Expression and purification of CFM2 and Filamin A repeat 10 domain

A thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Biochemistry at Massey University, Manawatu, New Zealand.

Ben Waite

2012

Abstract

Filamins are a group of proteins that interact with over 60 other proteins. Mutations to the Filamin A gene results in a spectrum of disorders including Otoplatodigital spectrum disorder type 1, Otoplatodigital spectrum disorder type 2, Frontometaphyseal dysplasia, Melnick-Needles syndrome and Periventricular Nodular Heterotopia. All cases of Melnick-Needles syndrome can be accounted for by mutations in repeat 10. Using a yeast-2 hybrid assay Professor S.P Robertson identified the protein FAM101A (the protein is alternatively named CFM2) that associated with Filamin A repeat 10. CFM2 was found to interact with itself in a yeast-2-hybrid screen, suggesting homo-dimerisation properties in addition to Filamin A repeats 10 and 21 binding affinity. If CFM2 dimerises and binds to repeat 10 and 21 it is possible that Filamin A's function will alter, thus altering the properties of the cytoskeleton. To investigate the interaction between Filamin A repeat 10 and CFM2, each was subcloned into an *E.coli* plasmid vector fused to a purification tag. Purification of CFM2 failed due to misfolding, this upholds later work that claims CFM2 cannot fold correctly without the presence of vertebrate Filamin. Filamin A repeat 10 purification went well but the fusion was unable to be concentrated without precipitating out of solution. Also the GST purification tag could not be cleaved without secondary cleavage products forming. Pull-down of C2C12 mouse fibroblast cell lysate using the GST-Filamin A repeat 10 fusion as the probe did not identify any other proteins that bind Filamin A repeat 10.

Acknowledgements

I would like to thank my supervisor Dr Andrew Sutherland-Smith for his support and guidance throughout my research.

A big thank-you to the post-graduates of the Institute of Molecular Biosciences, with special thanks to the post-graduates of Dr Andrew Sutherland-Smith and Dr Gill Norris. Being able to get help and share the success and failures of our respective projects makes the failures more bearable and the successes sweeter.

Thanks to my family. Your love and support has been a great motivator and enabler for me.

Abbreviations

AMP	ampicillin
LB	Luria broth
E.coli	*Escherichia coli*
MQ	milli-q
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
PCR	polymerase chain reaction
APS	ammonium persulfate
SDS	sodium dodecyl sulfate
DTT	dithiothreitol
DNase	deoxyribonuclease
cDNA	complementary deoxyribonucleic acid
TEMED	N'-tetramethylethylenediamine
Tris	tris (hydroxymethyl) aminomethane
BME	2-Mercaptoethanol
PPU	Precision plus protein unstained (Bio-rad)
FLNAR10	Filamin A repeat 10
IPTG	isopropyl β-D-1-thiogalactopyranoside
dNTP	deoxyribonucleotide triphosphate
MSC	mesenchymal stem cell
ABD	actin-binding domain
OPD1	Otoplatodigital spectrum disorder type 1
OPD2	Otoplatodigital spectrum disorder type 2
FMD	Frontometaphyseal dysplasia
MNS	Melnick-Needles syndrome
F-actin	filamentous actin
G-actin	globular actin
PVNH	Periventricular Nodular Heterotopia
EtBr	ethidium bromide
Ig	Immunoglobulin

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Cytoskeleton

The cytoskeleton is the cellular scaffold or skeleton present inside all eukaryotic cells. The cytoskeleton of non-muscle cells is not a rigid structure, but instead it is a pool of globular monomeric actin and filamentous actin dynamically polymerising. This dynamic shift from free to polymerised actin allows for large shifts in viscosity of the cytosol and changes in the elastic and mechanical resistance of the cell. By shifting the actin from free monomers to filament and by altering how bundled the filaments are cells are able to stretch, divide, contract and move. Branching of actin filaments is mediated by the ARP 2/3 complex (an ~220kDa complex of 7 proteins which initiates polymerization of new actin filaments). Actin achieves this cellular remodelling by depolymerisation at the negative end and polymerised more quickly adding monomers at the positive end; through this actin fibres can generate force. The cytoskeleton is thus vital to cell division and motility.

The control of monomeric globular actin (G-actin) to polymerised filamentous actin (F-actin) as well as the linkages between individual filaments is of direct importance to the activity, morphology and survival of the cell. Proteins with an actin binding domain (ABD), bind to F-actin. The binding can mediate anchoring to cadherin, a group of proteins involved in anchoring cells to one another; they also mediate the localization of proteins involved in signalling.

A group of F-actin promoting and cross-linking proteins contain a conserved ABD[2]. These proteins can form dimers or tetramers which allows them to cross-link F-actin, bundling the network at high concentrations and changing the mechanical properties at a macroscopic level[3]. One member of this group is unique; Filamin A (FLNA) has been shown to cross-link actin at right-angles [4-6]. These right angled networks are characteristic of the actin structure at the leading edge of a migrating cell [5]. The requirement of FLNA in the creation of orthogonal networks of actinin such a way for cells to produce useful locomotion has been demonstrated in human melanoma cells[7]. Filamin cross-linking is sensitive to concentration, at low molar concentrations

cross-linking is orthogonal (right angled) at high concentrations cross-linking is bundled [4].

1.2 Intramembranous Ossification

Intramembranous ossification is one of the two foetal processes during development through which bone tissue is created. Mesenchymal stem cells (MSC) are the precursor cells to the fibroblast cell lines. Some of these fibroblasts migrate to locations in a developing embryo and form aggregations called ossification centres where the cells will develop and differentiate into osteoblasts which produce bone from mesenchyme connective tissues (the intramembranous ossification pathway).

The craniofacial skeleton and the clavicle bone are both formed through osteoblast mediated intramembranous ossification [8], both are affected in the Otoplatodigital spectrum disorders(OPD1 and OPD2)which are linked by mutations in the FLNA gene.

1.3 Otoplatodigital syndrome spectrum disorders

These disorders include Otoplatodigital spectrum disorder type 1, Otoplatodigital spectrum disorder type 2, Frontometaphyseal dysplasia and Melnick-Needles syndrome (OPD1, OPD2, FMD and MNS). Due to the similarities in their phenotypes, these disorders have commonly being believed to have a similar genetic origin[9]. The only locus with all of the OPD spectrum disorders associated is the gene FLNA(found on the X chromosome)[10], which codes for a cytoskeletal protein.

Another disease which shows some overlap is Periventricular Nodular Heterotopia (PVNH), while this neuronal migration disorder is associated with epilepsy and cognitive disability there is no skeletal dysplasia. The mutational mechanism for PVNH appears to be a loss of FLNA expression, through frame shift resulting in truncation [11], resulting in non-migration of brain neurons. More than 25 FLNA mutations (mostly truncations) have been identified in humans that suffer from Periventricular Heterotopias, these mutations result in disruptions in cellular cytoskeleton and impairment of cell mobility.

Otoplatodigital syndrome spectrum disorders display broad range congenital defects from: craniofacial structures, brain, skeleton, viscera and urogenital tract. Mutations

associated with the OPD spectrum disorders maintain the reading frame unlike PVNH. OPD spectrum disorders do not have a higher than normal rate of seizures as seen in PVNH. From this it has been suggested that they are separate disorders. Importantly all cases of MNS can be accounted for by mutations in repeat 10, this mutation results in a gain-of-function phenotype (X-linked dominant) [10, 12, 13].

1.4 Filamins

Filamin A was originally purified from chicken gizzard [14] in 1975. Although the sequence was unknown at the time it was hypothesised that due to the high viscosity it demonstrated at high and low temperatures it probably had an extended structure. This was confirmed in the following 5 years [15, 16].

Filamin forms a homodimer by binding tail to tail through a dimerisation domain located at repeat 24 [17] (or repeat 6 in Dictyostelium Filamin(ddfilamin) [18]). The repeats consist of a β-sandwich [17, 19]. Each subunit of Filamin is 250kDa and forms a V shaped elongated structure [16, 20]. The N-terminal region of Filamin A (aa 1-264) contains an ABD, this domain consists of two calponin homology domains (CH1 and CH2) [21]. The CH1 domain contains a unique calmodulin binding site which competes for binding with F-actin [21]. Importantly two short hinge regions have been located in Filamin but not ddfilamin which has been postulated to be flexible [5], recombinant cells expressing FLNA missing the hinge region showed reduced flexibility under stress[22]. A homodimer of FLNA is represented in Figure 2. The structure of FLNA repeat 10 has been determined[1], the structure is shown in Figure 1.

The mutations that correlate with MNS and FMD would seem to disrupt hydrophobic packing within the loops but whether the loops disrupted are necessary for ligand binding or whether it is simply a lowering of stability or even impairment of translation is unclear. The loops in class A Immunoglobulin (Ig) repeats site the C and D β-strands to create a ligand binding site for unstructured motifs. FLNAR10 does not belong to the class A Ig repeats, which have been identified as ligand binding domains, as such it doesn't bind class A binding ligands. FLNAR10 is a class D Ig which have a two residue insert in the CD loop that causes a different conformation to occur, rationalising why class D Ig cannot bind the class A Ig ligands.

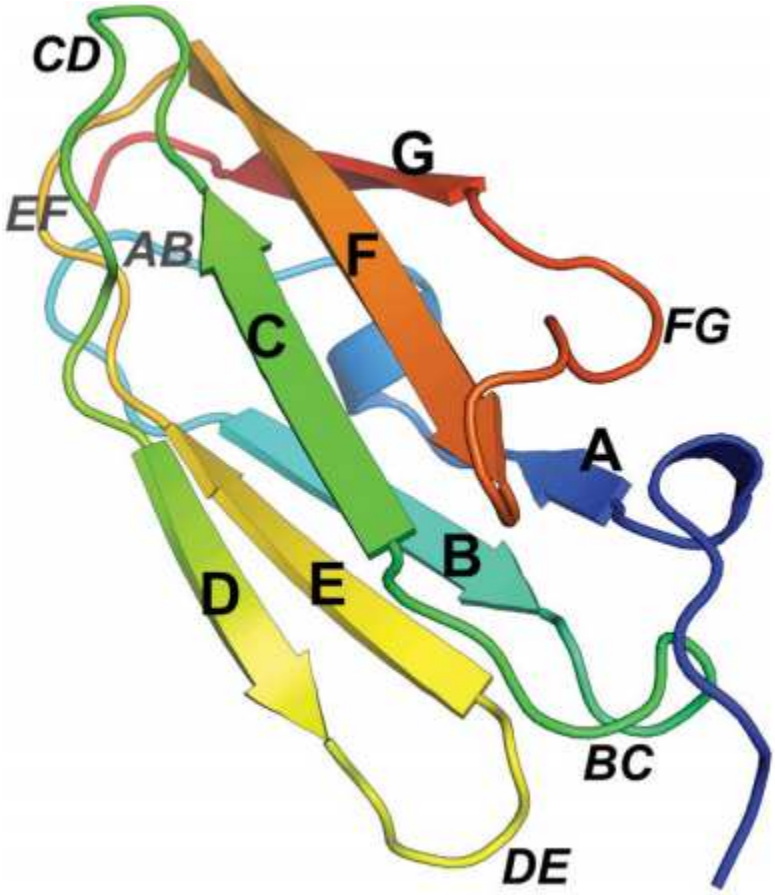


Figure 1: Structure of Filamin A repeat 10

Ribbon diagram of human FLANR10 as determined by [1]. B-strands are labelled A-G, loops are also labelled. Ribbon is coloured from the N-terminus (blue) to the C-terminus (red).

Filamins interact with over 60 other proteins[23], direct binding assays and yeast-2-hybrid assays have identified most of these proteins[5]. The functions of proteins that bind FLNA are incredibly varied i.e. GpIb/V/IX (VonWillebrand receptor) complex promotes cell spreading and Granzyme B participates in Granzyme B-mediated apoptosis.

Gelation of F-actin is a major feature of all isoforms of Filamin including ddfilamin. Being able to cross-link actin at right angles reduces the number of molecules required for cross-linking; by contrast Arp2/3 produces daughter filaments with a strict orientation [6, 17]. Filamin has been described, along with the α-actinin super family, as “actin saving proteins” as they reduce the molar requirement of actin necessary for gelation several fold [24].

While FLNA has a large number of proteins it is associated with as binding partners and MNS is associated with mutations in repeat 10 of FLNA, no protein had previously been found to associate with repeat 10. Using a yeast-2 hybrid assay Professor S.P Robertson (University of Otago) identified the protein FAM101A that associated with FLNAR10, the protein is alternatively named CFM2.

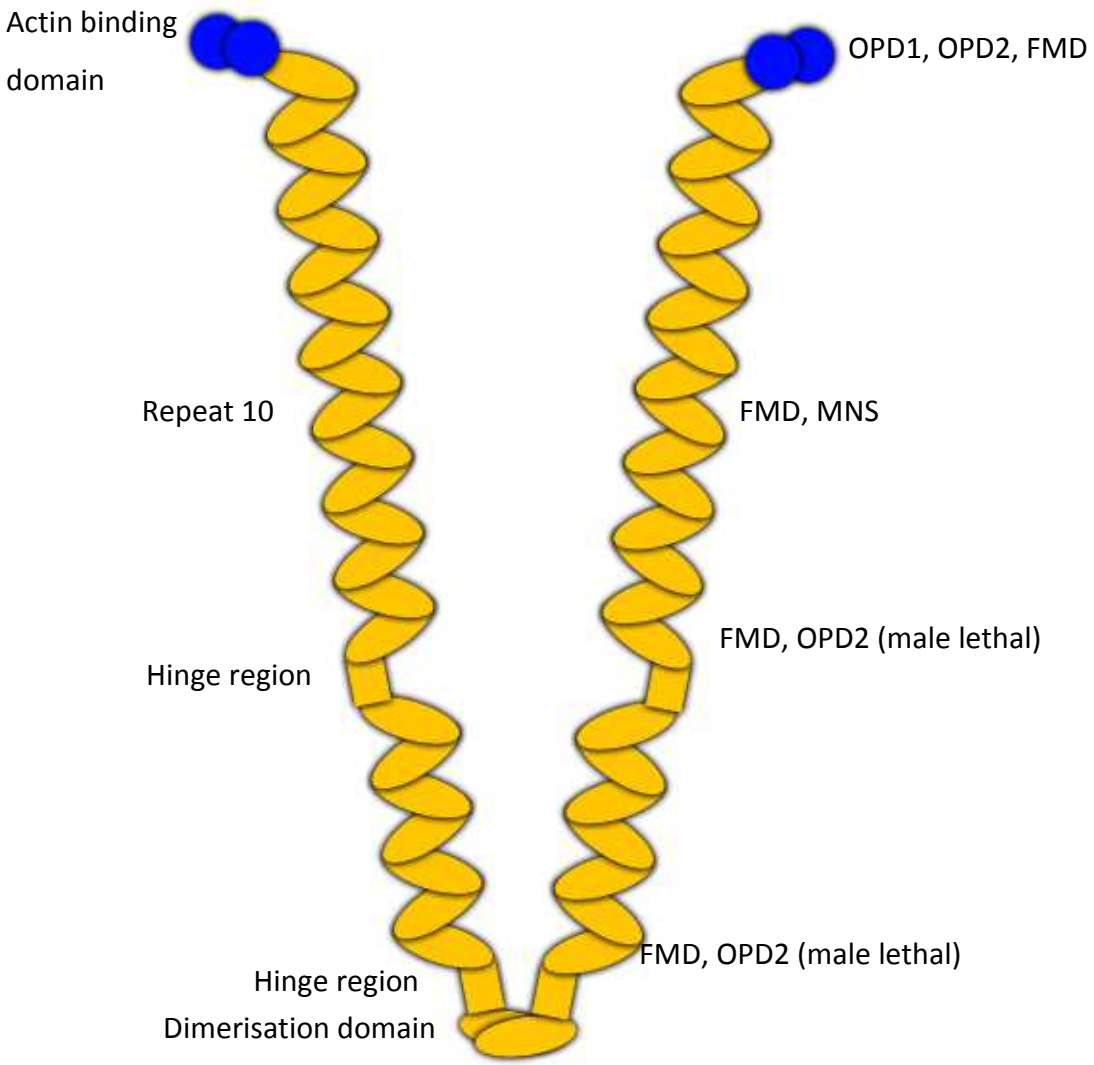


Figure 2: Diagram of Filamin A dimer

The blue regions represent the actin binding domains. On the right, clusters of mutations belonging to the OPD spectrum disorder listed. On the left protein domains are shown, repeats begin at 1 after the ABD.

1.5 CFM2

A yeast-2-hybrid analysis isolated an uncharacterised protein CFM2. Using CFM2 in a yeast-2 hybrid assay, it has been shown to bind to FLNA at repeat 10 (mutation hotspot of gain-of-function mutations) (unpublished from Professor S.P Robertson) and 21(binding site of Integrin β1)[25]. CFM2 was found to interact with itself in a yeast-2-hybrid screen, suggesting homo-dimerisation properties.

While CFM2 and Integrin β1 may compete for repeat 21 binding, it is currently unclear what effect this may have. Integrin β1 binds each Filamin protein dimer 14 times [26]. As CFM2 is only binding 1 of the 7 integrin β1 binding sites, it is unlikely to prevent clustering of integrin receptors. Other modifiers of integrin β1 binding or other proteins that bind FLNA repeat 21 may alter CFM2 binding to FLNA repeat 21, switching on or off CFM2 activity.

It has also been shown that CFM2 is expressed only in foetal skeletal cells that create bone via the intramembranous pathway.

The two isoforms of CFM2 (Uniprot Q6ZTI6 and Q6ZTI6-2) were considered for study at the beginning of this project. An alignment of both isoforms is shown in Table 1. The first and larger isoform has 81 more aminoacids at the N-terminal than the second isoform. To determine the predicted change to the secondary structure a Chou & Fasman algorithm prediction was performed [27]shown in Figure 3. As the two isoforms differ only in the beginning in a single descrete region and the secondary structure differs even on the regions that have identical aminoacid sequence it suggests that one of them may not be a product natively produced; that product is non-functional.

Table 1: Alignment of CFM2 isoforms

Sequences were aligned at http://www.uniprot.org/align/

```
1    MVGHLHLQGMEDSLKEQGREGLLDSPDSGLPPSPSPSPPFYSLAPGILDARAGGAGASSE   60  Q6ZTI6   F101A_HUMAN
1    ------------------------------------------------------------    0  Q6ZTI6-2 F101A_HUMAN

61   PPGPSEARAPPSQLPNPPASEMRPRMLPVFFGESIKVNPEPTHEIRCNSEVKYASEKHFQ  120  Q6ZTI6   F101A_HUMAN
1    --------------------MRPRMLPVFFGESIKVNPEPTHEIRCNSEVKYASEKHFQ   39  Q6ZTI6-2 F101A_HUMAN
                         ************************************

121  DKVFYAPVPTVTAYSETIVAAPNCTWRNYRSQLTLEPRPRALRFRSTTIIFPKHARSTFR  180  Q6ZTI6   F101A_HUMAN
40   DKVFYAPVPTVTAYSETIVAAPNCTWRNYRSQLTLEPRPRALRFRSTTIIFPKHARSTFR   99  Q6ZTI6-2 F101A_HUMAN
     ************************************************************

181  TTLHCSLGRPSRWFTASVQLQLCQDPAPSLLGPATL  216  Q6ZTI6   F101A_HUMAN
100  TTLHCSLGRPSRWFTASVQLQLCQDPAPSLLGPATL  135  Q6ZTI6-2 F101A_HUMAN
     ************************************
```

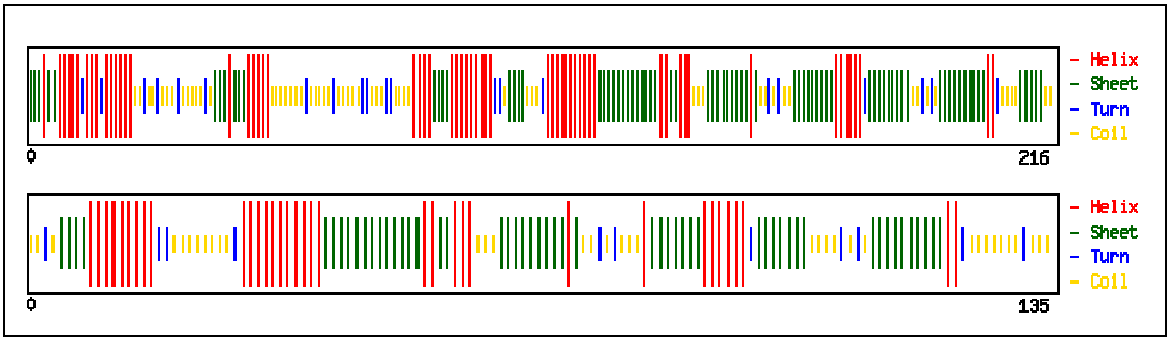


Figure 3: Chou & Fasman algorithm prediction of Secondary structure of CFM2 isoforms

Prediction performed at http://www.biogem.org/tool/chou-fasman/ using sequences shown in appendix (Table 20). Top Q6ZTI6 isoform. Bottom Q6ZTI6-2 isoform.

1.6 Filamin A repeat 10 - CFM2 interaction hypothesis

If CFM2 dimerises and binds to repeat 10 and 21 it is possible that FLNA cross-linking and activity will change; altering the properties of the cytoskeleton. The fact that the mutations that produce MNS are dominant (gain-of-function) mutations suggests the role of CFM2 is as a repressor of normal FLNA function and that this function must be turned off to produce normal cell migration necessary to the formation of intramembranous ossified bone. If this is achieved through the binding of CFM2 then the binding capability of FLNAR10 with CFM2 may be lost in mutated FLNA.

The proposed mechanism of action of CFM2 is shown in Figure 4. A homo-dimer of CFM2 binds repeats 10 and 21 of Filamin A, causing FLNA to fold over at the hinge region between repeats 16 and 17. The Filamin A monomer may lose its dimerisation function or actin binding function causing a change in the bundling of actin through effective Filamin A concentration changes. FLNA mediated clustering of integrin receptors may also be altered should CFM2 function result in large changes to FLNA structure. Filamins are cross-linking proteins and being the only group to cross-link orthogonal networks CFM2 activity may locally reduce cross-linking lowering mechanical stability or alter it in another way. A suggestion from a more recent paper than this project [28] suggested CFM2 would promote bundling of FLNA leading to parallel bundles of actin filaments. Should a mutation in repeat 10 result in the loss of CFM2 homo-dimer binding then this would explain the gain-of-function mutations that cause MNS.

During this project *E.coli* will be transformed with plasmids containing CFM2 or FLNA repeat 10. Both will have some form of affinity tag for purification purposes.

The aim will be to retrieve pure native folded CFM2 and FLNAR10, with these the homo-dimerisation properties of CFM2 can be tested and the binding interaction between CFM2 and FLNAR10 can be replicated using pull-down assays and co-purification by FPLC. Ideally X-ray crystallography of CFM2 and FLNAR10 binding sites could be obtained and the mechanism through which MNS mutations act could be investigated.

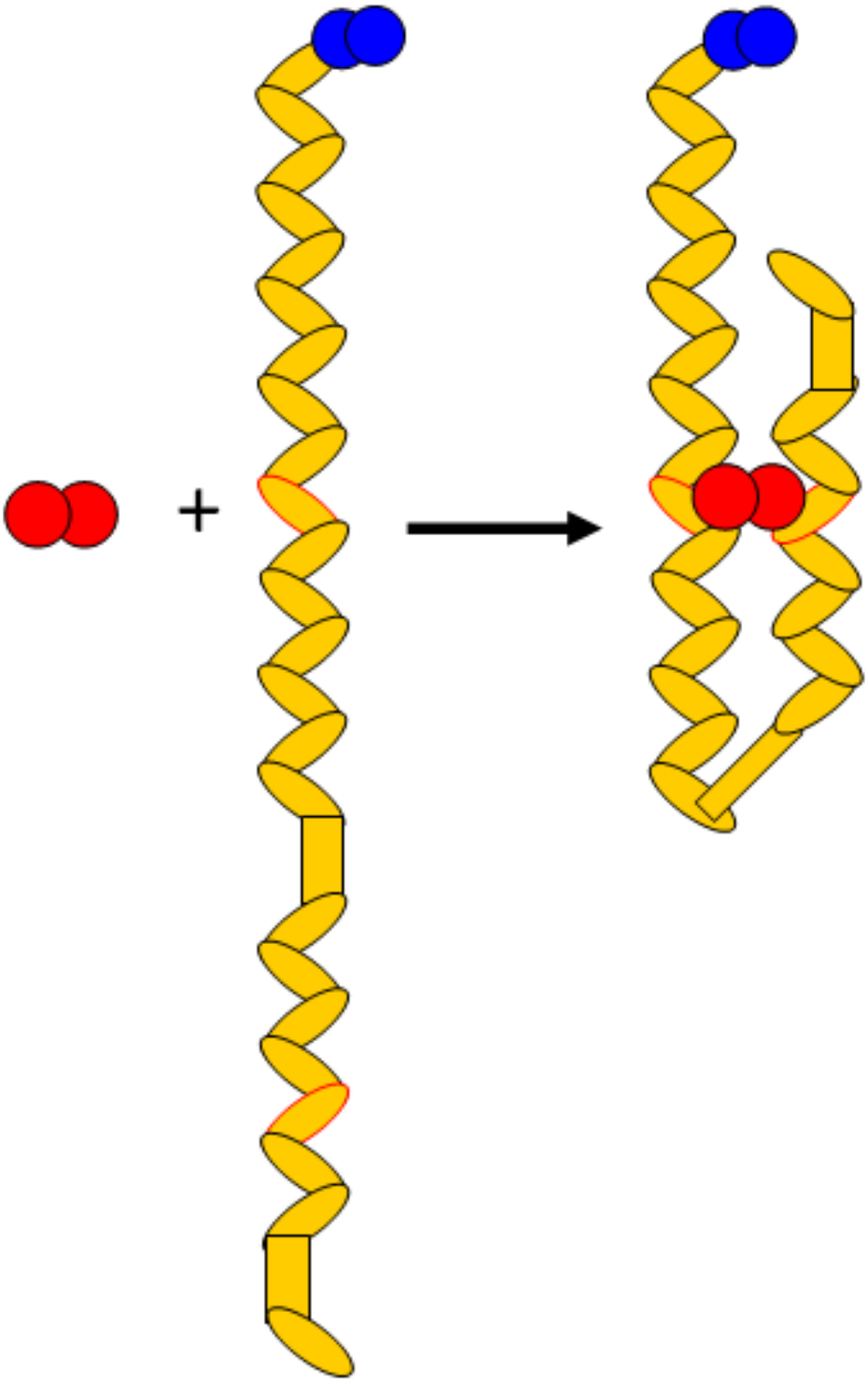


Figure 4: Diagram of proposed action of CFM2 on Filamin A

The CFM2 homo-dimer (shown in red) is proposed to interact with repeats 10 and 21 (shown as the red outlined regions). The proposed interaction is show on the right, the flexibility of the hinge region may allow the Filamin A protein to fold over itself into a form that loses activity.

2 Materials and Methods

2.1 Materials

2.1.1 DNA manipulation

Pwo polymerase buffer +$MgSO_4$, dNTPs, Pwo polymerase, full length FAM101A cDNA in pGEMT vector provided by S.P Robertson and forward and reverse primers (Sigma-Aldrich) were used in PCR for the purpose of subcloning. rTaq polymerase was used instead of Pwo polymerase in PCR for diagnostic testing.

NcoI, XhoI, BamHI, SalI, Buffer H (Roche) and EcoRI were used in restriction digestion experiments for diagnostic purposes and subcloning. Ligation buffer andT4 ligase were used with EcoRI during subcloning ligation procedures. pProEX HTB vector system (Invitrogen) and pGEX 4T3 vector system (GE Healthcare) were destination vectors for subcloning. Roche PCR purification kits were used in subcloning to purify DNA that had been treated with restriction enzymes.

Agarose, in Tris-Acetate-EDTA (TAE) buffer and sucrose loading buffer (Glycerol, bromophenol blue) were used to perform agarose gel electrophoresis of DNA compared to 1kb plus marker (Invitrogen). Ethidium bromide was used to visualize DNA in gels in a UV light box which were photographed in a “Gel Doc” light box using “Biologic” software (Quantity one).

2.1.2 Cell culturing

All cells were cultured in Luria Broth media (Invitrogen) at 25 g/L in milli-q water. Broth was sterilised by autoclaving at 121°C for 15 minutes. Luria Broth agar was made by adding 1% agarose powder prior to autoclaving. Ampicillin (Invitrogen) was added to a final concentration of 100 µg/mL to facilitate plasmid selection.

2.1.3 Protein manipulation

Native Buffer (5x) (250 mM NaH_2PO_4, 2.5M NaCl pH 8.0), RIPA buffer (10 mM NaH_2PO_4, 0.3M NaCl, 0.1% SDS, 1% NP40, 1% deoxycholate, 2 mM EDTA, pH 7.2) and PBS (137 mM NaCl, 2.7 mM KCl, 10mM NaH_2PO_4, 2.0 mM KH_2PO_4 pH 7.4) native buffers were used to maintain native protein folding of cell lysate and purified proteins. All buffers were sterilised by autoclaving at 121°C for 15 minutes.

SDS-PAGE sample buffer (35% glycerol, 0.125 M Tris-HCL pH 6.8, 2.5% SDS, 0.025% bromophenol blue), Coomassie stain (Coomassie brilliant blue R-250 0.0625% w/v, 10% acetic acid, 40% EtOH)Tris running buffer (25 mM Tris, 0.2 M glycine, 0.1% SDS) and Coomassie de-stain (20% EtOH, 10% acetic acid) were used in SDS-PAGE sample preparation and analysis.

His-Trap column (GE Healthcare) and imidazole (Sigma-Aldrich) were used in GST-fusion protein purification protocols. Superdex S200 10/300 HR column (GE Healthcare) and Superdex S70 10/300 HR column (GE Healthcare) were used in size exclusion chromatography.

2.1.4 General chemicals and equipment

A Nano Drop spectrophotometer (Thermo Scientific) was used to determine protein concentration. The sonicator S-4000 (QSONICA) and French press were used during cell lysis. SORVALL Evolution RC centrifuge using GS3 fixed angle rotor (at 6000g) and SS34 fixed angle rotor (at 30000g) and Minispin Plus centrifuge (Eppendorf) were used with PCR purification kits and to pellet cells or cell debris.

Sequencing performed by the Alan Wilson Centre Genome Service, Massey University, Palmerston North using AB13730 capillary equipment.

Table 2: Primers used in this project

Name	Annealing temperature (Tm^0C)	Template	Sequence 5`-3`	Orientation (sense (S) / anti-sense (A))
HCFM2LF1	67	HCFM2-non-truncated	TAA ACC ATG GCA GTG GGC CAC CTG CAT CTG	S
HCFM2LF2	66.9	HCFM2-non-truncated	TAA ACC ATG GCA GTG GGC CAC CTG CAT CT	S
HCFM2LF3	71.8	HCFM2-non-truncated	TAA ACC ATG GCA GTG GGC CAC CTG CAT CTG CAG GGC ATG	S
HCFM2SF1	70.6	hCFM2 / HCFM2-non-truncated	TAA ACC ATG GCA AGG CCC CGG ATG CTG CCA GT	S
HCFM2B1	68	hCFM2 / HCFM2-non-truncated	AAG CTT CTC CAG TTA TCG TGT CCA TCC TCC CGG	A
FLNAR10F2	65.7	FLNA	AAG GAT CCG ACG CAT CCA AAT GCT CAG GC	S
FLNAR10F3	85.6	FLNA	AAG GAT CCG ACG CAT CCA AAG TCA AGT GCT CAG GCC C	S
FLNAR10B2	69.8	FLNA	CAA TCC GTC GAC TTA CGC AGG TTC CAC CTG CAG CTT GCT	A
pGEX3`	64.0	pGEX	CCG GGA GCT GCA TGT GTC AGA GG	A
pGEX5`	65.7	pGEX	GGG CTG GCA AGC CAC GTT TGG TG	S

Table 3: Plasmids used in this project

Name	Plasmid Map	Expressed protein
pProEX HTb	Figure 6	-
pGEX 4T3	Figure 24	-
pREP4-FLNA	Figure 23	-
pProEX HTb -CFM2SF1-non-truncated	Figure 6	(6)His(N terminal)- Q6ZTI6-2
pGemT-FAM101A-truncated	Figure 5	-
pGemT-FAM101A	Figure 5	-
pGEX 4T3-FLNAR10	Figure 24	GST-FLNAR10 $_{(1159\text{-}1252)}$

Table 4: *E.coli* strains used in this project

Strain	Reference	Genotype
Top10	[29]	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (Str^R) endA1 nupG*
BL21 (DE3)	[30]	*F- ompT hsdSB (rB-mB-) gal dcm*

2.2 Methods

2.2.1 Competent cells

Competent cells are more easily able to incorporate foreign DNA via transformation. Cells that were not already competent are prepared by re-suspension of *E.coli* Top10 cells growing at log phase in calcium chloride on ice for 30 minutes.

2.2.1.1 Handling

Competent cells are stored in 100 µL aliquots at -70°C. All competent cells were handled in a PC2 laboratory according to Novagen.pET System Manual 11th Ed page 11.

2.2.1.2 Disposal

All unused competent cells from a 100 µL aliquot were disposed of after sterilisation using liquid autoclave followed by incineration.

2.2.1.3 Competent Cell Transformation

All competent cell transformations were done using heat shock at 42°C. Transformations were handled in a PC2 laboratory according to Novagen.pET System Manual 11th Ed page 12.

Selection of transformants was done by plating transformed cells onto LB agar plates containing AMP at 0.1 mg/mL.

2.2.2 Polymerase Chain Reaction

2.2.2.1 PCR amplification

Pwo polymerase was used for the amplification of coding regions where the purpose would be for preparation i.e. subcloning. This is because Pwo polymerase is ~10-18 fold more accurate than Taq polymerase owing to its 3`-5` exonuclease proofreading activity.

Each reaction was performed in 0.2 mL PCR tube with a final volume of 50 µL. 5 µL Pwo buffer x10 (+$MgSO_4$) (Roche), 1.25 µL dNTPs (10 mM), 5 µL of Forward primer (10 pmol/µL), 5 µL of Back primer (10 pmol/µL), 0.5 µL Pwo Polymerase (5U/µL), 2 µL (2 ng/µL) template DNA, filtered sterile MQ H_2O (to 50µL). PCR tubes were mixed using

flicking and centrifuged briefly. PCR tubes were heat cycled using the steps outlined in Table 4.

Table 5: PCR run cycle*

Step	Temperature	Length of step	Description
1	94°C	2 min	Initial denaturation
2	94°C	1 min	Denature
3	Primer Dependent	1 min	Annealing
4	72°C	1 min 30 seconds	Extension (Return to step 2 x29)
5	72°C	5 min	Final Extension
6	10°C	Pause	End Step

2.2.3 DNA analysis using ethidium bromide staining of agarose gel

Agarose powder was mixed with 35 mL of 1x TAE buffer (40 mM Tris acetate, 1 mM EDTA) to an agarose percentage of 1.0% or 1.25% as required. The agarose TAE solution was heated using the microwave until all of the agarose was melted. After the solution cooled to ~60°C the gel was poured into the casting tray and the comb was added to form the sample wells.

Gels were submerged in 1x TAE buffer (40 mM Tris acetate, 1 mM EDTA). Samples were loaded in sample wells and run at 100volts (capped) and 400mA for 40 minutes.

During and after ethidium bromide (EtBr) staining of agarose gels gloves, lab coats and glasses were worn at all times, movement of the stained gel occurred in a sealed container labelled for EtBr gel movement. Agarose gels were placed in EtBr bath for 25 minutes. Gels were destained in water to lower background EtBr for 25 minutes.

EtBr stained gels were visualised in a UV light box and photographed in a “Gel Doc” light box using “Biologic” software (Quantity one).

EtBr stained gels and all gloves used to touch equipment related to EtBr-agarose gels were discarded in EtBr specific bins labelled for safe disposal.

2.2.4 DNA Sub-cloning

DNA sub-cloning is a set of techniques used to move a gene of interest from a parent vector e.g. plasmid, to a destination vector. In this project sub-cloning was used to move complete genes or parts of genes from a plasmid used for high copy number maintenance (such as pGEMT) to a plasmid with inducible expressivity and an N-terminal affinity tag for the purpose of purification of the protein of interest.

