bind to E2 and AGM not susceptible to HCV. Phe186 in human CD81 EC2 is Leu in AGM CD81; the mutation of Phe186 to Leu in human CD81 completely abolishes E2 binding (Higginbottom et al. 2000). Small molecule inhibitors, designed to mimic this region of CD81, have been made and shown to inhibit the interaction with HCV providing useful ideas for vaccination development (VanCompernolle et al. 2003).

Clearly, CD81 is a cellular receptor for HCV but recent data suggests that CD81 alone is not sufficient for HCV entry and the relevance of the CD81-HCV E2 association for viral entry is still in question. There are a number of other candidate receptors for HCV in addition to CD81, namely; scavenger receptor class B type 1 (SR-B1) (Scarselli et al. 2002), dendritic cell-specific intracellular adhesion molecule 3 grabbing nonintegrin (DC-SIGN) (Lozach et al. 2003; Pohlmann et al. 2003) and the low density lipoprotein receptor (LDLR) (Wunschmann et al. 2000). One study has established a need for both CD81 and SR-B1 for binding and infectivity of HCV in hepatocytes, thus proposing a co- receptor role for CD81 (Bartosch et al. 2003). However, in another study the binding of HCV to three hepatocyte-derived cell lines was not blocked by anti-CD81 (Sasaki et al. 2003). It has also been observed that different HCV strains utilise CD81 molecule to different extents (Roccasecca et al. 2003; Kronenberger et al. 2004); (McKeating et al. 2004). Another study using a more diverse range of HCV glycoproteins revealed that mutation Asp196Glu in CD81 EC2 reduced the binding efficiency of HCV pseudotypes expressing CH35, HJC4 and C6a1 glycoproteins by 70-96% (McKeating et al. 2004). This data would suggest that all HCV pseudotypes require CD81 for infection of HepG2 cells but that the interactions between CD81 EC2 and HCV glycoproteins differ between sub-species. Perhaps the primary role of CD81 is as an essential component of the tetraspanin enriched microdomain (TEM) involved in HCV infection rather then as an actual receptor.

1.6.2 *Retroviruses*

To date, tetraspanins have been implicated in three distinct retroviral infections and interestingly, a different member of the tetraspanin family is involved in each case.

1.6.2.1 HTLV-1

Cell-cell fusion via syncytium formation is an important factor in viral spread and is thought to be one of the main transmission mechanisms of Human T cell Leukaemia Virus (HTLV-1) (Clapham et al. 1983). In an attempt to identify cell surface molecules involved in HTLV-1 induced syncytium formation, a panel of mAbs were tested: M38 (IgG1), M101 (IgG1), M104 (IgG1) and C33 (IgG2a) specifically inhibited T-cell syncytium formation (Fukudome et al. 1992). Of these, 3 (C33 and M104 and M101) were found to precipitate CD82.

When CD82 was cotransfected with HTLV-1 glycoproteins in COS-1 cells, a dose-dependent inhibition of cell fusion was noted compared to cells transfected with HLTV-1 glycoproteins alone (Pique et al. 2000). It was also confirmed that the presence of CD82 did not affect production, maturation or incorporation of HTLV-1 envelope glycoproteins into virions and it has been proposed that CD82 exerts its effects by a direct steric hindrance of the envelope-receptor interaction, thus inhibiting cell-cell fusion and syncytium formation (Pique et al. 2000).

1.6.2.2 *HIV-1*

By the use of immunogold electron microscopy, alterations in CD63 expression were described in Human Immunodeficiency Virus type-1 (HIV-1) infected cells compared to non-infected cells. Molecules expressed on the cell surface of H9 cells and blood mononuclear cells (MNC) were compared before and after infection with HIV-1. CD63, usually predominantly expressed in intracellular compartments, is upregulated on the surface of HIV-1 infected cells as well as being selectively incorporated into budding structures and newly synthesised virus particles (Meerloo et al. 1992; Meerloo et al. 1993). These findings were later confirmed when the virus membrane was examined for the presence of cellular proteins, CD63 was amongst the cellular proteins acquired by HIV (Orentas et al. 1993). In another study, vesicle fractions from non-infected H9 and PBMC cells were shown to contain CD5, HLA-DR, HLA-DQ, CD30 and CD44 with only traces of CD63 and CD43. However, in vesicles from infected cells both CD63 and CD43 were present and HLA-DQ was excluded (Gluschankof et al. 1997).

Macrophages act as virus reservoirs during HIV-1 infection. Previous work has described unidentified organelles within macrophages that harbour the virus particles. Expression analysis revealed high levels of CD63 and MHC II and low levels of Lamp 1, hence these organelles resemble the compartments in macrophages where MHC II molecules undergo the final stages of maturation. These major histocompatibility type II compartments (MIIC) are thought to be preferential sites for virus accumulation in macrophages. CD63 is also enriched on multivesicular bodies (MVB) and it is presumed that HIV-1 particles mimic the mechanisms utilised

to make natural internal vesicles (Raposo et al. 2002). Further evidence for the involvement of CD63 in HIV-1 infection of macrophages has been provided recently by the O'Brien group who found that anti-CD63 mAbs inhibited HIV-1 infection in macrophages (von Lindern et al. 2003). In these studies, only HIV-1 strains utilising the CCR5 co-receptor were inhibited and this phenomenon was only witnessed in macrophages and not T-cells (von Lindern et al. 2003).

1.6.2.3 *FIV*

Tetraspanins have been linked to a third retroviral infection by the discovery that an anti-CD9 mAb, vpg15, can inhibit Feline Immunodeficiency Virus (FIV) infection in a feline lymphoma cell line (Willett et al. 1994). By pre-incubating the cells with vpg15 or control antibodies followed by addition of FIV, it was observed that anti-CD9 antibody delayed reverse transcriptase production by approximately 10 days compared to control antibodies, without completely blocking reverse transcriptase production. When the antibody was removed this inhibition was lifted (de Parseval et al. 1997). The ability of FIV to infect CD9 $^{-/-}$ cells (Hohdatsu et al. 1996) implies that CD9 is not essential for FIV infection, or at least, that other tetraspanins can take over this role. Re-introduction of CD9 into these cells increased infectivity (Willett et al. 1997), suggesting that the presence of CD9 facilitates the viral life cycle. It has been proposed that CD9 is functioning in the intracellular transportation of viral proteins to the cell membrane but further work is required to define its role in more detail.

1.6.2.4 *Paramyxovirus*

Another CD9 mAb, K41, has been discovered to inhibit the infection Canine Distemper virus (CDV), a morbillivirus of the paramyxoviridae family. A plaque reduction assay was carried out to determine whether K41, or a panel of commercial CD9 antibodies, could inhibit uptake of the virus and/or syncytium formation. Both the size and number of plaques were reduced in Vero and HeLa cells pre-incubated with anti-CD9 (Loffler et al. 1997). No direct binding of CD9 to CDV glycoproteins has been detected, suggesting that CD9 is not the cellular receptor for CDV. Additional investigations have revealed similarities of CD9 function in FIV and CDV. CDV can still bind to the surface of cells in the presence of K41 suggesting that CD9 is not involved in the attachment of CDV to cells prior to infection (Schmid et al. 2000). Treatment with the antibody also had no effect on the production of viral DNA but it was demonstrated that CD9 had an important role in CDV–induced cell-cell fusion but not virus-cell fusion. There are a number of examples in the literature of tetraspanins being involved in cell fusion events, perhaps the most extensively studied example being the involvement of CD9 in mammalian oocyte fusion (see Chapter 1.5). Also, our unpublished data suggests that soluble EC2 domains of both CD9 and CD63 can inhibit Con A-induced fusion of human blood monocytes (see Appendix).

1.9 AIMS

It is clear from this review of the literature that there is much work to be done in elucidating the exact molecular mechanisms of tetraspanin function. The main tools currently utilised to study tetraspanins are monoclonal antibodies, gene knock-out mice and overexpression of tetraspanins in cell lines by transfection of DNA. Each of these methods has their uses, but also a number of possible weaknesses: The use of monoclonal antibodies can be misleading due to Fc receptor binding and steric hindrance by the bulky immunoglobulin chains; gene knock-out mice are costly and time-consuming to generate and ineffective when functional overlap occurs between tetraspanins, whereas transgene overexpression could lead to amplification or activation of cellular events that would not normally occur in WT cells. Several groups have previously successfully used recombinant EC2 proteins in order to study tetraspanin function (Flint et al. 1999; Higginbottom et al. 2000; Higginbottom et al. 2003; Takeda et al. 2003; Zhang et al. 2004). Few primary ligands for tetraspanins have so far been identified and the interactions between the tetraspanins and primary partner proteins that have been identified, have not been fully characterised. Information of this kind would give valuable insights to the functioning of tetraspanins at the molecular level. The major aims, therefore, were to clone the two recently identified primary ligands for CD9 and CD81, EWI-F and EWI-2 (Charrin et al. 2001; Clark et al. 2001; Stipp et al. 2001), express them as soluble proteins, and using recombinant EC2 proteins, characterise critical amino acids in the EC2 domains from CD9 and CD81 that may be involved in this interaction. It was also planned to characterise the domains in the EWI proteins that engage with CD9 and CD81 by cloning individual Ig-like domains as well as full length proteins. Other

primary tetraspanin ligands to be analysed were PSG17 and fibronectin – these two proteins were identified as binding partners for CD9 (Longhurst et al. 2002; Waterhouse et al. 2002) and it was planned to study the interactions of both proteins with CD9 EC2 using the GST-EC2 constructs.

A secondary aim was to optimise the large scale production of the existing EC2 constructs and to promote these recombinant proteins as alternative, cheap, quick and relatively easily produced tools to study tetraspanin function by collaborating with other tetraspanin researchers. It was also planned to increase the repertoire of the recombinant proteins by cloning EC2 domains from two further tetraspanins, CD82 and CD231 (Chapter 3).

Another part of the work was to attempt to solve the structure of a more complex EC2 domain than CD81, by X-ray crystallography (Chapter 4). CD63 fitted into this category and was also chosen because of its role in HIV infection.

CHAPTER 2

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 General materials

General laboratory reagents were from Sigma, BDH or Melford, and were AnalaR / equivalent grade.

2.1.2 Solutions and buffers

Preparation of commonly used buffers is summarised in Table 2.a. Preparation of media for the growth of mammalian and bacterial cells is shown in 2.b, and the antibiotic stock solution preparations are shown in 2.c.

a.i

Buffer	Preparation
BSS (Balanced Salt Solution)	43.36g NaCl, 1.83g KCl, 8.19g D-sorbitol, 3g $K_2PO_4.3H_2O$, 0.7g KH_2PO_4, 12.09g Hepes. pH 7.2 – 7.4, make up to 5L in H_2O.
Destaining Solution	70ml glacial acetic acid, 830ml H_2O.
ELISA Binding Buffer	1.59g Na_2CO_3, 2.93g $NaHCO_3$. pH 9.6 make up to 1L in H_2O.
Glutathione Elution Buffer	0.154g reduced glutathione dissolved in 50ml of 50 mM Tris HCl, pH8.0
Hank's Balanced Salt Solution (HBSS)	Purchased from Invitrogen (cat. No. 14170-088), contains/L: 400mg KCl, 600mg KH_2PO_4, 350mg $NaHCO_3$, 8 g NaCl, 48mg Na_2HPO_4, 1 g D-Glucose, 10mg Phenol Red.
1 x PBS (Phosphate Buffered Saline)	10 x stock/2L – 160 g NaCl, 4 g KCl, 28.8 g Na_2HPO_4, 4.6 g KH_2PO_4, pH 7.2 – 7.4.
1 x PBS Tween	As above + 0.01% Tween 20 to the 1 x solution.
TAE	50 x stock/L – 242 g Tris-base, 57.1ml glacial acetic acid, 100ml 0.5 M EDTA, pH 8.0

a.ii

Buffer	Preparation
SDS-PAGE Running Buffer	10 x stock/L – 30g Tris-base, 144g glycine, 100ml 10 % SDS.
SDS-PAGE Loading Buffer (non-reduced)	2ml glycerol, 2ml 10 % SDS, 0.25mg bromophenol blue, 2.5ml stacking gel 4 X buffer (see below). Add H_2O to a final volume of 100ml.
SDS –PAGE Loading Buffer (reduced)	As above with 0.5ml β - mercaptoethanol.
Separating Buffer	18.17g Tris-base, 4ml 10 % SDS. Adjust to pH 8.8 using 10M HCl and make up to 100ml with H_2O.
Stacking Buffer	6.06g Tris-base, 4ml 10 % SDS. Adjust to pH 6.8 using 10M HCl and make up to 100ml with H_2O.
Staining Solution	250ml isopropanol, 100ml glacial acetic acid, 650ml H_2O, 2.5 g Coomassie brilliant blue.

a.iii

Buffer	Preparation	
Buffers for Native Protein Purification on –NTA-Ni^{2+}	Wash Buffer 1L	6.90g $NaH_2PO_4.H_2O$, 17.54g NaCl, 1.36g Imidazole. Adjust to pH 8.0 using NaOH.
	Elution Buffer 1L	6.90g $NaH_2PO_4.H_2O$, 17.54g NaCl, 17.00g Imidazole. Adjust to pH 8.0 using NaOH
Buffers for Denatured Protein Purification on –NTA-Ni^{2+}	Lysis Buffer A 1 L	100mM (13.8g) $NaH_2PO_4.H_2O$, 10mM (1.2g) Tris-base, 6 M (573g) guanidine hydrochloride. Adjust to pH 8.0 using NaOH.
	Lysis Buffer B 1 L	100mM (13.8g) $NaH_2PO_4.H_2O$, 10mM (1.2g) Tris base, 8 M (480.5g) Urea. Adjust to pH 8.0 using NaOH.
	Wash Buffer C 1 L	100mM (13.8g) $NaH_2PO_4.H_2O$, 10mM (1.2 g) Tris-base, 8M (573g) Urea. Adjust to pH 6.3 using HCl.
	Elution Buffer D 1 L	100mM (13.8g) $NaH_2PO_4.H_2O$, 10mM (1.2g) Tris-base, 8 M (573g) Urea. Adjust to pH 5.9 using HCl.
	Elution Buffer E 1 L	100mM (13.8g) $NaH_2PO_4.H_2O$, 10mM (1.2g) Tris-base, 8 M (573g) Urea. Adjust to pH4.5.

b.

Media	Preparation
LB	1 L (in water): 10g tryptone, 5 g yeast extract, 10 g NaCl, adjust to pH 7.5 with 1M NaOH and autoclave
LB-Agar	1 L (in water) : 10g tryptone, 5 g yeast extract, 10 g NaCl, 15 g agar, adjust to pH 7.5 with 1M NaOH and autoclave
SOC	1 L (in water) : 16 g tryptone, 10 g yeast extract, 5 g NaCl, 5ml (2M) MgSO4, 5ml MgCl2, 10ml (2M) glucose.

Table 2a and b. (a.i) Commonly used buffers. (a.i.i) Buffers used in SDS-PAGE, (a.i.i.i) Buffers used for purification of His6-tagged proteins on nickel columns. (b) Media used for bacterial cell culture

c.

Antibiotic	Stock	Working Concentration
Carbenicillin	50mg/ml in 50% ethanol	50 µg /ml
Kanamycin	50mg/ml in H_2O	50 µg /ml

Table 2.c. Antibiotics

2.1.3 Sterilisation

Bacterial media, glassware, pipette tips and centrifuge tubes were sterilised by autoclaving (120 ° C, 20 minutes, 15 or 20lb/sq.in.). Antibiotics used in bacterial and mammalian culture were sterilised before use by passing through 0.4μm filters.

2.1.4 Molecular biology and cell culture reagents

Antibodies used throughout the project are listed in Table 2.d and other molecular biology reagents/kits are listed in 2.e

2.d

Name	Target antigen	Reported species specificity	Isotype	Label	Source
ALMA-1	CD9	Human	Mouse IgG1	None	F. Lanza, Strasbourg, France.
H5C6	CD63	Human	Mouse IgG	None	D. Azorsa, NIH.
Tspan1/2	CD82	Human	Mouse IgG	None	E.Rubinstein, France
BL2	CD82	Human	Mouse IgG1	None	Serotec, MCA1311
Anti-CD231	CD231	Human	Mouse IgG1	None	Oncogene Research, OP175100UG
14A2.H1	CD151	Human	Mouse IgG1	None	L.Ashman, Adelaide, Au.
1.3.3.2.2	CD81	Human	Mouse IgG1	None	Ancell
EAT-2	CD81	Human	Mouse IgG1	None	S.Levy, Stanford, US
RGS-His6 antibody	RGS-His6	Human	Mouse IgG1	None	Qiagen, 34610
Anti-GST	GST	N/A	Goat IgG	none	Amersham Biosciences, 27-4577-01
Anti-GST-HRP	GST	N/A	Goat IgG	HRP	Abcam, Ab6649
Anti-c-myc	Myc	Human	Rabbit, polyclonal	none	Abcam, Ab1529
Goat Anti-Mouse IgG	Mouse IgG	Mouse	Goat	HRP	Bio-Rad, 172-1011
Rabbit Anti-Goat	Goat IgG	Goat	Rabbit	HRP	Dako, PO449
Donkey Anti-Rabbit	Rabbit IgG	Rabbit	Donkey	none	Abcam, Ab6701

Table 2.d. Antibodies

2.e

Name	Catalogue number	Source
BugBuster Protein Extraction Reagent	70584	Novagen
Complete Mini EDTA free PCI	1836170	Roche
Coomassie Protein Assay Reagent	1856209	Pierce
ECL Western Blotting Reagent	RPN2106	Amersham Biosciences
Full Range Rainbow Molecular weight markers	RPN800	Amersham Biosciences
L-Glutathione Reduced	40750	Biochemika
Glutathione Sepharose 4B beads	27-4574-01	Amersham Biosciences
Hyperfilm ECL	RPN3103K	Amersham Biosciences
Multigrade Developer	1757855	Ilford
Ni-NTA Superflow	1018611	Qiagen
Nitrocellulose, Hybond N+	RPN303B	Amersham Biosciences
Nucleospin Plasmid Kit	635988	Macherey-Nagel
Polypropylene Columns	34964	Qiagen
Qiaex II Gel Extraction Kit	20021	Qiagen
Siliconized 1.5ml tubes	T4816-250EA	Sigma
Silver Staining Plus Kit	161-046(1-4), 161-0448	BioRad
Snakeskin Pleated Dialysis Tubing, 3,500 MWCO	68035	Pierce
Unifix, liquid fixer	521 1412	Kodak

Table 2e. Molecular Biology Reagents

2.1.5 Oligonucleotide production

Oligonucleotides were purchased from Sigma Aldrich or MWG.

2.1.6 Automated DNA sequencing

DNA sequencing was either performed at the in-house sequencing facility using an ABI 377 or ABI 310 Sequencing machine or sent to Lark Technology Inc., UK.

2.1.7 Instrumentation

Centrifugation of mammalian cells used a Harrier 15/80, MSE. A bench top Sigma 113 or a refrigerated Sigma 1K15 was used to centrifuge small samples. A Sigma 3K15 was used to pellet bacterial cells after lysis and larger bacterial cultures were centrifuged in a MSE Falcon 6/300.

Electroporation used BioRad Genepulser apparatus. For bacterial transformation, the apparatus was connected to a pulse controller.

For PCR, the DNA engine, Dyad Peltier Thermal Cycler was used.

Bacterial fermentation was carried out in a New Brunswick Scientific 20L Fermenter, BF4500.

Sterile mammalian cell culture manipulations were performed in a class II microbiological safety cabinet, the Bio 2+, Envair.

Spectrophotometery (for DNA and protein concentration and measuring OD of bacterial cultures) used a Unicam UV/visible dual beam spectrophotometer. OD measurements for 96 well plates (ELISAs, Bradford assays) used an EL800 Microplate Reader, Bio-Tek Instruments Inc. Visualisation of DNA in agarose gels and protein in SDS gels used a MultiImage Light Cabinet and the AlphaImager Software, Alpha Innotech Corporation.

BioRad apparatus was used for running protein and DNA gels.

A Soniprep 150 sonicator, MSE, was used to lyse bacterial cells.

2.2 MOLECULAR BIOLOGY - DNA WORK

2.2.1 Primer design

Nucleotide sequences were identified on the NCBI website and primers were designed to amplify the region of interest. Primers were generally 15-20bps long with a Tm around 55°C (roughly determined by counting G and C base pairs as 4°C and A and T as 2°C).

2.2.2 General PCR using Taq polymerase

Unless otherwise stated, PCR using Taq Polymerase (including Reddy Mix and High Fidelity Reddy Mix), were performed in the following manner:

Forward Primer	0.5µl
Reverse Primer	0.5µl
DNA template	1–5µl
Reddy Mix	6.0µl
dH_2O	to a total volume 12µl

Generally, the following conditions were used for PCR cycles:

Step	Time	Temperature	Cycles
Initial Denaturation	2 minutes	94°C	1 x
Denaturation	1 minute	94°C	25 x
Annealing	1 minute	50-60°C	
Extension	1 minute	72°C	
Final extension	10 minutes	68-72°C	1 x

2.2.3 TOPO-TA cloning

Specific primers were designed to amplify the DNA region of interest from a full length DNA template purchased from IMAGE or Kazusa.

For TOPO-TA cloning, 5' primers had a TA overhang to allow insertion into TOPO TA vector. High Fidelity Reddy Mix (Abgene) was used for the PCR. 0.5–4µl of the fresh PCR product was mixed with the TOPO vector and incubated for 5 minutes at RT. The plasmid was now ready for transformation into TOP10 Chemically Competent Cells (Invitrogen).

2.2.3.1 Transformation

All transformation procedures were carried out on ice. 10µl of competent cells were transferred to a pre-chilled 1.5ml tube containing 10-30µg DNA from the TOPO Cloning reaction and mixed gently by tapping. This was left on ice for 30 minutes before heat shocking the cells at 42°C for 30 seconds. The cells were immediately placed back on ice and left for 5 minutes. 200-1000µl of SOC medium was added and the cells were left to grow for 1 hour at 37°C, shaking at 220RPM. 20–200µl of cells were plated on agar plates containing the appropriate antibiotic (kanamycin or carbenicillin) and in the presence of X-Gal.

2.2.3.2 Colony screening

A number of positive (white) colonies and a few negative controls (blue) colonies were picked and resuspended in the following PCR mix:

9.9µl 1:1 Reddy mix

1µl Forward primer 10nM

1µl Reverse primer 10nM

Primers specific to the insert or internal vector primers were used.

The general PCR cycle for screening colonies was:

Step	Time	Temperature	Cycles
Initial Denaturation	2 minutes	94°C	1 x
Denaturation	1 minute	94°C	25 x
Annealing	1 minute	55°C	
Extension	1 minute	72°C	
Final extension	10 minutes	72°C	1 x

The amplified products were mixed 5:1 with DNA loading buffer and run on 1 % agarose gels containing 0.001 % ethidium bromide made up in TAE buffer, generally for 40 minutes at 90V. The gels were visualised on the U.V. light box and the presence of a correctly sized band confirmed that the transformation had worked and the insert was incorporated in the vector.

2.2.3.3 Small scale plasmid DNA preparation

A metal loop was flamed, cooled on agar, and used to pick single colonies containing the desired insert which were then suspended in 5ml LB containing the appropriate antibiotic. These were left to grow overnight at 37°C, RPM 220. The following morning the dense cultures were centrifuged at 4000 g for 5 minutes to pellet the bacterial cells and the supernatant was discarded. Nucleospin Plasmid Miniprep Kit (Macherey – Nagel) was used for all minipreps. The cells were resuspended in 250 µl of physiological buffer A1 before being lysed by addition of an equal volume of Lysis Buffer, A2. The solution was mixed gently and left to incubate at room temperature for 5 minutes. 300µl of neutralising buffer A3 was added to the lysate

before centrifugation at 10,000g to pellet the SDS precipitate and cell debris. The supernatant was loaded into a QuickPure column on which the plasmid DNA binds to the column's silica membrane. The DNA was washed by centrifuging wash buffer through the column and then discarding the flow-through. A buffer to remove endonuclease activity is used and then the DNA is washed twice with an 80 % ethanol wash to clean the DNA and then dried to remove residual ethanol which might inhibit enzymatic reactions. Finally, the DNA is eluted 25µl of 1mM Tris, pH8.0.

2.2.3.4 Quantification of DNA concentration

DNA has an maximal absorption at 260nm and an OD reading of 1 at this wavelength corresponds to 50µg of DNA/ml. DNA was diluted 500 fold in water, the absorbance at 260nm was recorded and the concentration of DNA was calculated.

2.2.4 DNA electrophoresis

Generally, 1 % agarose gels with 0.001 % ethidium bromide made up in TAE buffer were used. Gels ran in TAE buffer for 40 minutes at 90 V. Conditions were altered slightly for very small or very large DNA fragments.

2.2.5 DNA sequencing

The M13 Forward and Reverse primers from were used for sequencing products in Invitrogen vectors unless otherwise stated. The majority of sequencing was carried

out using BigDye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosciences), and the sequencing reaction performed by the Division of Genomic Medicine's Sequencing Facility, F Floor, The Medical school. DNA inserts larger than 800bp were sent to Lark Technology, Inc. for sequencing.

For In-house sequencing the following sequencing reaction was set up:

10–40ng plasmid DNA

3.2 pmol primer

2.0µl BigDye 3.1

4.0µl buffer

H_2O to a total of 20µl

And the following sequencing program performed:-

Step	Time	Temperature	Cycles
Initial Denaturation	1 minutes	96°C	1 x
Denaturation	10 seconds	96°C	25 x
Annealing	5 seconds	50–60°C	
Extension	4 minutes	60°C	
Cooling	10 minutes	4°C	1 x

To precipitate the DNA 50µl of 100 % ethanol and 3.0µl of CH3COONa pH 5.2 was added to each tube and this was either incubated on ice for 30 minutes, or frozen at -20°C overnight. The DNA was pelleted by centrifugation at 15 000g, 4°C for 20 minutes and then washed with 100µl, 80 % ethanol before a final spin for 10 minutes. The ethanol was carefully pipetted off and then the pellet was air dried for 30 minutes to 1 hour. The DNA was then ready to take to the DNA Sequencing Facility to be read.

2.2.6 Restriction enzyme digestion of DNA

Restriction enzymes were from Roche or Promega and were stored at –20°C in glycerol to prevent freezing, thereby retaining enzymatic activity. High glycerol concentrations are also known to modify endonuclease activity so care was taken to keep the volume of diluted enzyme one tenth or below the total reaction volume. Standard digestions used 2 units of enzyme for every µg of DNA to ensure complete digestion and digestions were generally performed at 37° C for 1-24 hours using the appropriate buffer. Following restriction digest the DNA was run on agarose gels to check the size of the insert.

2.2.7 DNA purification from agarose gels

The kit used for batch purification of DNA products from agarose gels was the Qiaex II Suspension (Qiagen) and the experiments were carried out according to the manufacture's instructions.

2.2.8 DNA ligation

Ligation of DNA into vectors used Ligase enzyme (Roche). A 1:1 ratio of DNA insert to vector is required and amounts were estimated qualitatively by looking at agarose gels of restriction digestion DNA. Various ratios of vector to insert were prepared as well as a re-ligation control with no insert. The ligation reaction was typically carried out for 2 hours at 14°C before the DNA was transformed into DH5α

competent cells (Invitrogen) or Omnimax competent cells (Invitrogen) for propagation of the plasmid. Colonies were PCR screened and positive colonies were mini-prepped and sequenced as previously described. Correctly sequenced constructs were then ready to be used for protein expression.

2.3 MOLECULAR BIOLOGY – PROTEIN WORK

2.3.1 Protein expression in bacterial systems

2.3.1.1 Expression vectors

pGEX-2KG vector was used for production of tetraspanin EC2 domains with an N-terminal GST tag. Cloning of EWI full length proteins used Gateway™ vectors pET DEST42 and pcDNA-DEST40 (Invitrogen) for bacterial and mammalian expression respectively (see Chapter 5 for specific Gateway™ Methods).

Cloning of individual Ig domains of EWI proteins used the pET32c and pET26b vectors (Novagen).

2.3.1.2 Bacterial strains for protein expression

BL21 Codon Plus Competent cells and Origami cells (Novagen) were used for protein expression.

2.3.1.3 Transformation

See 2.2.3.1

2.3.1.4 Inoculation

A metal loop was flamed, cooled and used to pick single colonies that were resuspended in 10ml of LB containing the appropriate antibiotic. The cultures were grown overnight at 37°C with 220 RPM shaking to produce a dense inoculate.

2.3.1.5 Determining optimal protein expression conditions.

This method gave 54 conditions and 6 controls to determine which set of conditions was best for producing the protein of interest.

Two 10ml overnight cultures were used to inoculate 400ml of autoclaved LB. One flask was incubated at 37°C and one at 25°C, both shaking at 220 RPM. The OD_{600} was monitored and when ODs 0.2, 0.4 and 0.8 were reached 10 x 10ml samples were taken at each OD and from each temperature. Three samples from each set were induced with 1.0mM, 0.5mM and 0.1mM IPTG, the final sample being the non-induced control. The samples were left to produce protein for 2 hours, 4 hours or overnight before centrifugation at 4500g for 10 minutes to pellet the bacteria. The pellets were then frozen at –20°C prior to analysis.

2.3.1.6 Biomass scale – up

Once the optimal conditions to express the protein were determined, a scale-up of protein production was performed.

Each 10ml overnight culture was used to inoculate 400ml of autoclaved LB, with added antibiotics, in 2L conical flasks (this gives the bacterial cells optimal room for acration). The cultures were then grown at the optimal temperature and until the desired OD_{600} was reached before induction of protein production using the pre-determined concentration of IPTG. The cultures were then left to grow for the appropriate time.

2.3.1.7 Harvesting cells

Cultures were transferred to 750ml centrifuge tubes and centrifuged at 4500 g for 20 minutes. The supernatant was removed, the cell pellet washed by resuspending in ice cold 1 x PBS and then the mixture was transferred into 50ml tubes and centrifuged as before. Again, the supernatant was removed and the pellet stored at – 20 ° C.

2.3.1.8 Biomass fermenter

The Biomass fermenter holds 25L of culture and so dense 500ml inoculates were required. This was achieved by setting up the 10ml of overnight culture, as before, and then transferring 1ml into 500ml of LB and leaving this to grow for a further 24 hours.

The inoculate was loaded into the fermenter by use of a peristaltic pump and the correct antibiotics were injected through a port in the fermenter. Samples were collected through a small valve to monitor the OD. When the correct OD was reached, IPTG was injected in the same way as the antibiotics. The cells were harvested by collecting the culture in 1L centrifuge tubes and centrifugation at 4000 g for 20 minutes, followed by washing the pellet in PBS and transferring to a 50ml tube before centrifuging again. The cell pellets were stored as before.

2.3.1.9 Defrosting the cell pellet

Cell pellets were always defrosted on ice with frequent vortexing to help speed up the process.

2.3.2 Isolation of soluble GST-tagged protein by affinity purification

2.3.2.1 Lysing the cell membrane and removal of the insoluble fraction

The cell pellet was re-suspended in 10 – 20ml of BugBuster (Novagen) with added protease inhibitors; leupeptin, aprotinin, and pepstatin (to a final concentration of 10μg/ml) and 1mM PMSF. The cells were lysed by gently rolling for 1 hour at room temperature. The nuclear fraction and cell components were removed by centrifugation at 14 000g for 10 minutes, and the supernatant contains the soluble protein fraction.

2.3.2.2 Storage of columns

Glutathione Sepharose beads were stored at 4°C in 1 x PBS with 20 % ethanol in paraffin sealed columns. After elution of protein from the beads, excess PBS was run through the column to wash the beads before filling the column with 20 % ethanol solution. Stored in this way the beads lasted for 2 or more years and were continually reused for purification of the same protein construct.

2.3.2.3 Preparation of beads

The glutathione Sepharose beads were washed 3 times in an excess of PBS to remove any residual ethanol that can interfere with enzymatic reactions. This was carried out by vortexing fresh beads until they were completely resuspended before adding the desired amount (typically 1.33ml beads per 1.6 L culture volume) to a 25ml tube. Each wash entailed filling the 25ml tube with PBS, centriguging at 500 g for 3 minutes and then carefully pouring off the supernatant.

2.3.2.4 Adherence to beads

The standard conditions for allowing adherence of GST fusion proteins to the glutathione beads was 1 hour at RT or overnight at 4°C.

2.3.2.5 Loading the column

15ml columns (Qiagen) were used routinely for protein purification. The columns were pre-washed with detergent, typically Triton X-100, and the beads were loaded into the columns and allowed to settle.

2.3.2.6 Washing the column

To remove unbound protein and other contaminants from the sample, 3 column volumes of PBS were washed through.

2.3.2.7 Elution of protein

Protein was eluted by addition of 0.5ml Glutathione Elution Buffer fractions. The eluates were collected in 1.5ml tubes and assayed for protein content using the Bradford Assay. Fractions were continued to be collected until the protein content could no longer be detected, typically 8–15 fractions.

2.3.3 Purification of His6-tagged proteins under native conditions

2.3.3.1 Soluble protein release

After thawing on ice the pellet was resuspended in 5ml BugBuster (Novagen) for every 1g of cell pellet and PCI (Roche). This was left to lyse on a rotator at RT for 1 hour before centrifugation (13, 000g, 20 minutes, 4°C) to pellet the cell debris.

2.3.3.2 Periplasmic protein purification

This method was used to purify periplasmic proteins from 10ml culture volumes and was scaled up accordingly for larger preparations.

The cell pellet was defrosted slowly on ice and then resuspended in 0.5ml 30 mM Tris-HCl pH 8.0, 20 % sucrose with 1µl EDTA, pH 8.0 and appropriate protease inhibitors. The cells were mixed for 10 minutes at RT and then centrifuged for 10 minutes at 10,000g at 4°C, and the supernatant was discarded. In order to break the cell wall and release the periplasmic proteins, the cell pellet was thoroughly resuspended in ice cold 5mM $MgSO_4$ for 10 minutes with gentle stirring. Centrifugation to pellet the cells and to isolate the periplasmic fraction in supernatant was performed at 13000g for 20 minutes at RT.

2.3.3.3 Binding to and elution from the Ni-NTA resin

The binding capacity of the Ni-NTA resin is 5–10mg/ml and the volume of beads was adjusted accordingly for each purification. The beads were washed 3 times in wash buffer by centrifuging at 500g for 5 minutes, before being mixed with the cell lysate for 1 hour at RT. After this incubation period the lysate – Ni-NTA mixture was loaded into a column and the flow through collected for SDS-PAGE analysis.

The column was washed with 2 column volumes of wash buffer and the flow-through collected for SDS-PAGE analysis. 0.5ml fractions of elution buffer were used to elute the protein from the beads and this process was continued until Bradford assay of the eluted fractions failed to show presence of protein.

2.3.4 Purification of His6-tagged protein under denaturing conditions

Thawed cell pellets were resuspended in 5ml of Buffer B for every 1g of cell pellet. The lysis reaction was left to take place at RT, with stirring for 1 hour or until the solution became translucent. The lysate was centrifuged at 10 000g for 30 minutes to pellet the cell debris and the cleared lysate (supernatant) was mixed with 1- 2ml of washed Ni-NTA beads (2.2.3). Binding of His6-tagged protein to the Nickel beads took place for 1 hour at RT. The lysate- resin mixture was carefully loaded into a column and the flow-through collected. This was saved for SDS-PAGE analysis. Buffer C was used to wash the column twice, 4ml each time and then the protein was eluted using 0.5ml fractions of Buffer E. The protein content of each fraction was assayed in a Bradford Assay.

2.3.5 Estimation of protein concentration using the Bradford assay

In a 96 well plate, 200µl of Bradford Reagent was mixed with 1µl of the protein sample of interest or the same volume of BSA standards, typically the range of concentration of BSA standards was 10mg/ml – 0.5mg/ml. After thorough mixing, the absorbance of the wells at 595nm wavelength was read on the plate reader. The concentrations of the standards were entered, a standard curve was automatically created and the concentrations of the unknown samples were calculated.

2.3.6 Protein dialysis

Proteins generally needed dialysing to remove the glutathione or imidazole from the elution buffers following purification on glutathione Sepharose or NTA-Ni^{2+} beads respectively. Protein fractions of similar concentrations were pooled and carefully pipetted into a short piece of dialysis tubing sealed at one end with a plastic clip. A second plastic clip was used to seal the open end and the sample was placed in 10L of PBS and left to equilibrate for 4 hours at 4°C with stirring. After 4 hours the PBS was replaced and then repeated twice more to be sure that the concentrations of eluting agent were negligible. SDS-PAGE was used to accurately determine protein concentrations following dialysis.

2.3.7 SDS-PAGE

Unless otherwise stated, 12.5 % acrylamide were used and were prepared according to the following formulations:

Separating gel	**Stacking gel**
2.4ml dH_2O	1.0ml dH_2O
3.0ml 30% acrylamide solution	300µl 30 % acrylamide solution
1.9ml separating buffer	444µl stacking gel
112µl Ammonium Persulphate (APS)	28µl 10 % APS
5.0µl TEMED	5.0µl TEMED

The separating gel was poured into the BioRad plates, held into position by the clamp. Once set, the stacking gel was poured on top and a 10 or 15 lane comb was positioned into the stack. When the second gel was set, the comb was removed and the gels were positioned in the tank and immersed in running buffer.

Samples were boiled in the appropriate loading buffer (generally reduced loading buffer unless a structurally sensitive antibody was being used in a Western blot). SDS-PAGE gels were run at 60V for 30 minutes or until the proteins had passed through the stacking gel, and then 130V for 60 minutes.

2.3.7.1 Coomassie staining

The gels were carefully taken out of the clamps and stained by immersing in an excess of Coomassie Brilliant Blue for one hour. De-staining the gels to remove non-specifically bound Coomassie was performed by discarding the Coomassie, adding an excess of de-staining solution and a piece of sponge to absorb the dye and help to speed up the process. For optimum results de-staining was carried out overnight before analysis the following morning.

2.3.7.2 Silver staining

Silver staining was performed as per the manufacturers' instructions. Firstly the gels were fixed in Fixative Enhancer Solution made by mixing 50 % v/v methanol, 10 % v/v acetic acid, 10 % v/v Fixative Enhancer Concentrate*, and 30 % v/v distilled water. The gels were fixed for 20 minutes or overnight with gentle agitation. The

gels were washed by rinsing twice in excess distilled water for 10 minutes each time with gentle agitation. To stain and develop the gel the following components were mixed in a glass beaker with a Teflon coated magnetic stirrer 5 minutes before use: 35ml deionised water, 5ml Silver Complex Solution*, 5ml Reduction Moderator Solution*, 5ml Image Development Reagent* and immediately before use 50ml of Development Accelerator Solution* at room temperature was added. This was poured over the gels and gently agitated for 10 – 20 minutes until desired staining was reached. The reaction was stopped by placing the gels in 5 % acetic acid for 15 minutes and then rinsing in water.

* solutions from Silver Stain Plus Kit, BioRad 161-0449

2.3.8 Western Blotting

SDS-PAGE gels were arranged in the Western blotting clamp with nitrocellulose membrane on the anode side. Protein transfer conditions were 250mA for one hour on ice, or 30mA for six hours.

2.3.8.1 Probing the membranes with antibodies

Membranes were blocked in PBS-T containing 5 % w/v milk powder for 1 – 2 hours at RT and primed with primary antibody at the manufacturers' suggested concentration diluted in blocking buffer. Incubation was at RT for 1 hour followed by three, five minute washes in PBS-T. The membranes were then incubated with secondary antibody labelled with horse radish peroxide (HRP), used at 1/3000 (the dilution recommended by manufacturer) in blocking buffer. Incubation was at RT for

1 hour. Thorough washing in PBS-T took place to remove unbound and non-specifically bound secondary antibody. Control blots were performed by omitting the primary antibody to determine any non-specific or background binding of secondary antibody.

2.3.8.2 Developing the Western blot

A 1:1 ratio of ECL reagents (Amersham) were mixed and immediately poured over the membrane. After 1 minute the membrane was taken out, drained, wrapped in clear plastic and taped into a film cassette protein side up. A sheet of autoradiography film was immediately placed over the membrane and exposed for 30 seconds to five minutes. The film was immersed in developer solution until bands became visible, washed in excess water and fixed in fixative solution until the film turned transparent. The fixative solution was washed briefly in water and the film left to air dry.

2.3.9 General ELISA method

ELISAs were performed in flat-bottomed Maxisorb 96 well plates (Nunc) unless otherwise stated. All washes were performed by filling the wells with 1 x PBS-T, flicking off and blotting on some paper towel. All incubations were for 1- 2 hours at room temperature.

Generally the wells were coated with 100µl of a 10µg/ml protein solution in binding buffer and incubated at 4°C overnight. Next morning, the excess protein solution was flicked off and the wells were washed before blocking with 200µl of blocking buffer

for 1–2 hours at RT. The wells were washed 3 times and incubated with primary antibody diluted in blocking buffer (see individual experiments for dilutions). Following washing the secondary antibody labelled with HRP at 1/3000 dilution was added and after the incubation period the plate was washed 5 times and was then ready to be developed. For each 96 well plate one TMSO tablet (Sigma) was completely dissolved in 1ml of DMSO before adding 9ml citrate phosphate buffer and 2µl H_2O_2. 100µl of this reagent is added to each well. The reaction is left to take place (typically 1-10 minutes) by shaking at RT. When a substantial level of colour was obtained in the control well the reaction was quenched by addition of 15 µl H_2SO_4. An absorbance reading at 450nm for each well was read on an ELISA plate reader.

2.4 CELL CULTURE METHODS

2.4.1 Cell Passage, harvest and cell count

Unless otherwise stated, all cells were grown in DMEM and 10 % FCS supplemented with 100µg/ml penicillin and 100µg/ml streptomycin in a 37°C incubator with 5 % CO_2.

Adherent cells were harvested by adding 2ml 5mM EDTA in HBSS (without calcium and magnesium) or Cell Dissociation Solution (Sigma) followed by a 3–5 minute incubation before gently pipetting the solution across the Petri dish to remove cells. Culture medium was added to a total volume of 10ml and the cells were isolated by

centrifugation at 400g for 4 minutes. The cell pellet was resuspended in fresh medium and the cells were plated as required.

A haemocytometer was used to count cells. The density (cells/ml) of a cells suspension was calculated by counting the number of cells in the central 5 x 5 squares. 100 cells counted was equal to 10^6 cells/ml.

2.4.2 Freezing of cells

Confluent cells on a 100 mm Petri dish (typically 10 x 10^6 cells) were harvested and re-suspended in 1ml of freezing mix comprising of 10 % v/v DMSO and 90 % FCS. Cell freezing chambers were used to slowly freeze the cells (1°C per minute) before transferring to liquid nitrogen dewars for storage.

2.4.3 Thawing of cells

Cells were defrosted rapidly by immersing in a 37 ° C water bath immediately after being taken from the storage dewar. Once the cell suspension had just thawed it was poured into fresh media, centrifuged to pellet the cells as described earlier, before being resuspended in appropriate media and plated.

2.4.4 Transfection of Plasmid DNA

2.4.4.1 Electroporation of plasmid DNA

CHO cells at 70–90% confluency were harvested in cell dissociated solution. One plate (~1x10^6 cells) was used for each electroporation event and after pelleting, the cells were resuspended in 0.8ml and added to a pre-chilled 1.5ml tube containing 50μg of chilled DNA. The DNA and cells were then transferred to pre-chilled cuvettes and incubated for 10–30 minutes on ice. The cells were then pulsed with 290 V, 960 μF and then transferred back onto ice. After 30 minutes incubation on ice the cells were divided between 5 plates containing the normal growth media. The cells were left to grow for 24 hours before addition of the selective media containing 40-100μg/ml geneticin.

2.4.4.2 Lipofectamine transfection procedure

Cells were plated on 6 well plates (Nunc) to 50–60 % confluency and left to adhere overnight. The DMEM used for the remainder of the protocol is without serum.

Preparing DNA per well:

A 1μg of the DNA of interest was mixed with 100μl of DMEM and 6μl of plus reagent and incubated for 15 minutes at RT.

B Add 4μl of Lipofectamine to 100μl of DMEM, mix and leave for 2 minutes at RT.

Solutions A and B were combined, vortexed and incubated at RT for 15 minutes.

Preparing cells:

The cell media was removed before adding 800µl of DMEM (-FCS) with 200µl of the DNA complex prepared as above. After 4 – 5 hours incubation (37°C, 5 % CO_2) the supernatant was removed and 2.5ml DMEM (+FCS) was added. Cells were grown for 48 hours before being assayed for protein expression. For stable constructs, 40–100 µg/ml of Geneticin was added to the cells after the initial 48hrs and then left for 2 – 3 weeks with half media changes as and when required.

2.4.4.3 Calcium phosphate transfection procedure

This procedure was for cells growing on 100 mm plates and was performed on cells that were 30 – 60 % confluent. Fresh media was added to each plate and left to re-condition for 1-2 hours.

Solution A. In a 1.5ml tube, 10 – 20µg of DNA was mixed with 62µl of $CaCl_2$ and made up to 500µl with dH_2O.

Solution B. 500µl 2 x PBS in a 15ml tube

The 15ml tube containing solution B was vortexed continuously whilst solution B was added in a drop wise manner. The solution was incubated at RT for 30 minutes in which time it tuned slightly opaque due to the formation of calcium phosphate-DNA co-precipitation. After a final vortex the solution was carefully added to the plates whilst swirling gently. Cells were harvested after 48 hours.

2.4.5 Isolation of protein from mammalian cells

Cells were harvested as previously described and resuspended in 50µl of cell lysis buffer containing protease inhibitors. The cell suspension was incubated for 10 minutes on ice, at RT or at 37°C to lyse cells. The cell lysate was centrifuged at 10000g for 10 minutes to pellet the nuclei and the supernatant was transferred to a fresh tube and assayed for protein content by the techniques described above.

2.4.6 Whole cell ELISA method

Confluent cells were harvested and counted as described in 2.3.1. A cell suspension of density 10^6 cells/ml was made up and 100µl plated per well in 96 well plates (method adapted from (Schober et al. 2002). The cells were left to adhere overnight and the following day antibody probing and development was carried out in the same was as described in 2.3.9.

2.4.7 Fluorescent activated cell sorting (FACS)

All centrifugation steps are at 400g for 4 minutes unless otherwise stated.

All washes are carried out as follows: The cells were centrifuged and resuspended in 100µl PBS with gentle pipetting up and down before centrifugation as before.

Sub-confluent cells were dissociated in 50mM EDTA in HBSS and centrifuged at 400g for 4 minutes. Cells were counted and resuspended at 10^6 cells per ml in BSA (Basic salt solution and 0.2 % w/v azide) and then plated at a concentration of 10^5

cells per well of a 96 well plate (100µl / well). Cells were centrifuged and resuspended in primary antibody diluted in HBSS with 1 % w/v BSA to block. Cells were then incubated on ice for 30–60 minutes and then washed twice. A 1/1000 dilution of appropriate secondary antibody, conjugated to FITC and diluted in HBSS with 1 % w/v BSA, was added and incubation was for 1 hour on ice in the dark. The cells were washed twice as before, resuspended in 300 µl HBSS and read immediately on a FACS calibre using the Cell Quest program and FL1 basic settings. Five thousand cells per experiment were counted and results were analysed using WinMDI software.

CHAPTER 3

RECOMBINANT TETRASPANIN EC2 PRODUCTION AND FUNCTIONAL STUDIES

3.1 INTRODUCTION

3.1.1 Tetraspanin EC2 domains

The tetraspanin EC2 domains are arguably the most intriguing and well researched region of this family of transmembrane proteins (see Chapter 1.2). Tetraspanin EC2 recombinant proteins have been used to study the properties and functions of this domain here in Sheffield (Higginbottom et al. 2000; Higginbottom et al. 2003; Garcia-Lopez et al. 2005), and elsewhere (Zhu, G. Z. et al. 2002; Ellerman et al. 2003).

3.1.2 Production of tetraspanin EC2 domains using the pGEX system

In order to assist the production and purification of proteins in *E. coli*, several vectors have been developed that express the desired protein as a fusion with another more easily purified protein. Fusion proteins can also prevent proteins from forming insoluble aggregates and also can enhance the stability of the protein (for review see (Cabrita et al. 2004). In Sheffield, tetraspanin EC2 proteins are expressed with N-terminal GST tags using the pGEX vector system (Amersham). The pGEX expression system was developed in 1988 (Smith et al. 1988). The same group had

previously expressed an enzymatically active, 26kDal, GST (Sj26) from *S. japonicum* in *E. coli* and other groups had demonstrated that mammalian GST isozymes can be purified by affinity chromatography on immobilised glutathione, followed by competitive elution with excess free glutathione (Simons et al. 1977; Simons et al. 1981). This property was found to also be true for Sj26 (Smith et al. 1988) which suggested to the designers that a vector with a Sj26 fusion protein would be extremely beneficial for the rapid, single-step purification of foreign polypeptides linked to Sj26. In addition, the GST tag improves the solubility of the fusion protein so that non-denaturing purification procedures can be used. The first vector of this type, pGEX1, was successfully used to express stable eukaryotic-GST fusion proteins (Smith et al. 1988). A more advanced pGEX vector, pGEX2T, was then produced encoding the cleavage recognition site for thrombin immediately 5' to the multiple cloning site to allow for removal of the GST fusion protein if desired (Smith et al. 1988). Further engineering of the pGEX vector was carried out to create pGEX-KG, which included a glycine-rich linker region containing the sequence PGISGGGGG located immediately following the thrombin cleavage site (Fig. 3.3). This glycine linker acts to increase the efficiency of thrombin cleavage by allowing greater accessibility of thrombin to its recognition sequence (Guan et al. 1991).

The desirable properties of optional GST removal, ease of purification and increased solubility of fusion protein when tagged with GST, were all factors that lead to the decision to use this vector system to express the tetraspanin EC2 constructs.

3.1.3 AIMS

The previous success of using recombinant EC2 proteins and the need for additional tools to study these complicated transmembrane proteins were motives for us to: A) optimise the current methods of production of our existing EC2 proteins; B) to attempt to clone and express the EC2 domains from CD82 and CD231; C) to use these proteins to study the interactions with primary tetraspanin ligands; D) to promote these proteins as novel and useful tools in the study of tetraspanins by collaborating with fellow tetraspanin researchers.

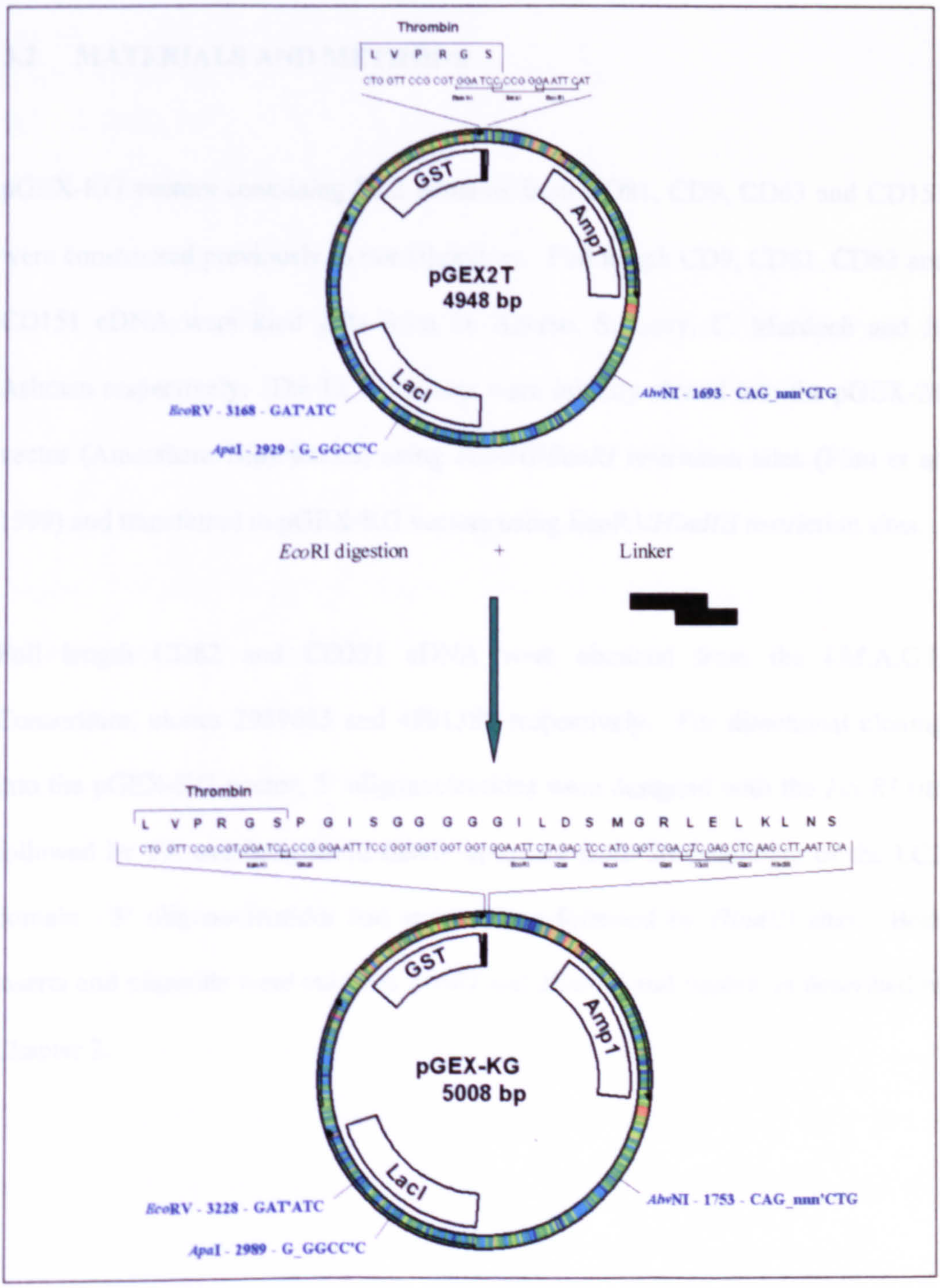


Figure 3.1. pGEX-2T vector conversion. The commercially available vector, pGEX-2T, was adapted by inserting an extended linker region and an additional thrombin protease site (Guan et al. 1991). The *Hind*III and *Eco*RI restriction sites, used to clone the tetraspanin EC2 domains, are shown.

3.2 MATERIALS AND METHODS

pGEX-KG vectors containing EC2 domains from CD81, CD9, CD63 and CD151 were constructed previously in our laboratory. Full length CD9, CD81, CD63 and CD151 cDNA were kind gifts from D. Azorso, S. Levy, C. Murdoch and L. Ashman respectively. The EC2 domains were initially cloned into the pGEX-2T vector (Amersham Biosciences) using *BamHI/EcoRI* restriction sites (Flint et al. 1999) and transferred to pGEX-KG vectors using EcoR1/*HindIII* restriction sites.

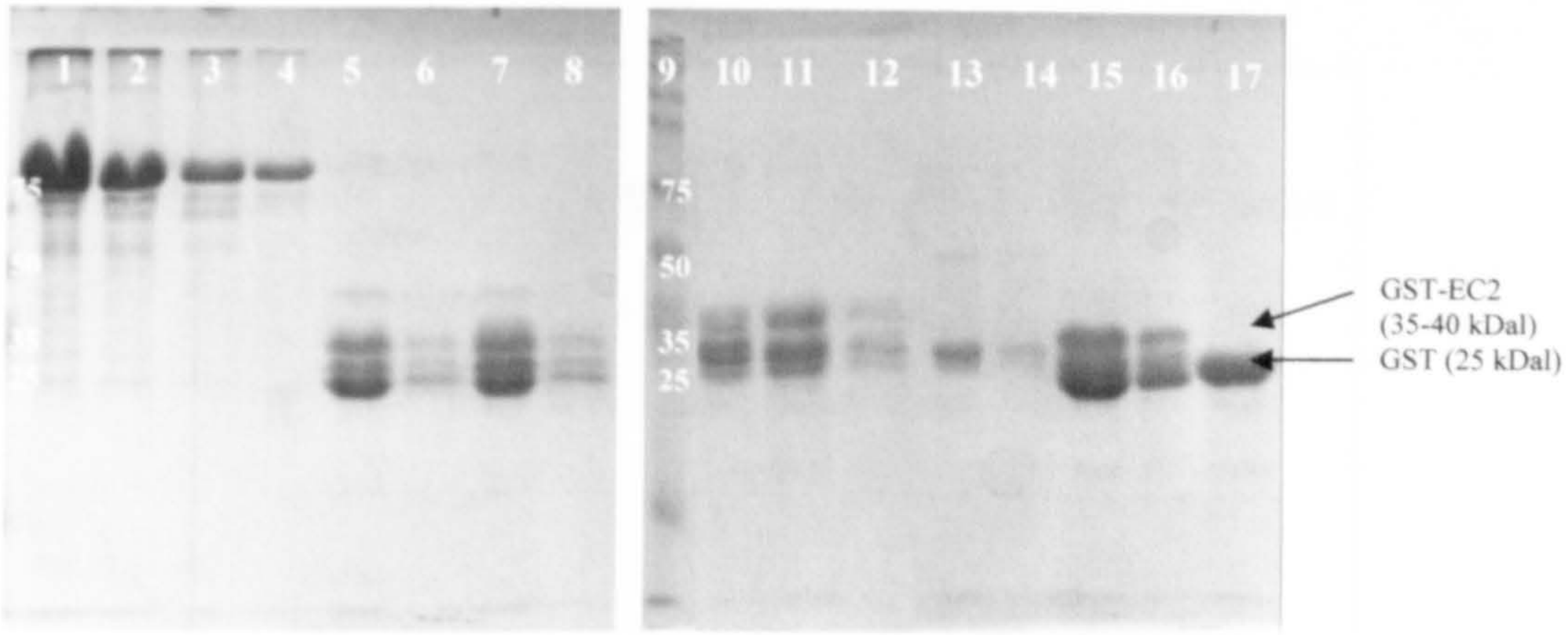
Full length CD82 and CD231 cDNA were obtained from the I.M.A.G.E Consortium, clones 2959683 and 4801388 respectively. For directional cloning into the pGEX-KG vector, 5' oligonucleotides were designed with the *EcoRI* site followed by TA overhang immediately upstream from the beginning of the EC2 domain. 3' oligonucleotides had stop codons followed by *HindIII* sites. Both inserts and plasmids were cut with *EcoRI* and *HindIII* and ligated as described in Chapter 2.

3.3 RESULTS

3.3.1 Production of recombinant EC2 domains

The production of the CD9, CD81, CD63 and CD151 tetraspanin EC2 proteins was continued for use in the current work and in collaborations (Chapter 7). 0.5ml aliquots of GST-EC2 proteins were eluted from the glutathione-Sepharose affinity columns and each aliquot was assayed for its protein content using Bradford reagent. Typically, the EC2 proteins were eluted in a total volume of 6ml. At this stage the samples were in elution buffer, 10mM reduced glutathione and 50mM Tris, pH8.0, and so required dialysis into PBS. Aliquots of similar protein concentration were pooled. For each sample, there were generally two pools; one that was of higher concentration and one of lower. Following dialysis, the concentration of each sample was determined by separating 5µl aliquots of each sample on SDS-PAGE gels along with known concentrations of BSA (Figure 3.2). Spot densitometry was then used to determine the protein concentration in each sample (Figure 3.2). Where possible, structurally sensitive antibodies were used to authenticate the tetraspanin proteins (Figure 3.3). It is apparent from Figures 3.2 and 3.3 that the GST-EC2 constructs are not 100% pure. In each sample there is some free GST (26kDal) the GST-EC2 protein (~35kDal) and what is presumably a degradation product (~30kDal). It is also noticible in the western blots that there is some free EC2 protein at around 10kDal that is recognised by the structurally sensitive antibodies.

A.



B.

Lane	Content	[Protein] mg/ml	Lane	Content	[Protein] mg/ml
1	BSA 10mg	10	9	Molecular weight marker	n/a
2	BSA 5.0mg	5	10	Human CD151 medium	6.2
3	BSA 2.5mg	2.5	11	Human CD151 high	7.4
4	BSA 1.0mg	1.0	12	Human CD151 low	3.4
5	Mouse CD9 high	5.9	13	Human CD63 high	3.1
6	Mouse CD9 low	1.4	14	Human CD63 low	1.9
7	Human CD9 high	8.2	15	Human CD81 high	12.1
8	Human CD9 low	2.8	16	Human CD81 low	3.3
			17	GST	2.8

Figure 3.2. Analysis of recombinant tetraspanin EC2s by SDS-PAGE. Following affinity purification the tetraspanin EC2 recombinant proteins (high and low concentrations pools) were dialysed against PBS and then separated on SDS-PAGE gels run under reducing conditions and stained with Coomassie (A) to determine the concentration of each sample. Table B shows the contents of each lane and the concentration determined by spot densitometry.

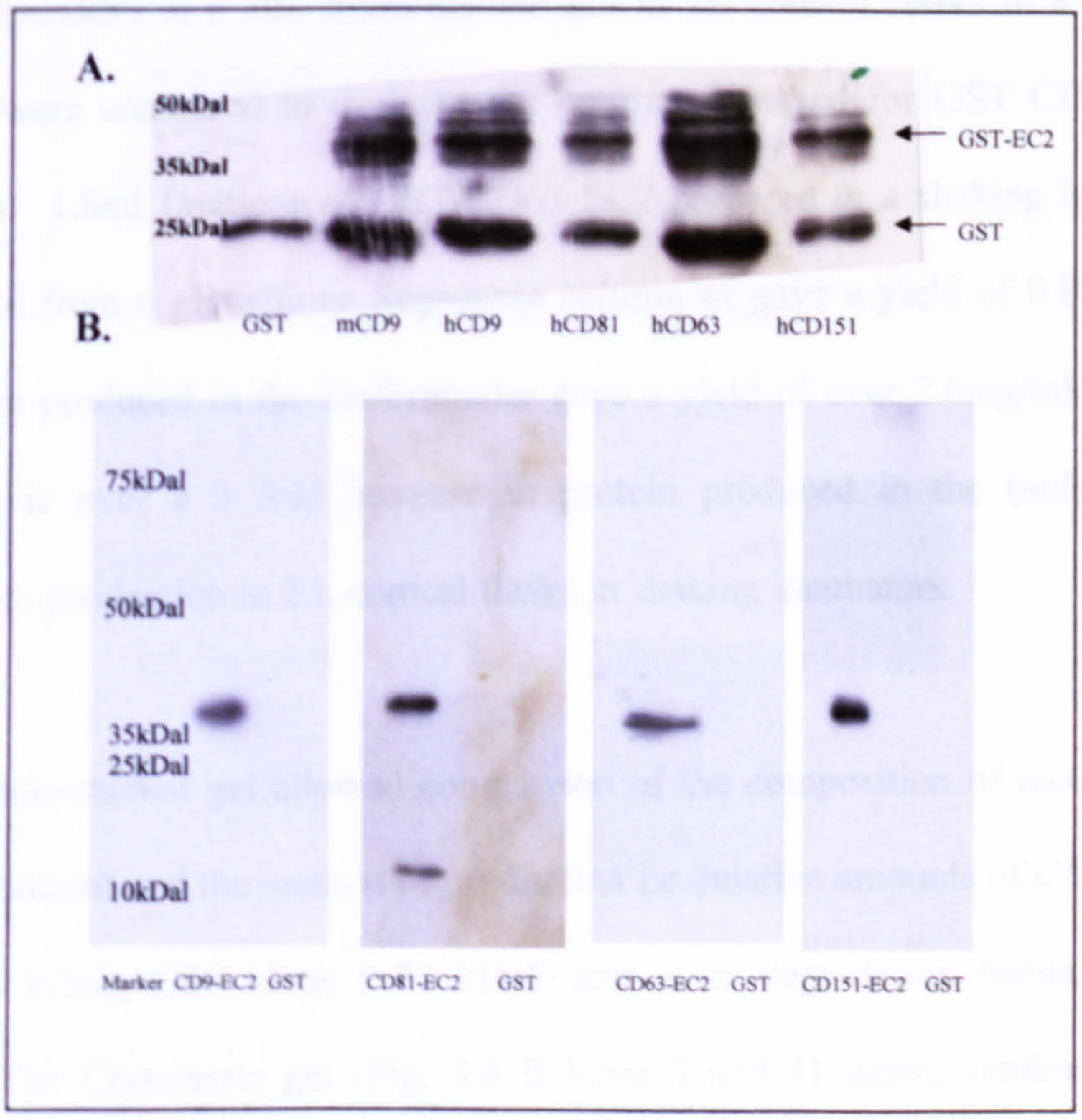


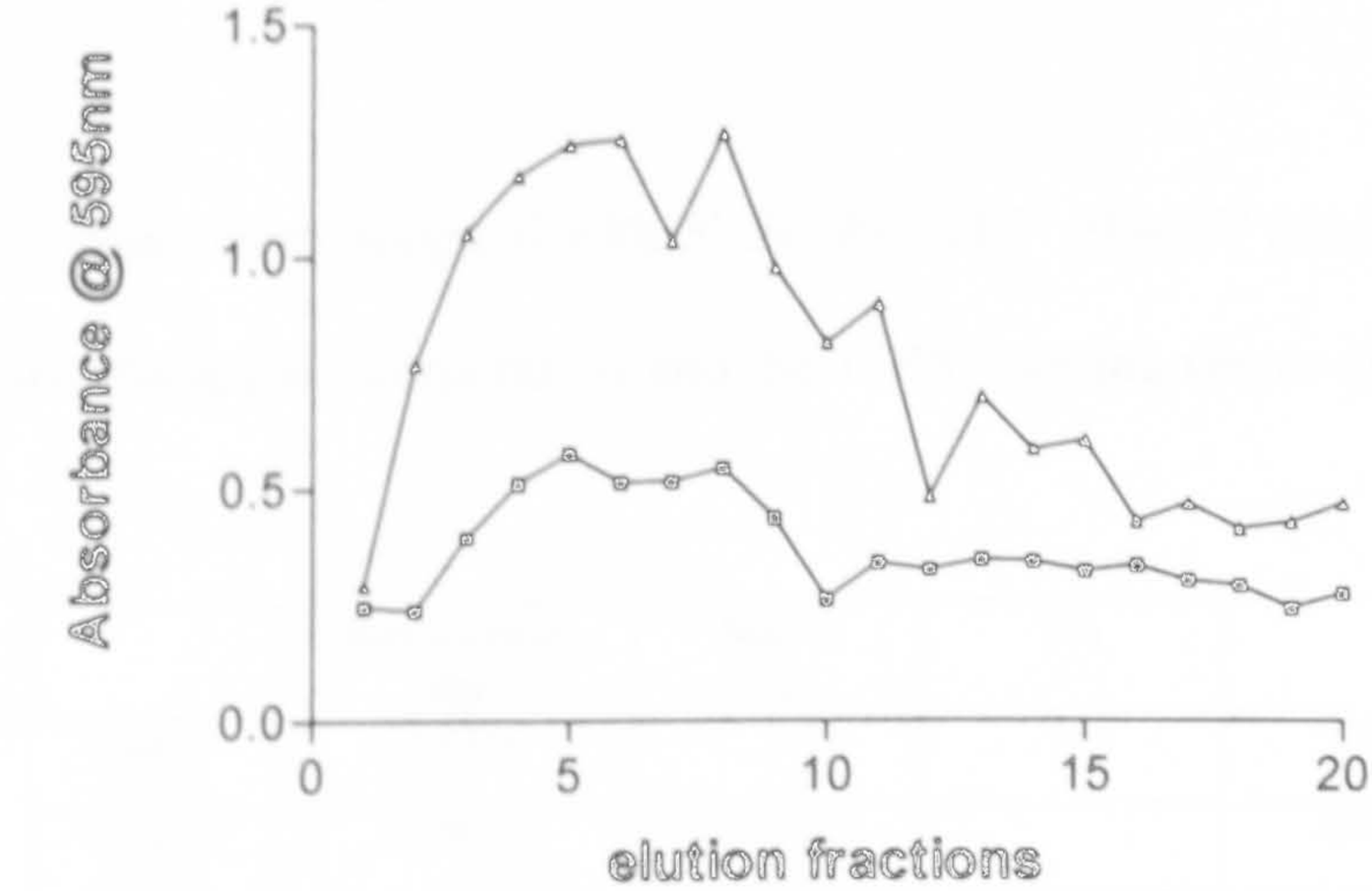
Figure 3.3 Tetraspanin recombinant protein analyses using Western blotting. 15 μg of the tetraspanin proteins were loaded on SDS-PAGE gels and run in the presence of β-mercaptoethanol prior to blotting. Membranes were blocked in 5 % w/v milk in PBS-T. In A, blots were probed with 1/3000 dilution of anti-GST-HRP diluted in blocking buffer. Figure B shows Western blots of tetraspanin EC2 constructs run on SDS-PAGE gels in the absence of reducing agent. Specific mAbs were used: ALMA-1 (hCD9), H5C6 (CD63), 1.3.3.2.2 (CD81) and 14A2 (CD151) each at 1/1000 dilution in blocking buffer followed by 1/3000 dilution of anti-mouse IgG diluted in blocking buffer.

