CHAPTER 3: INVESTIGATION OF THE PROTEIN-PROTEIN INTERACTIONS OF FAM/USP9X

INTRODUCTION AND CHAPTER SUMMARY

Introduction

FAM is now known to interact with a large and growing list of proteins, some of which are substrates like AF-6, β-catenin, epsin 1, and itch, and others that are accessory and/or regulatory proteins for FAM function (DCX). Many of these interactions are not fully characterised with binding sites in FAM either unknown or partially mapped to relatively large 500-900 amino acid stretches. The structural organisation of functional protein-protein interaction domains within FAM is also poorly characterised. The purpose of this investigation therefore, was to analyse the protein-protein interactions of FAM using two approaches; i) to more finely characterize and map the binding sites on FAM for known protein interactions and ii) to identify and characterize new structural features of FAM that may be involved in protein-protein interactions.

The group of proteins that are known to interact with FAM have quite a diverse range of functions within the cell, from cell-cell adhesion and cell polarity, to endocytosis and protein trafficking. Most of what is known about the cellular consequences of interactions with these FAM binding proteins has come from studies where the levels of full-length FAM protein have been altered (Chen et al., 2003; Mouchantaf et al., 2006; Murray et al., 2004). The major limitation of this approach is that by altering total levels of FAM in the cell, the function and/or stability of all of FAM's substrates and binding partners may be affected. This makes interpretation of a specific FAM-substrate interaction very difficult. Much more information could be gained from studies in which the ability of FAM to interact with each protein is disrupted in isolation.

In order to undertake such studies it would first be necessary to identify the amino acids in FAM that are responsible for each individual interaction. Subsequently this information could be used in a number of ways to specifically disrupt interaction with each FAM binding partner *in vivo*. This could be done by generating an antibody against the binding region that could then be microinjected into cells. In a similar manner, small molecule inhibitors could be designed to the surface of the binding site however this would require three-dimensional structural information about the interface surface. This would necessitate the identification of an independently folding domain containing the binding site and is beyond the scope of this project. Alternatively, if the binding site could be mapped to a small enough region of FAM, the amino acids responsible for the bulk of the interaction could be

identified through mutational screens. This would allow the generation of a mutant of the full length FAM protein that ablates or decreases the strength of binding to one substrate or interacting protein while leaving the remainder of the protein intact and able to function as normal.

As discussed above, very little information about the structure of the FAM protein was available at the time that this research was begun. Domain homology database searches, while predicting the region corresponding to the DUB catalytic activity of FAM, were unable to provide any additional information about structural features. One study that was aimed at establishing a standard procedure for using the suite of programs that are becoming available to aid in the prediction of protein structure, helped to provide some additional insight into FAM structure. This study utilised the large number of proteins found on the male-specific region of the human Y-chromosome, many of which not only have unknown function, but also have unknown protein structure (Ginalski et al., 2004). One of the proteins analysed in this screen was the human Y homologue of FAM, USP9Y. Analysis of the domain/3D fold structure of USP9Y proved to be quite difficult and large sections of the protein were proposed to be largely unstructured however a number of previously unknown features were predicted. Aside from the known catalytic core region, three long α-helical regions located on either side of the catalytic core region were identified and it was postulated that these may form right handed superhelical structures. Much more surprisingly, the computational analysis predicted a β-grasp domain fold structure in the N-terminal region of FAM (Ginalski et al., 2004). The β-grasp fold structure describes the three-dimensional configuration that ubiquitin and the ubiquitin-like proteins adopt. β-grasp fold domains can also be found in a diverse range of proteins where they appear to function as a protein-protein interaction domains and have been termed Ubiquitin-like (UBL) domains (Hartmann-Petersen and Gordon, 2004).

FAM functions in the ubiquitin-proteasome pathway, being a deubiquitylating enzyme that has demonstrated the ability to modulate the protein levels and half-lives of its substrates (Chen et al., 2003; Pantaleon et al., 2001; Taya et al., 1999; Taya et al., 1998). However, FAM also has links to ubiquitin-mediated protein trafficking, through its interaction with proteins know to be involved in these processes and its subcellular localisation (Chen et al., 2003; Friocourt et al., 2005; Mouchantaf et al., 2006; Murray et al., 2004). UBL domain containing proteins have been shown to be involved in both processes and for this reason attempts were made to identify proteins able to bind to the putative UBL domain (UBLD) of FAM (Alberti et al., 2002; Fallon et al., 2006; Heir et al., 2006; Kamitani et al., 2001; Luders et al., 2000; Sakata et al., 2003; Tanaka et al., 2003).

Chapter Summary

This chapter describes experiments that were designed to more finely map the interaction sites on FAM for known binding partners and also to investigate the potential protein-protein interaction properties of the predicted UBLD of FAM. To investigate known interactions of FAM, three independent and commonly applied *in vitro* assays were chosen with the aim of identifying regions of FAM that were able to interact with three of its known binding partners, β-catenin, epsin 1, and itch. At the time that this work was begun, these were the only proteins that were known to bind to FAM whose interaction interface had not been completely characterized; the binding site on FAM for AF-6 had been partially characterized in affinity chromatography experiments and found to lie in the catalytic core region of the protein (bounded by amino acids 1476-1918) (Taya et al., 1998). Although it had been observed that the catalytic core region of FAM was also able to interact with the armadillo repeat region of β-catenin, there was no published data on the ability of the remainder of the FAM protein to interact with β-catenin (Taya et al., 1999). Due to the fact that FAM's catalytic core is likely to be essential for its normal function, and that any changes within this region are likely to cause the enzyme to become non-functional, binding sites outside of this region were investigated in preference.

A yeast two-hybrid assay was used to investigate the interaction of FAM with β-catenin, epsin 1, and itch. In addition, the ability of FAM to interact with the cytoplasmic domain of E-cadherin was investigated using this assay. In each case, an interaction was not observed despite receiving positive results from control strains that were included in each assay. As an alternative to the yeast two-hybrid system, a mammalian two-hybrid system was also investigated, but technical difficulties caused this system to be abandoned. Attempts were also made at expressing a set of seven putative FAM "domains" predicted from partial proteolysis experiments (P. Khut, Masters Graduate, University of Adelaide, SA) (Figure 3.11a), as GST fusion proteins in bacteria. These GST fusion proteins were to have been used as bait to identify novel interacting proteins and to map binding sites of proteins to much smaller, 200-500 amino acid stretches of FAM. Unfortunately sufficient quantities of soluble bacterially expressed fusion proteins were unable to be obtained for any of the seven putative FAM domains.

To investigate the potential protein-protein interaction capabilities of the putatative FAM UBLD, *In vitro* binding assays were used. These assays were designed to identify novel FAM-UBLD binding proteins through both an unbiased and a best candidate approach. In the

unbiased approach, an immobilised GST-FAM-UBLD fusion protein was used as bait to identify novel binding proteins for FAM from cell and tissue lysates. Although some bands that appeared to be unique to the GST-FAM-UBLD lane were identified, sequencing of these bands failed to reveal any proteins of particular interest. The best candidate approach to identifying novel FAM interactors applied a similar procedure to that performed for the unbiased approach. In contrast to the unbiased approach however, bound proteins that were resolved by SDS-PAGE were subjected to western analysis. Given the association of FAM with protein trafficking pathways, antibodies raised against a number of protein trafficking pathway proteins containing ubiquitin binding motifs/domains were used. This approach tentatively identified the three members of the GGA family, GGA1, GGA2, and GGA3, as potential FAM-UBLD interacting proteins.

Repeats of the initial *in vitro* binding assay met with limited success as did attempts at optimizing the conditions of the pulldown. Given however, that FAM has previously shown strong Golgi staining by coimmunofluorescence studies (Murray et al., 2004), has been found to partially co-localise with the GGAs (personal communication, M. Scherer, PhD candidate, University of Adelaide, SA), and that the GGA proteins are known to be ubiquitylated by the E3 ligase hVPS18 (Shiba et al., 2004; Yogosawa et al., 2006), an interaction between the GGA proteins and FAM was plausible. Attempts were also made to confirm this interaction both *in vivo* and *in vitro*. Coimmunoprecipitation experiments were performed to determine if the full length proteins interacted in an *in vivo* setting. Although these experiments proved inconclusive, *in vitro* binding experiments using the isolated domains of GGA3 indicated that GGA3 is able to bind to both the full length FAM protein and the UBLD of FAM. Furthermore, these experiments suggested that the VHS domain of GGA3 is not likely to be involved in this interaction.

RESULTS

Yeast Two-Hybrid:

The yeast two-hybrid assay system was chosen because, unlike methods using bacterial expression and *in vitro* transcription and translation, it was concluded that there was likely to be a higher chance that expressed fragments of the FAM protein would fold correctly due to the presence of eukaryotic chaperones. Assessment of an interaction with the ProQuest yeast two-hybrid system (Invitrogen) makes use of three reporter systems in a similar manner to other commercially available yeast two-hybrid kits. The first, and probably the most

sensitive, is expression of the *HIS3* reporter gene. The *HIS3* reporter gene that is stably incorporated into the MAV203 yeast strains' genome, encodes imidazole glycerol phosphate dehydratase. This enzyme is used by *S. Cerevisiae* in the biosynthesis of histidine and has been shown to be specifically inhibited by 3-Amino-1,2,4-Triazole (3AT) in a dose dependent manner (Kishore and Shah, 1988). The MAV203 yeast strain has been designed to express a low basal level of the *HIS3* gene product. The concentration of 3AT required to inhibit this basal level is first determined in cells where an interaction is not taking place. Once this basal level is determined even small increases in expression generated by a weak interaction are readily detectable as growth of the co-transformed yeast cells on plates lacking histidine but containing the minimal determined concentration of 3AT. The second reporter allows expression of the *URA3* gene product in the presence of a protein-protein interaction. The *URA3* gene product permits the synthesis of uracil when yeast are grown on media lacking uracil, and thereby allows growth in these conditions. The last reporter gene is *LacZ* and produces β-galactosidase which can act on the chromogenic substrate 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) to form a blue colour. Five positive control strains, control strains A to E, are also included in the assay system to provide indicators for very weak to very strong interaction strengths (Table 3.1).

Table 3.1: Control yeast strains used in yeast two-hybrid experiments

Yeast Strain	Resident Plasmids	cDNA Insert	Interaction Strength
Control Strain A	pPC97	no insert	**None**
	pPC86	no insert	
Control Strain B	pPC97-RB	human RB amino acids 302-928	**Weak**
	pPC86-E2F1	human E2F1 amino acids 342-437	
Control Strain C	pPC97-*CYH*2s-dDP	*Drosophila* DP amino acids1-377	**Moderately Strong**
	pPC86-dE2F	*Drosophila* E2F amino acids 225-433	
Control Strain D	pPC97-Fos	rat cFos amino acids 132-211	**Strong**
	pPC86-Jun	mouse cJun amino acids 250-325	
Control Strain E	pCL1 (encoding full length GAL4)	GAL4 amino acids 1-881	**Very Strong**
	pPC86	no insert	

β-catenin

To determine if regions of the FAM protein outside of the catalytic core (amino acids 1497-1918) were able to interact with β-catenin, four overlapping fragments of FAM (Taya et al., 1998) were cloned into the yeast two-hybrid GAL4 activation domain (AD) fusion vector pDest22 (Invitrogen). These four FAM fragments were designated N1 (amino acids 1-674), N2 (amino acids 669-1213), C1 (amino acids 1210-2100), and C2 (2097-2554). A fragment spanning the catalytic core (amino acids 1476-1918) and designated FAM-Cat was also used as a positive control (Figure 1.6 c). β-catenin is known to contain a C-terminal transactivation domain that is able to interact with the TATA-binding protein (Hecht et al., 1999; van de Wetering et al., 1997), a protein that is present in *S. Cerevisiae* (Schmidt et al., 1989). To determine if the transactivation domain of β-catenin was active in *S. Cerevisiae*, the cDNA for full length β-catenin was cloned into the yeast two-hybrid GAL4 DNA binding domain (DBD) fusion vector, pDest32 (Invitrogen). An initial screen was carried out to determine if the β-catenin DBD fusion protein was able to self-activate and to determine the optimal concentration of 3AT to use in future assays. When cells were co-transformed with the expression vector for the β-catenin DBD fusion protein, pDEST32 β-catenin, and an empty GAL4 AD expression vector, pEXP AD502, transformed cells exhibited growth on plates containing all concentrations of the *HIS3* inhibitor 3AT (Figure 3.1). This indicated that the transactivation domain of β-catenin was indeed active in *S. Cerevisiae*.

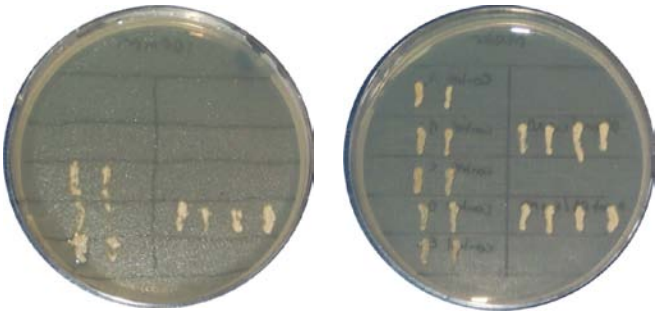
To avoid this problem, the minimal FAM binding region (FBR) of β-catenin consisting of armadillo repeats 3-10 (amino acids 184-535) (Taya et al., 1999), was cloned into pDest32 in place of the full length protein. Once again, an initial screen was undertaken to determine if the β-catenin FBR/DBD fusion protein could self-activate or form a complex with the Gal4 AD. Self activation was also tested for each of the expression vectors for the five FAM fragment/Gal4 AD fusion proteins, and an expression vector for a full length FAM/Gal4 AD fusion. Also tested for self-activation were an EGFP/Gal4 AD fusion, as a negative control, and an E-cadherin cytoplasmic domain/Gal4 AD fusion, as a positive control for interaction with the β-catenin FBR/DBD fusion protein. The empty Gal4 DBD expression vector pDBleu was co-transformed with each to test for the ability of each fusion protein to form an inappropriate complex with the Gal4 DBD in isolation. No self-activation was evident as transformants did not grow at any concentration of 3AT (Appendix A.1). A concentration of 10mM 3AT was chosen to assess interaction with the *HIS3* reporter gene.

Figure 3.1: ***Self activation and 3AT titration trial for β-catenin DBD***. The capacity of a β-catenin/Gal4 DNA binding domain fusion protein to self-activate was assessed. Yeast transformed with the expression construct (pDest32 β-catenin) were plated on media lacking histidine and containing increasing concentrations of the histidine biosynthesis inhibitor 3AT. A master plate **F)** was used to replica plate transformed and control yeast strains onto test plates with 10 mM **A)**, 25 mM **B)**, 50 mM **C)**, 75 mM **D)**, or 100 mM **E)** 3AT. Growth of pDest32 β-catenin transformed yeast was seen at all five concentrations of 3AT. **G)** Arrangement of experimental and control yeast strains on master and test plates.

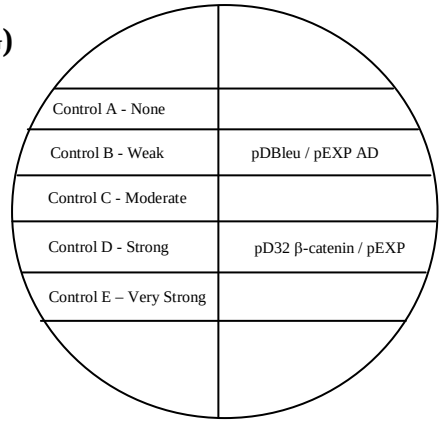
A) –His + 10mM 3AT **B)** –His + 25mM 3AT

C) –His + 50mM 3AT **D)** –His + 75mM 3AT

E) –His + 100mM 3AT **F)** Master Plate



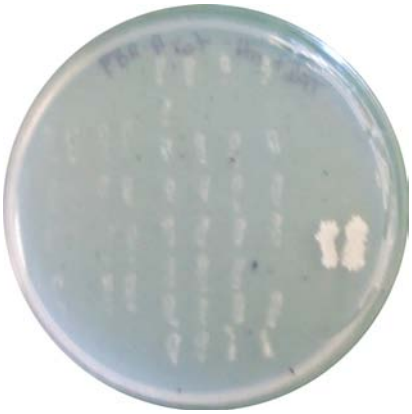
G)



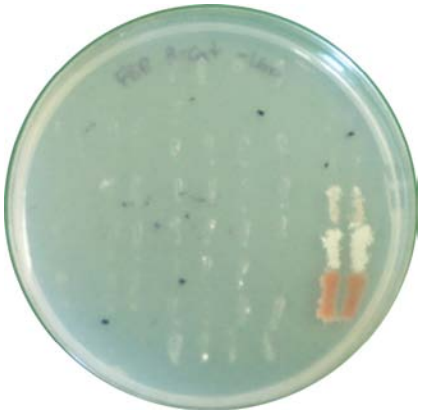
When the three reporter systems were used to assess the ability of the five FAM fragments and the full length FAM protein to interact with the β-catenin FBR, all three reporters returned a negative result (Figure 3.2). Although there is some growth seen on both the SC –Leu –Trp –His +3AT and the SC –Leu –Trp –Ura plates, a similar amount of growth is seen for control A. As control A consists of yeast transformed with base vectors containing no cDNA inserts, reconstitution of a functional transcriptional complex is not possible in this yeast strain. The observed growth of FAM or FAM fragment/β-catenin FBR transformed cells and control strain A is also far below the growth seen on the SC –Leu –Trp –Ura, for the very strong, strong, and moderate interaction strength controls E, D, and C respectively. This strongly suggests that no interaction between FAM or FAM fragments and β-catenin FBR has occurred, although a very weak interaction cannot be ruled out. Interestingly, the previously well characterized and very strong interaction in a yeast two-hybrid system (Simcha et al., 2001) between the cytoplasmic domain of E-cadherin and the armadillo repeat region of β-catenin (β-catenin FBR), also returned a negative result in all three assay systems. This result was quite puzzling because all of the co-transformed yeast colonies that were used in the assay showed quite strong growth on the SC –Leu –Trp selection plate (Figure 3.2 d) suggesting that both plasmids were present. Furthermore, the three positive controls, controls C, D, and E, that were included in the assay all returned positive results from at least two of the reporter genes. One possible explanation that was investigated was that the proteins were unable to be expressed. Eukaryotic expression of the five FAM fragments and of the full length protein had previously been successfully demonstrated in mammalian tissue culture cell lines with mammalian expression vectors for V5 tagged fusions of these proteins (L. Jolly, PhD candidate, The University of Adelaide, SA; and (Murray et al., 2004). In contrast, expression of the β-catenin FBR and E-cadherin cytoplasmic domain cDNAs had not previously been tested. An attempt was made at expressing myc/his tagged β-catenin FBR, E-cadherin cytoplasmic domain, and EGFP cDNAs in human embryonic kidney (HEK) 293T cells but proved inconclusive (data not shown). Lack of expression of the fusion proteins in yeast due to errors in the cDNA seemed unlikely however, as each of the cDNAs were DNA sequenced in their respective donor entry vectors subsequent to their sub-cloning by recombination reactions.

Figure 3.2: ***A Yeast two-hybrid assay between a β-catenin FBR-DBD fusion and FAM fragment-AD fusions failed to identify any interactions.*** A yeast two hybrid assay was carried out to investigate and map the site of β-catenin binding on FAM. MAV203 yeast cells were co-transformed with expression plasmids for a β-catenin FBR-DBD fusion protein (pDest32 β-catenin FBR) and one of five expression plasmids for FAM fragment-AD fusion proteins (pDest22 N1, N2, C1, C2, or FAM Cat [FC]). Controls consisting of MAV203 cells co-transformed with pDest32 β-catenin FBR and pDest22 EGFP (negative control), pDest22 E-cadherin cytoplasmic domain (ECT- positive control), or pDest22 FAM (FLF – positive control) were also included. Cells were grown on a master plate **D)** before being replica plated onto test plates consisting of solid synthetic complete media (SSCM) –His +10 mM 3AT **A)**, SSCM –Ura **B)**, or complete media overlaid with a nitrocellulose filter for a β-galactosidase assay (X-gal) **C)**. All three test plates returned negative experimental results. **E)** Diagram showing the arrangement of experimental and control yeast strains on master and test plates. *FBR = FAM binding region*

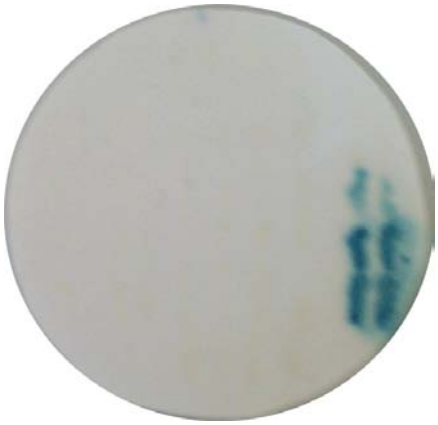
A) –His +10 mM 3AT



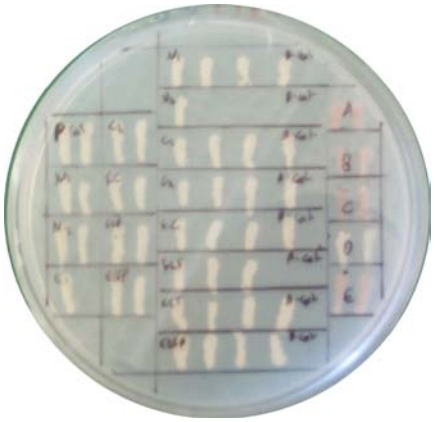
B) –Uracil



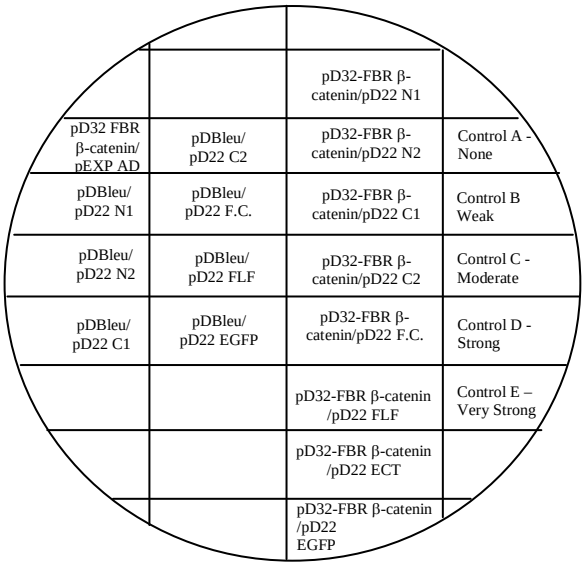
C) X-gal



D) Master Plate



E)



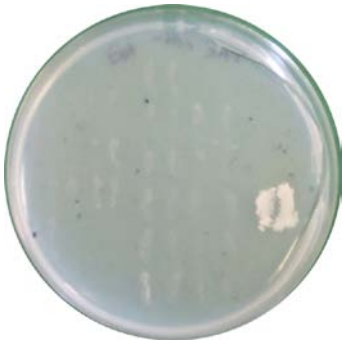
E-cadherin Cytoplasmic Domain

While this work was being undertaken, evidence was emerging that FAM had a role to play in the trafficking of the immature E-cadherin/β-catenin complex (Murray et al., 2004). It was also known that mature junctional E-cadherin/β-catenin complexes could be ubiquitylated by the E3 ligase hakai, with this ubiquitylation occurring on both β-catenin and E-cadherin. Furthermore, this ubiquitylation was observed to promote endocytosis of the complex ultimately resulting in its destruction at the lysosome (Fujita et al., 2002). It was postulated that one of the roles of the FAM deubiquitylating enzyme was to oppose the hakai induced ubiquitylation of the complex, thereby stabilising the junctional complex and cell-cell adhesion (Fujita et al., 2002). Whilst it was known that FAM could interact with, and lead to a stabilization of β-catenin through its enzymatic activity, it wasn't established whether any part of FAM was also able to interact with E-cadherin (Taya et al., 1999).

For this reason, the yeast two-hybrid system was used to investigate the possibility of a direct interaction between FAM and the cytoplasmic domain of E-cadherin (designated ECT). An initial titration trial was conducted to identify the minimal concentration of 3AT required to inhibit basal levels of *HIS3* expression in transformed cells and to test for self-activation of the ECT/DBD fusion (Appendix A.2). Self-activation was not observed and the yeast two-hybrid assay was conducted using all three reporter assays with *HIS3* reporter activity tested on a 25mM 3AT containing plate. As was seen for β-catenin, all three reporter systems returned a negative result (Figure 3.3). Once again, the moderate, strong, and very strong positive controls (controls C, D, and E respectively) returned positive results from the X-Gal and Ura^- assays, with control D returning a positive result from all three assays.

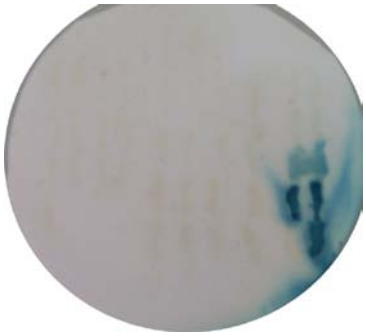
Figure 3.3: ***A Yeast two-hybrid assay between an E-cadherin cytoplasmic domain-DBD fusion and full length FAM or FAM fragment-AD fusions failed to identify any interactions***. A yeast two hybrid assay was carried out to investigate the potential for an interaction between FAM and the E-cadherin cytoplasmic domain (ECT) and to map the site of ECT binding on FAM. MAV203 yeast cells were co-transformed with expression plasmids for an ECT-DBD fusion protein (pDest32 ECT) and one of five expression plasmids for FAM fragment-AD fusion proteins (pDest22 N1, N2, C1, C2, or FAM Cat [FC]) or an expression plasmid for a full-length FAM-AD fusion protein. Cells were grown on a master plate **D)** before being replica plated onto test plates consisting of solid synthetic complete media (SSCM) –His +25 mM 3AT **A)**, SSCM –Ura **B)**, or complete media overlaid with a nitrocellulose filter for a β-galactosidase assay (X-gal) **C)**. All three test plates returned negative experimental results. **E)** Diagram showing the arrangement of experimental and control yeast strains on master and test plates.

A) –His +25 mM 3AT

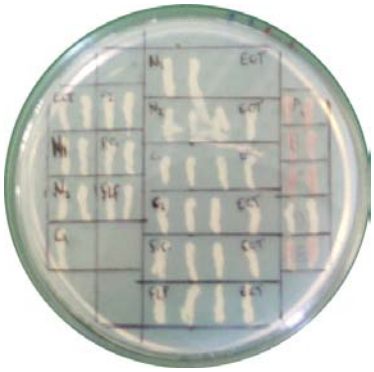


B) –Uracil

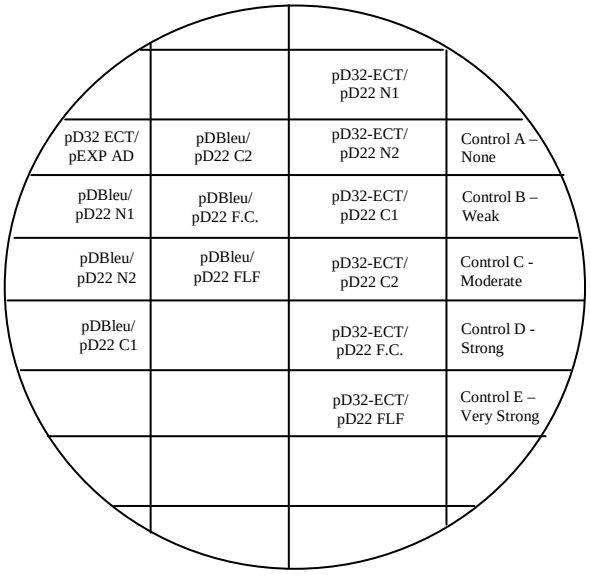
C) X-gal



D) Master Plate



E)



Epsin 1 & Itch

An interaction between epsin 1 and FAM was initially investigated by using an epsin 1/Gal4 DBD fusion as bait with the five FAM fragment/Gal4 AD fusion vectors, and the full length FAM/AD fusion vectors acting as prey. EGFP-Gal4 AD was again used as a negative control. Acting on past experience, the assay was undertaken with all three reporter systems using a concentration of 25mM 3AT for inhibition of basal levels of *HIS3* expression. Surprisingly, positive results were returned in all three reporter systems for every yeast clone that was transformed with the expression plasmid for the epsin1/Gal4 DBD, including the negative controls in which this vector was co-transformed with the EGFP/AD, and the empty pEXP AD502 expression vectors (Figure 3.4). This self-activation of epsin 1 may be explained by the ability of the Eps15 interactor NH_2-terminal homology (ENTH) domain of epsin 1 to interact with the transcription factor promyelocytic leukemia Zn^{2+} finger protein (PLZF) (Hyman et al., 2000).

To overcome the self-activation of epsin 1 in the yeast two-hybrid system, the cDNAs for epsin 1, full length FAM, and the five FAM fragments were cloned into the reverse expression plasmids. That is, an epsin 1/AD fusion protein was used as bait while FAM and the FAM fragments were expressed as Gal4 DBD fusions to use as prey. Once again a titration trial was performed to determine if epsin 1 or the FAM/FAM fragments were able to self-activate in this orientation. Indeed expressing epsin 1 as an AD fusion did appear to abolish the self-activation as no growth was observed on plates over 50mM 3AT (Figure 3.5 a & c). The interaction of FAM/FAM fragments with epsin 1 was subsequently analysed through the use of the *LacZ* reporter gene X-Gal assay. In this assay, no staining was observed for the FAM protein or individual fragments (Figure 3.5 a). An independent positive control consisting of a bait and a prey plasmid that were known to interact in a yeast two-hybrid system (a kind donation from C. Bracken, PhD graduate, The University of Adelaide, SA) did however produce a positive result (Figure 3.5 a).

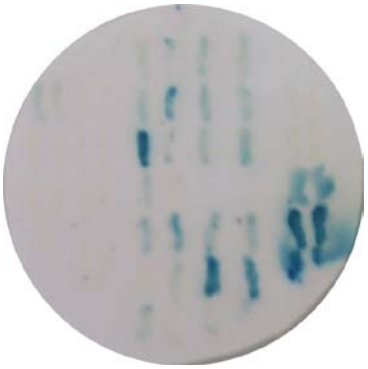
Yeast two-hybrid experiments were also undertaken to determine if any part of FAM could be shown to interact with the E3 ligase Itch. A 3AT titration trial was undertaken (data not shown) and the optimal concentration of 3AT determined to be 25mM with no self-activation observed for the Ich/DBD fusion. Experimental results were negative for all three reporter systems (Figure 3.5 b & d and data not shown). In an effort to understand why the β-catenin, E-cadherin, epsin 1, and itch yeast two-hybrid studies had all returned negative results, even when interactions had been shown previously (as in the case of β-catenin and

Figure 3.4: ***A Yeast two-hybrid assay between an Epsin1 DBD fusion and full length FAM or FAM fragment-AD fusions failed to identify any interactions.*** A yeast two hybrid assay was carried out to investigate and map the site of Epsin1 binding on FAM. MAV203 yeast cells were co-transformed with expression plasmids for an Epsin1-DBD fusion protein (pDest32 Epsin1) and one of five expression plasmids for FAM fragment-AD fusion proteins (pDest22 N1, N2, C1, C2, or FAM Cat [FC]). Controls consisting of MAV203 cells co-transformed with pDest32 Epsin1 and pDest22 EGFP (negative control) or pDest22 FAM (FLF – positive control) were also included. Cells were grown on a master plate **D)** before being replica plated onto test plates consisting of solid synthetic complete media (SSCM) –His +25 mM 3AT **A)**, SSCM –Ura **B)**, or complete media overlaid with a nitrocellulose filter for a β-galactosidase assay (X-gal) **C)**. All three test plates returned positive experimental results for any yeast cells transformed with pDest32 Epsin1, including negative controls. **E)** Diagram showing the arrangement of experimental and control yeast strains on master and test plates.

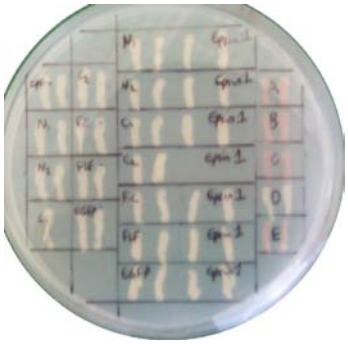
A) –His +25 mM 3AT

B) –Uracil

C) X-gal



D) Master Plate



E)

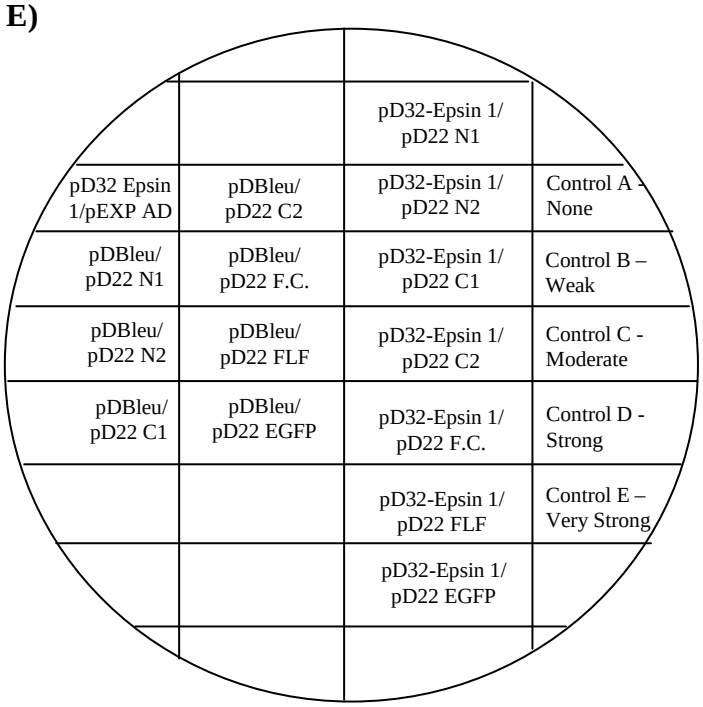
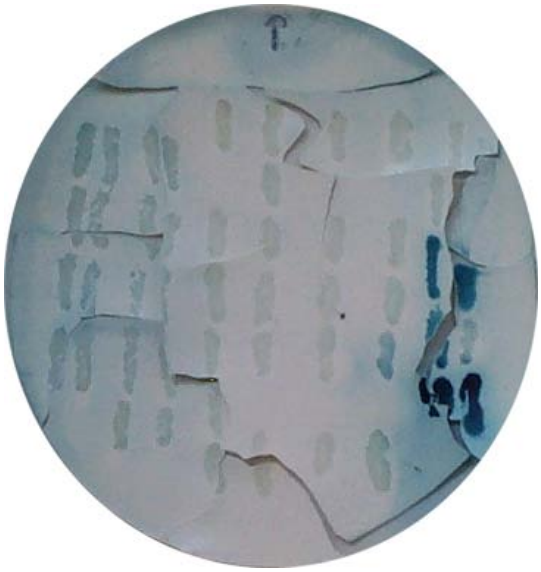
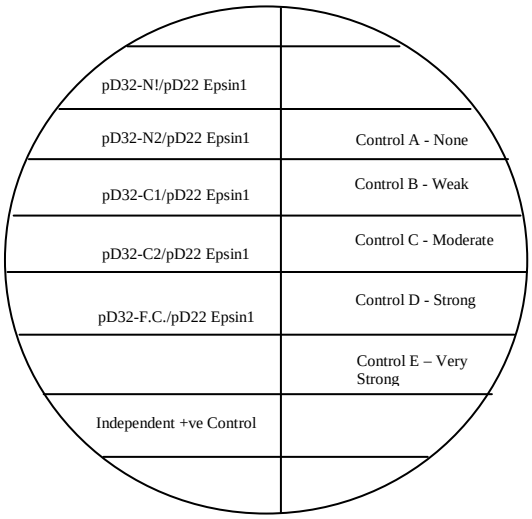


Figure 3.5: ***A reverse epsin1 yeast two-hybrid and a yeast two-hybrid between an Itch-DBD fusion and full length FAM or FAM fragment-AD fusions failed to identify any interactions***. Yeast two-hybrid assays were carried out to investigate and map the site of Epsin and Itch binding on FAM. MAV203 yeast cells were co-transformed with either expression plasmids for an epsin-AD fusion protein (pDest22 Epsin1) and one of five expression plasmids for FAM fragment-DBD fusion proteins (pDest32 N1, N2, C1, C2, or FAM Cat [FC]) or expression plasmids for an Itch-DBD fusion protein (pDest32 Itch) and one of five expression plasmids for FAM fragment-AD fusion proteins (pDest22 N1, N2, C1, C2, or FAM Cat [FC]) or an expression plasmid for a full-length FAM-AD fusion protein. Controls consisting of MAV203 cells co-tranformed with pDest32 Itch and pDest22 EGFP (negative control) or pDest22 FAM (FLF – positive control) were also included. **A)** A β-galactosidase assay was performed to test for interaction between epsin1 and the five FAM fragments. No interaction was seen however an independent positive control displayed a positive result. **B)** A β-galactosidase assay was performed to test for interaction between Itch and the five FAM fragments or full length FAM but no interaction was seen. **C)** Diagram showing the arrangement of experimental and control yeast strains on the test plate **D)** Diagram showing the arrangement of experimental and control yeast strains on the test plate.

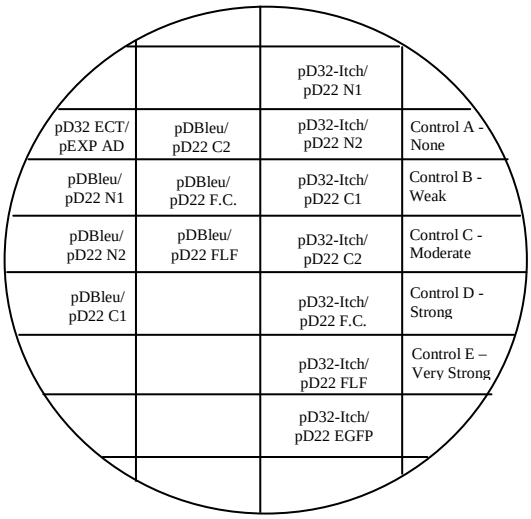
A) Epsin Reverse Two-Hybrid: β-Gal Assay **B)** Itch Two-Hybrid: β-Gal Assay



C)



D)



E-cadherin, an attempt was made at western analysis of yeast cell extracts to determine if the fusion proteins were being expressed. Two different lysis protocols were attempted, as described in chapter 2. Antibodies raised against the Gal4 activation domain, DNA binding domain, and β-catenin were used to detect expression of the fusion proteins in the yeast cells however all three antibodies failed to identify immunoreactive bands despite the fact that protein was extracted as demonstrated by coomassie blue staining (data not shown and Appendix B).

The detection of an interaction between the independent positive control bait and prey fusion proteins indicated that vectors were being efficiently introduced into the yeast cells. In addition, the viability of the pDEST32 vector was indicated by the ability of wild-type β-catenin and epsin 1 proteins expressed from this vector to self-activate. Taken together with the knowledge that the independent control bait and prey proteins were expressed from a different vector system, it was speculated that there may have been a problem with expression of fusion proteins from the pDEST22 expression vector. To test for this possibility, yeast cells transformed with pDEST22 EGFP were visualized under a fluorescent microscope. GFP fluorescence was observed indicating that expression from the pDEST22 vector was possible (data not shown).

Mammalian Two-Hybrid

Mammalian two-hybrid analysis of protein-protein interaction uses much the same principles as a yeast two-hybrid. In this system however, there is only one reporter system. This consists of a vector, pG51EBLuc, expressing firefly luciferase from a synthetic promoter containing Gal4 responsive elements. There are several advantages of the mammalian two-hybrid assay over the yeast two-hybrid system. The system is more representative of the environment in which the protein-protein interactions to be studied, FAM with β-catenin, epsin 1, and itch, occur. The system is also less time consuming and has the added benefit that western analysis of expressed proteins is much simplified. To use the mammalian two-hybrid system for the analysis of the interaction of FAM with its substrates/binding partners, the expression vectors for the GAL4-DBD and VP16-AD fusion proteins first needed to be converted into Gateway compatible vectors. The Gateway conversion of these vectors was carried out as described in chapter 2 to yield the Gateway compatible mammalian two-hybrid expression vectors pEFBOS-GAL4 DBD GW, and pEFBOS-VP16 AD GW. These two expression vectors allow high level mammalian expression of either GAL4 DBD, or VP16 AD N-terminally tagged fusion proteins. Expression is driven by an elongation factor 1α (EF)

promoter. β-catenin, β-catenin FBR, and AF-6 FBR were used as bait for the mammalian two-hybrid assay, and were sub-cloned into pEFBOS-GAL4 DBD GW.

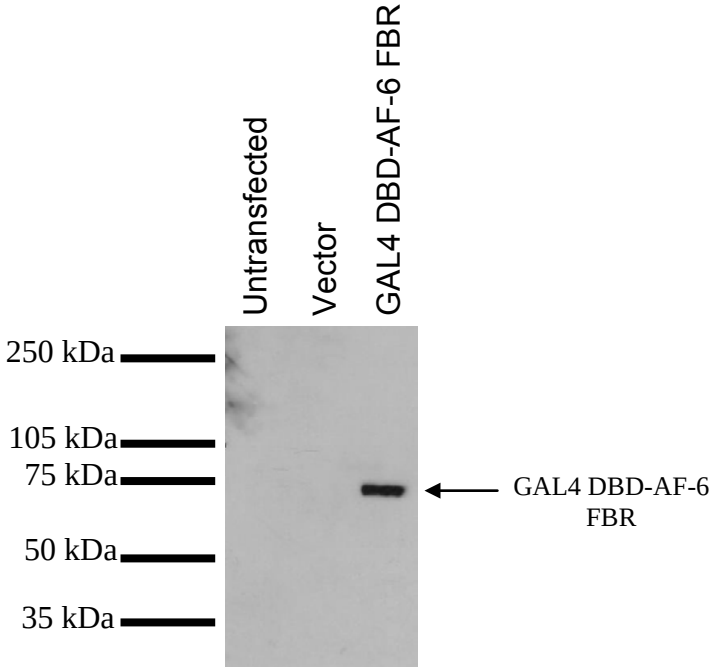
Initial experiments to determine the efficacy of the mammalin two-hybrid system utilized the known interaction of the catalytic region of FAM with each of these bait proteins. Due to difficulties with expression of the FAM catalytic (FAM Cat) region bounded by amino acids 1476-1918, a second cDNA encoding a fragment of the FAM protein that also incorporates the core cys and his box catalytic regions of the protein (amino acids 1553-1954), and previously used to characterize the enzymatic activity of FAM on epsin 1 (Chen et al., 2003), was sub-cloned into the prey expression plasmid pEFBOS-VP16 AD GW. A transient transfection into HEK 293T cells was performed to determine if expression was possible from the pEFBOS GAL DBD GW expression vector (no antibody was available for use against the VP16 AD fusion). Protein expression was indeed observed for the AF-6 FBR/GAL4 DBD fusion with detection by an anti-Gal4-DBD antibody of a band at the predicted molecular weight of ~80 kDa (Figure 3.6 a). This confirmed the viability of at least the pEFBOS GAL4 DBD GW expression vector. Despite this, control dual luciferase assays, using the Gateway converted mammalian two-hybrid system, that were designed to confirm the detection of known positive interactions between the β-catenin or AF-6 FAM binding regions and the catalytic core of FAM, proved inconclusive. In at least two assays, it appeared that luciferase activity was detected in cells co-transfected with the GAL4-DBD-AF-6 FAM binding region and VP16-AD-FAM Cat (1553-1954) expression vectors (Figure 3.6 b & d). However, the absolute values obtained for luciferase luminosity were found to be very low (Figure 3.6 c & e). At these low levels of luminosity even small differences between the transfection control Renilla luciferase reporter and the experimental Firefly luciferase reporter would become exaggerated and provide misleading information about the relative luciferase activity. For this reason, any readings taken with this system would be unreliable. No induction of luciferase activity was observed from cells co-transfected with the GAL4-DBD β-catenin FAM binding region and VP16-AD-FAM Cat (1553-1954) expression vectors (Figure 3.6 b & d).