Using restriction endonuclease enzymes dsDNA can be cut specifically. The restriction endonuclease chosen performs a staggered cut which allows for orientation specific placement of the gene of interest in the destination vector.

As it would be highly unlikely to find a specific restriction site only once and in the correct place on both the destination vector and the gene of interest, PCR was used to introduce an overhang, which included the restriction site, via the primers used.

Using primers that extend past the edge of the DNA to be amplified can present issues with the 3`-5` exonuclease activity digesting the primers overhang. To reduce this Pwo polymerase was introduced just before the PCR tube was placed in the thermocycler.

Once PCR was performed to generate a dsDNA with the two distinct restriction sites at either end, the PCR product was digested with both restriction endonuclease enzymes.

Digestion was done in a PCR tube with 5 µL Buffer H (Roche)x10 or React3 (Invitrogen)x10, 1 µL of each restriction endonuclease (10U/µl), 8 µL of PCR DNA or 3 µL vector DNA (2 ng/µL), H_2O to final volume of 50 µL. Digestion took place at 37°C for 3 hours.

2.2.5 Ligation

Using an ATP dependent ligase two strands of DNA can be ligated to complete the sub-cloning of a gene into a destination vector. Ligation is the mechanism which catalyses the formation of phosphodiester bonds between close 3`hydroxyl and 5`phosphate ends in dsDNA, fusing two restriction-digest products together in this case. Ligation works with blunt ended DNA but it is a much slower speed reaction compared with sticky ended DNA ligation.

Ligation was performed in a PCR tube. Varying molar ratios of double digested vector to double digested PCR product were added e.g. 1:10, 1:5, 1:3 1:1, totalling 10µL, approximating 20ng of DNA. 6µL of ligation buffer (5x)(Roche) was used, 1 µL T4 ligase (Roche) (10 U/µL) and H_2O to a final volume of 30 µL. Ligation was performed at 37°C for 20 minutes or overnight at 21°C.

2.2.6 Plasmid Isolation

Plasmids were used to transform Top10 *E.coli* competent cells. Cells were spread on LB agar plates containing ampicillin (AMP) at 0.1 µg/mL. Plates were incubated at 37°C for ~12 hours. Individual colonies picked were used to inoculate 10 mL LB containing AMP at 0.1 µg/mL. 10 mL inoculated broth were incubated at 37°C for ~12 hours.

10 mL cultures were spun down repeatedly (at $13X10^3$ rpm for 4 minutes per cycle) to collect the cell pellet in a 1.5 mL micro centrifuge tube.

"High Pure Plasmid Isolation Kit" from Roche was used on the cell pellets. This kit relies on lysing the cell wall with alkaline SDS without disturbing the chromosomal DNA, which can then be removed with the cell wall debris and denatured proteins as pellet. The supernatant is then passed over glass fibres in the presence of guanidine hydrochloride causing the plasmid DNA and not chromosomal DNA to be selectively retained. The plasmid DNA is then washed and eluted in a low salt buffer or water.

2.2.7 DNA sequencing

DNA sequencing was done by the Massey Genome Service in the Alan Wilson centre. The reaction type was BigDye Terminator v3.1.

2.2.8 Protein Expression

BL21 *E.coli* were transformed with expression vectors. Plasmid levels were maintained using AMP based selection. Growth to log phase and protein expression was generally done at 37°C, though tested at lower temperatures (30°C, 22°C and 25°C)to try and increase protein solubility.

All expression vectors utilised a lac operon based system. Transcription and subsequent expression was initiated with isopropylthio-β-galactoside (IPTG) to a final concentration of 0.1 mM.

Different expression time lengths were tried to allow recovery of the largest amount of soluble protein of interest. Cultures were pelleted by centrifugation; the pellets were stored at -20°C for protein purification.

All cultures depleted of cells from centrifugation were autoclaved before disposal.

2.2.9 Cell Lysis

Lysis of *E.coli* for retrieval of expressed proteins was done in 37 mL of varying buffers (i.e. PBS pH 7.3 or RIPA) per pellet (from 500 mL of culture). Cell walls were disrupted using either sonication or high-shear mechanical disruption using a French press.

Cell disruption using French press was performed at 1000MPa; lysate was passed through twice resting on ice between passes. Lysate was sonicated for 30 seconds at 20-40% (of 400W) between 2 and 4 times with cooling on ice between passes. Sonicated lysate was spun down to remove cell debris at 30000 x g at 4°C for 20 minutes.

2.2.10 Protein size and mass analysis using SDS-PAGE

SDS-PAGE allows the separation of proteins in gels based on size alone. SDS binds to hydrophobic regions causing the denaturation of the protein. SDS has a negatively charged sulphate group which binds to any positive side-chains giving all proteins in the solution an overall negative charge, this means all proteins will migrate towards a positive pole in an electric field.

Polyacrylamide is a polymer of acrylamide monomers; the gel contains a mesh of pores and tunnels which hinder proteins migrating through it. The larger proteins take a longer time to move through the poly-acrylamide gel. The difference in travel times for proteins based on size creates bands in the gel of proteins of the same size. By comparing the protein bands against proteins of known size in another lane, the size can be estimated. In this project 12.5% and 15% polyacrylamide gel were used. Protein samples and gels were prepared and run according to [31].

2.2.10.1 Protein visualisation of SDS-PAGE

Silver staining and Coomassie staining were used in this project to allow visualisation of the protein in SDS-PAGE gels.

Silver staining is 10-100x more sensitive than Coomassie for the detection of proteins; silver staining was the preferred method for protein detection where protein concentrations were low. Silver is deposited onto the proteins in the gel after electrophoresis. The initial deposition of silver (known as the latent image) catalyses the reduction and plating out of further silver ions when the developer (a moderate reducing agent) is added. The silver accumulates at sites of protein bands. The protocol used in this project was performed according to [32].

Coomassie blue staining while less sensitive than silver staining is much quicker and the solutions remain stable for much longer. Coomassie blue binds to proteins at positively charged amine groups and through Van der Waals attraction. Gels were stained in 0.25% Coomassie Blue for 2-24hours and destained in 10% acetic acid, 80% EtOH, 10% H_2O for 2-24 hours.

All gels were photographed using “Gel Doc” light box using Biologic software (Quantity one).

2.2.11 GST-fusion protein pull-down

The glutathione-s-transferase (GST) gene fusion system was used in this project to isolate an expressed recombinant fusion protein of GST and FLNAR10. GST is an enzyme that binds glutathione; in this experiment glutathione is covalently linked to sepharose.

E.coli was transformed using an inducible expression plasmid. The plasmid, pGEX 4T3, has the GST protein sequence followed by a short linker sequence which contains a thrombin cleavage site. Following the cleavage site the protein of interest sequence is located. When expressed the fusion protein can be specifically concentrated from the cell's other proteins.

Sepharose beads with covalently bound glutathione were used to immobilise the fusion protein allowing the undesired proteins to be washed away.

E.coli expressing the fusion protein was lysed under native conditions to retain tertiary structure required for GST binding to glutathione. The filtered supernatant of the lysate was passed slowly through a poured column (GE Healthcare) containing 2 mL of glutathione covalently bound to sepharose beads (equibrilated to the native buffer of the lysate). After binding of the lysate the column was washed with 20-40x the column bead volume of native buffer (PBS pH 7.3). The cleaned bound proteins were then eluted with native buffer (PBS pH 7.3 or RIPA) containing 10-20 mM reduced glutathione.

2.2.12 Histidine-tag protein pull-down

The 6xHis-tag system can be used as an affinity pull-down assay or affinity purification system. A nickel-NTA-agarose column has a micro-molar affinity for 5x or greater histidine in a row. Within the sequence methionine followed with six histidine was added to the N` terminal amino acids of the protein of interest by subcloning in frame into the pProEX HTb plasmid multiple cloning site.

10 mL pre-packed columns were equibrilated with native buffer containing 20 mM imidazole. Cells containing the Fusion protein with 6xHis-tag were lysed according to 2.2.9 Cell Lysis. Cell lysate was put through a 200µm filter to remove particulates. Filtered cell lysate was loaded onto the column using a peristaltic pump; the flow-through was passed over the column three times to ensure most of the histidine-tagged protein had bound to the column. Imidazole was used to competitively elute the fusion protein from the column.

The concentration of imidazole was increased either as a step gradient or as a linear gradient. The flow-through from the column was passed through a flow cell spectrophotometer to measure protein concentration (A^{280}). The flow-through was collected as 4 mL samples for SDS-PAGE analysis.

2.2.13 GST fusion protein probe pull-down

Using GST fusion proteins the bait protein (FLNAR10) linked to GST can be used to probe cell lysate i.e. C2C12 mouse fibroblast cells, for proteins that interact with the probe protein. The proteins that interact with the probe can be purified from the cell lysate along with any proteins that interact with the probe bound proteins by passing it over an affinity column specific to GST (usually agarose linked to glutathione).

The experiment used in this thesis was performed according to [33]. Cell lines to be probed (C2C12 mouse myoblast cells and *E.coli* BL21 cell expressing CFM2) were pelleted then re-suspended in 750 µL of RIPA or PBS. Cell walls were disrupted using sonication at 10% for 12 seconds twice. The lysates were spundown at 12000 x g for 20 minutes at 4°C.

Each supernatant was incubated with 0.5µg of GST and 50 µL of 50% glutathione-sepharose beads in 1.5 mL tubes with end over end mixing for 2 hours at 4°C to deplete GST binding proteins from the lysate.

The depleted supernatant was removed from the bead lysate mixture using centrifugation at 12000 g for 2 minutes at 4°C. The depleted supernatant was split evenly into two tubes. 7µg of GST-FLNAR10 fusion was added with 50 µL of 50% glutathione-sepharose bead slurry to the first tube. 5µg of GST was added with 50 µL of 50% glutathione-sepharose bead slurry was added to the second tube as a control for proteins that bind GST regardless of the FLNAR10 protein being missing. Both tubes were incubated end over end at 4°C for 2 hours.

The tubes were centrifuged at 12000 x g for 2 minutes at 4°C and the supernatant was removed, this should contain all of the proteins from the depleted lysate samples except those that specifically interact with FLNAR10 or GST. The beads were washed 3 times with 1 mL of the lysate buffer to remove any unbound proteins.

3.1 Results and Discussion

3.1.1 Sub-cloning CFM2

Here, the rationale behind the selection of vectors for the sub-cloning of CFM2 and the resulting difficulties with cloning PCR products and transformation of the finished ligated product is explored.

The role, structure and interactions of CFM2 protein are unknown. Expression of CFM2 in an expression system can allow for easier recovery of pure protein as the addition of purification tags can be added N-terminal or C-terminal of the CFM2 region. CFM2 is a eukaryotic protein; to express it requires the use of either an expression system capable of correctly splicing the introns from the pre-mRNA or to use a cDNA copy from cellular mRNA which had already been correctly spliced. CFM2 cDNA containing the full length gene without the introns, cloned into the pGEMT vector was available.

The pGEMT vector (Figure 5) was unsuitable for the expression of CFM2. The pGEMT vector is used for cloning PCR products in bacteria. As the pGEMT vector does not contain the necessary translation initiation sites or purification tag the CFM2 gene had to be subcloned into another vector.

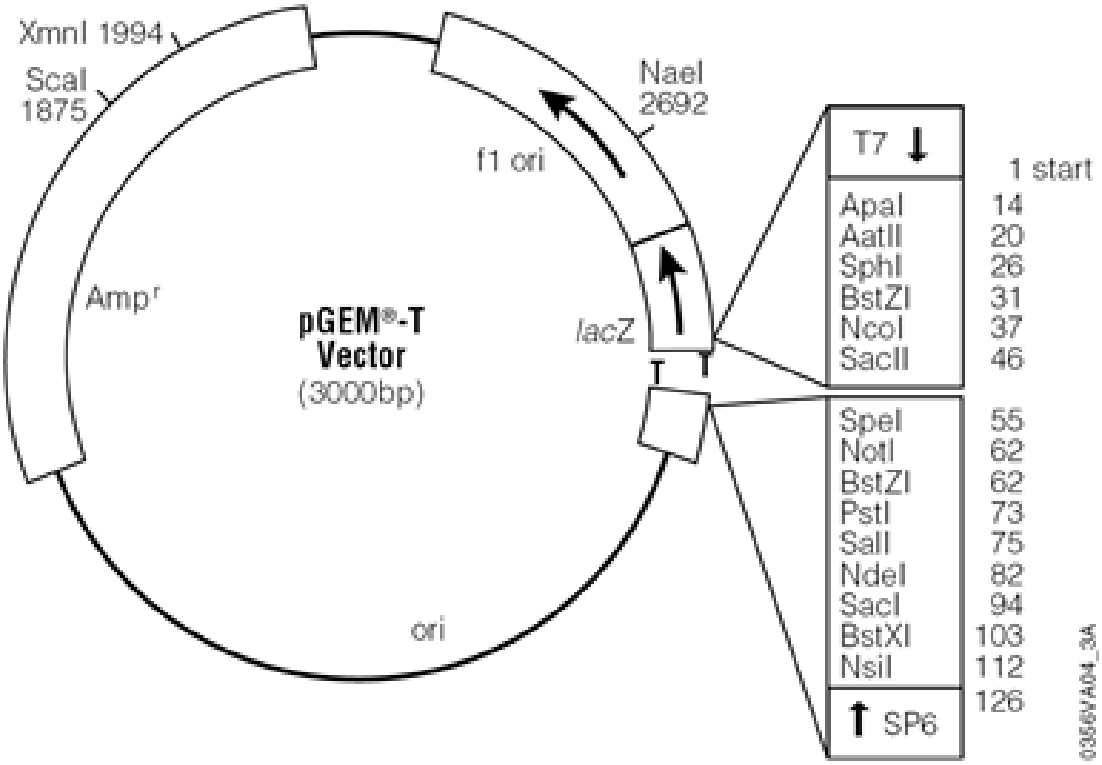


Figure 5: pGEM T Vector Map[34]

CFM2's proposed mechanism of function requires that it dimerises. The dimerisation of CFM2 is supported by the original yeast-2 hybrid assay that identified CFM2 as a potential binding partner to FLNA repeat 10. The destination vector for CFM2 had to meet certain requirements; other than the common requirements such as selection for plasmid maintenance, dimerisation should not occur in the fusion protein for reasons other than the protein of interests normal binding capabilities. This meant a GST-based system was passed over in favour of the polyhistidine-tag system as GST can dimerise. Had the GST based system been chosen, then later purification techniques like size exclusion chromatography would not be able to replicate the dimerisation of CFM2 as easily or convincingly.

The destination vector chosen was pProEX HTB, a 6xHis-tag system with inducible expression via the lac-operon and ampicillin resistance based selection / plasmid maintenance (Figure 6).

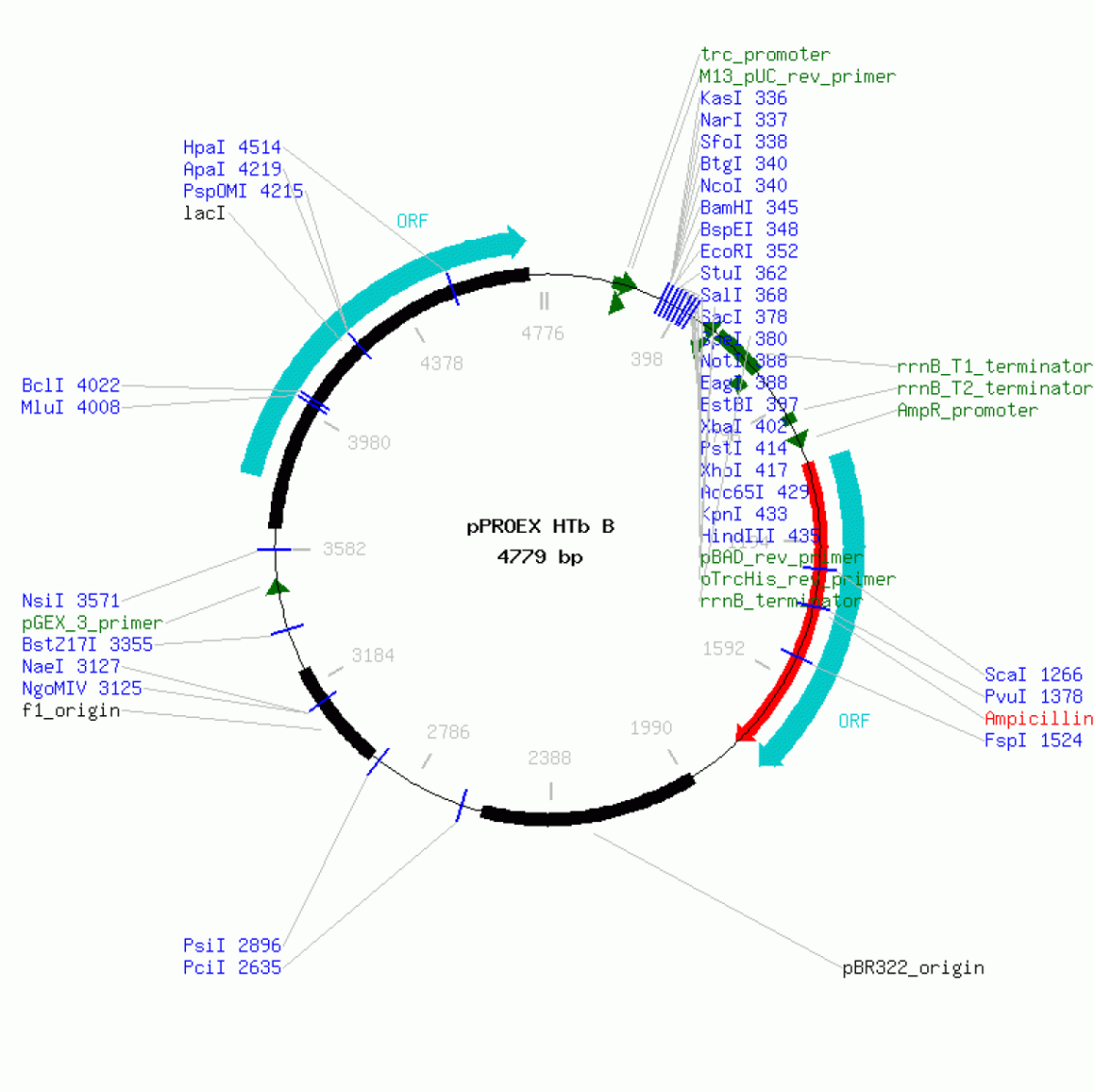


Figure 6: pProEX HTB Vector Map[35]

Two pGEM T vectors were supplied by Prof S.P Robertson (University of Otago), one containing CFM2, the other contained a version of CFM2 with an additional stop codon that would result in a truncated protein; both were human CFM2 cDNA. CFM2 has two isoforms (Uniprot numbers Q6ZTI6 and Q6ZTI6-2) in this project the HCFM2LF2 and HCFM2SF1 primers were used to isolate those isoforms (the Q6ZTI6-2 isoform is missing amino acids 1-81), which one is biologically most relevant was unknown. Primers were designed to amplify the Q6ZTI6 and Q6ZTI6-2 isoforms of CFM2 with overhangs added to the primers to create the appropriate cleavage sites (XhoI and NcoI) for restriction enzymes. By using these cleavage sites in the PCR product (CFM2)

and in the destination vector, correct frame and orientation of the cDNA was controlled.

PCR was performed according to 2.2.2.1 PCR amplification. The templates used were the pGEM-T: FAM101A-truncated and pGEM-T: FAM101A plasmid. Both were placed through PCR in two reactions using the back primer HCFM2B1 in combination with the forward primer HCFM2LF2 or HCFM2SF1. The temperature during the annealing step was set at 54°C.

The PCR products were examined using agarose gel electrophoresis according to 2.2.3 DNA analysis using ethidium bromide staining of agarose gel.

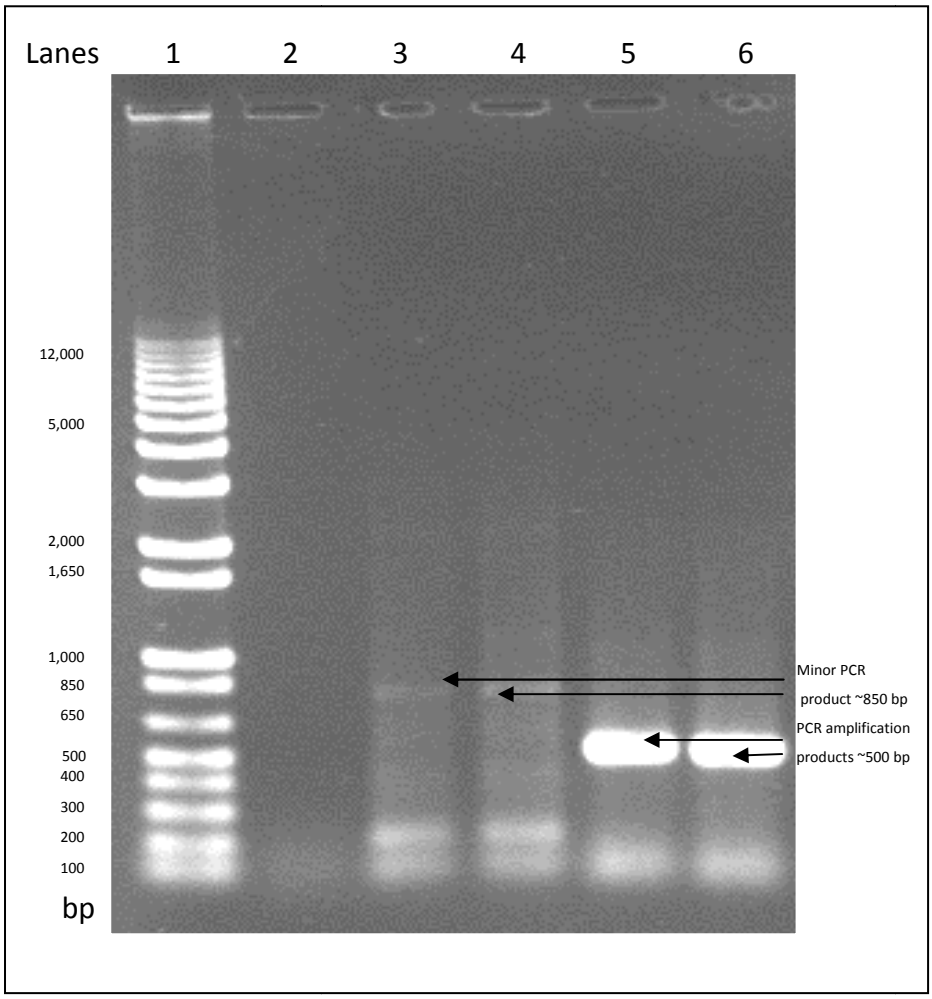


Figure 7: Amplification products of pGEM-T: FAM101A-truncated and pGEM-T: FAM101A with HCFM2LF2 and HCFM2SF1 forward primers

2µL of pGEM-T: FAM101A-truncated and pGEM-T: FAM101A (2 ng/µL) were amplified using PCR. Each template was amplified in two reactions. In each reaction 5 µL of HCFMFB1 (10 pmol/µL) was used as the back primer. In one reaction 5 µL of HCFMLF2 (10 pmol/µL) was used as the forward primer, in the other reaction 5 µL of HCFMSF1 (10 pmol/µL) was used as the forward primer. PCR products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. Expected product sizes were ~730bp for the long isoform Q6ZTI6 (truncated or not) and ~480 for the short isoform Q6ZTI6-2 (truncated or not). DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes 1 5 µL 1kb plus (Invitrogen) DNA ladder
2 3 µL Negative control (HCFM2LF2 and HCFM2B1 primers, no template)
3 3 µL pGEM-T: FAM101A, HCFM2LF2 forward primer
4 3 µL pGEM-T: FAM101A-truncated, HCFM2LF2 forward primer
5 3 µL pGEM-T: FAM101A, HCFM2SF1 forward primer
6 3 µL pGEM-T: FAM101A-truncated, HCFM2SF1 forward primer

PCR produced an amplification product with the combination of HCFM2SF1 forward primer and HCFM2B1 reverse primer with both pGEM-T: FAM101A and pGEM-T: FAM101A-truncated templates (see Figure 7 lanes 5 & 6). Amplification of pGEM-T:FAM101A-truncated and pGEM-T:FAM101A templates with HCFM2LF2 and B1 primers produced very weak concentrations of products and two bands around 200 and 100 bp of unknown origin (see Figure 7 lanes 3&4).

In order to reduce the presence of the two unknown bands and increase the proportion of the larger product, PCR was repeated using a touchdown method. Touchdown PCR started with an annealing temperature at 64°C which lowered each cycle for 10 cycles, PCR then continued for a further 20 cycles at 54°C. Primers will anneal most specifically at higher temperatures, by gradually decreasing the annealing temperature step it creates conditions where the most complementary sequences will bind first. As the primers were created to the sequence of the FAM101A gene the pairings with the highest complementarity should be between the primers and the sequence that is to be amplified. After the first few cycles have passed the proportion of the sequence of interest to other interfering bindings will be largely in favour of the sequence of interest, as PCR is an exponential amplifying process the sequence of interest will out replicate the non-specific interactions that were found in lane 3 & 4 of Figure 7.

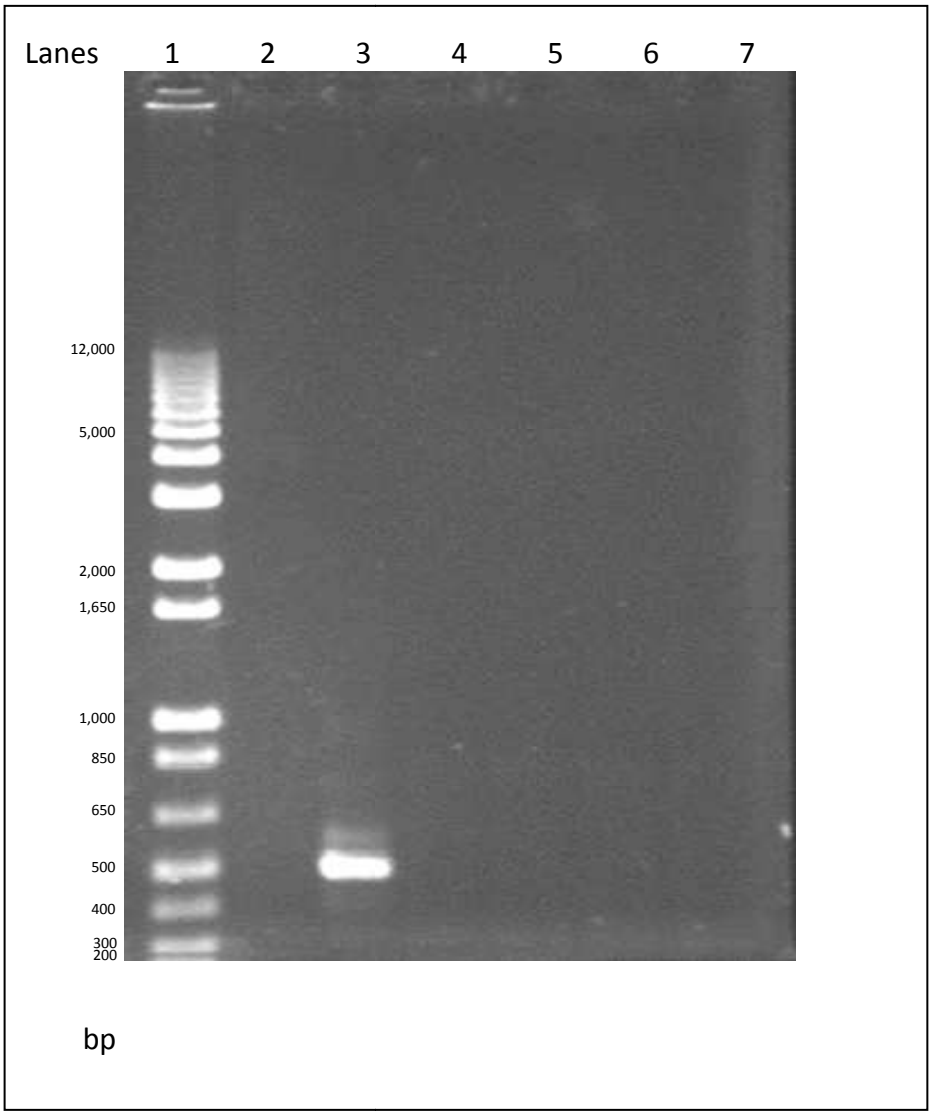


Figure 8: Touchdown PCR amplification products of pGEM-T: FAM101A and pGEM-T: FAM101A-truncated with HCFM2LF2 and HCFM2SF1 forward primers

pGEM-T: FAM101A and pGEM-T: FAM101A-truncated (2ng and 4ng per PCR reaction) were amplified using PCR. In each reaction 5µL of HCFMFB1 (10 pmol/µL) was used as the back primer, 5µL of HCFMLF2 (10 pmol/µL) was used as the forward primer. In the positive control 5µL of HCFMSF1 (10 pmol/µL) was used as the forward primer and 2ng of pGEM-T: FAM101A was used as the template. PCR products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**3µL Negative control (HCFM2SF1 and HCFM2B1 primers, no template)**
	3	**3µL Positive control (HCFM2SF1 and HCFM2B1 primers, pGEM-T:FAM101A)**
	4	**3µL pGEM-T:FAM101A 4ng, HCFM2LF2 forward primer**
	5	**3µL pGEM-T:FAM101A-truncated 4ng, HCFM2LF2 forward primer**
	6	**3µL pGEM-T:FAM101A 2ng, HCFM2LF2 forward primer**
	7	**3µL pGEM-T:FAM101A-truncated 2ng, HCFM2LF2 forward primer**

Touchdown PCR failed to recreate the band seen in Figure 7 lanes 5 & 6. To try and obtain the long isoform of the FAM101A gene amplified during PCR another primer was created HCFM2LF3. With the nine extra nucleotides over the HCFM2LF1 primer it

was hoped the higher melting temperature (71.8°C vs 66.9°C) would produce better annealing and consequently increase the amplification of the long isoform of the FAM101A gene during PCR.

Another Touchdown PCR cycle was run using the same settings as those used in Figure 7 with the addition of a reaction with the HCMF2LF3 primer as the forward primer. Polymerase was added just before cycling began in order to prevent degradation of the primers and to reduce non-specific amplification prior to PCR. The DNA polymerase enzyme (in this case Pwo) was added to each tube immediately prior to the cycling of the machine starting. This is particularly important when using an enzyme with 3`-5` exonuclease activity and primers that contain overhanging nucleotides as was the case in these reactions as the exonuclease proofreading activity can destroy the primers over time.

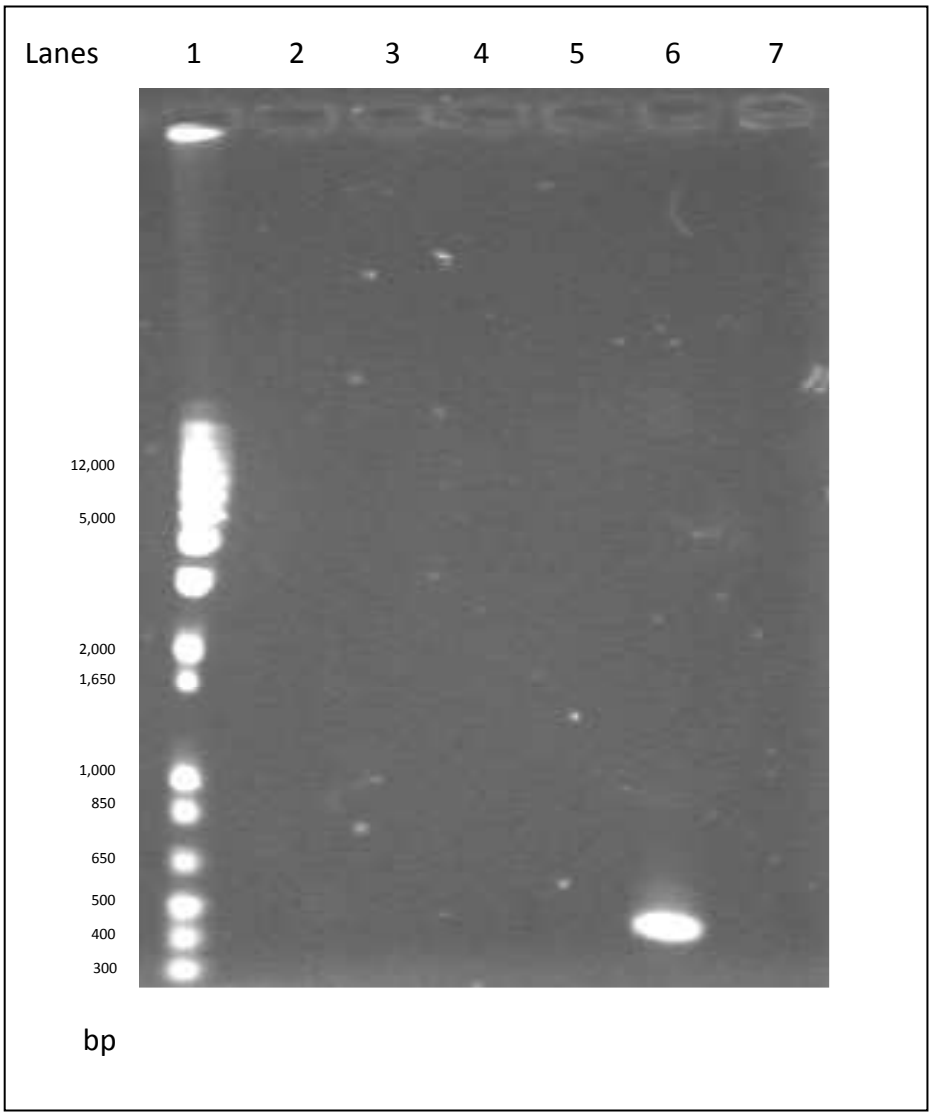


Figure 9: Touchdown PCR amplification products of pGEM-T: FAM101A and pGEM-T: FAM101A-truncated with HCFM2LF2, HCFM2LF3 and HCFM2SF1 forward primers

pGEM-T: FAM101A and pGEM-T: FAM101A-truncated (2ng PCR reaction) were amplified using PCR. In each reaction 5µL of HCFMFB1 (10 pmol/µL) was used as the back primer, 5µL of HCFMLF2 (10 pmol/µL) was used as the forward primer or 5µL of HCFMLF3 (10 pmol/µL). In the positive control 5µL of HCFMSF1 (10 pmol/µL) was used as the forward primer and 2ng of pGEM-T: FAM101A was used as the template. In the negative control 5µL of HCFMSF1 (10 pmol/µL) was used as the forward primer and no template was added. PCR products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes 1 5µL 1kb plus (Invitrogen) DNA ladder
2 3µL pGEM-T:FAM101A 2ng, HCFM2LF3 forward primer
3 3µL pGEM-T:FAM101A 2ng, HCFM2LF2 forward primer
4 3µL pGEM-T:FAM101A-truncated 2ng, HCFM2LF3 forward primer
5 3µL pGEM-T:FAM101A-truncated 2ng, HCFM2LF2 forward primer
6 3µL Positive control (HCFM2SF1 and HCFM2B1 primers, FAM101A template)
7 3µL Negative control (HCFM2SF1 and HCFM2B1 primers, no template)

Even under “hot start” touchdown conditions the HCFM2LF1, HCFM2LF2 and HCFM2LF3 forward primers failed to produce PCR products. As HCFM2SF1 forward primer succeeded in producing a PCR product the next aim was to ligate this product into the expression vector pProEX HTb.

3.1.1.1 Sub-cloning CFM2 PCR product into pProEX HTb

Using the product from PCR of pGEM-T: FAM101A and the SF1 and B1 primers (non-truncated CFM2 Short isoform), the vector plasmid and PCR product were digested using NcoI and XhoI.

The digest protocol was setup in reactions as follows in Table 6.

Table 6: Double digest of pProEX HTb and PCR product

	Negative Control	**Vector +NcoI**	**Vector +XhoI**	**Vector +NcoI +XhoI**	**PCR product +NcoI +XhoI**
Buffer H (10x)(µL)	5	5	5	5	5
DNA(µL)	3 pProEX	3 pProEX	3 pProEX	8 pProEX	8 PCR product
NcoI(µL)	-	1	-	1	1
XhoI(µL)	-	-	1	1	1
H_2O(µL)	42	41	41	35	35

Tubes were incubated for three hours at 37°C.

The digestion products were analysed using agarose gel electrophoresis.