3.3.2 Protein expression in a 20 L Biofermenter

Protein expression in a 20L biofermenter and in 2L conical flasks in a shaking incubator were compared to evaluate the optimum method for GST CD63-EC2 production. 1.0ml fractions of GST-CD63 EC2 prepared in a shaking incubator were eluted from a glutathione Sepharose column of gave a yield of 0.85mg/ml. The protein produced in the biofermenter gave a yield of over 2.6mg/ml Fig. 3.4 A). This is over a 3 fold increase in protein produced in the biofermenter compared to production in 2 L conical flasks in shaking incubators.

A Coomassie-stained gel allowed comparison of the composition of recombinant protein produced and the method of production i.e. relative amounts of GST-CD63 EC2 (~36 kDal), GST alone (~25 kDal) and other degradation bands that are present. The Coomassie gel (Fig. 3.4 B lanes 3 and 4) shows similar banding pattern between whole cell lysates and GST-CD63 EC2 produced by both methods, showing that production in the biofermenter does not affect the quality of the recombinant protein (top band only:- ~42 % GST-CD63 produced in the biofermenter versus ~ 40.5 % in shaking incubator). For functional assays, the concentration of EC2 protein was determined using all bands above the 25kDal GST alone band. The samples were used in functional assays with the contaminating GST present, but a GST control was always used to rule out any possibility that the free GST was having any affect.

A.



B.

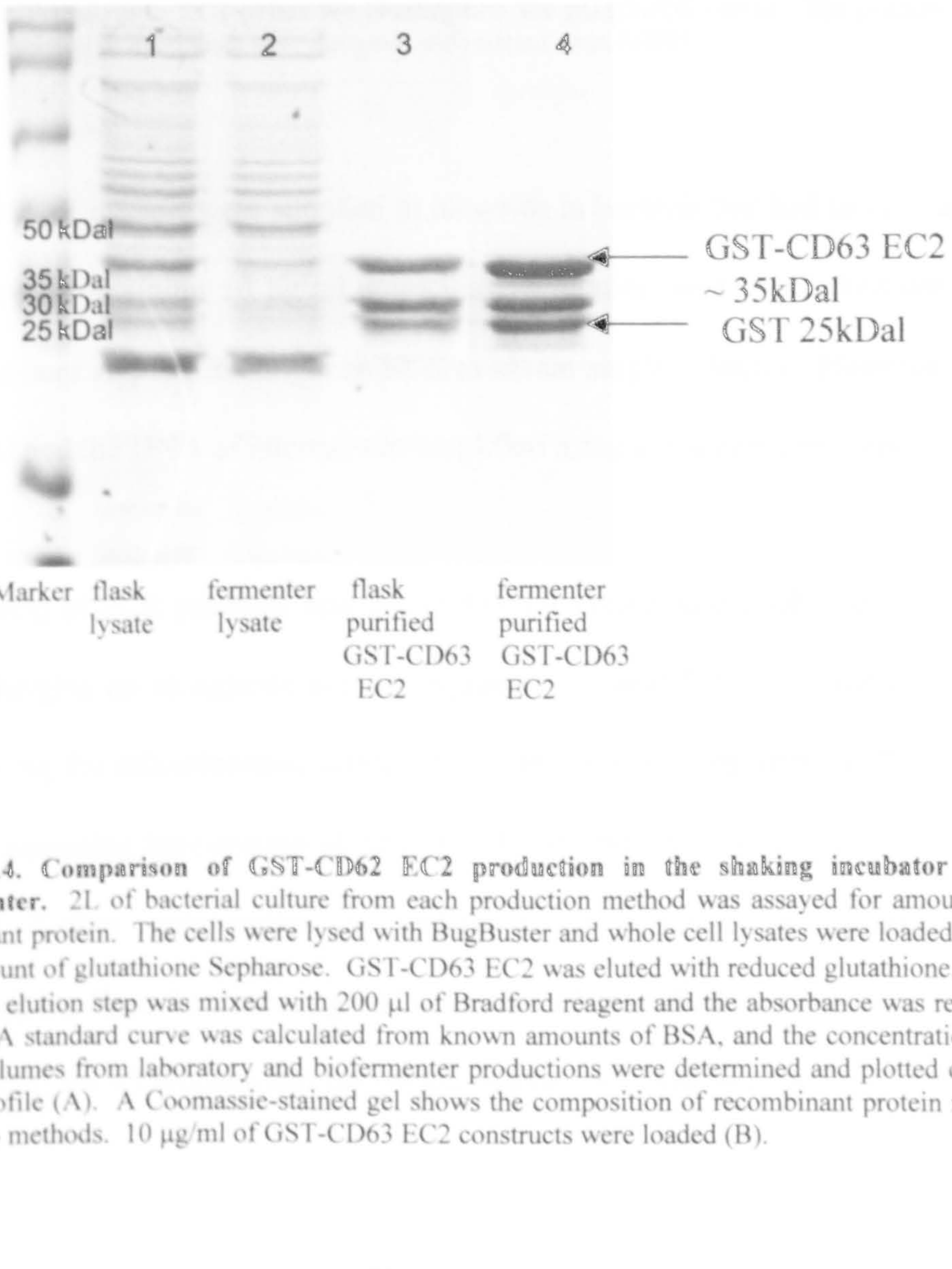


Figure 3.4. Comparison of GST-CD62 EC2 production in the shaking incubator and biofermenter. 2L of bacterial culture from each production method was assayed for amount of recombinant protein. The cells were lysed with BugBuster and whole cell lysates were loaded onto equal amount of glutathione Sepharose. GST-CD63 EC2 was eluted with reduced glutathione. 2 µl from each elution step was mixed with 200 µl of Bradford reagent and the absorbance was read at 595 nm. A standard curve was calculated from known amounts of BSA, and the concentration of elution volumes from laboratory and biofermenter productions were determined and plotted on an elution profile (A). A Coomassie-stained gel shows the composition of recombinant protein made by the two methods. 10 µg/ml of GST-CD63 EC2 constructs were loaded (B).

3.3.3 Cloning CD82 and CD231 EC2 domains

Primers were designed with 5' *EcoRI* and 3' *HindIII* restriction sites and a 5' TA overhang for incorporation into the TOPO cloning vector (Table 3.1).

	Restriction Site	Start/ stop	TA	Primer
CD82 F (5'-3')	GAATTC	-	TA	TACTTCAACATGGGCAAGC
CD82 R (5'-3')	AAGCTT	TCA		GTTCTCCTGCAGCAGCCACG
CD231 F (5'-3')	GAATTC	-	TA	TTTGTGTTTCGTCATGAGTC
CD231 R (5'-3')	AAGCTT	TCA		GTTAGTCTCCATGAAACTAGTTACC

Table 3.1. Tetraspanin EC2 prime for cloning into the pGEX-KG vector. The primers listed with their respective overhangs were designed and ordered from MWG.

The cDNA of interest were supplied in plasmids in bacteria that had been streaked on LB agar containing 50 µg/ml ampicillin. Colonies were re-streaked onto the same medium and left overnight at 37°C to obtain single colonies. Plasmids were extracted and the DNA of interest was amplified using the specific primers.

For cloning of PCR products into the TOPO-TA vector, single DNA bands at the correct heights on an agarose gel are required. A band for CD82 was generated easily using the recommended conditions for the proofreading enzyme Platinum® Pfx; an annealing temperature of 55°C for 30 seconds/cycle (Table 3.2). CD231 required further optimisation by changing the annealing temperature and increasing the time for annealing and extension. When the annealing temperature was at 50°C for 45 seconds and the period of extension was increased to 4 minutes,

a single band at the right size was obtained for CD231 EC2 (Table 3.2 A & B and Figure 3.5).

A.

PCR mixture	Volume/µl
10 x pfx amplifying buffer	5.0
10mM dNTP mix	1.5
50mM $MgSO_4$	1.0
1:1 mixture primers	1.5
Template DNA	1.0
Platinum pfx DNA pol	1.0
dH_2O	To 50 µl

B.

PCR conditions	Temp (°C) CD82	Temp (°C) CD231	Time (s) CD82	Time (s) CD231
DENATURE	94	94	15	15
ANNEAL	50-55	55-60	30	45
EXTEND	68	68	60	240
CYCLES	25 CYCLES			

Table 3.2. PCR conditions for amplification of CD231 and CD82 EC2. The composition of the reaction buffer for the PCR of each construct is listed in Table A. The conditions that produced single, discrete bands at the correct height are outlined in Table B.

The remainder of the PCR reaction mixture that gave single, discrete bands at the correct size on the agarose gel was used in the TOPO cloning reaction. The TOPO cloning reaction was carried out and transformed into TOP10 cells. 10–50µl of each transformation was spread onto pre-warmed LB agar plates containing 10µg/ml X-gal (for blue/white screening) and 50µg/ml of carbenicillin or 50µg/ml kanamycin (the TOPO vector has resistance to both antibiotics allowing for 2

rounds of selection). Around 40 white colonies had grown on both CD231 and CD82 transformants.

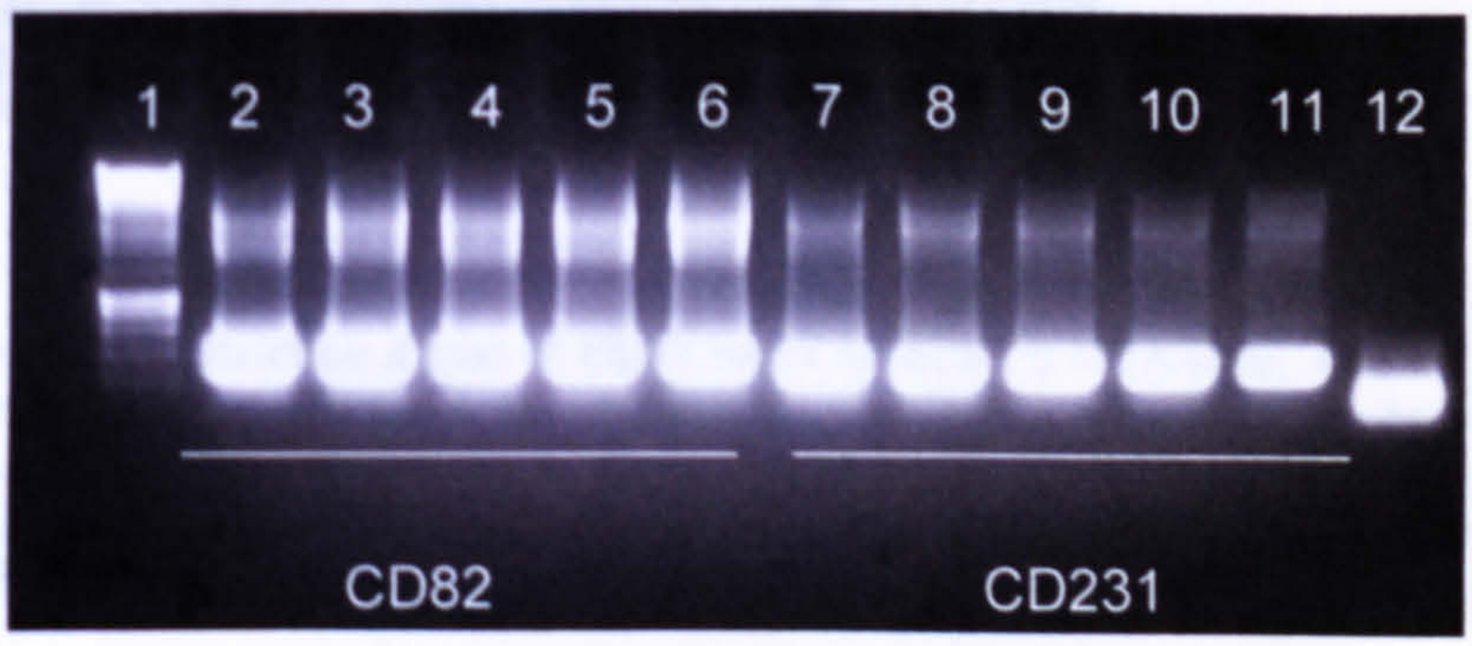


Figure 3.5. PCR amplification of the EC2 domains from CD82 and CD231. The PCR reaction was performed on a temperature gradient with Tm temperatures ranging from 50-55 ° C for CD82 and 55-60°C for CD231, at 1°C intervals. For both domains, all five temperatures gave single discrete bands at the right height (~350-400bp). In lane 1, there is a DNA marker and lane 12 is a no-template control.

Six white colonies were picked from each plate, grown overnight in LB containing the antibiotic that was not used in the LB agar, plasmids were extracted and digested with *Eco*RI to verify the presence of the insert (Fig. 3.6). The TOPO vector contains *Eco*RI restriction sites immediately up- and down-stream from the PCR product cloning site. The inserts were sequenced in the TOPO™ vectors to ensure that the full insert is present and of the correct sequence and orientation.

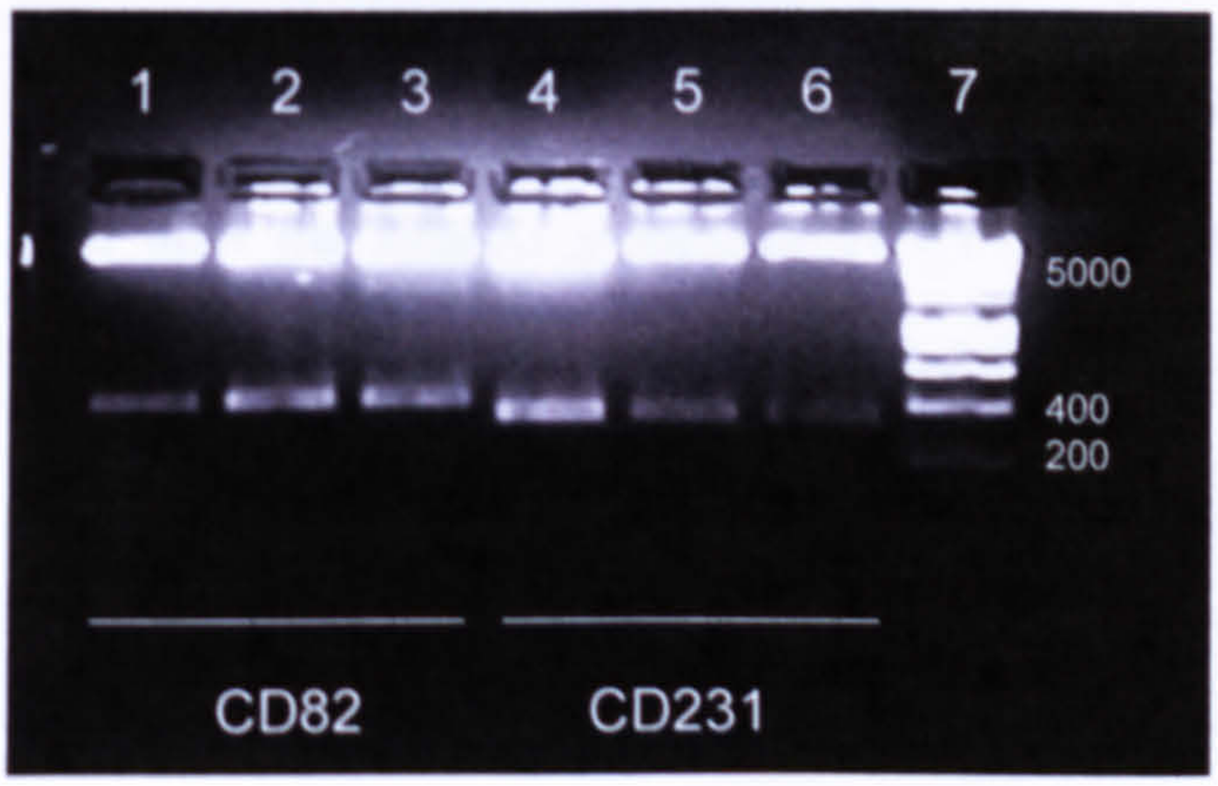


Figure 3.6. Restriction digest of CD82 and CD231 EC2 DNA from TOPO-TA. After cloning the inserts into the TOPO-TA vector (Invitrogen), colonies were screened for the presence of the interest before being sequenced to check the correct direction and sequence of the cloned DNA. TOPO-TA vector has *EcoRI* restriction sites flanking the inserted DNA. The miniprepped DNA was digested with *EcoRI* for 3 hours at 37˚C and reaction mixtures were run on a 1 % agarose gel.

After identification of correct inserts in the TOPO vector, a double digest of both TOPO-clone and pGEX-KG vector was performed. The digestion reactions were set up according to Table 3.3.

Component	Volume/µl	Component	Volume/µl
TOPO Clone	10	pGEX-KG	8
Multicore	2	Multicore	2
Water	6.5	Water	8.0
HindIII	1	*HindIII*	1
EcoRI	0.5	*EcoRI*	0.5

Table 3.3. Components of the HindIII, EcoRI double digest of CD231 and CD82 DNA from TOPO-TA and the insert in pGEX-KG expression vector. Correctly sequenced TOPO-TA vectors containing the CD231 and CD82 were digested with the components described in the two left hand columns and the pGEX vector as described in the two right hand columns. (*HindIII* was added alone to the TOPO-TA vector for 2 hours before addition of the *EcoRI*. This is to ensure correct cutting by the *EcoRI* as there are additional *EcoRI* sites in the TOPO-TA vector).

The digestion was left for 2 hours at 37˚C. *EcoRI* was added to the TOPO clone 1 hour after the *HindIII* to aid correct cutting as there are additional *EcoRI* sites flanking the insert. The reaction mixtures were run on high purity gels, the 5 Kb cut vector and 350-400bp EC2 DNA from CD82 and CD231 (as shown by *) were excised from the gel and the DNA purified (Figure. 3.7).

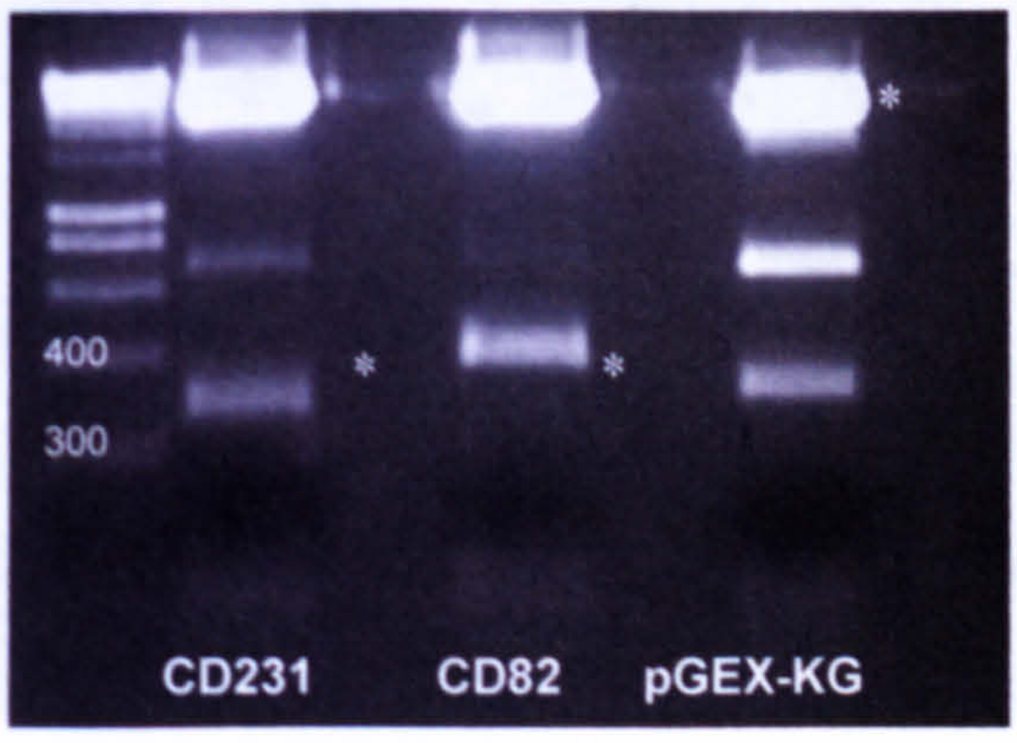


Figure 3.7. HindIII, EcoRI double digest of CD82 and CD231 DNA from TOPO-TA and the pGEX-KG expression vector . TOPO-TA vectors containing CD213 and CD82 EC2 inserts were digested with *HindIII* and *EcoRI* for a total of 3 hours at 37˚C. Reaction mixtures were run on high purity agarose gels and the 350-400bp EC2 inserts (*) as well as the cut pGEX vector (*) were excised and the DNA purified, ready for ligation.

The purified, cut vector DNA was ligated with CD82 and CD231 tetraspanin DNA for 2 hours at 14˚C. 5µl of each ligation reaction mixture was used to transform 100 µl DH5α competent cells. CD82 and CD231 colonies grew on the plates with insert: vector ratios of 3:1 and 1:1, respectively. Presence of the insert was examined by a direct PCR screen of the colonies using CD82 or CD231 specific primers. Positive colonies were picked, the plasmid was purified and the DNA digested with *EcoRI*/*HindIII* to confirm the presence of the inserts in the pGEX

vector (Figure. 3.8). The DNA was sequenced in the pGEX vector to ensure no ensuing mutations had occurred and to verify the correct orientation and reading frame of the insert.

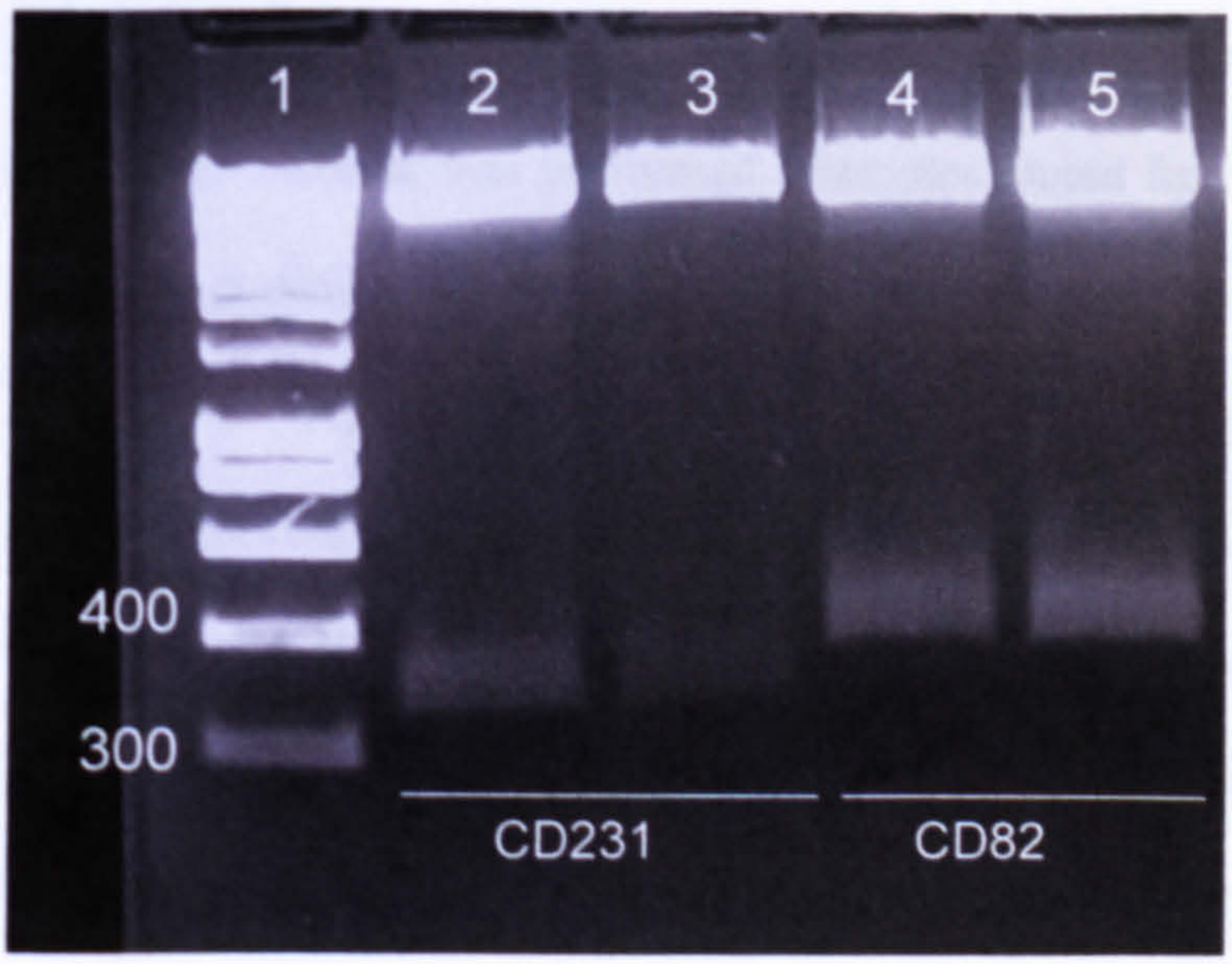
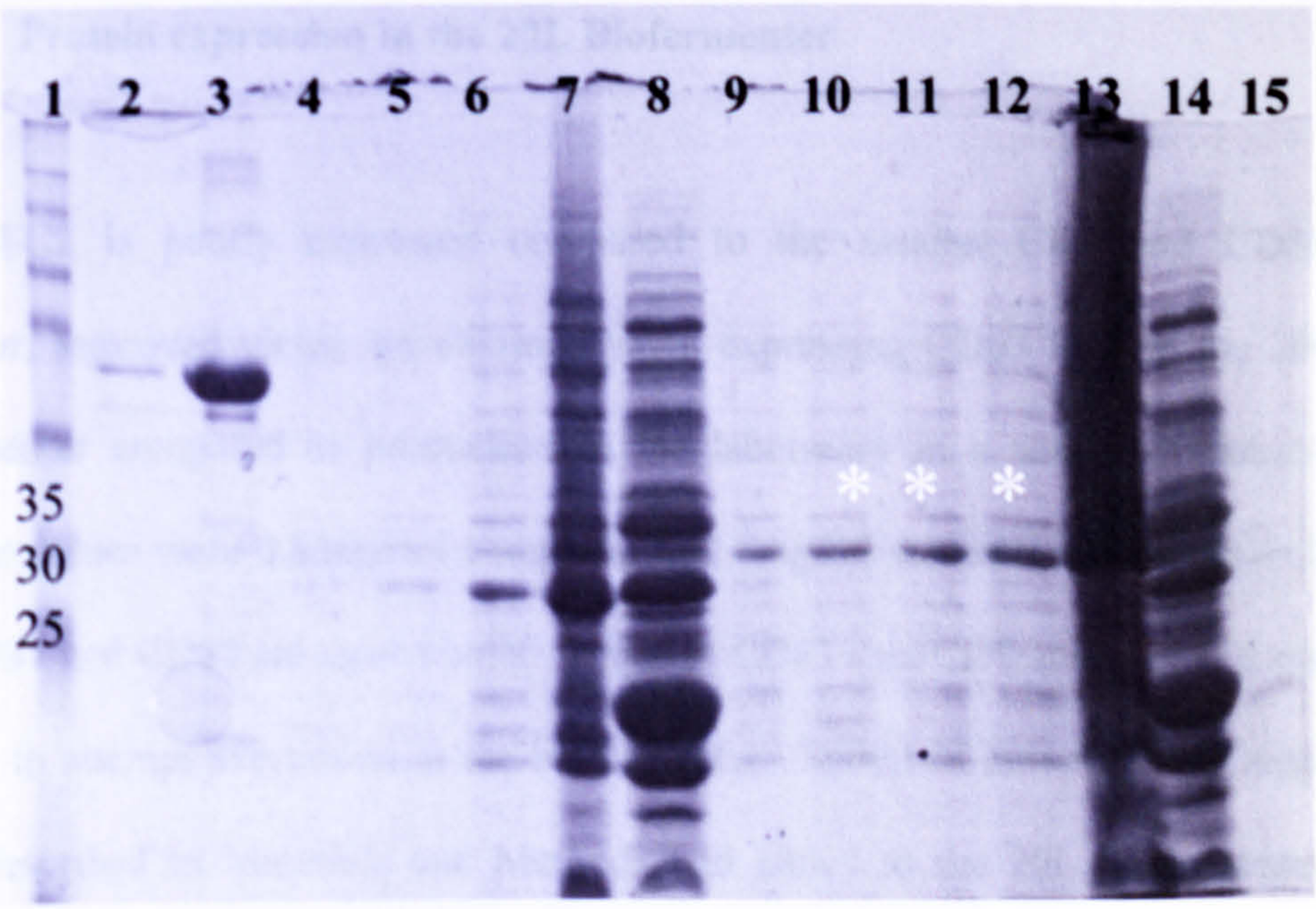


Figure 3.8. Double digest of pGEX-KG vector to confirm the presence of CD231 and CD82 inserts. Following ligation of the inserts excised from the TOPO-TA vector with the cut pGEX-KG vector, the ligation reactions were transformed into DH5α cells and the plasmids were purified. A double digest with *Hind*III and *Eco*RI was performed to confirm that the ligation had worked and the insert was present. The reactions were run on 1 % agarose gels.

3.3.4 Expressing CD82 and CD231 recombinant protein

A method to purify the EC2 domains from CD9, CD81, CD151 and CD63 has already been established (Flint et al. 1999). BL21 Codon Plus *E. coli* cells are transformed with the pGEX vector and a 10ml overnight inoculum added to 400ml LB and grown until OD_{600} = 0.8 – 1.0, before induction with 0.1mM IPTG. The bacterial cells are then left to express the protein for 4 hours before being harvested

by centrifugation. This, therefore, was the first method of choice when attempting CD231 and CD82 protein expression. Colonies grew on the antibiotic selection plates, suggesting that the transformation had worked and the plasmid had been taken up. Colonies were picked and grown as described for the other tetraspanins. The cell pellets were lysed with BugBuster and protein purification, using the glutathione Sepharose column, was performed. Samples eluted from the column with free, reduced glutathione were assayed for the presence of GST-tagged protein using Bradford Reagent. The first attempt showed no eluted protein. The whole expression procedure was repeated using freshly made LB-Agar plates, reduced glutathione and IPTG but there was still no production. It was then decided to decrease the temperature to 25°C and to use 1.5mM IPTG to induce. Samples from the purification procedure were run on an SDS gel and stained with Coomassie to test for the presence of correctly sized proteins (Figure. 3.9). There appeared to be evidence of some protein at around the correct weight (35 kDal) for CD82 in the eluted samples 2–4 although the Western blot for CD82 appeared negative in these lanes.



CD231 CD82

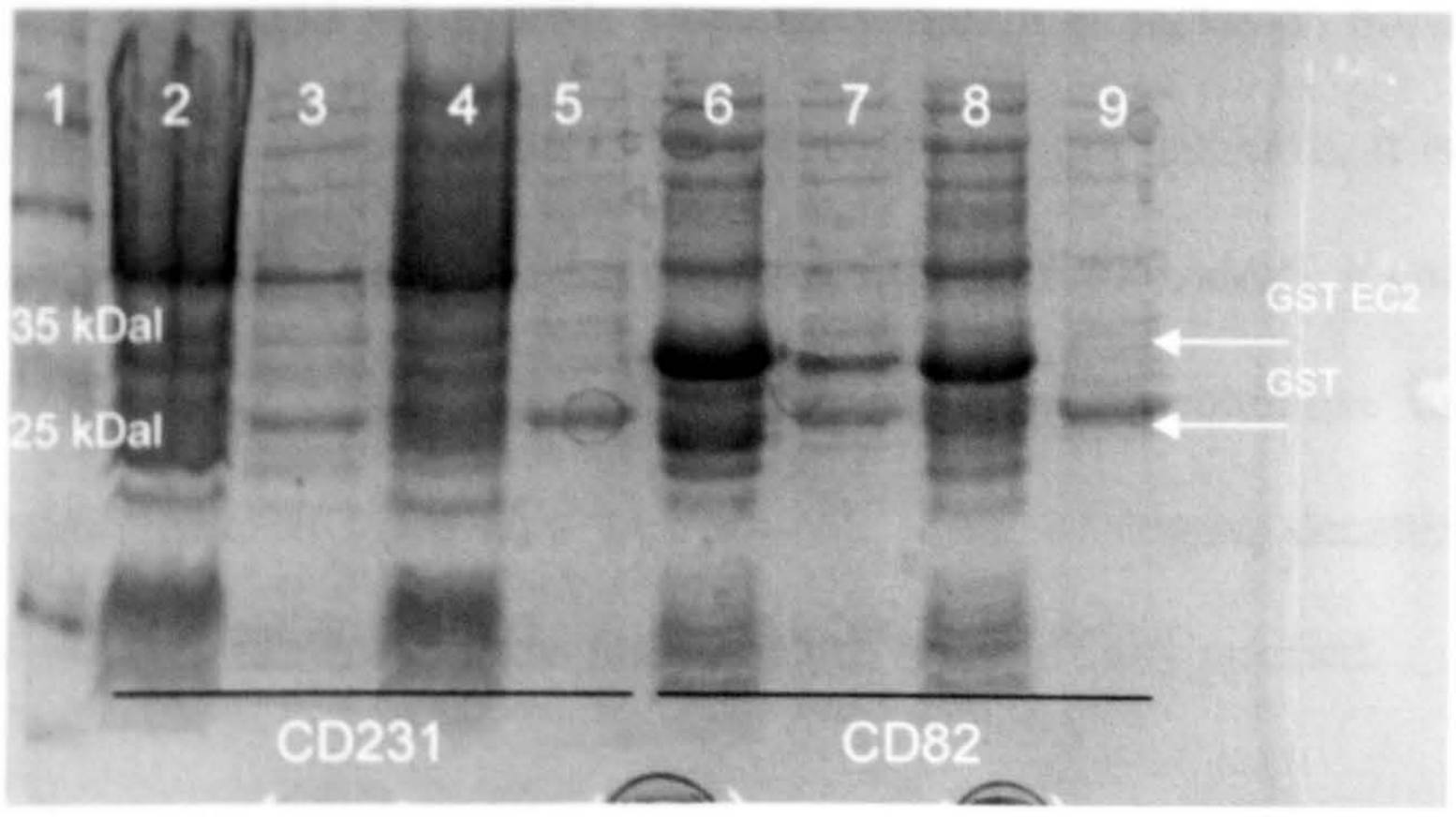
Lane	Content	Lane	Content
1	Marker	9	CD231 pellet
2	BSA 0.5mg	10	CD82 elute # 2
3	BSA 1.0mg	11	CD82 elute # 3
4	CD231 elute # 2	12	CD82 elute # 4
5	CD231 elute # 3	13	CD82 beads
6	CD231 elute # 4	14	CD82 Flow-through
7	CD231 beads	15	CD82 pellet
8	CD231 flow-through	16	-

Figure 3.9. Coomassie gel showing expression of CD231 and CD82 in shaking incubators at 25°C, with 1.5mM IPTG induction. Cells were lysed and GST-EC2 proteins were purified on glutathione Sepharose columns. Samples from various stages in the purification procedure were boiled in loading buffer and run on SDS-PAGE before staining in Coomassie.

3.3.5 Protein expression in the 20L Biofermenter

CD63 EC2 is poorly expressed compared to the smaller CD9 and CD81. However, improved yields are obtained when expressing CD63 EC2 in the 20L biofermenter compared to production in the laboratory in a shaking incubator (average values were 0.85mg/ml compared to 2.6mg/ml respectively, Figure 3.2). As CD231 and CD82 are more similar in size to CD63 than CD9 and CD81, it was decided to attempt expression in the biofermenter. 500ml of inoculate was made up as described in Materials and Methods and added to the 20L biofermenter. When OD_{600} reached 0.6, IPTG was added to a final concentration of 0.1mM proceeded by growth at 37°C for 4 hours.

There was a band at 25kDal in the CD231 and CD82 purified protein and glutathione Sepharose bead samples (Figure. 3.10), this is likely to be GST. There are no other clear, additional bands in the CD231 purified protein column (lane 3) but there is a band at 35kDal for CD82 (lane 7). However, this was not recognised by the commercial CD82 antibodies in a Western blot and yields were extremely low (0.2mg/L). It became evident that a more efficient method was required.



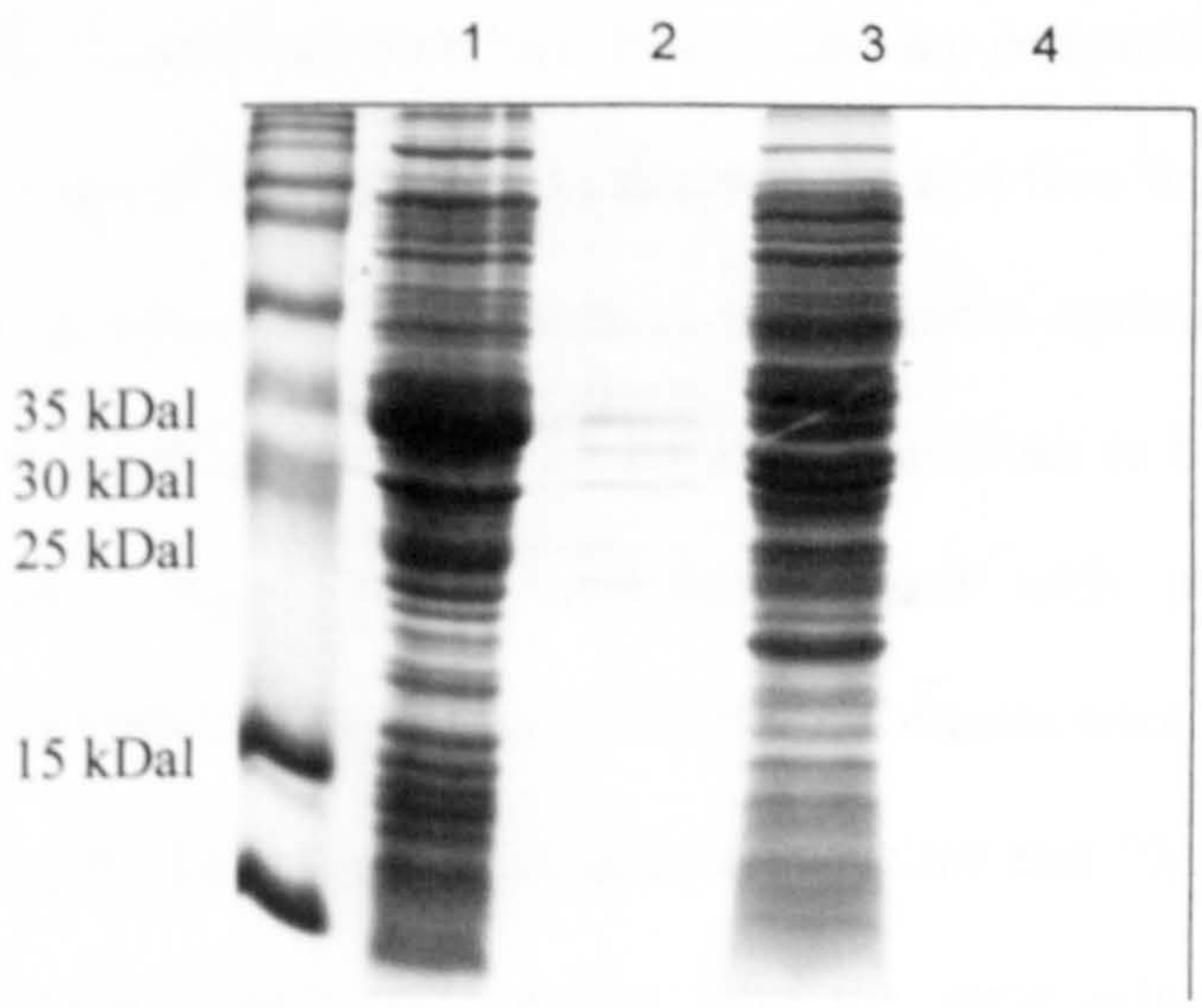
Lane	Contents	Lane	Contents
1	Marker	**6**	CD82 lysed cell pellet
2	CD231 lysed cell pellet	**7**	CD82 purified
3	CD231 purified	**8**	CD82 flow through
4	CD231 flow through	**9**	CD82 beads
5	CD231 beads	-	-

Figure 3.10. Coomassie stained gel showing CD231 and CD82 EC2 production in a 20L biofermenter. Cells were harvested and lysed using the method for protein expression in a shaking incubator. GST-tagged proteins were purified on glutathione-Sepharose beads. Samples from various stages in the purification procedure were boiled in loading buffer and run on SDS-PAGE before being Coomassie stained.

3.3.6 Protein purification from the insoluble fraction

It is common for complex recombinant protein expressed in bacterial cells to become trapped in insoluble aggregates known as inclusion bodies. Despite the fact that the other tetraspanin EC2 domains are not insoluble, it is possible CD82 and CD231 EC2 proteins have different properties. Indeed, it could explain why they are not being detected in BL21 cells. To solubilise inclusion bodies, guanadinum-HCl and urea are required. This, of course, denatured the proteins, which then needed to be re-folded before being affinity purified.

Figure 3.11 shows the whole insoluble fraction from BL21 cells transformed with CD82 (Lane 1) and CD231 (Lane 3), and the eluate from glutathione Sepharose columns loaded with the whole cell lysate, washed, and protein eluted with reduced glutathione. Although there are some bands in the eluate in lane 2, these were not recognised by the anti-CD82 antibody. There is no evidence of any GST-tagged protein in lane 4.



Lane	1	2	3	4
contents	CD82 insoluble fraction	CD82 elute from glutathione Sepharose column	CD231 insoluble fraction	CD231 elute from glutathione Sepharose column

Figure 3.11. Coomassie stained gel showing CD82 and CD231 purification under denaturing conditions. The insoluble fraction from *E.coli* transformed with CD82 and CD231 was isolated using 8 M Urea (lanes 1 and 3) and then GST-tagged proteins were purified on a glutathione-Sepharose column (lanes 2 and 4) and run on SDS-PAGE gels before protein staining by Coomassie took place.

3.3.7 The use of Origami cells to produce CD82 and CD231 EC2 domains

Origami cells have thioredoxin reductase (*trxB)* and glutathione reductase (*gor*) mutations that create a lower cytoplasmic redox potential, favouring the formation of disulphide binds. The mutations are carried on plasmids that are selectable on kanamycin and tetracycline; therefore these strains are compatible with the pGEX vector that is ampicillin resistant. As this was the first time that these cells were to be used for tetraspanin EC2 production, an expression-optimisation experiment was performed allowing a total of 54 different conditions to be screened (Table 3.4). Competent Origami cells were transformed with pGEX-KG vectors containing the CD82 EC2 or CD231 EC2 insert. Cultures were grown on LB agar containing 15µg/ml kanamycin, 12.5µg/ml tetracycline and 50µg/ml carbenicillin. Overnight cultures (10ml) were grown with the same selection and used to inoculate 400ml LB with selection antibiotics added.

The cultures were grown at 25°C or 37°C, and to a range of OD_{600} values (0.2, 0.4, 0.8) before being induced with 0.1, 0.5 or 1.0mM IPTG. Finally, the cultures were left to express the protein for 2 hours, 4 hours or overnight, before being harvested by centrifugation. Non-induced controls were included for background protein analysis purposes. The samples were lysed with BugBuster and centrifuged to leave the whole cell lysate. As a quick screen to check that some protein had been expressed, a range of samples were chosen at random for immunoblotting. 1µl of each sample was dotted on to the nitrocellulose three times. Using anti-GST antibody, it was observed that all of the samples tested had a greater amount of GST compared to the background (non-induced) sample (Figure 3.12).

Sample #	temp	OD600=	[IPTG]/mM	time/hrs	temp	Sample#
1	37	0.2	0.1	2	25	28
2	37	0.2	0.1	4	25	29
3	37	0.2	0.1	O/N	25	30
4	37	0.2	0.5	2	25	31
5	37	0.2	0.5	4	25	32
6	37	0.2	0.5	O/N	25	33
7	37	0.2	1	2	25	34
8	37	0.2	1	4	25	35
9	37	0.2	1	O/N	25	36
10	37	0.4	0.1	2	25	37
11	37	0.4	0.1	4	25	38
12	37	0.4	0.1	O/N	25	39
13	37	0.4	0.5	2	25	40
14	37	0.4	0.5	4	25	41
15	37	0.4	0.5	O/N	25	42
16	37	0.4	1	2	25	43
17	37	0.4	1	4	25	44
18	37	0.4	1	O/N	25	45
19	37	0.6	0.1	2	25	46
20	37	0.6	0.1	4	25	47
21	37	0.6	0.1	O/N	25	48
22	37	0.6	0.5	2	25	49
23	37	0.6	0.5	4	25	50
24	37	0.6	0.5	O/N	25	51
25	37	0.6	1	2	25	52
26	37	0.6	1	4	25	53
27	37	0.6	1	O/N	25	54

Table 3.4. Conditions for GST-231 EC2 and GST-82 EC2 production in Origami cells. 400mls of LB plus selection was inoculated with 10mls of overnight cultures of Origami cells transformed with pGEX-KG-CD82 EC2 or pGEX-KG-CD231 EC2. Flasks were left to shake at 37°C or 25°C until the above mentioned OD_{600} values were reached. For each set of conditions, 10ml of the culture was taken out and induced with the appropriate concentration of IPTG before being left to express the protein for the indicated time.

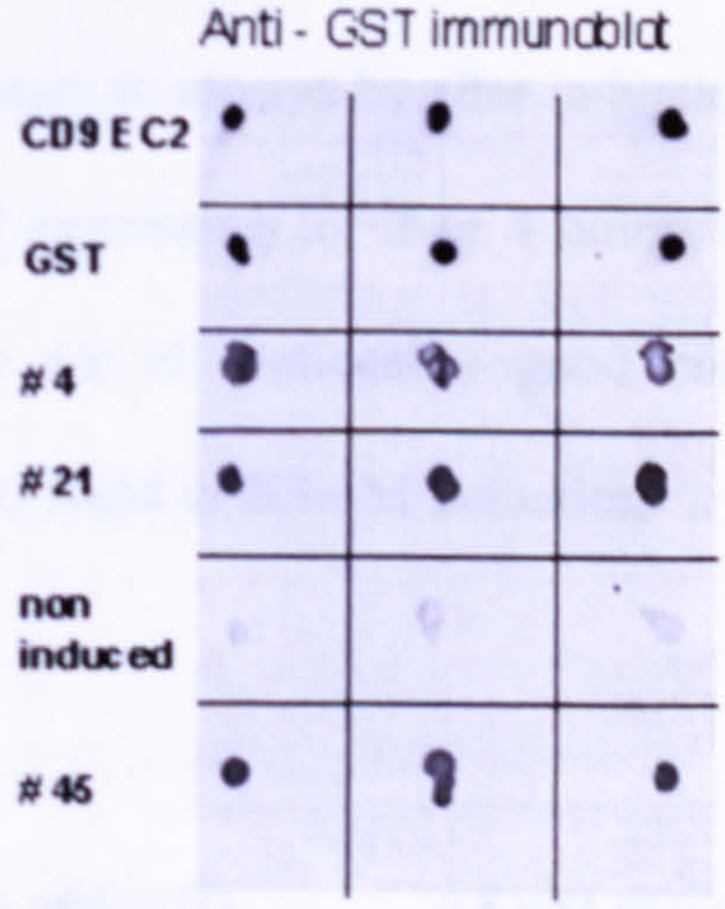


Figure 3.12. Dot blot showing GST-82 EC2 production in Origami cells. 400ml of LB plus selection was inoculated with 10ml of overnight cultures of Origami cells transformed with pGEX-KG-CD82 EC2 or pGEX-KG-CD231 EC2. Flasks were left to shake at 37˚C or 25˚C until the above mentioned OD_{600} values were reached. For each set of conditions, 10ml of the culture was taken out and induced with the appropriate concentration of IPTG before being left to express the protein for the indicated time. All of the samples were lysed with BugBuster and initially, a selection were immunoblotted onto nitrocellulose, blocked in 5 % w/v milk PBS-T and probed with anti-GST-HRP (1/3000 diluted in blocking buffer) to ensure that there was a level of GST-tagged proteins in the induced samples above the background levels (non-induced).

As there did appear to be GST protein present in the induced samples, it was decided to immunoblot the remaining samples to see if there was any correlation between the conditions used and the amount of protein produced. This is a crude method and was performed as a rough guideline only. In each square of the grid, 2 x 1µl drops were placed for each condition and blots were developed in the usual manner using an anti-GST antibody (Figure 3.13). A number of observations can be made from the blot. Firstly, it is clear that overnight expression at 37˚C has detrimental effects on protein production as shown by the diminished spots in the

overnight squares compared to the 2 and 4 hour results. This effect is not seen with proteins expression at 25˚C. For proteins expressed at both temperatures, induction at OD_{600} = 0.8 appears to be unfavourable as less GST-tagged protein is identified in these squares compared to induction at OD_{600} 0.2 and 0.4. The optimum expression appears to always be after induction at OD_{600} = 0.2. At 37˚C, using 0.5mM IPTG and expressing for 2 or 4 hours, or using 1.0mM IPTG and expressing for 2 hours are all noticeably good conditions, as are overnight expression at 25˚C with 0.1mM or 0.5mM induction.

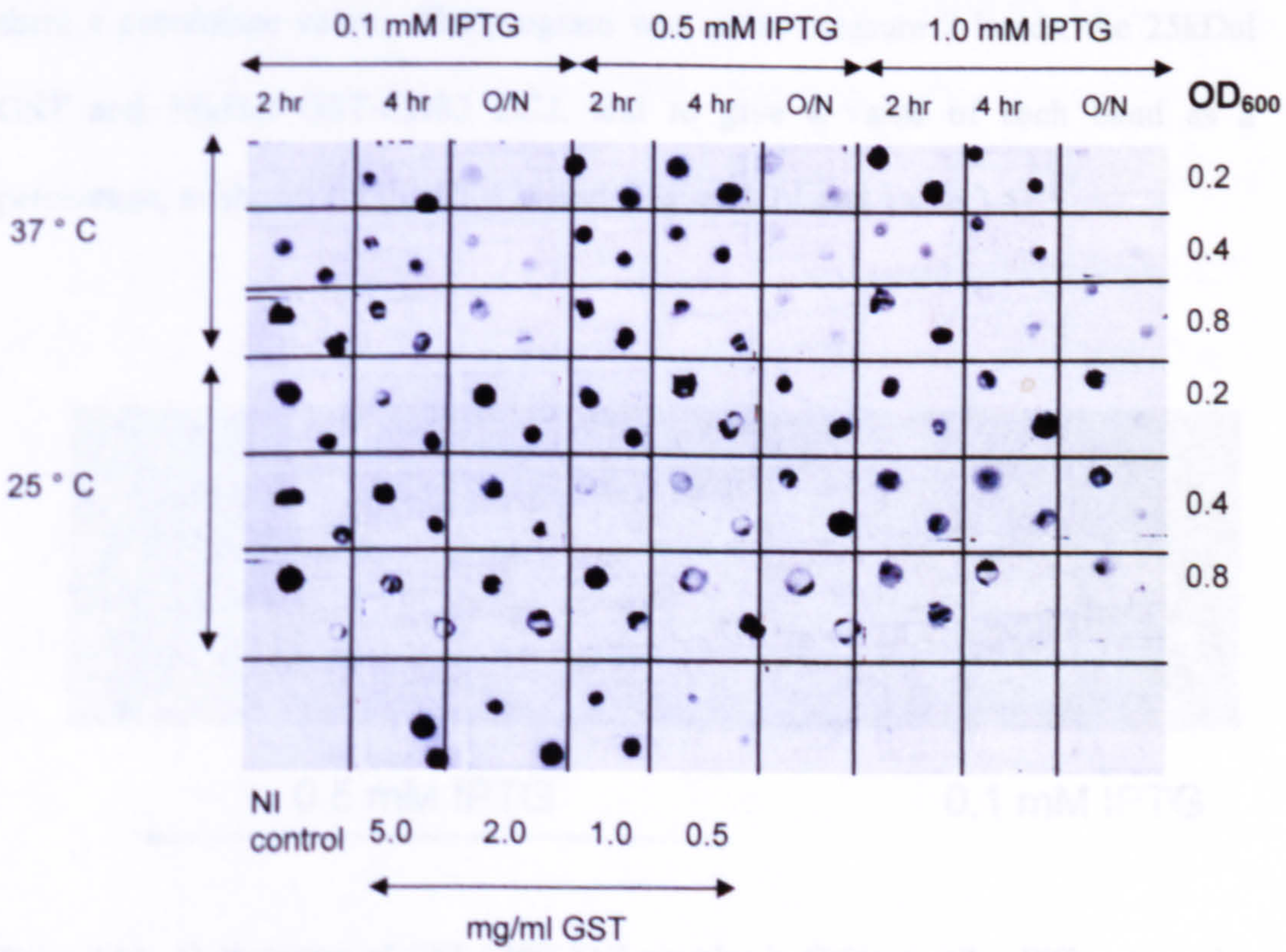


Figure 3.13. Dot blot showing production of GST-CD82 EC2 in Origami cells. All of the lysed samples from Table 3.4 were blotted onto nitrocellulose in duplicate. The nitrocellulose was then blocked in 5% milk in PBS-T and probed with anti-GST-HRP (1/3000 diluted in blocking buffer).

However, dot blots do not give any information about the quality of protein produced i.e. the ratio of GST-EC2 : GST alone. In order to gain this information SDS-PAGE gels are required. The candidate samples identified by the dot blot were mixed with glutathione-Sepharose beads in 1.5ml tubes. GST and GST-tagged protein were eluted with reduced glutathione in a batch purification procedure and aliquots were run on an SDS-PAGE gel and stained with Coomassie, Figure 3.14). Because of the increased degradation effects seen when growing at 37˚C we concentrated solely on the samples grown at 25˚C.

The densitometry program measures the density of all bands in each lane and gives them a percentage value. The program was set to measure 2 bands, the 25kDal GST and 35kDal GST-CD82 EC2, and to give a value of each band as a percentage, as shown for the CD82 band (Figure 3.14 and Table 3.5).

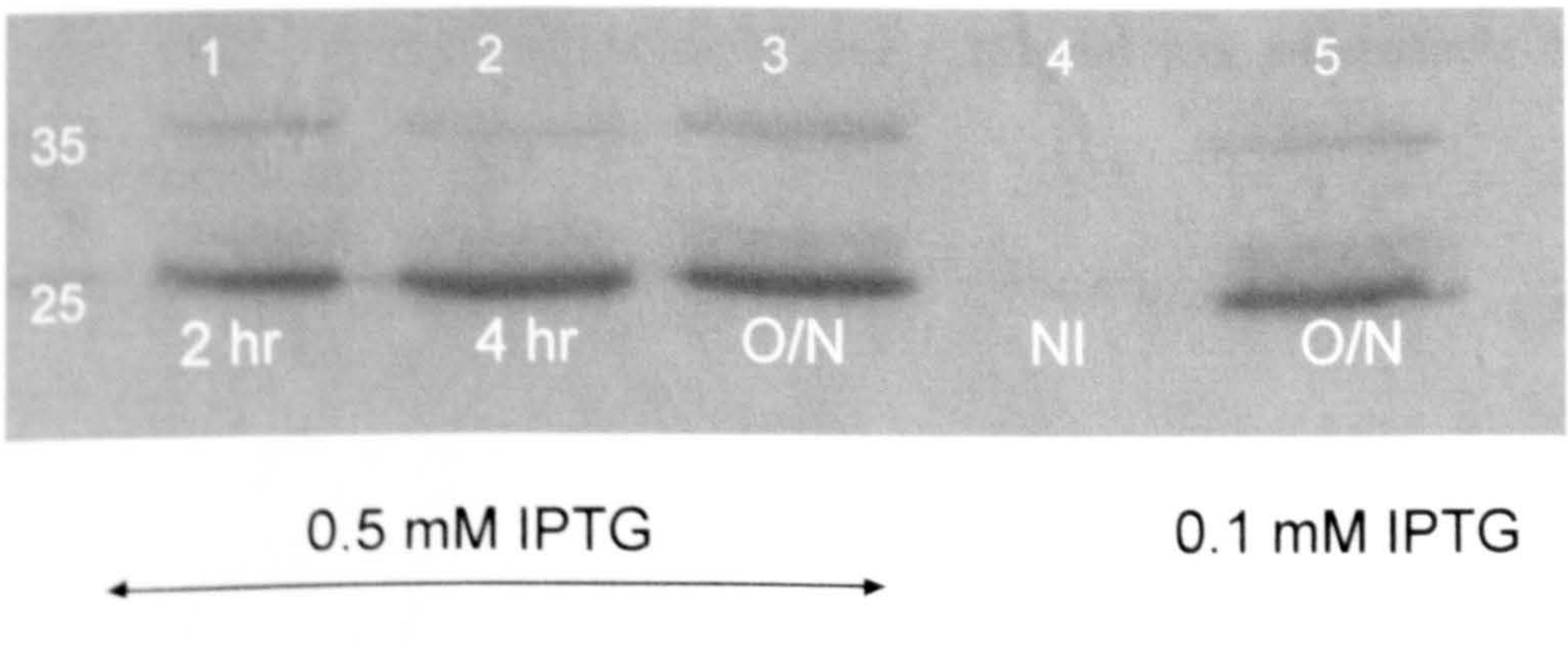


Figure 3.14. Optimisation of GST-CD82 EC2 proteins in Origami cells. Different samples picked out from the immunoblot (Fig. 3.13) were run on SDS-PAGE and Coomassie stained to look at the composition of the recombinant protein.

Lane	1	2	3	4	5
Value %	35.0	28.9	28.1	0.0	27.0

Table 3.5. Densitometry measurements of the percentage of GST-CD82 EC2 relative to GST taken from the coomassie in Figure 3.14. The densitometry of the bands in each lane were calculated using the Alpha-Imager. Percentage values of GST-CD82 EC2 are shown for lanes corresponding to those in Figure 3.14.

From this, it was evident that growing at 25°C, inducing with 0.5mM IPTG when the culture density reaches OD_{600} 0.2 and expressing the protein for 2 hours were the optimum conditions. Even though there was apparently more total protein produced using other conditions, the maximum amount of GST-CD82 EC2 relative to GST is desired. A Western blot probed with TS82 showed that the GST-CD82 EC2 was recognised by this antibody as well as a faint band at 10kDal that is presumably cleaved CD82 EC2 alone (Figure 3.15) and was undetectable on Coomassie-stained gels.

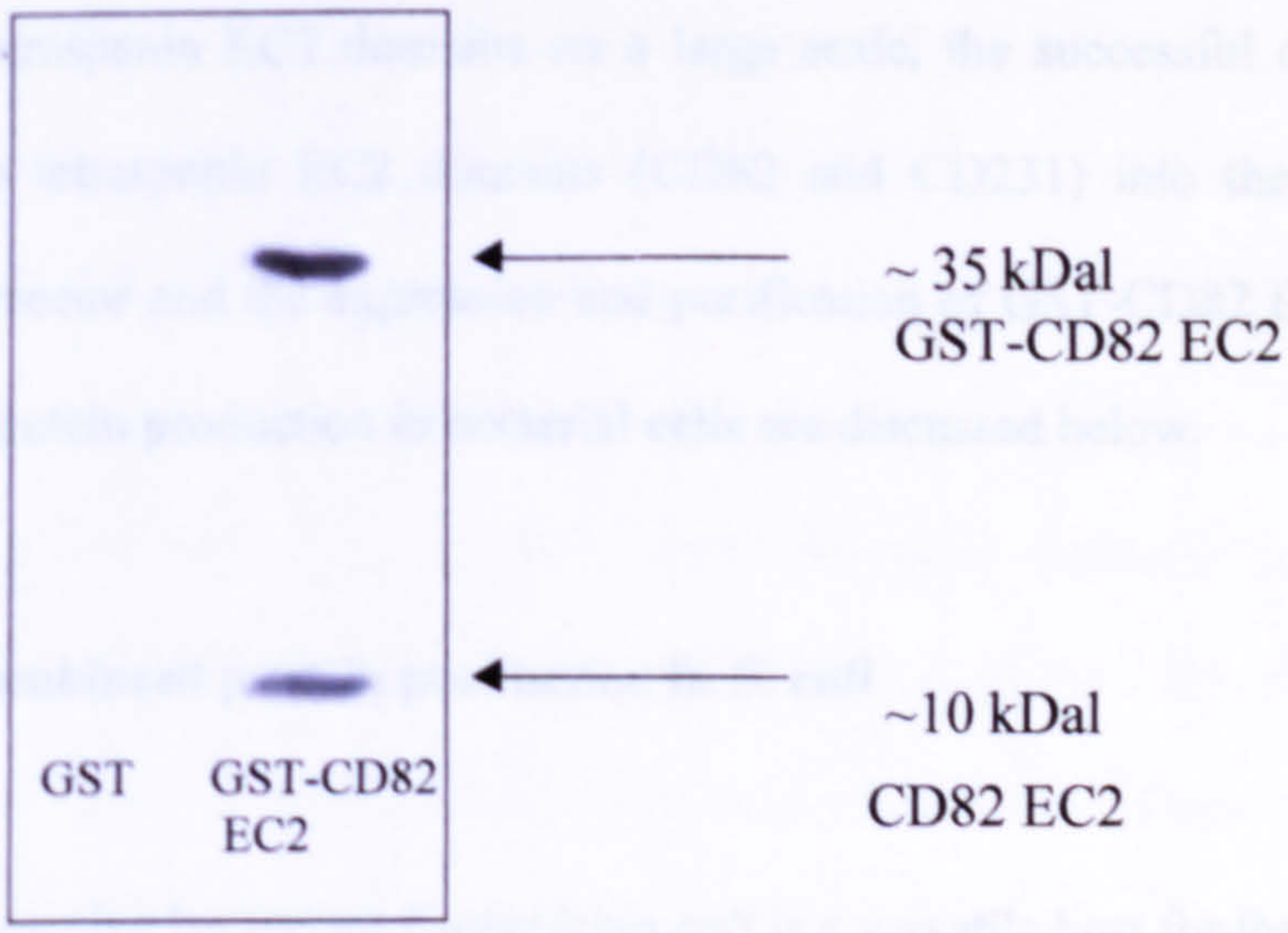


Figure 3.15. Western blot showing recognition of the GST-CD82 EC2 by the anti-CD82 antibody, TS82. Freshly produced GST-CD82 EC2 (10μg) and a GST (10μg) control were run on SDS-PAGE, transferred onto nitrocellulose and blocked in 5 % milk in PBS-T before being probed with TS82 (1/500 diluted in blocking buffer) followed by anti-mouse IgG-HRP (1/3000 in diluted in blocking buffer).

The same optimisation of GST-CD231 EC2 protein was performed in Origami cells but the immunoblots were negative for any GST-tagged protein produced despite presence of CD231 DNA in the Origami cells as shown by PCR of DNA from a plasmid purification of a sample of the culture medium taken after induction. CD231 specific primers were used and a band at the correct size was seen when the PCR product was run on an agarose gel, suggesting that although pGEX-CD231 EC2 is being taken up by the Origami cells, it cannot be expressed.

3.3 DISCUSSION

This chapter describes the successful development of a strategy to produce high levels of tetraspanin EC2 domains on a large scale, the successful cloning of a further two tetraspanin EC2 domains (CD82 and CD231) into the pGEX-KG expression vector and the expression and purification of GST-CD82 EC2. Some aspects of protein production in bacterial cells are discussed below.

3.3.1 Recombinant protein production in *E. coli*

The gram-negative bacterium *Escherichia coli* is a versatile host for the production of heterologous protein. Eukaryotic proteins can be expressed and purified on a large scale with relative ease in a host whose genetics are far better characterised than any other microorganism. However, there are limits to the potential of expressing eukaryotic proteins in *E. coli*, although with recent advances in biotechnology, the majority of these limits are gradually being overcome. All in all, *E. coli* is the initial host of choice when commencing recombinant protein production. Expression of CD9 and CD81 EC2 domains give the largest yields, probably because these extracellular loops are the smallest in the tetraspanin family, having just four cysteines that are predicted to form two disulphide bridges. The remainder of the EC2 domains made here are larger and more complicated with additional disulphide bridges that are necessary for protein folding. This would certainly help to explain the lower yields of CD63, C151 and CD82, and of the complete lack of expression of CD231.

3.3.2. Rare codon usage

One limitation in the expression of heterologous proteins in *E. coli* is the difference in tRNA usage between the organism from which the heterologous protein is derived and that of *E. coli*. If the tRNA required for high level synthesis of a heterologous protein is rare in *E. coli*, stocks may quickly become depleted and consequently the synthesis of the heterologous protein will be stalled. Due to this reason, the EC2 proteins that were previously produced in Sheffield (Higginbottom et al. 2000; Zhang et al. 2004), including the CD9 mutants, were expressed in BL21 CodonPlus competent cells (Stratagene). These cells contain extra copies of the rare tRNA encoding genes that are frequently used by GC-rich organisms, the *argU* and *proL* genes. These genes encode tRNAs that recognise the arginine codons AGA and AGG and proline CCC, respectively. The BL21 CodonPlus cells are efficient at expressing tetraspanins CD9, CD81, CD63 and CD151 but no expression was detected for CD82 and CD231. Given the fact that these two proteins contain additional disulphides, it was presumed that the folding was too complicated for this type of expression system and so it was decided to use Origami cells (Novagen). Origami cells are specially engineered cells that have a lower cytoplasmic redox potential that is more favourable for the formation of disulphide bridges (discussed in more detail in Chapter 6). This strain worked for the expression of CD82 EC2 but CD231 remained undetected. It is possible that CD231 requires more rare codons than CD82 and therefore using an Origami cell line (or similar) that contains, in addition to the lower cytoplasmic redox potential, more copies of the rare tRNA genes may aid CD231 expression. Due to the low redox potential in specially engineered cells such as Origami cells, it is possible

that nascent proteins form disulphide bridges too readily with incorrect partners, leading to protein aggregation. If this were the case then the best solution would be to attempt CD231 EC2 expression using CodonPlus cells and a periplasmic leader sequence (discussed in more detail in Chapter 6).

The percentage yields of GST-EC2 produced vary between constructs but in general 40-60% of the total protein is GST-EC2. The remainder is GST alone (25kDal) and there is always another unknown band at around 30kDal that one would presume to be a degradation product. Ideally we would like to make samples that contain 100 % GST-EC2 with no GST or degradation products. However, the system we have set up enables us to make large amounts of recombinant protein in *E. coli* in a quick, cheap and efficient manner at a sacrifice of having the GST contaminants. So long as a GST control is used alongside all experiments with these recombinant proteins, the effects of the GST alone can be ruled out. As already mentioned, the GST-EC2 recombinant proteins have already proved to be useful tools for studying tetraspanin protein functions. There are a number of advantages of using these recombinant proteins over antibodies; binding of the antibody Fc region to Fc receptors on whole cells can stimulate a range of immune responses that could mask the desired effects of the mAbs (Jefferis et al. 2002; Radaev et al. 2002); mAbs are large molecules compared to GST-EC2 recombinant proteins (150kDal compared to ~ 40kDal) and will cause more steric hindrance at the cell surface which may also lead to inaccurate results; as the recombinant tetraspanins are produced in *E. coli* they are cheaper and quicker to produce than mAbs and no animal sacrifice is required (this latter point is of particular interest to the sponsors of this work, The Humane Research Trust). The

GST-EC2 proteins are extremely useful for studying binding interactions of tetraspanins. They can be used directly in ELISA experiments to determine whether or not they bind to particular ligands (e.g. HCV binding to GST CD81-EC2 proteins but not mutant proteins (Higginbottom et al. 2000) and Chapter 4). The EC2 recombinant proteins can also be used in whole cell assays, generally to inhibit a tetraspanin function, e.g. the sperm-egg fusion assays (Higginbottom et al. 2003), the inhibition of monocyte fusion (V. Parthasarathy, F. Martin, A. Higginbottom, G.W.Moseley, P.N.Monk, L.J. Partridge – unpublished data), and the inhibition of HIV-1 infection in macrophages (Chapter 7). The exogenous tetraspanins could be binding to molecules on the cell surface that the endogenous proteins would normally bind to, thereby disrupting the formation of the tetraspanin web. Alternatively, the EC2 domains could be forming homo- or hetero-dimers with endogenous tetraspanins, therefore preventing their proper functioning. In addition to functional studies, these GST-EC2 recombinant proteins can also be used in structural studies (Kitadokoro et al. 2001; Kitadokoro et al. 2001).

Although the recombinant proteins have their uses in studying tetraspanin proteins, one has to appreciate that there are limitations. Single domains of proteins are quite unlikely to function in the same way as intact proteins and it is likely that the other tetraspanin domains are required for the proper functioning of the EC2 domain. It has already been shown the EC1 plays a role in the expression of EC2 (Masciopinto et al. 2001). Viewing the EC2 proteins as peptide inhibitors or simply, useful tools to study the binding properties of EC2 domains and to inhibit the function of EC2 domains in whole cell assays is a more accurate concept.

CHAPTER 4

LIGAND BINDING BY TETRASPANIN EC2 DOMAINS

4.1 INTRODUCTION

As previously described in Chapters 1 and 3, recombinant tetraspanin EC2 domains can be used as tools to study the interaction of EC2 domains and their binding partners. This chapter describes the binding experiments using our GST-EC2 domains and two proteins that have been reported to bind to CD9, PSG17 (Waterhouse et al. 2002) and fibronectin (Longhurst et al. 2002).

