The use of Affimers to investigate the role of the SH2 domain in the MAPK pathway

Thomas Young

201015867

Submitted in accordance with the requirements for the degree of Masters by Research

University of Leeds

Faculty of Biological Sciences

Signed: Thomas Young

Acknowledgements

My thanks go to Darren Tomlinson for allowing me to undertake this project, as well as for his guidance and support throughout the process. I would also like to thank all members of the Tomlinson lab, all of whom have at one point helped me. In particular I would like to thank Lewis Adams for his constant help and teaching of new techniques as well as carrying out the size exclusion chromatography. In addition, I would like to thank Christian Tiede for his help and guidance in crystallography and for help designing primers. I also have to thank Greg Billenness for his support and help throughout the year in getting results for this project.

To my fellow master's students, I have to say thank you for the friendship and support without which I might not have made it through this project. Hopefully nothing in this thesis deserves the wall of shame! Outside of the lab I have to thank my parents for supporting me financially while I studied, without this support I would have never been able to consider taking on this project. Thanks also go to Chi Trinh for his patience and help in designing optimisation trials for crystallisation as well as picking and diffracting my crystals.

Table of Contents

Table of Abbreviations

Abbreviation	Definition
ABL kinase	Abelson murine leukaemia viral oncogene homolog 1 kinase
Abs	Antibodies
BTK	Bruton's Tyrosine Kinase
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
Carb	Carbenicillin
CIP	Calf-intestinal Phosphatase
Crk1	Serine/Threonine Protein Kinase
DARPins	Designed Ankyrin Repeat Proteins
DNA	Deoxyribonucleic acid
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
ELK1	ETS like protein
ERK	Extracellular Signal Kinase
EtOH	Ethanol
FBS	Foetal Bovine Serum
FN3 domains	Fibronectin type 3 domains
GDP	Guanosine Diphosphate
Grb2	Growth Factor Receptor Bound Protein 2
GTP	Guanosine Triphosphate
HEK293	Human Embryonic Kidney Cells 293
HGF	Hepatocyte Growth Factor
HRV	Human Rhinovirus
hSUMO	Human SUMO
IG	Immunoglobulin
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ITC	Isothermal Calorimetry
JNK	Jun-amino Terminal Kinases
LB	Lysogeny Broth
MAPK	Mitogen activated protein kinase
MAPKK	MAPK Kinase
MAPKKK	MAPKK Kinase
MEK	Mitogen Activated Protein Kinase Kinase
Ni-NTA	Nickel-Nitrilotriacetic Acid
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PPI	Protein-Protein Interactions
RNA	Ribonucleic acid
RPM	Rotations Per Minute
RT-PCR	Reverse Transcription PCR
SAPK	Stress Activated Protein Kinases
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SH2	Src homology 2
SOE	Splice Overlap Extension
SOS	Son of Sevenless
SRF	Serum Response Factor
SUMO	Small Ubiquitin like Modifier
TKI	Tyrosine Kinase Inhibitors
TNBC	Triple Negative Breast Cancer
XLA	X-linked Agammaglobulinemia
ySUMO	Yeast SUMO

Abstract

Affimers are an antibody scaffold first developed in 2014 which are shown to have a high degree of specificity, reduced size (compared to antibodies) and good thermal stability, as well as being easy to produce. Previous work has generated a variety of Affimers which block protein-protein interactions essential to cell signalling. One of these pathways is the MAPK pathway which primarily controls cell proliferation. This pathway is of particular interest due to the prevalence of its overactivation in tumour cells causing constitutive cell division. In this project Affimers were used to inhibit the protein-protein interactions of SH2 domain containing proteins to determine if they had previously unknown activity within the MAPK pathway. This study was unsuccessful in identifying any such proteins. Progress was made towards developing an assay that can assess the effect of Affimers on the MAPK pathway. In addition to this crystals were obtained of an Affimer-SH2 complex, however when diffracted they did not produce high resolution data. Further knowledge of the MAPK pathway, how it is activated and how it uses protein-protein interactions to affect the cellular environment could lead to better treatments and patient outcomes.

Chapter 1 Introduction

1.1 Antibody Scaffolds

Antibodies (Abs) are theoretically able to target and interact with any antigen, this makes them invaluable to the field of biological sciences, with uses in therapeutics [1], diagnostics [2] and research [3]. However, antibodies have significant drawbacks, they cannot be produced in typical bacterial expression systems like *Escherichia coli* (*E.coli*), due to their complex structure and the presence of multiple vital disulphide bonds [4]. Instead they are produced in mammalian expression systems through the creation of hybridomas, this process is more time consuming and expensive than standard expression in *E. coli* [5]. Antibodies are also large proteins (150 kDa) which prevents them from accessing potential epitopes as well as preventing them crossing the cellular membrane [4]. Therefore, a wide range of non-antibody scaffolds have been developed, these proteins use similar principles to antibodies containing regions of variability which allow them to interact with any other protein.

Most antibody scaffolds have these major advantages compared to antibodies:

1. Reduced size allows them to reach targets inaccessible to antibodies
2. Able to be expressed in *E.coli* and mammalian cell lines
3. Ability to easily disrupt protein-protein interactions
4. Better stability making them easier to store and transport

The first non-antibody scaffold was developed by the Royal Institute of Technology in Sweden, by randomising the alpha helix of the z domain of a bacterial receptor they could bind the protein to different targets. They referred to this scaffold as an affibody[6]. Since then many new antibody scaffolds have been developed (a summary of the most relevant can be seen in Table 1) with a wide range of applications.

Table 1 Comparison of Antibody Scaffolds

Scaffold protein	Parent protein	Molecular Weight (kDa)	Expression System	T_m (°C)	Method of affinity generation	Location of protein interactions
Monoclonal Antibody (IgG)	N/A	150 [4]	Hybridomas [5]	61-71 [7]	V(D)J Recombination [4]	Two antigen binding sites containing 3 hypervariable loops of varying lengths [4]
Affibody	Z domain of protein A from *Staphylococcus aureus* [6]	6.5 [8]	*E.coli* [9]	Unknown	Phage display [9]	13 residues in two alpha helices typically used to bind Fc region [8]
Affimer	Phytocystatin [10]	10 [10]	*E.coli* [10]	101 [10]	Phage display [10]	Two loops of 9 residues [10]
Anticalin	Lipocalins [11]	< 20 [11]	*E.coli* [11]	>70 [11]	Phage display [12]	Four variable loops [11]
Avimer	A-domains of cell surface receptors [13]	4 [13]	*E.coli* [13]	50 [13]	Phage display combined with linking domains [13]	28 non-conserved residues [13]
DARPin	Ankyrin repeat proteins [14]	17.7 [15]	*E.coli* [14]	60-80 [16]	SRP phage display [17] or Ribosome display [14]	Variable interface formed by multiple ankyrin repeat units [14]
Knottin	Cysteine knot motif [18]	3.5 [18]	*E.coli* (periplasm) [19] and chemical synthesis [20]	83* [21]	Phage display [22]	Highly variable locations within knottin loops [23]–[25]
Monobody	Fibronectin type III domain [26]	10 [27]	*E.coli* [26]	57-63 [28]	Phage display [26]	Three variable loops [29]
FAB	Antibody [30]	48 [30]	*E.coli* (periplasm) [30]	Unknown	Digestion of existing monoclonal antibody [30]	Same 3 hypervariable loops on antibodies [30]
ScFv	Antibody [31]	35 [32]	*E.coli* [31]	62-66 [33]	Yeast surface display [34] and phage display	Same 3 hypervariable loops on antibodies [32]

Table 1 compares the properties of different antibody scaffolds compared to a monoclonal antibody. T_m is an estimate and varies based on the exact structure of the scaffold.

* Based on a naturally occurring knottin

1.1.1 DARPins

Designed ankyrin repeat proteins (DARPins) are antibody scaffolds formed from multiple ankyrin repeat domains, these domains form an L shaped structure from two α helices and a long loop [35]. Ankyrin repeats are common in eukaryotes, typically involved in protein-protein interactions [36]. Compatible repeat modules were first designed in 2003 [37] and developed for use in high affinity binding by 2004 [14]. Essential to the design of DARPins are the use of capping ankyrin repeats which prevent further interaction with other ankyrin repeats, limiting the overall size of the polyprotein.

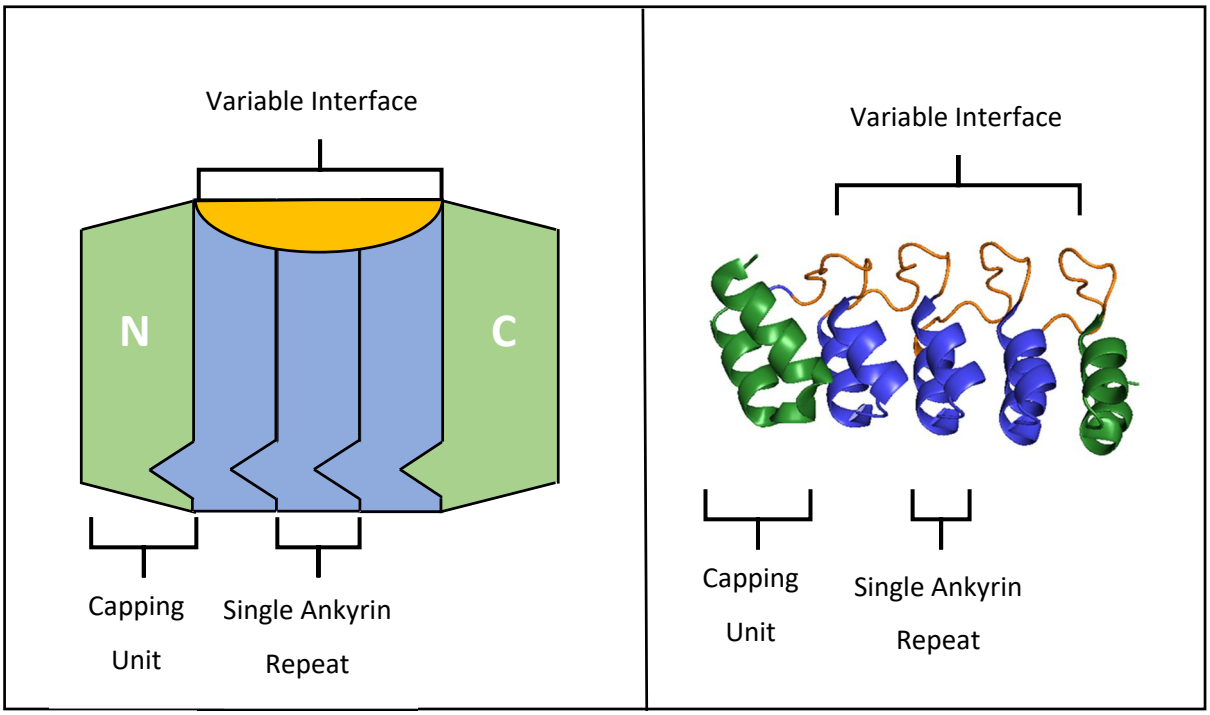


Figure 1. Schematic representation of a DARPin unit compared with its actual structure (PDB: 1MJ0). Showing the three ankyrin repeat units (blue) used to form the variable interface (orange) involved in protein-protein interactions. As well as the capping units (green) which limit the size of the DARPin. The C terminal capping unit is smaller than the N terminal capping unit.

DARPins have a wide range of applications in research, diagnostics and therapy [38], notably they have been developed to act as inhibitors of the mitogen activated protein kinase (MAPK) pathway by specifically targeting ERK kinase and blocking its activity in cells [39]. DARPins were also specific enough that they could distinguish between the phosphorylated and unphosphorylated forms of ERK, which is only differentiated by a small conformational change in the activation loop [39].

1.1.2 Monobodies

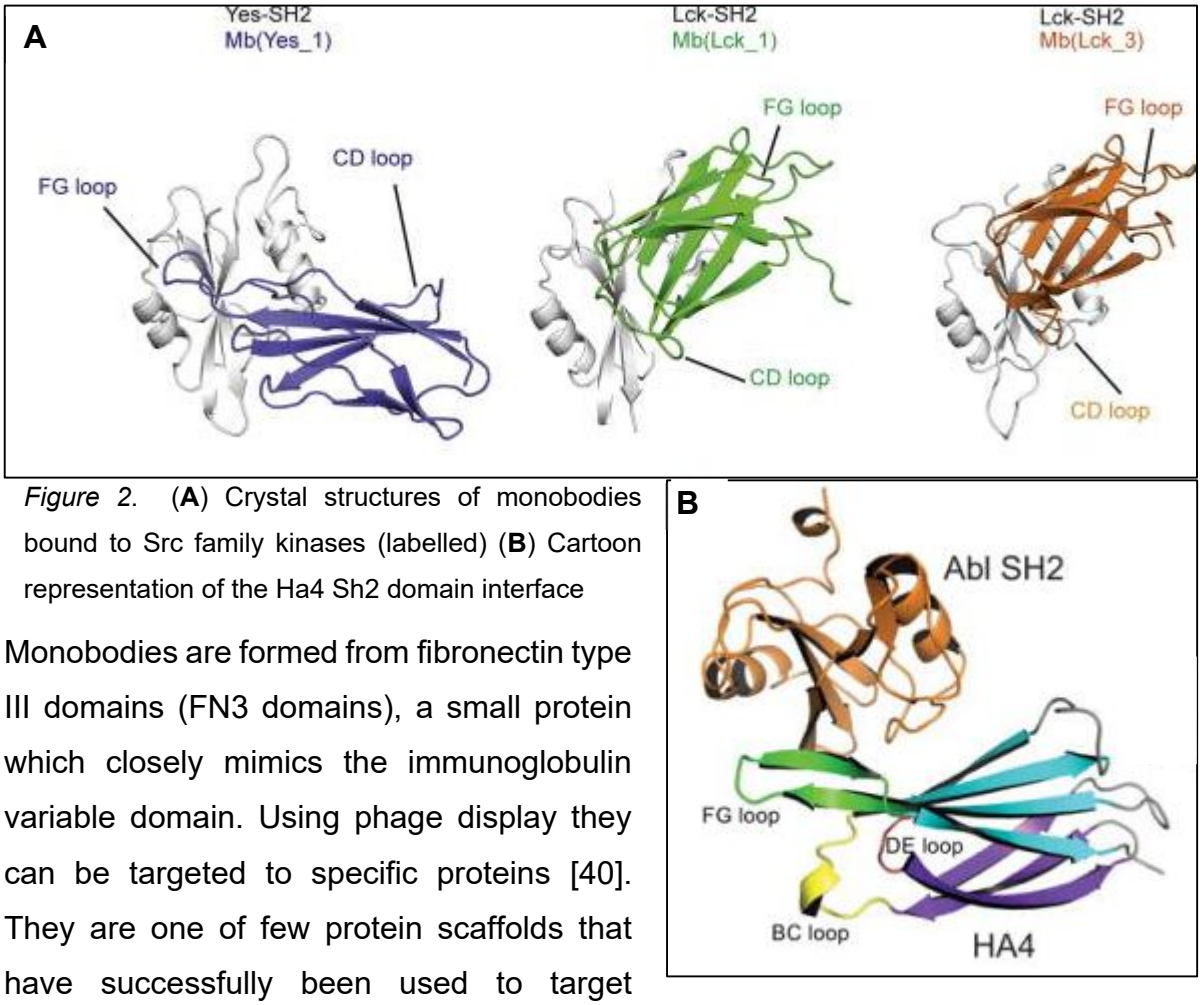


Figure 2. (**A**) Crystal structures of monobodies bound to Src family kinases (labelled) (**B**) Cartoon representation of the Ha4 Sh2 domain interface

Monobodies are formed from fibronectin type III domains (FN3 domains), a small protein which closely mimics the immunoglobulin variable domain. Using phage display they can be targeted to specific proteins [40]. They are one of few protein scaffolds that have successfully been used to target specific src homology 2 domains (SH2 domains). One study isolated monobodies targeting the SH2 domain of Abelson murine leukaemia viral oncogene homolog 1 kinase (ABL kinase). The monobody termed HA4 (see Figure 2B [41]) was able to be expressed in *E.coli* and was shown to inhibit Abl SH2 phosphopeptide interactions, with a Kd value of 7 nM. Importantly the group was able to demonstrate that HA4 was not cross reactive with 15 other SH2 domains, *in vivo* and *in vitro*. However it was shown to be cross reactive with Abl2 [41]. A more recent paper from 2017 produced monobodies which bound to the SH2 domains of Src family kinases (see Figure 2A [42]), interactome analysis of the monobodies found that they bound no other SH2 domain containing proteins. These two examples demonstrate that it is possible using non-antibody scaffolds to target specific SH2 domains with significantly reduced cross reactivity.

1.1.3 Affimers

Affimers (originally called adhirons) are a new scaffold developed in 2014, which has the potential to target and inhibit SH2 domains [43]. The Affimer scaffold was created from a consensus sequence of the phytocystatin gene and truncated to 92 amino acids long. It is formed of one α helix and four β strands, importantly the two variable regions (VR), 9 residues in length, are located on the loops between β strands (see Figure 3). The Affimer gene was also codon optimised to enhance its expression in *E.coli* [10].

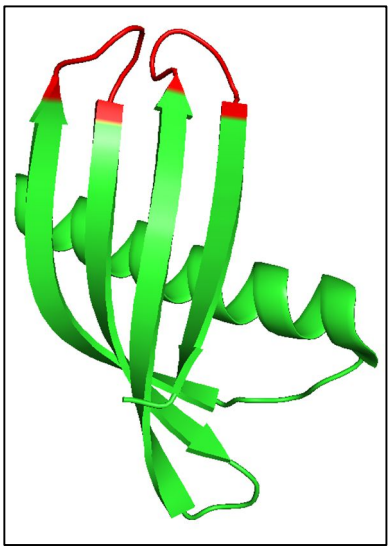


Figure.3 Cartoon representation of Affimer structure, the variable regions are in red. PDB ID:

Affimers have been shown to have wide variety of potential uses, owing to their ease of production, high thermal stability and target specificity. These uses include, inhibition of intracellular pathways, modulating ion channels, *in vivo* imaging, single particle tracking, western blotting and live cell imaging [43].

When Affimers were first developed they were tested by targeting yeast SUMO (ySUMO). Twenty-two distinct Affimers were isolated that bound ySUMO, they were found to have a thermal stability of between 85-95 °C. This demonstrates that the scaffold can maintain stability, even with changes in the variable regions. The binding of ySUMO Affimers 15, 20 and 22 was measured via isothermal calorimetry (ITC) and found to be 29.6, 33.3 and 47 nM respectively, demonstrating that Affimers can be developed with high binding affinity to a target. To show the selectivity and lack of cross-reactivity the ySUMO Affimers were tested for their binding to human SUMO (hSUMO). Analysis of their affinities showed that the ySUMO Affimers had a Kd 10000 fold higher when binding to hSUMO [10].