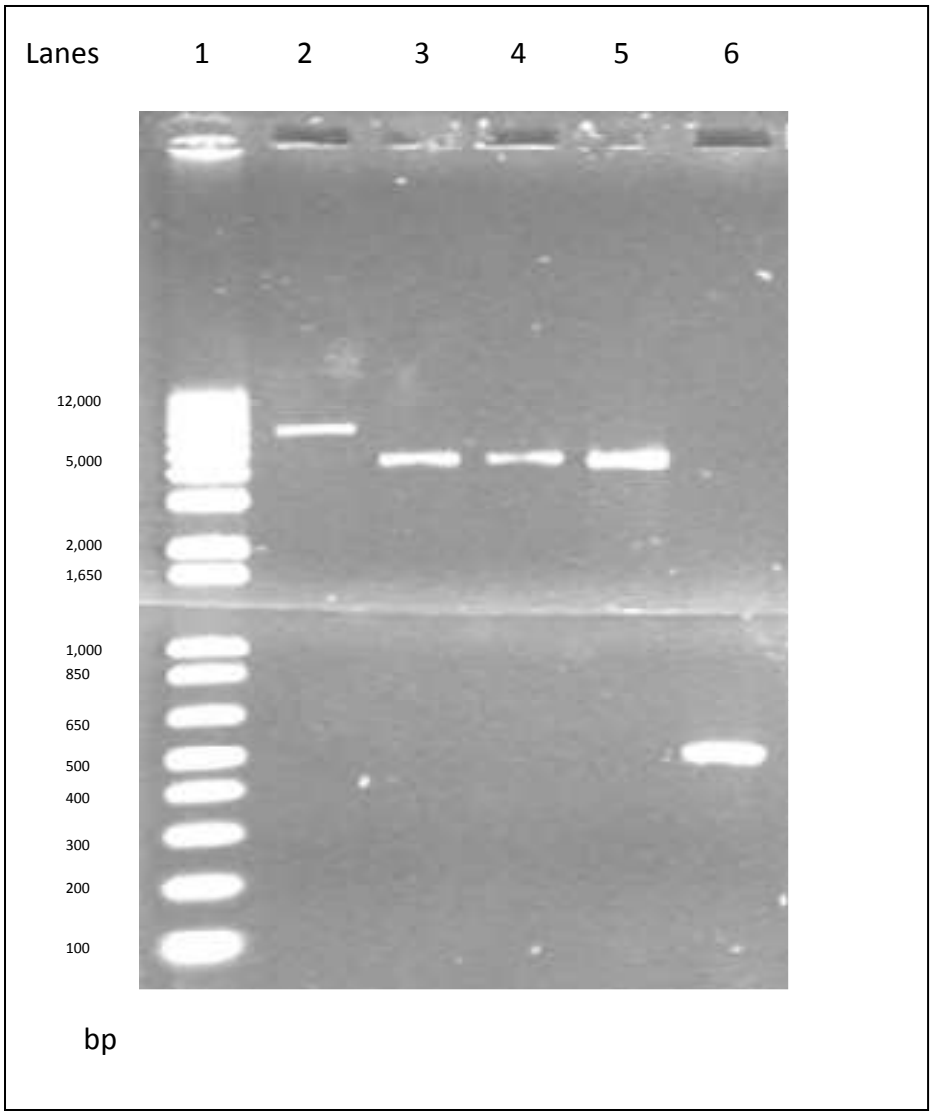


Figure 10: Double digest of pProEX HTb vector and PCR product (HCFM2-non-truncated SF1)

pProEX HTb vector (1µL) was digested with 1µL of NcoI and 1µL of XhoI in separate reactions. pProEX HTb vector (8µL) and PCR product (8µL) (HCFM2-non-truncated SF1) were each double digested with 1µL of XhoI and 1µL of NcoI. Digestion took place over 3 hours at 37°C. pProEX HTb (1µL) was treated to identical conditions as the previous reactions without the restriction enzymes as a negative control. Digestion products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL pProEX HTb, negative control**
	3	**5µL pProEX HTb, +NcoI**
	4	**5µL pProEX HTb, +XhoI**
	5	**5µL pProEX HTb, +NcoI +XhoI**
	6	**5µL HCFM2-non-truncated SF1 PCR product, +NcoI +XhoI**

As seen in Figure 10 the vector plasmid pProEX-HTb increases in mobility from lane 2 (negative control, uncut plasmid) to lane 3, 4 and 5 (cut plasmid and double cut plasmid). According to Figure 6 (vector map of pProEX-HTb) only one site is expected for the restriction enzymes NcoI and XhoI therefore the increase of mobility is expected to be not due to a reduced of size of the plasmid but the linearization of the plasmid. The increase in mobility suggests that XhoI and NcoI are digesting the vector as required.

Double digested PCR product (HCFM2-non-truncated SF1) and vector (pProEX HTb) were cleaned using a PCR purification kit (Roche). The double digested PCR product and vector were combined in a 1:1 volume ratio. As shown in Table 7 the negative control contained no ligase but did contain an additional restriction enzyme that should cut any residual non-cut plasmid that still contains the MCS region of DNA between the XhoI and NcoI sites, this tube was expected to produce no circular plasmid and therefore due to the reduced transformation efficiency of linear DNA, not produce colonies when *E.coli* are grown on a selective medium. The uncut vector control contained no ligase or EcoRI; this control was used to test the quality of the previous digestion, the number of colonies after transformation with this sample was expected to be very low because only linear DNA was seen in lanes 3 through 5 of Figure 10. The ligation reaction tube contained ligase and was expected to show an increase in the number of colonies present over the uncut vector control, as the possible products of the ligation may include previously uncut plasmids a fourth reaction was included. Ligation, reduced vector reaction included EcoRI with the ligase. The insert, double digested vector and ligation products which do not include the EcoRI site are not cut by EcoRI. Including EcoRI was expected to reduce the number of colonies by linearizing the base vector thereby increasing the proportion of vector with insert over vector alone.

Table 7: Ligation setup; pProEX HTb and HCFM2-non-truncated SF1

	Negative control	Uncut Vector Control	Ligation	Ligation, reduce vector only products
Vector PCR product (1:1 ratio v/v) (µL)	10	10	10	10
10x Ligation buffer(µL)	3	3	3	3
T4 Ligase(µL)	0	0	1	1
EcoRI(µL)	0.5	0	0	0.5
H_2O(µL)	16.5	17	16	15.5

Ligation was carried out in PCR tubes for three hours at 37°C.

Competent cells were transformed according to 2.2.1.3 Competent Cell Transformation using ligase reaction product; a positive control using pProEX HTb was included for transformation efficiency. Competent cells were spread onto AMP (0.1 mg/mL) containing LB-agar plates. Plates were incubated overnight at 37°C.

Only the positive control plate showed growth, the size colour and number of colonies was typical (~200-300 colonies).

To counter the possibility that the digest reactions had been contaminated the double digest protocol from Table 6 was repeated. A new ligation of the digest products was setup this time with varying molar ratios between the vector and insert in different reactions (see Table 8). By increasing the molar ratio of insert to vector it increases the chance of getting the correct ligation products.

Table 8: Ligation setup; pProEX HTb and HCFM2-SF1-non-truncated with ratios

	-Ligase +EcoRI	-Ligase -EcoRI	+Ligase -EcoRI	+Ligase +EcoRI	+Ligase +EcoRI	+Ligase +EcoRI
DNA molar ratio (vector: insert)	1:0	1:0	1:0	1:6	1:3	1:0.74
10x Ligation buffer (µL)	3	3	3	3	3	3
T4 Ligase (µL)	0	0	1	1	1	1
EcoRI (µL)	0.5	0	0	0.5	0.5	0.5
H_2O (µL)	16.5	17	16	15.5	15.5	15.5

Ligation was carried out in PCR tubes for three hours at 37°C.

Products of the ligation reaction were used to transform Top10 competent *E.coli* as per 2.2.1.3 Competent Cell Transformation. Competent cells were spread onto AMP (0.1 mg/mL) containing plates. Plates were incubated overnight at 37°C. Only the positive control (transformed with undigested pProEX HTb) and the ligation reaction with a molar ratio 1:0.74 Vector : Insert grew colonies. While the control had over 200 colonies the ligation product colonies only numbered 6.

Two colonies of Top10 *E.coli* were picked from the pProEX HTb+CFM2 ligation plate. Both were grown in 1 mL LB/AMP broth at 37°C and had their plasmid isolated in a mini prep, the steps are detailed in 2.2.6 Plasmid Isolation. To see if the colony plasmid mini preps contained the correct insert the colony mini preps were subjected to digestion by restriction enzymes XhoI and NcoI. If the correct insert was in the plasmid the restriction digest should liberate a dsDNA fragment of approximately 500bp. Mini prep plasmids pProEX HTb + CFM2 #1 and pProEX HTb + CFM2 #2 were digested by mixing reagents according to Table 9.

Table 9: Double digest mix for pProEX HTb+CFM2 plasmids #1 and #2

	Negative Control	Vector +NcoI	Vector +XhoI	Vector +NcoI +XhoI
Buffer H (10x)(µL)	5	5	5	5
DNA(µL)(pProEX HTb+CFM2 #1/#2)	5	5	5	5
NcoI(µL)	-	-	1	1
XhoI(µL)	-	1	-	1
H_20(µL)	40	39	39	38

Tubes were incubated for three hours at 37°C.

Restriction enzyme digestion of pProEX HTb+CFM2 #1 and #2 were visualised using gel electrophoresis. The gel was run and analysed according to 2.2.3 DNA analysis using ethidium bromide staining of agarose gel.

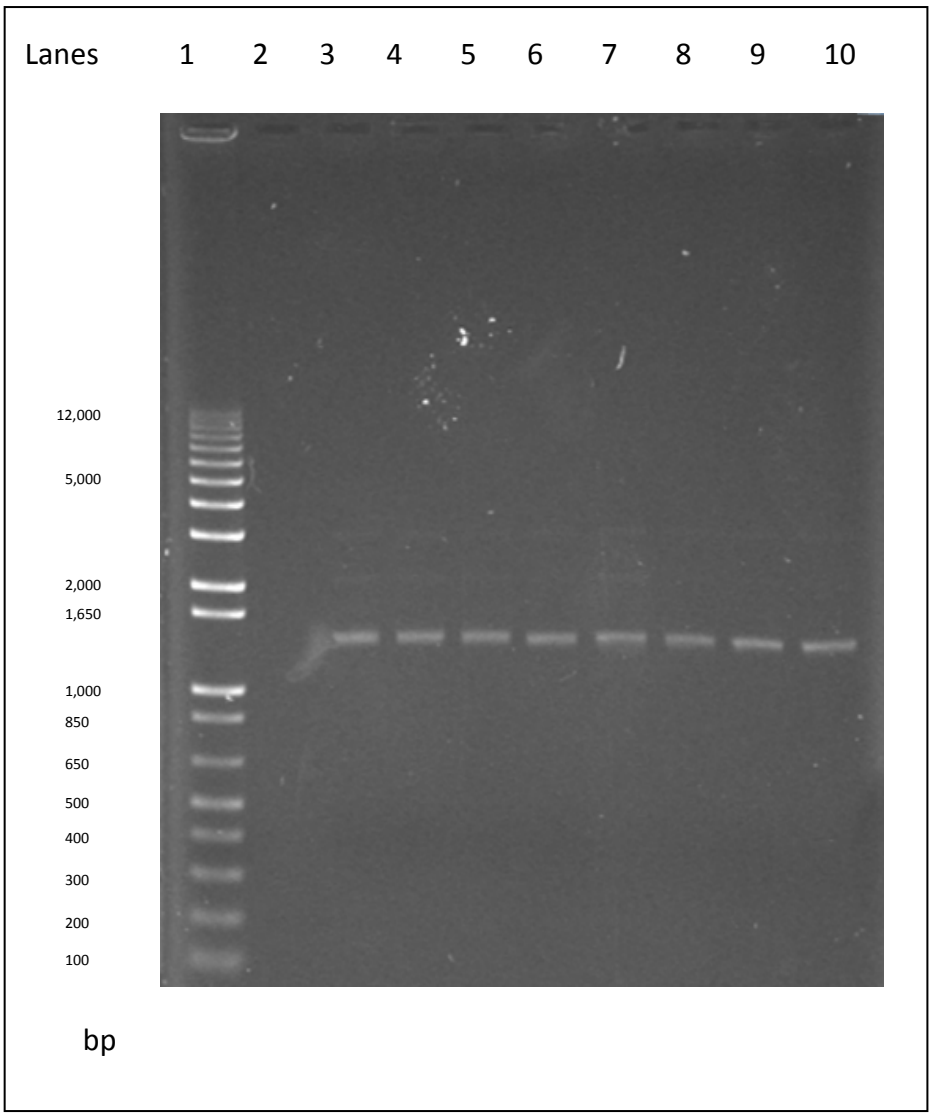


Figure 11: Double digest of pProEX HTb+CFM2 plasmid mini preps

pProEX HTb+CFM2 plasmid mini preps #1 and #2 (5µL) were digested with 1µL of NcoI and 1µL of XhoI in separate reactions. pProEX HTb+CFM2 plasmid mini preps #1 and #2 were each double digested with 1µL of XhoI and 1µL of NcoI. Digestion took place over 3 hours at 37°C. pProEX HTb+CFM2 plasmid mini preps #1 and #2 (5µL) were treated to identical conditions as the previous reactions without the restriction enzymes as a negative control. Digestion products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**-**
	3	**5µL pProEX HTb+CFM2 #2, +XhoI +NcoI**
	4	**5µL pProEX HTb+CFM2 #2, +NcoI**
	5	**5µL pProEX HTb+CFM2 #2, +XhoI**
	6	**5µL pProEX HTb+CFM2 #2, -XhoI –NcoI (uncut control)**
	7	**5µL pProEX HTb+CFM2 #1, +XhoI +NcoI**
	8	**5µL pProEX HTb+CFM2 #1, +NcoI**
	9	**5µL pProEX HTb+CFM2 #1, +XhoI**
	10	**5µL pProEX HTb+CFM2 #1, -XhoI –NcoI (uncut control)**

The bands present in Figure 11 suggest that either that both of the restriction enzymes did not work and that equal amounts of supercoiling was present in both plasmid samples (highly unlikely) or that the Buffer H, Sterile water and/or DNA were contaminated with another restriction enzyme.

In order to find out in another way whether the insert (HCFM2-non-truncated SF1) was present in the plasmid mini preps a PCR was setup using the HCFM2SF1 and HCFM2B1 primers as forward and back primers respectively. The PCR tubes were setup with reagents according to Table 10.

Table 10: Confirmation PCR of insert in plasmid mini prep #1 and #2

	Negative Control	**pProEX HTB+CFM2 #1**	**pProEX HTB+CFM2 #2**	**Positive Control (pGemT:FAM101a)**
Template DNA (µL) (2 ng/µL)	0	2	2	2
10xPwo Buffer (µL)	5	5	5	5
dNTPs (µL) (10 mM)	1.25	1.25	1.25	1.25
HCFM2SF1 forward primer (µL) (10 pmol/µL)	1	1	1	1
HCFM2SF1 back primer (µL) (10 pmol/µL)	1	1	1	1
Pwo (µL) (5U/µL)	0.5	0.5	0.5	0.5
H_2O (µL)	41.25	39.25	39.25	39.25

PCR cycle found at 2.2.2.1 PCR amplification.

The PCR products of the plasmid mini preps using HCFM2SF1 and HCFM2B1 as specific primers for the insert of interest were analysed using 2.2.3 DNA analysis using ethidium bromide staining of agarose gel. The results can be seen in Figure 12.

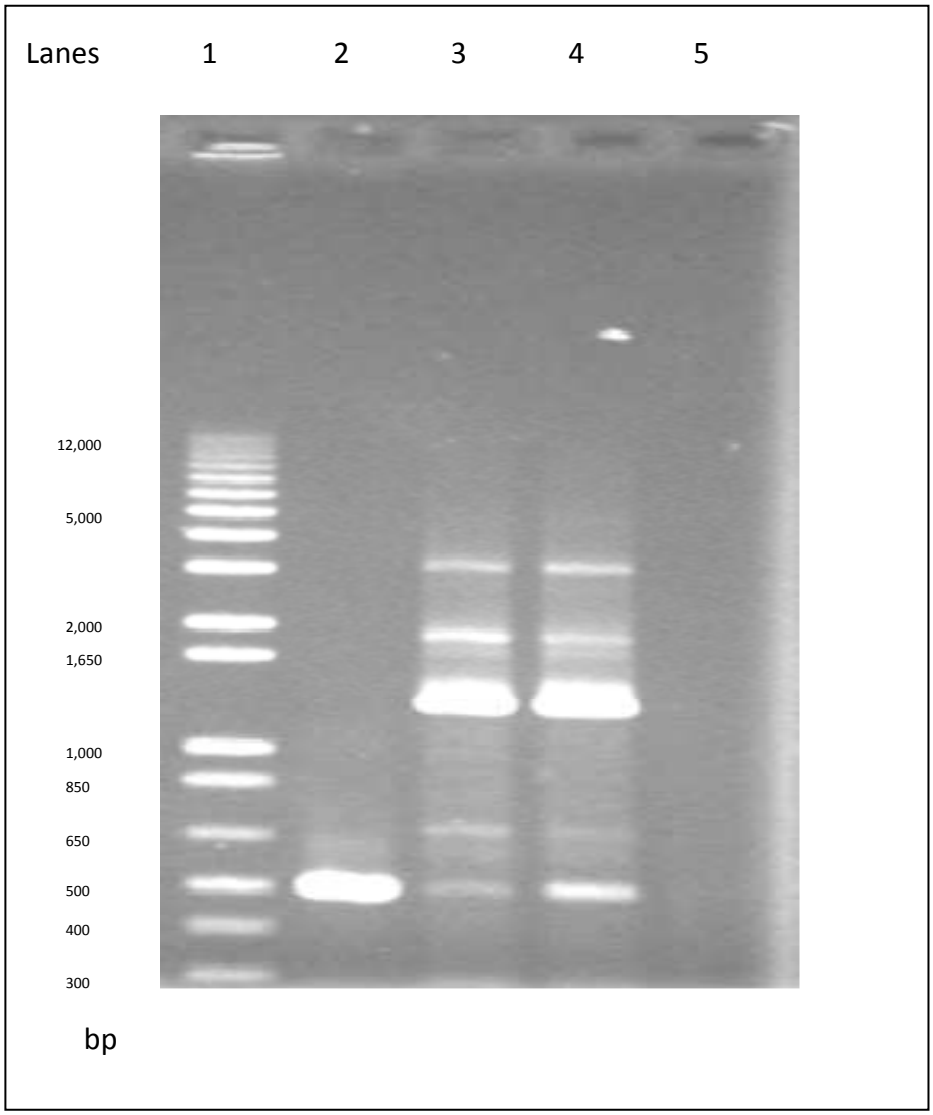


Figure 12: PCR amplification products of pProEX HTb+CFM2 #1 and #2 plasmids using HCFM2SF1 and HCFM2B1 primers

Plasmids from mini preps of colonies made with ligated pProEX HTb vector and HCFM2-non-truncated SF1 insert were amplified using PCR to determine if the insert was present in the plasmids. A negative control using no template DNA was included as well as a positive control where pGemT: FAM101A was used as the template DNA. The primers used were HCFM2SF1 and HCFM2B1 as the forward and reverse primers respectively; these were specific to the insert DNA. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes 1 5µL 1kb plus (Invitrogen) DNA ladder
2 5µL (Positive control) pGEM-T:FAM101A,+ SF1 and B1 primers
3 5µL pProEX HTb+CFM2 #2,+ SF1 and B1 primers
4 5µL pProEX HTb+CFM2 #1,+ SF1 and B1 primers
5 5µL (Negative control) No template,+SF1 and B1 primers

Mini-preps #1 and #2 showed a band of the same size as the positive control suggesting that the CFM2 gene is indeed inserted into the plasmids but the large number of unidentified bands suggests either that there are a large number of copies of the insert or more likely that there is a contaminant in the plasmid. The contaminant could be either introduced into the PCR tube before the PCR machine cycled or it could be more than one plasmid present in the original mini prep tubes.

Mini prep pProEX HTb + CFM2 #1 (HCFM2-non-truncated, SF1) was sequenced using forward primer HCFM2SF1 and reverse primer HCFM2B1. The sequence was of very poor quality which gave support to the conclusion that multiple templates were present.

To try and determine the size of the insert and whether they contained the CFM2 gene, colony PCR was performed using the M13R-48 and HTRVS primers using the procedure from 2.2.2.1 PCR amplification except a sterile pick of a single colony was used in place of template DNA. These primers are specific to the regions of pProEX HTb flanking the multiple cloning site. If the plasmid contains an insert there should be an increase in the size of the amplified region by ~480 bp larger that the positive control per insert copy. The HCFM2SF1 and HCFM2B1 were also used to screen each colony to determine whether any insert in the MCS of pProEX HTb plasmid is CFM2.

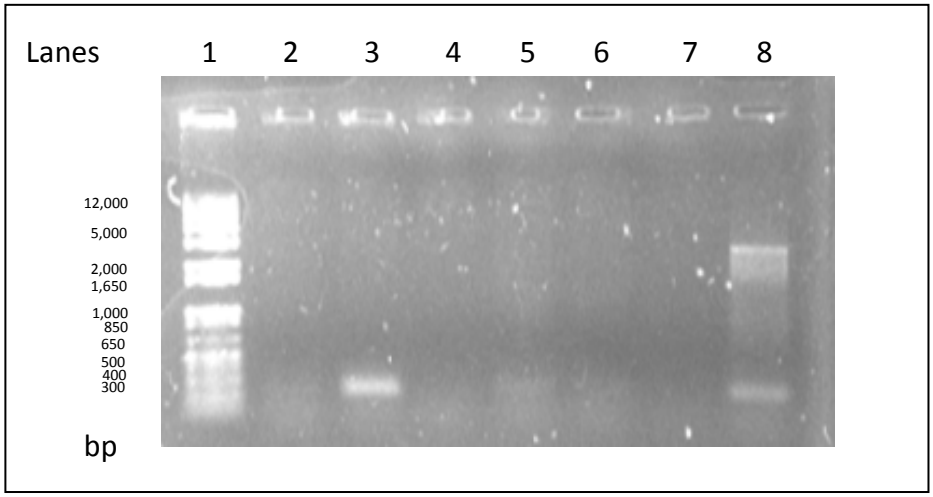


Figure 13: Colony PCR of pProEX HTb+CFM2 ligation product transformants using vector primers

Colonies made with ligated pProEX HTb vector and HCFM2-non-truncated SF1 insert were amplified using colony PCR using vector primers to determine if the insert was present in the plasmids. A negative control using no template DNA was included as well as a positive control where pProEX HTb was used as the template DNA. The primers used were M13R-48 and HTRVS as the forward and reverse primers respectively, these were specific to the region flanking the multiple cloning site. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL No Template, + HTRVS and M13R-48 primers (negative control)**
	3	**5µL pProEX HTb, + HTRVS and M13R-48 primers (positive control)**
	4	**5µL Colony #5,+ HTRVS and M13R-48 primers**
	5	**5µL Colony #4,+ HTRVS and M13R-48 primers**
	6	**5µL Colony #3, + HTRVS and M13R-48 primers**
	7	**5µL Colony #2,+ HTRVS and M13R-48 primers**
	8	**5µL Colony #1,+ HTRVS and M13R-48 primers**

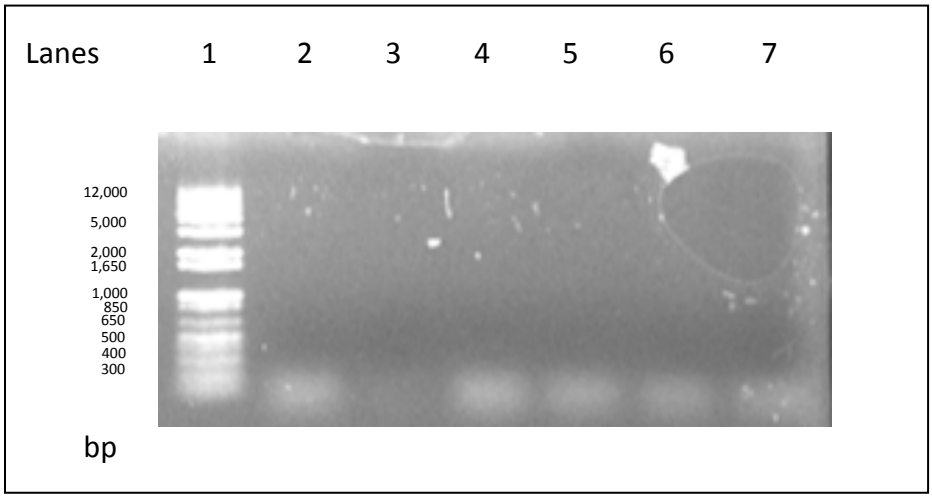


Figure 14: Colony PCR of pProEX HTb+CFM2 ligation product transformants using insert primers

Colonies transformed with ligated pProEX HTb vector and HCFM2-non-truncated SF1 insert were screened by colony PCR using insert primers to determine if the CFM2 insert was present in the plasmids. A negative control where pProEX HTb was used as the template DNA was included. The primers used were HCFM2SF1 and HCFM2B1 as the forward and reverse primers respectively; these were specific to the CFM2 insert. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL pProEX HTb (negative control)**
	3	**5µL Colony #5**
	4	**5µL Colony #4**
	5	**5µL Colony #3**
	6	**5µL Colony #2**
	7	**5µL Colony #1**

Colony PCR showed a band in Figure 13 lane 8 indicating the presence of an insert in the MCS of pProEX HTb, unfortunately another band was present in the same lane suggesting at least two different plasmids were present (or at least more than one site was recognised by the primers). The larger band (~5,000 bp) was much larger than the positive control combined with the PCR product HCFM2-non-truncated SF1 that was to be inserted and the smaller band was of the same size as the positive control. Whatever the nature of the larger band in Colony #1 the lack of a band in Figure 14 lane 7 which used colony #1 in colony PCR but used the CFM2 primers HCFM2SF1 and HCFM2B1 makes the presence of CFM2 in any of the plasmids unlikely.

Another round of digestion using the same method as Table 6 used a fresh Buffer H as it may have been contaminated. Ligation was performed using the same setup in Table 8 using the 1:0.74 Vector : Insert molar ratio. The ligation product was used to transform Top10 *E.coli* using 2.2.1.3 Competent Cell Transformation. Six colony picks

were taken and colony PCR was performed using the vector primers. To try and determine the size of the insert and whether they contained CFM2 gene colony PCR was performed using the M13R-48 and HTRVS primers using 2.2.2.1 PCR amplification, except a sterile pick of a single colony was used in place of template DNA.

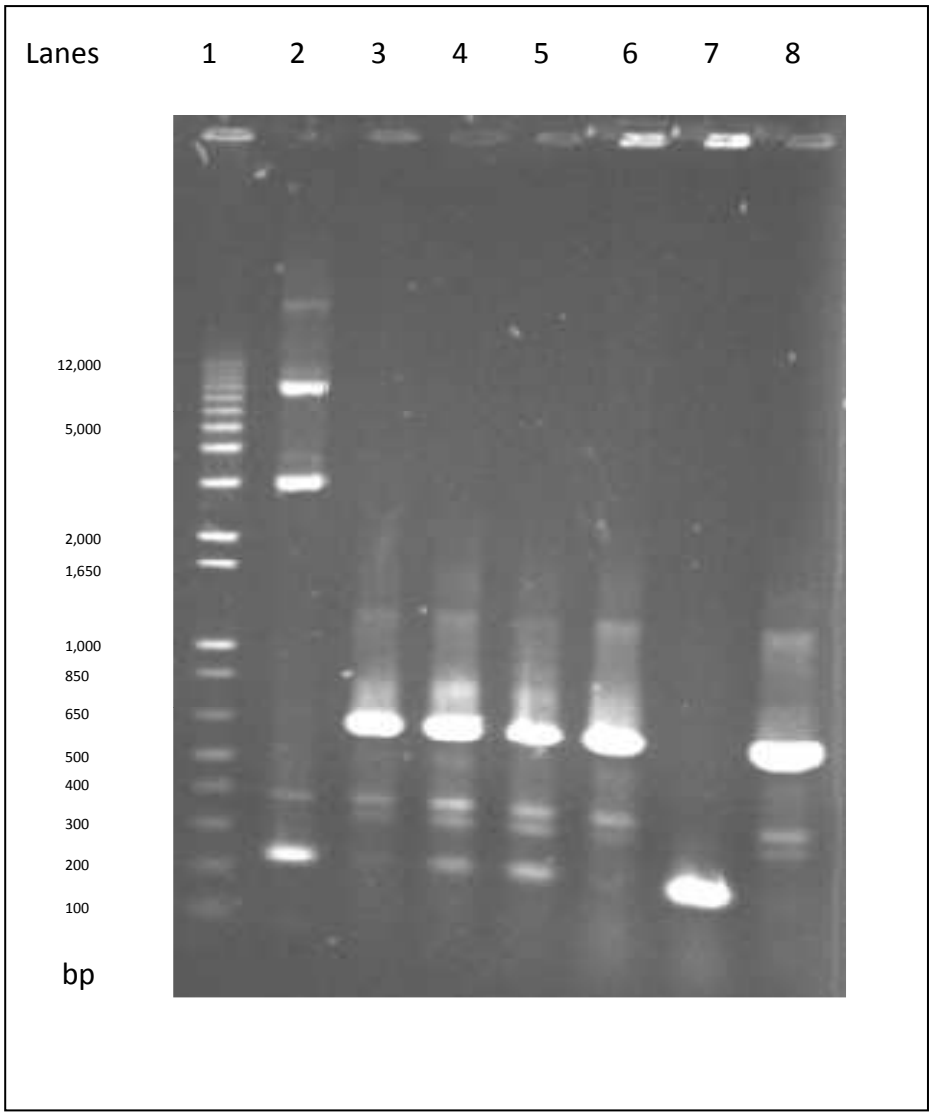


Figure 15: Colony PCR of pProEX HTb+CFM2 ligation product transformants using vector primers

Colonies transformed with ligated pProEX HTb vector and HCFM2-non-truncated SF1 insert were screened by colony PCR using vector primers to determine if the insert was present in the plasmids. A positive control where pProEX HTb was used as the template DNA was included. The primers used were M13R-48 and HTRVS as the forward and reverse primers respectively, these were specific to the region flanking the multiple cloning site in pProEX HTb. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualized by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL pProEX HTb (positive control)**
	3	**5µL Colony #6**
	4	**5µL Colony #5**
	5	**5µL Colony #4**
	6	**5µL Colony #3**
	7	**5µL Colony #2**
	8	**5µL Colony #1**

In the colony PCR shown in Figure 15 all of the colonies except colony #2 appear to have an insert. By taking the band shown in the positive control in lane 2 of approximately 200bp and subtracting the MCS region lost due to double digest (77bp) as shown in Table 11 then adding the HCFM2-non-truncated SF1 PCR product (~500bp) the expected band of ~623bp is created. Believing this band to be represented in colonies #1, 3-6 colony #1 was picked and a plasmid mini-prep was preformed as described in 2.2.6 Plasmid Isolation. The plasmid was sent for sequencing (see 2.2.7 DNA sequencing), in the 5`>3` and 3`>5` directions using HTRVS and M13R-48 primers. The chromatogram shown in Table 22 was viewed and cleaned using Sequence Scanner v1.0 from Applied Biosciences, the two sequence data files were aligned producing a contig using Serial Cloner 1.3-11.

Table 11: Predicted digest size of pProEX HTb using NcoI and XhoI

Restriction analysis of pProEX-HTB [Circular]
Incubated with XhoI + NcoI
2 fragments generated.
1: 4,099 bp - From XhoI[417] To NcoI[340]
2: 77 bp - From NcoI[340] To XhoI[417]

The pProEX HTb:CFM2 #1 contig was aligned using Serial Cloner 1.3-11 (Table 23) the insert in plasmid pProEX HTb:CFM2 #1 contains the HCFM2-non-truncated SF1 protein sequence.

Having successfully created a plasmid containing the CFM2 gene albeit only the HCFM2-non-truncated SF1 isoform and not the HCFM2-non-truncated LF1 isoform, the pProEX HTb -CFM2SF1-non-trucated colony #1 plasmid was used to transform BL21 competent *E.coli* (2.2.1.3 Competent Cell Transformation). The transformed cells were plated onto LB Agar plates containing AMP at 0.1 mg/mL then grown overnight at 37°C and colony picks from these plate transformations were used in expression experiments.

3.1.2 Expression of CFM2

Only the shorter isoform of CFM2 was able to be amplified using PCR. As the premature stop in the plasmid is not found in the NCBI FAM101A listing, expressing the HCFM2-non-truncated SF1 (short isoform missing the premature stop codon) was focused on.

A timed expression trial was setup using BL21 *E.coli* cells transformed with the pProEX HTb -CFM2SF1-non-truncated #1 plasmid. Transformants were plated onto LB Agar plates containing AMP at 0.1 mg/mL. Colony picks were used to inoculate 10 mL of sterile LB broth containing AMP (final concentration 0.1 mg/mL). The inoculated LB broth was grown at 37°C to approximately 0.6 OD as determined using visible light absorption at 600nm. At ~0.6 OD the broth was inoculated with IPTG to a final concentration 0.1 mM. The induced broth was incubated at 37°C while shaking to maintain aeration. 1.5 mL samples of the induced broth were taken at time intervals 0, 30, 120, 180 minutes and a final sample after overnight incubation.

The expression time trial samples were lysed using sonication (Sonicator S-4000 (QSONICA)) at 10% power for 10 seconds repeated 4 times each; lysis was done in MQ water. The lysed samples were spun down to separate the pellet and soluble fractions. Pellet and supernatant factions were boiled with SDS+DTT containing loading gel; 6 µL of each was loaded onto a 12.5% SDS-PAGE gel. The gel was run for 45 minutes at 200v; PPU was used as a standard marker for size. The gel was stained with Coomassie Blue for 20 minutes and destained overnight. The expected protein product size is approximately 18.6kDa.