4.1.1 AIMS

The aims of *this work were to use our recombinant GST-EC2 domains in ELISA* and FACS experiments to determine whether the interactions between CD9 and the reported ligands take place via the EC2 domain of CD9. In the event of CD9 binding to one or both ligands, it was also planned to investigate the structural requirement of CD9 in the interactions by using the CD9 EC2 mutants. Additionally, the specificity of the reported interactions for CD9 was to be investigated using the repertoire of tetraspanin EC2 proteins.

4.2 MATERIALS AND METHODS

Murine PSG17-Myc-His6 and control protein, murine PSG19-Myc-His6, were gifts from Gabriela S. Dveksler, Department of Pathology, Bethesda, US. PSG19-Myc-His6 had previously been shown not to bind to murine CD9 (personal communication with collaborator). Human plasma fibronectin was purchased from Invitrogen (F2006) and biotinylated human fibronectin was a gift from Simon Foster, Department of Molecular Biology and Biotechnology, University of Sheffield, UK.

4.3 RESULTS

4.3.1 Investigation into the binding between PSG17 and GST-CD9 EC2.

ELISA was used to investigate the binding between PSG17 and tetraspanin EC2 domains. It was previously reported that murine CD9 binds to PSG17 and due to the high sequence similarity between murine and human CD9, it was predicted that human CD9 EC2 should also bind to PSG17 since human CD9 EC2 inhibits mouse sperm/egg fusion to a similar extent as mouse CD9 EC2. To determine the concentration of tetraspanin EC2 proteins required to produce the best binding effect, mouse and human CD9 EC2 constructs were titrated against a fixed concentration (20 μg/ml) PSG17-Myc-His6, and bound PSG17 was detected using a rabbit anti-Myc antibody and an anti-rabbit IgG-HRP conjugated secondary antibody. Figure 4.1 shows the effects of concentration of tetraspanin EC2 constructs on the binding to PSG17.

From Figure 4.1, it is clear that maximum amount of binding of both mouse and human CD9 to PSG17 is seen at the highest concentration used, 20μg, and so this concentration of tetraspanin EC2 proteins was used throughout these experiments.

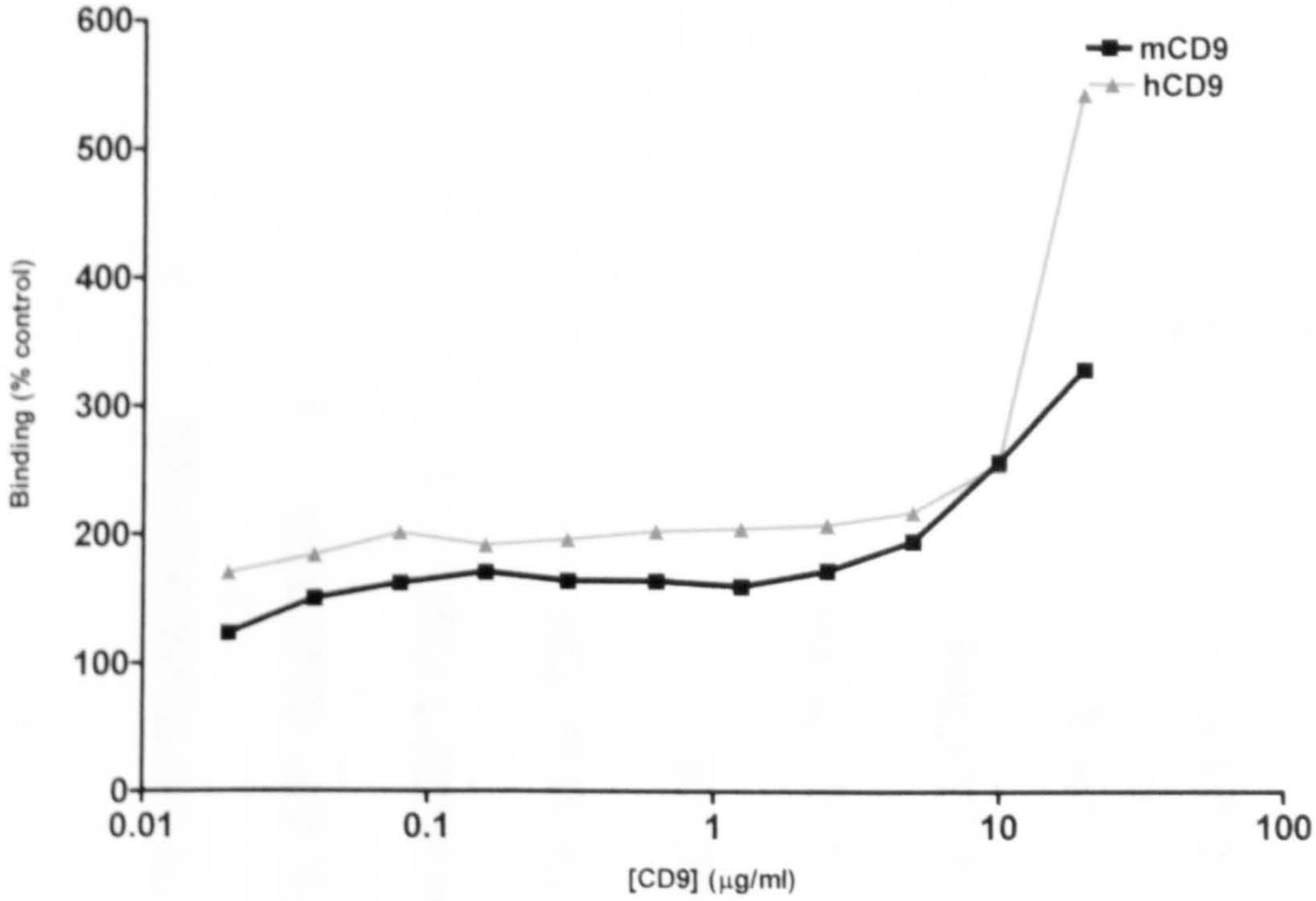


Figure 4.1. ELISA to determine the optimum concentration of tetraspanin EC2 domains to use in the binding assays. A range of concentrations of mouse and human CD9 (0-20μg/ml) diluted in binding buffer, were adhered to an ELISA plate overnight. After blocking in 5% milk in PBS-T, PSG17-Myc-His6 (at a fixed concentration of 20μg/ml) in blocking buffer was added. Following incubation and washing, binding was detected using an anti-Myc antibody (1/1000 dilution in blocking buffer) followed by anti-rabbit-HRP (1/3000 diluted in blocking buffer). Absorbance at 450nm was recorded on an ELISA plate reader. The graph shows the percentage increase in binding of mouse and human CD9 relative to the buffer control.

ELISAs were performed to determine whether the binding between human CD9 and PSG17 was significant, and to investigate whether the binding is specific for CD9 or whether other tetraspanins can also bind to PSG17. The results are presented in Figure 4.2 as percentage increase in binding compared to the buffer control from each experiment.

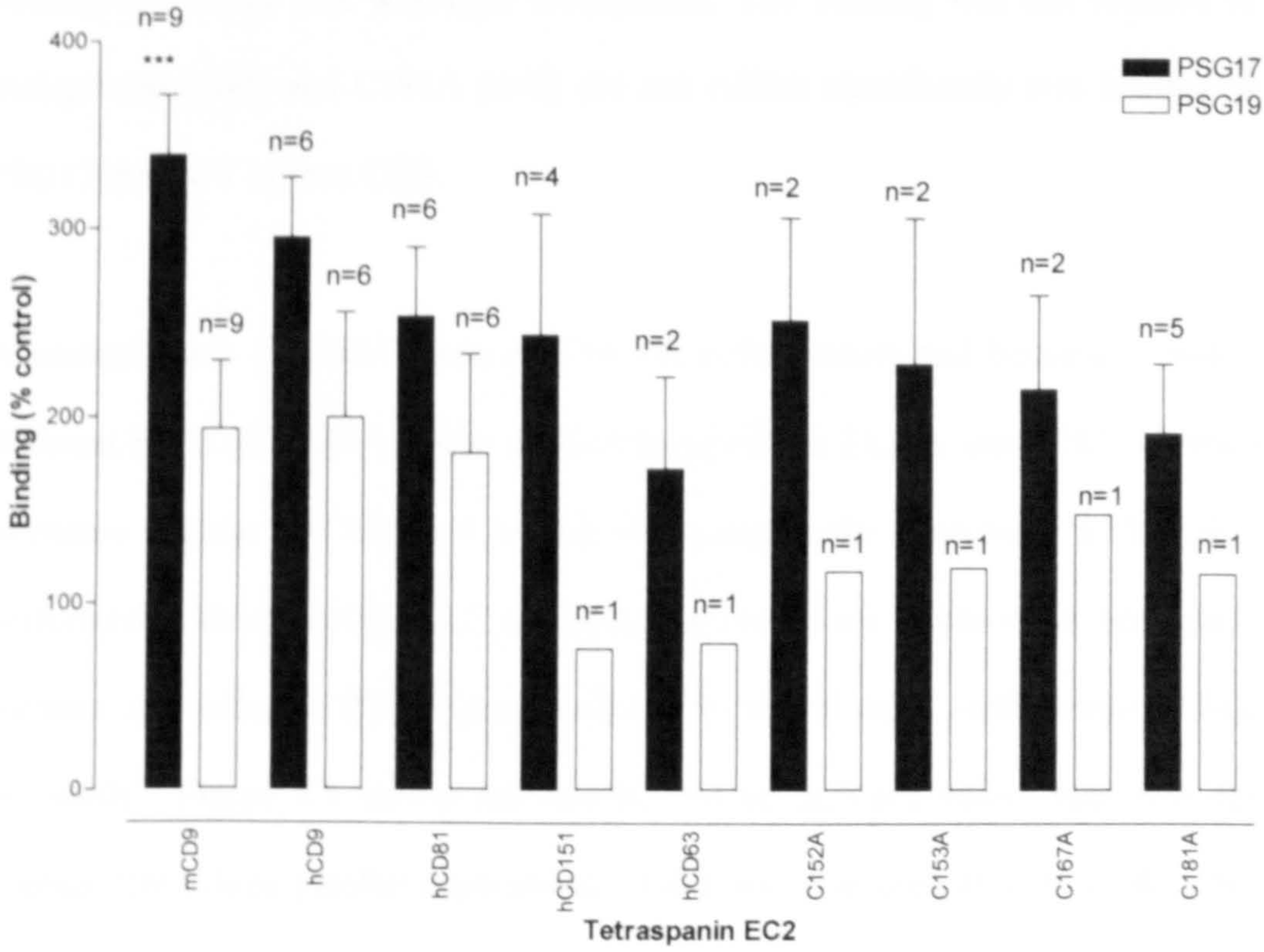


Figure 4.2. Bar chart showing the binding of PSG17 and PSG19 to a range of tetraspanin EC2 domains. Tetraspanin EC2 domains diluted in binding buffer, or buffer alone, were left to adhere to the surface of ELISA plates before blocking in PBS-T with 5 % milk and addition of PSG17-Myc-His6 or PSG19-Myc-His6 diluted to 20μg/ml in blocking buffer. After thorough washing, binding was detected by an anti-Myc antibody (1/1000 dilution in blocking buffer) followed by anti-rabbit-HRP (1/3000 dilution in blocking buffer). Absorbance at 450nm was recorded on an ELISA plate reader. Results are expressed as percentage increase in binding of PSG proteins to tetraspanins compared to the buffer control +/- SEM

All of the tetraspanin EC2 domains tested showed greater binding to PSG17 than the negative control, PSG19, and binding in all cases was greater than when buffer alone was substituted for tetraspanin EC2 domains. The trend in binding is as follows: mCD9>hCD9>hCD81>CD151. Statistical analysis between values

obtained for PSG17 and PSG19 revealed that only the binding of PSG17 to murine CD9 was significant. The human CD9 mutant constructs that lack the crucial cysteine residues involved in disulphide bridge formation had an apparent lesser binding to PSG17 that wild-type tetraspanins. The binding was not reduced to background levels and C181A (n=5) did not exhibit significantly less binding to PSG17 than WT human CD9.

A second batch of PSG17 was sent by the collaborators and because it was a different batch of PSG17, it was decided to repeat the ELISA using EC2 proteins to ensure that the new batch of protein was acting in the same way. ELISA was performed as described in 2.3.7, and again the results are displayed as percentage increase in binding to PSG17 compared to the control using buffer alone to coat the wells. Figure 4.3 shows the results, and again, each value represents the average from three parallel experiments. GST was also used and so results were compared to the background GST control values to check for any significance.

The new PSG17 protein showed a lower binding to the tetraspanin proteins than the previous batch (mean value for mCD9 binding to PSG17 previously was 339.6% compared to a value of 160.4% with the second batch of PSG17). In these experiments mCD9 (or the other tetraspanins tested), did not show significant binding to PSG17 when compared to the binding of GST alone.

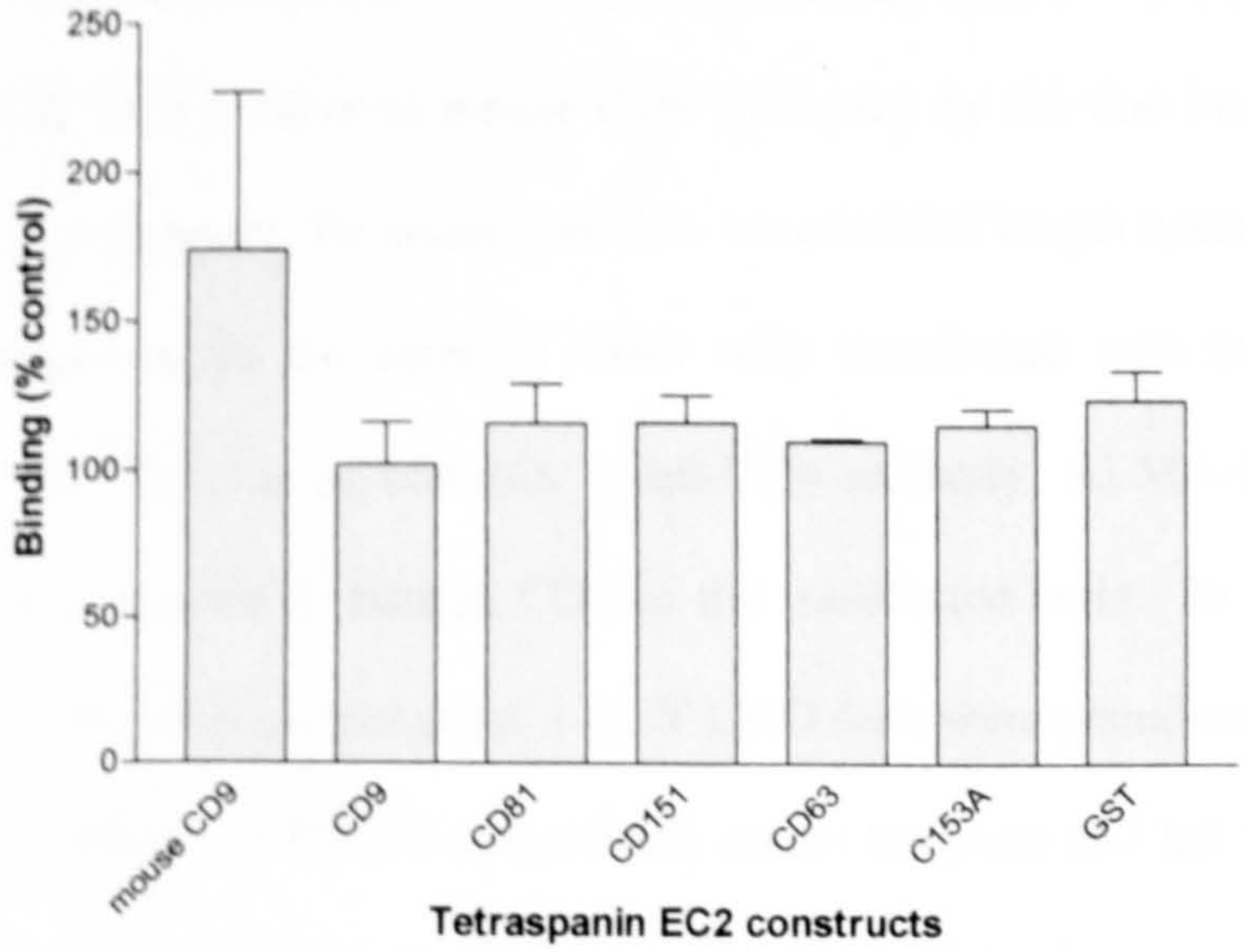


Figure 4.3. Bar chart showing the binding of a new batch of PSG17 to tetraspanin EC2 domains. 20μg/ml of tetraspanin EC2 domains diluted in binding buffer, or binding buffer alone, were left to adhere to the surface of ELISA plates before adding 20μg/ml PSG17-Myc-His6. After thorough washing and blocking in 5 % milk in PBS-T, binding was detected using anti-Myc antibody diluted 1/1000 in blocking buffer, followed by anti-rabbit-HRP diluted 1/3000 in blocking buffer. Absorbance at 450nm was recorded on an ELISA plate reader. Results are expressed as percentage increase in binding of PSG proteins to tetraspanins compared to the buffer control. (number of experiments =3, in triplicate).

4.3.2 Whole cell ELISA to investigate the binding between CD9 transfected CHO cells and PSG17

One possible explanation for the very low binding of both batches of PSG17 to human CD9 EC2 relative to mouse CD9 EC2 may be the that human CD9 EC2 does not fold properly. To investigate this, we used full length human CD9 instead of EC2 domains, in the form of CHO cells transfected with hCD9 and non-transfected CHO cells as controls. Anti-CD9 antibody, ALMA-1, was used to confirm the presence of human CD9 in the transfected cells. In each well of a sterile 96 well tissue culture plate, 1 x 10^4 CHO cells were plated and left to adhere overnight. PSG17 or PSG19 was titrated across the plate and left to bind for one hour followed by careful washing and detection of binding using an anti-Myc antibody. The results are displayed as: 1) the percentage increase in binding of PSG17 relative to PSG19 in CD9-transfected CHO cells and non-transfected cells (Figure 4.4); or 2) the percentage increase in binding of PSG17 and PSG19 to CD9-transfected cells relative to the binding to non-transfected cells (Figure 4.5). The results were analysed using an unpaired t-test but no significance was shown for either comparison.

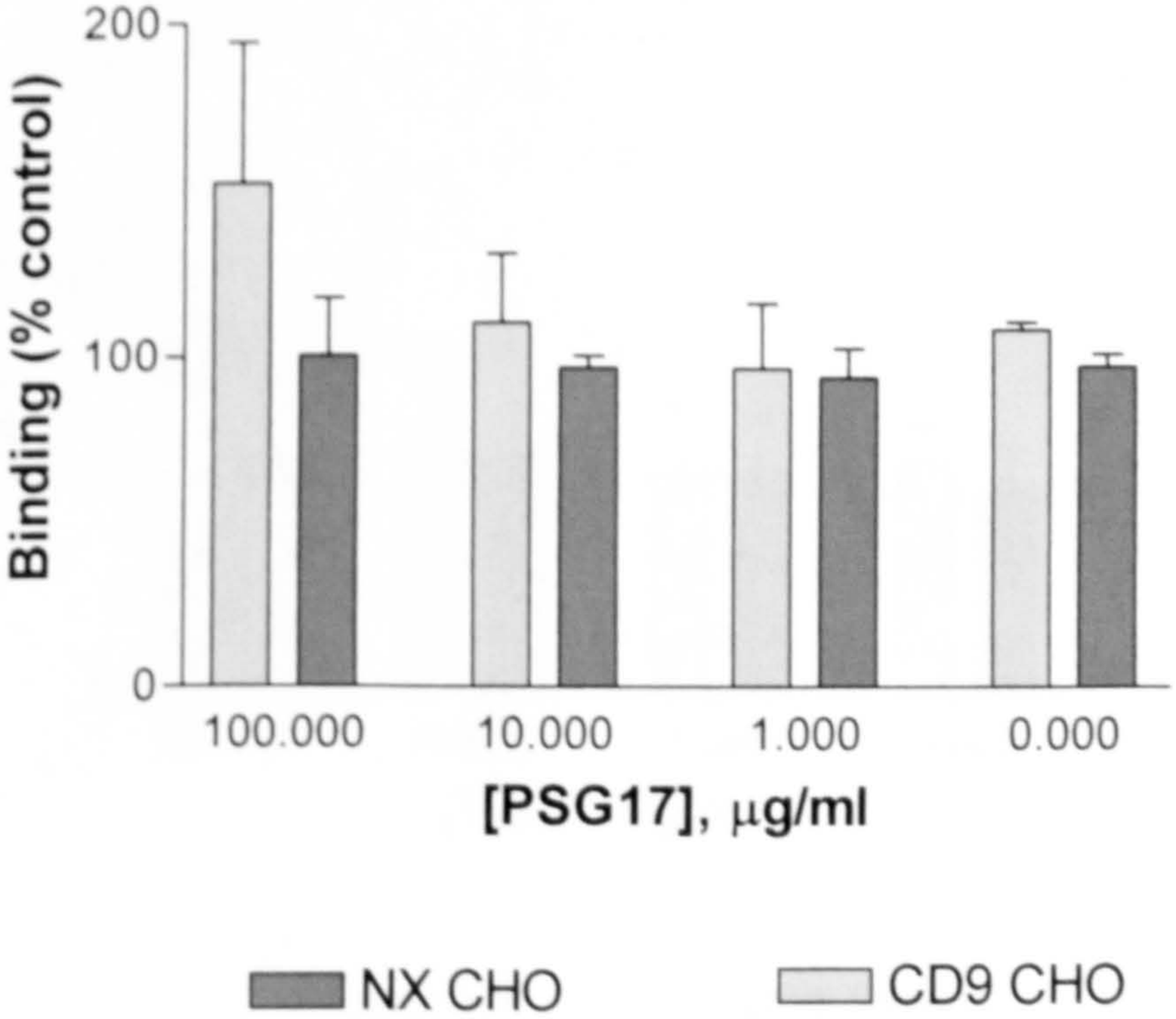


Figure 4.4. Bar chart to showing results from whole cell ELISA experiments investigating the binding of PSG17 relative to control protein in CD9-transfected CHO cells and non-transfected cells. CHO cells transfected with full length human CD9 or non-transfected cells were incubated with different concentrations of PSG17 or PSG19 protein (0-100μg/ml), diluted in HBSS and incubated on ice for 1 hour. After thorough washing, binding was detected using an anti-Myc antibody diluted 1/1000 in HBSS followed by anti-rabbit-HRP diluted 1/3000 in HBSS. Absorbance was recorded on an ELISA plate reader and the results are expressed in this bar chart as the percentage increase in binding of PSG17 relative to PSG19 in non-transfected (NX) and CD9 transfected CHO cells. (number of experiments=3, in triplicate).

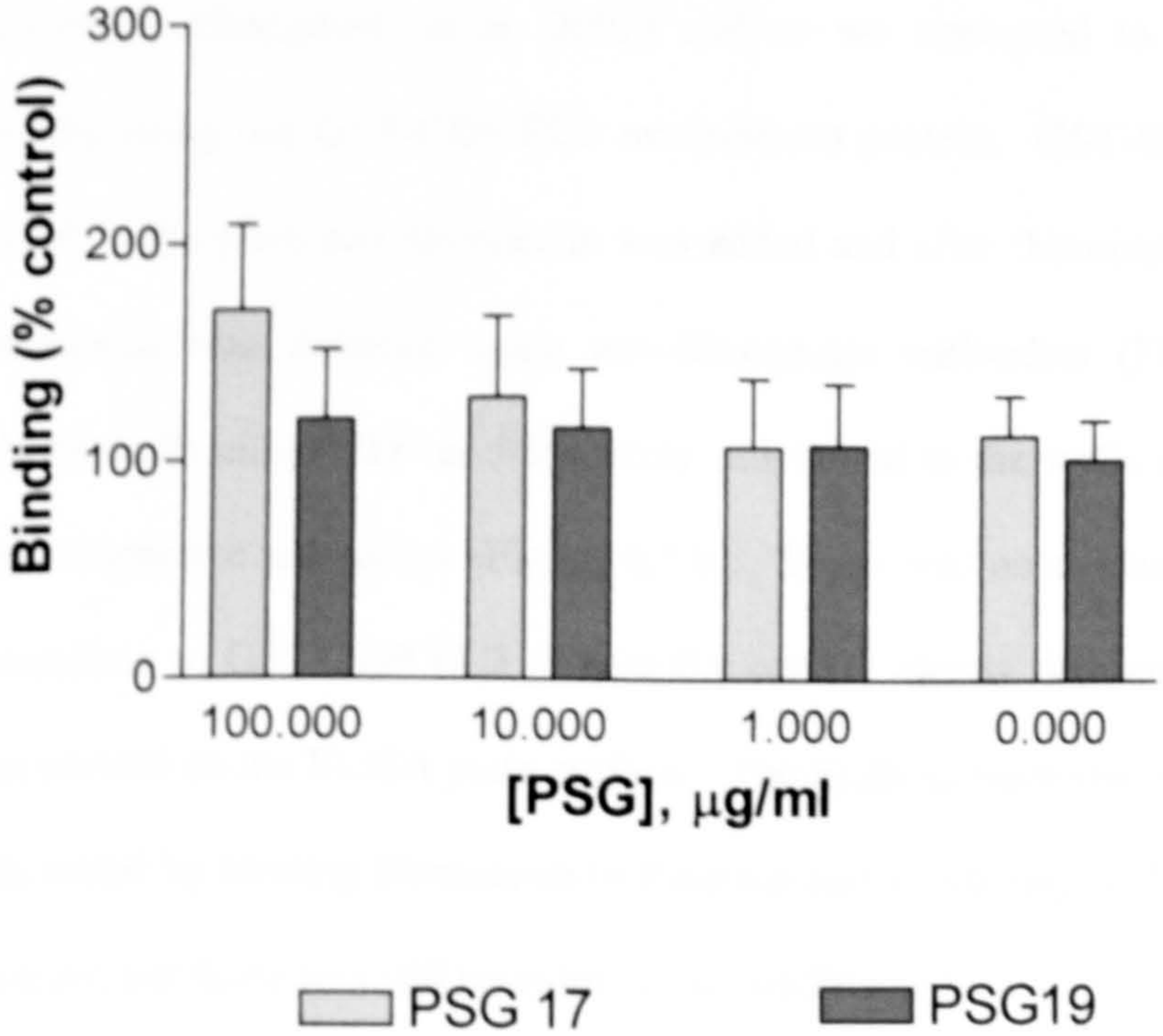
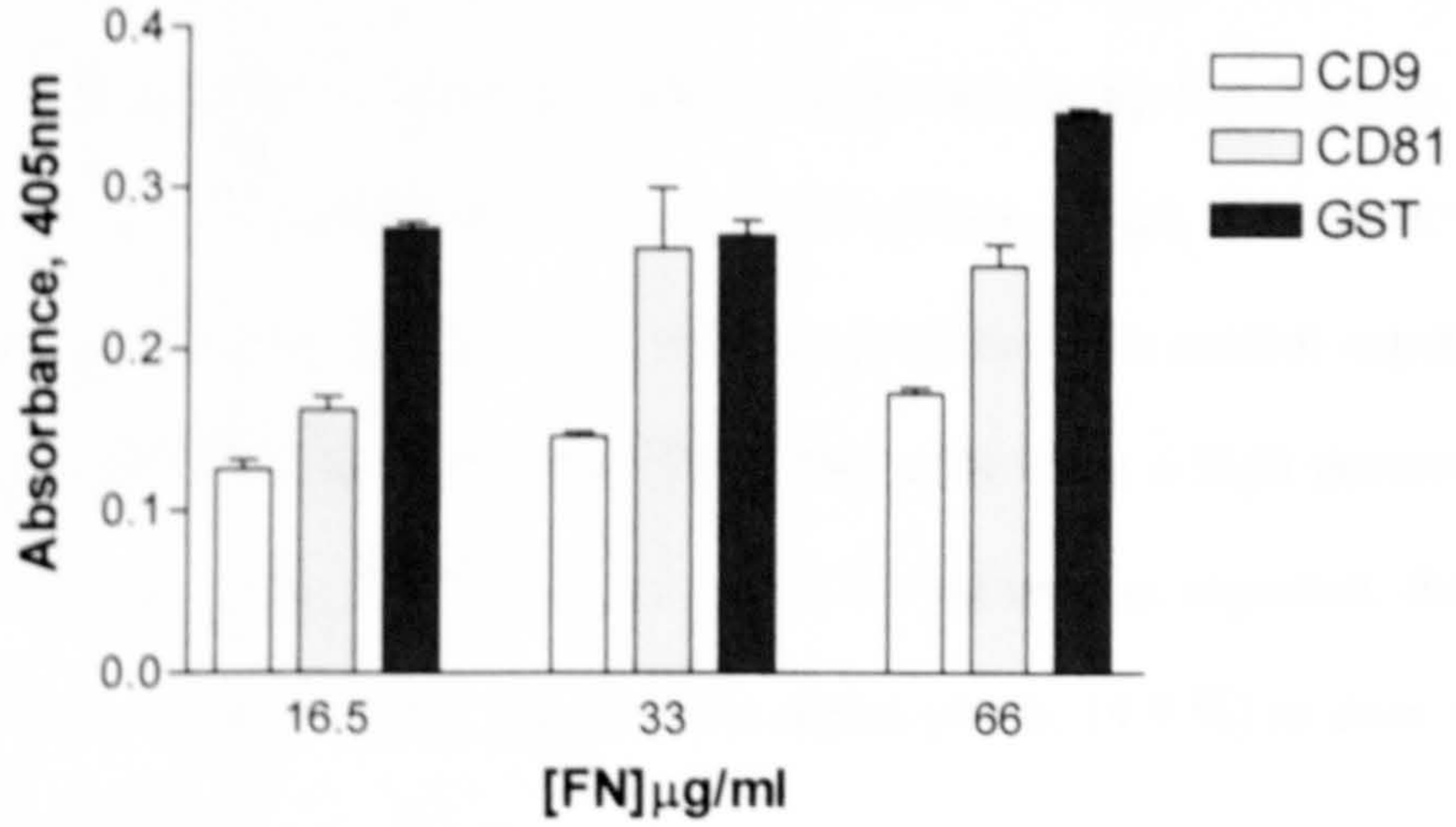


Figure 4.5. Bar chart showing the results from whole cell ELISA experiments investigating the binding of PSG17 and control protein to CD9-transefected and non-transfected cells. CHO cells transfected with full length human CD9 or non-transfected cells were incubated with differential concentrations of PSG17 or PSG19 control protein (0-100μg/ml) diluted in HBSS. After thorough washing, binding was detected using an anti-Myc antibody diluted to 1/1000 in HBSS followed by anti-rabbit-HRP diluted 1/3000 in HBSS. The results are expressed in this bar chart as the percentage increase in binding of PSG17 or PSG19 to CD9-transfected CHO cells relative to non-transfected (NX) CHO cells. (number of experiments = 3, in triplicate).

4.3.3 Investigation into the binding of CD9 EC2 and fibronectin by ELISA

An interaction between fibronectin and CD9 had been reported by Jennings and co-workers (Longhurst et al. 2002) and so we attempted to demonstrate this binding using our GST-CD9 EC2 recombinant protein. GST-EC2 proteins were bound to the plate and fibronectin was added and after thorough washing, bound fibronectin was detected using anti-fibronectin antibodies (Figure 4.7.a). In control wells either CD9 or fibronectin was bound to the wells and detected with their respective antibodies (Figure 4.7.b). There was no evidence of binding of fibronectin to GST-CD9 EC2 despite the control clearly demonstrating that CD9 was present on the ELISA plate surface. The ELISAs were also repeated the other way round by binding fibronectin to the plate and incubating with tetraspanin EC2 proteins, but there was still no evidence of binding.

A.



B.

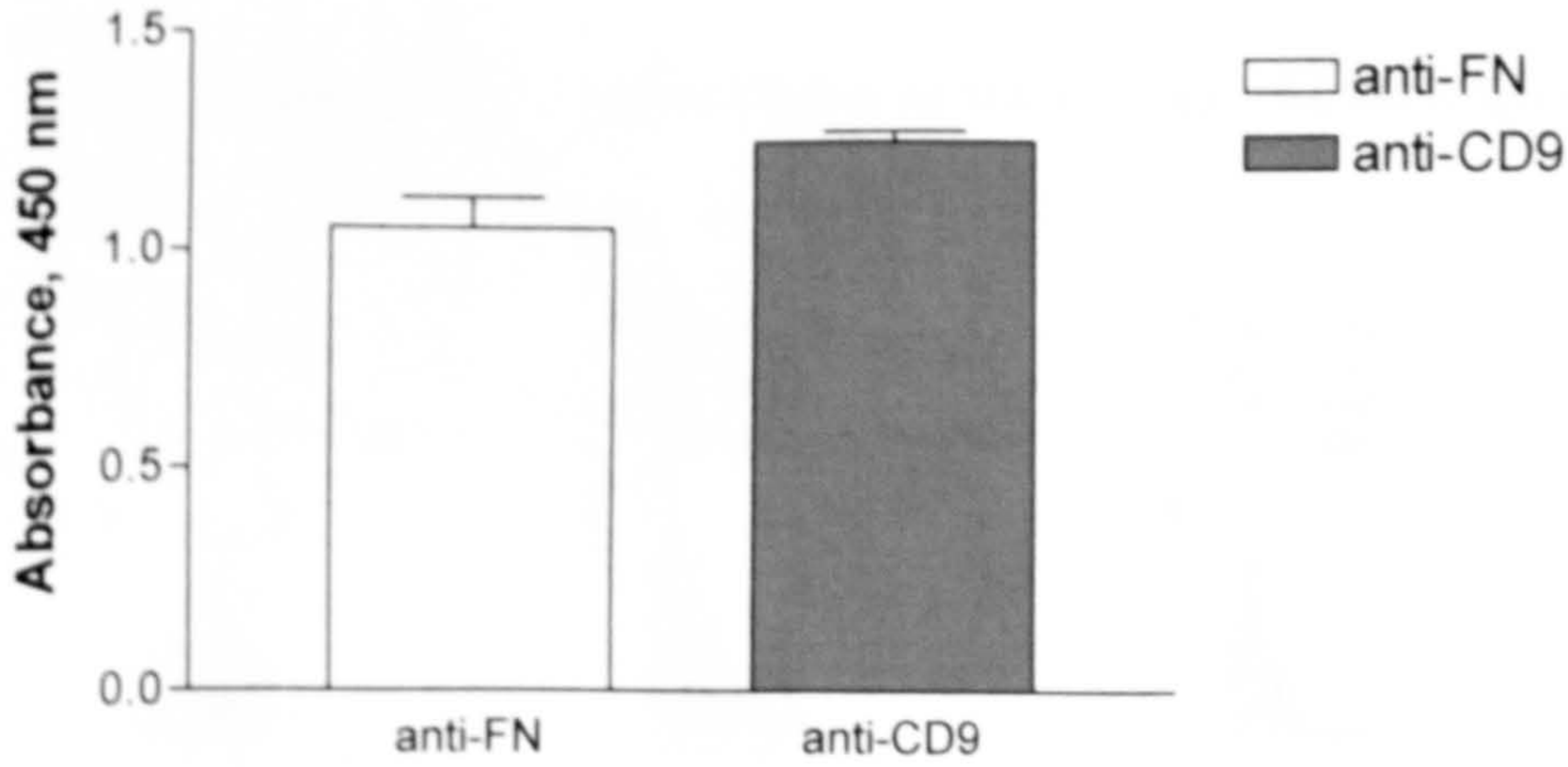


Figure 4.7. Bar charts showing ELISA analysis of fibronectin binding to GST-CD9 EC2. ELISAs were used to evaluate the binding of GST-EC2 CD9 to fibronectin. A) 10µg/ml of GST-CD9 EC2, GST-CD81 EC2 or GST alone were bound to the wells of an ELISA plate and after blocking in 5 % milk diluted in PBS-T, 16.5, 33 or 66µg/ml of FN were added. Binding was detected using an anti-FN antibody diluted to 1/1000 in blocking buffer and an anti-mouse-HRP secondary antibody diluted to 1/1000 in blocking buffer. Absorbance at 450nm was recorded on an ELISA plate reader B) GST-CD9 EC2 (10µg/ml) and FN (20µg/ml) were adhered to control wells. Anti-CD9 antibody (1/500 diluted in blocking buffer) was added to the CD9 wells and anti-FN antibody (1/1000 diluted in blocking buffer) was added to the FN wells, both followed by anti-mouse IgG-HRP (1/1000 in blocking buffer). Absorbance was recorded at 450nm on an ELISA plate reader. (number of experiments = 3, in triplicate).

4.3.4 FACS analysis of fibronectin binding to CD9 transfected CHO cells.

Chinese hamster ovary cells (CHO) transfected with full length human CD9 were tested for their ability to bind to biotinylated fibronectin by FACS assay. CD9-CHO and CHO cells transfected with a control protein, C5L2, were used. Figure 4.8 shows some examples of the fluorescence profiles of the control experiments. When probed with anti-CD9 mAb, CD9-CHO cells show a high percentage of cells in the gated fluorescence region (4.8.a, 97 %) and, as expected, there is a much lower percent of C5L2-CHO in this region (4.8.b, 19.9 %) as does an Ig-G control antibody (4.8.c, 6.7 %)

a. CD9-CHO &ALMA-1 **b.C5L2-CHO & ALMA-1** **c.CD9-CHO & IgG control**

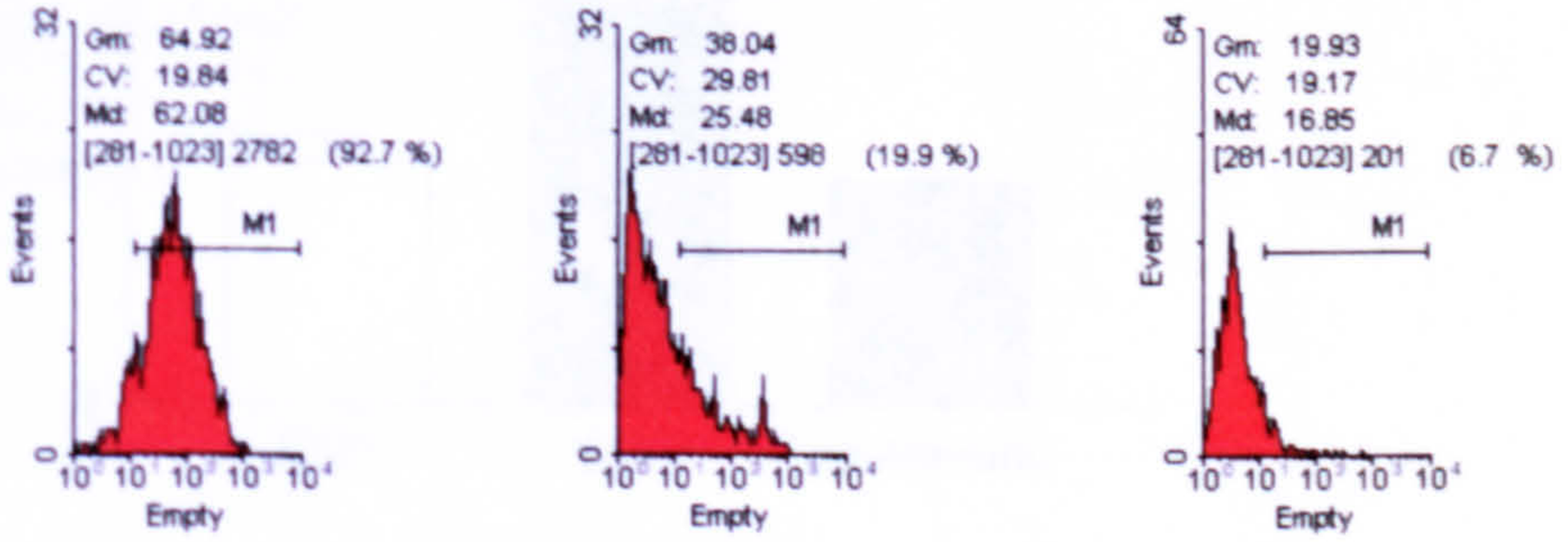


Figure 4.8. Fluorescent profiles of CD9 (a) and control transfected CHO cells (b) with anti-CD9 mAb, ALMA-1 and an IgG control antibody (c). Harvested CHO cells were incubated with ALMA-1 (1/500 in HBSS) or an IgG control antibody (1/500 in HBSS) for 1 hour on ice followed by washing and incubation with anti-mouse IgG-FITC (1/1000 in HBSS) for 1 hour on ice. Fluorescence was measured using the FL-1 Basic setting on a flow cytometer. 5000 cells per experiment were counted. The figure shows typical examples of the control experiments.

Figure 4.9 shows that there was no specific binding of fibronectin to CD9 transfected CHO cells compared to C5L2 transfected cells. Both populations of CHO cells were incubated with avidin-FITC to test whether there was a high background of avidin-FITC binding to C5L2 transfected cells and although avidin-FITC did appear to be "sticky", there was an average of 25 % avidin-FITC background for C5L2-CHO but a 40 % avidin-FITC background was observed for CD9-CHO. These results further highlighted that there was no specific binding between CD9-CHO and fibronectin in these assays.

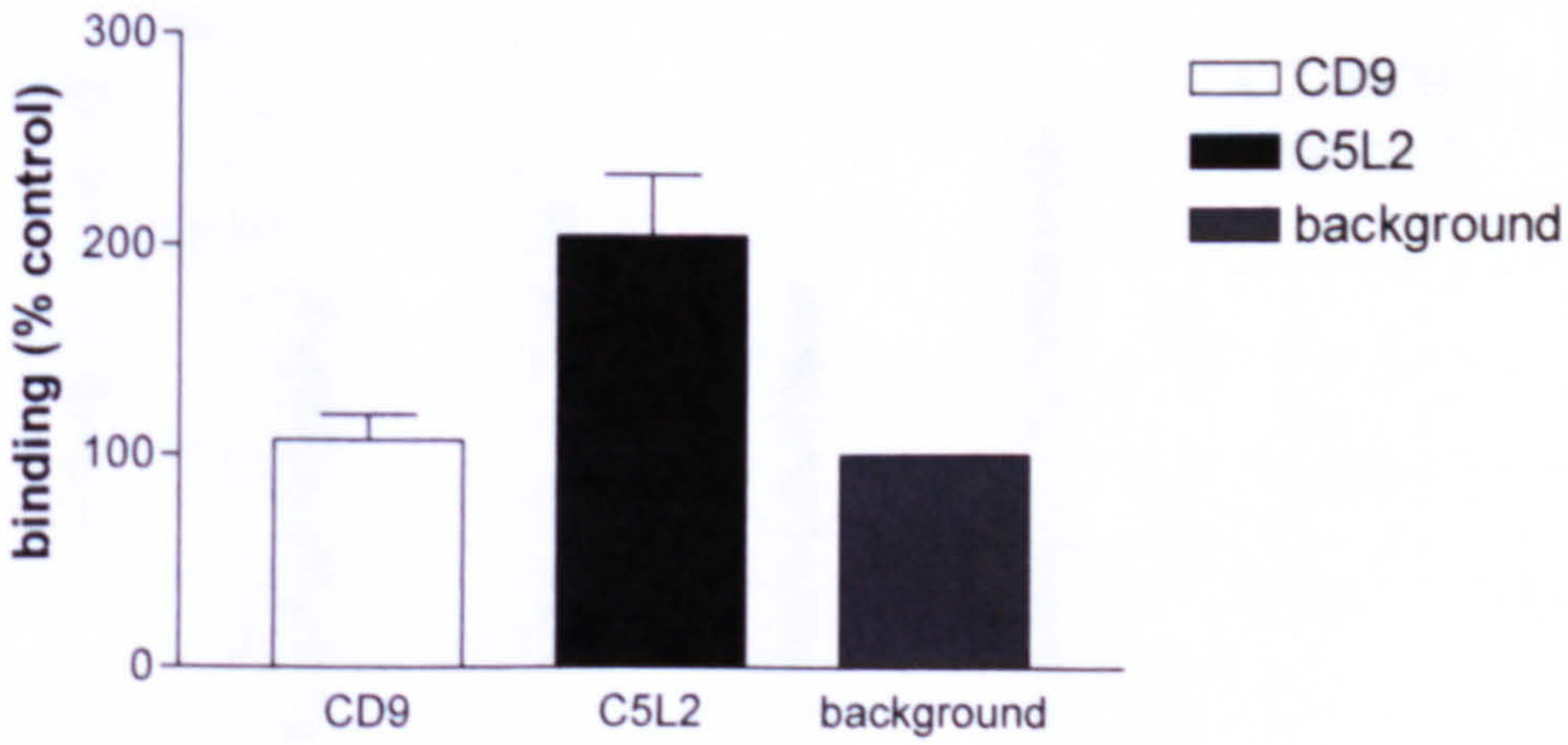


Figure 4.9. FACS assays to study binding of biotinylated fibronectin to full length CD9-transfected CHO cells. Harvested CHO cells transfected with CD9 were incubated with 50μg/ml biotinylated fibronectin in HBSS and incubated on ice for 1 hour. Binding was detected using avidin-FITC (1/1000 in HBSS). CHO cells transfected with C5L2 (a G-protein coupled receptor involved in the complement cascade) were used as controls and results are expressed as the percentage increase in binding relative to background (cells incubated with avidin-FITC only). Number of experiments = 2, in quadruplicate.

Further FACS experiments were performed using unlabelled fibronectin at concentrations ranging from 10-200μg/ml. This was done to try and avoid the high background binding of avidin-FITC to CHO cells. Following incubation with

fibronectin, cells were probed using a mouse anti-human fibronectin mAb and anti-mouse IgG-HRP conjugated secondary antibody. Control experiments where both populations of CHO cells were incubated with secondary antibody alone were included to establish non-specific, background fluorescence levels. Binding was calculated as the percentage change of cells in the fluorescent region compared to background fluorescence levels (taken as 100 %). Figure 4.10 shows that there was no increase in fibronectin binding to CD9-CHO relative to controls.

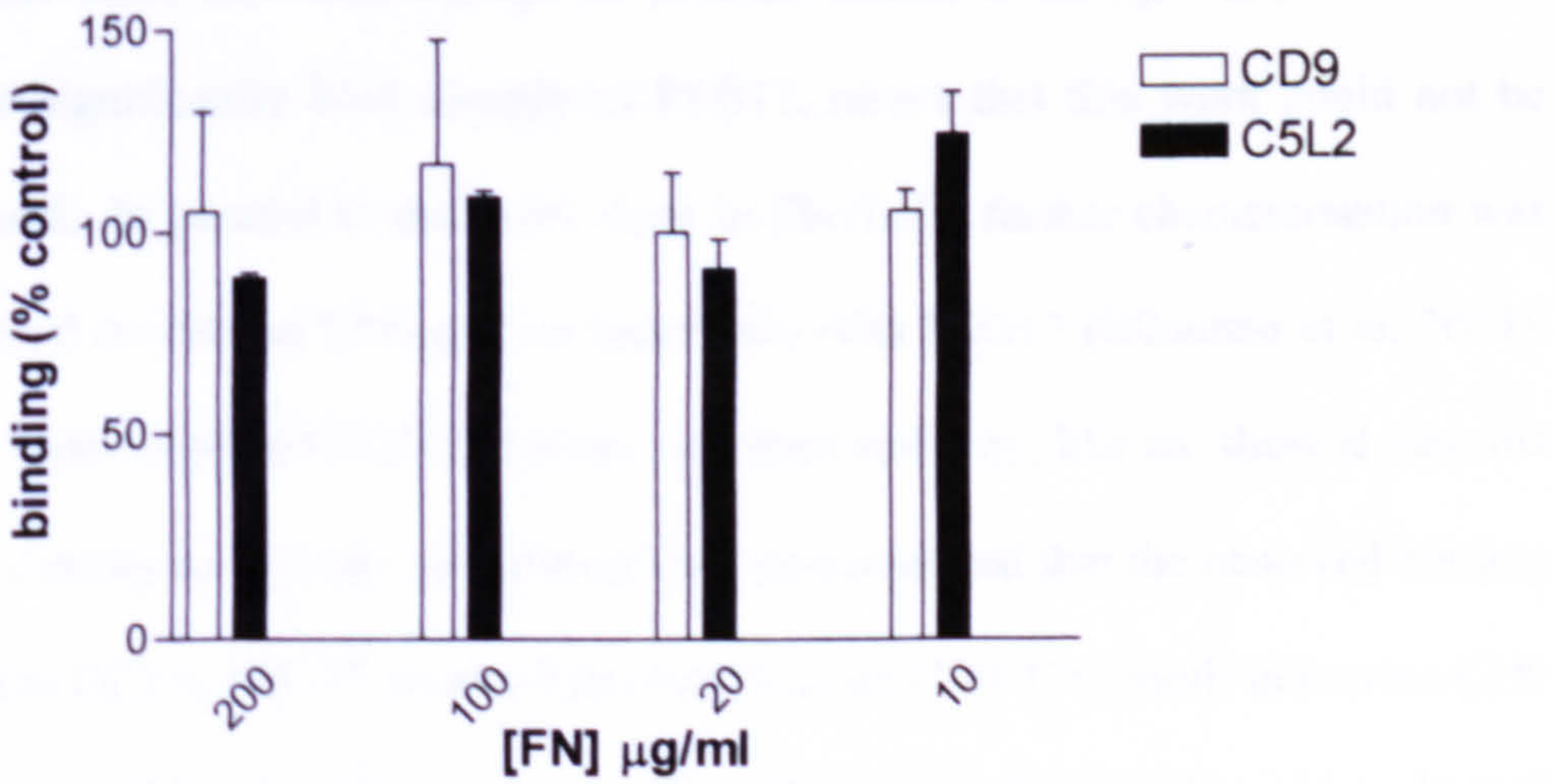


Figure 4.10. FACS assays to study binding of fibronectin to full length CD9-transfected CHO cells. FACS assays were performed as described in Figure 4.9 but using non-biotinylated fibronectin and at concentrations of 200–10 µg/ml. Background levels of secondary antibody alone were taken to be 100 % and the relative percentage binding of fibronectin was calculated for each experiment. (n=2, in quadruplicate).

4.4 DISCUSSION

Waterhouse et al. demonstrated that PSG17 binds to cells expressing murine CD9 but was unclear whether CD9 was acting as a direct receptor or as a co-receptor for PSG17 (Waterhouse et al. 2002). Using our GST-EC2 recombinant proteins, we set out to investigate the role of CD9 EC2 in this interaction. We demonstrated that PSG17-Myc-His6 (a murine CEA subfamily member) did indeed bind directly to our mouse CD9 EC2 recombinant protein ($P>0.05$), but there was no significant binding to the homologous human CD9 in our assays. The variation between different batches of PSG17-Myc-His6 provided by our collaborators together with the observation that other tetraspanin proteins tested, including human CD9-EC2, did not significantly bind directly to PSG17, meant that this work could not be continued. In parallel to this work done in Sheffield, further characterisation was performed on murine CD9 and its interaction with PSG17 (Ellerman et al. 2003). In this report murine GST-CD9 alone was used and they, like us, showed that this bound directly to PSG17. In addition, they demonstrated that the observed binding of mCD9 EC2 to PSG17 involved the Ser-Phe-Gln (173-175) motif in murine CD9 that is essential in sperm-egg fusion (Zhu, G. Z. et al. 2002). They also showed that pre-incubation of murine eggs with PSG17 before addition of sperm inhibited sperm-egg fusion, suggesting a role for PSG17 and therefore the CEA family in oocyte fusion (Ellerman et al. 2003). There was increased binding of human CD9, CD81 and CD151 EC2 to PSG17 compared to both GST and buffer controls, although these results were not significant. It is a possibility that human CD9 EC2 may not be properly folded relative to mouse EC2 and that the recognition sequence is masked or not formed thereby preventing binding between human

CD9 EC2 and PSG17. However, human CD9 EC2 shows potent activity in other systems e.g. CD9 EC2 inhibits monocyte giant cell formation from human monocytes but mouse CD9 EC2 is ineffective (Takeda et al. 2003). The fact that the same human CD9 EC2 is also functional in sperm-egg binding assays (Higginbottom et al. 2003) suggests that a different binding site is involved in this interaction. Many PSGs are predicted to bind to integrins because they contain the Arg-Gly-Asp (RGD) motif. Intriguingly, PSG17 does not contain the RGD motif. Therefore, it is a strong possibility that whereas many PSGs bind to integrins and are, perhaps, associated with tetraspanins via a TEM, murine PSG17, with the absence of an RGD motif, binds directly to murine CD9. There is no human ortholog of PSG17 and so it is possible that the direct involvement of CD9 in PSG functioning is a pathway present only in murine cells.

Although it has been reported that CD9 EC2 is responsible for the direct interaction with fibronectin (Cook et al. 1999; Longhurst et al. 1999; Cook et al. 2002), when we attempted to repeat these experiments, no binding was detected using our GST-EC2 CD9 recombinant proteins in ELISAs, nor could we detect binding between fibronectin and CHO cells transfected with CD9 in FACS experiments. Although we tried to replicate the experiments performed by Jennings and co-workers as closely as possible, it is feasible that slight differences in conditions and/or reagents used could account for the observed differences. Alternatively, it is possible that the results by Jennings and co-workers have been misinterpreted and fibronectin is actually associating with beta integrins and not CD9. The initial work by Jennings and co-workers was done using CHO cells that had been transfected with full length CD9 and mock transfected cells as a

control (see Chapter 1.3). Adhesion and spreading of CHO cells on FN was modulated in CD9 transfected cells compared to mock transfected cells, despite no apparent change in β1 and β3 integrins expression levels (Cook et al. 1999). However, CD9 could be functioning here by aiding the binding of integrins to ECM components by incorporating them into TEMs - a process that would not necessarily require an altered expression of integrin proteins, just a change in location. Indeed, it is possible that CD9 EC2 is involved in the process of binding to integrins and therefore in its absence, integrins are not integrated into the TEM and cannot, therefore, associate with fibronectin. This would certainly explain why we do not see binding between GST-CD9 EC2 and fibronectin in our ELISAs and perhaps differences in transfection levels of CD9 in our CHO cells may account for the contrast in our FACS results. Short peptides based on CD9 EC2 are reported to bind to fibronectin and partially inhibit the binding of full length CD9, although it is not stated in the original report whether these findings were significant (Longhurst et al. 2002).

CHAPTER 5

STRUCTURAL STUDIES OF CD9 AND CD63 EC2 DOMAINS

5.1 INTRODUCTION

5.1.1 X-ray crystallography

X-ray crystallography is the most common method used to solve the three-dimensional structures of proteins (Branden et al. 1999). It is a complicated procedure involving a large amount of mathematical manipulation but the main concept of X-ray crystallography is that the negatively charged electron clouds that orbit atoms have the ability to diffract X-rays. When a crystal is bombarded with X-rays, every atom in the molecule will individually diffract the X-rays and, in doing so, produce a spot on a film. The result is an electron diffraction pattern composed of diffraction spots from each atom within a molecule. Computer analysis and interpretation of the electron diffraction pattern reveals the positions of the electrons, and hence the atoms, within a molecule. Electron diffraction by single molecules is weak and so crystals are used to amplify the signals. Protein crystals contain many molecules precisely ordered and orientated in three dimensions. The smallest, unique part of a crystal is called the asymmetric unit and contains one or more copies of the protein. Rotation of one or more asymmetric units forms the unit cell, which is then translationally repeated to form the crystal lattice. Each unit cell diffracts X-rays identically and so the diffracted beams from all the unit cells within a crystal add together to produce strong, detectable X-ray beams (reviewed in (Branden et al. 1999; Rhodes 2000). However, the first major hurdle to be crossed before the 3D structure of a

macromolecule can be determined by X-ray diffraction from crystals, is to obtain a singular, perfect protein crystal which has dimensions of 0.2-0.4 mm in at least two of the three dimensions (Fig.5.1).

Figure 5.1 Typical observations in a crystallisation experiment. This figure is from Hampton Research (http://www.hamptonresearch.com/support). The figure shows the different types of observations one would expect to see in hanging drops after an initial screening experiment. Different crystal forms are shown ranging from a clear drop to different form of crystalline precipitate and eventually the ultimate goal, a single crystal.

Identifying favourable conditions for protein crystallisation to occur is a substantial task. There are many variables (pH, temperature, organic solvent etc. etc.) and the exact conditions required for crystalisation differ from protein to protein. In fact, protein crystallisation is considered by many scientists to be more

of an art than a science! To aid initial crystallisation experiments, a number of crystal screens are available. These screens generally contain around 50 formulations with varied precipitating agents, pH and divalent cations, that were initially selected from incomplete factorial experiments, where variables were tested at random for their ability to allow protein crystallisation (Carter et al. 1979; Carter et al. 1988). There are now a number of these kits commercially available that are excellent places to start when commencing protein crystallisation trials.

5.1.2 NMR

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful method for studying the biological properties of macromolecules (Branden et al. 1999). It is a complex set of techniques based on the properties of subatomic particles, namely protons and neutrons. Protons and neutrons possess a quantum property called spin. In many atoms that have even numbers of protons and neutrons (^{12}C), the spins from protons and neutrons pair together and there is no overall spin. However, if the number of protons or the number of neutrons is odd, then the nucleus possesses spin. The most common, naturally occurring atomic isotopes that possess spin are ^{1}H and ^{31}P; less abundant isotopes with a nuclear spin (^{13}C, ^{15}N) can be incorporated into a biomolecule by isotope labelling. In the presence of a magnetic field, the spins of the nuclei align in two or more different energy levels. The frequency of the radio waves required to produce the spin (or resonance) depends on the chemical environment of the atoms i.e. other atoms that are connected. As soon as the radio frequency excitation is stopped, the nuclei return to lower energy levels in a process called relaxation, and, in doing so,

transmit a radio frequency resonance response that is picked up by the NMR spectrometer. This response is then amplified and used to produce a spectrum that can be interpreted to determine information about the chemical environments of atoms within a molecule. In the case of 1D 1H NMR, the different resonance frequencies picked up in the spectrum represent the different chemical environments of the 1H atoms, and are known as chemical shifts. 1D 1H NMR is commonly used to determine whether a protein is folded. The 1H frequencies of all residues in unfolded proteins are very similar and the spectrum is recognised as having little dispersion. If there is some folding, different chemical environments of 1H atoms can be identified. Like crystallography, a highly concentrated and pure sample produces better results but crystals are not required, and a 1D spectrum to evaluate whether or not a protein has any structure can be performed in an hour (Branden et al. 1999).

5.1.3. AIMS

In 2001 the crystal structure of CD81 EC2 was solved (Kitadokoro et al. 2001). The EC2 domains of tetraspanins are variable, and structural information of another EC2 containing additional cysteines that are predicted to form disulphide bridges, was considered advantageous. The main aim of this work, therefore, was to attempt to determine the crystal structure of CD63 EC2 in order to gain a better insight to the function of CD63, especially its role in HIV infection (Chapter 7). CD63 EC2 contains a total of 6 cysteines that are predicted to form 3 disulphide bridges. Due to problems encountered when attempting to concentrate the cleaved CD63 EC2 (5.3.1), a collaboration with Dr Christopher Liu was established. This

collaborator provided lyophilised His6-CD9 EC2 and was willing to make some CD63-His6 EC2. Although the structure of CD9 EC2 was not one we originally planned to attempt to solve, since the protein was available in large quantities, we decided to have a preliminary attempt whilst awaiting the CD63 EC2 protein.

5.2 MATERIALS AND METHODS

5.2.1 GST cleavage from CD63 EC2

In addition to the reagents listed in Chapter 2, thrombin protease (Amersham Pharmacia Biotech, USA. Product number 27-0846-01) was also used in the work described in this chapter.

5.2.1.1 Small scale batch cleavage of GST from CD63 EC2

Siliconised 1.5ml tubes were used throughout these experiments to prevent the hydrophobic, cleaved CD63 EC2 from being absorbed onto the plastic. A known quantity of purified GST-CD63 EC2 was allowed to bind to glutathione Sepharose by incubation for 1 hour at RT. The beads were saturated with GST-CD63 EC2 to prevent non-specific binding of the cleaved CD63 EC2 (1ml of glutathione Sepharose beads binds 10mg of GST fusion protein). The glutathione Sepharose, with GST-CD63 EC2 attached, was pelleted by centrifugation (all centrifugation steps were 500 g for 2 minutes) and washed twice with 200ml of PBS. Thrombin protease at a concentration of 1 unit/µl was added to a final concentration of 10 or 1 unit(s) per mg of fusion protein, and the volume was made up in 200–500µl of PBS. The reaction was gently agitated at RT for 2 hours, 4 hours, 6 hours or overnight. Glutathione beads were removed by centrifugation. The supernatant, now containing CD62 EC2 (plus a negligible amount of thrombin), was removed into a fresh, siliconised tube. 5µl samples were taken at each step to monitor the

cleavage and purification process by Coomassie staining of SDS-PAGE gels and Western blotting.

5.2.2 Setting Crystals

Crystal screening kits were from Hampton Research, CA; Crystal Screen (50 conditions), Crystal Screen II (48), PEG ION (48), or from Nextal Biotech; the SM1 (96), the Classics (96), the pHClear (96), the PEGs (96), the $AmSO_4$ (96) and the MDPs (96). The Robot used to set crystals in 96 well plates was from Nextal Biotech.

5.2.3. Hanging drop vapour diffusion method for protein crystallography

To each well of a 24 well-plate, 500–700µl of the formulation was added. 1µl of protein was pipetted onto the surface of a siliconised coverslip and mixed together with 1µl of the formulation from the well. The rim of the well was coated with sealant (oil) and the coverslip was inverted over the well with the drop hanging over the pool of formulation (Fig 5.1). To ensure sealing, the coverslip was moved around a little in the oil. The drops were examined under a microscope every 2-4 days for the presence of crystals.

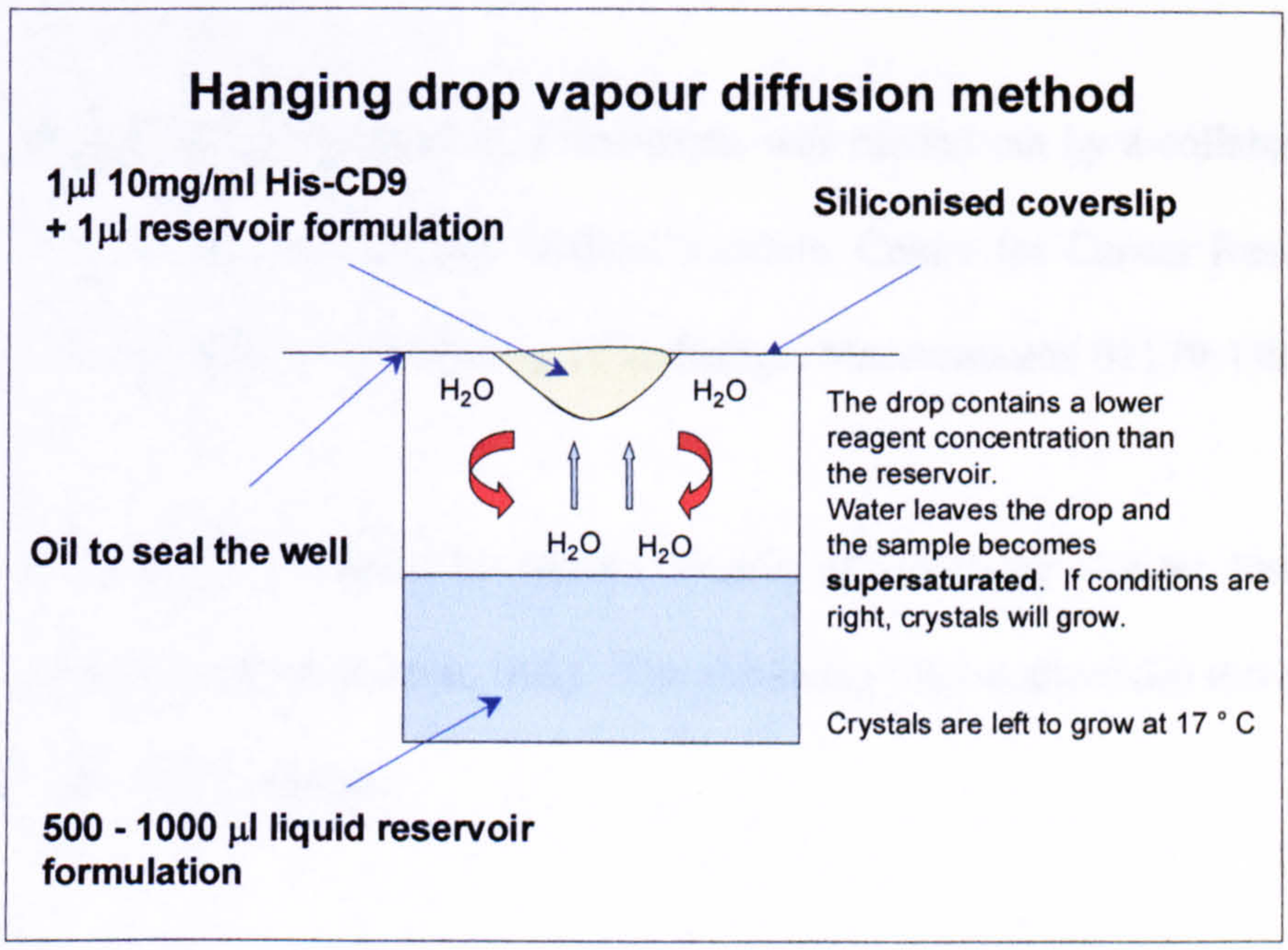


Figure 5.1. Hanging drop vapour diffusion method for protein crystallisation. This was the predominant method for crystal growth that was used. 24 well plates were used to set the crystals. 500–700µl per well of a commercial screening formulation, or a follow up trial formulation, was added. On a siliconised coverslip, 1µl of the same formulation was mixed with 1µl of the protein solution. The coverslip was inverted so that the drop was left suspended above the well. Oil was used to create and air-tight seal between the coverslip and the well.

5.2.4. Automated crystallisation screening

Automated crystallisation screening was performed (NeXtal suite) using 0.33µl of protein solution and 0.33µl of formulation from the reservoir in the well. This drop sits on a raised ledge in the well rather than hanging above it. Once the robot had set up each of the 96 wells in this manner, a plastic sheath was pressed over the plate to seal the wells.

5.2.5. Production of His6-tagged CD9 and CD63 EC2

[The production of His6-tagged EC2 constructs was carried out by a collaborator, Christopher Liu (Howard Hughes Medical Institute, Centre for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA)].

cDNAs were kindly provided by Martin Hemler (Dana-Faber Cancer Institute, Harvard Medical School, Boston, MA). The following oligonucleotides were used to amplify the EC2 domains:

Construct	Oligo (5' – 3')
CD9 F	CGCTAGCTAGCTCCCACAAGGATGAGGTGATTAA
CD9 R	CCGCTCGAGGTGGAATTTATTGTCGAAGACCTC
CD63 F	CGCTAGCTAGCAGAGATAAGGTGATGTCAGAGTTTAATAAC
CD63 R	CCGCTCGAGCTCATTTTTCCTCAGTTAGCCC

The EC2 domains were cloned into the *Nhe*I and *Xho*I restriction sites of the pET24a (Novagen) vector and sequenced to check for errors. BL21/pLysS competent cells (Novagen) were used for protein expression. The recombinant proteins were purified from inclusion bodies by guanadinium/urea extraction followed by partial purification by Ni-NTA chromatography (Qiagen). The solubilised protein was analysed by reverse-phase C18 HPLC with protein content monitored as absorbance at 229nm. HPLC protein peaks were collected and analysed for tetraspanin content by dot-blot. Tetraspanin protein was lyophilised and stored at -20°C before being sent to the UK.

5.2.6. NMR

It was decided to perform 1D 1H NMR on the CD63-His6 EC2 to ensure that the protein was folded. CD63-His6 EC2 was diluted to 4mg/ml in dH_2O and loaded into a Bruker DRX500. The 1D proton NMR spectrum was run with presaturation of the water and a hahn echo (2.6m total echo time) to improve convolution difference, 1second recycle time, 128 scans. The spectrum was processed with a convolution difference to further remove the water signal, and 5Hz line broadening.

5.3. RESULTS

5.3.1 Batch cleavage of GST from CD63 EC2

An initial experiment was performed on small quantities of GST-CD63 EC2 to determine the amount of thrombin and the length of time required to cleave GST from GST-CD63. 1 unit of thrombin should cut > 90 % of 100μg of a GST-fusion protein in PBS at 22°C for 16 hours (according to manufacturers' instructions). Optimum conditions for the cleavage of GST from CD63 were determined (Table 5.1 and Figure 5.2).