GST Affinity Binding Assay

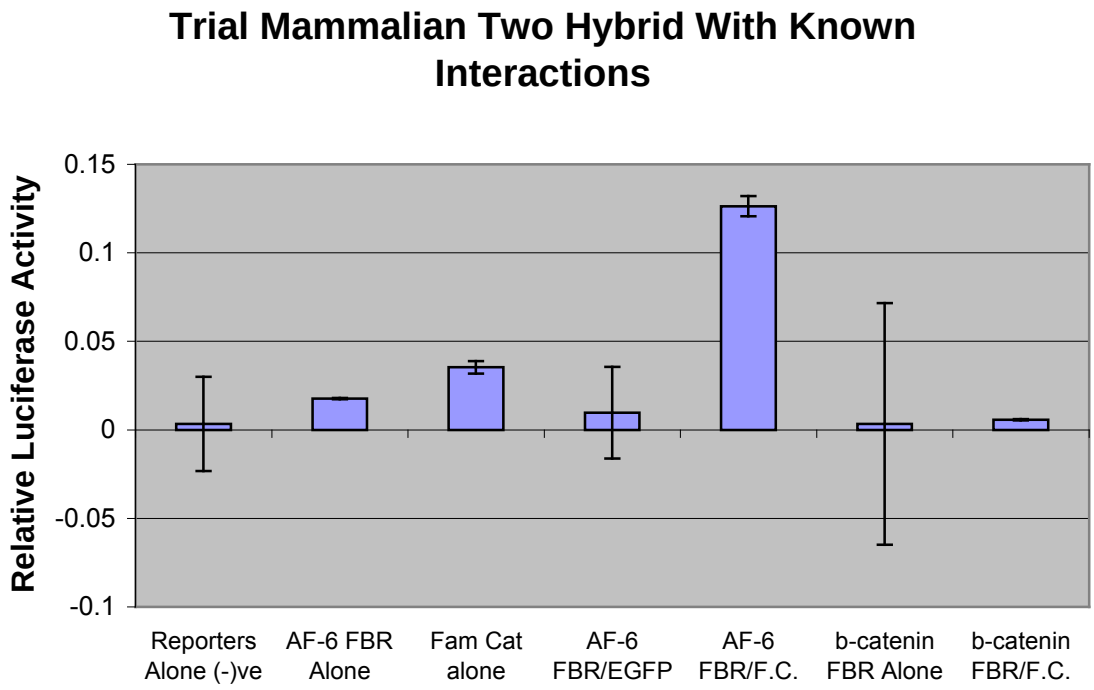
Probably one of the most common methods used to identify and characterize protein-protein interactions utilizes the strong interaction between the eukaryotic protein, glutathione-S transferase (GST) and its substrate, glutathione. Bacterially expressed fusion proteins of GST with a protein of interest can be used as bait to identify novel interactors, or to

Figure 3.6: ***A mammalian two-hybrid assay system could not be successfully applied to the identification of an interaction between Af-6 or β-catenin, and FAM***. **A)** HEK 293T cells were transiently transfected with either the pEFBOS GAL4 DBD GW base vector or with pEFBOS GAL4 DBD GW AF-6 FBR, or were mock transfected. Cells were incubated for 72 hours before being harvested, lysed, and subjected to western analysis with a monoclonal anti-Gal4 DBD antibody. Expression of the AF-6 FBR-Gal4 DBD fusion protein was confirmed by the presence of an anti-Gal4 DBD immunoreactive band at the predicted molecular weight of 70 kDa. **B)** Dual luciferase assays were performed to determine if any interaction between the known interacting protein fragments spanning either of the FAM binding regions (FBR) of AF-6, or β-catenin and the catalytic region of FAM, could be detected using the mammalian two hybrid system. Although some induction of the luciferase reporter system was detected in cells expressing AF-6-FBR and FAM Cat **B)** & **D)**, the average intensity of the luminosity recorded was found to be quite low, exaggerating the very small differences in readings from the transfection control luciferase reporter (renilla luciferase) and the experimental luciferase reporter (firefly luciferase), as can be seen when observing the raw data **C)** & **E)**. No induction of the luciferase reporter system was detected in cells expressing β-catenin-FBR and FAM Cat **B)**. (Plotted values are the mean ± SEM, n=3).

A)


Untransfected
Vector
GAL4 DBD-AF-6 FBR
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
GAL4 DBD-AF-6 FBR

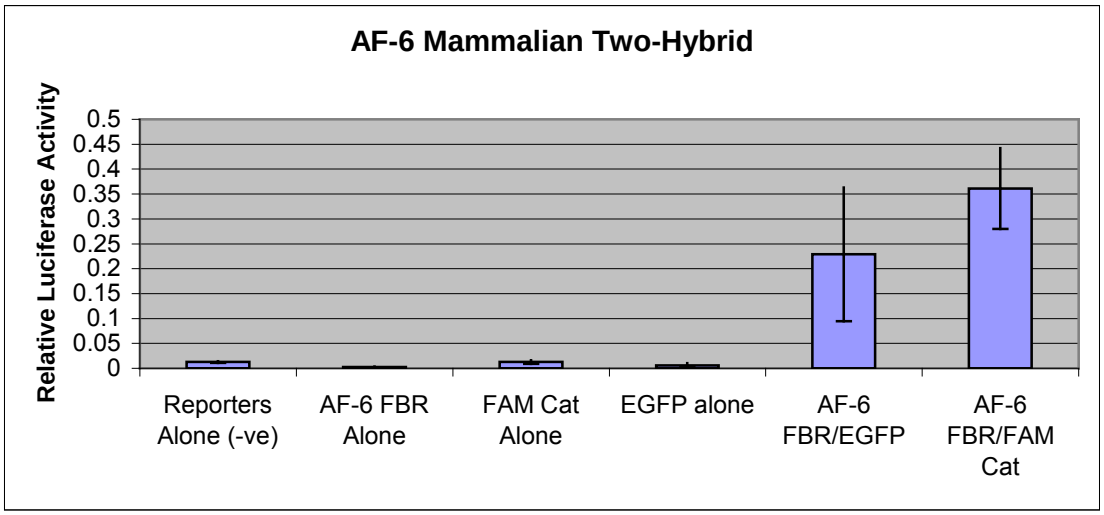
B)



C)

Sample	Firefly	Renilla	Ratio
Reporters Alone (-ve)	0.339	99.75	0.003
	0.313	86.42	0.004
	0.129	44.56	0.003
AF-6 FBR Alone	0.047	4.005	0.012
	0.033	1.951	0.017
	0.047	1.925	0.024
FAM Cat Alone	0.3	3.441	0.087
	0.046	8.031	0.006
	0.055	4.125	0.013
AF-6 FBR/EGFP	0.049	11.88	0.004
	0.034	8.484	0.004
	0.056	2.635	0.021
AF-6 FBR/F.C.	0.053	1.664	0.032
	0.294	1.135	0.259
	0.097	1.102	0.088
β-catenin	0.052	18.03	0.003
	0.063	15.13	0.004
	0.049	14.55	0.003
β-catenin FBR/F.C.	0.047	3.848	0.012
	0.042	25.88	0.002
	0.044	13.39	0.003

D)



E)

Sample	Firefly	Renilla	Ratio
Reporters Alone (-ve)	0.339	99.75	0.003
	0.313	86.42	0.004
	0.129	44.56	0.003
AF-6 FBR Alone	0.047	4.005	0.012
	0.033	1.951	0.017
	0.047	1.925	0.024
F. C. Alone	0.3	3.441	0.087
	0.046	8.031	0.006
	0.055	4.125	0.013
EGFP Alone	0.049	11.88	0.004
	0.034	8.484	0.004
	0.056	2.635	0.021
AF-6 FBR/EGFP	0.053	1.664	0.032
	0.294	1.135	0.259
	0.097	1.102	0.088
AF-6 FBR/F.C.	0.052	18.03	0.003
	0.063	15.13	0.004
	0.049	14.55	0.003

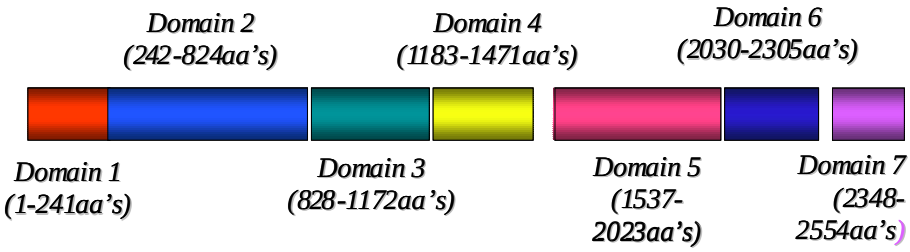
investigate known interactions, by immobilizing the fusion protein on agarose beads covalently linked to glutathione. Previous efforts to perform GST affinity chromatography experiments using the five fragments of FAM described by Taya *et al* (Taya et al., 1998), proved unsuccessful due to difficulties in obtaining sufficient quantities of soluble bacterially expressed fusion proteins (data not shown). This may have been due to the fact that the overlapping fragments that were used in the experiments described by Taya *et al* were fragments of the FAM protein whose boundaries had been chosen at random without any knowledge of domain boundaries or structural properties of the FAM protein. Such a design strategy may have produced proteins in which stabilising secondary and tertiary structures formed in the wild-type protein, were unable to be generated in the fragments. This may have led to an abundance of improperly folded proteins with exposed hydrophobic surfaces, producing aggregation and sedimentation of the fusion proteins when expressed in bacteria.

To try and combat this possibility, and to learn more about the structural organization of the FAM protein, a partial proteolysis screen was carried out by a colleague (P. Khut, Masters Graduate, The University of Adelaide, SA). In brief, the full length FAM protein was expressed in a baculovirus system as a GST-fusion, and column purified using GST affinity chromatography. The GST tag was cleaved to generate purified FAM protein. The FAM protein was then subjected to limited proteolysis by a suite of cysteine and serine proteases and the cleavage products of each protease were resolved by polyacrylamide gel electrophoresis. Major bands from each protealytic cleavage event were identified and gel slices prepared for analysis by mass spectrometry. From this analysis, a series of 7 sequential putative FAM domain sequences were proposed (Figure 3.7a) (P. Khut, Masters Graduate, The University of Adelaide). cDNAs encoding six of the seven putative FAM domains (difficulties were encountered in cloning FAM domain three) were cloned and introduced into a GST bacterial expression vector, pGEX 4T2 which had been converted into a Gateway destination vector. The purpose of this was to attempt to use these putative FAM domains to investigate the interactions of FAM with its binding partners through GST affinity chromatography.

Initial small scale inductions were carried out on each of the six FAM domain GST-fusion proteins to determine the extent of the solubility of the fusion proteins. Induction of pGEX 4T2 GW FAM domain (FD) -1, -4, and -6 expression vectors in BL21 cells at 37°C was expected to yield fusion proteins of 53, 60, and 58 kDa respectively. For these three expression vectors, no soluble or insoluble expressed protein was observed, nor was induction of protein expression observed (Figure 3.7 b, c, f, h, & i). Often proteins that are expressed in

Figure 3.7: ***Bacterial Inductions of GST fusions of predicted FAM domains 1, 2, 4, 5, 6, and 7 by the standard protocol either failed to produce any protein or failed to produce sufficient quantities of soluble protei****n.* **A)** Box diagram showing the amino acids spanned by each of the seven predicted FAM domains and the predicted molecular weights of their GST-fusion proteins. Protein inductions of GST fusion proteins were carried out by the standard protocol described in chapter 2 at 37°C for FD1 **B)**, FD2 **D)**, FD5 **G)**, FD6 **H)**, and FD7 **J)** or at 25 °C for FD1 **C)**, FD2 **E)**, FD4 **F)**, FD6 **I)**, and FD7 **K)**.

A)



Predicted Molecular Weight of Fusion Proteins

GST – FD1 – 53kDa

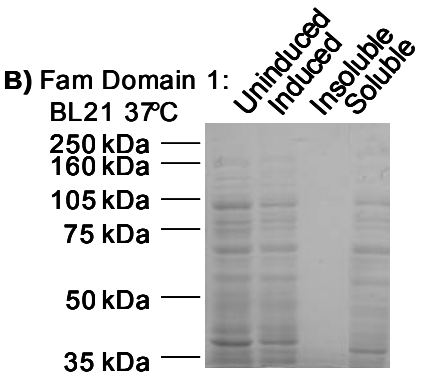
GST – FD2 – 95 kDa

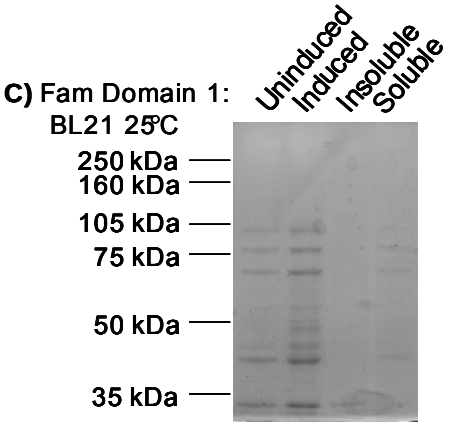
GST – FD4 – 60 kDa

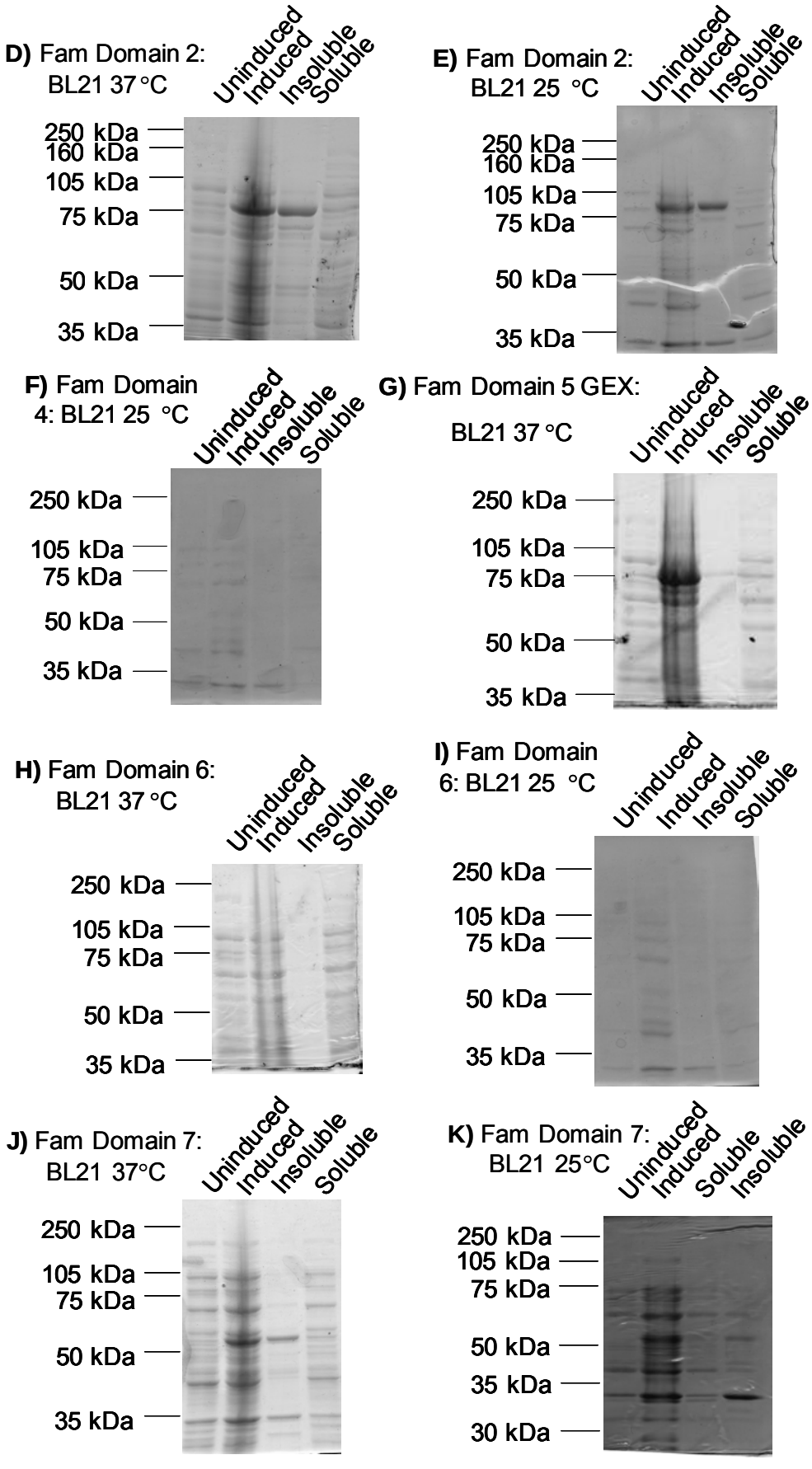
GST – FD5 – 83 kDa

GST – FD6 – 58 kDa

GST – FD7 – 50 kDa






D) Fam Domain 2: BL21 37 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
35 kDa
E) Fam Domain 2: BL21 25 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
35 kDa
F) Fam Domain 4: BL21 25 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
G) Fam Domain 5 GEX: BL21 37 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
H) Fam Domain 6: BL21 37 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
I) Fam Domain 6: BL21 25 °C
Uninduced
Induced
Insoluble
Soluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
J) Fam Domain 7: BL21 37°C
Uninduced
Induced
Insoluble
Soluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
K) Fam Domain 7: BL21 25°C
Uninduced
Induced
Soluble
Insoluble
250 kDa
105 kDa
75 kDa
50 kDa
35 kDa
30 kDa

bacteria from cDNAs containing a number of codons that are rarely used in *E. Coli*, suffer from poor induction due to translation stalling, premature translation termination, and translation frameshifting (Goldman et al., 1995; Kane, 1995; Kurland and Gallant, 1996). Researchers have found that introducing expression vectors for tRNAs that code for these rarely used codons can lead to increased yields and indeed protein induction where none was seen previously (Brinkmann et al., 1989; Rosenberg et al., 1993; Seidel et al., 1992). With this in mind, a sequence analysis was performed on the FAM coding region to identify if the gene uses any codons that are rare in *E. Coli*. It was found that there are rare codons used in all seven of the putative FAM domains with larger numbers occurring in domains four and five (4 of 515, and 8 of 717 amino acids respectively). For this reason, protein induction of domains one, four, and six was attempted in the rosetta DE3 strain of *E. Coli* (Novagen (Novey et al., 2001)). Rosetta DE3 is a strain of *E. Coli* that has been transformed with the pRARE vector that allows expression of tRNAs for seven of the rare codons that are rate limiting for heterologous protein expression in *E. Coli*. Inductions performed at either 25°C or 37°C failed once again to yield any expressed protein for all three FAM domains despite screening a number of colonies (Appendix C and data not shown [FD1])

Protein expression was observed for FAM domains two, five, and seven (Figure 3.7 d & e [FD2], g [FD5], and j & k [FD7]). In each case however, induction at either 25°C or 37°C appears to have resulted in the expressed protein being predominantly in the insoluble fraction. As discussed above, the codons used for eight of the amino acids in the cDNA sequence for FAM domain five are rarely used *E. Coli* codons. For this reason, all subsequent induction experiments performed on pGEX 4T2 GW FD5 were performed in Rosetta DE3 cells. A number of modifications to the standard induction protocol suggested by the manufacturer of the GST purification system were trialed in an attempt to improve the solubility of expressed FAM domain 2, 5, and 7 proteins (Amersham-Biosciences, 2002). Protein induction modifications included inducing protein expression with a lower concentration of IPTG (0.05 mM rather than 0.1 mM), addition of glucose to the media to prevent basal levels of protein expression prior to induction, and growing the cells to a higher optical density prior to induction, with a subsequent reduction in the induction period. Unfortunately, none of the above modifications to the expression protocol appeared to have any effect on the solubility of the expressed protein at either 25°C or 37°C (Appendix C).

Computational analysis of the 3D fold structure of the human Y-chromosome homologue of FAM, USP9Y, predicts that a large majority of the protein is unstructured (Ginalski et al., 2004). Indeed, a growing list of eukaryotic proteins are now known to have

unstructured domains including epsin 1 and AP180 (Kalthoff et al., 2002). Drawing on the knowledge that these unstructured domains tend to resist denaturation when heated, Kalthoff *et al* developed a technique to extract and partially purify these domains following expression in bacteria (Kalthoff, 2003). It was found that lysing the cells using high temperatures caused most of the ordered soluble bacterial proteins to denature and precipitate, while the unstructured epsin 1 and AP180 domains remained in solution. Furthermore, the bacterial proteases that would normally quickly degrade these easily accessible unstructured proteins when released into solution are also denatured in the process. In this way, the researchers were able to purify greatly improved quantities of expressed protein when compared to conventional lysis techniques such as sonication (Kalthoff, 2003). Thus, this technique provides a fast and simple way to extract and partially purify greater quantities of soluble expressed unstructured eukaryotic proteins from bacteria.

Since FAM shares 90% sequence homology with hUSP9Y (Appendix D), it was reasoned that it too is likely to have large stretches of unstructured protein. To investigate this possibility, protein inductions were performed on FAM domain 5 at 25°C and 37°C (Appendix C n & o respectively) and cells were lysed either by heat lysis or sonication lysis. Further protein inductions were performed at 37°C on FAM domains 2 (Appendix C e & f) for heat lysis and sonication respectively), and 7 also (Appendix C s & t for heat lysis and sonication respectively). Once again, no significant effect was seen on the solubility of the expressed protein.

Characterisation of the FAM UBLD and Bacterial Expression

Before beginning a functional analysis of the putative FAM ubl, a sequence alignment between the predicted Ubl domain in human USP9Y (amino acids 884-971) and the corresponding region in FAM (amino acids 886-970) was performed to determine if the FAM protein sequence contains the equivalent region. These sequences were submitted to the ClustalW multiple sequence alignment program (Lopez et al., 1997). There was a high degree of sequence identity (83%) and similarity (92%) between the equivalent regions (Figure 3.8 a). Next, the PHYRE (Protein Homology/analogy Recognition Engine) protein fold recognition server was used to computationally analyse and predict the 3D fold structure of the putative FAM Ubl (Kelley et al., 2005). The ten top scoring results showed matches to known structures for Ubl domain proteins, the highest match (75% estimated precision) being with a predicted protein LOC70980, the product of the 4931431F19Rik gene (unpublished data, http://www.pdb.org/pdb/explore.do?structureId=1WX8) and a predicted structure shown

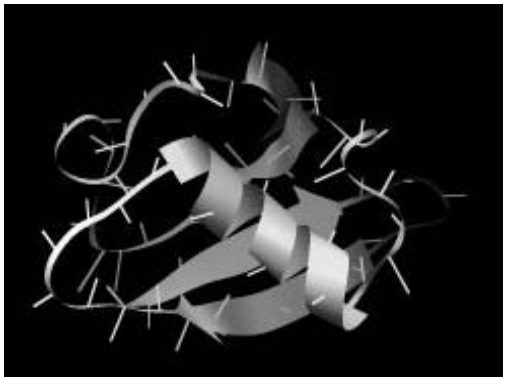
Figure 3.8: *Sequence analysis of the putative FAM UBLD*. A) Sequence alignment of the predicted Ubl domain region of human USP9Y with the equivalent region of FAM. Alignment using CLUSTALW (http://www.ebi.ac.uk/clustalw/), predicts 83% identity and 92% similarity between the sequences. **B)** Ribbon structure depicting the 3D fold structure of the Ubl region of FAM as predicted by the PHYRE threading server (Kelley et al., 2005). The Ubl region of FAM is predicted to take on a β-grasp, or ubiquitin-like fold structure. **C)** Sequence alignment of the putative FAM UBLD with ubiquitin. The two sequences share 20% homology and 45% similarity.

A)

```
hUSP9Y-UBLD          LSLIVRFPNQGRQVDELDIWFHTNDTIGSVRRCIVNRIKANVAHKKIELFVGGELIDSEN 60
FAM-UBLD             LSFIVRFPNQGRQVDDLEVWSHTNDTIGSVRRCILNRIKANVAHTKIELFVGGELIDPGD 60
                    **:************:*::* ************:*********.************. :

hUSP9Y-UBLD          DRKLIGQLNLKDKSLITAKLTQINFNMPSS 90
FAM-UBLD             DRKLIGQLNLKDKSLITAKLTQI-----SS 85
                    ********************** : : **
```

B)



C)

```
FAM-UBLD             LSFIVRFPNQGRQVDDLEVWSHTNDTIGSVRRCILNRIKANVAHTKIELFVGGELIDPGD 60
Ubiquitin            MQIFVKTLT-GKTIT-LEV--EPSDTIENVKAKIQD--KEGIPPDQQRLIFAGKQLEDG- 53
                    :.::*:  . *: :  ***  ...*** .*:  * :  * .:.  : .*:..*: :: *

FAM-UBLD             DRKLIGQLNLKDKSLITAKLTQISS 85
Ubiquitin            --RTLSDYNIQKESTLHLVLRLRGG 76
                      : :.: *::.:* :   *   ..
```

in Figure 3.8 b. The next highest match (65%) was with the UBL domain of parkin (Sakata et al., 2003). A sequence alignment was also performed between the putative FAM-UBLD and human ubiquitin. Linear amino acid sequence identity was quite low (20%) with moderate similarity (45%) (Figure 3.8 c). Many of the UBL proteins and UBL domains have been found to have quite low sequence conservation with ubiquitin however they show a quite striking similarity in their 3D fold structure (Hartmann-Petersen and Gordon, 2004; Pickart and Eddins, 2004).

Identifying Interactors

The computational approach described above predicts that FAM may indeed have a ubiquitin-like domain. To test the functional properties of this putative FAM UBLD in promoting protein-protein interactions, *in vitro* binding assays were used to identify novel interacting proteins through an unbiased and a best candidate approach. In order to achieve this goal, the cDNA for FAM-UBLD (nucleotides 2658-2910) was cloned into pDEST15, a Gateway compatible bacterial vector driving expression of GST-tagged fusion proteins from an inducible T7 promoter (Invitrogen). Ubiquitin is known to be a stable protein and so it was predicted that yields of soluble bacterially expressed FAM-UBLD would be quite high. Indeed a large proportion of GST-FAM-UBLD protein was observed in the soluble fraction (~50%) in small scale test inductions (data not shown) and large scale inductions (Figure 3.9 a). Large amounts of the protein (~5mg) were able to be purified which allowed subsequent affinity binding experiments to be performed (Figure 3.9 b).

Unbiased Approach

GST affinity binding assays were used to identify novel FAM-UBLD interacting proteins from three different sources, HEK 293T cells, mouse ES cells (mESC)s, and whole mouse brain extracts. An initial test assay from HEK 293T cell lysate was performed using five binding buffers ranging from low salt, low stringency buffers (10mM Tris pH 6.8, 0.5% Triton X-100, 50 or 100mM NaCl), to slightly higher stringency, mid range salt buffer (10mM Tris pH 6.8, 0.5% Triton X-100, 150mM NaCl) to high salt, moderate stringency buffers (10mM Tris pH 6.8, 0.5% Triton X-100, 200 or 300mM NaCl). This trial was performed to determine the optimal conditions under which the greatest number of novel bands could be seen when comparing the GST-FAM-UBLD bound eluates to the GST alone eluates (Figure 3.10). From this initial trial, it was determined that the greatest variance between bands in the GST control lane and bands in the GST-UBLD lane was observed when the binding buffer contained between 100 to 150mM NaCl (the concentration at which the buffer most closely resembles physiological conditions).

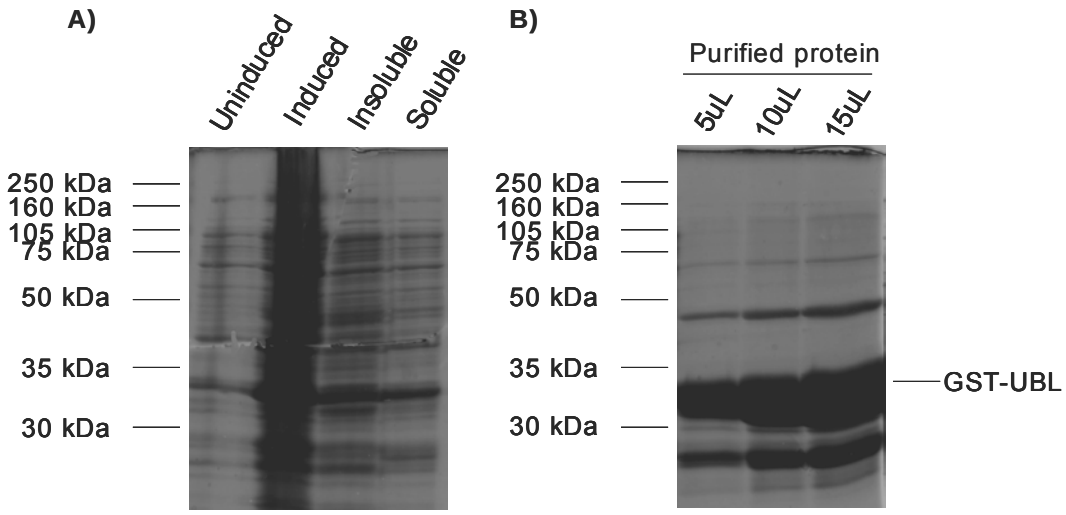


Figure 3.9: ***Bacterial expression and purification of GST-tagged FAM-UBLD.*** **A)** Bacterial expression of GST FAM-UBLD is suggested by the presence of a band at the predicted molecular weight for the GST-FAM-UBLD fusion protein in both the insoluble and soluble fractions with no such band seen in the uninduced sample. Approximately 50% of the expressed protein appears in the soluble fraction. **B)** Expressed fusion protein from **A)** was purified on glutathione agarose and 5, 10, or 15 μL of the sample was resolved by SDS PAGE. GST-FAM-UBLD was able to be purified in large quantities with a reasonable purity (some contaminating bands can be seen at higher molecular weights than that of GST-FAM-UBLD, however these are of a relatively low abundance when compared to GST-FAM-UBLD). The band seen at ~ 26 kDa is likely to correspond to isolated GST that remains after spontaneous breakdown of the GST-FAM-UBLD fusion protein.

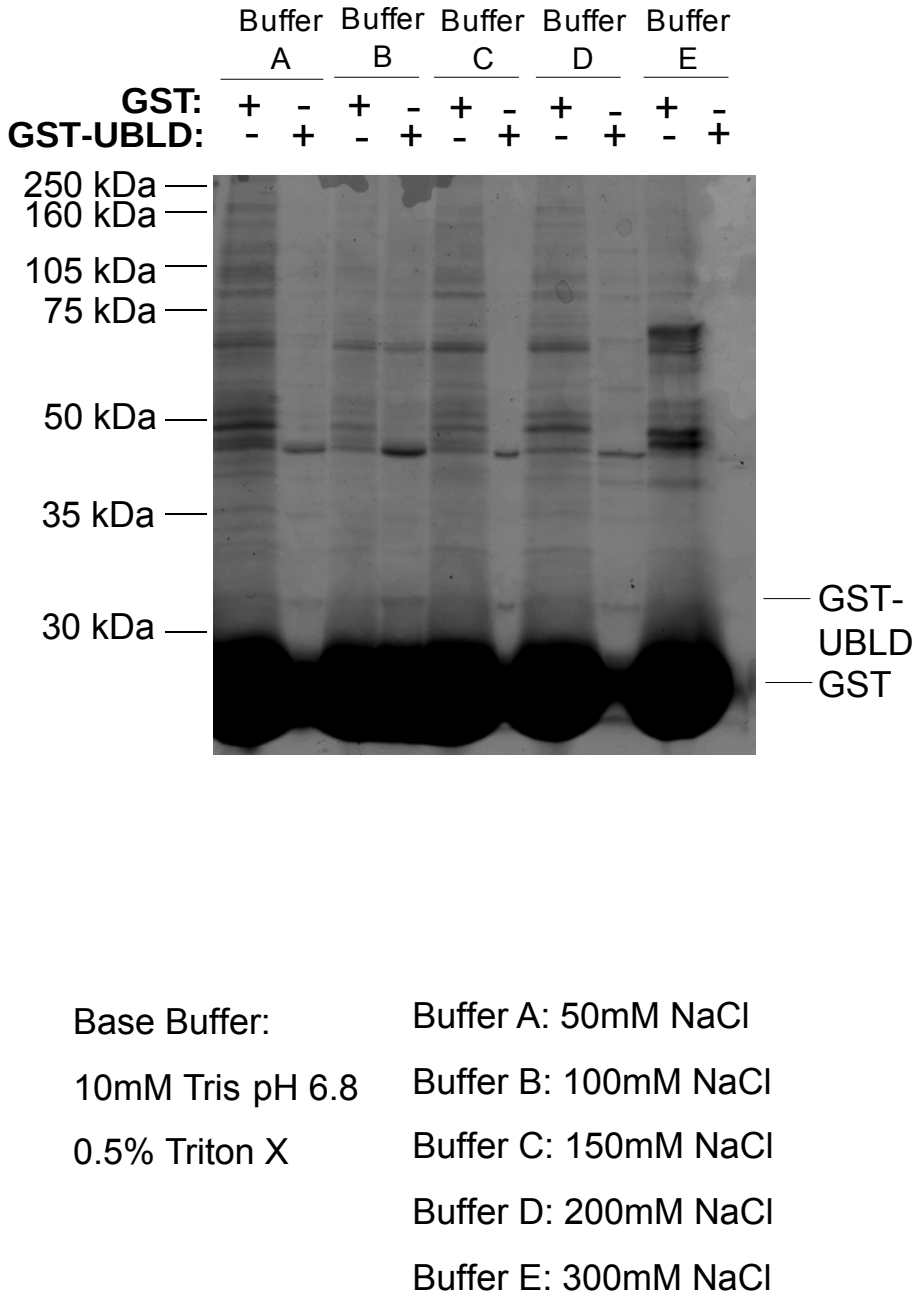


Figure 3.10: ***GST Affinity Chromatography – Unbiased approach.*** **A)** An initial trial experiment to determine the optimum buffer conditions for the affinity binding assay. 293T cell lysate was incubated with either GST alone or GST-FAM-UBLD in five different salt concentrations conferring a range of stringency levels from low to high stringency. Buffers B and C appeared to give the largest differences between the GST alone and GST-FAM-UBLD lanes and were used in subsequent experiments.

In the affinity binding assay the amount of each fusion protein used was normalized with 10mg of each being used in each assay. This assay was also performed with lysate from whole mouse brains, HEK 293T cells, and mouse ES cells. Protein binding to immobilised GST/GST-FAM-UBLD bait was carried out in buffer containing either 100 or 150mM NaCl. Approximately 10mg of 293T cell and mouse brain lysates and 5mg of mESC cell lysate were incubated with the fusion proteins. Five bands were identified, both by the author and by an independent researcher, on a coomassie stained gel containing the results of the binding affinity experiment using 293T cell lysate. The sizes of these bands were approximately 45 kDa, 52 kDa, 71 kDa, 95 kDa, and 140 kDa and appeared to be unique to the lanes loaded with GST-FAM-UBLD (data not shown). These bands were excised from the gel and sequenced by tandem mass spectrometry (Chris Bagley, Hanson Institute Protein Core Facility, Institute for Medical and Veterinary Science, Adelaide, SA). Each band contained multiple proteins with the most abundant protein in each band listed in Table 3.2.

Table 3.2: Most Abundant Protein in Bands Analysed by Mass Spectrometry

Approximate Size of Band	Most Abundant Protein
45 kDa	Putative tagatose 6 phosphate kinase gatZ from *E. Coli* (47,078.78 Da, accession #2507030) with 11 matched peptides
52 kDa	Human Tubulin Beta 5 (49,638.97 Da, accession #135471) with 16 matched peptides
71 kDa	Human 78 kDa glucose regulated protein precursor GRP 78 Immunoglobulin heavy chain binding protein BiP (72,288.435 Da, accession #14916999) with 14 matched peptides
95 kDa	Glycine dehydrogenase decarboxylating Glycine decarboxylating Glycine cleavage system P protein from *E. Coli* (104,309.753 Da, accession #417041) with 16 matched peptides
140 kDa	Carbamoyl phosphate synthase large chain from *E. Coli* (117,781.359 Da, accession #22095507) with 10 matches

Three of the bands that were sequenced (p45, p95, and p140) contained contaminating bacterial proteins as the predominant species and were likely copurified with the GST-FAM-UBLD fusion protein. The remaining two human proteins, β-tubulin 5 and GRP78 BiP, while possibly being binding partners of FAM, are also potentially false positives. GRP78 BiP is a stress induced chaperone that is likely to bind to many proteins that have exposed hydrophibic patches. β-tubulin 5 is quite possibly a binding partner for FAM. FAM localises to points of protein trafficking and has a punctate cytoplasmic staining that may correspond to vesicular

structures traveling along microtubules (Murray et al., 2004). However, tubulin proteins are often identified in GST affinity binding experiments as false positives due to their sticky nature, therefore this potential interaction was also not pursued further. In an effort to avoid contaminating bacterial proteins, the GST affinity binding experiment was performed again with an extra step included. Expressed bacterial proteins were purified, eluted from the GST affinity columns and buffer exchanged into 10mM Tris. 500µg of the purified proteins were then bound to fresh GST affinity columns and washed extensively with the medium stringency buffer described above. Total HeLa cell lysate (5mg) was then incubated with the columns in 100mM NaCl binding buffer. The columns were then washed, bound proteins eluted and proteins resolved by SDS PAGE. Although many protein bands were observed on the gel, upon comparison of the bands observed in the GST lane with the GST-FAM-UBLD lane, no bands appeared to be unique to the GST-FAM-UBLD lane (Figure 3.11). This indicated that despite pre-clearing the lysate on unbound glutathione agarose beads, a large proportion of mammalian proteins were binding either to the agarose beads themselves, to non-specific contaminating bacterial proteins on the columns, or to GST.

Generation of anti-FAM UBLD and anti-GST antibodies

The high solubility, stability, and ease of purification of GST-FAM-UBLD made it an ideal antigen for anti-FAM antibody production. Two rabbits (MPCH-1 and MPCH-2) were immunised with GST-FAM-UBLD and pre-bleed and test bleed samples taken at appropriate time points (http://www.imvs.sa.gov.au/vet/info_docs/antibody/antibody_index.html). Test bleed samples were tested for immunoreactivity against HEK 293T cell lysates before the serum was collected (data not shown). As the antibody was raised against GST-FAM-UBLD, it was reasoned that there would not only be anti-FAM UBLD antibodies in the serum, but also anti-GST antibodies. As the only means available to affinity purify the anti-FAM antibodies required the use of immobilised GST-FAM-UBLD, the serum first needed to be cleared of anti-GST antibodies to avoid co-purification of the two. In order to achieve this, glutathione affinity purified GST was crosslinked to cyanogen bromide (CNBr) activated sepharose, producing a column of immobilised GST capable of clearing the rabbit serum of anti-GST antibodies. Serum from immunised rabbits was first passed through the GST column. The anti-GST antibodies were then eluted from the GST column for use in bacterial expression experiments. Anti-GST antibodies were tested for titre and specificity by western

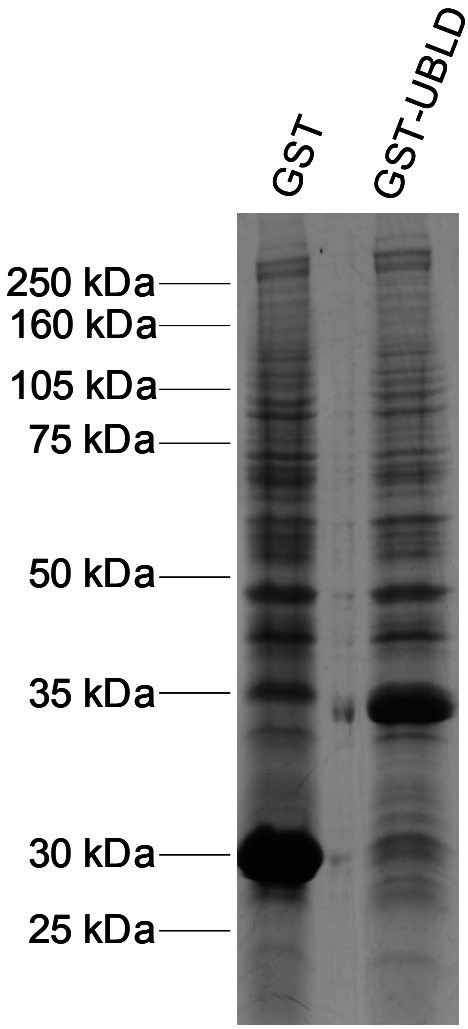


Figure 3.11: ***GST-affinity binding experiment with twice bound fusion proteins.*** Fusion proteins were purified on glutathione sepharose columns and then rebound and re-washed on fresh glutathione sepharose beads before the affinity binding experiment was performed. Total HeLa cell lysate was incubated with each fusion protein (GST or GST FAM-UBLD) before bound proteins were eluted and resolved by SDS PAGE. Bands were visualised by Coomassie Blue staining.

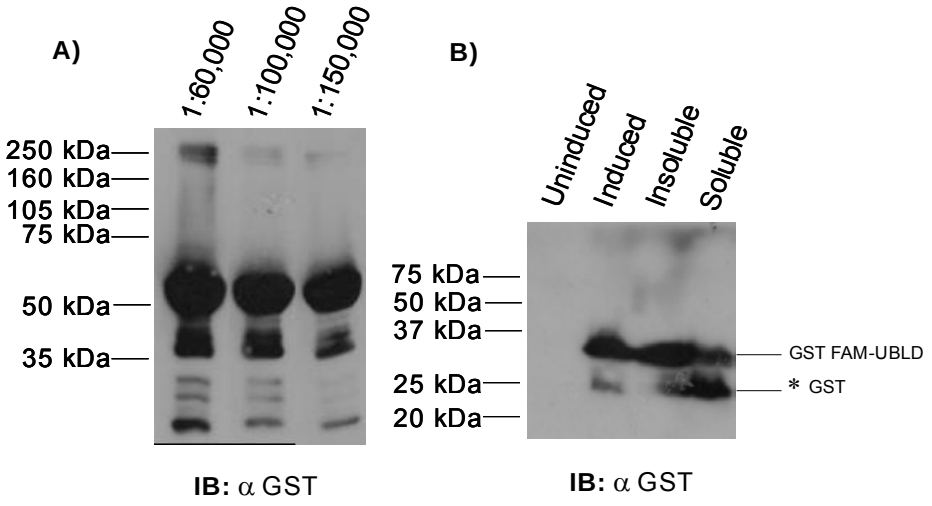
analysis of an affinity purified GST fusion protein (GST VHS-GAT). The optimal dilution factor for the anti-GST antibody for western analysis was found to be within the range of 1 in 150,000 to 1 in 200,000 (Figure 3.12 a). The αGST-antibody was then tested for specificity against an total bacterial cell lysate from a GST-FAM-UBLD fusion protein expression (Figure 3.12 b). The antibody displayed good specificity with bands appearing only in lanes where a GST fusion protein was expected (induced, insoluble, and soluble lanes) and only at the expected molecular weight for the GST fusion protein. A similar antibody affinity column was prepared with immobilised GST-FAM-UBLD and anti-FAM UBLD antibodies were purified from anti-GST antibody cleared serum using this column. Test westerns on total HEK 293T cell lysate with a dilution series of this antibody determined that the optimal concentration for the use of the anti-FAM UBLD antibody for western analysis was within the range of 0.5 to 2 μg/mL (Figure 3.12 c).