Affimers have been shown to be effective in the creation of novel biosensors, in 2018 a biosensor platform which can detect Her4 (a tumour biomarker) in serum was developed. A Her4 affimer was attached to a gold microarray able to detect Her4 with a better sensitivity compared to antibodies and with a better dynamic range. The sensor also didn't respond to non-specific serum proteins, so is less likely to produce a false positive [44]. Affimers have also shown to have possible uses in the treatment of arthritis. Two specific Affimers were developed which inhibited the binding of IgG to

the FcγRIIIa receptor preventing the release of tumour necrosis factor, reducing the inflammatory response. Importantly the Affimers didn't bind to the similar FcγRIIIb receptor one of few monospecific agents developed [45].

Affimers also have uses in research. Research into ubiquitination has been impeded by a lack of linkage-specific antibodies. Affimers were developed which were K6 and K33 linkage specific, this lead to the identification of E3 ligases, via Western blotting in combination with mass spectrometry, which assembled K6 and K33 linkages [46]. Affimers have also been utilised in fluorescent imaging, first demonstrated by creating GFP fused Affimers specific for F-actin in both fixed and live cells [47]. Affimers smaller size also allowed them to image densely packed regions not stained by antibodies [48].

Importantly, Affimers have been developed which specifically target SH2 domains, e.g in growth factor receptor bound protein 2 (Grb2). Affimers only bound to Grb2 and not Grb7, 14 or 10 and were able to disrupt protein-protein interactions[43]. It is their high specificity that makes Affimers an ideal candidate to study signalling pathways

1.1.4 Methods of library generation and isolation

The success of any method of affinity generation is reliant on the ability to generate a competent DNA library, the original Affimer library was generated using splice overlap extension (SOE) [10]. This method has the advantage of not requiring restriction sites [49].

Central to the use and design of antibody scaffolds is the ability to isolate them after selecting for specific binders. The most common technique used to achieve this is phage display (see Figure 4). First developed in 1989 by George Smith [50] it allows proteins to be displayed on the surface of a phage virus allowing it to bind to immobilised targets. Unbound phage can be washed away and bound phage can be used to infect *E.coli* amplifying the gene coding for a binding protein. Washing steps are often repeated multiple times to improve binding affinity [51]. This technology allows the screening of millions of potential proteins relatively easily and provides an easy way of obtaining the gene that codes for them. Phage display is also used in negative screens which allow the antibody scaffold to be selected for that does not bind a specific target by immobilising the target and collecting scaffolds that do not bind.

A less commonly used method is ribosome display in which fully folded proteins are prevented from leaving the ribosome due to the lack of stop codons, this allows the mRNA protein ribosome complex to be isolated by affinity selection. The mRNA can then be analysed to determine the sequence of the bound protein [52]. Ribosome display is able to take advantage of a larger library (10^{14}) compared to phage display (10^{11}) [53]. Despite this advantage phage display remains the most widespread method to generate high affinity proteins.

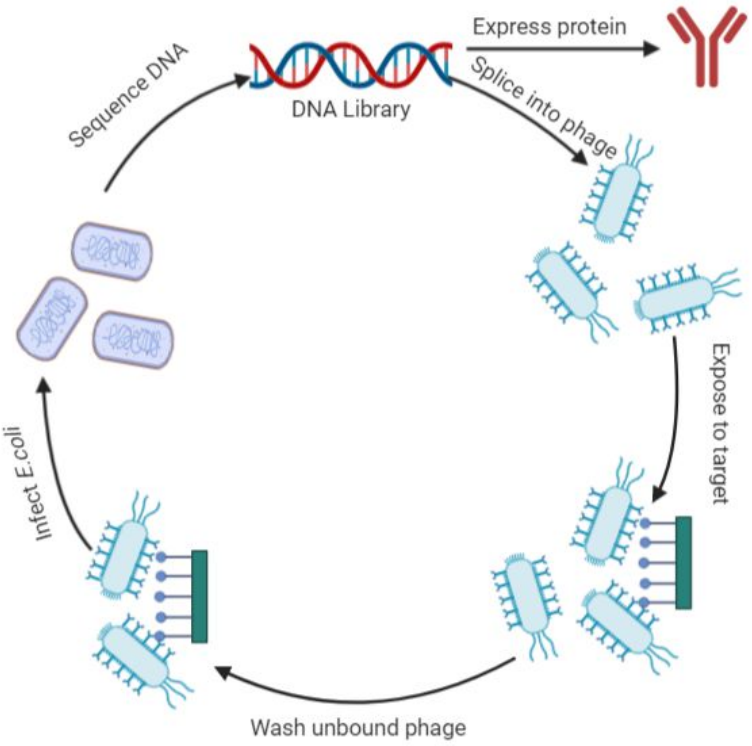


Figure 4. Process diagram of phage display used to create high affinity binding molecules. Created using BioRender

1.2 The MAPK pathway

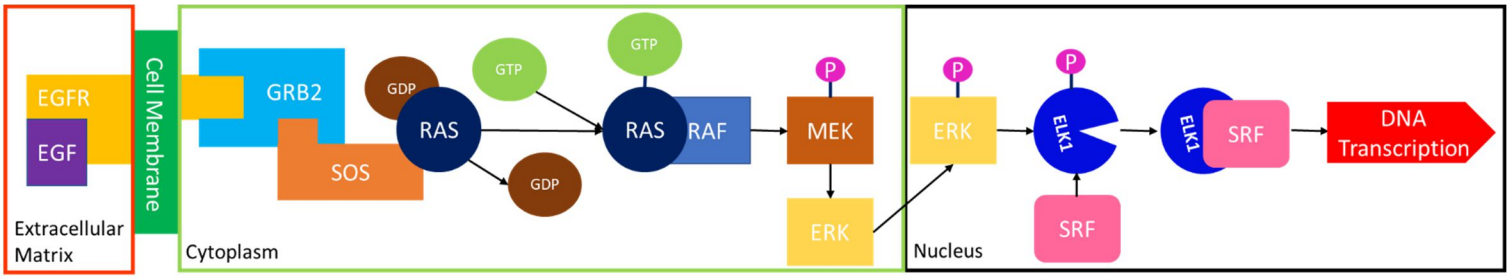


Figure 5. Simplified RAS/ERK pathway, EGF binding to the EGFR activates tyrosine kinase activity and Grb2 in complex with SOS binds to the phosphorylated tyrosines of EGFR which activates SOS allowing Ras-GDP to exchange its GDP for a GTP. RAS-GTP phosphorylates RAF which causes a kinase cascade leading to the phosphorylation of ERK via MEK. Phosphorylated ERK localises to the nucleus where it phosphorylates ELK1, ELK1 can then form a complex with SRF which acts as a transcription factor.
Abbreviations: Epidermal growth factor (EGF), Epidermal Growth Factor Receptor (EGFR), Growth Factor Receptor Bound Protein 2 (GRB2), Son of Sevenless (SOS), Guanosine Triphosphate (GTP), Guanosine Diphosphate (GDP) Mitogen activated protein kinase kinase (MEK), Extracellular Receptor Kinase (ERK), ETS like protein 1 (ELK1), Serum Response Factor (SRF)

Cells need a way to respond to changes in the extracellular environment, this is achieved by using specific cell signalling pathways. The MAPK pathway has a role in multiple processes within the body, including cell proliferation, the cardiovascular system and the immunological response [54], [55]. MAPK pathways are activated by specific extracellular signals and end in the activation of a MAPK protein which phosphorylates multiple transcription factors to drive change within the cell. There are three main MAPK families: Jun-amino terminal kinases (JNK), extracellular signal regulated kinases (ERK) and p38/stress activated protein kinases (SAPK). All MAPKs are activated by MAPK kinases (MAPKK) which are themselves activated by MAPK kinase kinases (MAPKKK) [56].

While there are multiple MAPK pathways this study focuses on the Ras/ERK pathway. The Ras/ERK pathway (see Figure 5) was first described in 1980 in Drosophila [57]. Since then it has been implicated in a significant number of cancers via constitutive activation of the receptors that stimulate the pathway and mutations of key regulatory proteins, such as Ras which is mutated in one third all of cancers (90% of which are K-Ras) [58]. This is due to the pathways control over cell proliferation, as constitutive activation will drive constant cell division. Multiple approved therapies now target the Ras/ERK pathway to help treat cancer [59]. For example Tyrosine kinase inhibitors

(TKI) that target EGFR are used to treat lung cancer, however the cancer is usually able to progress due acquired resistance to the TKI [60]. There is a lack of ERK inhibitor drugs compared to MEK. ERK inhibitors are becoming of particular interest due to their ability to overcome MEK inhibition resistance [61]. This highlights the importance of the MAPK pathway in oncology.

1.3 Protein-protein interactions

Protein-protein interactions (PPI) are highly specific contact between two or more proteins, they can be transient or more permanent, and are essential to biological processes. These interactions are often mediated by hydrogen bonds as well as hydrophobic and electrostatic effects. In the past they were often considered “undruggable” [62]. Since then improvements to characterisation of interaction domains, improved screening methods and the development of custom antibody scaffolds have made what previously seemed impossible a reality. Drugs such as Tirofiban, used in the treatment of myocardial infarction, inhibit the protein interactions of glycoprotein IIb/IIIa [63]. Maraviroc, used in the treatment of HIV inhibits the interactions of the CCR5 receptor preventing viral entry [64]. These examples and many others show that the inhibition of PPIs is now a viable option in the development of therapeutics.

One of the challenges that still exists when investigating PPIs is the frequent ubiquity of many protein interaction domains, their ability to interact with multiple proteins makes them biologically useful but also makes them hard to target. Often inhibitors will prevent the interaction between multiple proteins with the same domain. This provides a problem in research as it makes it harder to observe the effect of inhibition of a specific pathway and in therapeutics it increases the likelihood of side effects. Antibody scaffolds provide a solution to this, Affimers particularly are able to target domains in specific proteins with limited cross reactivity [10].

1.4 Growth factor receptor bound protein 2

Figure 6. The crystal structure of the Grb2 dimer, SH2 domains in white and SH3

Growth factor receptor bound protein 2 (Grb2) is an important adaptor protein in the MAPK pathway. It is 25 KDa and formed of two SH3 domains with a central SH2 domain, and is found in both monomeric and dimeric states[65] (Figure 6 [66]). Grb2 is the only protein in the RAS/ERK pathway that contains an SH2 domain. Grb2 has no intrinsic enzymatic activity, instead acting as an adaptor protein forming complexes with a variety of proteins. In the Ras/ERK pathway Grb2 is constitutively associated with the Son of Sevenless (SOS) protein, via Grb2's SH3 domains [67] and a proline rich sequence in SOS [68]. SOS is a guanine nucleotide exchange factor (GEF) that allows the exchange of GDP for GTP in Ras [69]. When EGFR is bound by EGF it undergoes autophosphorylation. Allowing Grb2 to bind to EGFR via Grb2's SH2 domain, which activates SOS GEF activity continuing the Ras/ERK signalling pathway [70].

Grb2 is unable to function when in its dimeric state but phosphorylation of Tyr160 (on Grb2) promotes the monomeric state in which Grb2 is active. This allows Grb2 to function as a control point in the Ras/ERK pathway as well as an adaptor protein. Control at Grb2 may be due to the background level of RTK phosphorylation in the absence of ligand, ensuring that the pathway is only activated when appropriate. This behaviour is important in preventing tumour development [71].

1.4.1 Grb2's General Role in Cancer

Grb2 has an obvious role in proliferative growth during cancer development as part of the MAPK pathway in controlling cell differentiation [72]. However, it has also been implicated in metastasis. Targeting metastasis is especially important as it is responsible for the majority of deaths due to cancer and studies using small molecule inhibitors of Grb2 have demonstrated reduced cell motility [73]. To achieve metastasis the cancer cell must be able to remodel the actin cytoskeleton, Grb2 has been shown

to interact with a range of proteins that associate with the actin cytoskeleton [74]. Hepatocyte Growth Factor (HGF) is a scatter factor that interacts with the receptor Met (a tyrosine kinase) to induce scattering [75]. During scattering cells dissociate from each other and the extracellular matrix (ECM) by disrupting E-cadherin mediated cell-cell adhesion [76]. Grb2 is able to interact with pY1356 in c-Met (sometimes via mediator Gab1), this leads to downstream signalling in multiple pathways which control both cell division and the regulation of other cytoskeletal associated proteins e.g FAK [77]. It is possible that inhibition of Grb2 interactions prevents this and reduces the motile abilities of the cell.

Another important hallmark of cancer is sustained angiogenesis which supplies the tumour with a constant source of oxygen and other important factors [78]. Studies have shown that inhibition of the SH2 domains in Grb2 reduce morphogenic events needed for angiogenesis in tumours [79]. This is because Grb2 is involved in multiple pathways which control angiogenesis [74].

Grb2 has also been a suggested target in breast cancer treatment. In all subtypes of breast cancer, the epidermal growth factor receptor (EGFR) is upregulated, this is most common in triple negative breast cancer (TNBC) which also lacks the estrogen receptor and HER2 [80]. Increased EGFR expression is also associated with lower rates of survival [81]. Multiple studies have attempted to develop therapeutically effective EGFR inhibitors, however they have only been successful in treating a small number of patients [82]. This highlights the need to develop different inhibitors that could potentially treat TNBC, such as ones which target Grb2. Grb2 has been shown to be upregulated in various breast cancer cell lines (MCF-7, MDA-MB-361 and -453), which likely increases their sensitivity to growth factors [83]. A variety of studies have also developed Grb2 inhibitors using peptidomimetics in which the SH2 domain of Grb2 was mimicked and modified, the resulting peptide was shown to inhibit EGF stimulation *in vitro* [84], [85]. However, none of these inhibitors were taken further into clinical trials.

1.5 The SH2 domain

The Src homology 2 domain (SH2 domain) is a common domain found in around 110 proteins in the human genome, the largest class of pTyr recognition domains[86]. It was previously thought to only exist in eukaryotic cells, however new research has

identified 93 SH2 domains In Legionella bacteria, distinct from those observed in eukaryotes [87]. SH2 domains are most frequently utilised during phosphotyrosine signalling in important processes such as cell growth and differentiation [88]–[90].The N-terminal SH2 domain followed by a kinase domain is highly conserved in all cytoplasmic tyrosine kinase family members [91]. The SH2 domain is usually found with other additional domains, forming N-terminal flanking regions. Most commonly this is an SH3 domain, but can also be FERM or F-Bar domains[92], in JAK kinases the SH2 domain is followed by two kinase domains [93]

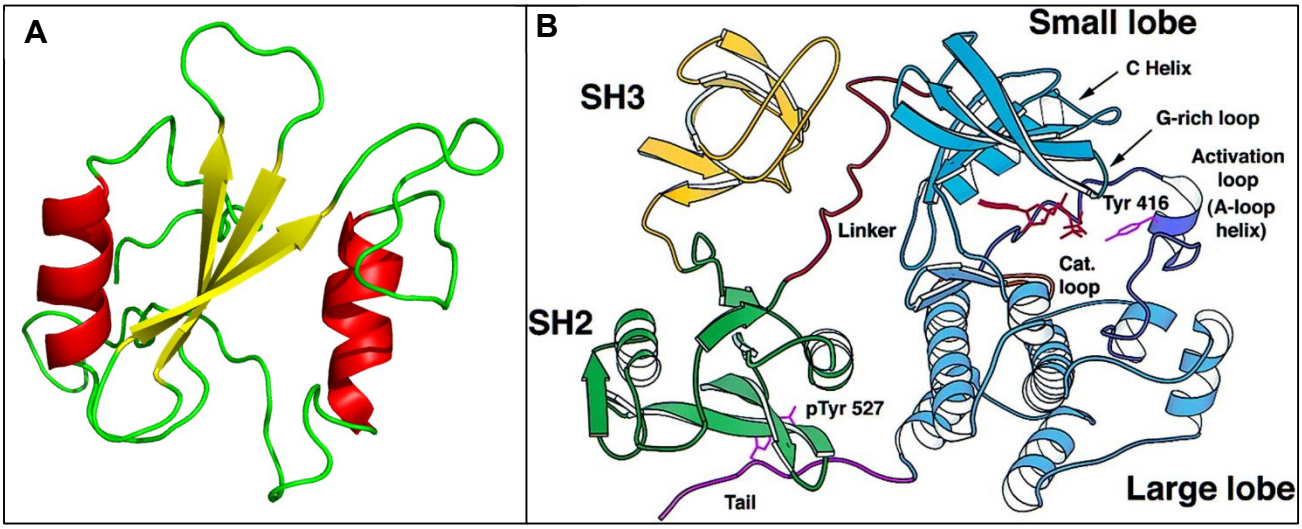


Figure 7. (**A**)The solution NMR structure of the SH2 domain of Grb2 at 1.09 Å. β sheets are in yellow, α helices in red. (**B**) Ribbon diagram of c-src in the inactive state in which pTyr 527 has bound to the SH2 domain and the A-loop helix has formed. SH2 and SH3 domains are in green and yellow respectively, the kinase domain is blue, and the phosphorylated tail section is in purple

SH2 domains are typically 100 amino acids long, comprised of one β sheet with α helices on each side (see Figure 7A [94]) [95], [96]. The SH2 domain is used to regulate the activity of tyrosine kinases, the mechanism for which has been determined by a number of crystallographic and molecular dynamics studies. In c-Src, phosphorylation of Tyr527 (located on the C-terminal tail) causes it to bind to the SH2 domain forming a compact inhibited state[97] (see Figure 7B [98]) this conformation is also stabilised by interaction with SH3 domains [99], [100]. In this state the SH2 and SH3 domains stabilise a loop helix which removes a key residue from the active site[95]. Dephosphorylation of Tyr527 converts the kinase into an active state and allows the subsequent phosphorylation of Tyr416 and the helix to adopt a disordered but catalytically active conformation[98], [101].

1.5.1 SH2 domains in protein-protein interactions

SH2 domains have long been known to control signal transduction in cell signalling pathways through the recognition of phosphotyrosine residues [102]. The majority of SH2 domains contain a conserved arginine residue (R175 in v-Src) which is essential for recognition of phosphotyrosine [103]. The ubiquity of SH2 domains in cell signalling poses an important question, how do SH2 domains recognise specific proteins in a pathway without interacting with other pY residues? A recent study used phosphotyrosine oligopeptides (with pY in the middle) to screen for the specificity of 70 SH2 domains in 70 different proteins. They identified 17 different classes of SH2 domain which preferentially bound to a specific pY containing sequence on the target protein [104]. The Grb2 SH2 domain was found to bind to sequences with an asparagine residue 2 residues from pY. On EGFR the Grb2 SH2 domain can bind to pY1173 [105] which has an asparagine residue only 3 residues from pY. Other studies have also identified secondary contacts which play a role in SH2 complex formation, in phospholipase-Cγ and FGFR1 secondary contacts are estimated to account for 20% of total binding affinity [106]. These studies demonstrate that the target sequence on an SH2 domain's ligand is vital in determining its specificity.