Lane	[thrombin] Units	Time	Buffer
1	0.1	2	PBS
2	0.1	4	PBS
3	0.1	6	PBS
4	0.1	O/N	PBS
5	1.0	2	PBS
6	1.0	4	PBS
7	1.0	6	PBS
8	1.0	O/N	PBS
9	0.1	6	Tris
10	1.0	6	Tris
11	0.0	6	PBS
12	0.0	6	Tris
13	0.0	-	PBS

Table 5.1. Optimisation of conditions to cut GST from GST-CD63 EC2 using thrombin. To determine optimum conditions, 100μg aliquots of GST-CD63 EC2 were cut with 1.0U (10U/mg) and 0.1 U of thrombin protease. The reactions were stopped (by freezing the samples) after 2 hours, 4 hours, 6 hours or overnight incubation at RT. GST-CD63 EC2 was typically dialysed into PBS, pH 7.4. To test whether this buffer was suitable for proteolysis to occur, a control of GST-CD63 EC2 in Tris, pH 8.0, was also included. The samples were made up to 100μl in the appropriate buffer.

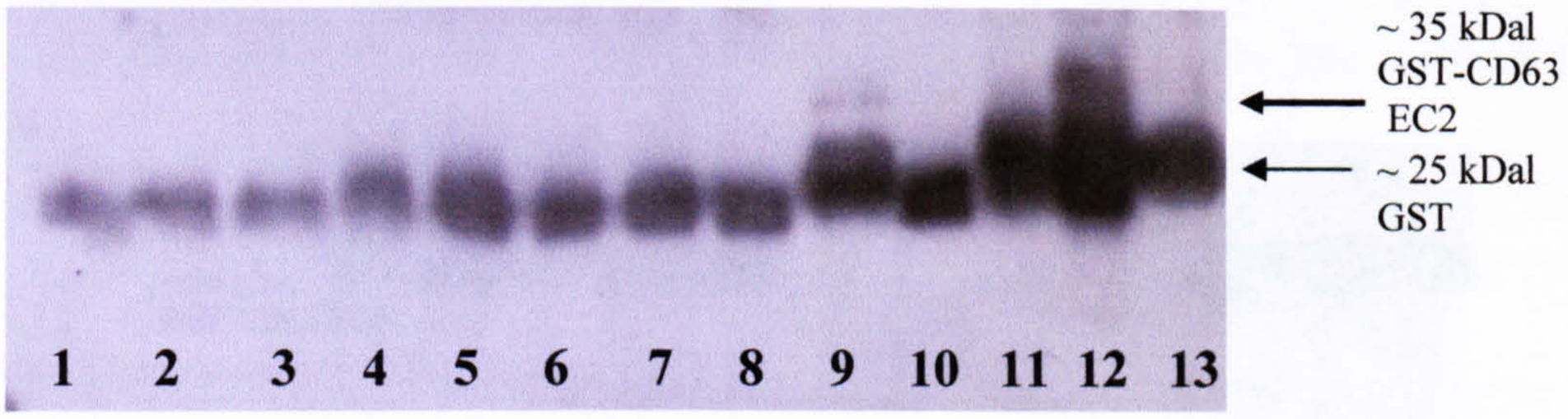
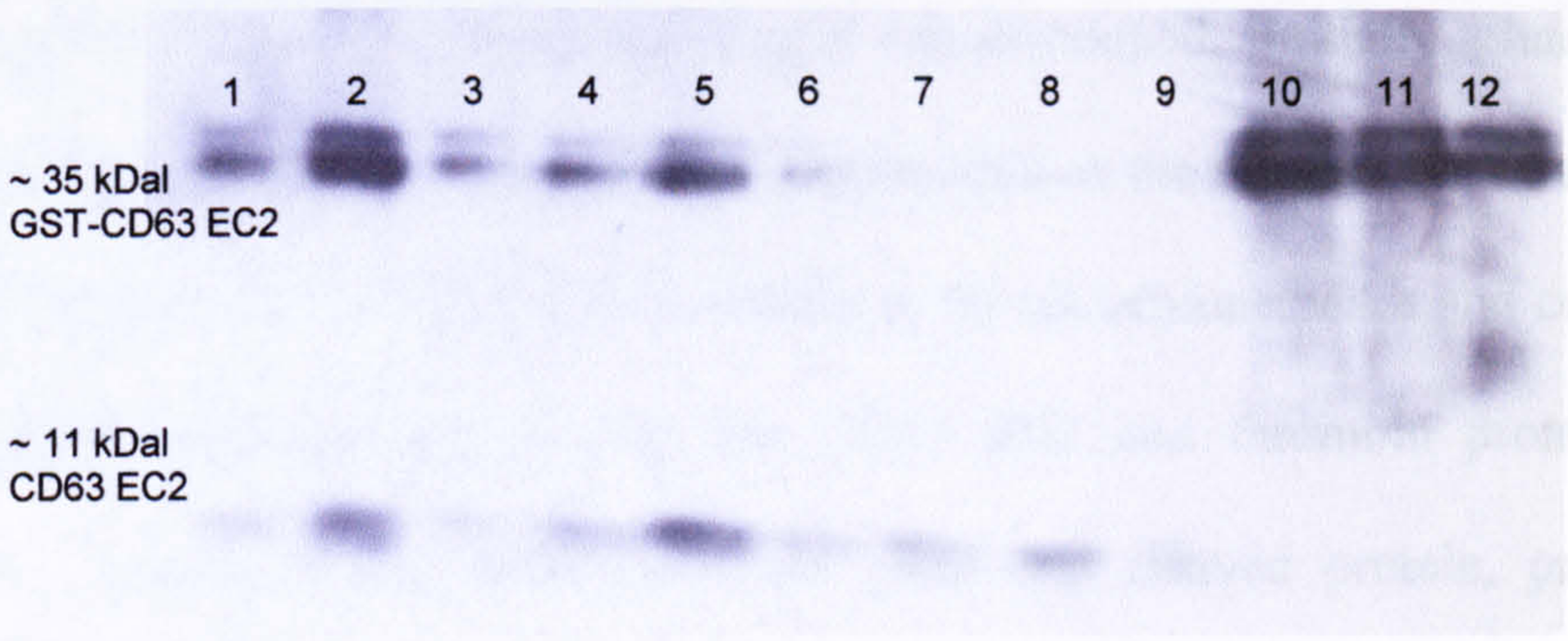


Figure 5.2. Western blot of samples from optimisation of GST-CD63 thrombin cleavage. 10µl of the samples corresponding to those in Table 5.1 were run on SDS-PAGE, blotted and blocked in 5 % milk in PBS-T and then probed with an anti-GST antibody, 1/1000 dilution in blocking buffer. Anti-mouse IgG-HRP was added (1/3000 diluted in blocking buffer) and the blot was developed using ECL reagent (Amersham). Lanes 11 and 12 show uncut GST-CD63 EC2 and lane 13 is GST alone. B.

The Western blot using anti-GST antibody shows 2 bands present in the uncut protein, the top band being GST-CD63 and the lower GST alone. All of the samples with thrombin present (lanes 1–10) show one band, GST alone, suggesting that the reaction has worked in all cases. The samples in PBS (lanes 1-8) appear to have cut more efficiently than the samples in Tris, and there does not appear to be a preference for 1.0 or 0.1 units of thrombin.

Different amounts of GST-CD63 EC2 were cut with thrombin protease to test the effectiveness of this method of thrombin cleavage on higher concentrations of protein (Figure 5.3).



Lane	1	2	3	4	5	6	7	8	9	10	11	12
Amount of protein (μg)	200	400	85	200	400	85	200	400	85	200	400	85
Reaction time (hours)	2	2	2	4	4	4	16	16	16	uncut	uncut	uncut

Figure 5.3. Western blot of samples from the second optimisation of GST-CD63 thrombin cleavage experiment. 0.1 U of thrombin was added to different concentrations of GST-CD63 (as described in the table) and the volume was made up to 100μl and left for 2 hours, 4 hours or 16 hours before stopping the reaction by freezing the samples. 20μl of each sample was run on SDS-PAGE and blotted. Following blocking in 5 % milk in PBS-T, anti-CD63 antibody, H5C6 (at 1/1000 dilution in blocking buffer) was added. After washing, anti-mouse IgG-HRP was added at 1/3000 dilution in PBS-T and the blot was developed in ECL reagent (Amersham).

As shown by the blot in Figure 5.3, 0.1 U of thrombin can cut up to 400μg of GST-CD63 EC2 and so this is the concentration used from now on. When the reaction was left for 16 hours, there was a clear reduction in the amount of CD63 EC2, probably due to protein degradation. However, at 4 hours there was still a substantial amount of uncut CD63 and so the reaction required at least 4 hours incubation. Interestingly, the Western blot also shows two bands at around the 35kDal mark. Perhaps one is a dimer or even a trimer of the 12kDal CD63 EC2 and the other is GST-CD63 EC2.

After the optimal cleavage conditions had been determined, the method for removing the cleaved GST from the sample was developed. Initially, glutathione Sepharose was added to the sample at the same time as the thrombin, as free GST and GST-CD63 EC2 would bind specifically to the glutathione beads and could be removed by centrifugation, leaving the CD63 EC2 and thrombin protease in solution. However, this method did not yield any cleaved protein, probably because the cut CD63 EC2 was binding non-specifically to the Sepharose beads. Instead, it was decided to first bind the intact GST-CD63 EC2 to the beads for one hour before adding the thrombin. Also, the amount of glutathione Sepharose used was carefully calculated so that it was completely saturated by the GST-CD63 EC2 protein, to reduce the non-specific attachment of cleaved CD63 EC2. The binding capacity of glutathione Sepharose beads is 10mg GST tagged protein per ml of slurry (according to the manufacturer), therefore, 100μg of GST-CD63 EC2 requires 10μl of slurry. After mixing the GST-CD63 with the beads for 1 hour followed by the addition of thrombin for 4 hours, the free CD63 EC2 was collected by centrifugation. The beads were then washed once in PBS. Samples of the supernatant, wash buffer and glutathione beads were evaluated by Western blotting and silver staining SDS-PAGE gels. The Western blot revealed that the majority of GST-CD63 EC2 had been cleaved using this method; there was very little CD63 attached to the beads or in the PBS wash (Figure 5.4). The silver stained gel also revealed that the sample was extremely pure, as no other protein bands were detected. The protein concentration of cleaved CD63 was 0.2mg/ml as determined by Bradford assay.

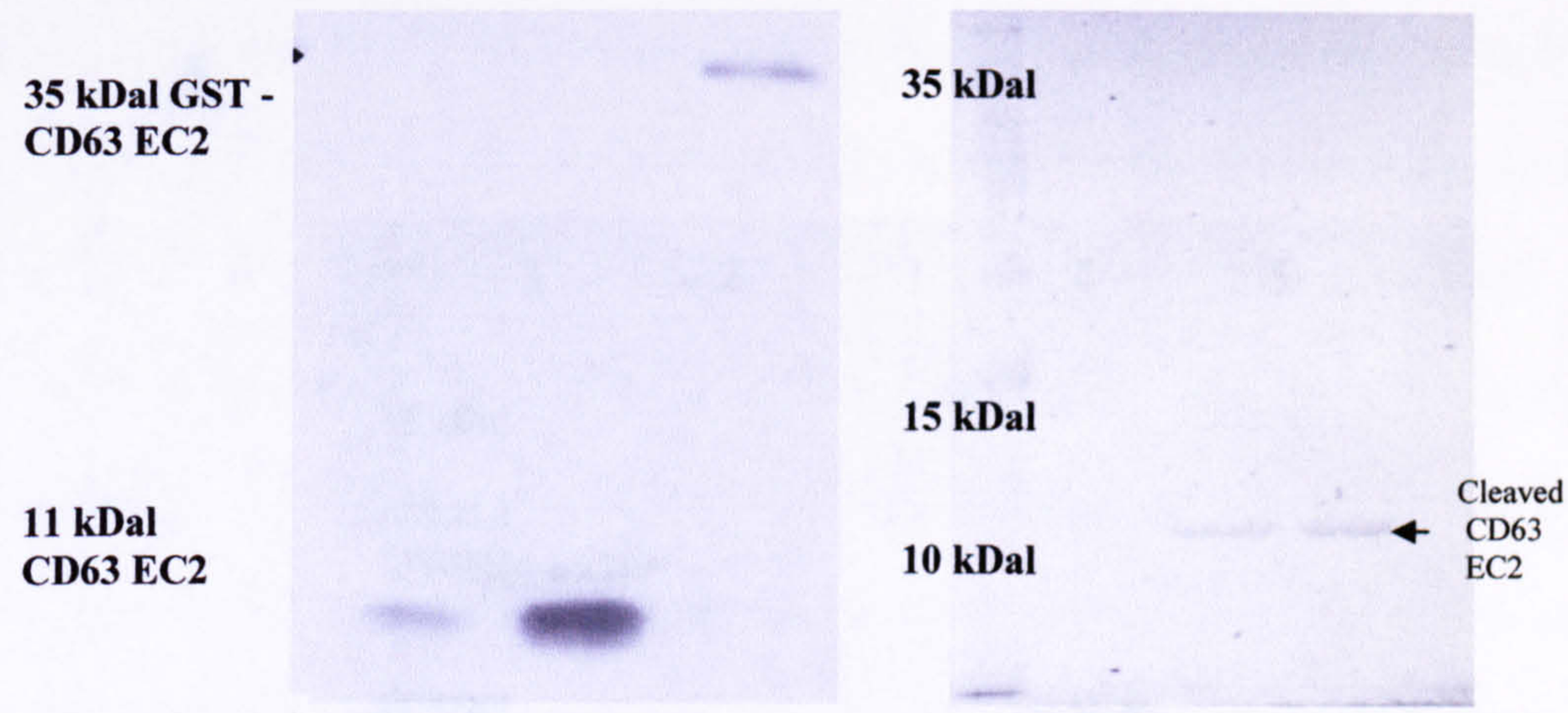
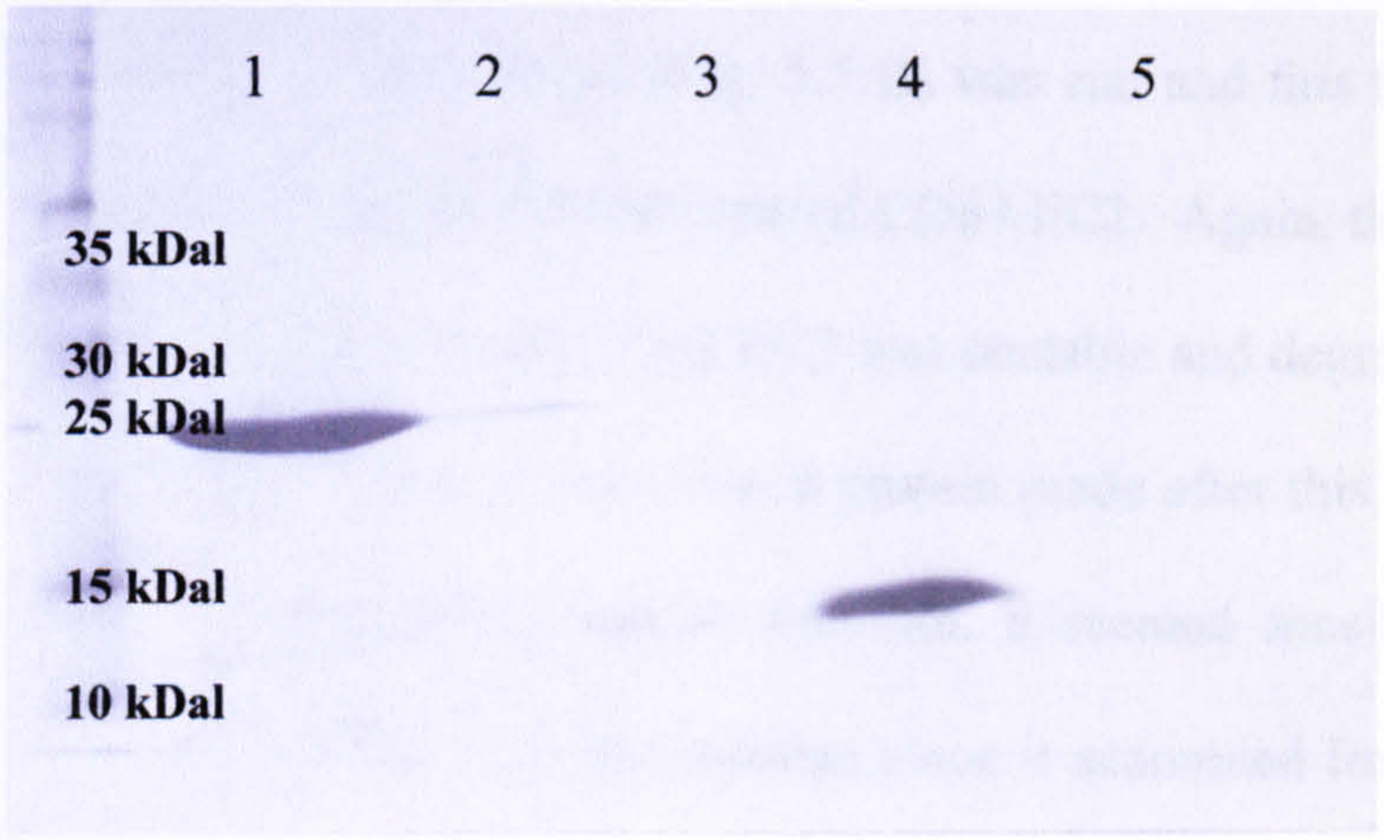


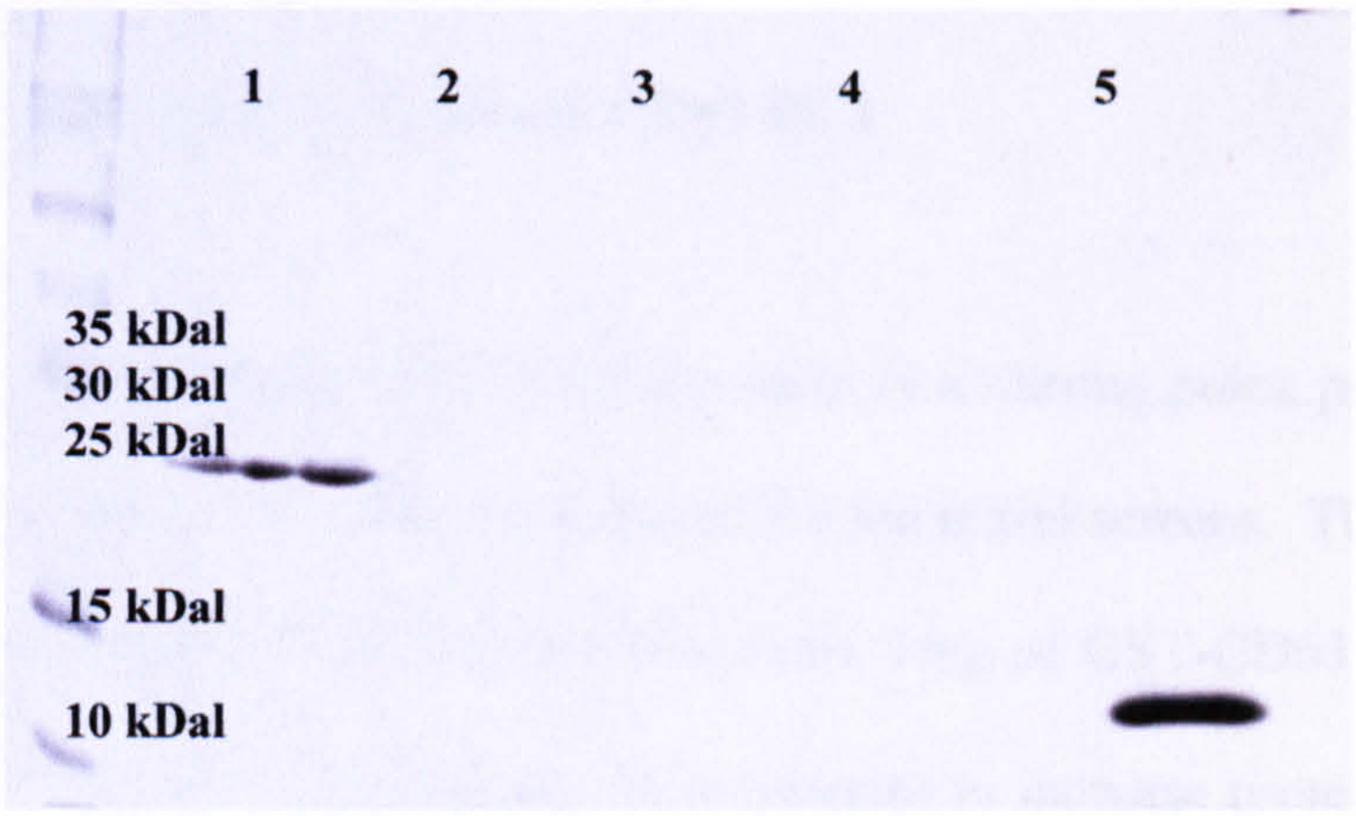
Figure 5.4. Western blot of samples used to optimise purification of cleaved CD63 EC2 from the reaction mixture containing GST. 100µg of GST-CD63 EC2 was mixed with 10µl of glutathione Sepharose slurry and left to bind for 1 hour at RT before addition of 0.1 U of thrombin. The beads were pellet by centrifugation at 400g for 2 minutes and the supernatant contained cleaved CD63. The beads were washed in PBS and centrifuged again, the wash buffer was saved and the beads were resuspended in 100µl PBS and saved for analysis. 20µl of the supernatant, wash buffer and beads were run on two SDS-PAGE gels. **A)** One gel was blotted and probed with H5C6 (as described in Figure 5.3). **B)** silver stain of the second gel.

After taking an aliquot of the purified, cleaved protein, the remainder of the sample was stored at 4˚C. The following week, the sample was used in an experiment to remove the thrombin using p-aminobenzamidine agarose. This agarose has a binding capacity of 7-10mg of thrombin per ml of slurry. The amount of thrombin present in the sample was in the ng range. A 10µl sample of the cut GST-CD63 EC2 was removed for analysis, and 5µl of p-aminobenzamidine agarose was added to the sample and mixed at 4˚C for 2 hours. After this incubation period the agarose was pelleted and the supernatant was carefully collected and analysed by SDS-PAGE stained with Coomassie, along with the sample taken before treatment with p-aminobenzamidine, with GST and a 10kDal protein used as markers (Figure 5.5).

A



B



Lane	1	2	3	4	5
A	GST	Before treatment	After treatment	10 kd protein control	empty
B	GST	Before treatment	After treatment	Empty	10 kd protein control

Figure 5.5. Effects on thrombin removal from cleaved CD63-EC2 using p-aminobenzamidine agarose. The solution of CD63-EC2 and thrombin protease from Fig. 5.4. was kept at 4˚C for 48 hours before the attempted removal of thrombin using p-aminobenzamidine agarose. Samples before treatment and after were run on SDS-PAGE and Coomassie (**A**) or silver-stained (**B**) along with a GST control and a 10kDal protein.

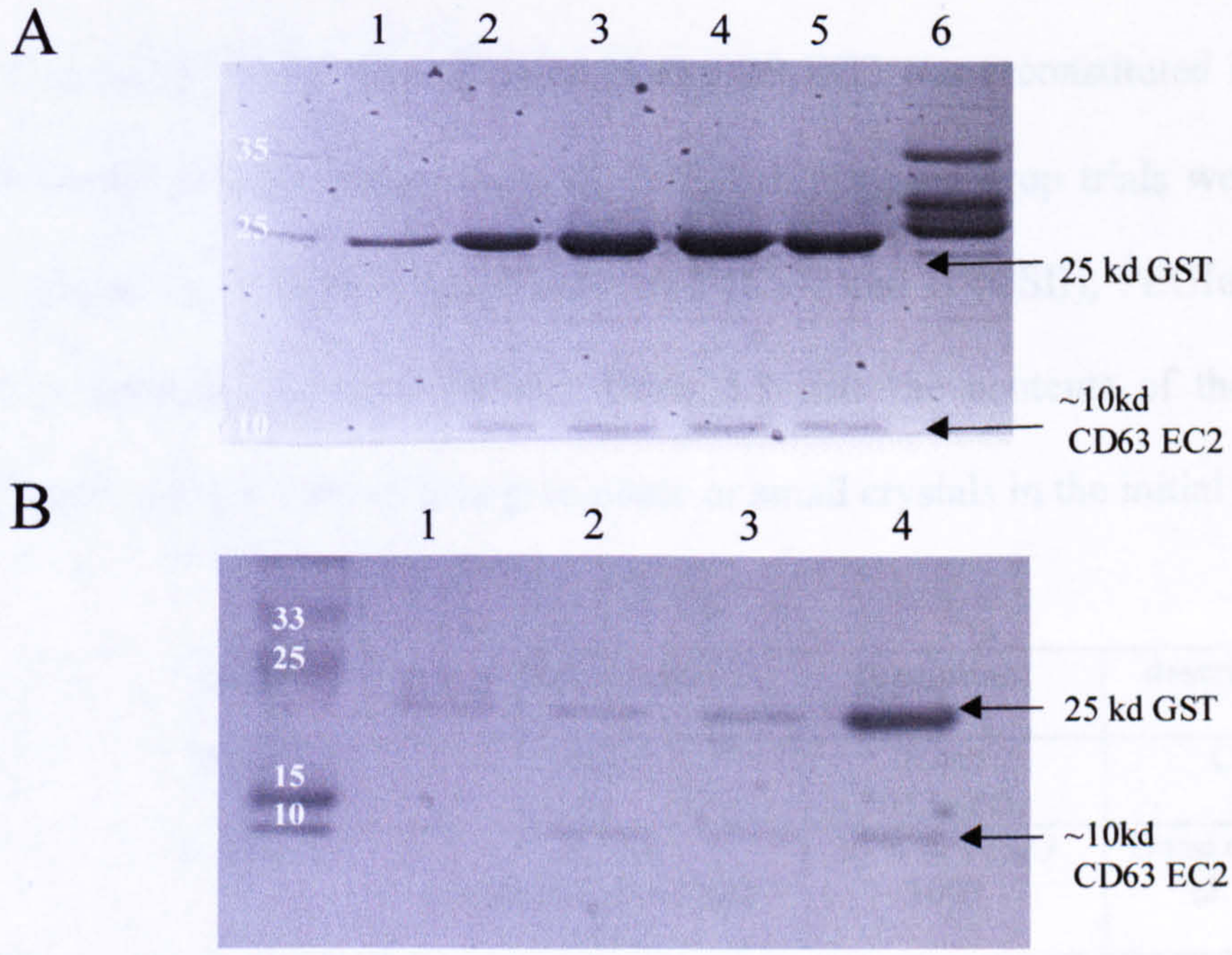
The first Coomassie stained gel (Fig.5.5 A) did not show any sign of a 10kDal protein either in the sample before treatment with p-aminobenzamidine, or in the post – treatment sample. A second gel (Fig. 5.5 B) was run and this time silver stained in an attempt to detect any 10kDal cleaved CD63 EC2. Again, this gel was negative. It appeared that the cleaved CD63 EC2 was unstable and degraded when stored at 4°C for a couple of days. Any cleaved protein made after this was made for immediate use. As the protein was so unstable, it seemed sensible not to attempt to remove the thrombin from the solution since it accounted for less than 0.1 % total protein.

5.3.2 Column purification of cleaved CD63 EC2

Crystal trials generally require 10mg/ml of protein as a starting point, preferably in water. However, only 100-200µl are required for the initial screens. The cleavage procedure is only around 5 % efficient (for every 1mg of GST-CD63 EC2, only about 50µg of CD63 EC2 is purified). In an attempt to increase protein yield for crystal trials, it was decided to use 15ml columns. The columns were siliconised and left to air-dry before being loaded with glutathione Sepharose saturated with 10mg of GST-CD63 EC2 (in approx. 15ml). At this point, in order to concentrate the final product, the non-bound volume in the column was drained out. This should be PBS as all of the GST-CD63 EC2 and GST alone should have bound to the glutathione Sepharose. The thrombin was added in 500µl PBS, and the cleavage reaction was performed in the column. After the reaction the column was drained into siliconised 1.5ml tubes and washed three times in PBS. Reduced glutathione was then added to Sepharose and collected after a 5 minute incubation.

If only GST eluted from the column then we would assume that the cleavage reaction had worked and all the CD63 EC2 present had been cut from GST.

Figure 5.6A shows the first Coomassie stained gel. The fraction taken after the cleavage event (lane 5) as well as washes (lanes 2-4), all contained the cleaved CD63 EC2 band, but they also contained high amounts of GST. In an attempt to remove this GST, 100µl of glutathione Sepharose was added to the eluted fractions, the theory being that this would bind and "mop up" the cleaved GST. Figure 5.6.B shows that although some of the GST had been removed there was still ~50 % contamination with GST. Due to these problems in purifying large quantities of cleaved CD63 it was decided to use His6CD63-EC2 for crystallisation trials.



Lane	1	2	3	4	5	6
A	eluted fraction	wash 3	wash 2	wash 1	cleaved CD63 EC2	uncut GST CD63 EC2
B	wash 2	wash 1	unbound	beads		

Figure 5.6. Silver stained SDS-PAGE showing samples from the optimisation of column purification and GST cleavage from CD63 EC2. 10mg of GST-CD63 EC2 was mixed for 1 hour at RT with 15ml glutathione Sepharose slurry in a pre-siliconised 15ml column. Thrombin protease (1U/mg) was added to the column and the reaction left for 4 hours at RT with gentle stirring. The column was drained and the eluted fraction and PBS washes were saved and run on SDS-PAGE and then silver stained (A). Glutathione beads were added to these samples in an effort to remove the GST. 100µl of glutathione Sepharose slurry was added and incubated for an hour at room temperature before centrifugation at 400g for 2 minutes to pellet the beads. PBS washes were performed in the same way and the samples run on SDS-PAGE and silver stained (B).

5.3.3 His6-CD9 EC2 crystallisation trials

A vial containing 2mg of lyophilised His6-CD9 EC2 was reconstituted in 200 µl dH_2O to give a final concentration of 10mg/ml. Hanging drop trials were set up using formulations from: Crystal Screens I (CSI) and II (CSII), PEGIon Screen (PI) and Structure Screen I (SSI). Table 5.3 lists the contents of the reagent formulation that gave crystalline precipitate or small crystals in the initial screen.

Screen #	Salt	Buffer / pH	precipitant	description
CS1 3	None	None	0.4M $NH_3H_2PO_4$	CP
CS1 15*	0.2 M NH_3SO_4	0.1 M $C_2H_6AsO_2Na.3H_2O$ pH 6.5	30 % w/v PEG 8000	Quasi crystals (5.7b)
CS1 46	0.2 M $CaC_4H_6O_4$	0.1 M $C_2H_6AsO_2Na.3H_2O$ pH 6.5	18 % w/v PEG 8000	CP
CSII 3	None	None	25 % v/v Ethylene Glycol	CP
CSII 16*	0.5 M NaCl	0.1 M $Na_3C_6H_5O_7.2H_2O$ pH 5.6	2 % w/v Ethylene Imine Polymer	CP (5.7a)
CSII 20	None	0.1 M MES pH 6.5	1.6 M MgSO4.7H2O	CP
CSII 24	0.05 M CsCl	0.1 M MES pH 6.5	30 % v/v Jeffamine M-600	CP
PI 20*	0.2 M $Mg(HCOO)_2.2H_2O$	pH 5.9	20 % w/v PEG 3350	Microcrystals (5.7c)
PI 24	0.2 M $LiC_2H_3O_2.2H_2O$	pH 7.8	20 % w/v PEG 3350	CP
PI 25	0.2 M $(CH_3CO_2)_2Mg.4H_2O$	pH 7.7	20 % w/v PEG 3350	CP
PI 29	0.2 M CH_3CO_2K	pH 7.8	20 % w/v PEG 3350	CP
PI 40	0.2 M $Na_2HPO_4.2H_2O$	pH 9.1	20 % w/v PEG 3350	CP
SSI 11*	0.2 M $CaCl_2$	0.2 M $CH_3COONa.3H_2O$ pH 4.6	20 % Propan-1-ol	Needle-like crystals (5.7d)

Table 5.3. His6-CD9 EC2 crystal screening. Crystal screen I and II (CSI and CSII), PegIon (PI) and Structure screen I were used for CD9 EC2 crystal trials. The tale lists the conditions that gave crystals (C) or crystalline precipitate (CP). * denotes conditions shown in Fig. 5.7.

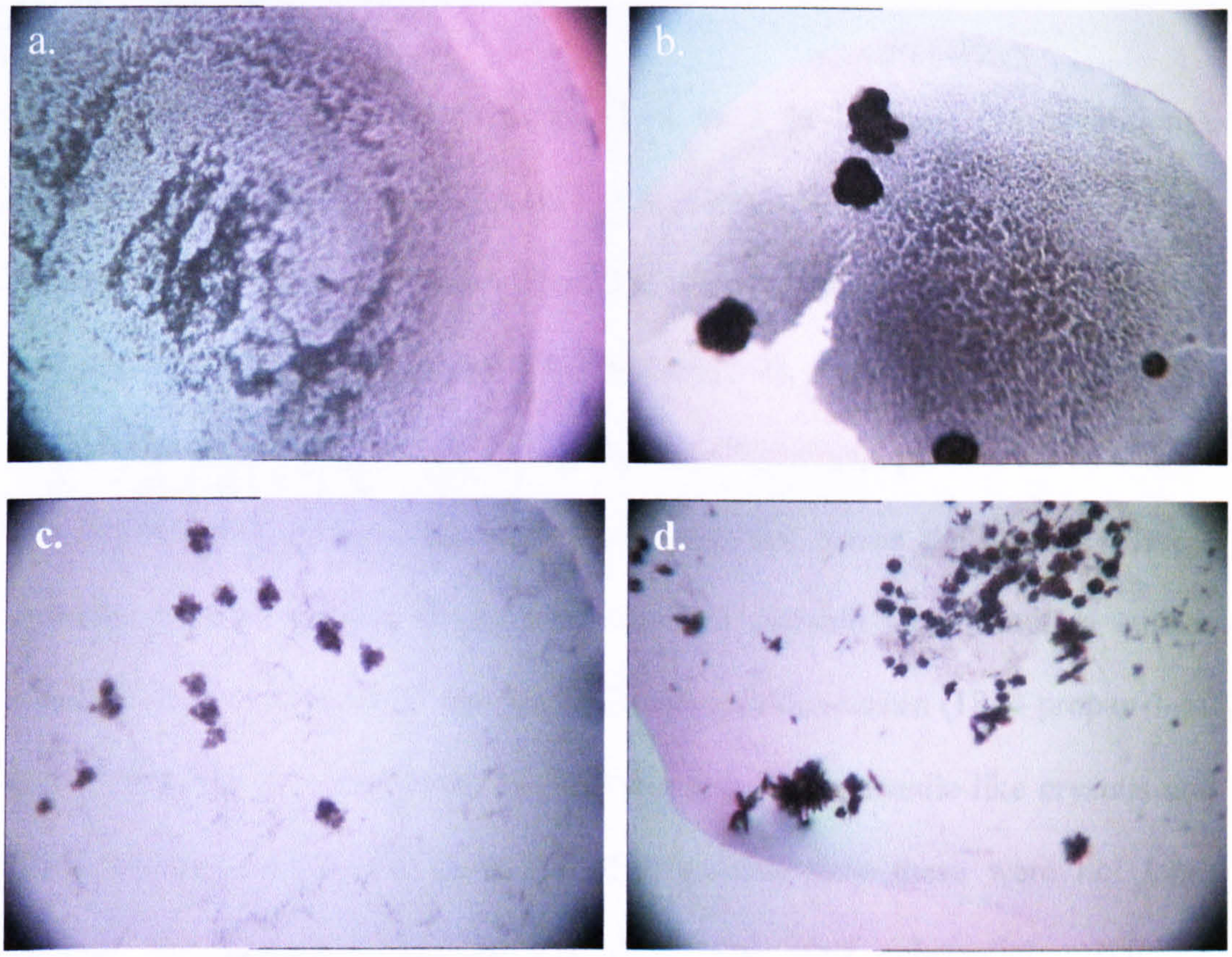


Figure 5.7. Results from the initial His6-CD9 EC2 screen. Crystal screening kits (Hampton) were used to determine crystallisation conditions for his6-CD9. 1µl of 10mg/ml His6CD9 in dH_2O was mixed with 1µl of each formulation and were set as hanging drops above 700µl formulation. The conditions in these well are listed in Table 5.3, a = CSII 16, b = CSI 15, c = PI 20, d = SSI 11. (all crystal pictures are 20 – 100 x magnification).

Figure 5.7 shows some of the more interesting wells identified in the first screen.

Figure 5.7.a is a good example of crystalline precipitate. Figures b–d contain more separate entities that resemble better crystal-like structures with a corresponding decrease in pH of formulation used. Figure b) shows Quasi-like crystals (pH 6.5), the bodies in c) resemble microcrystals (pH 5.9) and there are needle-like clusters in d), (pH4.6). The latter well looked the most promising and so further trials were set up based on the formulations of SSI 11.

5.3.3.1 CD9 EC2 crystallisation trials II

The variables were; 10–30 % propan-1-ol in 2 % intervals (11 conditions), CH_3CO_2Na, pH 4.0, 4.6 and 5.0, and $CaCl_2$ at concentrations of 0.1 M, 0.2 M or omitted, giving a total of 99 conditions (Table 5.4). The needle-like crystals were not seen again in this trial (Figure 5.8. and Table 5.4). Some crystals did grow in a tiny bubble within one well (26 % propan-1-ol, CH_3CO_2Na, pH 5.0, 0.2 M $CaCl_2$), but it was not possible to repeat the conditions that forced the growth of these crystals, and the crystals themselves were not suitable for diffraction studies (Figure.5.8.c). Another well was very interesting in this screen (18 % propan-1-ol, CH_3CO_2Na pH 4.6, 0.2M $CaCl_2$). This contained large needle-like crystals and Quasi/microcrystals in the same well but unfortunately these were not large enough for diffraction studies and there was also substantial precipitate contamination in this well (Figure 5.8.d).

#	% propan-1-ol	CH_3CO_2Na pH	$[CaCl_2]$ M	Description
1	10	4.0	0.2	CP (Fig5.8a)
2	12	4.0	0.2	Clear
3	14	4.0	0.2	P
4	16	4.0	0.2	P
5	18	4.0	0.2	P
6	20	4.0	0.2	P
7	22	4.0	0.2	P
8	24	4.0	0.2	CP
9	26	4.0	0.2	CP
10	28	4.0	0.2	CP
11	30	4.0	0.2	CP
12	10	4.6	0.2	CP
13	12	4.6	0.2	P
14	14	4.6	0.2	P
15	16	4.6	0.2	P
16	18	4.6	0.2	Large needle like crystals + Quasi/microcrystals + P (Fig. 5.8d)
17	20	4.6	0.2	CP
18	22	4.6	0.2	CP
19	24	4.6	0.2	CP
20	26	4.6	0.2	P
21	28	4.6	0.2	P
22	30	4.6	0.2	P
23	10	5.0	0.2	P
24	12	5.0	0.2	P
25	14	5.0	0.2	Clear
26	16	5.0	0.2	Clear
27	18	5.0	0.2	Clear
28	20	5.0	0.2	CP
29	22	5.0	0.2	CP
30	24	5.0	0.2	CP
31	26	5.0	0.2	Small crystals formed within a bubble, CP (Fig 5.8.c)
32	28	5.0	0.2	CP
33	30	5.0	0.2	CP
34	10	4.0	0.1	CP
35	12	4.0	0.1	CP
36	14	4.0	0.1	CP
37	16	4.0	0.1	P
38	18	4.0	0.1	P
39	20	4.0	0.1	P
40	22	4.0	0.1	P
41	24	4.0	0.1	P
42	26	4.0	0.1	CP
43	28	4.0	0.1	P
44	30	4.0	0.1	P
45	10	4.6	0.1	CP
46	12	4.6	0.1	P
47	14	4.6	0.1	P
48	16	4.6	0.1	CP
49	18	4.6	0.1	CP
50	20	4.6	0.1	Clear
51	22	4.6	0.1	CP
52	24	4.6	0.1	P
53	26	4.6	0.1	P
54	28	4.6	0.1	P
55	30	4.6	0.1	P
56	10	5.0	0.1	Infected
57	12	5.0	0.1	Infected
58	14	5.0	0.1	Infected
59	16	5.0	0.1	P
60	18	5.0	0.1	P

61	20	5.0	0.1	P
62	22	5.0	0.1	P
63	24	5.0	0.1	CP
64	26	5.0	0.1	CP
65	28	5.0	0.1	P
66	30	5.0	0.1	P
67	10	4.0	0.0	P
68	12	4.0	0.0	Hair
69	14	4.0	0.0	P
70	16	4.0	0.0	CP
71	18	4.0	0.0	CP
72	20	4.0	0.0	P
73	22	4.0	0.0	Clear
74	24	4.0	0.0	P
75	26	4.0	0.0	CP + outlines of globular regions (fig5.8b)
76	28	4.0	0.0	CP
77	30	4.0	0.0	CP
78	10	4.6	0.0	CP
79	12	4.6	0.0	CP
80	14	4.6	0.0	P
81	16	4.6	0.0	P
82	18	4.6	0.0	P
83	20	4.6	0.0	Infected
84	22	4.6	0.0	P
85	24	4.6	0.0	P
86	26	4.6	0.0	P
87	28	4.6	0.0	P
88	30	4.6	0.0	P
89	10	5.0	0.0	P
90	12	5.0	0.0	P
91	14	5.0	0.0	P
92	16	5.0	0.0	P
93	18	5.0	0.0	P
94	20	5.0	0.0	P
95	22	5.0	0.0	P
96	24	5.0	0.0	P
97	26	5.0	0.0	P
98	28	5.0	0.0	CP
99	30	5.0	0.0	CP

Table 5.4. CD9 EC2 further trials. The first set of further trials were based on SSI 11 conditions: 0.2 M $CaCl_2$, CH_3CO_2Na pH 4.6, 20% propan-1-ol. The variables are listed in the table, along with a description of the contents of the well taken 2 weeks after the trials were set up. P = Precipitate, CP = Crystalline Precipitate.

5.3.3.2 His6-CD9 EC2 crystallisation trials III

The remainder of the His6-CD9 EC2 was used in trials with two other conditions that were pulled out of the original screen, these being 20 % PEG 3350 and $Mg(HCO_2)_2$, pH 5.9 (Fig. 5.7.c), or CH_3CO_2K, pH 7.8. A concentration range of PEG 3350 from 10 – 25 % with $Mg(HCO_2)_2$ pH 5.8 and 6.4, or CH_3CO_2K pH 7.0 and 7.8, were used (Table 5.5). There was some crystal growth in the $Mg(HCO_2)_2$

pH 5.8 plate in well with 20 and 25 % PEG 3350 (Fig.5.8.e and f, respectively) but these were not large enough to test by X-ray diffraction.

	% PEG 3350	0.2 M $Mg(HCOO)_2$/ 0.2 M CH_3COOK	pH	Description
1	10	$Mg(HCO_2)_2$	5.8	P
2	15	$Mg(HCO_2)_2$	5.8	CP
3	20	$Mg(HCO_2)_2$	5.8	CP
4	25	$Mg(HCO_2)_2$	5.8	CP
5	10	$Mg(HCO_2)_2$	6.4	CP
6	15	$Mg(HCO_2)_2$	6.4	CP
7	20	$Mg(HCO_2)_2$	6.4	Small crystals around the edge of the well, P in centre (**Fig.5.8.e**).
8	25	$Mg(HCO_2)_2$	6.4	Multiple, small, needle-like crystals scattered throughout the well (**Fig.5.8.f**).
9	10	CH_3CO_2K	7.0	CP
10	15	CH_3CO_2K	7.0	CP
11	20	CH_3CO_2K	7.0	Clear
12	25	CH_3CO_2K	7.0	P
13	10	CH_3CO_2K	7.8	P
14	15	CH_3COOK	7.8	P
15	20	CH_3COOK	7.8	P
16	25	CH_3COOK	7.8	P

Table 5.5. Final trials with His6-CD9 EC2. Two further conditions picked out in the initial screen were:- 0.2 M $Mg(HCO_2)2$ pH 5.8, 20 % PEG3350 and 0.2 M CH_3CO_2K pH 7.8, 20 % PEG3350. These were chosen because they both contain PEG3350. The remainder of the His6-CD9 EC2 was used to screen around these conditions. Small crystals were formed in wells 7 and 8 (Fig.5.8. e) and f), respectively). Unfortunately, these crystals were too small but could provide useful information for future trials with this protein.

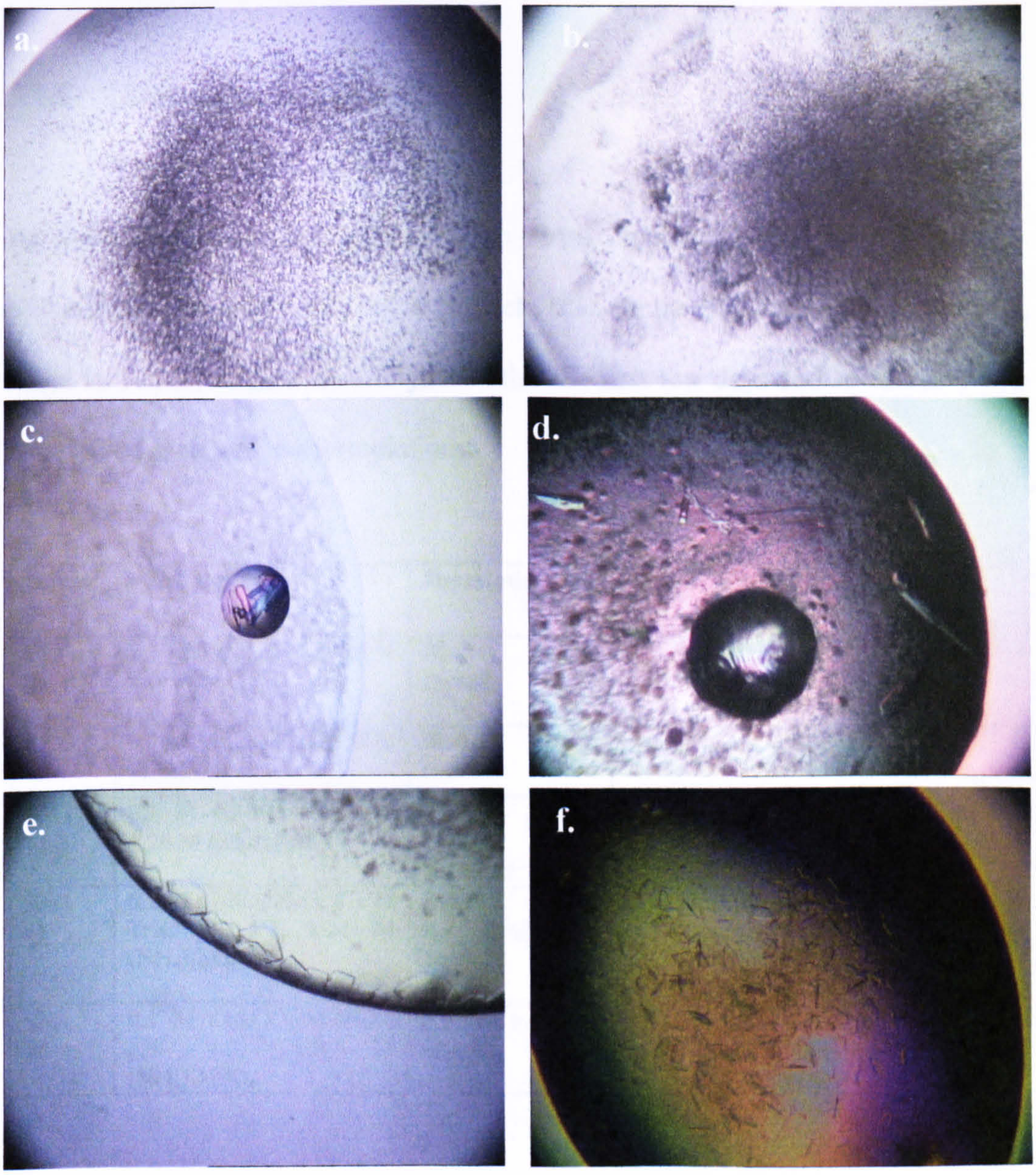


Figure 5.8. Further trials with His6 CD9-EC2. Figures a – d are the results of further trials centred around the conditions that gave the needle-like crystals, SSI (see Table 5.4. for description). Pictures e) and f) are taken from the results of the second further trials, centred on PEGIon (Table 5.5.).

5.3.3. CD63 crystallisation trials

An automated system was used to set up sitting drops of CD63-His6 EC2 and the 600 different formulations (Nextal Biotech, US). In the first screen, the conditions listed in Table 5.6 gave small crystals after the first few days and so further trials were based around these formulations.

Screen #	Formulation	Description	Further trials
PEG 11	0.1 M MES pH 6.5, 25 % w/v PEG 1000	Rod shaped crystal.	0.1 M MES, pH 5.0–8.0 in 0.5 intervals. 25 % and 35 % PEG1000
PEG 12	0.1 M MES, pH 6.5, 20 % PEG 2000	Rod shaped crystal.	As above but with PEG2000
SM1 11	0.1 M HEPES, pH 7.5, 1.26 M $(NH_4)_2SO_4$	2 rod shaped crystals. Attached.	0.1 M HEPES pH 6.0, 7.5, 8.0, 8.5. 0.6 M, 1.26 M, 2.5 M $(NH_4)_2SO_4$
SM1 12	0.2 M $Li_2SO_4.H_2O$, 0.1 M Tris, pH 8.5, 1.26 M $(NH_4)_2SO_4$	Single crystal but similar to SM1 11.	0.1 M Tris pH 7.0 – 9.0 in 0.5 intervals. 0.1 M, 0.2 M $Li_2SO_4.H_2O$ or without $Li_2SO_4.H_2O$. 0.6 M, 1.26 M or 2.5 M $(NH_4)_2SO_4$
SM1 25	0.1 M $CH_3CO_2Na.3H_2O$, pH 4.5, 1.0 M $(NH_4)_2HPO_4$	2 small, cuboidal crystals.	0.1 M $CH_3CO_2Na.3H_2O$ pH 3.0 – 5.5 in 0.5 intervals. 0.5 M, 1.0 M, 1.5 M, 2.0 M $(NH_4)_2HPO_4$

Table 5.6. Conditions from the first screen of CD63-His6. The robot was used to screen 6 formulation kits (96 formulations/kit) for the ability to produce CD63 EC2 crystals. The table lists the best conditions around which further trials were performed.

From these trials, crystals grew in the HEPES and 2.6M $(NH_4)2SO_4$ wells at pH 7.5, 8.0 and 8.5. However, when tested for X-ray diffraction, they produced a salt crystal diffraction pattern.

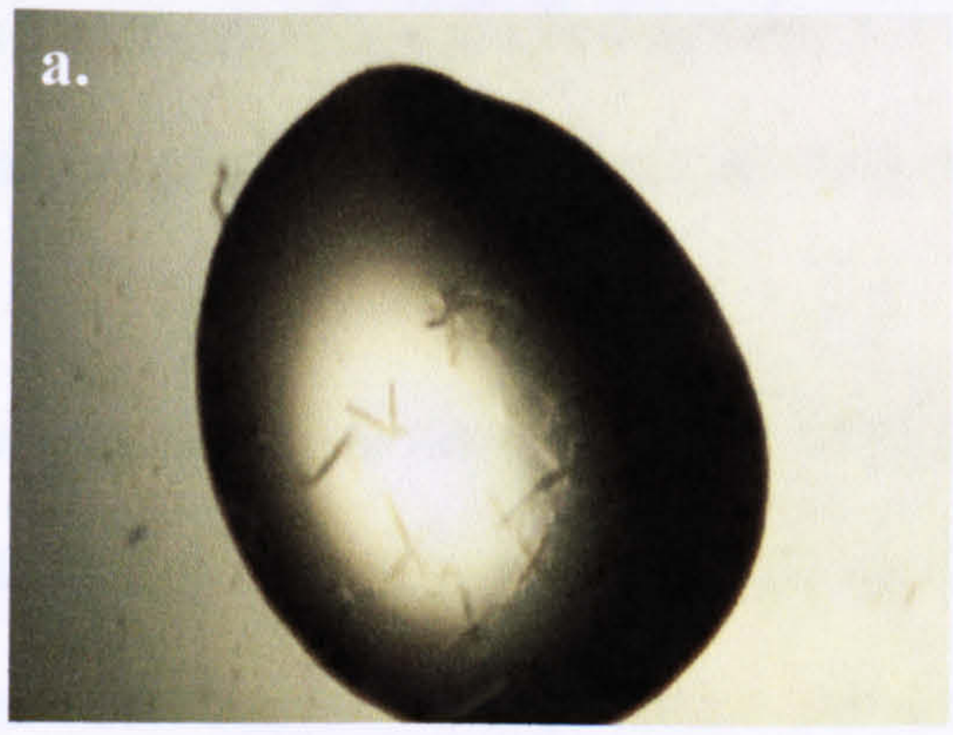
The PEG 12 screen grew a reasonably sized crystal and this was mounted for X-ray diffraction studies. The crystal turned out to be salt, but it was noted that, upon

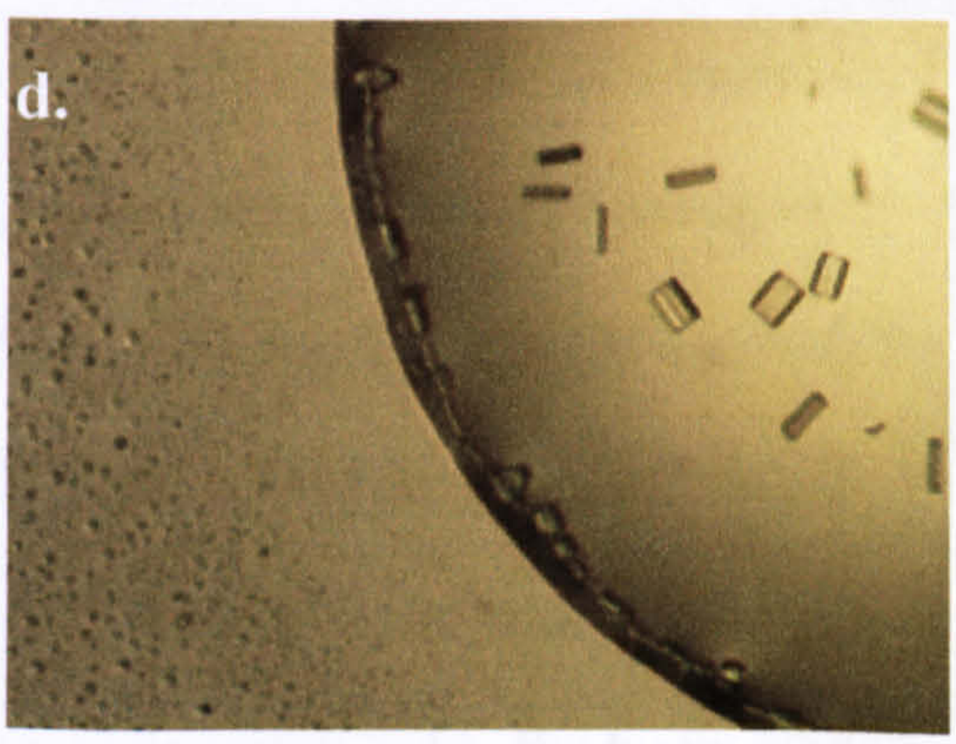
trying to mount the crystal, there was a skin across the drop that when pierced, released a large amount of precipitate. It appeared that the skin had developed preventing the drop from reaching equilibrium and therefore, preventing the formation of potential crystals. Using a fine acupuncture needle, the 0.33 µl drops were carefully tested for the presence of a skin and if one was found it was removed, the well resealed and the contents of each well left to reach equilibrium once more.

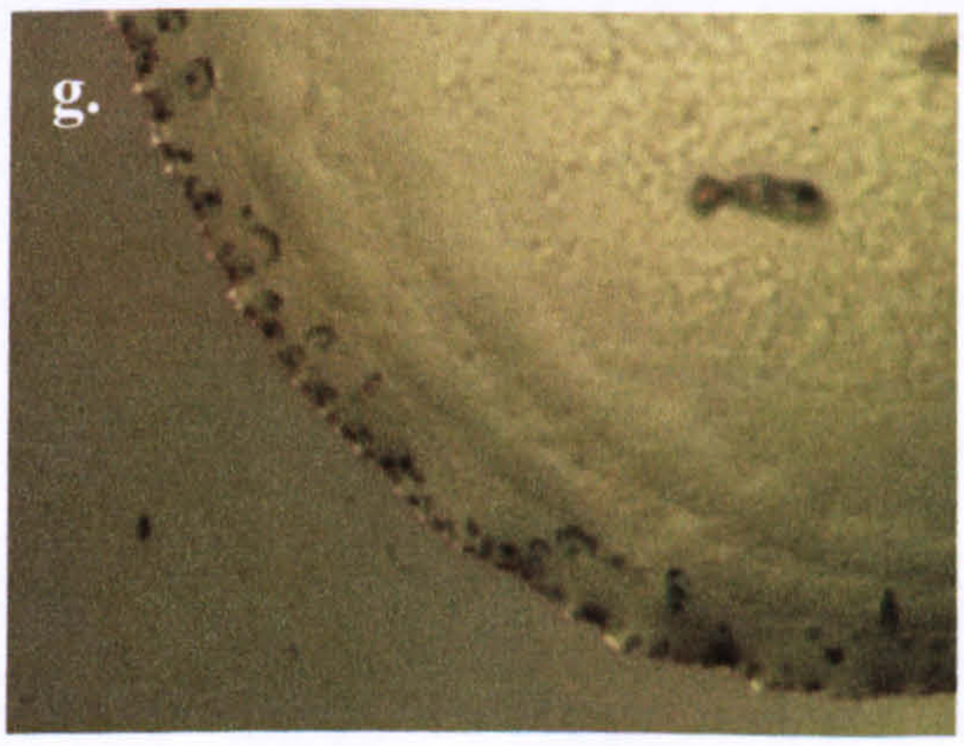
Two days later, the drops were viewed again, and the contents noted. Table 5.7 lists the conditions of the wells in which crystals grew and gives a description of the crystals formed. These crystals are shown in Figure 5.9. Fortunately, all of these crystals were large enough to diffract X-rays.

Screen #	Formulation	Description
PEG 1	0.2M NaF, pH 7.1, 20 % w/v PEG 3350	Small, cuboidal crystals around the perimeter of the drop **(Fig5.10.c and d)**
PEG 3	0.2 M NH_4F, pH 6.2, 20 % PEG 3350	Clusters of small crystals around the perimeter and some larger ones in the centre of the well. Jagged edges **(5.10.g and e)**
PEG40	0.2 M $Na_2HPO_4.2H_2O$, pH 9.8, 20% PEG 3350	Two crystals, crossed at one end. Rocket like shape **(5.10.f)**
PEG44	0.2 M $(NH_4)_2HPO_4$, pH 7.9, 20% PEG 3350	Fern-like crystal. Lots of tiny projections from two large, crossed over crystals **(5.10.h)**
CS1	0.1 M $CH_3COONa.3H_2O$, pH 4.5, 1.0 M $(NH_4)_2HPO_4$	Multiple, fern-like crystals **(5.10.a and b.)**

Table 5.7. Chemical formulations of wells where crystal forms were observed in the second screening experiment.







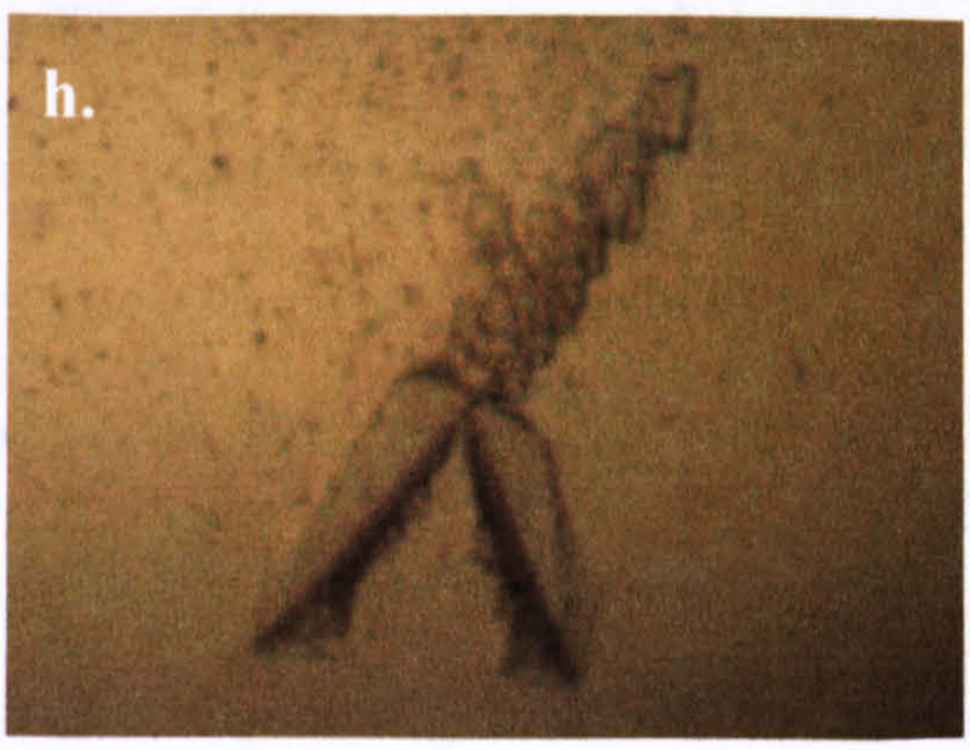


Figure 5.9. Crystals from screening CD63-His6 EC2. Magnification is between 20 and 100. For a description of the contents of the well see Table 5.7.

All of the crystals described in Table 5.7 were loaded onto the X-ray diffractor but they all produced the characteristic diffraction pattern from salt crystals.

As mentioned previously, around 50 % of the drops had developed a skin, thus preventing equilibrium between the contents of the drop and the reservoir being reached. The only way to remove the skin from the sitting drops produced by the robot was to peel back the film and use a needle to lift away the skin before re-sealing. This was a complicated procedure that took time, however, the drops were 0.33µl in volume and undoubtedly a large amount of evaporation would have taken place in the time taken to remove the skin. This may have been having some effect on crystal growth. It was therefore decided to manually set some crystal trials using the Hampton screens. These drops were larger (2µl) and individually sealed, thus reducing the exposure time if a skin should develop. Unfortunately, no crystal growth was detected by this method.

5.3.5 One dimensional ^{1}H spectrum of CD63-His6 EC2

The collaborator who made the CD63-His6 EC2 proteins used circular dichroism to verify the presence of alpha helices. To ensure that the protein had regained its structure after lyophilisation it was decided to perform 1D—NMR spectrometry on the sample. This was also done to confirm that protein crystals had not failed to grow simply because of lack of structure. 500µl of a 4mg/ml solution of CD63-His6 EC2 was used to obtain a one dimensional spectrum (Figure 5.11).

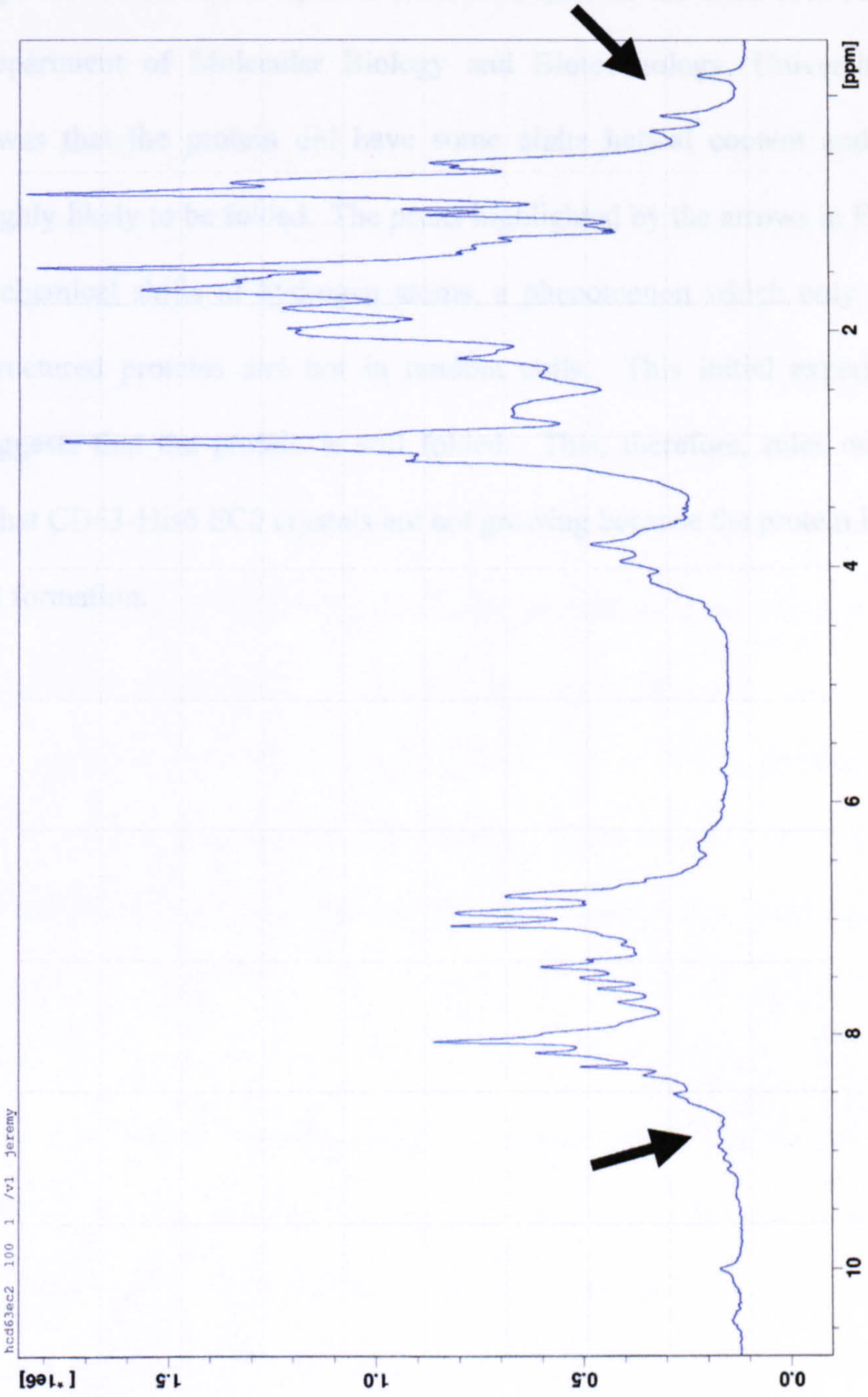


Figure 5.11. One dimensional ^{1}H spectrum of CD63-His6 EC2. 500µl of CD63-His6 EC2 at 4mg/ml in dH_2O was placed in a glass capillary tube and centrifuged briefly to ensure that all the protein was at the bottom of the tube. The tube was then loaded into a Bruker DRX500 and subjected to 500MHz proton frequency. The 1D proton NMR spectrum was run with presaturation of the water and a hahn echo (2.6m total echo time) to improve convolution difference,1s recycle time, 128 scans. The spectrum was processed with a convolution difference to further remove the water signal, and 5Hz line broadening.

From the spectra obtained, the opinion from an expert in the field (Dr. Jeremy Craven, Department of Molecular Biology and Biotechnology, University of Sheffield) was that the protein did have some alpha helical content and was therefore highly likely to be folded. The peaks highlighted by the arrows in Figure 5.11 show chemical shifts of hydrogen atoms, a phenomenon which only takes place in structured proteins and not in random coils. This initial experiment strongly suggests that the protein is still folded. This, therefore, rules out the possibility that CD63-His6 EC2 crystals are not growing because the protein is in a random coil formation.

5.4. DISCUSSION

The initial aim for this structural work was to cleave GST from GST-CD63 EC2, concentrate the purified, cleaved CD63 and use this protein in crystallisation trials. CD63 EC2 was successfully cleaved from GST and purified by batch purification. The cleaved product was considered extremely pure since only the 11kDal band corresponding to cut CD63 was visible on a silver stained SDS-PAGE of the sample. However, once GST was cleaved, CD63 became very unstable and proved to be problematic to work with. The problems that were encountered, such as apparent "sticking" of GST to plasticware, were probably due to the hydrophobic patches present in the EC2 of CD63. To avoid this problem, pre-siliconised plasticware was used but this did not prevent loss of cleaved CD63. Instead, CD63-His6 EC2, provided by a collaborator, was used for the crystallisation trials.

The prerequisite for X-ray crystallography is a single crystal with two or more dimensions being 0.2 mm or longer and this, in many cases, is the rate limiting step in structural studies. The conditions that cause a protein to crystallise are specific for the individual protein and variables include pH, ionic solvent and temperature. Commercially available sparse matrix screens have been developed to aid initial screening experiments and further developments of robots to set-up the screens using minute volumes of protein, means that hundreds, if not thousands of different screens can be performed each day. In many cases, however, proteins still do not crystallise. Recent data from various proteome structural genomics projects reveal the extent of this problem. The Human Proteome Structural

Genomics pilot project, (http://www.proteome.bnl.gov/progress.html) found that out of 120 soluble proteins expressed, only 19 (16 %) gave suitable crystals for diffraction studies. Other projects have resulted in comparable levels of success in gaining diffraction-quality crystals: the Berkeley Structural Genomics Centre (http://www.strgen.org/status/progresstotals.html) crystallised 19 % of the soluble expressed protein, a study by Lesley et al. (Lesley et al. 2002) showed crystallisation conditions for 31.2 % and only 3.1 % of proteins grew diffraction quality crystals in a study by the Joint Centre for Structural genomics (http://www.jcsg.org/scripts/prod/home.html).

