Characterisation of the interaction between FGFR2/EGL-15 and CRKL/CED-2

Dovile Milonaityte

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Molecular and Cellular Biology

July 2021

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to thank my supervisor Professor John Ladbury for the guidance and support throughout the project. I would also like to thank my co-supervisor Dr Patricija van Oosten-Hawle for her feedback and support. I am grateful to Cancer Research UK for funding my PhD. Thanks also go to Professor Ian Hope for the feedback on my nematode work, and for access to injection and Nomarski microscopes, without which the *C. elegans* chapter wouldn't exist.

I would also like to thank the current and former members of the Ladbury group for all their support, help and advice. In particular, I would like to give my eternal thanks to Dr Chi-Chuan Lin and Dr Kin Man Suen for all their encouragement and help with experiments, and the sea of coffee. This project would not be where it is now without their support.

My sincere thanks go to Dr Bharat Pokhrel and Dr Yannic Chen for their help and support with my early nematode work, and for all the delicious chicken lunches.

I would also like to thank my mom and my sister for their support through the years. Didžiulis ačiū už palaikymą. Special thanks go to Lukas for all his support and patience through the difficult year of the pandemic.

Abstract

A new type of Receptor Tyrosine Kinase (RTK) signalling has been uncovered, referred to as “tier 2” signalling. This signalling is mediated by an interaction between the proline-rich motifs in RTKs and the SH3 domains of binding proteins and can occur in absence of ligand stimulation. Investigating the impact of these interactions in various organisms can lead to better understanding of the complex cell signalling landscape.

In this study, an SH3-mediated interaction between FGFR2 and CRKL was established using biochemical and biophysical techniques. This interaction lead to increased FGFR2 autotransphosphorylation. Furthermore, FGFR2 was able to phosphorylate CRKL. Both proteins were shown to be important for cell migration under non-stimulating conditions. The downstream signalling landscape was investigated using Reverse Phase Protein Array (RPPA), which showed that CRKL helps mediate signalling through the AKT pathway even in absence of cell stimulation.

To determine whether this interaction is evolutionarily conserved, *Caenorhabditis elegans* model organism was employed. The animal contains FGFR orthologue EGL-15, and CRKL orhologue CED-2. A direct interaction between the EGL-15 PRNPLP proline-rich motif and CED-2 was found using Microscale Thermophoresis (MST), while a genetic interaction between *egl-15* and *ced-2* was established in the living animal using CRISPR/Cas9 generated genetic mutants. Separately, the SH3 domains of CED-2 were found to be essential for engulfment, Distant Tip Cell (DTC) migration, and normal brood size maintenance.

This study demonstrates that an interaction between FGFR2 and CRKL can influence cell signalling under non-stimulating conditions. The results also establish a similar interaction between *C. elegans* orthologues EGL-15 and CED-2, suggesting a novel role of EGL-15 in control of apoptotic germ cell engulfment.

Table of Contents

List of Tables

List of Figures

Abbreviations

µg – Microgram
µL – Microlitre
µm - Micrometre
µM – Micromolar
A – Alanine
ALK – Anaplastic lymphoma kinase
AO – Acridine orange
ATP – Adenosine triphosphate
BCAR – Breast cancer anti-estrogen resistance protein
BCL-XL – B-cell lymphoma-extra large
BCR – Breakpoint cluster region protein
BSA – Bovine serum albumin
$CaCl_2$ – Calcium chloride
CAN – Canal-associated neuron
CAS – CRISPR associated protein
CBL – Casitas B-lineage lymphoma
CDC – Cell division control protein
CED – CEll Death abnormality
CHK – Checkpoint kinase
CIP/WAF – CDK interacting protein
CLR – CleaR
CML – Chronic myelogenous leukaemia
CO_2 – Carbon dioxide
Co-IP – Co-immunoprecipitation
CRISPR – Clustered regularly interspaced short palindromic repeats
CRK(L) – CT10 regulator of kinase (like)
Da – Dalton
DAB – Disabled
DAG – Diacylglycerol
DAPI - 4′,6-diamidino-2-phenylindole
DMEM – Dulbecco’s modified eagle medium
DNA – Deoxyribonucleic acid
dNTP – Deoxynucleoside triphosphate
DOCK – Dedicator of cytokinesis
DPBS – Dulbecco’s phosphate-buffered saline

DTC – Distal tip cell
EGF – Epidermal growth factor
EGL – Egg laying defective
EGTA – Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
ELMO – Engulfment and cell motility
ERK – Extracellular signal-regulated kinase
EV – Empty vector
F – Phenylalanine
FAK – Focal adhesion kinase
FBS – Foetal bovine serum
FGF(R) – Fibroblast growth factor (receptor)
FRS – FGFR substrate
g – Gram
GAB1 – GRB2-associated binding protein
gDNA – Genomic DNA
GDP – Guanosine diphosphate
GEF – Guanine nucleotide exchange factor
GFP – Green fluorescent protein
GRB2 – Growth factor receptor-bound
GSK – Glycogen synthase kinase
GST – Glutathione S-transferase
GTP – Guanosine triphosphate
GTPγS – Guanosine 5'-[γ-thio]triphosphate
HBST – HEPES-buffered saline (+tween20)
HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGF – Hepatocyte growth factor
HRP – Horseradish peroxidase
HSPG – Heparin sulphate proteoglycan
IB - Immunoblot
IDT – Integrated DNA technologies
Ig – Immunoglobulin
IP_3 – Inositol-1,4,5-triphosphate
IPTG – Isopropyl β-D-thiogalactoside
JNK – C-Jun N-terminal kinase
KCl – Potassium chloride
kDa – Kilodalton

KPO_4 – Potassium phosphate

L – Leucine

L – Litre

LE – Long exposure

LED – Light-emitting diode

LET - LEThal

MAPK – Mitogen-activated protein kinase

MBP – Maltose binding protein

mg – Milligram

$MgCl_2$ – Magnesium chloride

$MgSO_4$ – Magnesium sulphate

mL – Millilitre

mM – Micromolar

mM – Millimolar

MOPS – 3-(N-morpholino)propane sulfonic acid

MS – Mass spectrometry

MST – Microscale thermophoresis

mTOR – Mammalian target of rapamycin

N - Asparagine

Na_2HPO_4 – Disodium phosphate

Na_3VO_4 – Sodium orthovanadate

NaCl – Sodium chloride

NaClO – Sodium hypochlorite

NaF – Sodium fluoride

NEB – New England Biolabs

NES – Nuclear export signal

ng – nanogram

NGM – Nematode growth media

nm – Nanometres

nM – Nanomolar

NSCLC – Non-small-cell lung carcinoma

PAK – P21 activated kinase

PCR – Polymerase chain reaction

PDK1 – Pyruvate dehydrogenase kinase

PI3-K – Phosphoinositide 3-kinase

PIP_2 – Phosphotidylinositol-4,5-bisphosphate

PIP_3 – Phosphotidylinositol-3,4,5-trisphosphate
PKC – Protein kinase C
PLCγ – Phospholipase C, gamma
PTB – Phosphotyrosine-binding
PTEN – Phosphatase and tensin homolog
PTM – Post-translational modification
PVDF – Polyvinylidene fluoride
R – Arginine
RAC – Ras-related C3 botulinum toxin substrate
RAF – Rapidly accelerated fibrosarcoma
RalGDS – Ral guanine nucleotide dissociation stimulator
RAP – Ras-related protein
RB – Retinoblastoma
RBD – Receptor-binding domain
RMS – Rhabdomyosarcoma
RNAi – RNA interference
RPM – Revolutions per minute
RPPA – Reverse phase protein array
rSAP – Shrimp alkaline phosphatase
RSK – Ribosomal S6 kinase
RTK – Receptor tyrosine kinase
RTKs – Receptor tyrosine kinases
S – Serine
SDS – Sodium dodecyl sulphate
SE – Short exposure
SEM – SEx Muscle abnormal
SH2 – SRC homology 2
SH3 – SRC homology 3
SHP – SH2 domain-containing phosphatase
shRNA – Short hairpin ribonucleic acid
SM – Sex myoblast
SOS – Son of sevenless
STAT – Signal transducer and activator of transcription
T – Threonine
TBS(T) – Tris-buffered saline (with Tween20)
TCEP – Tris(2-carboxyethyl)phosphine

TSPP – Tetrasodium pyrophosphate

V – Volt

W – Tryptophan

XIAP – X-linked inhibitor of apoptosis

Y - Tyrosine

YAP – Yes-associated protein

Chapter 1 Introduction

1.1 Receptor tyrosine kinases

One family of transmembrane receptors are receptor tyrosine kinases (RTKs): proteins that can catalyse the phosphorylation of tyrosine amino acid residues. The RTK family has 58 known members (Robinson et al., 2000). Canonically, these proteins recognise and bind an extracellular polypeptide ligand and become active by autophosphorylation of specific cytoplasmic tyrosine residues. The resulting phosphorylated tyrosine (pY) residues act as docking sites for downstream signalling proteins. Bound substrate proteins are then phosphorylated and activated by the RTK, which leads to downstream signalling cascade activity (Lemmon and Schlessinger, 2010).

1.1.1 General structure and function of RTKs

RTK family members share a set of common structural features: an extracellular domain which recognises polypeptide ligands, a transmembrane helix, and a cytoplasmic domain (Figure 1-1 A) (Yarden and Axel, 1988; Hubbard and Till, 2000). In order to recognise a wide variety of ligands, the RTK extracellular domain shows the most variety within the family. The extracellular region can contain cysteine-rich domains, EGF-like domains, fibronectin type III-like domains, immunoglobulin (Ig)-like domains, or other domain types (Lemmon and Schlessinger, 2010). The intracellular portion of the RTKs contain a juxtamembrane region, a tyrosine kinase domain, and a carboxy-terminal (C-terminal) region (Hubbard and Till, 2000). The overall RTK structure and some

signalling pathways are conserved in many organisms, from *Caenorhabditis elegans* nematode worms to humans (Robinson et al., 2000).

The main function of RTKs is to activate downstream signalling by phosphorylating target proteins. They do so by using their intrinsic enzymatic activity which involves the transfer the γ-phosphate from ATP to the tyrosine hydroxyl group of substrate proteins. The catalytic kinase domain of the RTK comprises an amino-terminal lobe containing a five-stranded β-sheet and one α-helix, and a carboxy-terminal lobe made up of α-helices (Hubbard et al., 1994; Hubbard, 1999). Ligand-induced receptor dimerisation leads to conformational changes which allow activation of the catalytic kinase domain (Schlessinger, 1988; Schlessinger, 2000). The activation loop, which in the inactive state blocks the active site of the catalytic domain, is *trans*-phosphorylated on key tyrosine residues by the other receptor molecule in the dimer, thus releasing inhibition (Hubbard et al., 1998; Lemmon and Schlessinger, 2010). The active kinase domain can then phosphorylate tyrosine residues within the other intracellular domains of the receptor, which provides docking sites for pY-recognising domains such as SRC homology 2 (SH2) and Phosphotyrosine-binding (PTB) domains that are present within RTK-binding proteins (Figure 1-1 B and C) (Anderson et al., 1990; Matsuda et al., 1990; Gustafson et al., 1995).

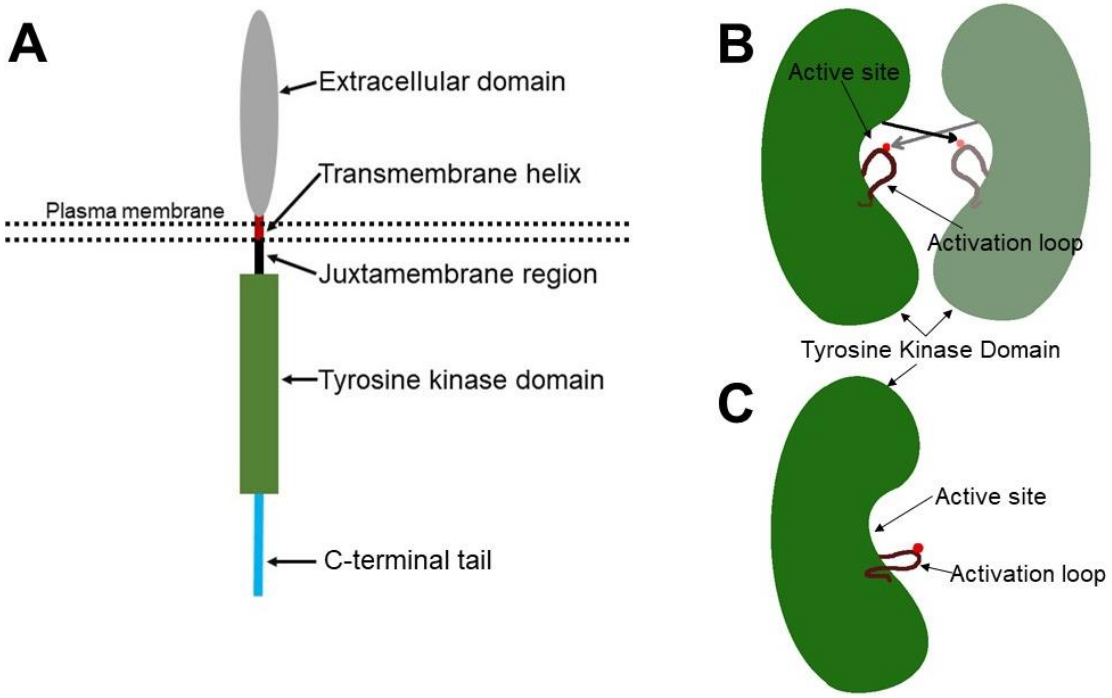


Figure 1-1. A schematic depiction of general RTK family protein structure and activation mechanism. A) A cartoon depiction of general features of RTKs. C-terminal tail can contain multiple proline-rich motifs. **B and C)** General schematic of the activation of the catalytic tyrosine kinase site of the receptor. Two inactive tyrosine kinase domains are brought together by ligand-induced receptor dimerisation. The nearby active site within the dimer phosphorylates a tyrosine residue (red) in a neighbouring activation loop. **C)** Phosphorylated activation loop no longer blocks the active site of the kinase domain, allowing further receptor activation by transphosphorylation, as well as phosphorylation of binding partner proteins.

In summary, ligand binding activates RTKs by enabling the dimerisation of the receptor and autotransphosphorylation of tyrosine residues. This allows downstream effector recruitment and activation of signalling cascades, which will be discussed in more detail in section 1.3.2.

1.2 Fibroblast growth factor receptor family

Fibroblast Growth Factor Receptors (FGFRs) are a family of RTKs which are stimulated by Fibroblast Growth Factor (FGF) ligands. FGFRs are essential during development, as they regulate cell proliferation, migration and differentiation, while in adults they assist tissue repair and participate in

neuronal signal transduction (Turner and Grose, 2010). There are 22 different FGFs which are highly conserved amongst vertebrates (Ornitz and Itoh, 2001), and 4 different FGFR receptors (FGFR1-4), with alternative splicing of FGFR1-3 generating “b” and “c” isoforms (Powers et al., 2000).

1.2.1 Structure of FGFRs

FGFR family proteins show a structure common to RTKs – they are comprised of an extracellular domain, a trans-membrane region, and an intracellular region (Johnson et al., 1991; Miki et al., 1992; Lemmon and Schlessinger, 2010). The extracellular domain of FGFRs contains three Ig-like domains, referred to as D1-D3. The FGF ligand binds to the region encompassing D2 and D3 domains (Plotnikov et al., 1999). Alternative splicing gives rise to two isoforms of FGFR1-3 with different D3 domains, known as IIIb and IIIc (Johnson et al., 1991). Splicing is regulated at tissue level, with IIIb isoforms found in epithelial lineages and IIIc isoforms found in mesenchymal cell lineages (Werner et al., 1992; Miki et al., 1992; Chellaiah et al., 1994). Alternative splicing alters ligand preferences for each isoform and allows crosstalk between epithelial and mesenchymal tissues (Dell and Williams, 1992). For example, FGF8 expressed in epithelium can activate the FGFR2IIIc isoform that resides in the mesenchyme, but it cannot bind the FGFR2IIIb isoform present in epithelial tissues (Orr-Urtreger et al., 1993). While FGF9 is needed in both the mesenchyme and the epithelium during lung development (del Moral et al., 2006). During limb bud development, FGF10 expressed in the limb bud mesenchyme bulge signals to the nearby epithelial cells of the ectoderm. The signal is received by the epithelial FGFR2IIIb isoform, which induces the formation of the apical ectodermal ridge (AER). FGF8, and later FGF4, 9, 17 are expressed in the developing AER. These ligands are

recognised by mesenchymal FGFR1IIIc and FGFR2IIIc to sustain mesenchyme proliferation in the limb bud and FGF10 expression (Ornitz and Itoh, 2015).

1.2.2 Signalling from FGFRs

Upon ligand binding, FGFR receptors dimerise non-covalently. FGF binding to FGFRs is assisted by heparin sulphate proteoglycans (HSPGs) present on the cell surface (Yayon et al., 1991). The heterotetrameric FGF-FGFR complex is stabilised by ligand-heparin, ligand-receptor, receptor-heparin and receptor-receptor interactions which combine to enable the autotransphosphorylation of kinase domains and activation of the receptor (Schlessinger et al., 2000).

FGFRs can activate several downstream signalling pathways (Figure 1-2). FGFR substrate 2 (FRS2) can bind the juxtamembrane domain of an active FGFR via its PTB domain (Ong et al., 2000). FRS2 is then phosphorylated by the receptor, which leads to recruitment of growth factor receptor-bound 2 (GRB2) and Src homology region 2 domain-containing phosphatase-2 (SHP2) (Hadari et al., 1998). GRB2 further mediates the recruitment of guanosine nucleotide exchange factor (GEF), son-of-sevenless (SOS) (Rozakis-Adcock et al., 1993). This leads to RAS activation, and eventually induction of MAPK/ERK signalling pathway (Dent et al., 1992; Kouhara et al., 1997). GRB2 can also associate with GRB2-associated binding protein 1 (GAB1), which in turn allows recruitment of PI3K and activation of the AKT pathway (Ong et al., 2001). RSK2, PLCγ, STAT, and p38 MAPK pathways can also be activated by FGFRs under certain conditions (Mohammadi et al., 1991; Tanaka et al., 1999; Hart et al., 2000; Kang et al., 2009). Due to their ability to activate cell growth promoting pathways, FGFR family members are prominent oncogenic proteins.

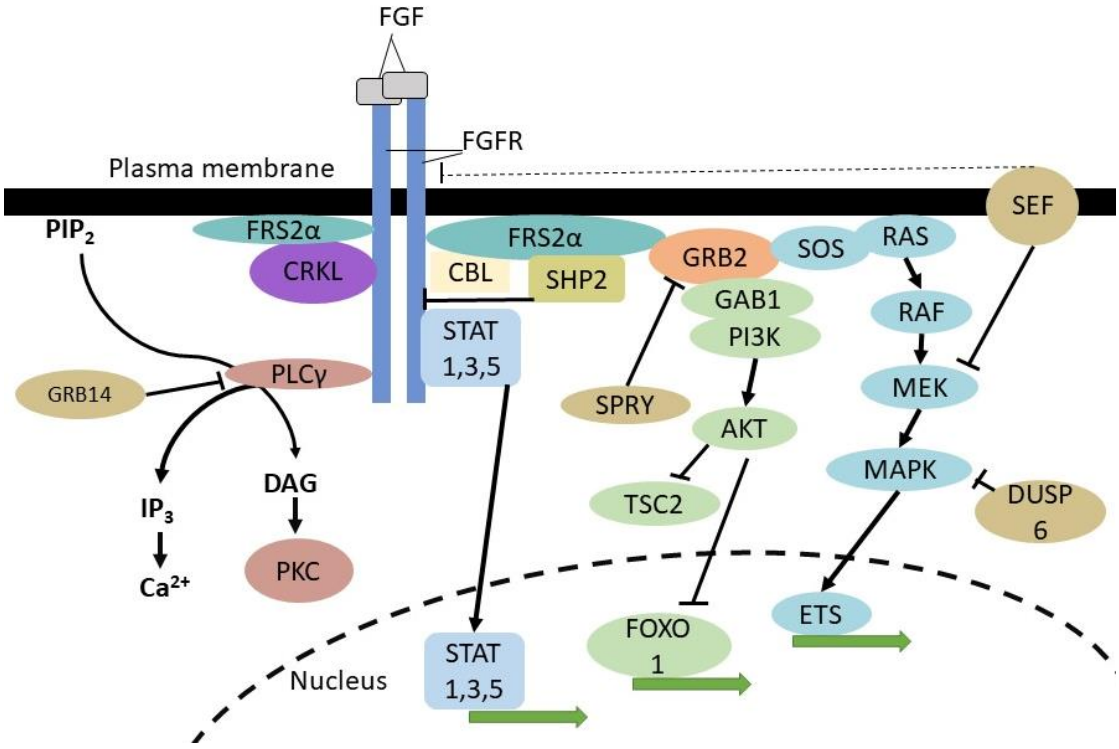


Figure 1-2. The most common FGFR regulated signalling pathways. FRS2α binds to FGFR directly, and recruits GRB2, leading to activation of MAPK and AKT pathways. PLCγ1 recruitment and activity leads to PIP_2 hydrolysis to IP_3 and DAG. The latter can recruit PKC to the plasmamembrane, thus enabling PKC activity, which induces MAPK pathway. CRKL can directly bind active FGFR, but downstream signalling is unknown. FGFRs can also activate STAT signalling pathway. Transcription factors activated downstream of FGFR include STAT transcription factors, and ETS transcription factor family (which includes Etv4, Etv5, Spry, Sef transcription factors). AKT signalling inhibits FOXO1 transcription factor, which promotes cell survival.

1.3 Fibroblast Growth Factor Receptor 2 (FGFR2)

FGFR2 is a member of the FGFR family. It shares the common family structure with other FGFRs. There are two alternatively spliced FGFR2 isoforms: FGFR2IIIb (epithelial) and FGFR2IIIc (mesenchymal), which show different preferences for FGF ligands (Table 1-1) (Orr-Urtreger et al., 1993; Ornitz et al., 1996; Ornitz and Itoh, 2015). This allows for specific crosstalk between epithelial and mesenchymal tissues during development and

organogenesis, including limb bud and lung development (Ornitz and Itoh, 2015).

Table 1-1 Ligand preferences of FGFR2 isoforms

Receptor isoform	FGFR2IIIb	FGFR2IIIc
Ligand	FGF1, FGF3, FGF7, FGF10, FGF22	FGF1, FGF2, FGF4, FGF5, FGF6, FGF8, FGF9, FGF15/19, FGF16, FGF17, FGF18, FGF20

1.3.1 Specific roles in normal cell signalling and diseases

FGFR2 plays a role in development, and aberrant receptor function can lead to congenital disorders. *FGFR2* gain of function missense mutations are found in Apert syndrome, Pfeiffer syndrome, Crouzon syndrome, Jackson-Weiss syndrome, Beare-Stevenson syndrome and Saethre-Chorzen syndrome (Jabs et al., 1994; Gorry et al., 1995; Park et al., 1995; Rutland et al., 1995; Przylepa et al., 1996; Wilkie, 1997; Paznekas et al., 1998). These congenital skeletal disorders are characterised by craniosynostosis (premature fusion of skull sutures) and short-limbed bone dysplasia. During normal development, FGFR2 mediates proliferation, differentiation and apoptosis of osteoprogenitor cells (N. Su et al., 2014). Premature cell differentiation caused by excess signalling downstream of mutant FGFR2 is suspected to cause this premature fusion (Azoury et al., 2017).

FGFR2 has also been found to play a role in a variety of cancers. Mutations have been found in invasive breast cancer, and missense mutations are present in melanoma and endometrial uterine cancer (Hunter et al., 2007; Katoh, 2008; Reintjes et al., 2013). These missense mutations often occur at the D3 Ig-like domain of the receptor, which alters ligand specificity – while normally a

mesenchyme-expressed FGFR2 receptor could only recognise epithelial FGF ligands, altered ligand specificity allows the receptor to recognise mesenchymal FGF, thus creating an autocrine loop where tight control of FGFR2 activation is lost. Alternatively, missense mutations are also detected in the tyrosine kinase domain, which allow receptor activation independent of ligands (Katoh, 2008). *FGFR2* gene amplification is found in triple-negative breast cancer and diffuse-type gastric cancer (Turner and Grose, 2010; Deng et al., 2012). Furthermore, in gastric cancer overexpression of *FGFR2* positively correlated with lymph node metastases (X. Su et al., 2014). In stomach cancer, FGFR2 variant “K-sam”, containing a C-terminal truncation, has been detected (Itoh et al., 1994; Hattori et al., 1996; Ueda et al., 1999). Excessive FGFR2 signalling is the major driver of cancers mentioned above. Several canonical FGFR2 signalling pathways will be discussed in section 1.3.2.

1.3.2 Canonical downstream signalling pathways mediated by FGFR2

1.3.2.1 MAPK1/2 pathway

FGFR2 can regulate cell proliferation, differentiation, migration and death via the MAPK pathway. Once FGFR2 is activated, GRB2 associates with FGFR2-bound, tyrosine-phosphorylated FRS2 via its SH2 domain, and recruits SOS via proline-rich motif – SH3 interactions (Kouhara et al., 1997). This allows SOS to act as a GEF to RAS (Chardin et al., 1993). Activated RAS recruits RAF and leads to its phosphorylation (Marais et al., 1995). Active RAF in turn leads to serine phosphorylation and activation of MEK1/2 (MAP2K1/2), which is followed by phosphorylation of ERK1/2 (MAPK1/2) (Pelech and Sanghera, 1992; Zheng and Guan, 1993). Active ERK1/2 activates cytosolic proteins like

RSK1/2 to perpetuate downstream signalling, and also translocates to the nucleus where it activates transcription factors that regulate the expression of genes needed for cell growth and survival (Yoon and Seger, 2006).

1.3.2.2 PI3K/AKT pathway

FGFR2 signalling can promote cell cycle progression and regulate the balance between cell growth and death via the PI3K/AKT/mTOR pathway (Ornitz and Itoh, 2015). The PI3K regulatory subunit p85 facilitates recruitment to the active receptor and thus to the plasma membrane, while the p110 subunit catalyses phosphatidylinositol-4,5-biphosphate (PIP_2) conversion to phosphatidylinositol-3,4,5-triphosphate (PIP_3) (Whitman et al., 1988). Increased PIP_3 concentration in the membrane leads to PDK-1 and AKT recruitment, resulting in AKT phosphorylation (Alessi et al., 1997; Franke et al., 1997). Active AKT in turn can activate a myriad of proteins that promote cell growth (Kane et al., 1999; Kandel et al., 2002; Basu et al., 2003).

1.4 Non-canonical “tier 2” signalling and its consequences

So far, only canonical FGFR2 signalling has been discussed, which involves an FGF ligand binding to the receptor in order to activate its tyrosine kinase activity. pY residues present on the RTK can recruit SH2 domain-containing proteins, leading to their activation by phosphorylation and facilitation of downstream signalling. However, FGFR2 has been shown to be capable of mediating signalling that occurs under non-stimulated conditions when no ligand is bound to the receptor. In this type of signalling, proline-rich motifs present in the C-terminal tail of FGFR2 play a central role.

The C-terminal region of FGFR2 has important regulatory functions. A FGFR2 splice variant with shortened C-terminal tail has been shown to enhance cell transformation (Cha et al., 2008). FGFR2 lacking the C-terminal tail was shown to be constitutively active, and it has been found in several breast and gastric cancer cases. (Gartside et al., 2009). The functional importance of the tail was shown later.

The C-terminal tail of FGFR2 is comprised of 58 amino acids and contains two proline-rich sequences which include at least one PXXP (where X is any amino acid) motif known to be recognised by SH3 domains. Dimeric GRB2 was shown to bind two FGFR2 molecules at their C-terminus and form a heterotetrameric complex. The interaction is mediated by GRB2 SH3 domain binding to the proline-rich motif which is found at the very C-terminus of the FGFR2 C-terminal tail. In this heterotetramer, phosphorylation of tyrosine residues within the FGFR2 C-terminal region is inhibited, thus abrogating recruitment of proteins via SH2 and PTB domains. The inhibition is released when FGFR2 is stimulated by an extracellular ligand. GRB2 is then phosphorylated and released (Lin et al., 2012).

The phospholipase PLCγ1 can compete for the GRB2 binding proline-rich site on the C-terminal tail of FGFR2. When the GRB2 concentration in cell is reduced, the PLCγ1 SH3 domain binds the FGFR2 proline-rich region and can be activated without phosphorylation under non-simulated conditions (no ligand is bound to FGFR2) (Timsah et al., 2014). Active PLCγ1 catalyses the hydrolysation of PIP_2 to inositol-1,4,5-triphosphate (IP_3) and diacylglycerol (DAG) (Bong and Rhee, 1995). The former assists in calcium ion release from cellular stores, and the latter leads to activation of protein kinase C (PKC) and other downstream signalling pathways (Klint and Claesson-Welsh, 1999; Tiong

et al., 2013). Under non-stimulated conditions and GRB2 depletion, active PLCγ1 reduces PIP_2 levels, which leads to PTEN inhibition and increase in AKT phosphorylation and activation. This promotes cell growth and migration, as well as tumour formation (Timsah et al., 2016).

1.5 Protein domains

CRKL adaptor protein, which will be discussed in section 1.6, contains SH2 and SH3 domains, and no catalytic domains. To better understand how CRKL can function, these domains will be discussed in this section.

1.5.1 SH2 domain

The SH2 domain can recognise short peptide sequences which contain a phosphotyrosine (pY) residue (Sadowski et al., 1986). The domain structure is comprised of two α-helices flanking an antiparallel β-sheet (Huang et al., 2008). An arginine residue within the first binding pocket located on the β-sheet is indispensable for binding, as it is required to form hydrogen bonds with phosphate oxygens of the pY, while the second binding pocket recognises adjacent residues to aid binding affinity and specificity (Zhou et al., 1993; Waksman and Kuriyan, 2004). The residues surrounding pY are very important, as they can greatly enhance domain affinity. Select residues in position range between -6 and +6 from the pY can confer SH2 binding specificity (Filippakopoulos et al., 2009).

1.5.2 SH3 domain

The SH3 domain can recognise short peptide sequences which contain at least one proline residue (Pawson, 1995). The SH3 domain is made up of β-strands which form either two antiparallel β-sheets or a β-barrel, with loops

playing a role in binding partner recognition. The peptide ligand forms a left-handed polyproline type II (PPII-type) helix containing three residues per turn. In order to bind proline-rich motifs two SH3 binding pockets are lined with hydrophobic aromatic residues, while the third and fourth binding pockets are more variable and aid in specificity (Feng et al., 1994). This allows for a huge variety of SH3 domains and the respective proline-rich motifs they recognise (Teyra et al., 2017). The SH3 domain can recognise and bind proline-rich motifs at two different orientations (Lim et al., 1994). There are two main classes of proline-rich motifs recognised by the SH3 domains. Class I motifs contain a positively charged residue (arginine or lysine) N-terminally from the prolines (+xxPxxP), while in class II motifs, the positively charged residue is C-terminal to the prolines (PxxPx+) (Kay et al., 2000). However, studies have shown that different SH3 domains can bind a big variety of different proline-rich motifs, and new classes of motifs are being defined (Teyra et al., 2017)

1.6 Human CRK family

CT-10 Regulator of Kinase (CRK) proteins are a family of adaptor proteins, first discovered in a chicken tumour, where a viral oncogene ($p47^{gag\text{-}crk}$) which increased global tyrosine phosphorylation, even though it lacked catalytic domains (Mayer et al., 1988). CRK proteins are found in many organisms, ranging from pre-metazoans, to *Caenorhabditis elegans* nematode worms, *Drosophila,* mice and humans (Shigeno-Nakazawa et al., 2016). CRK family proteins help regulate cell proliferation, migration and adhesion, apoptosis, gene expression, as well as phagocytosis of apoptotic cells and endocytosis of parasitic organisms, and are indispensable during development (Feller, 2001).

There are three CRK family member proteins in humans. CRK-I and CRK-II are alternatively spliced products of one gene (Matsuda et al., 1992), while CRKL is encoded by a separate gene (ten Hoeve et al., 1993). CRK-II and CRKL share a common SH2-SH3N-SH3C domain structure, with domains connected via linkers, whereas CRK-I lacks the SH3C domain (Figure 1-3) (Reichman et al., 1992). CRK family proteins help facilitate signalling from various receptors, such as RTKs and integrins. The majority of CRK-mediated signalling pathways are thought to follow the same scenario: CRK binds phosphorylated tyrosine residues of upstream proteins via the SH2 domain, while the SH3N domain binds downstream effectors (Birge et al., 2009).

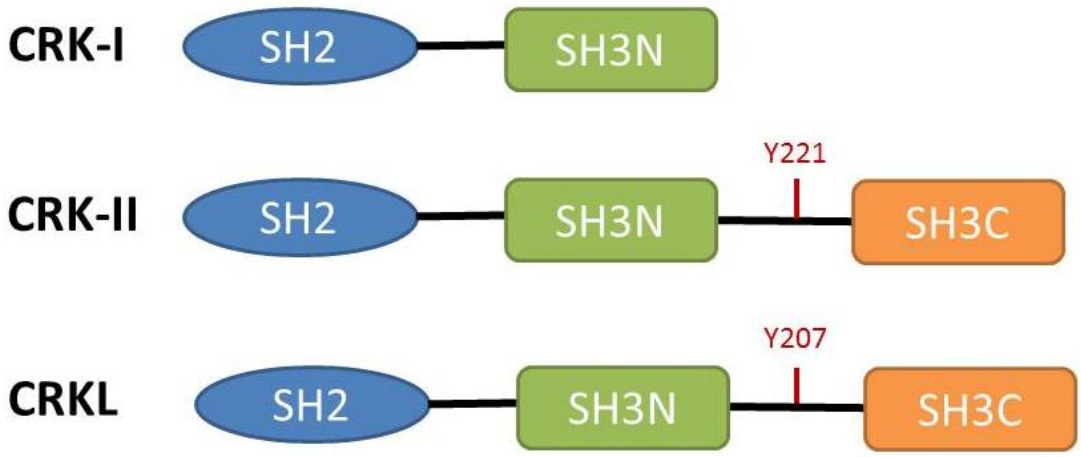


Figure 1-3. Cartoon representation of CRK family protein structures. Y221 and Y207 are commonly researched phosphorylation sites of CRK-II and CRKL.

Despite the common structure and pathways, proteins within the family mostly have unique roles and cannot fully compensate for each other during development. During mouse embryo development, single mutants of either CRKL or CRK-II show distinct lethal developmental defects (Guris et al., 2001; Park et al., 2006). CRK family proteins are essential for cardiac and neural crest development in the embryo, as well as development of neuromuscular synapse

- conditional knockouts in mice showed that together, CRK and CRKL are involved in neuronal migration within the hippocampus, cerebellum and cerebral cortex during brain development (Park and Curran, 2008). CRKL deletion in mice causes cardiovascular and craniofacial defects resulting in embryonic lethality, which were postulated to be caused by aberrant neuron behaviour after their migration from neural crest (Guris et al., 2001). In the adult organism some functions of CRK-II and CRKL can overlap, such as maintenance of cell structure or Reelin signalling pathway (Matsuki et al., 2008; Park and Curran, 2008).