1.5.2 The SH2 Domain in Grb2

The SH2 domain of Grb2 associates with EGFR after the receptor is phosphorylated, triggered by EGF binding. EGFR has multiple possible phosphorylation sites. Using deletion mutations it has been determined that Grb2 binds preferentially to pY1068 and pY1173 [107] with Y1068 more commonly phosphorylated *in vivo* [108]. It has also been demonstrated that Grb2 has a higher affinity for EGFR compared to other SH2 containing binding proteins [109]. In addition there is little difference in the binding kinetics of Grb2 and the isolated Grb2 SH2 domain, which suggests it is heavily reliant on the interaction between phosphotyrosine and the SH2 domain, not influenced by other residues [110].

In vivo fluorescent imaging has revealed that Grb2 takes around 5 minutes to reach 50% EGFR binding on the membrane and 10 minutes to reach equilibrium. This is in contrast to far western blotting which shows that SH2 binding sites become available much faster in around 1 minute [109]. One suggestion as to what causes this behaviour is that clustering of the EGFR increases membrane dwell time and therefore

the time taken for EGFR to cluster upon EGF activation increases how long it takes for Grb2 to reach equilibrium on the membrane [109], [110].

1.5.3 SH2 Domains in Disease

Due to their use in a wide range of signalling pathways, mutations in SH2 domains are associated with multiple diseases [111]. One well studied example is x-linked agammaglobulinemia (XLA) first described in 1952 [112], it is an x chromosome linked disorder that causes an increased likelihood of bacterial infection, hypergammaglobulinemia and reduced circulating B cells [113]. Most XLA patients die from bronchiectasis or chronic lung disease due to repeat infections [114]. XLA is caused by mutations in Bruton's tyrosine kinase (BTK), many of which are located in the SH2 domain, either affecting its stability or function [111], [115], [116]. Due to the lack of BTK signalling B cells fail to mature within the bone marrow and do not enter circulation [117]. Currently the only treatment for XLA is immunoglobulin (IG) therapy in which intravenous (IV) injections of IG are given to the patient every few weeks [118]. This therapy has multiple issues including the lack of IgA and IgM in serum injections, as it is possible to develop anti-IgA antibodies [119], this leaves mucosal surfaces (such as those in the lung) vulnerable to infection [120]. Therapy also lasts for life, significantly reducing quality of life for the patient and does not replace other B cell functions (only the production of antibodies) [118]. It is possible that with a further understanding of SH2 domains a more effective treatment could be developed.

1.6 Project aims

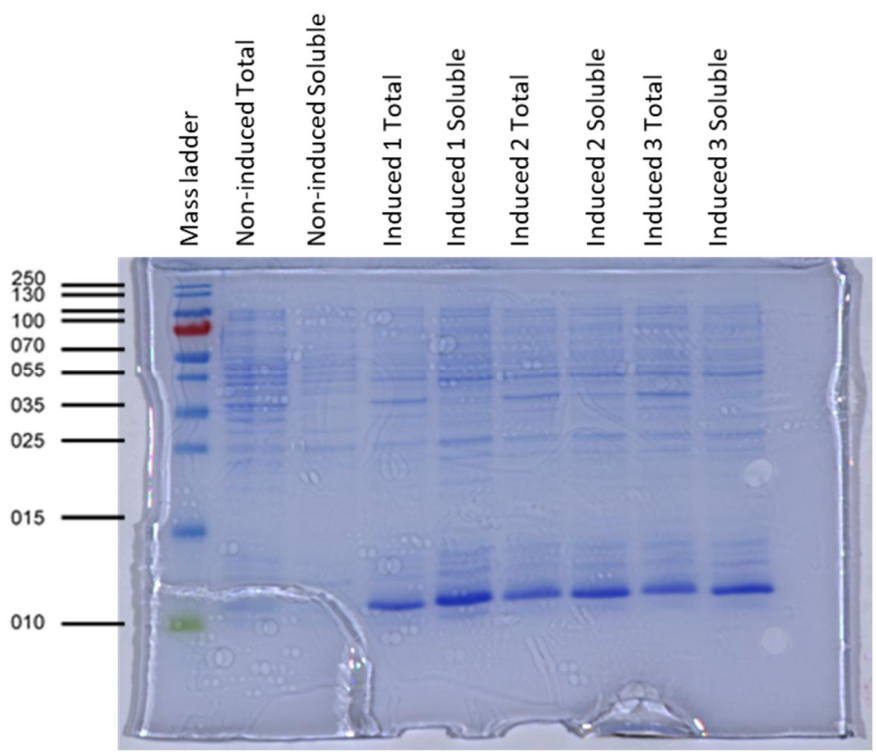
This project aimed to characterise the structural interaction between the SH2 domain of Grb2 and Affimers that inhibit its activity, using X-ray crystallography. Another objective was to use Affimers to target SH2 domains within specific proteins to determine if they have previously unknown activity within the MAPK pathway. Grb2 was used to develop an assay in which an Affimer disrupts the protein-protein interactions of the SH2 domain and the effect on the MAPK pathway is observed. This assay was then applied to Affimers which inhibit protein interactions of other SH2 domain containing proteins. The further characterisation of the MAPK pathway could lead to potential identification of novel targets in cancer therapy. An ideal outcome would be to identify a protein that has increased activity within the MAPK pathway in cancer cell lines compared to healthy cells

Chapter 2 Results

2.1 Crystallisation of the Grb2-SH2 Affimer complex

Prior work in the Tomlinson lab produced a range of Affimers which interacted with the SH2 domains of various proteins, including the Grb2 SH2 domain. A co-crystal structure between the Grb2 SH2 domain and an Affimer would provide insights as to how Affimers interact with and block the activity of Grb2 SH2 domains. Previous attempts in the Tomlinson lab had been made to crystallise the full Grb2 protein with an Affimer, which were unsuccessful likely due to the low concentration of Grb2 produced. An alternate approach was to express only the Grb2 SH2 domain and attempt to crystallise in complex with an Affimer.

2.1.1 Expression of the Grb2 SH2 domain

Temp (°C)	Expression
18°C	

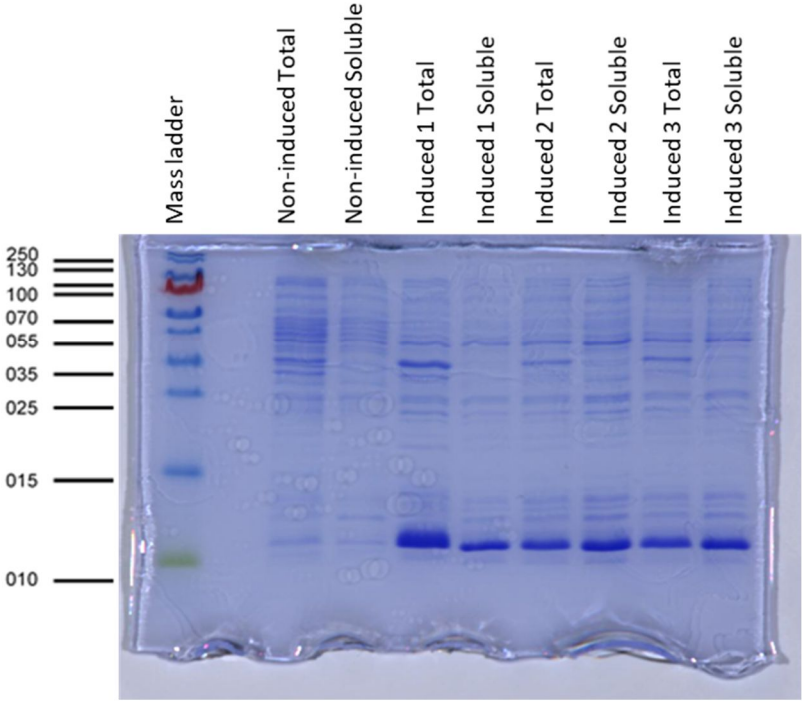
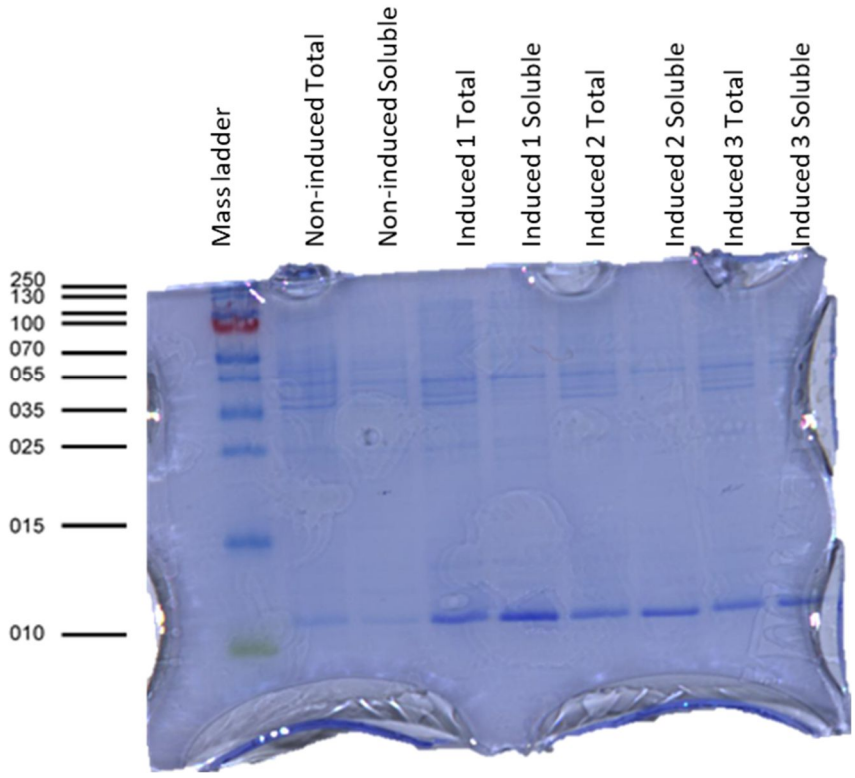
Temp (°C)	Expression
20°C	
22°C	

Figure.8 Showing 15% SDS PAGE gels of *E.coli* induced with IPTG to express SH2 domain. *E.coli* was grown in 50 mL cultures and samples were diluted 1 in 10 with SDS PAGE buffer before being run on the gel. The control had the SH2 domain containing plasimd but was not induced with IPTG. Expression was estimated by observing the size of the bands at around 10 kDa. Total fractions are

taken from samples of the whole cell lysate, soluble samples are taken from the soluble fraction of the cell lysate. Expression for each condition was repeated 3 times indicated by the number of the sample

To crystallise the SH2 Affimer complex high concentrations of both were required. Affimers already have a well-defined protocol and can easily be produced at high concentration in *E.coli*. To find the optimum conditions to produce the SH2 domain, expression trials were carried out in BL21 DE3 *E.coli* at three different temperatures 18, 20 and 22 °C. Expression was analysed by running the total and soluble lysates on an SDS PAGE gel, it was not possible to obtain an exact concentration of the SH2 domain as it was unlabelled.

Figure.8 shows that the 20 °C condition produced the largest bands at around 10 kDa (the mass of the SH2 domain) whereas 22 °C produced the thinnest bands suggesting it had the least expression of any of the conditions. The level of expression in all the conditions was consistent both in the total and soluble fractions as well as across the repeats. 20 °C was chosen as the temperature to express SH2 domain.

Figure.9 demonstrates that the SH2 domain could be produced at both 400 mL and 1.5 L culture volumes, simplifying large scale production.

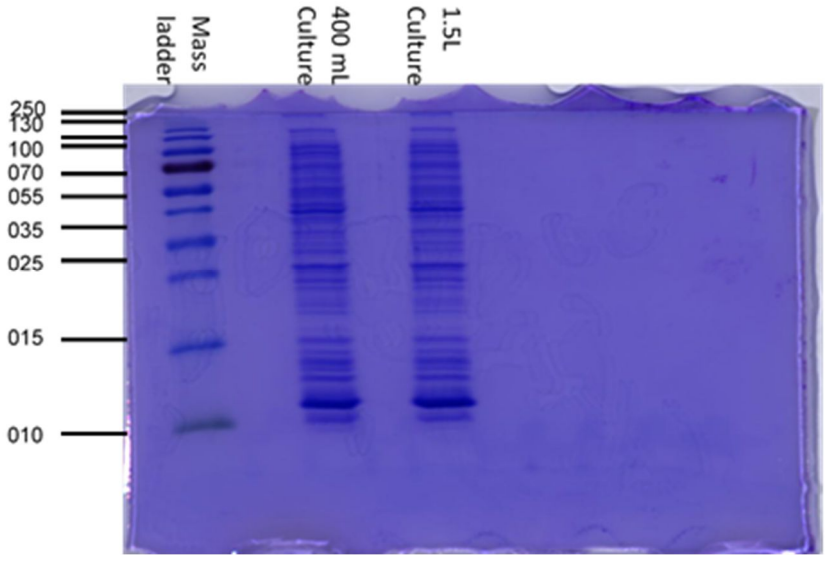


Figure.9 15% SDS PAGE gels of *E.coli* induced with IPTG and left overnight at 20 °C to express SH2 domain. *E.coli* was grown in 400 mL and 1.5 L cultures and gel samples were diluted to the same concentration

2.1.2 Modifying Affimers with an HRV 3C site

An HRV 3C site added into Affimers between the Affimer sequence and His tag would allow them to be cleaved from a Ni-NTA column using a HRV protease, this could potentially aid crystallisation. The Affimers D6 and F1 were chosen to be modified as they both showed good inhibition of the MAPK pathway in HEK293 cells (see Chapter 2.2), suggesting that they had a high binding affinity for Grb2 which would aid complex formation.

The Affimers were modified by PCR using custom primers (See methods 3.1.3) and addition of the site was confirmed by sequencing. Expression of the modified Affimers (referred to as D6H and F1H) was compared to expression of Affimers without the HRV 3C site by expressing the Affimers in *E.coli* and comparing on an SDS PAGE gel.

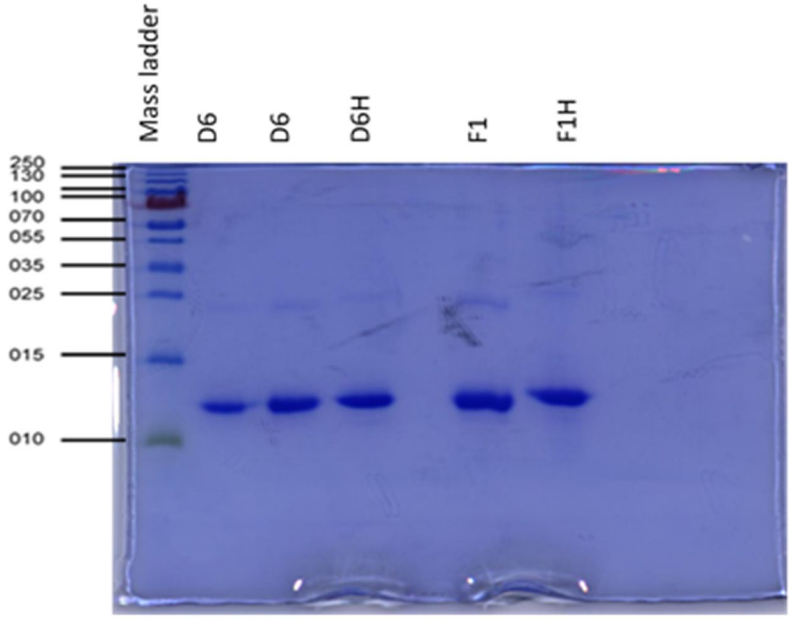


*Figure.10 15%*SDS PAGE comparing expression of Affimers with and without an HRV 3C site. All proteins were expressed and purified under the same conditions, samples were taken from purified fractions diluted the same and loaded onto gel. D6 was accidentally loaded twice.

Figure.10 shows that there is no difference between the expression of the Affimers with and without the HRV 3C site. The concentration of the purified samples was also measured using a Nanodrop and found to be 4-5 mg/ mL in all conditions.

2.1.3 Affimer:SH2 complex formation

Pull-down assays were performed using magnetic Ni-NTA beads to ensure that the Affimers could form a stable complex with the SH2 domain. Trials were then performed to determine the optimal ratio of Affimer to SH2 domain for complex formation.

Figure.11 shows that the pulldown was successful as all conditions produced both Affimer and SH2 domain. There was no notable difference between the Affimers with

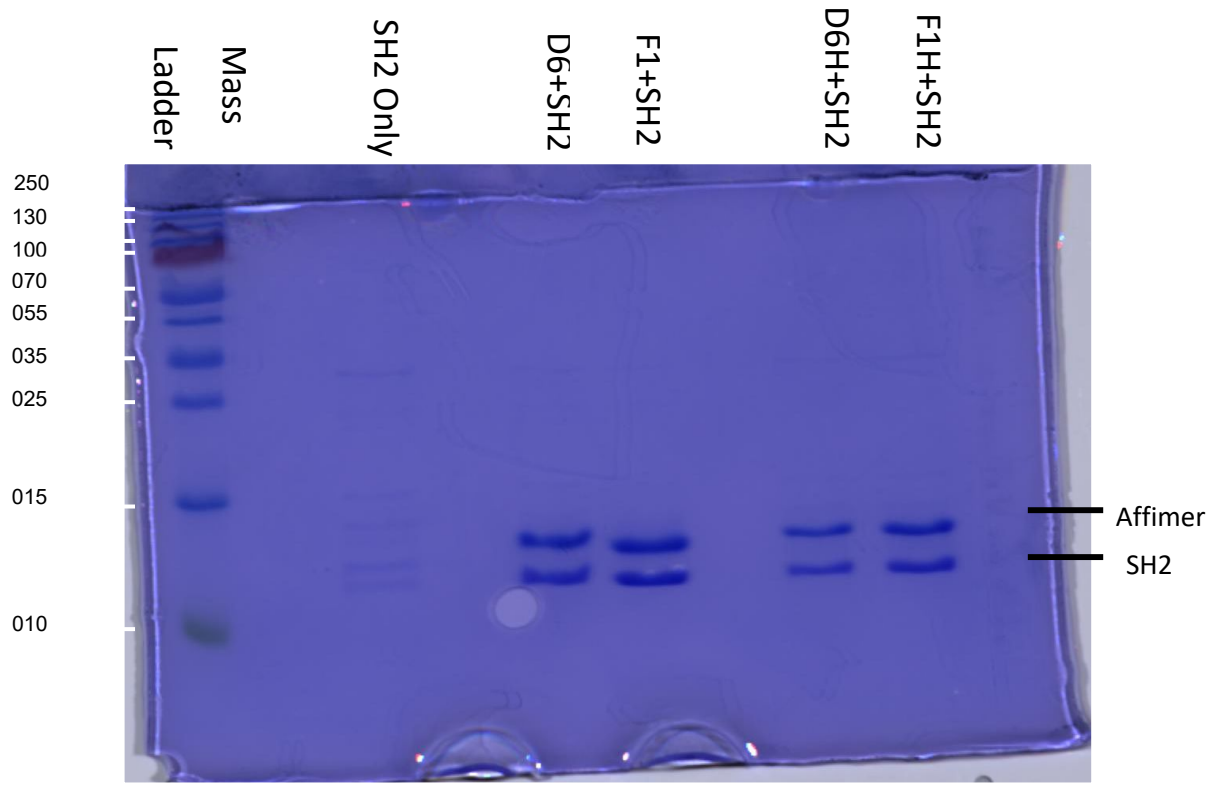


Figure.11 15% SDS PAGE gel showing results of pulldown on magnetic NI-NTA beads of Affimer in complex with SH2 domain. Samples were eluted from beads using a 100mM imidazole elution buffer and diluted before being run on a gel. SH2 only condition was run on beads but did not have any Affimer added. D6H and F1H Affimers have the added HRV 3C site.

and without the HRV 3C site as they pulled down the same amount of SH2 domain. The SH2 only condition also did not show a large band of SH2 domain, meaning it is not sticking to the beads on its own and the Affimer was able to form a stable complex with the SH2 domain. The F1 Affimer was then used in assays to determine the correct ratio of Affimer and SH2 lysate, the best ratio would have very little SH2 domain in the unbound fraction indicating that it has saturated the Affimer bound to beads.