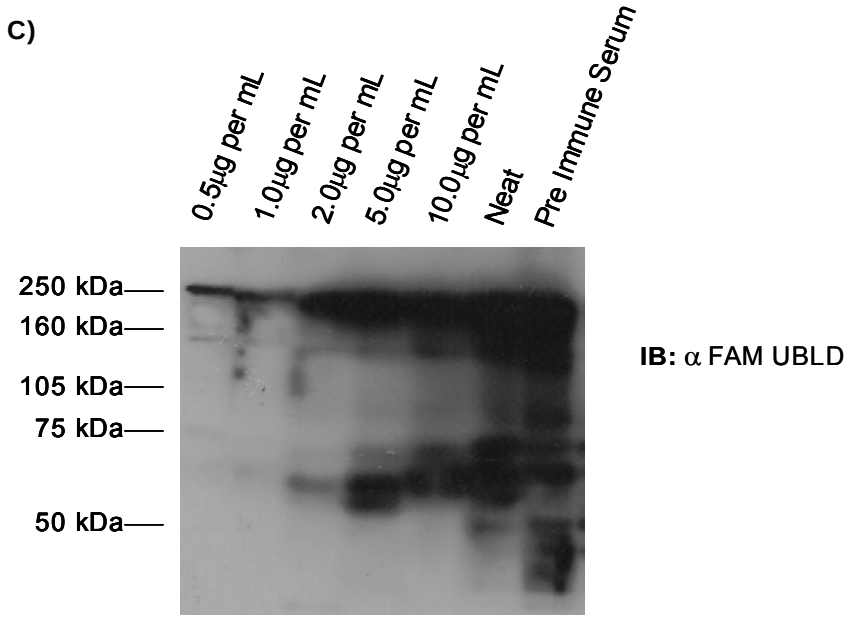
Best-Candidate Approach

The best candidate approach to identify novel proteins that interacted with the FAM-UBLD brought together the knowledge of the types of proteins that the UBL domains of other UBD proteins interacted with, and the experimentally determined intracellular localisation of FAM. As discussed above, UBL domains have not only been found to interact with the proteasome, but also with UIM containing proteins involved in the protein trafficking pathways of the cell. Given that FAM has shown co-localisation with a number of markers for various points on the protein trafficking pathway (Murray et al., 2004), there was a good chance that the UBLD of FAM may also be able to interact with ubiquitin binding proteins in these pathways. To determine if this was the case, GST affinity binding experiments were conducted as described above and the eluted bound proteins were resolved by SDS PAGE. Proteins were then transferred to nitrocellulose membranes and western analysis performed using antibodies raised against 6 protein trafficking pathway proteins that are known to bind ubiquitin. These were Stam, Hrs, Epsin 1, and the three GGA proteins, GGA1, GGA2, and GGA3. Stam and Hrs did not appear to be bound to either the GST or GST-FAM-UBLD columns. Bands at the predicted molecular weights of 60 and 86 kDa for Stam and Hrs respectively, were seen only in whole cell lysate control lanes and not in GST or GST-FAM-UBLD lanes (Figure 3.13 a & b).

There are conflicting reports on the apparent molecular weights of the GGA proteins in the literature. Some report the apparent molecular weight of GGA1 as less than 66 kDa (Kakhlon et al., 2006), others as over 75 kDa (Boman et al., 2000) and even 90 kDa (Yogosawa et al., 2006). GGA2 has been reported at an apparent molecular mass of 65-70

Figure 3.12: ***Assessment of α-GST and α-GST-FAM-UBLD antibody titre and specificity***. **A)** Dilutions of purified anti-GST antibody (1/60,000, 1/100,000, and 1/150,000) in 5% Milk PBS Tween-20 were tested for immunoreactivity against a GST-VHS-GAT fusion protein (50μg). **B)** Specificity of the α-GST antibody was tested by performing western analysis of a GST-FAM-UBLD bacterial expression. Immunoreactivity was observed only in lanes in which expressed protein would be expected (induced, insoluble, and soluble lanes), and only at the predicted molecular weight for the fusion protein. ∗ The band at 26kDa is likely to correspond to isolated GST that remains upon spontaneous degradation of the fusion protein. **C)** A range of purified anti-FAM-UBLD antibody concentrations (0.5 μg/mL, 1.0 μg/mL, 2.0 μg/mL, 5.0 μg/mL, 10.0 μg/mL, diluted in 5% milk PBS Tween-20, and undiluted [neat]) tested for immunoreactivity against whole cell lysate from HEK-293T cells. Minimal background was observed with an antibody concentration of between 0.5 and 2.0 μg/mL.


A)
1:60,000
1:100,000
1:150,000
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
35 kDa
IB: α GST
B)
Uninduced
Induced
Insoluble
Soluble
75 kDa
50 kDa
37 kDa
25 kDa
20 kDa
GST FAM-UBLD
* GST
IB: α GST


C)
0.5μg per mL
1.0μg per mL
2.0μg per mL
5.0μg per mL
10.0μg per mL
Neat
Pre Immune Serum
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
IB: α FAM UBLD

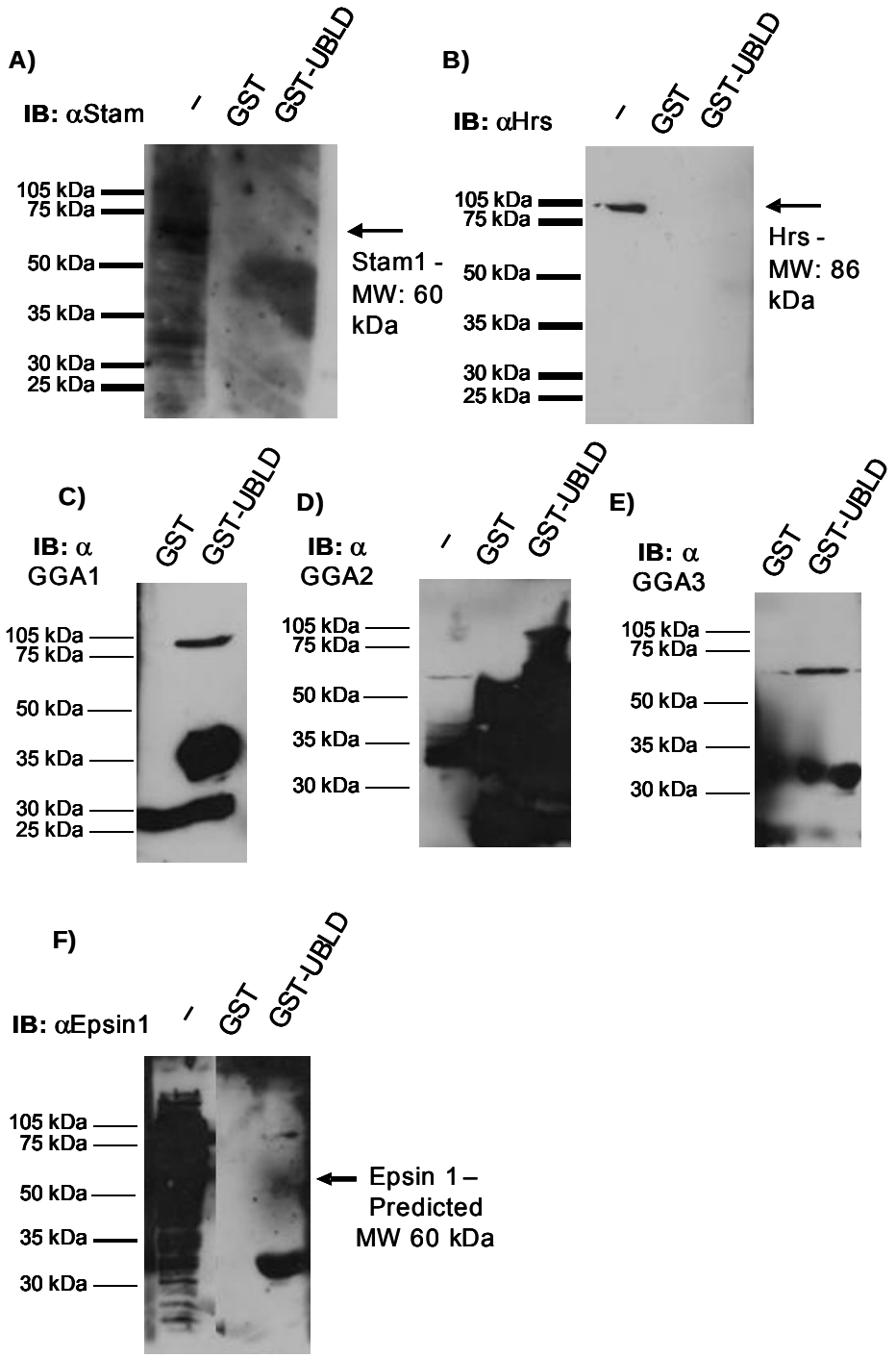
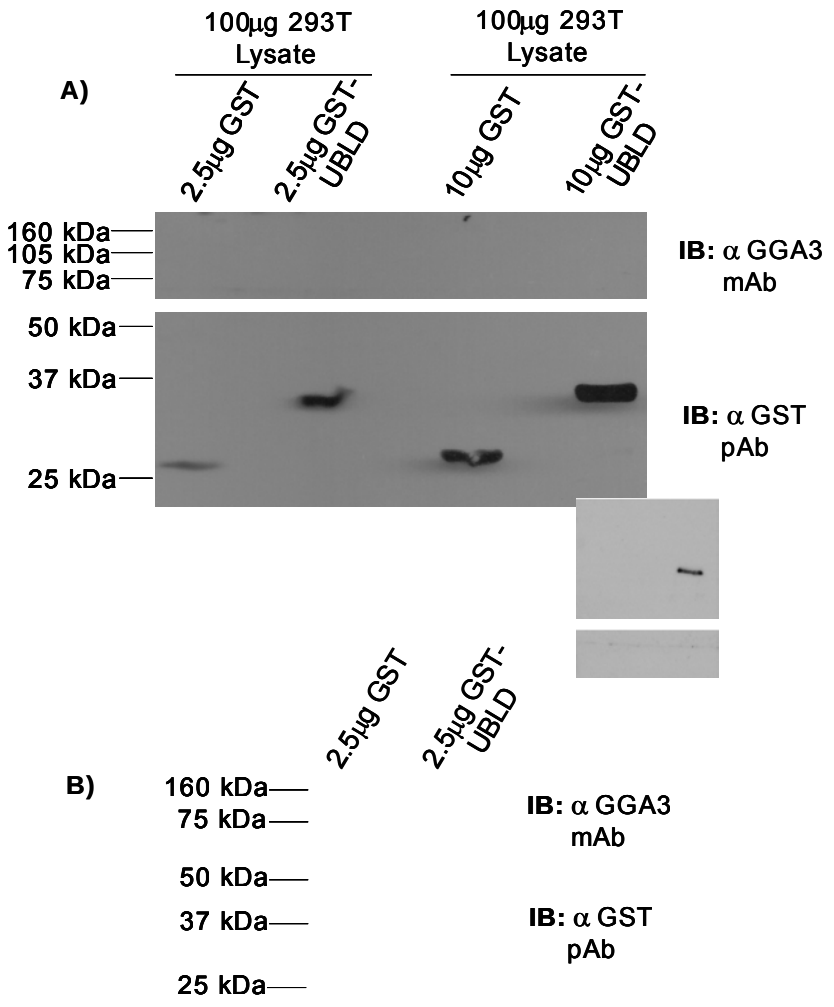


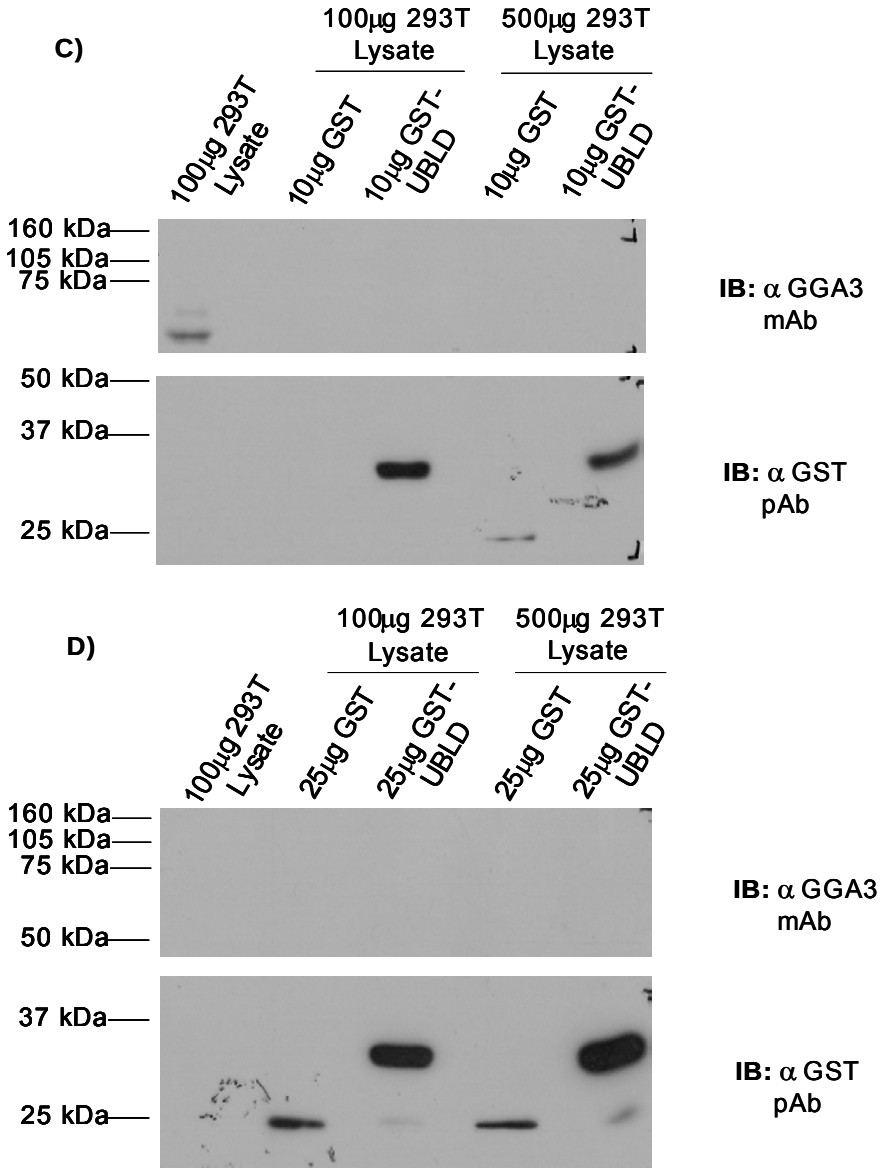
Figure 3.13: ***GST affinity chromatography – Best Candidate approach attempt I*** Total HEK 293T cell lysate (1mg) was incubated with 1mg of either immobilised GST alone or GST-FAM-UBLD overnight at 4°C. Bound proteins were eluted and each of the two elutions were split between six wells of a polyacrylamide gel. Lysate controls (100 µg of total HEK 293T cell lysate) were also loaded for Stam, Hrs, Epsin 1, and GGA2. Blots were cut into strips and then probed with **A)** αHrs, **B)** αStam, **C)** αGGA1, **D)** αGGA2, **E)** αGGA3, or **F)** αEpsin 1

kDa (Dell'Angelica et al., 2000; Yogosawa et al., 2006) while GGA3 has been reported at 90 kDa (Dell'Angelica et al., 2000). Bands were observed at ~ 90 kDa and 70 kDa in the GST-FAM-UBLD lanes only for blots probed with GGA1 and GGA3 antibodies respectively (Figure 3.13 c & e). Unfortunately, lysate alone controls were not included for these two interactors in the initial assay. A band at ~ 65 kDa was observed in the lysate control lane for the blot probed with anti-GGA2. A band at the same size may be visible in the GST-FAM-UBLD lane (Figure 3.13 d). Background staining however, obscured this part of the blot and also obscured the corresponding region of the GST alone lane making this result inconclusive. Similarly, a band at the predicted molecular weight of Epsin 1 (60 kDa) may have been observed in the GST-FAM-UBLD lane and not in the GST alone lane. However, the high level of background staining, and the large number of non-specific bands that appear in the lysate control lane make it difficult to draw any conclusions. Subsequent repeats of the experiment resulted in inconsistent results. These included apparent weak binding of two of the three GGA proteins to the GST-FAM-UBLD fusion protein specifically (Appendix E.1), or non-specific binding of the GGA proteins to both the GST-FAM-UBLD fusion protein and to GST alone (appendices E.2 and E.3). In every GST affinity experiment performed, a band at the correct molecular weight for Hrs was observed only in the lysate control lane (Figure 3.13 b, Appendix E.1 a, Appendix E.2 a, & Appendix E.3 a). Binding of Epsin 1 to the putative FAM UBLD was also never conclusively observed (Figure 3.13 f & Appendix E.1 e).

One possible explanation for the above results may have been the quality /purity of the antibodies used. The antibodies used to detect the GGA proteins had previously only been used in ELISA based assays and hadn't been tested for their specificity on westerns. To circumvent any potential failure of the antibodies to correctly detect the endogenous GGA proteins, the GST affinity binding experiments were repeated with two independent modifications. Firstly, a commercially available GGA3 monoclonal antibody was used in place of the sheep polyclonal. However, this antibody also proved unreliable in a number of GST affinity binding experiments including a series of experiments that were aimed at optimizing the conditions for the pulldown (data not shown and Figure 3.14 a-d). In these experiments, a number of variables were explored. This included using a much smaller quantity of affinity beads in a microfuge tube (10-50 μL of beads were used here whereas previous binding studies had been performed in column format with a bed volume of 0.2mls) (Figure 3.14 a & b). Affinity binding studies were also carried out using reduced quantities of purified fusion protein (2.5-25 μg of fusion protein), and concomitantly, reducing the quantity

Figure 3.14: ***GST affinity chromatography – Best candidate approach optimisation***. Binding assays were performed in a batch format in these studies as opposed to the column format of previous studies. Fusion proteins were bound to either 10 μl **A), C),** and **D)** or 50 μl **B)** glutathione sepharose beads in a microcentrifuge tube. In **A)**, 2.5 or 10 μg of each fusion protein was incubated with 100 μg of 293T cell lysate. In **B),** 2.5 μg of each fusion protein was incubated with 100 μg of 293T cell lysate. In **C),** 10 μg of each fusion protein was incubated with either 100 or 500 μg of 293T cell lysate. In **D),** 25 μg of each fusion protein was incubated with either 100 or 500 μg of 293T cell lysate. In each case blots were probed with a monoclonal α GGA3 antibody to detect the presence of co-purified endogenous GGA3 protein. Lysate alone controls were included in C & D however endogenous GGA3 protein was not detected.


A)
100µg 293T Lysate
100µg 293T Lysate
2.5µg GST
2.5µg GST-UBLD
10µg GST
10µg GST-UBLD
160 kDa
105 kDa
75 kDa
50 kDa
37 kDa
25 kDa
IB: α GGA3 mAb
IB: α GST pAb
B)
2.5µg GST
2.5µg GST-UBLD
160 kDa
75 kDa
50 kDa
37 kDa
25 kDa
IB: α GGA3 mAb
IB: α GST pAb


C)
100µg 293T Lysate
500µg 293T Lysate
100µg 293T Lysate
10µg GST
10µg GST-UBLD
10µg GST
10µg GST-UBLD
160 kDa
105 kDa
75 kDa
50 kDa
37 kDa
25 kDa
IB: α GGA3 mAb
IB: α GST pAb
D)
100µg 293T Lysate
500µg 293T Lysate
100µg 293T Lysate
25µg GST
25µg GST-UBLD
25µg GST
25µg GST-UBLD
160 kDa
105 kDa
75 kDa
50 kDa
37 kDa
25 kDa
IB: α GGA3 mAb
IB: α GST pAb

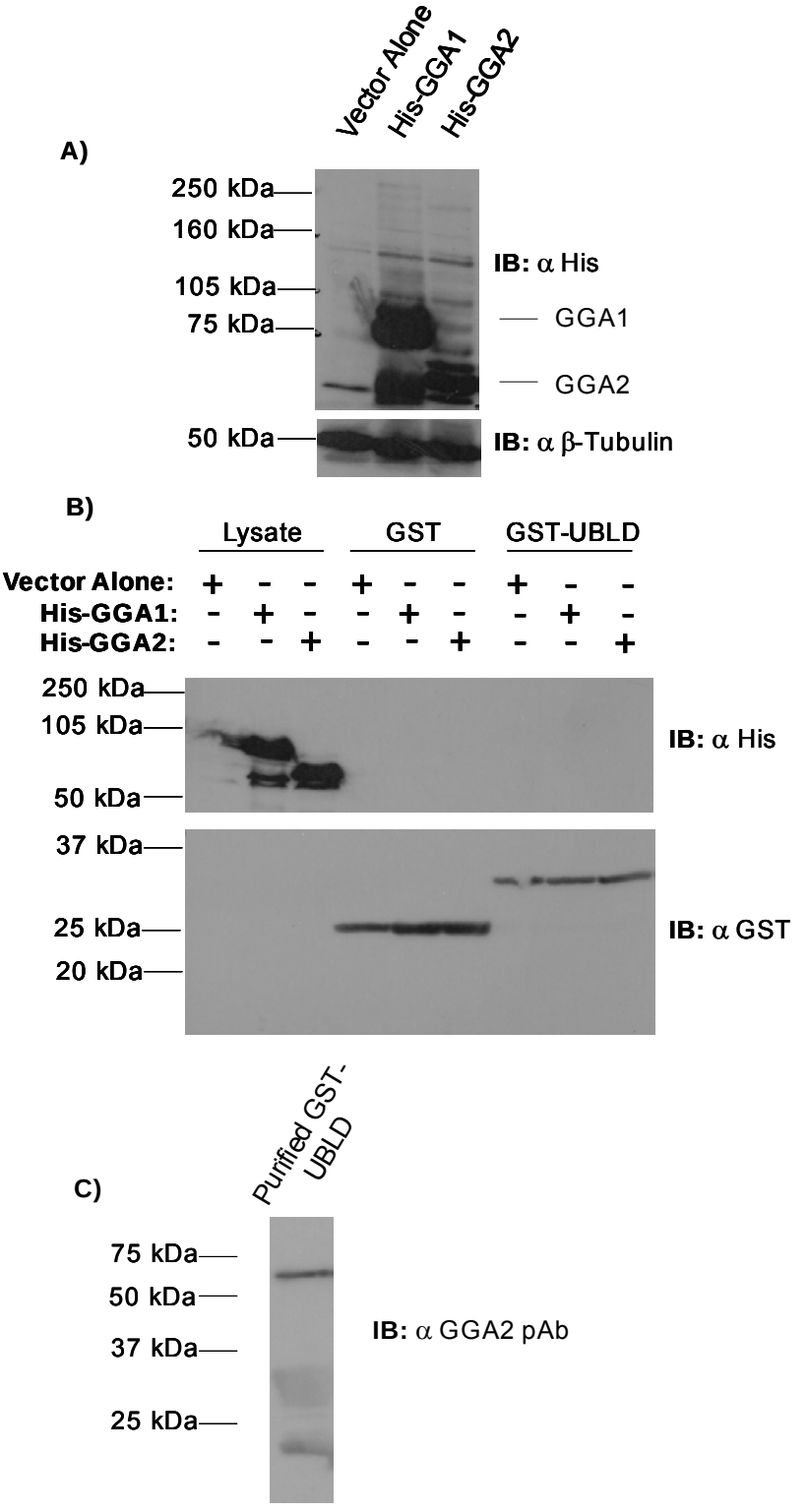
of total cellular lysate (lysate was reduced to 100-500 μg) thereby increasing the ratio of lysate to fusion protein (Figure 3.14 c & d). This optimization proved unsuccessful.

The second step that was taken to avoid antibody problems was to transiently transfect 293T cells with a vector driving expression of GGA1, and GGA2 with an amino-terminal 6xHIS tag from a CMV promoter to enable detection of the GGAs with an anti His antibody. Despite overexpression of the exogenous proteins to quite high levels (Figure 3.15 a), GST affinity binding experiments failed to detect any significant binding of either His-GGA1 or his GGA2 to either GST-FAM-UBLD or GST alone (Figure 3.15 b). The polyclonal GGA antibodies that were used in the previously described affinity binding assays were raised against bacterially expressed GST fusions of the GAT domain of each GGA. It seemed possible therefore, that the sheep that were immunised with the GST-GGAx-GAT domain fusion proteins were also immunised against a contaminating bacterial protein. These antibodies were column purified by a similar method described for the purification of the GST and FAM-UBLD antibodies that the author generated. It was possible therefore, that antibodies that were raised against this contaminating bacterial protein were also co-purified with the anti-GGA antibodies on the GST-GGA-GAT columns, and were cross reacting with the same or a similar bacterial protein that co-purified with the GST-FAM-UBLD fusion protein. To test this, 50ug of purified GST-FAM-UBLD was run on an SDS polyacrylamide gel, transferred and a western analysis performed with the GGA2 polyclonal antibody. An immunoreactive band was observed at the predicted molecular weight for GGA2 (Figure 3.15 c). This result either confirmed that there were non-GGA antibodies in the purified antibody solution or suggested that the polyclonal anti-GGA antibodies were able to non-specifically cross react with bacterial proteins.

Further exploration of the FAM /GGA interaction

Despite the inconsistent findings described above, the possibility of an interaction between FAM and the GGAs was still plausible. As discussed above, the GGA proteins are ubiquitylated and FAM has been seen to localise to regions within the cell corresponding to the localisation of the GGAs. It is possible that the interaction between the GGA proteins and FAM does not require the UBLD of FAM. To test for this possibility, co-immunoprecipitation experiments were conducted to determine if the endogenous proteins interact. Co-immunoprecipitation experiments were conducted using a rabbit polyclonal antibody directed against the 20 N-terminal amino acids of FAM and lysates from both HEK 293T cells, and whole mouse brain extracts. Bands were observed in the lysate control lanes at the predicted

Figure 3.15: ***GST-Pulldowns with overexpressed GGA proteins.*** **A)** 293T cells were transiently transfected with a vector driving expression of Myc/His tagged GGA1 and GGA2 fusion proteins. Cells were harvested 48hrs post transfection and western analysis was performed with a monoclonal α His antibody to determine the extent of overexpression. High levels of overexpression were observed for both proteins. **B)** GST-affinity binding assays were performed as described previously using 10 μg of each fusion protein and 100 μg lysate from 293T cells overexpressing His-tagged GGA1 and GGA2. No binding was observed for either fusion protein. **C)** 50 μg of purified GST-FAM-UBLD protein was run on an SDS polyacrylamide gel and western analysis performed with the polyclonal α GGA2 antibody. A band was seen at the size corresponding to the size of the band that was seen in previous GST-affinity binding experiments and assumed to be GGA2. This indicated that the antibody was cross reacting with a co-purified contaminating bacterial protein and was not detecting endogenous mammalian GGA2 in pulldown experiments.


A)
Vector Alone
His-GGA1
His-GGA2
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
IB: α His
GGA1
GGA2
IB: α β-Tubulin
B)
Lysate
GST
GST-UBLD
Vector Alone: + - - + - - + - -
His-GGA1: - + - - + - - + -
His-GGA2: - - + - - + - - +
250 kDa
105 kDa
50 kDa
IB: α His
37 kDa
25 kDa
20 kDa
IB: α GST
C)
Purified GST-UBLD
75 kDa
50 kDa
37 kDa
25 kDa
IB: α GGA2 pAb

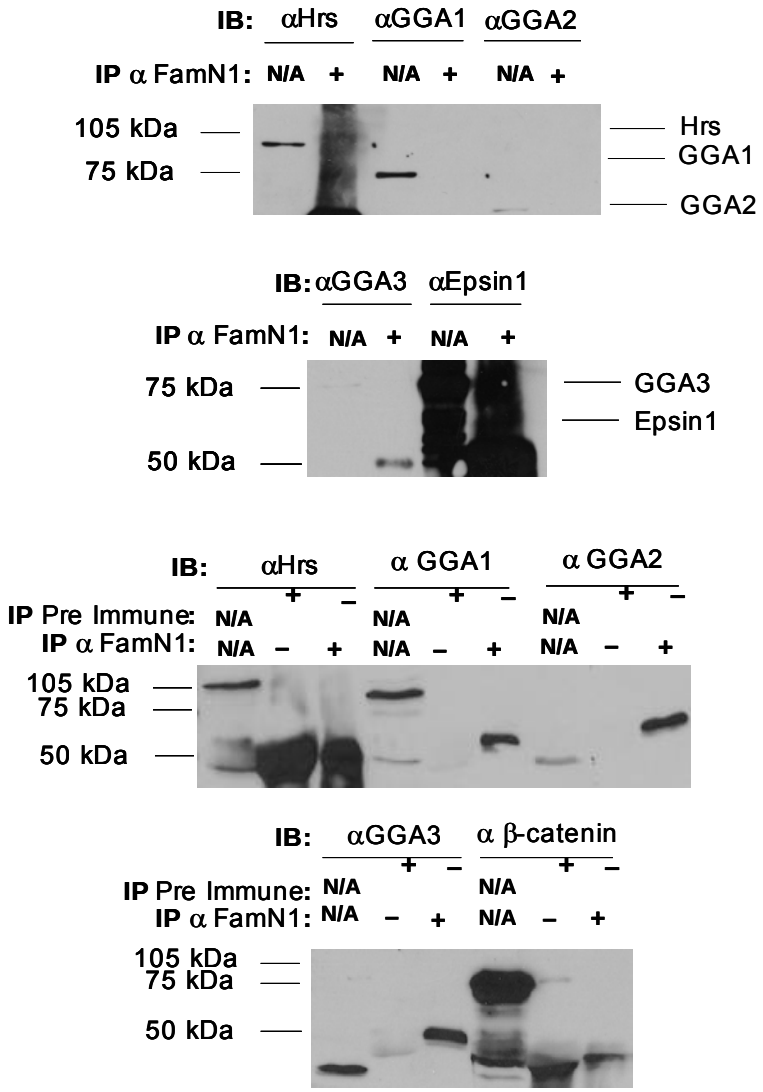


Figure 3.16: ***Assessing the ability of endogenous FAM and the GGA proteins to interact in vivo by co-immunoprecipitation*** **A)** Total 293T cell extracts were subjected to immunoprecipitation with α FAM N1 antibodies. Co-immunoprecipitation was not observed for any of the five proteins tested. **B)** Whole mouse brain cell extracts were subjected to similar immunoprecipitation with α FAM N1 antibodies with similar results.

molecular weights for the three GGA proteins and Hrs in the immunoprecipitation experiment from 293T cell extracts (Figure 3.16 a), and for GGA1, β-catenin, and Hrs in the immunoprecipitation experiment from whole mouse brain extracts (Figure 3.16 b). Similar bands were not observed however, in the anti-FAM immunoprecipitated lanes in each case (Figure 3.16 a & b).

GST affinity chromatography experiments were also performed using GST fusions of the isolated domains of GGA3 (Figure 3.17) to determine if any part of GGA3 could precipitate wild-type FAM from cell extracts. The results of these experiments proved to be inconclusive. In one experiment, an anti-FAM-UBLD antibody immunoreactive band was seen at the experimentally determined molecular weight of FAM (290 kDa) in the lanes of proteins that co-eluted with GST-VHS-GAT and Hinge-GAE (Figure 3.18 a). However, in this experiment, a band at the molecular weight of FAM was not seen in the lysate control lane. A band may also been seen in the VHS alone lane that is much fainter and it is hard to determine if this is simply background staining (Figure 3.18 a). It is possible that this staining would be as dark as that seen for VHS-GAT and Hinge-GAE if the amount of fusion protein used for VHS was similar to that used for the other fusion proteins. No binding was observed for GST alone or for GST-GAT (Figure 3.18 a). In the second experiment, a very faint band is seen at the molecular weight of FAM in the VHS-GAT and GAT lanes but not in the Hinge-GAE lane (Figure 3.18 b). Very strong staining is seen in the lysate control lane (Figure 3.18 b).

Lastly, the effect of overexpressing the GGA proteins on the cellular localisation of FAM was explored. Constructs driving the expression of 6XHis-Flag tagged fusions with each of the GGA proteins (pcDNA3.1 6XHis-Flag GGA-1, 2, or 3) were transiently transfected into COS1 cells in independent experiments. A high level of overexpression of the GGA proteins is associated with a disruption of the morphology of the Gogi apparatus including “frizzling” of the Golgi stacks, compaction, and fragmentation of the Golgi into vacuolar-like spherical structures (Boman et al., 2000; Poussu et al., 2000; Takatsu et al., 2000). This phenomenon was also apparent in the experiments described here. Cells that were transfected 6xHIS-Flag tagged GGA1 and GGA3 in particular, and to a lesser extent cells that were transfected with GGA2 displayed spherical vacuolar-like staining with an anti-Flag antibody (Figure 3.19). FAM staining in cells expressing exogenous GGA1, 2, and 3 was significantly altered when compared to untransfected or vector transfected control cells. In vector transfected or untransfected cells, FAM displayed the previously reported punctate cytoplasmic staining which is often combined with a perinuclear staining pattern that is more concentrated to one side of the cells nucleus. In GGA1 and GGA3 transfected cells in particular, and to a lesser

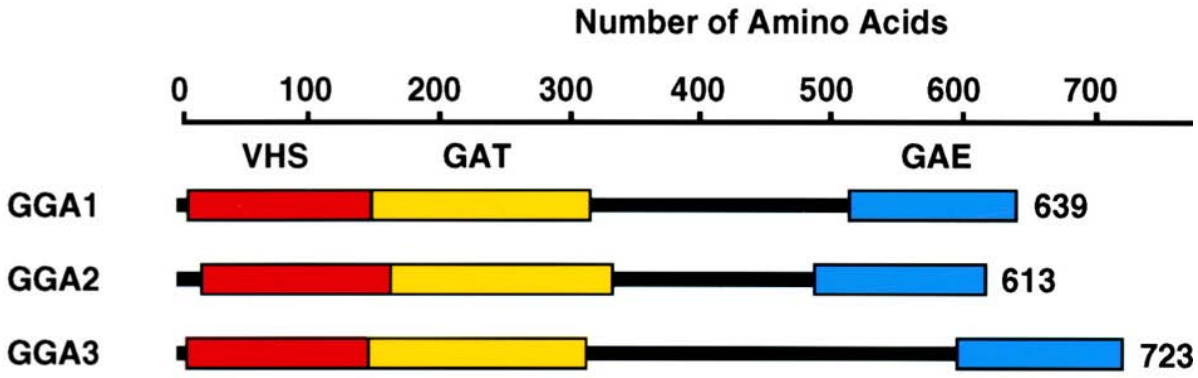
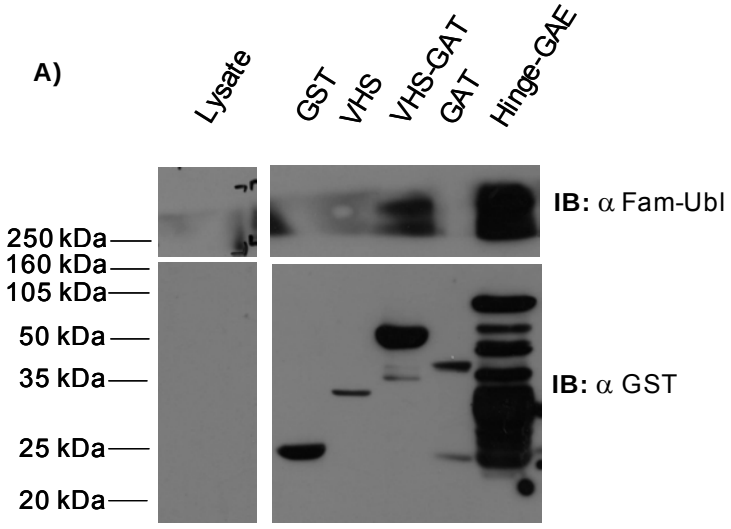
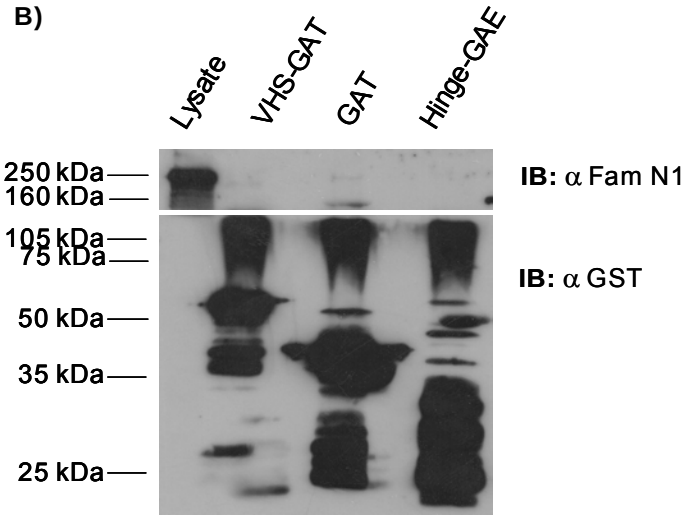


Figure 3.17: ***Domain organization of the three GGA (Golgi-associated, gamma-adaptin ear-like) proteins***. All three proteins have a common four domain structure. Situated at the N-terminus is a VHS (Vps27, Hrs, and Stam) domain. The VHS domain is responsible for binding to the DXXLL sorting motif of the Golgi cargo binding mannose-6-phosphate receptors enabling packaging of both the receptor and its cargo into clathrin coated vesicles. Following the N-terminal VHS domain is a GAT domain (GGA and Tom1) that is known to recruit the GGA proteins to the Golgi membrane via its interaction with activated GTP-bound Arf proteins. The GAT domain has also been found to associate with ubiquitin and has therefore been postulated to confer an additional level of specificity for cargo substrates. At the C-terminus resides a GAE (gamma-adaptin ear-like) domain. This domain, in combination with a flexible hinge region located between the GAT and GAE domains, interacts with and is responsible for recruiting clathrin and other adapter proteins that are required for formation of clathrin coated vesicles (adapted from (Dell'Angelica et al., 2000).

Figure 3.18: ***Reverse GST-Pulldown to test for the ability of the GGA proteins to interact with endogenous FAM***. GST affinity binding assays were performed using the column method **A)** or the batch method **B)** as described previously **A)** 500μg of total HeLa cell lysate was incubated on columns containing 100μg of immobilised GST, GST-GGA3-VHS, GST-GGA3-VHS-GAT, GST-GGA3-GAT, or GST-GGA3-Hinge-GAE. Bound proteins were eluted and subjected to SDS PAGE and western analysis with α FAM-UBLD and α GST antibodies. Some fusion proteins (GST-GGA3-VHS and GST-GGA3-GAT) were underloaded. α FAM-UBLD immunoreactive bands were seen in the GST-GGA3-VHS-GAT and GST-GGA3-Hinge-GAE lanes but not in the remaining lanes. An α FAM-UBLD immunoreactive band was not observed in the lysate control lane either making this experiment inconclusive. **B)** 500μg of total 293T cell extract was incubated with 10μg of each fusion protein immobilised on 50μL of glutathione sepharose beads. Bound proteins were eluted and subjected to SDS PAGE and western analysis with α FAM-N1 and α GST antibodies. Faint α FAM-UBLD immunoreactive bands were seen in the GST-GGA3-VHS-GAT and GST-GGA3-GAT lanes but not in the GST-GGA3-Hinge-GAE lane. A very strong α FAM-UBLD immunoreactive band was seen in the lysate control lane.


A)
Lysate
GST
VHS
VHS-GAT
GAT
Hinge-GAE
250 kDa
160 kDa
105 kDa
50 kDa
35 kDa
25 kDa
20 kDa
IB: α Fam-Ubl
IB: α GST

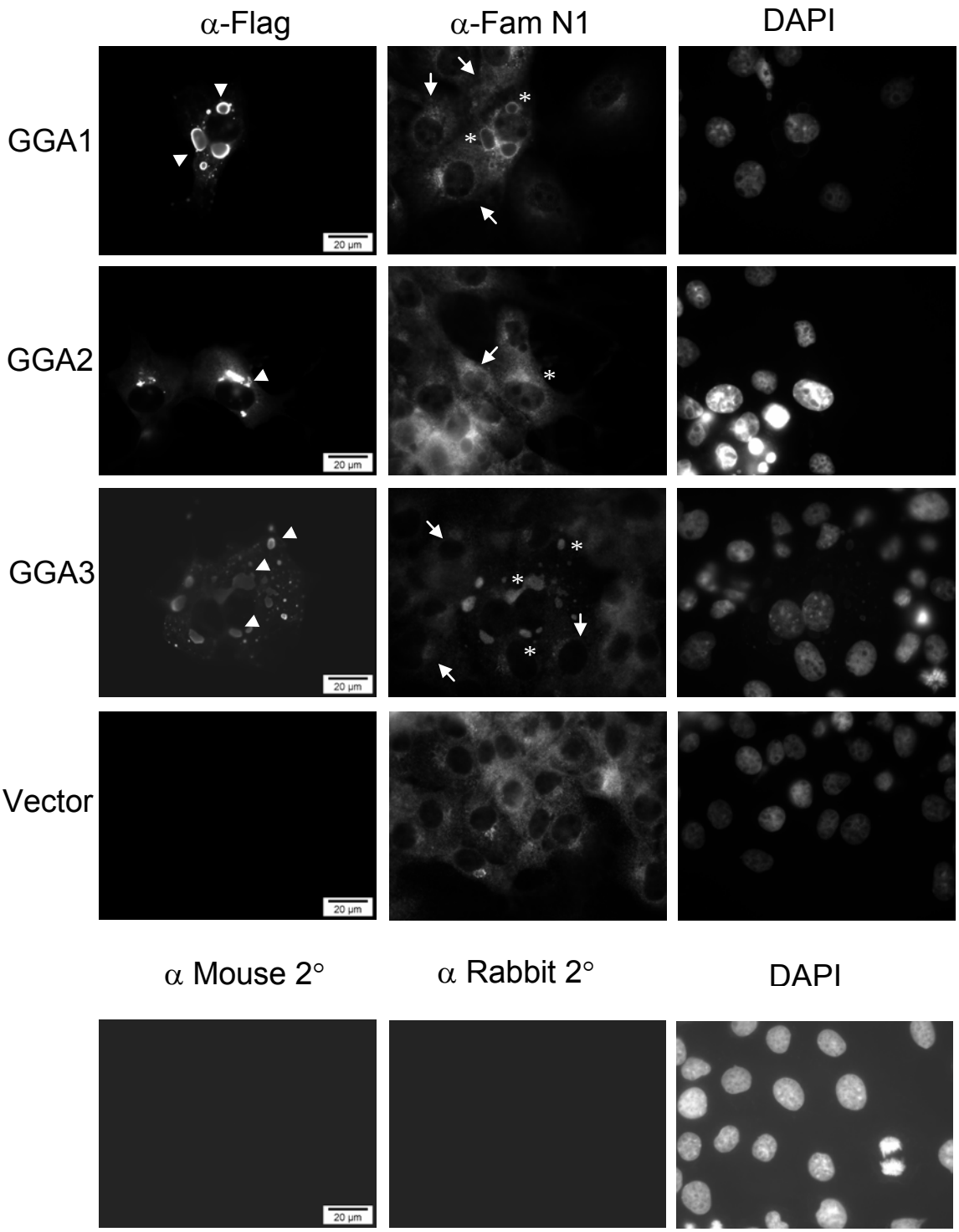

B)
Lysate
VHS-GAT
GAT
Hinge-GAE
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
35 kDa
25 kDa
IB: α Fam N1
IB: α GST

extent in GGA2 transfected cells, this staining pattern is present but an additional FAM distribution consisting of a ring of staining surrounding the anti-Flag immunoreactive globular structures is also present (Figure 3.19).