1.6.1 CRKL structure

CRKL shares the SH2-SH3N-SH3C structure of CRK family. CRKL has been found to be important in signalling pathways such as those activated by hematopoietic cytokine receptor activation, T-cell stimulation, and integrin-mediated adhesion (Birge et al., 2009). As an adaptor protein, its function is defined by its binding partners, which will be discussed next.

CRK family SH2 domain recognises and binds the pY-x-x-P motif (Zhou et al., 1993). The SH2 domain is the module which allows CRKL to detect active signalling “inputs” to propagate downstream signalling. The adaptor protein has been shown to bind a variety of proteins with pY residues, including RTK FGFR1, β-integrins, and CBL (Uemura and Griffin, 1999; Moon et al., 2006). Functional consequences of these interactions will be discussed in section 1.6.2.

The SH3N domain is the major proline-rich motif binding module of CRKL. The exact proline-rich motif sequence preference for CRKL has not been well defined, but for the SH3N domain of CRK-II it is P-x-x-P-x-K/R, which

belongs to Class II proline-rich motif family (see section 1.5.2) (Knudsen et al., 1994; Teyra et al., 2017). The SH3N domain has been shown to be essential for normal CRKL function as it facilitates binding to effectors such as DOCK1 and C3G (Birge et al., 2009). It has also been shown to be required for CRKL-mediated oncogenesis in non-small cell lung carcinoma (NSCLC) harbouring 22q11.21 amplification (a region where *CRKL* gene resides) (Cheung et al., 2011).

A stretch of 50 amino acids forms a linker between the SH3N and the SH3C domains. It contains tyrosine residues which are targets for phosphorylation by tyrosine kinases. The most notable and most researched phosphorylation site is Y207 (Figure 1-3), which was first discovered as a BCR/ABL phosphorylation site (Ten Hoeve et al., 1994). So far it is unclear how exactly tyrosine phosphorylation affects CRKL. It has been shown that removal of this phosphorylation site does not affect CRKL binding to binding partners such as SOS or C3G, which bind via the SH3N domain (De Jong et al., 1997). In the paralogue CRK-II, it has been shown that pY221 (equivalent to pY207) acts as an intramolecular binding site to the SH2 domain of the protein. Y221 phosphorylation is autoinhibitory towards CRK-II, as it occludes the SH3N proline-rich motif binding site and prevents the protein from propagating downstream signalling (Rosen et al., 1995). Structural studies have shown that in CRKL, when the SH2 domain is bound to pY207, it is unable to bind other proteins, while the SH3N domain is still able to form interactions with downstream signalling proteins (Kobashigawa and Inagaki, 2012). Regardless of the phosphorylation state of CRKL, the SH3C domain does not interact with the other domains (Jankowski et al., 2012).

The SH3C domain function of CRKL is the most elusive. It shows the canonical beta-barrel fold SH3 domain structure, with an antiparallel three-stranded β-sheet which packs orthogonally against an antiparallel β-sheet consisting of two strands (Harkiolaki et al., 2006). However, it is unable to bind PPII-type sequences, as the SH3C domain does not contain a strictly conserved tryptophan at position 38, which leads to a different binding groove shape and lack of hydrophobic stacking potential (Muralidharan et al., 2006; Jankowski et al., 2012). Furthermore, the domain does not contain negatively charged residues that normally promote binding to positively charged arginine or lysine residues found in many proline-rich motifs (Jankowski et al., 2012).

However, the SH3C domain has been shown to be functionally important. The domain is essential for the adhesion of hematopoietic cells and fibroblast transformation (Senechal et al., 1996). It was also shown to bind CD34, which is a transmembrane protein required for adhesion in lymphohematopoietic cells (Felschow et al., 2001). Later Harkiolaki et al shed some doubt on whether this interaction occurs, but the assays used might not have been optimal for detecting the CD34-CRKL interaction (Harkiolaki et al., 2006). The SH3C domain also contains a nuclear export signal (NES) sequence, which is unsurprising as CRKL has been detected in the nucleus as well as the cytoplasm (Rhodes et al., 2000). It is however unclear how exactly CRKL nuclear import/export is regulated. The paralogue CRK-II has been found to interact with CRM1 (Exportin-1), a nuclear export factor, via the SH3C domain (Smith et al., 2002).

Furthermore, the SH3C domain of CRKL acts as a protein dimerisation domain. SH3C domains from each monomer exchange single N-terminal β-strands to allow dimerisation. During the dimerisation the NES sequence is

exposed. It is unclear what is the physiological importance of NES exposure or CRKL dimerisation (Harkiolaki et al., 2006).

1.6.2 CRKL roles in signalling pathways and disease

1.6.2.1 Upstream receptor binding partners

As an adaptor protein, CRKL can propagate cell signalling by binding to receptors directly. There are several receptor tyrosine kinases which are known to bind CRKL directly to facilitate signalling (Figure 1-4; Figure 1-5). CRKL has been found to be an important downstream effector of anaplastic lymphoma kinase (ALK), which is a receptor tyrosine kinase (An et al., 2016). CRKL is required for ALK-mediated RAP1 activation and subsequent cell proliferation, but it is unclear which CRKL domains can bind the receptor (Schönherr et al., 2010). Unsurprisingly, CRKL is a key signalling protein in NSCLC mediated by EML4-ALK fusion protein that enables cancer cell proliferation and survival (An et al., 2016). An interaction between CRKL and EGFR has been detected as well, although signalling consequences of this interaction are unclear (Ronan et al., 2016). The CRKL SH3N domain can bind the proline-rich region of ABL non-receptor tyrosine kinase (Senechal et al., 1996; Senechal et al., 1998). Unsurprisingly, CRKL is also able to bind the BCR/ABL fusion protein, which is a key oncogenic driver in Philadelphia chromosome-positive CML (Uemura et al., 1997). CRKL has also been seen in FGFR family signalling. Specifically, in cells stimulated by FGF8, FGFR1 and FGFR2 can bind CRKL (Moon et al., 2006). The CRKL interaction with FGFR1 is dependent on pY463 site on the receptor, while FGFR2 binding site is unknown (Seo et al., 2009). In this context, CRKL allows activation of RAC1, CDC42 and PAK, which in turn assist

RAF1 and MEK1 activation by RAS. This enables CRKL-mediated, FGF8 induced cell migration (Figure 1-5) (Seo et al., 2009).

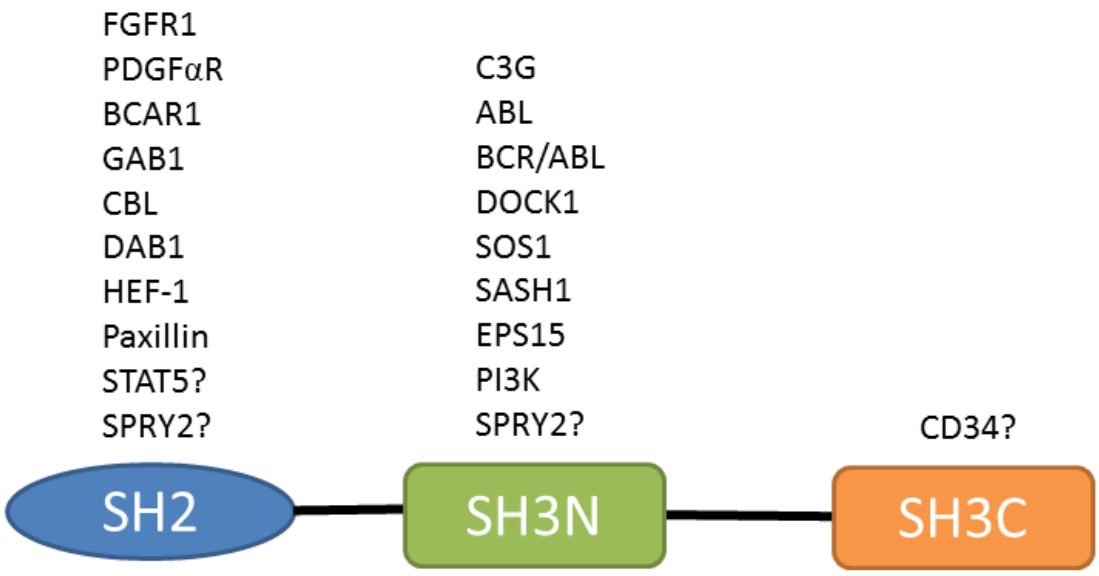


Figure 1-4. Known CRKL binding partners and their binding domains. Several interactions that are yet unclear are noted with a question mark.

1.6.2.2 Other upstream binding partners

In some cases, CRKL is recruited to the active receptor indirectly through scaffolding proteins (Figure 1-4). Scaffold proteins can bind receptors directly and recruit many downstream proteins due to abundance of phosphorylated amino acid sites, various motifs and other binding modules. This allows exponential amplification of downstream signalling. BCAR1 (p130Cas) is the major CRKL-binding scaffold protein. It contains 15 pY-x-x-P motifs, which is the consensus binding motif for CRKL SH2. BCAR1 is found to be highly phosphorylated in cancers, as it can be activated by many transmembrane receptors, integrin-mediated adhesion and even mechanical stretching (Defilippi et al., 2006). In theory, one hyper-phosphorylated BCAR1 could recruit many different CRK/partner complexes, to facilitate cell growth and migration, but it remains to be experimentally confirmed. It has been shown that BCAR1 is able

to recruit ABL to focal adhesions via CRKL (Salgia et al., 1995). BCAR1-bound CRKL can also recruit C3G, as indicated by RAP1 activation (Defilippi et al., 2006). Upon interacting with CRKL, C3G GEF activity is stimulated. C3G aids in removing GDP from RAP1, allowing it to spontaneously reload with GTP and become active (Gotoh et al., 1995). In turn active RAP1 can facilitate downstream signalling. RAP1 signalling can regulate cell growth, cytoskeletal reorganization and cell adhesion (Birge et al., 2009). For example, in hematopoietic cell lines upon erythropoietin and interleukin-3 stimulation, this interaction leads to activation of MAPK/ERK signalling, and subsequent activation of ELK-1 transcription factor. That in turns leads to *c-fos* expression and promotion of cell survival. In this context, only the SH2 and the SH3N domains of CRKL are essential (Nosaka et al., 1999). Interestingly, in some cases CRKL-dependent RAP1 activation is also dependent on SRC (Cheung et al., 2011). GAB1 is another scaffold protein which can recruit CRKL. When phosphorylated by hepatocyte growth factor (HGF)-stimulated c-MET receptor, GAB1 can bind CRKL, leading to C3G recruitment and RAP1 activation (Sakkab et al., 2000). DAB1, a protein essential for Reelin signalling during neuron migration and positioning in brain development, is known to recruit CRKL (Ballif et al., 2004). This leads to recruitment and activation of C3G, resulting in RAP1 downstream signalling activation, and this downstream signalling cascade is essential for Reelin function (Matsuki et al., 2008; Bock and May, 2016).

1.6.2.3 Downstream CRKL binding partners

Recruitment of CRKL to the plasma membrane proximal region by active receptors or scaffold proteins allows the adaptor protein to recruit several different downstream effectors in order to activate different signalling pathways

which result in cell growth, migration and adhesion (Figure 1-4). CRKL can associate with C3G via the SHN3 domain – proline-rich motif interaction, and this complex has been detected in numerous cell types (Feller, 2001).

CRKL can associate with another GEF complex comprised of DOCK1 (Dock180) and ELMO proteins via its SH3N domain. Once this tripartite complex is formed, ELMO is tyrosine-phosphorylated, enabling its GEF function. Active DOCK1/ELMO complex allows RAC1 to release GDP, which leads to GTP binding and subsequent RAC1 activity that promotes cell polarity establishment and migration (Birge et al., 2009).

CBL is another known CRKL binding partner (Ronan et al., 2016). In hematopoietic cells the interaction between CRKL and CBL occurs upon stimulation with thrombopoietin, stem cell factor, and interleukin-2 (Nosaka et al., 1999). It also occurs in BCR/ABL-transformed CML cells (Uemura et al., 1997). HGF-stimulated c-MET RTK can also phosphorylate CBL and facilitate CBL-CRKL interaction. It is thought that CRKL could present proteins to CBL for ubiquitination, and thus protein degradation (Feller, 2001). Several signalling pathways mentioned in these sections are summarised in Figure 1-5.

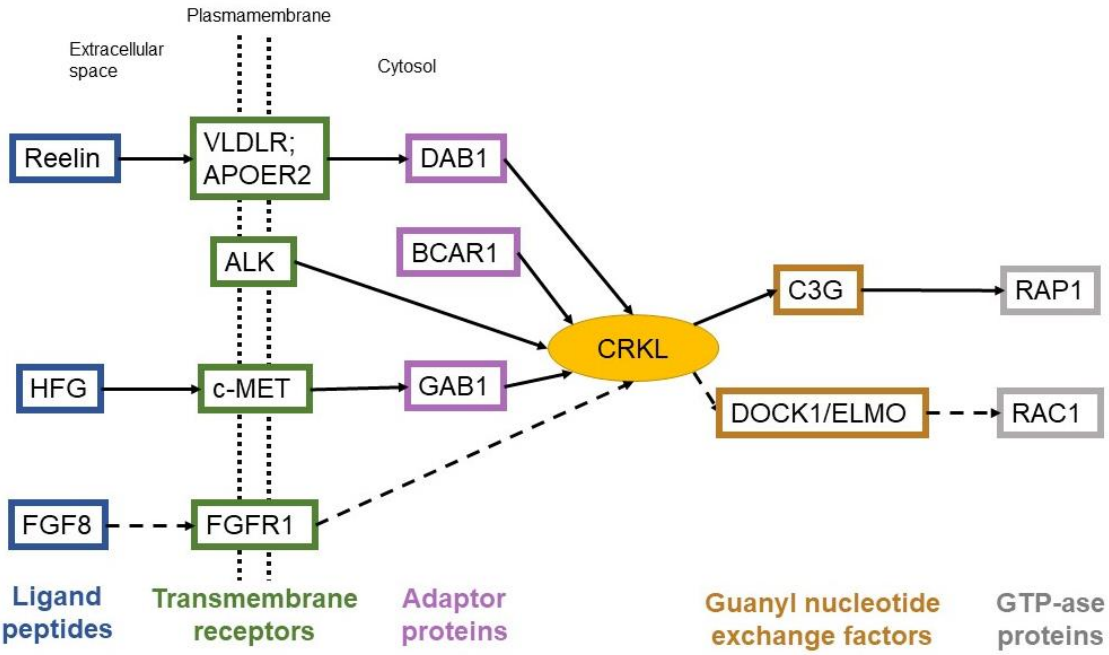


Figure 1-5. A cartoon depiction of several known CRKL signalling pathways. CRKL participates in cell signalling via RAP1 (solid arrows) or RAC1 (dashed arrows) signalling pathways. Different receptors can recruit CRKL directly or via other adaptor proteins. CRKL binds to C3G or DOCK1/ELMO guanyl nucleotide exchange factors, leading to their activation. Resulting RAP1 and RAC1 activity leads to cell proliferation and survival, as well as actin cytoskeleton rearrangement.

1.6.2.4 Negative regulation of CRKL

Not much is known about negative regulation of CRKL. SASH1 tumour suppressor has been shown to inhibit CRKL in cell lines from pancreatic, colorectal and hepatocellular cancer. It can bind the CRKL SH3N domain via a proline-rich motif, which leaves CRKL unable to bind downstream effectors such as C3G (Franke et al., 2019). CBL-B is also a potential negative CRKL regulator. It has been shown that the CRKL-CBL-B interaction reduces CRKL affinity for C3G binding (Liu, 2014).

1.6.2.5 CRKL roles in cancer

CRK family members can facilitate cell transformation. However, CRKL shows stronger oncogenic activity than CRK-II (Kobashigawa and Inagaki,

2012). CRKL has been found to be an essential gene in a number of cell lines derived from small and non-small cell lung carcinomas, CML, lymphocytic leukaemia and glioblastoma (Luo et al., 2008; Lin et al., 2015). Overexpression has been found in invasive ductal carcinoma (breast cancer) and ovarian cancer (Zhao et al., 2013), pancreatic cancer (Fu et al., 2015), and endometrial cancer (Cai et al., 2017). CRKL was shown to have a prominent role in rhabdomyosarcoma (RMS), where it shows high expression in RMS cell lines, xenografts and human tumour samples. (Yeung et al., 2013). CRKL has a very important role in CML, as deletion of CRKL-binding proline-rich motif in BCR/ABL, the main driver of CML, reduces fibroblast transforming activity by half (Senechal et al., 1996). CRKL is also very highly phosphorylated (pY207) in Philadelphia-positive CML patients (Oda et al., 1994). CRKL phosphorylation status is a prognostic marker in patients treated with imatinib, a BCR/ABL inhibitor (La Rosée et al., 2008).

Recently CRKL was found to be overexpressed in laryngeal squamous cell carcinoma (Kostrzewska-Poczekaj et al., 2020). Overexpression has been found to correlate with tumour stage and decreased patient survival (Bian et al., 2018; Kostrzewska-Poczekaj et al., 2020). CRKL expression is also higher in metastases compared to primary tumour (Lin et al., 2015). Interestingly, CRKL can facilitate oncogenic signalling even when protein levels are normal. In the case of ALK-rearranged NSCLC, where the upstream activator (ALK) of CRKL is overexpressed, CRKL contributes to cancer signalling even at normal levels, as over-active ALK was enough to result in increased signalling downstream of CRKL (An et al., 2016).

To summarise, the CRKL adaptor protein has roles in a variety of signalling pathways depending on cellular context. Regardless of the upstream

recruiter, the most common CRKL-recruited proteins are C3G and DOCK1/ELMO GEFs which activate RAP1 and RAC1 respectively. There is a plethora of unexplored upstream and downstream roles of CRKL which should be studied due to the prevalent role of CRKL in cancer.

1.7 Aims of the project

The “tier 2” signalling mediated by the RTK proline-rich motif interactions with SH3 domain containing proteins in the absence of extracellular ligand binding has been implicated to have roles in normal cell signalling and cancer (Timsah et al., 2014; Timsah et al., 2016). It is imperative to gain more insight and see whether the interplay between FGFR2, GRB2 and PLCγ1 in absence of extracellular stimulants is a unique feature of these proteins, or whether the “tier 2” signalling can also occur with other proteins in human cells. Proline-rich motifs are present in a number of receptor tyrosine kinases. If present at the C-terminal tail of the RTK, they are likely to be easily accessible to other proteins and act as “sticky arms” (Williamson, 1994). Thus, many receptors (such as EGFR, INSR, RET, etc) can potentially employ “tier 2” signalling. Furthermore, it is also unknown whether this type of signalling is evolutionarily conserved in other organisms. The focus of this project was to gain further insight into this “tier 2” signalling phenomenon.

The first aim of this project was to establish a proline-rich motif – SH3 domain-based interaction between CRKL adaptor protein and RTK proteins. First, CRKL binding to proline-rich C-terminal tails of 53 receptor tyrosine kinase proteins was investigated using a dot blot technique which evaluated direct interaction between purified receptor tyrosine kinase C-terminal tails and CRKL SH3 domains. Next, I focused on FGFR2 was to investigate further the

interaction between FGFR2 and CRKL using biophysical methods and pulldown experiments. The effects of this interaction on cell migration and downstream signalling were also investigated.

The second aim of this project addressed whether “tier 2” signalling is conserved in other organisms. This was done by assessing the *in vivo* consequences of the FGFR2 and CRKL interaction using the *Caenorhabditis elegans* model organism. The CRISPR/Cas9 gene editing technique was used to produce mutants of *egl-15* and *ced-2*, the sole FGFR and CRK orthologues in *C. elegans* respectively. The impact of deletion of the EGL-15 (FGFR2) proline-rich motifs and CED-2 (CRKL) SH3 domains were investigated to observe their effects on apoptotic germline cell engulfment, DTC migration and fecundity.

Chapter 2 Methods

2.1 Plasmid DNA preparation

Bacteria were grown in 5 mL of LB broth (Merck) with appropriate antibiotic overnight in a 37ºC shaking incubator. The next day bacteria were spun down at 15,000 RPM for 10 minutes in a tabletop centrifuge. The pellet was retained. QIAprep spin Miniprep kit (Qiagen) was used to extract plasmids according to manufacturer's instructions.

2.2 Bacteria transformation

20 µL of competent DH5α bacteria (New England Biolabs (NEB)) were thawed on ice. 100 ng of plasmid DNA was added, and the bacteria were incubated on ice for 30 minutes. Bacteria were then heat-shocked in 42ºC water bath for 45 seconds, and chilled on ice for 2 minutes. 500 µL of room-temperature LB broth media was added, and bacteria were incubated at 37ºC in a shaking incubator for 1 hour. Bacteria were then spread on a pre-warmed LB agar plate containing appropriate antibiotics and incubated at 37ºC overnight.

2.3 Site-directed mutagenesis

Mutations in plasmids were introduced using the PCR-based site-directed mutagenesis method. Primers were designed using PrimerX website: each primer contained 11-15bp regions flanking the mutation, with up to 60% G/C content (with the exception of one set of primers with 80% G/C content). Reagents, including plasmids and mutagenic primers were put together in a PCR tube (Table 2-1)

Table 2-1. Site-directed mutagenesis reaction mixture

Reagent	Volume (µL)
dNTP (20 mM) (NEB)	1
Nuclease-free water (IDT)	11
Betaine (Sigma)	5
10x reaction buffer (Agilent)	2.5
Plasmid template (250 ng/µL)	1
Forward primer (0.6 µM) (IDT)	1.5
Reverse primer (0.6 µM) (IDT)	1.5
Pfu Ultra II (Agilent)	2

PCR programmes and DNA primers (with mutation nucleotides in bold) used for site-directed mutagenesis can be found in Appendix A.

PCR was followed by an addition of 1 µL DpnI (NEB) restriction enzyme and at least 1 hour incubation at 37ºC to digest methylated parental DNA. Presence of plasmid DNA was detected by running the reaction on an agarose gel. The plasmid DNA was then transformed into DH5α *E. coli* bacteria.

2.4 Human cell maintenance

HEK293T and SkBr3 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco). H520 cell line was grown in RPMI medium (Gibco). All maintenance media was supplemented with 10% Foetal Bovine Serum (FBS) (Sigma-Aldrich) and 1% Penicillin-Streptomycin (Gibco). The cells were maintained in tissue culture flasks (Corning) and kept in a 37ºC humidified incubator with 5% CO_2 and passaged two times a week, at a ratio of 1:10 for HEK293T cell line, and a ratio of 1:4 for SkBr3 and H520 cell lines.

2.5 Knockdown cell line generation

The cells were seeded onto 6-well plates to achieve 50-70% confluency on the day of transfection. 2 ml of media (with FBS, but without penicillin-streptomycin) was added into wells, along with 2 µL of Polybrene (Santa Cruz) and 20 µL of viral particle solution containing either shRNA against CRKL or control shRNA (viral particles were kind gift from Dr Chi-Chuan Lin, University of Leeds, while proprietary shRNA constructs were purchased from Santa Cruz Biotechnology). The plates were incubated overnight. Next day, media was replaced with fully supplemented media. Transduced cells were selected using puromycin (2.5 µg/ml for SkBr3 cell line, 10 µg/ml for HEK293T cell line).

2.6 Transient transfection of human cell lines

Lipofectamine 2000 (Thermo Fisher) was used to transfect human cell lines according to manufacturer's instructions. In brief, 1-10 µg of plasmid DNA was mixed with equal volume of Lipofectamine 2000 in 400 µL OPTI-MEM (Gibco) medium and incubated at room temperature for up to 20 minutes before being added to cells which were at 80% confluency. The cells were incubated overnight. The media was replaced the next day and after confirming protein expression, the cells were used for experiments.

2.7 Human cell starvation/stimulation and lysis

To starve cells, 70%-90% confluent cells were grown in serum-free DMEM medium for 12-17 hours. The cells were then either kept as starved or stimulated with DMEM (+FBS) or FGF9 (Biotechne) (0.1 µg/ml) for 15 minutes. For basal conditions, cells were grown overnight with DMEM medium containing

1% FBS. The cell medium was then removed, and the cells were washed with ice-cold Dulbecco's Phosphate Buffered Saline (DPBS) (Sigma). Cell lysis buffer (20 mM HEPES pH 7.5 (Sigma-Aldrich), 150 mM NaCl (Sigma-Aldrich), 1% NP-40 (Sigma-Aldrich), 1 mM Na_3VO_4 (Sigma-Aldrich), 10% glycerol (Thermo), Pierce Protease and Phosphatase Inhibitor (Thermo)) was added to the plates, and the cells were scraped into an Eppendorf tube and left to lyse on ice for 10 minutes. The tubes were then spun down at 15,000 RPM for 10 minutes in a tabletop centrifuge to spin down the debris. The supernatant was transferred into a fresh Eppendorf tube and either used for experiments or stored at -80^{o}C.

2.8 Protein concentration measurement

Coomassie Assay reagent (Thermo Scientific) was used to determine protein concentration in cell lysates, and the absorbance was measured at 595 nm. Purified protein concentration was measured using Nanodrop 2000 (Thermo Scientific), and the absorbance was measured at 280 nm.

2.9 SDS-PAGE

20-50 µg of protein was used for analysis. The samples were mixed with 2x sample buffer (Biorad) and boiled in a 95^{o}C heat block for 10 minutes. The samples were loaded onto 4-20% mini-PROTEAN precast gels (Biorad), along with a size marker (Color Prestained Protein Standard, Broad Range. NEB). 1x SDS-PAGE running buffer (10x: 30.3 g Tris-Base (Sigma-Aldrich), 144 g glycine (Sigma-Aldrich), 10 g SDS (Sigma-Aldrich) in 1 L distilled H_2O; pH 8.3) was used to fill the tank. The gels were run at 140 V for 1 hour.

2.10 Western blot

Proteins on an SDS-PAGE gel were transferred onto a PVDF membrane (Biorad) using a wet transfer method. Sponges and filter papers were soaked in 1x transfer buffer (10x: 30.3 g Tris-Base (Sigma-Aldrich), 144 g glycine (Sigma-Aldrich) in 1 L distilled H_2O, pH 8.3; 1x: 100 ml 10x transfer buffer, 200 ml of methanol (Sigma-Aldrich); 700 ml distilled H_2O). PVDF membrane was activated by soaking in methanol. Transfer was run at constant 100 V for 1 hour. Alternatively, proteins were transferred onto a membrane with the iBlot2 (Thermo) dry blotting system using PVDF stacks. If needed, the membrane was stained with Ponceau S (Sigma) and visualised with G:BOX (Syngene). The stain was washed off with TBST (20 mM Tris (Sigma-Aldrich), 150 mM NaCl, 0.1% Tween 20 (ChemCruz)) and blocked using 5% BSA (Thermo) in TBST for 1 hour at room temperature. Primary antibody was diluted in 5% BSA in TBST (1:1000 ratio) and the membrane was incubated with primary antibody overnight at 4^{o}C with rocking. The next day membrane was washed by with TBST for 10 minutes on a rocker. After three washes membrane was incubated with a solution of HRP-conjugated secondary antibody (1:2000 ratio) in 5% BSA in TBST for 1 hour at room temperature. Membrane was then washed three times with TBST. Proteins were detected using Clarity Western ECL substrate (Biorad) to visualise HRP-conjugated secondary antibodies. G:BOX (Syngene) was used to capture images. Antibodies used for western blotting are listed in Appendix A.

2.11 Co-immunoprecipitation

1 µg of anti-CRKL rabbit antibody was added to the cell lysate containing 0.5-1 mg of protein in 500 µl total volume. 1 µg of normal rabbit IgG was used as IgG control. The samples were incubated overnight at 4^{o}C with rotation. The following day 20 µL of Protein A/G PLUS Agarose (Santa Cruz Biotechnology) was added to each tube and incubated at 4^{o}C with rotation for 3 hours. Beads were washed by adding 100 µl of cell lysis buffer and spinning down at 4^{o}C, 4000 RPM for 5 minutes in a tabletop centrifuge. The washing step was repeated three times. After washing and removing supernatant, 20 µL of sample loading buffer was added to each tube. Proteins were then eluted by boiling for 10 minutes and analysed using SDS-PAGE and western blot.

2.12 Protein expression and purification

Starter cultures of bacteria inoculated from glycerol stocks were grown overnight in a 37^{o}C shaking incubator with appropriate antibiotic. The following day the starter culture was inoculated into the final culture volume of LB broth (Merck) containing selective antibiotic. The flasks were incubated at 37^{o}C with shaking until optical density at 600 nm reached 0.6-1.0. Protein expression was induced by adding 0.1 mM IPTG. Following IPTG addition bacteria were incubated overnight in a 20^{o}C shaking incubator. The next day cells were pelleted and resuspended in cell lysis buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1mM β-mercaptoethanol (Sigma-Aldrich)). The lysate was then sonicated on ice, and debris was spun down at high speed for 30 minutes.

For small-scale protein purification, 50 ml of bacteria culture was used. To purify GST-tagged proteins, Glutathione Sepharose beads (GE healthcare)

were used. 400 µL of bead slurry was put in a column and washed with 1 ml of distilled water, followed by a wash with 2 ml of cell lysis buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1mM β-mercaptoethanol (Sigma-Aldrich)). Bacterial cell lysate was then added to the column, and the column was incubated for 1 hour with rotation at room temperature. All unbound protein was washed off with the cell lysis buffer. To elute GST-tagged protein, 1 ml of buffer B (20mM Tris pH 8.0, 150 mM NaCl, 20 mM glutathione (Sigma-Aldrich), 5 mM β-mercaptoethanol) was added to the column and allowed to elute. The eluate was then dialysed in 500 ml of buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich)) and left stirring at 4^{o}C for 4 hours. This was repeated once. Concentration of the purified protein was then determined.

For large-scale protein purification, 1 L of bacteria culture was used. Proteins were purified from bacteria cell lysate using Akta protein purification system (purification done by Dr Chi-Chuan Lin, University of Leeds).

For proteins used for MST, the GST tag was cleaved off using thrombin (1 unit/µL). Thrombin was added to purified protein and incubated overnight. The next day, proteins of interest were obtained using Akta protein purification system (work done by Dr Chi-Chuan Lin, University of Leeds).

2.13 GST- pulldowns

Glutathione Sepharose beads were used for GST-pulldowns. Beads were washed by adding five gel bed volumes of lysis buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1mM β-mercaptoethanol (Sigma-Aldrich)) and spun down in a tabletop centrifuge at 4000 RPM for 5 minutes. The wash was repeated two

more times, followed by addition of bacterial cell lysate and incubation overnight at 4^{o}C with rotation. The following day the beads were washed five times with the lysis buffer. For buffer exchange, mammalian cell lysis buffer was used to wash beads. Mammalian cell lysate containing 0.5-1 mg of protein was then added to the beads and incubated at 4^{o}C overnight with rotation. The beads were then washed three times with mammalian cell lysis buffer. Two gel bed volumes of 2xSDS loading buffer were added to the beads, and the tubes were put in a 95^{o}C heating block for 10 minutes to denature protein. The supernatant was analysed by SDS-PAGE and western blot.

2.14 In vitro kinase assay

Glutathione Sepharose beads were used for this assay. GST-CRKL protein was bound to the beads as described previously (2.13). For buffer exchange, a buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM β-mercaptoethanol, 0.005% Triton X-100 (Sigma-Aldrich)) was used to wash the beads once. 100 ng of purified FGFR2cyto (kind gift from Dr Chi-Chuan Lin) was added to each sample, along with 5mM ATP (Sigma-Aldrich) and 10 mM $MgCl_2$ (Sigma-Aldrich) for phosphorylation samples; and 1 µL of recombinant shrimp alkaline phosphatase (rSAP, NEB) for unphosphorylated control to ensure dephosphorylation. The tubes were incubated overnight at 4^{o}C with rotation. The following day the beads were washed three times with the buffer. The sample was then boiled and analysed by SDS-PAGE and western blot.

2.15 *In vitro* kinetic kinase assay

The assay was performed by Dr Chi-Chuan Lin, similar to *in vitro* kinase assay described above (2.14). FGFR2cyto concentration was 100 nM, CRKL concentration was 10 µM.

2.16 In vitro protein phosphorylation for mass spectrometry

GST-CRKL was immobilised onto Glutathione Sepharose beads and phosphorylated using FGFR2cyto (as described in 2.14) in a tabletop protein purification column. FGFR2cyto was then washed off. Non-phosphorylated and phosphorylated GST-CRKL was eluted and purified as described previously (2.12). The samples were concentrated by centrifugation with Amicon Ultra centrifugal filters (Merck), and pure protein was submitted to Leeds Mass Spectrometry facility for analysis (analysis carried out by Rachel George).

2.17 Dot blot

2 µL of purified MBP-tagged C-terminal tail peptides (kind gift from Dr Chi-Chuan Lin) were spotted onto a nitrocellulose membrane. The membrane was left to dry and then blocked with 3% BSA in HBST (140 mM NaCl (Sigma-Aldrich), 1.5 mM Na_2HPO_4 (Sigma-Aldrich), 50 mM HEPES pH 7.5 (Fisher)) for 30 minutes. Solution of 1 µg/ml prey protein in HBST was added onto the membrane and incubated for 2 hours at room temperature. The membrane was washed with HBST for 10 minutes three times. Primary antibody (in 3% BSA in HBST) was added, and the membrane was incubated at room temperature for 2 hours. The membrane was then washed again three times and incubated with a secondary antibody (in 3% BSA in HBST) for 1 hour at room temperature. After

three washes with HBST, the protein signal was detected using Clarity ECL substrate (Biorad) and imaged with G:Box (Syngene).

2.18 Microscale thermophoresis

For MST, buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 0.5 mM TCEP was used. In each tube, concentration of Atto^{488} labelled protein was constant (100 nM), while the second unlabelled protein was serially diluted over 16 dilution series. Solution with labelled and unlabelled proteins was transferred to capillaries (Monolith) and inserted into MST instrument (Monolith NT.115) for detection. Blue LED was used for MST, with power setting adjusted to obtain 200-1600 fluorescence counts. Each scan ran for 3 repeats.