It was found that the best binding was achieved with 1:5 or 1:10 F1 to SH2 domain lysate (see *figure.12*). Bound fractions are those eluted from the Ni-NTA beads using buffers with high concentrations of imidazole. Unbound fractions are taken from a sample of the Affimer/SH2 lysate after being mixed with the beads, representing the protein that was unable to bind to the beads. The unbound fractions of these two ratios both show a similar amount of F1 and SH2 domain. The bound fractions showed a significant increase in the amount of SH2 pulled down at the 1:10 and 1:5 ratios while the amount of Affimer did not change this suggests that the Affimer had become saturated by SH2 domain. Ratios lower than 1:5 or with more Affimer than SH2 were ineffective at forming the F1-SH2 complex. It was decided that the 1:10 ratio would be used in crystal trials moving forward as it was not hard to get large amounts of SH2 domain lysate and it ensured that the maximum amount of complex could form.

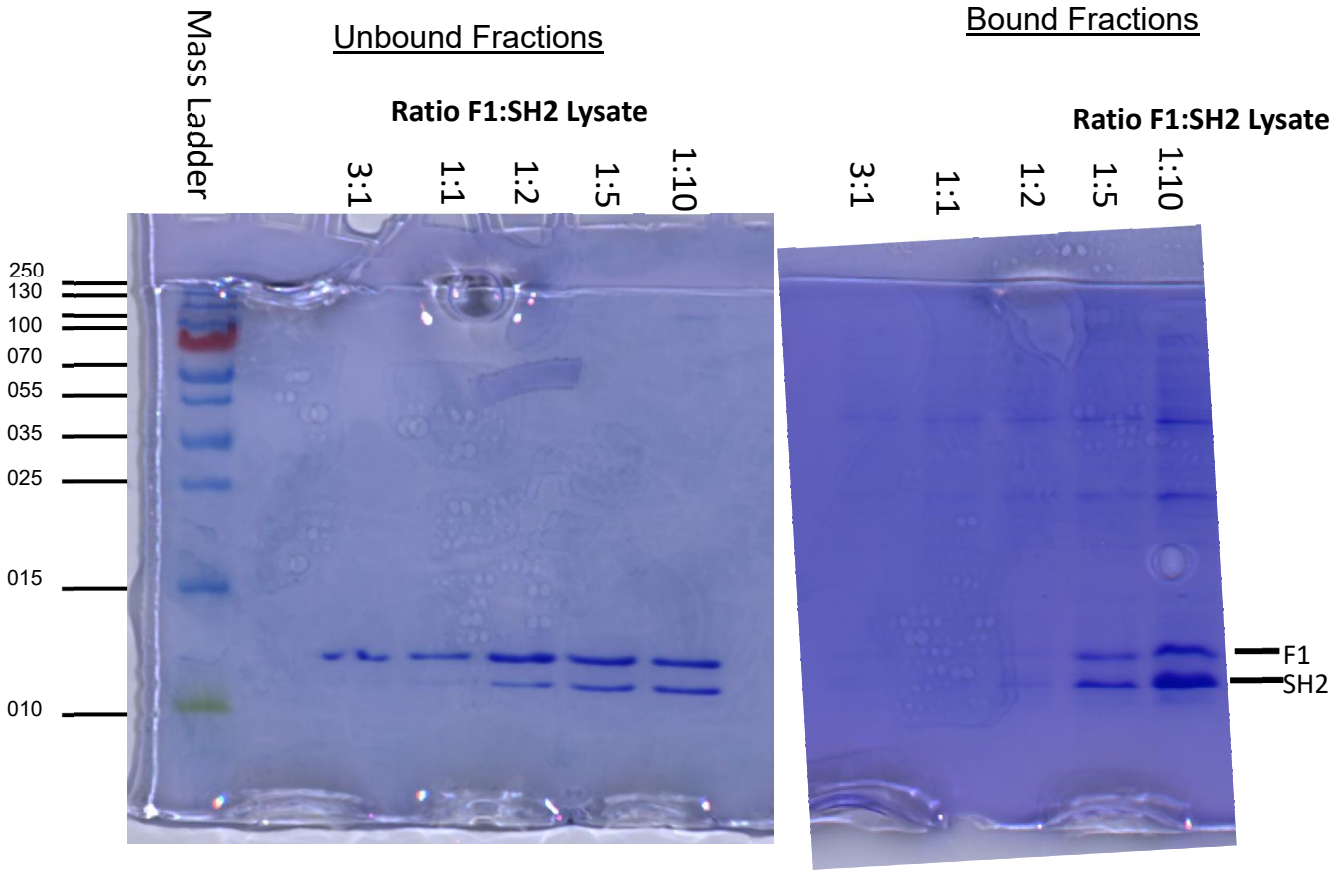


Figure.12 15% SDS PAGE gel showing bound and unbound fractions at different ratios of F1:SH2 domain lysate after being mixed with Ni-NTA beads. Bound fractions were eluted using a buffer with high concentration of imidazole to outcompete the His tagged Affimer. Unbound fractions are samples taken from the washing step (see 3.1.7)

2.1.4 Size exclusion chromatography

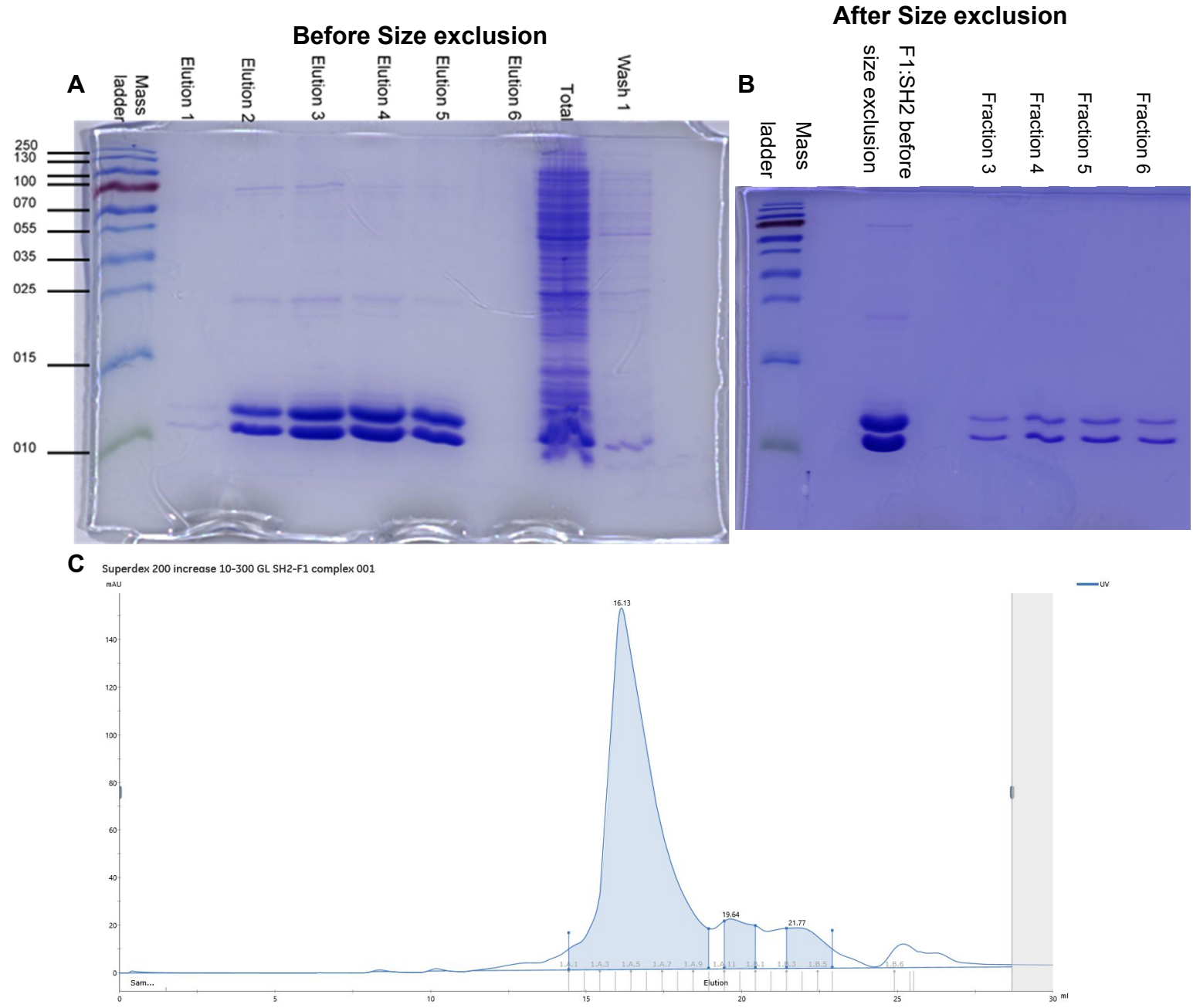


Figure.13 (**A**) 15% SDS PAGE gel of F1:SH2 complex after elution from a Ni-NTA column (**B**) fractions ran on a 15% SDS PAGE gel after being separated by size exclusion chromatography, fractions 3-6 are taken from the largest peak in the graph. The F1:SH2 complex is a sample of the complex before size exclusion (**C**) graph of size exclusion chromatography of the F1:SH2 complex.

A pure sample of complex is needed for crystallisation to occur and to reduce the probability of something other than the desired complex crystallising. Size exclusion chromatography ensures that the sample is not contaminated. A ratio of 1:10 Affimer SH2 lysate was purified using Ni-NTA beads then run on a size exclusion column.

The results showed that before size exclusion (*figure.13A)* there was some non-specific binding that was also eluted along with the F1:SH2 complex. After size exclusion (*figure.13B*) there is no longer any visible non-specific proteins on the gel, so the complex has been purified successfully. The size exclusion graph (*figure.13C)* showed one distinct peak which was the complex (as shown on the SDS PAGE of the fractions).

2.1.5 Crystal Trials

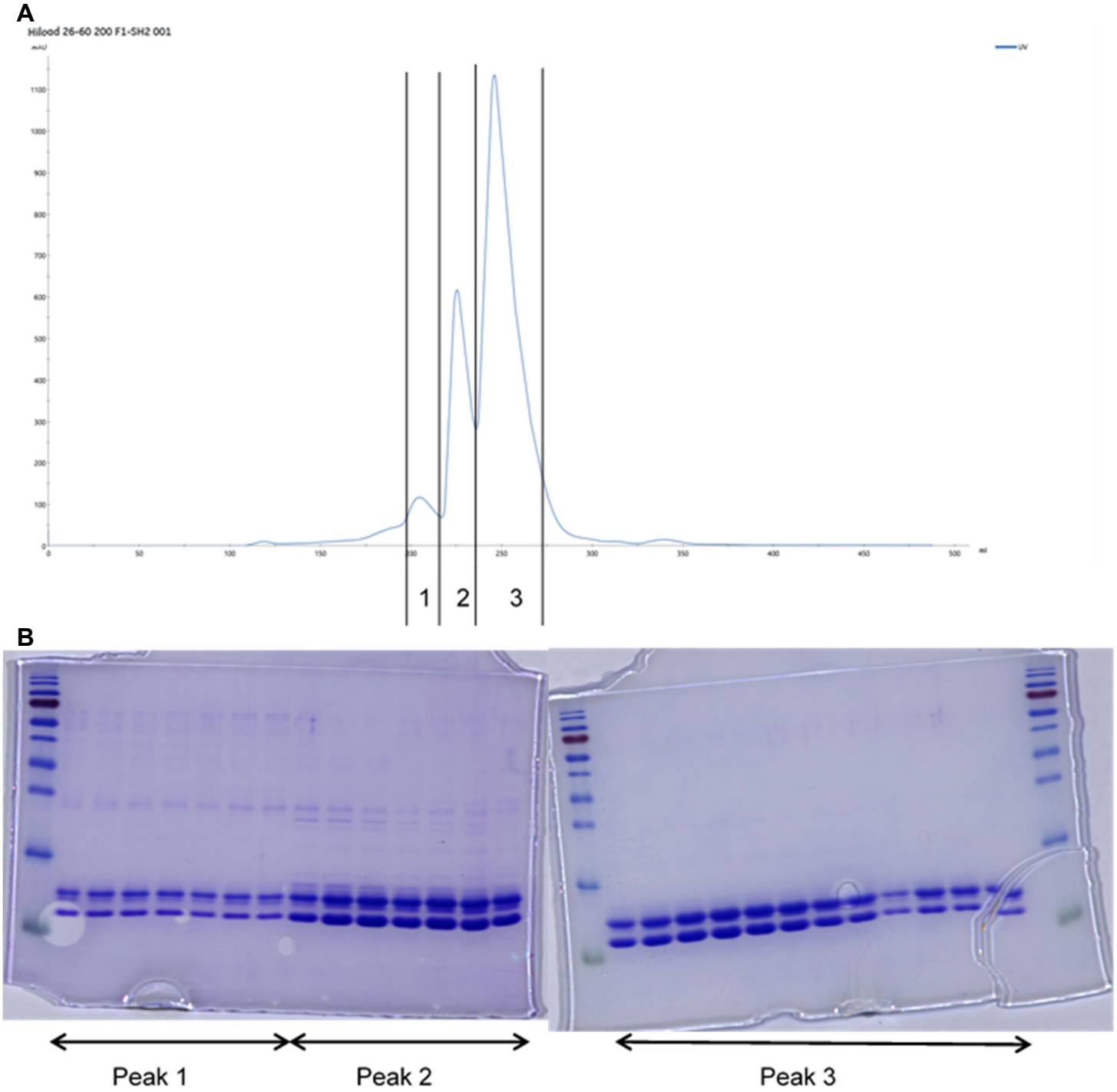


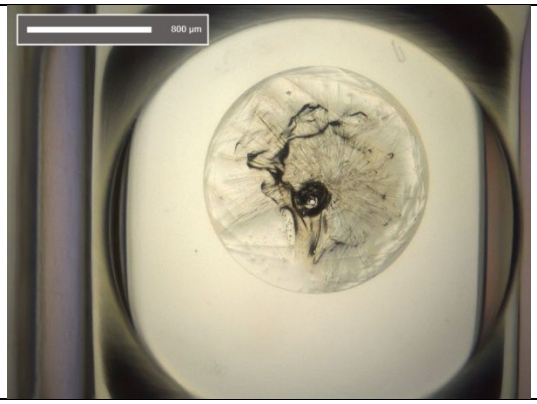
Figure.14 (**A**) Graph of size exclusion chromatography showing three main peaks (**B**) 15% SDS PAGE gel of size exclusion peaks, all peaks had F1:SH2 complex.

For the final trials 6 1.5L cultures of SH2 domain and 4 400 mL cultures of F1 were used to produce soluble lysate. In total 18.3 mL of F1 lysate was mixed with 183 mL of SH2 domain lysate and purified on a Ni-NTA column before being purified by size

exclusion. Figure14A shows that there were three main peaks eluted from the size exclusion column, when ran on a gel (*figure.14B*) it was found that all three peaks contained the F1:SH2 complex. The first peak contained the least complex with peaks 2 and 3 having the most. It is not clear why the complex formed three peaks instead of the one previously seen, peaks 1 and 2 do have more contaminants in them which may be the reason they came off of the column first. Peak 3 is largely clean with no non-specific proteins. Peaks 2 and 3 were concentrated and used in the crystal trials. Concentration of the complex was not measured due to lab shutdowns.

After purifying the complex crystal trials were set up using JSCG I, II, III, IV and Morpheus buffer blocks, also varying the protein to buffer ratio in each condition (either 1:1, 2:1, 1:2). The complex was then left in a crystal hotel to crystallise, while being monitored remotely. The complex was left for 130 days in total.

Table 2 Crystal producing conditions

Conditions	**Protein to Buffer Ratio**	**Image**
0.05 M MgCl2 (Salt) 0.1 M TRIS 8.5 pH (Buffer) 40 %v/v EtOH (Precipitant)	1:1	
0.2 M NaCl (Salt) 0.1 M Na Acetate 4.5 pH (Buffer) 40 %v/v PEG 300 (Precipitant)	1:1	

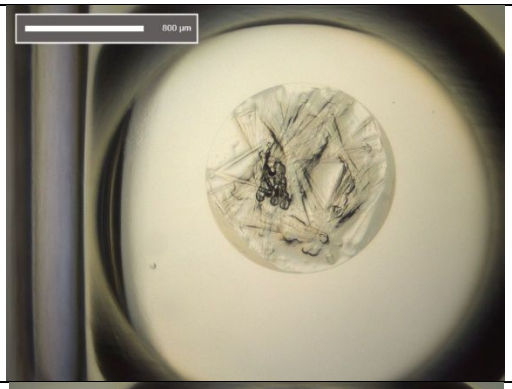
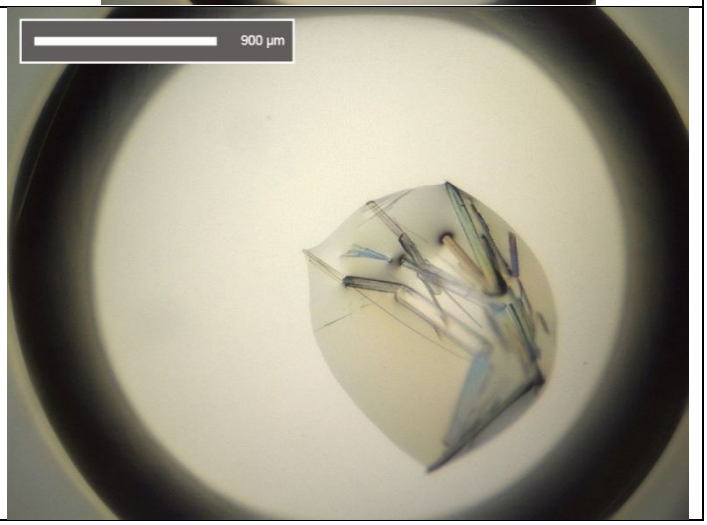
	2:1	
	1:2	
0.2 M LiSO4 (Salt) 0.1 M CAPS 10.5 pH (Buffer) 2 M (NH4)2SO4 (Precipitant)	1:2	

Table 2 Shows the conditions that were successful in producing potential crystals of the F1-SH2 complex. All images were taken 130 days after being plated.