FAMΔUBLD

In an effort to gain insight into the function of the putative UBLD of FAM, the region spanning predicted amino acid sequence for FAM-UBLD was deleted from the Gateway entry vector pDONR201-FAM, generating the deletion mutant designated FAMΔUBLD. The FAMΔUBLD cDNA was then sub-cloned into the Gateway pDEST40 mammalian expression vector. This vector drives expression of C-terminally V5 tagged FAMΔUBLD from a CMV promoter. HEK 293T cells were transiently transfected with this vector, pDEST40 FAM (wild-type FAM), or pDEST51 FAM C1556S (catalytically inactive cys mutant FAM). Expression of wild-type and cys mutant FAM were observed by western blotting with an anti-V5 antibody however expression of the UBLD deletion mutant was not observed (Figure 3.20). Smeared bands at lower molecular weights than the known molecular weight of FAM were observed in the two lanes loaded with FAMΔUBLD (Figure 3.20). These may be breakdown products of FAMΔUBLD indicating that deletion of the UBLD from FAM causes instability of the protein.

Figure 3.19: ***The effect of overexpression of the GGA proteins on the cellular localisation of FAM.*** COS1 cells grown on coverslips were transiently transfected with 10ug of pcDNA3.1 6xHis-Flag-GGA1, 2, 3 or vector alone. Cells were incubated for 48 hours before being fixed, permeabilized, and stained with a monoclonal anti-Flag antibody and a polyclonal anti-FAM antibody or secondary mouse and rabbit antibodies alone. As previously reported, overexpression of the GGA proteins (in particular GGA1 and GGA3) led to morphological disruption of the Golgi apparatus with the formation of globular vacuolar-like structures in transfected cells (arrowheads). In untransfected (those cells that don't show immunofluorescence with anti-Flag antibodies indicated by arrows) and vector control transfected cells, the distribution of FAM protein consists of a punctate cytoplasmic staining pattern, often with large concentrations of FAM staining towards one side of the nucleus. In transfected cells, this distribution is altered such that in addition to the staining pattern seen in untransfected cells, a ring is formed that surrounds the spherical structures which stain positive for the GGA proteins (asterisks). All images were taken at 100 X magnification.


α-Flag
α-Fam N1
DAPI
GGA1
GGA2
GGA3
Vector
20 µm
α Mouse 2°
α Rabbit 2°
DAPI

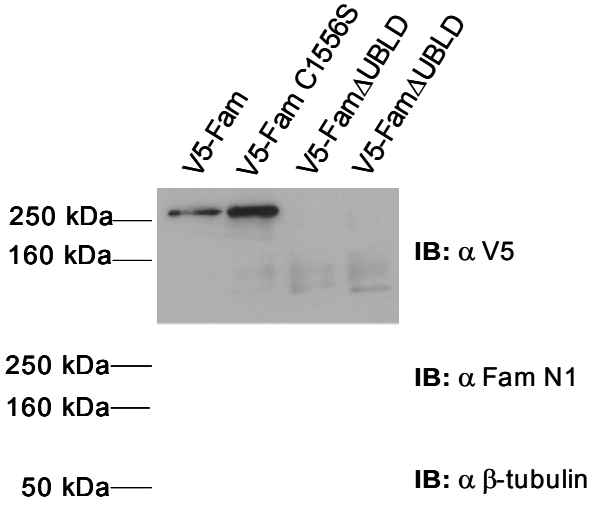


Figure 3.20: ***Test expression of FAMΔUBLD***. Expression vectors for wild-type, catalytically inactive, and UBLD deletion mutant V5 fusion proteins were transiently transfected into 293T cells. Cells were incubated for 72 hours to allow expression of the fusion proteins before being harvested and analysed by western blot. Blots were probed with αV5 antibody to detect exogenously expressed protein, α FAM N1 to determine overall FAM levels, and α β-tubulin as a loading control. No exogenously expressed UBLD deletion mutant FAM was observed although a smear of bands were observed at molecular weights smaller than the known molecular weight of FAM, possibly indicating degradation products of the expressed protein.

DISCUSSION

The aim of the experiments described above was to gain a better understanding of how FAM interacts with its many substrates and accessory proteins. This was approached in two ways. Studies were undertaken 1) to better characterize the binding sites on FAM for known substrates and interacting proteins and 2) to identify and characterize novel features of the FAM protein with the potential to act as protein-protein interaction domains. Unfortunately in the first approach no useful information was able to be gained about the interactions of FAM with known binding partners through the use of three widely used and successful binding assays. Despite this, these techniques may yet prove useful in investigations of the interactions of FAM in future experiments. In the sections below, a number of ways of improving the outcomes of these techniques in investigating FAM and its interactions will be discussed. A higher degree of success was achieved in the second approach with the identification and characterization of the ubiquitin-like domain of FAM. Investigations into the functional characteristics of this domain led to the identification of potential interactions of FAM with the three members of the GGA family of Golgi complex sorting/adapter proteins. Approaches to further strengthen the evidence for a FAM/GGA interaction will also be discussed in the sections to follow.

Yeast Two-Hybrid Analysis of Known Interactions

While no interactions were observed between any part of FAM and its known binding partners in the yeast two-hybrid studies discussed above, there were indications that the assay system had the potential to provide such information in this laboratory. Activation of the system was observed when yeast cells were transformed with an expression plasmid for wild-type β-catenin-GAL4 DBD, through the action of β-catenin's transactivation domains Activation of the system was also seen when epsin 1 was expressed in isolation as a GAL4 DBD fusion, possibly through the interaction of epsin 1 with endogenous yeast transcription factors. Positive control strains that were included in every assay returned positive results in each of the reporter systems tested. Furthermore, an independent positive control consisting of two proteins that were previously shown to interact in a yeast two-hybrid system also returned a positive result. The positive results obtained indicate that the selection media and reagents were adequate for detection of an interaction. It would be interesting however, to conduct the analysis again using a different set of expression vectors such as those that are commercially available from BD Biosciences.

Problems may have also been encountered in the transformation efficiency of the yeast cells. The transformation protocol that was used to introduce expression vectors into the yeast cells has much higher transformation efficiency when transformations are performed sequentially. That is, cells are first transformed with the bait vector and selected on single amino acid dropout media, and then subsequently transformed with the prey plasmid. In the experiments described in this chapter, both vectors were transformed into the yeast cells at the same time. This effectively reduces the transformation efficiency from ~ 10^5 colonies per μg DNA to ~ 10^4 colonies per μg DNA (BD-Biosciences, 2001). The effect of this is that if the product of one of the expression vectors is cytotoxic there is a greater selective pressure for clones in which spontaneous deletions in the coding region of the gene product have occurred (BD Biosciences, 2001). To address this issue, a "parental strain" could first be generated consisting of yeast cells transformed with the expression vector that is common to all tested interactions. In the case of the FAM binding region of β-catenin, the cytoplasmic domain of E-cadheren, and itch, this would consist of the GAL4 DBD (pDEST32) expression vector, while for epsin 1 this would consist of the GAL4 AD (pDEST22) expression vector. Following selection on appropriate synthetic complete dropout medium, the parental strain could then be transformed with each of the expression vectors for the control vectors, or the fragment of FAM to be tested.

Mammalian Two-Hybrid

The biggest failing in this assay resulted from the low level of expression from the reporter vector, even upon stimulation (personal communication, C. Bracken, PhD graduate, The University of Adelaide, SA). Many of the expression and reporter plasmids that were used in this assay were constructed in-house or modified from existing vectors. It is possible that the difficulties encountered in this assay system may have been overcome by using a commercially available mammalian two-hybrid kit. At the time that this work was begun, no commercial Gateway compatible mammalian two-hybrid kit was available. There are however, a number of other commercially available mammalian two-hybrid kits. It would be worthwhile trialing one of these systems to determine their usefulness in investigating the interactions of FAM with other proteins. The mammalian two-hybrid system would probably not be useful in the investigation of interactions of FAM with itch and epsin 1, as both mammalian proteins are known to be involved in signaling pathways which would be likely to interfere with the read-out of the mammalian two-hybrid reporter system.

GST Affinity Chromatography

Despite the fact that there were no obvious signs of soluble protein produced for any of the expressed FAM domains, under any of the conditions trialed, it is possible that there was in fact a small amount of soluble protein produced. At the time that this work was conducted, the only anti-GST antibody that was available in the laboratory produced very high background staining to the point that it was not useful for distinguishing expressed GST-fusion proteins from endogenous bacterial proteins. For this reason, none of the polyacrylamide gels that were run to determine the solubility of the fusion protein expressions, were checked by western analysis, a far more sensitive technique than coomassie staining. There are a number of commercially produced anti-GST antibodies available that would enable this to be done. Furthermore, in the course of producing an antibody raised against the putative UBLD of FAM, it was also possible to affinity purify a high titre, highly specific rabbit polyclonal anti-GST antibody from the same serum. Western analysis of the small scale inductions performed on the FAM domain expression vectors would have enabled the detection of even small amounts of soluble protein.

It is also possible to recover expressed proteins from the insoluble inclusion bodies that *E. Coli* produce. This can be done by first solubilising the protein from the inclusion bodies by incubating the insoluble material in one of a number of denaturing buffers. These can include 4-6 M guanidine hydrochloride, 4-8 M urea, alkaline > pH 9, organic solvents, or strong detergents such as 0.5-2% triton-X 100 or N-lauroylsarcosine (Sarcosyl) (Frangioni and Neel, 1993; Marston, 1986). The major problem with this however, is that following solubilisation, the proteins lose their natural fold structure due to the denaturing conditions. This means that the proteins must be refolded before any functional analysis can be performed on them. This relies on the removal of denaturants by desalting or by dilution/dialysis and may or may not allow refolding of the protein and formation of the correct intramolecular associations. There are no established conventions in this procedure and no guarantees that it will produce a correctly folded protein. This ultimately means that a lot of time must be spent optimizing the denaturing/refolding conditions with no certainty of a result at the end.

Lastly, the predicted FAM domain boundaries that were originally suggested have since been refined by analyzing the sequence of FAM using a number of online bioinformatics programs. These include, the JUFO server (a program for calculating the secondary structure of a protein from primary sequence data (Meiler et al., 2002), GlobPlot (a program that predicts globular domains and the intervening unstructured regions of a protein

(Linding et al., 2003), SAPS (Statistical Analysis of Protein Sequences; a program that evaluates a protein sequence for a wide variety of protein sequence properties by statistical criteria (Brendel et al., 1992), and the fold recognition program LOOPP (Learning Observing and Outputting Protein Patterns (Teodorescu et al., 2004). From this analysis four major domains were predicted and designated domain 1, domain 2, the catalytic core, and the c-terminal domain (P. Khut, Masters Graduate, The University of Adelaide). Domain 1 (amino acids 60-975) is predicted to contain three subdomains (defined as small structural units that posses the necessary information to undergo spontaneous non-covalent formation of a localised folding unit with native-like secondary and/or tertiary structural features (Wu et al., 1994)), subdomain 1.1 (amino acids 60-187), 1.2 (amino acids 219-867), and 1.3 (888-975). Subdomain 1.3 encompasses the putative UBLD of FAM. Domain 2 is predicted to span amino acids 1009-1549. The catalytic core is predicted here to encompass amino acids 1792-2261. The c-terminal domain encompasses amino acids 2039-2480. It is possible that these new FAM domain sequences may represent a closer approximation of the independent protein fold boundaries, and therefore be more likely to fold as a soluble protein when expressed in bacteria.

Identifying Interactors of FAM-UBLD

Unbiased approach

Although there were no unique FAM-UBLD interacting proteins identified by this approach, it is worthwhile to discuss ways to improve the assay to give greater stringency, and other methods that can be used for identifying novel interacting proteins. The major problem that was encountered with this GST-affinity binding assay was the presence of contaminating bacterial proteins and non-specific binding of proteins to the affinity columns. One way to remedy the issue of purity of expressed proteins from bacteria would be to use a tandem affinity purification (TAP) system. In this system, the fusion protein is expressed as an N- or C-terminal fusion with two affinity tags, a 6XHis tag and a GST tag for example, separated by a TEV (tobacco etch virus) cleavage site. This allows a two step purification process whereby the fusion protein is first bound to a nickel affinity column with unbound proteins removed in the washing steps. Elution is achieved by cleavage of the first affinity tag with TEV protease. The eluted fusion protein can then be bound to a glutathione affinity column with further removal of contaminating proteins by a second round of washes allowing the preparation of highly purified fusion protein for use in binding assays.

Strategies for decreasing the amount of non-specific binding of mammalian proteins to the columns could address the problem from two angles. Specificity could be improved prior to the affinity experiments by not only pre-clearing the lysate on unbound glutathione agarose beads, but by also using beads coated with total bacterial protein content. A total bacterial lysate column could be prepared by lysing untransformed bacteria and incubating the soluble cell lysate with CNBr activated sepharose beads. Specificity could also be improved in the wash stages following binding of total mammalian cell lysate with the fusion protein bound beads. The wash buffers trialed in the experiments described above ranged from quite mild to a medium stringency wash buffer containing a high salt content. Wash buffer stringency can be further increased by using a higher stringency detergent such as SDS, or by simply increasing the concentration of the detergent used. Interestingly, TAP was first developed to be used as a method to identify interacting protein complexes in mammalian cells (Puig et al., 2001). Using TAP in this manner would completely overcome the problem of contaminating bacterial proteins. Alternatively, novel interacting partners of the FAM-UBLD could be identified by yeast two-hybrid analysis. Library screening by this method is very commonly and successfully used to identify new protein-protein interactions in many laboratories.

Best candidate approach

Although this approach has been used successfully to identify proteins that interact with the UBL domains of other UDPs (Fallon et al., 2006; Heir et al., 2006; Regan-Klapisz et al., 2005) it met with limited success when applied to the identification of proteins that interact with the putative FAM UBLD. Part of the reason for this may be attributed to the polyclonal anti-GGA1, 2, and 3, anti-stam, and anti-epsin 1 antibodies that were used in the assay. Initial attempts at GST-affinity binding assays using these antibodies provided misleading information due to high background staining and low specificity. Despite extensive optimization and the use of alternative methods for the detection of the GGA proteins, the results obtained in the initial affinity binding assay were unable to be repeated. It is worth noting however, that the size of the band that was detected with the polyclonal anit-GGA1 antibody, both in the lysate control lanes and the GST-FAM-UBLD lanes, was consistently larger than the band observed with the anti-GGA2 and anti-GGA3 antibodies. It seems unlikely that there would be two contaminating bacterial proteins that co-purified with GST-FAM-UBLD and could not only be recognised by one antibody and not the other, but also resolve by SDS PAGE with the size differential predicted for GGA1 when compared to GGA2. It remains a possibility therefore, that the interaction between the putative FAM-UBLD domain and GGA1 is in fact real, if weak, and that only the observed interactions

between GGA2 and GGA3 with FAM-UBLD are artefactual. Indeed, it has been shown that GGA2 does not have the ability to bind to ubiquitin as GGA1 and GGA3 do (Shiba et al., 2004). For these reasons, it would be worthwhile to perform the experiments again with more specific anti-GGA antibodies.

When attempts were made at detecting the interaction by the use of exogenously expressed, epitope tagged GGA proteins, a negative result was obtained. There are a number of reasons why this may have been the case. It is possible that the N-terminal epitope tag on the GGA proteins interfered with the GGA/FAM-UBLD interaction. The use of a similar C-terminal tag would rule this out as a possibility. It is also possible that the negative result was a consequence of the disruption of the Golgi and the mechanics of the TGN due to high level expression of the GGA proteins. This may have been a result of 1) the presence or lack of a post-translational modification of the GGA proteins (eg. phosphorylation or ubiquitylation) possibly caused by the saturation of the modifying enzyme or an inability of the enzyme and the GGA proteins to localise to the same compartment, 2) competitive binding and occlusion of the FAM-UBLD binding site on the GGA proteins by another protein caused by a retention of this protein at the TGN or 3) the absence of a cofactor required for the GGA/FAM-UBLD interaction or other unknown mechanisms. It has been reported that a lower level of exogenous expression of the GGA proteins has a less extreme effect on TGN morphology although the subcellular distribution of other TGN and GGA interacting proteins is still altered (Boman et al., 2000; Dell'Angelica et al., 2000). It would be interesting to perform the affinity binding assay with cells expressing a more moderate level of exogenous epitope tagged GGA proteins.

FAM is known to interact with, and is likely to deubiquitylate epsin 1 in a calcium signaling dependent manner (Chen et al., 2003). While the catalytic core of FAM has been shown to be sufficient to bind epsin 1, investigations have not revealed if this region is solely responsible for the FAM/epsin 1 interaction. As discussed above, although Parkin is able to interact with Eps15 via its catalytic ring finger domain, the interaction between Parkin and Eps15 is entirely ablated if the UIM/UBL domain interaction is lost. It is possible that, as with Parkin and Eps15, the catalytic core mediated binding of FAM with epsin 1 may not be the only interaction that occurs. A UIM/UBL domain interaction may also occur between FAM and epsin 1 and, as with Parkin and Eps15, this interaction may be essential for binding of the two proteins to occur. The experiments described above failed to detect an interaction between the FAM-UBLD and epsin 1. However, the polyclonal anit-epsin antibody that was used in the assay had a very high background level and seemed to cross react with a number

of proteins as evident in the large number of bands seen in lysate control lanes. It would be interesting to perform the *in vitro* binding assays described above with a more specific epsin 1 antibody, or with overexpressed, tagged epsin 1. It is also possible that an interaction may only be observed if cells are stimulated, as with the enhancement of the interaction between Parkin and Eps15 upon EGF treatment. It would be interesting to observe the effect of stimulation of cells with EGF or treating them with with a calcium mobilizing agent such as ionomycin, on the interactions of FAM with other proteins.

Exploring the FAM/GGA interaction further

Coimmunoprecipitation experiments rely on a number of factors to produce results. Antibody antigen interactions need to be strong enough to immobilise the target protein while still being able to be disrupted to release the precipitated protein for western analysis, a property that is highly variable between antibodies. Some antibodies for instance, work well in western analysis or immunofluorescence, but don't perform quite so well in coimmunoprecipitation experiments. The interaction strength between the two proteins of interest is also a factor. A weak interaction may cause the coprecipitating protein to be lost in the wash stages. Factors affecting the interaction *in vivo* also come into play such as post translational modifications. For example, phosphorylation brought on by a signaling event is known to be essential for some protein-protein interactions to occur with no interaction seen in the absence of such an event. For these reasons, an *in vivo* interaction between FAM and the GGA proteins cannot be ruled out. Indeed, the fact that in at least two separate experiments, bacterial expressed and purified GST-fusion proteins of isolated GGA3 domains were able to purify endogenous FAM protein from total cellular lysates seems to support the idea of an interaction between FAM and the GGAs.

The effect of overexpression of the GGA proteins on the intracellular distribution of FAM was also examined. In each case (overexpression of GGA1, GGA2, and GGA3) the distribution of FAM intracellular localisation was altered such that FAM protein became concentrated in puncta in rings surrounding the globular vacuole-like structures which formed in transfected cells and that also stained positive for the exogenously expressed GGA proteins. This suggests that, at least in COS-1 cells, the GGA proteins are either directly or indirectly responsible for the transport of FAM to the cellular compartments in which it normally resides. Several possibilities exist to explain this. As discussed above, it is possible that high levels of overexpression of the GGA proteins saturates or titrates out an enzyme that is responsible for the post-translational modification of the GGA proteins. It is possible that

this post-translational modification may mediate the correct formation and/or budding of GGA coated TGN to endosomal/lysosomal transport vesicles. If the GGA proteins are directly involved in the transport of FAM, inhibition of the formation, budding, or transport of these carrier vesicles would also inhibit the transport of FAM. It may even be possible that it is the deubiquitylating activity of FAM this is being saturated and that deubiquitylation of the GGAs is required for these processes. One way to test this possibility would be to either overexpress or reconstitute FAM null cells with a catalytically inactive mutant of FAM. If FAMs deubiquitylating activity was required for correct GGA function, it might be expected that similar globular vacuolar-like structures that stain positive for the GGA proteins might form. If the GGA/FAM interaction is confirmed to be mediated by the UBLD of FAM, another way to test for this possibility would be to overexpress or reconstitute FAM null cells with a FAMΔUBLD or FAM UBLD mutant protein that is no longer able to interact with the GGA proteins.

A second possibility is that GGA overexpression disrupts FAM localisation indirectly by preventing the association of FAM with other proteins that promote trafficking from the TGN. GGA overexpression has previously been reported to lead to exclusion of AP-1 and the M6PRs from the Golgi (Boman et al., 2000; Dell'Angelica et al., 2000). It is possible that FAM interacts with one of these or another protein that is excluded from the TGN by GGA proteins and it is this interaction that is responsible for the correct localisation of FAM. If this is the case, then catalytically inactive FAM or a FAMΔUBLD/UBLD mutant may not lead to the formation of the GGA immunoreactive globular vacuolar-like structures formed with GGA overexpression. Instead, altering the levels of AP-1 or the M6PRs may lead to the formation of perinuclear rings of FAM staining similar to those seen with GGA overexpression.

Generation of anti-FAM UBLD and anti-GST antibodies

Two rabbits were immunised with GST-FAM-UBLD to generate an antibody to the UBLD of FAM for general laboratory use. Affinity columns coupled with GST alone and GST-FAM-UBLD were prepared and were subsequently used to purify GST antibodies, clear the FAM-UBLD immunreactive serum of GST antibodies, and to purify FAM-UBLD antibodies. In each case, the process was successful as antibodies with specific immunoreactivities towards GST and FAM, having low background levels, were isolated. The anti-GST and anti-FAM UBLD antibodies were subsequently used to detect bacterially expressed GST fusion proteins and endogenous FAM from mammalian cell lysates

respectively. The anti-FAM UBLD antibody was also tested for use in immunofluorescence with some limited success (data not shown).

FAMΔUBLD

Although attempts at expressing the UBLD deletion mutant of FAM in 293T cells were unsuccessful, this investigation may still be possible. Restriction digest analysis of the entry vector for FAMΔUBLD indicated that the region spanning the UBLD had been successfully removed from the FAM cDNA sequence. As discussed above, attempted expression of V5 epitope tagged FAMΔUBLD resulted in the presence of a number of lower molecular weight species that were anti-V5 immunoreactive. If these bands are indeed breakdown products of FAMΔUBLD, degradation of the protein may have occurred due to major structural faults in the protein caused by the deletion. Alternatively, it may be possible that the FAM-UBLD has a mechanistic role in maintaining the stability of the protein. Investigation of the effect on FAM function of deletion of the UBLD region may still be possible by deleting a smaller section of FAM. If the UBLD of FAM is actually a subdomain of a larger domain as suggested by P. Khut, it may be necessary to delete the entire region spanned by the predicted FAM domain 1 (amino acids 60-975). Alternatively, single point mutations could be introduced into the UBLD region based on the available data on the interaction of ubiquitin and other UBLDs with ubiquitin binding domains. The major site of interaction in many cases appears to revolve around a hydrophobic core formed by Leu 43, Ile 44, and Val 70 of ubiquitin, and the equivalent amino acids of the UBLDs of other UDPs (May et al., 2004 leucine 43; Mueller and Feigon, 2003; Sakata et al., 2003; Walters et al., 2002). A FAM L934A, F935A double mutant (corresponding to a ubiquitin L43A, I44A mutant) has been generated by the author and awaits analysis.

Concluding Remarks

Many of the techniques described in this chapter have been used previously to identify interactions with FAM. It is therefore worthwhile to consider those investigations with respect to the studies undertaken here.

In vitro binding studies have formed the basis for most research into the interactions of FAM to date. AF-6, β-catenin, and Epsin 1 interactions with FAM have all been investigated using affinity chromatography assays (Chen et al, 2003, Taya et al, 1998; Taya et al, 1999). In the case of AF-6 interactions, an *in-vitro* transcription/translation system was used to produce

four overlapping ^{35}S methionine labelled fragments of FAM spanning the entire molecule. These were mixed with immobilised GST-tagged bacterially expressed AF-6 protein.

There are several limitations to the experiments conducted here. Firstly, while an interaction with an 890 amino acid fragment of FAM spanning its catalytic domain was observed, there was no demonstration of the efficiency of protein production of the four fragments. It is therefore hard to judge whether the fragment that was observed to bind was simply the one most amenable to correct folding and expression. Secondly, the boundaries of the four fragments described were arbitrarily chosen and therefore unlikely to encompass independently folding units of the FAM protein. Thirdly, the *in vitro* transcription and translation system employed utilised bacterial transcription components coupled with a reticulocyte lysate. While eukaryotic chaperones and translational machinery are present in this mix, reports have suggested that the rapid transcription promoted by the bacterial T7 polymerase can lead to an uncoupling of transcription from translation and can result in non bio-active or incorrectly folded protein (Iskakova et al., 2006). This is especially the case with unstable or difficult to fold proteins. Considering the difficulties encountered in expressing these same fragments as GST fusion proteins described in this chapter, it seems likely that these four FAM fragments would fall into this category.

Studies into the interactions of FAM with β-catenin and Epsin 1, and further investigations of AF-6 interactions utilised bacterial expression of only one fragment of the FAM protein. This was performed using only GST or maltose binding protein fusions of the catalytic domain of FAM (amino acids 1498-1918) for β-catenin (Taya et al, 1998) and AF-6 (Taya et al, 1999) respectively or a slightly shifted GST fusion of this region (amino acids 1554-1953) for Epsin 1 (Chen et al, 2003) . The use of this one small fragment of FAM may have been warranted for investigations of AF-6/FAM interactions as the work with *in-vitro* transcribed and translated FAM fragments suggested that this was the site of AF-6 binding on FAM. In contrast, there is no evidence in the literature to suggest that regions outside of the catalytic domain of FAM are not involved in interactions with β-catenin and Epsin 1 therefore these investigations would seem to be incomplete. Investigations into the physical interaction of FAM with Itch reported in the literature are also relatively limited. As yet, it is only known that the full length FAM protein is able to interact with itch and no attempts have been made at mapping the site of their interaction on FAM (Mouchantaf et al, 2006). Other approaches to investigation of protein-protein interactions with FAM reported in the literature have included co immunoprecipitation experiments, library screening by yeast two-hybrid analysis and

western overlay experiments, all of which have their limitations. The experiments described in this chapter were conducted to address some of these limitations.

In-vitro Binding Studies

The work in this chapter described attempts at generating bacterially expressed GST fusion proteins spanning the entire length of the FAM protein based on a partial proteolysis analysis designed to identify the domain structure of FAM. These studies were conducted to address the issue of incomplete investigation of the interactions of β-catenin, epsin 1, and itch with FAM. The approach that we employed was taken to avoid difficulties caused by attempting to express protein fragments from FAM cDNA dissected in a random manner. A random dissection does not take into account domain structure. This could potentially result in misfolded proteins being expressed and thereby, causing regions of the FAM protein required for interactions to be overlooked. Success in these studies would have allowed a thorough investigation to be conducted into the binding sites on FAM for the known interactors β-catenin, epsin 1, and itch. This was essential to this project as the overall aim of these investigations was to generate FAM protein carrying mutations that specifically disrupted the interactions between FAM, and one of its binding partners. In order to do this, a thorough understanding of the way in which the entire protein physically interacts with all of its partners was required. *In vitro* binding studies were also utilised in this chapter to identify novel interactions of the UBL domain of FAM.

Yeast and Mammalian Two-Hybrid Analysis

Only one interaction of FAM has been identified/explored using two-hybrid analysis, that of the doublecortin-FAM interaction (Friocourt et al, 2001). This interaction was identified using the traditional form of yeast two-hybrid analysis, namely the screening of a library of cDNAs and cDNA fragments using a known cDNA (in this case the cDNA for DCX). It is likely that few FAM interactions have been discovered using this technique as a result of the large size of the FAM cDNA. cDNA libraries more often than not contain cDNA fragments of not more than 1-2kb in length due to the nature of the library generation protocols. Since the cDNA encoding FAM is in excess of 9kb, any cDNAs of FAM in commonly available libraries are likely to contain only fragments of the full-length cDNA and these fragments are not necessarily going to encode for an entire domain of FAM. As a result, many interactions will be missed in a library screen due to incorrectly folded or missing FAM domains.

In contrast, rather than attempting to identify novel interactions, two-hybrid analysis was used in this study to further explore known FAM interactions in an effort to refine the sites of these interactions on the FAM protein. This approach was also taken to address the issue of difficulties in expressing FAM fragments in bacterial or *in vitro* transcription/translation systems that do not guarantee correctly folded proteins will be generated. These investigations proved inconclusive with even previously demonstrated interactions of the catalytic domain of FAM with β-catenin and epsin 1 not returning positive results. There are a number of reasons why this may have been the case. Investigations of FAM cat interactions in the literature have reported a relatively low binding strength with an estimated K_d of 1 μM with β-catenin and 890 μM with AF-6 (Taya et al, 1999). In addition, the interaction is likely to be an enzyme substrate interaction and thereby transient in nature. It is possible therefore, that the interaction in yeast cells is not sufficiently strong or long enough to promote transcription of the reporter genes used in the two-hybrid assay. It is also possible that the catalytic core of FAM only recognises its substrates (in this case β-catenin or epsin 1) when they are ubiquitylated, a modification that may not occur when these proteins are expressed in yeast cells.

It remains the authors' opinion that two-hybrid analyses was and will be a powerful tool for studies examining the interactions of FAM. The technique is both sensitive and rapid (in the case of mammalian two-hybrid analysis). Both the yeast and the mammalian two-hybrid analyses are conducted in eukaryotic cells. This provides the full range of post-translational protein modification machinery and eukaryotic chaperone molecules that are likely to be required for the correct and efficient folding of the large FAM protein and its fragments. In addition, any potential interactions are allowed to occur in conditions that more closely resemble the *in vivo* environment in which such interactions would be likely to take place.

Co Immunoprecipitation

Co immunoprecipitation studies have often been used as a first step in identifying unknown interactors of a known protein. Many of the interacting partners of FAM were initially identified using this technique including AF-6 (Taya et al, 1998; Kanai-Azuma and Wood, Unpublished Data), β-catenin (Taya et al, 1999), and Epsin 1 (Chen et al, 2001; Chen et al, 2003). While co immunoprecipitation experiments provide some insight into the interaction *in vivo* they have the major limitation that they are not able to distinguish a direct interaction

between two proteins from an interaction mediated by a third party (protein or complex) (Clegg, 1998). For this reason, they are often used in conjunction with other techniques that test the interaction of the two proteins in isolation, as was the case with all of the examples cited above. Co immunoprecipitation was not used in this study to a large extent as much of the work was aimed at further refining the sites of known interactions requiring confidence that the interactions observed were direct interactions. Co immunoprecipitation was primarily used in this work in an attempt at secondary confirmation of an interaction between FAM and the GGA proteins.

Alternative Strategies to Investigate FAM Interactions:

Eukaryotic Linear Motif and Other Current Predicted Structural Features of FAM

The eukaryotic linear motif (ELM) server is a computational protein sequence analysis tool that searches a protein sequence for functional sites defined by short non-globular sequence motifs. Due to the short sequence nature of linear motifs and the fact that these functional sites are defined by a variable "consensus motif ", it is difficult to assign a significance value to any sites that are found in a protein sequence. For this reason, analysis with this tool is combined with contextual information about the protein to strengthen assessment of the likelihood that the identified sequences are functional. Contextual information can include subcellular localisation, clashes with predicted globular domains (by comparison with pfam and smart databases), and taxonomic data. It is also important to note that the likelihood of these short motif sequences appearing in a protein sequence increases with the size of the protein such that large proteins are more likely to have a higher abundance of hits. With this in mind, when the protein sequence of FAM was analysed on the ELM server, a number of interesting ELMs arose. Contextual information that was supplied with the analysis consisted of the subcellular localisation filters for the presence of FAM at the plasma membrane, the Golgi apparatus, the lysosome, endosomes, and in the cytoplasm. The taxonomic filter of *Mus Musculus* was also applied.

Four sequences of particular interest were identified that consisted of two clathrin-box motifs, one AP-2 α subunit binding motif, and a calmodulin binding motif. The clathrin-box motif interacts with the β-propeller structure found at the N-terminus of the clathrin heavy chain and is generally found on proteins that can be classified into four groups. The first group consists of cargo adaptor proteins and those involved in the initiation of coated vesicle

formation. The second group includes proteins that are involved in interactions with the actin cytoskeleton. The third and fourth groups include proteins that are involved in vesicle budding and disassembly of the clathrin coat. The clathrin box consensus motif is defined as L [IVLMF] x [IVLMF] [DE] where any of the residues indicated in the square brackets can reside at this position and where X represents any amino acid. Clathrin box motifs are found in a large number of proteins including the three AP complexes, the amphiphysins, the β-arrestins, the epsins and a number of other proteins involved in protein trafficking (Dell'Angelica, 2001). Often the consensus site is located in an unstructured or flexible hinge-like region of the protein (Gaidarov et al., 2001; Kirchhausen, 1999; Lafer, 2002). There are many more proteins that contain sequences that either fit or loosely fit the consensus motif but that have no defined roles in clathrin-dependent trafficking. There are also sequences matching the consensus motif in proteins that are known to bind to clathrin that are not functional clathrin binding sequences (Dell'Angelica, 2001). For this and other reasons, the predictive power of the consensus motif has been called into question. It is suggested therefore, that when investigating a novel clathrin-box motif, contextual information about the known localisation and function of the protein, and the positioning of the motif within the protein itself, should first be considered (Dell'Angelica, 2001). The two sequences identified in FAM, LIGLD at amino acids 768-772 and LVSLD at amino acids 2207-2211, fit well within the defined consensus motif and are located in regions of the protein that are predicted to be unstructured and contain no other recognised protein domains (Ginalski et al., 2004; Puntervoll et al., 2003).

The second motif of interest that was identified was an AP-2 α-subunit binding motif. These motifs are responsible for recruiting accessory endocytic proteins at the plasma membrane to sites of clathrin coated vesicle formation. The consensus motif for AP-2 α-subunit binding motifs is FxDxF where F represents phenylalanine, D represents aspartic acid, and x represents any amino acid (Brett et al., 2002). The AP-2 α-subunit binding motif that was identified in FAM lies towards the C-terminal end of the catalytic core (amino acids 1812-1816) and consists of the residues F^{1812}, N^{1813}, D^{1814}, Y^{1815}, and F^{1816}. Given FAM's subcellular localisation patterns, and its interaction with endocytic proteins like epsin1 and itch, this motif has a high likelihood of being functional.

The third motif that was identified was the IQ, or calmodulin binding motif. This motif is again a protein-protein interaction motif and allows interaction of the protein with members of the calmodulin family (Bahler and Rhoads, 2002; Rhoads and Friedberg, 1997). These interactions can either be Ca^{2+} dependent or independent. The motif begins with a

mostly invariable isoleucine (I) residue followed by an invariable glutamine residue (Q) which is where the motif derives its name. The rest of the motif proceeds as follows IQxxx[RK]xxxx(x)[RKQ]. The sequence that was identified in FAM was IQWIDRFIEELRTN and spans residues 543-562. The presence of a calcium binding protein interaction site on FAM is of interest as there is some evidence in the literature that calcium signalling is able to modulate FAM function. In untreated HeLa cells epsin 1 can be found in an ubiquitylated form (Chen et al., 2003). When calcium influx is stimulated in these cells by treatment with ionomycin, a general decrease in ubiquitylated proteins including epsin 1 is observed (Chen et al., 2003). This decrease in ubiquitylated proteins is not hindered by treatment of the cells with a proteasome inhibitor which suggests that the reduction in ubiquitylated proteins is caused by large scale deubiquitylation by DUBs (Chen et al., 2003). When these same cells are first transfected with siRNA constructs directed against FAM, the general deubiquitylation of proteins still occurs with calcium influx, however the deubiquitylation of epsin 1 is ablated (Chen et al., 2003). Taken together, this information suggests that calcium influx serves as a stimulus for FAM to deubiquitylate epsin 1. Furthermore, FAM is known to be phosphorylated (personal communication, R. Murray, Child Health Research Institute, WCH, North Adelaide, SA) and has been observed to run as a doublet on well resolved SDS polyacrylamide gels (Chen et al., 2003). This doublet resolves into a single band when cells are treated with ionomycin to produce a calcium influx, suggesting that calcium signalling may induce dephosphorylation of FAM. It is possible that calcium influx causes calcium dependent association of calmodulin or a calmodulin-like protein with FAM. The existence of this interaction and its physiological role are just speculation at this point but it may be that calcium stimulated calmodulin binding to FAM causes a conformational change in FAM that allows it to become active. This may involve a dephosphorylation event, exposure of other protein interaction domains of FAM, a change in its subcellular localisation, recruitment of FAM modifying or coactivator proteins via calmodulin, or a myriad of other possibilities.

Investigations into the functionality of these four motifs could provide valuable information about the cellular functions of FAM. It would first be necessary to identify whether these motifs are functional in FAM by determining if they are able to interact in isolation with their respective targets; clathrin box motifs bind to clathrin, AP-2 alpha binding motifs interact with the alpha subunit of AP-2, and the IQ motif binds to calmodulin and calmodulin-like proteins. Any of the assays for protein-protein interactions described above would be suitable for investigating the interaction of the two predicted clathrin-box motifs and the predicted AP-2 alpha binding motif. As the IQ motif can bind multiple members of

the calmodulin family, an investigation of this predicted motif in FAM would require a broader protein interaction assay that was capable of identifying novel interactions. A yeast two-hybrid library screen, co-immunoprecipitation, or a GST affinity binding experiment coupled with mass spectrometry would be more suited to investigating the IQ motif. Since many of these motifs are quite small and often rely upon one or two critical residues for their function, it would be a relatively simple matter to then mutate the critical residue and investigate whether this generates a non functional binding mutant. It would then be necessary to determine whether these motifs are functional in the context of the full length FAM protein. It would also be necessary to determine whether the binding mutants are still defective in binding their respective targets when introduced into the full length protein. Investigations could then be carried out on the effects of these binding mutants, in isolation or in combination, on the cellular localisation and/or function of FAM. These experiments would benefit greatly from $FAM^{-/-}$ cell lines, or highly efficient RNA interference constructs, as the ability of the binding mutants to functionally replace wild-type FAM could be more accurately judged in these circumstances. Functional investigations could take the form of transferrin and/or EGFR uptake in the case of the clathrin-box and AP-2 alpha binding motifs as these are most likely to have an effect on FAM on the endocytic pathway. Internalisation and/or stability of the E-cadherin/β-catenin complex, AF-6, and the stability of Itch and epsin 1 could also be investigated. The calcium induced changes in FAM function and phosphorylation status could be investigated in the case of the IQ motif mutant.

The ENSEMBL genome browser automated annotation server (release 42) also lists two previously unreported predicted globular domains for murine FAM (Curwen et al., 2004; Hubbard et al., 2007). These domains were predicted using the structural classification of proteins (SCOP) database (Murzin et al., 1995). One of the predicted domains consists of an armadillo (Arm) repeat region spanning amino acids 279-582 of FAM. Assuming a repeat length of ~ 40 amino acids, this would suggest that the region contains around 7 repeats. Arm repeat regions are well known as protein-protein interaction domains and can be found in a large number of proteins involved in a diverse range of cellular processes (Andrade et al., 2001). The 3-dimensional structure of the armadillo repeat region has been solved to a 2.1 Å resolution (Huber et al., 1997). The 12 arm repeats form a single continuous structural domain characterized by an extended hydrophobic core. Each individual arm repeat consists of three helices designated H1, H2, and H3 (Huber et al., 1997). The 12 repeats stack on top of each other with the short H3 helix sandwiched between each one. The H3 helices form a parallel row on one side of the structure with the H1 helices forming a similar row on the other side of the molecule. An average 30° rotation per arm repeat creates a regular right-handed

superhelix of helices with the overall effect being a spiraling two-sided ribbon (Huber et al., 1997; Shapiro, 2001). On the H3 side of the structure, a long shallow groove is created between the first 10 arm repeats that is ~ 95 Å long and 20 Å wide. This groove is rich in basic amino acid residues and forms the basis of charged based interactions of β-catenin with APC, Tcf/Lef transcription factors, and E-cadherin (Huber et al., 1997; Shapiro, 2001). Once again, the ability of the predicted arm repeat region of FAM to act as a functional protein-protein interaction domain could be investigated using a yeast two-hybrid library screen, co-immunoprecipitation experiments, or GST-affinity binding assays to determine which, if any, proteins interact with FAM through this domain. It would also be interesting to attempt crystallization of this domain to determine if it does indeed take on the structure of an arm repeat region.

The second globular domain predicted by SCOP and displayed on the ESEMBL website spans residues 2336-2466 and is a predicted annexin fold domain. The annexin family has between 1-20 members in all eukaryotic organisms (primates have 12). Many members of the annexin family are known to bind to negatively charged membrane phospholipids in a calcium dependent manner however there are members of the family that show negligible calcium binding capacity (Moss and Morgan, 2004). The cellular function of many of the annexins is as yet still unclear but several have clearly defined functions. Annexin A1 is involved in the inflammatory process, annexin A2 is implicated in vesicular traffic, and annexin A7 has a demonstrated role in the regulation of cell growth (Moss and Morgan, 2004). Crystal structures of many of the annexins have been published and from these structures it has been found that the general structure of the annexin core domain is relatively well conserved (Benz et al., 1996; Burger et al., 1996; Favier-Perron et al., 1996; Lewit-Bentley et al., 1992; Luecke et al., 1995; Sopkova et al., 1993; Weng et al., 1993). The core annexin folding domain is made up of four repeating units labeled from I to IV. Each unit (~ 70 amino acids) is constructed from 5 alpha helical structures (labeled from A to E) arranged such that they form less than two turns of a right-handed superhelix (Liemann and Huber, 1997). Four of the five helices are organized in an almost (anti)-parallel configuration with the fifth lying roughly perpendicular to the others. The four units are oriented in a relatively flat and roughly circular shape with a small curve producing one slightly convex and one slightly concave shape. Calcium binding occurs on the convex face of the molecule with each of the four units capable of binding calcium in an “endonexin fold” situated between helices A and B (Liemann and Huber, 1997).