2.19 Wound-healing assay

HEK293T cells growing in a 10 cm culture dish were transfected as described previously (2.6). The next day cells were seeded in a 96-well plate and left to settle overnight in serum-containing media. The following day, scratches were made in each well using Woundmaker96, and wells were washed once with DPBS. The media was changed to DMEM (+1% FBS) to mimic basal conditions. As the interest of this project was to look at FGFR2/CRKL signalling under non-stimulating conditions, some experiments were done under basal conditions. 1% FBS was used for these conditions, rather than total starvation in media without FBS, because *in vivo*, minor presence of growth factors is expected. 0.03 µM SU5402 (Sigma-Aldrich) RTK inhibitor was also used where noted (with equivalent volume of DMSO in "-SU5402" controls). Cell migration was captured for 12 hours using Incucyte system (Essen Bioscience) under normal cell maintenance incubation

conditions as described before (2.4). Image J was used to determine the area of the wound for data analysis.

2.20 RAP1 and RAC1 activity assay

Cell lysates were prepared as described (2.7). Active RAP1 and RAC1 kits (Cell Signalling Technology) were used for GTP-bound RAP1 and RAC1 detection according to manufacturer's instructions.

2.21 RPPA assay

SkBr3 cells were grown in 10 cm dishes until 80-90% confluent. Media was then replaced with DMEM (1% FBS) (for basal conditions), or DMEM (-FBS) for starved/stimulated conditions. 0.03 μM SU5402 (or equal volume of DMSO as control) was also added. The cells were left to incubate overnight. The next day, cells with stimulated conditions were incubated with FGF9 for 15 minutes. Cell lysates were obtained as described before (2.7), but a different cell lysis buffer was used (Table 2-2).

Table 2-2. RPPA cell lysis buffer

2x buffer solution	
Reagent	**Volume (ml)**
Triton X-100	5
0.5 M HEPES pH 7.4	50
0.5 M EGTA (Sigma) pH 7.5-8	1
1 M NaCl	75
1 M $MgCl_2$	0.75
1x lysis buffer	
Reagent	**Volume (ml)**
2x buffer solution	5
100 mM Na_3VO_4	0.1
100 mM TSPP (Sigma-Aldrich)	1
1 M NaF (Merck)	1
Pierce Protease and Phosphatase Inhibitor	1 tablet
Glycerol	1
Distilled H_2O	1.9

Cell lysates containing 2 mg/ml protein were used for the RPPA assay. The assay was done by Host and Tumour Profiling unit in Cancer Research UK Edinburgh Centre at University of Edinburgh, utilising Quanterix and Innopsys RPPA platforms. Antibodies used in this assay are listed in Appendix C.

2.22 RPPA data analysis

The initial data normalisation was done by Kenneth Macleod (Host and Tumour Profiling unit in Cancer Research UK Edinburgh Centre at University of Edinburgh), with protein-specific signal being normalised to total protein signal to produce a ratio of protein abundance. For phospho:total heatmap, ratio of phosphorylated to total protein was calculated from protein abundance data. For heatmap generation, ratio was converted to Z-score, with an average Z-score then generated for each antibody and condition from three repeats. Heatmaps were generated with R studio, using the pheatmap package.

Chapter 3 Investigating the interaction between FGFR2 and CRKL

3.1 Introduction

In absence of extracellular stimulation, “tier 2” signalling, based on interactions between RTK proline-rich motifs and SH3 domains of binding proteins, can regulate cell signalling, as demonstrated by interplay between FGFR2, GRB2 and PLCγ1 (Timsah et al., 2014). It is important to find more proteins which can behave in a similar way to gain more insight into this type of signalling. A good candidate for this Is human CRKL adapter protein. It is known to bind several tyrosine kinases via its SH3 domain, with ABL being the best documented (Senechal et al., 1996; Uemura et al., 1997). Downstream, CRKL can activate RAP1 and RAC1, GTPase proteins which are responsible for induction of cell growth, adhesion and motility (Gotoh et al., 1995; Hasegawa et al., 1996). Due to this, CRKL is a known oncogene with an important role in tumour metastasis. CRKL overexpression has been seen in many different cancer types and correlates with poor prognosis (see section 1.6.2.5). However, knowledge of the exact CRKL function and binding partners is lacking. It is crucial to uncover more CRKL binding partners, which could allow development of better cancer therapeutics. The goal of this Chapter is to find SH3 domain mediated CRKL interactions with RTKs, establish their biochemical features and physiological consequences.

3.2 A screen for CRKL SH3-binding RTKs

To identify whether CRKL was able to bind to a subset of RTKs, a screen using the dot-blot technique was performed. RTK C-terminal tails tagged with

maltose-binding protein (MBP) tag were cloned, expressed and purified by Dr Chi-Chuan Lin (University of Leeds). MBP-tagged peptides were dotted onto a nitrocellulose membrane and incubated with GST-tagged CRKL^{2xSH3} peptide (residues 115-303), which contains SH3N and SH3C domains of CRKL, but not the SH2 domain, to ensure that the potential interaction is mediated by the SH3 domains. Binding was then visualised using an anti-GST antibody. Dot intensity visualisation by western blotting showed that the CRKL^{2xSH3} construct can bind a variety of peptides with differing affinities (Figure 3-1). The CRKL^{2xSH3} construct bound to the C-terminal tails of a number of RTKs, but not to the MBP-tag alone (positions 50-53). The most prominent interaction was observed with LMTK1 (AATK), a protein which has been proposed to function in neuron development (Raghunath et al., 2000). The second strongest interactor was Ephrin receptor EphA4, which has roles in development (Richter et al., 2007); and with ROS1, another RTK involved in development (Acquaviva et al., 2009).

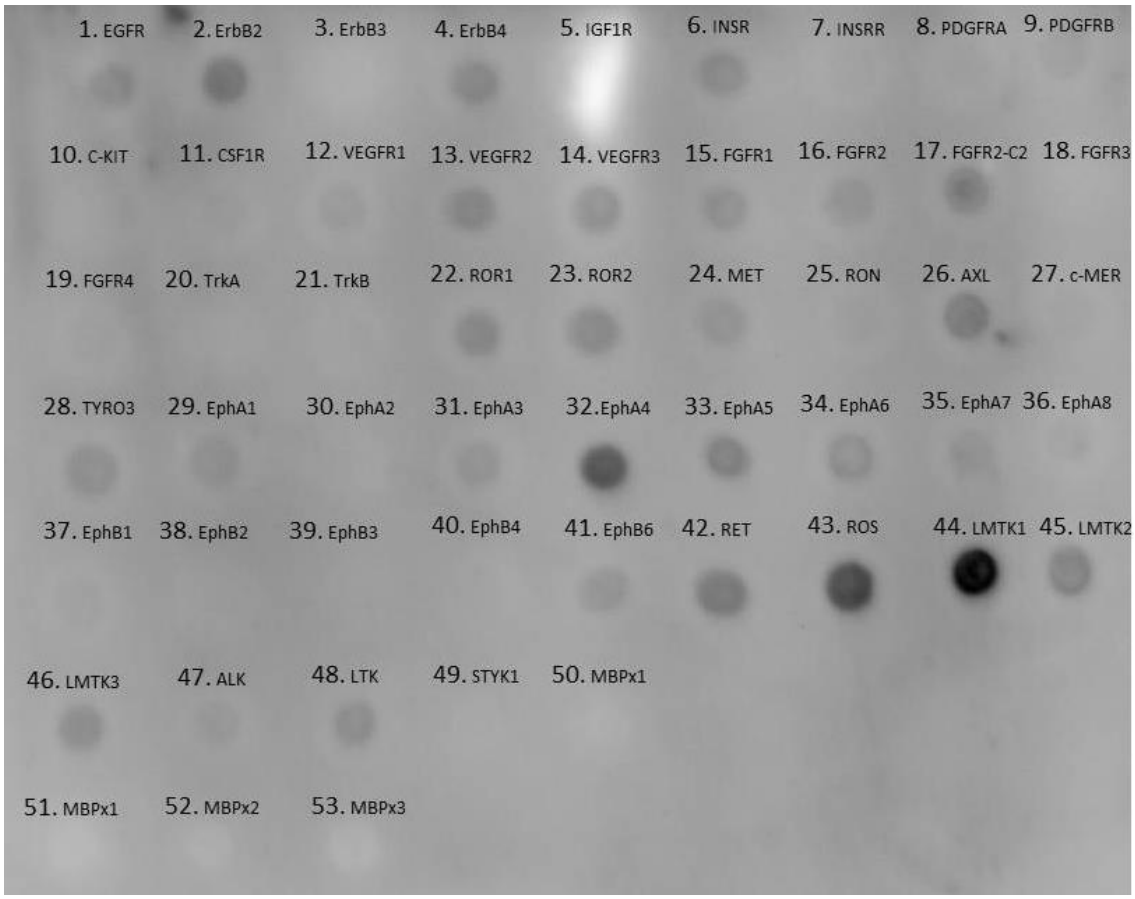


Figure 3-1. Dot blot screen of CRKL SH3 interactions. MBP-tagged C-terminal tails of protein kinases (numbered) were dotted onto the membrane and incubated with GST-CRKL^{2xSH3} construct. Binding was detected using anti-GST antibodies.

The dot-blot also revealed that CRKL can bind to the FGFR2 C-terminal tail, as well as the C-terminal tail of FGFR2-C2, which is an FGFR2 splice variant known as “K-sam C2”. This variant contains a shorter, modified C-terminal tail and is expressed in gastric cancer (Itoh et al., 1994; Hattori et al., 1996; Ueda et al., 1999). FGFR2 was chosen for further experiments for several reasons. It has been shown previously that the FGFR2 C-terminus is involved in cell signalling under basal conditions, as well as cancer (Gartside et al., 2009; Ahmed et al., 2010; Lin et al., 2012; Timsah et al., 2014). CRKL is also known to bind FGFR2, but the interacting domains have not been discerned (Seo et al., 2009). Furthermore, *Caenorhabditis elegans* orthologues of both proteins

are known to physically interact, which provided the opportunity for the potential use of the nematode as a model organism to investigate *in vivo* relevance of the interaction (Lo et al., 2010).

3.3 CRKL binds FGFR2 *in vitro*

In order to see how FGFR2 and CRKL interaction affects cell signalling, it was important to confirm that these proteins interact in cells. To identify a suitable cell line for experimentation the following lines were tested: H520 NSCLC cell line, SkBr3 breast cancer cell line, and HeLa cervical cancer cells. These cell lines were clinically relevant, as CRKL has been reported to be overexpressed in cancer types that these cells were derived from (Zhao et al., 2005; Kim et al., 2010; Zhao et al., 2013). The presence of endogenous FGFR2 and CRKL proteins in the cell lysates of H520, SkBr3 and HeLa cells was confirmed using western blotting (Figure 3-2). H520 and SkBr3 cells were chosen to be used in further experiments due to strong levels of FGFR2 and CRKL expression and their clinical relevance.

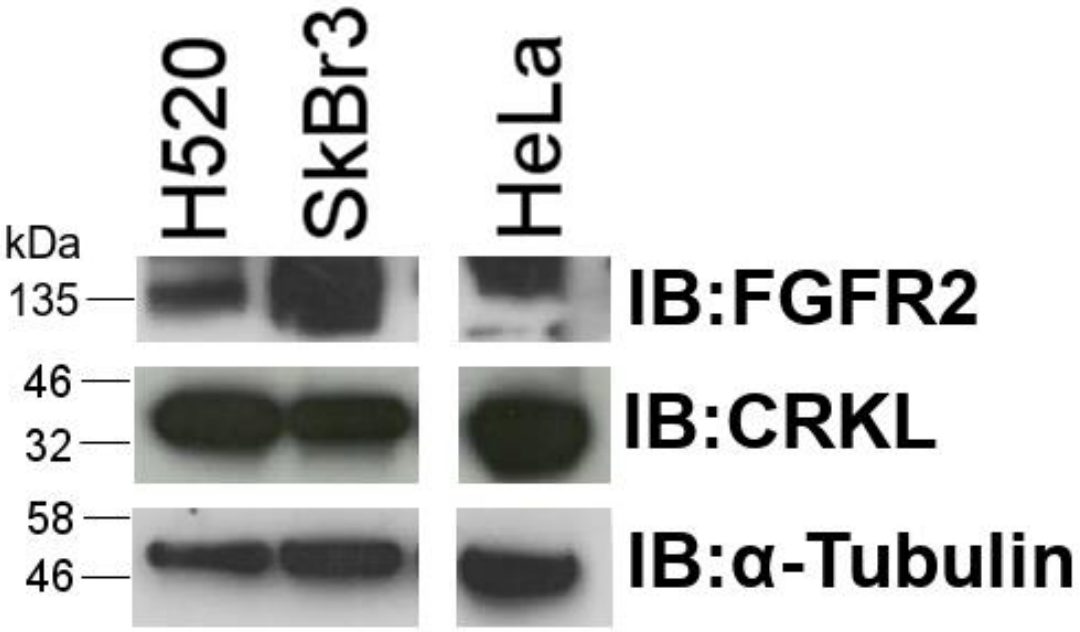


Figure 3-2. Expression levels of FGFR2 and CRKL in different cell lines. 50 µg of protein was run in each lane, and protein levels were detected by western blotting.

The interaction between full length CRKL and full length FGFR2 in cells was confirmed by co-immunoprecipitation (co-IP). Anti-CRKL antibody was incubated with whole cell lysate and then immobilised onto protein A/G agarose beads. IP eluent was then run on a western blot and probed for FGFR2 and CRKL. The anti-CRKL antibody precipitated CRKL in a complex with FGFR2 in both H520 and SkBr3 cells under serum-starved conditions or in the presence of FGF9 (Figure 3-3). Note that while in H520 cell lysates there are two bands detected by anti-FGFR2 antibody, SkBr3 cell lysates show one band, indicating that the protein could be differently post-translationally modified in different cell types. Furthermore, it was difficult to detect FGFR2 in the input controls. Although 20 µg of protein in the input sample was enough to detect CRKL, perhaps FGFR2 was not detected due to its low levels in the lysate, as transmembrane proteins are difficult to solubilise. The results indicate that

FGFR2 and CRKL can be found within the same complex in H520 and SkBr3 cells.

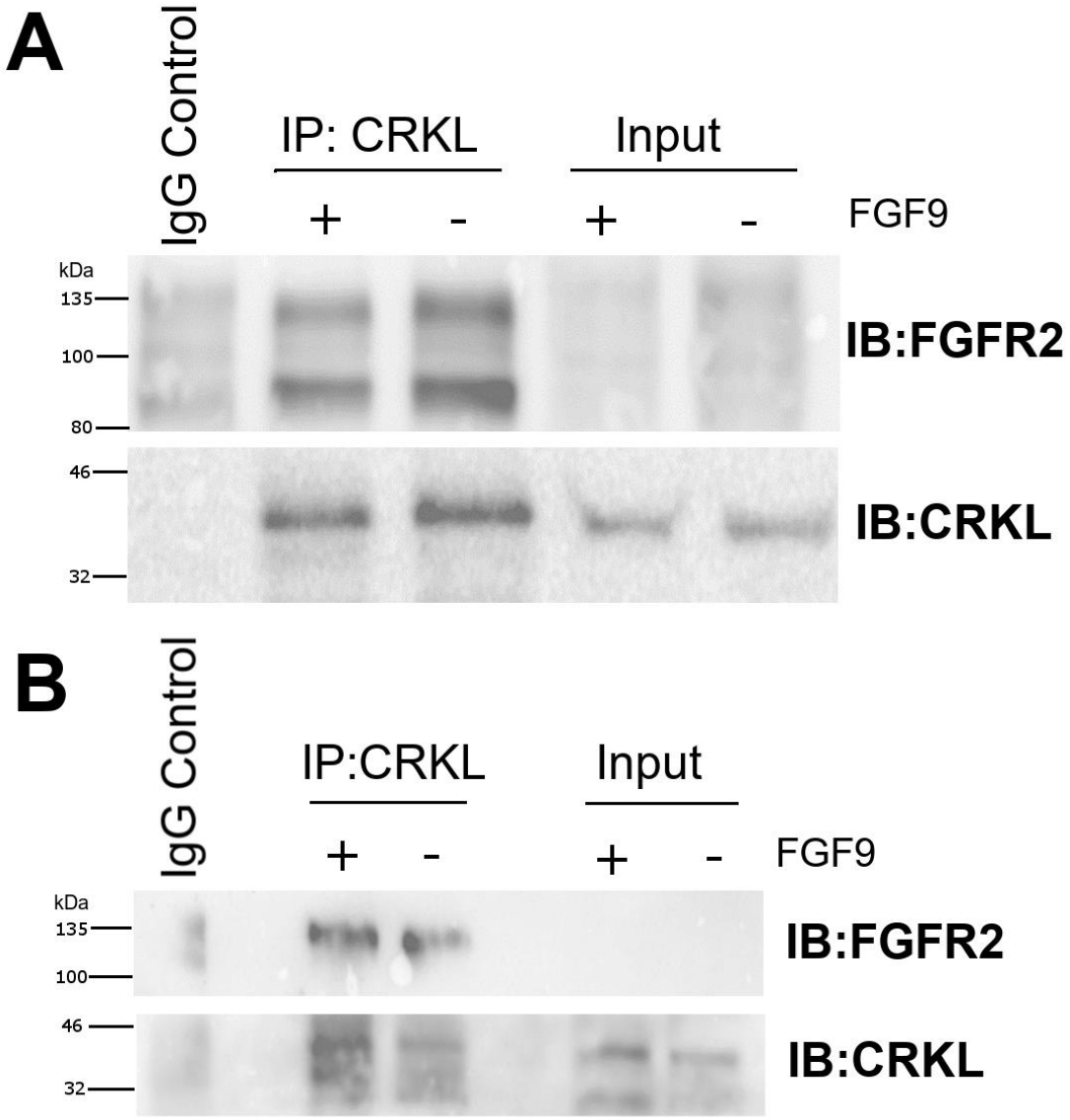


Figure 3-3. CRKL and FGFR2 form a complex in cells. Co-immunoprecipitation shows that anti-CRKL antibody pulls down CRKL from cell lysates, and FGFR2 along with it in starved and FGF9-stimulated (**A**) H520 cells and (**B**) SkBr3 cells. N = 2

3.4 CRKL^{2xSH3} can bind FGFR2 C58 with strong affinity

Microscale thermophoresis (MST) was employed to investigate the affinity between the CRKL SH3 domains and the FGFR2 C-terminal tail

containing proline-rich motifs. The FGFR2 peptide containing the last 58 residues of the C-terminal tail (hereby referred to as FGFR2 C58) was cloned, purified, and Atto^{488}-labelled by Dr Chi-Chuan Lin (University of Leeds). The GST-tagged CRKL^{2xSH3} construct was expressed in *E. coli*, purified and GST tag was cleaved off. The binding affinity between these two peptides was then measured. The MST data are shown in Figure 3-4. The CRKL SH3 domains can bind the FGFR2 C58 with the dissociation constant (K_d) of 2.510 µM ± 0.264 µM. The K_d is within the range of SH3-PxxP interactions (between 5 – 100 µM; Pawson et al, 1995), and stronger than previously documented FGFR2 interaction with PLCγ1 (≈ 40 µM, (Timsah et al., 2014)), but weaker than FGFR2 binding to GRB2 (≈ 0.1 µM, (Lin et al., 2012)). MST data, combined with co-IP results suggest that CRKL and FGFR2 form a complex in cells that is, at least in part, mediated by the SH3 domain interaction with the proline-rich motif.

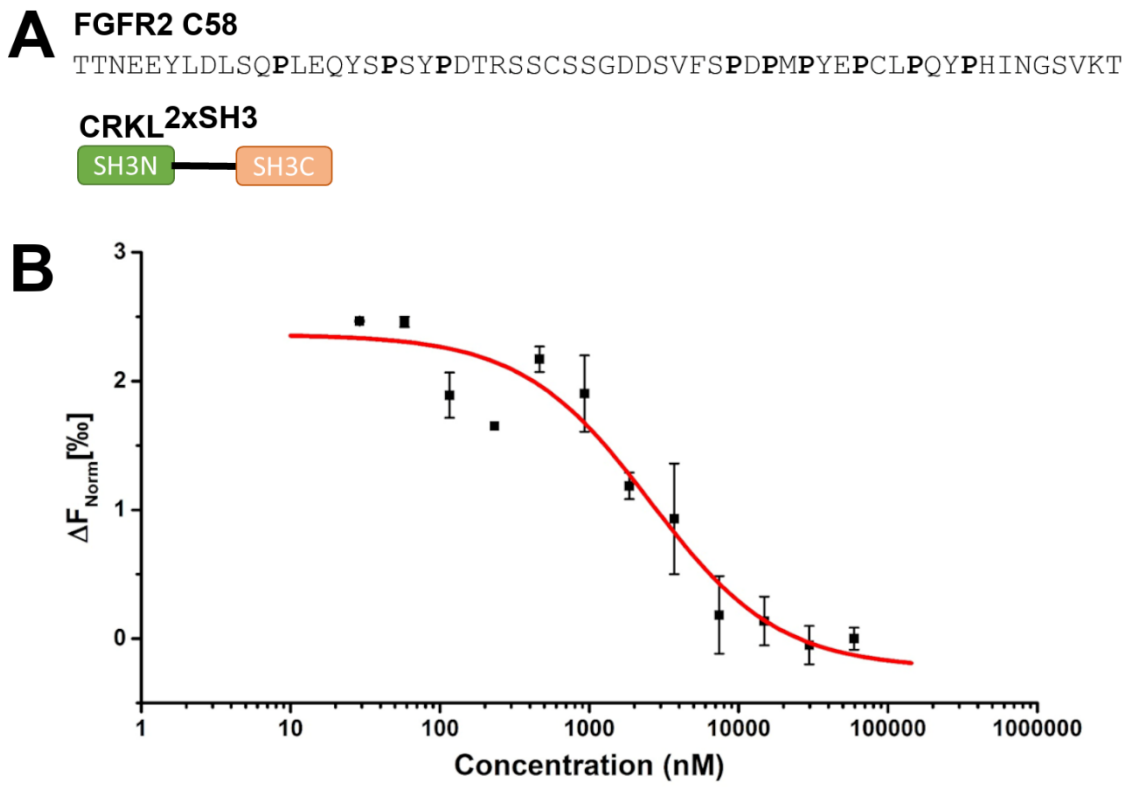


Figure 3-4. CRKL SH3 domains bind FGFR2 C-terminal tail. A) Peptide sequence of FGFR2 C-terminal tail (top) and domain structure of $CRKL^{2xSH3}$ construct (bottom) used in MST. **B)** MST binding curve of CRKL SH3 domain construct and FGFR2 C-terminal tail peptide. $Atto^{488}$-labelled FGFR2 C58 peptide was at a constant concentration of 100 nM, while unlabelled $CRKL^{2xSH3}$ was serially diluted (1:2) over 16 capillaries, with a maximum concentration of 362 µM. Each sample cappilary was scanned three times to obtain ΔF_{Norm}[‰], which was plotted against variable concentration of $CRKL^{2xSH3}$ and fitted into the law of mass action to obtain the dissociation constant (K_d). K_d = 2.510 µM ± 0.264 µM.

3.5 SH3C is not involved in binding to FGFR2

To establish which domain(s) of CRKL binds to FGFR2, several point mutations in a GST-CRKL full length construct were introduced (Figure 3-5). The mutants were designed to abolish binding of the individual domains through substituting residues which are known to be important in the recognition of ligands. The R21/39A mutations in the CRKL SH2 domain substitute the positively charged arginine residues which bind to phosphorylated tyrosine residues on the SH2 domain ligands. The W160L and W275L mutations

disrupts the hydrophobic proline-rich motif binding pocket of the SH3N and the SH3C domains respectively. While the Y207F mutation replaces a phosphorylatable tyrosine residue, which prevents SH2 domain binding in that region.

Figure 3-5. CRKL point mutant constructs. The mutations shown abolish the function of the respective domain, while the tyrosine mutation abolishes phosphorylation of Y207 residue.

The GST-CRKL constructs were expressed in *E. coli* and immobilised onto Glutathione Sepharose beads. The beads were incubated with the purified intracellular region of FGFR2 (FGFR2cyto) (cloning, expression and purification of FGFR2cyto was done by Dr Chi-Chuan Lin, University of Leeds). Binding of the GST-CRKL constructs to FGFR2cyto was investigated *in vitro* in the presence of phosphatase (to ensure no phosphorylation occurs) and visualised using western blotting (Figure 3-6). Only the GST-CRKL W275L mutation did not diminish pulldown of FGFR2, indicating that the SH3C domain is not involved in binding to the receptor. Since all the other mutations reduce the interaction between CRKL and FGFR2cyto, these data imply that the interaction can be mediated by multiple sites on CRKL. Hence, the CRKL SH2 and SH3N domains, as well as the intact Y207 site are required for binding to FGFR2.

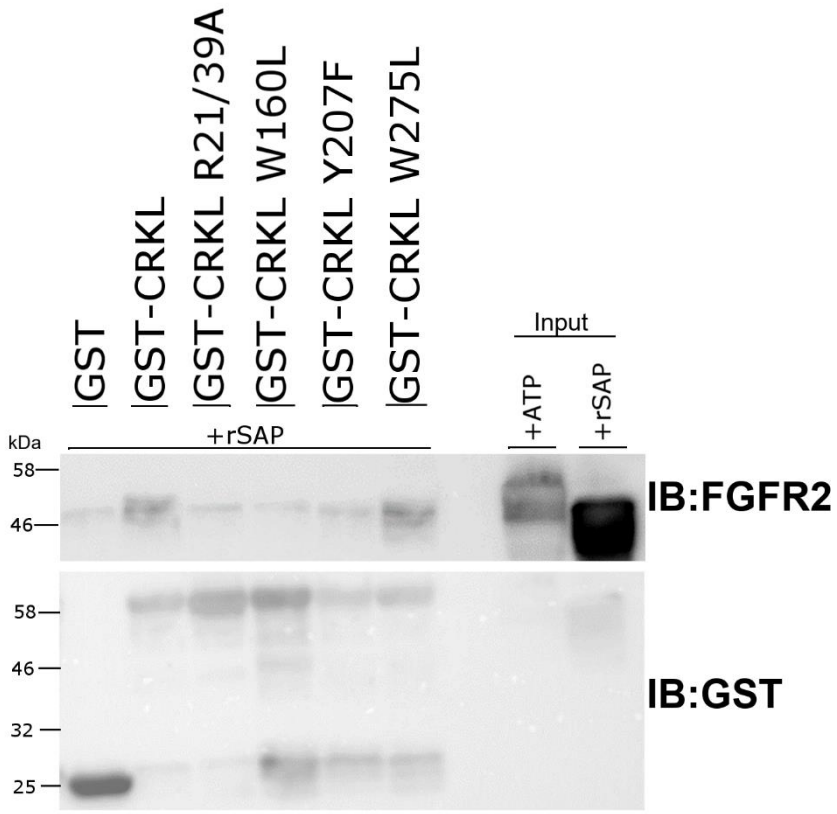


Figure 3-6. CRKL SH3C domain is not necessary for interaction with FGFR2. GST-tagged CRKL mutants were used to pull down FGFR2cyto construct containing the intracellular portion of the receptor. The reaction was performed under dephosphorylating conditions, by adding recombinant shrimp alkaline phosphatase (rSAP). R21/39A – SH2 mutant, W160L – SH3N mutant, Y207F – Y207 phosphosite mutant, W275L – SH3C mutant. N = 2.

3.6 FGFR2 C24 region is required for CRKL binding

In order to find the CRKL binding site on FGFR2 C-terminal tail, a GST pulldown assay was performed. GST-tagged FGFR2 C58 peptide was expressed in *E. coli* and immobilised onto Glutathione Sepharose beads. Two additional constructs were used, representing the first 24 residues (C24) and the last 34 residues (C34) of FGFR2 C-terminal tail to more precisely identify the region of interaction. The constructs were used to pull down full length CRKL from lysates of HEK293T and H520 cell lines, which express CRKL endogenously. The cells were starved overnight in DMEM media and left

starved or stimulated with DMEM containing 10% FBS for 15 minutes. The results are shown in Figure 3-7. The FGFR2 C-terminal tail can bind to CRKL from cell lysates under both starved and stimulated conditions. The results suggest that CRKL can bind the C58 construct, as well as C24 and C34 constructs.

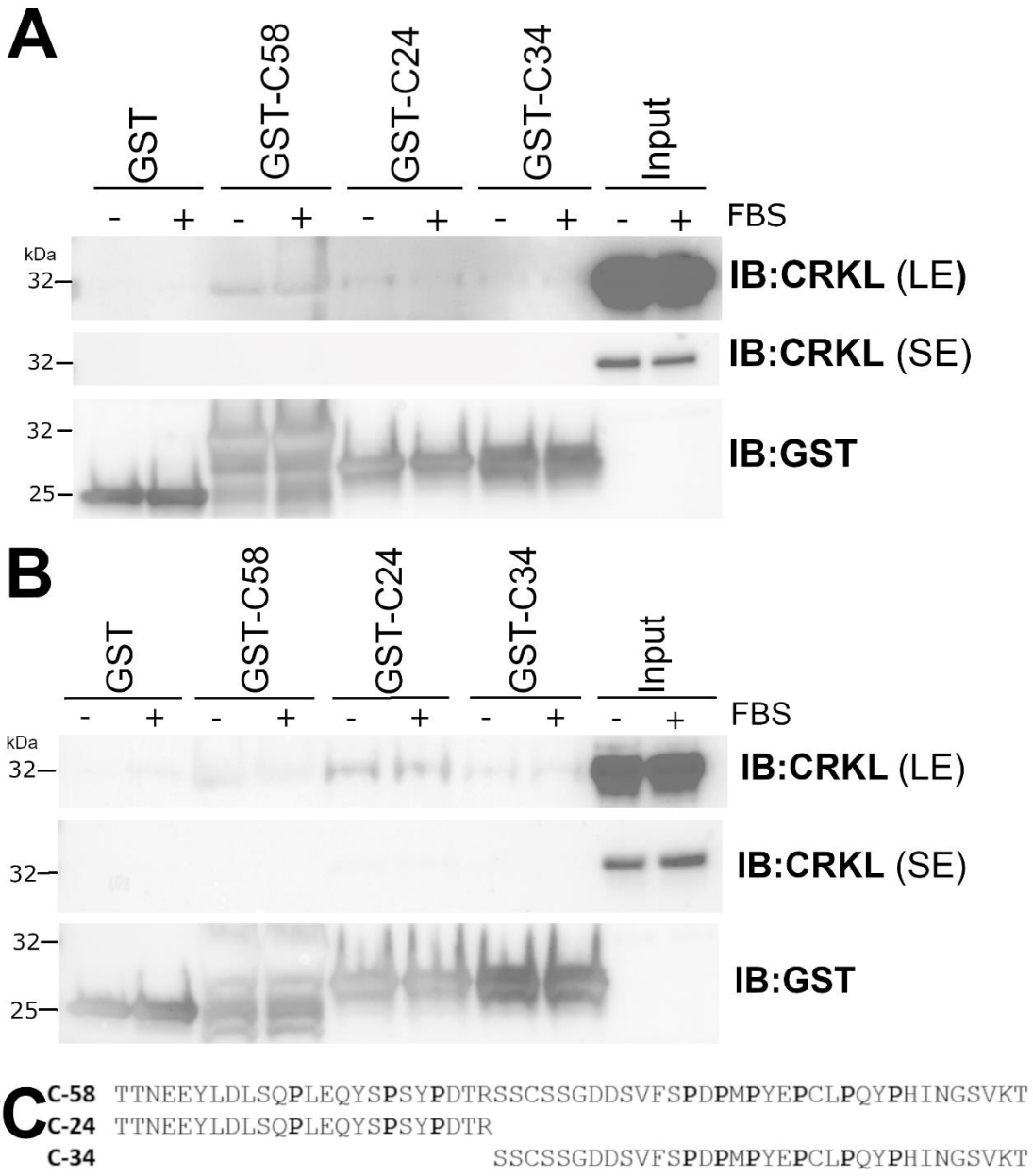


Figure 3-7. FGFR2 C24 construct is the primary CRKL binding site. GST-tagged FGFR2 C-terminal tail constructs were used to pull down CRKL from (**A**) H520 and (**B**) HEK293T cell lysates which were starved or stimulated with foetal bovine serum (FBS). **C**) Sequences of C58, C24 and C34 FGFR2 C-terminal tail constructs, with proline residues highlighted. N = 3 for each cell line. LE – long exposure, SE – short exposure.

3.7 CRKL can form dimers

Although the SH3C domain of CRKL does not mediate binding to FGFR2, it could still be functionally relevant. It has been shown before that CRKL can dimerise via its SH3C domain (Harkiolaki et al., 2006). MST was used to confirm these results using purified full-length CRKL and $CRKL^{2xSH3}$ constructs (Figure 3-8) (protein purification, tagging and MST was done by Dr Chi-Chuan Lin, University of Leeds). The MST data reveals that full length CRKL will interact with other CRKL molecules in solution. The profile of the binding curve obtained suggests that that CRKL self-assembly occurs to form tight dimers ($K_{d\ dimer}$ = 0.093 µM ± 0.008 µM). $CRKL^{2xSH3}$ construct also forms a tight dimer ($K_{d\ dimer}$ = 0.129 µM ± 0.0999 µM). The K_d dimerisation is stronger compared to previously predicted CRKL SH3C dimerisation affinity of 6.1 µM, as determined by sedimentation equilibrium analytical ultracentrifugation (Harkiolaki et al, 2006). The MST binding data is biphasic. A second binding curve showing a different behaviour (ΔF_{Norm}[‰] value increasing rather than decreasing over change in concentration) can be fit to the $CRKL^{2xSH3}$ binding data, giving a lower affinity K_d = 7.89 ± 1.19 µM. This binding occurs when protein concentration is increasing, suggesting that there might be a secondary binding event that could represent oligomerisation. Low abundancy of higher-order CRKL complexes have been seen before (Harkiolaki et al., 2006). The reason for the potential secondary binding event so far is unclear, but the K_d is higher than that for FGFR2 binding.