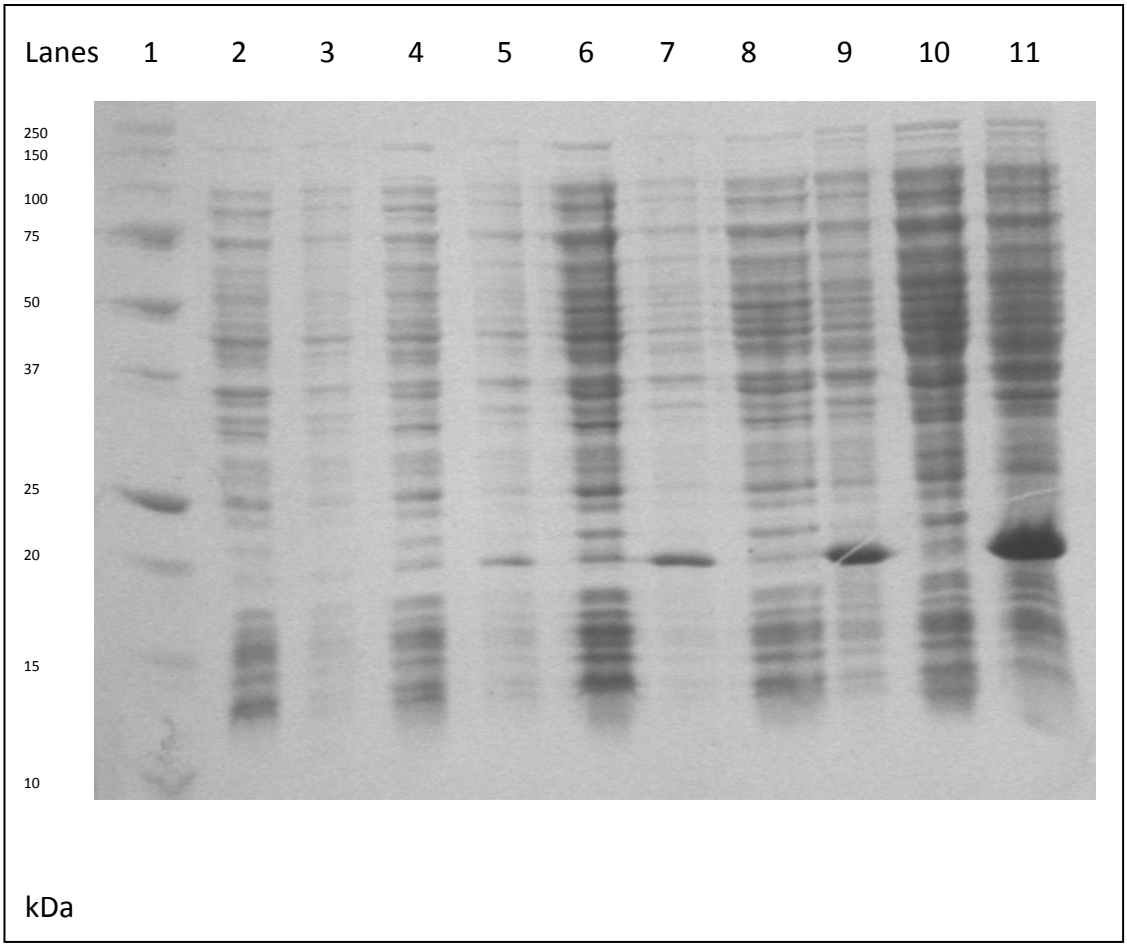


Figure 16: Time expression trial of CFM2 short isoform non-truncated

Expression was induced in BL21 *E.coli* were transformed with pProEX HTb: CFM2-short non-truncated Luria broth with IPTG (final concentration 0.1 mM) and incubated at 37°C. Samples (1.5 mL) were taken at T=0, 30, 120, 180 minutes and an overnight sample. Cells were lysed using sonication in MQ water. Lysed cells were separated into pellet and supernatant using centrifugation. Pellet and supernatant fraction were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 12.5% gel. Gel was run for 45 minutes at 200V. Gel was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**5µL Supernatant T=0min**
	3	**5µL Pellet T=0min**
	4	**5µL Supernatant T=30min**
	5	**5µL Pellet T=30min**
	6	**5µL Supernatant T=120min**
	7	**5µL Pellet T=120min**
	8	**5µL Supernatant T=180min**
	9	**5µL Pellet T=180min**
	10	**5µL Supernatant T=overnight**
	11	**5µL Pellet T=overnight**

From the time expression trial show in Figure 16 the expected protein increases in relative concentration once expression has been induced using IPTG. While the CFM2 protein is in the pellet and not the supernatant this may be expected due to the lysis buffer being water. As CFM2 has no known enzymatic activity and no other assays for determining that it is correctly folded exist the protein will not be able to be refolded from the insoluble pellet fraction as it would not be possible to know whether it is in its native state. In order to retain CFM2 in a native folded state different lysis buffers were used on an overnight expression of pProEX HTb CFM2 short isoform non-truncated in BL21 *E.coli*.

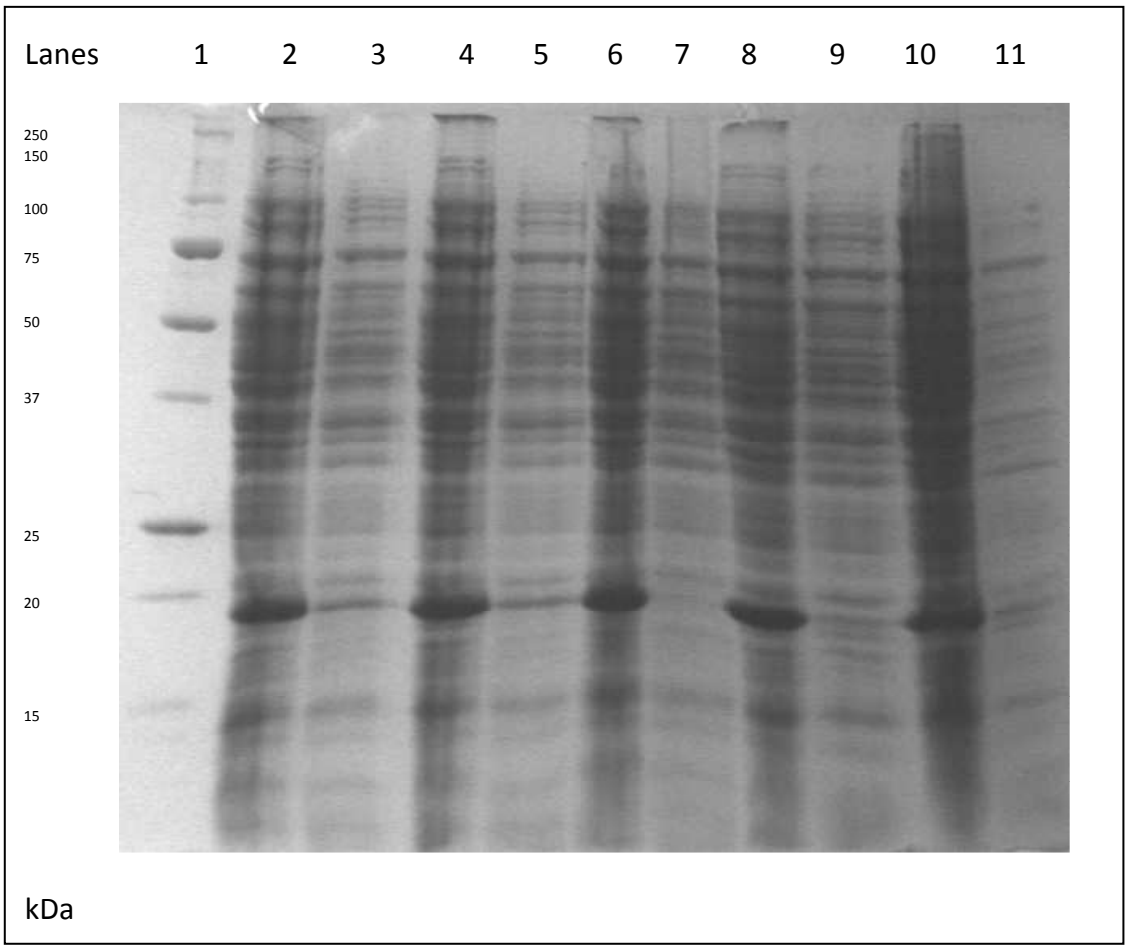


Figure 17: Lysis buffer trial using overnight expression of CFM2 short isoform non-truncated

BL21 *E.coli* were transformed with pProEX HTb: CFM2-short non-truncated and colony picks were used to inoculate sterile broth. The colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM) and incubated overnight at 37°C. Cells were lysed using sonication in various buffers. Lysed cells were separated into pellet and supernatant fractions using centrifugation. Pellet and supernatant fraction were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 12.5% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**5µL 1xNative Buffer +0.1 mM EDTA+DTT Pellet**
	3	**5µL 1xNative Buffer +0.1 mM EDTA+DTT Supernatant**
	4	**5µL 1xNative Buffer +0.1 mM EDTA Pellet**
	5	**5µL 1xNative Buffer +0.1 mM EDTA Supernatant**
	6	**5µL 1.0M Tris HCL pH 6.8 Pellet**
	7	**5µL 1.0M Tris HCL pH 6.8 Supernatant**
	8	**5µL 1.5M Tris HCL pH 8.8 Pellet**
	9	**5µL 1.5M Tris HCL pH 8.8 Supernatant**
	10	**5µL 1xNative Buffer Pellet**
	11	**5µL 1xNative Buffer Supernatant**

Solubility of the protein of interest (CFM2) increased the most under lysis conditions using Native buffer and 0.1 mM EDTA as shown in Figure 17. The pellet fraction still contained a large proportion of the CFM2 protein, higher concentrations of EDTA and Tween 20 were used as Lysis buffer additives to try and increase the proportion of soluble to insoluble CFM2.

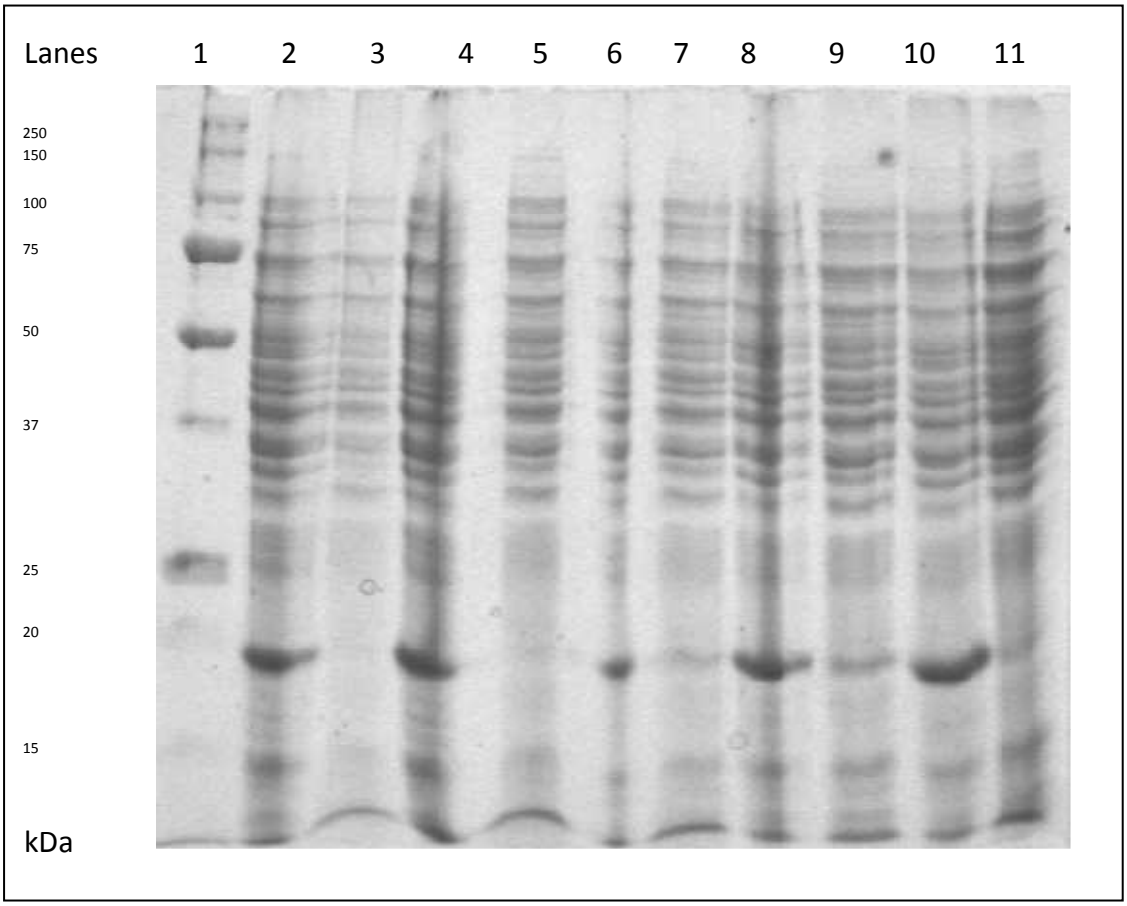


Figure 18: Lysis buffer trial 2 using overnight expression of CFM2 short isoform non-truncated

BL21 *E.coli* were transformed with pProEX HTb: CFM2-short non-truncated and colony picks were used to inoculate sterile broth. The colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM) and incubated overnight at 37°C. Cells were lysed using sonication in various buffers. Lysed cells were separated into pellet and supernatant using centrifugation. Pellet and supernatant fraction were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 12.5% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**5µL 1xNative Buffer +0.5% Tween 20 Pellet**
	3	**5µL 1xNative Buffer +0.5% Tween 20 Supernatant**
	4	**5µL 1xNative Buffer +0.25% Tween 20 Pellet**
	5	**5µL 1xNative Buffer +0.25% Tween 20 Supernatant**
	6	**5µL 1xNative Buffer +5 mM EDTA Pellet**
	7	**5µL 1xNative Buffer +5 mM EDTA Supernatant**
	8	**5µL 1xNative Buffer +1 mM EDTA Pellet**
	9	**5µL 1xNative Buffer +1 mM EDTA Supernatant**
	10	**5µL 1xNative Buffer Pellet**
	11	**5µL 1xNative Buffer Supernatant**

The solubility of CFM2 did not appear to increase due to Tween 20 as shown in Figure 18. The increase in solubility if any due to 5 mM EDTA over the solubility due to 1 mM EDTA was decided not to be high enough to warrant using it in the lysis of the BL21 *E.coli* containing CFM2 as nickel-NTA-agarose columns lose their binding capacity when in the presence of excess chelating agents like EDTA. While the greater proportion of CFM2 is in the insoluble phase Native buffer 1x + 1 mM EDTA was used as the lysis buffer.

3.1.3 Purification of CFM2

The CFM2 lysis buffer while being poor at producing large proportions of soluble CFM2 was expected to produce enough to begin designing a purification procedure. BL21 *E.coli* transformed with pProEX HTB CFM2 short isoform non-truncated inoculated 500 mL of Luria Broth containing AMP at 0.1 mg/mL and, after incubation at 37°C, at ~0.6 OD A_{600} was induced using IPTG up to a final concentration of 0.1 mM. After overnight induction at 37°C cells were pelleted at 6000 x g for 20 minutes at 4°C. The 500 mL cell pellets were re-suspended using 35 mL of Native buffer + 1 mM EDTA.

The cells were lysed in a French press twice and again using sonication for 30 seconds 3x at 30% (of 500W). The lysate debris was removed by centrifugation at 30000 x g for 20 minutes at 4°C.

The 10 mL nickel based His-Trap (GE Healthcare) column was washed with 20x column volume of Native buffer +10 mM imidazole. As imidazole competes for the histidine binding sites of the column, this low concentration of imidazole was expected to lower the non-specific binding of *E.coli* proteins preventing the column from becoming over loaded, thereby increasing the proportional concentration and total amount of CFM2 on the column. The column loading, washing and elution steps are described in Table 12. Fractions from each step were collected for SDS-PAGE analysis.

Table 12: His-Trap Column loading washing and elution protocol

Step	Flow rate	Buffer
Loading Lysis supernatant	1 mL/min	1xNative buffer +1 mM EDTA and *E.coli* supernatant (~30 mL)
Wash	1 mL/min	1xNative buffer + 10 mM imidazole (20 mL)
Gradient	1 mL/min	1xNative buffer + 10 mM imidazole to 500 mM imidazole (over 87.5 mL)
Elution	1 mL/min	1xNative buffer + 500 mM imidazole (20 mL)

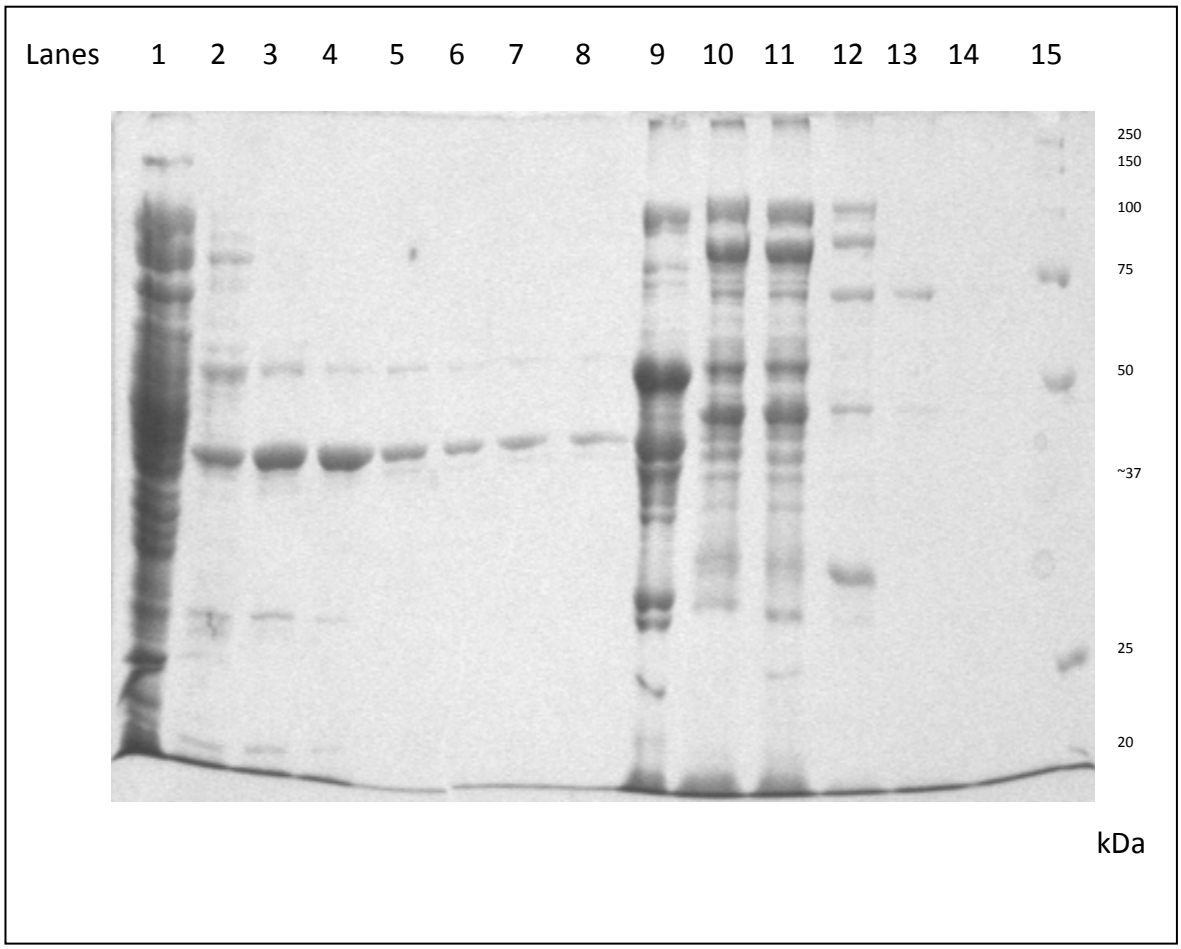


Figure 19: His-Trap purification of BL21 *E.coli* expressing CFM2 short isoform non-truncated

BL21 *E.coli* were transformed with pProEX HTb: CFM2-short non-truncated and colony picks were used to inoculate sterile broth. The colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM) and incubated overnight at 37°C. Cells were lysed using sonication in Native buffer with 1 mM EDTA. Lysed cells were separated into pellet and supernatant using centrifugation. The supernatant fraction was loaded onto a His-Trap according to Table 12; the loaded column was washed and eluted of proteins using imidazole. Fractions of each different stage were collected and samples were loaded with SDS-PAGE loading buffer and loaded into an SDS-PAGE 12.5% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the right in kDa.

Lanes	**1**	**5µL Column loading Flowthrough**
	2	**5µL Wash step fraction 1**
	3	**5µL Wash step fraction 2**
	4	**5µL Wash step fraction 3**
	5	**5µL Elution Gradient fraction 5**
	6	**5µL Elution Gradient fraction 6**
	7	**5µL Elution Gradient fraction 7**
	8	**5µL Elution Gradient fraction 8**
	9	**5µL Elution Gradient fraction 9**
	10	**5µL Elution Gradient fraction 10**
	11	**5µL Elution Gradient fraction 11**
	12	**5µL Elution Gradient fraction 12**
	13	**5µL Elution Gradient fraction 13**
	14	**5µL Elution Gradient fraction 14**
	15	**5µL PPU**

The fractions chosen for analysis by SDS-PAGE were those that contained a reasonable level of protein (within the detectible range of Coomassie at 280nm stained SDS-PAGE gel ~ 50 ng) as determined by nano drop spectrometry. Also almost all proteins will be competed from the column between 20 mM and 200 mM imidazole. Due to the gel having very poor resolution in the lower size range where the CFM2 protein should be, all Figure 19 can show is that a large number of *E.coli* proteins remain on the His-Trap even at fairly high imidazole concentrations. While the fact *E.coli* protein being retained rules out the possibility of a single step purification for CFM2, if the difference in size between CFM2 and any *E.coli* proteins that come off the column at the same Imidazole concentration as it were big enough, then size exclusion chromatography could be used to separate the *E.coli* proteins from the CFM2 protein.

To narrow down the concentration of imidazole at which CFM2 is eluted from the Ni^{2+}-NTA-sepharose beads, smaller 1.5 mL tube based purification was used. A fresh 10 mL broth containing BL21 *E.coli* transformed with pProEX: CFM2 (short isoform, non-truncated) was grown to 0.6ODA_{600} then expression was induced by adding IPTG to a final concentration of 0.1 mM. The cells were induced overnight at 37°C. The cells were pelleted, lysed and loaded under the same conditions as those in Figure 19.

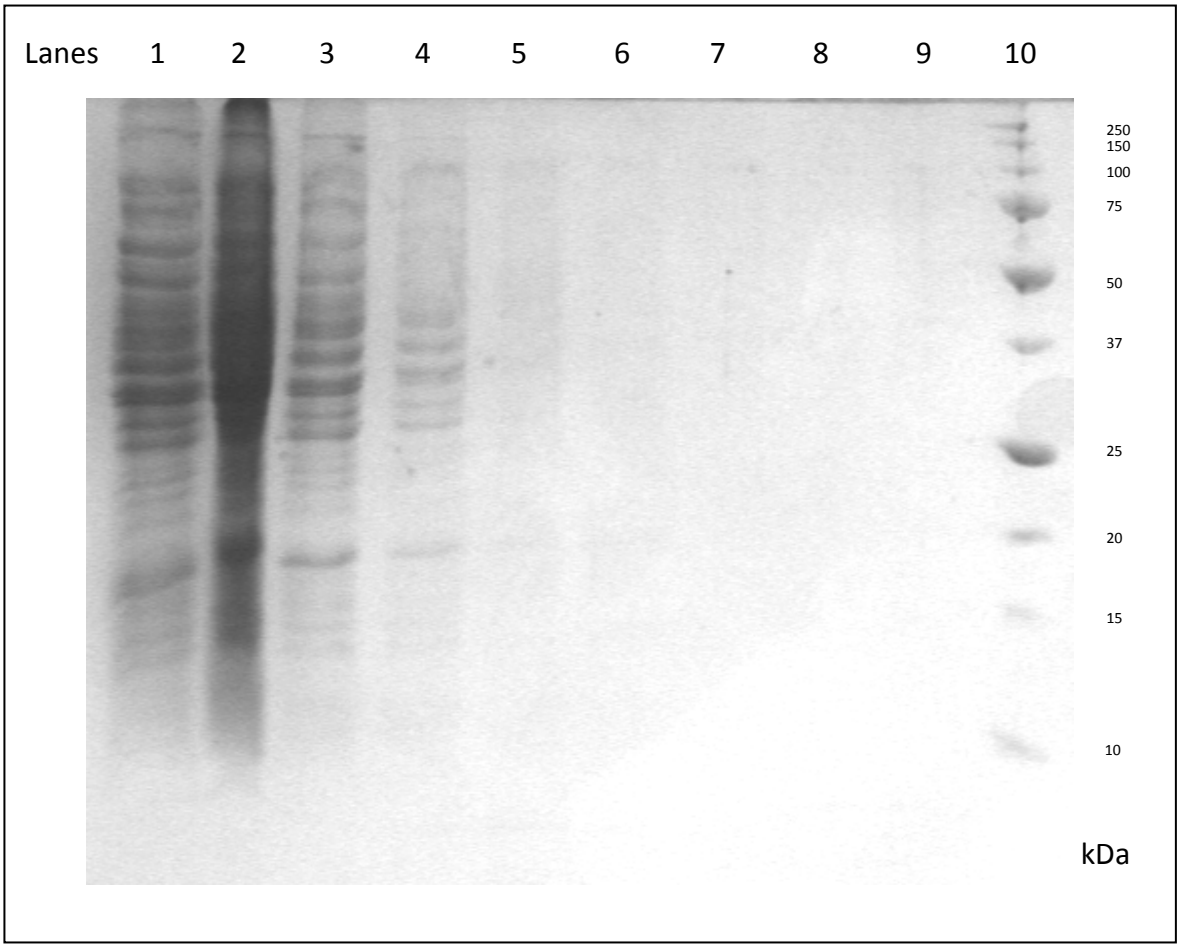


Figure 20: His Bead 1.5 mL Tube purification of BL21 *E.coli* expressing CFM2 short isoform non-truncated

BL21 *E.coli* were transformed with pProEX HTb: CFM2-short non-truncated and colony picks were used to inoculate 10 mL of sterile broth. The colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM) and incubated overnight at 37°C. Cells were lysed using sonication in Native buffer with 1 mM EDTA. Lysed cells were separated into pellet and supernatant using centrifugation. The supernatant fraction was loaded onto a His-Trap; the loaded column was washed and eluted of proteins using imidazole. Fractions of each different stage were collected and samples were loaded with SDS+DTT and loaded into an SDS-PAGE 15% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the right in kDa.

Lanes	**1**	**5µL Lysis Supernatant**
	2	**5µL Lysis Pellet**
	3	**5µL Unbound supernatant from beads**
	4	**5µL Elution 20 mM imidazole #1**
	5	**5µL Elution 20 mM imidazole #2**
	6	**5µL Elution 50 mM imidazole #1**
	7	**5µL Elution 50 mM imidazole #2**
	8	**5µL Elution 500 mM imidazole #1**
	9	**5µL Elution 500 mM imidazole #2**
	10	**5µL PPU**

In Figure 20 the expected product band just below 20kDa in size was not concentrated by the 1 mL His-Trap tube. It appears as though the 6xHis-CFM2 fusion protein did not bind to the column. As in this experiment the 6xHistidine encoded on the N-terminal end of the fusion protein sequence, it is unlikely that the histidine tag is not being transcribed or translated as this would be due to a frame shift. This would almost certainly result in a truncated fusion protein which would be obvious in the size of the band expressed. Due to the nature of how His-Trap columns work it usually wouldn't matter if the protein is unfolded as long as the conditions required to bind the 6xHistidine tag were correct. One option is that the histidine groups may be folded in such a way that they are hidden from the column, such as inside the protein. Another might include a protease removing the tag from the fusion protein.

To control for the possibility of the histidine groups being folded into the protein, preventing His-Trap based purification, would be to unfold the protein by lysing the cells in denaturing conditions. As CFM2 has no solved structure or any assay that can test for correct folding, refolding would not necessarily regain the proteins native structure making this approach unhelpful.

Another full scale preparation (500 mL broth) of BL21 *E.coli* expressing CFM2 short isoform non-truncated was made. Lysis and purification was performed using an identical setup to Figure 19.

Table 13: His-Trap loading extended washing and elution setup

Step	Flow rate	Buffer	Fractions
Loading lysis supernatant	1 mL/min	1xNative buffer +1 mM EDTA and *E.coli* supernatant (~30 mL)	1 through 4
Wash	1 mL/min	1xNative buffer + 10 mM imidazole (20 mL)	4 through 13
Wash Gradient	1 mL/min	1xNative buffer + 10 mM imidazole to 20 mM imidazole (over 120 mL)	14 through 53
Gradient	1 mL/min	1xNative buffer + 20 mM imidazole to 250 mM imidazole (over 140 mL)	54 through 89

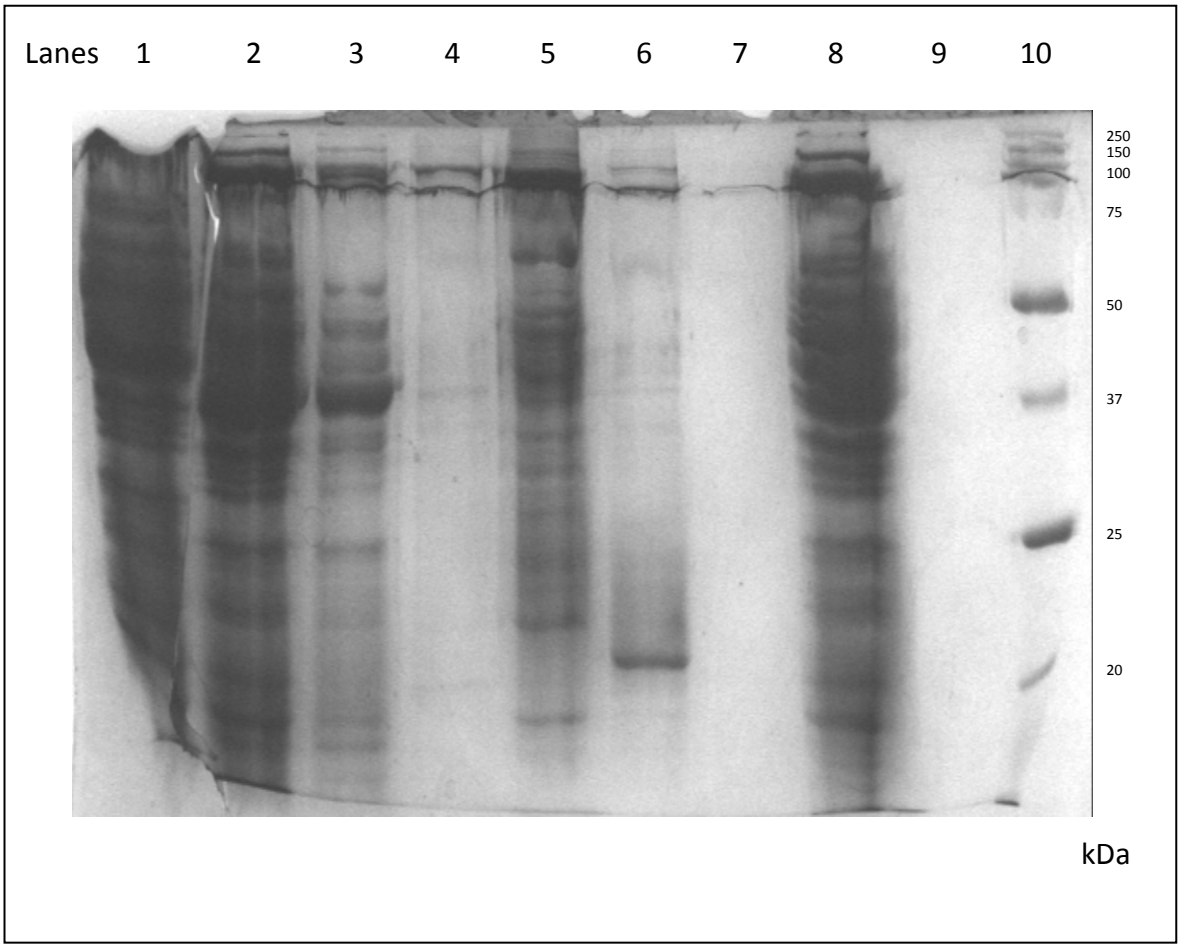


Figure 21: His-Trap purification of BL21 *E.coli* expressing CFM2 short isoform non-truncated

BL21 *E.coli* were transformed with pProEX HTb: CFM2-SF1 non-truncated and colony picks were used to inoculate 10 mL of sterile broth. The colony broth was grown to ~0.6 OD at 37°C. The 10 mL broth was used to inoculate 500 mL of sterile broth. The 500 mL colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM) and incubated overnight at 37°C; cells were pelleted at 6000g for 20 min at 4°C. Cells were lysed using sonication in 37 mL of Native buffer with 1 mM EDTA. Lysed cells were separated into pellet and supernatant using centrifugation. The supernatant fraction was loaded onto a His-Trap according to Table 13; the loaded column was washed and eluted of proteins using Native buffer + imidazole. Fractions of each different stage were collected and samples were loaded with SDS+DTT and loaded into an SDS-PAGE 15% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the right in kDa.

Lanes	**1**	**5µL Wash step fraction 1**
	2	**5µL Wash step fraction 2**
	3	**5µL Wash step fraction 3**
	4	**5µL Elution Gradient fraction 66**
	5	**5µL Elution Gradient fraction 67**
	6	**5µL Elution Gradient fraction 68**
	7	**5µL Elution Gradient fraction 77**
	8	**5µL Flow-through**
	9	**-**
	10	**5µL PPU**

An error occurred during purification resulting in the stepping of the ratio of imidazole to go from a linear gradient to the highest concentration. This step in imidazole resulted in a peak of protein concentration which was analysed in Figure 21 as fractions 66 to 68. In lane 6 of Figure 21 a single band is present at approximately 20kDa. The estimated size of 6xHis-CFM2 protein product was 18.6kDa. As CFM2 has not being purified before, antibodies have not been raised to it meaning a specific confirmation of the band as belonging to CFM2 requires mass spectrometry or Western Blot of the His-tag. The elution fraction containing the CFM2 candidate (fraction 68) was stored at 4°C.

In order to run the fraction in a mass spectrometer the sample would need to be as pure as possible. A concentrator with a 10kDa cut off was used to concentrate the sample from 4 mL to 400 µL; the concentrator was topped with lysis buffer to remove the imidazole which can interfere with readings. All proteins below 10kDa were not retained by the concentrator.

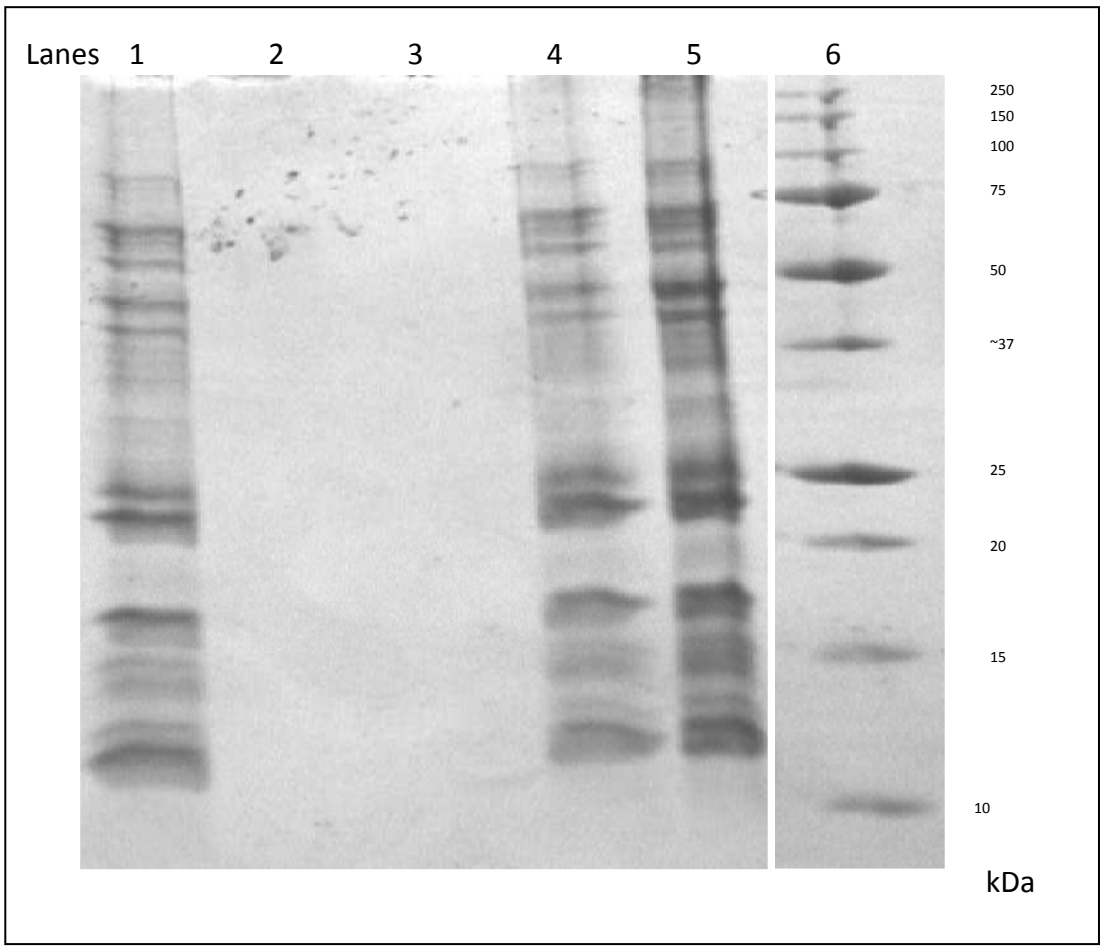


Figure 22: Concentration and purification of Fraction 68 using 10 kDa cut-off concentrator

Fraction 68 from Figure 21 lane 6 was concentrated using a 10 kDa cut-off concentrator. Buffer was exchanged with lysis buffer to remove imidazole. Fractions of each different stage were collected and samples were loaded with SDS+DTT and loaded into an SDS-PAGE 15% gel. The gel was run for 45 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the right in kDa. Lane between 5 and 6 was not related to this experiment, it was removed for clarity.