The relatively small nature of the annexin fold domain identified in FAM (130 amino acids) suggests that if this is a functional domain in FAM, it most likely only constitutes a single five helix unit rather than the entire annexin fold domain. It is possible that the function of the domain in this context is related to the calcium sensitivity demonstrated by FAM. Perhaps the domain undergoes a conformational change upon calcium signaling within the cell which activates FAM, exposes a hidden functional site, or promotes the interaction of FAM with a substrate or accessory protein. The study of this domain structure within FAM could be undertaken in a number of ways. X-ray crystallography of the isolated domain would determine whether it is an independently folding unit and whether it requires the presence of calcium for stability. Yeast two-hybrid library screens, co-immunoprecipitation, and/or GST-affinity binding assays would determine if it was a protein-protein interaction domain of FAM. It is unlikely that this domain would retain the phospholipid binding activity of annexin due to the absence of the three remaining units, but this possibility could also be tested by chemical cross-linking experiments and electron microscopy of the domain co-crystallized with lipids. Mutation of the critical residues within the calcium “endonexin fold” loop could also be performed and the effects on normal FAM cellular function and response to calcium signaling could be examined.

CHAPTER 4: INVESTIGATION OF THE ROLE OF FAM IN LYSOSOMAL PROTEIN TRAFFICKING

CHAPTER SUMMARY

The role of ubiquitin in protein trafficking and sorting on the endocytic route is now well established. Much less is known however about the regulation of the biosynthetic route by ubiquitin. In yeast, the correct trafficking of at least four proteins (Csp1p, Phm5p, Hmx1p, and Gap1p) from the trans-Golgi network (TGN) has some requirement for their ubiquitylation (Helliwell et al., 2001; Reggiori and Pelham, 2001). This ubiquitylation is likely to occur somewhere between leaving the TGN and arriving at endosomal/lysosomal system, and is probably mediated by the action of the Hect domain containing E3 ligase Rsp5p (Helliwell et al., 2001; Katzmann et al., 2001). Little is known about the role of ubiquitin and ubiquitin ligases in trafficking from the mammalian Golgi apparatus, although the two closest mammalian homologues of Rsp5p, Nedd4 and Itch/AIP4 have been observed to localise to the Golgi apparatus (Angers et al., 2004; Shirk et al., 2005). Both mammalian homologues are also known to be involved in the endocytosis and sorting of a number of plasma membrane proteins including Notch and ENaC (Qiu et al., 2000; Staub et al., 1997; Wilkin et al., 2004). Furthermore, GGA1 and GGA3, two proteins that are intimately involved in packaging and sorting of proteins at the TGN, are both known to bind to ubiquitin and are themselves ubiquitylated (Shiba et al., 2004; Yogosawa et al., 2006). Thus there is a growing body of evidence to suggest that ubiquitin plays an important role in sorting of cargo and regulation of trafficking machinery at the Golgi apparatus.

Less still is known about the role of deubiquitylating enzymes in protein trafficking in general let alone from the TGN. At present, only three deubiquitylating enzymes have known associations with protein trafficking, yeast Ubp4/Doa4 and its mammalian homologue Usp8/UBPY, AMSH, and FAM (Amerik et al., 2000; McCullough et al., 2004; Murray et al., 2004). Doa4 is thought to be primarily responsible for recovering conjugated ubiquitin from endosomal cargo that are destined for the lumen of the yeast vacuole (Amerik et al., 2000; Losko et al., 2001). Doa4 is recruited to the invaginating vesicles of maturing multi-vesicular bodies by the ESCRT-III complex (Amerik et al., 2000; Losko et al., 2001). Its mammalian counterpart, Usp8/UBPY, is known to interact with Hrs-binding protein (Hbp/STAM2). Hrs and Hbp are the mammalian equivalents of Vps27 and Hse1 respectively, the ESCRT-III complex components that bind to Doa4 (Kato et al., 2000). AMSH is a deubiquitylating enzyme of the JAMM family and is also known to bind to Hbp/STAM (Tanaka et al., 1999). AMSH has been shown to deubiquitylate EGFR and loss of the enzyme leads to an increase in degradation of the internalised EGF receptor (McCullough et al., 2004).

Of the three DUBs that have been associated with protein trafficking, FAM is the only one to have shown significant localisation to the Golgi apparatus (Murray et al., 2004). The potential interaction between FAM and the GGA proteins described in chapter 3, and the observation that FAM staining is altered in fibroblasts taken from patients with lysosomal storage diseases (unpublished data), suggests that FAM is involved in the sorting processes that occur at the TGN. In addition, FAM is associated with trafficking of the E-cadherin/β-catenin complex (Murray et al., 2004). For these reasons, an investigation into the role of FAM in protein trafficking and sorting at the TGN was carried out. FAM has the potential to affect trafficking of proteins at the TGN both in a CIS manner; by directly interacting with and/or deubiquitylating the proteins that are sorted, or in a TRANS manner; by interacting with and/or deubiquitylating the TGN sorting machinery (Hicke and Dunn, 2003). In this investigation, expression of exogenous FAM and reduction of endogenous FAM levels by shRNA were used to observe the effect of altering FAM levels on trafficking and processing of eleven lysosomal proteins. The eleven lysosomal proteins studied were, acid sphingomyelinase (ASA), N-sulphatase/sulphamidase (N-Sulph), α-L-Iduronidase (IdUA), acid-α glucosidase (GAA), β-glucoronidase (β-Gluc), α-galactosidase (α-Gal), aryl sulphatase A (ASA), Iduronate-2-sulphatase (I2S), 4-sulphatase (4S), Lamp1, and Saposin-C - one of two products of prosaposin (Sap-C). All of these proteins, with the exception of Lamp1 and Sap C, are trafficked to the lysosome from the TGN via the mannose-6-phospate receptors and the GGA proteins, and hence can have no physical interaction with FAM (Lobel et al., 1989; Nielsen et al., 2001). Lamp1 can reach the lysosome both directly from the Golgi where it has been seen to enter AP-3 coated vesicles, and indirectly by first being directed to the cell surface and then internalised through the endo-lysosomal system (Akasaki et al., 1995; Le Borgne et al., 1998). Targeting of Lamp1 occurs through a tyrosine based sorting mechanism. Sap C has also been reported to reach the lysosome by two routes. On one route, the protein is initially transported to the cell surface before being redirected back to the lysosome and is targeted by a lipid raft sorting mechanism (Hiesberger et al., 1998). Alternatively, Sap C can be sorted directly to the lysosome through interaction with sortilin, a transmembrane Golgi protein that may act in conjunction with the GGA proteins (Lefrancois et al., 2003; Ni and Morales, 2006; Nielsen et al., 2001; Takatsu et al., 2001). As discussed in chapter 1, all lysosomal proteins undergo a high degree of post-translational modification that starts in the rough endoplasmic reticulum and continues through to their packaging and transport at the TGN (Dittmer and von Figura, 1999; Hasilik et al., 1980; Lazzarino and Gabel, 1988; Rohrer and Kornfeld, 2001; Tabas and Kornfeld, 1980; Varki and Kornfeld, 1980; Varki and Kornfeld, 1981).

It was postulated that if FAM was involved in trafficking from the Golgi apparatus, changes in FAM levels may affect both the maturation/processing and overall cellular levels of some or all of these lysosomal proteins. In addition, of the eleven lysosomal proteins that were examined, all but Lamp1 and Sap C are luminal proteins. Thus FAM, which is a cytoplasmic protein, cannot interact directly with these proteins. It was therefore hypothesised that any observed effects on the trafficking or processing of these proteins by an alteration in FAM levels must be caused indirectly or in a TRANS manner. For this reason, three cellular and molecular biological techniques were applied to determine the effect of modulating FAM levels on traffic of lysosomal proteins. An ELISA based analysis was applied to simultaneously measure the secreted and total cellular levels of all eleven lysosomal proteins (Meikle et al., 2006). From this work, it became evident that the levels of three of the lysosomal proteins, Lamp1, ASA, and 4S, were significantly altered upon expression of exogenous FAM with the opposite effect seen upon knock-down of endogenous FAM levels.

As the ELISA based assay measured only the total cellular lysosomal protein levels, no distinction could be made using this method, between unprocessed immature protein and functional mature protein. The extensive post-translational modification of lysosomal proteins results in the presence of multiple molecular weight species of the proteins which appear as a banded pattern upon western analysis. Often the mature forms of lysosomal proteins resolve upon protein electrophoresis at lower molecular weights due to editing of the sugar chains and proteolytic cleavage events. Therefore, western analysis was performed probing for Lamp1, ASA, 4S, and a fourth protein, I2S, whose levels did not appear to be affected, with the aim of determining if altering FAM protein levels had an effect on the processing of these proteins. Immunofluorescence studies were also conducted with the aim of identifying site/s of potential blockage of trafficking of the lysosomal proteins and to determine if there were any gross changes in their localisation and distribution when FAM levels were altered.

RESULTS

Effect of Altered FAM Levels on Lysosomal Proteins – Multiplex ELISA Analysis

FAM Overexpression

An initial transient transfection was carried out to determine the optimal amount of FAM expression vector to transfect and the optimal incubation period required to achieve maximal expression of exogenous protein. HEK 293T cells were transfected with either 10μg of pDEST40-V5 or 10, 15, or 20μg of pDEST40-V5-FAM vector and incubated for 48, 72, or 96 hours (Figure 4.1). The quantity of DNA transfected did not appear to have any affect on the level of expression of exogenous FAM as determined by immunoblot with anti-V5 antibodies (Figure 4.1). Optimal expression of exogenous FAM was seen 96 hours after transfection (Figure 4.1). Three independent transfections of 293T cells were carried out using 10μg of either pDEST40-V5 or pDEST40-V5-FAM. Cells were incubated for 96 hours before being harvested and processed for use in the multiplex ELISA system. Western analysis detected the V5-epitope in all three transfections confirming expression of exogenous FAM (Figure 4.2a).

Total Cellular Protein Levels

Statistical analysis (student's T-test) identified significant reductions in total cellular levels of protein for five of the eleven lysosomal proteins in cells overexpressing FAM. Protein levels of Heparin N-sulphatase (N-Sulph), Lamp1, acid α glucosidase (GAA), aryl sulphatase A (ASA), and N-acetylgalactosamine-4-sulphatase were 25%, 15%, 16%, 19%, and 21% lower respectively, in FAM transfected cells compared with cells transfected with empty vector ($p<0.05$ N-sulph, $p<0.01$ for the remaining, Figure 4.2 b). Iduronidase (IdUA) and Iduronate-2-sulphatase (I2S) also showed reductions however these were not significant. In contrast, a statistically significant increase was observed for α-galactosidase A (A-gal) with total protein levels that were 35% higher in FAM transfected cells ($p<0.01$, Figure 4.2 b)

Secreted Protein Levels

As discussed in chapter one, many lysosomal proteins utilize two routes to reach the lysosome. All eleven proteins used in this ELISA assay enter the lysosome both via the direct route from the Golgi to the endo/lysosomal system and by being trafficked to the cell surface, secreted, and re-internalised via the endocytic route. It is possible that FAM may be involved in one or both pathways. For this reason, secreted lysosomal protein levels were also

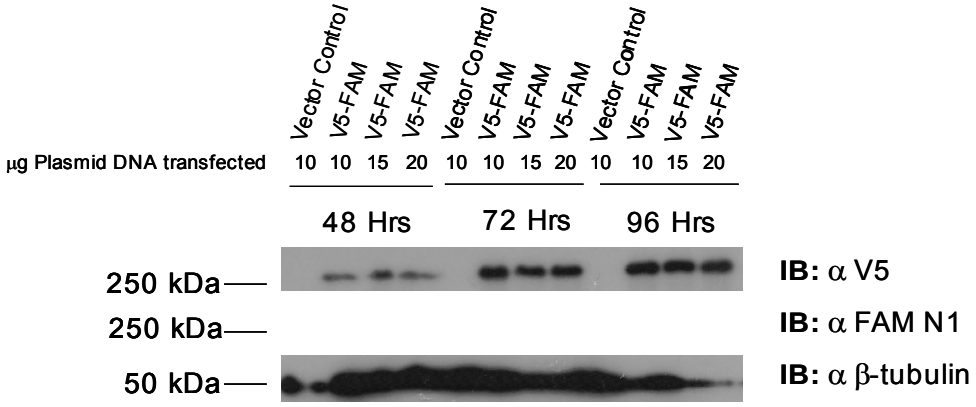
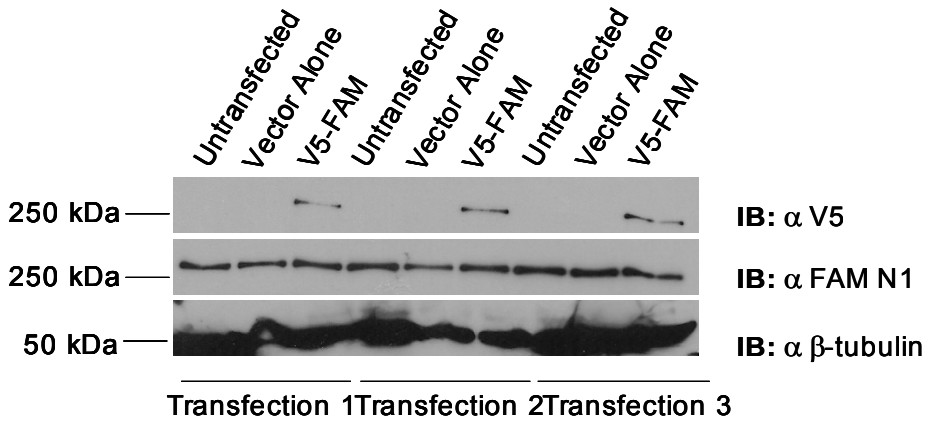


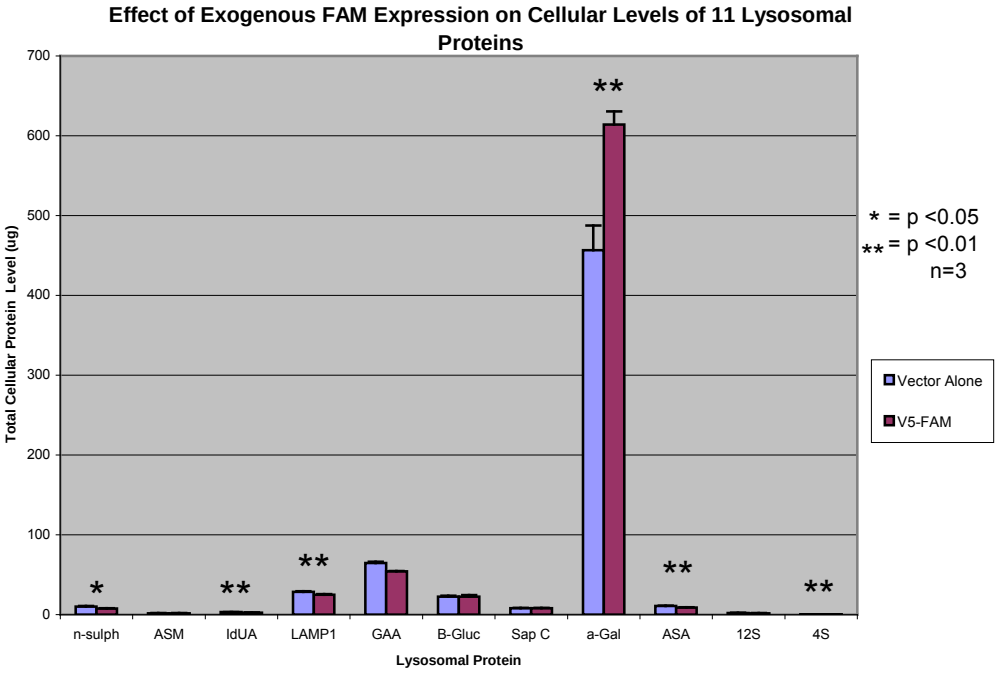
Figure 4.1: ***Optimisation of transient transfection conditions for exogenous FAM expression.*** 293T cells were either mock transfected (10ug of empty vector) or transiently transfected with 10, 15, or 20 µg of pDest40 V5 FAM plasmid and incubated for 48, 72, or 96 hours. Cells were harvested, lysed, and analysed by western blot with anti-V5, anti-FAM, and anti-β-tubulin antibodies. Anti-V5 antibodies showed optimal expression of exogenous FAM occurred at 96 hours post transfection.

Figure 4.2: ***Investigating the effect of transient exogenous expression of FAM on the total cellular and secreted protein levels of eleven lysosomal proteins.*** **A)** Three independent transient transfections of HEK 293T cells were carried out with either 10 μg of pDest40 (vector alone) or pDest40 FAM expression vector (V5-FAM). Mock transfected cells were also included. Cells were incubated for 96 hours before being harvested, lysed, and analysed by western blot with anti-V5, anti-FAM, and anti-β-tubulin antibodies. Western analysis with an Anti-V5 antibody indicated exogenous FAM was expressed. Immunoblot with anti-FAM N1 antibodies did not indicate an overall increase in FAM levels. Immunoblot with anti-β-tubulin antibodies indicated similar loading of all lanes. **B)** Total cellular protein concentration of eleven lysosomal proteins in empty vector transfected cells (blue) and V5 epitope tagged FAM expressing cells (red) (plotted values are the mean ± SEM (n = 3)). **C)** Percentage of secreted lysosomal protein in V5-FAM expressing cells normalised to vector control transfected cells (plotted values are the mean ± SEM (n = 3)). * = $p < 0.05$, ** = $p < 0.01$

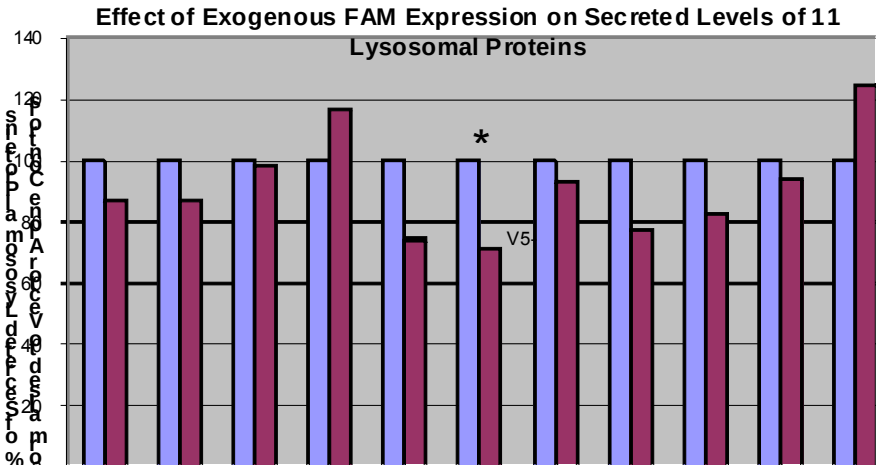
A)


Untransfected
Vector Alone
V5-FAM
Untransfected
Vector Alone
V5-FAM
Untransfected
Vector Alone
V5-FAM
250 kDa
250 kDa
50 kDa
IB: α V5
IB: α FAM N1
IB: α β-tubulin
Transfection 1
Transfection 2
Transfection 3

B)


Effect of Exogenous FAM Expression on Cellular Levels of 11 Lysosomal Proteins
Total Cellular Protein Level (ug)
700
600
500
400
300
200
100
0
n-sulph
ASM
IdUA
LAMP1
GAA
B-Gluc
Sap C
a-Gal
ASA
12S
4S
Lysosomal Protein
* = p <0.05
** = p <0.01
n=3
Vector Alone
V5-FAM

C)


Effect of Exogenous FAM Expression on Secreted Levels of 11 Lysosomal Proteins
140
120
100
80
60
40
20
*
V5-

determined by collecting the media in which the cells were incubated. It should be noted that the media was changed 16 hours post transfection and hence any accumulation or depletion of lysosomal proteins in the media occurred over a shorter, 80 hour time course. N-sulph, ASM, IdUA, GAA, β-gluc, Sap-C, α-Gal, ASA, and I2S all showed lower protein levels in the media of FAM overexpressing cells compared to cells transfected with empty vector. However, the only statistically significant result was that of β-gluc which had 29% lower levels in FAM transfected cells ($p<0.05$, Figure 4.2 c).

Decreased FAM Levels – FAM Knock-down

To determine the effect of a reduction in cellular levels of FAM protein three independent shRNA triggers were transfected into 293T cells with expression driven by a retroviral vector, pRetro Super (pRS) (a kind gift from René Bernards, The Netherlands Cancer Institute, Division of Molecular Carcinogenesis, The Netherlands) (Brummelkamp et al., 2002; Brummelkamp et al., 2002). The three triggers designated pRS-1800, pRS-3601, and pRS-USP9X, were directed against the sequence for the human homologue of FAM, USP9X, and have previously been used to successfully demonstrate a significant knock-down of FAM protein levels in HEK 293T cells (Mouchantaf et al., 2006). The optimal incubation period post transfection to observe a knockdown of FAM levels has been reported to be 72 hours (Mouchantaf et al., 2006). A similar observation was made in an initial test transfection to determine the efficacy of the three triggers (Figure 4.3). It was decided however, that a series of time-points would be used for the experiment to gain a greater insight into the dynamics of any changes in lysosomal protein content as the levels of endogenous FAM protein decreased over time. For ELISA analysis, cells were transfected and then harvested 24, 48, 72, and 96 hours later. It was found that, in contrast to the first trial transfection, only the shRNA construct designated pRS 1800 was effective in reducing endogenous FAM levels (Figure 4.4). FAM levels began decreasing 48 hours post transfection with a maximum knockdown observed at 96 hours post transfection (Figure 4.4). Despite several attempts, the remaining two shRNA constructs yielded no obvious reduction in endogenous FAM levels (data not shown). In contrast pRS 1800 produced consistent levels of FAM knockdown in at least three experiments. Multiplex ELISA analysis was performed on cell lysate from cells transfected with the pRS 1800 FAM knockdown construct and vector control cells.

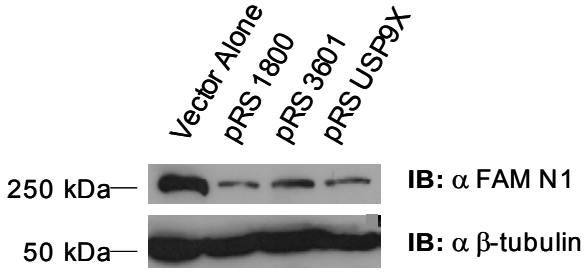


Figure 4.3: ***Determining the efficacy of shRNA triggers directed against FAM***. HEK 293T cells were transiently transfected with the base expression vector pRetro Super (pRS), or one of three expression vectors for three independent shRNA triggers directed against FAM (pRS1800, pRS3601, or pRSUSP9X). Cells were incubated for 96 hours levels of FAM were analysed by western blot At 96 hours post transfection, all three shRNA triggers were able to significantly reduce endogenous FAM protein levels. β-tubulin antibodies indicated equal loading of protein in each lane.

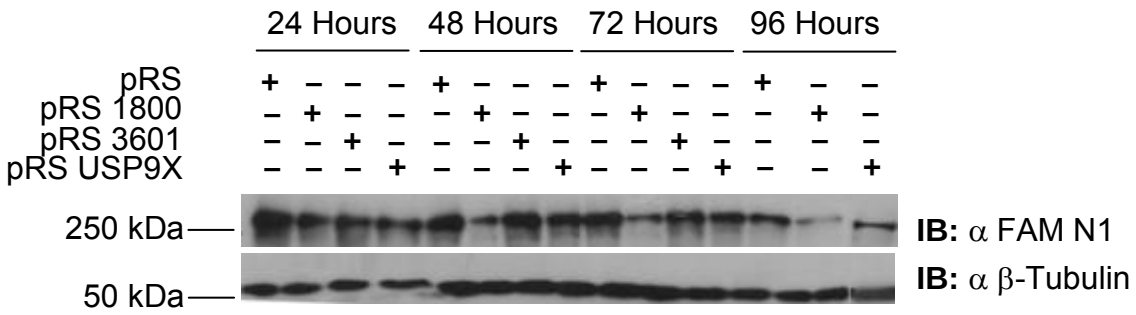


Figure 4.4: ***Time course of FAM knockdown***. HEK 293T cells were transiently transfected with the base expression vector pRetro Super (pRS), or one of three expression vectors for independent shRNA triggers directed against FAM (pRS1800, pRS3601, or pRSUSP9X). Cells were incubated for 24, 48, 72, or 96 hours post transfection, before analysis by western blot with anti-FAM antibodies. Anti-FAM antibody staining indicated a reduction in endogenous FAM protein levels at the 48, 72 and 96 hour time points for cells transfected with the 1800 shRNA trigger expressing construct. A slight reduction in endogenous FAM protein levels can be seen at the 48, 72 and 96 hour time points for cells transfected with the USP9X shRNA trigger expressing construct. No such reduction in endogenous FAM protein levels was observed for cells transfected with the 3601 shRNA trigger expressing construct. β-tubulin antibodies indicated equal loading of protein in each lane.

Total Cellular Protein Levels

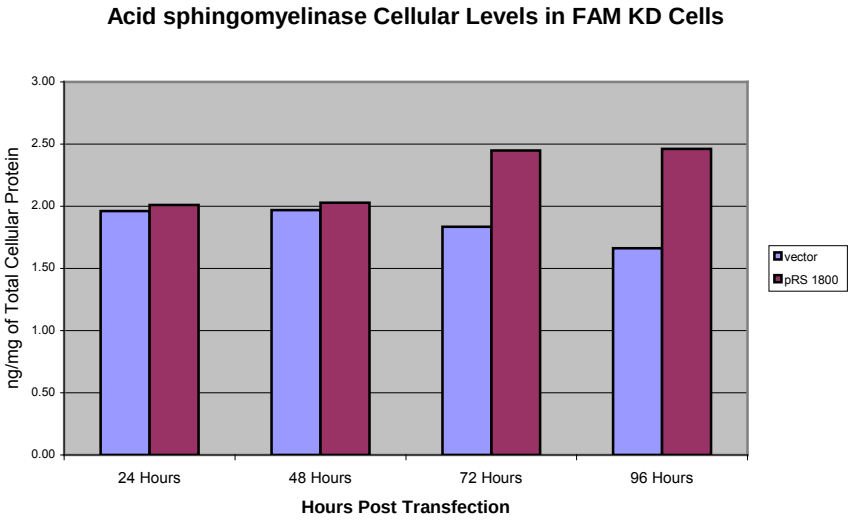
Trends in the changes in cellular protein levels over the time course of the experiment are described first for the following reasons. 1) To establish whether there were variations in the cellular levels of lysosomal proteins as cells became more confluent over the course of the experiment. 2) To add power to the analysis by comparing the "normal" changes over time in vector transfected cells to the changes over time observed with altered FAM levels. 3) To determine whether there were fluctuations in cellular protein levels from day to day that may have constituted noise in the system. In addition, to simplify the analysis it was assumed that since both FAM protein and the eleven lysosomal proteins are likely to have half-lives measured in days, any changes in lysosomal protein levels must have occurred in cells treated with FAM shRNA, after the existing cellular pool of FAM was degraded. This would suggest that any differences between cellular lysosomal protein levels in vector control cells and FAM shRNA treated cells observed at the early 24 and 48 hour time points, must be a result of differences existing in the starting populations of cells or due to noise in the system.

While it might have been expected that total cellular levels of each of the eleven lysosomal proteins would remain constant over the time course with each cell having a prescribed amount of each protein, this did not appear to be the case for all of the proteins. This was evidenced when the "normal" changes over time in vector control cells were examined. Cellular levels of several of the proteins, N-sulph, IdUA, GAA, Sap C, and ASA, appeared to increase over the course of the experiment (Figure 4.5 a, c, e, g, and i). In contrast, total cellular levels of ASM, Lamp-1, β-gluc, and α-Gal appeared to decrease over time (Figure 4.5 b, d, f, and h). Protein levels of the two remaining proteins, I2S and 4S, fluctuated over the time course but remained relatively stable (Figure 4.5 i and j). To aid in the analysis of the data it was assumed that the fluctuations seen in the relatively stable levels of I2S and 4S, and that any changes in protein levels that oppose the general increasing or decreasing trends of the remaining proteins were due to noise in the system. Fluctuations in I2S and 4S protein levels, and trend opposing changes in the remaining proteins, never exceeded a level of 10% difference. Therefore, based upon the assumptions described above, it was further assumed that any differences between protein levels in FAM knockdown cells and vector control cells that were less than 10% were unlikely to be significant. A greater than 10% difference was also chosen for a significance cut off as changes in protein levels that were found to be significant in the FAM overexpression study also exceeded 10%.

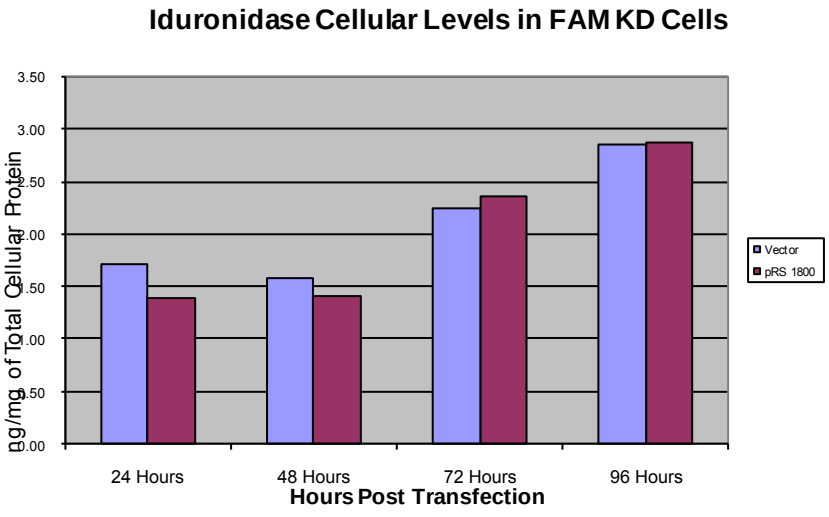
Figure 4.5: ***Effect of reducing endogenous FAM levels on total cellular protein levels of eleven lysosomal proteins***. HEK 293T cells were transiently transfected with the base expression vector pRetro Super (pRS), or the expression vector for the 1800 shRNA trigger directed against FAM (pRS1800). Cells were incubated for 24, 48, 72, or 96 hours. **A)** to **K)** Total cellular protein levels (ng/mg of total cellular protein) of N-sulphatase **A)**, Acid sphingomyelinase **B)**, Iduronidase **C)**, Lamp-1 **D)**, Acid α-glucosidase **E)**, β-glucosidase **F)**, Saposin C **G)**, α-Galactosidase **H)**, Aryl sulphatase A **I)**, Iduronate-2-sulphatase **J)**, and 4-sulphatase **K)** in vector control (blue) or FAM shRNA trigger 1800 (red) expressing cells at 24, 48, 72, and 96 hours post transfection. **L)** Percentage difference in total cellular protein levels of eleven lysosomal proteins in cells with reduced FAM protein levels normalized to vector control treated cells at 24 (blue), 48 (red), 72 (yellow) and 96 (green) hours post transfection.

A)

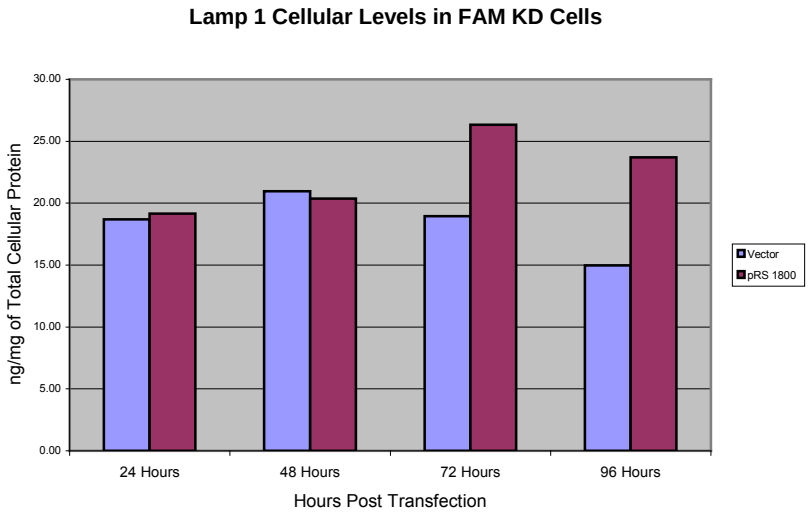
B)


Acid sphingomyelinase Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
3.00
2.50
2.00
1.50
1.00
0.50
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
vector
pRS 1800

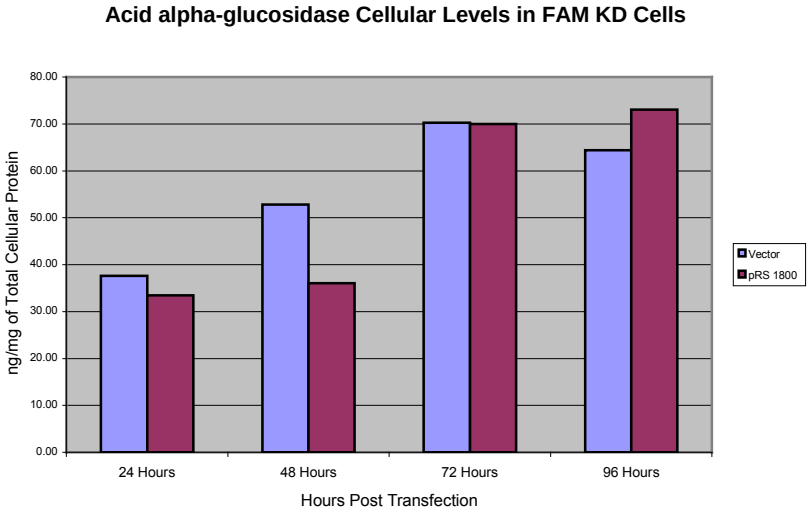
C)

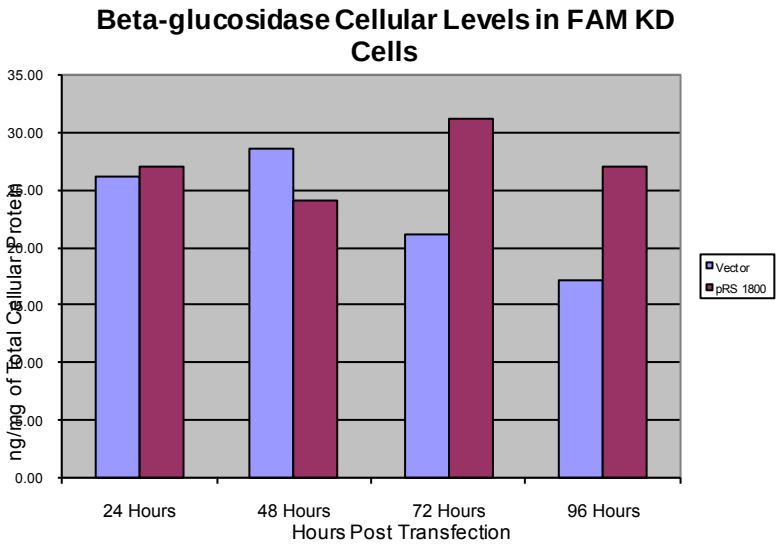

Iduronidase Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

D)


Lamp 1 Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
30.00
25.00
20.00
15.00
10.00
5.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

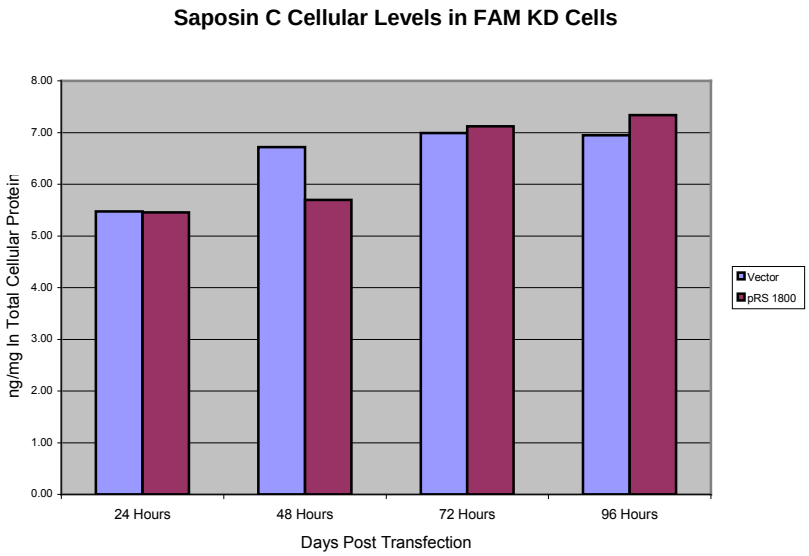
E)


Acid alpha-glucosidase Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

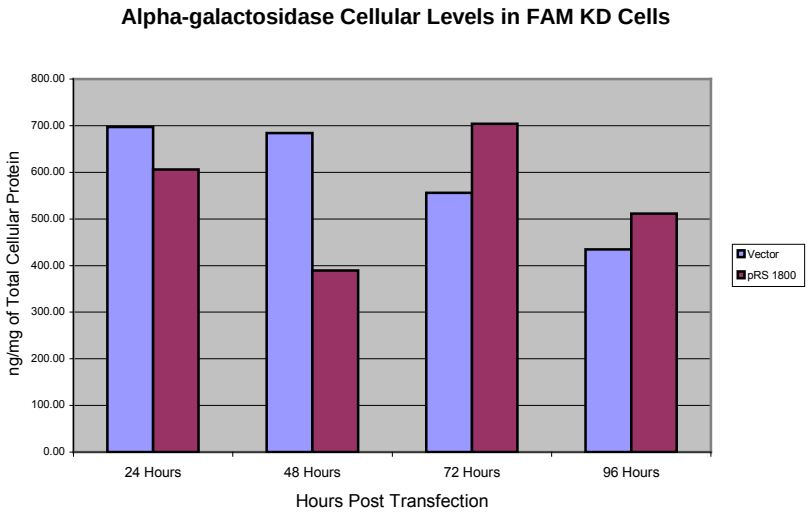

Beta-glucosidase Cellular Levels in FAM KD Cells
35.00
30.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

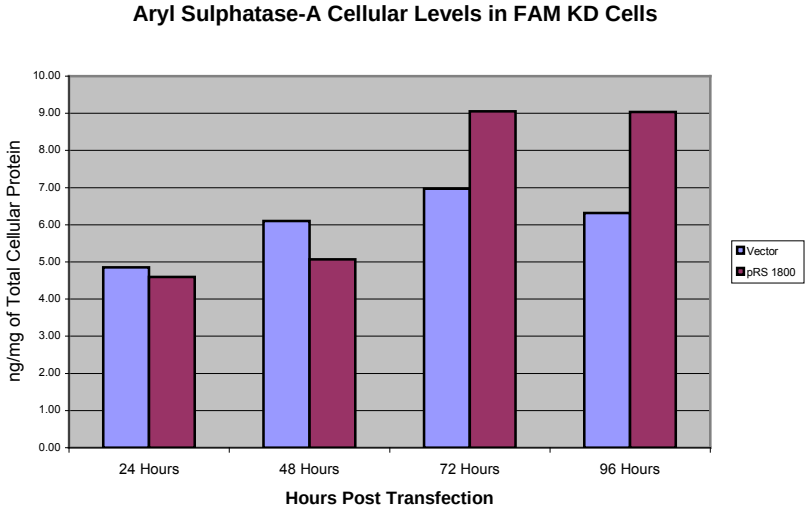
F)

G)


Saposin C Cellular Levels in FAM KD Cells
ng/mg In Total Cellular Protein
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Days Post Transfection
Vector
pRS 1800

H)


Alpha-galactosidase Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
800.00
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

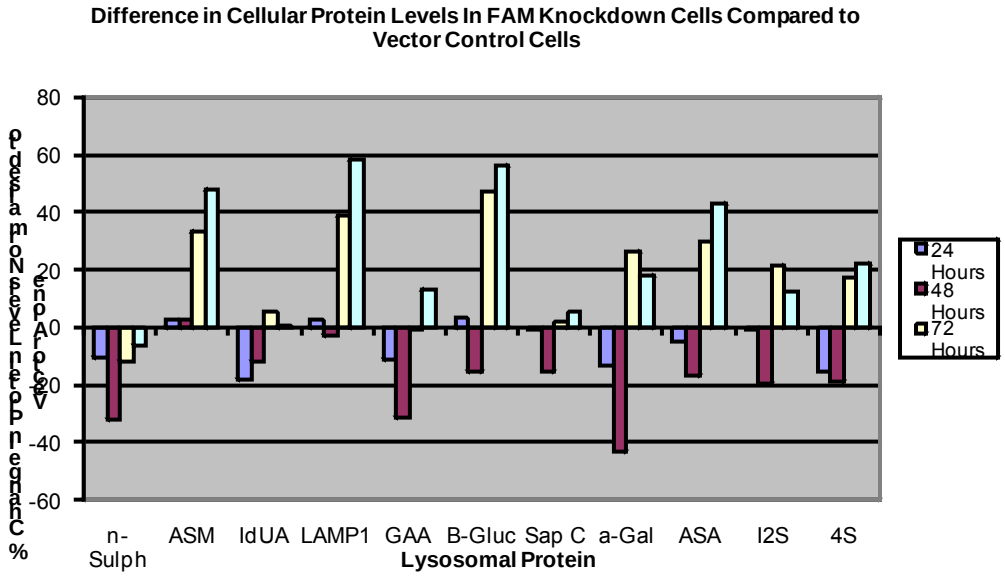

Aryl Sulphatase-A Cellular Levels in FAM KD Cells
ng/mg of Total Cellular Protein
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

I)

J)

K)

L)


Difference in Cellular Protein Levels In FAM Knockdown Cells Compared to Vector Control Cells
Percent Increase or Decrease Compared to Vector
80
60
40
20
0
-20
-40
-60
n-Sulph
ASM
IdUA
LAMP1
GAA
B-Gluc
Sap C
a-Gal
ASA
I2S
4S
Lysosomal Protein
24 Hours
48 Hours
72 Hours

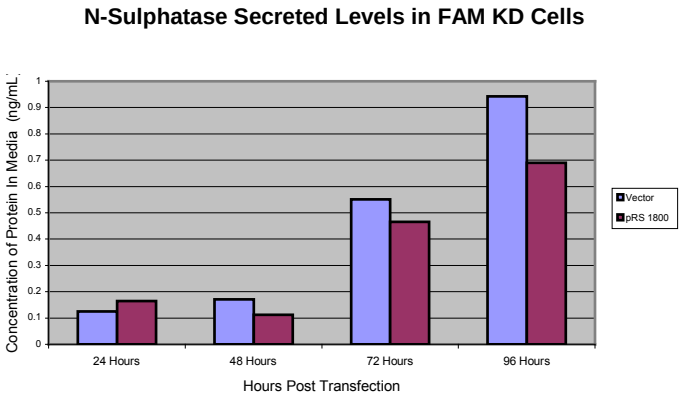
In contrast to the trends seen in vector control cells, ten of the eleven lysosomal proteins showed an increase in cellular protein levels in cells with reduced FAM levels, with α-Gal, whose levels only fluctuated, being the only protein that differed from this trend. As a result, there were three proteins for which cellular protein level changes over the time course directly opposed those seen in vector control cells (ASM, Lamp-1, and β-Gluc). All three proteins had cellular protein levels at the 72 and 96 hour time points that were greater than 10% higher than the levels seen in vector control cells (33% & 48%, 38% & 58%, and 46% & 56% higher than vector controls at the 72 & 96 hour time points for ASM, Lamp-1 and β-Gluc respectively; Figure 4.5 l). Of the five proteins whose levels increased over the time course in vector control cells, the corresponding increases in cells with reduced FAM levels for two (N-sulph and ASA) were much more pronounced. From the 48 to 72 hour time point, cellular levels of N-sulph increased by 54% (40% higher than the increase seen in control cells). Similarly, from the 72 to 96 hour time point N-sulph levels rose by a further 16% (double the increase seen in control cells). For ASA, cellular levels rose by 78% in cells with reduced FAM levels between the 48 to 72 hour time points (64% higher than the increase seen in control cells). Additionally, whereas protein levels dropped by 9% in control cells from the 72 to 96 hour time point, levels remained stable in cells with reduced FAM levels. Finally, although cellular levels of both I2S and 4S remained relatively stable in vector control cells, both proteins showed a sharp increase at the 72 hour time point in cells treated with the FAM 1800 shRNA, with final cellular concentrations at the 96 hour time point that were over 10% higher in treated cells. Overall these results suggested that if FAM levels were reduced, the total cellular pool of at least eight (ASM, Lamp-1, GAA, β-Gluc, α-Gal, ASA, I2S and 4S) lysosomal proteins became significantly increased

Secreted Protein Levels

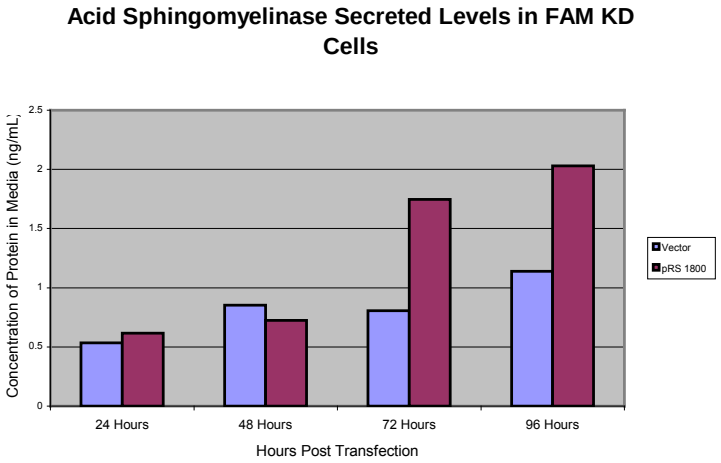
As might be expected, protein levels secreted into the media of vector control transfected cells increased over the course of the experiment for all eleven lysosomal proteins (Figure 4.6 a – k; blue bars). This was also true for cells in which FAM levels had been reduced (Figure 4.6 a – k; red bars) and there did not appear to be any affect of altered FAM levels on lysosomal protein secretion. If the same assumption applied above was made; that any difference in secreted protein levels between cells treated with FAM 1800 shRNA and vector control transfected cells that was above 10% was significant, there were four proteins whose secreted levels were significantly reduced, (N-sulph, β-Gluc, Sap-C, and 4S) and two whose levels were significantly increased in shRNA treated cells (ASM and I2S)

Figure 4.6: ***Effect of reducing endogenous FAM levels on secreted protein levels of eleven lysosomal proteins***. HEK 293T cells were transiently transfected with the base expression vector pRetro Super (pRS), or the expression vector for the 1800 shRNA trigger directed against FAM (pRS1800). Cells were incubated for 24, 48, 72, or 96 hours. **A)** to **K)** Secreted protein levels (ng/mL of media) of N-sulphatase **A)**, Acid sphingomyelinase **B)**, Iduronidase **C)**, Lamp-1 **D)**, Acid α-glucosidase **E)**, β-glucosidase **F)**, Saposin C **G)**, α-Galactosidase **H)**, Aryl sulphatase A **I)**, Iduronate-2-sulphatase **J)**, and 4-sulphatase **K)** in vector control (blue) or FAM shRNA trigger 1800 (red) expressing cells at 24, 48, 72, and 96 hours post transfection. **L)** Percentage difference in total cellular protein levels of eleven lysosomal proteins in cells with reduced FAM protein levels normalized to vector control treated cells at 24 (blue), 48 (red), 72 (yellow) and 96 (green) hours post transfection.