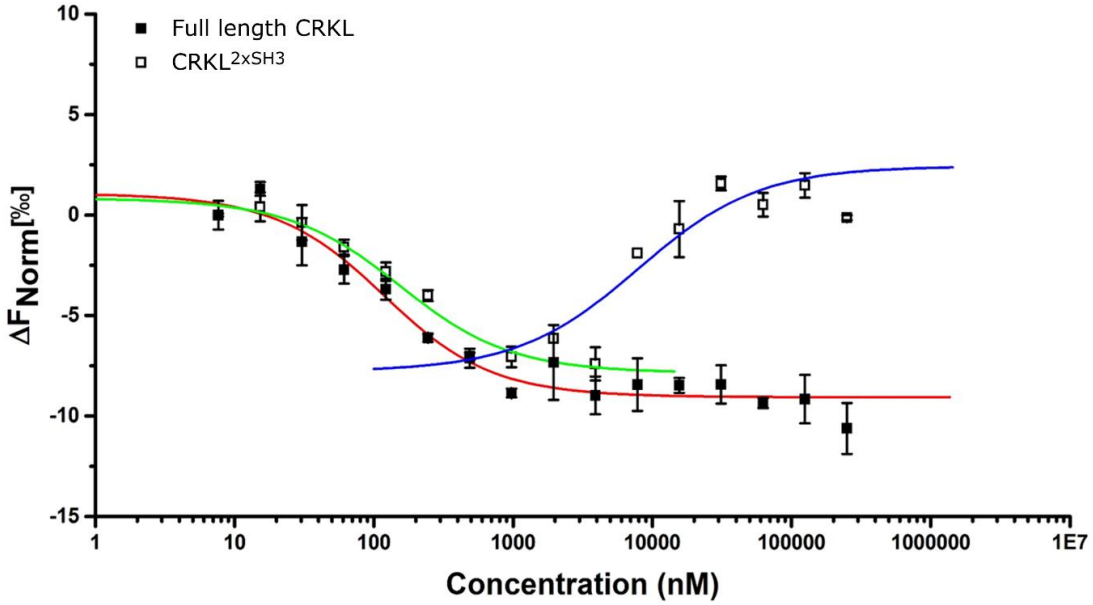


Figure 3-8. CRKL can form dimers and potentially oligomers. MST was used to determine the dimerisation of CRKL. Each sample capillary was scanned three times to obtain ΔF_{Norm}[‰], which was plotted against variable concentration of $CRKL^{2xSH3}$ and fitted into the law of mass action to obtain the dissociation constant (K_d). Full-length CRKL dimer (red) has K_d of 0.093 ± 0.008 µM, $CRKL^{2xSH3}$ construct dimer (green) K_d = 0.129 ± 0.0999 µM, $CRKL^{2xSH3}$ construct potential oligomer (blue) K_d = 7.89 ± 1.19 µM.

3.8 Dimeric CRKL enhances autophosphorylation of FGFR2

The K_d of dimerisation for CRKL (Figure 3-8, K_d = 0.093 ± 0.008 µM) is stronger than that of CRKL-FGFR2 interaction (Figure 3-4, K_d = 2.510 µM ± 0.264 µM), therefore, depending on protein concentrations, CRKL can bind FGFR2 as a dimer. This interaction of the CRKL adaptor protein could mimic what has previously been observed for the interaction between FGFR2 and GRB2. In the absence of extracellular stimuli dimeric GRB2 forms a heterotetramer with an FGFR2 dimer, holding the receptor in a dimeric state and preventing full activation in absence of ligand activation. Once FGFR2 is activated, it phosphorylates GRB2 to facilitate protein release (Lin et al., 2012). In order to investigate the importance of CRKL dimerisation for FGFR2 function, an *in vitro* kinase assay was performed to measure autophosphorylation of

FGFR2cyto. FGFR2cyto construct contains all intracellular parts of the FGFR2 receptor: juxtamembrane domain, tyrosine kinase domain, and the C-terminal tail, as well as a His tag. FGFR2cyto was incubated with either full-length CRKL, CRKL R21/39A mutant lacking functional SH2 domain, or $CRKL^{2xSH3}$ construct. ATP/$MgCl_2$ was present to allow receptor phosphorylation. FGFR2 phosphorylation was visualised by western blotting using pY99 antibodies, which detect all phosphoryolated tyrosine residues, and quantified using densitometry (Figure 3-9) (assay done by Dr Chi-Chuan Lin, University of Leeds). The CRKL concentration in these reactions was 10 µM, which is well above the apparent $K_{d\ dimer}$ for the constructs, ensuring CRKL dimerisation in solution as well as interaction with FGFR2. It is worth noting that at this concentration, $CRKL^{2xSH3}$ could also be oligomeric rather than dimeric . FGFR2 phosphorylation was quantified as a percentage of phosphorylated protein compared to total protein (the latter represents 100%). In the absence of CRKL, FGFR2 phosphorylation does not exceed 30% of total protein. When FGFR2 is incubated with full-length CRKL, phosphorylation reaches 58% within 5 minutes of ATP addition, 68% after 10 minutes, 77% after 20 minutes, and 88% after 40 minutes. Using CRKL SH2 domain mutant (R21/39A) did not markedly change the phosphorylation compared to wild-type CRKL. 61% of FGFR2 was phosphorylated after 5 minutes, 68% phosphorylated after 10 minutes, 73% phosphorylated after 20 minutes, and 74% phosphorylated after 40 minutes. To show that the increase in phosphorylation of FGFR2 was due to presence of CRKL SH3 domains, the $CRKL^{2xSH3}$ construct was used. Phosphorylation of FGFR2 in these conditions reached 64% after 5 minutes, 77% after 10 minutes, 85% after 20 minutes and 88% after 40 minutes after the addition of ATP. Taken together, these results show that the presence of CRKL, which is

potentially in a dimeric state, can strongly enhance overall FGFR2 phosphorylation, and that it is the two SH3 domains of CRKL that are required for this function.

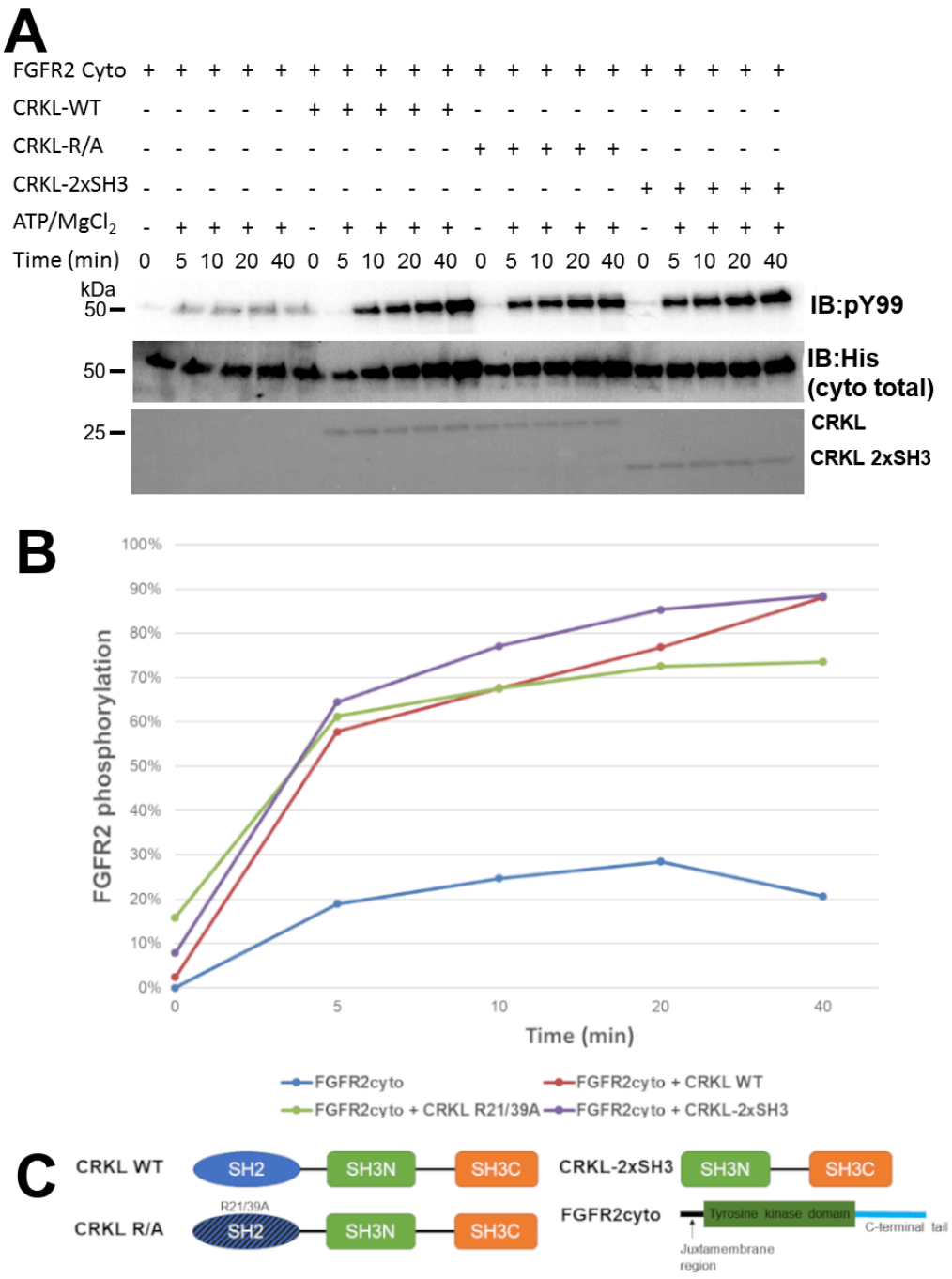


Figure 3-9. CRKL enhances FGFR2 phosphorylation. Timed *in vitro* kinase assay was performed to measure FGFR2 phosphorylation. Full length CRKL, CRKL R21/39A (CRKL-R/A) and CRKL^{2xSH3} constructs were used for the assay, along with FGFR2cyto. (**A**) CRKL/CRKL^{2xSH3} was visualised by Ponceau staining, while western blotting was used to visualise total (IB:His) and phosphorylated (IB:pY99) FGFR2cyto. (**B**) Percentage of FGFR2 phosphorylation was quantified using densitometry. N = 2. (**C**) Constructs used in this assay. CRKL WT – full length CRKL protein, CRKL R/A – full length protein containing R21/39A point mutations to disable SH2 domain binding, CRKL-2xSH3 – construct containing SH3N and SH3C domains of CRKL, FGFR2cyto – cytoplasmic portion of FGFR2 protein, containing a juxtamembrane domain (black line), tyrosine kinase domain (green box) and C-terminal tail (cyan line).

3.9 CRKL is phosphorylated by FGFR2

CRKL enhances FGFR2 phosphorylation. The next step was to see whether active FGFR2 can in turn phosphorylate CRKL. CRKL has several known phosphorylation sites, but the most investigated site is Y207 due to its importance in CML (Oda et al., 1994; Ten Hoeve et al., 1994; La Rosée et al., 2008). An *in vitro* kinase assay was performed, where GST-tagged CRKL and the single mutant constructs shown in Figure 3-5, were immobilised onto Glutathione Sepharose beads and incubated with FGFR2cyto and ATP/$MgCl_2$ (for phosphorylating conditions) or shrimp alkaline phosphatase (rSAP) (for dephosphorylating conditions). The assay was analysed by western blotting. This allowed assessment of CRKL Y207 phosphorylation in the presence of the receptor, as well as binding of phosphorylated or non-phosphorylated CRKL to FGFR2 (Figure 3-10). In the presence of ATP/$MgCl_2$, CRKL SH2 domain mediated binding to phosphorylated FGFR2. These results suggest that when FGFR2 is active and phosphorylated, CRKL can bind phosphotyrosine sites, as well as proline-rich motifs. CRKL Y207 phosphorylation was undetectable in the Y207F mutant, as expected. When FGFR2 phosphorylation was absent, only the CRKL SH3C domain was not essential for binding, because under dephosphorylating conditions (+rSAP; addition of recombinant shrimp alkaline phosphatase), GST-$CRKL^{W275L}$ bound to FGFR2 as strongly as wild-type GST-CRKL. The $CRKL^{Y207F}$ mutant was unable to bind FGFR2 under either condition. This suggests that the Y207 residue, regardless of its phosphorylation status, could somehow be required for CRKL binding to FGFR2. This could be due to Y207 being required for the formation of correct CRKL tertiary structure which would enable SH2 or SH3 domains binding to FGFR2. The results suggest that CRKL could have different binding site

preferences on FGFR2, which is dependent on both FGFR2 and CRKL phosphorylation status.

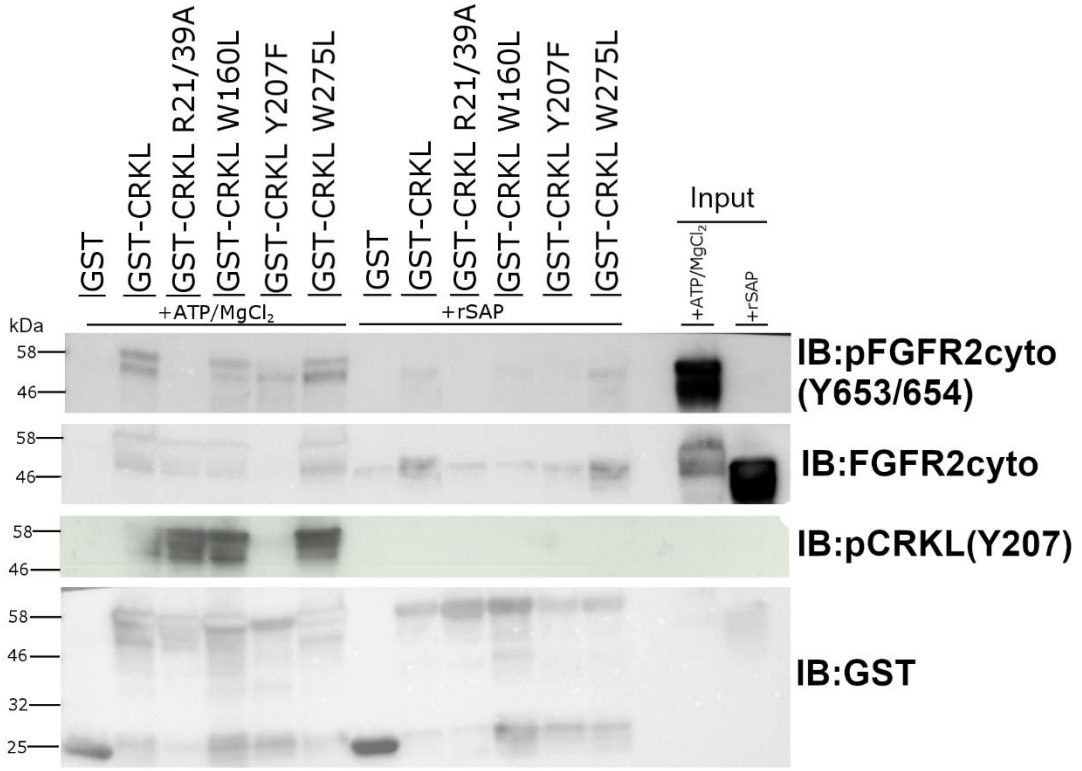


Figure 3-10. *In vitro* kinase assay to evaluate CRKL and FGFR2 phosphorylation. GST-tagged CRKL and point mutants were used to determine which CRKL domains are required for interaction with phosphorylated (pY653/654 in full-length protein) or non-phosphorylated FGFR2cyto construct, and if the interaction is required for CRKL Y207 phosphorylation. The experiment was done under phosphorylating (+ATP/$MgCl_2$) and dephosphorylating (+rSAP) conditions. R21/39A – SH2 mutant, W160L – SH3N mutant, Y207F – Y207 phosphosite mutant, W275L – SH3C mutant. N = 2.

After investigating the Y207 phosphorylation site using antibodies, mass spectrometry was used to identify other possible FGFR2 phosphorylation sites on CRKL. An *in vitro* kinase assay method was used to phosphorylate GST-CRKL by the FGFR2cyto construct (with addition of ATP and $MgCl_2$). After phosphorylation, the GST-CRKL protein was eluted and used for intact protein mass measurement as well as protease digestion followed by mass

spectrometry (MS/MS) (MS done by Biomolecular Mass Spectrometry facility, University of Leeds). The mass of full-length proteins was measured, which enabled the comparison of the mass of unphosphorylated GST-CRKL and phosphorylated GST-CRKL allowing determination of the total number of phosphorylatable sites.

The results of intact protein mass measurement are shown in Figure 3-11. There are two peaks present in the non-phosphorylated protein sample (Figure 3-11A). Peak A has a mass of 59641 Da, while peak B has a mass of 59773 Da. The difference in mass between the peaks is 131.26 Da. The lower molecular weight peak could be GST-CRKL protein that has lost the N-terminal methionine residue. In the presence of FGFR2cyto, ATP and $MgCl_2$, there are four peaks found in the mass measurement (Figure 3-11B). A portion of GST-CRKL is not phosphorylated in this sample, as peaks A and C have the same molecular mass as peaks A and B in the non-phosphorylated sample. There is an 80 Da shift between peaks A and B, as well as between peaks C and D. This suggests that there is one residue on GST-CRKL which has been phosphorylated, adding 80 Da to its molecular mass.

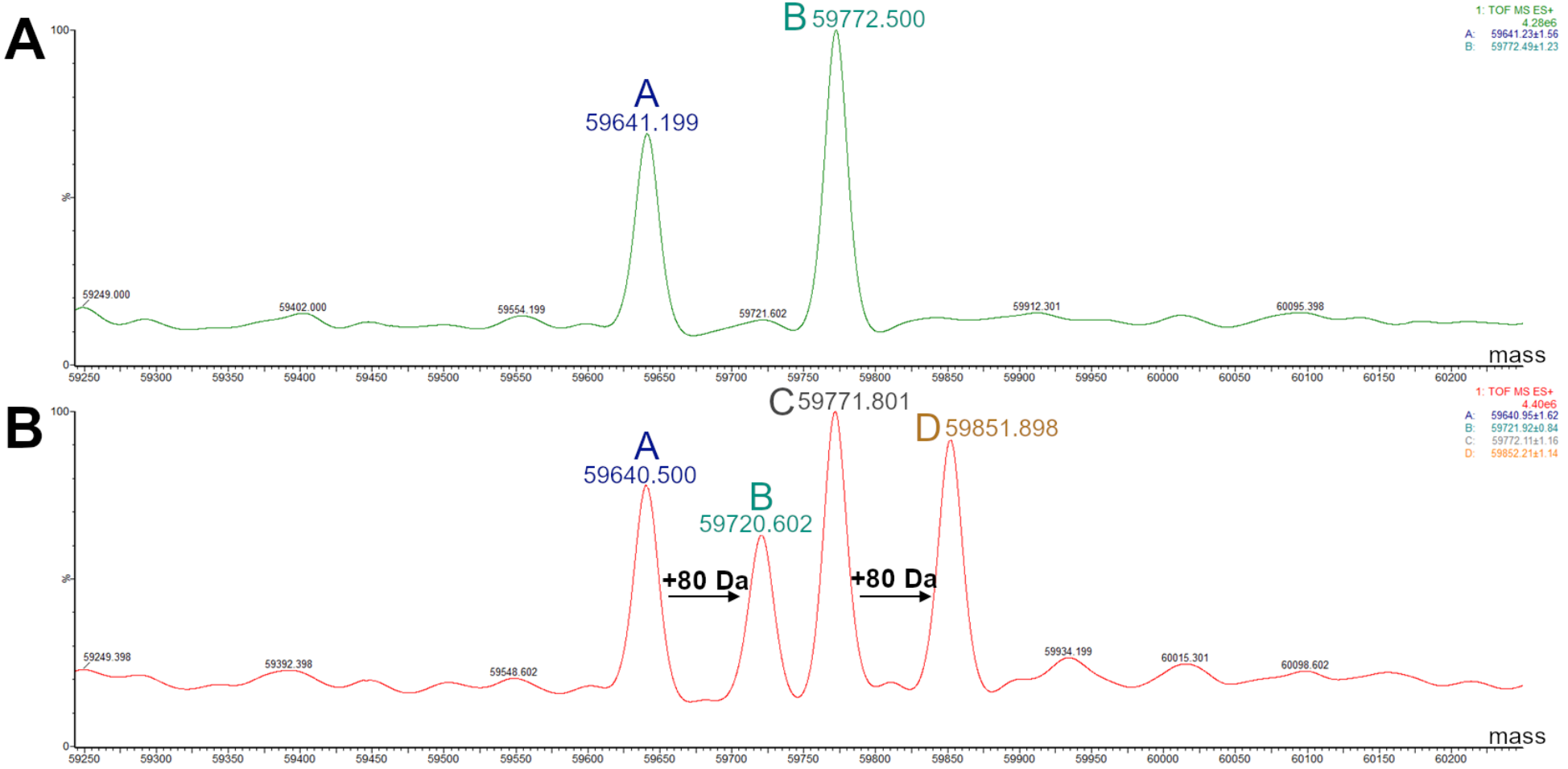


Figure 3-11. GST-CRKL intact protein mass measurement. Peaks showing the mass of (**A**) non-phosphorylated, and (**B**) phosphorylated GST-CRKL construct.

Next, the precise sites of phosphorylated tyrosine residues were mapped. The purified GST-CRKL protein was digested by proteases, and the resulting peptides were analysed by MS/MS. Peptides from digestion showed good coverage along the whole protein (Appendix B). Three sites were found on GST-CRKL, corresponding to CRKL residues Y48, Y198 and Y207 (Figure 3-12). Y48 resides in the SH2 domain, while Y198 and Y207 both are in the linker region connecting the two SH3 domains. As determined by intact protein measurement (Figure 3-11B), it is likely that only one of the residues could be phosphorylated on one molecule of CRKL at a time. All three residues of CRKL are previously unreported FGFR2 phosphorylation sites.

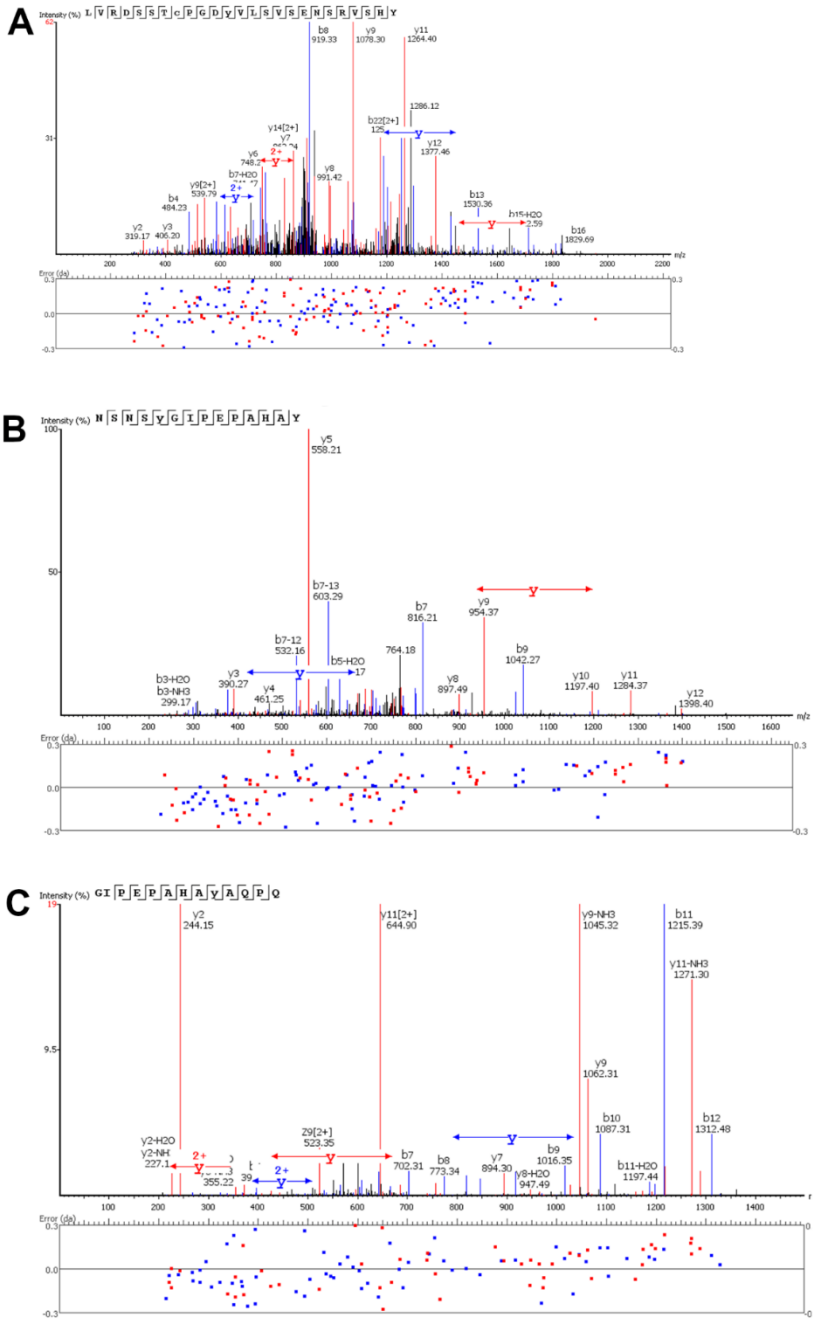


Figure 3-12. Mapping of GST-CRKL phosphorylation sites. Protein digestion followed by MS/MS was used to detect post-translational modifications in GST-CRKL construct. Each peak represents the mass of a charged peptide obtained, while colour denotes the charge for the peptide (b/blue if the charge is on the N-terminus, y/z/red if the charge is on the C-terminus). Mass differences between peaks are then used to determine the identity of the amino acid sequence in reference to CRKL sequence. Y with arrows represents the distance between two peaks which indicate the presence of phosphorylated tyrosine, while 2+Y arrows represent non-phosphorylated tyrosine residue. (**A**) phosphorylation of Y48 site; (**B**) phosphorylation of Y198 site, (**C**) phosphorylation of Y207 site. MS result generation, analysis and image generation were performed by Rachel George, University of Leeds Mass Spectrometry facility.

3.10 CRKL and RTK signalling assists cell migration

CRKL is an important regulator of cell migration (Feller, 2001; Birge et al., 2009; Lin et al., 2015). In order to see if the CRKL interaction with FGFR2 is important for this function, a wound healing assay was performed. HEK293T cells were used, in which native CRKL was knocked down using shRNA targeting several CRKL regions. A scrambled shRNA was used as control (Ci cells). Knockdown efficiency was quantified by western blot densitometry (Figure 3-13). Densitometry results show that CRKL expression was reduced in HEK293T CRKL shRNA cells (Figure 3-13B). Expression of the paralogue protein CRK-II was not affected strongly in CRKL knockdown cells (Figure 3-13B), ensuring that any effects seen in the following experiments are only due to differences in CRKL expression. Other potential off-target effects were not investigated.

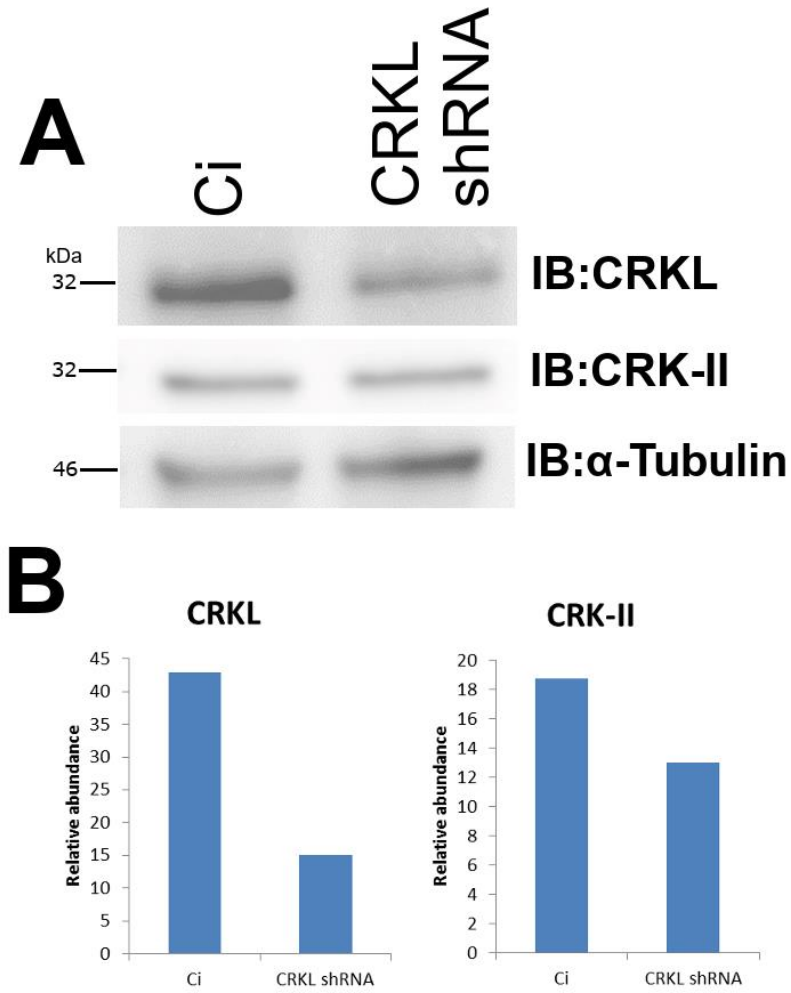


Figure 3-13. CRKL knockdown quality control in HEK293T. CRKL expression was knocked down by short hairpin RNA. The expression levels were detected by (**A**) western blotting and (**B**) quantified by densitometry by normalising to α-Tubulin. N = 1

As HEK293T cells do not express FGFR2, experimental cells were transfected with FGFR2-GFP construct for transient expression, while cells expressing GFP alone were used as a control. The experiment was done under basal conditions, using DMEM media containing 1% FBS. Wound closure was measured after 12 hours and the percentage of wound closure was calculated (Figure 3-14). Ci GFP control cells established a baseline for expected wound closure, with 30.88% of the wound closed after 12 hours. When CRKL is knocked down in these cells, the wound closure percentage dropped to 21.95%,

which is lower, indicating that CRKL has an important role in HEK293T cell migration. SU5402 inhibitor was also used in this assay. It can competitively bind the ATP-binding pocket of FGFR2 and several other RTKs (VEGFR2, FGFR1, PDGFRβ and EGFR). This inhibits canonical RTK signalling and downstream protein phosphorylation (Mohammadi et al., 1997; Cunningham et al., 2020). After treatment with SU5402, Ci GFP and CRKL shRNA GFP cells show a wound closure of 10.77% and 9.24% respectively. The results show that CRKL is partially needed for HEK293T cell migration under basal conditions, and that this migration is mediated by canonical RTK signalling, as there is no difference between Ci GFP and CRKL shRNA GFP cell migration when cells are treated with SU5402. When FGFR2 is introduced to the cells, Ci FGFR2-GFP cells show a wound closure of 34.37% under basal conditions, while CRKL shRNA FGFR2-GFP cells show a wound closure of 27.65%. As with GFP control cells, the difference suggests that CRKL is partially responsible for cell migration in FGFR2-GFP-containing cells. When cells are treated with SU5402, Ci FGFR2-GFP cells show 15.10% wound closure, while CRKL shRNA FGFR2-GFP cells show 10.99% wound closure, which is lower than that of untreated cells. The results show that CRKL is involved in HEK293T cell migration, but it is unclear whether FGFR2 involvement in cell migration is significant.

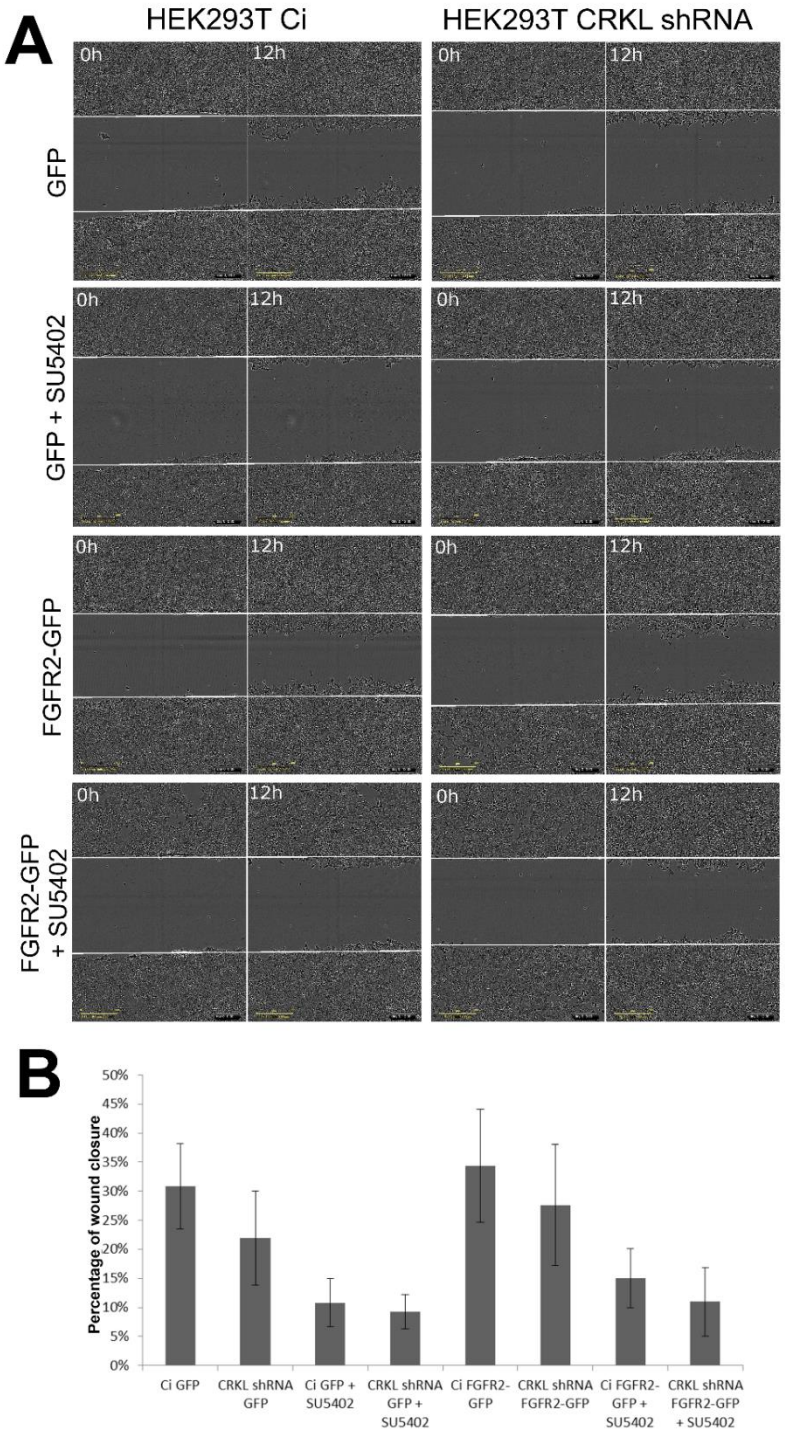


Figure 3-14. Roles of CRKL and FGFR2 in cell migration under basal conditions. Migration of HEK293T cells with or without CRKL and/or FGFR2 was measured. Ci – control shRNA cells, GFP – cells transfected with GFP, CRKL shRNA – cells with CRKL knockdown, FGFR2-GFP – cells transfected with FGFR2-GFP. SU5402 was also used as an FGFR2 (and other RTK) inhibitor. Experiment was performed under basal conditions (1% FBS). **A)** Representative images of the wound healing assay. **B)** Quantification of the wound closure percentage. N = 2.