These first crystal trials were successful in producing multiple conditions with crystals (see *Table 2*), there does not appear to be a pattern in the conditions that produced the crystals. Crystals were observed after 130 days as shutdowns made the lab inaccessible. All three of these conditions were taken forward into optimisation trials to try and obtain as much crystal material as possible.

2.1.6 Optimisation Trials

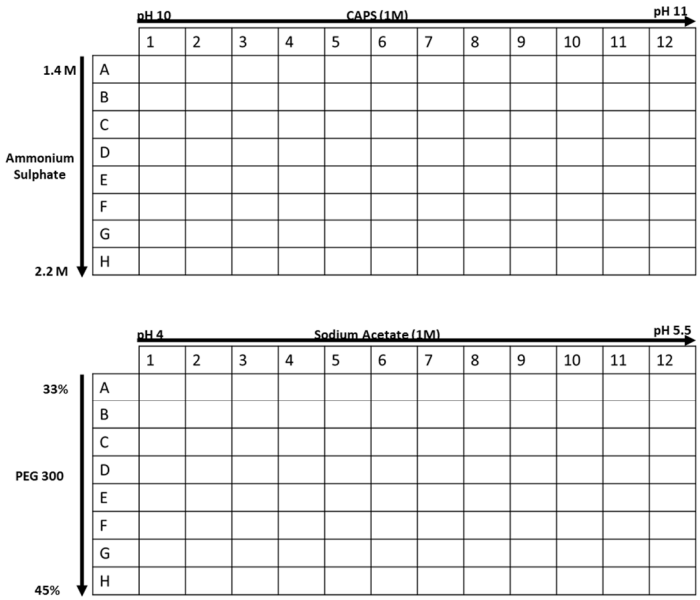


Figure.15 Layout of buffer solutions used in crystal optimisation trials. pH varied by columns, precipitant concentration varied by rows.

A new sample of F1SH2 complex was produced (same procedure as previously described) and used in the optimisation trials. Crystal plates were designed according to the maps in *figure 15,* varying the concentration of precipitant along the rows and the pH of buffer along the columns. The minimum and maximum (pH and concentration) values were chosen by altering the successful conditions positively and negatively to put the original conditions in the middle of the plate. The concentration of salt was kept constant in all conditions at the concentration used in the original trials.

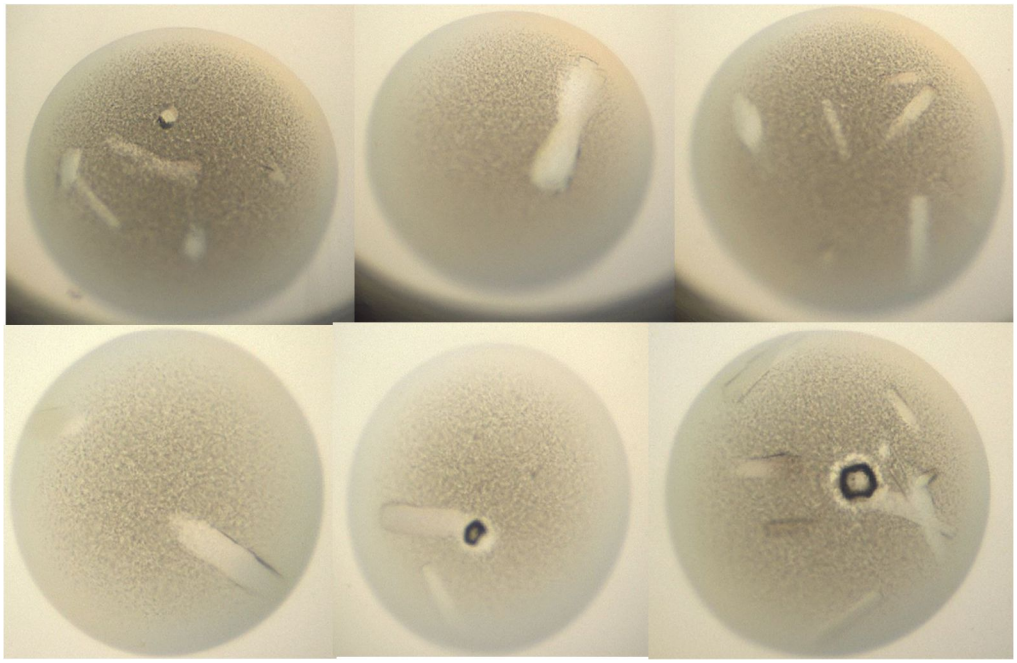


Figure.16 A selection of crystals grown in optimisation trials, images were taken after 21 days. Conditions were: 0.2 M NaCl, PEG 300, sodium acetate (varying concentrations)

The only condition to produce crystals again was the NaCl/PEG 300/Sodium acetate condition (see *figure.16*). This could be because the crystals have not been left for as long as they were originally, so they have not had time to grow. It is also possible that the two samples were at different enough concentrations for it to affect its ability to crystallise under certain conditions. The concentrations of the complex used in each trial was not obtained due to time constraints so this cannot be known for sure. The crystals that have grown appear to be of a good size and shape increasing the chance that they will be able to provide useful structural data. All of these crystals were sent to be diffracted.

2.1.7 Diffraction of F1-SH2 crystals

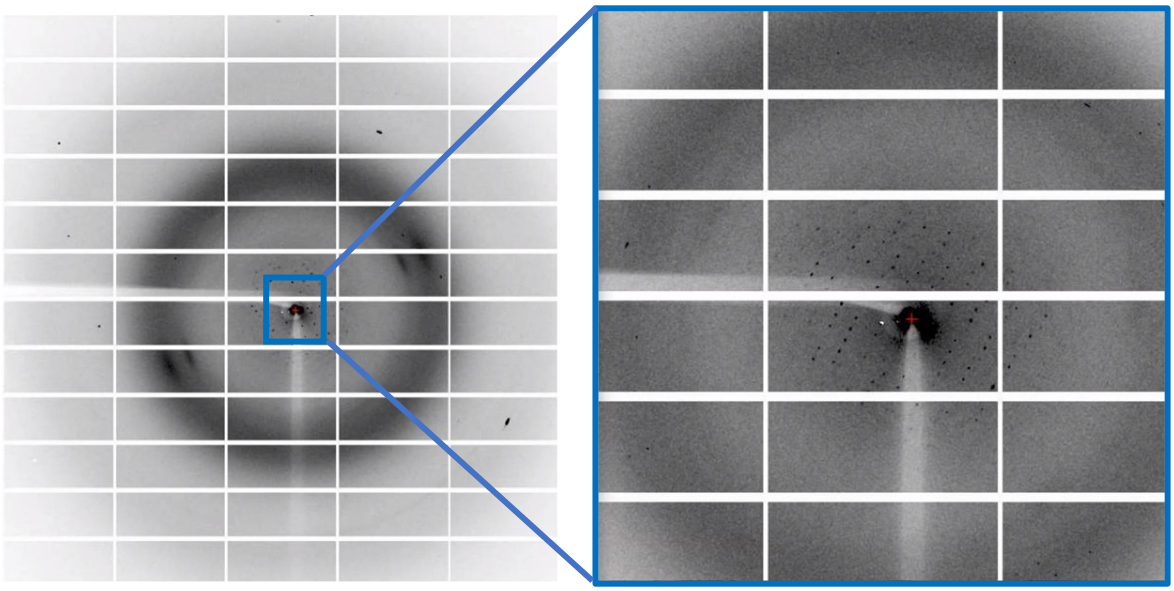


Figure.17 (**A**)Diffraction image of SH2-F1 complex crystal 22, this was the highest resolution diffraction image obtained, image taken at 0° at 100% beam intensity for 0.1 s (**B**) Zoomed in view of diffraction image showing middle with lattice

In total 32 crystals were sent to Diamond light source to be diffracted (diffraction and beam operation carried out by Dr Chi Trinh). Crystals were diffracted at 100% beam intensity for 0.1 s at 0°,45° and 90°. Diffraction images of all the crystals showed unstructured spots close to the centre, indicating that the crystals were protein but did not diffract well enough to produce a structure. The best resolution obtained was 6.5 Å, not enough to gain meaningful structural data, shown in *figure 17*. Diffraction patterns also varied at different angles indicating that the crystals were not singular. Patterns were also not good enough to accurately calculate the unit cell size.

These results show that while we were able to grow protein crystals, they were not of good enough quality to obtain structural data. However, they are a good start and further optimisation trials should be carried out to produce better crystals.

Potentially the reason the crystals did not diffract well is that the His tag on the Affimer made the structure too disordered. This could be solved by using the previously produced Affimers with an HRV 3C site that allows the His tag to be removed. These Affimers have already been trialled to ensure that they do form complex with the Grb2 domain, but it is not confirmed that they inhibit its activity. As the HRV site does not affect the variable domains it is unlikely that the modification affects their behaviour. It should also be considered that this method would likely reduce the amount of complex purified which could affect crystallisation. This approach was not trialled in this study due to time constraints.

2.2 Effect of SH2 targeted Affimers on the MAPK pathway

The only known SH2 domain containing protein to have activity in the MAPK pathway is Grb2. This project aimed to determine whether any other SH2 containing proteins had previously unknown activity in the MAPK pathway by using Affimers which blocked protein-protein interactions. Affimers are particularly well suited to this task due to their high level of specificity, only targeting the SH2 domain of one protein. This limited amount of cross interactivity allows a specific protein to be identified with a good degree of confidence.

In total 32 different proteins were targeted by 119 different Affimers (Affimers produced by previous work in the lab). If any of these proteins were involved in the MAPK pathway then the Affimers should block their interactions and alter the activity of the pathway.

2.2.1 Luciferase assay time trial

To measure the effect on the MAPK pathway it must be stimulated with epidermal growth factor (EGF), therefore a time course was initially performed to determine the optimum time to stimulate the pathway. An alanine Affimer was used as a negative control, all the residues in the hypervariable regions were alanine which should prevent any interactions with other proteins. The Grb2 D6 Affimer acts a positive control as Grb2 has known activity in the MAPK pathway and the D6 Affimer is known to block Grb2 interactions. A second positive control is the K6 Affimer which interacts and blocks the activity of Ras, this provides a comparison as to what reduced activity in the MAPK pathway looks like.

To measure the activity of the MAPK pathway a Luciferase assay was used. HEK293 cells were transfected with two plasmids, one plasmid contained a constitutively expressed Renilla luciferase the other a Firefly luciferase expressed in the presence of serum response factor. Each of these luciferase enzymes produces light in the presence of different chemicals allowing their signals to be measured independently. Constitutively expressed Renilla acts as a transfection/transcription control factor to ensure that the level of transfection in a cell is not affecting the results.

The Firefly luciferase is used to measure the activity of the MAPK pathway which (among other effects) produces serum response factor. A ratio of Renilla and Firefly

luminescence therefore acts as a measure of the activity of the MAPK pathway. The higher the ratio the more active the pathway.

The Affimers K6 and D6 were used in this time course. K6 is an Affimer that blocks the activity of Ras (the upstream of Grb2) and D6 is a Grb2 SH2 domain targeted Affimer. As shown in *figure 18* in the presence of the alanine Affimer the activity of the MAPK pathway increased over time, reaching a peak at 8 hours. It is not clear whether this was the limit or whether activity would have continued to increase had the stimulation continued for longer. From this we can conclude that the Affimer alone has no effect on the MAPK pathway and any changes can only be from specific interactions between protein domains. The D6 and K6 Affimers both showed a similar reduction in MAPK activity which did not change by a large amount compared to the unstimulated condition. From these results it was concluded that 6 hours was the best stimulation time as the pathway had reached a significant level of activity in the negative control condition

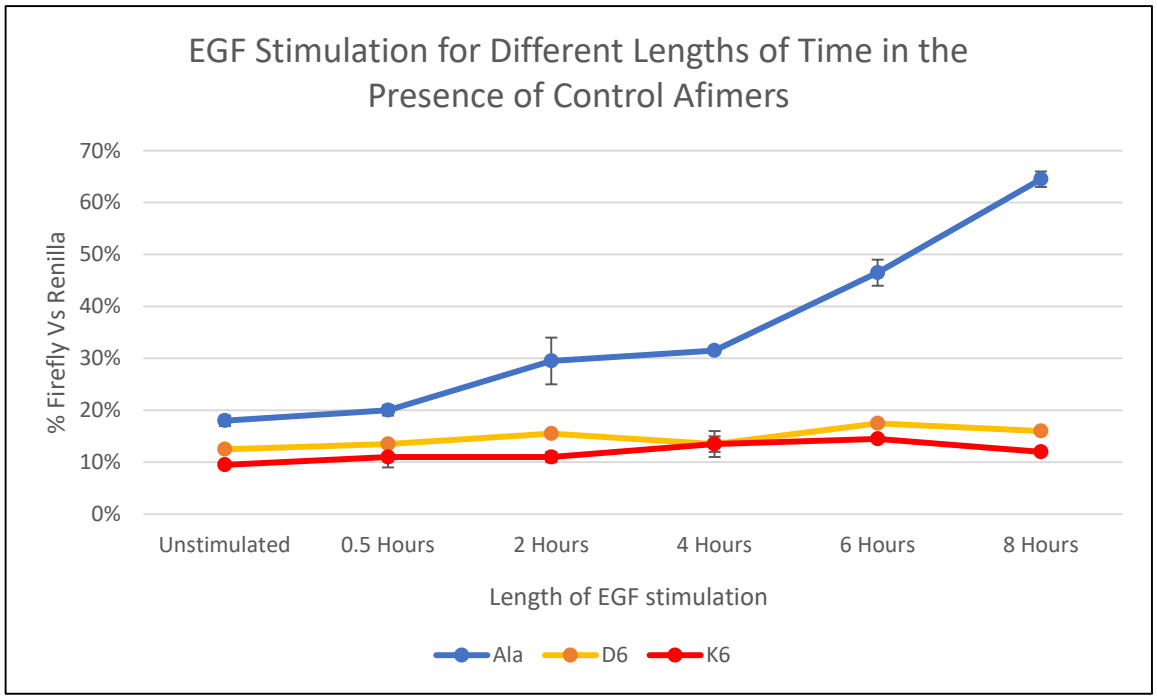


Figure.18 Graph showing the relationship between the length of EGF stimulation and the ratio of Firefly and Renilla luminescence, in the presence of alanine, D6 and K6 Affimers. Ratio values are N=3 averages. The alanine Affimer is a negative control, D6 is a Grb2 positive control, K6 is a Ras positive control. Error bars show standard deviation. The unstimulated condition did not have any EGF added to the cells

2.2.2 Trial of robust Z scores statistical test

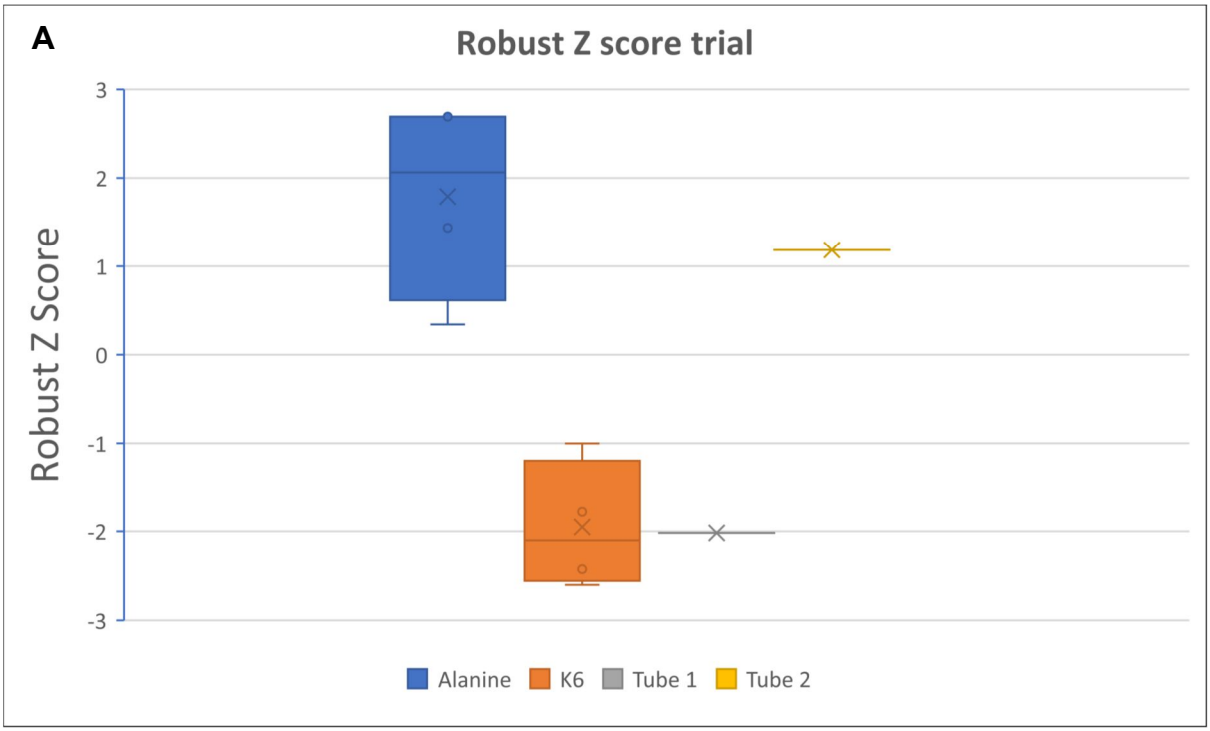


*Figure.19(**A**)* Box and whiskers plot of the robust Z scores of the ratio between Firefly and Renilla luminescence in each condition. More variation can be seen in the K6 and Alanine conditions as there were multiple repeats. K6 is a positive control.