In the case of crystallising a protein domain, as we did here, it is possible that the protein itself could be a good variable. Indeed, earlier studies have shown that if it is not possible to crystallise the target protein, homologous proteins from another species could serve as a good alternative. The difference of just a couple of amino acids can change the properties of a protein and allow crystallisation to occur in the standard screens. Strategic point mutations and truncations or deletions have also been used to encourage crystallisation, as have addition of fusion proteins to the target protein, (reviewed in (Dale et al. 2003). Producing new constructs that include or exclude amino acids on either side of the predicted domain structure would be a wise next step, although time constraints did not permit these experiments in this project.

Affinity tags are used primarily to aid purification of the protein but can also act to stabilise and solubilise the target protein. There are conflicting reports about the usefulness of tags in protein crystallisation. However, at the Centre for Genomics

and Proteomics at the UCLA, 60 % of the protein crystals that diffracted, and 57 % of the structures solved, used proteins with C-terminal His6-tags (Goulding et al. 2003) so it is probably advisable to continue the use of His6-tagged protein for any further work.

1	pH and buffer
2	Ionic strength
3	Temperature
4	Concentration and nature of precipitant
5	Concentration of macromolecule
6	Purity of macromolecule
7	Additives, effectors and ligands
8	Organism source of macromolecule
9	Substrates, enzymes, inhibitors
10	Reducing/oxidising environment
11	Metal and other specific ions
12	Rate of equilibrium and rate of growth
13	Detergents
14	Gravity, convection and sedimentation
15	Vibrations and sound
16	Volume of crystallisation sample
17	Presence of amorphous or articulate material
18	Surfaces of crystallisation vessels
19	Proteolysis
20	Contamination by microbes
21	Pressure
22	Electric and magnetic fields
23	Handling by investigator and cleanliness
24	Viscosity of mother liquid
25	Heterogeneous nucleating reagents

Table 4.8. Factors that could affect protein crystal growth (McPherson 1990).

As the probability of growing protein crystals is enhanced as the concentration of protein is increased, the recommended starting concentration is 10-100mg/ml. The concentration of GST-CD63 EC2 used was only 5mg/ml making it less likely that crystals would grow.

The 1D-NMR spectrometry graph revealed that the protein had structure, probably containing alpha-helices. This data was promising and 2D-NMR techniques could be a real possibility to continue the work in the future.

CHAPTER 6

CLONING AND EXPRESSION OF EWI PROTEINS

6.1 INTRODUCTION

6.1.1 EWI proteins

In 2001, three independent groups identified two novel primary binding partners for CD9 and CD81, namely EWI-F and EWI-2 (Charrin et al. 2001; Clark et al. 2001; Stipp et al. 2001). The EWI proteins are a subfamily of the IgSF, but the other 2 members of the EWI family do not bind to CD9 or CD81 (Clark et al. 2001; Stipp et al. 2001). EWI-F and EWI-2 consist of 6 and 4 Ig-like domains, respectively. It has been demonstrated that these proteins exist in association with CD9 or CD81 in the presence of Triton X-100, thus fulfilling the requirements of 'primary tetraspanin binding proteins' (Chapter 1.2). Each of the three studies pinpointed the EC2 domain from CD9 and CD81 as being responsible for the interaction with EWI proteins. However, critical amino acids in the EC2 domains and EWI proteins that are involved in the association had yet to be identified.

6.1.2 AIMS

We decided to use our tetraspanin EC2 constructs to attempt to further characterise the interactions between CD9 and CD81 and the two ligands, EWI-2 and EWI-F. Using WT and mutant CD9 and CD81 GST-EC2 recombinant proteins, we aimed to identify critical amino acids in these regions that are involved in the interaction.

In order to gain a complete understanding of the involvement of EWI proteins with tetraspanin complexes, it was necessary to make numerous constructs of the EWI proteins: Constructs of the four members of the human and mouse EWI family were required for expression as full length, native protein in mammalian cells as well as soluble, recombinant protein expression in bacterial cells. It was also planned to identify the Ig-like domains in EWI-2 and EWI-F responsible for binding to CD9 and CD81, and so individual Ig-like domain expression in bacteria was also planned. The Gateway™ System (Invitrogen Life Technologies, CA) was chosen to aid the cloning procedure (see 6.1.3.).

6.1.3. Gateway™ Technology

Invitrogen developed Gateway™ Technology as a highly flexible and powerful approach to cloning multiple proteins based on bacterial λ recombination (Gujral 2003). Briefly, the DNA of interest is amplified by PCR and recombined into a donor vector to create an entry clone, which is in turn mixed with one of a number of destination vectors in a 5 minute bench-top reaction, to produce an expression vector that is ready to be transformed or transfected.

In λ recombination, *att*B and *att*P sites recombine to generate *att*L and *att*R recombinant sites, and vice versa. Bacteriophage λ has three enzymes that catalyse the recombination events: Integrase (Int), Excisionase (Xis), and *E. coli* Integration Host Factor (IHF). These enzymes recognise the *att* sites and act to bring together these sites, cleave them and covalently attach the DNA.

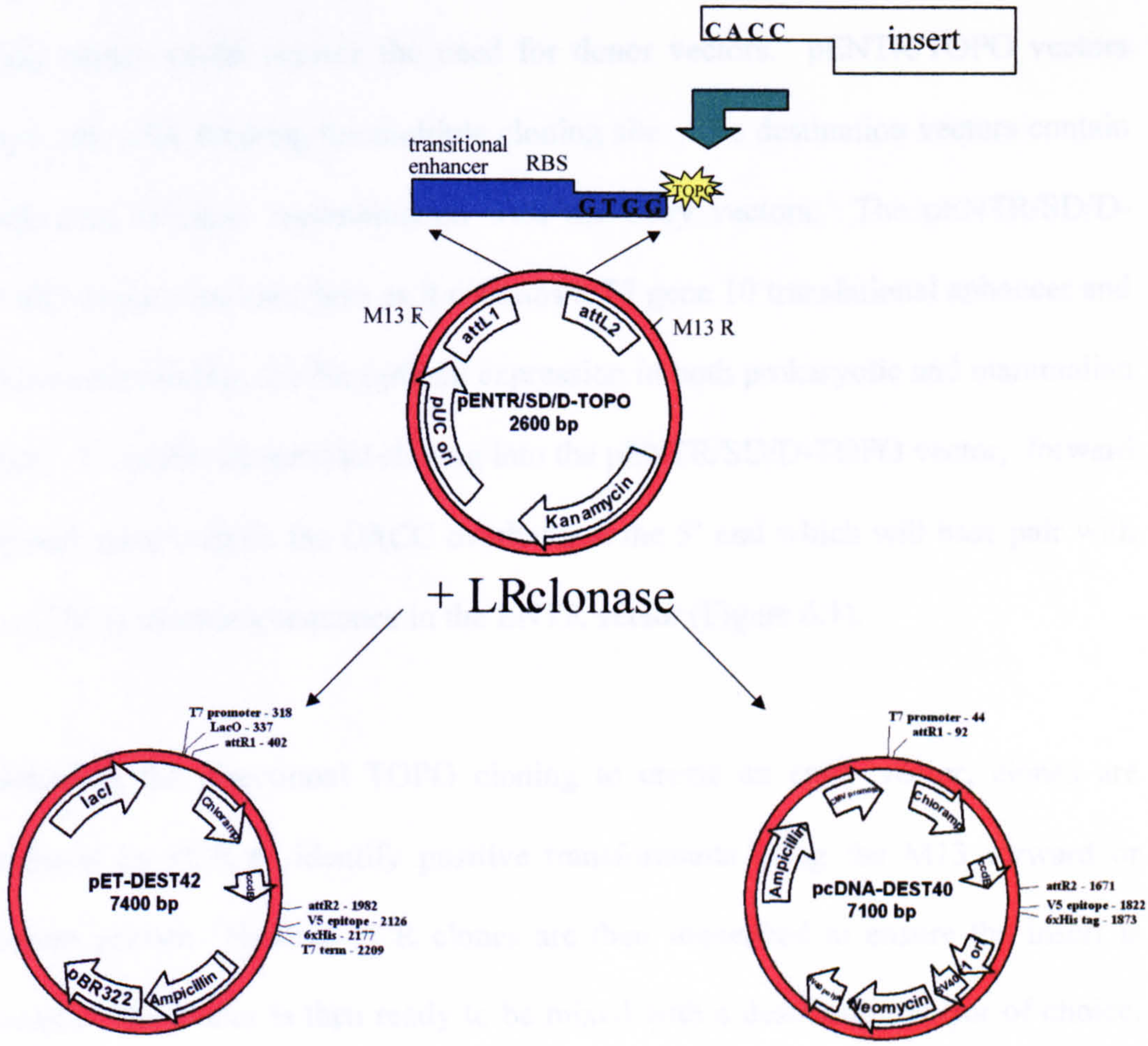


Figure 6.1. Gateway™ Cloning Reaction. To allow entry of the insert into the Gateway™ system, a CACC overhang is designed at the 5' end of the forward primer. This will combine with the GTGG overhang of the pENTR/SD/D-TOPO vector in a reaction catalysed by TOPO isomerase (yellow star). Following transformation of the pENTR/SD/D-TOPO reaction mixture with fresh PCR product and plasmid isolation from correctly sequenced constructs, the LR clonase reaction is performed. Here, the cloned insert is shuffled into either destination vector by recombination of the *att*L sites in the entry vector with *att*R sites in pET-DEST42 or pCDNA-DEST40 in a reaction catalysed by LR clonase. The diagrams of the destination vectors show selection genes, promoter sites and carboxy-terminal tags.

The first recombination reaction is to create an *att*L-containing entry clone. The quickest entry into the Gateway™ system uses pENTR/TOPO vectors to create entry clones which replace the need for donor vectors. pENTR/TOPO vectors have *att*L sites flanking the multiple cloning site. The destination vectors contain *att*R sites to allow recombination with the entry vectors. The pENTR/SD/D-TOPO vector was used here as it contains a T7 gene 10 translational enhancer and a ribosome binding site for optimal expression in both prokaryotic and mammalian hosts. To enable directional cloning into the pENTR/SD/D-TOPO vector, forward primers must contain the CACC overhang at the 5' end which will base pair with the GTGG overhang sequence in the ENTR vector (Figure 6.1).

Following the directional TOPO cloning to create an entry vector, clones are analysed by PCR to identify positive transformants using the M13 forward or reverse primer. Positive PCR clones are then sequenced to ensure the insert is correct. The vector is then ready to be mixed with a destination vector of choice. There are a number of Gateway™ destination vectors available for bacterial, mammalian and insect expression. For this study the bacterial vector pET-DEST42 (cat # 12276-010) and mammalian vector pcDNA-DEST40 (cat # 12274-015) were used (Fig 6.1). Both have C terminal V5-His6-tags and were chosen for the following reasons:

1) Due to an absence of antibodies, purification and identification of the proteins via the His6-tags was necessary.

2) The recombinant tetraspanins are GST-fusion proteins and so making EWI constructs as His6-tagged proteins would not interfere with the detection of tetraspanins.

3) Since the carboxy-terminal in the native protein contains a TM and IC region, a tag on this side was chosen. It is more likely that an N-terminal tag could interfere with the interaction with CD9 and CD81.

Once the DNA has been cloned into the destination vectors, an expression vector is formed which can then be either transformed into *E. coli* or transfected into mammalian cell lines.

6.2 MATERIALS AND METHODS

Full length EWI-F cDNA (clone KIAA1436) was ordered from Kazusa DNA Research Institute, China. EWI-2 was ordered from the IMAGE Consortium, clone 3613821. The primers for full length cloning were designed and purchased from Sigma-Aldrich (Table 6.1). These primers are suitable for protein expression in both eukaryotic and prokaryotic cells. When the ATG initiation codon immediately follows the CACC necessary for directional cloning, then the ATG is properly spaced from the RBS within the entry vector, allowing proper translation of the PCR product in prokaryotic cells – both the 5' primers were designed in this way. For proper eukaryotic translation, a Kozak consensus sequence is required. A number of sequences are possible for this but a G or A position at -3 and a G at +4 is critical for function e.g. (G/A)NNATGG. The 5' primers also fitted this description and the critical (G/A) and G residues are highlighted in bold type in Table 6.1. Stop codons were omitted as inclusion of the C-term His6-tags were required.

Oligo name	overhang	sequence
EWI-2 5'	C ACC ATG	GGC GCC CTC AGG C
EWI-2 3'		CAC AAA TAG GGT GTC CAG G
EWI-F 5'	C ACC ATG	GGG CGC CTG GCC T
EWI-F 3'		GGG ATA CTT GAA GGC GTT C

Table 6.1. Oligonucleotides designed for full length, soluble protein expression in the Gateway™ system. The C-term transmembrane region was not included so that the protein could be expressed as a soluble protein. 5' primers required the CACC overhang to allow recombination into the entry vector, immediately followed by the initiation codon, ATG. This format makes the insert in the same reading frame as the ribosome binding site in the entry vector, thus allowing expression in prokaryotic cells. The A at position -3 (A of ATG is taken as position + 1) and the G at + 4 (both highlighted in bold), create a Kozak consensus sequence to allow expression in eukaryotic cells.

6.2.1 LR clonase reaction

The LR clonase reaction was performed to recombine the entry vector containing the gene of interest with the destination vectors. The following components were added to a 1.5ml tube and mixed at room temperature:

Component	volume
Entry clone (150ng)	1-5 µl
Destination vector (150ng/µl)	1 µl
5X LR clonase reaction buffer	2 µl
TE buffer, pH 8.0	To 8 µl

The LR clonase was thawed on ice and 2µl was then added to each reaction tube and vortexed briefly to mix. The reaction was then incubated at RT for one hour before being terminated by addition of proteinase K and incubation at 37°C for 10 minutes. The destination vector now contained the gene of interest and was transformed into DH5α competent cells following the protocol given in Chapter 2.

[N.B. From here onwards the protocols were the same as the protocols given in the general Methods and Materials with regards to isolating DNA and identifying positive clones by PCR and restriction digests. All inserts were sequenced before being transfected or transformed into expression hosts.]

6.3 RESULTS

6.3.1 Prokaryotic expression using the Gateway™ system

In order to obtain a Gateway™ entry clone, the first requirement is a blunt end PCR product that runs on an agarose gel as a single, discrete band of the correct size. Thermostable, proofreading polymerase, Platinum® Pfx, (Invitrogen), was used initially. A gradient PCR block was used to test a range of Tm temperatures from 55–65°C. EWI-2 PCR product gave single, discrete bands at all the temperatures tested (Figure 6.2). However, EWI-F produced several bands that perhaps corresponded to the different immunoglobulin-like domains (Figure 6.2). PCR conditions were optimised and eventually a single band was produced (Figure 6.2) by using the conditions depicted in Table 6.2.

PCR mix components (µl)	EWI-2	EWI-F
Reddy mix	10	10
Forward primer	0.25	1.0
Reverse primer	0.25	1.0
DNA template	0.5	2.0
dH_2O	9.0	6.0
Annealing temp (°C)	55-65	65

Table 6.2. PCR conditions used to obtain single bands. In order to clone the inserts into the TOPO vector, PCR products should run as single, discrete bands at the correct size. The table lists the conditions used to obtain these.

The PCR products were run on high purity agarose gels and the DNA purified from the gel. For cloning into the pENTR TOPO vectors it is recommended to use 1-5ng of a 1 kb PCR product or 5-10ng of a 2 kb PCR product. The purified EWI-2 DNA was estimated to be at 20ng/µl and EWI-F at 30ng/µl. The DNA was therefore diluted 1:10 and 4 µl used in each TOPO reaction. The TOPO reaction

using the pENTR/SD directional TOPO vector and the cloning reaction was transformed into One Shot cells by using the chemical transformation described in Materials and Methods. 12 colonies from each construct were picked and named (EWI-)2 (1-12) and (EWI-)F (1-12) and the plasmid DNA was isolated. Commercially available M13 forward and reverse primers were used to analyse the plasmids by PCR to confirm the presence of the insert. Positive samples were sequenced using the M13 forward and M13 reverse primers for each sample. The sequences are only reliable for the first 900bp so it was necessary to design internal primers to get a full read. The oligonucleotides used for internal sequencing are listed in Table 6.3

Construct	Starting position	Forward/reverse	oligo
EWI-2	856	Forward	CTATGCTGAGCGATTGG
EWI-F	851	Forward	CCAAGAATGTGTCTGTGGCT
EWI-F	1768	Reverse	CAGCCATATGAGAACAGAG

Table 6.3. Internal primer design for EWI DNA. In order to read the whole DNA sequences it was necessary to design internal primers. The table lists the oligonucleotides and the starting positions.

6.3.2 Sequencing results

The entry plasmids EWI-2:4 and EWI-F:2 were found to contain the correct inserts. The LR clonase reaction was performed to transfer the gene of interest from the entry clone to a destination vector in order to generate an expression clone. The first destination vector used was pET-DEST42. The reaction mixture from the clonase experiment was then ready to transform into competent *E. coli.* DH5α cells were used and positive colonies were identified firstly by PCR

screening using the insert-specific primers, and secondly, by sequencing of samples that were positive in the PCR using M13 forward and M13 reverse primers. From this second screen, EWI-2:2 and EWI-F:3 were sequenced and found to be correct and so were used to transform BL21 cells. Two colonies were picked and inoculated overnight so that optimisation of protein conditions could be carried out. For each condition listed in Table 6.4, 50µl of inoculum was added to 1ml LB with 50µg/ml carbenicillin and grown for 2 hours. For each expression construct, a non-induced control was grown for 4 hours. Whole cell lysates were analysed on Coomassie stained gels, (Figure 6.3).

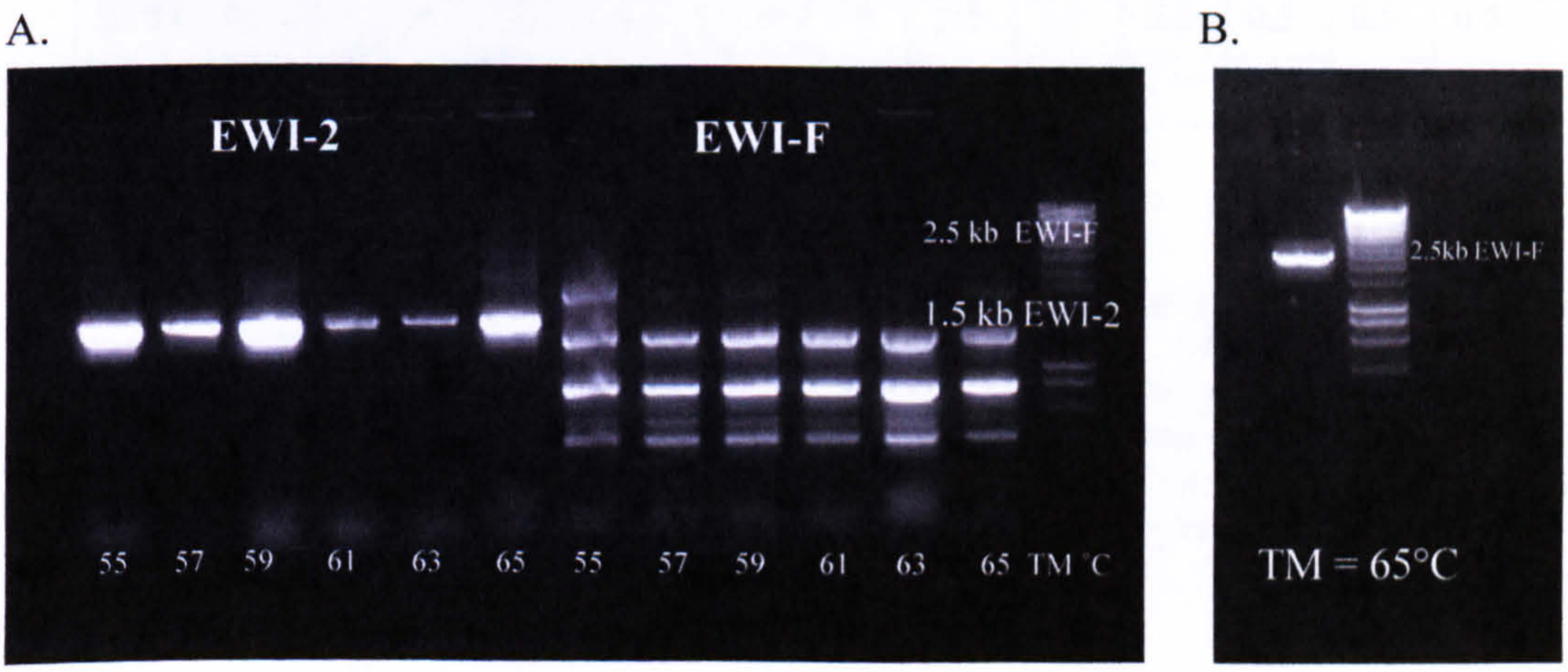
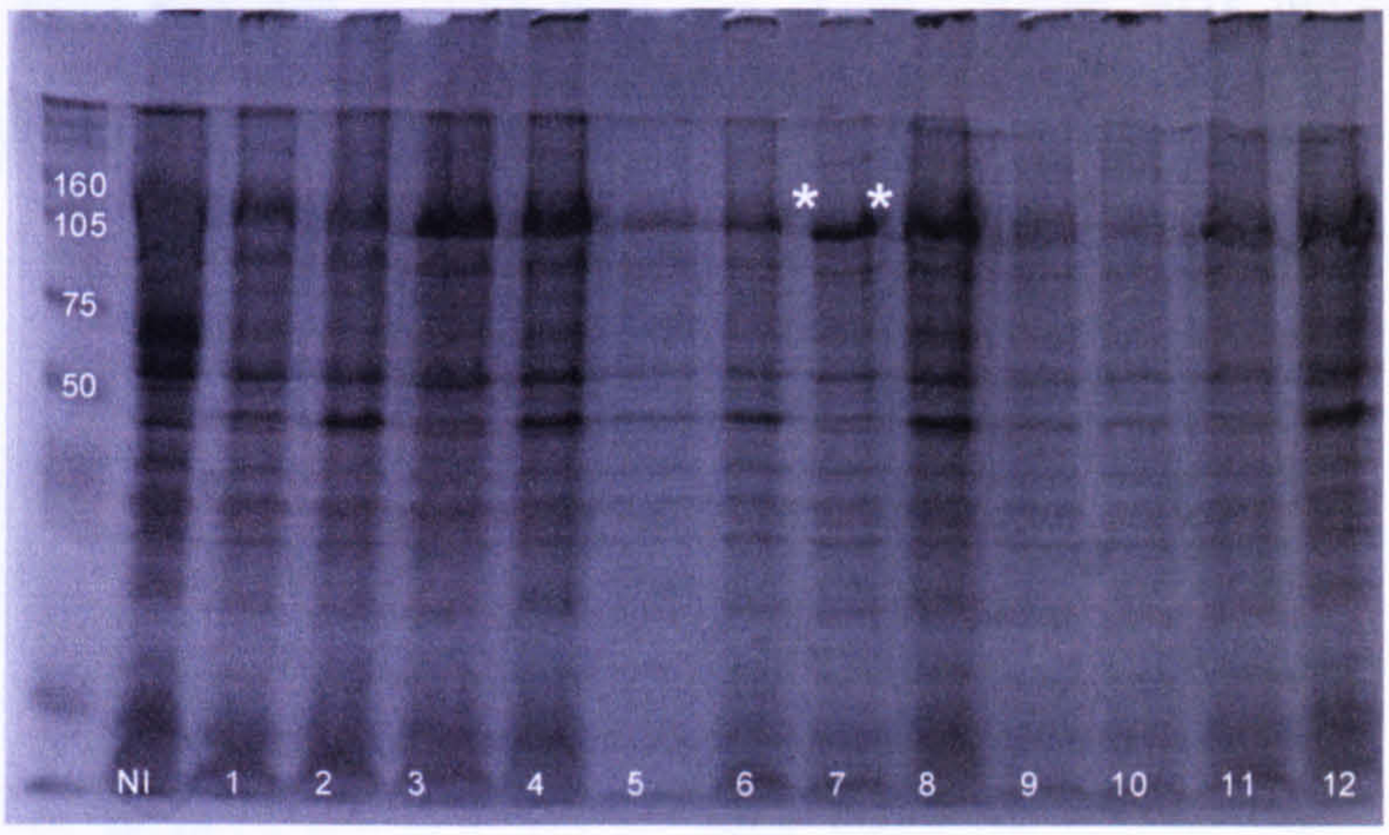


Figure 6.2. PCR amplification of EWI insert from cDNA clones. The first PCR reaction used Platinum® Pfx and gradient block with TM between 55-65°C in 2°C intervals. EWI-2 gave single bands at the correct height (1.5 kb) at all annealing temperatures but these conditions were not right for the production of EWI-F (2.5 kb) and a banding pattern was obtained (A). The polymerase was changed to HighFidelity (Abgene) and PCR conditions were altered to obtain a single, discrete band at 2.5 kb for EWI-F (B).

There was no evidence of any protein 70kDal corresponding to full length EWI-2. However, in the EWI-F optimisation from colony 2, one band at the correct size was seen (Figure 6.3). The conditions for band number 7 were: induction with 1.0mM IPTG and growth for 4 hours.

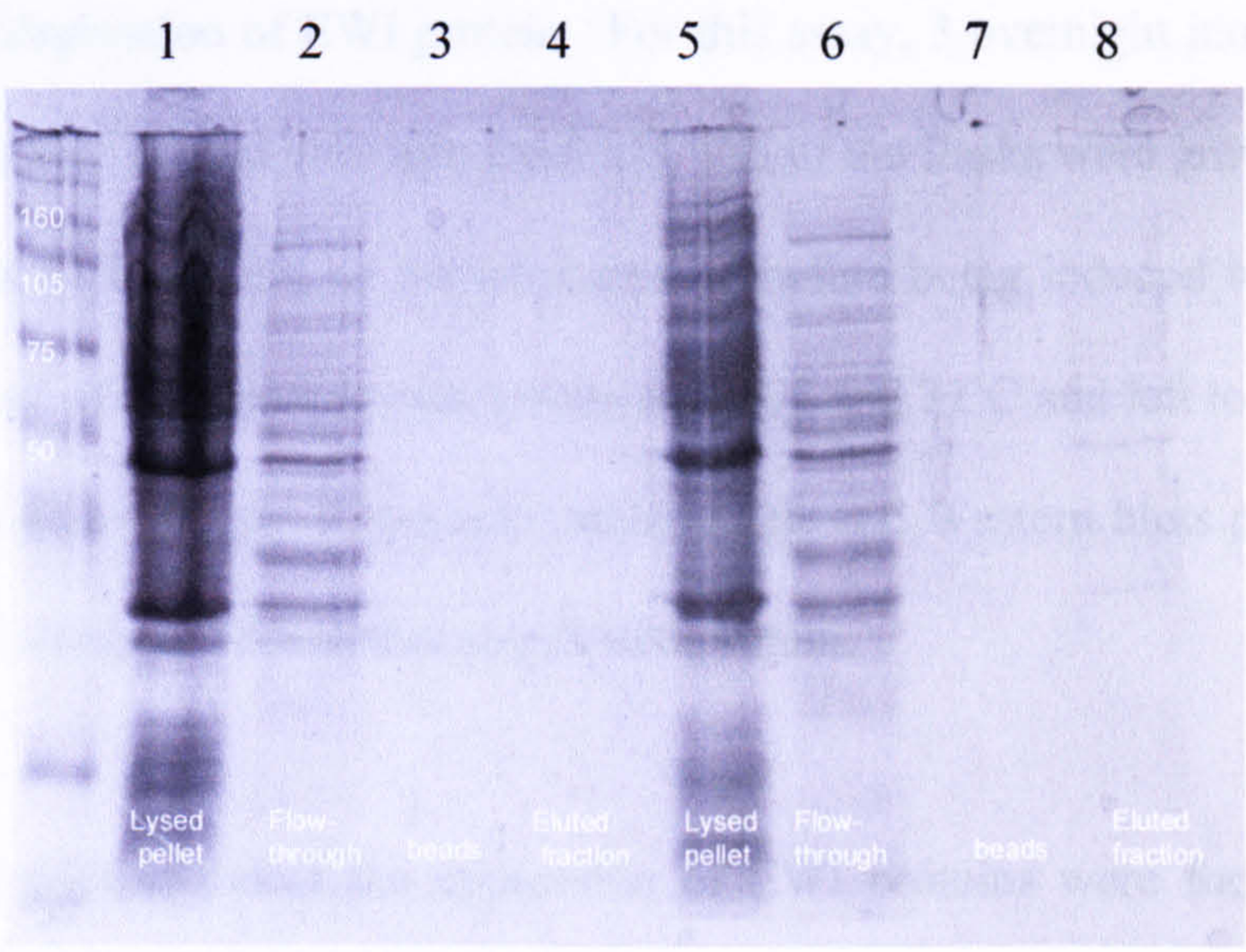


Lane #	1	2	3	4	5	6	7	8	9	10	11	12
[IPTG] M	0.1	0.1	0.1	0.1	1.0	1.0	1.0	1.0	0.5	0.5	0.5	0.5
time	2 hr	3 hr	4 hr	O/N	2 hr	3 hr	4 hr	O/N	2 hr	3 hr	4 hr	O/N

Figure 6.3. Coomassie stained SDS-PAGE gel showing whole cell lysates from the trial expression of EWI-F in BL21 cells. Correctly sequenced, full length EWI-2 and EWI-F in the pET-DEST 42 vectors were used to transform BL21 cells. Different concentrations of IPTG were used to induce protein expression and then the cells were left to express for 2 hours, 3 hours, 4 hours or overnight. Cells were harvested and whole cell lysates were run on SDS gels next to a non-induced control and stained with Coomassie. The Coomassie stained gel shown here is of EWI-F. [* = 135kDal band that could be EWI-F].

The same EWI-F plasmid was used to re-transform some fresh BL21 cells in order to scale up the protein production using these identified conditions. EWI-2 was also expressed in the same way to check the possibility that protein concentration was too low for detection in the small scale trial. Bacterial pellets weighing > 5 g were obtained from both preparations and the protein was purified by the denaturing His6-tag purification method. Samples eluted from the column were immediately tested for protein concentration but found not to contain protein. To verify that the protein had not been produced but then lost during the purification

procedure, samples from the lysed cell pellet, column flow-through, beads and eluted fractions were run on SDS-PAGE gels. Coomassie staining did not show correctly sized proteins (Figure 6.4) and Western-blots were negative for His6-tagged proteins.



Lane	1	2	3	4	5	6	7	8
Contents	lysed pellet	flow-through	beads	eluted fraction	lysed pellet	flow through	beads	eluted fraction
EWI	**EWI-F**				**EWI-2**			

Figure 6.4. Coomassie stained SDS-PAGE showing various stages of the EWI-F and EWI-2 purification procedure. Whole cell lysates from BL21 cells transfected with His6-tagged EWI-F or EWI-2 were loaded onto Ni-NTA beads (Qiagen). Samples from the purification procedure were run on an SDS-PAGE gel and stained in Coomassie.

6.3.3 Investigation on the effects of temperature on the production of EWI protein

Up to this point, all of the protein induction experiments had been carried out at 37°C. It was decided to investigate the effects of lowering the growth temperature during expression of EWI protein. For this assay, 3 overnight inoculums for each protein were diluted 1:40 into fresh LB. All of the flasks were grown for the first 2 hours at 37°C to bulk up bacterial density before being induced with 1mM IPTG. One flask of each protein was grown at 37, 25 and 22°C and left to grow for 2, 4 or 6 hours or overnight. Coomassie stained gels and Western blots probed with anti-His6 antibody did not detect any cloned protein.

All attempts to detect the expression of EWI proteins were focused on soluble protein expression or protein being secreted from the cells. It is possible that the proteins were being expressed but accumulating in insoluble inclusion bodies within the cell that were not solubilised during native purification. Lysing the cells using guanidinium/urea and thereby unfolding the proteins, will release the contents of inclusion bodies. However, Coomassie stained gels and Western blots of samples prepared in this way were negative and so EWI protein was not accumulating within inclusion bodies.

6.3.4 EWI protein expression in mammalian cells

Eukaryotic protein expression was the obvious next step after having exhausted the possibilities in the prokaryotic system. Although the yields of protein produced

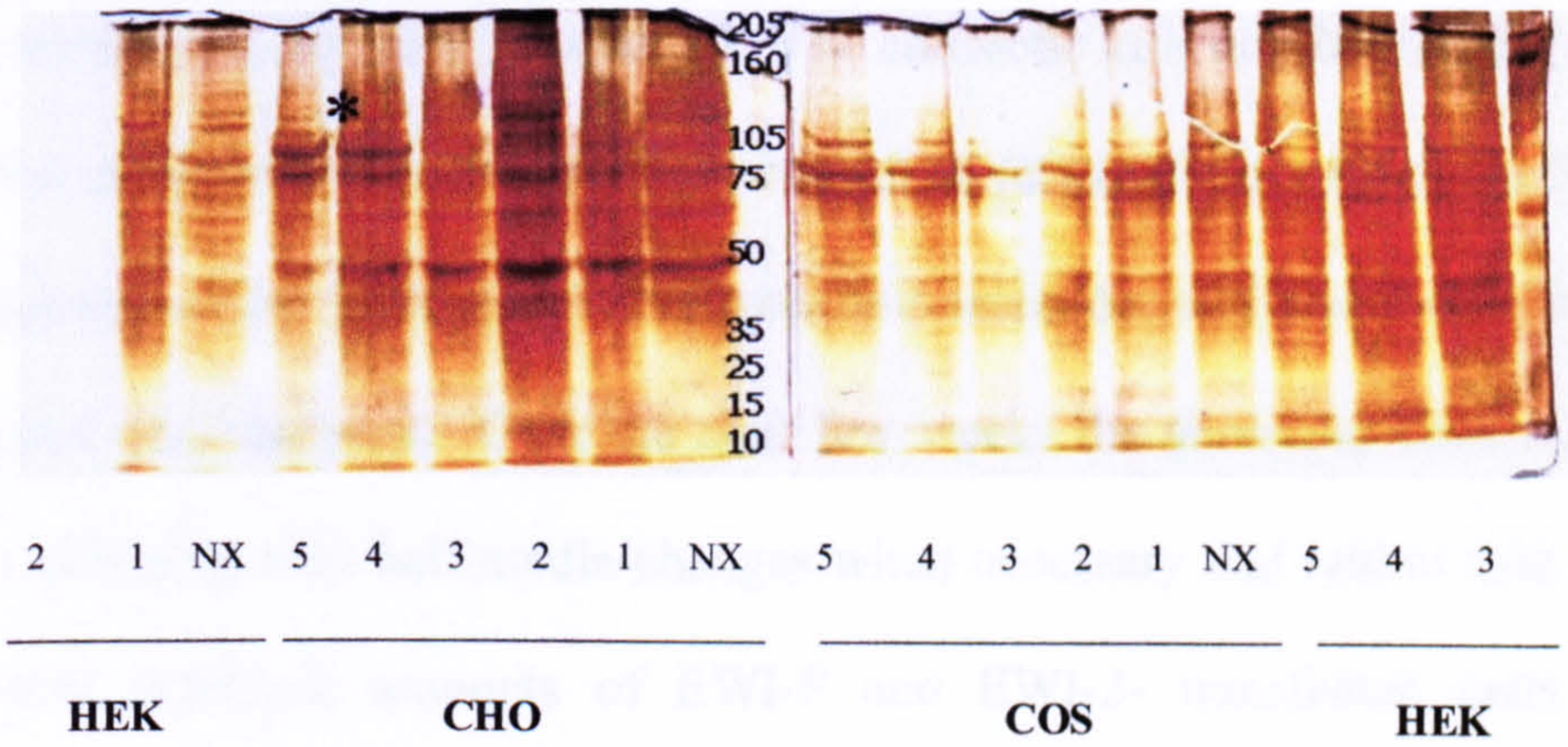
are generally lower from eukaryotic cells, these organisms have more sophisticated methods for producing proteins that have complex folding. Since the EWI proteins are of human origin, is it possible that they are too complex to be produced in prokaryotic cells. The same entry vectors used for the prokaryotic cloning were used again, this time in an LR clonase reaction with destination vector pcDNA-DEST40 (Figure 6.1). This vector is for cloning and expression of C-terminal fusion proteins in mammalian cells. Following the LR clonase reaction the reaction mixture was transformed into DH5α cells. Plasmid DNA was isolated in the usual way and analysed for the presence of the insert. A PCR screen, performed using the insert-specific primers, confirmed that the cloning had worked.

6.3.5 Lipofectamine transfection

In the first attempt at transfecting the EWI DNA, Lipofectamine Plus (Invitrogen) was used. Three cell lines: COS; CHO and HEK cells were transfected with 500ng and 1000ng of each DNA and an additional 2500ng of the larger EWI-F. After 24 hours the CHO and HEK cells had grown to ~ 90 % confluency. The COS cells, however, had only reached around 40 % confluency. After 48 hours, the supernatant was removed for protein assays. The cells were harvested and lysed for Western-blotting and silver staining. For each cell type there was a non-transfected control and 5 different transfections carried out (Table 6.5).

The supernatant was used in an immunoblot to detect His6-tagged protein. However, the immunoblot and Western blot did not recognise any His6-tagged

protein. Silver-stained gels of the cell lysates showed evidence of some EWI-F (105 kDal) in CHO cell lysates, sample number 4 and 5 in Figure 6.5. A protein of this molecular weight does not appear to be present in the non-transfected control.



#	1	2	3	4	5
DNA	EWI-2	EWI-2	EWI-F	EWI-F	EWI-F
[DNA]ng	500	1000	500	1000	2500

Figure 6.5. Expression of full length EWI-F in CHO, COS and HEK cells. Lipofectamine reagent (Invitrogen) was used to transfect full length EWI-F or EWI-2 DNA into CHO, COS and HEK cells. Whole cell lysates were run on SDS-PAGE gels and silver stained.

After preparing more DNA, the transfections were repeated using 2.5μg of each DNA and this time harvesting cells 2, 3, 4, 5 and 6 days following transfection but again no recombinant protein was detected with anti-His6 antibody. A final attempt at transient Lipofectamine transfection used up to 3000ng of DNA but no expression of the desired protein was detected.

6.3.6 Stable transfection using Lipofectamine

Stable transfection can also be achieved by using the same plasmid vector, by selection with the antibiotic, geneticin. CHO cells were used initially with the same transfection protocol as for transient transfection, but 48 hours post-transfection, geneticin was added. The antibiotic kills non-transfected cells that do not carry the plasmid containing resistance to geneticin. After the first week of growth in the presence of the antibiotic selection, cell death of non-transfected cells was observed. Over the next few weeks the surviving cells were cultured, performing only half media-changes when necessary and within four weeks there were sufficient amounts of EWI-F and EWI-2- transfected cells to test for expression (the fact that the cells grew in selective media suggests that transfection of the resistance containing plasmid had occurred). One plate of each cells transfected with each insert was left to propagate and the remaining cells were harvested in trypsin and lysed in order to assay for protein expression. Gel analysis and Coomassie staining did not show any bands of transfected proteins in the cells nor in the supernatant, and the negative result was confirmed with anti-His6 and anti-V5 probed Western blots.

A new batch of COS-7 cells were tested for transfection and expression of EWI proteins since these are chimpanzee cells that are more similar to human cells and should therefore be better equipped to express human protein. Both stable and transient expression experiments were performed in tandem, using Lipofectamine plus reagent. After 48hrs the cell lysates and supernatants from the transient transfection tested negative for EWI protein. Following addition of geneticin, any

non-transfected cells were left to be killed by the antibiotic and any transfected cells were cultured for 21 days. The assays were negative for transfected protein expression in both the stable and transiently transfected cells.

6.3.7 Calcium phosphate transfection of EWI-2 and EWI-F

Calcium phosphate transfection was attempted to rule out the possibility of Lipofectamine being incompatible with this DNA for whatever reason. 48 hours post-transfection the cells were harvested in the usual way. An immunoblot of the supernatants and a Western blot of the cell lysates was negative.

6.3.8 DNA electroporation of EWI-2 and EWI-F

Since chemical transformation experiments had proved unsuccessful, the decision was made to use electroporation methods to transfect the cells with DNA. CHO cells were used since they appeared to be the most robust. Following the electroporation event, the cells were left to grow for 24 hours before addition of geneticin, the selective agent, to a total concentration of 200μg/ml. Three days after the transfection, some plates were confluent and so were sub-cultured and the concentration of geneticin was increased to 400μg/ml. Cells on other plates were patchier in growth and seemed to be responding better to the antibiotic. More geneticin was added to these plates to a total volume of 400μg/ml but they were not sub-cultured. Half media changes were performed as necessary and after 3 weeks the remaining cells were harvested, lysed and tested for His6-tagged EWI protein. No protein was detected in the lysed cell pellets or in the media.

6.3.9 EWI domain expression in the pET system

Due to the unsuccessful attempts to express full length EWI proteins in Gateway™ vectors it was decided to attempt to clone into a different expression system. The pET32c vector was chosen (Figure 6.6) as this vector was available in the laboratory and is designed to be used in conjunction with Origami cells – a specially adapted *E. coli* strain, capable of forming disulphide bridges in the cytoplasm (see Discussion). It was also decided to clone and express double Ig-like domains rather than full length protein (Figure 6.7).

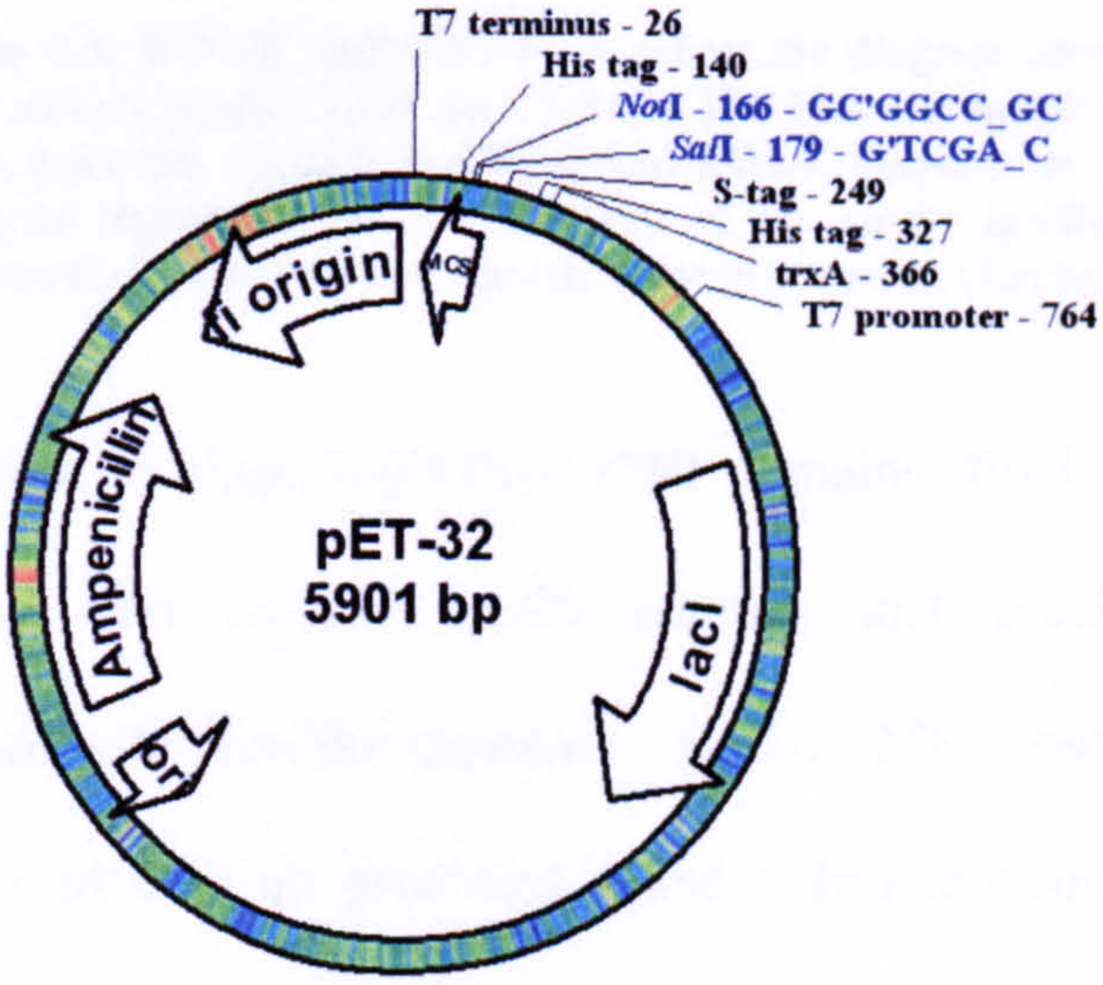


Figure 6.6. pET32c vector. The restriction sites used for cloning the insert are shown in blue (Not1 and Sal1). N- and C-terminal tags are also highlighted.

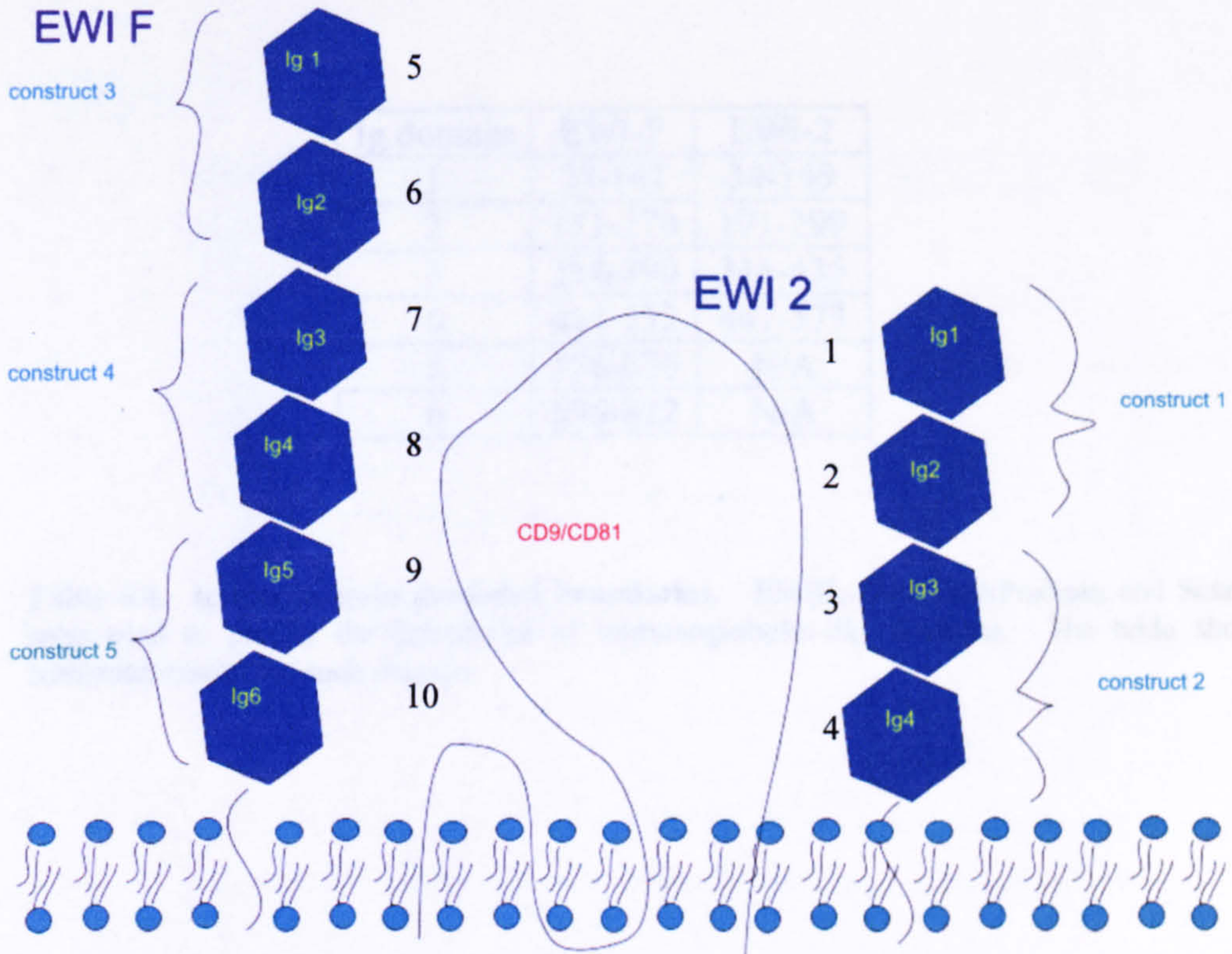


Figure 6.7. EWI-2 and EWI-F. A schematic diagram showing EWI-2 and EWI-F surrounding their primary ligand, CD9 (or CD81). The blue hexagonal shapes represent Ig-like domains, of which there are 4 and 6 in EWI-2 and EWI-F, respectively. The black lettering adjacent to the hexagons represents the nomenclature of the single Ig-like domain cloning, and the construct numbers (1-5) represent the nomenclature given when cloning double Ig-like domains.

In order to clone individual EWI domains, firstly, the full length sequence was pasted into on-line search engines that predicted the boundaries of the immunoglobulin-like domains. EMIL-EMI, InterProScan and ScanProsite were used and they all predicted 4 and 6 Ig-like domains for EWI-2 and EWI-F as shown in the InterPro annotation (Figure 6.8A and 6.9A, respectively and Table 6.4).

Ig domain	EWI-F	EWI-2
1	28-141	34-149
2	151-270	171-299
3	284-390	311-435
4	414-532	447-573
5	556-676	N/A
6	696-822	N/A

Table 6.4. Ig-like domain predicted boundaries. EMBL-EMI, InterProScan and ScanProsite were used to predict the boundaries of immunoglobulin-like domains. The table shows the combined results for each domain.

However, the boundaries for each domain differed slightly (Table 6.4). Figures 6.8B and 6.9B show the final predicted domains for EWI-2 and EWI-F, respectively, as a result of the 3 searches pooled together. The forward oligonucleotide starting positions for each domain are shown. The pET-32c vector offers a number of restriction enzyme sites in its multiple cloning site. In order to determine which restriction enzymes would be compatible with my protein of interest, an on-line tool that searches for restriction sites within the chosen DNA was used (http:www4.carolina.com/cgi-bin/). The full length DNA sequences of EWI-2 and EWI-F were cut and pasted into this site and the results revealed that Sal1 and Not1 did not cut within these sequences, and so these restriction sites were chosen for cloning the EWI proteins into pET32c.

EWI-2

A.

B.

| Domain 1 →
MGALRPTLLPPSLPLLLLLMLGMGCWAREVLVPEGPLYRVAGTAVSISCNVTGYEGPAQQ

NFEWFLYRPEAPDTALGIVSTKDTQFSYAVFKSRVVAGEVQVQRLQGDAVVLKIARLQAQ

| Domain 2 →
DAGIYECHTPSTDTRYLGSYSGKVELRVLPDVLQV*SAAPPGPRGRQAPTSPPRMTVHEGQ*

ELALGCLARTSTQKHTHLAVSFGRSVPEAPVGRSTLQEVVGIRSDLAVEAGAPYAERLAA

|

*GELRLGKEGTDRYRMVVGGAQAGDAGTYHCTAAEWIQDPDGSWAQIAEKRAVLAH*VD**VQT**
Domain 3 →
LSSQLAVTVGPGERRIGPGEPLELLCNVSGALPPAGRHAAYSVGWEMAPAGAPGPGRLVA

QLDTEGVGSLGPGYEGRHIAMEKVASRTYRLRLEAARPGDAGTYRCLAKAYVRGSGTRLR

| Domain 4 →
*EAA***SARSRPLPVHVREEGVV***LEAVAWLAGGTVYRGETASLLCNISVRGGPPGLRLAASWW*

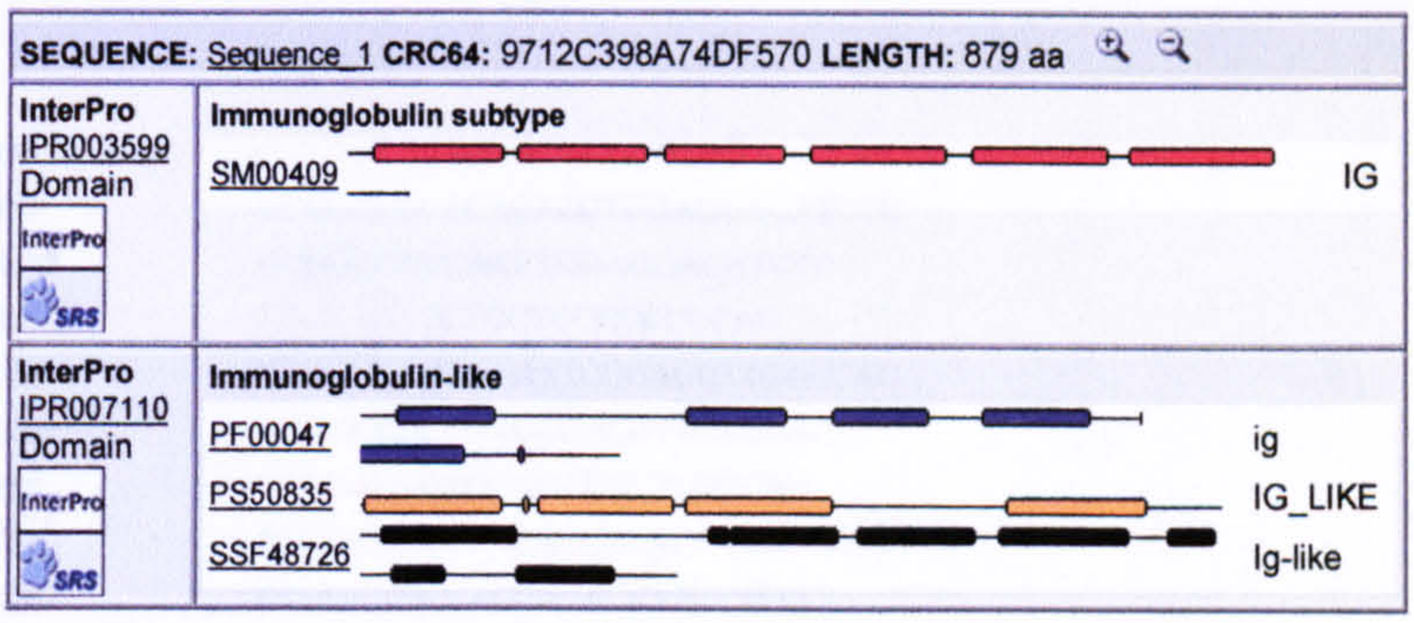
VERPEDGELSSVPAQLVGGVGQDGVAELGVRPGGGPVSVELVGPRSHRLRLHSLGPEDEG

VYHCAPSAWVQHADYSWYQAGSARSGPVTVYPYMHALDTLFVPLL

Figure 6.8. Immunoglobulin-like domain predictions and primer design for EWI-2. Figure A shows the results from one of the Ig-like domain searches using InterPro, and shows that there are 4 Ig-like domains (red). The other coloured domains predication are less probable matches. B. shows amino acid sequence with each domain highlighted.

EWI-F

A.



B.

| Domain 1 →
MGRLASRPLLLALLSLALCRGRVVRVPTATLVRVVGTELVIPCNVSDYDGPSEQNFDWSF

SSLGSSFVELASTWEVGFPAQLYQERLQRGEILLRRTANDAVELHIKNVQPSDQGHYKCS

| Domain 2 →
*TPSTDATVQGNYEDTVQVKVL***ADSLHVG***PSARPPPSLSLREGEPFELRCTAASASPLHTH*

LALLWEVHRGPARRSVLALTHEGRFHPGLGYEQRYHSGDVRLDTVGSDAYRLSVSRALSA

| Domain 3 →
*DQGSYRCIVSEWIAEQGNWQEIQEKAVEVA***TV**VIQPSVLRAAVPKNVSVAEGKELDLTCN

ITTDRADDVRPEVTWSFSRMPDSTLPGSRVLARLDRDSLVHSSPHVALSHVDARSYHLLV

| Domain 4 →
RDVSKENSGYYYCHVSLWAPGHNRSWHKVA**EAVSSPAGVGVTWL***EPDYQVYLNASKVPGF*

ADDPTELACRVVDTKSGEANVRFTVSWYYRMNRRSDNVVTSELLAVMDGDWTLKYGERSK

QRAQDGDFIFSKEHTDTFNFRIQRTTEEDRGNYYCVVSAWTKQRNNSWVKSKDVFSKPVN

| Domain 5 →
*IFW***A**LEDSVLVVKARQPKPFFAAGNTFEMTCKVSSKNIKSPRYSVLIMAEKPVGDLSSPN

ETKYIISLDQDSVVKLENWTDASRVDGVVLEKVQEDEFRYRMYQTQVSDAGLYRCMVTAW

| Domain 6 →
SPVRGSLWREAATSLSNPI**EIDFQT***SGPIFNASVHSDTPSVIRGDLIKLFCIITVEGAAL*

DPDDMAFDVSWFAVHSFGLDKAPVLLSSLDRKGIVTTSRRDWKSDLSLERVSVLEFLLOV

Figure 6.9. Immunoglobulin-like domain predictions and primer design for EWI-F. Figure A shows the results from one of the Ig-like domain searches, and shows that there are 6 Ig-like domains (red) The other colours in A show less probable Ig-like domain matches. B. shows amino acid sequence with each domain highlighted.

Forward primers were designed with the *SalI* restriction site at the 5'end (GTCGACA). The pET32c vector has N-terminal tags and so there was no requirement for an initiation codon. The 3'primers were designed with the *NotI* restriction site at their 3' end (GCGGCCGC) followed by a stop codon for

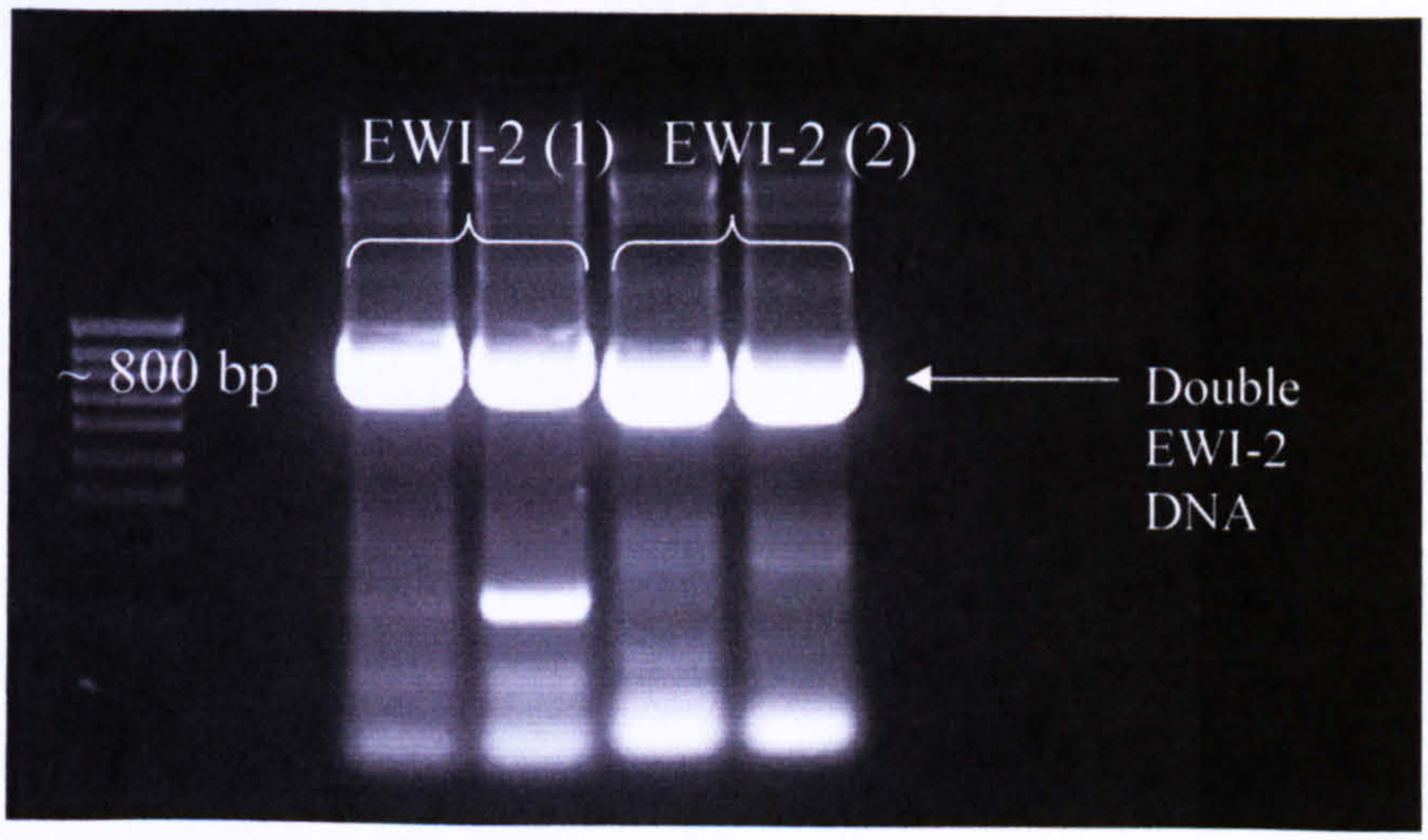
termination and no addition of a C-term His6-tag (primer 3) or without the stop codons to allow transcription of the C-terminal His6-tag, (primer b3). The stop codons used were the same as the ones in the native proteins, EWI-2 uses TCA and EWI-F, CTA (Figure 6.10).

FM15	GTCGACAATGGGCGCCCTCAGG
FM13	GCGGCCGCTCACACCTGGAGGACATCTG
FM1b3	GCGGCCGCCACCTGGAGGACATCTG
FM25	GTCGACATCTGCTGCCCCCCCAG
FM23	GCGGCCGCTCAGATCCGACGTTCACCAG
FM2b3	GCGGCCGCGATCCGACGTTCACCAG
FM35	GTCGACAGTGCAGACGCTGTCCAG
FM33	GCGGCCGCTCACACCACACCTTCCTCCC
FM3b3	GCGGCCGCCACCACACCTTCCTCCC
FM45	GTCGACAAGTGCCCGTTCCCGGC
FM43	GCGGCCGCTCAAGGCACAAATAGGGTGTCC
FM4b3	GCGGCCGCAGGCACAAATAGGGTGTCC
FM55	GTCGACAATGGGGCGCCTGGCCT
FM53	GCGGCCGCCTAGCCCACGTGCAGGGAG
FM5b3	GCGGCCGCGCCCACGTGCAGGGAG
FM65	GTCGACAGCCGACTCCCTGCACG
FM63	GCGGCCGCCTACACGGTGGCAACTT
FM6b3	GCGGCCGCCACGGTGGCAACTT
FM75	GTCGACAACCGTGGTGATCCAGCC
FM73	GCGGCCGCCTATAGCCAGGTCACACCCA
FM7b3	GCGGCCGCTAGCCAGGTCACACCCA
FM85	GTCGACAGTGTCTTCCCCAGCTGG
FM83	GCGGCCGCCTATGCCCAAAATATGTTAACAGG
FM8b3	GCGGCCGCTGCCCAAAATATGTTAACAGG
FM95	GTCGACAGCATTAGAAGATTCCGTGC
FM93	GCGGCCGCCTAGGTTTGGAAGTCTATCTCAA
FM9b3	GCGGCCGCGGTTTGGAAGTCTATCTCAA
FM105	GTCGACAGAGATAGACTTCCAAACCTC
FM103	GCGGCCGCCTAGGGATACTTGAAGGCGTT
FM10b3	GCGGCCGCGGGATACTTGAAGGCGTT

Figure 6.10. Oligonucleotide design for EWI domain expression. Forward primer names end in 5 and reverse primers end in 3 (with a "stop" codon) or 3b (without a "stop" codon) depending on whether or not a C-terminal His6-tag is required. For shortened domain nomenclature, refer to Fig. 6.1. *Sal*1 restriction sites are shown in red and *Not*1 in blue.

Double Ig-like domains were PCR amplified using the Highfidelity Reddy mix (Abgene), the specific primers and correctly sequenced full length EWI DNA in TOPO vectors as the template. The PCR products were run on 1 % high purity agarose gels, Figure 6.11 and Table 6.5.

A.



B.

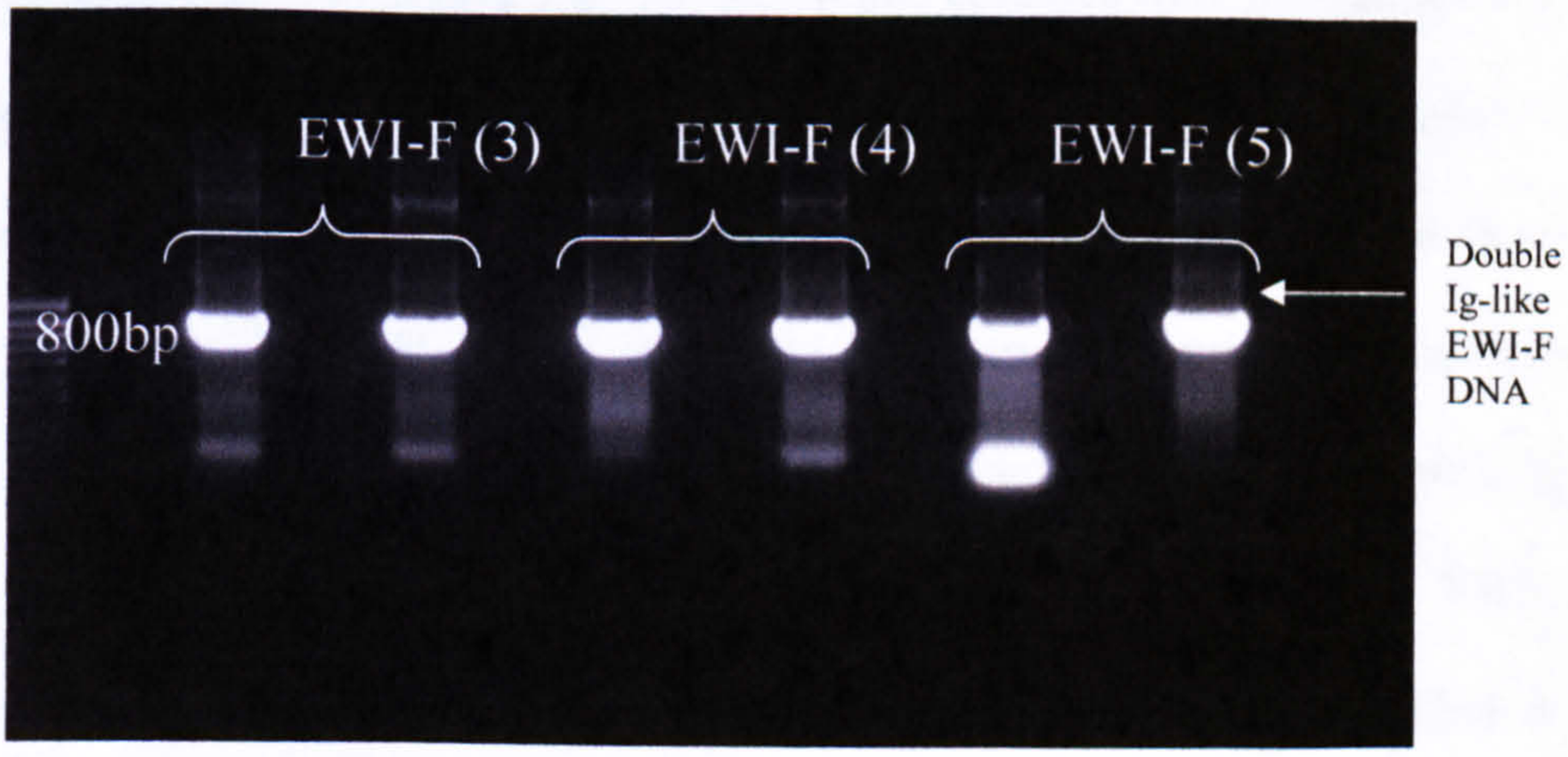


Figure 6.11. PCR amplification of EWI-2 and EWI-F double Ig-like domains. Full length EWI DNA from Gateway™ experiments was used as a template and the double Ig-like domains were amplified by PCR. Bands were observed at the correct height (800bp). N.B. DNA from EWI-2 (1) and EWI-F (5) with unidentified, smaller sized bands were not used again for the remainder of the cloning experiments.