3.11 RAP1 and RAC1 activation in HEK293T cells

To activate downstream signalling pathways, CRKL binds to C3G and DOCK1/ELMO GEFs in order to activate RAP1 and RAC1 GTPases, respectively (Knudsen et al., 1994; Tanaka et al., 1994; Sakkab et al., 2000). It is unclear whether FGFR2 can activate RAP1 or RAC1, but bFGF (FGF2), a ligand for the receptor, has been shown to activate RAC1 when promoting fibroblast migration (Kanazawa et al., 2010). The first step to investigate what occurs downstream of FGFR2/CRKL was to evaluate RAP1 and RAC1 activation. Commercial pulldown kits to detect active, GTP-bound RAP1 and RAC1 proteins were utilised. HEK293T cells were used, in which CRKL expression was knocked down using shRNA, with control cells (Ci) containing scrambled shRNA. Cells were also transfected with FGFR2-GFP-containing plasmid, or GFP-containing plasmid as a control, but there was a mixed population of GFP and non-GFP cells. The experiment was performed under basal conditions (DMEM +1% FBS). SU5402 inhibitor was used to inhibit FGFR2 signalling. Active RAP1 was detected using a commercial kit, which employs a GST-tagged RalGDS RBD domain that selectively binds active, GTP-bound RAP1. Cell lysate was incubated with either GTPγS (a GTP analogue that is highly resistant to hydrolysis) as a positive control, and GDP as a negative control to force RAP1 to be active or inactive, respectively. The pulldown experiment was visualised by western blotting, and the relative abundance of GTP-RAP1 (compared to total RAP1 in input) was determined using densitometry (Figure 3-15).

In HEK293T Ci control cells, RAP1 showed higher relative activation in cells containing FGFR2-GFP compared to GFP only (0.76 ± 0.51 vs 0.42 ±

0.26, respectively). The presence of SU5402 had no significant effects on these cells. When CRKL was knocked down, active RAP1 abundance in HEK293T CRKL shRNA GFP cells was 0.55 ± 0.35, while adding SU5402 inhibitor resulted in lower relative GTP-RAP1 abundance of in HEK293T CRKL shRNA cells containing GFP. HEK293T CRKL shRNA cells expressing FGFR2-GFP and treated with SU5402 showed no significant difference in active RAP1 abundance. The levels of RAP1 activation between repeats was highly variable, as it can be seen from high standard deviation and showed no statistical significance. The results could indicate that RAP1 activation in these cells is highly transient.

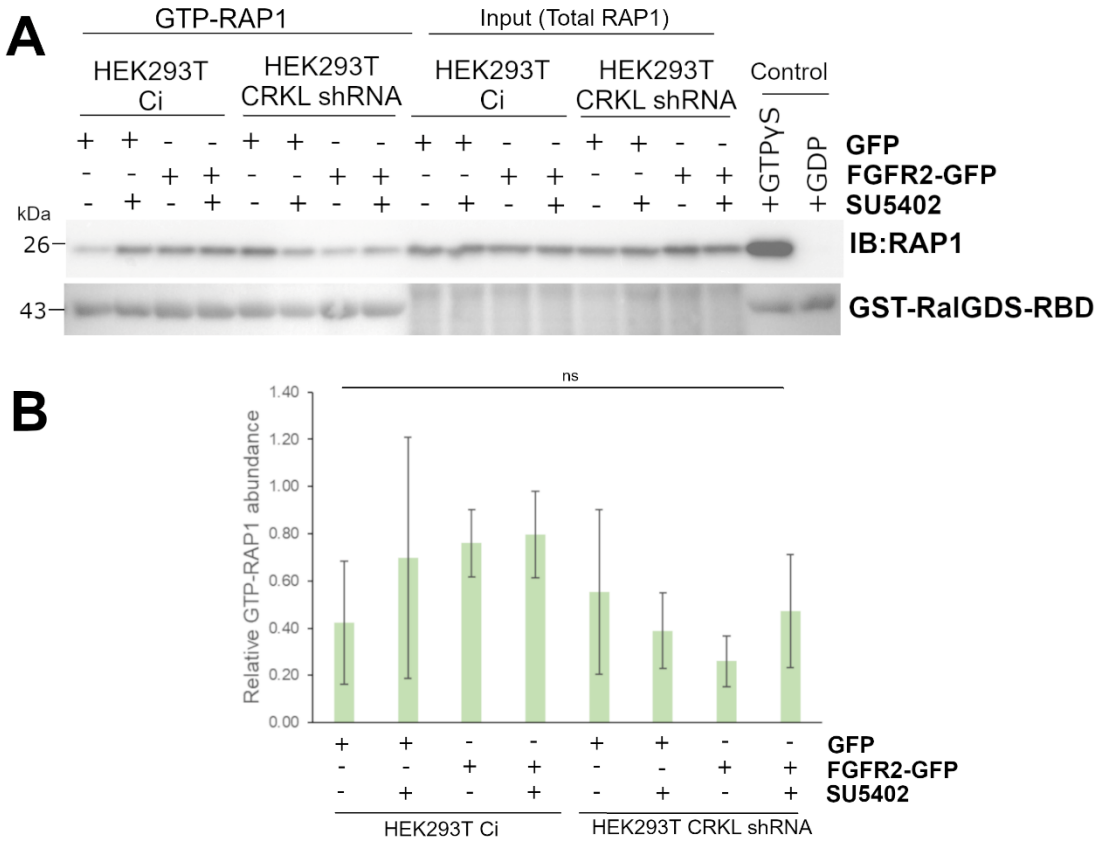


Figure 3-15. RAP1 activation in HEK293T cells. Ci – control shRNA cells, GFP – cells transfected with GFP, CRKL shRNA – cells with CRKL knockdown, FGFR2-GFP – cells transfected with FGFR2-GFP. **A)** RBD domain of RalGDS was used to pull down GTP-bound (active) RAP1 and run on a western blot, followed by visualisation using anti-RAP1 antibodies. **B)** Band intensity was quantified using densitometry, and relative GTP-RAP1 abundance, including standard deviation, was calculated. N = 3. Ns – not significant.

To evaluate RAC1 activation, a commercial kit was used, which utilised a GST-tagged PAK1 PBD domain that selectively bound active GTP-bound RAC1 (Figure 3-16). GTP-bound RAC1 was visualised and quantified as described above. In HEK293T Ci GFP cells, relative RAC1 activation was 0.71, while treatment of these cells with SU5402 increased relative RAC1 activation to 1.23. In HEK293T Ci cells containing FGFR2-GFP, relative abundance of GTP-RAC1 was 1.34, which was reduced to 0.70 after SU5402 treatment. HEK293T CRKL shRNA cells expressing GFP with or without SU5402 treatment showed

relative active RAC1 abundance of 0.41 and 0.43, respectively. HEK293T CRKL shRNA cells expressing FGFR2-GFP showed relative active RAC1 abundance of 0.31, while treatment of these cells with SU5402 resulted in relative GTP-bound RAC1 abundance of 0.45. Unfortunately, due to time constraints a third repeat could not be performed, thus so far, no conclusions can be drawn about RAC1 activation.

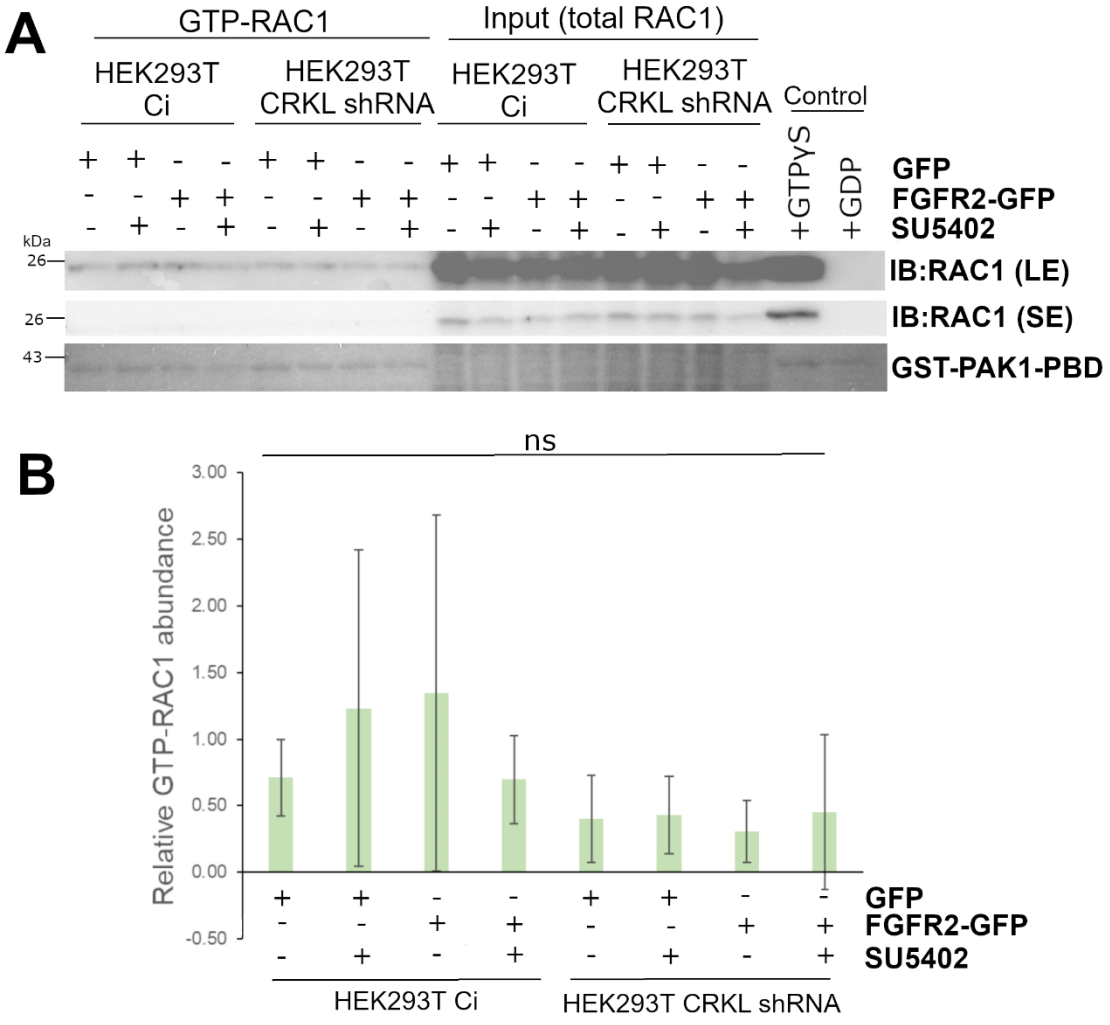


Figure 3-16. RAC1 activation assay in HEK293T cells. Ci – control shRNA cells, GFP – cells transfected with GFP, CRKL shRNA – cells with CRKL knockdown, FGFR2-GFP – cells transfected with FGFR2-GFP. **A)** PBD domain of PAK1 was used to pull down GTP-bound (active) RAC1 and run on a western blot, followed by visualisation using RAC1 antibodies. **B)** Band intensity was quantified using densitometry. N = 2. Ns – not significant. LE – long exposure, SE – short exposure.

3.12 RAP1 and RAC1 signalling in SkBr3 breast cancer cells

RAP1 and RAC1 activation by FGFR2/CRKL was also investigated in SkBr3 cells. First, CRKL expression was knocked down to generate SkBr3 CRKL shRNA cells (Figure 3-17). CRKL protein level was visualised using western blotting and quantified by densitometry. Obtained SkBr3 CRKL shRNA cells express ~24% less CRKL than SkBr3 Ci control cells. CRK-II levels were also checked to ensure that shRNA targets CRKL selectively. Interestingly, knocking down CRKL in SkBr3 cells lead to higher CRK-II expression (by ~28%).

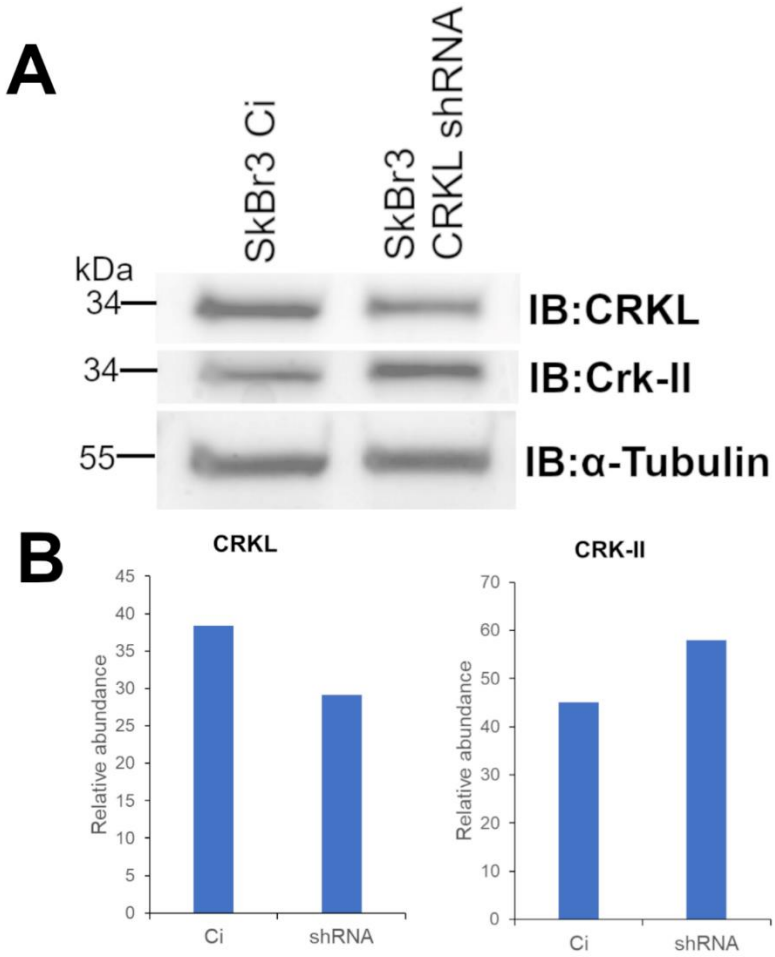


Figure 3-17. CRKL knockdown quality control in SkBr3 cells. CRKL expression was knocked down by short hairpin RNA. The expression levels were tested by (**A**) western blotting and (**B**) quantified by densitometry. N = 1

RAP1 activation in SkBr3 cells was detected and quantified as previously described (Figure 3-18). First, RAP1 activation was evaluated in cells grown under basal conditions (DMEM 1% FBS), as with HEK293T cells (Figure 3-18, A and B). SU5402 acts as an FGFR2-specific inhibitor in these cells, as it is the only inhibitor target present in this cell line. In SkBr3 Ci cells relative average abundance of GTP-bound RAP1 was 8.79 (-SU5402) and 5.91 (+SU5402). When CRKL was knocked down, RAP1 activation decreased to 0.63 in untreated SkBr3 CRKL shRNA cells. Relative GTP-bound RAP1 abundance was slightly higher in same cells treated with SU5402 (1.53). The results were not statistically significant, likely due to high variance between measurements.

RAP1 activation was also investigated in SkBr3 cells that were either starved, or stimulated with FGF9 (FGFR2-activating ligand) (Figure 3-18, C and D). In starved SkBr3 Ci cells, average relative GTP-RAP1 abundance was 1.27, which increased to 2.48 after SU5402 treatment. In FGF9-stimulated SkBr3 Ci cells, active RAP1 abundance was 2.11, while in same cells treated with SU5402 the abundance was 1.92. Knocking down CRKL reduced RAP1 activation. Starved SkBr3 CRKL shRNA cells showed average GTP-RAP1 abundance of 0.05, while the same cells treated with SU5402 showed active RAP1 abundance of 0.14. Relative abundance of GTP-RAP1 in SkBr3 CRKL shRNA cells stimulated with FGF9 was 0.27, and the same cells treated with SU5402 showed an average abundance of active RAP1 of 0.39. The results were not significant, but they could point to a conclusion that CRKL is likely to be required for RAP1 activation in SkBr3 cells under starved and stimulating conditions, but the relationship with the FGFR2 receptor is unclear..

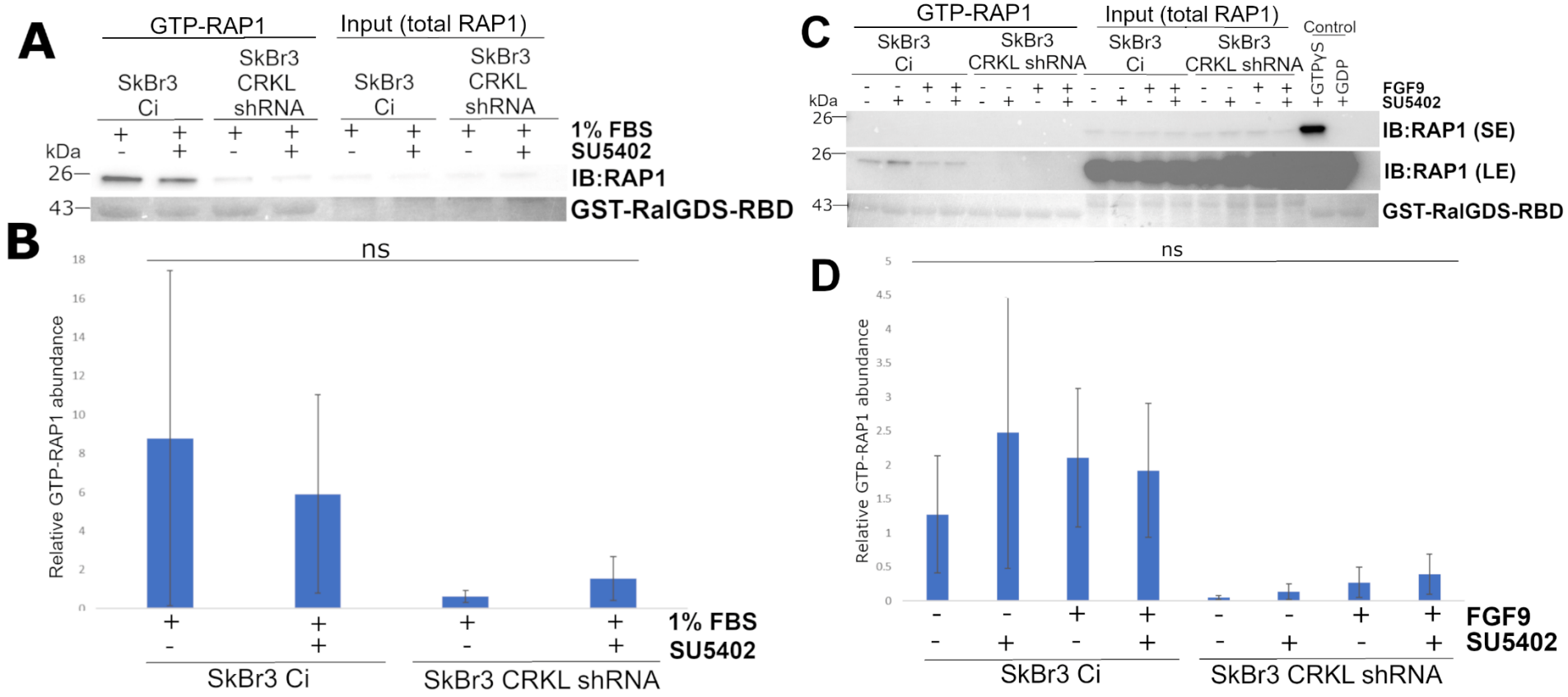


Figure 3-18. RAP1 activation in SkBr3 cells. A) RBD domain of RalGDS was used to pull down GTP-bound (active) RAP1 from SkBr3 cells (Ci – control; CRKL shRNA – CRKL knockdown) grown under basal conditions and run on a western blot, followed by visualisation using RAP1 antibodies. **B)** Band intensity was quantified using densitometry. N = 2. **C)** Pulldown was also done in cells that were either starved or stimulated with FGF9. **D)** Densitometry was used to quantify relative abundance of active protein. N = 2. Ns – not significant, * - $p \leq 0.05$.

RAC1 activation was also investigated in SkBr3 cells (Figure 3-19). Although RAC1 was expressed in these cells, no active GTP-bound RAC1 could be detected. The results suggest that RAC1 is not activated in SkBr3 cells under experimental conditions used.

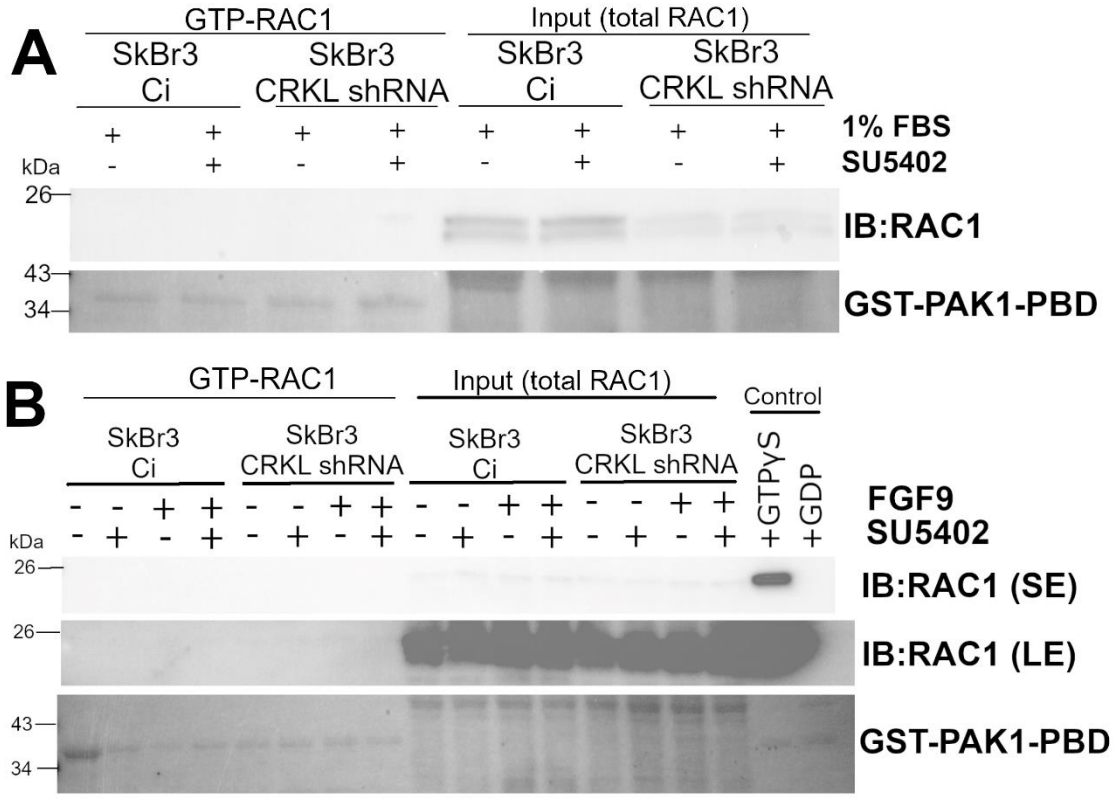


Figure 3-19. RAC1 activation in SkBr3 cells. PBD domain of PAK1 was used to pull down GTP-bound (active) RAC1 and run on a western blot, followed by visualisation using anti-RAC1 antibodies. The experiment was done in SkBr3 cells (Ci – control, CRKL shRNA – CRKL knockdown) grown under (**A**) basal conditions in DMEM media containing 1% FBS or (**B**) starved in serum-free media and left starved or stimulated with FGF9.

3.13 CRKL-dependent signalling in SkBr3 breast cancer cells

FGFR2 has previously been shown to play important roles in multiple intracellular signalling pathways (see section 1.3). In order to investigate if CRKL also functions in these pathways, Reverse Phase Protein Array (RPPA)

was employed. This assay is based on the dot-blot method, and utilises antibodies to detect levels of total and phosphorylated proteins in cell lysate. A wide selection of antibodies was adopted, which allowed inclusion of not only pathways in which FGFR2 or CRKL participate, but also other pathways not known to be regulated by these proteins to explore the possibility that they can be regulated by FGFR2/CRKL signalling. SkBr3 Ci (control) and SkBr3 CRKL shRNA (CRKL knockdown) cells were used. The cells were grown under starved (-FGF9), stimulated (+FGF9) or basal (1% FBS) conditions and either left untreated, or treated with SU5402 FGFR2 inhibitor. Signal for relative protein expression (normalised to total protein signal) from three biological repeats was converted to a Z-score and presented as a clustered heatmap (Figure 3-20). SkBr3 Ci and SkBr3 CRKL shRNA cells, regardless of treatment, formed two distinct clusters representing the two cell types, which suggests that CRKL presence has an effect on these signalling pathways.

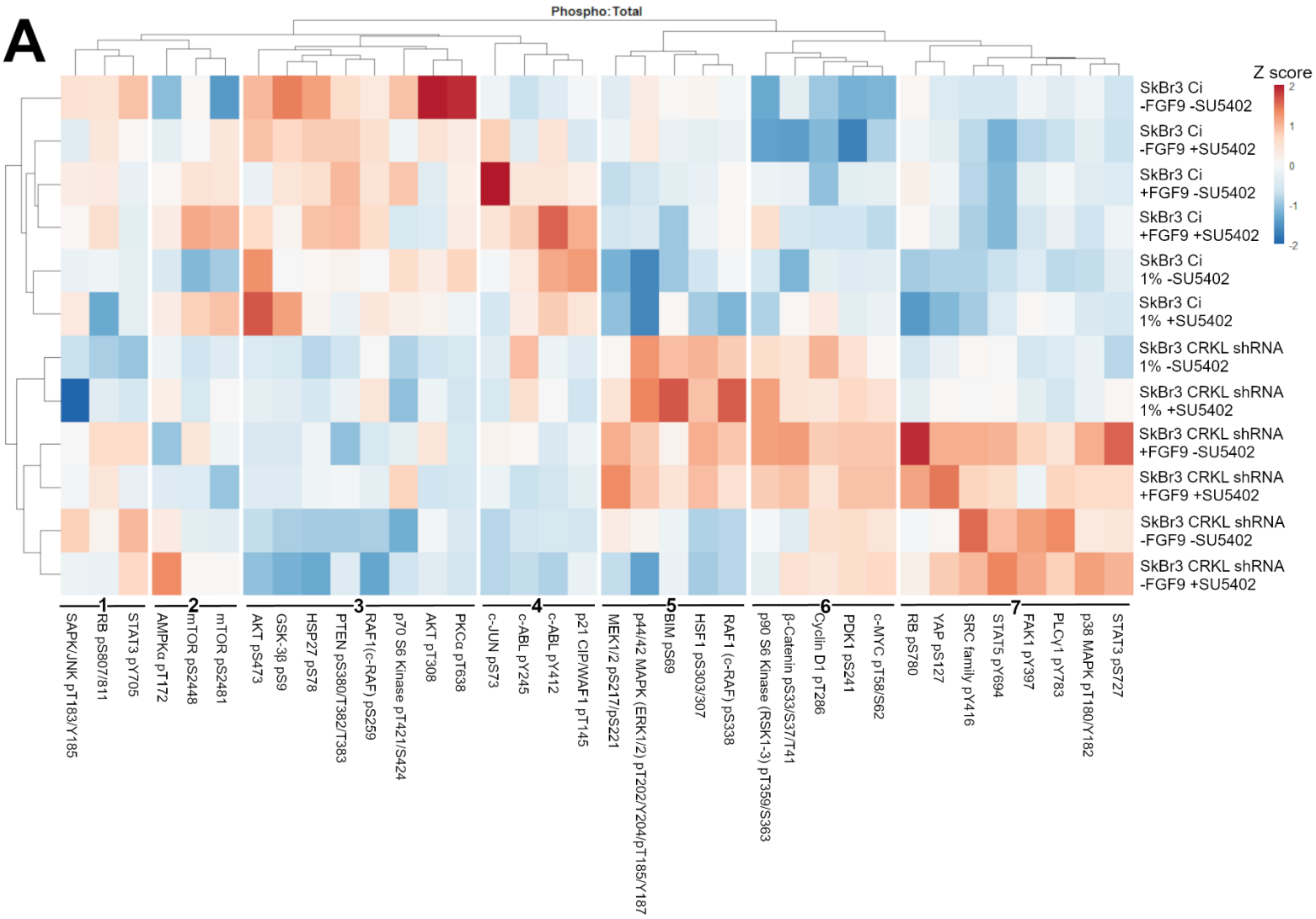
The relative protein phosphorylation heatmap (Figure 3-20A) revealed that overall, SkBr3 Ci cells have increased phosphorylation in three separate groups of proteins. The first group (2) contains mTOR (pS2481 and pS2448) and AMPKα (pT172). In this group, mTOR shows higher phosphorylation in control cells under most conditions, while AMPKα shows only a small increase in Ci cells under a few conditions, and also higher phosphorylation in SkBr3 CRKL shRNA cells under a few conditions.The next cluster (3) consists of AKT (pS308 and pS473), GSK-3-β (pS9), HSP27 (pS78), PTEN (pS380/T382/T383), RAF1 (c-RAF) (pS259), p70 S6 Kinase (pT412/S424), and PKCα (pT638). These proteins show an increased phosphorylation under all SkBr3 Ci cell treatment conditions. The last group (4) of proteins showing increased phosphorylation levels in SkBr3 Ci cells (except for -FGF9 -SU5402 cells)

contains c-JUN (pS73), c-ABL (pY245 and pY412), and p21 CIP/WAF1 (pT145). A group of proteins shows increased phosphorylation in both SkBr3 Ci and CRKL shRNA cells under several different conditions, and it (1) consists of SAPK (JNK) (pT183/Y185), RB (pS807/S811), and STAT3 (pY705). In SkBr3 CRKL shRNA cells, there are three separate clusters of proteins that show increased protein phosphorylation levels. The first group (5) contains MEK 1/2 (pS217/221), p44/42 MAPK ERK1/2 (pT202/158, pY204/187), Bim (pS69), HSF1 (pS303/307), and RAF1 (c-RAF) (pS338). Proteins from this cluster show increased phosphorylation under 1% and +FGF9 conditions with or without SU5402. The next group (6) of proteins show increased phosphorylation in SkBr3 CRKL shRNA cells under all treatments: p90 S6 Kinase (RSK1-3) (pT359/S363), β-catenin (pS33/S37/T41), Cyclin D1 (pT286), PDK1 (pS241), c-MYC (pT58/S62). When cells are starved or stimulated (-FGF9 or +FGF9), relative phosphorylation of Rb (pS780), YAP (pS127), Src family proteins (pY416), STAT5 (pY694), FAK1 (pY397), PLCγ1 (pY783), p38 MAPK (pT180/Y182), and STAT3 (pS727) is increased, forming a separate cluster (7). In this heatmap, SkBr3 Ci -FGF9 -SU5402 cells form one cluster. Another group is formed by all other SkBr3 Ci cells, which can be divided further into two different sub-groups – one consisting of SkBr3 Ci -FGF9 +SU5402 cells, as well as SkBr3 Ci +FGF9 cells with or without SU5402 treatment, and another sub-group consisting of SkBr3 Ci cells grown under basal conditions (1% FBS) with or without SU5402 treatment. SkBr3 CRKL shRNA cells form more defined clusters, where cells were grouped by their growth conditions (1%, +FGF9 or -FGF9), with SU5402 treatment having only minor effects on protein phosphorylation.

A heatmap portraying Z-scores for total protein (or phospho-protein) levels follows a similar trend as the phospho:total heatmap, where overall, SkBr3 Ci and SkBr3 CRKL shRNA cells form two distinct groups where proteins show mirroring increases and decreases in protein expression and phosphorylation (Figure 3-20B). Proteins in the first cluster (1), containing mTOR, Profilin, BIM, p38 (MAPK), CHK1 (pS345), PKA substrate (pRRXS/T), FAK1, SMAD2/3 (pS465/S423/S467/S425), PLCγ1, NFκB p105/p50, GSK-3-β, RAF1 (c-RAF), HSP27, BAD (pS112), AKT, ILK1, p44/42 MAPK (ERK1/2), AKT substrate (pRXXS/T), S6 ribosomal protein (pS235/S236), GRB2, β-Tubulin, SHP2 (pY542), HSF1, PDK1, SHC (pY239/240), show increased proteins levels in SkBr3 CRKL shRNA grown under starved (-FGF9) or stimulated (-FGF9) conditions with or without SU5402. The next group (2) consists of proteins that show increased levels in SkBr3 CRKL shRNA cells grown under all conditions: RB, p70 S6 kinase, PTEN, BCL-XL, ATM/ATR Substrate pS/T, SAPK (JNK1/JNK2), RAS, B-RAF, and PKCα. The third cluster (3) contains c-JUN (N-term), c-ABL, Calmodulin, GAPDH, and p21 CIP/WAF1. These proteins show a small increase in their level in SkBr3 CRKL shRNA cells grown under all conditions, as well as SkBr3 Ci cells grown under select few conditions. The next group (4) of proteins show increased levels in SkBr3 Ci cells regardless of treatment. These proteins include c-MYC, Aurora A/B/C (pT288/T232/T198), STAT1 (pS727), and STAT3. SMAD1/5 (pS463/S465), Cortactin, BID, and p53 form the next cluster (5): these proteins show increased levels in SkBr3 Ci and CRKL shRNA cells under basal conditions (1%), as well as few other select conditions. The next group of proteins (6), consisting of AMPKα, GSK-3α/β (pS21/S9), MNK1 (MKNK) (pT197/202), PKC (pan, pS660), MMP21, RhoA, and STAT6 (pY641), show increased levels in starved (-FGF9) SkBr3 Ci and CRKL

shRNA cells untreated with with SU5402, among few other select conditions depending on protein. The following cluster (7) contains NFκB p65 (pS536), FOXO3a, Cyclin D1, and Ubiquitin. These proteins mostly show an increased level in non-stimulated (-FGF9) SkBr3 Ci cells with or without SU5402. The last group (8) of proteins show increased levels in SkBr3 Ci cells regardless of treatment. The group encompasses FOXO1, XIAP, CRKL, PI3-Kinase p110α, YAP1, STAT5, β-Catenin, COOL1/βPIX, CHK2 (pT68), β-actin, SRC, IRS1 (pS636/639), LKB1, MEK1/2, Histone H2A.X (pS139), p90 S6 kinase (Risk1-3). Here, SkBr3 Ci -FGF9 +/-SU5402 cells form a separate group, while SkBr3 Ci +FGF9 +/-SU5402 and SkBr3 Ci 1% +/-SU5402 cells form another cluster, further sub-divided into cells treated with +FGF9 or 1%. While for SkBr3 CRKL shRNA cells, cells grown under 1% FBS form a separate cluster, while SkBr3 CRKL shRNA +FGF9 +/-SU5402 and SkBr3 CRKL shRNA -FGF9 +/-SU5402 cell form another group, further sub-divided into cells grown with or without FGF9.