Lanes	**1**	**Elution tube 68 (pre 10 kDa filtration)**
	2	**Elution tube 68 (post 10 kDa filtrate)**
	3	**Elution tube 68 (post 10 kDa filtrate buffer exchanged)**
	4	**Elution tube 68 (post 10 kDa filtration concentrate)**
	5	**Elution tube 68 (post 10 kDa filtration pellet in concentrate)**
	6	**5µL PPU**

Shown in lanes 1, 4 & 5 of Figure 22, the proteins present in the stored fraction 68 had degraded. The single band that was hoped to represent CFM2 protein was lost. Possible reasons for the degradation include the presence of protease enzymes in the fraction, instability in the CFM2 protein due to incomplete or improper folding.

Further attempts at purifying CFM2 resulted in a lack of separation of the CFM2 sized protein band from the host *E.coli* proteins as see before in Figure 20, suggesting a lack of access to the 6x Histidine tag. In some cases the CFM2 sized band in the cell lysate was missing. Work by Postdoctoral Fellow Dr Greg Sawyer has also shown problems transforming CFM2 without mutations being present.

With the difficulty in transforming CFM2 into *E.coli*, the protein being mostly insoluble, the low stability and lack of separation of CFM2 from host proteins (possibly due to poor access for the 6xHis-tag to the substrate) this part of the project was put on hold.

It may be that one or more folding proteins or other conditions that *E.coli* lack compared to human is causing insolubility of the CFM2 protein. This may be the explanation for the low mass and slower growth of the *E.coli* transformed with the CFM2 expressing plasmid compared to the cloning vector pGemT-FAM101A. That the yeast 2-hybrid assay identified CFM2 (FAM101a) as a binding partner suggests eukaryote based expression systems may be able to produce correctly folded CFM2 more reliably.

3.1.4 Sub-cloning FLNAR10

The rationale behind the selection of the expression vector for FLNAR10 and the difficulties involved with sub-cloning FLNAR10 into the expression vector is described here.

The gene for FLNA is a eukaryotic gene; unlike prokaryotic genes introns are present which need to be removed before the gene can be translated correctly into protein.

The FLNA gene present in the vector was provided from a cDNA library, this means that the gene has already been transcribed into RNA; the introns removed, then reverse transcribed back into cDNA.

As shown in Figure 23 the pREP4 vector has an AMP resistance gene (β-lactamase) for maintaining selection and a pUC origin for high copy number in *E.coli*.

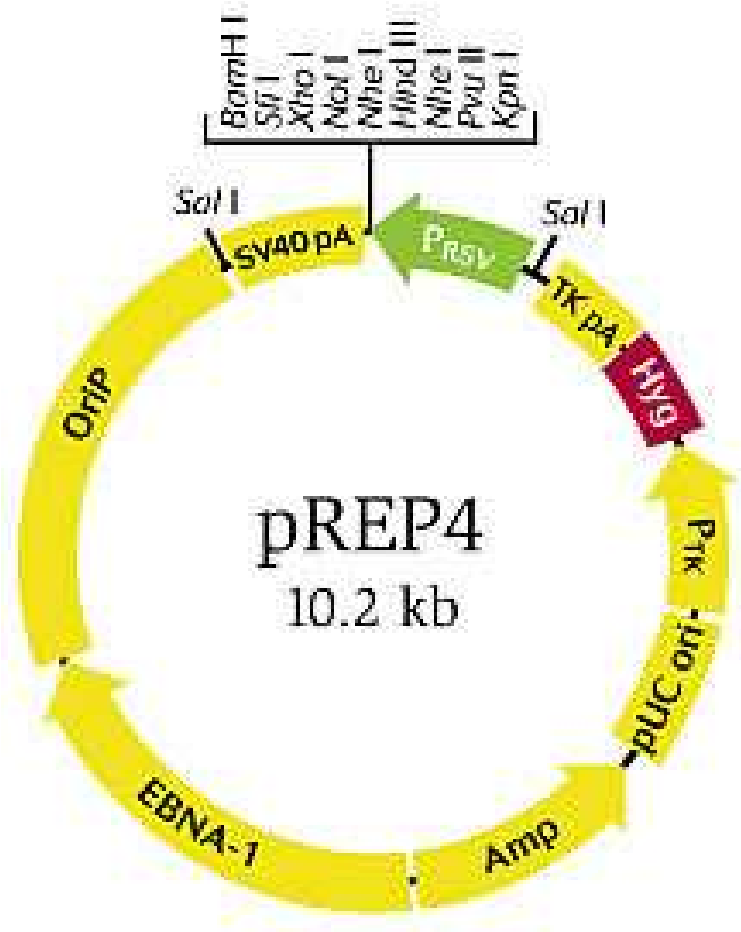


Figure 23: pREP4 Vector Map[36]

The pREP4 vector is primarily a mammalian expression vector. The pREP4 vector contained the entire FLNA gene; working with *E.coli* repeat 10 of FLNA needed to be subcloned into an *E.coli* expression system.

As crystallisation studies were hoped to be performed on the FLNA10 (thought to bind CFM2) a purification tag was desired to assist affinity chromatography.

The destination vector chosen was pGEX 4T3, a GST tag system with inducible expression via the lac-operon and ampicillin resistance based selection / plasmid maintenance (Shown in Figure 24).

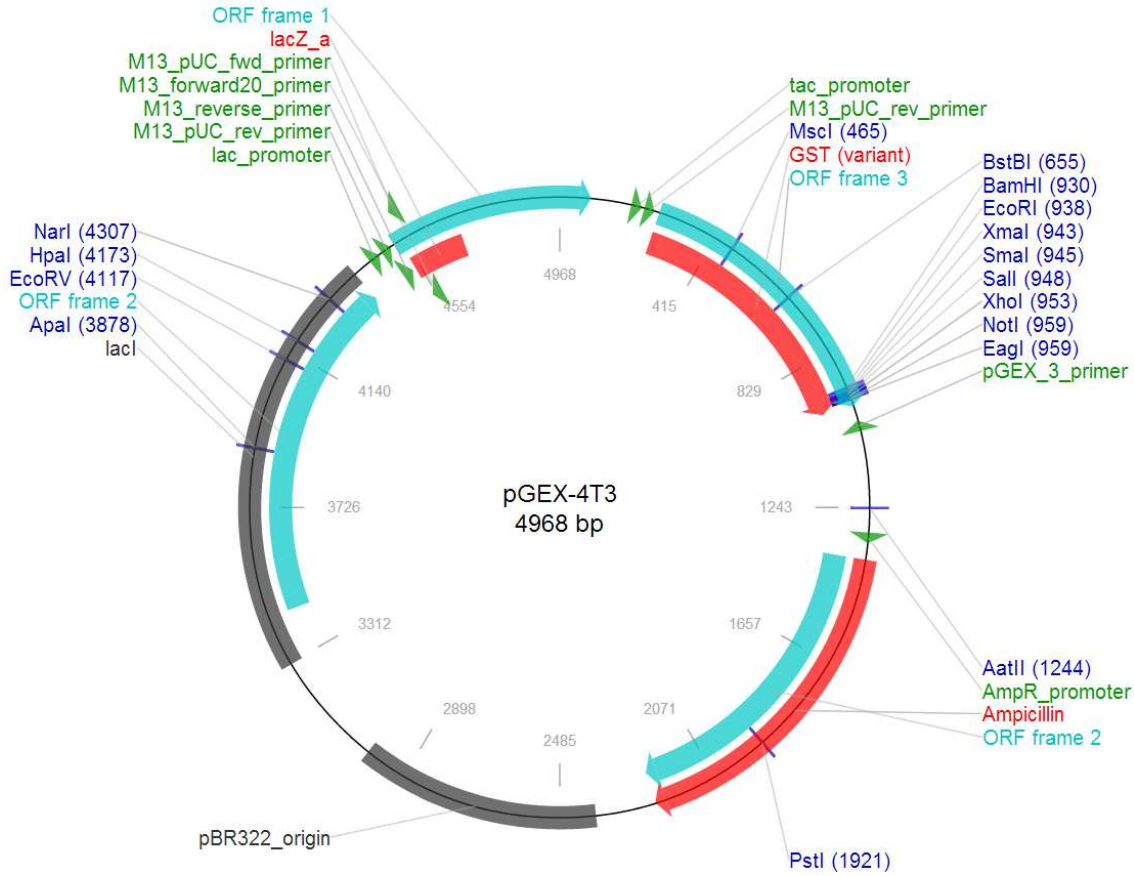


Figure 24: pGEX 4T3 Vector Map[37]

Between the GST tag and the protein of interest a thrombin protease site is present. This allows for the removal of the GST tag during the purification procedure which maybe important for crystallizing FLNAR10 as a single domain rather than a more flexible fusion protein.

Primers were designed to amplify a region of FLNAR10 gene which codes for amino acids 1159-1252. The primers had to have an overhang which created restriction digest sites on the end of the FLNAR10 amplicon.

```
>HgaICviKI-1
Hpy99IUnbI
>MmeIFmuI
Hpy188IHaeIII
DpnISau96I
BstKTIBmgT120I
MboI>BseMII
ChaI<FokIDdeI
Asi256I<StsI>BspCNI
BamHI<BtsCI<Bpu10I
NlaIV>SfaNIBsp1286I
BstYI>BscAIBsiHKAI
    || | ||  ||            | ||   |
  aaggatccGACGCATCCAAAGTCAAGTGCTCAGGCCC  < 37
           10        20        30

Tm= 85.6°C
```

Figure 25: Restriction map of FLNAR10F3 (5` to 3`) Showing restriction enzymes cutting maximum 1 time provided by Serial Cloner 1.3-11

```
>HgaI
Hpy99I
Hpy188I
>MmeI
ChaI
BstKTIMslI
MboI<StsI
DpnI<FokI
Asi256I<BtsCI>BspCNI
BamHI>SfaNI>BseMII
NlaIV>AlfIDdeI
BstYI>BscAIBlpI
    || | ||  ||        ||
  aaggatccGACGCATCCAAATGCTCAGC  < 28
           10        20

Tm= 65.7°C
```

Figure 26: Restriction map of FLNAR10F2 (5` to 3`) Showing restriction enzymes cutting maximum 1 time provided by Serial Cloner 1.3-11

```
Cac8I
AluI
>BbvI
TseI
EsaBC3IFnu4HI
TaqIBisI
Hpy166IIBlsI
SalIBthCI
HincIIHpyCH4V
AccIPstI
Hpy99I>AarICviKI-1
<TspGWINlaIVSfcI
     | |||           |    |  ||| ||
  caatccgtcgacttaCGCAGGTTCCACCTGCAGCTTGCT  < 39
           10        20        30

Tm= 69.8°C
```

Figure 27: Restriction map of FLNAR10B2 (5` to 3`) Showing restriction enzymes cutting maximum 1 time provided by Serial Cloner 1.3-11

As shown in Figure 25and Figure 26 the forward primers produce an overhang (represented by the lower case lettering) which provides a restriction digest site (BamHI) compatible with the pGEX 4T3 plasmid, the back primer produces a restriction digest site (SalI) compatible with the pGEX 4T3 plasmid.

The primers FLNAR10F2 and FLNAR10B2 were used to amplify repeat 10 of the human FLNA cDNA in the pREP4: FLNA plasmid. PCR was run according to Table 5 with annealing temperature set to 65°C.

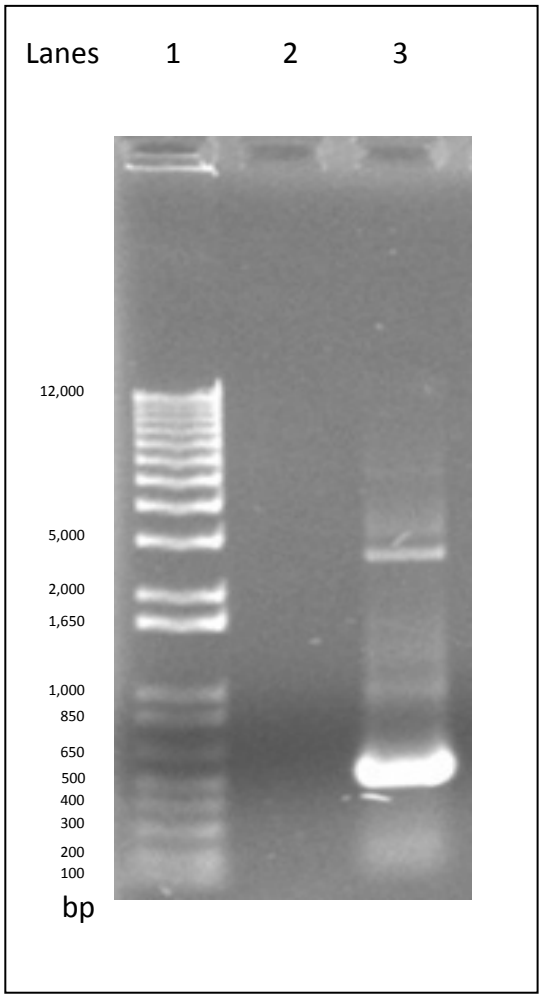


Figure 28: Amplification products of pREP4: FLNA using repeat 10 domain primers

2µL of pREP4: FLNA (2 ng/µL) was amplified using PCR. 5µL of FLNAR10F2 (10 pmol/µL) was used as the forward primer, FLNAR10B2 was the reverse primer. PCR products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualized by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes 1 5µL 1kb plus (Invitrogen) DNA ladder
2 3µL pREP4: FLNA amplification
3 3µL Positive control for polymerase (pGEMT: HCFM2 template, HCFM2SF1 forward primer, HCFM2B1 reverse primer)

As the band representing FLNAR10 is missing from Figure 28 lane 2, the secondary forward primer (FLNAR10F3) was used in a gradient PCR.

In gradient PCR a midpoint temperature is selected which is the temperature of the heating block in the middle during the annealing step. To the left hand side the annealing temperature increases by 5°C, on the right it lowers by 5°C. By varying the annealing temperature a range of selectivity and amplification will increase the odds of producing a pure highly amplified product.

Gradient PCR was performed with only the midpoint of 55°C and the two end points filled. PCR was otherwise loaded and run according to 2.2.2.1 PCR amplification and Table 5: PCR run cycle* respectively.

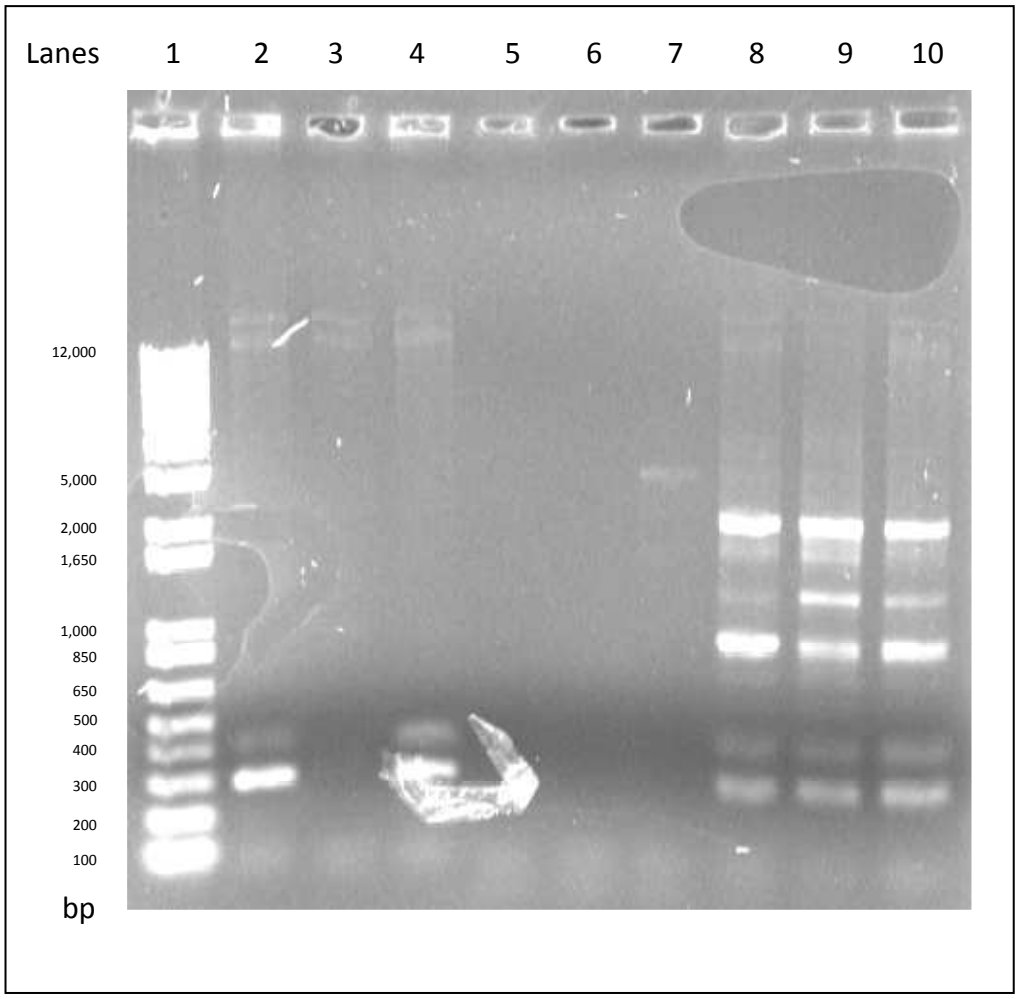


Figure 29: Gradient PCR of pREP4: FLNA plasmid clones

Gradient PCR was performed on pREP4: FLNA plasmid samples #1 through #3. Each plasmid was run according to 2.2.2.1 PCR amplification and Table 5: PCR run cycle* with an annealing temperature of 49°C, 55°C and 59°C. PCR products were run on 1% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualized by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**3µL FLNA#3 59°C annealing temperature**
	3	**3µL FLNA#3 55°C annealing temperature**
	4	**3µL FLNA#3 49°C annealing temperature**
	5	**3µL FLNA#2 59°C annealing temperature**
	6	**3µL FLNA#2 55°C annealing temperature**
	7	**3µL FLNA#2 49°C annealing temperature**
	8	**3µL FLNA#1 59°C annealing temperature**
	9	**3µL FLNA#1 55°C annealing temperature**
	10	**3µL FLNA#1 49°C annealing temperature**

Table 14: Alignment of FLNAR10fragment with FLNAR10F3 and FLNAR10B2

```
Seq_1  1     --------GACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTG  52
                     ||||||||||||||||||||||||||||
Seq_2  1     aaggatccGACGCATCCAAAGTCAAGTGCTCAGGCCC----------------------  37

Seq_1  53    GGGAGGTGGGCCAATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTG  112

Seq_2  38    ------------------------------------------------------------  37

Seq_1  113   AGATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCA  172

Seq_2  38    ------------------------------------------------------------  37

Seq_1  173   CGCACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACG  232

Seq_2  38    ------------------------------------------------------------  37

Seq_1  233   GCGGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAGGTGGAACCTGCG  282
                                       ||||||||||||||||||||||
Seq_2  38    -------------------------AGCAAGCTGCAGGTGGAACCTGCGtaagtcgacggattg
```

As shown in Figure 29 plasmids pREP4: FLNA #3 and pREP4: FLNA #1 were successfully amplified using the forward and reverse primers FLNAR10F3 and FLNAR10B2. According to the predicted primer binding sites the amplicon should be ~300bp in size (see Table 14). The expected fragment is present in lanes 2, 4, 8, 9 and 10. The extra bands which appear are of unknown origin, to reduce the risk of contamination the PCR products from lanes 2 and 4 were used for digest and ligation reactions.

PCR product shown in lane 2 of Figure 29 and pGEX 4T3 was purified and subjected to a double digest reaction. The reaction was set up according to Table 15.

The double digest products from Table 15 were ligated according to Table 16. The ratio of insert (FLNAR10) to destination plasmid (pGEX 4T3) was varied in order to spread the chances of getting the correct number of inserts in the vector (1:1).

After ligation 10 µL of the ligation reactions were used to transform 30 µL of competent cells (Top10) which were plated onto LB agar plates under AMP selection. The transformation protocol used was 2.2.1.3 Competent Cell Transformation

Table 15: Double digest of pGEX 4T3 and PCR product FLNAR10

	PCR product (of pREP4:FLNA and FLNAF3 and FLNAB2)	pGEX 4T3 plasmid
Buffer H (10x)(µL)	3	3
DNA(µL) (2 ng/µL)	8 FLNAR10 PCR product	8 pGEX 4T3 plasmid
BamHI(µL) (10U/µl)	1	1
SalI(µL) (10U/µl)	1	1
H_20(µL)	17	17

Tubes were incubated for three hours at 37°C.

Table 16: Ligation of Double digested pGEX 4T3 vector and FLNAR10 insert

Reaction Tube	**1**	**2**	**3**	**4**
FLNAR10 ddigest (µL)	5	3	1	0.5
pGEX 4T3 ddigest (µL)	5	7	9	9.5
10x Ligation Buffer (µL)	6	6	6	6
T4 DNA ligase (µL)	1	1	1	1
H_2O (µL)	13	13	13	13

Tubes were incubated for three hours at 37°C.

Colonies of Top10 transformants were picked from ~150 colonies and used in colony PCR to find one with the correct number of inserts. While the presence of the stop codon at the end of the insert (put in by the primer overhang) means that extra insert copies in the same plasmid would not alter the fusion protein of GST-FLNAR10, by selecting a colony with the correct sized insert reduces the chance a contaminant piece of DNA or large scale mutation has occurred.

Eight colonies were picked and colony PCR was performed on each colony twice.PCR using vector primers were used to measure the total size of the insert in the vector. PCR using vector 5` primer and insert 3` primer (FLNAR10B2) were used to count the number of FLNAR10 inserts present. PCR was run according to 2.2.2.1 PCR amplification using Taq polymerase. The cycle on the Thermocycler was run according to Table 5: PCR run cycle* with the annealing temperature set to 64°C.

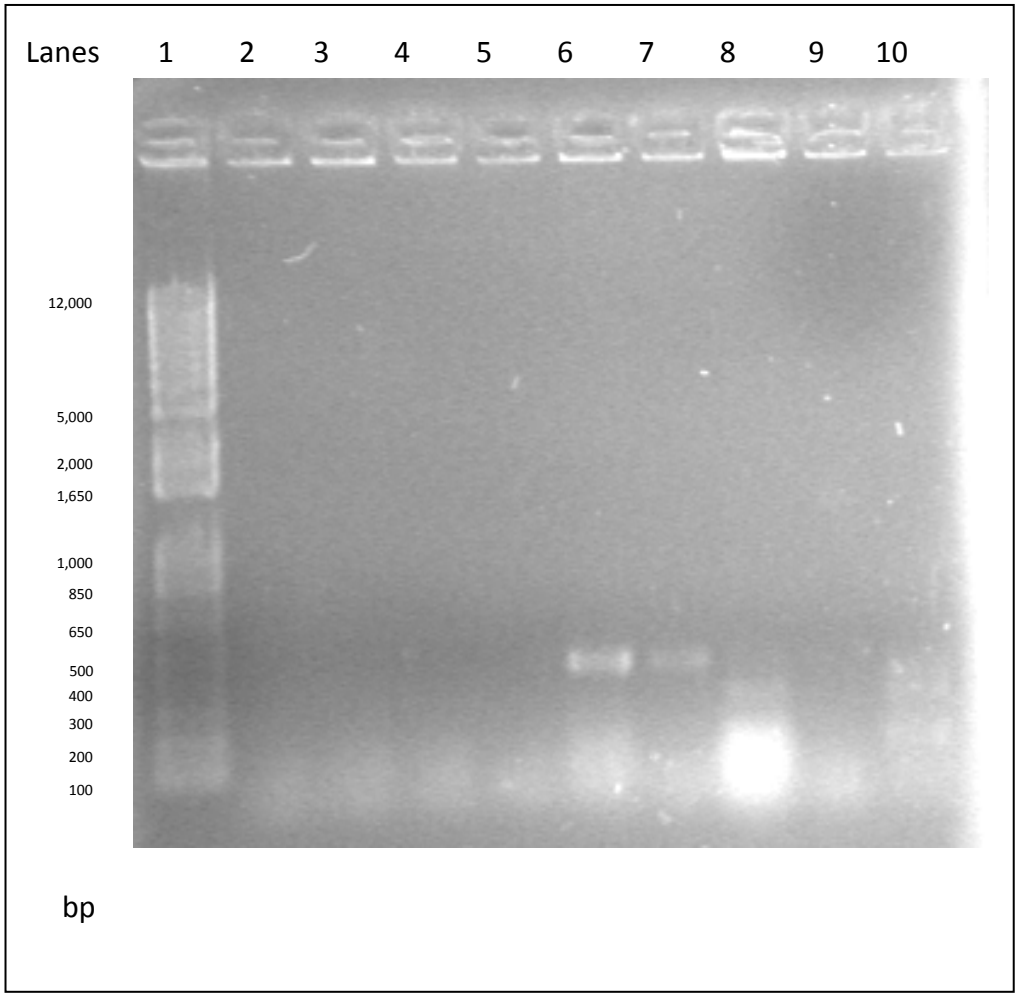


Figure 30: Colony PCR of pGEX 4T3 + FLNAR10 ligate transformants using vector primer

Colonies were picked from agar plated with Top10 *E.coli* transformed with ligated pGEX 4T3 + FLNAR10. Colony picks were amplified using PCR to determine how many copies of the insert were present in the plasmids. A positive control was included; pGEX 4T3 plasmid was used as the template DNA. The primers used were pGEX 5` and pGEX 3` as the forward and reverse primers respectively, these were specific to the vector. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL Colony 1**
	3	**5µL Colony 2**
	4	**5µL Colony 3**
	5	**5µL Colony 4**
	6	**5µL Colony 5**
	7	**5µL Colony 6**
	8	**5µL Colony 7**
	9	**5µL Colony 8**
	10	**5µL pGEX 4T3 plasmid (positive control)**

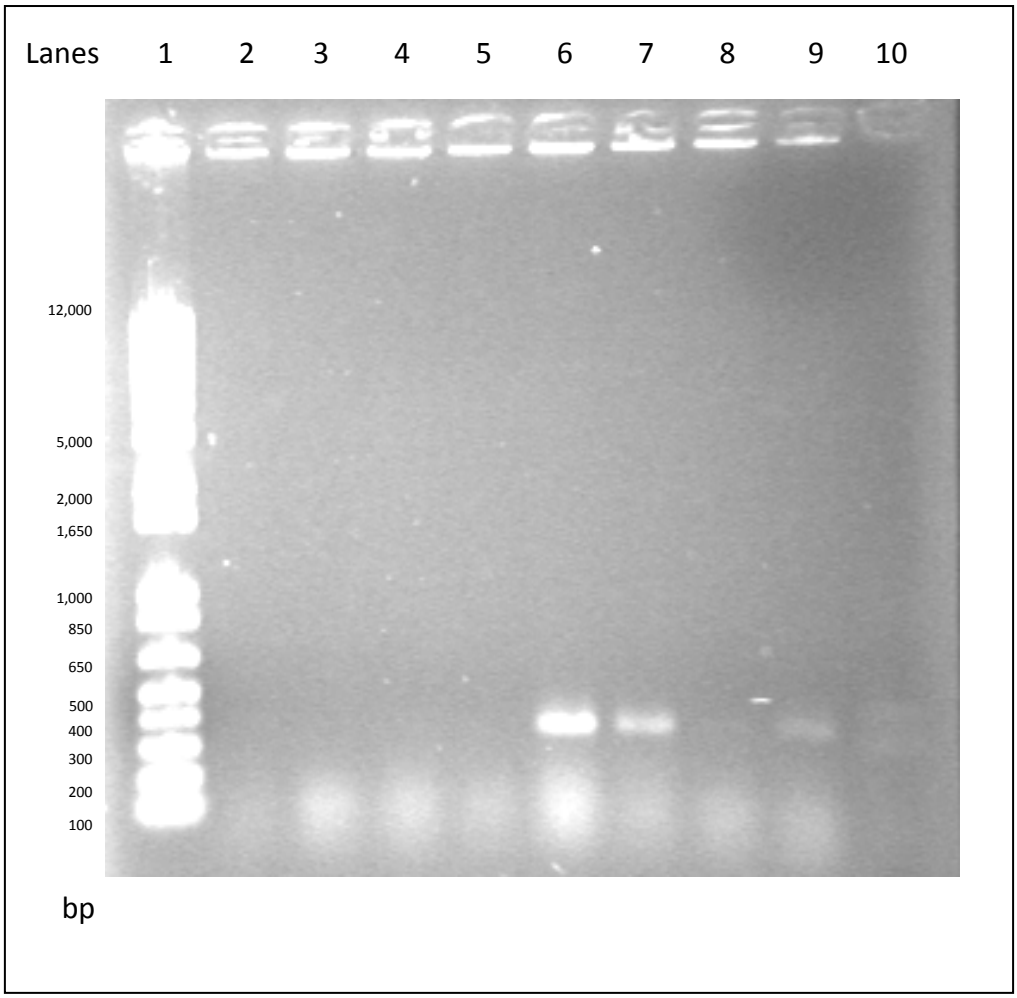


Figure 31: Colony PCR of pGEX 4T3 + FLNAR10 ligate transformants using Vector and Insert primers

Colonies were picked from agar plated with Top10 *E.coli* transformed with ligated pGEX 4T3 + FLNAR10. Colony picks were amplified using PCR to determine how many copies of the insert were present in the plasmids. A negative control pGEX 4T3 plasmid was used as template DNA. The primers used were pGEX 5` and FLNAR10B2 as the forward and reverse primers respectively, these were specific to the vector 5` of the MCS and the 3` end of the insert. PCR products were run on 1.25% agarose in 1x TAE buffer at 100V for 40 minutes. DNA was visualised by incorporating ethidium bromide into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	**1**	**5µL 1kb plus (Invitrogen) DNA ladder**
	2	**5µL Colony 1**
	3	**5µL Colony 2**
	4	**5µL Colony 3**
	5	**5µL Colony 4**
	6	**5µL Colony 5**
	7	**5µL Colony 6**
	8	**5µL Colony 7**
	9	**5µL Colony 8**
	10	**5µL pGEX 4T3 plasmid (negative control)**

Table 17: Alignment of pGEX 4T3:FLNAR10 (Sequence 1) with pGEX 5` and pGEX 3` primers (Sequence 2)

```
Seq_1  1     gggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggatctggttccgcg  60
             ||||||||||||||||||||||
Seq_2  1     gggctggcaagccacgtttggtg-----------------------------------  23

Seq_1  61    tggatccGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGG  120

Seq_2  24    ------------------------------------------------------------  23

Seq_1  121   GGAGGTGGGCCAATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTGA  180

Seq_2  24    ------------------------------------------------------------  23

Seq_1  181   GATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCAC  240

Seq_2  24    ------------------------------------------------------------  23

Seq_1  241   GCACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGG  300

Seq_2  24    ------------------------------------------------------------  23

Seq_1  301   CGGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAgctggaacctgcgtaagtcgactc  360

Seq_2  24    ------------------------------------------------------------  23

Seq_1  361   gagcggccgcatcgtgactgactgacgatctgcctcgcgcgtttcggtgatgacggtgaa  420

Seq_2  24    ------------------------------------------------------------  23

Seq_1  421   aacctctgacacatgcagctcccgg  445
               |||||||||||||||||||||||
Seq_2  23    --cctctgacacatgcagctcccgg  1
```

Table 18: Alignment of pGEX 4T3:FLNAR10 (Sequence 1) with pGEX 5` and FLNAR10B2 primers (Sequence 2)

```
Seq_1  1     -ggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggatctggttccgcg  59
              ||||||||||||||||||||
Seq_2  1     gggctggcaagccacgtttggtg-----------------------------------  23

Seq_1  60    tggatccGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGG  119

Seq_2  24    ------------------------------------------------------------  23

Seq_1  120   GGAGGTGGGCCAATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTGA  179

Seq_2  24    ------------------------------------------------------------  23

Seq_1  180   GATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCAC  239

Seq_2  24    ------------------------------------------------------------  23

Seq_1  240   GCACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGG  299

Seq_2  24    ------------------------------------------------------------  23

Seq_1  300   CGGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAgctggaacctgcgtaagtcgac---  357
                                  |||||||||||| ||||||||||||||||||||
Seq_2  39    -------------------------AGCAAGCTGCAGGTGGAACCTGCGTAAGTCGACGGA  4

Seq_1  358   ---  357

Seq_2  3     TTG  1
```

The predicted PCR product of the FLNAR10 insert in the pGEX 4T3 vector using the vector primers pGEX 5` and pGEX 3`is shown in Table 17. The predicted size of the PCR product is approximately 450 bp. The predicted PCR product of the FLNAR10 insert in the pGEX 4T3 vector using the vector primer pGEX 5` and insert primer FLNAR10B2 is shown in Table 18. The predicted size of the PCR product is approximately 350 bp. Colony PCR from colony picks number 5 & 6 have the predicted sized band in Figure 30and Figure 31 lanes 6 and 7.

As both colonies 5 & 6 had only a single band in Figure 30 and Figure 31 and the bands were of the expected size, both were grown as a plasmid stock according to 2.2.6 Plasmid Isolation in the 2.2 Methods section.

Plasmid samples of both were sent for sequencing according to 2.2.7 DNA sequencing (chromatograms in Table 24); both were aligned according to the predicted ligation products between pGEX 4t3 and FLNAR10. Colony #5 and #6 were found to have the correct insert in the correct frame (Table 25 &Table 26). While the pGEX 5` primer was used during sequencing (meaning only the MCS and insert would be sequenced,) the GST region was not sequenced. As the only relevant feature of the GST region is its ability to assist in the purification of the fusion protein, this will be tested by attempting purification using GST binding beads.

Having successfully subcloned FLNAR10 into pGEX 4T3, plasmids from Colony #5 and #6 were used for the expression and purification trials.

3.1.5 Expression of FLNAR10

To ensure the purification of GST-FLNAR10 fusion protein under the most favourable conditions, the expression variables were tested to find those that produced the largest total mass and proportional concentration of soluble fusion protein.

A time expression trial for protein expression was performed. The pGEX 4T3:FLNAR10 plasmids from colonies #5 and #6 were used to transform BL21 *E.coli* according to 2.2.1.3 Competent Cell Transformation. Colonies from transformation were used to inoculate 10 mL of sterile LB broth under selection using AMP at 0.1 mg/mL. Inoculated tubes were grown overnight at 37°C to approximately 0.6 ODA_{600}. IPTG was added to each 10 mL broth up to 0.1 mM and each was incubated at 37°C. 1.5 mL of each broth was taken at time intervals post IPTG induction.

Each time interval sample was pelleted for -20°C storage until the final overnight samples were collected. Pellets were re-suspended in 200 µL of PBS pH 7.3 and sonicated for 5 seconds at 20%.

Whole cell lysate and supernatant only fractions were loaded onto a 15% SDS-PAGE. Gel was run at 200V for 40 minutes. Gel was stained with Coomassie for 20 minutes and destained overnight.