A statistical test was needed to ensure that the assay was sensitive enough to distinguish between the variation in negative controls and an Affimer actually affecting the pathway. Robust Z scores show deviation from the median, any Affimers that can be said to inhibit the pathway would therefore show a deviation from the median similar to a positive control. The positive control used was K6 which has been shown to inhibit the activity of Ras. Two blind tubes were also included in the experiment one which had another positive control and another that had a negative control. The plate layout used can be seen in *figure.19B*, it uses a high number of negative controls to ensure that the Z scores are accurate. As *figure 19A* shows the assay was sensitive enough to distinguish between the two tubes. Tube 1 was correctly identified as containing the positive control and tube 2 the negative control. This confirms that the assay is sensitive enough to significantly distinguish between random variation and Affimers with a definite effect on the pathway.

2.2.3 MAPK pathway Affimer screening

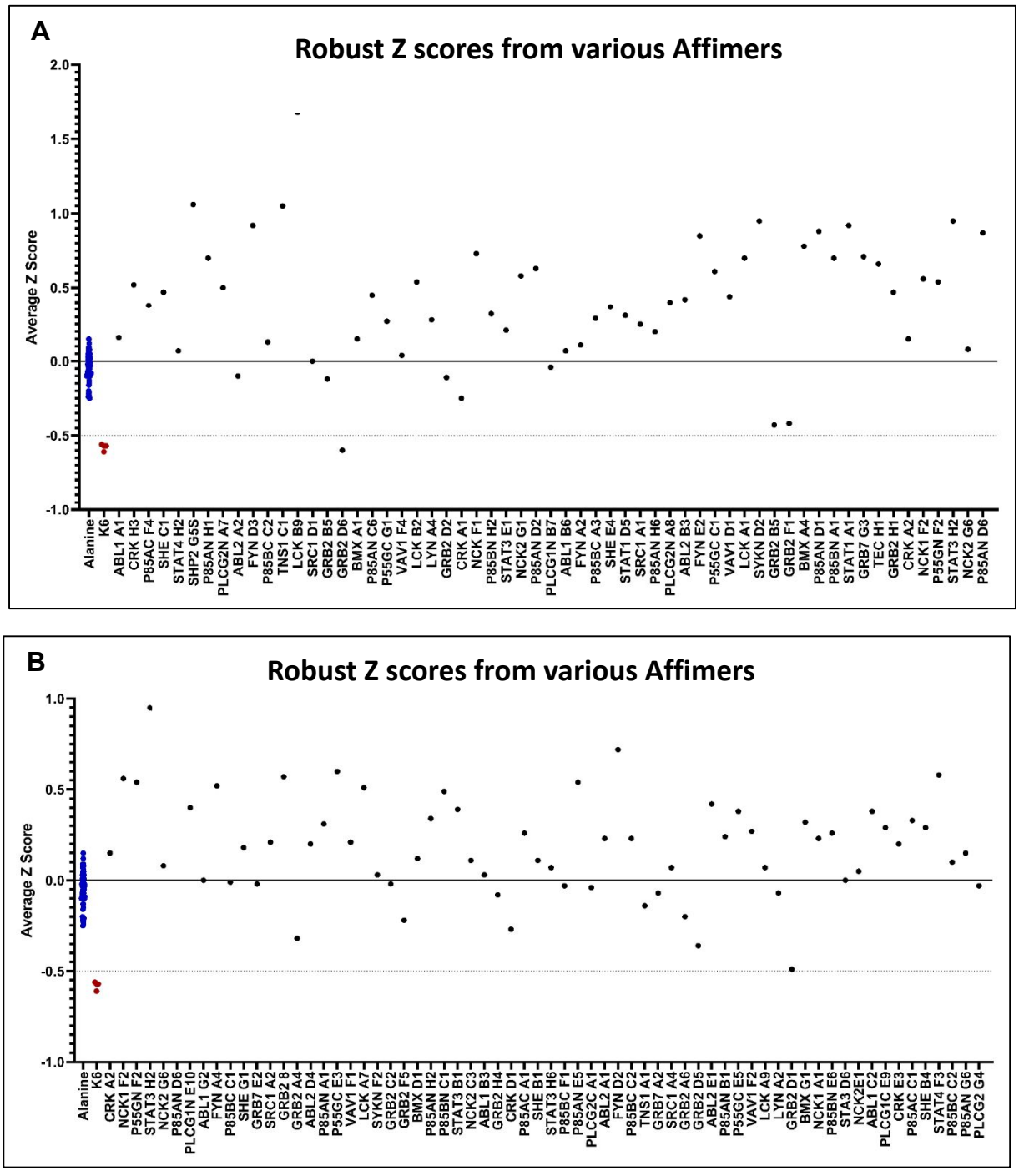


Figure.20 Graph of robust Z scores in the presence of various Affimers, the negative control (an alanine Affimer) is in blue, the positive control (K6 Affimer) is in red. The positive and negative controls plotted represent the robust Z score for each plate screened. Robust Z scores represent deviation from the median. All Affimer conditions are averages taken from two repeats. N = 1

In total 102 different Affimers were screened for their effect on the MAPK pathway targeting the SH2 domains of 20 different proteins. The results from *figure 20* show that the majority of Affimers screened did not reduce the activity of the MAPK pathway.

It also shows that there was a wide range of variation present in the Alanine negative control. This suggests that the activity of the MAPK pathway in this assay varies enough that it is difficult to determine whether or not an Affimer is having an effect. This could be due to issues with the protocol that caused variations in the behaviour of the pathway or problems that caused the measuring of the firefly and Renilla luciferase to be inaccurate.

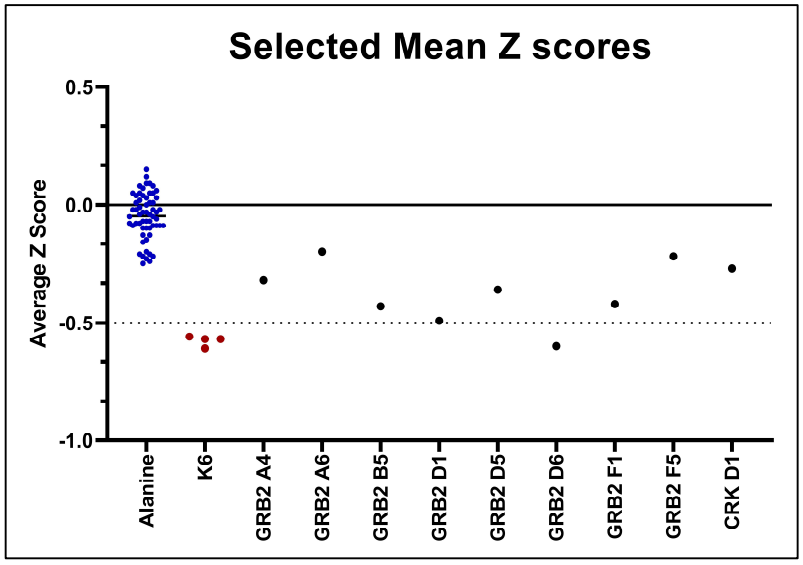
Many of the Affimers screened appear to show an increased activity in the MAPK pathway where their robust Z score is higher than the negative control. This could be due to a natural variation in the MAPK activity, but this variation was not seen in the negative controls. So, it is more likely that the results are simply not reliable enough due to the lack of repeats It is unlikely that the Affimers have been able to increase activity as they block protein-protein interactions

Some of the Affimers screened did show a negative effect on the pathway, as highlighted in *figure 21A*. However all but one of these Affimers targeted Grb2 which is already known to have activity in the MAPK pathway, although it does at least show some consistency. The only other Affimer with a negative effect is Crk1 D1, although this is a very small effect and not large enough to say that the Affimer is definitely having an effect on the pathway. Shown in figure 21B the variable regions of the Affimers do have some similarity as Grb2 D1, D6 and F1 all have a WYxN pattern in the same position and a P residue on the end of variable region 1. These Affimers produce the strongest negative effect on the MAPK pathway so it is possible that this pattern is important in blocking the protein-protein interactions of the SH2 domain

The role of Crk1 in the MAPK pathway is not well understood, previous studies have shown that it plays a role in the p38 MAPK pathway [121] inducing the production of Src. Additionally it has been found to activate JNK in HEK293 cells [122], although evidence suggests that this was through its SH3 domain [123]. These studies along with our results show that further study of the role of Crk1 in the MAPK pathway is needed.

Overall, these results are not significant enough to draw any definite conclusions as to the effects of these Affimers on the MAPK pathway. Further study is needed with more repeats to ensure reliability and possibly a more sensitive assay able to detect smaller changes in MAPK activity.

A



B

Protein	Affimer	Average Z score	Variable region sequence
Grb2	A4	-0.32	HVLWENAGP HTRYEYFVY
	B5	-0.43	NKEQRHWSE FQYVNWPVP
	A6	-0.20	HEYPMHQHN PLFMNVPLP
	D1	-0.49	KWYMNTMFP EPGRFNEML
	D6	-0.60	PWYQNVPYP REERNMNAM
	F1	-0.42	RWYVNVSLP DNMDNMNKI
	F5	-0.22	DWWEAGVFM WNEINYMFD
	D5	-0.36	FSHAYMNVV VGGGGN
Crk1	D1	-0.27	QSAWTGEQH EVHYPFSFS

Figure.21 (**A**) Highlighted robust Z scores in comparison to Alanine negative and K6 positive controls. (**B**) A table of Affimers with negative robust Z scores, the protein they target and their variable region (1 and 2) sequences. Hydrophilic residues are in blue and hydrophobic in red, areas of similarity have been highlighted

Chapter 3 Materials and Methods

3.1 Crystallisation of the Grb2-SH2 Affimer complex

3.1.1 Purification and transformation of Affimer DNA

Affimer DNA was provided by Greg Billenness of the Tomlinson lab. The Affimer sequence was contained in a pET11a plasmid and transformed into XL1-Blue supercompetent cells. 10 µL of cells were incubated on ice with 1 µL of Affimer pET11a (varying concentrations of Affimer DNA were used, all were above 100 ng/µL) for 30 minutes, then heat shocked for 45 seconds at 42 °C and further incubated on ice for 2 minutes. The cells were then incubated with shaking at 37 °C 230 rpm in 200 µL of 2YT media. 100 µL of the transformation mix was plated onto carbenicillin (100 µg/mL) (carb) lysogeny broth (LB) plates and incubated overnight at 37 °C. Plates were removed from the incubator the next day and stored at 4 °C.

The plasmid was then purified using the QiAprep Spin Miniprep Kit (QIAGEN, Cat. No. 27106). Sequencing was carried out by Genewiz and results were compared to known Affimer sequences.

3.1.2 Agarose gels

Agarose was dissolved in 20 mL TAE buffer at the desired concentration and microwaved until agarose was completely dissolved. Agarose solution was then mixed with a DNA dye at a ratio of 1:10000 and cooled in a mould along with a comb to produce sample loading wells. 5 µL DNA samples were mixed with loading buffer and loaded into wells. The gel was then run at 100 V for 30 min – 1 hr or until the dye had reached the end of the gel. UV images were then used to confirm the presence of the DNA bands in the gel

3.1.3 Addition of HRV site to pET11a Affimer

The pET11a vector was digested with (50 units) *Nhe1* and *Not1* restriction enzymes for 5 hours at 37 °C, 1 µL of Calf-intestinal alkaline phosphatase (CIP) (NEB cat no. M0290) was added and left at 37 °C for 30 minutes to prevent re-ligation of the DNA fragment. The sample was then cleaned up using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Cat. No. 740609.250) and run on a 0.7% agarose gel

(see 3.1.2). The digested vector was then extracted using the same clean up kit and stored at – 20°C.

HRV forward primer: GAA GGA GAT ATA CAT ATG G HRV reverse primer: ATT CGC GGC CGC CGG ACC CTG AAA CAG AAC TTC CAG AGC GTC ACC AAC CGG TTT G
Figure 22. Primers used to introduce HRV site, blue indicates the HRV sequence, red the Not1 restriction site

Affimer DNA was PCR amplified using primers which introduced the HRV site (see *Figure 22*) between the Affimer and the His tag. Denaturation was carried out at 98 °C, annealing at 54 °C and extension at 72 °C for 30 cycles. *DpnI* digest was then carried out for 1 hr at 37 °C using 0.5 µL *DpnI* and DNA was purified using Fastgene® Gel/PCR Extraction Kit (Cat no. P4-0508). The amplified sequence was then digested with *Nhe1* and *Not1* enzymes using the same method.

5 ng of digested insert and 30 ng of digested vector were combined with 4 µL 10X *T4 DNA ligase* and left to incubate overnight at 4 °C. A vector only control with no insert was also used. The modified Affimer DNA was then transformed into XL1-Blue supercompetent cells and purified using previously described methods (3.1.1).

3.1.4 Production and purification of His-tagged Affimers

Purification methods were the same for Affimers with and without an HRV 3C site. Affimer-pET11a plasmids were transformed into BL21 DE3 *E.coli* using the previously described methods (3.1.1). Single colonies were picked from an LB carb (100 µg/mL) plate and grown overnight in LB carb (100 µg/mL) media at 37 °C while shaking at 230 rpm. Overnight culture was then inoculated into LB carb (100 µg/mL) media (50 or 400 mL) and grown at 37 °C while shaking until OD_{600} reached ~0.7. Cultures were then induced with IPTG to a final concentration of 0.1 mM and incubated overnight shaking at 37 °C. Cells were then harvested by centrifugation at 4500 xg for 15 minutes, if needed pellets were stored at -20 °C until ready to be used.

Pellets were thawed and resuspended in 1/50th of original culture volume lysis cocktail (0.1 mg/mL lysozyme, 1% Triton x-100, 10 U/mL Benzonase Nuclease Purity, 1 X Halt Protease Inhibitor cocktail and lysis buffer consisting 20 mM Tris and 100 mM NaCl) resuspended pellets were incubated at room temperature for 1 hr on a fixed speed

rotator. Non-specific proteins were denatured by incubation in a 50 °C water bath for 20 minutes. Lysis suspension was then centrifuged at 4500 xg for 20 mins pellets from culture volumes >400 mL were also spun a second time at 12000 xg). Ni-NTA resin (300 µL slurry for 50 mL pellets or 800 µL slurry for 400 mL pellets) was washed with (1 mL for 50 mL pellets or 5 mL for 400 mL pellets) wash buffer (20 mM Tris and 100 mM NaCl) and then centrifuged to sediment the resin and the wash buffer was removed. Supernatant (containing soluble protein) was then incubated with the resin on a fixed speed rotator for 1-2 hrs. Resin was then washed several times to remove unbound protein until wash buffer run off A_{280} measured <0.09. Bound Affimer was then eluted with elution buffer (20 mM Tris, 100 mM NaCl and 300 mM imidazole) in 1 mL fractions. Protein was then run on an SDS-PAGE gel to confirm purification and expression of the correct protein. When needed Affimers were dialysed using Slide-A-Lyzer Dialysis Cassettes, 7K MWCO, 0.5/3 mL (Thermo Scientific, Cat. No. 66373) into 5 L of suitable buffer by leaving overnight at 4 °C.

3.1.5 SDS-PAGE gel

Separating gel was created by combining 2.4 mL of gel buffer (1.5M tris and 0.4% SDS) 5.2 mL 30% acrylamide stock, 2.4 mL deionised H_2O, 150 µL 10% APS and 6µL TEMED. Separating gel mix was then left in gel mould until set. Stacking gel was created using 1.2 mL stacking gel buffer (0.4M Tris and 0.4% SDS), 0.8 mL 30% acrylamide, 2.8 mL deionised H_2O, 100 µL 10% APS and 4 µL TEMED. Stacking gel mix was added to the top of the gel mould and teeth were added to create wells. Gel was then left until set. Gels were stored in SDS-Buffer at 4 °C overnight for no more than 2 days.

Samples loaded onto gels were mixed with sample buffer and 20% mercaptoethanol and boiled at 95 °C for 5 minutes. Gels were run at 140 V for around 45 minutes or until sample buffer had run off the gel. Gels were left in Coomassie blue stain for 45 minutes and then destained overnight in destain solution (25% mercaptoethanol and 7.5% acetic acid).

3.1.6 Production of Grb2-SH2 domain

SH2-pET11a was transformed into BL21 DE3 cells using the previously described (3.1.1) method and expressed using IPTG induction as previously described (3.1.4), but incubated overnight at 20 °C. The cell pellet was resuspended in previously

described lysis cocktail (3.1.4) and incubated at 1-2 hrs at room temperature on a fixed speed rotator. Then centrifuged at 4500 xg and the soluble fraction was separated from the pellet. The presence of protein was confirmed by SDS-PAGE gel (3.1.5).

3.1.7 F1-SH2 complex formation

Purified Affimer protein was mixed with Grb2-SH2 domain at a 1:10 ratio, in trials purified Affimer was mixed with soluble fraction Grb2-SH2 domain, however for the final crystallisation trial Affimer soluble fraction (instead of purified Affimer) was mixed with Grb2-SH2 domain soluble fraction. In complex formation trials magnetic Ni-NTA beads were washed 1 mL with wash buffer and incubated with Affimer SH2 mix for 1 hr on a fixed speed rotator. Beads were then washed with wash buffer until wash buffer run off measured A_{280} < 0.09 and complex was eluted using elution buffer and run on an SDS-PAGE gel to confirm that the pull down was successful. During crystallisation attempts 8 mL Ni-NTA resin was used instead of magnetic beads.

3.1.8 Crystallisation of F1-SH2 complex

6 1.5 L cultures were used to produce the SH2 domain and 4 400 mL cultures were used to produce the F1 Affimer. 18.3 mL of F1 Affimer soluble fraction was mixed with 183 mL of SH2 domain soluble fraction and left overnight at 4 °C shaking. 8.3 mL of Ni-NTA resin was used to purify the F1-SH2 complex using previously described method (see 3.1.4 and 3.1.7). Resin was then spun at 1000 xg and unbound fraction discarded. Resin was washed until run off A_{280} measured <0.09 and complex was eluted with elution buffer in 2 mL fractions. Elutions were then pooled and concentrated by centrifugation in a concentrator at 2000 xg in 5 min intervals until the volume was around 10 mL, concentrated complex was then run on size exclusion column and fractions collected. Size exclusion fractions were then run on an SDS-PAGE gel and desired fractions were pooled and concentrated to be used in crystal trials (final concentration was not obtained due to lab shutdowns)

Trials were then carried out using JCSG I, II, III, IV and Morpheus buffer kits with 1:1 2:1 and 1:2 ratios of complex to buffer. Crystal growth was then monitored using remotely accessed images.

3.1.9 Crystal optimisation trials

F1 Affimer soluble fraction and SH2 domain soluble fraction were produced using the methods previously described and complex was also formed using the previously described method (3.1.1-3.1.7). Concentration of the sample was not measured due to time constraints. Buffers used in optimisation trials were designed and prepared using automated machines. Condition one was produced using 3.5 M ammonium sulphate, 1M lithium sulphate and 1 M CAPS, pH 9.5 and 11.5. Condition two was produced using 1 M magnesium chloride, 95% ethanol (%v/v) and 1 M Tris pH 7 and 9. Condition three was produced using 1 M sodium chloride, 100% polyethene glycol 300 and 1 M sodium acetate pH 4 and 5.5.