The results show that overall, SkBr3 cell signalling landscape shows changes influenced by presence of CRKL, while different growth conditions (1% FBS, +/-FGF9) or FGFR2 signalling/inhibition (-/+ SU5402) also can influence SkBr3 Ci and SkBr3 CRKL shRNA cell signalling.


A
Phospho:Total
Z score
2
1
0
-1
-2
SkBr3 Ci -FGF9 -SU5402
SkBr3 Ci -FGF9 +SU5402
SkBr3 Ci +FGF9 -SU5402
SkBr3 Ci +FGF9 +SU5402
SkBr3 Ci 1% -SU5402
SkBr3 Ci 1% +SU5402
SkBr3 CRKL shRNA 1% -SU5402
SkBr3 CRKL shRNA 1% +SU5402
SkBr3 CRKL shRNA +FGF9 -SU5402
SkBr3 CRKL shRNA +FGF9 +SU5402
SkBr3 CRKL shRNA -FGF9 -SU5402
SkBr3 CRKL shRNA -FGF9 +SU5402
1
SAPK/JNK pT183/Y185
RB pS807/811
STAT3 pY705
2
AMPKα pT172
mTOR pS2448
mTOR pS2481
3
AKT pS473
GSK-3β pS9
HSP27 pS78
PTEN pS380/T382/T383
RAF1(c-RAF) pS259
p70 S6 Kinase pT421/S424
AKT pT308
PKCα pT638
4
c-JUN pS73
c-ABL pY245
c-ABL pY412
p21 CIP/WAF1 pT145
5
MEK1/2 pS217/pS221
p44/42 MAPK (ERK1/2) pT202/Y204/pT185/Y187
BIM pS69
HSF1 pS303/307
RAF1 (c-RAF) pS338
6
p90 S6 Kinase (RSK1-3) pT359/S363
β-Caterin pS33/S37/T41
Cyclin D1 pT286
PDK1 pS241
c-MYC pT58/S62
7
RB pS780
YAP pS127
SRC family pY416
STAT5 pY694
FAK1 pY397
PLCγ1 pY783
p38 MAPK pT180/Y182
STAT3 pS727

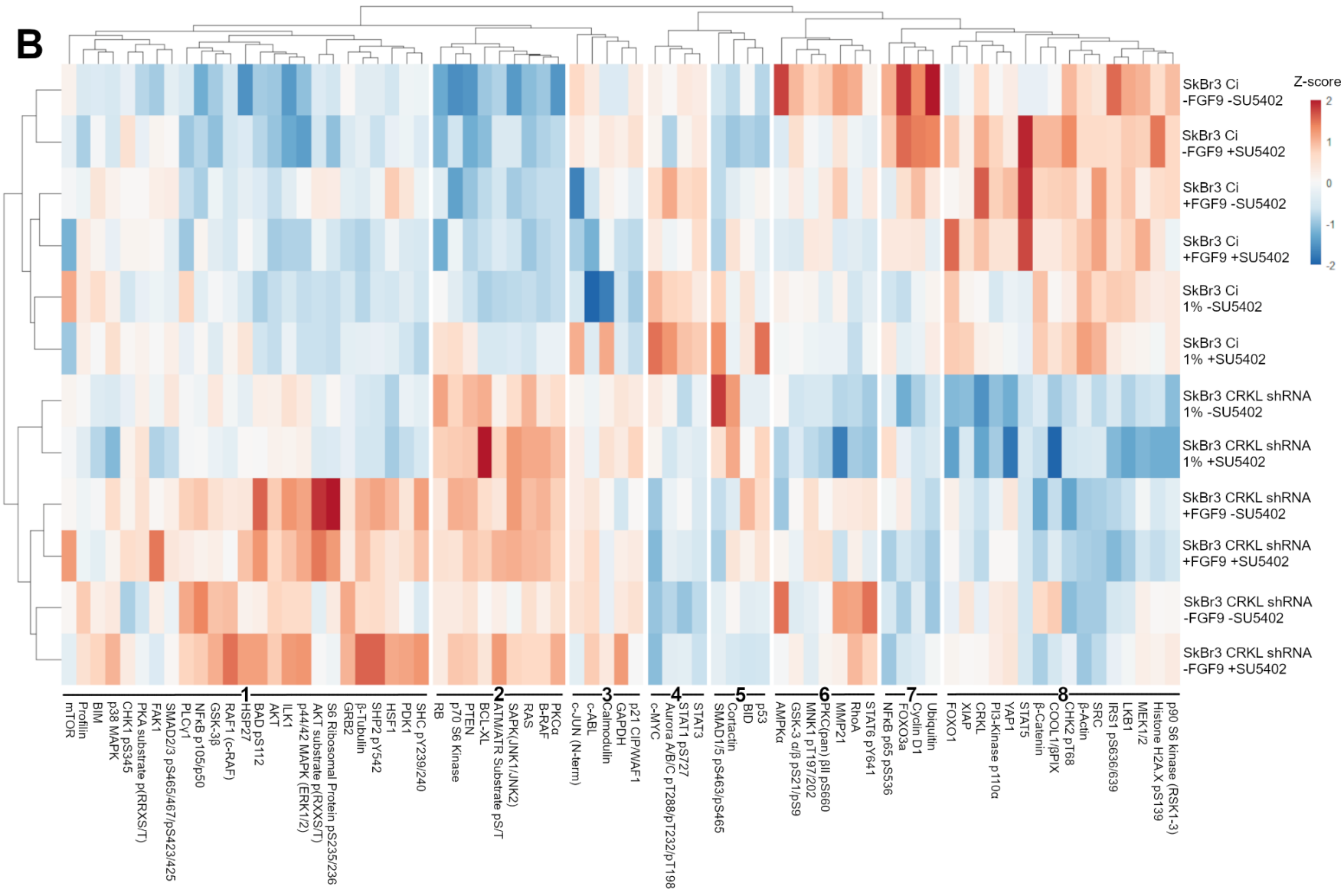

B
Z-score
2
1
0
-1
-2
SkBr3 Ci -FGF9 -SU5402
SkBr3 Ci -FGF9 +SU5402
SkBr3 Ci +FGF9 -SU5402
SkBr3 Ci +FGF9 +SU5402
SkBr3 Ci 1% -SU5402
SkBr3 Ci 1% +SU5402
SkBr3 CRKL shRNA 1% -SU5402
SkBr3 CRKL shRNA 1% +SU5402
SkBr3 CRKL shRNA +FGF9 -SU5402
SkBr3 CRKL shRNA +FGF9 +SU5402
SkBr3 CRKL shRNA -FGF9 -SU5402
SkBr3 CRKL shRNA -FGF9 +SU5402
1
2
3
4
5
6
7
8
mTOR
Profilin
BIM
p38 MAPK
CHK1 pS345
PKA substrate p(RRXS/T)
FAK1
SMAD2/3 pS465/467/pS423/425
PLCγ1
NFκB p105/p50
GSK-3β
RAF1 (c-RAF)
HSP27
BAD pS112
AKT
ILK1
p44/42 MAPK (ERK1/2)
AKT substrate p(RXXS/T)
S6 Ribosomal Protein pS235/236
GRB2
β-Tubulin
SHP2 pY542
HSF1
PDK1
SHC pY239/240
RB
p70 S6 Kinase
PTEN
BCL-XL
ATM/ATR Substrate pS/T
SAPK/JNK1/JNK2
RAS
B-RAF
PKCα
c-JUN (N-term)
c-ABL
Calmodulin
GAPDH
p21 CIP/WAF1
c-MYC
Aurora A/B/C pT288/pT232/pT198
STAT1 pS727
STAT3
Cortactin
SMAD1/5 pS463/pS465
BID
p53
AMPKα
GSK-3 α/β pS21/pS9
PKC(pan) βII pS660
MNK1 pT197/202
MMP21
RhoA
STAT6 pY641
NFκB p65 pS536
FOXO3a
Cyclin D1
Ubiquitin
FOXO1
XIAP
CRKL
PI3-Kinase p110α
YAP1
STAT5
β-Catenin
COOL1/βPIX
CHK2 pT68
β-Actin
SRC
IRS1 pS636/639
LKB1
MEK1/2
Histone H2A.X pS139
p90 S6 kinase (RSK1-3)

Figure 3-20. Changes in cell signalling between SkBr3 Ci and SkBr3 CRKL shRNA cells under different treatments. RPPA was utilised to detect phosphorylated and total protein levels in three biological repeats, and Z-score was generated to produce the heatmaps. **A)** Heatmap depicting a ratio between phosphorylated and total protein to determine changes in protein phosphorylation. **B)** Heatmap depicting total protein levels, and levels of phosphorylated proteins for which a total protein antibody was not available (thus it shows overall abundance of phosphorylated proteins relative to total protein).

3.14 Discussion

3.14.1 CRKL SH3 domains can bind various RTK C-terminal tails

CRKL adaptor protein is a known oncogene that is found to be overexpressed in multiple cancers (Senechal et al., 1996; Luo et al., 2008; Yeung et al., 2013; Zhao et al., 2013; Fu et al., 2015; Lin et al., 2015; Cai et al., 2017). Presence of CRKL is associated with more aggressive and invasive cancer stages (Lin et al., 2015; Bian et al., 2018; Kostrzewska-Poczekaj et al., 2020). However, not much is known about how CRKL is activated. Filling this knowledge gap would allow development of new cancer therapeutics. Furthermore, it has been established that SH3 domains bind directly to the proline-rich motifs in RTKs to mediate intracellular signalling (Ahmed et al., 2010; Lin et al., 2012). SH3 domains of the CRKL adapter protein are known to be crucial for cell transformation (Senechal et al., 1998; Uemura and Griffin, 1999). The purpose of this Chapter was to establish the importance of CRKL SH3 domain interaction with FGFR2 proline-rich motif in intracellular signalling.

3.14.2 Dot blot reveals multiple CRKL binding partners

To screen for potential receptor tyrosine kinases which CRKL SH3 domains could bind, a dot-blot technique was utilised. It allowed to probe for interactions between RTK C-terminal tails and purified CRKL-2xSH3 domain construct. The strongest CRKL interaction was with LMTK1 (AATK). The kinase is proposed to play a role in neuronal differentiation and apoptosis induction (Raghunath et al., 2000). However, LMTK1 is a non-receptor tyrosine kinase, thus it was not taken forward for further investigation. Second strongest CRKL

interactor was ROS1, an orphan receptor that is a known oncogene most often found as a fusion protein (Acquaviva et al., 2009). CRKL also showed very strong interaction with EphA4, an RTK which can detect membrane-bound ligands in neighbouring cells to perpetuate contact-dependent signalling and is important in nervous system development (Richter et al., 2007). Investigating an interaction between CRKL and ROS1 would be particularly interesting. ROS1 is an established oncogene, with its C-terminal portion being fused to other proteins, forming fusions such as FIG-ROS, SLC-ROS, TPM3-ROS1 (Acquaviva et al., 2009). It could be that CRKL is one of the downstream effector proteins of ROS1, allowing activation of oncogenic pathways for cell proliferation, migration and invasion. Investigating this interaction could pave way for more ROS1-positive lung cancer therapeutics. Overall, the dot-blot revealed a multitude of potential CRKL SH3-domain binding partners, revealing many interactions that can be investigated in the future.

3.14.3 CRKL can bind FGFR2 C-terminal tail

FGFR2 was identified as one of the interactors with the SH3 domains of CRKL in the RTK dot-blot screen. It was chosen to be taken forward for further experiments for several reasons. The C-terminal tail of this receptor has already been shown to be important for cell signalling (Ahmed et al., 2010), and a potential interaction between FGFR2 and CRKL has already been established in humans (Seo et al., 2009), as well as *C. elegans* model organism (Lo et al., 2010), but not looked into further in either case.

The experiments demonstrated that CRKL and FGFR2 can bind via SH3 domain – proline-rich motif interaction. FGFR2 and CRKL were found to be in the same complex in cells (Figure 3-3) and were shown to be able to physically

interact with moderate affinity in the context SH3 domain – proline-rich motif interactions (Figure 3-4, K_d = 2.510 µM ± 0.264 µM). It appears that for this interaction CRKL SH3C domain is dispensable (Figure 3-6; Figure 3-10). However, these results are not conclusive. In the absence of FGFR2 or CRKL phosphorylation, CRKL SH2 domain was still required for interaction with FGFR2, as the R21/39A mutant lacking the functional SH2 domain showed weaker binding to FGFR2 compared to full length CRKL (Figure 3-6). Loss of Y207 phosphorylation site on CRKL also affected binding to FGFR2 under non-phosphorylating conditions, as the CRKL Y207F mutant construct showed weaker binding to FGFR2 compared to wild-type CRKL (Figure 3-6). It could be that these mutations impaired normal CRKL protein folding, thus leaving other domains non-functional. While for CRKL, the SH3 domain binding site on FGFR2 is likely to be the proline-rich motif within the C24 peptide (Figure 3-7). Confidence in this CRKL binding region is reinforced by dot-blot results (Figure 3-1), where $CRKL^{2xSH3}$ construct could bind FGFR2-C2 variant, which still retains the “C24” region of the tail. The exact protein regions used for CRKL-FGFR2 interaction should be investigated further. To establish the FGFR2 proline-rich region to which CRKL can bind, full-length CRKL protein and C58 FGFR2 peptides, in which individual proline residues are mutated to alanine, could be used in MST to determine binding. MST should also be used to establish which CRKL domain is required for FGFR2 C-terminal tail binding. Full-length CRKL could be used, along with R21/39A mutation to prevent SH2 domain binding, combined with either W160L mutation to disable SH3N, or W275L mutation to disable SH3C domains. Using these double mutants, along with FGFR2 C58 peptide, would give a definite answer on which SH3 domain of CRKL mediates the interaction with FGFR2 proline-rich motif. Furthermore, it

would be interesting to visualise the FGFR2 and CRKL interaction in living cells. For that, fluorescently-tagged FGFR2 and CRKL proteins could be expressed in human cells and Fluorescence Resonance Energy Transfer (FRET) could be used to establish co-localisation using confocal microscopy. The effect of aforementioned FGFR2 and CRKL mutations on the interaction and its localisation in living cells could also be investigated using the same method.

3.14.4 CRKL can dimerise and enhance FGFR2 phosphorylation

The finding that CRKL SH3C domain is not required for FGFR2 binding is unsurprising. It has been postulated that the SH3C domain is likely to be unable to bind known proline-rich motifs due to the lack of a conserved tryptophan, as well as negatively charged residues within the binding pocket. Instead, the SH3C domain was found to act as a dimerisation domain for CRKL proteins (Harkiolaki et al., 2006). CRKL can dimerise as measured by MST (Figure 3-8), and the dimerisation affinity (K_d = 0.093 ± 0.008 µM for full-length protein) was found to be stronger than that reported before (K_d = 6.1 µM) (Harkiolaki et al., 2006). Potential CRKL oligomers were also detected in the MST experiment (K_d = 7.89 ± 1.19 µM), which also corroborates previous report (Harkiolaki et al., 2006). So far it is unclear whether CRKL oligomerisation occurs in cells. If oligomerisation does occur, CRKL oligomers could potentially act as signalling hubs to bring different proteins within close proximity to enhance signalling activation.

It appears that CRKL binds to FGFR2 as a dimer, as K_d for CRKL dimerisation is stronger than the K_d for CRKL-FGFR2 binding (0.093 µM and 2.510 µM respectively). If CRKL bound to FGFR2 is a dimer, it is then likely that

CRKL dimer can bring two molecules of FGFR2 together. Previously, GRB2 was shown to dimerise and form a heterotetramer with two FGFR2 molecules (Lin et al., 2012). Dimeric CRKL enhanced autotransphosphorylation between FGFR2 receptors (Figure 3-9). This establishes CRKL as a potentially important regulator of FGFR2 phosphorylation. One way this could happen is that CRKL dimer would form via the SH3C domains, while each SH3N domain would bind a separate molecule of FGFR2 at the C-terminal tails to bring two receptor molecules within close proximity to enhance receptor phosphorylation (Figure 3-21). The next step is to further characterise CRKL-induced FGFR2 phosphorylation. In FGFR2-GRB2 heterotetramer, GRB2 enhances phosphorylation of a subset of FGFR2 tyrosine residues, but sterically hinders any downstream signalling activity. GRB2 is released only when an extracellular ligand binds and fully activates FGFR2 (Lin et al., 2012). It would be interesting to see whether CRKL has the same function as GRB2. Alternatively, CRKL dimer could enhance ligand bound FGFR2 phosphorylation efficiency leading to stronger receptor activation. If that is the case, downstream kinase studies could be performed to see whether enhanced receptor phosphorylation leads to enhanced downstream (MAPK and PI3K/AKT pathway) signalling. This newly discovered CRKL function can have potential roles in cancer. If CRKL enhances FGFR2 activity, overexpression of the adapter protein could lead to increased forced dimerisation and trans-activation of FGFR2, which would then lead to canonical downstream FGFR2 signalling and increased cell growth and invasion (Schlessinger, 2014; Ségaliny et al., 2015). It would also be interesting to see whether this CRKL-enhanced FGFR2 dimerisation could activate the receptor in absence of stimulus. To investigate this, CRKL could be overexpressed in cells which also contain FGFR2 mutant lacking functional

ligand recognition domain to prevent ligand induced FGFR2 activation while keeping the kinase domain functional. Downstream MAPK/ERK and AKT signalling activation could be then measured as an output of FGFR2 function.

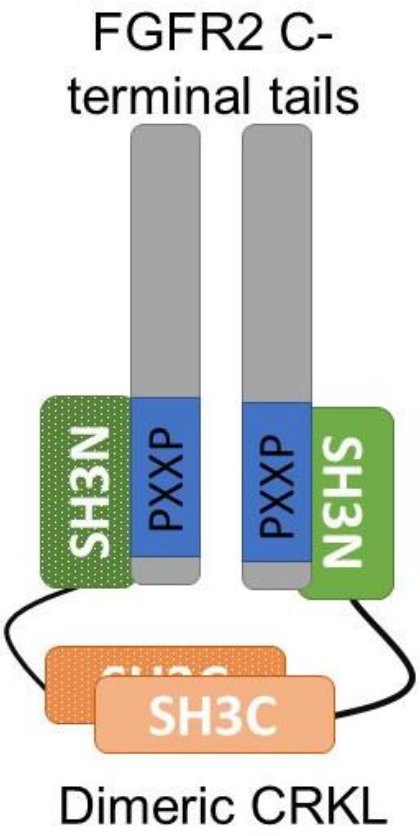


Figure 3-21. Dimeric CRKL can enhance FGFR2 activation. CRKL dimerises via SH3C domains, while individual SH3N domains interact with proline-rich motif within the FGFR2 tail. This binding brings two FGFR2 molecules within close proximity, which enhances FGFR2 phosphorylation.

3.14.5 CRKL Y48, Y198 and Y207 residues are phosphorylated by FGFR2

CRKL has several tyrosine residues that can be phosphorylated, but the most documented one is Y207, which is phosphorylated in CML by BCR/ABL (Oda et al., 1994). Using an *in vitro* kinase assay, FGFR2 was found to phosphorylate CRKL Y207 (Figure 3-10). This kinase assay was combined with GST-pulldown study to see which CRKL domains are required for binding to

dephosphorylated or phosphorylated FGFR2. Functional CRKL SH2 domain was required to bind phosphorylated FGFR2, while the SH3 domains were not necessary (Figure 3-10). The results show that CRKL can bind active (phosphorylated) FGFR2 receptor via both SH3 and SH2 domains, suggesting that CRKL has the potential to bind both inactive and active (phosphorylated) FGFR2 receptor, much like GRB2.

CRKL tyrosine phosphorylation on other residues by FGFR2 was detected using mass spectrometry. CRKL was found to be phosphorylated on Y48, Y198 and Y207 residues (Figure 3-12). It is likely that only one site can be phosphorylated at one time, as indicated by intact protein mass measurement (Figure 3-11). It is currently unclear how CRKL phosphorylation could affect its function. Y48 site is within the SH2 domain. Perhaps its phosphorylation could distort the domain structure and prevent SH2 from binding other pY residues. Y198 and Y207 residues are within the linker between SH3 domains. Their phosphorylation could be a part of adaptor protein function of CRKL. The adaptor protein can bring multiple proteins together in close proximity to activate downstream signalling. Having pY residues would allow CRKL to recruit SH2 domain-containing proteins, while SH2 and SH3 domains already allow it to recruit pY-containing and proline-rich motif-containing proteins, respectively. Alternatively, pY residues within the linker could act as a ligand for CRKL's own SH2 domain. Such self-binding occurs in CRK-II, and it was shown to block SH3N binding pocket, thus acting an inhibition mechanism (Rosen et al., 1995). However, it was shown that the SH3N binding pocket is not blocked in CRKL if SH2 is bound to phosphorylated tyrosine (Jankowski et al., 2012). Perhaps instead phosphorylation could be a method for regulating CRKL signalling. Once CRKL binds an upstream receptor, it would be phosphorylated. Resulting

SH2 binding to pY of the linker between the SH3 domains, could lead to a conformational change in which SH3N is primed for enhanced proline-rich motif binding. This aspect of CRKL signalling is an interesting potential venue to explore. Determining the structure of CRKL phosphorylated on Y48, Y198 and Y207 residues could give insight on how phosphorylation could affect CRKL functions, including binding to FGFR2 and other upstream and downstream binding partners, as well as dimerisation. Biological relevance of phosphorylation should also be investigated. Once downstream signalling induced by CRKL is defined, effect of loss of any of the phosphorylation sites could be investigated.

3.14.6 CRKL and FGFR2 role in cell migration and downstream GTPase activation

As CRKL is known to be involved in cell migration, wound closure was measured as a phenotypical outcome of CRKL/FGFR2 interaction. It was found that CRKL is partially responsible for RTK-mediated migration of HEK293T cells, as cell migration was reduced in CRKL shRNA cells compared to Ci cells when cells contained either GFP or FGFR2-GFP, with no SU5402 inhibitor present. When the cells were treated with SU5402 inhibitor, wound closing was reduced to similar levels under all conditions, suggesting that SU5402 could inhibit not only FGFR2, but also other RTKs which help regulate cell migration. Furthermore, the experiment was performed under basal conditions (1% FBS), indicating that cell migration can occur without strong receptor stimulation.

RAP1 and RAC1 signalling pathways, both of which can be activated by CRKL, are major contributors to cell migration (Bauer et al., 2007; Y.L. Zhang et al., 2017). Their activation was measured under the same conditions as

HEK293T cell migration (Figure 3-15, Figure 3-16). The results were highly variable, and no statistical significance was observed. This means that migration phenotype cannot be connected to RAP1 or RAC1 signalling pathways. These results pointed out a few flaws in the method used. For these experiments, HEK293T cells were used. They do not express FGFR2, thus FGFR2-GFP was transiently expressed in cells. The level of FGFR2-GFP expression could not be finely controlled and might have been highly variable between different conditions and repeats, which could have resulted in a high variance in RAP1 and RAC1 activation outcomes. A complete CRKL knockout should also be considered, which could be done using CRISPR/Cas9 system. Furthermore, SU5402 is an inhibitor that is not particularly specific to FGFR2 – it can also inhibit VEGFR2, FGFR1 (present in HEK293T at high levels), PDGFRβ and EGFR. This could have had unintended effects RAP1 and RAC1 activation. These issues could be remedied by adjusting the assay conditions: a HEK293T cell line stably expressing FGFR2 could be generated, and to ensure FGFR2 tyrosine kinase activity inhibition, FGFR2 Y656/657F kinase dead mutant could be used instead of SU5402 inhibitor. However, there is also a possibility that CRKL is not solely responsible for RAP1 and RAC1 activation in these cells, as HEK293T CRKL shRNA cells did not show a strong reduction in GTP-RAP1 and GTP-RAC1 (Figure 3-15, Figure 3-16).

To remedy shortcomings of using HEK293T cells, the SkBr3 breast cancer cell line, which expresses FGFR2 endogenously, was also used. A wound healing assay was attempted, but these cells did not migrate well, and had a low CRKL knockdown efficacy. Active RAP1 pulldown assay revealed that overall, RAP1 activation is highly variable, suggesting that it is transient. RAP1 was found to be GTP-bound in SKBR3 Ci cells under all conditions:

starved (-FGF9), basal (1% FBS) and stimulated (+FGF9), with SU5402 inhibitor also having minor effects on RAP1 activation (Figure 3-18). Data from SkBr3 Ci cells suggests that RAP1 can be activated via signalling that is not purely dependent on cell stimulation. Under all conditions, SkBr3 CRKL shRNA cells show severely reduced RAP1 activation. The differences in RAP1 activation between control (SkBr3 Ci) and CRKL knockdown cells (SkBr3 CRKL shRNA) show statistical significance under starved (-FGF9) and stimulated (+FGF9) conditions, similar to the results shown in HEK293T cells. It is still clear that reduction of CRKL expression strongly impacts RAP1 activation. The results indicate that in SkBr3 cells, CRKL could be required for RAP1 activation, and that this activity is likely to be stimulation independent. RAP1 activation in cells grown under basal (1% FBS) conditions show no significant differences, perhaps due to the possibility that RAP1 could be activated via other pathways that do not involve CRKL. It is also unclear whether FGFR2 is involved in CRKL-mediated RAP1 activation. If FGFR2 is indeed involved, it would be likely that FGFR2-CRKL interaction would be mediated by proline-rich motif – SH3 domain interaction, as RAP1 activation does not seem to require stimulation. To see whether FGFR2 is required for CRKL to activate RAP1, a SkBr3 cell line stably expressing FGFR2 mutant lacking the C58 C-terminal tail, which normally contains CRKL-binding proline-rich motifs (or a more precise mutation once the exact CRKL binding site on FGFR2 has been established) could be generated and used for active RAP1 pulldown experiments.

RAC1 activation was also investigated in SkBr3 cells. Although RAC1 is expressed in SkBr3 cells, no active RAC1 was pulled down (Figure 3-19). The results suggest that RAC1 is not activated in SkBr3 cells by FGF9.

These results allow a few conclusions to be drawn. First, even a weak CRKL knockdown can result in strong reduction of active RAP1 levels, indicating that in SkBr3 cells CRKL is almost solely responsible for RAP1 signalling. CRK-II can also activate RAP1 (Ling et al., 2003), but higher expression of this adaptor protein in SkBr3 CRKL shRNA cells did not lead to higher abundance of GTP-RAP1, suggesting that this protein is not involved in RAP1 activation. High variability in active RAP1 in both HEK293T and SkBr3 cells suggests that perhaps RAP1 activation is very transient. Overall, measure of RAP1 or RAC1 activation does not seem to be a suitable method of investigating signalling consequences of FGFR2-CRKL interaction.

3.14.7 Role of CRKL in regulating downstream signalling pathways

RPPA allowed to evaluate the effects of CRKL signalling on a variety of signalling pathways in SkBr3 cells. There were clear differences in clustering between control (SkBr3 Ci) and CRKL knockdown (SkBr3 CRKL shRNA) cells. Broad overview of the change in phosphorylated and total protein levels suggested a model showing major differences in signalling pathways between SkBr3 Ci and SkBr3 CRKL shRNA cells (Figure 3-22). In cells containing CRKL, signalling appears to be biased towards AKT signalling pathway (Figure 3-22A). Furthermore, elevated levels of phosphorylated GSK-3-β and PTEN were detected as well at the PTM sites associated with reduced protein function, thus preventing these proteins from inhibiting AKT signalling (Vazquez et al., 2000; Beurel et al., 2015). STAT5 expression is also increased in these cells, which contributes to increased PI3-K p110α expression (Schmidt et al., 2014). However, AKT activation is dependent on growth conditions. AKT appears to be most active in SkBr3 Ci -FGF9 -SU5402 cells, as both T308 and S473 sites are

phosphorylated. Interestingly, in SkBr3 Ci cells stimulated with FGF9, both with or without SU5402, the T308 site of AKT is not highly phosphorylated, suggesting that in these cells AKT is not fully active. In non-starved SkBr3 Ci cells, grown either under basal conditions (1% FBS) or stimulated with FGF9, c-ABL is also phosphorylated, and thus is active (Brasher and Van Etten, 2000). c-ABL can also potentially contribute to AKT activity (Atfi et al., 2005). These cells also show increased levels of p21 CIP/WAF phosphorylation (caused by AKT signalling), which promotes DNA synthesis (Li et al., 2002). SkBr3 Ci cells show increased expression of proteins which promote cell growth and survival, including SRC, STAT5, YAP, XIAP and Cyclin D1. The RPPA results for SkBr3 Ci cells suggest that cell signalling normally associated with cell stimulation can also be activated under low (basal, 1% FBS) stimulation or starvation (-FGF9) conditions.

The results from SkBr3 CRKL shRNA cells suggest that reduced CRKL expression affects cell signalling. Under basal (1%) or stimulating (+FGF9) conditions, these cells show an upregulation of MAPK/ERK signalling pathway (Figure 3-22B), as evident by increased phosphorylation of c-RAF, MEK1/2, ERK1/2, and RSK1-3. Under +FGF9 conditions (but not basal conditions), this signalling is further reinforced by increased levels of phosphorylated SHP2 and SHC2, as well as increased GRB2 expression. Because MAPK pathway is not activated under starved conditions, and inhibition of canonical FGFR2 signalling does not have a strong effect on protein phosphorylation levels, it could be that this signalling is mediated by activation of other receptors. For example, it could be FGFR3, which can also be activated by FGF9. Interestingly, PLCγ1 signalling is increased in both starved (-FGF9) and stimulated (+FGF9) SkBr3 CRKL shRNA cells (Figure 3-22C). Due to SU5402 having no severe effect on

PLCγ1 activation under starved or stimulated conditions, and the fact that PLCγ1 is more active under starved conditions, it could be that in the absence of CRKL, the phospholipase could be activated via non-canonical signalling. This suggests that CRKL could act similarly to GRB2 in attenuating aberrant phospholipase activity (Timsah et al., 2014). Changes seen in SkBr3 CRKL shRNA cells shows several pro-survival mechanisms: inactivation of pro-apoptotic BIM, suppression of DNA damage pathway by reducing RB activity, as well as increased expression of BCL-XL anti-apoptotic protein and inhibition of pro-apoptotic CHK2.

Next, it would be interesting to see whether FGFR2 is required for the cell signalling pathways found up- or down-regulated in the RPPA assay. A mutant FGFR2 constructs could be introduced into SkBr3 Ci and CRKL shRNA cells, which would have kinase-dead FGFR2, FGFR2 lacking the CRKL-binding C-terminal tail, or both mutations. This would allow to see if signalling pathways dependent on CRKL are also dependent on FGFR2, and whether they are dependent on canonical or non-canonical FGFR2 signalling.

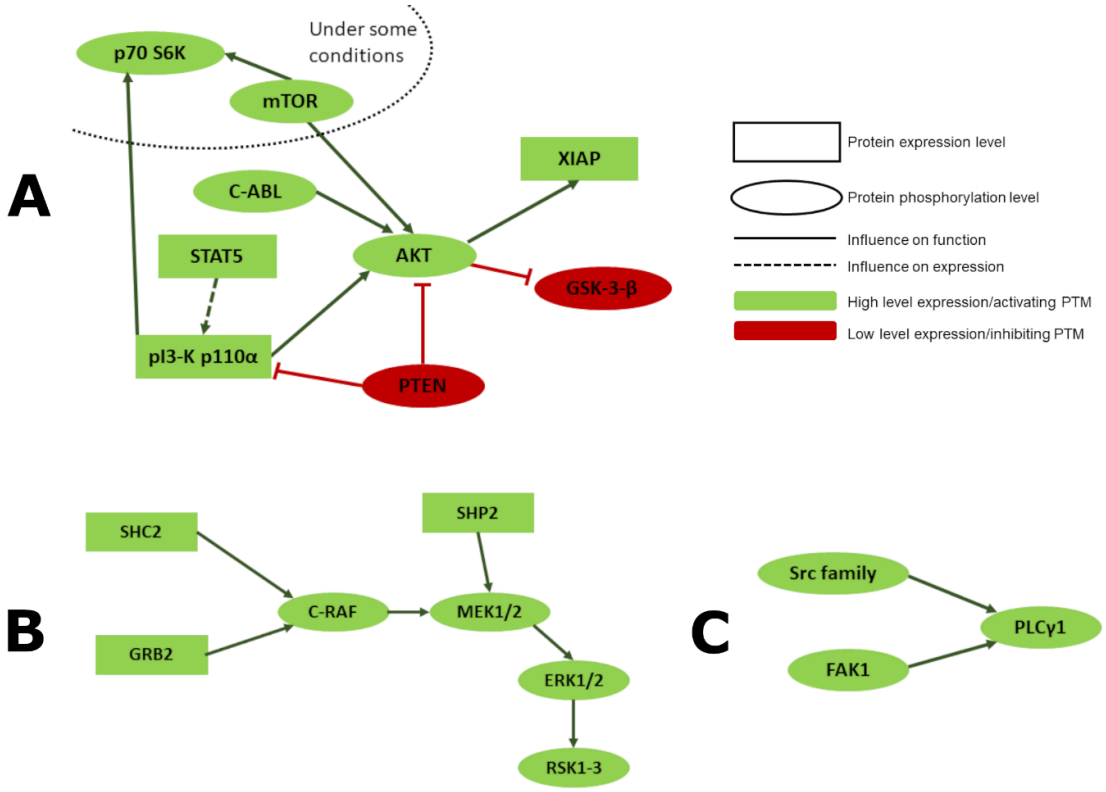


Figure 3-22. Cell signalling pathways activated in SkBr3 cells. A) Proteins within the AKT signalling pathway show increased phosphorylation and expression levels in SkBr3 cells in the presence of higher CRKL levels. **B)** Proteins within the MAPK/ERK signalling pathway that show increased phosphorylation and expression levels in SkBr3 CRKL shRNA cells grown under basal (1% FBS) and stimulated (+FGF9) conditions. **C)** Src family, FAK1 and PLCγ1 phosphorylation is increased in SkBr3 CRKL shRNA cells grown under starved (-FGF9) and stimulated (+FGF9) conditions. (PTM – post-translational modification)

Overall, the RPPA results suggest that CRKL participates in signalling via the AKT pathway. Once CRKL levels are reduced, AKT signalling is reduced, while MAPK/ERK signalling is upregulated. Furthermore, AKT signalling can occur when cells are starved, while MAPK/ERK signalling requires stimulation. These findings are not surprising, as CRKL-induced AKT activation has been described before (J. Zhang et al., 2017). SkBr3 Ci and SkBr3 CRKL shRNA cell lines promote growth and cell survival by upregulating the expression or activation of different sets of pro-survival proteins. Interestingly, RPPA findings suggest that role of CRKL in SkBr3 cell signalling is

opposite to that of CRKL in NSCLC (AALE cell line). Lung cancer cells overexpressing CRKL showed RAP1-dependent upregulation of MAPK/ERK pathway, with no changes to AKT signalling (Cheung et al., 2011). RAP1 can indeed activate MAPK pathway by activating B-RAF in a cell-type specific manner (Ohtsuka et al., 1996). This suggests that the role of CRKL in cell signalling could be dependent on cellular context as well.