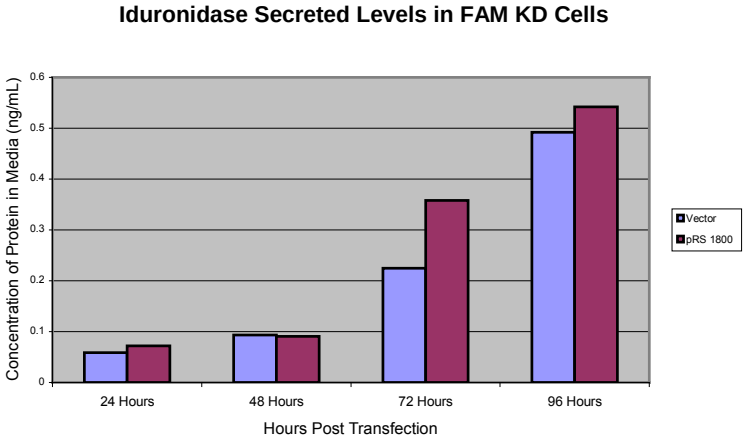
A)


N-Sulphatase Secreted Levels in FAM KD Cells
Concentration of Protein In Media (ng/mL)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

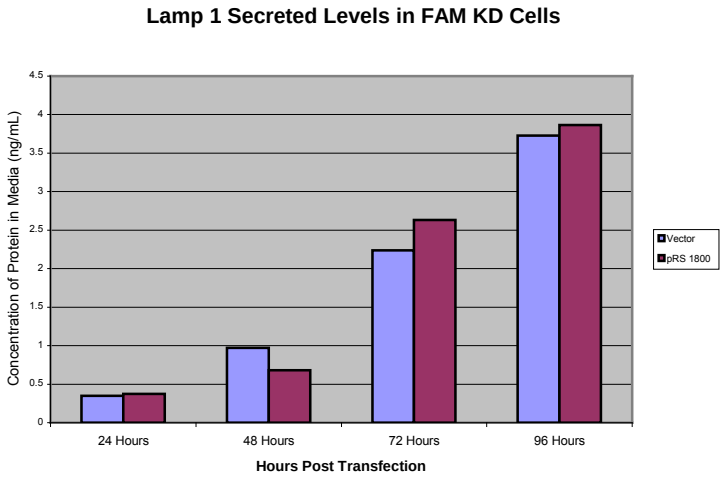
B)


Acid Sphingomyelinase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
2.5
2
1.5
1
0.5
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

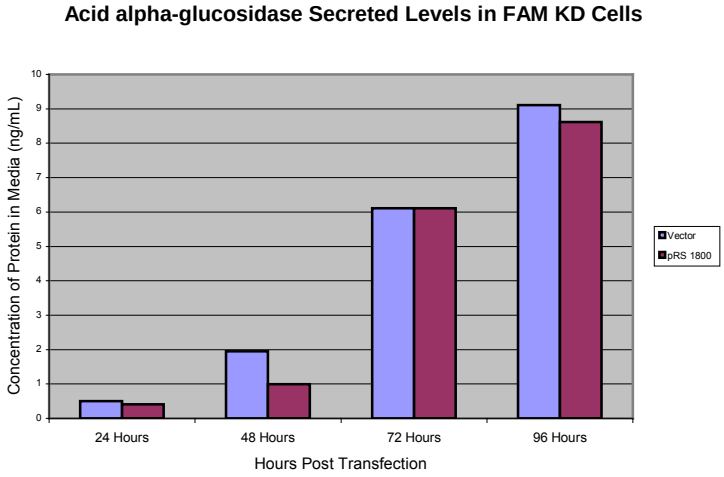
C)


Iduronidase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
0.6
0.5
0.4
0.3
0.2
0.1
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

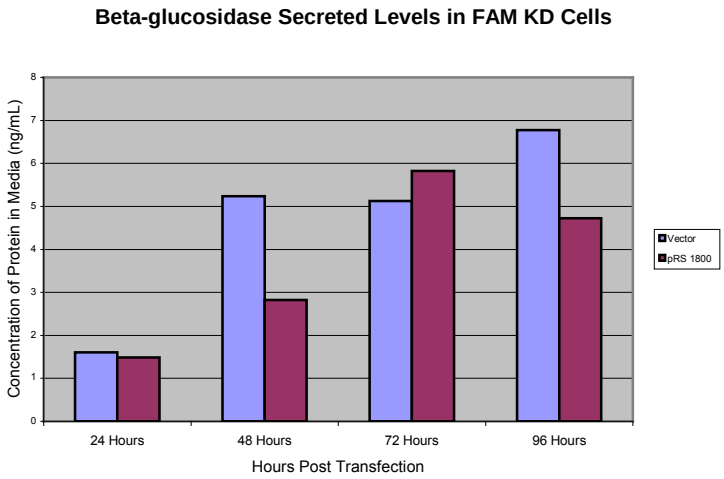
D)


Lamp 1 Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

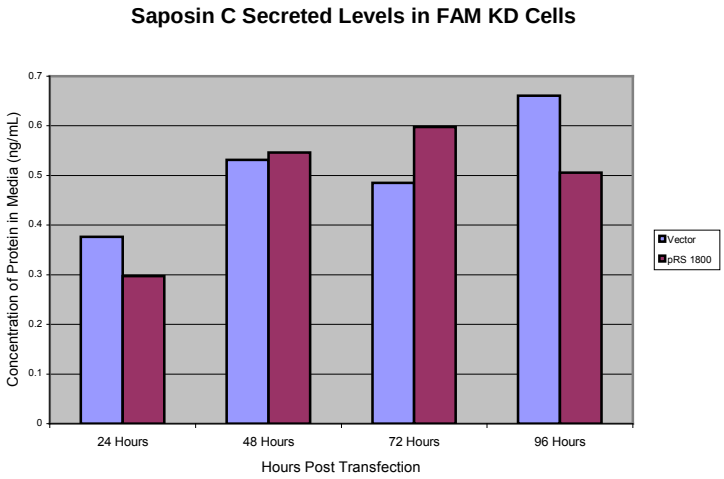
E)


Acid alpha-glucosidase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
10
9
8
7
6
5
4
3
2
1
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

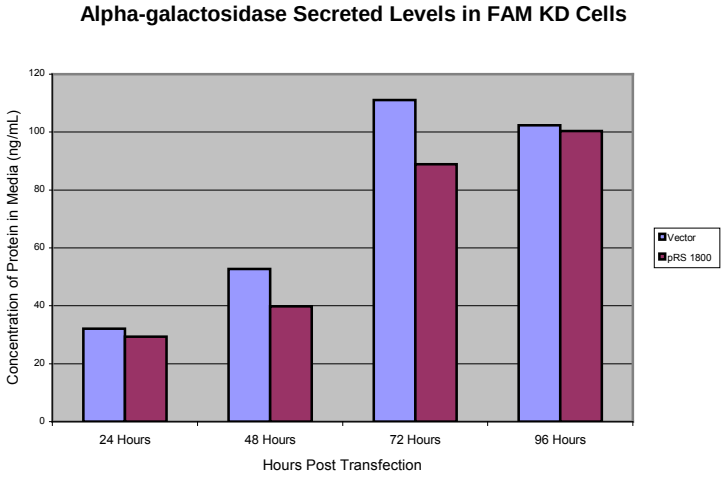
F)


Beta-glucosidase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
8
7
6
5
4
3
2
1
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

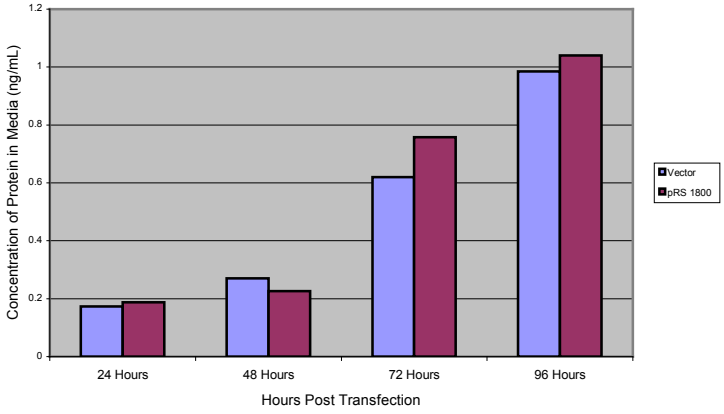
G)


Saposin C Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

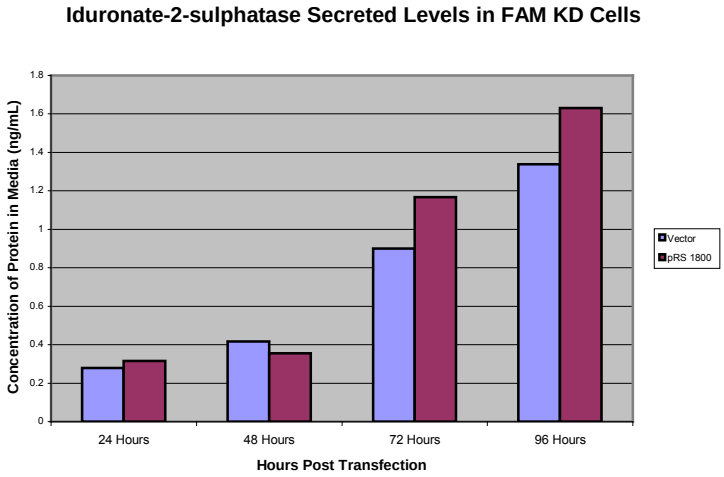
H)


Alpha-galactosidase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
120
100
80
60
40
20
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

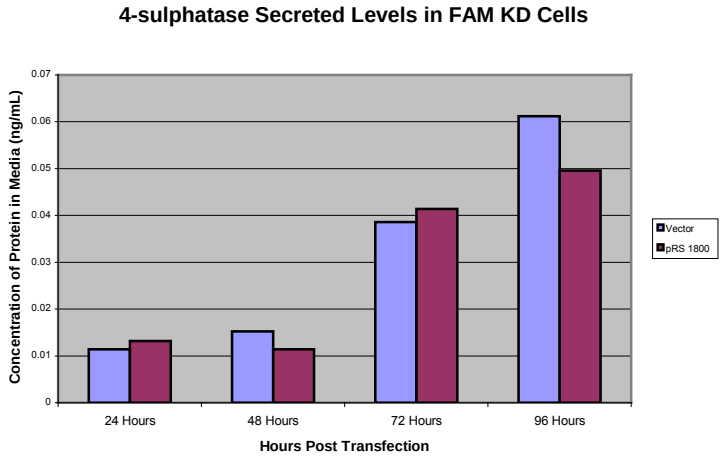
I)


Aryl sulphatase A Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
1.2
1
0.8
0.6
0.4
0.2
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

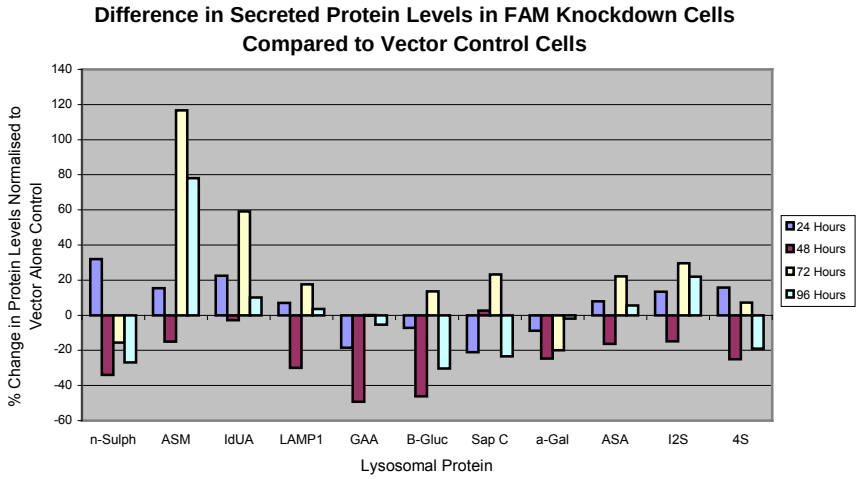
J)


Iduronate-2-sulphatase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

K)


4-sulphatase Secreted Levels in FAM KD Cells
Concentration of Protein in Media (ng/mL)
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
24 Hours
48 Hours
72 Hours
96 Hours
Hours Post Transfection
Vector
pRS 1800

L)


Difference in Secreted Protein Levels in FAM Knockdown Cells
Compared to Vector Control Cells
% Change in Protein Levels Normalised to
Vector Alone Control
140
120
100
80
60
40
20
0
-20
-40
-60
n-Sulph
ASM
IdUA
LAMP1
GAA
B-Gluc
Sap C
a-Gal
ASA
I2S
4S
Lysosomal Protein
24 Hours
48 Hours
72 Hours
96 Hours

(Figure 4.6 l). However, there are many confounding factors that must be considered when analyzing secreted protein levels (discussed below).

Summary

Table 4.1 summarizes the overall results for cells that were incubated for the optimal time post transfection (4 days) to obtain the greatest expression of exogenous FAM, or the greatest reduction in endogenous FAM levels. From these results, alterations in cellular protein levels of the eleven lysosomal protein caused by altered FAM protein levels, do not appear to correlate with changes in secreted protein levels. That is, a FAM mediated increase or decrease in a cellular lysosomal protein level does not consistently result in an increase or decrease in secreted levels of that protein. There does however, appear to be a general correlation between overexpression of FAM and a reduction in the total cellular protein levels of lysosomal proteins with these reductions reaching statistical significance (by a student's t-test) in at least five cases (N-sulph, Lamp-1, GAA, ASA, and 4S). A trend towards reduced total IdUA and I2S levels was also seen although this change was not found to be significant. Furthermore, the reverse also appears to be true, with a reduction in cellular FAM protein levels correlating with a general increase in total cellular lysosomal protein levels with marked increases (over 25%) in at least four proteins (ASM, Lamp-1, β-gluc, and ASA) and lower (over 10%) increases in an additional four proteins (GAA, α-Gal, I2S, and 4S). These increases may or may not be significant however they are higher than the "normal" variations observed in vector control treated cells.

Sorted Cell Multiplex-ELISA analysis

Despite the fact that the above transient transfections only yielded a transfection efficiency of approximately 50%, some significant and repeatable effects on the levels of some lysosomal proteins were observed. It was hypothesized therefore, that if the percentage of transfected cells could be increased, the observed effects may be amplified. To achieve this goal, the above transfections were repeated with two differences. Firstly, the cells were co-transfected with the appropriate shRNA or FAM expression vector as described above, however an additional vector driving the expression of V5 epitope tagged enhanced green fluorescent protein, pDest51 EGFP, was also included. This facilitated the enrichment of transfected cells by GFP fluorescence flow cytometric cell sorting. Secondly, since approximately half of the cells were expected to be lost following cell sorting due the absence of GFP fluorescence, a greater number of cells were initially transfected. This was done to

Table 4.1: Summary of results obtained from multiplex ELISA of transient transfections

	Increased FAM		Decreased FAM	
	Cellular	Secreted	Cellular	Secreted
N-Sulph	−−	x	x	x
ASM	x	x	+++	+++
IdUA	x	x	x	++
Lamp1	−−	++	+++	x
GAA	−−	x	++	x
B-gluc	x	x	+++	x
Sap C	x	x	x	x
a-Gal	+++	−−	++	x
ASA	−−	x	+++	x
I2S	x	x	++	++
4S	−−	x	++	x

++ = >10% higher levels after four days

+++ = >25% higher levels after four days

−− = >15% lower levels after four days

x = Variable Changes over the time course or No Change in Levels Compared to Vector Control

ensure that approximately the same number of cells were seeded and harvested at each of the time points. Secreted protein levels were not examined for the experiments involving sorted cells due to the fact that the levels of protein in the media were unlikely to be representative of what was occurring in the cells. This is because the media in which these cells were incubated was changed twice during the course of the experiment; once 16 hours post transfection and a second time at 48 hours post transfection, when cells were harvested for sorting. Analysis of secreted protein levels was also excluded as there appeared to be no correlation between the observed cellular and secreted protein levels in the experiments described above.

Overexpression of FAM

In these experiments 293T cells were cotransfected with either pDest40 or pDest40 FAM, and the EGFP selection marker expression vector. A catalytically inactive mutant of FAM bearing a cysteine to serine mutation (C1566S) and expressed as V5 epitope tagged fusion protein, was also included in these experiments. The cys mutant of FAM was included to provide an indication of whether any changes in lysosomal protein levels observed were

dependent on the deubiquitylating activity of FAM. Transfected cells were incubated for 24 hours before being harvested in normal growth media and cell sorted on the basis of EGFP fluorescence. Only mid to high EGFP fluorescent cells were selected to avoid auto fluorescing untransfected cells from being collected. The cells were then reseeded and grown for a further 48, 72, or 96 hours before being harvested and processed for western and ELISA analysis. Samples for the 72 hour time point were discarded as the yield of protein recovered was too low for either western or multiplex-ELISA analysis.

An increase in total FAM levels was observed at the 120 hour post-transfection time point for both wild-type and cys mutant FAM as determined by immunoblotting with an anti-FAM N1 antibody (Figure 4.7). The western blot that was probed with anti-FAM antibody was stripped and re-probed with anti-V5 antibody to detect exogenous expression of FAM protein. As a result some residual staining from the anti-FAM antibody can still be seen in vector control lanes. Despite this, exogenous expression of wild-type FAM was evident at both time points, as a slightly stronger anti-V5 reactive band that can be seen in the lanes loaded with lysate from pDest40 FAM transfected cells (Figure 4.7). No such band can be observed in the lanes loaded with the cys mutant FAM possibly indicating that exogenous FAM cys mutant was not expressed and therefore, analysis of FAM cys mutant transfected cell ELISA data was not included (Figure 4.7). Lysates from both the 96 and 120 hour post transfection time point were analysed by multiplex ELISA however only the results from the 96 hour time point will be discussed below as this is the only remaining time point that overlaps with the data obtained for unsorted FAM-V5 transfected cells.

Total Cellular Protein Levels

Unexpectedly, the results obtained from the multiplex ELISA analysis of sorted FAM overexpressing cells disagreed with those obtained from the simple transient transfection experiments. Whereas in the initial transient transfection experiments, the levels of many of the lysosomal proteins were significantly reduced in FAM overexpressing cells 96 hours post transfection, at the same time point in cell populations enriched for transfected cells, total cellular lysosomal protein levels were increased (Figure 4.8 a - j). When percentages for the relative increases were examined, many of these increases would also be considered significant according to the criteria outlined above (Figure 4.8 k).

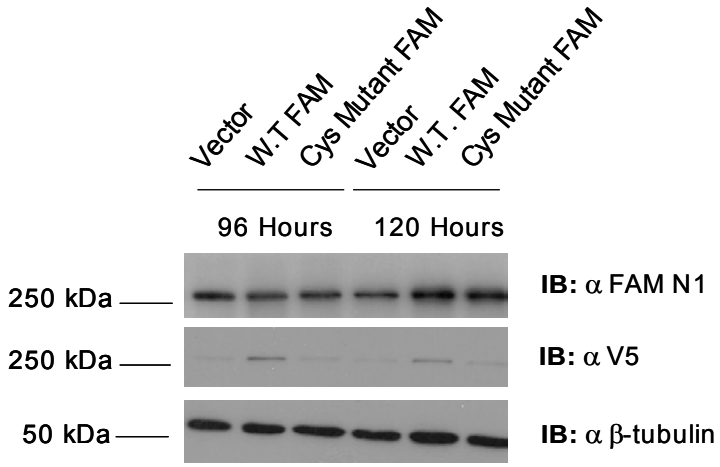
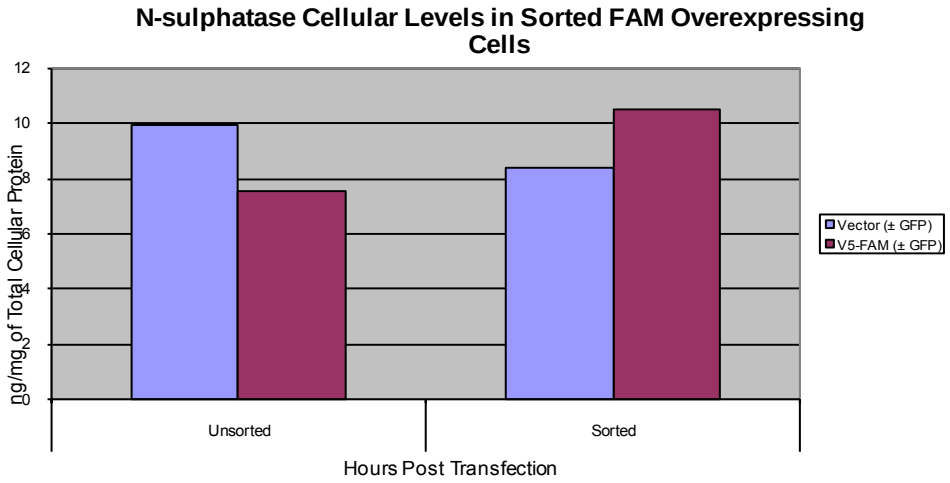


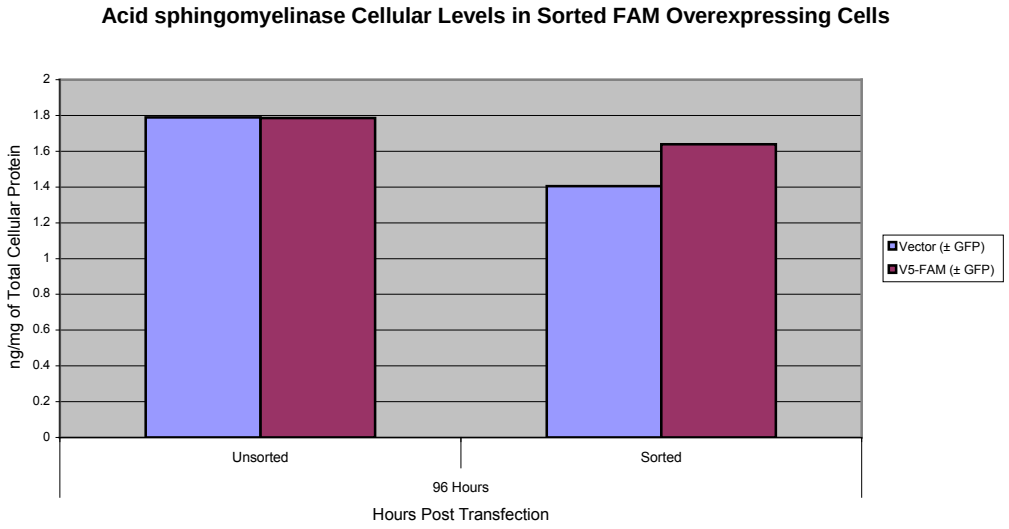
Figure 4.7: ***Exogenous expression of FAM protein in GFP sorted transient transfections***. HEK 293T cells were transiently co-transfected with the base expression vector pDest40, or expression vectors for V5 epitope tagged wild-type (pDest40 V5 FAM) or catalytically inactive cysteine mutant (pDest51 V5 C1566S FAM) and an expression vector for EGFP (pDest51 V5 EGFP). Cells were incubated for 24 hours before being cell sorted on the basis of EGFP fluorescence, reseeded and grown for a further 3 or 4 days. Western analysis with Anti-V5 antibody detected exogenous expression of wild-type FAM (W.T. FAM) protein at the 96 and 120 hour time points. Anti-FAM-N1 antibodies indicated an increase in total cellular FAM levels at the 5 day time point in cells transfected with pDest40-V5-FAM and pDest51-V5-C1566S FAM when compared to cells transfected with empty vector. Anti-β-Tubulin antibodies indicated equal loading of proteins in each lane.

Figure 4.8: ***Effect of using a more homogeneous population of FAM overexpressing cells on the cellular protein levels of eleven lysosomal proteins.*** HEK 293T cells were transiently co-transfected with either the base expression vector pDest40, or an expression vector for V5 epitope tagged wild-type (pDest40 V5 FAM) and an expression vector for EGFP (pDest51 V5 EGFP). Cells were incubated for 24 hours before being cell sorted on the basis of EGFP fluorescence, reseeded and grown for a further 3 days. **A)** to **J)** total cellular protein levels (ng/mg of total cellular protein) of N-sulphatase **A)**, Acid sphingomyelinase **B)**, Iduronidase **C)**, Lamp-1 **D)**, Acid α-glucosidase **E)**, β-glucosidase **F)**, Saposin C **G)**, Aryl sulphatase A **H)**, Iduronate-2-sulphatase **I)**, and 4-sulphatase **J)**, in vector control (blue) or FAM overexpressing (red) cells at 96 hours post transfection. Corresponding levels of each protein for unsorted cells are shown. **L)** Percentage difference in total cellular protein levels of eleven lysosomal proteins in cells expressing exogenous FAM protein normalized to vector control treated cells at 96 hours post transfection.

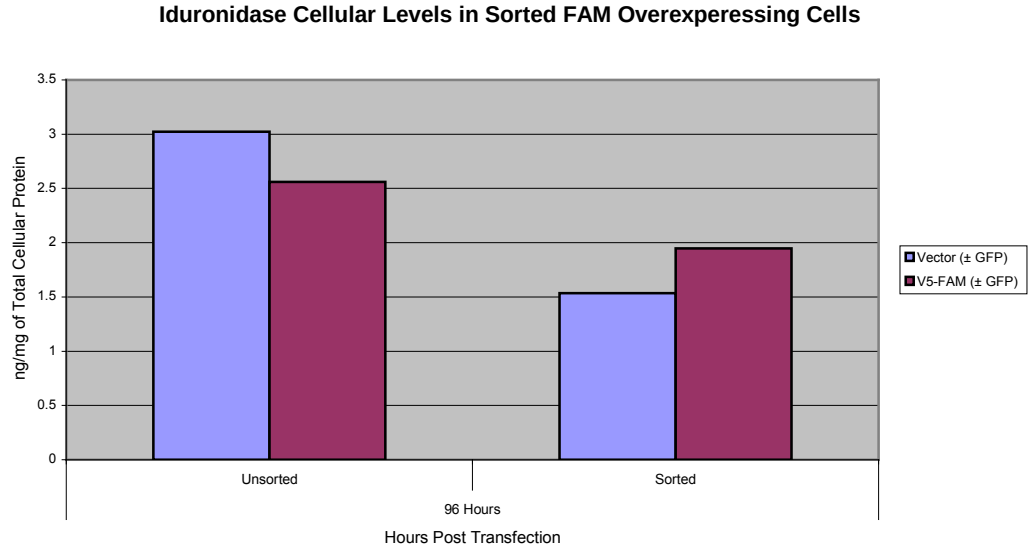
A)


N-sulphatase Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
12
10
8
6
4
2
0
Unsorted
Sorted
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

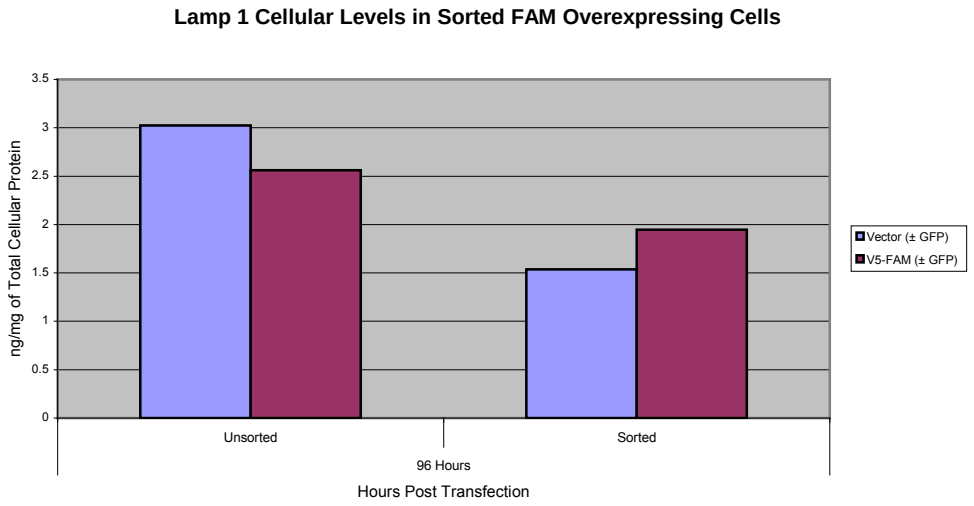
B)


Acid sphingomyelinase Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

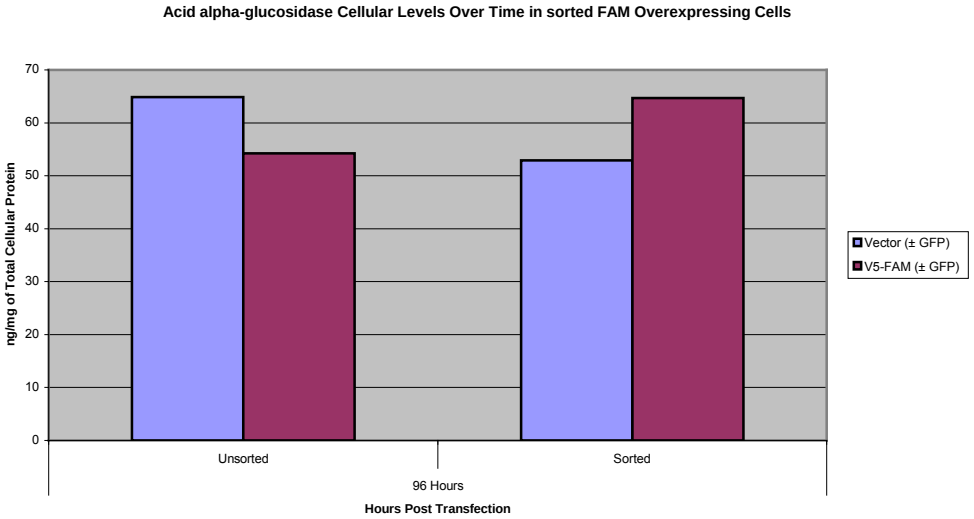
C)


Iduronidase Cellular Levels in Sorted FAM Overexperessing Cells
ng/mg of Total Cellular Protein
3.5
3
2.5
2
1.5
1
0.5
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

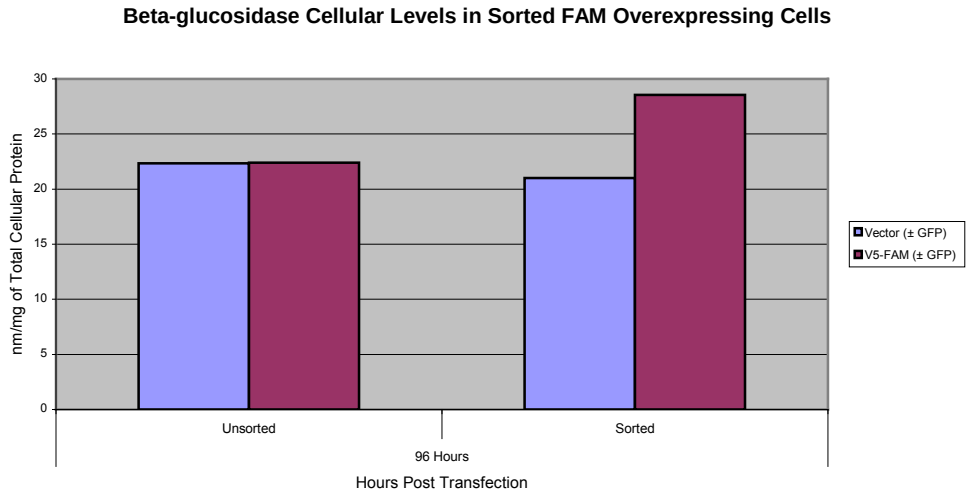
D)


Lamp 1 Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
3.5
3
2.5
2
1.5
1
0.5
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

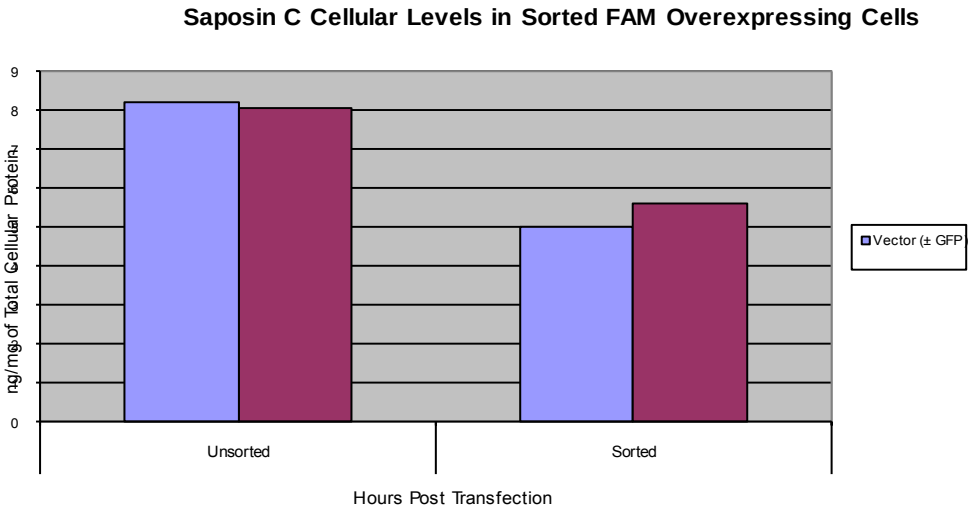
E)


Acid alpha-glucosidase Cellular Levels Over Time in sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
70
60
50
40
30
20
10
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

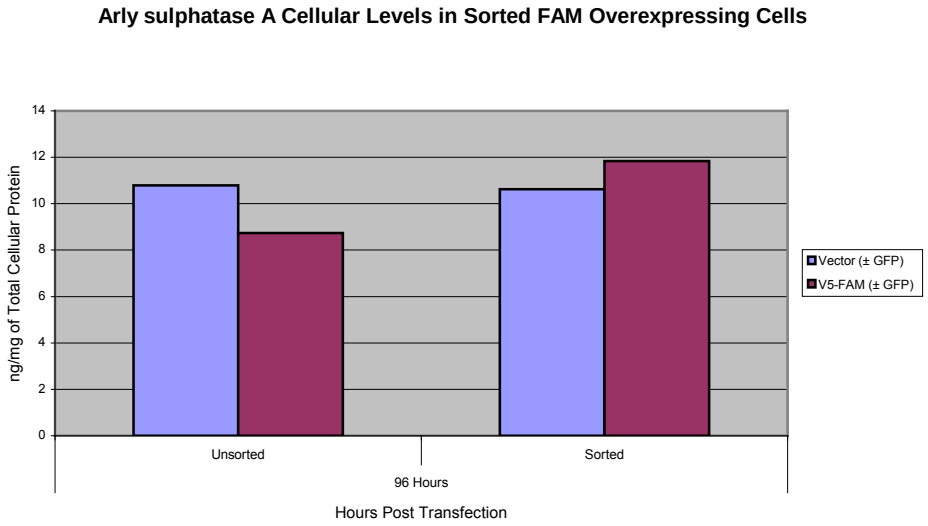
F)


Beta-glucosidase Cellular Levels in Sorted FAM Overexpressing Cells
nm/mg of Total Cellular Protein
30
25
20
15
10
5
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

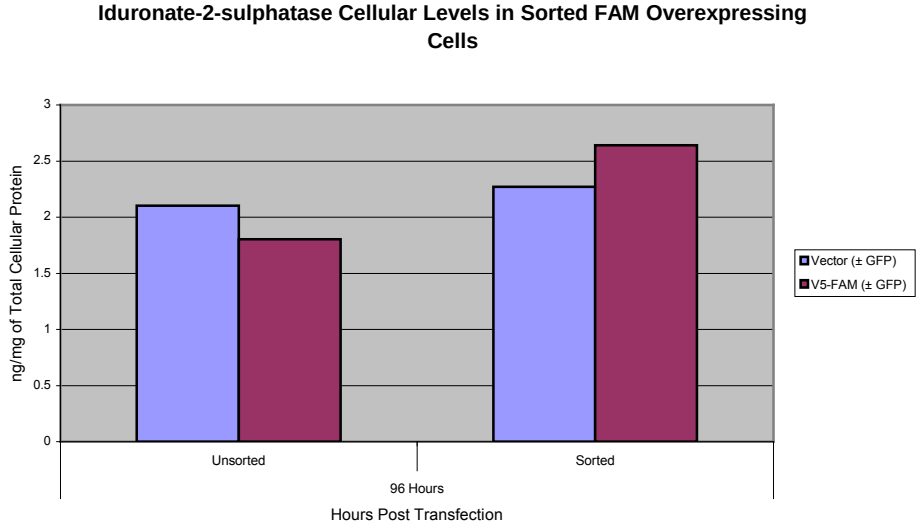
G)


Saposin C Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
0
1
2
3
4
5
6
7
8
9
Unsorted
Sorted
Hours Post Transfection
Vector (± GFP)

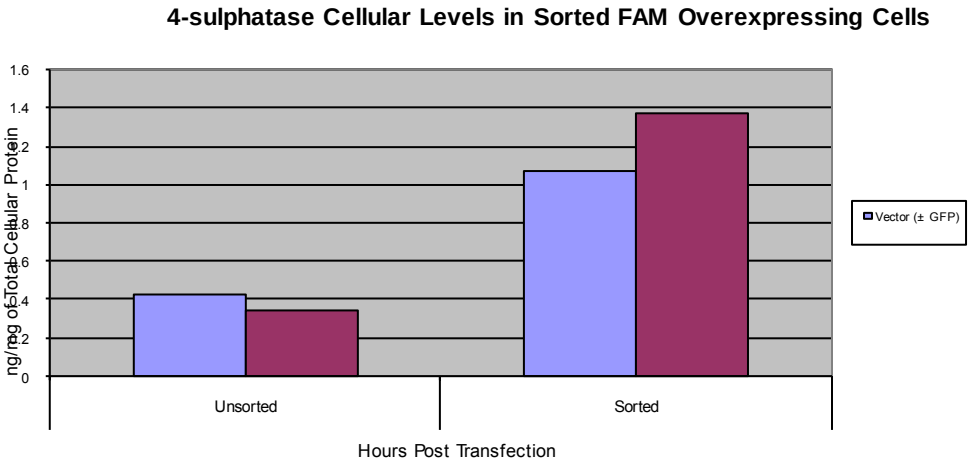
H)


Arly sulphatase A Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
0
2
4
6
8
10
12
14
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

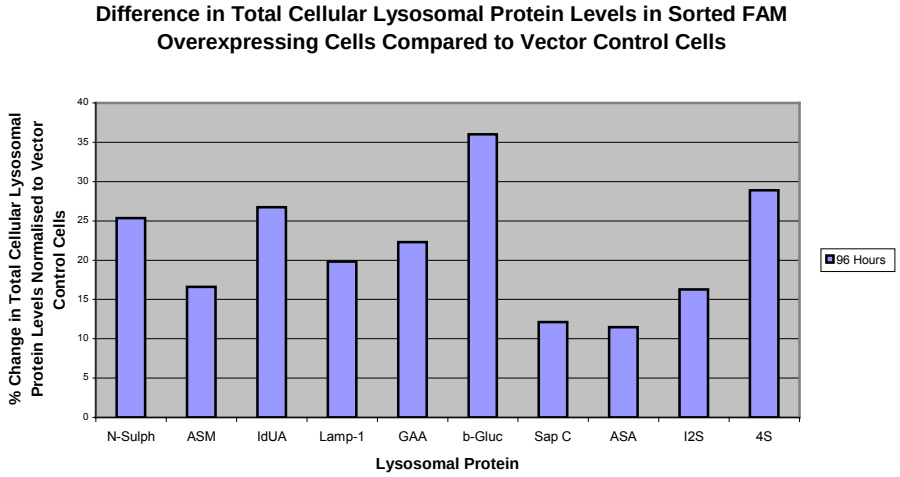
I)


Iduronate-2-sulphatase Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
0
0.5
1
1.5
2
2.5
3
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
V5-FAM (± GFP)

J)


4-sulphatase Cellular Levels in Sorted FAM Overexpressing Cells
ng/mg of Total Cellular Protein
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Unsorted
Sorted
Hours Post Transfection
Vector (± GFP)

K)


Difference in Total Cellular Lysosomal Protein Levels in Sorted FAM Overexpressing Cells Compared to Vector Control Cells
% Change in Total Cellular Lysosomal Protein Levels Normalised to Vector Control Cells
40
35
30
25
20
15
10
5
0
N-Sulph
ASM
IdUA
Lamp-1
GAA
b-Gluc
Sap C
ASA
I2S
4S
Lysosomal Protein
96 Hours

Reduced FAM Levels – FAM Knockdown

To enrich for cells transfected with the FAM shRNA expressing construct, a similar strategy to that described above for FAM overexpression experiments was used. 293T cells were co-transfected with either pRetro Super or pRetro Super 1800 and pDest51 EGFP or with pDest51 EGFP alone. Cells were incubated for 24 hours before being harvested into PBS and cell sorted based on GFP fluorescence. Sorted cells were then reseeded and grown for a further 72, 96 or 120 hours before being harvested and processed for multiplex-ELISA and western analysis. Western analysis with an anti-FAM antibody revealed that a reduction in endogenous FAM levels was achieved at all three time points (Figure 4.9). Levels of endogenous FAM, while still lower than the levels in vector and GFP alone controls, were beginning to recover at the 120 and 154 hour time points (Figure 4.9). Once again, as the only time point that overlapped with the earlier FAM knockdown study was the 96 hour time point; this is the only data that will be discussed below.