PCR mix components (µl)	EWI-2	EWI-F
Reddy mix	12.5	12.5
Forward primer	2	2
Reverse primer	2	2
DNA template	6	2
dH_2O	2.5	6.5
Annealing temp (˚C)	55	60

Table 6.5. PCR conditions for initial cloning of Ig-like double domains. The table lists the conditions that gave single, discrete bands at ~ 800bp for each EWI double domain.

Bands at 800bp were cut, the DNA was purified from the gel and the A-addition reactions were performed before TOPO-cloning. The plates contained X-gal and so blue/white screening could be carried out. White colonies were picked and PCR screened using specific oligonucleotides before being digested with *Eco*R1 to confirm the presence of the insert in the TOPO vector (Figures 6.12 and 6.13). Positive colonies were sequenced using the T7 terminator primer. Correctly sequenced DNA and pET32c vector DNA were cut using *Sal*I and *Not*I restriction enzymes. It was decided to continue with a couple of clones to begin with. EWI-F (5) and EWI-2 (2), the 2 membrane distal domains, were chosen because they are predicted to be the domains involved in the binding to CD9 and CD81.

*Sal*I is a less efficient cutter and so was incubated with the DNA for 4 hours before addition of *Not*1 and incubation for a further hour. After the incubation period loading buffer was added and the samples were run on 1 % high purity agarose gels under normal conditions (Figure 6.14). Positive bands at around 800bp were cut out as well as the ~ 6000bp band of vector DNA. The DNA was purified from the gel and ligations of the pET32c vector and EWI double Ig-like domain DNA were performed.

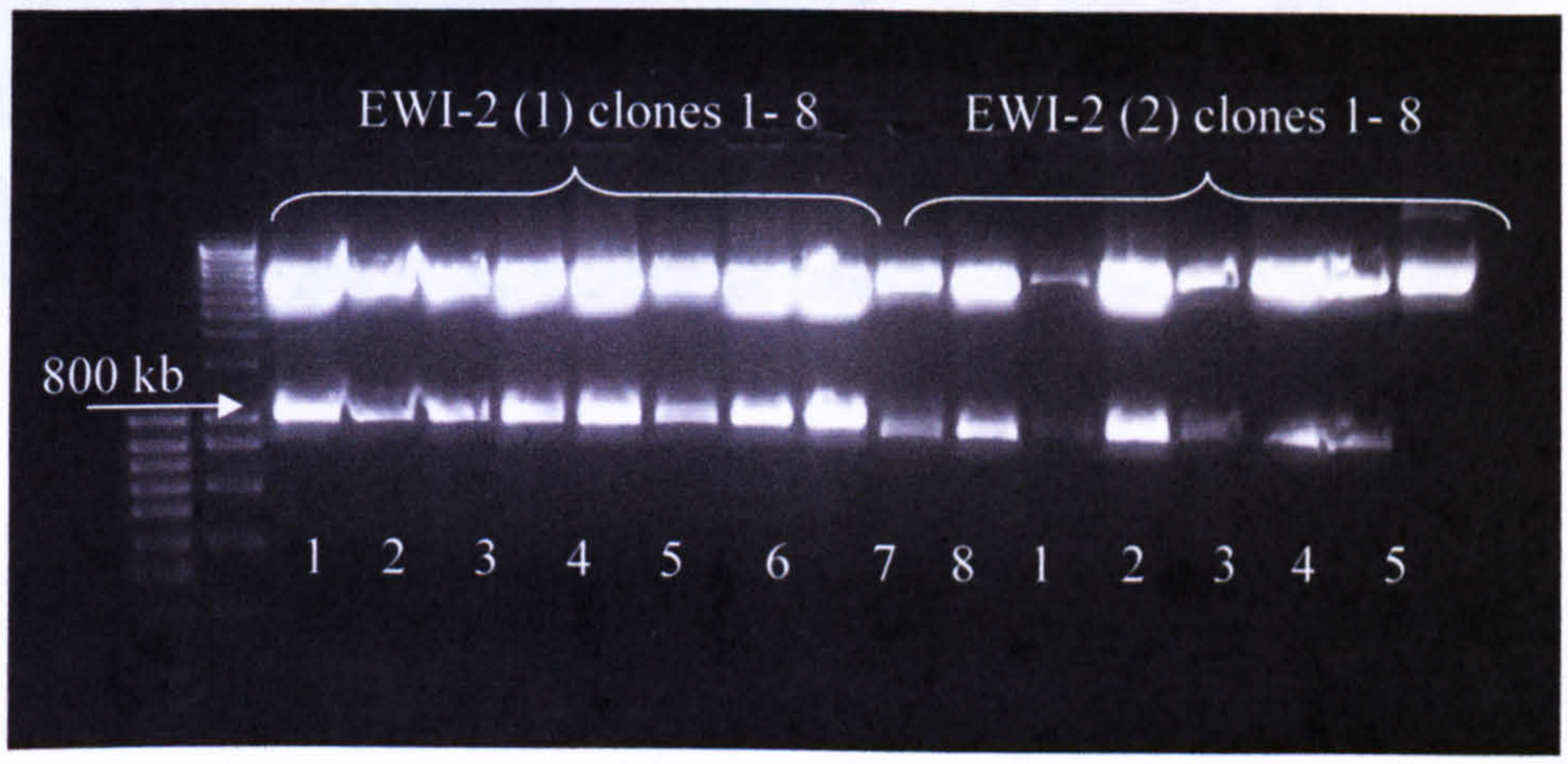


Figure 6.12. *Eco*R1 digest of EWI-2 double domains. The double Ig-like DNA was cloned into the TOPO vector, transformed into TOP10 cells and the plasmid DNA from white colonies that grew was purified. The miniprepped DNA was cut with *EcoRI* (3 hours at 37˚C) to confirm the presence of the insert.

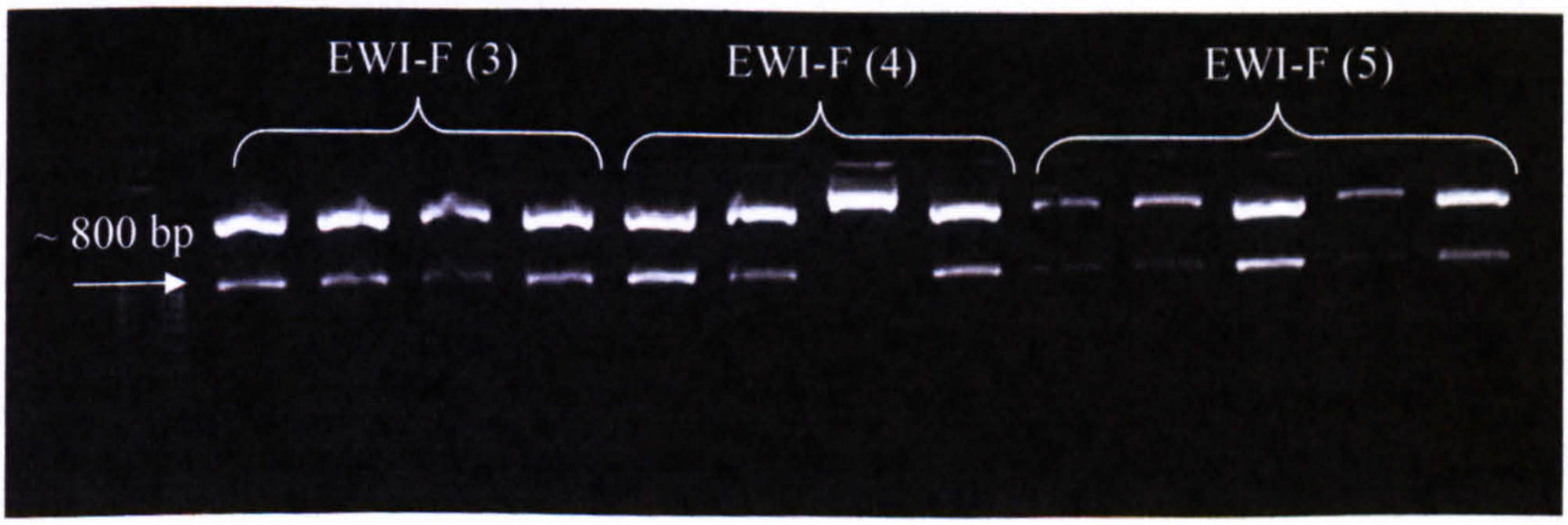


Figure 6.13. *Eco*R1 digest of EWI-F double domains. The double Ig-like DNA was cloned into the TOPO vector, transformed into TOP10 cells and the plasmid DNA from white colonies that grew was purified. The miniprepped DNA was cut with *EcoRI* (3 hours at 37˚C) to confirm the presence of the insert.

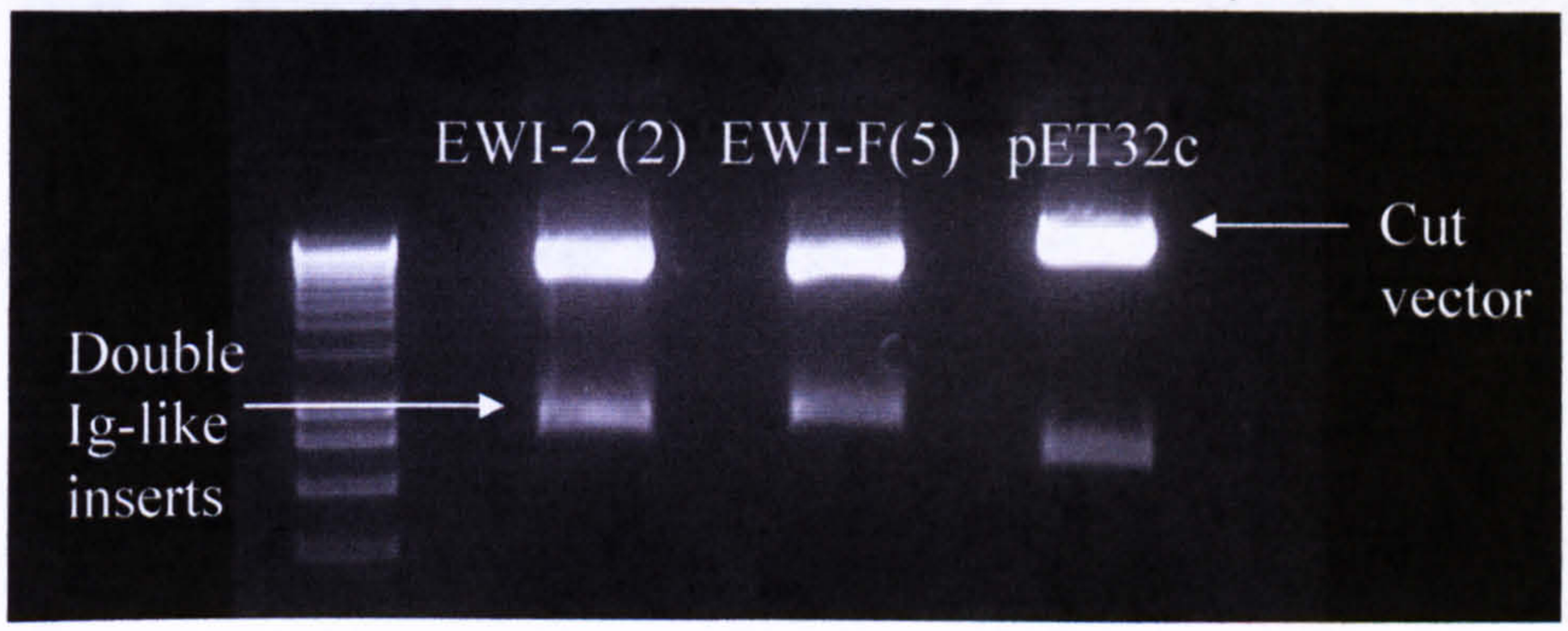


Figure 6.14. Double digest of EWI inserts in TOPO vector and pET32c vector. DNA was first cut with *Sal*1 for 3 hours followed by addition of *Not*1 for 1 hour. The reaction mixtures were run on high purity agarose gels and the EWI double domains (800bp) along with the pET32 c vector (~ 6000bp) were purified from the gel.

The ligation reaction was transformed into Omnimax™ cells (Invitrogen) and plated on carbenicillin selective plates in the presence of X-gal. White colonies were PCR screened using specific primers. Positive colonies were picked, the plasmid was purified and another double digest was performed to confirm the presence of the inserts (Figure 6.15) before being transformed into the Origami or BL21 for protein expression.

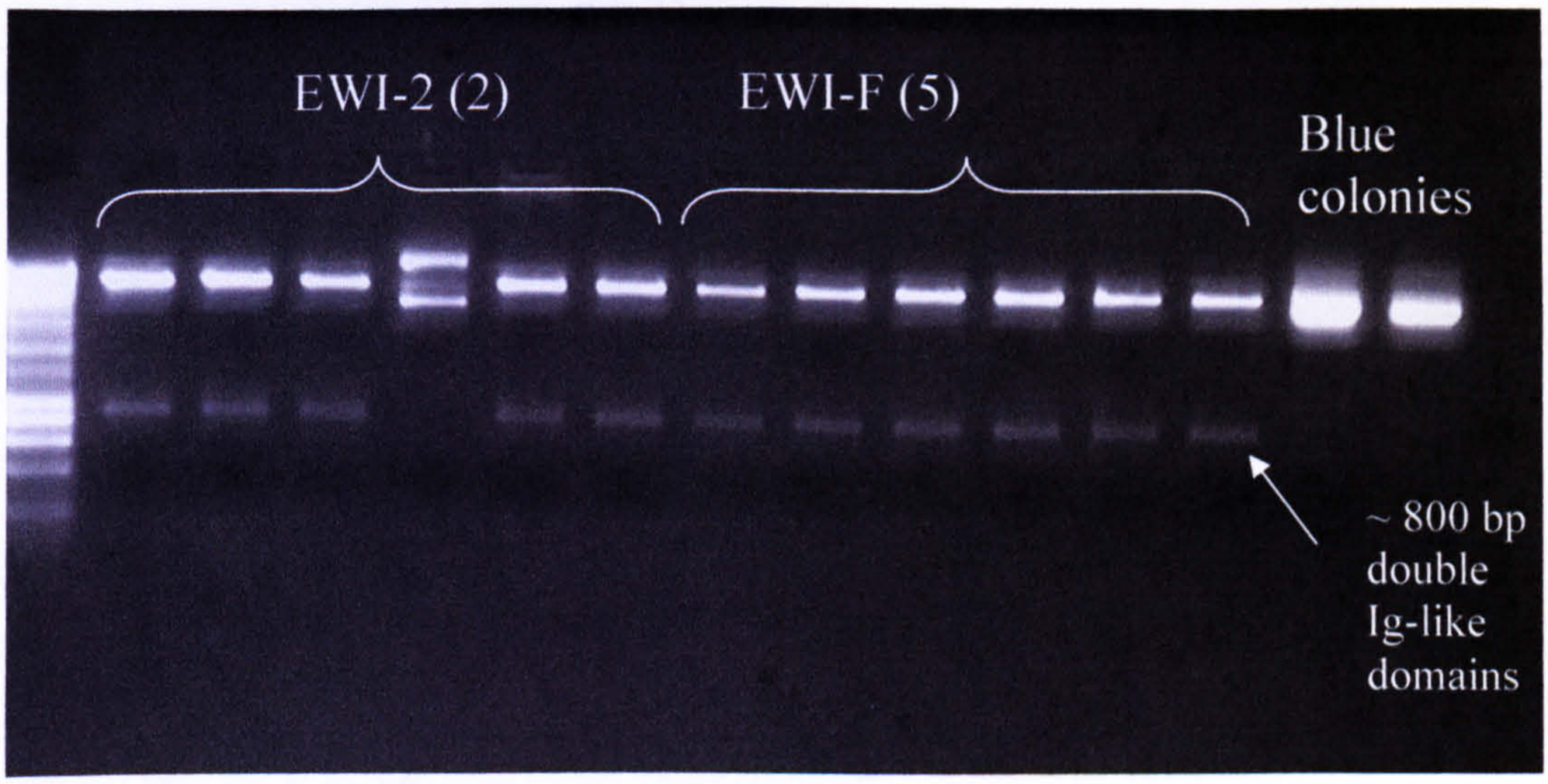


Figure 6.15. Restriction digest of ligated clones. Following ligation, 6 white colonies and one blue colony from each clone were picked and digested with *Sal*1 and *Not*1 to confirm the presence of the insert.

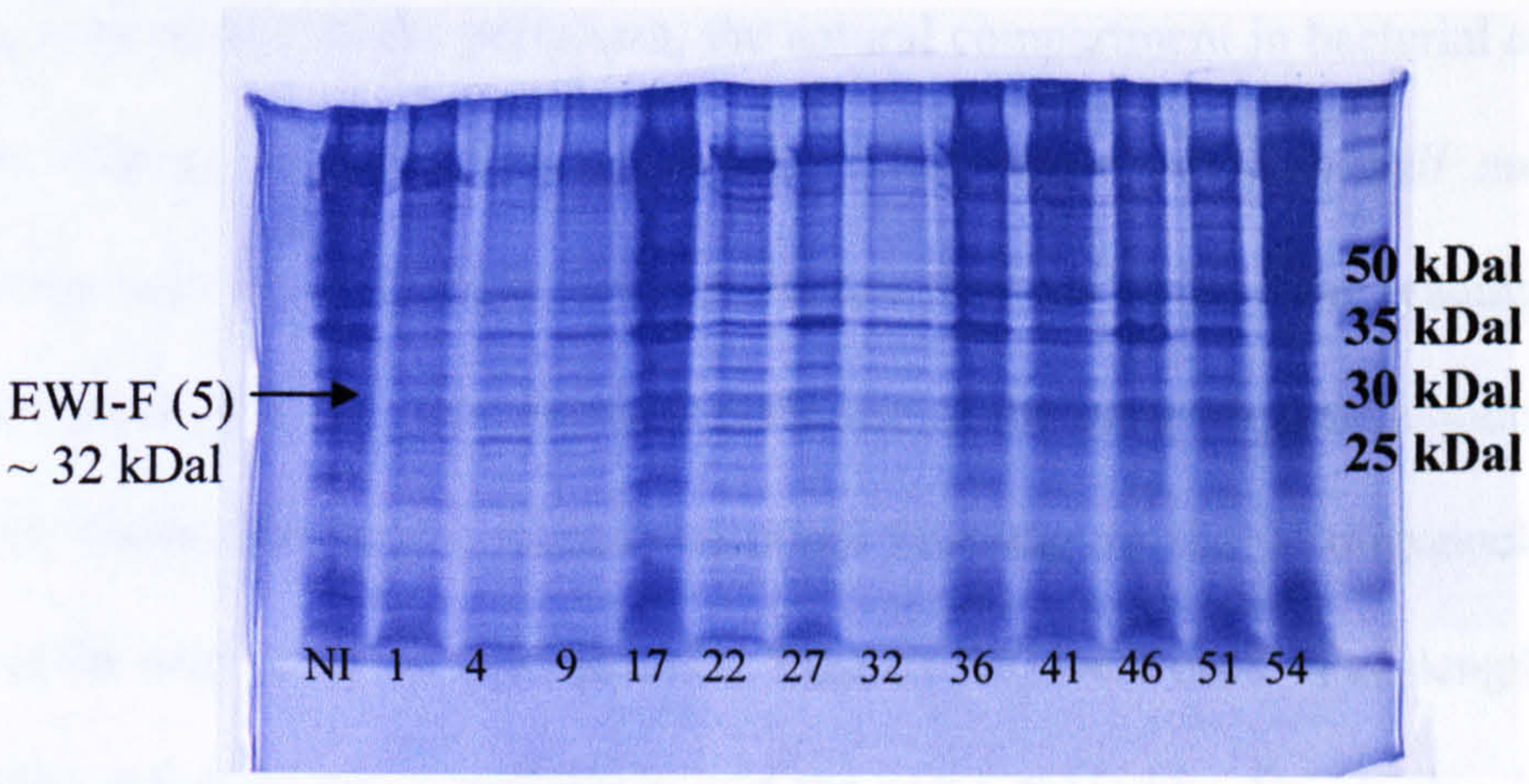
Expression of EWI-F (5) was attempted first. Origami cells were transformed with pET-32c-EWI-F (5), and colonies were picked, resuspended in 10ml of LB with selection and grown overnight. Overnight cultures were added to flasks containing 400ml of fresh LB with selection. One flask was grown at 25˚C and the other at 37˚C. The OD of the cultures was measured against a blank of LB alone and when cultures reached 0.2, 0.4 and 0.8 OD, 5ml samples were taken out and induced with 0.1, 0.5 or 1.0mM IPTG. Cultures were grown for 2 hours, 4 hours or overnight giving a total of 54 different growth conditions as well as non-induced controls for each OD reading at each temperature.

The cells were pelleted and lysed in 200µl of BugBuster and protease inhibitor cocktail. 100µl of whole cell lysate for each sample were loaded onto nitrocellulose membranes using a 96 well manifold. The membranes were probed

with anti-His6 antibodies but there was no recognition of any His6-tagged protein. Coomassie stained gels of whole cells lysates did not reveal any protein bands at the right size that were upregulated in comparison to the non-induced control (Figure 6.16). To ensure that the Origami cells were not ejecting the plasmid DNA after induction, induced and non-induced cells were grown in standard conditions (37°C, 0.5mM IPTG, 4 hour growth). The cells were pelleted and the plasmid DNA was extracted. Using specific primers to PCR amplify the DNA it was shown that the plasmid DNA was still present in the induced control.

The inserts were transformed into BL21 cells to see if this strain could express the double Ig-like domains. A full expression screen, as described for expression in Origami cells, was carried out but no protein was detected.

A.



B.

	OD_{600} = 0.2			OD_{600} = 0.4			OD_{600} = 0.8			
37 ° C	1	2	3	10	11	12	19	20	21	**0.1 mM**
	4	5	6	13	14	15	22	23	24	**0.5 mM**
	7	8	9	16	17	18	25	26	27	**1.0 mM**
25 ° C	28	29	30	37	38	39	46	47	48	**0.1 mM**
	31	32	33	40	41	42	49	50	51	**0.5 mM**
	34	35	36	42	44	45	52	53	54	**1.0 mM**

[IPTG]

Figure 6.16. SDS-PAGE analysis of pET-32c-EWI-F (5) expression in Origami cells. A large protein expression screen was performed, using the conditions are listed in **B**. Cells were pelleted by centrifugation at 400g for 3 minutes before being lysed by the addition of 200µl of BugBuster. 100µl of the lysate was used in an immunoblot with anti-His6 antibody (1/1000 dilution in blocking buffer) followed by anti-mouse IgG-HRP (1/3000 in blocking buffer) (Figure not shown). 20µl of the remaining cell lysate from a selection of conditions were run on SDS-PAGE and Coomassie stained (**A**).

6.3.10 Protein expression using pET26b

The pET26b vector has a pelB leader sequence for targeting of the newly synthesised peptide to the periplasm, the natural compartment in bacterial cells for protein folding to occur (Figure 6.17). This vector contains *SalI* and *NotI* restriction sites in the multiple cloning site and they are in the same reading frame as the ones in pET32c, thus permitting the use of the same oligonucleotides. The pET26b vector has only a carboxy-terminal His6-tag so the 3' oligonucleotides without the stop codon (FM b3 primers, Figure.6.10) were used. Full length DNA in TOPO plasmids were used again and the same procedure for cloning into the pET32c vector was used for the pET26b. Construct 5 was the simplest to clone and was transformed into BL21 cells. A large protein expression analysis was carried out as described previously, and the periplasmic fraction was isolated by osmotic shock. 100µl of each sample was loaded onto nitrocellulose membrane with the aid of a manifold. When the blot was probed with the anti-His6 antibody, lysates from one of the conditions produced positive spots (Fig. 6.18). These corresponded to overnight growth at 25°C and induction at a culture OD_{600} of 0.8. The concentration of IPTG used for induction appeared not to be a factor in this assay.

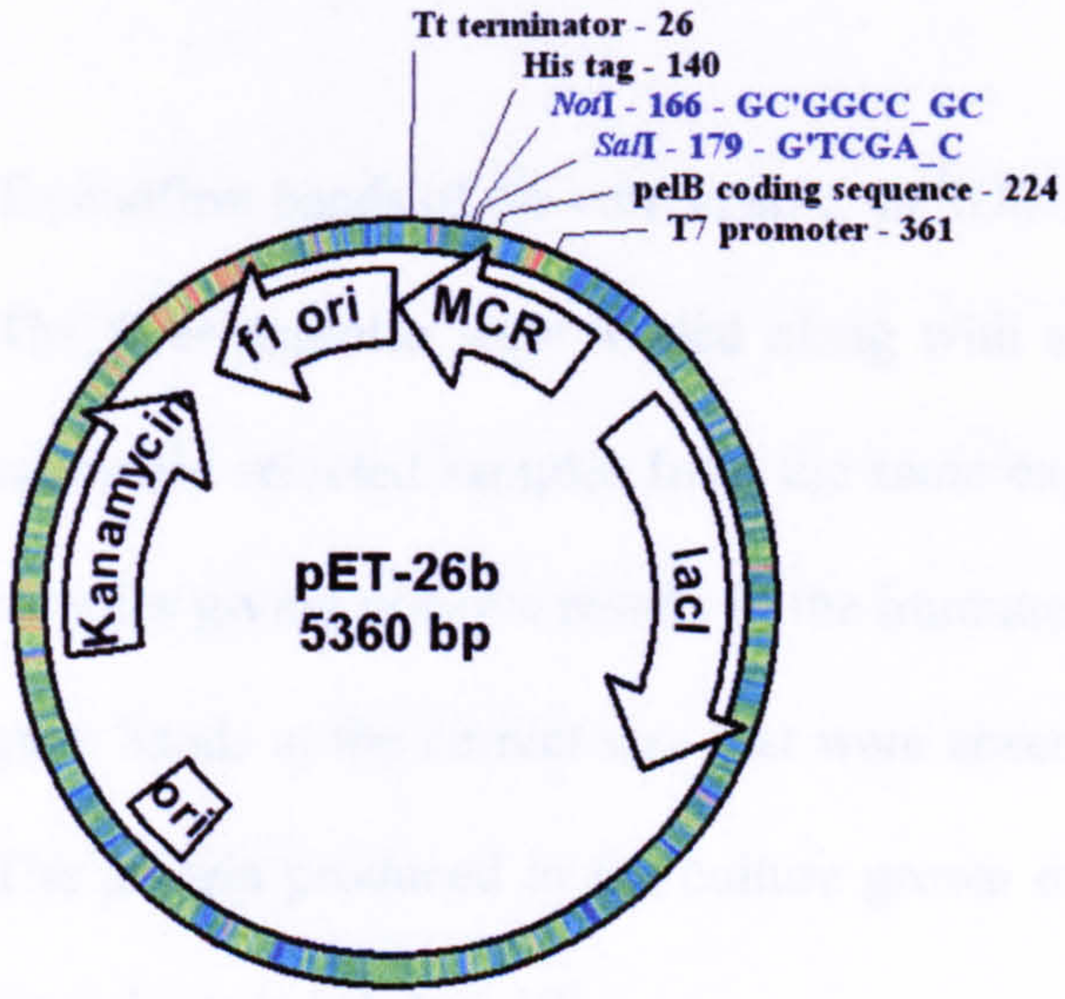


Figure 6.17. pET26b vector. The restriction sites used for cloning the insert are shown in blue (*Not*1 and *Sal*1). N- and C-terminal tags are also highlighted.

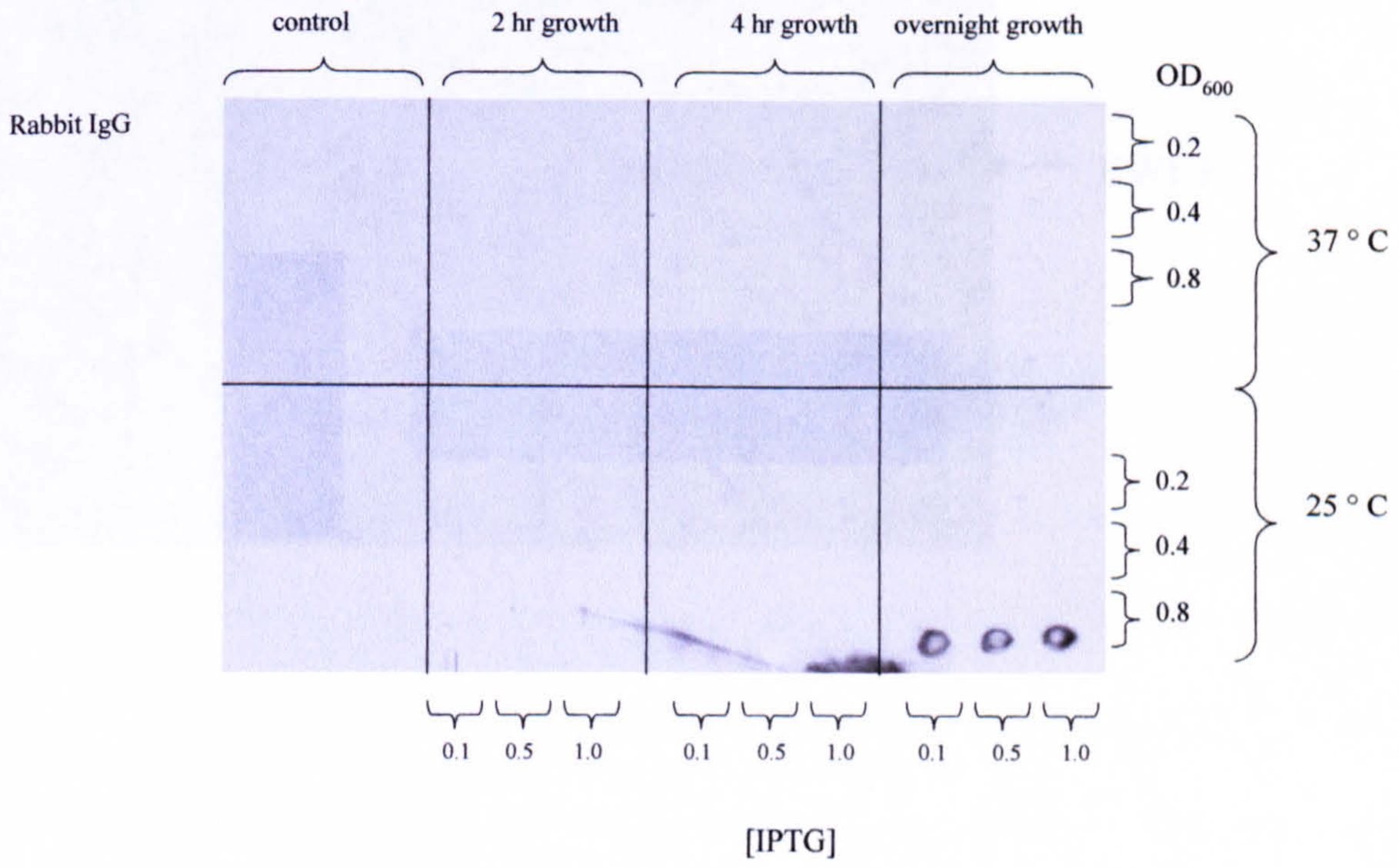


Figure 6.18. Immunoblot of periplasmic fraction of BL21 cells transformed with pET-26b-EWI-F (5). A large protein expression optimisation experiment was performed, the variables being temperature, concentration of IPTG used to induce the plasmid expression, length of growth time and OD 600 of the culture before induction. The periplasmic fractions were isolated and immunoblotted with an anti-His6 antibody. Three spots appeared, corresponding to cultures grown at 25°C overnight after being induced at OD600 = 0.8 and induction with [IPTG] = 0.1, 0.5 or 1.0mM.

To confirm bands of the correct size, an SDS-PAGE gel was run and silver stained. The three samples were loaded along with a non-induced control and five other randomly selected samples from the same experiment. Surprisingly, not only the samples giving positive results in the immunoblot, but also all of the others tested, gave bands at the correct size that were absent in lane 4, the non-induced control. The protein produced in the culture grown overnight at 37°C (lane 4) appears to have degraded (Fig. 6.19).

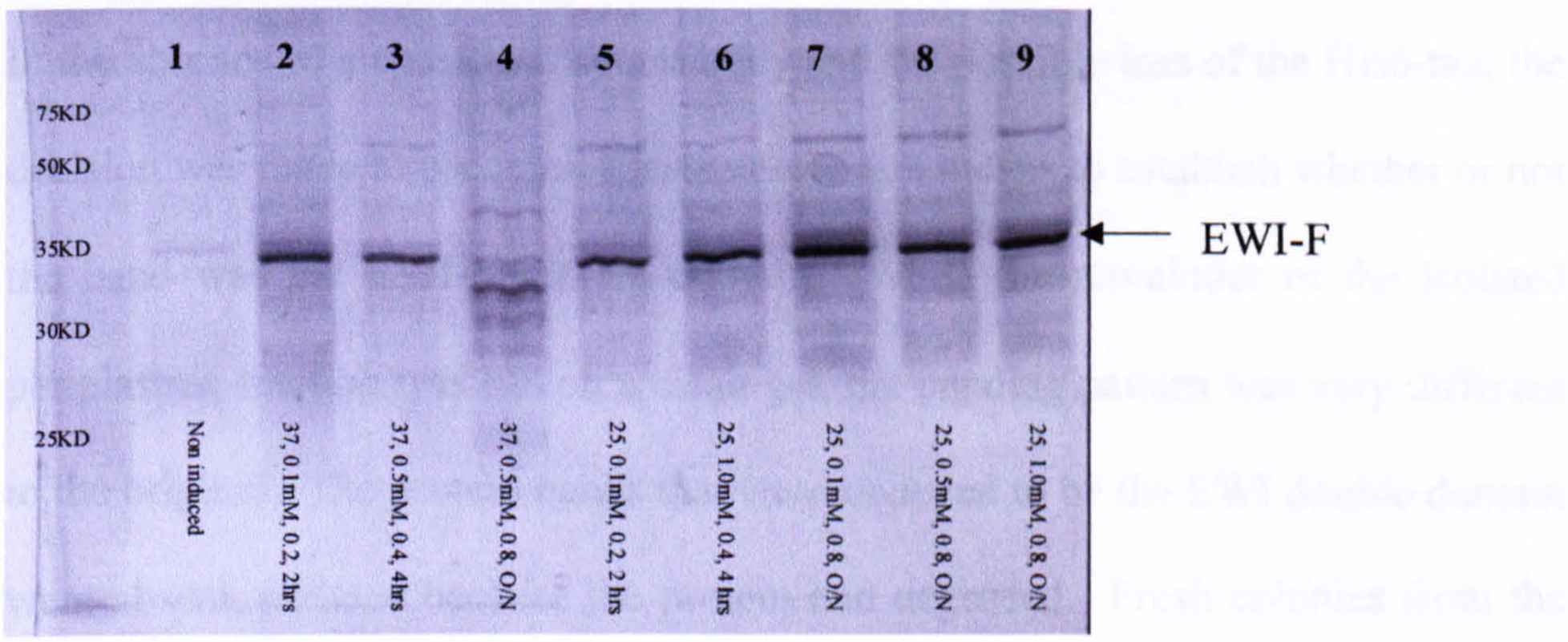


Figure 6.19. Silver-stained gel of the 3 samples that were recognised by the anti-His6 immunoblot. Lanes 7-8 are the samples recognised by the immunoblot and the dark band is at the right height to be EWI-F (5) double domain (~ 35 kDal).

To further verify the results, a Western blot was performed the following day using an anti-His6 antibody. The anti-His6 antibody blot was only affirmative for the positive control and not for any of the samples. One can only presume that the protein tag is unstable and is degraded in time (the reminder of the periplasmic fractions had been through a freeze-thaw cycle in between the original immunoblot and the Western blot).

6.3.6 Mass Spectrometry

In the absence of a specific EWI antibody and the possible loss of the His6-tag, the decision was made to use mass spectrometry as a means to establish whether or not the band was the EWI-F protein domain. When the remainder of the isolated periplasmic fraction was run on a clean gel, the banding pattern was very different to the original. The protein bands that were believed to be the EWI double domain were absent, perhaps because the protein had degraded. Fresh colonies from the same transformation reaction were picked and following overnight inoculation, cultures were grown in what appeared to be optimum conditions; 25°C, induced with 0.5mM IPTG when $OD_{600} = 0.8$ and then grown overnight. The periplasmic fraction was isolated as before and the concentration of protease inhibitors used was doubled. Mass spectrometry requires samples to be run on clean SDS-PAGE gels to cut down on contamination which could obscure the results. Instead of using the general Coomassie, Brilliant Blue G-Colloidal concentrate was used. The first attempt to make more protein was unsuccessful – there was no band in the induced that was absent from the non-induced at the correct weight. A second attempt gave better results, although still not as clear as the initial experiment.

There appeared to be far more background periplasmic protein in this fraction (Figure 6.20).

The band highlighted (*) in Figure 6.20 was carefully excised and placed into a sterile, 1.5ml microcentrifuge tube. The gel was covered with dH_2O and sent to the Mass Spectrometry Sequencing Department at the Aberdeen Proteome Facility, University of Aberdeen. The peptide searches were performed on human databases only. A probable identification was recorded but it was not EWI-F but a G protein-coupled receptor, GPR35. This protein was found on two independent searches, MS-Fit and Mascot and had 25 % sequence coverage to one particular peptide fragment. The molecular weight of GPR35 is 34kDal which is similar to that of EWI-F (6) but BLAST searches between GPR35 and EWI-F showed no significant homology.

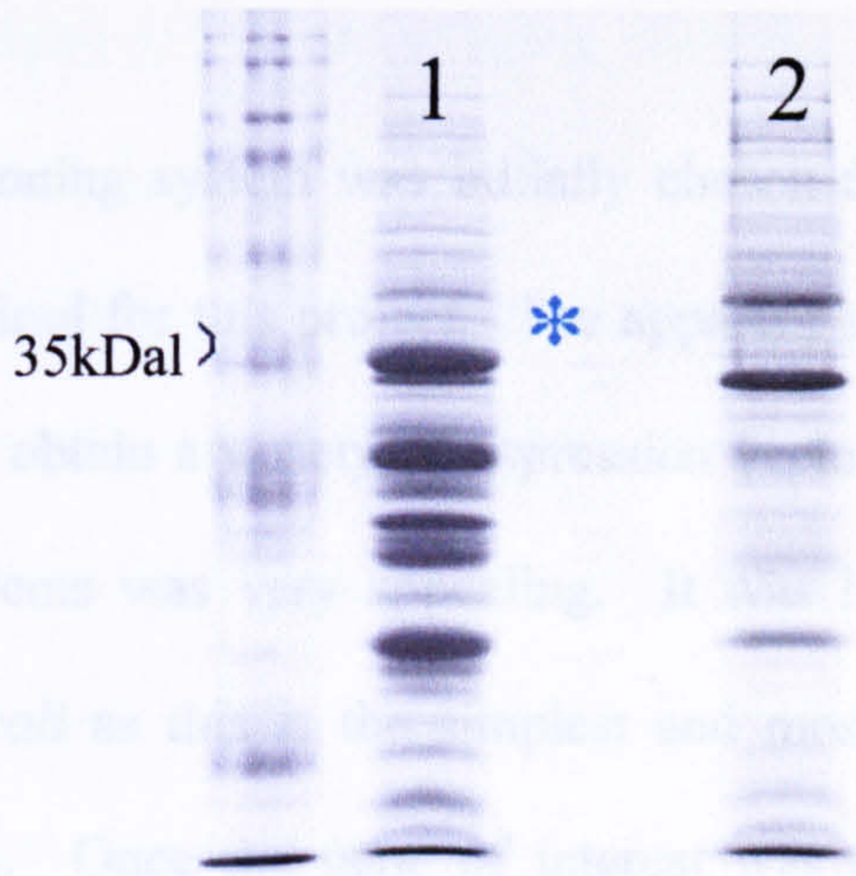


Figure 6.20. Protein identification by mass spectrometry. As no EWI-F antibodies were available, it was decided to use mass spectrometry to identify whether, or not, the protein at 35kDal was indeed 2 Ig-like domains from EWI-F. The SDS-PAGE gel was made an run taking care to keep all apparatus clean to prevent contamination of the protein sample. The 35kDal band present in the induced sample lane (lane 1) but absent in the non-induced control lane (lane 2) was carefully excised, placed into a 1.5ml tube and covered with dH_2O. It was then sent to the mass spectrometry sequencing facility in Aberdeen.

6.4 DISCUSSION

6.4.1 Expression of full length proteins

The Gateway™ cloning system was initially chosen due to the large number of cloning steps required for this project. The apparent simplicity of sub-cloning to quickly and easily obtain a variety of expression vectors for protein expression in different host systems was very appealing. It was initially decided to attempt expression in *E. coli* as this is the simplest and most cost effective method of protein production. Once the gene of interest was cloned into the expression vectors, protein expression experiments were carried out using a range of conditions in an attempt to optimise any protein production. The first experiments were carried out at 37°C and did not produce His6-tagged protein at the correct weight to correspond to EWI protein and, in addition, there were no bands at the right size in a Coomassie gel of the cell lysates in the induced samples when compared to the non-induced controls. The next logical method when attempting protein production in *E. coli* is to lower the bacterial growth temperature which prevents protein aggregation and favours production of soluble protein (Van Heeke et al. 1989); (Lin et al. 1987). It is the general consensus that lowering the temperature of bacterial growth slows down the translational process allowing the expressed protein more time to fold properly. In the protein expression experiments using the Gateway™ system, the effect of lowering the temperature did not favour soluble protein expression.

Although decreasing the temperature did not allow soluble protein expression, it was still possible that the proteins formed inclusion bodies. Inclusion bodies are

insoluble aggregates of mis-folded proteins that are frequently formed by over-expressing proteins in *E.coli* (Cabrita et al. 2004; Sorensen et al. 2004). If lowering the bacterial growth temperature does not prevent the aggregation of proteins into inclusion bodies, it is possible to extract the proteins under denaturing conditions (in guanadinium hydrochloride and urea) and then refold the proteins *in vitro* by dialysis against a re-folding buffer. However, conditions that allow re-folding to occur are often time consuming to optimise and yields for purification of protein from aggregates are low. If yields are sufficient, it is possible to avoid inclusion body formation and increase soluble, active protein formation by co-overexpression with molecular chaperones, such as heat shock proteins (hsp) (Thomas et al. 1996). For full length EWI expression in the Gateway™ system, there was neither evidence of His6-tagged proteins in aggregates nor any refolded proteins following guanidinium/urea denaturation and re-folding.

The decision to move on to EWI expression in mammalian cells was made because it seemed probable at this stage that the EWI IgSF protein was just too large and complex to be produced in a bacterial host. The original Gateway™ entry vector was used to but this time it was recombined with a mammalian expression vector containing C-terminal V5 and His6-tags. Initially, CHO and HEK cells were used because they were readily available in our laboratory. In addition, EWI production in COS-7 cells was tested as this cell line is frequently used for the production of human proteins (Helfrich et al. 2000; Schneider et al. 2005; Zhang et al. 2005). It was decided to try transient transfection to optimise conditions. Lipofectamine reagent was chosen as this has been successful in previous transfections using these cells. The manufacturers' protocol was followed for initial transfection

experiments and then it was adapted by changing the amount of DNA transfected and varying the time period for this incubation with Lipofectamine, but EWI protein expression was not detected in the whole cell lysates or in the supernatant. Following on from this, stable transfection using Lipofectamine reagent, transient DNA transfection with Calcium phosphate and both stable and transient transfection by electroporation were all tried several times, but all attempts failed to produce the full length EWI proteins. There are a number of reasons why these attempts failed: The protein could be toxic to the cell and so was being rapidly degraded; the vector may have too weak a promoter driving expression; the protein folding is too complex for the host cells to synthesise. At this stage, all feasible methods had been tested and so it was decided to continue the project by expressing protein domains rather than the full length molecules. Additionally, hidden costs in the Gateway™ system, low colony yields and a large number of "false positive colonies" that did not contain the gene of interest, all contributed to the decision to change to a different cloning strategy after full length EWI protein expression was unsuccessful in *E. coli* and mammalian cell lines.

6.4.2 Domain expression

It was concluded from the above mentioned experiments that the proteins would not express in prokaryotic nor eukaryotic hosts, simply due to their large size and complex folding patterns that involve multiple disulphide bridges. Therefore, it was decided to take the project forwards by attempting double Ig-like domain expression, rather than full length, using specially adapted *E. coli* cells and plasmid vectors, that are designed specifically for expression of mammalian proteins with

multiple disulphide bridges. There are numerous examples in the literature of Ig and Ig-like protein domains being produced in *E. coli*, as recently reviewed (McMahan et al. 2003; Anderson et al. 2004).

6.4.2.1 Disulphide bridges

Disulphide bridges are a common and necessary feature of many proteins; correct folding and stability is largely dependent on their formation. Newly synthesised proteins are initially subjected to the highly reduced environment of the cytoplasm and it is not until they are transported to oxidising compartments that the majority of disulphide bridges are formed (reviewed in (Bardwell 1994). Such compartments are the mammalian endoplasmic reticulum and the bacterial periplasm. Not only is the periplasm an oxidising environment, it is also contains a number of proteins that facilitate disulphide bridge formation. The first periplasmic protein with disulphide isomerase activity to be discovered was DsbA in 1985 (Edman et al. 1985). A further two proteins of this type, DsbC and DsbD (Missiakas et al. 1993), were later discovered. These proteins catalyse the formation of disulphide bridges and are characterised by the presence of a Cys-X-X-Cys motif involved in the thiol–disulphide bridge exchange (Holmgren 1985). Mutations in these proteins impair formation disulphide bridges (Bardwell et al. 1991; Kamitani et al. 1992).

In order to overcome the problem of disulphide bridge formation and folding of recombinant proteins in *E. coli*, two main strategies have been developed. The first is to mimic the periplasmic environment in the cytoplasm. This is done by

mutating proteins that control the redox potential of the cytoplasm and is discussed in 6.4.3. The second method is to direct the newly synthesised, recombinant protein to the periplasm by the using a specific leader sequence (see 6.4.4)

6.4.3 Protein expression of pET 32c EWI double domains in Origami cells

The reducing environment of the cytoplasm, which is unfavourable for disulphide bridge formation, is maintained by two enzymes; thioredoxin reductase and glutathione/glutaredoxin reductase. In an attempt to create a more oxidising environment in this compartment, Mossner and colleagues made a thioredoxin null mutant (Mossner et al. 1999). This knockout of the *trxB* gene product allowed the formation of disulphide bridges within proteins in the cytoplasm but with poor yields (Mossner et al. 1999). Double knockout mutations of the *trxB* and *gor* (the gene encoding glutathione reductase) increased the yields of soluble protein production in the cytoplasm (Bessette et al. 1999). There are now a number of commercially available host strains that carry these mutations on plasmids that require selection. These strains are becoming increasingly popular for expression of mammalian proteins in *E. coli* (Levy et al. 2001; Santala et al. 2004; Sorensen et al. 2005). The host strain chosen here was Origami (Novagen), a *trxB/gor* mutant. The recommended vector to be used in conjunction with these cells is the pET32 vector.

Correctly sequenced EWI-2 and EWI-F double Ig-like domains were cloned successfully into the pET32c vector with the aid of on-line database searches to predict Ig-like domain boundaries. Oligonucleotides were designed for each Ig

domain but it was decided to initially attempt double, rather than single domain expression, as antibody fragments consisting of 2 Ig domains have been commonly expressed in bacteria (Levy et al. 2001); (Kipriyanov et al. 1997; Mukherjee et al. 2004). The pET32c vector allows an optional C-terminal His6-tag in addition to the N-terminal tag and so primers were designed with and without stop codons leaving open the option of expressing either construct. EWI-F (5) was the first construct to be successfully cloned and so its expression was tested first. Once again, a large optimisation of protein expression experiment was performed with transformed, but non-induced cultures acting as the negative controls; no EWI protein expression was observed. Since the construct in the pET32c expression vector was made, expression in BL21 codon plus cells (Stratagene) was attempted but to no avail.

Although in many cases the mutations in Origami and similar cell types are beneficial, there are some drawbacks. In the oxidised environment of double *trxB/gor* mutants, proteins with larger numbers of cysteines are more likely to form random, non-native disulphide bridges. To prevent this, a periplasmic enzyme DsbC has been engineered without its leader sequence and so is able to reside in the cytoplasm of the double mutant *E. coli* where is was able to facilitate correct protein folding (Bessette et al. 1999). Certain Origami cells have the modified DsbC in the cytoplasm to aid disulphide bridge formation further. In addition, cloning vectors are available from Novagen with DsbC tags to aid the folding of the native protein in the cytoplasm. The DsbC tag can be cleaved afterwards and the advantage is that the enzyme is there, actually attached to the recombinant protein to assist the folding. Time constraints in this project did not permit these

methods to be tested but either of these methods would be worth attempting if the work were to be continued and protein expression in the cytoplasm was desired.

6.4.4 Periplasmic expression of EWI proteins

As a final attempt to express these proteins the decision was made to try periplasmic expression by means of a vector containing a signal sequence that would direct the native protein to the periplasm, the natural environment for disulphide bridge formation in the bacterial cell (reviewed in (Missiakas et al. 1997). Directing nascent proteins to the bacterial periplasm is a relatively common method used for the production of mammalian proteins in bacteria (Battistoni et al. 1995), especially antibody fragments (Kipriyanov et al. 1997; Simmons et al. 2002; Mukherjee et al. 2004). The pET26b vector that carried the *pel*B leader sequence was used here because it has a C-terminal His6-tag and the same restriction enzymes (*Not*I, *Sal*I) in its multiple cloning site so that the gene in the pET32c vector could be shuffled across easily without the need for new oligonucleotides. These cloning steps were carried out successfully. After translocation across the cytoplasmic membrane, the pelB leader sequence is removed due to the presence of a signal peptidase cleavage site immediately downstream from the leader sequence. Expression of the EWI-F (5) was tested first and the results indicated that the protein was present in the periplasmic fraction in the majority of the conditions tested and not in the non-induced controls. However, it appeared that most of the EWI-F (5) produced had lost their His6-tags because they were not picked out in Western blots using anti-His6 antibodies. The samples that were initially recognised in an immunoblot by an

anti-His6 antibody (Figure 6.18) also appeared to lose their Hig6 tag over time suggesting that the tag was extremely labile.

In the absence of an EWI-F antibody, or to be more precise, an antibody directed against the two membrane proximal Ig-like domains, it was decided to identify the protein by mass spectrometry.

The mass spectrometry results (6.3.4) were somewhat confusing given that the probable match to a human GPR35 was highly unlikely to be due to sample contamination and the protein domain that was cloned, EWI-F (6), had no similarity to this probable match. Further more, no colleagues in the laboratory were working with GPR35 so cross contamination was also ruled out. When the results were queried with the Proteome Facilities Co-ordinator, we were assured that no mix-up of samples could possibly have taken place. The protein production was scaled up in an attempt to get more EWI-F (5) protein to send to be analysed by mass spectrometry elsewhere. However, the results were not repeated in the first two attempts and eventually time limitations of the project prevented a second protein domain sample being sent for mass spectrometry analysis. Furthermore, problems were encountered in cloning the remaining EWI double Ig-like domains into the pET26b vector, which meant that this work could not be completed. However, given the encouraging results from the initial protein screening experiment, (Figure 6.19), double Ig-like domain expression in the periplasm would be a good starting place to continue this project in the future.

CHAPTER 7

FUNCTIONAL ANALYSIS OF TETRASPANIN EC2 DOMAINS: EC2 DOMAINS ARE POTENT INHIBITORS OF THE INFECTION OF MACROPHAGES BY HIV-1

7.1 INTRODUCTION

There are an estimated 39.4 million people worldwide living with HIV/AIDS and since 1981 the virus has accounted for more than 20 million lives. In 2004, 4.9 million people acquired the virus and so the epidemic lives on. Since 2001 there has been an increase in the global AIDS funding from US$ 2.1 billion to an estimated US$ 6.1 billion in 2004 (www.unaids.org).

Entry of the retrovirus HIV (*genus Lentivirus*) into cells occurs in three distinct steps: the attachment of virions to cell surface CD4 by envelope gp120 protein, the subsequent interplay of the conformationally altered gp120 with cellular CCR5 and CXR4 co-receptors, and, finally, envelope gp41 protein-mediated fusion to target cells (Moore et al. 2003). The HIV genome consists of two copies of RNA which are reverse transcribed to DNA to form the pre-integration complex (PIC) containing the viral matrix (MA), integrase and Vpr proteins as well as several host proteins. The PIC is imported into the nucleus and incorporated into the host chromosomes, where it can become latent for 3-10 years (Greene 2004). The subtypes of HIV, HIV-1 and HIV-2, have the same modes of transmission and are

associated with similar opportunistic infections, but immunodeficiency seems to develop more slowly in HIV-2 (Marlink et al. 1994). There are different strains of HIV-1, the most common being the R5 and X4 strains that have evolved to utilise the chemokine receptors CCR5 and CXCR4, respectively.

In 2003 it was reported that anti-CD63 mAbs were able to inhibit HIV infection in macrophages (von Lindern et al. 2003), and so it was decided to investigate the inhibitory properties of our tetraspanin GST-EC2 fusion proteins in HIV-1 infection of MDM. This work was done in collaboration with Cecilia Cheng-Mayer's laboratory (Aaron Diamond Aids Research Centre, New York). My part in the work was to provide all of the tetraspanin EC2 constructs for the study. CD63 EC2 without the GST tag was also made to use in this study in addition to the structural studies (Chapter 5), because we wanted to confirm that the effects we were observing were not simply due to the presence of GST. Parthasarathy Varadjaran (Department of Molecular Biology and Biotechnology, University of Sheffield) performed the EC2 FITC labelling experiments.

7.2 MATERIALS AND METHODS

7.2.1 Preparation of PBMC and monocyte-derived macrophage cultures

PBMCs were prepared by Ficoll gradient centrifugation, stimulated with phytohemagglutinin (PHA) (3μg/ml) in RPMI 1640 medium containing 10 % FCS and 20U IL-2 (kindly provided by Chiron Corp., Emeryville, CA). Monocytes were enriched by centrifugation of PBMCs through a 46 % Percoll cushion. Cells were then resuspended in RPMI 1640 medium supplemented with 10 % FCS and 5 % human AB serum. Monocytes were allowed to adhere overnight, harvested and replated in 96well at 7 X 10^4 cells/well for infection. For microscopic studies, cells were replated in 35mm glass bottom No 1.5 poly-d-lysine dishes (Mat-Tek Corp., Ashland, MA) at 7.5 x 10^5 cells/plate. Monocytes were allowed to differentiate for 5-7 days before use.

7.2.2 Generation of Viruses

Luciferase reporter viruses pseudotyped with different envelope glycoproteins were generated by transcomplementation, as described previously (Connor et al. 1996). The reporter viruses were derived from the HIV-1 pNL4-3 proviral DNA, in which the *env* gene was deleted and a firefly luciferase cassette was inserted in place of the *nef* gene. The *env* constructs used were: pEnv162P3 expressing an CCR5-using HIV-1 envelope (Hsu et al. 2003), pEnvA2 expressing an CXCR4-using HIV-1 envelope (Chakrabarti et al. 2002), and a plasmid expressing VSV-G (pVSV-G, kindly provided by J. McKeating, University of Birmingham, UK)

(Pohlmann et al. 2003). Because of the lack of a proviral *env* gene, pseudotyped viruses were capable of only a single round of replication. The viruses were generated by lipofection of 1.5ug each of pNL-LucE-R+ plasmid and of a pEnv vector in 293T cells plated at 7 x 10^5 cells per well in six-well plates. The lipofection was performed with the DMRIE-C reagent according to the manufacturers' recommendations (Gibco-BRL, Gaithersburg, Md.). Cell culture supernatants were harvested 72 hours post-transfection, centrifuged at 800 g, filtered through 0.45μm-pore-size filters, and stored at -70°C until use. The viral content was quantified by a p24 Gag ELISA, (Abbott Laboratories, Chicago, Ill.).

To generate Vpr-eGFP viruses, 293T cells were co-transfected with a full length NL4-3 proviral genome in which the envelope gene has been replaced with that of R5-SHIV$_{SF162P3}$ (Hsu et al. 2003) and an expression plasmid for Vpr-EGFP (kindly provided by N. Landau, Salk Institute, La Jolla, CA.). Culture supernatants were harvested 48 h later, centrifuged at 800 x g, filtered through 0.45 μm-pore-size filters and concentrated by centrifugation through a 20 % sucrose cushion. The pelleted virions were resuspended in Hanks balanced salt solution, aliquoted and stored at -70^{0}C until use. The viral content was quantified by a p24 Gag ELISA.

7.2.3 Virus infection inhibition assays

Inhibition assays were performed as follows: macrophages (7 x 10^4 cells per well) or PBMC (10^6 cells per well) in 96-well plates were pretreated with 50 μl of serial dilutions of GST or GST-tetraspanin fusions for 30-60 minutes at 37°C. Control cells received Hanks BSS alone. An equal volume containing 5ng p24 Gag

equivalent of each of the pseudotype viruses was then added and incubated for 3 hours at 37°C. At the end of the incubation period, 100µl of macrophage or PBMC culture medium was added and the culture maintained for 72 hours before being tested for luciferase activity. Cells were lysed and incubated with the luciferase assay reagents according to the manufacturers' instructions (Promega, Madison, Wis.). The luciferase activity was measured in a Dynex MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantily, Va.). All infections were performed in duplicate.

7.2.4 Labelling of cells with virus expressing Vpr-eGFP

Macrophages in Mat-Tek plates were treated with 50µl of Hanks, GST or GST-CD63 at 4µg per ml for 60 minutes at 37°C. 50µl of Vpr-EGFP virus (containing 200ng p24 Gag equivalent) was then added and incubated for 2 hours at 37°C. At the end of the incubation period, infected macrophages were washed extensively and fixed with 3 % formaldehyde in PBS ON at 4°C. Next day, cells were stained with ConA-Rhodamine (1ng/ml) at RT for 15 minutes and washed 5x with PBS. DAPI was then added and images were obtained with Delta Vision. Viral uptake was measured by counting fluorescent particles inside cells.

7.3 RESULTS

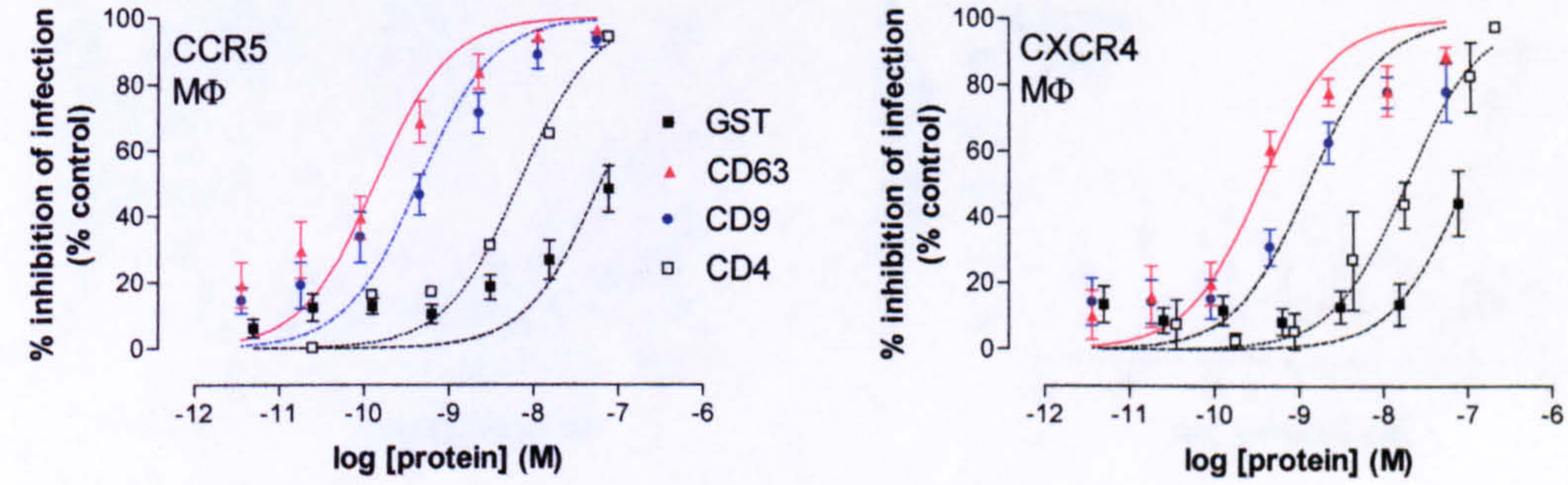
Following the recent publication by the Von Lindern laboratory (von Lindern et al. 2003), we were intrigued to see if our EC2 constructs had any inhibitory potential. It is reasonably well documented that tetraspanin EC2 proteins can inhibit tetraspanin-dependent functions such as fertilisation (Higginbottom et al. 2003), leukocyte adhesion (Barreiro et al. 2005) and monocyte fusion (Takeda et al. 2003), which were initially identified by the use of anti-tetraspanin antibodies. We therefore tested a range of tetraspanin EC2-GST proteins for their inhibitory activity of R5 and X4 HIV-1 strains. Soluble CD4 was used as a positive control and GST alone as a negative. Monocyte derived macrophages (MDM) or peripheral blood mononuclear cells (PBMC) were pre-incubated with EC2 proteins and washed before addition of virus. All of the tetraspanin EC2 proteins inhibited both R5 and X4 and R5/X4 chimeric virus infection in MDM (Fig 7.1). Indeed, in MDM, all the tetraspanin EC2 proteins were better inhibitors than CD4. There appeared to be a hierarchy of potency with CD63 EC2 being the most potent inhibitor of R5 and X4 virus, followed closely by CD9 EC2. CD81 EC2 and CD151 EC2 were 5 – 10 fold less potent, and mCD9 EC2 10 – 50 fold less potent, than CD63 EC2. GST alone did show some inhibitory activity but in each case this was at notably higher concentrations than the tetraspanin EC2.

Von Lindern and colleagues described inhibition using anti-CD63 antibodies in MDM but not in PBMC. Using the EC2 proteins we were able to detect inhibition of R5 virus in PBMC, albeit to a lesser extent that in the MDM. There was no specific inhibition of X4 virus in PBMC by any of the tetraspanins tested, even at

high concentrations of 10μM. Soluble CD4 could completely inhibit infection, confirming that the assay was still viable (Figure 7.2). Mouse CD9 is seen to have an inhibitory effect in MDM but interestingly the effect is not significant in PBMC.

To examine the role of the GST fusion partner protein, GST was cleaved from CD63-GST EC2. The pGEX-2KG contains an extended linker region between the GST tag and the protein of interest. The linker region contains a thrombin protease site. Re-adherence of purified CD63-GST to glutathione Sepharose beads, followed by on-bead cleavage with thrombin enables elution of purified CD63 EC2 whilst the GST remains adhered to the beads. GST was undetectable on silver stained SDS-PAGE gels (see chapter 5).

A.



B.

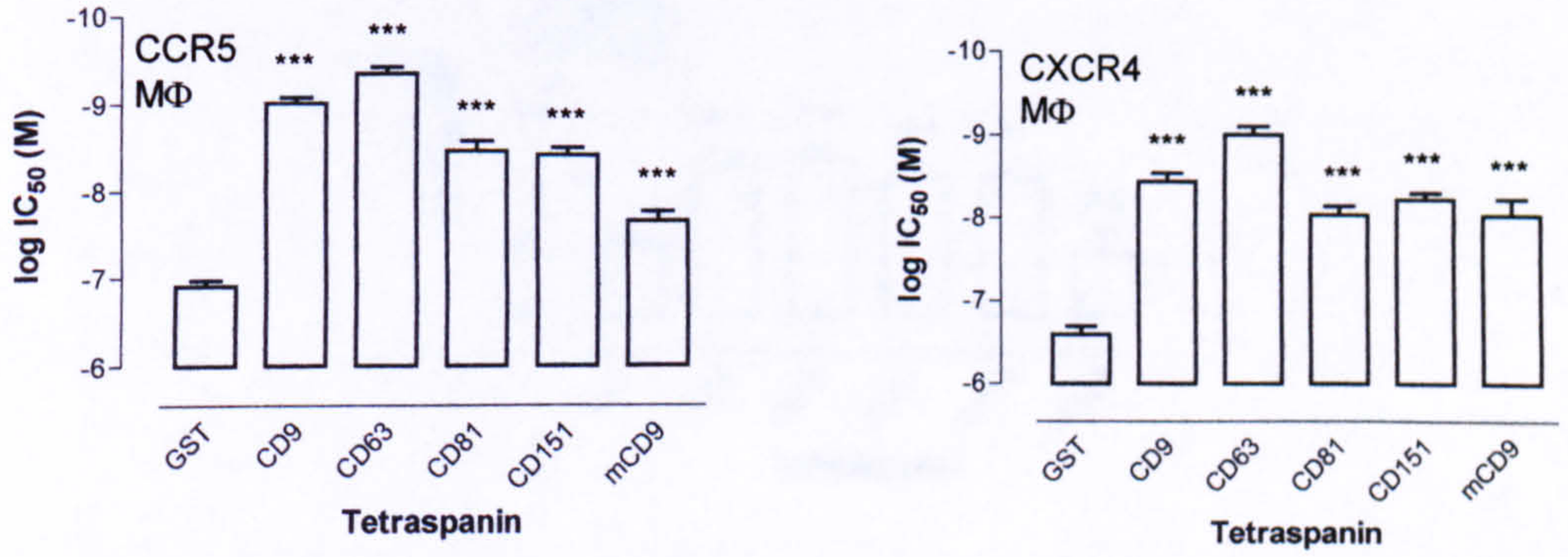
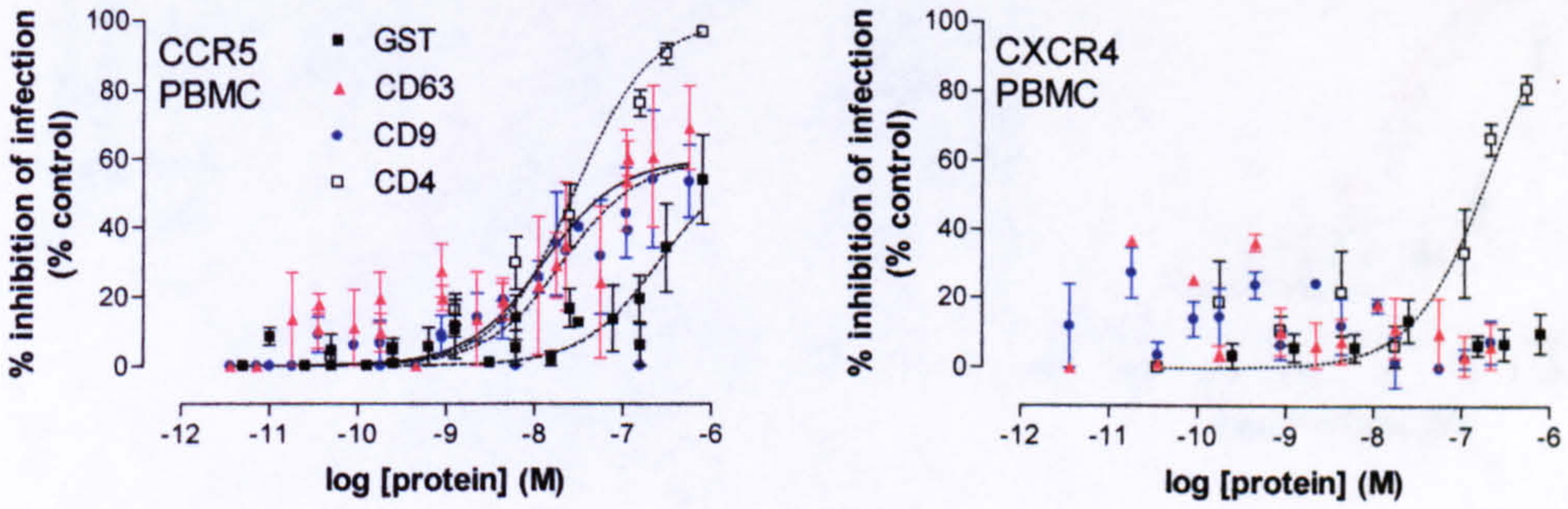


Figure 7.1 the effects of recombinant human tetraspanin large extracellular domain-GST proteins on infection of MDM by CCR5 and CXCR4-tropic HIV-1. MDM were isolated as described in Materials and Methods and treated with different concentrations of recombinant tetraspanin large extracellular domain-GST fusion proteins, GST alone or with Ig-CD4 for 30 – 60 minutes prior to the addition of HIV-1 virions expressing CCR5 or CXCR4 specific env protein. Infection was measured after 3 days as chemiluminescence from luciferase expressed under the control of the HIV-1 LTR promoter. Infection of MDM (A and B) pre-treated with human CD9, CD63, CD81, CD151 or mouse CD9 large extracellular domain-GST fusion proteins, GST alone or Ig-CD4. Results are shown as percentage inhibition of infection relative to untreated control cells and are the means of at least 3 separate experiments (A) or log IC_{50} values in (B). Significance of difference from GST control was assessed by two tailed t-test; *** $p < 0.001$.

A.



B.