The results propose a hypothesis on how the “CRKL-dependent signalling switch” could occur (Figure 3-23). In SkBr3 cells, CRKL is required for RAP1 activation. Active RAP1 can activate PI3-K/AKT signalling (Tsygankova et al., 2001), and inhibit c-RAF function and subsequent MAPK/ERK pathway activation (Kitayama et al., 1989). This is enforced by the RPPA results, where in SkBr3 Ci cells c-RAF in more phosphorylated at S259, which enables 14-3-3 to bind c-RAF and inhibit its function (Light et al., 2002), while in SkBr3 CRKL shRNA cells c-RAF is phosphorylated at S338, which is an activating PTM (Chong et al., 2001). When CRKL is knocked down, RAP1 is not activated, and in turn PI3-K/AKT signalling is inactive, while MAPK/ERK signalling is upregulated due to unhindered c-RAF activity. Alternatively, the change between MAPK/ERK and PI3-K/AKT signalling pathway could be due to increased CRK-II expression in CRKL knockdown cells. Thus, to confirm the hypothesis, perhaps a new variant of CRKL knockdown cells should be generated, in which CRK-II expression is unchanged.

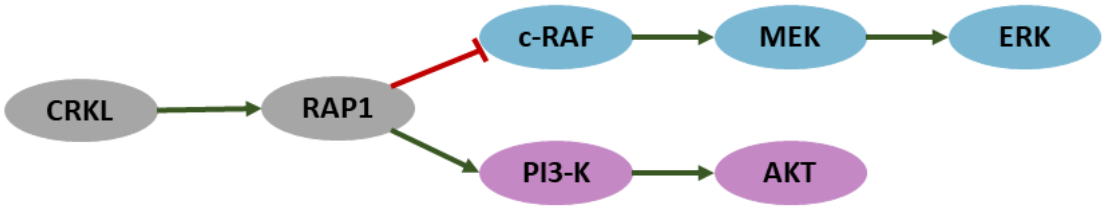


Figure 3-23. CRKL-regulated cell signalling switch. CRKL can activate RAP1 (possibly via C3G). Active RAP1 can in turn activate PI3-K/AKT signalling pathway, while simultaneously inhibiting c-RAF, and thus the MAPK/ERK signalling pathway. Gray – CRKL/RAP1, blue – c-RAF/ERK signalling axis, pink – AKT signalling axis, green arrow – activation, red line – inhibition.

The RPPA assay acts as a starting point in investigating the signalling consequences of FGFR2-CRKL interaction – it revealed which proteins can be affected by CRKL presence under starved, stimulated and basal conditions. This would allow to refine the RPPA antibody list for future experiments. Next, to determine how FGFR2 proline-rich motif and CRKL SH3 domain direct interaction can affect cell signalling, RPPA should be repeated with additional cell lines that include mutant FGFR2 lacking the intact CRKL SH3 binding domain. Assays such as colony formation and MTT survival assay could be used to see whether signalling via CRKL-RAP1-AKT pathway gives SkBr3 cells a survival advantage (or disadvantage) over SkBr3 cells lacking CRKL and signalling via the MAPK/ERK pathway. It would also be interesting to see whether CRKL-induced AKT signalling also occurs in other cell lines.

3.15 Conclusions

This Chapter focused on the interaction between the FGFR2 proline-rich motif and CRKL SH3 domains, and its consequence for downstream signalling. FGFR2 proline-rich motif within the C24 peptide is the potential CRKL binding

site, while the CRKL SH3C domain is not required for this interaction. Furthermore, CRKL was shown to dimerise, and dimeric CRKL was shown to enhance FGFR2 phosphorylation. Active FGFR2 in turn can phosphorylate CRKL at three distinct sites, although the function of phosphorylated CRKL is still unknown and require further investigation. FGFR2-CRKL signalling that is independent of FGFR2 kinase activity was also found to be a likely contributor to HEK293T cell migration. In HEK293T cells, role of CRKL in RAP1 and RAC1 is unclear, while in SkBr3 cells CRKL was important for RAP1, but not RAC1 activity. Investigating downstream signalling in SkBr3 cells revealed that loss of CRKL leads to a change in active signalling pathways.

Chapter 4 Interaction between *Caenorhabditis elegans* EGL-15 (FGFR) and CED-2 (CRKL)

4.1 Introduction

4.1.1 *Caenorhabditis elegans (C. elegans)*

The nematode worm *C. elegans* is a small, non-parasitic, free living organism. The animal is easy to maintain, and its transparency allowed mapping of the cell lineage by microscopy (Sulston et al., 1983). *C. elegans* was also the first multicellular organism to have its whole genome sequenced, revealing that up to 80% of human genes are conserved in the animal (Markaki and Tavernarakis, 2010). Hence, it has been invaluable in facilitating our understanding in various aspects of human biology. *C. elegans* have a short lifespan and fast reproductive cycle, which is hugely beneficial for generating mutant animals as well as research fields that require long observation time (e.g. ageing) (Zhang et al., 2020). Egg development starts inside the mother and continues after it has been laid. Once hatched, the animals undergo four larval stages (L1 – L4). If the animals are starved during the L2 stage, they can switch their developmental programme from L3 to “dauer” larva, where development is arrested. Once they encounter a food source, the development continues into the L4 stage (Byerly et al., 1976). Adult *C. elegans* are self-fertilising hermaphrodites, allowing easy propagation of genetic mutants. Male nematodes arise at a low frequency (Honda, 1925), enabling genetic crosses to generate strains containing multiple mutations or transgenes. These features establish *C. elegans* as a great model organism to investigate molecular biology (Brenner, 1974).

4.1.2 *C. elegans* reproductive system

The hermaphrodite reproductive system (Figure 4-1) is comprised of three parts: the germline, the somatic gonad and the egg-laying apparatus (Kimble and Hirsh, 1979). The hermaphrodite germline contains male and female gametes. While oocytes are produced continually, there is a limited supply of sperm that has been produced during the L4 stage (Ward et al., 1981). The somatic gonad, which controls gametogenesis, includes the uterus, the spermatheca, the spermatheca-uterine valve, the distal tip cells and the gonadal sheath (Hirsh et al., 1976; McCarter et al., 1997). The egg-laying apparatus consists of the uterus, the vulva and their respective muscles and neurons (Hirsh et al., 1976; Sulston and Horvitz, 1977). Several reproductive system features relevant to this work will be discussed below.

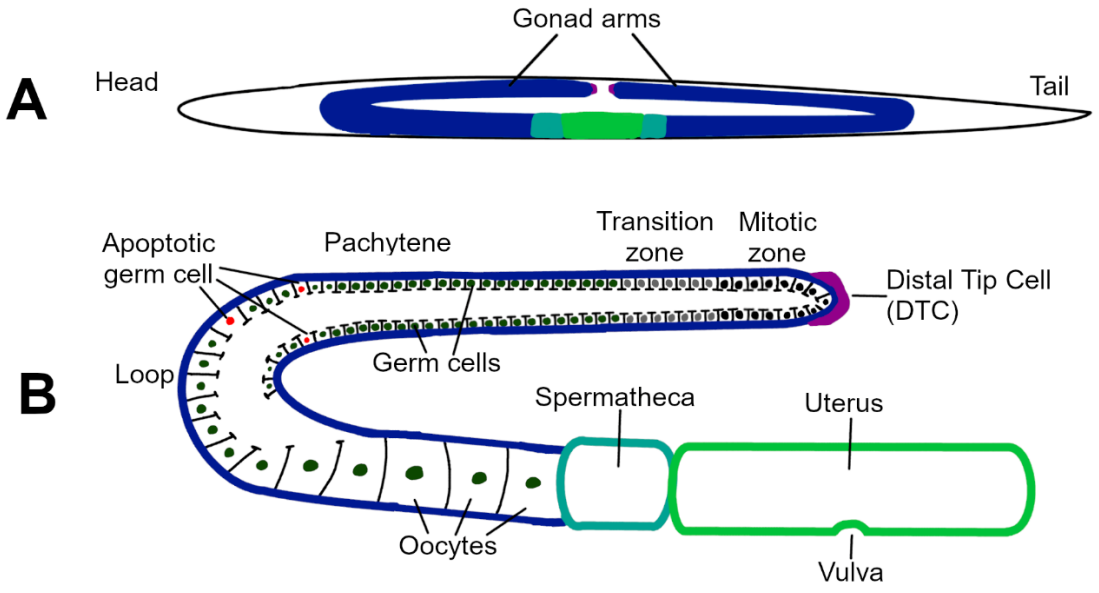


Figure 4-1. Caenorhabditis elegans adult hermaphrodite reproductive system. A) The gonad in the intact animal shows two U-shaped gonad arms (dark blue), connected by the uterus. **B)** Closer look at the reproductive system. DTC, loop region containing apoptotic germ cells, and vulval muscles are key features relevant to this work.

4.1.2.1 Distal Tip Cell

The tip of each gonad arm has a somatic cell called the Distal Tip Cell (DTC) (Figure 4-1) (Hall et al., 1999). It is required for mitosis promotion and inhibition of meiosis of the germ cells during development and in adulthood (Kimble and White, 1981; Crittenden et al., 2003). The DTC is also responsible for leading the elongation of gonad arms during development (Figure 4-2A) (Cecchetelli and Cram, 2017). Both gonad arms start elongation at the ventral side of the animal, near the site of the future vulva. The DTC leads arm elongation towards anterior or posterior of the animal during L2/L3 stages (Phase 1). During the mid-L3 stage, the arms make a turn towards the dorsal side, followed by a second turn (Phase 2). During the L4 stage the gonad arms migrate along the dorsal side back towards the centre of the body and ceases migration upon reaching it, resulting in a characteristic U-shape of the gonad (Phase 3). Several signalling pathways linked to cytoskeletal arrangement in the DTC are required for correct migration (Figure 4-2B) (Nishiwaki, 1999; Cram et al., 2006; Cecchetelli and Cram, 2017).

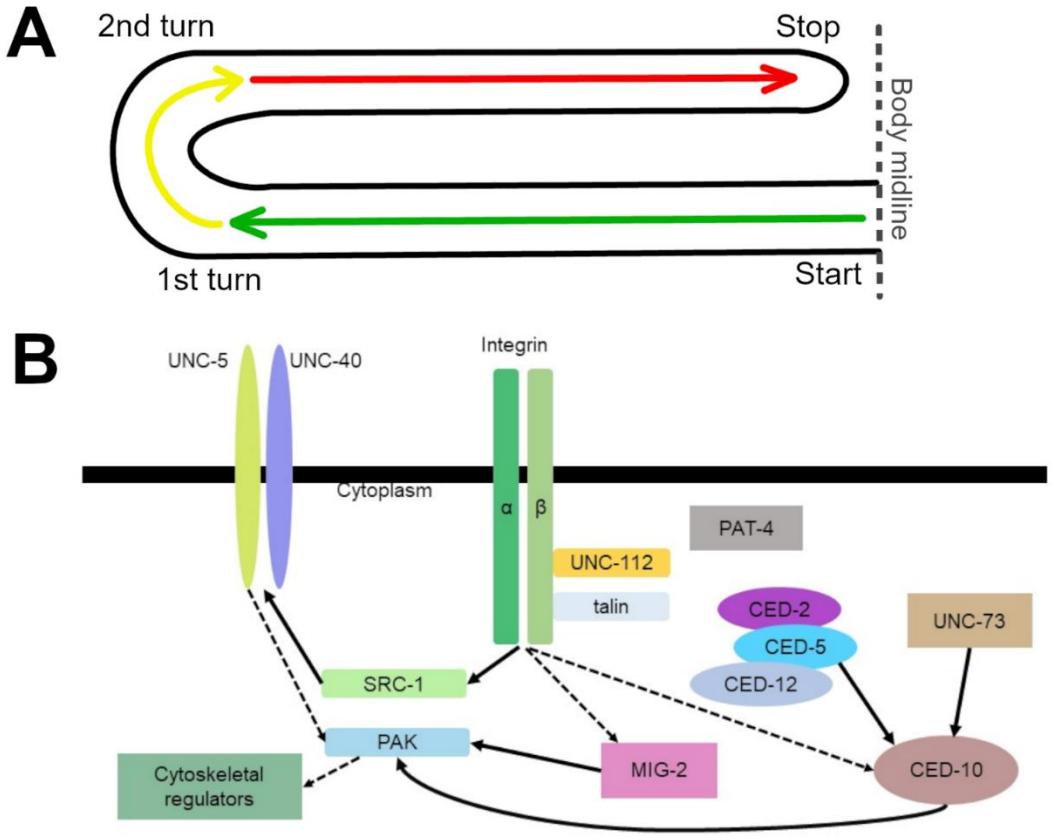


Figure 4-2. DTC migration in C. elegans. A) Gonad migration during development occurs in three steps to yield the characteristic U-shape. Phase 1 is green, phase 2 is yellow, phase 3 is red. **B)** Proteins essential for DTC migration. Solid arrows are direct interactions, while dashed arrows represent indirect interactions.

4.1.2.2 Programmed germ cell death

Programmed cell death occurs in *C. elegans* as a part of normal development. 113 somatic cells undergo apoptosis in the embryo, with 18 additional cells doing so in L2 stage (Wang and Yang, 2016). In the adult animal, half of the germline cell population undergoes cell death (programmed or stress-induced) close to the gonad loop (Gumienny et al., 1999). Germline cell apoptosis is thought to have two functions: to act as quality control, and to act as nurse cells, providing additional cytoplasm and organelles to developing oocytes (Gumienny et al., 1999; Huelgas Morales and Greenstein, 2018).

There are three distinct programmed cell death phases. The cell commits to death during the specification phase. The apoptotic programme is then activated during a killing phase, followed by the execution phase during which the apoptotic cells are engulfed by nearby cells (Horvitz, 1999). In the germline, engulfment is performed by gonadal sheath cells belonging to the somatic gonad, which surround the gonad arm (Gumienny et al., 2001). Engulfing cells recognise apoptotic signals exposed on the cell surface, leading to activation of cell signalling which results in actin cytoskeleton rearrangement and phagosome maturation (Wang et al., 2007; Fullard et al., 2009; Neukomm et al., 2011). There are two partially redundant engulfment pathways (Figure 4-3). One is mediated by *ced-1, ced-6* and *ced-7*, while another pathway centres around *ced-2, ced-5* and *ced-10* (Gumienny et al., 2001). An additional pathway, *abl-1* negatively regulates corpse clearance by inhibiting *abi-1* (Hurwitz et al., 2009).

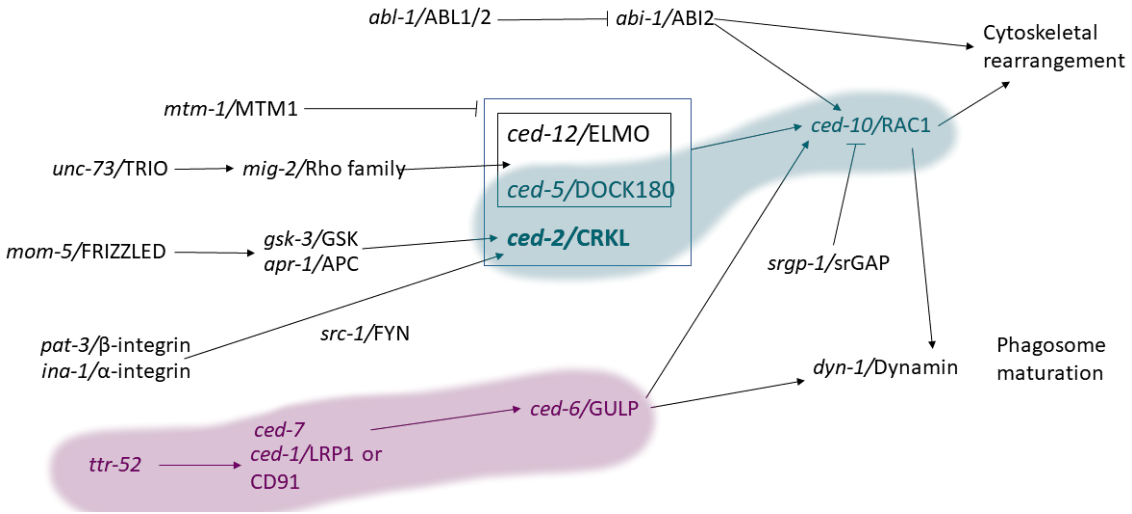


Figure 4-3. Apoptotic germline cell engulfment pathway. *C. elegans* genes are in lower case italics, while their human orthologues are named after "/". *Ced-1/ced-6/ced-7* pathway is highlighted in pink, while *ced-2/ced-5/ced-10* pathway is highlighted in blue.

4.1.2.3 Vulval muscles

Vulva muscles are an essential part of the egg-laying apparatus. During development, at the L1 stage, the mesoderm (M) blast cell divides and gives rise to two (left and right) sex myoblast (SM) cells. During the L2 stage the SMs migrate towards the site of the future vulva present in the centre of the developing gonad. After reaching their final position, SMs divide and form vulval and uterine muscles (Sulston and Horvitz, 1977). The muscles are an essential part of the egg-laying process – in mutants lacking developed vulval muscles the eggs are trapped inside the hermaphrodite, resulting in a “bag of worms” phenotype (Burdine et al., 1998).

4.1.3 EGL-15 (FGFR)

C. elegans have a single FGFR, which is encoded by the *egl-15* gene. *egl-15* is an essential gene which has multiple roles in development (DeVore et al., 1995). Protein structure is highly conserved between EGL-15 and mammalian FGFRs – it comprises three Ig repeats in the extracellular region, a transmembrane portion, and an intracellular region containing a juxtamembrane region, a split tyrosine kinase domain, and a C-terminal tail containing proline-rich motifs (Borland et al., 2001). As with its human FGFR orthologues, EGL-15 is subjected to alternative splicing in several regions. Mutually exclusive splicing of exon 5 results in two isoforms, called EGL-15(5A) and EGL-15(5B) (Goodman et al., 2003). These isoforms are functionally different and will be discussed below. The EGL-15 C-terminal tail is also subject to alternative splicing, producing several different tail variants (WormBase, release WS280). EGL-15 structure is summarised in Figure 4-4. There are several key differences between the proteins. First, in EGL-15, alternative splicing leads to

different variants of an EGL-15 – specific insert between IG1 and IG2 Ig-like repeats, while in FGFR2, it is the D3 (third Ig-like repeat) that is subjected to alternative splicing. However functionally, different variants have the same function – alternative splicing changes the receptors affinity for different ligands. The second key difference lies in the C-terminal tail. EGL-15 generally contains a longer C-terminal tail than FGFR2. While the exact position of proline-rich motifs, and the motifs themselves are not generally conserved (see Appendix D for the alignment), the very C-terminus of both proteins does contain multiple proline-rich motifs which could act as binding regions for SH3 domain - containing proteins, and thus the functional role of the C-terminal tails of these proteins is conserved.

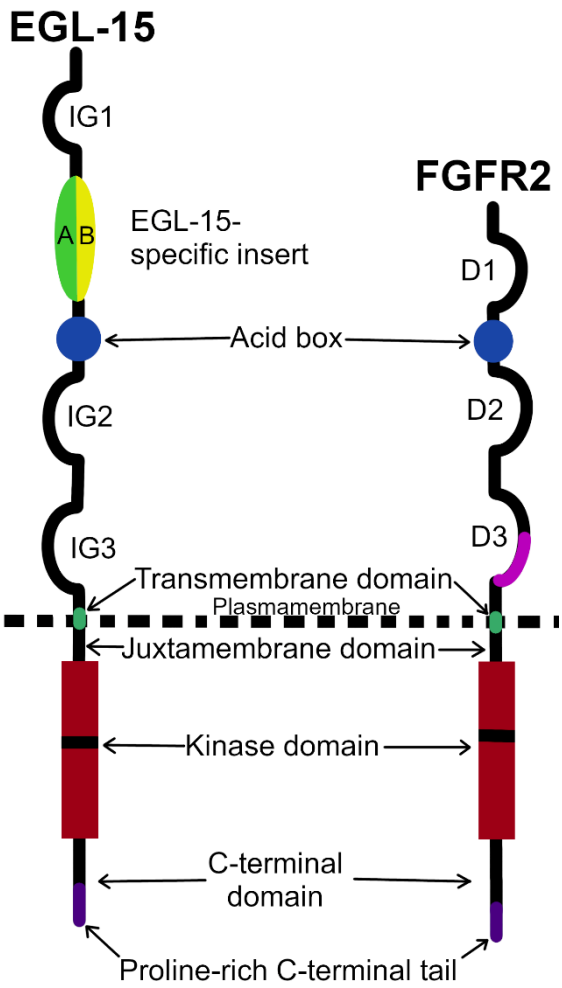


Figure 4-4. EGL-15 and FGFR2 protein structure. EGL-15 extracellular domain contains of three Ig-like repeats (IG 1-3), with an EGL-15-specific insert and an acid box between IG1 and IG2. EGL-15-specific insert has two variants (5A and 5B), produced by alternative splicing of exon 5, which determines ligand specificity. FGFR2 extracellular domain also contains three Ig-like repeats (D1-3), with D3 being subjected to alternative splicing at the pink area, and an acid box. Intracellular portion of both proteins consists of a juxtamembrane domain, a kinase domain, and a C-terminal domain containing phosphorylatable tyrosine residues important for signalling, as well as a C-terminal tail containing proline-rich motifs.

4.1.3.1 EGL-15 isoforms

The EGL-15(5A) isoform is involved in cell migration. The isoform is primarily found in the M lineage, which gives rise to SMs (Kuroyanagi et al., 2006; Lo et al., 2008). EGL-15(5A) is essential for the migration of SM cells during gonad development (Goodman et al., 2003). At the end of the L1 stage,

two SM cells migrate to the centre of the gonad (Sulston and Horvitz, 1977), where they divide and give rise to vulva muscles essential for egg-laying. The FGF orthologue EGL-17 is a ligand for EGL-15(5A), and acts as a chemoattractant for the migrating SMs (Burdine et al., 1998; Chen and Stern, 1998; Branda and Stern, 2000). SEM-5, an orthologue of human GRB2, is required for downstream signalling in SM migration (Clark et al., 1992). It can directly bind EGL-15 to activate downstream signalling. SEM-5 binds Y1009 and Y1087 phosphorylated tyrosine residues present in the EGL-15 Type-I C-terminal tail via the SH2 domain. This leads to activation of LET-60, an orthologue of RAS (DeVore et al., 1995; Chen and Stern, 1998), and downstream MAPK signalling cascade (Figure 4-5). EGL-15(5A) is also required for migration and branching of several neurons - it assists in the outgrowth of several axons during development and maintains correct axon position during adulthood. Interestingly, this function requires *let-756*, which is typically the ligand for the EGL-15(5B) isoform (Bülow et al., 2004). Furthermore, EGL-15(5A) and its ligand EGL-17 are important in regulating canal-associated neuron (CAN) migration, although the downstream pathway regulating this function is not yet clear, as *let-60/*RAS, which normally signals downstream of EGL-15, is not required for this function (Fleming et al., 2005).

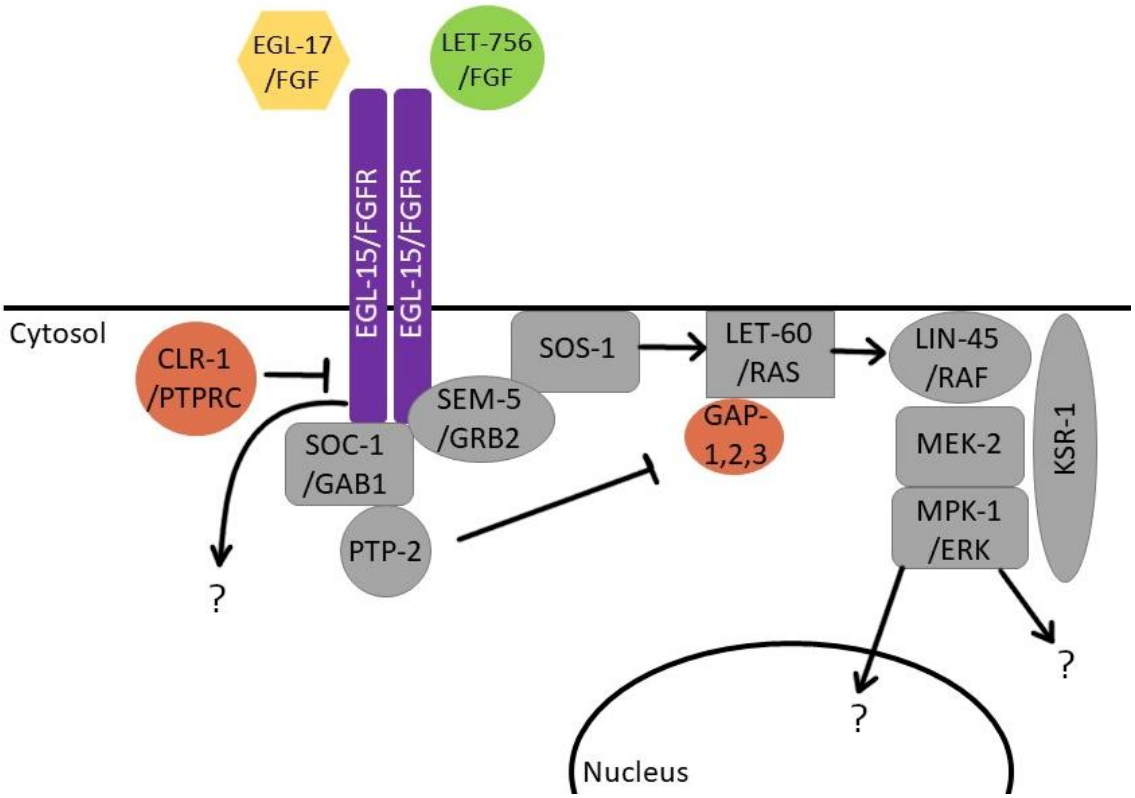


Figure 4-5. Known EGL-15 RTK signalling pathway. EGL-15 can be activated by two different FGF ligands. Direct binding of SEM-5 SH2 to pY of EGL-15 triggers the downstream MAPK signalling cascade, where SEM-5 recruits SOS-1, leading to activation of LET-60/RAS. LET-60 then activates downstream LIN-45/MEK-2/MPL-1 signalling cascade, which is assisted by KSR-1 scaffolding protein. SOC-1 is also recruited to active EGL-15, in turn recruiting PTP-2 to assist in EGL-15 signalling. Negative regulators of EGL-15 signalling are in orange.

The EGL-15(5B) isoform is essential for *C. elegans* development and is expressed in the hypodermis (Lo et al., 2008). EGL-15(5B) is activated by the LET-756 FGF ligand, and loss of either the ligand gene, or the 5B isoform is lethal (Goodman et al., 2003). Due to this phenotype, the exact signalling pathway of the 5B isoform has not been found. It is however, known that SEM-5 is required for this function, but it could be recruited to EGL-15 via a different binding site than pY1009/pY1087 (Lo et al., 2010). The LET-756, EGL-15 and SEM-5 signalling pathway is also required for muscle arm extension during development, which leads to formation of the neuromuscular junction between

body wall muscles and motor axons (Dixon et al., 2006), as well as fluid homeostasis (Sasson and Stern, 2004). CLR-1 receptor tyrosine phosphatase is a negative regulator of EGL-15, which prevents over-activation of the RTK receptor, and in turn overactivation of LET-60, to maintain fluid levels within the pseudocoelomic space, although the exact mechanism of how this is achieved is unknown (Kokel et al., 1998; Schutzman et al., 2001; Huang and Stern, 2004).CED-2

CED-2 is the sole CRK-family orthologue in *Caenorhabditis elegans*. Its SH2-SH3N-SH3C structure is conserved, showing equal similarity to both human CRK-II and CRKL proteins (Figure 4-6). The adaptor protein is required for several processes requiring actin cytoskeleton remodelling. *ced-2* is essential for DTC migration during gonad development, midbody degradation, and engulfment of apoptotic germline cell corpses (Reddien and Horvitz, 2000; Chai et al., 2012).

```
CED-2          MTTNGFDPFEWRSFYFPGMSREEAHKLLGEPQVSIGTFLMRDSS-RPGEYSLTVREADEGNAVCHYLIERGEPKE-----
CRKL           MSSARFDSSDRSAWYMGPVSRQEAQTRLQGQ--RHGMFLVRDSSTCPGDYVLSVSE---NSRVSHYIINSLP--------
CRK-II         MAG-NFDSEERSSWYWGRLSRQEAVALLQGQ--RHGVFLVRDSSTSPGDYVLSVSE---NSRVSHYIINSSGPRPPVPPS
               *:   **. :  ::*   :**:**   *       * **:****  **:* *:* *   .. *.**:*:

CED-2          -----DGTAAAGVKIANQSFPDIPALLNHFKMRVLTEASLLAAYK-----------------KPIIEVVVGTFKFTGER
CRKL           ---------NRRFKIGDQEFDHLPALLEFYKIHYLDTTTLIEPAPRYPSPPMGSVSAPNLPTAEDNLEYVRTLYDFPGND
CRK-II         PAQPPPGVSPSRLRIGDQEFDSLPALLEFYKIHYLDTTTLIEPVSRS-------RQGSGVILRQEEAEYVRALFDFNGND
                           .:*.:*.*  :****:.:*:: *  ::*: .                   :   * *   :.* *:

CED-2          ETDLPFEQGERLEILSKTNQDWWEARNALGTTGLVPANYVQIQMEFHN-DRTS-KGASQSSIGSSGGGAERF---SSAST
CRKL           AEDLPFKKGEILVIIEKPEEQWWSARNKDGRVGMIPVPYVEKLVRSSP------HGKHGNRNSNSYGIPEPAHAYAQPQT
CRK-II         EEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMIPVPYVEKYRPASASVSALIGGNQEGSHPQPLGGPEPG-PYAQPSV
                 ****::*: * * .*.:::**.*.:  *  *::*. **:             *   .   .. * .*     :....

CED-2          SSD--------------NIELQPRLPAKAKVTFDRVPNAYDPTQLRVKKGQTVLVTQKMSNGMYKAELDGQIGSVPHTYL
CRKL           TTPLPAVSGSPGAAITPLPST-QNGPVFAKAIQKRVPCAYDKTALALEVGDIVKVTRMNINGQWEGEVNGRKGLFPFTHV
CRK-II         NTP--------------LPNL-QNGPIYARVIQKRVPNAYDKTALALEVGELVKVTKINVSGQWEGECNGKRGHFPFTHV
               .:                .   . *  *:.  .*** *** * * :: *: * **:   .* ::.* :*: * .*.*::

CED-2          RFTAVSE------
CRKL           KIFDPQNPDENE-
CRK-II         RLLDQQNPDEDFS
               ::   .:
```

Figure 4-6. Alignment of CED-2, CRKL and CRK-II proteins. Overall SH2-SH3-SH3 structure is conserved with good similarity, while linker regions are more variable between proteins.

4.1.3.2 CED-2 in DTC migration

CED-2 is needed in the DTC for correct gonad arm migration (Figure 4-2). Although the exact mechanism has not been found, loss-of-function alleles of *ced-2* show a partially penetrant DTC migration defect (Reddien and Horvitz, 2000). In this phenotype, the gonad arm, after making two turns during phase 2, makes an additional turn and instead of migrating towards the centre of the animal during phase 3, it migrates towards the anterior/posterior end.

4.1.3.3 CED-2 in apoptotic cell engulfment

CED-2 plays an essential role in the apoptotic cell engulfment pathway (Figure 4-3). There are several genes which encourage the CED-2/CED-5/CED-12 complex formation. Two pathways have been shown to directly lead to CED-2 activation. MOM-5/Frizzled becomes active upon recognising an apoptotic cell. The active receptor potentially enables APR-1/APC to interact with CED-2, while also directs GSK-3 signalling towards the CED-10/RAC pathway (Cabello et al., 2010; Neukomm et al., 2011). PAT-3 and INA-1 integrins can recruit CED-2 through SRC-1 and modulate CED-2 activity in cell corpse engulfment (Hsu and Wu, 2010; Neukomm et al., 2011). Downstream, CED-2 interacts with CED-5 (human DOCK1 orthologue), which in turn binds CED-12 (human ELMO orthologue) (Gumienny et al., 2001). The CED-2/CED-5/CED-12 GEF complex then activates CED-10 (human RAC1 orthologue) by allowing GTP loading, to enable engulfment (Reddien and Horvitz, 2000).

4.2 Aims

The aim of this Chapter was to see whether “tier 2” signalling could be investigated in an animal model. In order to do that, the interaction between the *C. elegans* EGL-15 proline-rich motifs and the CED-2 SH3 domains was investigated using biochemical and phenotypic assays. The first goal was to determine which proline-rich motif of EGL-15 is responsible for CED-2 binding. Next, precise proline-rich motif and SH3 domain deletions were generated using CRISPR/Cas9 to investigate the phenotypic consequences of this interaction.

4.3 Materials and methods

4.3.1 *Caenorhabditis elegans* maintenance

N2 (Bristol) strain was used as wild-type. The animals were maintained by growing them on Nematode Growth Media (NGM) plates (3 g/L NaCl (Sigma-Aldrich), 17 g/L agar (Sigma-Aldrich), 2.5 g/L peptone (Sigma-Aldrich), 1 mM $CaCl_2$ (Sigma-Aldrich), 5 mg/L cholesterol (Sigma-Aldrich), 1mM $MgSO_4$ (Sigma-Aldrich), 25 mM KPO_4 (Sigma-Aldrich) buffer) seeded with OP50 *E. coli* strain as food source and kept at 20^{o}C.

All strains used in this project are listed in Appendix A.

4.3.2 Generation of mutant *C. elegans* using CRISPR/Cas9

Mutant animal strains were generated using the CRISPR/Cas9 gene editing system (Arribere et al., 2014). crRNA targeting the gene region of interest was designed using “Custom Alt-R CRISPR-Cas9 guide RNA” online tool from Integrated DNA Technologies (IDT) (targeting sequences can be found in Appendix A). To form a Cas9:crRNA:tracrRNA ribonucleoprotein complex (RPN), equimolar amounts of crRNA (IDT) and tracrRNA (IDT) were mixed and incubated at 95^{o}C for 5 minutes, followed by incubation at room temperature for 5 minutes. Then, 27 µM of Cas9 protein (expressed and purified by Dr Kin Man Suen or purchased from IDT) and 27 µM of crRNA:tracrRNA duplex were mixed and incubated at room temperature for 5 minutes. The final injection mix contained 17.5 µM Cas9 protein, 17 µM crRNA:tracrRNA duplex (against target gene and *dpy-10* gene), 6 µM ssDNA target repair template, 0.5 µM *dpy-10* ssDNA repair template (repair sequences can be found in Appendix A). *dpy-10* was used as a co-injection marker to help finding the jackpot brood (Paix et al., 2015).