F1 SH2 complex was then left in crystal plates in a crystal hotel at protein to buffer ratios of 1;1, 1:2 and 2:1. Buffer layouts can be found in figure 15. Crystal were left for over 30 days before being picked and sent to Diamond light source for diffraction. Diffraction was carried out at 100% beam intensity for 0.1 s at 0°, 45° and 90°. Picking and diffraction was carried out by Dr Chi Trinh.

3.2 Effect of Affimers on MAPK pathway

3.2.1 Transfection of HEK293 cells

HEK293 cells were cultured in DMEM containing 10% foetal bovine serum (FBS) growing in T75 flasks at 37 °C. Cell lines were confirmed to be 70-90% confluent before transfecting. 100 ng of Affimer DNA, 1 µL of luciferase DNA, (100 ng/µL) and 15 µL of lipofectamine (200 nL) was added to each well in a 96-well plate (total volume 31 µL). DNA-Lipofectamine mix was then left on a shaker at room temp for 20 minutes. Cells were then suspended in DMEM (10% FBS) and concentration was measured using a ThermoFisher Countess and diluted to 1.35 $x10^5$ cells/mL. 75 µL of cells were then added to each well (around 10000 cells per well) using a multichannel pipette, the plate was then left for 20 minutes before leaving in 37 °C incubator for 48 hrs. Expression of Affimers was confirmed using fluorescence microscopy as all Affimers were tagged with green fluorescent protein (GFP).

3.2.2 Luciferase Assay

The same assay method was used in all experiments. Serum starved media was pre-warmed at 37 °C before being used. Cells were transfected in 96-well plates using

methods previously described (3.2.1) before being left for 48 hrs. Cells were observed via microscope to ensure that they were growing correctly and not infected .Cells were serum starved for 1 hr at 37 °C then 10 µL of 0.3 µg/mL epidermal growth factor (EGF) was added to each well. EGF was not added to any unstimulated conditions. Cells were incubated at 37 °C for 6 hrs then media was removed and replaced with serum free media. Luciferase assay was carried out according to manufacturer's instructions (Promega cat: E1500).

Chapter 4 Conclusions

This project aimed to determine the co-crystal structure of an Affimer in complex with the Grb2 SH2 domain and investigate the effects of different SH2 targeted Affimers on the MAPK pathway.

Our results did show a small amount of inhibition of the MAPK pathway when cells were transfected with Crk1 targeted Affimers. Which is backed up by other studies which found similar results [121]–[123]. However, our data is not enough to draw a confident conclusion as to how Crk1 affects the MAPK pathway. Further study is needed using more sensitive techniques such as RT-PCR which allow for the quantitative detection of specific changes in cellular RNA,

While the luciferase assay results are promising the low number of repeats used in the study, make the results highly unreliable. The results from the time course and trial statistical test do show that the assay could potentially be used to investigate the activity of signalling pathways, but significant optimisations and repeats are needed to ensure significant and reliable data.

Had a definite candidate been found to have previously undetected activity in the MAPK pathway it could have represented a new drug target and provided more insight into how cells regulate proliferation. The next step (once activity in the MAPK pathway had been confirmed) would have been to compare the results in the HEK293 cells to common cancerous cell lines. If this had revealed a difference in activity e.g Crk1 more active in the MAPK pathway in the cancerous cell line, then it could point to Crk1 being an oncogene.

In addition. this project was successful in producing the F1-SH2 complex in high enough concentration to produce crystals, these crystals were confirmed to be protein by their diffraction patterns. As none of the crystals diffracted to a high enough resolution, a structure could not be obtained. However, this represents a good first step and further work is needed to optimise crystal growth conditions.

In conclusion this thesis made good progress towards achieving its aims, ensuring that future work will likely be successful.

References

[1] H. M. Shepard, G. L. Phillips, C. D Thanos, and M. Feldmann, “Developments in therapy with monoclonal antibodies and related proteins,” *Clin Med (Lond)*, vol. 17, no. 3, pp. 220–232, Jun. 2017, doi: 10.7861/clinmedicine.17-3-220.

[2] M. Z. Siddiqui, “Monoclonal Antibodies as Diagnostics; an Appraisal,” *Indian J Pharm Sci*, vol. 72, no. 1, pp. 12–17, 2010, doi: 10.4103/0250-474X.62229.

[3] R. D. Goldman, “Antibodies: indispensable tools for biomedical research,” *Trends in Biochemical Sciences*, vol. 25, no. 12, pp. 593–595, Dec. 2000, doi: 10.1016/S0968-0004(00)01725-4.

[4] Charles Janeway, Kenneth Murphy, and Casey Weaver, “IgG antibodies consist of four polypeptide chains,” in *Janeway’s Immunobiology*, 9th ed., Garland Science, 2017, pp. 141–145.

[5] R. Y. Zhang and W. D. Shen, “Monoclonal antibody expression in mammalian cells,” *Methods Mol. Biol.*, vol. 907, pp. 341–358, 2012, doi: 10.1007/978-1-61779-974-7_20.

[6] M. Högbom, M. Eklund, P.-Å. Nygren, and P. Nordlund, “Structural basis for recognition by an in vitro evolved affibody,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 6, p. 3191, Mar. 2003, doi: 10.1073/pnas.0436100100.

[7] A. W. Vermeer and W. Norde, “The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein.,” *Biophys J*, vol. 78, no. 1, pp. 394–404, Jan. 2000.

[8] J. Löfblom, J. Feldwisch, V. Tolmachev, J. Carlsson, S. Ståhl, and F. Y. Frejd, “Affibody molecules: Engineered proteins for therapeutic, diagnostic and biotechnological applications,” *FEBS Letters*, vol. 584, no. 12, pp. 2670–2680, Jun. 2010, doi: 10.1016/j.febslet.2010.04.014.

[9] K. Nord, J. Nilsson, B. Nilsson, M. Uhlén, and P. A. Nygren, “A combinatorial library of an alpha-helical bacterial receptor domain,” *Protein Eng.*, vol. 8, no. 6, pp. 601–608, Jun. 1995, doi: 10.1093/protein/8.6.601.

[10] C. Tiede *et al.*, “Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications,” *Protein Eng. Des. Sel.*, vol. 27, no. 5, pp. 145–155, May 2014, doi: 10.1093/protein/gzu007.

[11] A. Skerra, “Alternative binding proteins: Anticalins – harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities,” *The FEBS Journal*, vol. 275, no. 11, pp. 2677–2683, 2008, doi: 10.1111/j.1742-4658.2008.06439.x.

[12] G. Beste, F. S. Schmidt, T. Stibora, and A. Skerra, “Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 96, no. 5, pp. 1898–1903, Mar. 1999, doi: 10.1073/pnas.96.5.1898.

[13] J. Silverman *et al.*, “Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains,” *Nature Biotechnology*, vol. 23, no. 12, pp. 1556–1561, Dec. 2005, doi: 10.1038/nbt1166.

[14] H. K. Binz *et al.*, “High-affinity binders selected from designed ankyrin repeat protein libraries,” *Nature Biotechnology*, vol. 22, no. 5, pp. 575–582, May 2004, doi: 10.1038/nbt962.

[15] A. Kohl, H. K. Binz, P. Forrer, M. T. Stumpp, A. Plueckthun, and M. G. Gruetter, “Designed to be stable: Crystal structure of a consensus ankyrin repeat

protein," *Proc.Natl.Acad.Sci.USA*, vol. 100, pp. 1700–1705, 2003, doi: 10.2210/pdb1mj0/pdb.
[16] T. Merz, S. K. Wetzel, S. Firbank, A. Plückthun, M. G. Grütter, and P. R. E. Mittl, "Stabilizing Ionic Interactions in a Full-consensus Ankyrin Repeat Protein," *Journal of Molecular Biology*, vol. 376, no. 1, pp. 232–240, Feb. 2008, doi: 10.1016/j.jmb.2007.11.047.
[17] D. Steiner, P. Forrer, and A. Plückthun, "Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display," *J. Mol. Biol.*, vol. 382, no. 5, pp. 1211–1227, Oct. 2008, doi: 10.1016/j.jmb.2008.07.085.
[18] J. R. Kintzing and J. R. Cochran, "Engineered knottin peptides as diagnostics, therapeutics, and drug delivery vehicles," *Current Opinion in Chemical Biology*, vol. 34, pp. 143–150, Oct. 2016, doi: 10.1016/j.cbpa.2016.08.022.
[19] J. K. Klint *et al.*, "Production of Recombinant Disulfide-Rich Venom Peptides for Structural and Functional Analysis via Expression in the Periplasm of E. coli," *PLoS One*, vol. 8, no. 5, May 2013, doi: 10.1371/journal.pone.0063865.
[20] R. J. Clark and D. J. Craik, "Chapter three - Engineering Cyclic Peptide Toxins," in *Methods in Enzymology*, vol. 503, K. D. Wittrup and G. L. Verdine, Eds. Academic Press, 2012, pp. 57–74.
[21] V. Herzig and G. F. King, "The Cystine Knot Is Responsible for the Exceptional Stability of the Insecticidal Spider Toxin ω-Hexatoxin-Hv1a," *Toxins*, vol. 7, no. 10, pp. 4366–4380, Oct. 2015, doi: 10.3390/toxins7104366.
[22] G. P. Smith, S. U. Patel, J. D. Windass, J. M. Thornton, G. Winter, and A. D. Griffiths, "Small binding proteins selected from a combinatorial repertoire of knottins displayed on phage11Edited by J. Karn," *Journal of Molecular Biology*, vol. 277, no. 2, pp. 317–332, Mar. 1998, doi: 10.1006/jmbi.1997.1621.
[23] S. J. Moore, C. L. Leung, H. K. Norton, and J. R. Cochran, "Engineering Agatoxin, a Cystine-Knot Peptide from Spider Venom, as a Molecular Probe for In Vivo Tumor Imaging," *PLoS One*, vol. 8, no. 4, Apr. 2013, doi: 10.1371/journal.pone.0060498.
[24] S. Gunasekera *et al.*, "Engineering Stabilized Vascular Endothelial Growth Factor-A Antagonists: Synthesis, Structural Characterization, and Bioactivity of Grafted Analogues of Cyclotides," *J. Med. Chem.*, vol. 51, no. 24, pp. 7697–7704, Dec. 2008, doi: 10.1021/jm800704e.
[25] L. Dardevet *et al.*, "Chlorotoxin: A Helpful Natural Scorpion Peptide to Diagnose Glioma and Fight Tumor Invasion," *Toxins*, vol. 7, no. 4, pp. 1079–1101, Apr. 2015, doi: 10.3390/toxins7041079.
[26] A. Koide, C. W. Bailey, X. Huang, and S. Koide, "The fibronectin type III domain as a scaffold for novel binding proteins," *J. Mol. Biol.*, vol. 284, no. 4, pp. 1141–1151, Dec. 1998, doi: 10.1006/jmbi.1998.2238.
[27] A. L. Main, T. S. Harvey, M. Baron, J. Boyd, and I. D. Campbell, "The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions," *Cell*, vol. 71, no. 4, pp. 671–678, Nov. 1992, doi: 10.1016/0092-8674(92)90600-h.
[28] C. J. Bond, J. C. Marsters, and S. S. Sidhu, "Contributions of CDR3 to VHH Domain Stability and the Design of Monobody Scaffolds for Naive Antibody Libraries," *Journal of Molecular Biology*, vol. 332, no. 3, pp. 643–655, Sep. 2003, doi: 10.1016/S0022-2836(03)00967-7.
[29] L. S.-L. Cheung, D. J. Shea, N. Nicholes, A. Date, M. Ostermeier, and K. Konstantopoulos, "Characterization of Monobody Scaffold Interactions with

Ligand via Force Spectroscopy and Steered Molecular Dynamics," *Sci Rep*, vol. 5, Feb. 2015, doi: 10.1038/srep08247.
[30] M. Better, C. P. Chang, R. R. Robinson, and A. H. Horwitz, "Escherichia coli secretion of an active chimeric antibody fragment," *Science*, vol. 240, no. 4855, pp. 1041–1043, May 1988, doi: 10.1126/science.3285471.
[31] J. S. Huston *et al.*, "Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli.," *Proc Natl Acad Sci U S A*, vol. 85, no. 16, pp. 5879–5883, Aug. 1988.
[32] S. Dornieden *et al.*, "Characterization of a Single-Chain Variable Fragment Recognizing a Linear Epitope of Aβ: A Biotechnical Tool for Studies on Alzheimer's Disease?," *PLOS ONE*, vol. 8, no. 3, p. e59820, Mar. 2013, doi: 10.1371/journal.pone.0059820.
[33] S. Jung, A. Honegger, and A. Plückthun, "Selection for improved protein stability by phage display11Edited by J. A. Wells," *Journal of Molecular Biology*, vol. 294, no. 1, pp. 163–180, Nov. 1999, doi: 10.1006/jmbi.1999.3196.
[34] T. F. Chen, S. de Picciotto, B. J. Hackel, and K. D. Wittrup, "Chapter Fourteen - Engineering Fibronectin-Based Binding Proteins by Yeast Surface Display," in *Methods in Enzymology*, vol. 523, A. E. Keating, Ed. Academic Press, 2013, pp. 303–326.
[35] P. Michaely, D. R. Tomchick, M. Machius, and R. G. W. Anderson, "Crystal structure of a 12 ANK repeat stack from human ankyrinR," *EMBO J*, vol. 21, no. 23, pp. 6387–6396, Dec. 2002, doi: 10.1093/emboj/cdf651.
[36] S. G. Sedgwick and S. J. Smerdon, "The ankyrin repeat: a diversity of interactions on a common structural framework," *Trends Biochem. Sci.*, vol. 24, no. 8, pp. 311–316, Aug. 1999, doi: 10.1016/s0968-0004(99)01426-7.
[37] P. Forrer, M. T. Stumpp, H. K. Binz, and A. Plückthun, "A novel strategy to design binding molecules harnessing the modular nature of repeat proteins," *FEBS Letters*, vol. 539, no. 1–3, pp. 2–6, 2003, doi: 10.1016/S0014-5793(03)00177-7.
[38] A. Plückthun, "Designed Ankyrin Repeat Proteins (DARPins): Binding Proteins for Research, Diagnostics, and Therapy," *Annual Review of Pharmacology and Toxicology*, vol. 55, no. 1, pp. 489–511, 2015, doi: 10.1146/annurev-pharmtox-010611-134654.
[39] L. Kummer *et al.*, "Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries," *Proc Natl Acad Sci U S A*, vol. 109, no. 34, pp. E2248–E2257, Aug. 2012, doi: 10.1073/pnas.1205399109.
[40] "The Fibronectin Type III Domain as a Scaffold for Novel Binding Proteins - PubMed." https://pubmed.ncbi.nlm.nih.gov/9837732-the-fibronectin-type-iii-domain-as-a-scaffold-for-novel-binding-proteins/ (accessed Jan. 18, 2020).
[41] "A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. - PubMed - NCBI." https://www.ncbi.nlm.nih.gov/pubmed/20357770 (accessed Jan. 18, 2020).
[42] "Selective Targeting of SH2 Domain–Phosphotyrosine Interactions of Src Family Tyrosine Kinases with Monobodies." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5417323/ (accessed Jan. 18, 2020).
[43] C. Tiede *et al.*, "Affimers proteins are versatile and renewable affinity reagents," p. 35.

[44] P. Zhurauski *et al.*, “Sensitive and selective Affimer-functionalised interdigitated electrode-based capacitive biosensor for Her4 protein tumour biomarker detection,” *Biosensors and Bioelectronics*, vol. 108, pp. 1–8, Jun. 2018, doi: 10.1016/j.bios.2018.02.041.
[45] J. I. Robinson *et al.*, “Affimer proteins inhibit immune complex binding to FcγRIIIa with high specificity through competitive and allosteric modes of action,” *Proc Natl Acad Sci U S A*, vol. 115, no. 1, pp. E72–E81, Jan. 2018, doi: 10.1073/pnas.1707856115.
[46] M. A. Michel, K. N. Swatek, M. K. Hospenthal, and D. Komander, “Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling,” *Mol Cell*, vol. 68, no. 1, pp. 233-246.e5, Oct. 2017, doi: 10.1016/j.molcel.2017.08.020.
[47] A. Lopata *et al.*, “Affimer proteins for F-actin: novel affinity reagents that label F-actin in live and fixed cells,” *Sci Rep*, vol. 8, no. 1, p. 6572, 26 2018, doi: 10.1038/s41598-018-24953-4.
[48] G. Carrington, D. Tomlinson, and M. Peckham, “Exploiting nanobodies and Affimers for superresolution imaging in light microscopy,” *Mol Biol Cell*, vol. 30, no. 22, pp. 2737–2740, Oct. 2019, doi: 10.1091/mbc.E18-11-0694.
[49] R. M. Horton, Z. L. Cai, S. N. Ho, and L. R. Pease, “Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction,” *Biotechniques*, vol. 8, no. 5, pp. 528–535, May 1990.
[50] G. P. Smith, “Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface,” *Science*, vol. 228, no. 4705, pp. 1315–1317, Jun. 1985, doi: 10.1126/science.4001944.
[51] L. Ledsgaard, M. Kilstrup, A. Karatt-Vellatt, J. McCafferty, and A. H. Laustsen, “Basics of Antibody Phage Display Technology,” *Toxins (Basel)*, vol. 10, no. 6, Jun. 2018, doi: 10.3390/toxins10060236.
[52] J. Hanes and A. Plückthun, “In vitro selection and evolution of functional proteins by using ribosome display,” *PNAS*, vol. 94, no. 10, pp. 4937–4942, May 1997, doi: 10.1073/pnas.94.10.4937.
[53] J. Bazan, I. Całkosiński, and A. Gamian, “Phage display—A powerful technique for immunotherapy,” *Hum Vaccin Immunother*, vol. 8, no. 12, pp. 1817–1828, Dec. 2012, doi: 10.4161/hv.21703.
[54] S.-H. Yang, A. D. Sharrocks, and A. J. Whitmarsh, “MAP kinase signalling cascades and transcriptional regulation,” *Gene*, vol. 513, no. 1, pp. 1–13, Jan. 2013, doi: 10.1016/j.gene.2012.10.033.
[55] Y. Sun, W.-Z. Liu, T. Liu, X. Feng, N. Yang, and H.-F. Zhou, “Signaling pathway of MAPK/ERK in cell proliferation, differentiation, migration, senescence and apoptosis,” *Journal of Receptors and Signal Transduction*, vol. 35, no. 6, pp. 600–604, Nov. 2015, doi: 10.3109/10799893.2015.1030412.
[56] B. D. Cuevas, A. N. Abell, and G. L. Johnson, “Role of mitogen-activated protein kinase kinase kinases in signal integration,” *Oncogene*, vol. 26, no. 22, pp. 3159–3171, May 2007, doi: 10.1038/sj.onc.1210409.
[57] J. B. Duffy and N. Perrimon, “The Torso Pathway in Drosophila: Lessons on Receptor Tyrosine Kinase Signaling and Pattern Formation,” *Developmental Biology*, vol. 166, no. 2, pp. 380–395, Dec. 1994, doi: 10.1006/dbio.1994.1324.
[58] I. A. Prior, P. D. Lewis, and C. Mattos, “A comprehensive survey of Ras mutations in cancer,” *Cancer Res*, vol. 72, no. 10, pp. 2457–2467, May 2012, doi: 10.1158/0008-5472.CAN-11-2612.