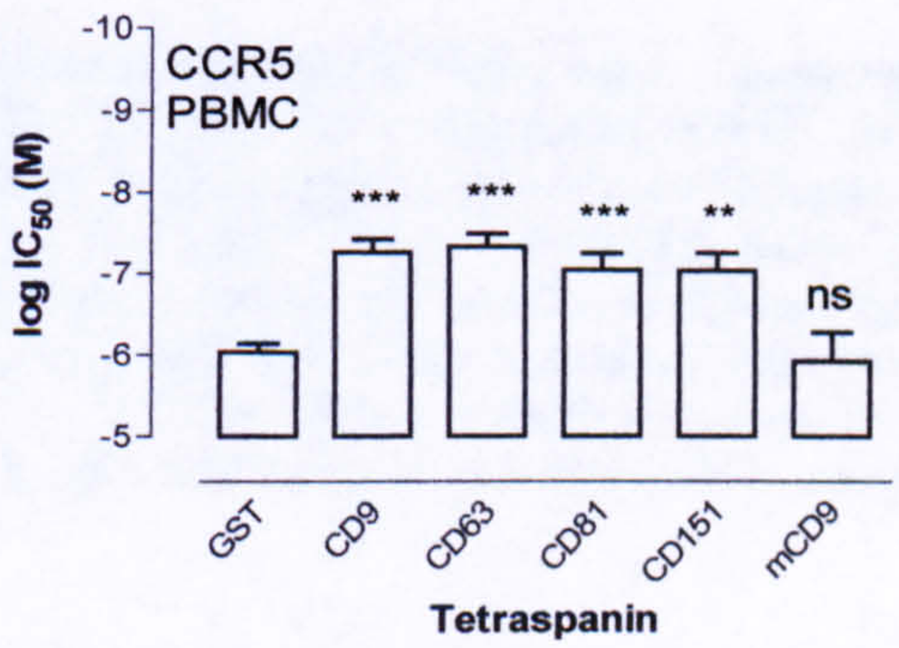


Figure 7.2 The effects of recombinant human tetraspanin large extracellular domain-GST proteins on infection of PBMC by CCR5 and CXCR4-tropic HIV-1. PBMC were isolated as described in Materials and Methods and treated with different concentrations of recombinant tetraspanin large extracellular domain-GST fusion proteins, GST alone or with Ig-CD4 for 30 – 60 minutes prior to the addition of HIV-1 virions expressing CCR5 or CXCR4 specific env protein. Infection was measured after 3 days as chemiluminescence from luciferase expressed under the control of the HIV-1 LTR promoter. Infection of PBMC pre-treated with human CD9, CD63, CD81, CD151 or mouse CD9 large extracellular domain-GST fusion proteins, GST alone or Ig-CD4. Results are shown as percentage inhibition of infection relative to untreated control cells and are the means of at least 3 separate experiments (A) or log IC_{50} values in (B). Significance of difference from GST control was assessed by two tailed t-test; *** $p < 0.001$.

Inhibition of R5 and X4 infection was not prevented by the removal of GST in MDM. In R5 virus assays the cleaved CD63 EC2 was 10 fold more active than GST alone however, is was around 6 fold less active than GST-CD63 EC2 (Figure 7.3). X4 infection was inhibited to the same extend with cleaved CD63 EC2 and GST-CD63 EC2 (Figure 7.3).

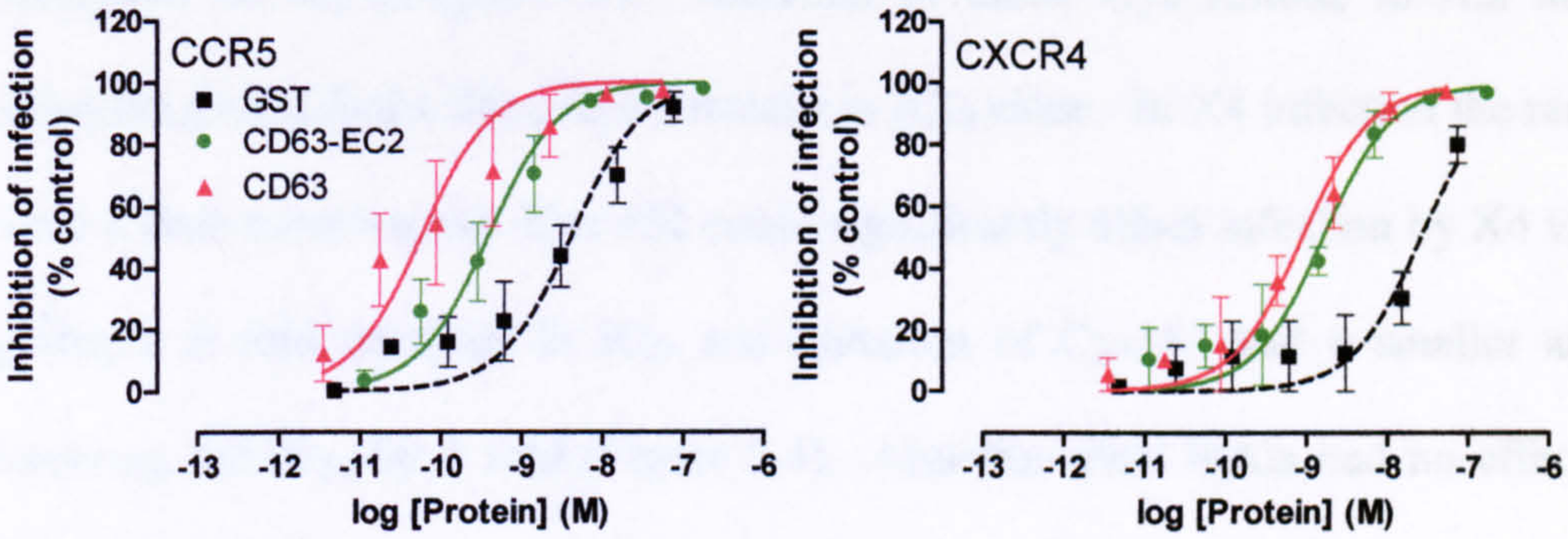


Figure 7.3 The effects of removal of GST from recombinant CD63 EC2 domain on the inhibition of infection by MDM by CCR5-tropic HIV-1. MDM were isolated as described in Materials and Methods and treated with different concentrations of recombinant human CD63-EC2 GST fusion protein (CD63-EC2), CD63-EC2 with GST removed (CD63) by thrombin cleavage or GST alone for 30-60 minutes prior to the addition of virions expressing CCR5-specific HIV-1 env protein. Infection was measured after 3 days as chemiluminescence from luciferase expressed under the control of the HIV-1 LTR promoter. Results are shown as percentage inhibition of infection relative to untreated control cells are the means of at least 3 separate experiments ± SEM.

Critical amino acid residues in the EC2 domains have been highlighted in the past as being essential for a particular function. For example, in sperm–egg fusion assays the residue Phe176 in CD9 EC2 was identified as being critical for CD9-EC2 inhibition of mouse fertilisation, since mutation of this residue to alanine completely removed the inhibitory action of the CD9 EC2 (Zhu, G. Z. et al. 2002; Higginbottom et al. 2003). Mutations of cysteine residues that are predicted to form disulphide bridges in this EC2 domain also bring about an abolishment of inhibition, presumably because the disulphides are not formed and the EC2 proteins loses any structure it may have had. In an attempt to understand the mode of action of the EC2 proteins in the HIV inhibition assays, we used the CD9 mutants to determine whether a similar mechanism was responsible here. The cysteine mutants that had the greatest effect in the mouse fertilisation assays, Cys152Ala and Cys153Ala, and the mutation Phe176Ala were tested. Although

significant, there was only a small decrease in inhibition by either Cys mutant compared to WT (Figure 7.4). Mutation of either Cys residue to Ala in R5 infection assays had a three fold decrease in IC_{50} value. In X4 infection the results were a little more varied. Cys 152 could significantly affect infection by X4 virus, giving a 9 fold decrease in IC_{50} and mutation of Cys153 had a smaller affect lowering the IC_{50} by 4 fold (Figure 7.4). Mutation Phe176Ala had no effect on inhibition for either virus (Figure 7.4).

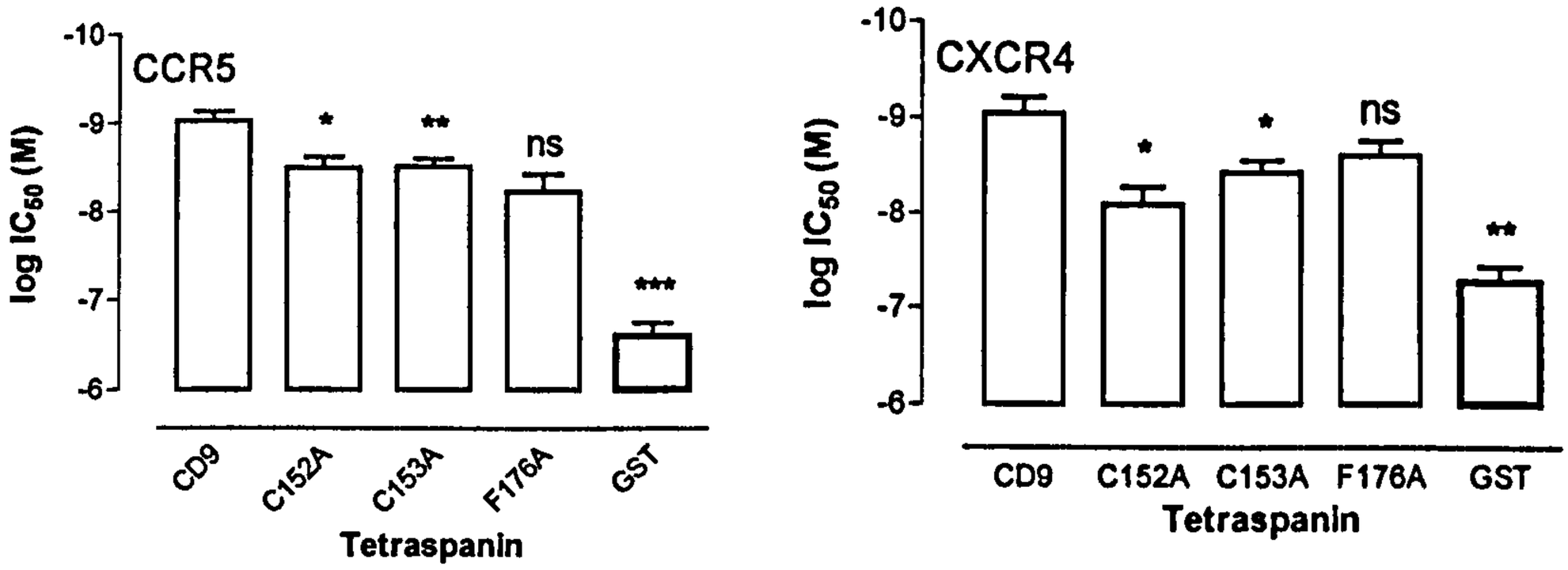
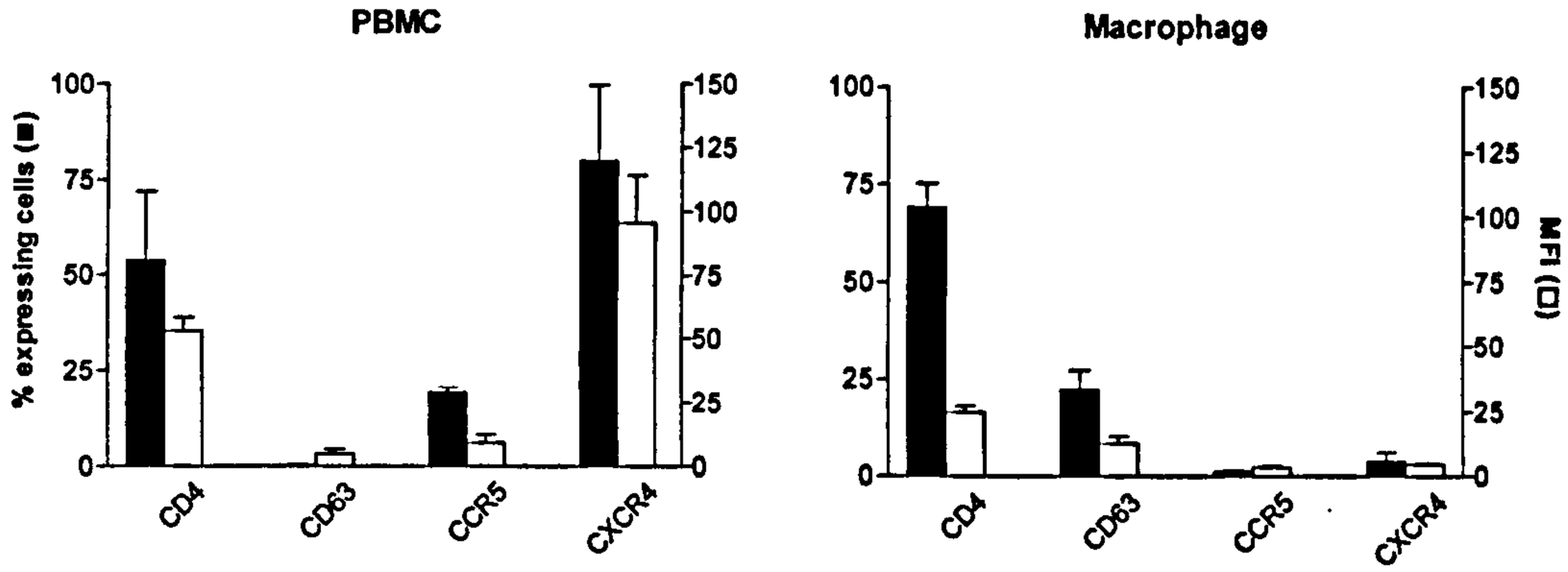


Figure 7.4 The effects of mutation of residues in the CD9 EC2 domains on the potency of inhibition of HIV-1 infection in MDM. MDM isolated as described in Materials and Methods were treated with different concentrations of wild-type or mutant recombinant human CD9 EC2 – GST or GST alone for 30-60 minutes prior to the addition of virions expressing specific HIV-1 env protein. Infection was measured after 3 days as chemiluminescence from luciferase expressed under the control of the nef promoter. Results are shown as IC_{50} values and significance of difference from GST control was assessed by two-tailed t-test; ***$p>0.001$; **$p>0.01$; *$p>0.05$

It is possible that treatment with GST-CD63 EC2 imposes inhibitory effects on virus infection by altering the expression levels of receptor/co-receptor at the plasma membrane. It is also a possibility that the difference in inhibitory effects in MDM and PBMC is due to differential expression levels of CD63, CD4 and HIV-1 coreceptors on these cell types. To address these possibilities, cells were incubated with and without 0.11μM GST or GST-CD63 EC2 (Figure 7.5.B) and the

expression levels of CD4 and HIV-1 coreceptors were determined by FACS analysis. Our results were in agreement with previous reports (Wu et al. 1997; Naif, H. M. et al. 1998), MDM express lower levels of CD4 and HIV-1 coreceptors than PBMC and the percentage of cells expressing HIV-1 coreceptors is lower in MDM (Figure 7.5.A). In contrast to these findings, the level of CD63 expression as well as the percentage of cells expressing CD63 is higher in MDM than PBMC (Figure 7.5.A). Treatment with CD63 did not decrease expression levels of any of the proteins (Figure 7.5.B). Levels of the cell surface marker CD14 and co-receptor CXCR4 were unaltered whereas CD63, CCR5 and CD4 were actually slightly increased (Figure 7.5.B). There was no detection of GST suggesting that the elevated level of CD63 in treated compared with untreated cells is due to the redistribution of endogenous CD63 and not GST-CD63 (Figure 7.5.B).

A.



B.

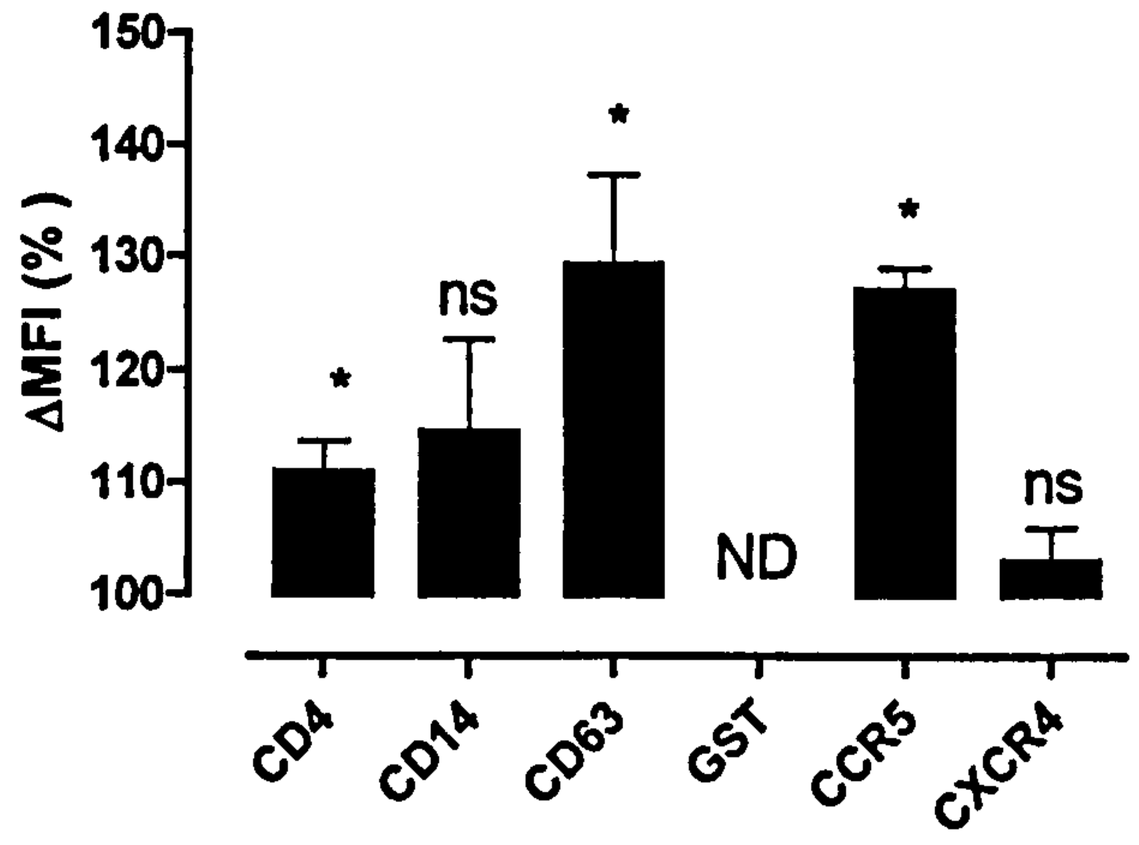


Figure 7.5 The effects of treatment of MDM with recombinant CD63 EC2 domain on membrane protein expression. Adherent MDM were incubated with 0.11μM CD63 EC2 for 1 hour and the cells were extensively washed and harvested by scraping. After incubation with appropriate primary antibodies and secondary antibody labelled with FITC, fluorescence was quantified by flow cytometry. (A) Percent and mean fluorescence intensity (MFI) of CD4, CD63, CCR5 and CXCR4 on untreated MDM and PBMC. (B) Difference in expression of CD4, CD14, CD63, GST, CCR5 and CXCR4 in cells treated with GST-CD63 EC2 compared to treatment with GST alone. Results are shown as mean SEM of three separate experiments performed in duplicate. Significance was assessed by one sample t-test; * p>0.5. N.D, Not detected.

It is generally accepted that the predominant route of entry of HIV-1 in T cells is by receptor/co-receptor dependent fusion at the plasma membrane but it is now thought that the preferred route of entry in macrophages is receptor/co-receptor independent endocytosis. To determine the mechanism of entry of the virus in our assay we decided to use a VSV pseudotype virus. This is the HIV-1 core virus

expressing VSV coat proteins. VSV is known to enter cells by endocytosis in a receptor independent mechanism. Surprisingly, all of the tetraspanin EC2 proteins inhibited the pseudotype infection in a similar manner to R5 inhibition (Figure 7.6). Like R5 infection, VSV was also partly inhibited in PBMCs. In contrast to R5 inhibition, mCD9 EC2 was ineffective in VSV infection (Figure 7.6).

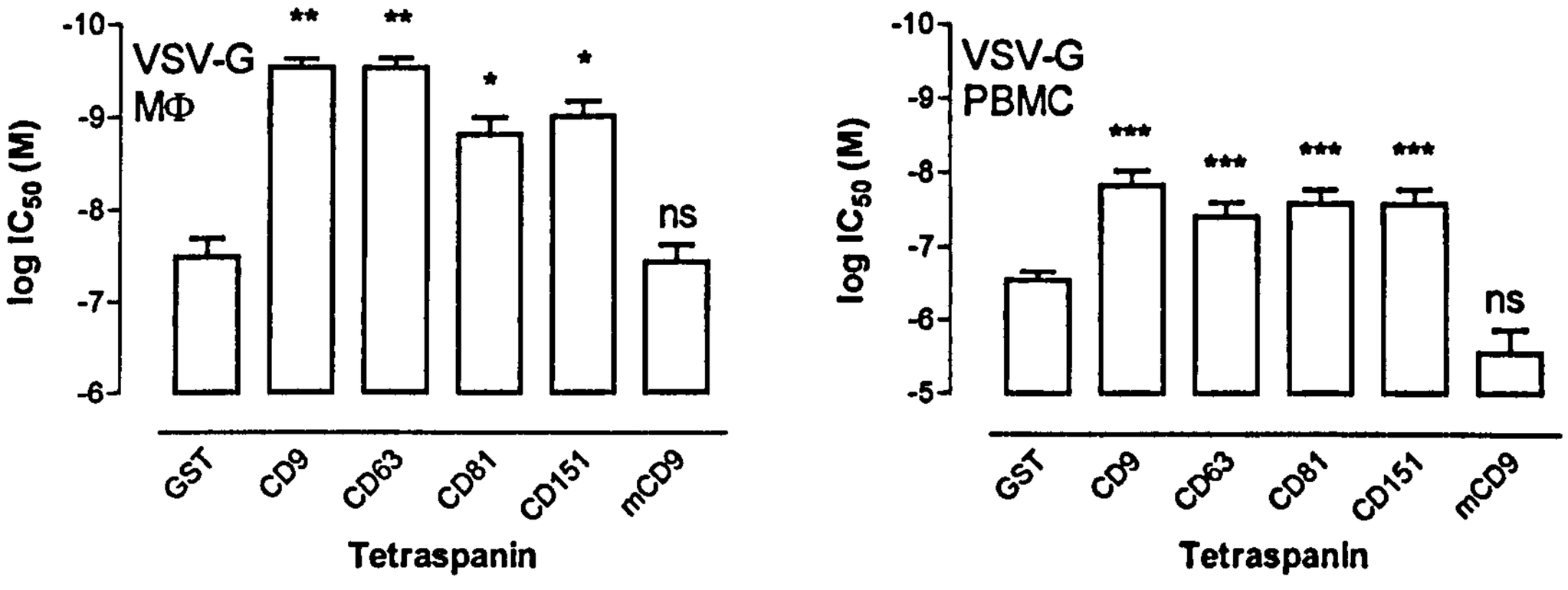
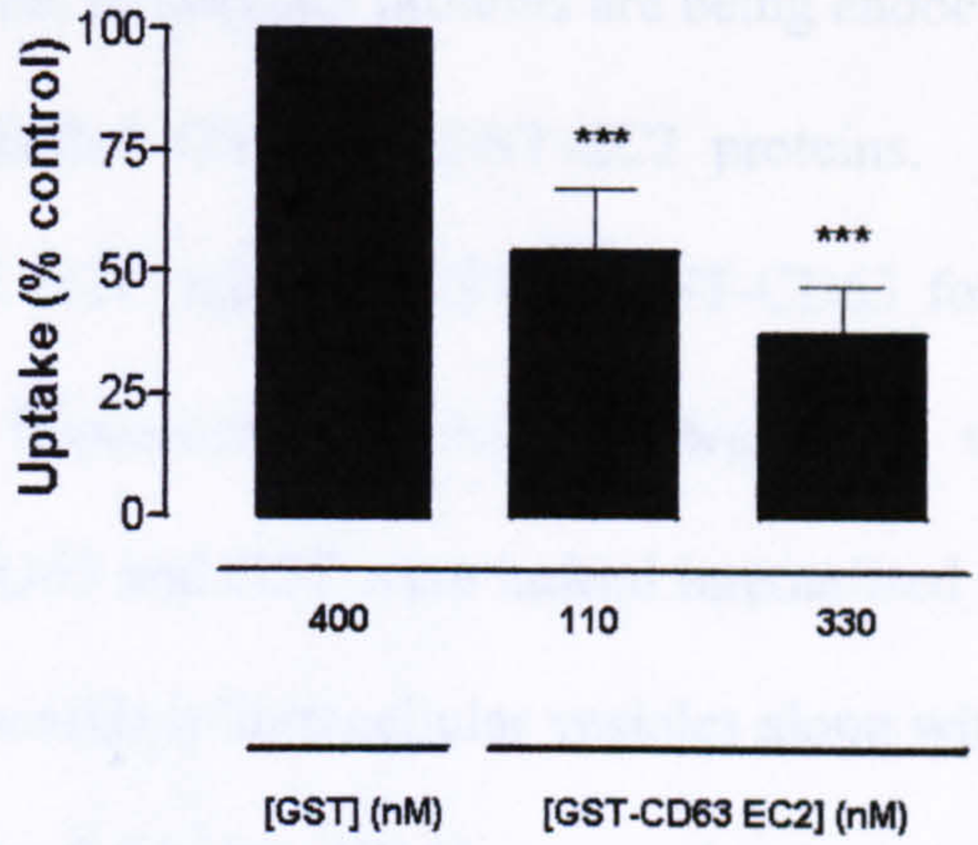


Figure 7.6 The effects of recombinant human tetraspanin large extracellular domain-GST proteins on infection of MDM by VSV-pseudotype HIV-1. MDM were isolated as described in Materials and Methods and treated with different concentrations of recombinant tetraspanin large extracellular domain-GST fusion proteins, GST alone or with Ig-CD4 for 30 – 60 minutes prior to the addition of HIV-1 virions expressing VSV-G glycoprotein. Infection was measured after 3 days as chemiluminescence from luciferase expressed under the control of the HIV-1 LTR promoter. Infection of MDM pre-treated with human CD9, CD63, CD81, CD151 or mouse CD9 large extracellular domain-GST fusion proteins or GST alone. Results are shown as log IC_{50} values in and significance of difference from GST control was assessed by two tailed t-test; *** $p < 0.001$; **$p > 0.01$; *$p > 0.05$.

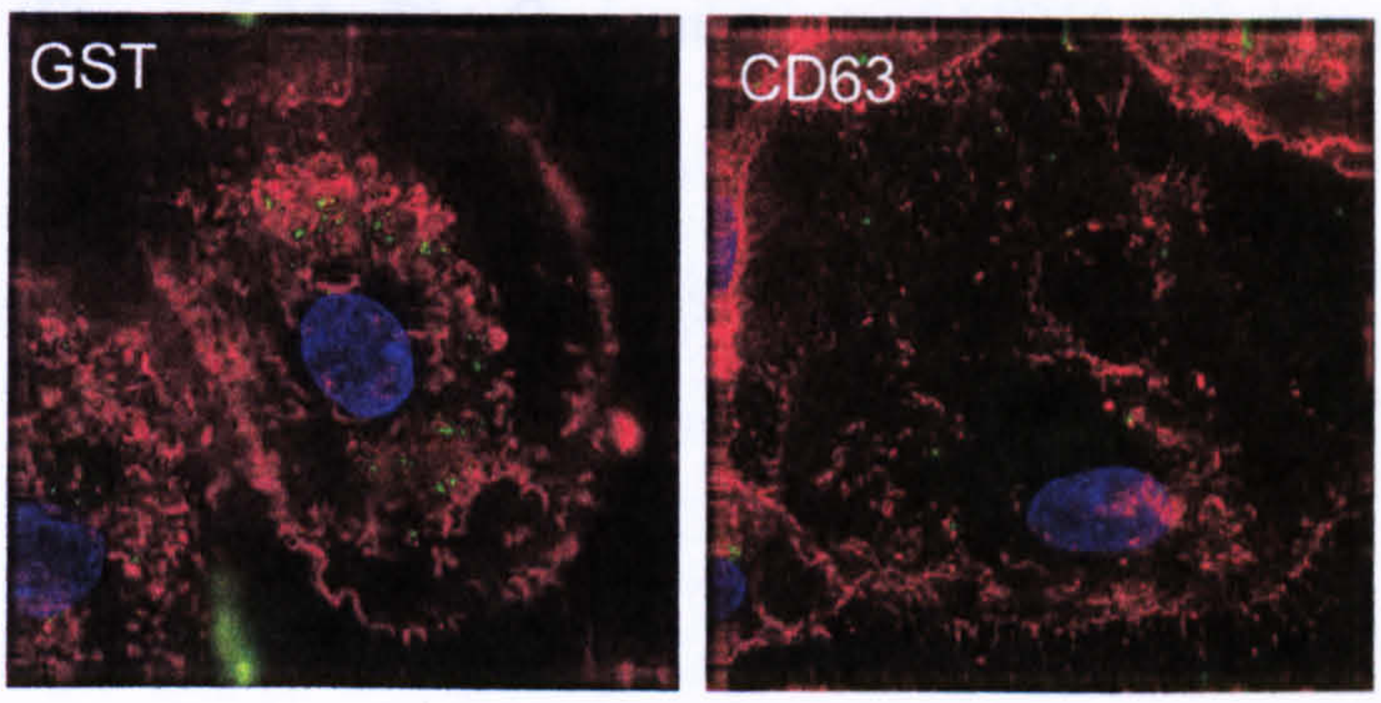
To determine whether the tetraspanin EC2-dependent inhibition is occurring at the stage of binding/uptake of virus or at a later stage of the virus life cycle, we used fluorescently labelled virus. MDM were pre-incubated with 0.4μM GST or different concentrations of GST-EC2 (0.11μM and 0.33μM) for 30 minutes prior to the addition of R5 virus tagged to a vpr-eGFP fusion protein. Infection was left to proceed for 10, 30, 60 or 120 minutes before being terminated by washing and

fixation. CD63 does significantly inhibit viral uptake at both concentrations tested (Figure 7.7.A and B). To asses the time of inhibition, MDM were incubated with CD63-EC2 or GST control upon inoculation or immediately after. The inhibitory effect by CD63-EC2 was most potent when GST-CD63 EC2 is added to MDM at the same time as the virus, and is significantly reduced at all other time points suggesting that the EC2s are inhibiting at the viral entry stage (Figure 7.7.C). The inhibition seen by GST alone is independent of the time of addition which suggests that the mechanism of this less potent inhibition is unrelated to that seen by the tetraspanin EC2 domains.

A.



B.



C.

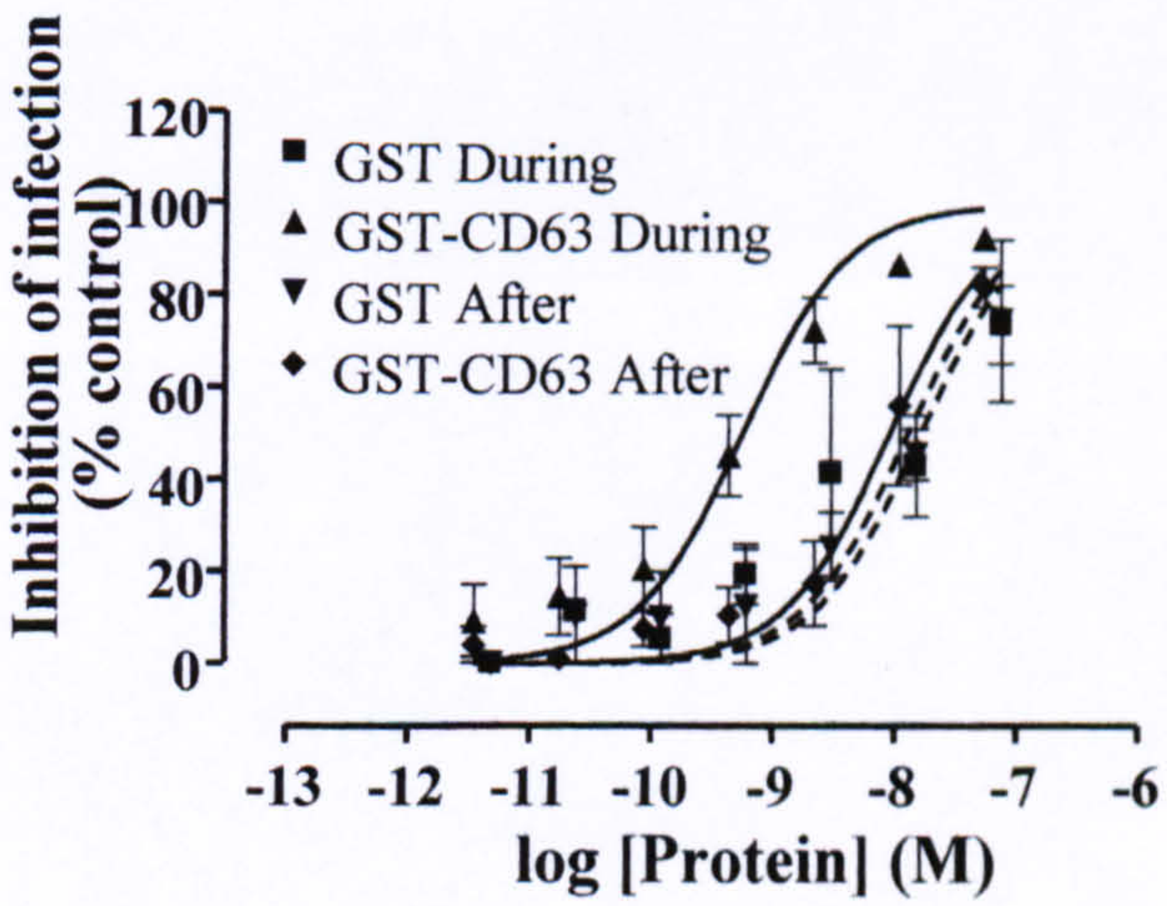
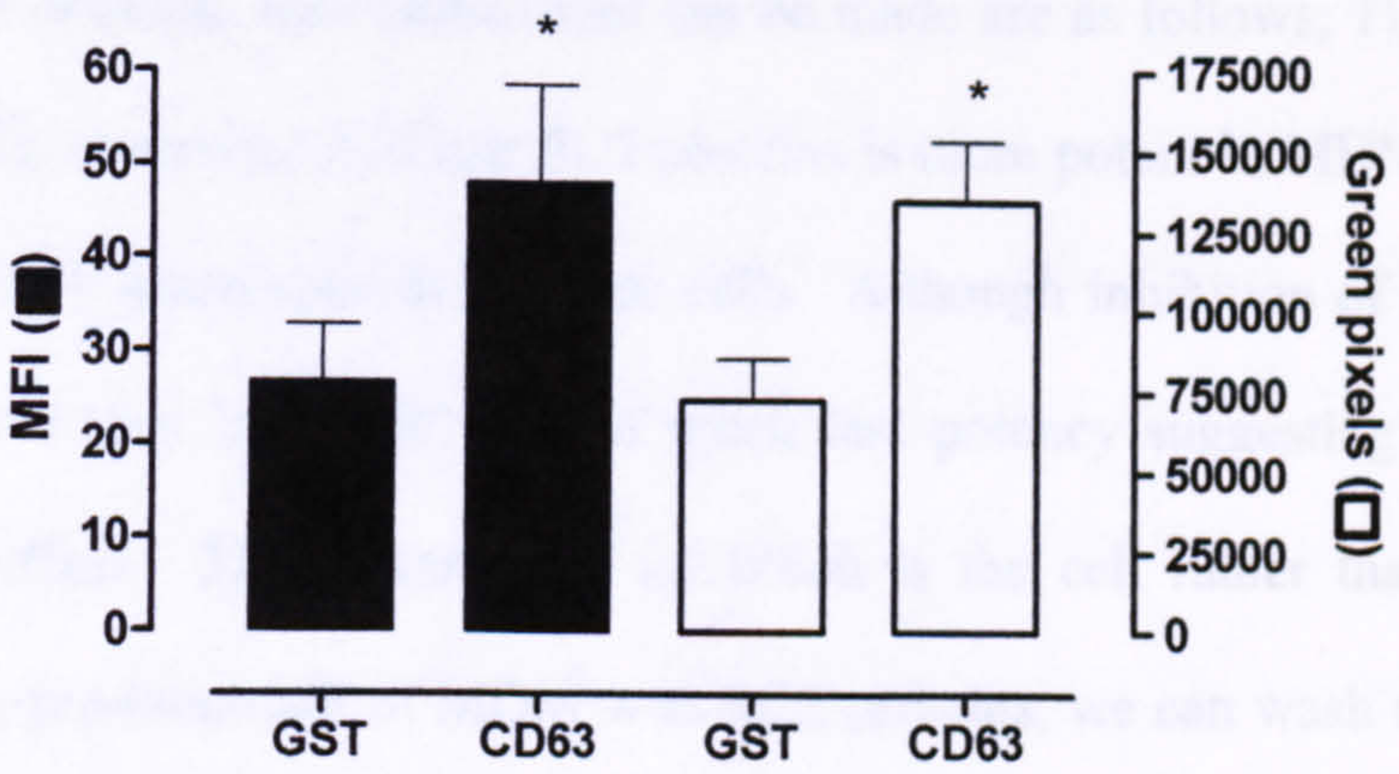


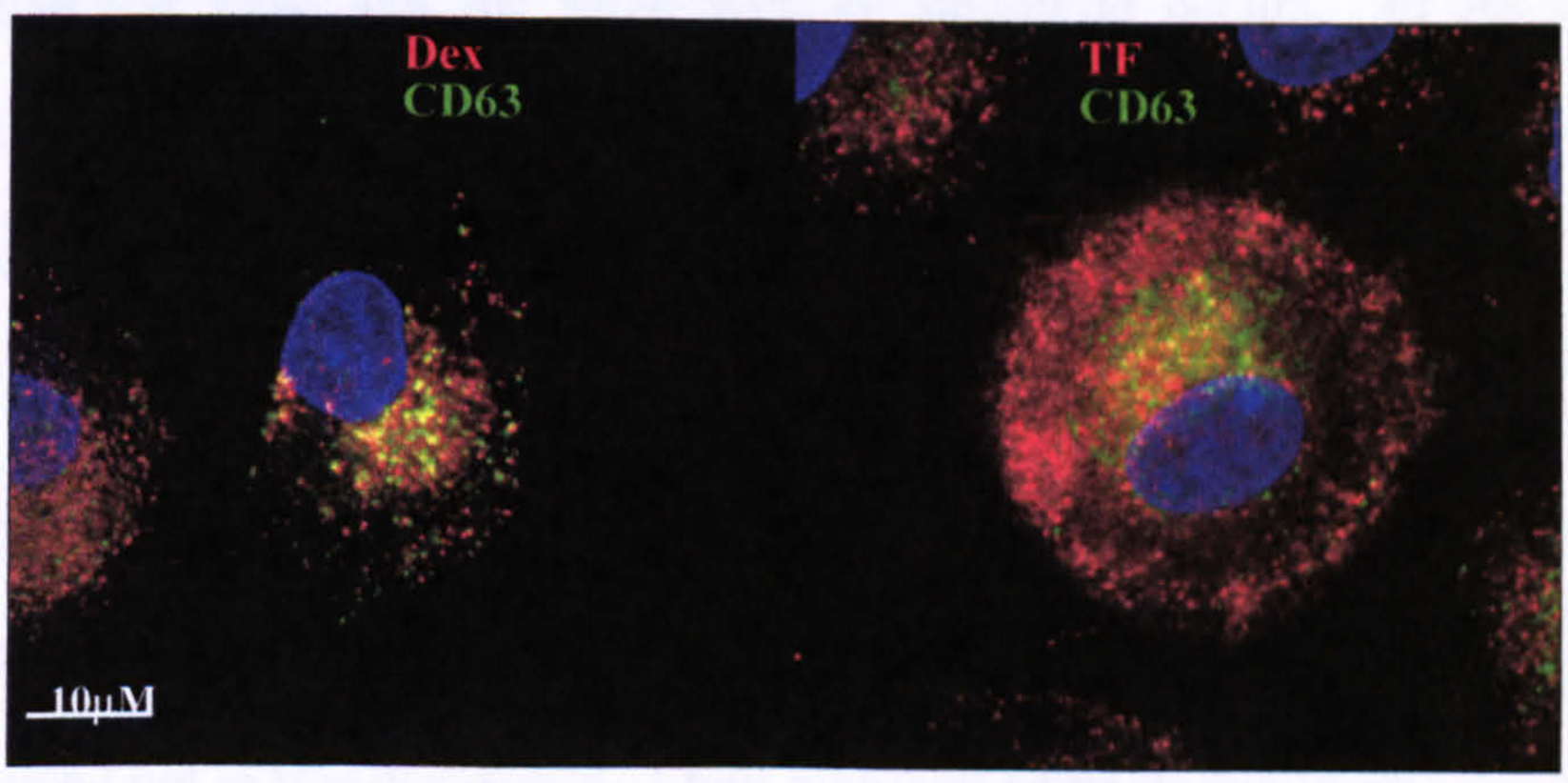
Figure 7.7 The effects of treatment with high concentrations of recombinant CD63 EC2 domains on the uptake of CCR5-tropic HIV-1 by MDM. MDM were pre-incubated with 0.4μM GST or 0.11μM or 0.33μM GST-CD63 EC2 for 30min prior to addition of Vpr-eGFP R5 virus, and infection terminated after 10, 30 or 60min after virus addition. Virus uptake was quantified by flow cytometry of cells in suspension (A) or by counting of green particles in micrographs (examples shown in B) of adherent cells. As CD63-EC2 had an inhibitory effect on the numbers of green particles associated with cells at all of these time points, data were pooled. Results are the means SEM of at least three experiments. The differences between each dose of CD63 EC2 and the GST only control (=100) were significant (*** $p < 0.01$) by one sample t-test. C) MDM were incubated with R5 virus for 1 hour at 37°C. Serial dilutions of GST alone or GST-CD63 EC2 were present either during virus inoculation or added immediately after the virus was removed by washing.

To test whether the tetraspanin proteins are being endocytosed by the cells we used fluorescently labelled GST and GST-EC2 proteins. MDM were pre-incubated with 0.11μM of FITC labelled GST or GST-CD63 for one hour at 37°C. Flow cytometry and fluorescence microscopy was used to measure cell-associated fluorescence. CD63 and GST were indeed internalised and shown by fluorescence microscopy to localise in intracellular vesicles along with endocytosed dextran and transferrin (Figure 7.8A, B). CD63 appeared to be internalised at a faster rate than GST alone (Figure 7.8A, B). Co-incubation of MDM with Vpr-eGFP virus and CD63-EC2 labelled with AlexaFluor647, rare instances of virus and tetraspanin co-localisation could be detected after 20 minutes (Fig 7.8.C).

A.



B.



C.

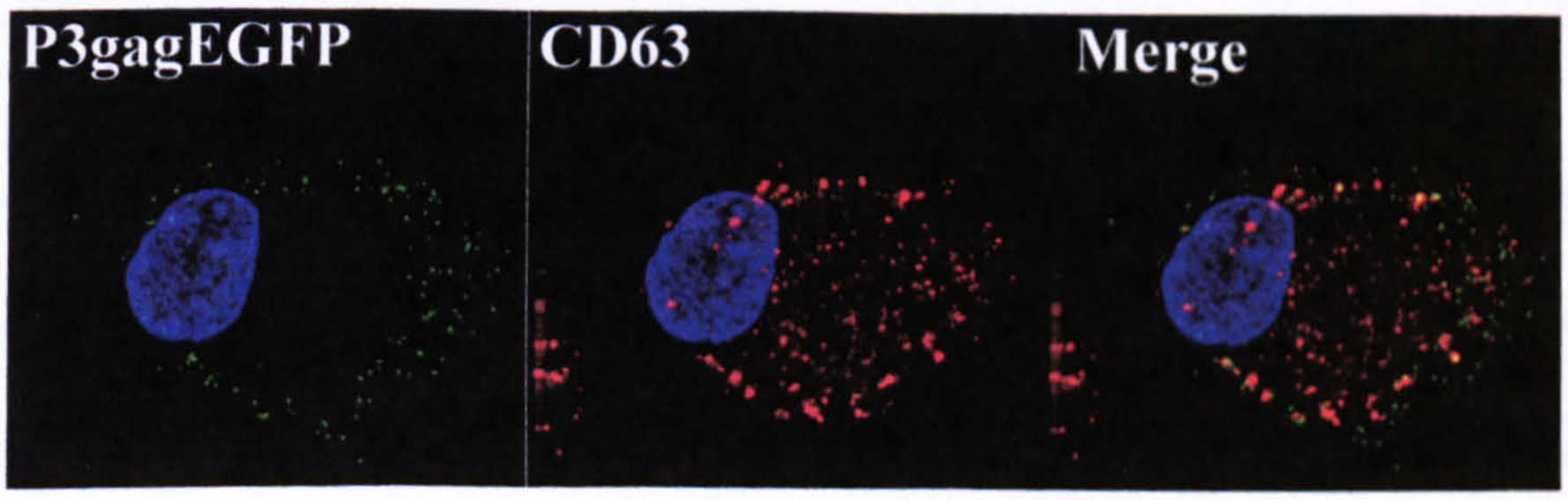


Figure 7.8 Uptake of fluorescently-labelled CD63 EC2 in MDM. (A) MDM were incubated with 0.11 M of fluorescently-labelled GST or GSTCD63 EC2 protein for 1hour at 37°C and cell-associated fluorescence measured by flow cytometry (MFI) or by analysis of images using ImageJ software (green pixels). The data are the means of 4 separate experiments ± SEM and significance of differences from GST controls assessed by two-tailed t-test; * $p<0.05$. (B) 0.11M GST-CD63 EC2-FITC proteins were added to MDM together with dextran-rhodamine (Dex) or transferrin-AlexaFluor647 (TF) for 30 minutes at 37°C. Representative images are shown. (C) MDM were incubated with 0.11M of GST-CD63 EC2 labelled with AlexaFluor647 and P3-Gag-eGFP virus for 20 minutes at RT. Microphotographs illustrating the patterns of expression were taken after fixation and representatives are shown.

7.4 DISCUSSION

The most obvious, conclusions that can be made are as follows; 1) The inhibition that we are observing with our EC2 proteins is more potent in MDM. 2) Inhibition is not HIV-1 strain specific in these cells. Although inhibition of R5 but not X4 infection is seen in PBMC it is of much less potency suggesting a macrophage specific effect. 3) The target of inhibition is the cell rather than the virus as following pre-treatment of MDM with EC2 proteins, we can wash the cells before addition of the virus and the same level of inhibition occurs. 4) As all the tetraspanin EC2 proteins that we tested could significantly inhibit infection, we are likely to be inhibiting the functioning of a tetraspanin enriched microdomain (TEM), rather than a specific tetraspanin–ligand interaction. 5) Because the VSV-pseudotyped virus could be inhibited with equal potency, the inhibition is likely to be independent of receptor/co-receptor. The mechanism of inhibition remains to be elucidated but in this discussion, I will highlight some possibilities.

A number of lentiviruses including human, simian and feline immunodeficiency viruses (HIV, SIV and FIV respectively) have acquired a wide tropism and an expanded range of target cells including cells of the monocyte/macrophage lineage and CD4+ T lymphocytes. Lentiviruses may cause depletion and immunodeficiency of CD4+ T cells but it is the ability of the virus to infect macrophages and other antigen presenting cells (APC) that determines the persistence and pathogenesis of infection. Macrophages play a central role in the immune response by either directly destroying invading pathogens or secreting cytokines that can activate the adaptive immune response thereby providing an

important cross-talk between the innate and adaptive immune systems. In the case of lentiviruses, macrophages appear to do neither of these roles and instead they provide a safe reservoir, harbouring the virus. Aside from the resistance of macrophages to the cytopathic effects of HIV, there are a number of differences between the HIV life cycle in macrophages and T cells. Firstly, different strains of HIV-1 have different tropisms for T cells and macrophages. HIV-1 strains that predominantly infect T cells are CXCR4 dependent (X4 strains) and those infecting macrophages utilise the CCR5 coreceptor (R5 strains). CXCR4 is expressed in macrophages yet there is much debate in the literature as to whether or not X4 utilising strains can infect these cells (Naif, H.M. et al. 1998; Verani et al. 1998; Valentin et al. 2000). Some have found that X4 strains can efficiently enter macrophages but are restricted at the post-entry level and can not replicate (Schmidtmayerova et al. 1998). O'Brien's group found that the inhibition they observed with anti-CD63 antibodies in primary isolates was specific to R5 utilising virus (von Lindern et al. 2003), whereas using EC2 proteins, the inhibition was independent of co-receptor. This could be an effect of EC2 proteins vs. mAb or primary isolates vs. laboratory adapted strains. The EC2 proteins have been preliminarily tested by O'Brien's laboratory and appear to be having a similar but less potent effect as with the laboratory adapted HIV-1 strains (William O'Brien personal communication).

Another possible mechanism is that endogenous tetraspanins act to regulate the expression levels of virus receptors at the cell surface, rather like the *Drosphila* tetraspanins SUN. This tetraspanin down-regulates rhodopsin in the presence of intense or prolonged exposure to light (Xu et al. 2004). In untreated cells the

expression of CD4 and coreceptors may be regulated by endogenous tetraspanins. Exogenous EC2 domains, by disrupting the formation of TEM, could prevent upregulation signals thereby decreasing levels of receptor/coreceptor expression which would in turn inhibit infection. We tested expression levels of receptor and coreceptors in cells that had been pre-treated with GST-CD63 EC2 or GST alone and did not find any decrease in expression of CD4, CCR5 nor CXCR4 in the CD63-treated cells.

7.4.1 VSV pseudotype and endocytosis

The differences between HIV-1 infection in macrophages and T cells are not limited to cell binding properties. It is becoming the general consensus that although HIV-1 assembly and budding takes place at the plasma membrane in T cells, it is thought that intracellular vesicles are the main sites for assembly and budding of virus particles in macrophages (Raposo et al. 2002; Pelchen-Matthews et al. 2003; Ono et al. 2004). Evidence for this has come from studies of newly synthesised virus membranes that found high levels of lysosomal marker proteins including the lysosomal associated proteins LAMP-1 and LAMP–2 as well as CD63 (Meerloo et al. 1992; Meerloo et al. 1993). CD63 was found in these studies to be selectively upregulated into the budding virion membrane and a later study confirmed CD63 selective incorporation into virions as well as the tetraspanins CD81 and CD82 (Pelchen-Matthews et al. 2003). Indeed, in this study anti-CD63 antibodies were shown to precipitate over 95% of infections particles from the medium (Pelchen-Matthews et al. 2003). These reports of virus and tetraspanins in endosomal compartments prompted us to consider endocytosis

as a mechanism of entry in our virus assays as oppose to receptor-mediated fusion at the plasma membrane.

Vesicular stomatitis virus (VSV) is a member of the Vesiculovirus genera of the *Rhabdovirus* family. It was first reported in 1971 that VSV enters cells by endocytosis and the basic pH-dependent mechanism has been well documented (Heine et al. 1971; Fan et al. 1978; Matlin et al. 1982); (Schlegel et al. 1981). Its single-stranded, negative sense RNA encodes 5 major proteins, glycoprotein (G), matrix protein (M), nucleoprotein (N), large protein (L,) and phosphoprotein (P). The glycoprotein (G) is of most interest here because it is responsible for the virus attachment to the host cell as well as the fusion of the virus envelope with the endosomal membrane (Fan et al. 1978; Matlin et al. 1982). Protein G contains an internal fusion peptide responsible for mediating the fusion event at acidic pH (Fan et al. 1978; Fredericksen et al. 2002). Various conformational changes take place when the virus encounters the low pH environment of endocytic vesicles, allowing exposure of the protein G and therefore membrane fusion to take place. VSV attachment and entry, therefore, does not require CD4 or the two chemokine receptors that HIV strains have adapted to exploit. To determine the dependency of inhibition that we observed with our EC2 proteins on CD4 and the coreceptors, we used a VSV-HIV pseudotype virus; that is a functional HIV virus that expresses VSV Glycoprotein G and not HIV glycoproteins. When used in the same assay as the R5 and X4 pseudotypes, it was observed that this virus was also inhibited in the same manner as the R5 pseudotype. Due to these findings we hypothesise that HIV-1 uses an endocytic pathway to enter MDM. This hypothesis also accounts for the partial inhibition of R5 strain and the complete

lack of inhibition of X4 strain in PBMC because the X4 strain is "fusogenic" and predominantly fuses with the plasma membrane and is not taken up by endocytosis (Marechal et al. 2001).

Virtually all cells use endocytosis to take up exogenous materials for nutritional value, to take up pathogens in immune mechanisms or simply to monitor the extracellular environment. There are four main strategies for the uptake of exogenous material by mammalian cells: 1) Clathrin-mediated endocytosis was the first pathway to be discovered and remains the best characterised pathway as reviewed by Schmid (Schmid 1997). This mechanism of endocytosis allows internalisation of transmembrane proteins and their bound ligands. Firstly the clathrin molecules and transmembrane proteins cluster into clathrin coated pits that eventually pinch off resulting in the formation of vesicles, 85-110 nm in diameter. Vesicles then release their cargo into endosomes, where receptors are either recycled or sent to lysosomes for degradation (Schmid 1997). The clathrin molecules, along with the different adaptor proteins, shape the vesicle and determine the selected cargo to be taken up (Schmid 1997). 2) Non-clathrin mediated endocytosis is another form of receptor-mediated endocytosis and includes calveolae. These vesicles are the smallest of the 4 types (50 to 80 nm). They are enriched with calveolin, cholesterol and spingolipids and participate in the entry of some viruses (i.e. simian virus), bacteria and macromolecules. 3) Phagocytosis involves the internalisation of large particles (>500 nm), it is an actin-dependent and usually clathrin-independent mechanism that takes place primarily in specialised phagocytic cells such as macrophages and neutrophils. Phagocytosis is critical in Metazoa for the uptake and degradation of infectious particles and apoptotic cells. There are 3 types of phagocytosis; Fc receptor

mediated, complement receptor mediated and mannose receptor mediated as reviewed by Aderem and Underhill, 1999 (Aderem et al. 1999). 4) Macropinocytosis was first discovered by Warren Lewis in 1931 (Lewis 1931). He used time-lapse microcinematography and revealed, by speeding up the movements of macrophages, membrane ruffling events that led to the formation of intracellular vesicles. Unlike clathrin-coated micropinocytosis that takes place uniformly across the cell surface, macropinocytosis usually occurs at the boundaries of membrane projections of a spread cell. Macropinosomes are heterogeneous in size and are considered to be more similar to non-clathrin mediated vesicles (Swanson et al. 1995). There are numerous examples in the literature of viruses exploiting endocytic pathways to gain entry to cells (i.e. human adenovirus is internalised by both clathrin mediated endocytosis and macropinocytosis (Meier et al. 2004). There have been two reports that describe the uptake of HIV-1 by macropinocytosis (Marechal et al. 2001; Liu et al. 2002). One report implicated macropinocytosis as the method of choice for HIV-1 to cross the blood-brain barrier. It describes the uptake of HIV-1 by brain microvascular endothelial cells (BMVEC) by macropinocytosis in a mechanism dependent on lipid rafts. The internalisation is also dependent on mitogen activated protein kinase (MAPK) signalling since inhibition of the MAPK/Erk pathway inhibited virus entry (Liu et al. 2002). The second report is specific to macrophages and so is of more interest here. Electron microscopy was used to identify the presence of HIV-1 virus particles at the plasma membrane of macrophages as well as inside clathrin coated pits and in macropinosomes (Marechal et al. 2001). It was demonstrated the macrophages were not selective in their internalisation. Non-infectious HIV-1 virions lacking *env* proteins

(HIVΔ*env*), HIV-1 virions coated with *env* and VSV-G glycoproteins were all internalised. However, the HIVΔ*env* remained in the lumen of intracellular vesicles and were not delivered to the cytosol for infection to take place. This implies that without the env proteins, although the HIV can enter macrophages via macropinocytosis, subsequent steps of the viral life cycle can not occur (Marechal et al. 2001).

There are numerous membrane fusion events within the endocytic pathway that could be inhibited with by the EC2 proteins. Interestingly, the results presented in Figure 7.7 suggest that partial inhibition takes place at the plasma membrane during binding and entry of the virus. Fluorescent microscopy studies of MDM incubated with GST or CD63 prior to infection with fluorescently labelled virus, showed a significant reduction of virus particles in the cytoplasm of CD63 treated MDM compared to GST treated cells. In addition, the inhibitory effect of CD63 was only observed when CD63-EC2 was added to the MDM prior to, or immediately after infection. Taken together, these data strongly suggest that the EC2 domains are inhibiting a receptor independent endocytic event that takes place at the plasma membrane.

7.4.2 Concluding remarks

Although this work does yield potentially interesting data on the role of tetraspanins in HIV infection of macrophages, there is still much work to be done and many questions that are currently unanswered. GST appears to have a substantial amount of anti-HIV activity which is, in some cases, comparable to that of CD4 (Figure 7.1). Glutathione *S*–transferases are a family of enzymes that play a key role in the detoxification if substances such as products of oxidative stress.

They act by catalysing the reaction between glutathione and a substrate to form a sulphur-substituted glutathione. The GST tags used in the Amersham pGEX system are from *Schistoma japonicum.* Interestingly, a high activity of the GST control in the inhibition of sperm-egg fusion was also observed (Higginbottom et al. 2003). It was clear from these experiments that, whatever the mechanism of inhibition is, it is certainly different from the mechanism used by the recombinant CD63 EC2 – the latter mechanism was lost when CD63 EC2 was added post-infection whereas the inhibition seen by GST was the same when added to MDM before and after viral infection.

Clearly, further characterisation of the effects seen with the EC2 proteins using primary HIV-1 isolates is needed here. It would also be of immense interest to determine the critical amino acids in CD63 EC2 that are required for the inhibition, and to design small peptides inhibitors based on the critical region of CD63 EC2. Also, it must be determined whether the inhibition seen by CD63 and the other tetraspanins is a general inhibition of the endocytic pathway in macrophages. Given the fact that a much less potent inhibition of HIV-1 infection is observed in PBMC, and that there appears to be some cellular localisation of CD63 EC2 with CD4 and co-receptors (Figure 7.8), it is possible that HIV-1 associates with CD4 and co-receptors at the plasma membrane of macrophages, but membrane fusion at the plasma membrane does not occur and the virus is internalised by endocytosis. Clearly, electron microscopy would be the best way to get a clearer picture of the events during inhibition of HIV by CD63 EC2.

CHAPTER 8

GENERAL DISCUSSION

8.1 Summary and perspectives

A major part of this work was to continue and optimise the production of recombinant tetraspanin EC2 domains in *E.coli*. The total yield of recombinant protein was increased without effecting the quality or timescale. This was successfully achieved by developing a method for producing large quantities of recombinant protein using a 20L biofermenter. The same media and antibiotic selection was used making the cost per litre of culture unaltered, yet the yield per litre was more than tripled.

It was also planned to clone and express the EC2 domains of CD231 and CD82. The cloning of both constructs was carried out successfully, but problems were encountered when it came to the expression of these larger EC2 domains. Neither protein would express in the bacterial host that was used to make the other recombinant EC2 proteins despite altering the general protocol for EC2 production by changing the concentrations of induction agent, the growth temperature and the density that the cultures reached before induction. Amongst the possible explanations for this was that the disulphide bridges in the protein were not forming properly leading to degradation, and so it was decided to use an alternative *E.coli* strain that promotes the formation of disulphide bridges in the bacterial cytoplasm. GST-CD82 was expressed and purified in this strain of *E. coli*, Origami (Novagen) and was detected by a widely used anti-CD82 mAb, Tspan2. This was an exciting achievement due to the greatly studied role of CD82 as a

suppressor of metastasis (reviewed in (Jackson et al. 2005), and the recombinant GST-CD82 EC2 protein will hopefully be used in functional studies in the future. Although CD231 was cloned into pGEX-KG successfully and the sequence was correct, expression was not detected in BL21 or Origami cells. Using an expression vector with a periplasmic leader sequence to direct protein folding to the periplasm would be the next obvious method to use here. Failing that, protein expression in a mammalian cell would be advisable, though yields would undoubtedly be lower. Although there were some problems encountered when producing recombinant EC2 proteins, on the whole, the method established in our laboratory, allows rapid expression and purification of high yields of recombinant EC2 proteins that have already proved to be an extremely useful tool in the study of tetraspanin proteins (Higginbottom et al. 2000; Higginbottom et al. 2003). Here, recombinant murine CD9 EC2 was shown to bind specifically to PSG17 using ELISA as previously reported by Dveksler and co-workers (Waterhouse et al. 2002). However, as our human CD9 did not bind to PSG17, we were unable to use our range of mutant human CD9 EC2 constructs to identify critical amino acids utilised in this binding. Of relevance to this work, and in support of murine CD9 binding directly to plasma proteins, a second report describing an external ligand for CD9 has recently been published. Here, it was shown that murine CD9 can function as an alternate interleukin 16 (IL-16) receptor (Qi et al. 2005). CD4 is a putative IL-16 receptor but haematopoietic cells isolated from CD4 null mice were shown to be as responsive to IL-16 as wild-type mice, and so the concept of a second IL-16 receptor was initiated (Qi et al. 2002). CD9 and other tetraspanins were postulated to be involved in the IL-16 receptor complex given their expression on haematopoietic cells and their association with CD4 and it was

demonstrated that CD9 is likely to be the major IL-16 receptor on murine mast cells (Qi et al. 2005). Although murine and human CD9 are close homologues sharing 77% homology in the EC2 domain, and can be interchanged in the sperm-egg fusion assays (Kaji et al. 2002), there is a growing amount of evidence here to suggest that this may not always be the case. In addition to human CD9 not being able to bind PSG17, murine CD9 could not inhibit monocyte fusion (Appendix and (Takeda et al. 2003). It would be interesting to investigate the interaction between IL-16 and CD9 in human cell lines and using human and mouse GST-CD9 EC2.

In addition, we showed that our recombinant CD9 EC2 did not bind to fibronectin in ELISA and FACS assays as previously reported (Longhurst et al. 2002). Our data highlights the requirement for additional proteins in the interaction between the cell and ECM; in the case of interaction with fibronectin, integrin molecules are likely to directly bind to fibronectin and CD9 is likely to be required for the accurate and intricate organisation of an involved TEM. This assumption is based on the fact that fibronectin contains 11 integrin binding sites and it is well documented that certain integrins associate with CD9 (Baudoux et al. 2000; Kawakami et al. 2002; Gutierrez-Lopez et al. 2003). To date only one group has described a direct association between CD9 and a component of the ECM.

Structural information about a protein or protein domain is an invaluable tool when attempting to determine the functional significance of a protein and the molecular mechanisms employed by a protein. Until now, the only structural information available for members of the tetraspanin family is the crystal structure of CD81 EC2 (Kitadokoro et al. 2001; Kitadokoro et al. 2001). Although tetraspanins are a

family of closely related proteins, further structural information for other members of the family would no doubt aid future studies. We attempted to solve the structure of CD63 EC2 by X-ray crystallography but were halted at the first hurdle; no protein crystals grew. 1D-NMR studies did reveal that the EC2 domain had alpha helical content and plans are in place to continue this structural study using NMR spectroscopy. The recent advances in crystallography screening methods which allow smaller and smaller proteins volumes to be screened against thousands of different crystallising conditions in a matter of hours, should speed up the process of gaining crystals to use in X-ray crystallography studies. Coupled with the advances in molecular replacement and structural modelling software, the crystal structures of more tetraspanin EC2 domains (and some of the larger EC1 domains) should be a realistic possibility within the next few years.

Probably the most exciting usages of the EC2 domains in terms of beneficial potential, is the ability to inhibit HIV-1 infection in macrophages (Chapter 7) and the inhibition of monocyte fusion (Appendix). If it is to be believed that the EC2 recombinant proteins inhibit cellular events such as fusion and viral entry, by disrupting the formation and organisation of TEMs, then the full length EC2 protein would be a poor drug to prevent HIV-1 infection in macrophages since every cell line expresses tetraspanins and is therefore likely to contain these organised clusters of molecules on their cell surface and so, disrupting the functioning of them all could have dire consequences. However, a small peptide modelled on the EC2 domain that specifically inhibits the infection of HIV-1 in human MDM would have quite the opposite effect. Macrophages harbour HIV-1 virus leaving them hidden from the host's immune attack. In this sheltered

environment they can propagate freely and be released *en-masse* to cause heightened destruction. A small peptide drug based on the CD63 EC2 domain which specifically inhibits the admission of HIV-1 into macrophages, could be used in conjunction with other therapies in anti-HIV vaccinations.

The inhibition of monocyte fusion seen by the EC2 domains of tetraspanins CD9 and CD63 (see Appendix) could also make these proteins potential therapeutic agents for quite different disorders. Osteoclasts are multinucleated giant cells that are derived from the monocyte/macrophage lineage, and in certain types of inflammatory arthritis (e.g. rheumatoid arthritis), these cells are thought to contribute to bone destruction (Suzuki et al. 2001). Anti-CD9 antibodies have already been shown to inhibit osteoclast formation (Tanio et al. 1999) and so it is feasible that small peptides based on the EC2 domains could have the potential to be developed as drugs to treat the symptoms of inflammatory diseases such as rheumatoid arthritis.

Since tetraspanins have been described in cell fusion events both here and in the literature (Kaji et al. 2000; Takeda et al. 2003), it is possible that they are involved in the control of numerous cellular fusion events. Using the recombinant tetraspanin EC2 proteins as tools, fusion events such as tumour formation, fusion of other virus membranes with host cell membranes, endocytic processes and multinucleated cell formation could be studied for any involvement of tetraspanin proteins. Any involvement of tetraspanins in such processes could be investigated quickly by pre-incubating cells with recombinant EC2 domains before addition of virus or fusogenic stimuli and viewing any differences in fusion patterns under a

microscope. The most probable role of tetraspanins in such events, should they be discovered, is the *in cis* organisation of a TEM to regulate a fusion event.

In summary, this thesis has highlighted the importance and the diverse usage of tetraspanin EC2 domains and has demonstrated for the first time that functional, recombinant EC2 proteins can be made in large quantities in biofermenters, and used as tools to study the role of tetraspanins in a variety of biological assays. The production of the EC2 proteins in this study has led to the identification of tetraspanin EC2 domains that are able to inhibit HIV infection and monocyte fusion, giving these recombinant proteins enormous and exciting potential as future therapeutic agents.

References:

Adachi, M., T. Taki, Y. Ieki, C. L. Huang, M. Higashiyama and M. Miyake (1996). "Correlation of KAI1/CD82 gene expression with good prognosis in patients with non-small cell lung cancer." Cancer Res 56(8): 1751-5.

Aderem, A. and D. M. Underhill (1999). "Mechanisms of phagocytosis in macrophages." Ann Rev Immunol 17: 593 - 623.

Anderson, D. C. and D. E. Reilly (2004). "Production technologies for monoclonal antibodies and their fragments." Curr Opin Biotechnol 15: 456-62.

Andria, M. L., G. S. Barsh and S. Levy (1992). "Expression of TAPA-1 in preimplantation mouse embryos." Biochem Biophys Res Commun 186(3): 1201-6.

Andria, M. L., C. L. Hsieh, R. Oren, U. Francke and S. Levy (1991). "Genomic organization and chromosomal localization of the TAPA-1 gene." J Immunol 147(3): 1030-6.

Ashman, L., Fitter, S, Sincock, PM, Nguyen, L, Cambareri, AC (1997). Leukocyte Typing VI. T. Kishimoto, Oxford University Press: 681-683.

Banerjee, S. A., M. Hadjiargyrou and P. H. Patterson (1997). "An antibody to the tetraspan membrane protein CD9 promotes neurite formation in a partially alpha3beta1 integrin-dependent manner." J Neurosci 17(8): 2756-65.

Bardwell, J. C. (1994). "Building bridges - disulphide bond formation in the cell." Mol Microbiol 14(2): 199-205.

Bardwell, J. C., K. McGovern and J. Beckwith (1991). "Identification of a protein required for disulfide bond formation in vivo." Cell 67(3): 581-9.

Barreiro, O., M. Yanez-Mo, M. Sala-Valdes, M. D. Gutierrez-Lopez, S. Ovalle, A. Higginbottom, P. N. Monk, C. Cabanas and F. Sanchez - Madrid (2005). "Endothelial tetraspanin microdomains regulate leukocyte firm adhesion during extravasation." Blood 105(7): 2852-61.

Bartosch, B., A. Vitelli, C. Granier, C. Goujon, J. Dubuisson, S. Pascale, E. Scarselli, R. Cortese, A. Nicosia and F. L. Cosset (2003). "Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor." J Biol Chem 278(43): 41624-30.

Battistoni, A., A. P. Mazzetti, R. Pertruzzelli, M. Muramatsu, G. Federici, G. Ricci and M. Lo Bello (1995). "Cytoplasmic and periplasmic production of human placental glutathione transferase in *Escherichia coli*." Protein Exp Pur 6: 579-587.