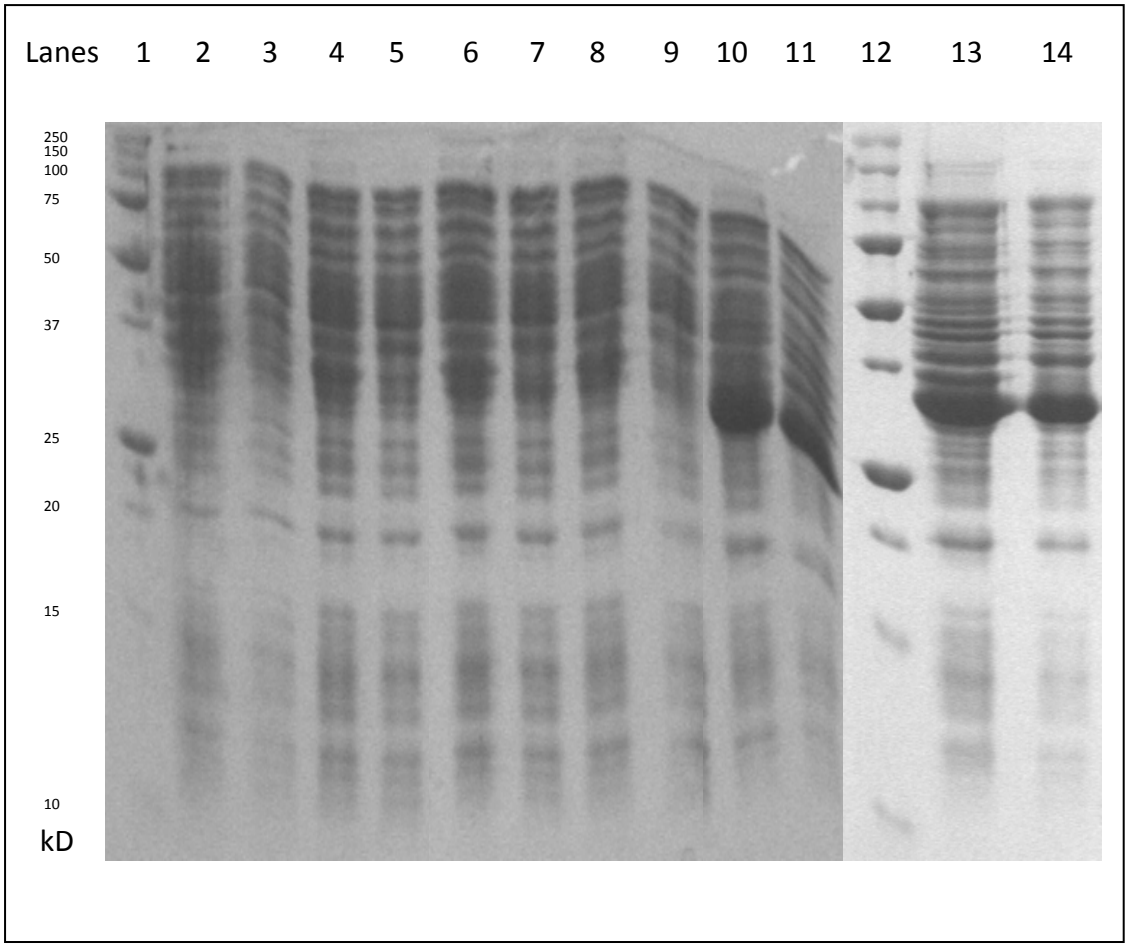


Figure 32: Time expression trial of GST-FLNAR10 fusion protein

BL21 *E.coli* were transformed with pGEX 4T3:FLNAR10 from colony #5 and #6, colony picks were used to inoculate 10 mL of sterile broth. The colony broth was grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM). Samples (1.5 mL) were taken at T=0, 45 minutes and an overnight sample. Cells were lysed using sonication for 5 seconds at 20% in 1.5 mL of PBS pH 7.3. 200 µL samples of whole cell lysate were collected; the rest was separated into pellet and supernatant using centrifugation. Supernatant and whole cell lysate samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa. Lanes not related to this experiment were removed, a second gel was used for the overnight expression of pGEX 4T3:FLNAR10 #6 (lane 12-14).

Lanes	**1**	**5µL PPU**
	2	**5µL #5 Whole cell T=0**
	3	**5µL #5 Supernatant T=0**
	4	**5µL #6 Whole cell T=0**
	5	**5µL #6 Supernatant T=0**
	6	**5µL #5 Whole cell T=45minutes**
	7	**5µL #5 Supernatant T=45minutes**
	8	**5µL #6 Whole cell T=45minutes**
	9	**5µL #6 Supernatant T=45minutes**
	10	**5µL #5 Whole cell T=Overnight**
	11	**5µL #5 Supernatant T=Overnight**
	12	**5µL PPU**
	13	**5µL #6 Whole cell T=Overnight**
	14	**5µL #6 Supernatant T=Overnight**

In lanes 11 and 14 of Figure 32, the overnight induction supernatant fractions of GST-FLNAR10 fusion expression show an increase in a protein band between the standard markers 37 kDa and 25 kDa. The fusion protein is predicted to have a size of 36kDa and is shown in Table 19. The empty vector expressing only GST with the thrombin cleavage site should only be 27.75kDa.

As the sequencing confirms the correct insert and frame and the band that appears during the time expression trial is very near the appropriate size, this was taken as confirmation that the correct fusion product was being formed.

To find out if a lower incubation temperature during induction produces a higher mass ratio of fusion protein to *E.coli* protein a temperature expression trial was setup. pGEX 4T3:FLNAR10#6 transformed *E.coli* inoculated 4 10 mL sterile LB broth, selection was maintained using AMP at 0.1 mg/mL. Broth was grown to ~0.6OD. IPTG was added to each 10 mL broth up to 0.1 mM, tubes were incubated at 22, 25, 30 and 37°C respectively.

1.5 mL of each broth was taken after overnight incubation, post IPTG induction.

1.5 mL samples were pelleted then re-suspended in 200 µL of PBS pH 7.3 and sonicated for 5 seconds at 20%.

Whole cell lysate and supernatant only fractions were loaded onto a 15% SDS-PAGE. Gel was run at 200V for 40 minutes. Gel was stained with Coomassie for 20 minutes and destained overnight.

Table 19: GST-FLNAR10 predicted fusion protein sequence

MW: 35.99kDa MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRY IADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGD HVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHP PKSDLVPRGSDASKVKCSGPGLERATAGEVGQFQVDCSSAGSAELTIEICSEAGLPAEVYIQDHGDGTHTITYI PLCPGAYTVTIKYGGQPVPNFPSKLQLEPA*

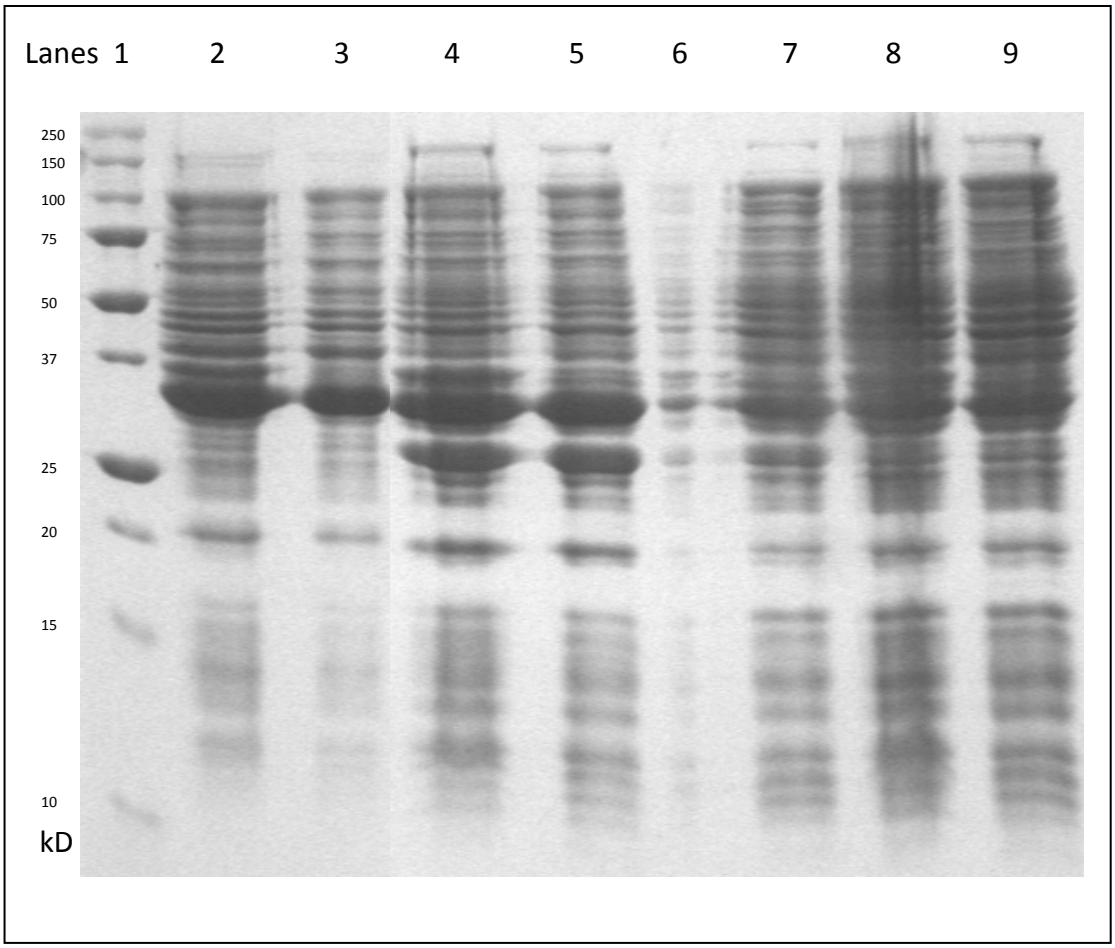


Figure 33: Temperature expression trial of GST-FLNAR10 fusion protein

Colony picks from pGEX 4T3:FLNAR10#6 transformed BL21 *E.coli* were used to inoculate 4 sets of 10 mL of sterile LB broth. The inoculated broths were grown to ~0.6 OD at 37°C. Expression was induced with IPTG (final concentration 0.1 mM). Samples (1.5 mL) were taken after overnight expression at 22, 25 30 and 37°C respectively. Pellets re-suspended in 1.5 mL of PBS pH 7.3 were lysed using sonication for 5 seconds at 20%. 200 µL samples of whole cell lysate were collected; the rest was separated into pellet and supernatant using centrifugation. Supernatant and whole cell lysate samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa. Lanes not related to this experiment were removed.

Lanes	**1**	**5µL PPU**
	2	**5µL 37°C Whole cell**
	3	**5µL 37°C Supernatant**
	4	**5µL 22°C Whole cell**
	5	**5µL 22°C Supernatant**
	6	**5µL 25°C Whole cell**
	7	**5µL 25°C Supernatant**
	8	**5µL 30°C Whole cell**
	9	**5µL 30°C Supernatant**

Shown in Figure 33 the alternate temperature to 37°C for expression (30°C, 25°C or 22°C), showed no improvement in fusion protein vs. *E.coli* total protein yield or in protein solubility. There is a band of ~27kDa present in the 22°C trial that is not present in the others, as it doesn't appear to be one of the expected proteins from fusion (except perhaps a GST without FLNAR10) I decided not to pursue that protein bands expression.

Using the expression conditions and lysis conditions from this section large scale expression (500 mL) was started for purification of the GST-FLNAR10 fusion protein.

3.1.6 Purification of FLNAR10

3.1.6.1 Preparation of Lysate

10 mL of 0.6OD A_{600} LB broth inoculated with BL21 *E.coli* containing pGEX 4T3-FLNAR10 was used to inoculate 500 mL of sterile LB broth. Selection for the plasmid was maintained using AMP at 0.1 mg/mL.

The inoculated 500 mL broth was grown overnight at 37°C until it reached ~0.6OD. Expression of the fusion protein GST-FLNAR10 was then induced using IPTG at 0.1 mM. Expression was carried out overnight at 37°C.

The 500 mL broth containing the expressed fusion protein GST-FLNAR10 was pelleted. The pellet was re-suspended in 35 mL PBS pH 7.3. The cells were lysed using a French press twice followed by sonication for 15 seconds at 30% twice. The sonicated lysate was spun-down in a centrifuge at 30000 x g for 50 minutes at 4°C to remove cell debris. The supernatant was filtered using a 450nm syringe filter to remove any further cell debris.

3.1.6.2 GST trap purification of FLNAR10

A 1 mL high pressure GST column from GE Healthcare was equibrilated with 8 mL PBS pH 7.3 at 1 mL/minute.

35 mL of filtered supernatant from lysed *E.coli* was passed over the column at 1 mL/minute. A sample of lysate was held back for analysis. The flow-through was collected for analysis.

The column was washed with 20x the column volume (20 mL) of PBS pH 7.3 at 1 mL/minute to remove any proteins not bound to the column. A sample of the wash step flow-through was collected for analysis.

The column was eluted using 2x the column volume (2 mL) PBS pH 7.3 containing 10 mM reduced glutathione at 0.5 mL/minute. The eluate was collected and stored at 4°C.

A sample of the eluate was put aside for analysis. The rest of the eluate was concentrated using a 10kDa cut off Vivaspin Centrifugal Concentrator.

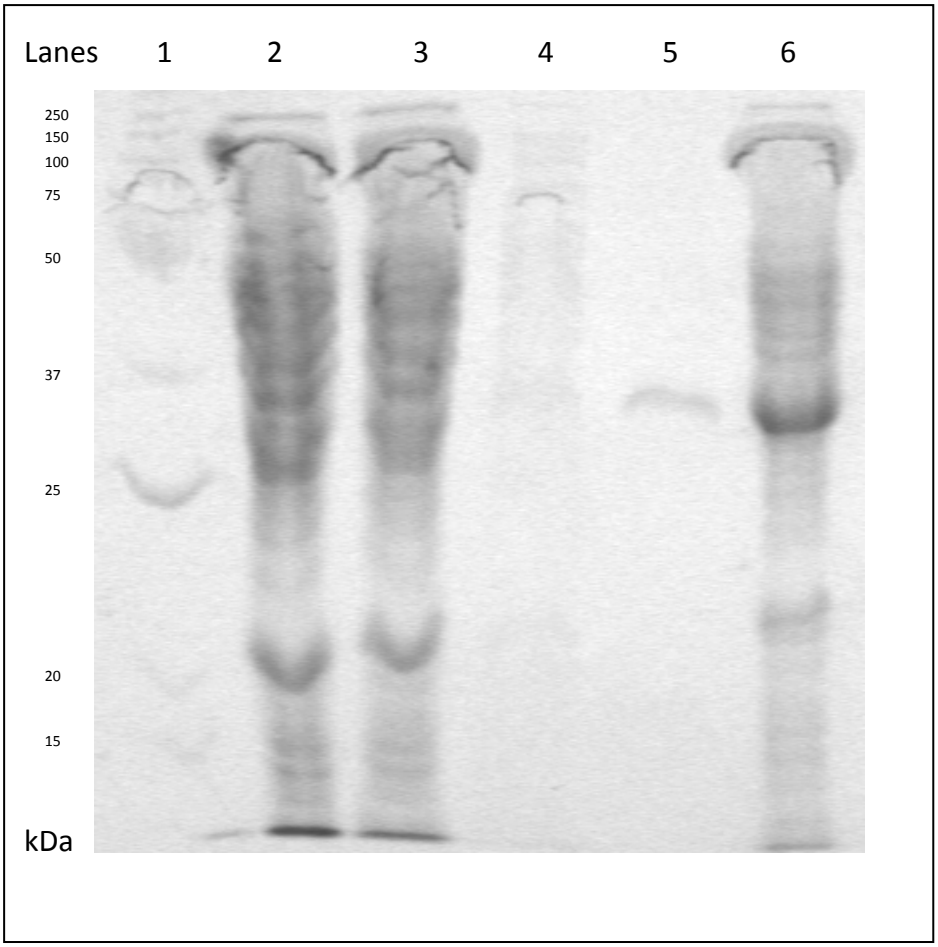


Figure 34: GST column purification of GST-FLNAR10 fusion protein from BL21 *E.coli* lysate

Lysate from BL21 *E.coli* expressing GST-FLNAR10 was passed over a GST column. Column was eluted using 10 mM reduced glutathione in PBS pH 7.3 Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**5µL Lysate supernatant**
	3	**5µL Flow-through**
	4	**5µL Wash**
	5	**5µL Eluate**
	6	**5µL Concentrated Eluate**

After concentration (see lane 6, Figure 34) it was apparent that the sealed column (opposed to a self packed column) may be retaining *E.coli* proteins after the wash step. Purification attempts after this one were done using a 5 mL self packed column containing 1 mL of 50% glutathione bound agarose resin.

3.1.6.3 Thrombin digest of GST-FLNAR10

The thrombin protease site between GST and FLNAR10 allows for cleavage of the fusion protein by thrombin. Digestion of the fusion protein may be preferential when attempting crystallisation.

Digestion while the fusion protein is bound to the glutathione-sepharose bead affinity column would allow separation of the GST and FLNAR10 units. To optimise thrombin cleavage digestion trials were setup with glutathione-sepharose bead affinity column purified GST-FLNAR10 fusion protein.

10 units of thrombin were added to 3 mL of eluted fusion protein. This was split into three tubes and incubated at different temperatures (4, 21 and 37°C). Samples of each reaction were taken at time intervals of 40 minutes, 120 minutes, 24 hours and 3 days.

Samples were analysed using 15% SDS-PAGE gel.

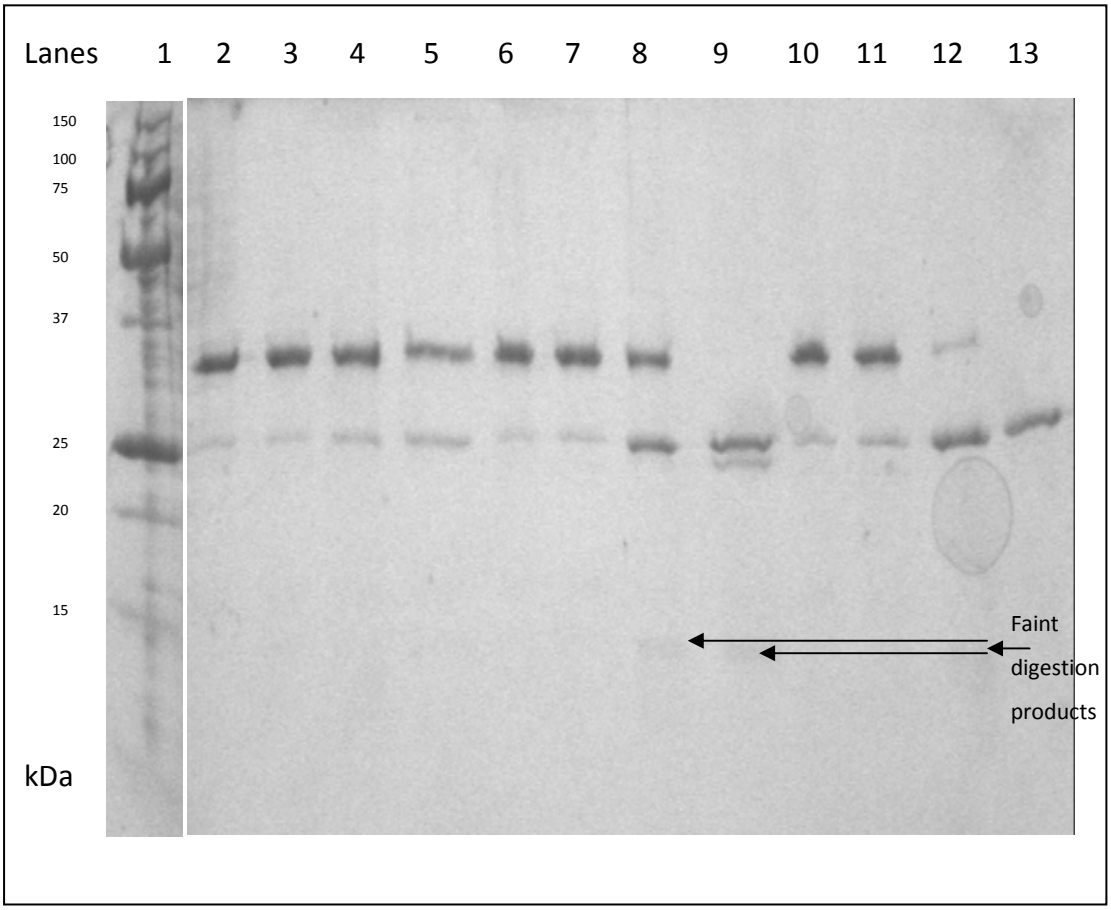


Figure 35: Thrombin digest of GST-FLNAR10 fusion protein from BL21 *E.coli* lysate

Eluate of GST-FLNAR10 from a GST column was digested using thrombin. Digestion took place in three reaction tubes at differing temperatures (4, 21 and 37°C). Samples were taken at time intervals (40 min, 120 min, 24 hours and 3 days) samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa. Lanes not related to this experiment were removed. Areas in lanes 8, 9 and 12 pointed to highlight faint bands <15 kDa in size.

Lanes	**1**	**5µL PPU**
	2	**5µL T=40min 4°C**
	3	**5µL T=120 min 4°C**
	4	**5µL T=24 hours 4°C**
	5	**5µL T=3d 4°C**
	6	**5µL T=40min 21°C**
	7	**5µL T=120 min 21°C**
	8	**5µL T=24 hours 21°C**
	9	**5µL T=3d 21°C**
	10	**5µL T=40min 37°C**
	11	**5µL T=120 min 37°C**
	12	**5µL T=24 hours 37°C**
	13	**5µL T=3d 37°C**

After digestion with thrombin (pointed to in lanes 8, 9 and 12 of Figure 35) a faint band appears below the 15kDa marker. The predicted size of FLNAR10 protein is ~9.7kDa. While the size of the band appears larger than expected, this may be due to inaccuracies in the markers size or a higher positive charge on the protein. Regardless the GST-FLNAR10 fusion shifts from ~34kDa to the expected 25kDa, the faint band that appears should be at most the difference between the GST fusion and GST alone.

There was some secondary digestion of the GST (lane 9 Figure 35). As the secondary digest product only appeared after all of the fusion protein was digested, it suggested that the first products formed from digestion were the GST and FLNAR10 by digestion of the thrombin cleavage site.

To try and get a better gel of the products formed during thrombin digestion a new purification of lysate was set up.

The lysate was prepared as described in 3.1.6.1 Preparation of Lysate and loading onto the column according to 3.1.6.2 GST trap purification of FLNAR10 with the exception that instead of digesting the product after elution, thrombin digestion was done on the column. Bead samples were taken at 60 minutes, 13 hours and 15 hours.

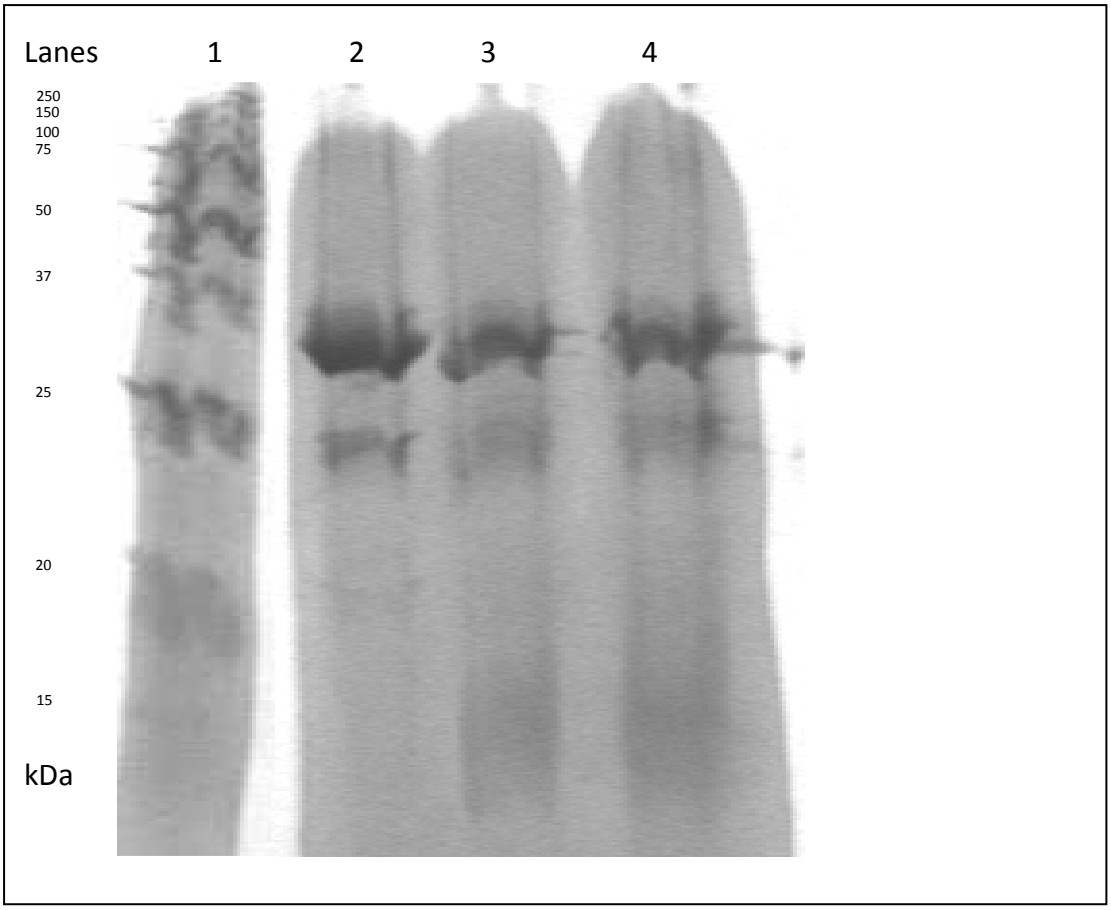


Figure 36: Thrombin digest of GST-FLNAR10 fusion protein on GST column

GST-FLNAR10 on a GST column was digested using 10 units of thrombin. Digestion took at 21°C. Bead containing eluate samples were taken at time intervals (60 mins, 13 hours and 15 hours) samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was silver stained. Molecular sizes of bands are indicated on the left in kDa. Lanes not related to this experiment were removed.

Lanes	**1**	**5µL PPU**
	2	**8µL T=60 mins**
	3	**8µL T=13 hours**
	4	**8µL T=15 hours**

Due to the poor quality of Figure 36 the column beads were sampled. As the GST tag provides the affinity for the bound glutathione, the digested products should be able to be separated. The GST should remain bound to the column and the FLNAR10 should elute using only PBS.

The column was eluted using 2 mL of PBS. The eluate and beads were sampled for SDS-PAGE analysis. The remaining eluate was concentrated using a Vivaspin 5kDa MW cut-off concentrator, this concentrate was also sampled for analysis.

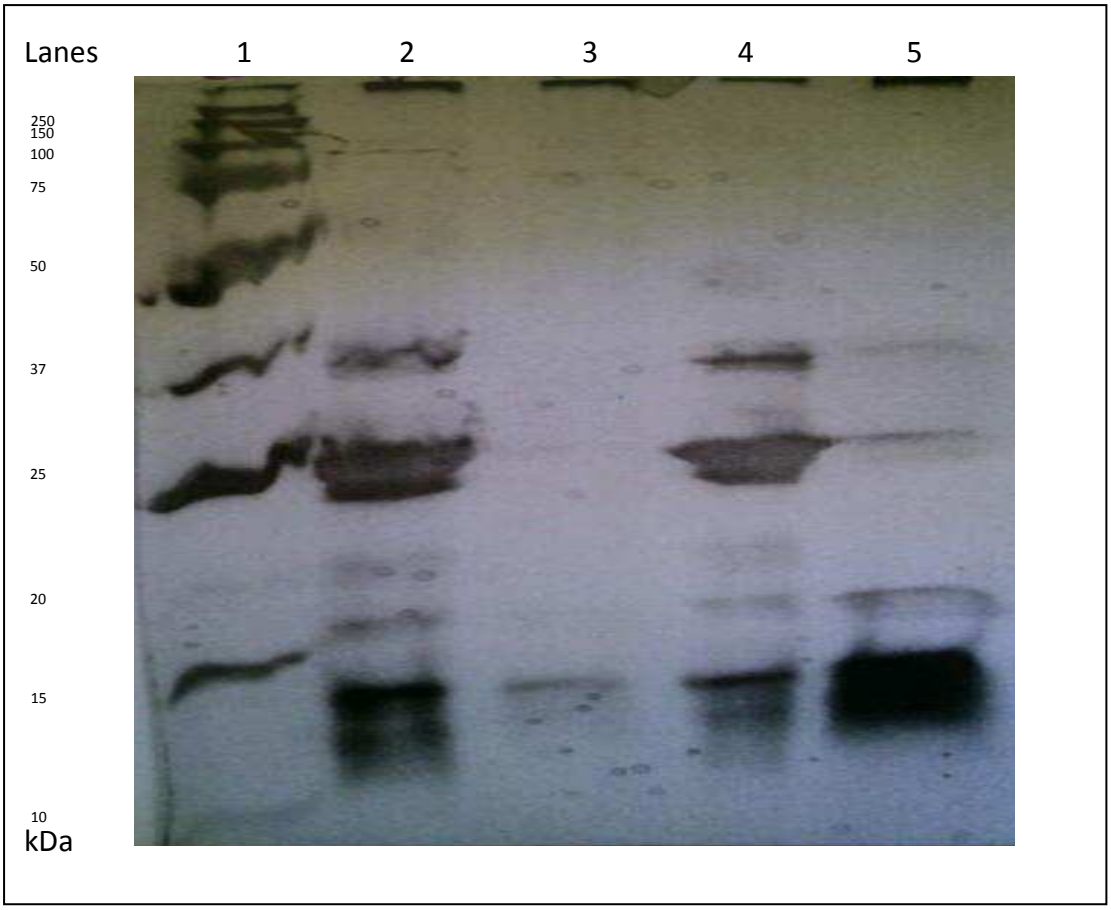


Figure 37: Thrombin digest products of GST-FLNAR10 fusion protein on GST column

GST-FLNAR10 on a GST column was digested using 10 units of thrombin. Digestion took at 21°C. Samples of the digested bound fusion protein, eluate of digested fusion protein, eluted column beads and concentrated eluate were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was silver stained. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**8µL Column beads pre-elution**
	3	**8µL Eluate**
	4	**8µL Column beads post-elution**
	5	**8µL Concentrated Eluate**

3.1.6.3 AKTA purification of GST-FLNAR10 thrombin digest products

The digestion of GST-FLNAR10 created a double band of proteins in the 10-15kDa MW area seen in lanes 2-5 of Figure 37. The concentrated eluate (lane 5) contains a large fraction of both of the proteins.

To purify these two proteins the FPLC AKTA would be loaded with concentrated eluate from a column digest of the fusion protein. A fresh column digest was performed according to 3.1.6.1 Preparation of Lysate and 3.1.6.2 GST trap purification of FLNAR10. Digestion occurred on the column using 10 units of thrombin at 21°C overnight.

The eluate was concentrated down to 400 µL using a Vivaspin20 5kDa MW concentrator. While the column used (Superdex 75 10/300) recommends a sample volume of less than 250 µL, the difference in size between these two bands means separation will not be possible regardless. The rationale behind using the AKTA is that any other proteins of sufficiently different sizes will be purified out.

The s75 10/300 column was run using PBS pH 7.3 at 0.6 mL/min. Fractions of 3 mL were collected. Peak fractions were analysed using SDS-PAGE gel.

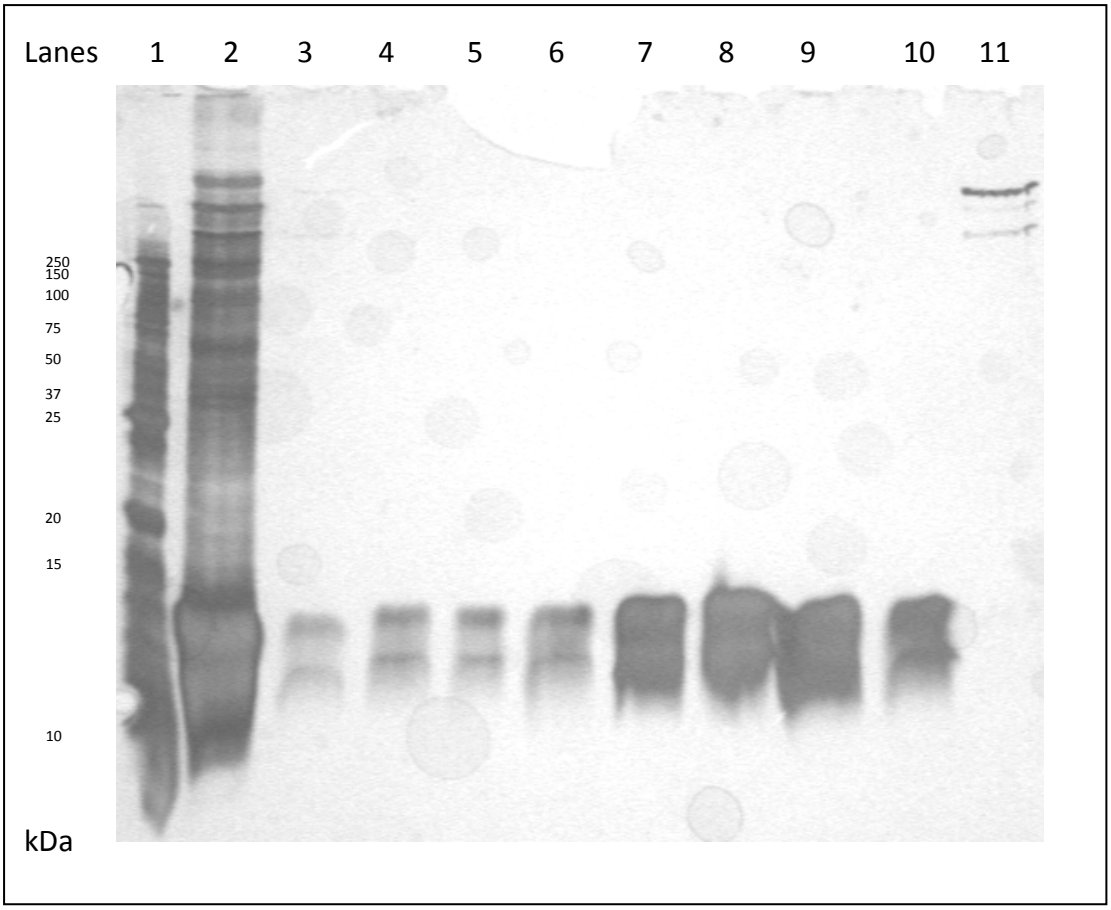


Figure 38: Purification of thrombin digest products using FPLC

Concentrated eluate thrombin digest products of GST-FLNAR10 were passed through a Superdex s75 10/300 FPLC column. Protein peaks from FPLC purification were sampled. Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was silver stained. Molecular sizes of bands are indicated on the left in kD, due to the poor discrimination in the marker bands the positioning of the sizes (indicated on the left) are only estimates.

Lanes	**1**	**5µL PPU**
	2	**4µL Concentrated digestion eluate**
	3	**4µL Peak Fraction 1**
	4	**4µL Peak Fraction 2**
	5	**4µL Peak Fraction 3**
	6	**4µL Peak Fraction 4**
	7	**4µL Peak Fraction 5**
	8	**4µL Peak Fraction 6**
	9	**4µL Peak Fraction 7**
	10	**4µL Peak Fraction 8**
	11	**4µL Peak Fraction 9**

Size exclusion chromatography with concentrated eluent produced a large sharp peak. Analysis using SDS-PAGE in Figure 38 identifies it as the two unknown digest products of the GST-FLNAR10 fusion by thrombin. As the two proteins are too close together to be separated using size exclusion chromatography, thrombin digestion was used to try and determine if one of the bands was a FLNAR10 with the thrombin site undigested from it, or if the smaller band is FLNAR10 that had been digested further.

A 50 µL sample from lane 8 (fraction 6) Figure 38 was mixed with 2 units of thrombin and incubated over 12 hours at 21°C.

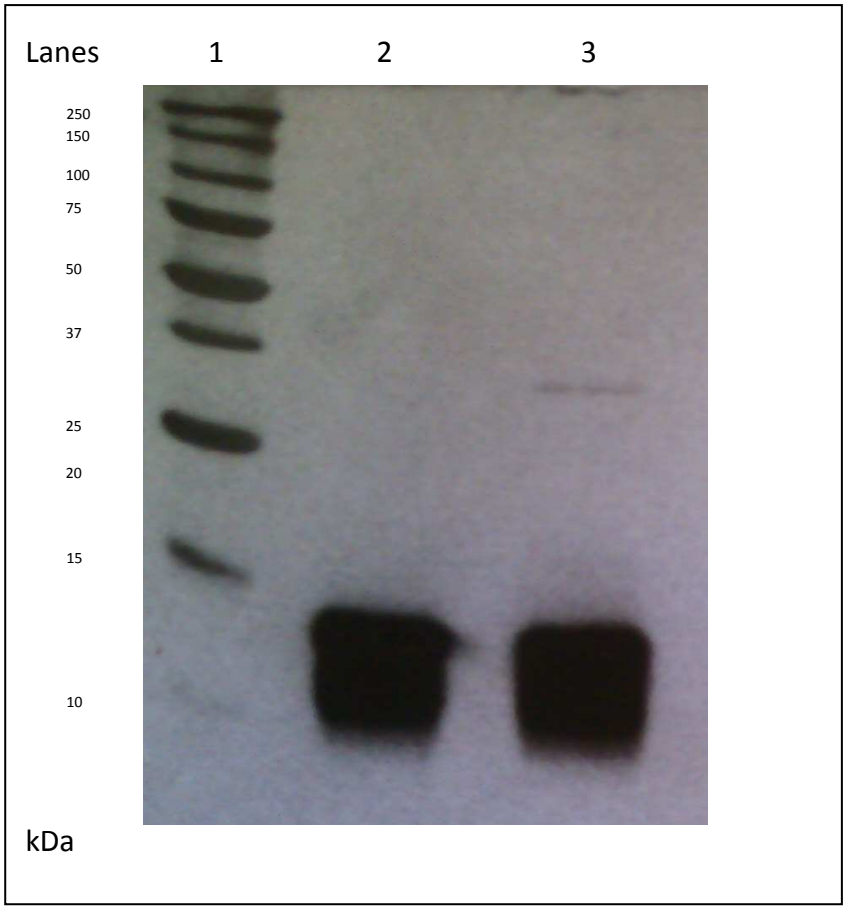


Figure 39: Purification of thrombin digest products using FPLC

Eluate collected from column bound GST-FLNAR10 digested using Thrombin was subject to FPLC purification using a Superdex s75 10/300 FPLC column. Lane 8 (fraction 6) Figure 38 was mixed with thrombin over 12 hours at 21°C to see if one of the bands could be digested to match the other. Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. The gel was run for 40 minutes at 200V and was silver stained. Molecular sizes of bands are indicated on the left in kD; due to the poor discrimination of the 20 kDa marker (indicated on the left) the 20 kDa marker may be placed incorrectly.