Total Cellular Protein Levels

In sorted FAM knockdown cells, ASM, Lamp1, Sap C, ASA, and 4S were found to have a significant increase in their total cellular protein levels at 96 hours when protein levels were normalized to vector alone controls in agreement with the results described above (Figure 4.10 b, d, h, j, and k). GAA, β-Gluc, and I2S also had raised levels in sorted FAM knockdown cells (Figure 4.10 e, f, g, and i) but the normalized difference from vector controls for these three proteins was below the 10% cut-off assumed above (Figure 4.10 k). An unexpected result was observed when the total cellular levels of the ten lysosomal proteins examined in this experiment were measured in cells expressing the eGFP sorting marker alone. Given that all of the plasmids used in these experiments were prepared in the same manner, transfection of cells with either the pDEST51-V5-eGFP vector alone or in combination with the pRS base vector should have produced similar results. Instead, the total cellular levels of some of the lysosomal proteins were found to be as high or higher in eGFP alone expressing cells as the levels seen in FAM knockdown cells (Figure 4.10 b, c, f, and g). If eGFP is considered to be an inert protein that should have no effect on cells, this result would seem to indicate that there was a much larger degree of variation or noise in the system than previously reasoned, and would call into question the results obtained in this and earlier experiments. While this may have been the case, arguments against this conclusion will be discussed below.

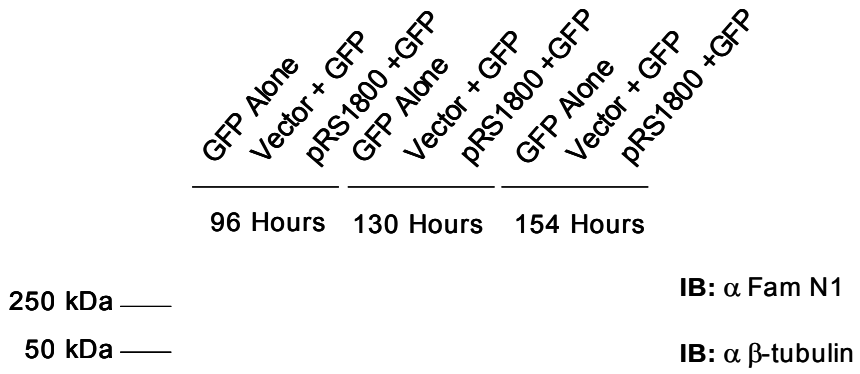
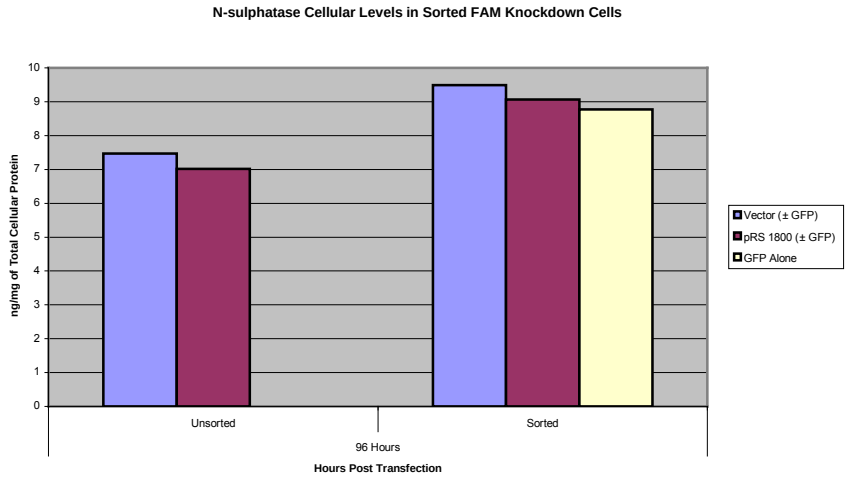


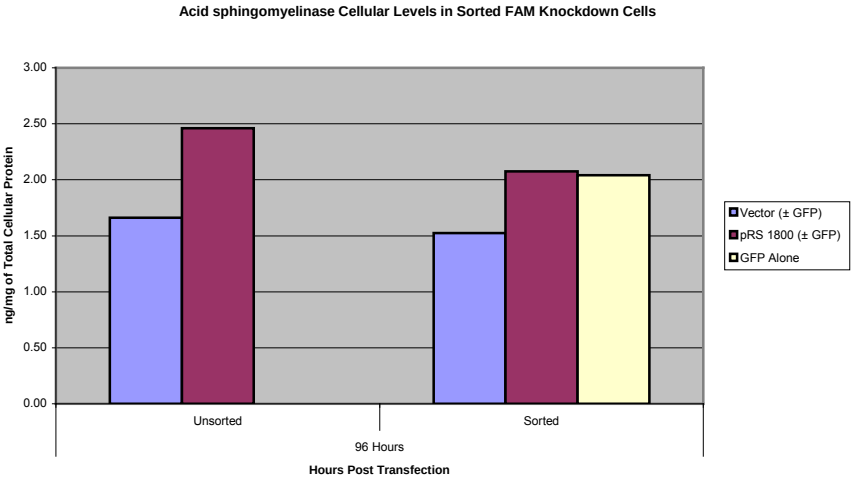
Figure 4.9: ***Determining the level of FAM knockdown in GFP sorted 1800 shRNA transfected cells.*** HEK 293T cells were transiently cotransfected with either the base expression vector pRetro Super (pRS), or an expression vector for the FAM shRNA 1800 trigger (pRS 1800) and an expression vector for enhanced GFP (pDEST51-V5-eGFP). Cells were incubated for 24 hours before being cell sorted on the basis of EGFP fluorescence, reseeded and grown for a further 3, 4 or 5 days. Total cellular protein was then analysed by western blot with anti-FAM, and anti-β-tubulin antibodies. . Western analysis with Anti-FAM antibody showed a reduction in endogenous FAM protein levels at the 96, 130 and 154 hour time points for cells transfected with the 1800 shRNA trigger expressing construct. Anti-β-tubulin antibodies indicate equal loading of protein in all lanes. The absence of β-tubulin staining in the last lane may indicate a bubble in the transfer set up which was confirmed by ponceau S staining.

Figure 4.10: ***Effect of using a more homogeneous population of FAM shRNA treated cells on the cellular protein levels of eleven lysosomal proteins.*** HEK 293T cells were transiently cotransfected with either the base expression vector pRetro Super (pRS), or an expression vector for the FAM shRNA 1800 trigger (pRS 1800) and an expression vector for enhanced GFP (pDEST51-V5-eGFP). Cells were incubated for 24 hours before being cell sorted on the basis of EGFP fluorescence, reseeded and grown for a further 72 hours. **A)** to **J)** Total cellular protein levels (ng/mg of total cellular protein) of N-sulphatase **A)**, Acid sphingomyelinase **B)**, Iduronidase **C)**, Lamp-1 **D)**, Acid α-glucosidase **E)**, β-glucosidase **F)**, Saposin C **G)**, Aryl sulphatase A **H)**, Iduronate-2-sulphatase **I)**, and 4-sulphatase **J)**, in vector control cells (blue), FAM overexpressing cells (red), or GFP alone (yellow) transfected cells at 96 hours post transfection. Corresponding levels for unsorted cells at the 96 hour time point are shown of each protein. **K)** Percentage difference in total cellular protein levels of eleven lysosomal proteins in cells with reduced FAM protein levels normalized to vector control treated cells at 96 (blue) hours post transfection.

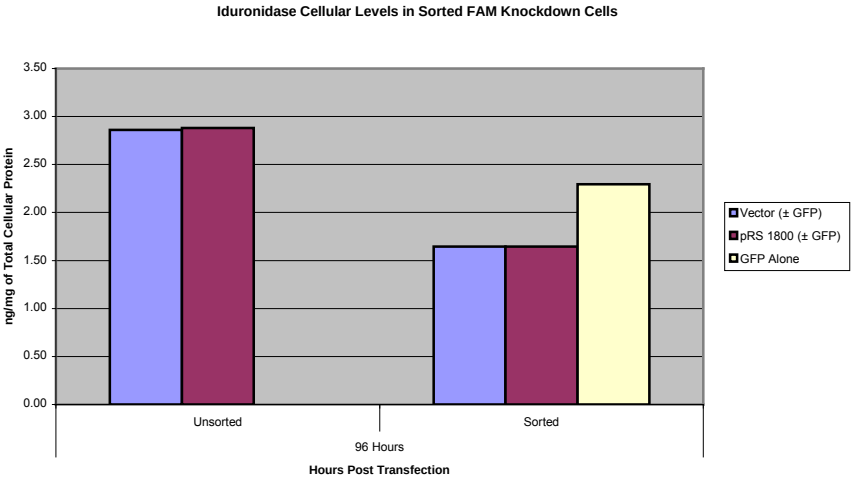
A)


N-sulphatase Cellular Levels in Sorted FAM Knockdown Cells
ng/mg of Total Cellular Protein
10
9
8
7
6
5
4
3
2
1
0
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

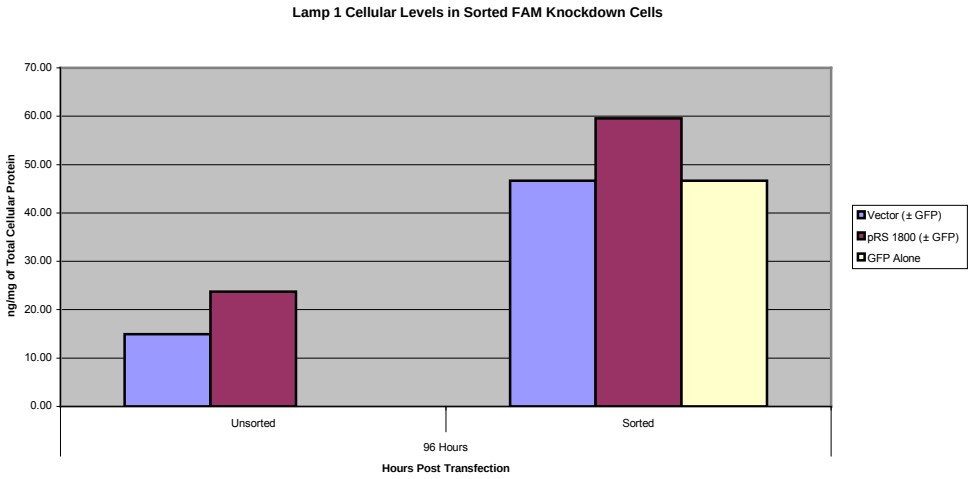
B)


Acid sphingomyelinase Cellular Levels in Sorted FAM Knockdown Cells
ng/mg of Total Cellular Protein
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

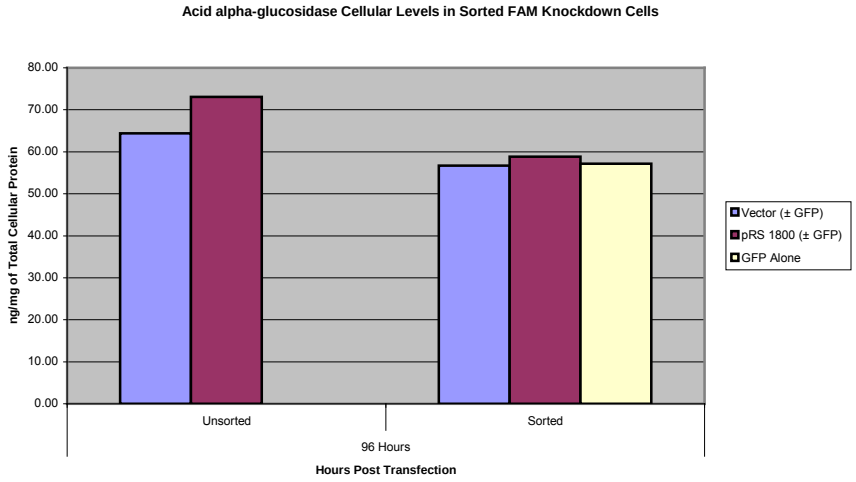
C)


Iduronidase Cellular Levels in Sorted FAM Knockdown Cells
ng/mg of Total Cellular Protein
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

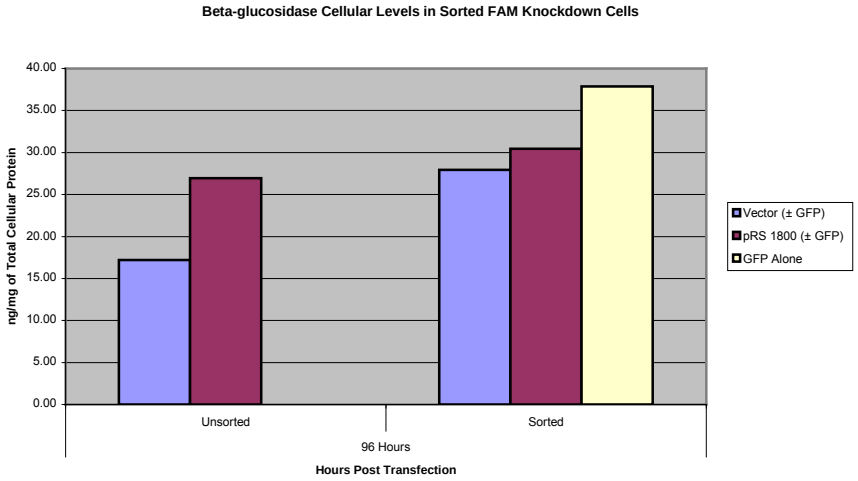
D)


Lamp 1 Cellular Levels in Sorted FAM Knockdown Cells
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

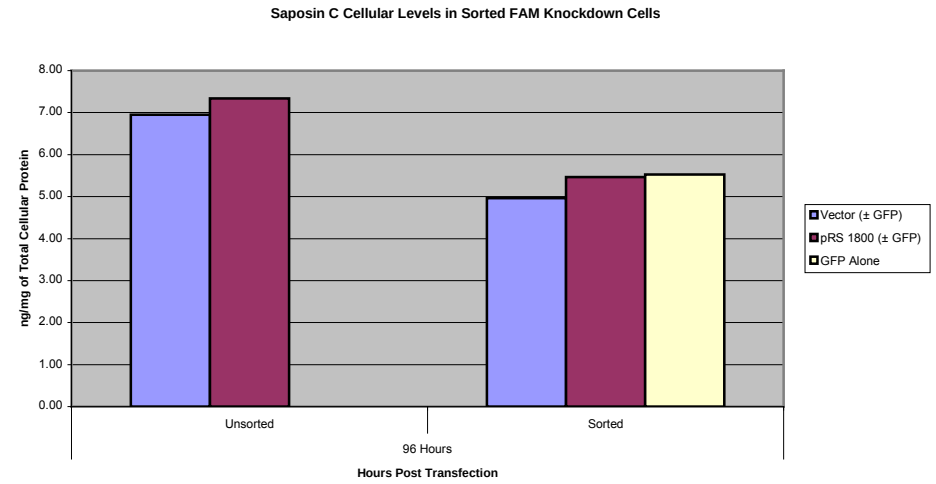
E)


Acid alpha-glucosidase Cellular Levels in Sorted FAM Knockdown Cells
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

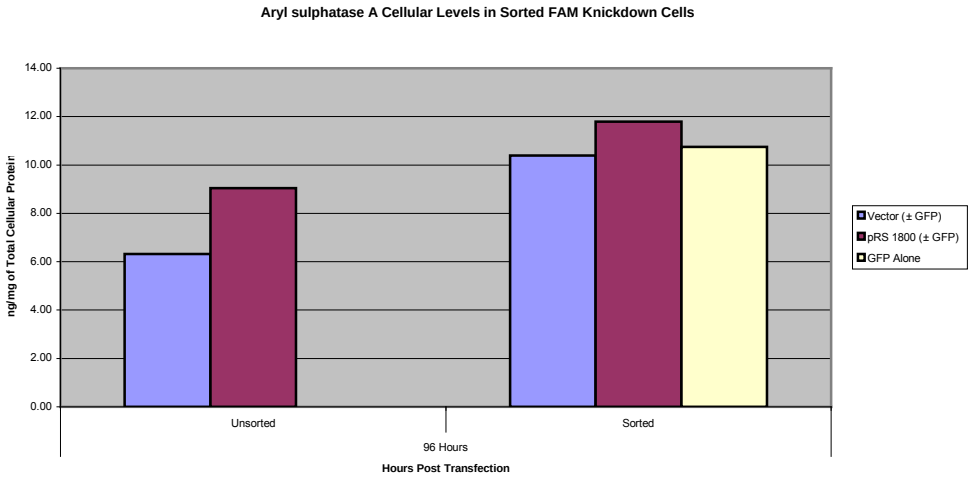
F)


Beta-glucosidase Cellular Levels in Sorted FAM Knockdown Cells
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

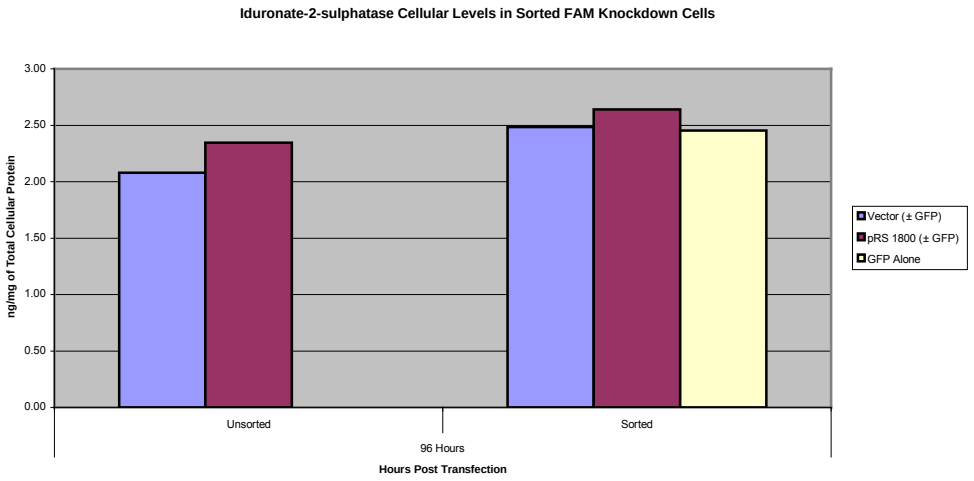
G)


Saposin C Cellular Levels in Sorted FAM Knockdown Cells
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

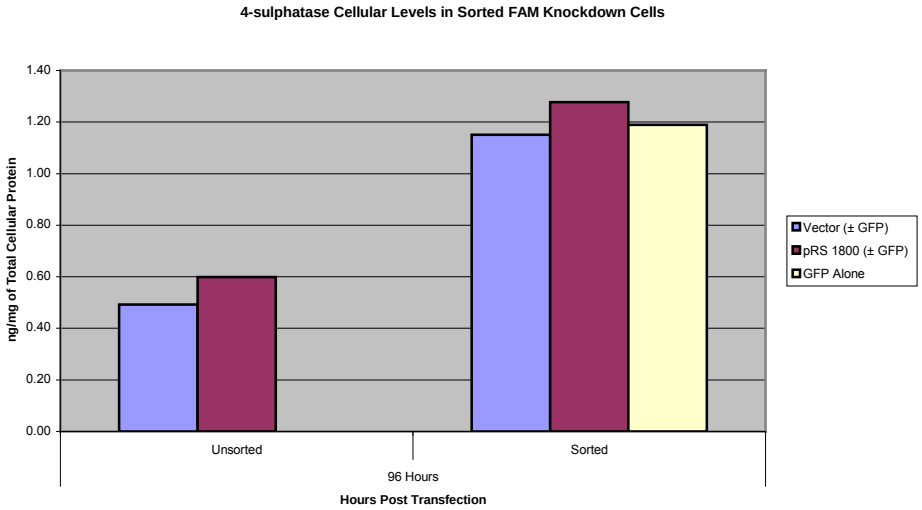
H)


Aryl sulphatase A Cellular Levels in Sorted FAM Knickdown Cells
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

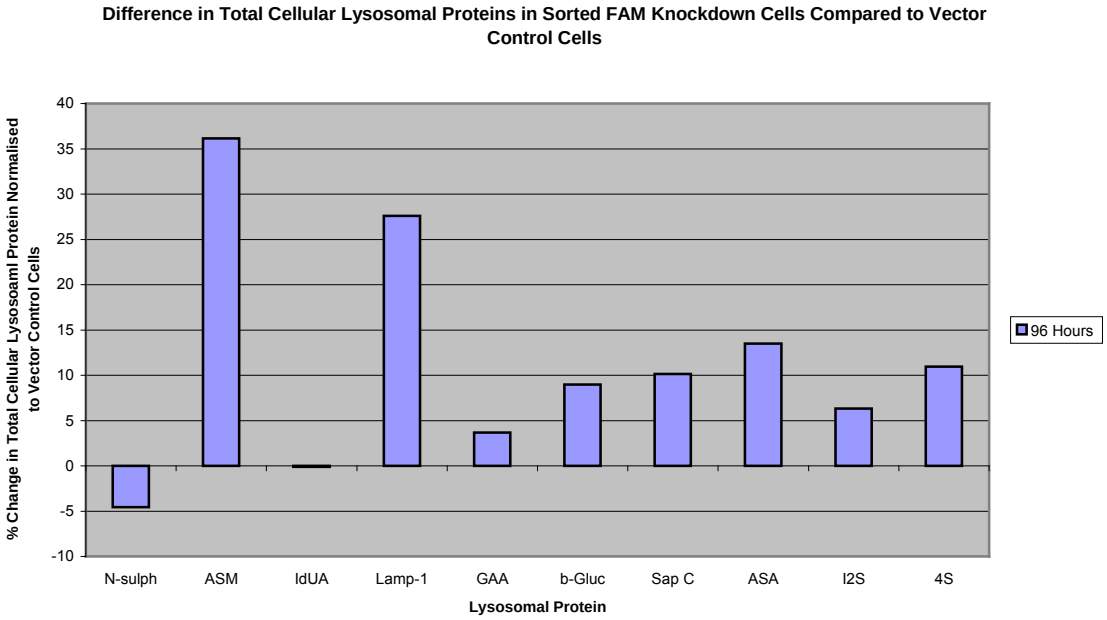
I)


Iduronate-2-sulphatase Cellular Levels in Sorted FAM Knockdown Cells
3.00
2.50
2.00
1.50
1.00
0.50
0.00
ng/mg of Total Cellular Protein
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

J)


4-sulphatase Cellular Levels in Sorted FAM Knockdown Cells
ng/mg of Total Cellular Protein
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
Unsorted
Sorted
96 Hours
Hours Post Transfection
Vector (± GFP)
pRS 1800 (± GFP)
GFP Alone

K)


Difference in Total Cellular Lysosomal Proteins in Sorted FAM Knockdown Cells Compared to Vector Control Cells
% Change in Total Cellular Lysosoaml Protein Normalised to Vector Control Cells
40
35
30
25
20
15
10
5
0
-5
-10
N-sulph
ASM
IdUA
Lamp-1
GAA
b-Gluc
Sap C
ASA
I2S
4S
Lysosomal Protein
96 Hours

Summary

Table 4.2 summarizes the total cellular protein levels of the eleven lysosomal proteins examined above at the 96 hour time point. Comparing the results obtained for both sets of experiments it appeared that enriching for the population of transfected cells did not enhance the differences between control cells and FAM overexpression/knockdown cells. Results obtained from sorted cells expressing exogenous FAM protein are contrary to those obtained from the simple transfection experiments. There was a much higher degree in similarity in the results obtained from sorted FAM knockdown cells and those obtained from the simple transient transfection experiments. An overall reduction in N-sulph cellular protein levels was seen in FAM knockdown cells in both sets of experiments. Similar degrees of increases in the cellular protein levels of the remaining lysosomal proteins were also seen in FAM knockdown cells from both sets of experiments. The inclusion of cells that were transfected with the eGFP reporter vector alone raised some doubts about the normal variation in cellular lysosomal protein levels and therefore the significance of the noted increase in levels associated with a reduction in FAM levels, especially given that the results were obtained from one replicate. There are however reports in the literature that suggest that eGFP is not, as assumed, an inert protein, especially when it is expressed at high levels (Hanazono et al., 1997; Liu et al., 1999). With reference to these reports, some hypotheses to explain the unexpected results seen in both experiments involving the sorting of transfected cells using GFP fluorescence will be discussed below.

Effects of Altered FAM Levels on Lamp1, ASA, I2S, and 4S

Of the 11 lysosomal proteins that were investigated, four were chosen for further analysis. Lamp1, aryl sulphatase A (ASA), and 4-sulphatase (4S) were chosen as their cellular levels were decreased in FAM gain of function cells and increased in FAM loss of function cells. The opposing effects of an increase or a decrease in FAM on the total cellular protein levels of these proteins was consistent with these proteins being potential candidates for proteins whose trafficking was affected by alterations in FAM levels. Iduronate-2-sulphatase was also examined as its protein levels did not appear to be significantly affected by alteration of cellular FAM protein levels. Although the multiplex ELISA provides information about the total level of these lysosomal proteins found within the cell, it does not distinguish between immature and mature forms of the protein. It also does not provide any information about the site of any potential block or other disruption in the trafficking of the proteins caused by altering FAM effects. As a result, two additional experiments were conducted in attempt to address these issues.

Table 4.2: Summary of results obtained from multiplex ELISA of sorted cells

	Increased FAM		Reduced FAM	
	Unsorted	Sorted	Unsorted	Sorted
N-Sulph	−−	+++	X	X
ASM	X	++	+++	+++
IdUA	X	+++	X	X
Lamp1	−−	++	+++	+++
GAA	−−	++	++	X
B-gluc	X	+++	+++	X
Sap C	X	++	X	++
a-Gal	+++	N/A	++	N/A
ASA	−−	++	+++	++
I2S	X	++	++	X
4S	−−	+++	++	++

++	= >10% higher levels after four days
+++	= >25% higher levels after four days
−−	= >15% lower levels after four days
X	= Variable Changes over the time course or No Change in Levels Compared to Vector Control

Firstly, western analysis was performed to determine the effects of altering FAM protein levels on the maturity and processing of the four selected lysosomal proteins. Newly synthesized lysosomal proteins undergo a number of post translational modifications on their journey to the lysosome. All soluble lysosomal proteins are highly *N*-glycosylated on asparagine residues in the rough endoplasmic reticulum (ER). These glycosylations are then edited and processed in the trans-Golgi in a process which leads to the production of exposed mannose-6-phosphate (M6P) residues. These M6P residues serve as the recognition substrate for either of the two known mannose-6-phosphate receptors (M-6-PRs) and allow sorting of lysosomal proteins to the lysosome. Membrane resident lysosomal proteins, such as the Lamp proteins, also undergo extensive glycosylation and editing en-route to the lysosome. The extensive post-translational modification of lysosomal proteins results in the presence of multiple molecular weight species of these proteins which appear as a banded pattern upon western analysis.

Western analysis of selected lysosomal protein levels and maturity

Western analysis to examine the banding/smearing pattern of Lamp1 in FAM shRNA transfected cells revealed a difference in the lower molecular weight species. At the 72 and 96 hour post transfection time points, cells with reduced FAM levels displayed a greater abundance of lower molecular weight species of Lamp-1 than those seen in control cells (Figure 4.11 a). This observation was made in three independent experiments. A subtle difference was also noted in the banding pattern observed for ASA. The top-most band of a doublet seen at approximately 60 kDa (arrows in Figure 4.11 a) appeared to be more pronounced in FAM shRNA treated cells than in vector control cells. This phenomenon was observed in at least two experiments. Immunoreactivity using antibodies raised against I2S and 4S was quite low and, no differences between vector control cells and FAM overexpressing or knockdown cells was observed (data not shown). The banding/smearing pattern of Lamp1 was investigated in the FAM overexpression experiment in which transfected cells were selected for by GFP fluorescence cell sorting. It was postulated that in contrast to what was seen when FAM levels were reduced in cells, expression of exogenous FAM might lead to lower levels of the lower molecular weight species of Lamp1. However, western analysis did not show any difference between the banding/smearing pattern of lamp in vector transfected cells and wild-type, or cys mutant FAM transfected cells (Figure 4.11 b). Staining with anti-ASA, anti-I2S, and anti-4S was unsuccessful in this experiment (data not shown).

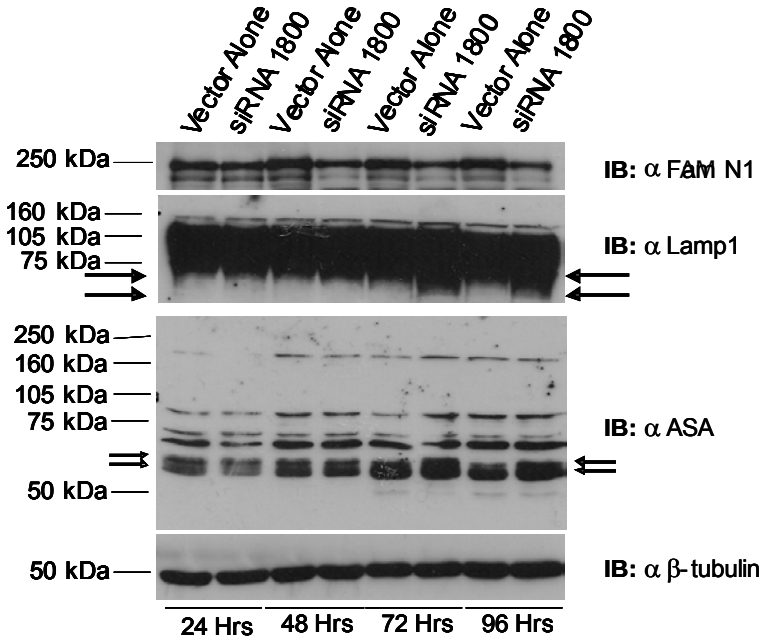
Immunofluorescence analysis of lysosomal protein localisation upon modulation of FAM levels

FAM Overexpression

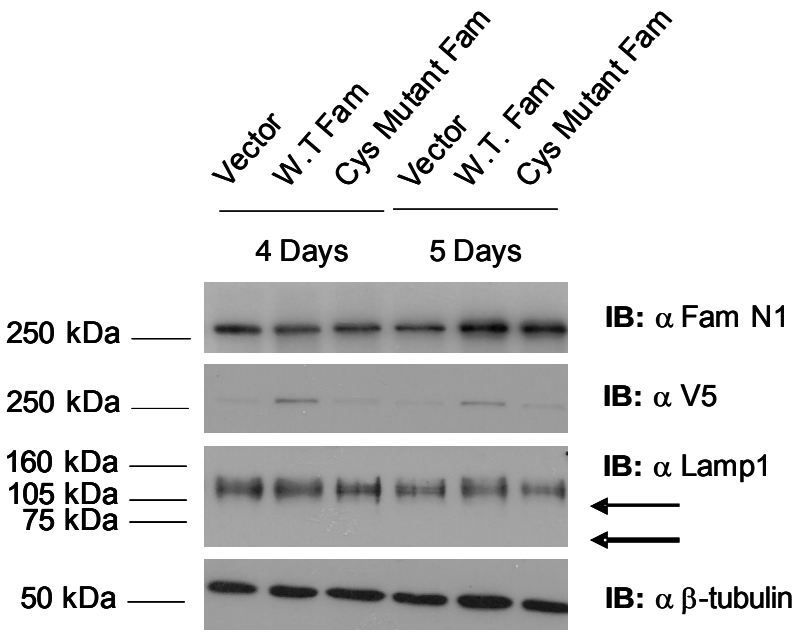
The effect of expression of exogenous FAM on the intracellular localisation of Lamp1, ASA, I2S, and 4S was investigated by transfecting HEK 293T cells with pDest40, pDest40 FAM, or pDest51 FAM C1566S. The catalytically inactive Cys mutant of FAM was included in these experiments to determine if any changes in the sub cellular distribution of the lysosomal proteins investigated that may have been observed, was due to the deubiquitylating

Figure 4.11: ***Effects of altered FAM levels on maturity and processing of selected lysosomal proteins.*** **A)** HEK 293T cells were transiently transfected with the pRS base vector alone or an expression vector for the FAM shRNA 1800 trigger. Cells were incubated for 24, 48, 72 or 96 hours before being harvested, lysed, and analysed by western blotting with antibodies raised against FAM , Lamp1, ASA, and β-tubulin. A reduction in endogenous FAM levels at the 48, 72, and 96 hour time points was detected by western blotting (top panel). Western blotting with a sheep polyclonal anti-Lamp1 antibody detected a broadening of the banding/smearing pattern of anti-Lamp1 reactive bands (indicated by the two bars) in the lysate of FAM shRNA 1800 transfected cells at the 72 and 96 hour time points. The same broadening of the banding/smearing pattern of anti-Lamp1 reactive bands was not seen in cells transfected with the vector control. Western blotting with a sheep polyclonal anti-ASA antibody detected an increase in the intensity of a band at approximately 65 kDa in the lysate from FAM shRNA 1800 transfected cells at the 72, and 96 hour time points. The same band in vector control transfected cells did not show any change in intensity over the time course of the experiment. **B)** HEK 293T cells were transiently cotransfected with the pDest40 base vector alone, an expression vector for V5 epitope tagged wild-type FAM (pDest40 V5 FAM), or V5 epitope tagged Cys mutant FAM (pDest 51 V5 C1566S FAM) and an expression vector for EGFP (pDest51 EGFP). Cells were incubated for 24 hours before being cell sorted on the basis of EGFP fluorescence and reseeded. Cells were grown for a further 2, 3, or 4 days before being harvested, lysed, and analysed by western blotting with antibodies raised against V5, FAM, Lamp1, and β-tubulin. Exogenous expressed wild-type FAM was detected at the 3, 4, and 5 day time points using an anti-V5 antibody. Exogenous expression of cys mutant FAM was not able to be detected. There were no obvious changes in the processing or maturity of Lamp1 in wild-type FAM expressing cells that could be detected with an anti-Lamp1 sheep polyclonal antibody.

A)


Vector Alone
siRNA 1800
Vector Alone
siRNA 1800
Vector Alone
siRNA 1800
Vector Alone
siRNA 1800
250 kDa
IB: α Fam N1
160 kDa
105 kDa
75 kDa
IB: α Lamp1
250 kDa
160 kDa
105 kDa
75 kDa
50 kDa
IB: α ASA
50 kDa
IB: α β-tubulin
24 Hrs
48 Hrs
72 Hrs
96 Hrs

B)

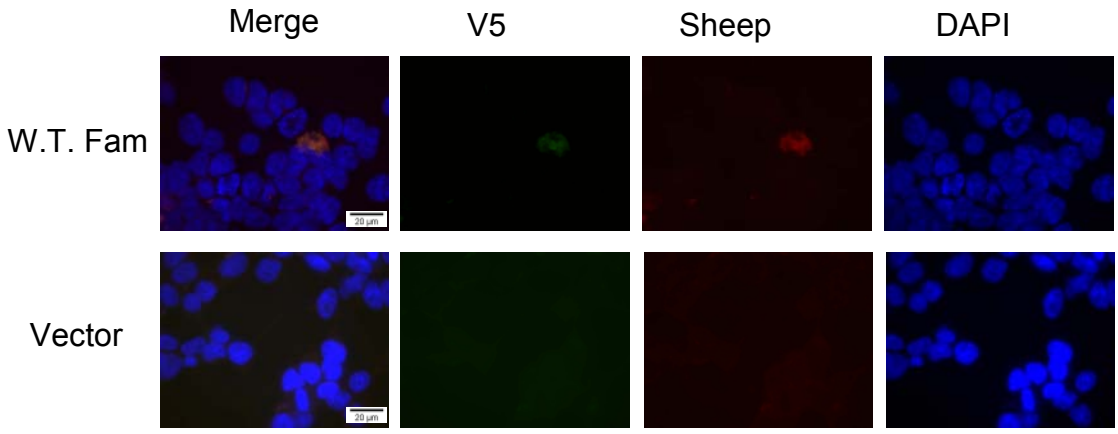

Vector
W.T Fam
Cys Mutant Fam
Vector
W.T. Fam
Cys Mutant Fam
4 Days
5 Days
250 kDa
IB: α Fam N1
250 kDa
IB: α V5
160 kDa
105 kDa
75 kDa
IB: α Lamp1
50 kDa
IB: α β-tubulin

activity of FAM. Cells were transfected, grown for 24 hours to allow for maximum transfection, and then reseeded onto glass coverslips to be grown for a further 3 days. Immunofluorescence staining was performed using primary sheep antibodies raised against each of the four lysosomal proteins to be studied, and an anti sheep Alexa Fluor 555 secondary antibody. The absence of any specific staining in secondary alone control stained cells supports the assumption that the staining pattern seen with each of the antibodies raised against Lamp1, ASA, I2S, and 4S was not due to background staining (Figure 4.12 a). Exogenous expression of wild-type and cys mutant FAM was evident from the anti-V5 immunoreactive staining observed in cells transfected with the expression constructs (green in Figure 4.12 b-f). In 293T cells, FAM showed a diffuse but punctate cytoplasmic staining and was excluded from the nucleus (red staining in Figure 4.12 b). Often an area of localised higher intensity staining was observed to one side of the nucleus and was presumed to be Golgi staining based on previously reported co-immunofluorescence and co-localisation with the GGA proteins (Figure 3.19 and Murray et al., 2004).

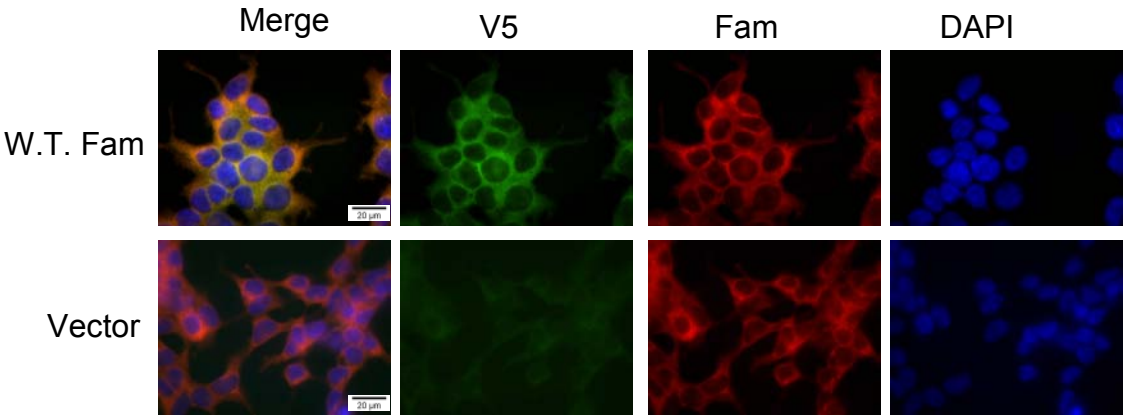
Lamp1 displayed a punctate staining in the cytoplasm with larger vesicular-like structures primarily in a perinuclear pattern. These large round structures are likely to correspond to late-endosomes/lysosomes as described previously (Carlsson et al., 1988; Parkinson-Lawrence et al., 2005). No co-localisation was seen between exogenously expressed FAM (green in Figure 4.12 c) and Lamp1 (red in Figure 4.12 c). No obvious difference in the sub cellular distribution of Lamp1 was observed in cys mutant FAM expressing 293T cells (red in Figure 4.12 c – note, data for co-staining of wild-type FAM with Lamp1 was lost). The staining pattern observed for ASA was similar to that observed for Lamp1 and corresponded to the sub cellular localisation of the protein previously described (Stein et al., 1989; Tanaka et al., 1983). Once again no co-localisation was observed between FAM (green in Figure 4.12 d) and ASA (red in Figure 4.12 d). As with Lamp1, neither exogenous expression of wild-type nor cys mutant FAM had any effect on the sub cellular localization of ASA. I2S and 4S both displayed a punctate cytoplasmic staining pattern with generally smaller puncta than those seen for Lamp1 and ASA staining (Figure 4.12 e & f). Both staining patterns were consistent with the sub cellular localization of I2S and 4S described previously (Coronado-Pons et al., 2004, Parenti, 1987 #864; Peters et al., 1991; Scriver et al., 2001). As with Lamp1 and ASA, little or no co-localisation was seen between either I2S or 4S, and FAM (Figure 4.12 e & f). No change in sub cellular localization was seen for either of I2S or 4S in cells overexpressing wild-type or cys mutant FAM (Figure 4.12 e & f).