Baudoux, B., D. Castanares-Zapatero, M. Leclercq-Smekens, N. Berna and Y. Poumay (2000). "The tetraspanin CD9 associates with the integrin alpha6beta4 in cultured human epidermal keratinocytes and is involved in cell motility." Eur J Cell Biol 79(1): 41-51.

Berditchevski, F. (2001). "Complexes of tetraspanins with integrins: more than

meets the eye." J Cell Sci 114(Pt 23): 4143-51.
Berditchevski, F., E. Gilbert, M. R. Griffiths, S. Fitter, L. Ashman and S. J. Jenner (2001). "Analysis of the CD151-alpha3beta1 integrin and CD151-tetraspanin interactions by mutagenesis." J Biol Chem 276(44): 41165-74.
Berditchevski, F. and E. Odintsova (1999). "Characterization of integrin-tetraspanin adhesion complexes: role of tetraspanins in integrin signaling." J Cell Biol 146(2): 477-92.
Berditchevski, F., E. Odintsova, S. Sawada and E. Gilbert (2002). "Expression of the palmitoylation-deficient CD151 weakens the association of alpha 3 beta 1 integrin with the tetraspanin-enriched microdomains and affects integrin-dependent signaling." J Biol Chem 277(40): 36991-7000.
Berditchevski, F., K. F. Tolias, K. Wong, C. L. Carpenter and M. E. Hemler (1997). "A novel link between integrins, transmembrane-4 superfamily proteins (CD63 and CD81), and phosphatidylinositol 4-kinase." J Biol Chem 272(5): 2595-8.
Berditchevski, F., M. M. Zutter and M. E. Hemler (1996). "Characterization of novel complexes on the cell surface between integrins and proteins with 4 transmembrane domains (TM4 proteins)." Mol Biol Cell 7(2): 193-207.
Bessette, P. H., F. Auslund, J. Beckwith and G. Georgiou (1999). "Efficient folding of proteins with multiple disulphide bonds in the *Escherichia coli* cytoplasm." Proc Natl Acad Sci U S A 96: 13703-13708.
Bienstock, R. J. and J. C. Barrett (2001). "KAI1, a prostate metastasis suppressor: prediction of solvated structure and interactions with binding partners; integrins, cadherins, and cell-surface receptor proteins." Mol Carcinog 32(3): 139-53.
Bonifacino, J. S. and E. C. Dell'Angelica (1999). "Molecular basis for the recognition of tyrosine-based sorting signals." J Biol Chem 145: 923-26.
Boucheix, C., P. Benoit, P. Frachet, M. Billard, R. E. Worthington, J. Gagnon and G. Uzan (1991). "Molecular cloning of the CD9 antigen. A new family of cell surface proteins." J Biol Chem 266(1): 117-22.
Boucheix, C., Duc, GH, Jasmin, C, Rubinstein, E (2001). Tetraspanins and malignancy. Expert Rev Mol Med. 31: 1-17.
Boucheix, C. and E. Rubinstein (2001). "Tetraspanins." Cell Mol Life Sci 58(9): 1189-205.
Bowen, J. A. and J. S. Hunt (2000). "The role f integrins in reproduction." Proc Soc Exp Biol Med 223(4): 331-343.
Branden, C. and J. Tooze (1999). Introduction to Protein Structure. New York, Garland Publishing Inc.
Cabrita, L. and S. P. Bottomley (2004). "Protein expression and refolding - A practical guide to getting the mot out of inclusion bodies." Biotechnology Annual Review 10: 31-50.
Cannon, K. S. and P. Cresswell (2001). "Quality control of transmembrane domain assembly in the tetraspanin CD82." Embo J 20(10): 2443-53.
Carter, C. W., J. E. T. Baldwin and L. Frick (1988). "Statistical design of experiments for protein crystal growth and the use of a precrystallization assay." J Crystal Growth(90): 60 - 73.

Carter, C. W. and C. W. Carter (1979). "Protein crystalisation using incomplete factorial experiments." J Biol Chem 254(23): 12219-12223.

Chakrabarti, L. A., T. Ivanovic and C. Cheng-Mayer (2002). "Properties of the surface envelope glycoprotein associated with virulence of simian-human immunodeficiency virus SHIV(SF33A) molecular clones." J Virol 76(4): 1588-99.

Charrin, S., F. Le Naour, V. Labas, M. Billard, J. P. Le Caer, J. F. Emile, M. A. Petit, C. Boucheix and E. Rubinstein (2003). "EWI-2 is a new component of the tetraspanin web in hepatocytes and lymphoid cells." Biochem J 373(Pt 2): 409-21.

Charrin, S., F. Le Naour, M. Oualid, M. Billard, G. Faure, S. M. Hanash, C. Boucheix and E. Rubinstein (2001). "The major CD9 and CD81 molecular partner. Identification and characterization of the complexes." J Biol Chem 276(17): 14329-37.

Charrin, S., S. Manie, M. Billard, L. Ashman, D. Gerlier, C. Boucheix and E. Rubinstein (2003). "Multiple levels of interactions within the tetraspanin web." Biochem Biophys Res Commun 304(1): 107-12.

Charrin, S., S. Manie, M. Oualid, M. Billard, C. Boucheix and E. Rubinstein (2002). "Differential stability of tetraspanin/tetraspanin interactions: role of palmitoylation." FEBS Lett 516(1-3): 139-44.

Charrin, S., S. Manie, C. Thiele, M. Billard, D. Gerlier, C. Boucheix and E. Rubinstein (2003). "A physical and functional link between cholesterol and tetraspanins." Eur J Immunol 33(9): 2479-89.

Chen, M. S., K. S. Tung, S. A. Coonrod, Y. Takahashi, D. Bigler, A. Chang, Y. Yamashita, P. W. Kincade, J. C. Herr and J. M. White (1999). "Role of the integrin-associated protein CD9 in binding between sperm ADAM 2 and the egg integrin alpha6beta1: implications for murine fertilization." Proc Natl Acad Sci U S A 96(21): 11830-5.

Cherukuri, A., R. H. Carter, S. Brooks, W. Bornmann, R. Finn, C. S. Dowd and S. K. Pierce (2004). "B Cell Signaling is Regulated by Induced Palmitoylation of CD81." J Biol Chem 279(30): 31973 - 31982.

Cherukuri, A., T. Shoham, H. W. Sohn, S. Levy, S. Brooks, R. Carter and S. K. Pierce (2004). "The tetraspanin CD81 is necessary for partitioning of coligated CD19/CD21-B cell antigen receptor complexes into signaling-active lipid rafts." J Immunol 172(1): 370-80.

Chi-Rosso, G., P. J. Gotwals, J. Yang, L. Ling, K. Jiang, B. Chao, D. P. Baker, L. C. Burkly, S. E. Fawell and V. E. Koteliansky (1997). "Fibronectin type III repeats mediate RGD-independent adhesion and signaling through activated beta1 integrins." J Biol Chem 272(50): 31447-52.

Claas, C., C. S. Stipp and M. E. Hemler (2001). "Evaluation of prototype transmembrane 4 superfamily protein complexes and their relation to lipid rafts." J Biol Chem 276: 7974- 7984.

Clapham, P., K. Nagy, R. Cheingsong-Popov, M. Exley and R. A. Weiss (1983). "Productive infection and cell-free transmission of human T-cell leukemia virus in a nonlymphoid cell line." Science 222(4628): 1125-7.

Clark, K. L., A. Oelke, M. E. Johnson, K. D. Eilert, P. C. Simpson and S. C. Todd (2004). "CD81 associates with 14-3-3 in a redox regulated palmitoylation dependent manner." J Biol Chem.

Clark, K. L., Z. Zeng, A. L. Langford, S. M. Bowen and S. C. Todd (2001).

"PGRL is a major CD81-associated protein on lymphocytes and distinguishes a new family of cell surface proteins." J Immunol **167**(9): 5115-21.

Connor, R. I., K. E. Sheridan, C. Lai, L. Zhang and D. D. Ho (1996). "Characterization of the functional properties of env genes from long-term survivors of human immunodeficiency virus type 1 infection." J Virol **70**(8): 5306-11.

Cook, G. A., C. M. Longhurst, S. Grgurevich, S. Cholera, J. T. Crossno, Jr. and L. K. Jennings (2002). "Identification of CD9 extracellular domains important in regulation of CHO cell adhesion to fibronectin and fibronectin pericellular matrix assembly." Blood **100**(13): 4502-11.

Cook, G. A., D. A. Wilkinson, J. T. Crossno, Jr., R. Raghow and L. K. Jennings (1999). "The tetraspanin CD9 influences the adhesion, spreading, and pericellular fibronectin matrix assembly of Chinese hamster ovary cells on human plasma fibronectin." Exp Cell Res **251**(2): 356-71.

Dale, G. E., C. Oefner and A. D'Arcy (2003). "The protein as a variable in protein crystallisation." J Structural Biology **142**: 88-97.

de Parseval, A., D. L. Lerner, P. Borrow, B. J. Willett and J. H. Elder (1997). "Blocking of feline immunodeficiency virus infection by a monoclonal antibody to CD9 is via inhibition of virus release rather than interference with receptor binding." J Virol **71**(8): 5742-9.

Dedhar, S. and V. Gray (1990). "Isolation of a novel integrin receptor mediating Arg-Gly-Asp-directed cell adhesion to fibronectin and type I collagen from human neuroblastoma cells. Association of a novel beta 1-related subunit with alpha v." J Cell Biol **110**(6): 2185-93.

Delaguillaumie, A., J. Harriague, S. Kohanna, G. Bismuth, E. Rubinstein, M. Seigneuret and H. Conjeaud (2004). "Tetraspanin CD82 controls the association of cholesterol-dependent microdomains with the actin cytoskeleton in T lymphocytes: relevance to co-stimulation." J Cell Sci **117**: 5269-82.

Domanico, S. Z., A. J. Pelletier, W. L. Havran and V. Quaranta (1997). "Integrin alpha 6A beta 1 induces CD81-dependent cell motility without engaging the extracellular matrix migration substrate." Mol Biol Cell **8**(11): 2253-65.

Edman, J. C., L. Ellis, R. W. Blacher, R. A. Roth and W. J. Rutter (1985). "Sequence of protein disulphide isomerase and implications of its relationship to thioredoxin." Nature **317**(6034): 267-70.

Ellerman, D. A., C. Ha, P. Primakoff, D. G. Myles and G. S. Dveksler (2003). "Direct binding of the ligand PSG17 to CD9 requires a CD9 site essential for sperm-egg fusion." Mol Biol Cell **14**(12): 5098-103.

Fan, D. P. and B. M. Sefton (1978). "The entry into host cells of Sindbis virus, vesicular stomatitis virus and sendai virus." Cell Adhes Commun **15**: 985-992.

Flint, M., C. Maidens, L. D. Loomis-Price, C. Shotton, J. Dubuisson, P. Monk, A. Higginbottom, S. Levy and J. A. McKeating (1999). "Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81." J Virol **73**(8): 6235-44.

Fradkin, L. G., J. T. Kamphorst, A. DiAntonio, C. S. Goodman and J. N. Noordermeer (2002). "Genomewide analysis of the Drosophila

tetraspanins reveals a subset with similar function in the formation of the embryonic synapse." Proc Natl Acad Sci U S A **99**(21): 13663-8.

Fredericksen, B. L., B. L. Wei, J. Yao, T. Luo and J. V. Garcia (2002). "Inhibition of endosomal/lysosomal degradation increases the infectivity of humn immunodeficiency virus." J Virol **76**(22): 11440-11446.

Fukudome, K., M. Furuse, T. Imai, M. Nishimura, S. Takagi, Y. Hinuma and O. Yoshie (1992). "Identification of membrane antigen C33 recognized by monoclonal antibodies inhibitory to human T-cell leukemia virus type 1 (HTLV-1)-induced syncytium formation: altered glycosylation of C33 antigen in HTLV-1-positive T cells." J Virol **66**(3): 1394-401.

Garcia-Lopez, M. A., O. Barreiro, A. Garcia-Diez, F. Sanchez-Madrid and P. F. Penas (2005). "Role of Tetraspanins CD9 and CD151 in Primary Melanocyte Motility." J Invest Dermatol **125**(5): 1001-9.

Gluschankof, P., I. Mondor, H. R. Gelderblom and Q. J. Sattentau (1997). "Cell membrane vesicles are a major contaminant of gradient-enriched human immunodeficiency virus type-1 preparations." Virology **230**(1): 125-33.

Goldberg, A., Molday, RS. (1996). Defective subunit assembly underlies a digenic form of retinitis pigmentosa linked to mutations in peripherin/rds and rom-1. PNAS. **26:** 13726-30.

Goulding, C. W. and L. J. Perry (2003). "Protein production in *Escherichia coli* for structural studies by X-ray crystallography." J Structural Biology **142**: 133-143.

Gourgues, M., P. H. Clergeot, C. Veneault, J. Cots, S. Sibuet, A. Brunet-Simon, C. Levis, T. Langin and M. H. Lebrun (2002). "A new class of tetraspanins in fungi." Biochem Biophys Res Commun **297**(5): 1197-204.

Greene, W. C. (2004). "The brightening future of HIV therapeutics." Nat Immunol **5**: 867-71.

Guan, K. and J. E. Dixon (1991). "Eukaryotic protein expressed in *E. coli:* An improvedhrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase." Anal Biochem **192**: 262-267.

Gujral, A. (2003). Rapid and efficient cloning into multiple systems for expression and functional analysis. Focus. Invitrogen Life Technologies. **25.1**.

Gutierrez-Lopez, M. D., S. Ovalle, M. Yanez-Mo, N. Sanchez-Sanchez, E. Rubinstein, N. Olmo, M. A. Lizarbe, F. Sanchez-Madrid and C. Cabanas (2003). "A functionally relevant conformational epitope on the CD9 tetraspanin depends on the association with activated beta1 integrin." J Biol Chem **278**(1): 208-18.

Hadjiargyrou, M., Z. Kaprielian, N. Kato and P. H. Patterson (1996). "Association of the tetraspan protein CD9 with integrins on the surface of S-16 Schwann cells." J Neurochem **67**(6): 2505-13.

Heine, J. W. and C. A. Schnaitman (1971). "Entry of vesicular stomatitis virus into L cells." J Virol **8**(5): 786-795.

Helfrich, W., H. J. Haisma, V. Magdolen, T. Luther, V. J. Bom, J. Westra, R. van der Hoeven, B. J. Kroesen, G. Molema and L. de Leij (2000). "A rapid and versitile method for harnessing scFv antibody fragments with various biological effector functions." J Immunol Methods **237**((1-2)):

131-45.
Hemler, M. E., B. A. Mannion and F. Berditchevski (1996). "Association of TM4SF proteins with integrins: relevance to cancer." Biochim Biophys Acta **1287**(2-3): 67-71.
Higginbottom, A., E. R. Quinn, C. C. Kuo, M. Flint, L. H. Wilson, E. Bianchi, A. Nicosia, P. N. Monk, J. A. McKeating and S. Levy (2000). "Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2." J Virol **74**(8): 3642-9.
Higginbottom, A., Y. Takahashi, L. Bolling, S. A. Coonrod, J. M. White, L. J. Partridge and P. N. Monk (2003). "Structural requirements for the inhibitory action of the CD9 large extracellular domain in sperm/oocyte binding and fusion." Biochem Biophys Res Commun **311**(1): 208-14.
Hohdatsu, T., H. Hirabayashi, K. Motokawa and H. Koyama (1996). "Comparative study of the cell tropism of feline immunodeficiency virus isolates of subtypes A, B and D classified on the basis of the env gene V3-V5 sequence." J Gen Virol **77 (Pt 1)**: 93-100.
Holmgren, A. (1985). "Thioredoxin." Annu Rev Biochem **54**: 237-71.
Horejsi, V. and C. Vlcek (1991). "Novel structurally distinct family of leucocyte surface glycoproteins including CD9, CD37, CD53 and CD63." FEBS Lett **288**(1-2): 1-4.
Hotta, H., Ross, A.H., Huebner, K., Isobe, M., Wendeborn, S., Chao, m.V., Ricciardi, R.P., Tsujimoto, Y., Croce, C.M and Koprowski, H. (1988). "Molecular cloning and characterisation of an antigen associated with early stages of melanoma tumour progression." Cancer Res **48**(11): 2955 - 62.
Houle, C. D., X. Y. Ding, J. F. Foley, C. A. Afshari, J. C. Barrett and B. J. Davis (2002). "Loss of expression and altered localization of KAI1 and CD9 protein are associated with epithelial ovarian cancer progression." Gynecol Oncol **86**(1): 69-78.
Hsu, M., J. Zhang, M. Flint, C. Logvinoff, C. Cheng-Mayer, C. M. Rice and J. A. McKeating (2003). "Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles." Proc Natl Acad Sci U S A **100**(12): 7271-6.
Huang, C. I., N. Kohno, E. Ogawa, M. Adachi, T. Taki and M. Miyake (1998). "Correlation of reduction in MRP-1/CD9 and KAI1/CD82 expression with recurrences in breast cancer patients." Am J Pathol **153**(3): 973-83.
Ikeyama, S., M. Koyama, M. Yamaoko, R. Sasada and M. Miyake (1993). "Suppression of cell motility and metastasis by transfection with human motility-related protein (MRP-1/CD9) DNA." J Exp Med **177**(5): 1231-7.
Israels, S. J., E. M. McMillan-Ward, J. Easton, C. Robertson and A. McNicol (2001). "CD63 associates with the alphaIIb beta3 integrin-CD9 complex on the surface of activated platelets." Thromb Haemost **85**(1): 134-41.
Jackson, P., A. Marreiros and P. J. Russell (2005). "KAI1 tetraspanin and metastasis supressor." Int J Biochem Cell Biol **37**(3): 530-4.
Jefferis, R. and J. Lund (2002). "Interaction sites on human IgG-Fc for FcgammaR: current models." Immunol Lett **82**(1-2): 57-65.
Jones, M., Mason, DY (1997). in Leukocyte Typing VI, ed. T. Kishimoto et al.

Oxford University Press: 681-683.
Kaji, K., S. Oda, S. Miyazaki and A. Kudo (2002). "Infertility of CD9-deficient mouse eggs is reversed by mouse CD9, human CD9, or mouse CD81; polyadenylated mRNA injection developed for molecular analysis of sperm-egg fusion." Dev Biol **247**(2): 327-34.
Kaji, K., S. Oda, T. Shikano, T. Ohnuki, Y. Uematsu, J. Sakagami, N. Tada, S. Miyazaki and A. Kudo (2000). "The gamete fusion process is defective in eggs of Cd9-deficient mice." Nat Genet **24**(3): 279-82.
Kamitani, S., Y. Akiyama and K. Ito (1992). "Identification and characterization of an Escherichia coli gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme." Embo J **11**(1): 57-62.
Kanetaka, K., M. Sakamoto, Y. Yamamoto, S. Yamasaki, F. Lanza, T. Kanematsu and S. Hirohashi (2001). "Overexpression of tetraspanin CO-029 in hepatocellular carcinoma." J Hepatol **35**(5): 637-42.
Kawakami, Y., K. Kawakami, W. F. Steelant, M. Ono, R. C. Baek, K. Handa, D. A. Withers and S. Hakomori (2002). "Tetraspanin CD9 is a "proteolipid," and its interaction with alpha 3 integrin in microdomain is promoted by GM3 ganglioside, leading to inhibition of laminin-5-dependent cell motility." J Biol Chem **277**(37): 34349-58.
Kipriyanov, S. M., G. Moldenhauer and M. Little (1997). "Hight level production of sluble single chain antibodies in small-scale *Escherichia coli* cultures." J Immunol Methods **200**: 69-77.
Kitadokoro, K., D. Bordo, G. Galli, R. Petracca, F. Falugi, S. Abrignani, G. Grandi and M. Bolognesi (2001). "CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs." Embo J **20**(1-2): 12-8.
Kitadokoro, K., G. Galli, R. Petracca, F. Falugi, G. Grandi and M. Bolognesi (2001). "Crystallization and preliminary crystallographic studies on the large extracellular domain of human CD81, a tetraspanin receptor for hepatitis C virus." Acta Crystallogr D Biol Crystallogr **57**(Pt 1): 156-8.
Klassen, H., M. R. Schwartz, A. H. Bailey and M. J. Young (2001). "Surface markers expressed by multipotent human and mouse neural progenitor cells include tetraspanins and non-protein epitopes." Neurosci Lett **312**(3): 180-2.
Kovalenko, O. V., X. Yang, T. V. Kolesnikova and M. E. Hemler (2004). "Evidence for specific tetraspanin homodimers: inhibition of palmitoylation makes cysteine residues available for cross-linking." Biochem J **377**(Pt 2): 407-17.
Kronenberger, B., C. Sarrazin, W. P. Hofmann, M. von Wagner, E. Herrmann, C. Welsch, R. Elez, B. Ruster, A. Piiper and S. Zeuzem (2004). "Mutations in the putative HCV-E2 CD81 binding regions and correlation with cell surface CD81 expression." J Viral Hepat **11**(4): 310-8.
Lagaudriere-Gesbert, C., Conjeaud, H (1997). Leukocyte Typing VI. T. Kishimoto, Oxford University Press: 681-683.
Le Naour, F., E. Rubinstein, C. Jasmin, M. Prenant and C. Boucheix (2000). "Severely reduced female fertility in CD9-deficient mice." Science **287**(5451): 319-21.
Lesley, C. A., P. Kuhn, A. Godzik, A. M. Deacon, I. Matthews, A. Kreusch, G.

Spraggon, H. E. Klock, D. McMullan, T. Shin, J. Vincent, A. Robb, L. S. Brinen, M. D. Miller, T. M. McPhillips, M. A. Miller, D. Scheibe, J. M. Canaves, C. Guda, L. Jaroszewski, T. L. Selby, M. A. Elsliger, J. Wooley, S. S. Taylor, K. O. Hodgson, I. A. Wilson, P. G. Schultz and R. C. Stevens (2002). "Structural genomics of the thermatoga maritima proteome implemented in a high-throughput structure determination pipeline." Proc Natl Acad Sci U S A **99**: 11664-11669.

Levy, R., R. Weiss, G. Chen, B. L. Iverson and G. Gergiou (2001). "Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli trx*B *gor* mutants via the coexpression of molecular chaperones." Protein Exp Pur **23**: 338-347.

Levy, S., V. Q. Nguyen, M. L. Andria and S. Takahashi (1991). "Structure and membrane topology of TAPA-1." J Biol Chem **266**(22): 14597-602.

Levy, S., Todd, SC, Maecker, HT (1998). CD81 (TAPA-1): a molecule involved in signal
transduction and cell adhesion in the immune system. Annu Rev Immunol.

Lewis, W. (1931). "Pinocytosis." Johns hopkins hospital bull **49**: 17-27.

Lijovic, M., G. Somers and A. G. Frauman (2002). "KAI1/CD82 protein expression in primary prostate cancer and in BPH associated with cancer." Cancer Detect Prev **26**(1): 69-77.

Lin, T. C., J. Ruch, E. K. Spicer and W. H. Konigsberg (1987). "Cloning and expression of T4 DNA polymerase." Proc Natl Acad Sci U S A **84**(20): 7000-7004.

Lin, T. M., S. P. Halbert and W. N. Spellacy (1974). "Measurement of pregnancy associated plasma proteins during human gestation." J Clin Invest **54**: 567-582.

Linder, M. E. and R. J. Deschenes (2003). "New Insights into the Mechanisms of Protein Palmitoylation." Biochemistry **42**(15): 4311 - 4320.

Liu, N. Q., A. S. Lossinsky, W. Popik, X. Li, C. Gujuluva, B. Kriederman, J. Roberts, T. Pushkarsky, M. Bukrinsky, M. Witte, M. Weinard and M. Fiala (2002). "Human immunodefficiency virus type 1 enters brain microvascular endothelia by macropinocytosis dependent on lipid rafts and the mitogen activated protein kinase signaling pathway." J Virol **76**(13): 6689 - 6700.

Loffler, S., F. Lottspeich, F. Lanza, D. O. Azorsa, V. ter Meulen and J. Schneider-Schaulies (1997). "CD9, a tetraspan transmembrane protein, renders cells susceptible to canine distemper virus." J Virol **71**(1): 42-9.

Longhurst, C. M., J. D. Jacobs, M. M. White, J. T. Crossno, Jr., D. A. Fitzgerald, J. Bao, T. J. Fitzgerald, R. Raghow and L. K. Jennings (2002). "Chinese hamster ovary cell motility to fibronectin is modulated by the second extracellular loop of CD9. Identification of a putative fibronectin binding site." J Biol Chem **277**(36): 32445-52.

Longhurst, C. M., M. M. White, D. A. Wilkinson and L. K. Jennings (1999). "A CD9, alphaIIbbeta3, integrin-associated protein, and GPIb/V/IX complex on the surface of human platelets is influenced by alphaIIbbeta3 conformational states." Eur J Biochem **263**(1): 104-11.

Lozach, P. Y., H. Lortat-Jacob, A. de Lacroix de Lavalette, I. Staropoli, S. Foung, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J. L. Virelizier, F. Arenzana-Seisdedos and R. Altmeyer (2003). "DC-SIGN and L-SIGN are high affinity binding receptors for hepatitis C virus

glycoprotein E2." J Biol Chem **278**(22): 20358-66.
Marechal, V., M. Prevost, C. Petit, E. Perret, J.-M. Heard and O. Schwartz (2001). "Human Immunodeficiency virus type 1 entry into macrophages by macropinocytosis." J Virol **75**(22): 11166 - 11177.
Marlink, R., P. Kanki, I. Thior, K. Travers, G. Eisen, T. Siby, I. Traore, C. C. Hsieh, M. C. Dia and E. H. Gueye (1994). "Reduced rate of disease develpment after HIV-2 infection as compared to HIV-1." Science **265**: 1587-90.
Martin, F., D. M. Roth, D. A. Jans, C. W. Pouton, L. J. Partridge, P. N. Monk and G. W. Moseley (2005). "Tetraspanins in viral infections: a fundemental role in viral biology." J Virol: 10839-51.
Masciopinto, F., S. Campagnoli, S. Abrignani, Y. Uematsu and P. Pileri (2001). "The small extracellular loop of CD81 is necessary for optimal surface expression of the large loop, a putative HCV receptor." Virus Res **80**(1-2): 1-10.
Matlin, K. S., H. Reggio, A. Helenius and K. Simon (1982). "Pathway of vesicular stomititis virus entry leading o infection." J Mol Biol **156**: 609-631.
McKeating, J. A., L. Q. Zhang, C. Logvinoff, M. Flint, J. Zhang, J. Yu, D. Butera, D. D. Ho, L. B. Dustin, C. M. Rice and P. Balfe (2004). "Diverse hepatitis C virus glycoproteins mediate viral infection in a CD81-dependent manner." J Virol **78**(16): 8496-505.
McMahan, R. H., L. Watson, R. Meza-Romero, G. G. Burrows, D. N. Bourdette and A. C. Buenafe (2003). "Production, characterisation, and immunogenicity of a soluble rat single chain T cell receptor specific for an encephalitogenic peptide." J Biol Chem **278**(33): 30961-70.
McPherson, A. (1990). "Current approaches to macromolecular crystallisation." Eur J Biochem **189**: 1-23.
Meerloo, T., H. K. Parmentier, A. D. Osterhaus, J. Goudsmit and H. J. Schuurman (1992). "Modulation of cell surface molecules during HIV-1 infection of H9 cells. An immunoelectron microscopic study." Aids **6**(10): 1105-16.
Meerloo, T., M. A. Sheikh, A. C. Bloem, A. de Ronde, M. Schutten, C. A. van Els, P. J. Roholl, P. Joling, J. Goudsmit and H. J. Schuurman (1993). "Host cell membrane proteins on human immunodeficiency virus type 1 after in vitro infection of H9 cells and blood mononuclear cells. An immuno-electron microscopic study." J Gen Virol **74 (Pt 1)**: 129-35.
Meier, O. and U. F. Greber (2004). "Adenovirus endocytosis." J Gene Med **6**: 153 - 161.
Miller, B. J., E. Georges-Labouesse, P. Primakoff and D. G. Myles (2000). "Normal fertilization occurs with eggs lacking the integrin alpha6beta1 and is CD9-dependent." J Cell Biol **149**(6): 1289-96.
Missiakas, D., C. Georgopoulos and S. Raina (1993). "Identification and characterization of the Escherichia coli gene dsbB, whose product is involved in the formation of disulfide bonds in vivo." Proc Natl Acad Sci U S A **90**(15): 7084-8.
Missiakas, D. and S. Raina (1997). "Protein folding in the bacterial periplam." J Bacteriology **179**(8): 2467-2471.
Miyado, K., G. Yamada, S. Yamada, H. Hasuwa, Y. Nakamura, F. Ryu, K. Suzuki, K. Kosai, K. Inoue, A. Ogura, M. Okabe and E. Mekada

(2000). "Requirement of CD9 on the egg plasma membrane for fertilization." Science **287**(5451): 321-4.
Moore, J. P. and R. W. Doms (2003). "The entry of entry inhibitors: a fusion of science and medicine." Proc Natl Acad Sci U S A **100**: 10598-602.
Mossner, E., M. Huber-Wunderlich, A. Reitsch, J. Beckwith, R. Glockshuber and F. Aslund (1999). "Importance of redox potential for the in vivo function of the cytoplasmic disulphide reductant thioredoxin from *Escherichia coli*." J Biol Chem **274**: 25254-25259.
Mukherjee, K. L., D. C. D. Rowe, N. A. Watkins and D. K. Summers (2004). "Studies of Single-Chain Antibody expression in quiescent *Escherichia coli*." App Env Micro **70**(5): 3005-3012.
Murdoch, J. N., K. Doudney, D. Gerrelli, N. Wortham, C. Paternotte, P. Stainer and A. J. Copp (2003). "Genomic organisation and embryonic expression of igsf8, an immunoglobulin superfamily member implicated in development of the nervous system and organ epithelia." Mol Cell Neurosci **22**(1): 62 - 74.
Naif, H. M., S. Li, M. Alali, A. Sloane, L. MWu, M. Kelly, G. Lynch, A. Lloyd and A. L. Cunningham (1998). "CCR5 expression correlates with susceptibility of maturing monocytes to himan immunodefficiency virus type 1 infection." J Virology **72**: 830-836.
Naif, H. M., S. Li, M. Alali, A. Sloane, L. Wu, M. Kelly, G. Lynch, A. Lloyd and A. L. Cunningham (1998). "CCR5 expression correlates with susceptibility of maturing monocytes to human immunodeficiency virus type 1 infection." J Virol **72**(1): 830-6.
Ono, A. and E. O. Freed (2004). "Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body." J Virol **78**(3): 1552-63.
Ono, M., K. Handa, S. Sonnino, D. A. Withers, H. Nagai and S. Hakomori (2001). "GM3 ganglioside inhibits CD9-facilitated haptotactic cell motility: coexpression of GM3 and CD9 is essential in the downregulation of tumor cell motility and malignancy." Biochemistry **40**(21): 6414-21.
Orentas, R. J. and J. E. Hildreth (1993). "Association of host cell surface adhesion receptors and other membrane proteins with HIV and SIV." AIDS Res Hum Retroviruses **9**(11): 1157-65.
Pankov, R. and K. M. Yamada (2004). "Fibronectin at a glance." J Cell Sci **151**: 3861-3863.
Pelchen-Matthews, A., B. Kramer and M. Marsh (2003). "Infectious HIV-1 assembles in late endosomes in primary macrophages." J Cell Biol **162**(3): 443-55.
Penas, P. F., A. Garcia-Diez, F. Sanchez-Madrid and M. Yanez-Mo (2000). "Tetraspanins are localized at motility-related structures and involved in normal human keratinocyte wound healing migration." J Invest Dermatol **114**(6): 1126-35.
Pileri, P., Y. Uematsu, S. Campagnoli, G. Galli, F. Falugi, R. Petracca, A. J. Weiner, M. Houghton, D. Rosa, G. Grandi and S. Abrignani (1998). "Binding of hepatitis C virus to CD81." Science **282**(5390): 938-41.
Pique, C., C. Lagaudriere-Gesbert, L. Delamarre, A. R. Rosenberg, H. Conjeaud and M. C. Dokhelar (2000). "Interaction of CD82 tetraspanin proteins with HTLV-1 envelope glycoproteins inhibits cell-to-cell

fusion and virus transmission." Virology 276(2): 455-65.
Pohlmann, S., J. Zhang, F. Baribaud, Z. Chen, G. J. Leslie, G. Lin, A. Granelli-Piperno, R. W. Doms, C. M. Rice and J. A. McKeating (2003). "Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR." J Virol 77(7): 4070-80.
Qi, J. C., R. L. Stevens and R. Wadley (2002). "IL-16 regulation of human mast cells/basophils and their susceptibility to HIV-1." J Immunol 168: 4127-4134.
Qi, J. C., J. Wang, S. Mandadi, K. Tanaka, B. D. Roufogalis, M. C. Madigan, K. Lai, F. Yan, B. H. Chong, R. L. Stevens and S. A. Krilis (2005). "Human and mouse mast cells use the tetraspanin CD9 as an alternative interleukin 16 receptor." Blood.
Radaev, S. and P. Sun (2002). "Recognition of immunoglobulins by Fcgamma receptors." Mol Immunol 38(14): 1073-83.
Raposo, G., M. Moore, D. Innes, R. Leijendekker, A. Leigh-Brown, P. Benaroch and H. Geuze (2002). "Human macrophages accumulate HIV-1 particles in MHC II compartments." Traffic 3(10): 718-29.
Rhodes, G. (2000). Crystallography Made Clear. London, Academic Press.
Rivas, A., C. L. Ruegg, J. Zeitung, R. Laus, R. Warnke, C. Benike and E. G. Engleman (1995). "V7, a novel leukocyte surface protein that participates in T cell activation. I. Tissue distribution and functional studies." J Immunol 154: 4423-4433.
Roccasecca, R., H. Ansuini, A. Vitelli, A. Meola, E. Scarselli, S. Acali, M. Pezzanera, B. B. Ercole, J. McKeating, A. Yagnik, A. Lahm, A. Tramontano, R. Cortese and A. Nicosia (2003). "Binding of the hepatitis C virus E2 glycoprotein to CD81 is strain specific and is modulated by a complex interplay between hypervariable regions 1 and 2." J Virol 77(3): 1856-67.
Rous, B. A., B. J. Reaves, G. Ihrke, J. A. G. Briggs, S. R. Gray, D. J. Stephens, G. Banting and J. P. Luzio (2002). "Role of adaptor complex AP-3 in targeting wild-type and mutated CD63 to lysosomes." Mol Biol Cell 13: 1071-1082.
Rubinstein, E., F. Le Naour, C. Lagaudriere-Gesbert, M. Billard, H. Conjeaud and C. Boucheix (1996). "CD9, CD63, CD81, and CD82 are components of a surface tetraspan network connected to HLA-DR and VLA integrins." Eur J Immunol 26(11): 2657-65.
Rubinstein, E., A. Ziyyat, M. Prenant, E. Wrobel, J. P. Wolf, S. Levy, F. Le Naour and C. Boucheix (2005). "Reduced fertility of female mice lacking CD81." Dev Biol.
Ruegg, C. L., A. Rivas, N. D. Madani, J. Zeitung, R. Laus and E. G. Engleman (1995). "V7, a novel leukocyte surface protein that participates in T cell activation. II. Molecular cloning and characterization." J Immunol 154: 4434.
Santala, V. and U. Lamminmaki (2004). "Production of biotinylated single-chain antibody fragment in the cytoplasm of Escherichia coli." J Immunol Methods 284(1-2): 165-175.
Sasaki, M., K. Yamauchi, T. Nakanishi, Y. Kamogawa and N. Hayashi (2003). "In vitro binding of hepatitis C virus to CD81-positive and -negative human cell lines." J Gastroenterol Hepatol 18(1): 74-79.
Scarselli, E., H. Ansuini, R. Cerino, R. M. Roccasecca, S. Acali, G. Filocamo,

C. Traboni, A. Nicosia, R. Cortese and A. Vitelli (2002). "The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus." Embo J **21**(19): 5017-25.

Schlegel, R., M. Willingham and I. Pastan (1981). "Monesin blocks endocytosis of vesicular stomatitis virus." Biochem Biophys Res Commun **102**(3): 992-998.

Schmid, E., A. Zurbriggen, U. Gassen, B. Rima, V. ter Meulen and J. Schneider-Schaulies (2000). "Antibodies to CD9, a tetraspan transmembrane protein, inhibit canine distemper virus-induced cell-cell *fusion but not virus-cell* fusion." J Virol **74**(16): 7554-61.

Schmid, S. L. (1997). "Clathrin-coated vesicle *formation and protein sorting:* an integrated process." Ann Rev Biochem **66**: 511-548.

Schmidtmayerova, H., M. Alfano, G. Nuovo and M. Bukrinsky (1998). "Human immunodeficiency virus type 1 T-lymphotropic strains enter macrophages via a CD4- and CXCR4-mediated pathway: replication is restricted at a post-entry level." J Virol **72**: 4633-4642.

Schneider, C., W. E. Boeglin and A. R. Brash (2005). "Human cyclo-oxygenase-1 and an alternative splice variant: contrasts in expression of mRNA, protein and catalytic activities." Biochem J **385**(1): 57-64.

Schober, J. M., N. Chen, T. M. Grzeszkiewicz, I. Jovanovic, E. E. Emeson, T. P. Ugarova, R. D. Ye, L. F. Lau and S. C. Lam (2002). "Identification of integrin alpha(M)beta(2) as an adhesion receptor on peripheral blood monocytes for Cyr61 (CCN1) and connective tissue growth factor (CCN2): immediate-early gene products expressed in atherosclerotic lesions." Blood **99**(12): 4457-65.

Seigneuret, M. (2006). "Complete predicted three-dimensional strucutre of the facilitator transmembrane protein and hepititis C virus receptor CD81: Conserved and variable structural domains in the tetraspanin superfamily." Biophsical Journal **90**: 212-227.

Seigneuret, M., A. Delaguillaumie, C. Lagaudriere-Gesbert and H. Conjeaud (2001). "Structure of the tetraspanin main extracellular domain. A partially conserved fold with a structurally variable domain insertion." J Biol Chem **276**(43): 40055-64.

Sengoku, K., N. Takuma, T. Miyamoto, M. Horikawa and M. Ishikawa (2004). "Integrins are not involved in the process of human sperm-oolemma lfusion." Hum Reprod **19**(3): 639-44.

Serru, V., F. Le Naour, M. Billard, D. O. Azorsa, F. Lanza, C. Boucheix and E. Rubinstein (1999). "Selective tetraspan-integrin complexes (CD81/alpha4beta1, CD151/alpha3beta1, CD151/alpha6beta1) under conditions disrupting tetraspan interactions." Biochem J **340 (Pt 1)**: 103-11.

Silvie, O., Rubinstein, E, Franetich, JF, Prenant, M, Belnoue, E, Renia, L, Hannoun, L, Eling, W, Levy, S, Boucheix, C, Mazier, D (2003). Hepatocyte CD81 is required for Plasmodium falciparum and Plasmodium yoelii sporozoite infectivity. Nat Med. **9**: 93-6.

Simmons, L. C., D. Reilly, L. Klimowski, T. S. Raju, G. Meng, P. Sims, K. Hong, R. L. Shields, L. A. Damico, P. Rancatore and D. G. Yansura (2002). "Expression of full-length immunoglobulins in *Escherichia coli*: rapid and efficient production of aglycosylated antibodies." J Immunol Methods **263**: 133-147.

Simons, D. C. and D. L. Vander Jagt (1981). "Purification of glutathione S - transferases by glutathione-affinity chromatography." Methods Enzymol **77**: 235 - 237.

Simons, K. and E. Ikonen (1997). "Functional rafts in cell membranes." Nature **387**: 569-72.

Simons, P. C. and D. L. Vander Jagt (1977). "Purification of glutathione S-transferases from human liver by glutathione-affinity chromotography." Anal Biochem **82**: 334-341.

Sincock, P., Mayrhofer, G, Ashman, LK (1997). Localization of the transmembrane 4 superfamily (TM4SF) member PETA-3 (CD151) in normal human tissues: comparison with CD9, CD63, and alpha5beta1 integrin. J Histochem Cytochem. **45:** 515-25.

Smith, D. B. and K. S. Johnson (1988). "Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase." Gene **67**: 31-40.

Snyder, S. K., D. H. Wessner, J. L. Wessells, R. M. Waterhouse, L. M. Wahl, W. Zimmermann and G. S. Dveksler (2001). "Pregnancy-specific glycoproteins function as immunomodulators by inducing secretion of IL-10, IL-6 and TGF-β by human monocytes." AJRI **45**: 205-216.

Sorensen, H. P. and K. K. Mortensen (2004). "Advanced genetic strategies for recombinant protein expression in *Escherichia coli*." J Biotechnology **115**: 113-138.

Sorensen, H. P. and K. K. Mortensen (2005). "Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*." Microbial Cell Factories **4**: 1-8.

Sterk, L. M., C. A. Geuijen, J. G. van den Berg, N. Claessen, J. J. Weening and A. Sonnenberg (2002). "Association of the tetraspanin CD151 with the laminin-binding integrins alpha3beta1, alpha6beta1, alpha6beta4 and alpha7beta1 in cells in culture and in vivo." J Cell Sci **115**(Pt 6): 1161-73.

Stipp, C. S., T. V. Kolesnikova and M. E. Hemler (2001). "EWI-2 is a major CD9 and CD81 partner and member of a novel Ig protein subfamily." J Biol Chem **276**(44): 40545-54.

Stipp, C. S., T. V. Kolesnikova and M. E. Hemler (2003). "EWI-2 regulates alpha3beta1 integrin-dependent cell functions on laminin-5." J Cell Biol **163**(5): 1167-77.

Stipp, C. S., T. V. Kolesnikova and M. E. Hemler (2003). "Functional domains in tetraspanin proteins." Trends Biochem Sci **28**(2): 106-12.

Stipp, C. S., Kolesnikova, T.V., Helmer M.E. (2003). "Functional domians in tetraspanin proteins." Trends Biochem Sci **28**(2): 106 - 12.

Suzuki, Y., Y. Tsutsumi, M. Nakagawa, H. Suzuki, K. Matsushita, M. Beppu, H. Aoki, Y. Ichikawa and Y. Mizushima (2001). "Osteoclast-like cells in an *in vitro* model of bone destruction by rheumatoid arthritis." Rheumatology **40**: 673-682.

Swanson, J. A. and C. Watt (1995). "Macropinocytosis." Trends in cell biology **5**: 424-428.

Takahashi, Y., D. Bigler, Y. Ito and J. M. White (2001). "Sequence-specific interaction between the disintegrin domain of mouse ADAM 3 and murine eggs: role of beta1 integrin-associated proteins CD9, CD81, and CD98." Mol Biol Cell **12**(4): 809-20.

Takeda, Y., I. Tachibana, K. Miyado, M. Kobayashi, T. Miyazaki, T. Funakoshi, H. Kimura, H. Yamane, Y. Saito, H. Goto, T. Yoneda, M. Yoshida, T. Kumagai, T. Osaki, S. Hayashi, I. Kawase and E. Mekada (2003). "Tetraspanins CD9 and CD81 function to prevent the fusion of mononuclear phagocytes." J Cell Biol **161**(5): 945-56.

Tanio, Y., H. Yamazaki, T. Kunisada, K. Miyake and S. I. Hayashi (1999). "CD9 molecule expressed on stromal cells is involved in osteoclastogenesis." Exp Hematol **27**(5): 853-9.

Thomas, J. G. and F. Baneyx (1996). "Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat shock proteins." J Biol Chem **271**(19): 11141-11147.

Todd, S. C., V. S. Doctor and S. Levy (1998). "Sequences and expression of six new members of the tetraspanin/TM4SF family." Biochim Biophys Acta **1399**(1): 101-4.

Tole, S. and P. H. Patterson (1993). "Distribution of CD9 in the developing and mature rat nervous system." Dev Dyn **197**(2): 94-106.

Tomlinson, M. G., T. Hanke, D. A. Hughes, A. N. Barclay, E. Scholl, T. Hunig and M. D. Wright (1995). "Characterization of mouse CD53: epitope mapping, cellular distribution and induction by T cell receptor engagement during repertoire selection." Eur J Immunol **25**(8): 2201-5.

Tomlinson, M. G., A. F. Williams and M. D. Wright (1993). "Epitope mapping of anti-rat CD53 monoclonal antibodies. Implications for the membrane orientation of the Transmembrane 4 Superfamily." Eur J Immunol **23**(1): 136-40.

Valentin, A., H. Trivedi, W. Lu, L. G. Kostrikis and G. N. Pavlakis (2000). "CXCR4 mediates entry and productive infection of syncytia-inducing (X4) HIV-1 strains in primary isolates." Virology **269**: 294-304.

Van Heeke, G. and S. M. Schuster (1989). "Expression of human asparagine synthetase in *Escherichia coli*." J Biol Chem **264**(10): 5503-5509.

VanCompernolle, S. E., A. V. Wiznycia, J. R. Rush, M. Dhanasekaran, P. W. Baures and S. C. Todd (2003). "Small molecule inhibition of hepatitis C virus E2 binding to CD81." Virology **314**(1): 371-80.

Verani, A., E. Pesenti, S. Polo, E. Tresoldi, G. Scarlatti, P. Luso, A. G. Siccardi and D. Vercelli (1998). "CXCR4 is a functional coreceptor for infection of human macrophages by CXCR4 dependent primary HIV-1 isolates." J Immunol **161**: 2084 - 2088.

von Lindern, J. J., D. Rojo, K. Grovit-Ferbas, C. Yeramian, C. Deng, G. Herbein, M. R. Ferguson, T. C. Pappas, J. M. Decker, A. Singh, R. G. Collman and W. A. O'Brien (2003). "Potential role for CD63 in CCR5-mediated human immunodeficiency virus type 1 infection of macrophages." J Virol **77**(6): 3624-33.

Waterhouse, R., C. Ha and G. S. Dveksler (2002). "Murine CD9 is the receptor for pregnancy-specific glycoprotein 17." J Exp Med **195**(2): 277-82.

Watt, F. M. and K. J. Hodivala (1994). "Cell adhesion. Fibronectin and integrin knockouts come unstuck." Curr Biol **4**(3): 270-2.

Willett, B., M. Hosie, A. Shaw and J. Neil (1997). "Inhibition of feline immunodeficiency virus infection by CD9 antibody operates after virus entry and is independent of virus tropism." J Gen Virol **78 (Pt 3)**: 611-8.

Willett, B. J., M. J. Hosie, O. Jarrett and J. C. Neil (1994). "Identification of a

putative cellular receptor for feline immunodeficiency virus as the feline homologue of CD9." Immunology **81**(2): 228-33.

Wong, G. E., X. Zhu, C. E. Prater, E. Oh and J. P. Evans (2001). "Analysis of fertilin alpha (ADAM1)-mediated sperm-egg cell adhesion during fertilization and identification of an adhesion-mediating sequence in the disintegrin-like domain." J Biol Chem **276**(27): 24937-45.

Wright, M., G. W. Moseley and A. B. Vanspriel (2004). "Tetraspanin micrdomains in immune cell signalling and malignant disease." Tissue Antigens **64**: 533-542.

Wright, M., Tomlinson, M (1994). The ins and outs of the transmembrane-4 superfamily. Immunology Today. **15:** 588-94.

Wu, L., W. A. Paxton, N. Kassam, N. Ruffing, J. B. Rottman, N. Sullivan, H. Choe, J. Sodroski, W. Newman, R. A. Koup and C. R. Mackay (1997). "CCR5 levels and expression pattern correlate with infectability by macrophage-tropic HIV-1, in vitro." J Exp Med **185**(9): 1681-91.

Wunschmann, S., J. D. Medh, D. Klinzmann, W. N. Schmidt and J. T. Stapleton (2000). "Characterization of hepatitis C virus (HCV) and HCV E2 interactions with CD81 and the low-density lipoprotein receptor." J Virol **74**(21): 10055-62.

Xu, H., S. J. Lee, E. Suzuki, K. D. Dugan, A. Stoddard, H. S. Li, L. A. Chodosh and C. Montell (2004). "A lysosomal tetraspanin associated with retinal degeneration identified via a genome-wide screen." Embo J **23**(4): 811-22.

Yang, X., C. Claas, S. K. Kraeft, L. B. Chen, Z. Wang, J. A. Kreidberg and M. E. Hemler (2002). "Palmitoylation of tetraspanin proteins: modulation of CD151 lateral interactions, subcellular distribution, and integrin-dependent cell morphology." Mol Biol Cell **13**(3): 767-81.

Yokosaki, Y., E. L. Palmer, A. L. Prieto, K. L. Crossin, M. A. Bourdon, R. Pytela and D. Sheppard (1994). "The integrin alpha 9 beta 1 mediates cell attachment to a non-RGD site in the third fibronectin type III repeat of tenascin." J Biol Chem **269**(43): 26691-6.

Zemni R, B. T., Vinet MC, Sefiani A, Carrie A, Billuart P, McDonell N, Couvert P, Francis F, Chafey P, Fauchereau F, Friocourt G, des Portes V, Cardona A, Frints S, Meindl A, Brandau O, Ronce N, Moraine C, van Bokhoven H, Ropers HH, Sudbrak R, Kahn A, Fryns JP, Beldjord C, Chelly J (2000). A new gene involved in X-linked mental retardation identified by analysis of an X;2 balanced translocation. Nat Genet. **24:** 167-70.

Zhang, J., G. Randall, A. Higginbottom, P. Monk, C. M. Rice and J. A. McKeating (2004). "CD81 is required for hepatitis C virus glycoprotein-mediated viral infection." J Virol **78**(3): 1448-55.

Zhang, S., Y. Yang and Y. Shi (2005). "Characterization of human SCD2, an oligomeric desaturase with improved stability and enzyme activity by cross-linking in intact cells." Biochem J **388**(1): 135-42.

Zhang, X. A., A. L. Bontrager and M. E. Hemler (2001). "Transmembrane-4 superfamily proteins associate with activated protein kinase C (PKC) and link PKC to specific beta(1) integrins." J Biol Chem **276**(27): 25005-13.

Zhang, X. A., W. S. Lane, S. Charrin, E. Rubinstein and L. Liu (2003). "EWI2/PGRL associates with the metastasis suppressor KAI1/CD82

and inhibits the migration of prostate cancer cells." Cancer Res **63**(10): 2665-74.

Zhou, B., L. Liu, M. Reddivari and X. A. Zhang (2004). "The Palmitoylation of Metastasis Suppressor KAI1/CD82 Is Important for Its Motility- and Invasiveness-Inhibitory Activity." Cancer Research **64**: 7455 - 7463.

Zhu, G. Z., B. J. Miller, C. Boucheix, E. Rubinstein, C. C. Liu, R. O. Hynes, D. G. Myles and P. Primakoff (2002). "Residues SFQ (173-175) in the large extracellular loop of CD9 are required for gamete fusion." Development **129**(8): 1995-2002.

Zhu, X. and J. P. Evans (2002). "Analysis of the roles of RGD-binding integrins, alpha(4)/alpha(9) integrins, alpha(6) integrins, and CD9 in the interaction of the fertilin beta (ADAM2) disintegrin domain with the mouse egg membrane." Biol Reprod **66**(4): 1193-202.

APPENDIX

TETRASPANINS CD63 AND CD9 ARE INVOLVED IN THE FORMATION OF MULTINUCLEATED GIANT CELLS BY HUMAN MONOCYTES.

[1]Varadarajan Parthasarathy, [1,2]Francine Martin, [2]Adrian Higginbottom, [1]Helen Murray, [1§]Gregory W. Moseley, [2]Robert C. Read, [2]Peter N. Monk and [1]Lynda J. Partridge

Running title: Tetraspanins inhibit fusion of monocytes

Department of Molecular Biology and Biotechnology, University of Sheffield, S10 2TN, UK[1] and Division of Genomic Medicine, University of Sheffield Medical School, Sheffield S10 2RX, UK[2].

[§]**Current address:** Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3080, Australia.

Introduction

Certain tetraspanins have been implicated in cell fusion processes. Most notably, CD9 has been shown to play a critical role in sperm-egg fusion, since oocytes from CD9 knockout mice are unable to fuse with sperm, resulting in infertility (Kaji et al. 2000; Le Naour et al. 2000; Miyado et al. 2000). Fusion is restored by the expression of ectopic CD9 (Kaji et al. 2002; Zhu et al. 2002) and it appears that the tetraspanin CD81 can compensate for loss of CD9 (Kaji et al. 2002; Zhu et al. 2002). Monoclonal antibodies (mAbs) to the EC2 region of CD9, as well as recombinant proteins corresponding to CD9 EC2, inhibit fusion (Chen et al. 1999; Zhu et al. 2002; Higginbottom et al. 2003). CD9 and CD81 are also implicated in muscle cell fusion (Tachibana et al. 1999) and CD9, CD81 and CD82 have been linked with virus-induced syncitium formation (reviewed in (Martin et al. 2005)).

Most recently a role for CD9 and CD81 in the formation of multinucleated giant cells has been proposed (Takeda et al. 2003). Antibodies to CD9 and CD81, but not CD63, were found to enhance Con A induced fusion of human monocytes and mouse alveolar macrophages in vitro. In addition, macrophages from CD9 and CD81 null mice showed enhanced formation of multinucleated cells in vitro and CD9/CD81 double knockout mice showed spontaneous formation of multi-nucleated giant cells. This led the authors to speculate that CD9 and CD81 normally act to negatively regulate monocytes/macrophage fusion (Takeda et al. 2003). Here we report the use of recombinant tetraspanin EC2 domains to more fully investigate the role of tetraspanins in

monocyte fusion. We show for the first time that the EC2 domains of tetraspanin CD63 but not CD81, CD82 or CD151 can inhibit monocyte fusion leading to multi-nucleated giant cell formation. Thus it is likely that both CD63 and CD9 but not CD81 have roles in the fusion process, suggesting that the role of tetraspanins is more complex than previously thought.

Materials and Methods

Monocyte fusion assay

Peripheral blood monocytes were derived from peripheral whole blood of healthy, anonymized human volunteers by Ficoll-Hypaque density centrifugation as described elsewhere (Stevanin et al., 2005).

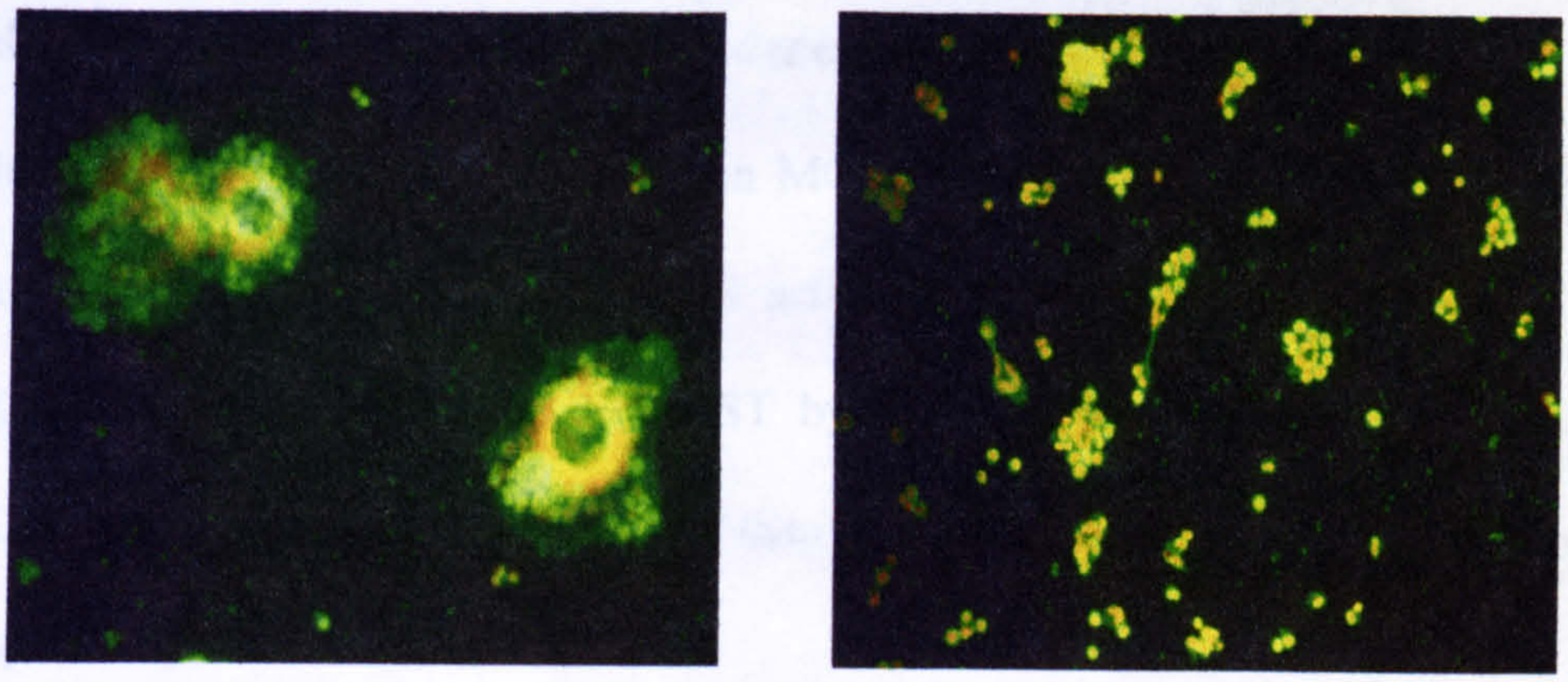
The study was approved by the South Sheffield Research Ethics Committee (protocol number SSREC/02/299). Mononuclear cells were seeded at $5x10^5$cells.chamber^{-1} in 0.5ml RPMI-1640-10% FCS in an 8 chambered slide (Lab-Tek®). After overnight culture, non-adherent cells were removed by washing three times and adherent cells cultured in RPMI-1640 + 10% FCS in the presence/absence of 10μg.ml^{-1} Con A (Sigma) for 72 hrs at 37°C. Recombinant tetraspanin EC2s or GST protein were added to some assays. The cells were washed with PBS, fixed and permeabilised with acetone (5 mins, room temperature), rehydrated with PBS then labelled with FITC-anti-CD52 and the nuclei counter-stained with propidium iodide. Fusion rates were determined essentially as described previously (Most et al. 1990) by counting the number of nuclei in fused cells (>3 nuclei per cell) and unfused cells in 6 randomly chosen fields using a Nikon Eclipse E400 immunofluorescence microscope (fusion rate = no. of nuclei in fused cells/ number of nuclei in total (%)). In some cases, the numbers of nuclei in MGCs in 6 random fields was recorded.

Results

Effects of GST-EC2 tetraspanins on MGC formation

Soluble forms of tetraspanin EC2 regions have been shown to have biological activity in a number of systems including sperm-egg fusion (Zhu et al. 2002; Higginbottom et al. 2003), hepatitis C virus binding (Flint et al. 1999; Higginbottom et al. 2000; Petracca et al. 2000), and leukocyte transmigration (Barreiro et al. 2005). We therefore investigated a range of recombinant human tetraspanin EC2 domains produced as GST fusion proteins (GST only, GST-CD9 EC2, CD63 EC2, CD81 EC2, CD82 EC2, CD151 EC2) for their effects on MGC formation. The effects of GST mouse CD9 EC2 (GST-mCD9), and a human CD9 EC2 with a mutation in one of the conserved cysteines involved in disulfide bond formation (GST-CD9C153A) were also investigated (Fig. 1). GST alone had no effect on cell fusion, and neither did the soluble tetraspanin EC2 proteins GST-CD81, GST-CD82 or GST-CD151. The lack of an effect with GST-CD81 EC2 is unexpected in view of previous observations that anti-CD81 antibodies enhance MGC formation (Takeda et al., 2003) Recombinant GST-CD63 and GST-CD9, however, showed significant inhibition of MGC formation. Our results are consistent, however, with those of Takeda and co-workers on MGC formation. The effect of the GST-CD9 appears to be sequence-specific, since mouse CD9-EC2, which has 77% identity with human CD9 EC2, had no effect. Interestingly, however, a CD9-EC2 containing a C153A mutation had a slight but significant negative effect on MGC formation. This mutant is not recognised by conformation-specific antibodies and would not be expected to contain a correctly folded subloop in its EC2 region.

A.



B.

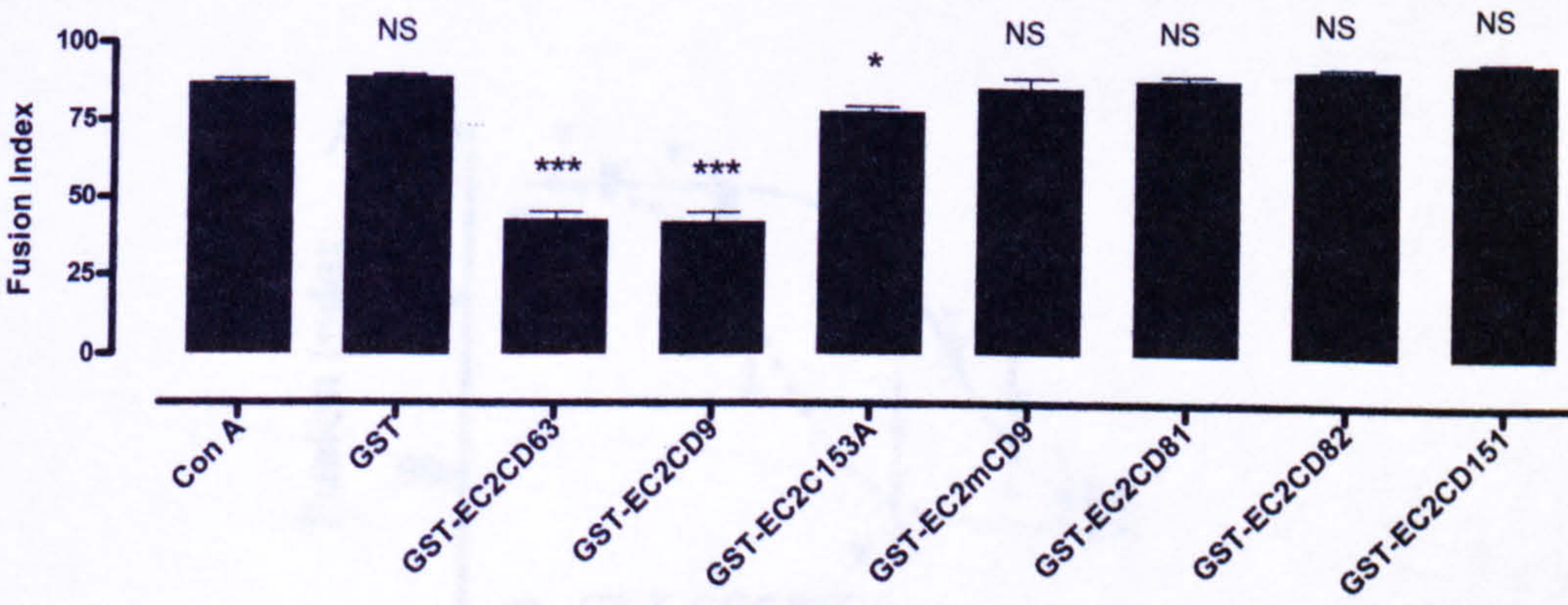


Figure 1. **Effects of different GST-EC2 tetraspanins on Con A induced monocyte fusion**. Monocytes in Lab-Tek® chamber slides were stimulated with Con A (10μg.ml^{-1}) for 72hrs in the presence or absence of 20μg/ml GST-EC2 tetraspanins or GST control. A) shows ConA-stimulated monocytes lablled with propidium iodide and anti-CD52 antibody in the absence (left panel) and presence (right panel) of GST-CD63 EC2. B) shows data obtained using a range of EC2 domains. The results with Con A, GST, GST-CD63 EC2 and GST-C153A EC2 are the means +/- SEM of 17 experiments in duplicates carried out with monocytes from 12 blood donors; the results of GST-CD9 EC2 are the means +/- SEM of 12 experiments in duplicate using monocytes from 12 blood donors and the results of other GST-EC2 tetraspanins are the means +/- SEM of 3 experiments in duplicate with monocytes from 3 different blood donors. The level of significance between the Con A control and different treatment was tested by unpaired t test, *** $P<0.0001$; * $P<0.01$; NS $P>0.05$.

Inhibition of MGC formation is not dependent on the presence of GST

Although GST alone had no effect on MGC formation, we considered that the fusion protein might affect tetraspanin EC2 activity. To investigate this, we cleaved GST-CD63EC2 and removed the free GST by affinity chromatography (Fig 2). Cleaved CD63EC2 was even more potent than GST-CD63EC2 (IC_{50} 1.6nM and 59nM, respectively).

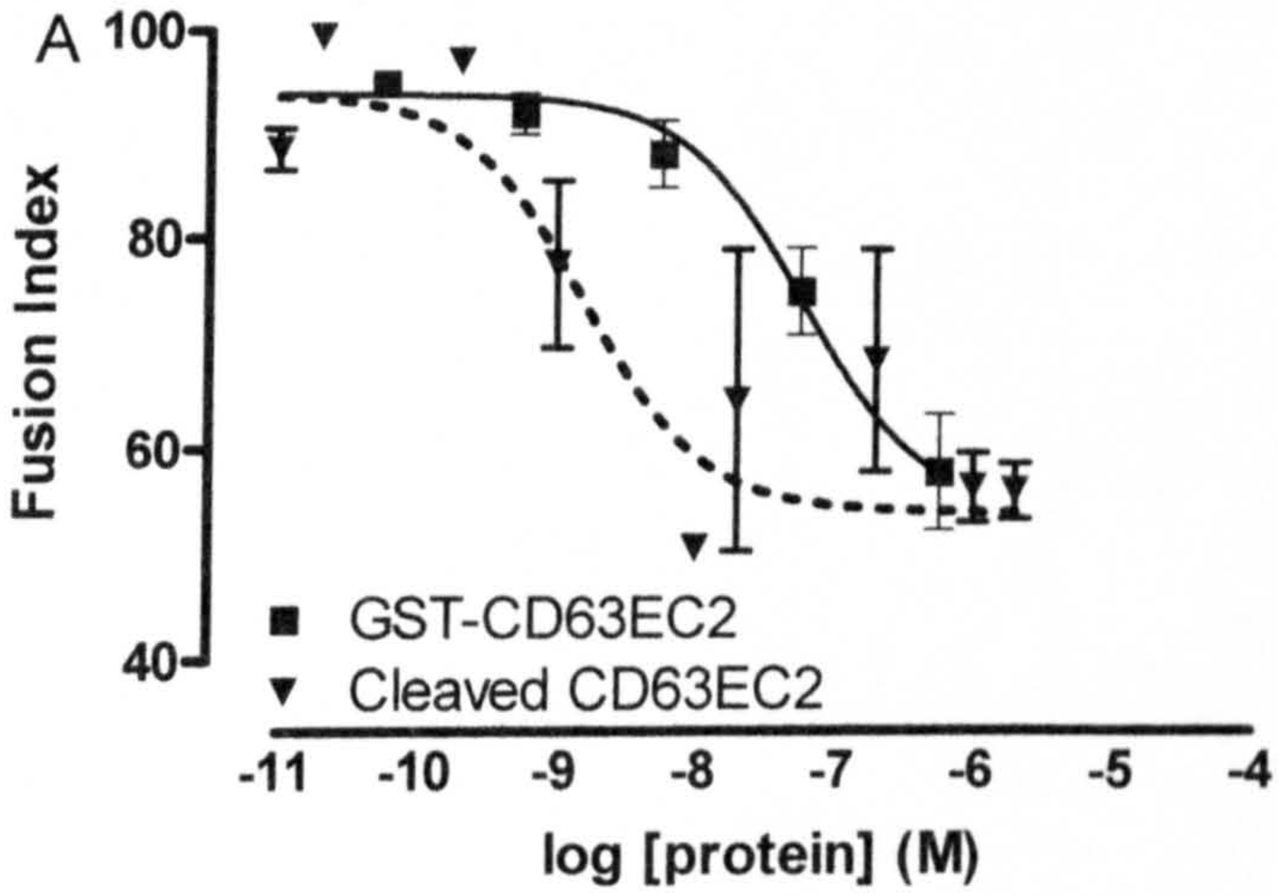


Figure 2. **Effects of different GST-CD63 EC2 and cleaved CD63 EC2 on Con A induced monocyte fusion**. Monocytes in Lab-Tek® chamber slides were stimulated with Con A ($10\mu g.ml^{-1}$) for 72hrs in the presence of GST-CD63 EC2 or cleaved CD63 EC2 and the monocyte fusion index was calculated.

Conclusions

Overall, our results indicate roles for tetraspanins CD9 and CD63 in MGC formation. Since inhibitory effects were observed with recombinant EC2 regions, we can conclude that these regions of the CD9 and CD63 molecules are involved. The effects are tetraspanin-specific, since no inhibition was observed with recombinant EC2s of CD81, CD82 or CD151. In addition, no inhibition was observed with murine CD9 EC2, which shows 77% sequence identity in this region.

.