Lanes	**1**	**5µL PPU**
	2	**4µL Fraction 6**
	3	**4µL Fraction 6 + thrombin**

No disappearance of the top band was seen, no laddering from slow digestion. At best guess the top of the two bands is FLNAR10, the only thrombin recognition site is between GST and FLNAR10 making it the preferred site to be digested first. Going by the initial size of the top band in lane 2 of Figure 36 relative to the lower band, the top band is the dominant product of the two to form. Unfortunately as there are no antibodies specific to repeat 10 of FLNA the identity of both bands cannot be confirmed. While mass spectrometry would be able to identify each they would first need to be separated.

Even though the preferred state for crystallization is FLNAR10 alone, protein structures have been determined while fused to affinity tags [38]. Due to the difficulties in removing the GST tag without getting secondary products which cannot be removed the GST tag removal was abandoned.

3.1.7 Pull-down experiments with FLNAR10

Having access to *E.coli* expressing CFM2 and C2C12 cells which are a mouse myoblast cell line, GST-FLNAR10 fusion protein and GST protein were purified for use in GST fusion protein probe pull-down experiments.

Even though the agarose-glutathione slurry can bind 5mg of protein per mL and the column used in the purification was 2 mL in volume, the final mass of protein collected was less than 1mg. It may be the fusion protein is falling out of solution whenever the protein concentration gets above 1 mg/mL. This would explain why a 1L preparation of *E.coli* lysate containing GST-FLNAR10 only just produced over the 0.8mg the 500 mL GST prep did. Purified GST and GST-FLNAR10 fusion proteins were of very high purity (see lanes 5 and 10, Figure 40).

Using the purified GST and GST-FLNAR10 fusion protein, pull-down experiments were performed. The experiments were run according to2.2.13 GST fusion protein probe pull-down. Four experiments were setup, C2C12 mouse myoblast cells and *E.coli* BL21 cell expressing CFM2 were used to provide the lysate to be probed, and two different lysate buffers were used on each cell line, PBS and RIPA.

By comparing the supernatant with the washed beads, the proteins bound specifically can be determined. By comparing the GST control probe with GST-FLNAR10 fusion probe beads, proteins that bind the GST-FLNAR10 fusion from the C2C12 mouse myoblast cells indicates proteins that may be functionally linked. Proteins that bind the GST-FLNAR10 fusion from the *E.coli* BL21 cell expressing CFM2 lysate would likely be the same size as the CFM2 protein as only the CFM2 protein should bind to FLNAR10.

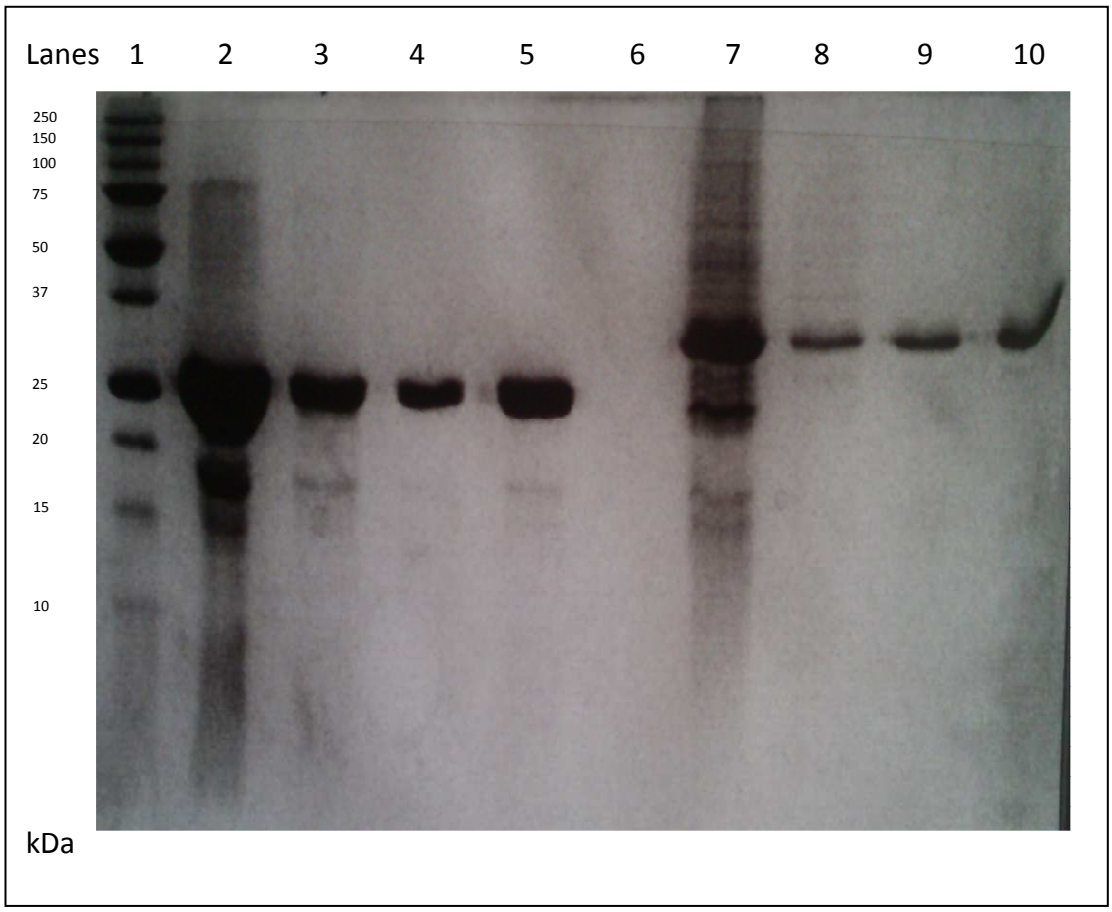


Figure 40: GST affinity purification of GST-FLNAR10 and GST

BL21 *E.coli* transformed with pGEX 4T3 or pGEX 4T3:FLNAR10 were grown overnight in 1L broth, then induced to express using IPTG (0.1 mM) overnight at 37°C. Cells were lysated and purified according to 3.1.6.1 Preparation of Lysate and 3.1.6.2 GST trap purification of FLNAR10. Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. Gel was run for 40minutes at 200V. Gel was stained with Coomassie for 20 minutes and destained overnight. Molecular sizes of bands are indicated on the left in kDa.

Lanes	**1**	**5µL PPU**
	2	**5µL GST (on beads)**
	3	**5µL GST column eluate**
	4	**5µL GST eluate**
	5	**5µL GST eluate concentrate**
	6	**-**
	7	**5µL GST-FLNAR10 (on beads)**
	8	**5µL GST-FLNAR10 column eluate**
	9	**5µL GST-FLNAR10 eluate**
	10	**5µL GST-FLNAR10 eluate concentrate**

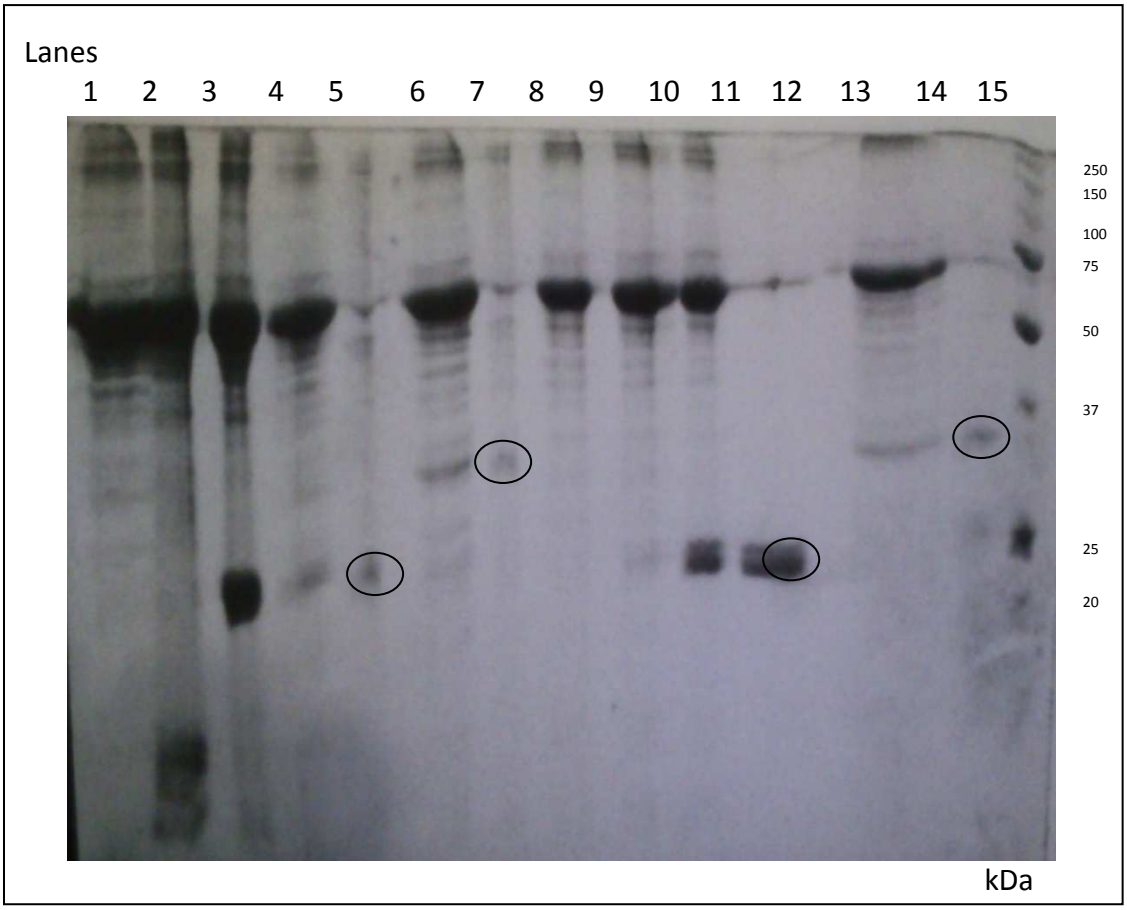


Figure 41: Fusion protein probe pull-down of C2C12 mouse myoblast cells

C2C12 mouse myoblast cells were lysed according to 2.2.13 GST fusion protein probe pull-down using PBS buffer and RIPA buffer. The lysate was pre-cleared using GST bound to glutathione cross-linked to agarose. Cleared lysate was probed using the negative control GST and the test fusion protein probe GST-FLNAR10. Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. Gel was run for 40 minutes at 200V. Gel was stained with Coomassie blue. Molecular sizes of bands are indicated on the right in kDa. Lanes 1-7 were done in PBS, lanes 8-14 were done in RIPA buffer.

Lanes	**1**	**5µL Supernatant (PBS buffer)**
	2	**5µL Pellet (PBS buffer)**
	3	**5µL Lysate depletion step beads (PBS buffer)**
	4	**5µL GST control eluent (PBS buffer)**
	5	**5µL GST control beads (PBS buffer)**
	6	**5µL Fusion eluent (PBS buffer)**
	7	**5µL Fusion beads (PBS buffer)**
	8	**5µL Supernatant (RIPA buffer)**
	9	**5µL Pellet (RIPA buffer)**
	10	**5µL Lysate depletion step beads (RIPA buffer)**
	11	**5µL GST control eluent (RIPA buffer)**
	12	**5µL GST control beads (RIPA buffer)**
	13	**5µL Fusion eluent (RIPA buffer)**
	14	**5µL Fusion beads (RIPA buffer)**
	15	**5µL PPU**

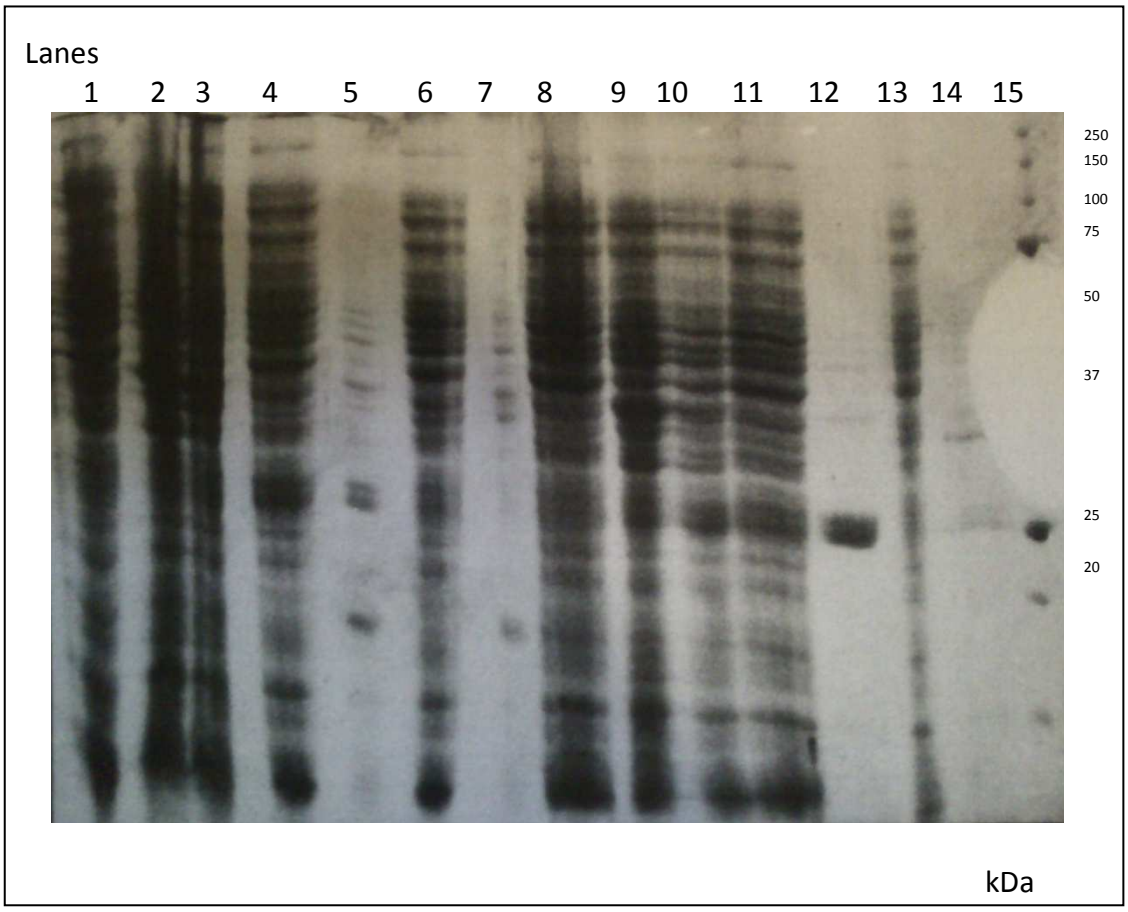


Figure 42: Fusion protein probe Pull-down of *E.coli* expressing CFM2 - cell lysate

BL21 *E.coli* cells expressing CFM2 were lysed according to 2.2.13 GST fusion protein probe pull-down using PBS buffer and RIPA buffer. The lysate was pre-cleared using GST bound to glutathione cross-linked to agarose. Cleared lysate was probed using the negative control GST and the test fusion protein probe GST-FLNAR10. Samples were boiled in SDS+DTT with indicator dye and loaded into an SDS-PAGE 15% gel. Gel was run for 40 minutes at 200V. Gel was stained with Coomassie Blue. Molecular sizes of bands are indicated on the right in kDa. Lanes 1-7 were done in PBS, lanes 8-14 were done in RIPA buffer.

Lanes	**1**	**5µL Supernatant (PBS buffer)**
	2	**5µL Pellet (PBS buffer)**
	3	**5µL Lysate depletion step beads (PBS buffer)**
	4	**5µL GST control eluent (PBS buffer)**
	5	**5µL GST control beads (PBS buffer)**
	6	**5µL Fusion eluent (PBS buffer)**
	7	**5µL Fusion beads (PBS buffer)**
	8	**5µL Supernatant (RIPA buffer)**
	9	**5µL Pellet (RIPA buffer)**
	10	**5µL Lysate depletion step beads (RIPA buffer)**
	11	**5µL GST control eluent (RIPA buffer)**
	12	**5µL GST control beads (RIPA buffer)**
	13	**5µL Fusion eluent (RIPA buffer)**
	14	**5µL Fusion beads (RIPA buffer)**
	15	**5µL PPU**

Comparing lanes 5 vs. 7 of Figure 41 the bands present that are not GST or the fusion protein respectively are those bound to the GST, the FLNAR10 or the column. Bands circled in lanes 7 and 14 of Figure 41 are the fusion protein probe, those in lanes 5 and 11 of Figure 41 are the GST probe. There was no increase in any of the bands from lane 5 to 7 which would have indicated a protein that bound GST-FLNAR10 fusion protein over the GST alone. The same can be said for lanes 12 vs. 14, which is the same probe assay but the cells were lysed in the RIPA buffer. While the C2C12 cells are the progenitors to muscle cells not bone, had another protein being found it would have been the only one other than CFM2 identified as a binding partner.

The same comparison of lanes 5 vs. 7 and 12 vs. 14 of Figure 42 show no increase in a band not belonging to the probe protein (GST-FLNAR10) or control protein (GST). While it was hoped that the fusion protein probe would bind to the CFM2 protein the *E.coli* cells were meant to produce, the fact that most of it is insoluble indicated that it was probably not folding correctly and so would not bind to FLNAR10.

4.1 Summary

4.1.1 CFM2

Initial transformations of *E.coli* with CFM2 expressing plasmids had very low success rates. Transformations of CFM2 expressing plasmids done by Post Doctorate Dr Greg Sawyer also had low success rate with premature stop-codon mutations present in the successful transformants. Only the short isoform of CFM2 was able to be amplified using PCR and therefore subcloned into *E.coli* expression plasmid pProEX HTb.

The very high level of insoluble protein when CFM2 was produced by *E.coli* in this project explains the low growth rates of *E.coli* transformed with CFM2 producing plasmid. The insolubility of CFM2 in *E.coli* can be explained by the observations [28] that RefilinA and RefilinB (FAM101A and FAM101B respectively) are stabilised upon interaction with Filamins in eukaryotes.

Only vertebrate Filamins are known to co-immunoprecipitate with Refilins. The low stability of 6xHis-CFM2 protein product is predicted by ExPASy ProtParam tool at ~64 on the instability index (a value greater than 40 indicates the protein is probably unstable).

RefilinA and RefilinB have the DSG motif which mediates the degradation of short lived proteins [25]. With this in mind any expression system for Refilin should try and account for this.

4.1.2 FLNAR10

Filamin A repeat 10 fused to N-terminal GST subclonned into pGEX 4T3 was transformed successfully into BL21 *E.coli*. An expression protocol was setup that produced a large amount of protein per *E.coli* protein weight. Purification under protease inhibiting conditions using glutathione-sepharose bead affinity column produced a very pure fusion protein solution at ~1 mg/mL. Using Vivaspin 10kDa cut off tube the fusion protein was unable to be concentrated beyond ~1 mg/mL, it may be aggregating but no properties of the protein or buffer suggests a cause for the aggregation. Thrombin digestion of the fusion GST-FLNAR10 initially produced the

expected two bands but due to the digestion time or instability due to incorrect buffers multiple digestion products were created. The size of the digest products meant FPLC could not separate the digestion products. According to the paper [1] published after this projects experimentation was done, pure FLNAR10 was able to be produced, the buffer used was similar [26] but for the addition of 10% glycerol which is thought to prevent protein aggregation by inhibiting protein unfolding through interacting with hydrophobic surfaces [39].

4.1.3 FLNAR10 pull-down

GST-FLNAR10 fusion protein was used as a probe on C2C12 mouse fibroblast cell lysate and *E.coli* BL21 (expressing CFM2) lysate. RIPA and PBS buffers were used as lysis buffers in parallel experiments, RIPA buffer because of the high detergent level to help with cell disruption and PBS as a comparison in case artefacts occurred.

While CFM2 is not known to be expressed in C2C12 cells the lysate was probed for any other proteins that may interact with FLNAR10. No such protein was found. Using *E.coli* BL21 (expressing CFM2) lysate and probing with GST-FLNAR10 fusion was hoped to confirm CFM2-FLNAR10 interaction. Unfortunately CFM2 did not get pulled down by the fusion.

The lack of CFM2 pull-down by GST-FLNAR10 could have been for many reasons. The most likely is that as stated in the paper [28] “Refilins [FAM101A and FAM101B] are stabilized upon interaction with Filamins”. While they give no supporting papers or evidence for this, the transformation and insolubility problems in *E.coli* by Post Doctoratal Fellow Dr Greg Sawyer and myself make this credible. Another reason could be CFM2 only weakly binds FLNAR10.

The yeast-2-hybrid screen which identified repeats 10 and 21 of FLNA as binding partners may be much more sensitive than the detection of Coomassie staining used in the pull-down experiment. The FLNA repeats 10 and 21 are of different classes (D and A respectively) [26] which have been shown to bind the other’s ligands weakly at best.

4.1.4 Future work

Due to the solubility problems with CFM2 in the prokaryotic expression system based in *E.coli* probably being due to a lack of Filamin A to stabilise the protein, further work should attempt to use an expression system that is eukaryotic.

An obvious choice would be a yeast system. Since a yeast-2 hybrid assay originally identified CFM2s' binding partners CFM2 is soluble and functional in such a system. Due to the potential for the CFM2 product to be toxic to the yeast a tightly controlled promoter would be recommended.

Another option would be a baculovirus infected insect cell as the host cells lyse and die during each infection cycle anyway the toxicity to the cells is less important.

By using a cell line that has a Filamin network that may be affected by over-expressed CFM2 immunostaining of Filamin A in normal and CFM2 over expressing cell lines may be interesting.

With the addition of glycerol to PBS buffer Filamin A repeat 10 purification appears more stable, with purified CFM2 from a eukaryotic expression system the characteristics and nature of their interactions may be more fully explored.

With the potential to purify CFM2 with use of a eukaryotic expression system, crystallography may be able to identify a structure for CFM2.

5 References

1. Page, R.C., J.G. Clark, and S. Misra, *Structure of filamin A immunoglobulin-like repeat 10 from Homo sapiens.* Acta Crystallographica. Section F, Structural Biology and Crystallization Communications, 2011. **67**(Pt 8): p. 871-6.
2. Hartwig, J.H., *Actin-binding proteins. 1: Spectrin super family.* Protein Profile, 1995. **2**(7): p. 703-800.
3. Wagner, B., et al., *Cytoskeletal Polymer Networks: The Molecular Structure of Cross-Linkers Determines Macroscopic Properties.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(38): p. 13974-13978.
4. Honda, M., et al., *Morphogenesis of liposomes encapsulating actin depends on the type of actin-crosslinking.* Journal of Molecular Biology, 1999. **287**(2): p. 293-300.
5. Stossel, T.P., et al., *Filamins as integrators of cell mechanics and signalling.* Nature Reviews Molecular Cell Biology, 2001. **2**(2): p. 138-45.
6. Wang, K. and S.J. Singer, *Interaction of filamin with f-actin in solution.* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**(5): p. 2021-5.
7. Flanagan, L.A., et al., *Filamin A, the Arp2/3 Complex, and the Morphology and Function of Cortical Actin Filaments in Human Melanoma Cells.* The Journal of Cell Biology, 2001. **155**(4): p. 511-517.
8. Olsen, B.R., A.M. Reginato, and W.F. Wang, *Bone development.* Annual Review of Cell and Developmental Biology, 2000. **16**: p. 191-220.
9. Verloes, A., et al., *Fronto-otopalatodigital osteodysplasia: Clinical evidence for a single entity encompassing Melnick-Needles syndrome, otopalatodigital syndrome types 1 and 2, and frontometaphyseal dysplasia.* American Journal of Medical Genetics, 2000. **90**(5): p. 407-422.
10. Robertson, S.P., *Otopalatodigital syndrome spectrum disorders: otopalatodigital syndrome types 1 and 2, frontometaphyseal dysplasia and Melnick-Needles syndrome.* European Journal of Human Genetics, 2007. **15**(1): p. 3-9.
11. Eksioglu, Y.Z., et al., *Periventricular heterotopia: an X-linked dominant epilepsy locus causing aberrant cerebral cortical development.* Neuron, 1996. **16**(1): p. 77-87.
12. Robertson, S.P., et al., *Localized mutations in the gene encoding the cytoskeletal protein filamin A cause diverse malformations in humans.* Nature Genetics, 2003. **33**(4): p. 487-491.
13. Robertson, S.P., et al., *Frontometaphyseal dysplasia: Mutations in FLNA and phenotypic diversity (vol 140A, pg 1726, 2006).* American Journal of Medical Genetics Part A, 2006. **140A**(24): p. 2840-2840.

14. Wang, K., J.F. Ash, and S.J. Singer, *Filamin, a new high-molecular-weight protein found in smooth muscle and non-muscle cells.* Proceedings of the National Academy of Sciences of the United States of America, 1975. **72**(11): p. 4483-6.

15. Hock, R.S., G. Davis, and D.W. Speicher, *Purification of human smooth muscle filamin and characterization of structural domains and functional sites.* Biochemistry, 1990. **29**(40): p. 9441-51.

16. Tyler, J.M., J.M. Anderson, and D. Branton, *Structural comparison of several actin-binding macromolecules.* The Journal of Cell Biology, 1980. **85**(2): p. 489-95.

17. Fucini, P., et al., *The repeating segments of the F-actin cross-linking gelation factor (ABP-120) have an immunoglobulin-like fold.* Nature Structural Biology, 1997. **4**(3): p. 223-30.

18. Fucini, P., et al., *Molecular architecture of the rod domain of the Dictyostelium gelation factor (ABP120).* Journal of Molecular Biology, 1999. **291**(5): p. 1017-23.

19. Nakamura, F., et al., *Structural basis of filamin A functions.* The Journal of Cell Biology, 2007. **179**(5): p. 1011-1025.

20. Gorlin, J.B., et al., *Human endothelial actin-binding protein (ABP-280, nonmuscle filamin): a molecular leaf spring.* The Journal of Cell Biology, 1990. **111**(3): p. 1089-105.

21. Nakamura, F., et al., *Ca2+ and calmodulin regulate the binding of filamin A to actin filaments.* Journal of Biological Chemistry, 2005. **280**(37): p. 32426-33.

22. Gardel, M.L., et al., *Prestressed F-actin networks cross-linked by hinged filamins replicate mechanical properties of cells.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(6): p. 1762-7.

23. Pollard, T.D., et al., *Structure of Actin-binding proteins - Insights about function at atomic-resolution.* Annual Review of Cell Biology, 1994. **10**: p. 207-249.

24. Ito, T., A. Suzuki, and T.P. Stossel, *Regulation of water flow by actin-binding protein-induced actin gelatin.* Biophysical Journal, 1992. **61**(5): p. 1301-5.

25. Gay, O., et al., *RefilinB (FAM101B) targets FilaminA to organize perinuclear actin networks and regulates nuclear shape.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(28): p. 11464-9.

26. Ithychanda, S.S., et al., *Identification and characterization of multiple similar ligand-binding repeats in filamin implication on filamin-mediated receptor clustering and cross-talk.* Journal of Biological Chemistry, 2009. **284**(50): p. 35113-35121.

27. BioGem. *CFSSP: Chou & Fasman Secondary Structure Prediction Server*. 2012 [cited 2012 10/8]; Available from: http://www.biogem.org/tool/chou-fasman/.

28. Gay, O., F. Nakamura, and J. Baudier, *Refilin holds the cap.* Communicative & Integrative Biology, 2011. **4**(6): p. 791-5.

29. technologies, L. *One Shot® TOP10 Chemically Competent E.coli*. 2013 [cited 2012 10/8]; Available from: http://products.invitrogen.com/ivgn/product/C404010.

30. technologies, L. *One Shot® BL21 Star™ (DE3) Chemically Competent E.coli*. 2013 [cited 2012 10/8]; Available from: http://products.invitrogen.com/ivgn/product/C601003.

31. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

32. Protana. *Silver staining protocol compatible with mass spectrometry, excellent signal-to-noise ratio*. [cited 2011 10/7]; Available from: www.lamondlab.com/pdf/silver_staining.pdf.

33. *Detection of protein-protein interactions using the GST fusion protein pull-down technique.* Nature Methods, 2004. **1**(3): p. 275-276.

34. Promega. *pGEM®-T Vector Systems*. 2012 [cited 2012 14/9]; Available from: http://www.promega.com/products/pcr/pcr-cloning/pgem_t-vector-systems/.

35. addgene. *pPROEX HTb B*. 2012 [cited 2012 10/9]; Available from: http://www.addgene.org/vector-database/3836/.

36. technologies, L. *pREP4 - Life Technologies*. 2013 [cited 2012 10/8]; Available from: http://products.invitrogen.com/ivgn/product/V00450.

37. Healthcare, G. *pGEX vectors*. 2012 [cited 2012 13/9]; Available from: http://www.gelifesciences.com/webapp/wcs/stores/servlet/productById/en/GELifeSciences-nz/28954552.

38. Smyth, D.R., et al., *Crystal structures of fusion proteins with large-affinity tags.* Protein Science, 2003. **12**(7): p. 1313-1322.

39. Vagenende, V., M.G.S. Yap, and B.L. Trout, *Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol.* Biochemistry, 2009. **48**(46): p. 11084-11096.