[59] A. A. Samatar and P. I. Poulikakos, “Targeting RAS–ERK signalling in cancer: promises and challenges,” *Nature Reviews Drug Discovery*, vol. 13, no. 12, pp. 928–942, Dec. 2014, doi: 10.1038/nrd4281.
[60] E. Martinelli, F. Morgillo, T. Troiani, and F. Ciardiello, “Cancer resistance to therapies against the EGFR-RAS-RAF pathway: The role of MEK,” *Cancer Treatment Reviews*, vol. 53, pp. 61–69, Feb. 2017, doi: 10.1016/j.ctrv.2016.12.001.
[61] G. Hatzivassiliou *et al.*, “ERK inhibition overcomes acquired resistance to MEK inhibitors,” *Mol. Cancer Ther.*, vol. 11, no. 5, pp. 1143–1154, May 2012, doi: 10.1158/1535-7163.MCT-11-1010.
[62] M. R. Arkin and J. A. Wells, “Small-molecule inhibitors of protein–protein interactions: progressing towards the dream,” *Nature Reviews Drug Discovery*, vol. 3, no. 4, Art. no. 4, Apr. 2004, doi: 10.1038/nrd1343.
[63] D. Caldeira and H. Pereira, “Adjuvant antithrombotic therapy in ST-elevation myocardial infarction: A narrative review,” *Revista Portuguesa de Cardiologia*, vol. 38, no. 4, pp. 289–297, Apr. 2019, doi: 10.1016/j.repc.2018.05.020.
[64] S. S. Lieberman-Blum, H. B. Fung, and J. C. Bandres, “Maraviroc: a CCR5-receptor antagonist for the treatment of HIV-1 infection,” *Clin Ther*, vol. 30, no. 7, pp. 1228–1250, Jul. 2008, doi: 10.1016/s0149-2918(08)80048-3.
[65] M. G. Myers *et al.*, “Role of IRS-1-GRB-2 complexes in insulin signaling.,” *Mol Cell Biol*, vol. 14, no. 6, pp. 3577–3587, Jun. 1994.
[66] S. Maignan, J. P. Guilloteau, N. Fromage, B. Arnoux, J. Becquart, and A. Ducruix, “Crystal structure of the mammalian Grb2 adaptor.,” *Science*, vol. 268, no. 5208, pp. 291–293, Apr. 1995, doi: 10.1126/science.7716522.
[67] Y. Qu *et al.*, “SUMOylation of Grb2 enhances the ERK activity by increasing its binding with Sos1,” *Mol. Cancer*, vol. 13, p. 95, Apr. 2014, doi: 10.1186/1476-4598-13-95.
[68] C. B. McDonald, K. L. Seldeen, B. J. Deegan, and A. Farooq, “Structural basis of the differential binding of the SH3 domains of Grb2 adaptor to the guanine nucleotide exchange factor Sos1,” *Arch. Biochem. Biophys.*, vol. 479, no. 1, pp. 52–62, Nov. 2008, doi: 10.1016/j.abb.2008.08.012.
[69] P. Chardin *et al.*, “Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2,” *Science*, vol. 260, no. 5112, pp. 1338–1343, May 1993, doi: 10.1126/science.8493579.
[70] J. den Hertog and T. Hunter, “Tight association of GRB2 with receptor protein-tyrosine phosphatase alpha is mediated by the SH2 and C-terminal SH3 domains,” *EMBO J.*, vol. 15, no. 12, pp. 3016–3027, Jun. 1996.
[71] Z. Ahmed *et al.*, “Grb2 monomer–dimer equilibrium determines normal versus oncogenic function,” *Nat Commun*, vol. 6, no. 1, pp. 1–11, Jun. 2015, doi: 10.1038/ncomms8354.
[72] Cosmic, “GRB2 Gene - COSMIC.” https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=GRB2 (accessed Oct. 30, 2019).
[73] B. Gay *et al.*, “Effect of potent and selective inhibitors of the Grb2 SH2 domain on cell motility,” *J. Biol. Chem.*, vol. 274, no. 33, pp. 23311–23315, Aug. 1999, doi: 10.1074/jbc.274.33.23311.
[74] A. Giubellino, T. R. Burke, and D. P. Bottaro, “Grb2 Signaling in Cell Motility and Cancer,” *Expert Opin Ther Targets*, vol. 12, no. 8, pp. 1021–1033, Aug. 2008, doi: 10.1517/14728222.12.8.1021.

[75] C. Birchmeier, W. Birchmeier, E. Gherardi, and G. F. V. Woude, “Met, metastasis, motility and more,” *Nat Rev Mol Cell Biol*, vol. 4, no. 12, pp. 915–925, Dec. 2003, doi: 10.1038/nrm1261.
[76] J. F. Lai, S. C. Kao, S. T. Jiang, M. J. Tang, P. C. Chan, and H. C. Chen, “Involvement of focal adhesion kinase in hepatocyte growth factor-induced scatter of Madin-Darby canine kidney cells,” *J. Biol. Chem.*, vol. 275, no. 11, pp. 7474–7480, Mar. 2000, doi: 10.1074/jbc.275.11.7474.
[77] G. Hartmann, K. M. Weidner, H. Schwarz, and W. Birchmeier, “The motility signal of scatter factor/hepatocyte growth factor mediated through the receptor tyrosine kinase met requires intracellular action of Ras,” *J. Biol. Chem.*, vol. 269, no. 35, pp. 21936–21939, Sep. 1994.
[78] D. Hanahan and R. A. Weinberg, “The Hallmarks of Cancer,” *Cell*, vol. 100, no. 1, pp. 57–70, Jan. 2000, doi: 10.1016/S0092-8674(00)81683-9.
[79] J. V. Soriano *et al.*, “Inhibition of angiogenesis by growth factor receptor bound protein 2-Src homology 2 domain bound antagonists,” *Mol. Cancer Ther.*, vol. 3, no. 10, pp. 1289–1299, Oct. 2004.
[80] M. L. Burness, T. A. Grushko, and O. I. Olopade, “Epidermal growth factor receptor in triple-negative and basal-like breast cancer: promising clinical target or only a marker?,” *Cancer J*, vol. 16, no. 1, pp. 23–32, Feb. 2010, doi: 10.1097/PPO.0b013e3181d24fc1.
[81] J. R. Sainsbury, J. R. Farndon, G. K. Needham, A. J. Malcolm, and A. L. Harris, “Epidermal-growth-factor receptor status as predictor of early recurrence of and death from breast cancer,” *Lancet*, vol. 1, no. 8547, pp. 1398–1402, Jun. 1987, doi: 10.1016/s0140-6736(87)90593-9.
[82] K. Nakai, M.-C. Hung, and H. Yamaguchi, “A perspective on anti-EGFR therapies targeting triple-negative breast cancer,” *Am J Cancer Res*, vol. 6, no. 8, pp. 1609–1623, Aug. 2016.
[83] R. J. Daly, M. D. Binder, and R. L. Sutherland, “Overexpression of the Grb2 gene in human breast cancer cell lines,” *Oncogene*, vol. 9, no. 9, pp. 2723–2727, Sep. 1994.
[84] P. Furet *et al.*, “Structure-Based Design and Synthesis of High Affinity Tripeptide Ligands of the Grb2-SH2 Domain,” *J. Med. Chem.*, vol. 41, no. 18, pp. 3442–3449, Aug. 1998, doi: 10.1021/jm980159a.
[85] C. García-Echeverría *et al.*, “Potent Antagonists of the SH2 Domain of Grb2: Optimization of the X+1 Position of 3-Amino-Z-Tyr(PO3H2)-X+1-Asn-NH2,” *J. Med. Chem.*, vol. 41, no. 11, pp. 1741–1744, May 1998, doi: 10.1021/jm970856n.
[86] B. A. Liu, K. Jablonowski, M. Raina, M. Arcé, T. Pawson, and P. D. Nash, “The Human and Mouse Complement of SH2 Domain Proteins—Establishing the Boundaries of Phosphotyrosine Signaling,” *Molecular Cell*, vol. 22, no. 6, pp. 851–868, Jun. 2006, doi: 10.1016/j.molcel.2006.06.001.
[87] T. Kaneko *et al.*, “Identification and characterization of a large family of superbinding bacterial SH2 domains,” *Nat Commun*, vol. 9, no. 1, p. 4549, 31 2018, doi: 10.1038/s41467-018-06943-2.
[88] S. J. Parsons and J. T. Parsons, “Src family kinases, key regulators of signal transduction,” *Oncogene*, vol. 23, no. 48, pp. 7906–7909, Oct. 2004, doi: 10.1038/sj.onc.1208160.
[89] X. Liu and T. Pawson, “Biochemistry of the Src protein-tyrosine kinase: regulation by SH2 and SH3 domains,” *Recent Prog. Horm. Res.*, vol. 49, pp. 149–160, 1994.

[90] B. A. Liu and P. D. Nash, “Evolution of SH2 domains and phosphotyrosine signalling networks,” *Philos Trans R Soc Lond B Biol Sci*, vol. 367, no. 1602, pp. 2556–2573, Sep. 2012, doi: 10.1098/rstb.2012.0107.
[91] B. A. Liu, E. Shah, K. Jablonowski, A. Stergachis, B. Engelmann, and P. D. Nash, “The SH2 Domain–Containing Proteins in 21 Species Establish the Provenance and Scope of Phosphotyrosine Signaling in Eukaryotes,” *Sci Signal*, vol. 4, no. 202, p. ra83, Dec. 2011, doi: 10.1126/scisignal.2002105.
[92] T. P. Luc E.M Marengere, “Structure and function of SH2 domains,” *Journal of Cell Science*, pp. 97–104, 1994.
[93] C. Haan, S. Kreis, C. Margue, and I. Behrmann, “Jaks and cytokine receptors--an intimate relationship,” *Biochem. Pharmacol.*, vol. 72, no. 11, pp. 1538–1546, Nov. 2006, doi: 10.1016/j.bcp.2006.04.013.
[94] M. M. Senior *et al.*, “The three-dimensional solution structure of the SRC homology domain-2 of the growth factor receptor-bound protein-2,” *J. Biomol. NMR*, vol. 11, no. 2, pp. 153–164, Feb. 1998, doi: 10.1023/A:1008250309874.
[95] W. Xu, S. C. Harrison, and M. J. Eck, “Three-dimensional structure of the tyrosine kinase c-Src,” *Nature*, vol. 385, no. 6617, pp. 595–602, Feb. 1997, doi: 10.1038/385595a0.
[96] L. Tong, T. C. Warren, J. King, R. Betageri, J. Rose, and S. Jakes, “Crystal structures of the human p56lck SH2 domain in complex with two short phosphotyrosyl peptides at 1.0 A and 1.8 A resolution,” *J. Mol. Biol.*, vol. 256, no. 3, pp. 601–610, Mar. 1996, doi: 10.1006/jmbi.1996.0112.
[97] J. C. Williams *et al.*, “The 2.35 A crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions,” *J. Mol. Biol.*, vol. 274, no. 5, pp. 757–775, Dec. 1997, doi: 10.1006/jmbi.1997.1426.
[98] W. Xu, A. Doshi, M. Lei, M. J. Eck, and S. C. Harrison, “Crystal Structures of c-Src Reveal Features of Its Autoinhibitory Mechanism,” *Molecular Cell*, vol. 3, no. 5, pp. 629–638, May 1999, doi: 10.1016/S1097-2765(00)80356-1.
[99] M. A. Young, S. Gonfloni, G. Superti-Furga, B. Roux, and J. Kuriyan, “Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation,” *Cell*, vol. 105, no. 1, pp. 115–126, Apr. 2001, doi: 10.1016/s0092-8674(01)00301-4.
[100] G. Superti-Furga, S. Fumagalli, M. Koegl, S. A. Courtneidge, and G. Draetta, “Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src,” *EMBO J.*, vol. 12, no. 7, pp. 2625–2634, Jul. 1993.
[101] S. W. Cowan-Jacob *et al.*, “The Crystal Structure of a c-Src Complex in an Active Conformation Suggests Possible Steps in c-Src Activation,” *Structure*, vol. 13, no. 6, pp. 861–871, Jun. 2005, doi: 10.1016/j.str.2005.03.012.
[102] M. F. Moran *et al.*, “Src homology region 2 domains direct protein-protein interactions in signal transduction.,” *Proc Natl Acad Sci U S A*, vol. 87, no. 21, pp. 8622–8626, Nov. 1990.
[103] H. Huang *et al.*, “Defining the Specificity Space of the Human Src Homology 2 Domain,” *Molecular & Cellular Proteomics*, vol. 7, no. 4, pp. 768–784, Apr. 2008, doi: 10.1074/mcp.M700312-MCP200.
[104] M. Tinti *et al.*, “The SH2 Domain Interaction Landscape,” *Cell Reports*, vol. 3, no. 4, pp. 1293–1305, Apr. 2013, doi: 10.1016/j.celrep.2013.03.001.
[105] A. G. Batzer, D. Rotin, J. M. Ureña, E. Y. Skolnik, and J. Schlessinger, “Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor.,” *Molecular and Cellular Biology*, vol. 14, no. 8, pp. 5192–5201, Aug. 1994, doi: 10.1128/MCB.14.8.5192.

[106] A. Stein and P. Aloy, “Contextual Specificity in Peptide-Mediated Protein Interactions,” *PLoS One*, vol. 3, no. 7, Jul. 2008, doi: 10.1371/journal.pone.0002524.
[107] A. G. Batzer, D. Rotin, J. M. Ureña, E. Y. Skolnik, and J. Schlessinger, “Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor.,” *Molecular and Cellular Biology*, vol. 14, no. 8, pp. 5192–5201, Aug. 1994, doi: 10.1128/MCB.14.8.5192.
[108] J. Downward, P. Parker, and M. D. Waterfield, “Autophosphorylation sites on the epidermal growth factor receptor,” *Nature*, vol. 311, no. 5985, pp. 483–485, Oct. 1984, doi: 10.1038/311483a0.
[109] J. A. Jadwin *et al.*, “Time-resolved multimodal analysis of Src Homology 2 (SH2) domain binding in signaling by receptor tyrosine kinases,” *Elife*, vol. 5, p. e11835, Apr. 2016, doi: 10.7554/eLife.11835.
[110] D. Oh, M. Ogiue-Ikeda, J. A. Jadwin, K. Machida, B. J. Mayer, and J. Yu, “Fast rebinding increases dwell time of Src homology 2 (SH2)-containing proteins near the plasma membrane,” *PNAS*, vol. 109, no. 35, pp. 14024–14029, Aug. 2012, doi: 10.1073/pnas.1203397109.
[111] I. Lappalainen, J. Thusberg, B. Shen, and M. Vihinen, “Genome wide analysis of pathogenic SH2 domain mutations,” *Proteins: Structure, Function, and Bioinformatics*, vol. 72, no. 2, pp. 779–792, 2008, doi: 10.1002/prot.21970.
[112] O. C. Bruton, “Agammaglobulinemia,” *Pediatrics*, vol. 9, no. 6, pp. 722–728, Jun. 1952.
[113] D. Suri, A. Rawat, and S. Singh, “X-linked Agammaglobulinemia,” *Indian J Pediatr*, vol. 83, no. 4, pp. 331–337, Apr. 2016, doi: 10.1007/s12098-015-2024-8.
[114] E. S. Resnick, E. L. Moshier, J. H. Godbold, and C. Cunningham-Rundles, “Morbidity and mortality in common variable immune deficiency over 4 decades,” *Blood*, vol. 119, no. 7, pp. 1650–1657, Feb. 2012, doi: 10.1182/blood-2011-09-377945.
[115] M. Fiorini *et al.*, “BTK: 22 novel and 25 recurrent mutations in European patients with X-linked agammaglobulinemia,” *Hum. Mutat.*, vol. 23, no. 3, p. 286, Mar. 2004, doi: 10.1002/humu.9219.
[116] S. R. Tzeng *et al.*, “Stability and peptide binding specificity of Btk SH2 domain: molecular basis for X-linked agammaglobulinemia.,” *Protein Sci*, vol. 9, no. 12, pp. 2377–2385, Dec. 2000.
[117] V. Seda and M. Mraz, “B-cell receptor signalling and its crosstalk with other pathways in normal and malignant cells,” *European Journal of Haematology*, vol. 94, no. 3, pp. 193–205, 2015, doi: 10.1111/ejh.12427.
[118] B. Shillitoe and A. Gennery, “X-Linked Agammaglobulinaemia: Outcomes in the modern era,” *Clinical Immunology*, vol. 183, pp. 54–62, Oct. 2017, doi: 10.1016/j.clim.2017.07.008.
[119] L. J. Maarschalk-Ellerbroek, I. M. Hoepelman, and P. M. Ellerbroek, “Immunoglobulin treatment in primary antibody deficiency,” *International Journal of Antimicrobial Agents*, vol. 37, no. 5, pp. 396–404, May 2011, doi: 10.1016/j.ijantimicag.2010.11.027.
[120] U. Baumann, S. Miescher, and C. Vonarburg, “Immunoglobulin replacement therapy in antibody deficiency syndromes: are we really doing enough?,” *Clinical & Experimental Immunology*, vol. 178, no. S1, pp. 83–85, 2014, doi: 10.1111/cei.12521.

[121] T. Watanabe *et al.*, “Adaptor protein Crk induces Src-dependent activation of p38 MAPK in regulation of synovial sarcoma cell proliferation,” *Mol Cancer Res*, vol. 7, no. 9, pp. 1582–1592, Sep. 2009, doi: 10.1158/1541-7786.MCR-09-0064.
[122] F. Dolfi, M. Garcia-Guzman, M. Ojaniemi, H. Nakamura, M. Matsuda, and K. Vuori, “The adaptor protein Crk connects multiple cellular stimuli to the JNK signaling pathway,” *PNAS*, vol. 95, no. 26, pp. 15394–15399, Dec. 1998, doi: 10.1073/pnas.95.26.15394.
[123] S. Tanaka, T. Ouchi, and H. Hanafusa, “Downstream of Crk adaptor signaling pathway: Activation of Jun kinase by v-Crk through the guanine nucleotide exchange protein C3G,” *Proc Natl Acad Sci U S A*, vol. 94, no. 6, pp. 2356–2361, Mar. 1997.

Word Count: 12687 (not including references and annotations)