Figure 4.12: ***Exogenous expression of FAM in HEK 293T cells has no effect on intracellular distribution or localisation of selected lysosomal proteins*****.** HEK 293T cells were transiently transfected with pDest40 base vector, or an expression vector for V5 epitope tagged wild-type (pDest40 V5 FAM), or cys mutant (pDest51 V5 C1566S FAM) FAM. Cells were grown for 24 hours before being split onto glass coverslips and incubated for a further 3 days. Dual indirect immunofluorescence was performed using polyclonal antibodies directed against the V5 epitope and Lamp1, ASA, I2S, or 4S in combination with anti-mouse alexa 488 (V5 - Green) and anti-sheep alexa 555 (lysosomal proteins - Red) secondary antibodies. **A)** Images show the staining patterns seen when cells were stained with anti-V5 antibody (green), secondary anti-sheep alexa 555 alone (red) and DAPI in wild-type FAM and vector control transfected cells. Images were overlaid in the merge image to visualize any co-localisation **B)** Images show the staining pattern observed when vector transfected cells were probed with anti-V5 (green), anti-FAM (red), and DAPI. Images were overlaid in the merge image to visualize any co-localisation **C)**, **D)**, **E)**, & **F)** Images show the staining pattern observed when cells were probed with anti-V5, anti-Lamp1 **C)**, ASA **D)**, I2S **E)**, or 4S **F)**, and DAPI in wild-type FAM, cys mutant FAM, and vector control transfected cells. Images for each cell type were merged to visualize any co-localisation.

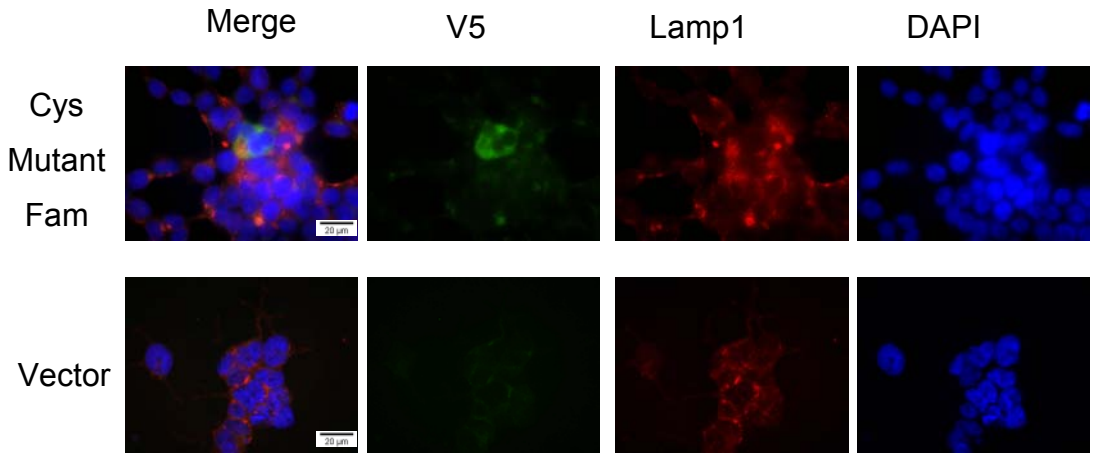
A)


Merge
V5
Sheep
DAPI
W.T. Fam
Vector

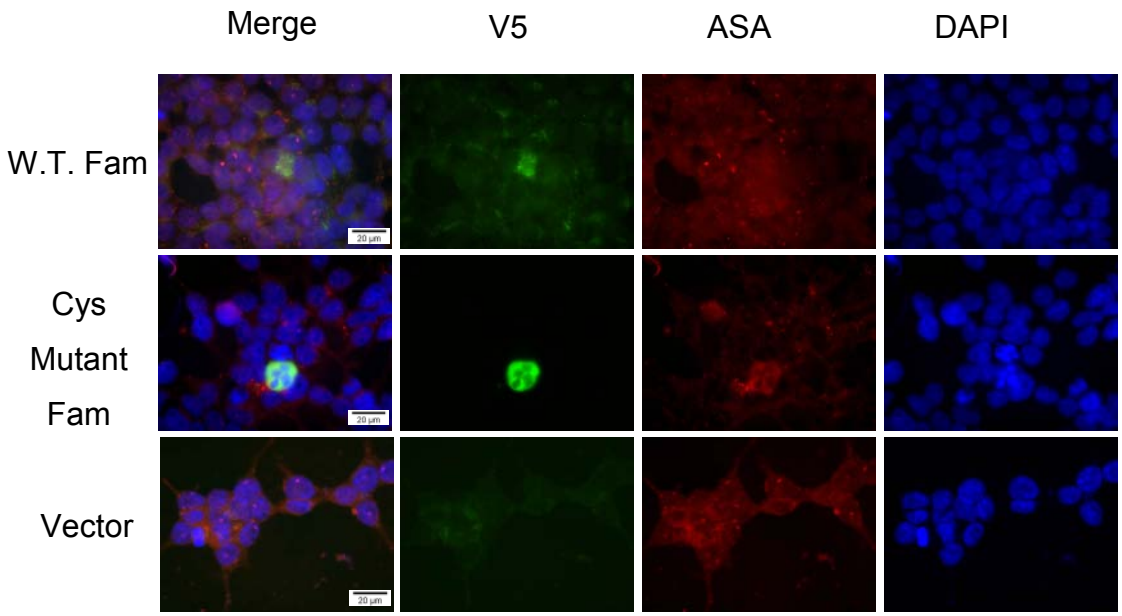
B)


Merge
V5
Fam
DAPI
W.T. Fam
Vector

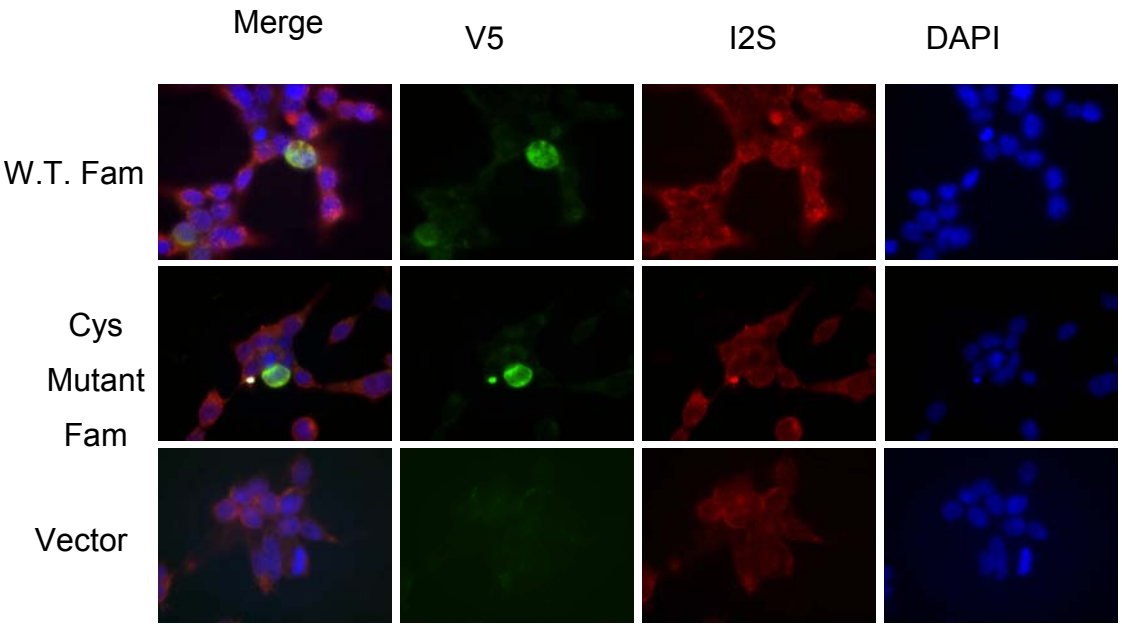
C)

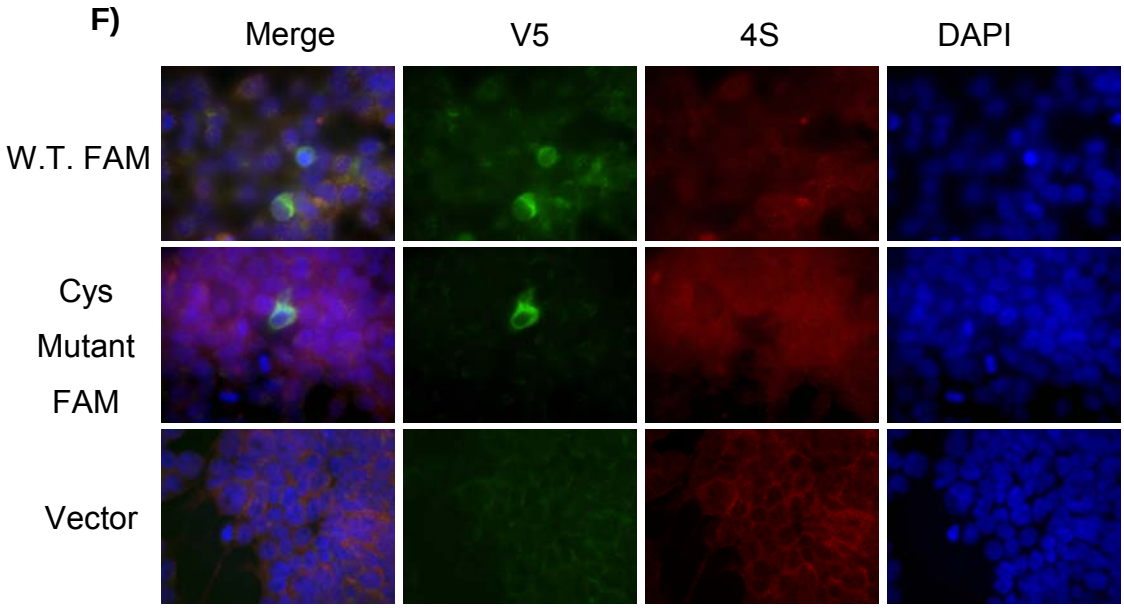

Merge
V5
Lamp1
DAPI
Cys
Mutant
Fam
Vector

D)


Merge
V5
ASA
DAPI
W.T. Fam
Cys
Mutant
Fam
Vector

E)


Merge
V5
I2S
DAPI
W.T. Fam
Cys
Mutant
Fam
Vector


F)
Merge
V5
4S
DAPI
W.T. FAM
Cys
Mutant
FAM
Vector

Reduced FAM Levels – FAM Knockdown

To examine if a reduction in FAM protein levels could have an effect on the cellular localisation and distribution of Lamp1, ASA, I2S, or 4S, 293T cells were transfected with either pRetro Super, or pRetro Super 1800. In each case a vector driving expression of V5-tagged EGFP from an elongation factor promoter, pDest51 EGFP, was co-transfected with the expression vector to allow identification of transfected cells. The cells were grown for 24 hours to ensure maximum levels of transfection before being reseeded on glass coverslips and grown for a further 3 days. Immunofluorescent staining was performed using primary sheep antibodies raised against each of the four lysosomal proteins to be studied, and an anti sheep Alexa Fluor 555 secondary antibody. Once again, the absence of any specific staining in secondary alone control stained cells supports the assumption that the staining pattern seen with each of the antibodies raised against Lamp1, ASA, I2S, and 4S is not simply background staining (Figure 4.13 a). A reduction of endogenous FAM levels in transfected cells was evident from the lower intensity or absence of anti FAM immunoreactivity (red staining in Figure 4.13 b) in transfected cells (indicated by EGFP fluorescence, green staining in Figure 4.13 b). Lamp1, ASA, I2S, and 4S all showed similar staining patterns in untransfected cells to those seen in the FAM overexpression experiments described above (Figure 4.13 c-f). In EGFP positive cells, no gross changes in cellular localisation and/or distribution were observed for any of the lysosomal proteins studied (Figure 4.13 c-f).

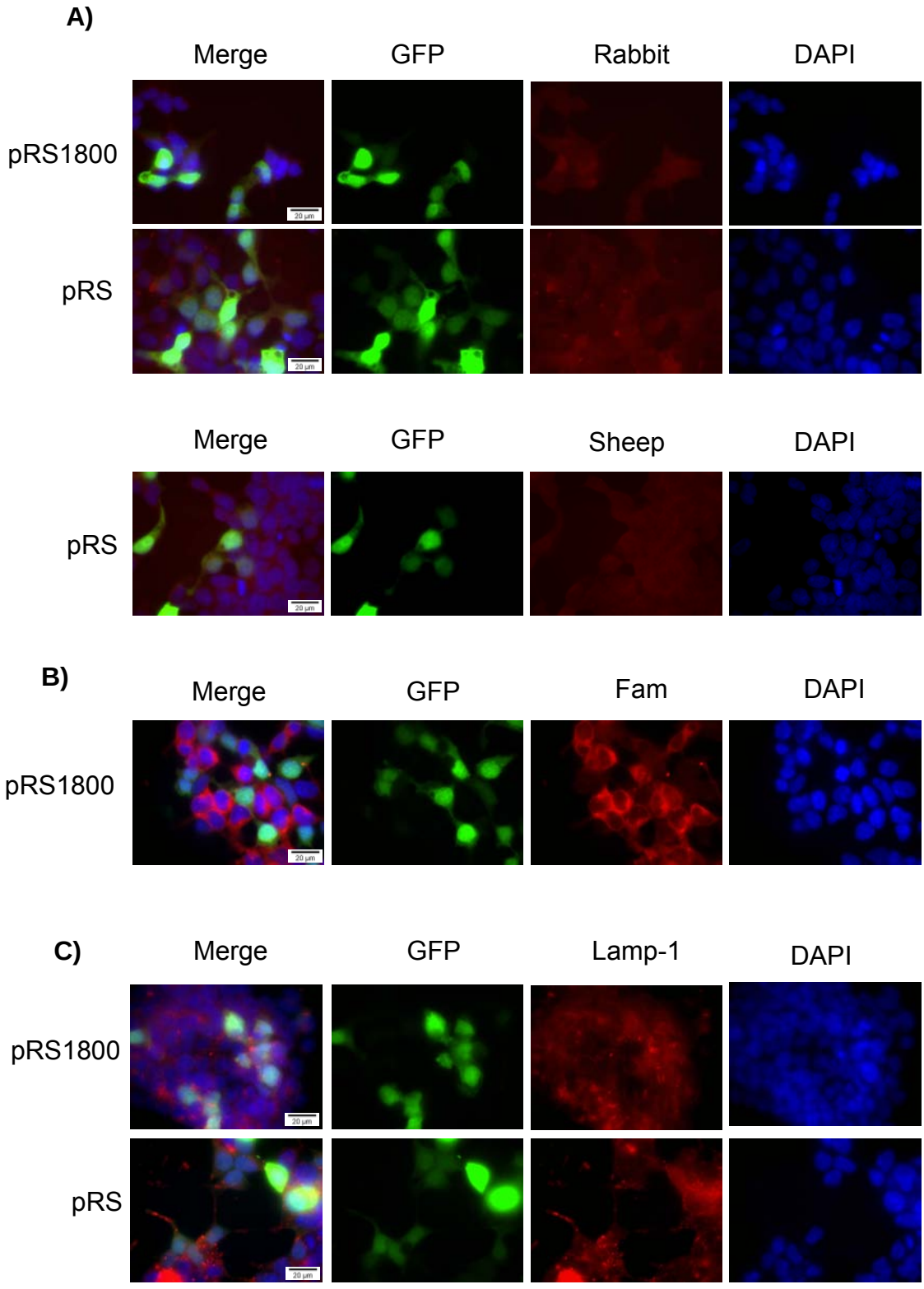
DISCUSSION

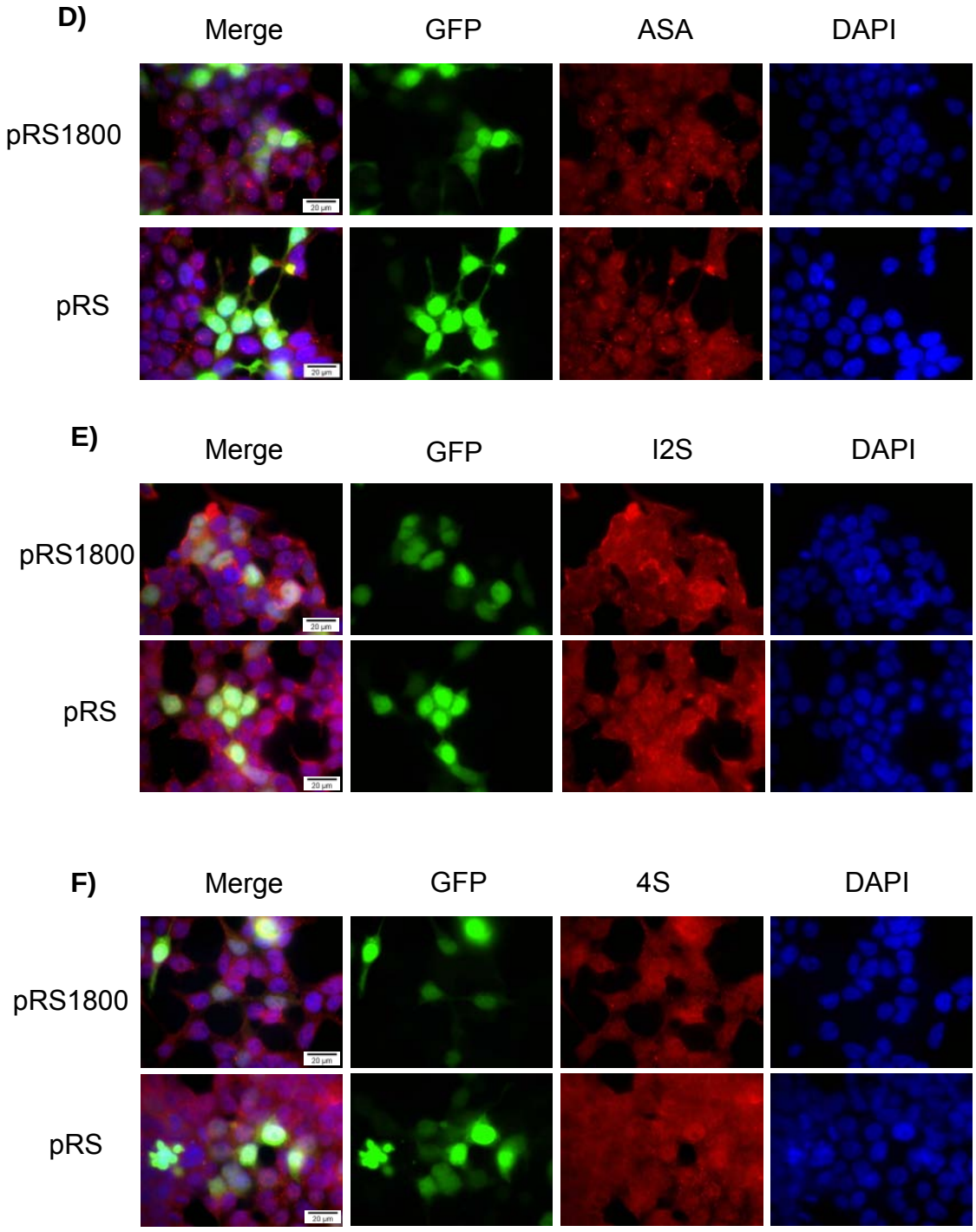
Effect of Altered FAM Levels on Lysosomal Proteins – ELISA Analysis

FAM Overexpression

Data presented in this chapter showed that, at least by the use of simple transient transfection experiments, altering the levels of FAM in cells can cause changes in the total cellular levels and in some cases, the level of processing of lysosomal proteins. Of the eleven lysosomal proteins that were investigated, five showed significant reductions in total cellular protein levels when exogenous FAM was expressed and eight showed significant (based on the assumptions described above) increases when endogenous levels of FAM were reduced by

Figure 4.13: ***shRNA knockdown of FAM in HEK 293T cells has no effect on intracellular distribution or localisation of selected lysosomal proteins***. HEK 293T cells were transiently transfected with pRS base vector, or an expression vector for the FAM shRNA trigger 1800. Cells were grown for 24 hours before being split onto glass coverslips and incubated for a further 3 days. Indirect immunofluorescence was performed using polyclonal antibodies directed against Lamp1, ASA, I2S, or 4S in combination with anti-sheep alexa 555 (lysosomal proteins - Red) secondary antibody. **A)** Images show the staining patterns seen when cells were stained secondary anti-sheep alexa 555 alone, or secondary anti-rabbit Cy3 (red) and DAPI (blue) in FAM shRNA trigger 1800, and vector control transfected cells. Transfected cells were identified by EGFP fluorescence (green). Images were overlaid in the merge image to aid in visualization **B)** Images show the staining pattern observed when FAM shRNA trigger 1800 transfected cells were probed with anti-FAM (red), and DAPI (blue). Transfected cells were identified by EGFP fluorescence (green) Images were overlaid in the merge image to aid in visualization **C), D), E), & F)** Images show the staining pattern observed when cells were probed with anti-Lamp1 **C)**, ASA **D),** I2S **E),** or 4S **F)** (red), and DAPI (blue) in FAM shRNA trigger 1800, and vector control transfected cells. Transfected cells were identified by EGFP fluorescence (green). Images for each cell type were merged to aid in visualization.


A)
Merge
GFP
Rabbit
DAPI
pRS1800
pRS
Merge
GFP
Sheep
DAPI
pRS
B)
Merge
GFP
Fam
DAPI
pRS1800
C)
Merge
GFP
Lamp-1
DAPI
pRS1800
pRS


D)
Merge
GFP
ASA
DAPI
pRS1800
pRS
E)
Merge
GFP
I2S
DAPI
pRS1800
pRS
F)
Merge
GFP
4S
DAPI
pRS1800
pRS

treatment with a FAM specific shRNA. Of the proteins that were unaffected by altered FAM levels, one (Iduronidase) is transported to the lysosome by the mannose-6-phosphate receptors while the other (Saposin-C) reaches the lysosome via an alternative mechanism involving sortilin. The remaining proteins (N-sulphatase, ASM, β-gluc, ASA, GAA, α-Gal, I2S, and 4S) whose cellular levels were affected by altered FAM levels, are all transported to the lysosome by the mannose-6-phosphate receptors, a mechanism that is known to be dependent on the action of the GGA proteins. In addition, while Lamp-1 does not utilize the mannose-6-phosphate transport signal, it has been reported to be transported in AP-1 coated vesicles that are associated with the GGA proteins (Honing et al., 1996). It is possible that FAM may have indirectly affected trafficking of these proteins through its interactions with components of the trafficking machinery, such as the GGA proteins.

One model that may explain the results of these experiments comes from investigations of the intracellular trafficking of the lysosomal-associated protein transmembrane 5 (LAPTM5). Trafficking of LAPTM5 to the lysosome requires association with both the E3 ligase NEDD4 and GGA3 (Pak et al., 2006). The PY motifs (L/PPxY) of LAPTM5 interact with NEDD4-WW domains (Pak et al., 2006). Ubiquitylated GGA3 is also recruited to the complex either already bound to NEDD4 (and possibly ubiquitylated by it) or after the association of LAPTM5 with NEDD4 (Pak et al., 2006). This allows a direct interaction of the UIM domain LAPTM5 with the ubiquitin moiety conjugated to GGA3 and packaging and transport of LAPTM5 to the lysosome in clathrin coated vesicles (Pak et al., 2006). This trafficking shows a requirement for ubiquitylation of GGA3 (Pak et al., 2006) which may suggest that GGA3 remains ubiquitylated and traffics with LAPTM5 to the late endosome/lysosome. It is also possible that deubiquitylation of GGA3 may be required for its retrograde transport from the late endosome/lysosome to the Golgi apparatus and that deubiquitylation of the GGA proteins is mediated by FAM. Consequently, if FAM protein levels were reduced in cells, the GGA proteins may not be able to return to the Golgi apparatus. This in turn might trap lysosomal proteins at the TGN where they would be exposed to glycosylation modifying enzymes for a longer period and hence become more highly processed as observed above. Many of these proteins may then default to the secretory pathway as seen in I-cell disease (Reitman et al., 1981). A further consequence of a block in the retrograde transport of the GGA proteins the Golgi would be the eventual reduction in the lysosomal levels of lysosomal proteins. Conversely, an increase in FAM levels may promote a faster return of the GGAs from the late endosome/lysosome and more rapid transport of lysosomal proteins.

To investigate this model, several experiments could be performed. Firstly, the ubiquitylation status of the GGA proteins needs to be investigated in cells with altered FAM levels. If FAM is responsible for the deubiquitylation of the GGA proteins, exogenous expression of FAM might be expected to lead to a decrease in the abundance of ubiquitylated forms of the GGA proteins and conversely, a reduction in endogenous FAM protein levels would result in an increase in ubiquitylated forms of the GGAs. Additionally, the catalytically dead cysteine mutant of FAM might be able to act as a dominant negative form of FAM. Consequently, transfection of cells with this mutant may lead to a partial or full phenocopy of treatment of cells with FAM shRNA. However, FAM has demonstrated the ability to bind to many proteins and therefore it would be difficult to determine if any observed affects were due to direct regulation of the GGA proteins by FAM. If the interaction of FAM with the GGA proteins could be better characterized, a GGA binding mutant of FAM could be utilized to determine loss of the FAM/GGA interaction could phenocopy the effects on lysosomal protein levels observed when FAM protein levels are reduced in the cell. A GGA binding mutant of FAM would also allow a clearer understanding of the ability of FAM to deubiquitylate the GGA proteins to be gained. Another investigation of the role of GGA ubiquitylation in lysosomal protein trafficking could focus on the GGA proteins themselves. The major site of ubiquitylation of GGA3 by the RING-H2 hVPS18 E3 ligase has been identified as lysine 258 (Yogosawa et al., 2006). It might be possible therefore, to partially phenocopy the effects of treating cells with FAM shRNA by overexpressing a K258R mutant of GGA3 in cells. This investigation would be hampered however, by the potential of the remaining two GGA proteins to compensate for the loss of GGA3 activity.

A second model to explain the results described above may involve the action of the GGA proteins in packaging of cargo at the TGN via GAT domain-ubiquitin interactions. It has been observed that overexpression of the GGA proteins causes a dramatic increase in the amount of ubiquitylated proteins that localise to the Golgi suggesting that the ubiquitin system may be involved in sorting and possibly trafficking on the biosynthetic pathway (Shiba et al., 2004). Furthermore, GGA1 and GGA3 have the ability to bind to ubiquitylated proteins, a property that shows a requirement for the GAT domain of these proteins (Shiba et al., 2004). In addition, ubiquitylation of GGA1 and GGA3 by hVPS18 has been found to inhibit the interaction of the GAT domain with ubiquitin (Yogosawa et al., 2006). It is possible that the ubiquitin binding ability of the GAT domain, and consequently the non-ubiquitylated form, of the GGA proteins is required for the formation of complexes at the

Golgi that mediate the packaging and sorting of cargo destined for the lysosome. If this were the case, then the deubiquitylating activity of FAM may be required to allow the formation of GGA packaging/sorting complexes and thereby indirectly affect the trafficking of proteins from the Golgi to the lysosome. Once again, if deubiquitylation of the GGA proteins was unable to occur due to loss of FAM protein, lysosomal proteins may be prevented from leaving the Golgi apparatus, undergo a higher degree of processing, and their traffic may default to the secretory pathway.

Both of the models described above can be used to explain the results obtained above. In cells with reduced FAM levels, lysosomal protein deficiencies caused by blockade of lysosomal proteins at the Golgi could potentially result in a feedback mechanism whereby uptake and/or *de novo* synthesis of lysosomal proteins is up regulated to compensate. In both models, there are many possible explanations for the higher cellular levels of lysosomal proteins in this situation. In the simple transient transfection experiments described above, approximately 50% of the cells that were analysed were untransfected. Therefore, one possibility is that the increase in cellular lysosomal protein levels that was observed was due primarily to untransfected cells (which began with normal cellular levels of lysosomal proteins) internalising from the media the excess secreted proteins from transfected cells. As transfected cells may themselves have been able to take up the secreted proteins, the net result would have been an increase in total cellular lysosomal proteins in untransfected cells, a return to normal levels in transfected cells, and a net total increase in protein levels across the cellular pool. Alternatively, lysosomal protein levels in untransfected cells may have remained the same with an up regulation of lysosomal protein synthesis, secretion, and re-uptake in transfected cells accounting for the net increase in lysosomal protein levels. Due to the complexity and the dynamic nature of the cellular protein trafficking pathways, there are many more possible explanations for the observations made in the experiments described above.

Similarly, there are many possible explanations for the reduced cellular lysosomal protein levels that were observed in cells expressing exogenous FAM. One explanation could involve the post-translational processing and modification of lysosomal proteins. If the presence of additional FAM protein within the cell accelerates the transport of proteins to the late endosome/lysosome, either by increasing the rate of retrograde transport of the GGAs to the Golgi or by promoting the formation of more GGA mediated packaging/sorting/transport complexes by deubiquitylation, then it is possible that these lysosomal proteins reach the lysosome before they have been processed to their mature form. As such, these immature forms of the proteins may be non-functional and consequently degraded in the lysosome. Pulse chase experiments using ^{35}S methionine labeling of newly synthesised proteins would

provide an indication as to whether the half lives of lysosomal proteins are affected by exogenous FAM expression. There are also bioassays available to determine the functionality of many lysosomal proteins. These assays could be used to determine if there is a difference in lysosomal protein activity in cells expressing exogenous FAM compared to untreated cells. An alternative explanation may be that the expression of excess FAM promoted the targeting of lysosomal proteins to the secretory pathway before they had reached maturity. As a result, the mannose-6-phosphate sorting signal on these proteins may not have been exposed and they may not be able to re-enter the cell resulting in a deficiency of the proteins in cells expressing exogenous FAM. To determine if this was the case, binding assays could be performed using the media from transfected cells (containing the potentially immature secreted lysosomal proteins) and immobilised mannose-6-phosphate receptor. This assay would require a pure population of FAM expressing cells to be utilized.

There were several limitations to the multiplex ELISA system used here to investigate the effects of altered FAM levels on the trafficking of lysosomal proteins. Firstly, as discussed, the assays were performed on cells that were transiently transfected which resulted in a population of cells that consisted of a heterogeneous mixture of transfected and untransfected cells. For this reason, the potential existed that changes in lysosomal protein levels in transfected cells may have been masked by the absence of such changes in untransfected cells. In addition, as many of the lysosomal proteins studied can be secreted from the cell and re-enter the same or neighbouring cell/s, the potential for cross talk between transfected and untransfected cells could not be discounted. This assay system would benefit from a more homogenous population of transfected cells. There are a number of ways in which this could be achieved. One of these methods was explored above. Transfected cells in a transient transfection experiments can be enriched for by utilizing a selection marker (such as GFP expression, or expression of a cell surface marker that is normally absent from the cell line) and flow cytometry. The major disadvantage of this method is that the sorting process is time consuming. This limits the number of samples that can be used in each experiment. In the case of the experiments described above, this limitation prevented replicates from being included in the assay which precluded the use of statistical analysis. Given that experimental variation was observed over the course of the experiments described above, statistical analysis is required to validate data from the multiplex ELISA system. In addition, this method of obtaining a relatively pure population of transfected cells resulted in the acquisition of data that conflicted with the data obtained from simple transient transfection experiments. Although the possibility cannot be ruled out that this data was correct, one possible explanation of the conflicting results may have arisen from the use of GFP as a sorting marker. There are reports in the literature that high level overexpression of GFP can be toxic

to cells (Hanazono et al., 1997; Liu et al., 1999). In addition, unmodified GFP has been reported to interact with the actin binding site of myosin, preventing the actin-myosin interaction (Agbulut et al., 2007). Given the importance of actin-myosin interactions in the motility of trafficking vesicles, it is possible that the high level expression of GFP in the experiments described above resulted in a disruption of actin-myosin interactions and thereby interfered with direct protein trafficking from the Golgi apparatus to the lysosome (Allan et al., 2002; Fath, 2005; Krendel et al., 2007). This may be an explanation for what occurred in the sorted FAM overexpressing cell experiments described above. GFP mediated disruption of the actin-mysosin interaction may have resulted in a masking of the effect of exogenous FAM expression, instead producing data which resembled that observed with a reduction of FAM protein. Ultimately this may mean that enrichment of transfected cells by GFP mediated cell sorting cannot be used in future investigations of FAM function.

Accordingly, a second way to ensure a more homogenous population of transfected cells is the generation of stably transfected cell lines. Several attempts have been made in this laboratory at generating stable cell lines that overexpress FAM protein (M. Scherer, PhD candidate, and S. M. Millard, PhD graduate, The University of Adelaide). In each case the attempts have proven to be unsuccessful and it appears that cells will not tolerate constantly increased levels of FAM protein. An alternative may be to generate a stable cell line that allows inducible overexpression of FAM protein. This would permit homogeneous overexpression of FAM protein for the time periods required for the analysis to be performed. In a similar manner, a stable cell line that allows inducible expression of short interfering RNA triggers for the destruction of FAM message could be used to generate a homogeneous population of cells with reduced FAM protein levels.

A further limitation of the multiplex ELISA system is that, as discussed above, secreted lysosomal proteins can re-enter the cell through the endocytic pathway. As the multiplex ELISA system only measures the levels of total cellular protein, it cannot distinguish between newly synthesised protein that is en-route to the lysosome or has already entered the endosomal/lysosomal system, and protein that has exited the cell and has reentered through the endocytic pathway. As a consequence, if alteration of FAM levels does affect trafficking of lysosomal proteins from the Golgi apparatus, it is not possible with the current assay to determine if FAM is affecting transport via the direct (TGN to lysosome) or indirect (TGN to cell surface to endocytic pathway) routes. Furthermore, if alteration of FAM levels affects the direct route, trafficking of lysosomal proteins may default to the indirect

route (as in I-cell disease) and as a consequence, any blockage in the production or accumulation of lysosomal proteins in the cell would be masked by the protein re-entering and supplementing the cell on the indirect route. One way to overcome this complication would be to regularly change the media on the cells, removing any secreted proteins in the process. Alternatively, most of the proteins analysed in this assay enter the endocytic pathway through interaction with the cation independent mannose-6-phosphate receptor (CI-MPR). Therefore, a second improvement to the assay would be to block cellular re-uptake of lysosomal proteins by adding free mannose-6-phosphate to the media. This would have the effect of saturating the binding sites of cell surface CI-MPRs preventing binding of the exposed mannose-6-phosphate internalisation signal of the lysosomal proteins. A third possibility would be to combine these two methods.

Despite the considerations discussed above, the aim of the experiments described in this chapter was to determine if alteration of FAM levels can have a global (TRANS) affect on cellular protein trafficking and on protein trafficking from the Golgi apparatus in particular. While it cannot be said conclusively that alteration of FAM levels has an affect on trafficking from the Golgi, the experiments described in this chapter do suggest that these alterations can have a global affect on lysosomal protein levels and maturity. Furthermore, it can be said that this affect is likely to occur at the level of the trafficking/sorting machinery rather than cargo as FAM is unable to directly interact with any of the proteins investigated. Given the cellular localisation of FAM, the observed changes in FAM localisation in lysosomal storage disease patient cells or upon GGA overexpression, and the potential interaction of FAM with the GGA proteins, there is evidence that FAM does in fact affect trafficking at the Golgi apparatus. A number of experiments that could be conducted in future to further elucidate the role of FAM in TGN sorting and/or trafficking will be discussed below.

Fluorescence microscopy was undertaken to determine if increased or decreased FAM protein levels can have an effect on the sub cellular localisation and distribution of the four lysosomal proteins that showed changes in cellular protein levels and, in the case of Lamp1 and ASA, processing. No large scale changes in sub cellular localisation or distribution of any of the four lysosomal proteins could be detected by epi-fluorescence microscopy. However, since 293T cells have quite a large nuclear to cytoplasm ratio, and given that the sensitivity of epi-fluorescence microscopy is quite low, disruption of the cellular distribution of these proteins due to altered FAM levels cannot be discounted. There are several ways in which the investigation could be improved to more clearly establish if changes in cellular localisation/distribution do or do not occur. Firstly, the studies could be conducted in a cell line in which the nuclear to cytoplasm ratio is smaller such as Cos7 or CHO cells. Secondly,

fixed and stained cells could be analysed by confocal microscopy to allow for a more sensitive detection of the areas of the cell in which the proteins are localised. Confocal microscopy also has the advantage that fluorescence is measured one focal plane at a time allowing a three dimensional picture of the cell to be obtained.

This investigation could also benefit from other forms of microscopy to determine the effects of altered FAM levels on the sub cellular distribution of lysosomal proteins. Electron microscopy using gold-labelled antibodies raised against lysosomal proteins of interest would allow a much higher resolution picture of the distribution of these proteins to be obtained in cells with altered FAM levels. Electron microscopy would also allow a greater understanding of the precise sub cellular distribution of FAM in cells to be gained if similar studies were done using gold-labelled anti-FAM antibodies. Live cell imaging in combination with fluorescence recovery after photo bleaching (FRAP) is another technique that could provide a clearer picture of the effects of altered FAM levels on the trafficking of lysosomal proteins. By using this method, the routes taken to the lysosome by lysosomal proteins of interest could be investigated in real-time in normal cells and in cells where FAM levels have been altered. Furthermore, there are many biomarkers now available for tracking changes in the trafficking pathways of the cell. LysoTracker and LysoSensor probes (Invitrogen) are available and can be used to label the lysosome. This enables real-time imaging of the transport of GFP-tagged proteins from the Golgi apparatus to the late endosome/lysosome. Used in combination with exogenous expression or reduced expression (using shRNA) of FAM, this would allow a better understanding of the role of FAM in trafficking of proteins from the Golgi apparatus to the lysosome.

Future investigations of the function of FAM in protein trafficking should also make use of the information available about the Rab GTPase proteins and dominant negative mutants of these proteins. The family of Rab GTPases are the largest division (there are over 60 members in mammals) of the Ras GTPase superfamily. The nucleotide-dependent switch that is common to the superfamily, is utilized by the Rab proteins in association with a large list of effector molecules, to regulate the four processes required at each stage of protein trafficking; vesicle budding, vesicle transport, vesicle docking, and vesicle fusion (Grosshans et al., 2006). The precise function of the majority of the Rab GTPases has yet to be determined (Somsel Rodman and Wandinger-Ness, 2000). Five members of the family Rab4, Rab5, Rab7, Rab11 and Rab9, have been well characterized and can be used in investigations of intracellular protein trafficking at all points along the pathway.

Rab5 is essential for the packaging of ligands into clathrin coated pits at the cell surface (Christoforidis et al., 1999; McLauchlan et al., 1998) and the homotypic fusion of endocytic vesicles with the early endosome (Barbieri et al., 1996; Bucci et al., 1992; Gorvel et al., 1991). A GTP binding mutant (dominant negative) of Rab5, Rab5S34N, reduces endocytosis by 50% and blocks trafficking of proteins at the early endosome without affecting the recycling pathway (Bucci et al., 1992). Rab4 acts as the regulator of recycling at the early endosome by targeting and packaging ligands into both recycling vesicles on the fast recycling pathway or to the recycling endosome on the slow pathway (Daro et al., 1996; McCaffrey et al., 2001; Mohrmann and van der Sluijs, 1999). A dominant negative mutant of Rab4, Rab4S22N, prevents the recycling of ligands by blocking their exit from the early endosome (McCaffrey et al., 2001). Traffic to and from the intermediate compartment on the slow recycling route, the recycling endosome, is controlled by Rab11 (Wilcke et al., 2000). The dominant negative mutant Rab11 S25A blocks the transit of proteins from the recycling endosome to the plasma membrane (Calhoun et al., 1998; Duman et al., 1999). In addition, the Rab11 S25A mutant appears to block both transport from the recycling endosome to the Golgi apparatus and reciprocally, from the Golgi to the recycling endosome (Chen and Wandinger-Ness, 2001; Wilcke et al., 2000). Rab7 function is essential for transport from the early endosome to the late endosome (Feng et al., 1995; Mukhopadhyay et al., 1997). Overexpression of dominant negative mutants of Rab7, N125I and T22N, results in a significant inhibition of transport from the early endosome to the late endosome (Feng et al., 1995; Mukhopadhyay et al., 1997). Finally Rab9 appears to be involved in retrograde transport of proteins from the late endosome to the Golgi apparatus. A dominant negative mutant of Rab9, Rab9 S21N, prevents recycling of the mannose-6-phosphate receptors from the late endosome to the Golgi resulting in a decrease in the efficiency of lysosomal sorting (Riederer et al., 1994). Interestingly, cells overexpressing Rab9 S21N compensated for the loss of lysosomal trafficking from the Golgi apparatus by up regulating the *de novo* synthesis and uptake of secreted lysosomal proteins (both soluble and membrane associated), a phenomenon observed in cells with reduced FAM levels (Riederer et al., 1994). In addition, overexpression of wild-type or expression of constitutively active (active site glutamine Q to L, GTPase deficient) mutants of each of the Rab proteins described above (except Rab9 Q66L mutant whose effects have not been investigated), has been found to increase the rate of exit from the cellular compartment controlled by each respective Rab protein described here (Bucci et al., 1992; McCaffrey et al., 2001; Mukhopadhyay et al., 1997; Wilcke et al., 2000).

The Rab4, 5, 7, 9 and 11 proteins could be used in investigations of FAM function in protein trafficking in two ways. Firstly they could be used to better characterise which cellular compartments FAM can be found associated with. This could be achieved in real-time, with co-expression of GFP/CFP/YFP-labeled fusions of the wild-type Rab proteins and CFP/YFP/GFP-labeled fusions of FAM, or statically, by blocking entry into and/or exit from each compartment by expression of the dominant negative mutants of each Rab protein. Secondly, they could be used to track the trafficking of lysosomal proteins to each compartment when FAM levels are altered. This could be achieved by co-expression of GFP /CFP/YFP-fusions of each of the Rab proteins individually, with CFP/YFP/GFP fusions of selected lysosomal proteins in the presence of increased or decreased FAM levels. Finally, given that expression of the dominant negative mutant of Rab9 was seen to have similar effects on lysosomal protein levels as were seen with a reduction in FAM levels, an investigation could be carried out to determine whether overexpression of wild-type Rab9 or Rab9 Q66L GTPase deficient mutant results in a similar decrease in total cellular lysosomal protein levels observed with exogenous expression of FAM.

CONCLUSION

The work presented in this thesis has resulted in the identification and partial characterization of a novel domain in the deubiquitylating enzyme FAM/USP9X. The ubiquitin-like domain is located in the N-terminal half of protein (outside of the catalytic core of the enzyme) and appears to be involved in protein-protein interactions. By investigating the abilities of this domain to mediate protein-protein interactions, a novel interaction between FAM and the Golgi adapter proteins GGA1, GGA2, and GGA3 was identified. While it remains to be determined if this interaction occurs *in vivo*, the known properties and the co-localisation of these proteins in cells suggest strongly that they will be associated with each other.

The identification of this interaction provided strengthening evidence of the involvement of FAM in protein trafficking including on the biosynthetic route, a function for FAM that has been suggested from investigations into its role in regulating β-catenin, epsin-1, and the E3 ligase itch. The identification of this novel interaction also prompted further investigation of the role of FAM in trafficking of proteins from the Golgi apparatus. These experiments investigated the effect on the cellular protein levels of eleven lysosomal proteins, of increasing or decreasing the levels of FAM protein in the human fibroblast cell line, HEK

293T. It was found that exogenous expression of FAM resulted in an overall decrease in protein levels of many lysosomal proteins, while a reduction in cellular levels of FAM resulted in a general increase of the same proteins. Interestingly, a loss of Rab9 function resulting in an inability of the mannose-6-phosphate receptors to return to the Golgi apparatus from the late endosome was shown to produce a similar increase in lysosomal protein levels to that observed in FAM knockdown experiments. This suggests strongly that FAM function may be required for transport of proteins from the Golgi to the late endosome. While the precise mechanism of FAM function in this context remains to be elucidated, uncovering the role of the association of FAM with the GGA proteins is very likely to provide many key answers.