6 Appendix

Table 20: FASTA protein sequences of CFM2 isoforms

>sp|Q6ZTI6|F101A_HUMAN Protein FAM101A OS=Homo sapiens GN=FAM101A PE=2 SV=3

MVGHLHLQGMEDSLKEQGREGLLDSPDSGLPPSPSPSPPFYSLAPGILDARAGGAGASSE

PPGPSEARAPPSQLPNPPASEMRPRMLPVFFGESIKVNPEPTHEIRCNSEVKYASEKHFQ

DKVFYAPVPTVTAYSETIVAAPNCTWRNYRSQLTLEPRPRALRFRSTTIIFPKHARSTFR

TTLHCSLGRPSRWFTASVQLQLCQDPAPSLLGPATL

>sp|Q6ZTI6-2|F101A_HUMAN Isoform 2 of Protein FAM101A OS=Homo sapiens GN=FAM101A

MRPRMLPVFFGESIKVNPEPTHEIRCNSEVKYASEKHFQDKVFYAPVPTVTAYSETIVAA

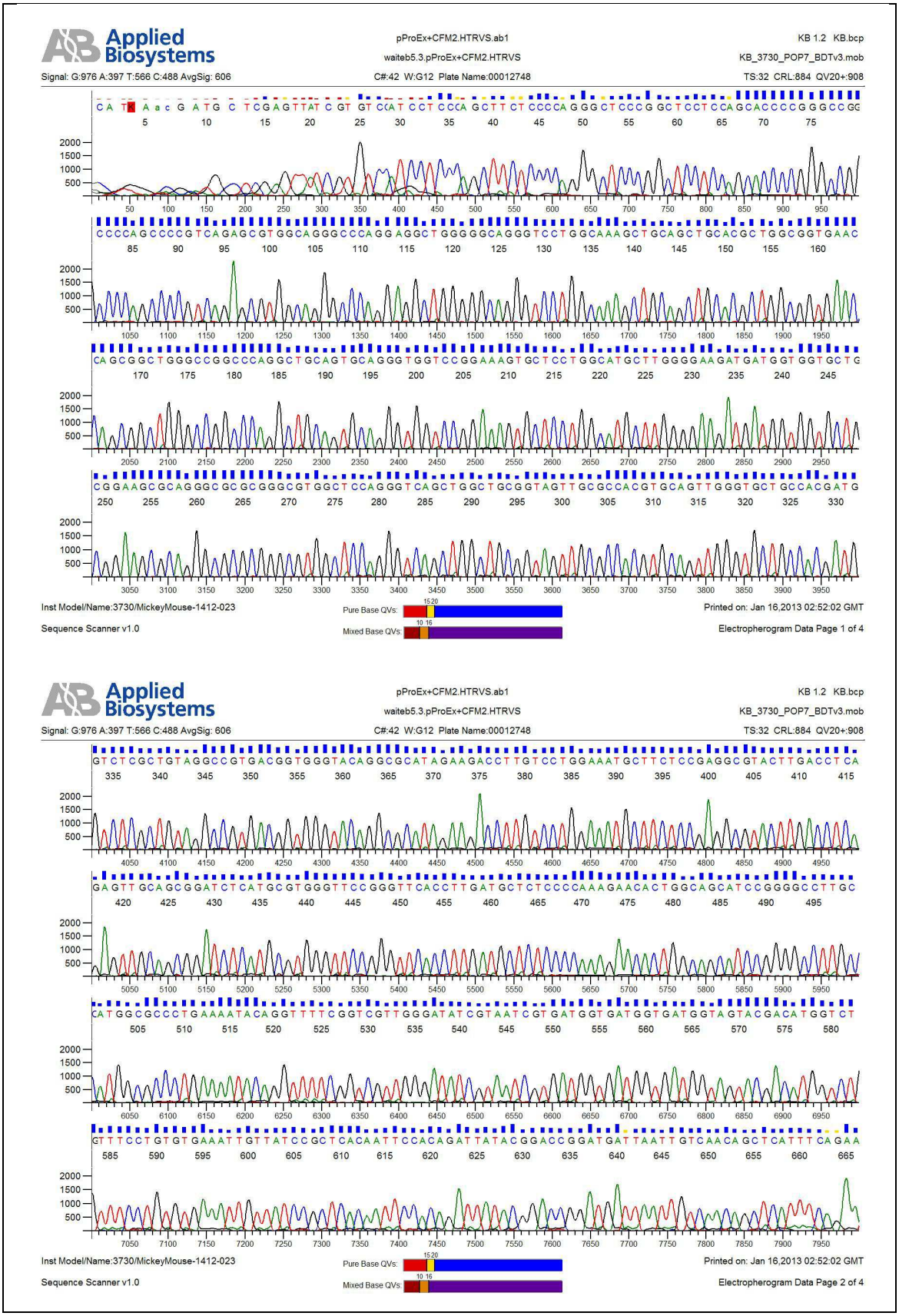
PNCTWRNYRSQLTLEPRPRALRFRSTTIIFPKHARSTFRTTLHCSLGRPSRWFTASVQLQ

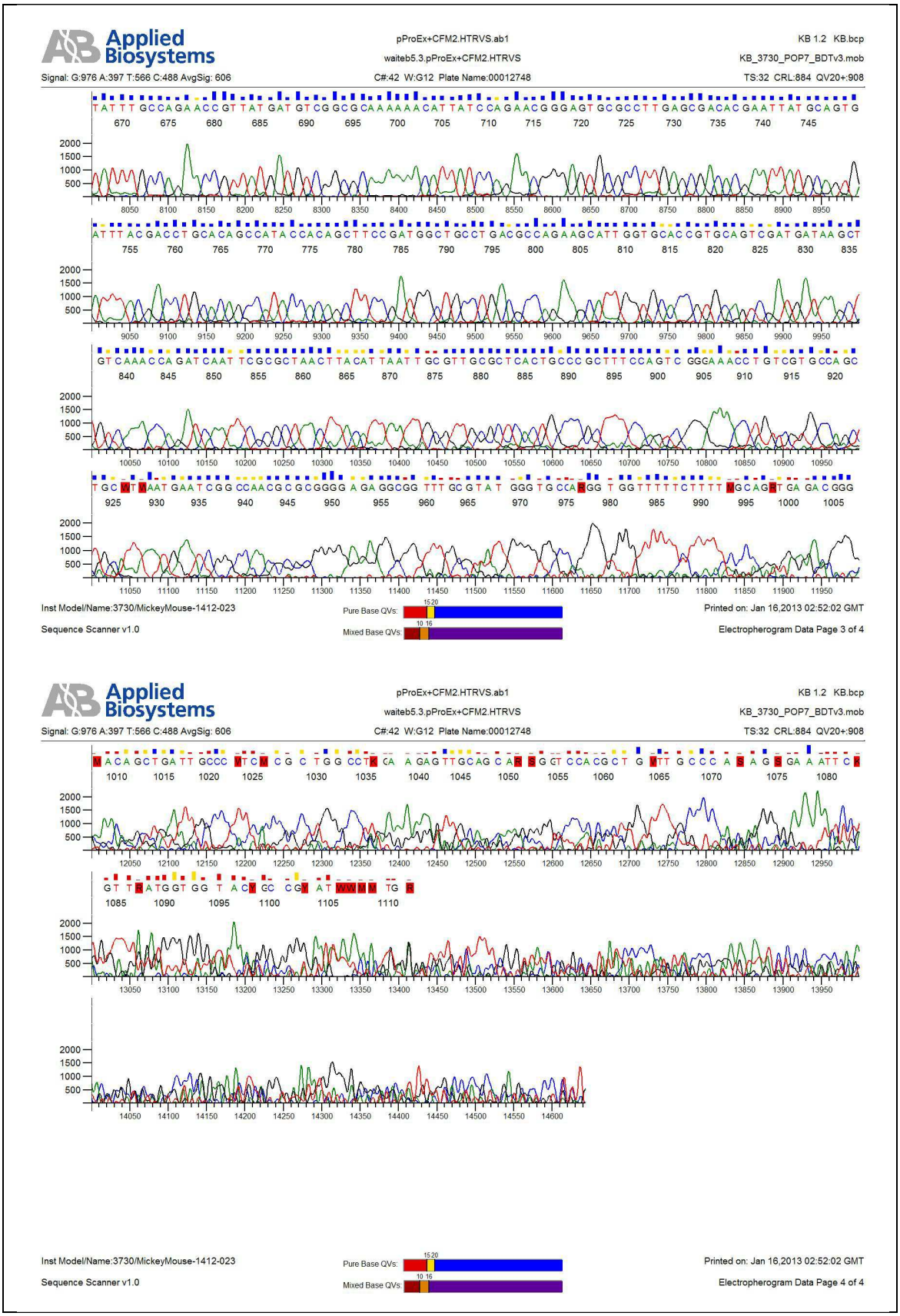
LCQDPAPSLLGPATL

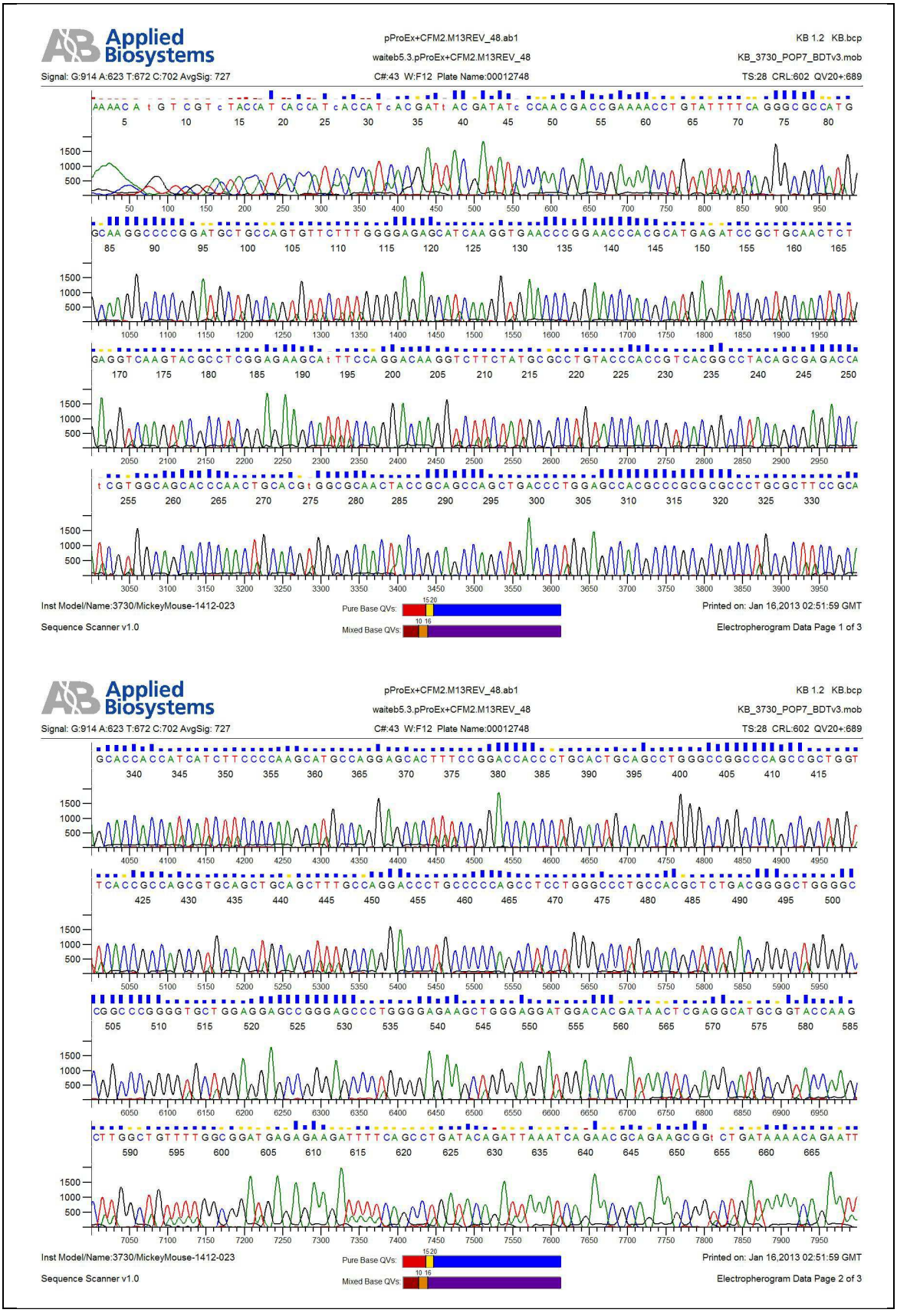
Table 21: FAM101A sequence in pGEMT vector

5`-ATG GTG GGC CAC CTG CAT CTG CAG GGC ATG GAG GAC AGC CTG AAG GAG CAG GGC CGG GAG GGC TTG CTG GAC AGC CCC GAC TCC GGG CTG CCC CCC AGC CCC AGC CCC AGC CCG CCC TTC TAC TCC CTG GCG CCC GCG ATC CTC GAC GCG CGC GCG GGG GGC GCC GGC GCC TCC GAG CCC CCG GGA CCC AGA GAG GCC AGA ACG CCC CCC TCC CAA CTC CCA AAT CCC CCG GCG TCG GAG ATG AGG CCC CGG ATG CTG CCA GTG TTC TTT GGG GAG AGC ATC AAG GTG AAC CCG GAA CCC ACG CAT GAG ATC CGC TGC AAC TCT GAG GTC AAG TAC GCC TCG GAG AAG CAT TTC CAG GAC AAG GTC TTC TAT GCG CCC GTA CCC ACC GTC ACG GCC TAC AGC GAG ACC ATC GTG GCA GCA CCC AAC TGC ACG TGG CGC AAC TAC CGC AGC CAG CTG ACC CTG GAG CCA CGC CCG CGC GCC CTG CGC TTC CGC AGC ACC ACC ATC ATC TTC CCC AAG CAT GCC AGG AGC ACT TTC CGG ACC ACC CTG CAC TGC AGC CTG GGC CGG CCC AGC CGC TGG TTC ACC GCC AGC GTG CAG CTG CAG CTT TGC CAG GAC CCT GCC CCC AGC CTC CTG GGC CCT GCC ACG CTC-3`

Table 22: Chromatogram of pProEX+CFM2 sequencing using HTRVS and m13REV primers




Applied Biosystems
pProEx+CFM2.HTRVS.ab1
KB 1.2 KB.bcp
waiteb5.3.pProEx+CFM2.HTRVS
KB_3730_POP7_BDTv3.mob
Signal: G:976 A:397 T:566 C:488 AvgSig: 606
C#:42 W:G12 Plate Name:00012748
TS:32 CRL:884 QV20+:908
Inst Model/Name:3730/MickeyMouse-1412-023
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 02:52:02 GMT
Sequence Scanner v1.0
Electropherogram Data Page 3 of 4
Applied Biosystems
pProEx+CFM2.HTRVS.ab1
KB 1.2 KB.bcp
waiteb5.3.pProEx+CFM2.HTRVS
KB_3730_POP7_BDTv3.mob
Signal: G:976 A:397 T:566 C:488 AvgSig: 606
C#:42 W:G12 Plate Name:00012748
TS:32 CRL:884 QV20+:908
Inst Model/Name:3730/MickeyMouse-1412-023
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 02:52:02 GMT
Sequence Scanner v1.0
Electropherogram Data Page 4 of 4


Applied Biosystems
pProEx+CFM2.M13REV_48.ab1
waiteb5.3.pProEx+CFM2.M13REV_48
C#:43 W:F12 Plate Name:00012748
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
TS:28 CRL:602 QV20+:689
Signal: G:914 A:623 T:672 C:702 AvgSig: 727
Inst Model/Name:3730/MickeyMouse-1412-023
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 02:51:59 GMT
Electropherogram Data Page 1 of 3
Applied Biosystems
pProEx+CFM2.M13REV_48.ab1
waiteb5.3.pProEx+CFM2.M13REV_48
C#:43 W:F12 Plate Name:00012748
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
TS:28 CRL:602 QV20+:689
Signal: G:914 A:623 T:672 C:702 AvgSig: 727
Inst Model/Name:3730/MickeyMouse-1412-023
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 02:51:59 GMT
Electropherogram Data Page 2 of 3

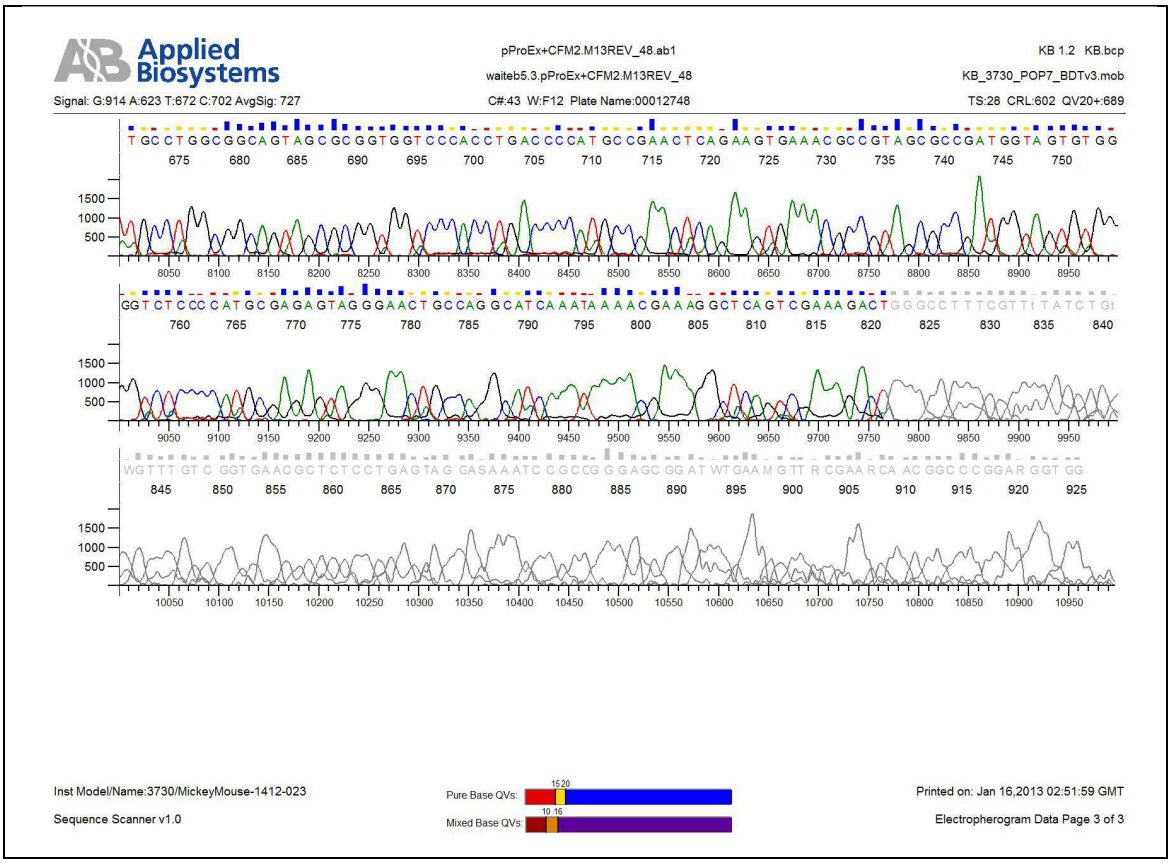
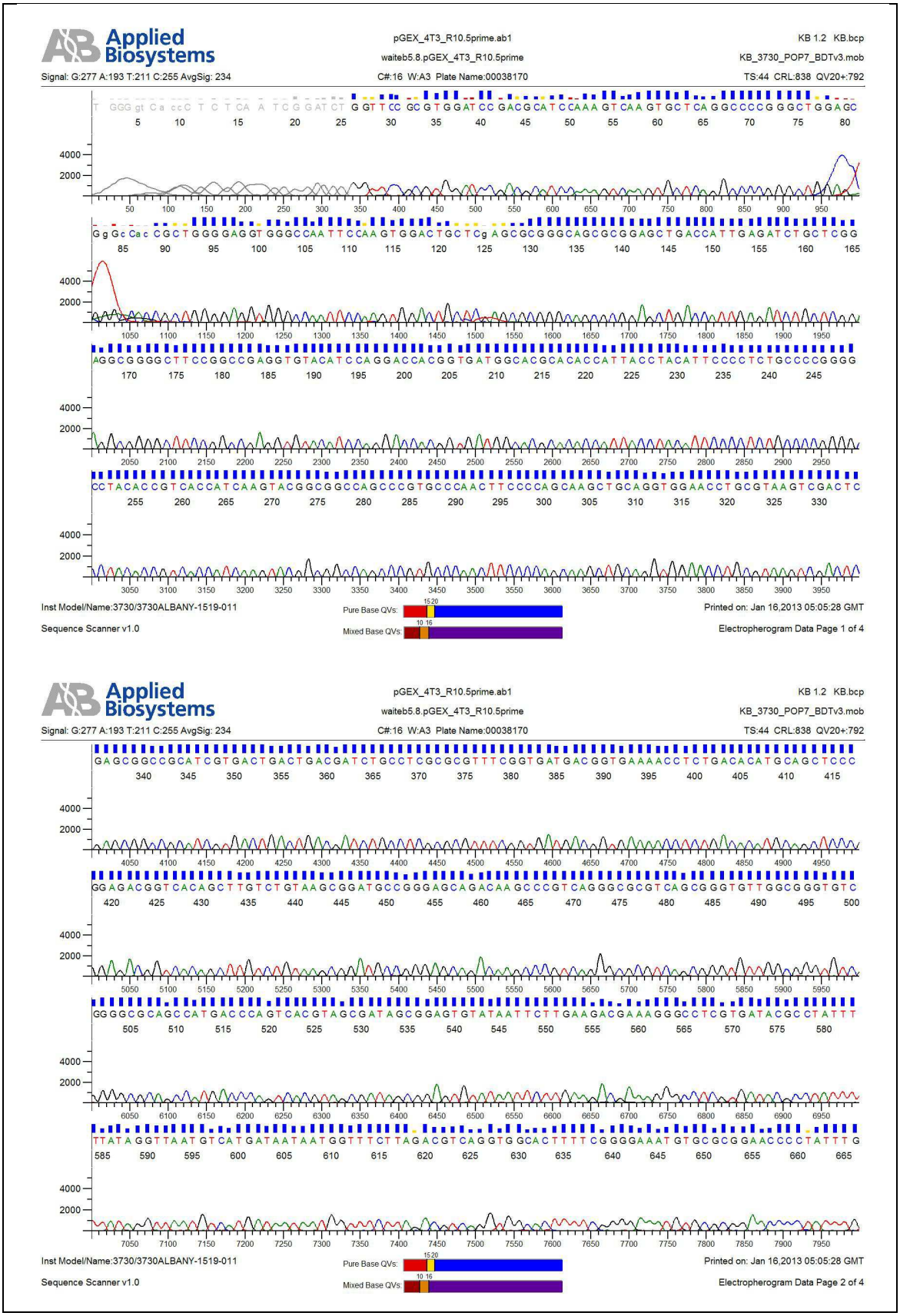


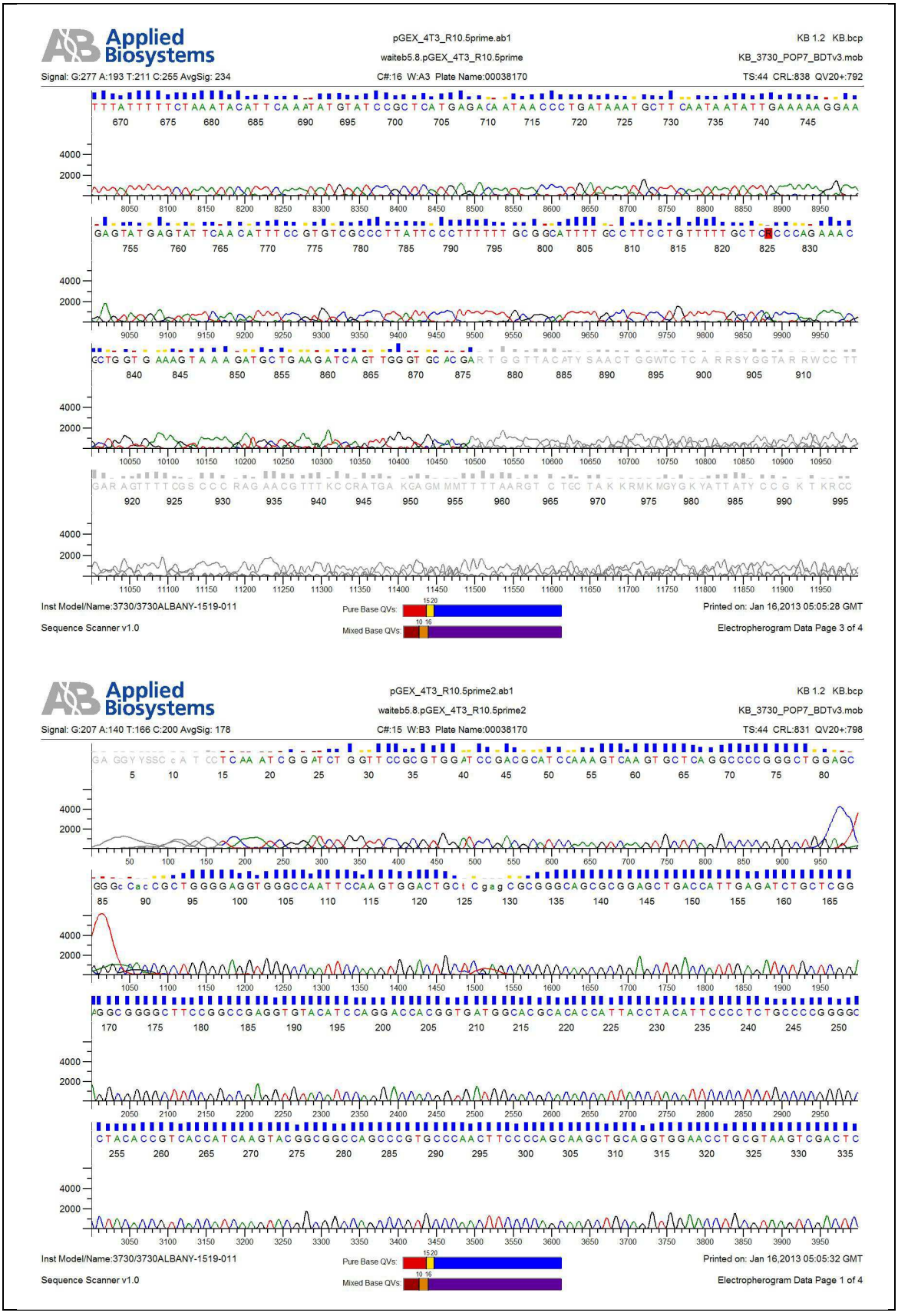
Table 23: Alignment of (Sequence 1) FAM101A mRNA with (Sequence 2) pProEX+CFM2 sequenced consensus

```
Similarity : 434/486 (89.30 %)

             T  L  S  L  *  F  Y   P  C  P  G  L  I  K  C  S  E  D  I  N
Seq_1  1     ACACTCAGTCTCTGATTTT-ATCCCTGCCCTGGCTTAATTAAGTGCTCTGAAGACATAAA  59
             |  |  ||  | |   ||| |   |          | |         |   |   |
Seq_2  1     ------A-TGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGA  53
                   S  Y  Y  H  H  H  H  H  H  D  Y  D  I  P  T  T  E

              R  F  F  V  Y  S  S  L  G  E  A  G  R  K  G  L  F  S  *  K
Seq_1  60    TAGGTTTTTTGTTTACTCAAGCTTGGGAGAAGCTGGCAGAAAAGGCCTTTTCAGCTAAAA  119
              |   | | | || |      | |||  |
Seq_2  54    AAACCTGTATTTTCAGGGCGCCATGGCA------------------------------  81
              N  L  Y  F  Q  G  A  M  A

              L  L  T  A  G  R  C  E  A  A  S  F  Q  P  G  R  S  L  S  A
Seq_1  120   GCTCCTTACAGCCGGAAGGTGTGAAGCGGCAAGTTTCCAGCCGGGAAGAAGCCTCTCAGC  179

Seq_2  82    ------------------------------------------------------------  81


              V  G  V  F  A  R  S  C  E  P  P  S  Q  L  P  N  P  P  A  S
Seq_1  180   CGTAGGCGTCTTTGCCCGGAGCTGTGAGCCCCCCTCCCAACTCCCAAATCCCCCGGCGTC  239

Seq_2  82    ------------------------------------------------------------  81


              E  M  R  P  R  M  L  P  V  F  F  G  E  S  I  K  V  N  P  E
```

```
Seq_1  240   GGAGATGAGGCCCCGGATGCTGCCAGTGTTCTTTGGGGAGAGCATCAAGGTGAACCCGGA  299
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  82    -------AGGCCCCGGATGCTGCCAGTGTTCTTTGGGGAGAGCATCAAGGTGAACCCGGA  134
                    R  P  R  M  L  P  V  F  F  G  E  S  I  K  V  N  P  E


              P  T  H  E  I  R  C  N  S  E  V  K  Y  A  S  E  K  H  F  Q
Seq_1  300   ACCCACGCATGAGATCCGCTGCAACTCTGAGGTCAAGTACGCCTCGGAGAAGCATTTCCA  359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  135   ACCCACGCATGAGATCCGCTGCAACTCTGAGGTCAAGTACGCCTCGGAGAAGCATTTCCA  194
              P  T  H  E  I  R  C  N  S  E  V  K  Y  A  S  E  K  H  F  Q


              D  K  V  F  Y  A  P  V  P  T  V  T  A  Y  S  E  T  I  V  A
Seq_1  360   GGACAAGGTCTTCTATGCGCCCGTACCCACCGTCACGGCCTACAGCGAGACCATCGTGGC  419
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Seq_2  195   GGACAAGGTCTTCTATGCGCCTGTACCCACCGTCACGGCCTACAGCGAGACCATCGTGGC  254
              D  K  V  F  Y  A  P  V  P  T  V  T  A  Y  S  E  T  I  V  A


              A  P  N  C  T  W  R  N  Y  R  S  Q  L  T  L  E  P  R  P  R
Seq_1  420   AGCACCCAACTGCACGTGGCGCAACTACCGCAGCCAGCTGACCCTGGAGCCACGCCCGCG  479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  255   AGCACCCAACTGCACGTGGCGCAACTACCGCAGCCAGCTGACCCTGGAGCCACGCCCGCG  314
              A  P  N  C  T  W  R  N  Y  R  S  Q  L  T  L  E  P  R  P  R


              A  L  R  F  R  S  T  T  I  I  F  P  K  H  A  R  S  T  F  R
Seq_1  480   CGCCCTGCGCTTCCGCAGCACCACCATCATCTTCCCCAAGCATGCCAGGAGCACTTTCCG  539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  315   CGCCCTGCGCTTCCGCAGCACCACCATCATCTTCCCCAAGCATGCCAGGAGCACTTTCCG  374
              A  L  R  F  R  S  T  T  I  I  F  P  K  H  A  R  S  T  F  R


              T  T  L  H  C  S  L  G  R  P  S  R  W  F  T  A  S  V  Q  L
Seq_1  540   GACCACCCTGCACTGCAGCCTGGGCCGGCCCAGCCGCTGGTTCACCGCCAGCGTGCAGCT  599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  375   GACCACCCTGCACTGCAGCCTGGGCCGGCCCAGCCGCTGGTTCACCGCCAGCGTGCAGCT  434
              T  T  L  H  C  S  L  G  R  P  S  R  W  F  T  A  S  V  Q  L


              Q  L  C  Q  D  P  A  P  S  L  L  G  P  A  T  L  *  R  G  W
Seq_1  600   GCAGCTTTGCCAGGACCCTGCCCCCAGCCTCCTGGGCCCTGCCACGCTCTGACGGGGCTG  659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  435   GCAGCTTTGCCAGGACCCTGCCCCCAGCCTCCTGGGCCCTGCCACGCTCTGA--------  486
              Q  L  C  Q  D  P  A  P  S  L  L  G  P  A  T  L  *
```

Table 24: Chromatogram of pGEX 4T3-FLNAR10 #5 and #6 plasmid sequencing using pGEX 5` primer




Applied Biosystems
pGEX_4T3_R10.5prime.ab1
waiteb5.8.pGEX_4T3_R10.5prime
C#:16 W:A3 Plate Name:00038170
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
TS:44 CRL:838 QV20+:792
Signal: G:277 A:193 T:211 C:255 AvgSig: 234
Inst Model/Name:3730/3730ALBANY-1519-011
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 05:05:28 GMT
Electropherogram Data Page 3 of 4
Applied Biosystems
pGEX_4T3_R10.5prime2.ab1
waiteb5.8.pGEX_4T3_R10.5prime2
C#:15 W:B3 Plate Name:00038170
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
TS:44 CRL:831 QV20+:798
Signal: G:207 A:140 T:166 C:200 AvgSig: 178
Inst Model/Name:3730/3730ALBANY-1519-011
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 05:05:32 GMT
Electropherogram Data Page 1 of 4

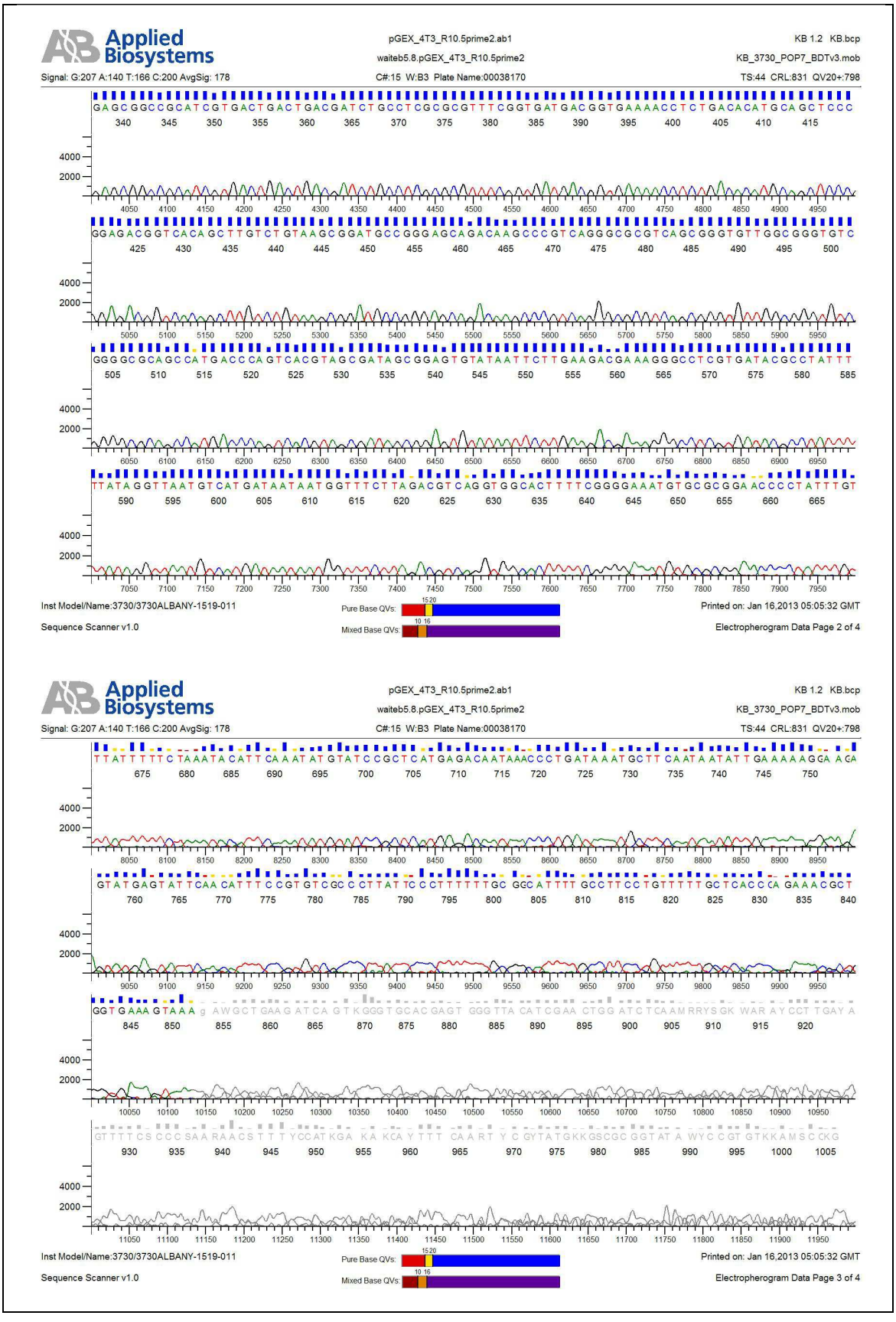

Applied Biosystems
pGEX_4T3_R10.5prime2.ab1
KB 1.2 KB.bcp
waiteb5.8.pGEX_4T3_R10.5prime2
KB_3730_POP7_BDTv3.mob
Signal: G:207 A:140 T:166 C:200 AvgSig: 178
C#:15 W:B3 Plate Name:00038170
TS:44 CRL:831 QV20+:798
Inst Model/Name:3730/3730ALBANY-1519-011
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 05:05:32 GMT
Sequence Scanner v1.0
Electropherogram Data Page 2 of 4
Applied Biosystems
pGEX_4T3_R10.5prime2.ab1
KB 1.2 KB.bcp
waiteb5.8.pGEX_4T3_R10.5prime2
KB_3730_POP7_BDTv3.mob
Signal: G:207 A:140 T:166 C:200 AvgSig: 178
C#:15 W:B3 Plate Name:00038170
TS:44 CRL:831 QV20+:798
Inst Model/Name:3730/3730ALBANY-1519-011
Pure Base QVs:
Mixed Base QVs:
Printed on: Jan 16,2013 05:05:32 GMT
Sequence Scanner v1.0
Electropherogram Data Page 3 of 4

Table 25: Alignment of (Sequence 1) Sequencing of Colony #5 with (Sequence 2) pGEX-4T3:FLNAR10

```
Similarity : 307/320 (95.94 %)

                                                                V  P  R
Seq_1  1     ------------------------------------------------GGTTCCGCGT  10
                                                               ||||||||||
Seq_2  1     ggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggatctggttccgcgt  60
             G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D  L  V  P  R


             G  S  D  A  S  K  V  K  C  S  G  P  G  L  E  R  A  T  A  G
Seq_1  11    GGATCCGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGgGcCacCGCTGGG  70
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    ggatccGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGGG  120
             G  S  D  A  S  K  V  K  C  S  G  P  G  L  E  R  A  T  A  G


             E  V  G  Q  F  Q  V  D  C  S  S  A  G  S  A  E  L  T  I  E
Seq_1  71    GAGGTGGGCCAATTCCAAGTGGACTGCTCgAGCGCGGGCAGCGCGGAGCTGACCATTGAG  130
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GAGGTGGGCCAATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTGAG  180
             E  V  G  Q  F  Q  V  D  C  S  S  A  G  S  A  E  L  T  I  E


             I  C  S  E  A  G  L  P  A  E  V  Y  I  Q  D  H  G  D  G  T
Seq_1  131   ATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCACG  190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   ATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCACG  240
             I  C  S  E  A  G  L  P  A  E  V  Y  I  Q  D  H  G  D  G  T


             H  T  I  T  Y  I  P  L  C  P  G  A  Y  T  V  T  I  K  Y  G
Seq_1  191   CACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGGC  250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   CACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGGC  300
             H  T  I  T  Y  I  P  L  C  P  G  A  Y  T  V  T  I  K  Y  G


             G  Q  P  V  P  N  F  P  S  K  L  Q  V  E  P  A  *  V  D  S
Seq_1  251   GGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAGGTGGAACCTGCGTAAGTCGACTCG  310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  301   GGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAggtggaacctgcgtaagtcgac---  357
             G  Q  P  V  P  N  F  P  S  K  L  Q  V  E  P  A  *  V  D


             S  G  R  X
Seq_1  311   AGCGGCCGCA  320

Seq_2  358   ----------  357
                   X  X
```

Table 26: Alignment of (Sequence 1) Sequencing of Colony #6 with (Sequence 2) pGEX-4T3:FLNAR10

```
Similarity : 318/340 (93.53 %)

                                                  K  S  D  L  V  P  R
Seq_1  1     ---------------------------------------AAATCGGATCTGGTTCCGCGT  21
                                                    |||||||||||||||||||||
Seq_2  1     ggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggatctggttccgcgt  60
             G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D  L  V  P  R


             G  S  D  A  S  K  V  K  C  S  G  P  G  L  E  R  A  T  A  G
Seq_1  22    GGATCCGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGcCacCGCTGGG  81
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    ggatccGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGGG  120
             G  S  D  A  S  K  V  K  C  S  G  P  G  L  E  R  A  T  A  G


             E  V  G  Q  F  Q  V  D  C  S  S  A  G  S  A  E  L  T  I  E
Seq_1  82    GAGGTGGGCCAATTCCAAGTGGACTGCtCgagCGCGGGCAGCGCGGAGCTGACCATTGAG  141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GAGGTGGGCCAATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTGAG  180
             E  V  G  Q  F  Q  V  D  C  S  S  A  G  S  A  E  L  T  I  E


             I  C  S  E  A  G  L  P  A  E  V  Y  I  Q  D  H  G  D  G  T
Seq_1  142   ATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCACG  201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   ATCTGCTCGGAGGCGGGGCTTCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCACG  240
             I  C  S  E  A  G  L  P  A  E  V  Y  I  Q  D  H  G  D  G  T


             H  T  I  T  Y  I  P  L  C  P  G  A  Y  T  V  T  I  K  Y  G
Seq_1  202   CACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGGC  261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   CACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCATCAAGTACGGC  300
             H  T  I  T  Y  I  P  L  C  P  G  A  Y  T  V  T  I  K  Y  G


             G  Q  P  V  P  N  F  P  S  K  L  Q  V  E  P  A  *  V  D  S
Seq_1  262   GGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAGGTGGAACCTGCGTAAGTCGACTCG  321
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  301   GGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAggtggaacctgcgtaagtcgac---  357
             G  Q  P  V  P  N  F  P  S  K  L  Q  V  E  P  A  *  V  D


             S  G  R  I  V  T  X
Seq_1  322   AGCGGCCGCATCGTGACTG  340

Seq_2  358   ------------------  357
                        X  X
```