4.3.3 *C. elegans* microinjection

The CRISPR/Cas9 injection mix was injected into *C. elegans* adult hermaphrodite gonads. The injection mix was loaded into pulled glass capillary needles. Animals were mounted on a 2% agarose pad which contained a drop of Halocarbon oil 700 (Sigma-Aldrich) on top. Zeiss Axiovert 10 microscope was used for injections. For each strain, around 20 hermaphrodites were injected. The injected animals were recovered via addition of recovery buffer (5 mM HEPES (Fisher) pH 7.2, 3 mM $CaCl_2$ (Sigma-Aldrich), 3 mM $MgCl_2$ (Sigma-Aldrich), 66 mM NaCl (Sigma-Aldrich), 2.4 mM KCl (Sigma-Aldrich), 4% glucose (Sigma-Aldrich)) on top of mounted hermaphrodites immediately after injection. The animals were then allowed to recover for 10 minutes, after which M9 buffer (3 g KH_2PO_4 (Sigma-Aldrich), 6 g Na_2HPO_4 (Sigma-Aldrich), 5 g NaCl (Sigma-Aldrich), 1 ml 1 M $MgSO_4$ (Sigma-Aldrich), distilled H_2O to 1 litre) was added. The nematodes were then left until they were swimming briskly, then more M9 was added to aid recovery. Recovered animals were singled onto seeded NGM plates.

4.3.4 *C. elegans* CRISPR/Cas9 mutant screening

Once the progeny of injected nematodes reached L4 stage, the plate containing the most animals with “dumpy” and “roller” phenotypes was dubbed as the “jackpot brood”. 96 F1 progeny from the “jackpot brood” were singled onto seeded NGM plates and allowed to lay eggs overnight. The following day genomic DNA was extracted from F1 hermaphrodites using proteinase K digestion method.

4.3.5 *C. elegans* genomic DNA (gDNA) extraction

Genomic DNA was extracted using a proteinase K digestion method. The reaction mixture consisted of 7.5 µL nuclease free water (IDT), 2 µL 5xHF buffer (NEB), 0.5 µL Proteinase K (NEB). A single adult hermaphrodite was transferred into each reaction tube. The samples were then frozen at -80^{o}C, followed up by an incubation at 65^{o}C for 30 minutes, and 95^{o}C for 20 minutes. The gDNA was then used for PCR.

4.3.6 *C. elegans* polymerase chain reaction

PCR was used for genotyping CRIPSR/Cas9 mutants. The PCR mix consisted of 8.38 µL nuclease-free water, 2.5 µL 5xHF buffer (NEB), 0.625 µL each PCR primer (IDT), 0.25 µL of 2.5mM dNTP stock (NEB), 0.125 µL Phusion DNA polymerase (NEB), 1 µL of *C. elegans* gDNA. The list of primers and PCR programmes used for mutant screening can be found in Appendix A. The PCR products were run on 2% agarose gels, at 100 V for 30-50 minutes. Each mutant strain was back-crossed with N2 three times.

4.3.7 Genetic crosses

To generate double mutant strains, males were generated from *egl-15* mutant strain (as the gene is in the X chromosome). To increase the frequency of males, a few plates containing 10 L4 stage hermaphrodites each were heat-shocked at 35^{o}C for 3 hours. The plates were then kept in the 20^{o}C incubator and checked daily for appearance of male progeny. For the genetic crosses, 9 *egl-15* mutant males were put onto an NGM plate containing food, along with 3 *ced-2* mutant hermaphrodites (L4 stage) and left to mate for two days, after which the males were removed. 16 F1 progeny were singled onto NGM plates, and once they reached adulthood, they were allowed to lay eggs overnight. The

next day, gDNA from F1 parents was extracted (as described in 4.3.5) and subjected to screening for *egl-15* and *ced-2* mutations (as described in 4.3.6 and Appendix A). 16 F2 progeny from F1 animals heterozygous for mutations in both genes were singled onto NGM plates and allowed to self-fertilize. The F2 mothers were left to lay eggs overnight. The next day F2 animals were subjected to screening for *egl-15* and *ced-2* mutations. This time, progeny from F2 hermaphrodites homozygous for both mutations were retained as a new double mutant strain. Mutant homozygosity was checked in F3 animals to confirm the generation of a new double mutant strain.

4.3.8 *C. elegans* RNA interference

HT115 *E. coli* bacteria (*C. elegans* RNAi collection (Ahringer)) expressing either dsRNA control or dsRNA against genes of interest were inoculated in LB broth (Merck) containing 1 mM ampicillin (Alfa Aesar) and grown overnight. Expression of dsRNA was induced the next morning with 0.5 mM IPTG for 4 hours. Induced bacterial cultures were then seeded onto NGM plates supplemented with 1 mM ampicillin and 0.5 mM IPTG.

For RNAi experiments, adult hermaphrodites were put on RNAi plates and left to lay eggs for a few hours. The mothers were then removed, and experimental animals were fed with RNAi from hatching before being taken forward for staining.

4.3.9 *C. elegans* acridine orange staining and microscopy

A 35 mm NGM plate containing non-starved adult hermaphrodites was stained with 500 µL 0.02 mg/ml acridine orange solution (Thermo Fisher) in M9. The plate was incubated at 20^{o}C in the dark for 2 hours. Stained animals were then transferred onto a fresh NGM plate containing food and incubated at 20^{o}C

for 2 hours to destain. Nematodes were then mounted on a 5% agar pad (on a glass slide) and immobilised in a drop of 1 mM levamisole (Merck). Animals were visualised using Zeiss Axioplan microscope, and images were taken with Retiga 2000R (Q Imaging) camera.

4.3.10 *C. elegans* brood size assay

Gravid *C. elegans* adults were dot-bleached to generate a semi-synchronised population. Once the animals reached L4 stage they were singled onto seeded NGM plates and allowed to lay eggs for two days. The mothers were then transferred to new seeded NGM plates to finish laying eggs. The F1 progeny were left to grow until L4-adulthood stages and counted.

4.3.11 Plasmid DNA preparation

Bacteria were grown in 5 ml of LB broth with appropriate antibiotic overnight in a 37^{o}C shaking incubator. The next day bacteria were spun down at 15,000 RPM for 10 minutes in a tabletop centrifuge. The pellet was retained. QIAprep spin Miniprep kit (Qiagen) was used to extract plasmids according to manufacturer instructions.

4.3.12 EGL-15 pro-3 cloning

An oligonucleotide encompassing the PRNPLP proline-rich motif was ordered from IDT. Complementary primers were used to amplify the gene fragment and to introduce BamHI and NotI restriction sites. The PCR reaction consisted of 0.4 µL 10 mM dNTP, 1 µL 10 µM forward primer, 1 µL 10 µM reverse primer, 200 ng oligo template, 2 µL Phusion DNA polymerase (NEB), nuclease-free water up to 20 µL. The PCR product was run on an agarose gel to confirm successful PCR reaction. The DNA bands were then extracted using QIAquick gel extraction kit (Quiagen) according to the manufacturer's

instructions. The purified PCR product, as well as pGEX-4T-2 plasmid DNA were digested with BamHI and NotI. Reaction consisted of 100 ng oligonucleotide or 300 ng of plasmid, 5 µL Cutsmart buffer (NEB), 1 µL BamHI-HF (NEB), 1 µL NotI-HF (NEB), nuclease-free water up to 50 µL. Digestion reactions were purified by gel-extraction using QIAquick gel extraction kit (Qiagen). Ligation was done using NEB quick ligation kit. The reaction contained: 1:3 molar ratio of vector:insert, 10 µL quick ligase reaction buffer, 1 µL quick ligase, nuclease-free water up to 20 µL. The mix was incubated at room temperature for 5 minutes and transformed into DH5α bacteria (NEB).

4.3.13 *ced-2* cloning (work done by Dr Kin Man Suen)

C. elegans mRNA was a gift from Dr Kin Man Suen. cDNA was reverse-transcribed using First Strand cDNA synthesis protocol (E6300) (NEB) with random primers. *ced-2* was then PCR-amplified using primers listed in Appendix A. *ced-2* was cloned into a pGEX vector using BamHI and NotI sites.

4.3.14 Bacteria transformation

20 µL of competent DH5α bacteria (NEB) were thawed on ice. 5 µL of ligation reaction was added, and bacteria were incubated on ice for 30 minutes. Bacteria were then heat-shocked in 42^{o}C water bath for 45 seconds, and chilled on ice for 2 minutes. 500 µL of room-temperature LB broth was added, and bacteria were incubated in 37^{o}C shaking incubator for at least 1 hour. Bacteria were then spread on a pre-warmed LB agar plate containing 100 µg/mL ampicillin and incubated at 37^{o}C overnight.

4.3.15 *C. elegans* synchronisation

C. elegans were synchronised by using hypochlorite solution. Gravid adults were washed off a plate with 1ml of sterile water. 1.5 ml of hypochlorite

solution (3 ml of 4.99% NaClO (Sigma-Aldrich), 1 ml of 10 M NaOH (Sigma-Aldrich), 6 ml sterile distilled H_2O) was then added, and the animals were left to dissolve for 5 minutes to release eggs. The tubes were spun down at 2000 RPM for 2 minutes in a centrifuge and supernatant was removed. 5 ml of sterile distilled water was added, and tubes were spun down again. This wash was repeated twice. After two washes with sterile water, the eggs were then washed once with M9. Finally, the falcon tubes were filled up to 10 ml with M9 and left to incubate overnight at 20^{o}C with rocking. Next morning, the solutions were checked for hatched larvae.

4.3.16 *C. elegans* lysis

Synchronised hatched larvae were put onto seeded NGM plates. Synchronised adult hermaphrodites were washed off the plate with M9. Tubes were spun at 2000 RPM for 1 minute, M9 was removed and fresh M9 was added to wash. This was repeated several times until M9 was clear. For last wash, cold TBST (20 mM Tris (Sigma-Aldrich), 150 mM NaCl, 0.1% Tween 20 (ChemCruz)) was added instead. The animals were resuspended in as little TBST as possible and frozen by dropping the mix into liquid nitrogen. The frozen balls containing the hermaphrodites were then ground into powder with mortar and pestle. Frozen "*C. elegans* powder" was then thawed on ice, and the solution was sonicated to break down the animals further. The lysate was then spun down at 15,000 RPM for 10 minutes, and supernatant was collected.

4.3.17 Protein concentration measurement

Coomassie Assay reagent (Thermo Scientific) was used to determine protein concentration in lysates, and the absorbance was measured at 595 nm.

4.3.18 Protein expression and purification

C. elegans protein expression and purification was performed as described in section 2.12.

4.3.19 GST- pulldown

Glutathione Sepharose beads were used for GST-pulldowns. Beads were washed by adding five gel bed volumes of lysis buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1mM β-mercaptoethanol (Sigma-Aldrich)) and spun down in a tabletop centrifuge at 4000 RPM for 5 minutes. The wash was repeated two more times, followed by addition of bacterial cell lysate and incubation overnight at 4^{o}C with rotation. The following day the beads were washed five times with the lysis buffer. For buffer exchange, buffer that was used for *C. elegans* lysis (TBST (20 mM Tris (Sigma-Aldrich), 150 mM NaCl, 0.1% Tween 20 (ChemCruz)) was used to wash beads. *C. elegans* lysate containing 1 mg of protein was then added to the beads and incubated at 4^{o}C overnight with rotation. The beads were then washed three times with *C. elegans* lysis buffer. Two gel bed volumes of 2xSDS loading buffer were added to the beads, and the tubes were put in a 95^{o}C heating block for 10 minutes to denature protein. The supernatant was analysed by SDS-PAGE and western blot.

4.3.20 SDS-PAGE

Samples for analysis were mixed with 2x sample buffer (Biorad) and boiled in a 95^{o}C heat block for 10 minutes. The samples were loaded onto 4-10% Bis-Tris precast gels (Invitrogen). NuPage MOPS SDS running buffer (Novex) was used to fill the tank. The gels were run at 150 V for 45 minutes.

4.3.21 Western blotting

Proteins from the SDS-PAGE gel were transferred using wet transfer method onto a PVDF (Biorad) membrane. 1x Transfer buffer (10x: 30.3 g Tris-Base (Sigma-Aldrich), 144 g glycine (Sigma-Aldrich) in 1 L distilled H_2O, pH 8.3; 1x: 100 ml 10x transfer buffer, 200 ml methanol (Sigma-Aldrich), 700 ml distilled H_2O) was used to fill the tank, and proteins were transferred at 100 V for 1 hour. To visualise GST, the membrane was stained with Ponceau S (Sigma) and visualised using G:BOX (Syngene). The stain was washed off with TBST, followed by blocking with 5% BSA (Fisher Scientific) in TBST. The membrane was then incubated with primary antibody (in 5% BSA TBST; 1:1000 ratio) overnight. Next day, the blot was washed three times with TBST for 10 minutes, followed by 1 hour incubation with secondary antibody (in 5% BSA TBST). Finally, the blot was washed three times with TBST for 10 minutes before imaging. For imaging, the protein bands were visualised using HRP-conjugated secondary antibodies, which were detected by Clarity Western ECL (Biorad) using G:BOX (Syngene).

4.3.22 Microscale thermophoresis (work done by Dr Chi-Chuan Lin)

For MST, buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 0.5 mM TCEP was used. GST-CED-2 was expressed in *E. coli* and purified, followed by GST-tag cleavage with thrombin. EGL-15 peptides were purchased from Genscript and labelled with $Atto^{488}$. In each tube, the concentration of labelled protein was constant (100 nM), while the second protein was serially diluted over 16 dilution series. Solution with labelled and unlabelled proteins was transferred to capillaries (Monolith) and inserted into Monolith NT.115

machine for detection. Blue LED was used for MST, with power setting adjusted be within 200-1500 fluorescence counts. Each scan ran for 3 repeats.

4.4 Results

4.4.1 *C. elegans egl-15* mutant generation

After establishing that the human FGFR2 proline-rich motif and CRKL SH3 domains directly interact in Chapter 3, *C. elegans* model organism was used to investigate whether the interaction is evolutionarily conserved. These nematodes have a single FGFR family kinase called EGL-15, as well as a single adaptor protein CED-2, which is orthologous to human CRKL and CRK-II. Mutations for these proteins were introduced in N2 (wild-type Bristol strain) using CRISPR/Cas9. Gene editing using CRISPR/Cas9 in *C. elegans* is an extremely robust technique to generate and screen for gene deletions or small precise mutations with relative ease (Arribere et al., 2014). CRISPR/Cas9 allows gene editing at a native gene level, ensuring that mutant protein is expressed at regular level and in correct tissues.

Protein sequences of the EGL-15 C-terminal tail and CED-2 full-length protein are shown in Figure 4-7. Each of the three proline-rich motifs were deleted in EGL-15, while whole SH3 domains were deleted in CED-2. Mutations were detected by PCR amplification followed by restriction digestion, or by size-difference alone, and sequenced to confirm correct mutation. The methods for mutation screening are summarised in Appendix A. All strains were generated successfully, and nucleotide deletions introduced were accurate (Figure 4-8). All single mutant strains had no visible debilitating phenotypes and normal development timing. The CRISPR/Cas9 generated strains and alleles are summarised in Table 4-1.

A EGL-15 C-terminal tail, type IV

LTMTNETIEGSQEFNDQFFSERSTASGPVSPMESFQKKRKHRPLSA**PVNLPSEP**QHTICDDYESNFSVE**PPNDP**NHLYCN

DNMLKNHIITPETSQLYIHKVLNEPIGNGYVRQDKLARAVSGVANQSLDSALGSPAWPSYDRPSNKASCLDQTHQYYNTT

SKIQYLHFTFDDPDCMTRSRDSAIFEESYHPNYIQSHPLYSKIIIKKNMT**PRNPLP**TKETIV

B CED-2

MTTNGFD**PFEWRSFYFPGMSREEAHKLLGEPQVSIGTFLMRDSSRPGEYSLTVREADEGNAVCHYLIERGEPKEDGTAAAG**

VKIANQSFPDIPALLNHFKMRVLTEASLLAAYKKPIIEVVVG**TFKFTGERETDLPFEQGERLEILSKTNQDWWEARNALGT**

TGLVPANYVQIQMEFHNDRTSKGASQSSIGSSGGGAERFSSASTSSDNIELQ**PRLPAKAKVTFDRVPNAYDPTQLRVKKGQ**

TVLVTQKMSNGMYKAELDGQIGSVPHTYLRFTAVSE

Figure 4-7. Protein sequences. A) The sequence of EGL-15 type IV C-terminal tail (present in isoforms H and P as per WormBase). The proline-rich regions which were deleted are highlighted in bold. **B)** CED-2 full length sequence. The SH2 domain is highlighted in blue, the SH3N domain is highlighted in green, the SH3C domain is highlighted in orange. The SH3N and SH3C domains were deleted to generate CED-2 strains used for further experiments.

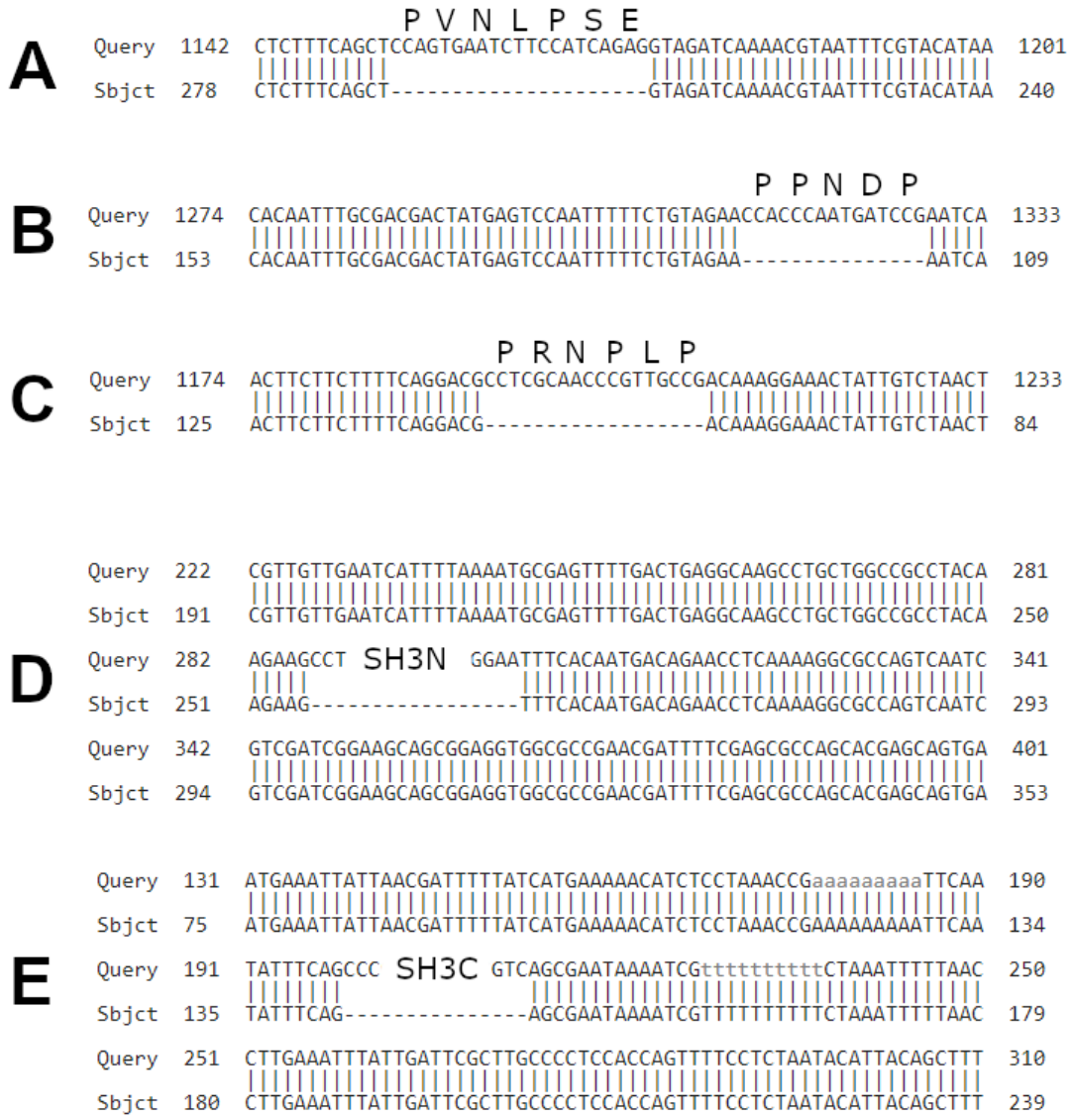


Figure 4-8. Sequencing results of generated mutants. All mutants were generated using CRISPR/Cas9 combined with homology-directed repair. **A)** EGL-15 PVNLPSE motif deletion to generate *egl-15(jez10)* allele, **B)** EGL-15 PPNDP motif deletion to generate *egl-15(jez8)* allele, **C)** EGL-15 PRNPLP motif deletion to generate *egl-15(jez20)* allele, **D)** CED-2 SH3N domain deletion to generate *ced-2(jez45)* allele, **E)** CED-2 SH3C domain deletion to generate *ced-2(jez50)* allele.

Table 4-1. CRISPR/Cas9 generated strains bearing egl-15 and ced-2 mutations.

Strain name	Genotype	Mutation
LAD28	*egl-15(jez10)*	EGL-15(ΔPVNLPSE)
LAD20	*egl-15(jez8)*	EGL-15(ΔPPNDP)
LAD23	*egl-15(jez20)*	EGL-15(ΔPRNPLP)
LAD24	*ced-2(jez45)*	CED-2(ΔSH3N)
LAD25	*ced-2(jez50)*	CED-2(ΔSH3C)
LAD26	*ced-2(jez45) IV; egl-15(jez20) X*	CED-2(ΔSH3N); EGL-15(ΔPRNPLP)
LAD27	*ced-2(jez50) IV; egl-15(jez20) X*	CED-2(ΔSH3C); EGL-15(ΔPRNPLP)

4.4.2 Direct EGL-15 and CED-2 interaction

The goal of this Chapter was to establish whether the C-terminal tail of EGL-15 interacts with CED-2, as well as to determine the potential role of this interaction in *C. elegans*. The EGL-15 C-terminus is subject to alternative splicing, producing a variety of different tail variants (WormBase, release WS280). There are three different proline-rich motifs that can be present within the tail. To determine which EGL-15 proline-rich motif binds CED-2, an MST assay was employed (work done by Dr Chi-Chuan Lin). Full-length GST-tagged CED-2 protein (cloned by Dr Kin Man Suen) was expressed in *E. coli*, purified, and the tag was cleaved off using thrombin. Atto^{488}-tagged EGL-15 peptides, each containing a single proline-rich motif, were purchased from GenScript. Peptide sequences can be found in Appendix A. Binding of full-length CED-2 protein to each of the proline-rich motif was measured using MST (Figure 4-9). PVNLPSEP and PPNDP proline-rich motif-containing peptides were unable to bind CED-2. Only the peptide containing the PRNPLP proline-rich motif of EGL-15 can bind CED-2 (K_d = 3.5 ± 0.8 µM).

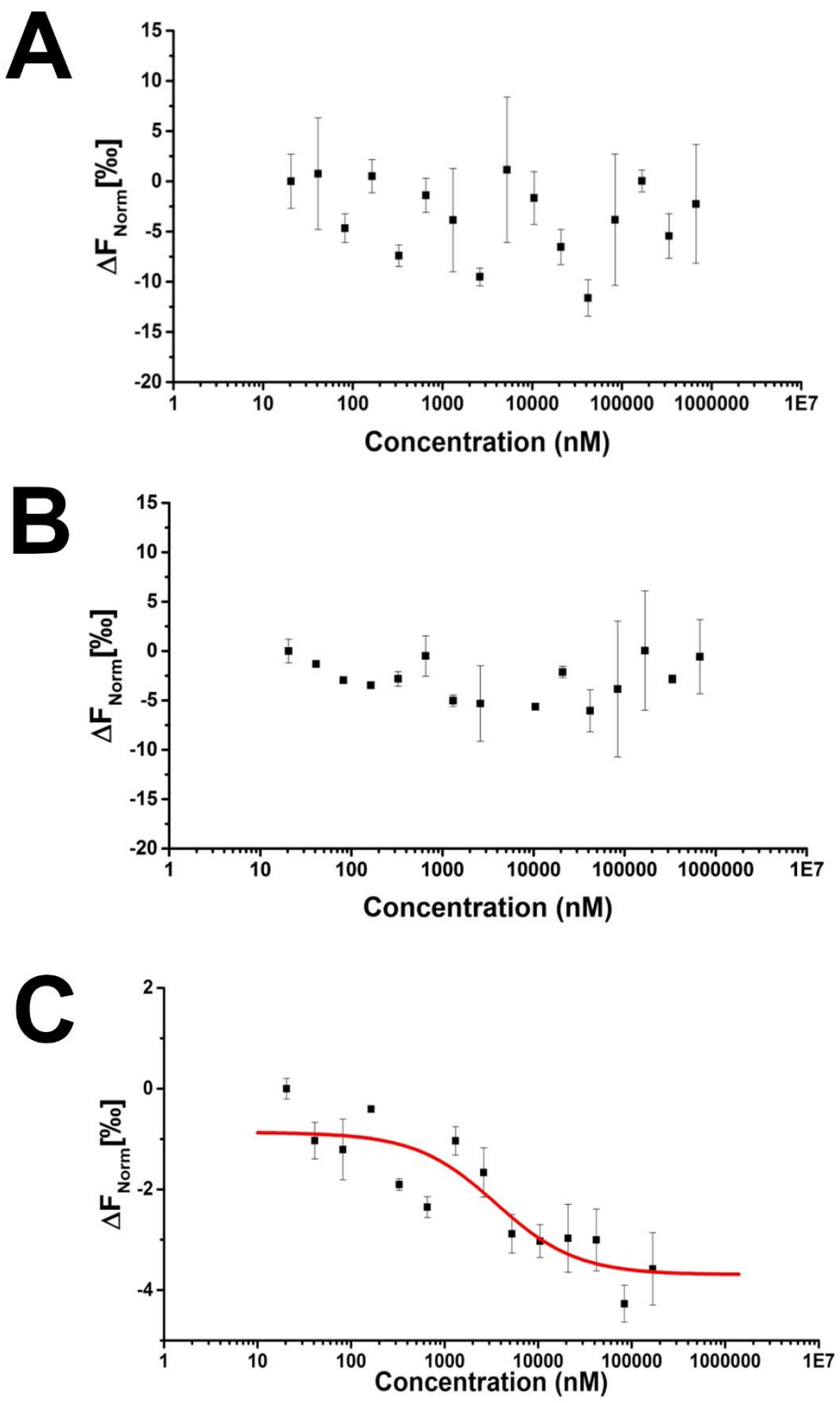


Figure 4-9. CED-2 binds EGL-15 PRNPLP proline-rich motif. Binding between CED-2 and EGL-15 proline-rich motif containing peptides was measured using MST. CED-2 is unable to bind EGL-15 PVNLPSEP **(A)** or EGL-15 PPNDP **(B)** containing peptides. **C**) CED-2 binds the EGL-15 PRNPLP containing peptide, K_d = 3.5 ± 0.8 µM.

To further establish an interaction between CED-2 and EGL-15, a pulldown assay was performed. A peptide containing EGL-15 PRNPLP proline-rich motif (GST-Pro3) was cloned, expressed in *E. coli* and purified (the peptide sequence can be found in Appendix A). The peptide was then immobilised onto Glutathione Sepharose beads, followed by an incubation with whole animal lysates. The pulldown was then analysed using western blotting with anti-CED-2 antibody (Figure 4-10; work for this figure was done by Jasmine Bensilum under my supervision). The polyclonal antibody produces multiple bands on a blot, indicating that it can potentially recognise multiple proteins, thus *ced-2(e1752)* strain was used as a negative control. The *e1752* allele contains a premature stop codon within the SH3N domain of CED-2 and does not produce a full-length protein. Unfortunately, the GST control pulls down CED-2 from whole animal lysate non-specifically. Due to this, even if there is a band representing CED-2 in the GST-Pro3 lane, it cannot be stated that CED-2 can bind EGL-15 peptide under these experimental conditions.

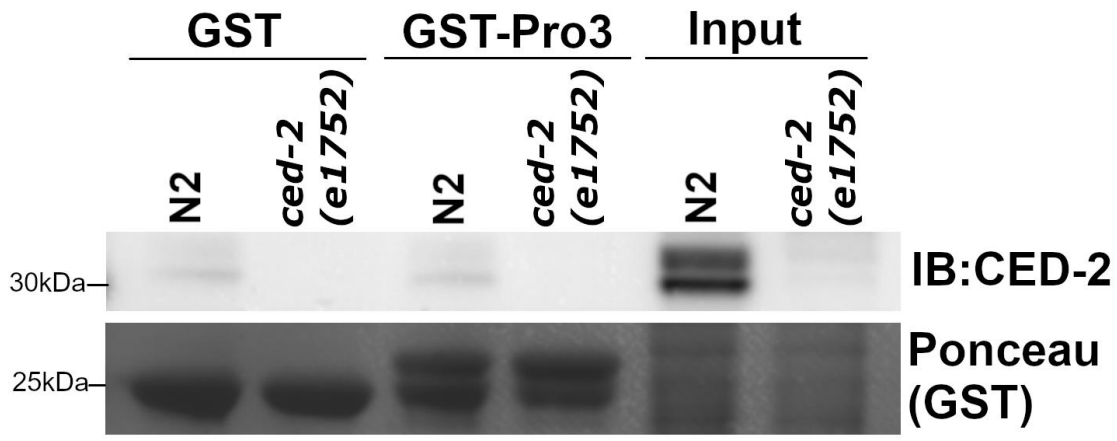


Figure 4-10. GST pulldown to show EGL-15 Pro-3 peptide (containing PRNPLP motif) and CED-2 interaction. Anti-CED-2 antibodies were used to probe for CED-2, while Ponceau stain was used to visualise GST-tagged peptides. Unfortunately, CED-2 is pulled down by GST control, rendering this experiment inconclusive.

4.4.3 EGL-15 PRNPLP motif and CED-2 SH3 domains have roles in engulfment of apoptotic germline cells

The PRNPLP proline-rich motif was identified as the binding site of CED-2 by MST (Figure 4-9), therefore it was chosen to be investigated further. *C. elegans* strain LAD23, which contains the *egl-15(jez20)* allele, was used for the following experiments in this Chapter (EGL-15(ΔPRNPLP) in figures). Strains harbouring *ced-2(n1994)* and *ced-2(e1752)* alleles were obtained from CGC to use as negative controls, as both alleles are *ced*-2 loss-of-function mutants lacking intact SH3 domains.

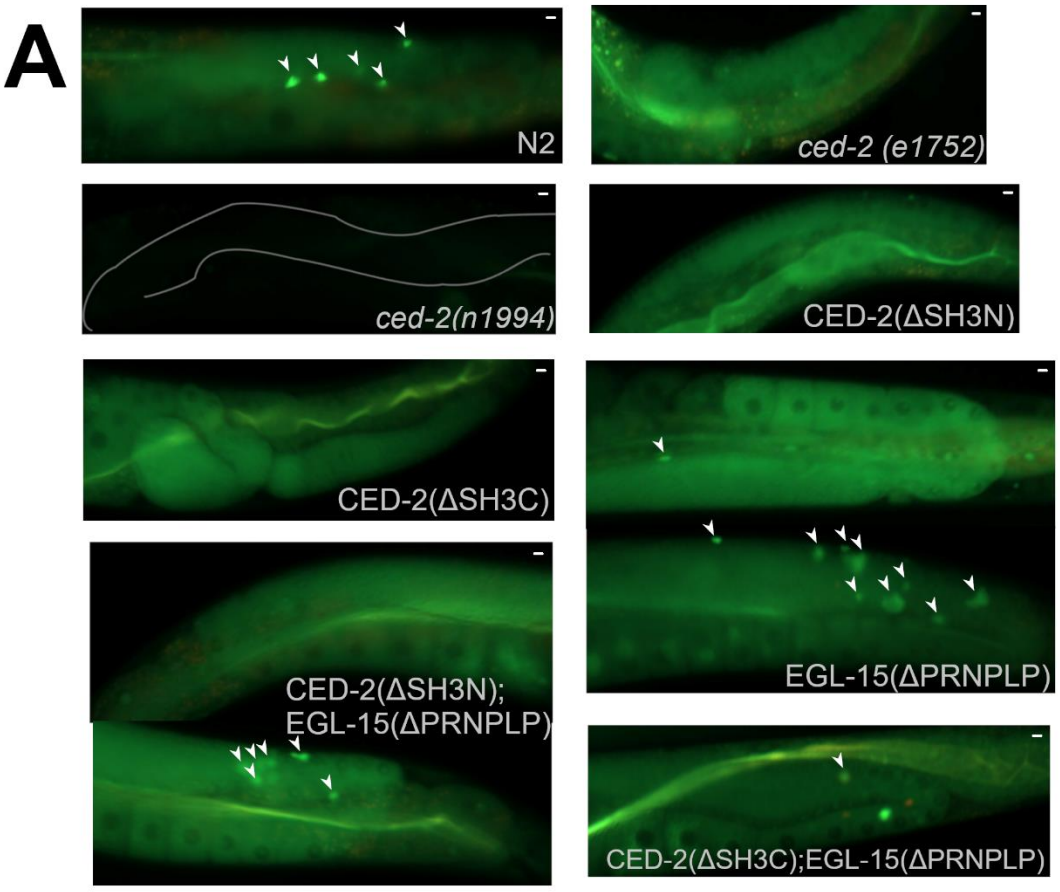
EGL-15 has established roles in fluid homeostasis and vulva development. The *egl-15* impairment or loss-of-function phenotypes are obvious and debilitating (e.g., lethality and “bag of worms”; section 4.1.3). However, EGL-15(ΔPRNPLP) animals did not show any of these “classical” *egl-15* mutant phenotypes. To investigate the outcome of the interaction between EGL-15 and CED-2, the approach was first to look at CED-2 related phenotypes in mutant nematodes. CED-2 is essential for engulfment of cells undergoing physiological apoptosis (Figure 4-3). In adult animals, CED-2 present in gonadal sheath cells is required for signalling pathways that enable those cells to engulf apoptotic germline cells (Hedgecock et al., 1983). The protein is also needed for correct DTC migration during *C. elegans* gonad development (Reddien and Horvitz, 2000). Thus, germline cell engulfment and gonad morphology were evaluated in *C. elegans*. Acridine orange (AO) dye was used to visualise engulfed germline cell corpses (Figure 4-11). In apoptotic cells, AO binds double stranded DNA and produces green fluorescence (excitation at 502 nm, emission at 525nm), and it preferentially stains cells which are being engulfed, thus it was used as an engulfment-specific dye. This can be seen as bright green spots in the death

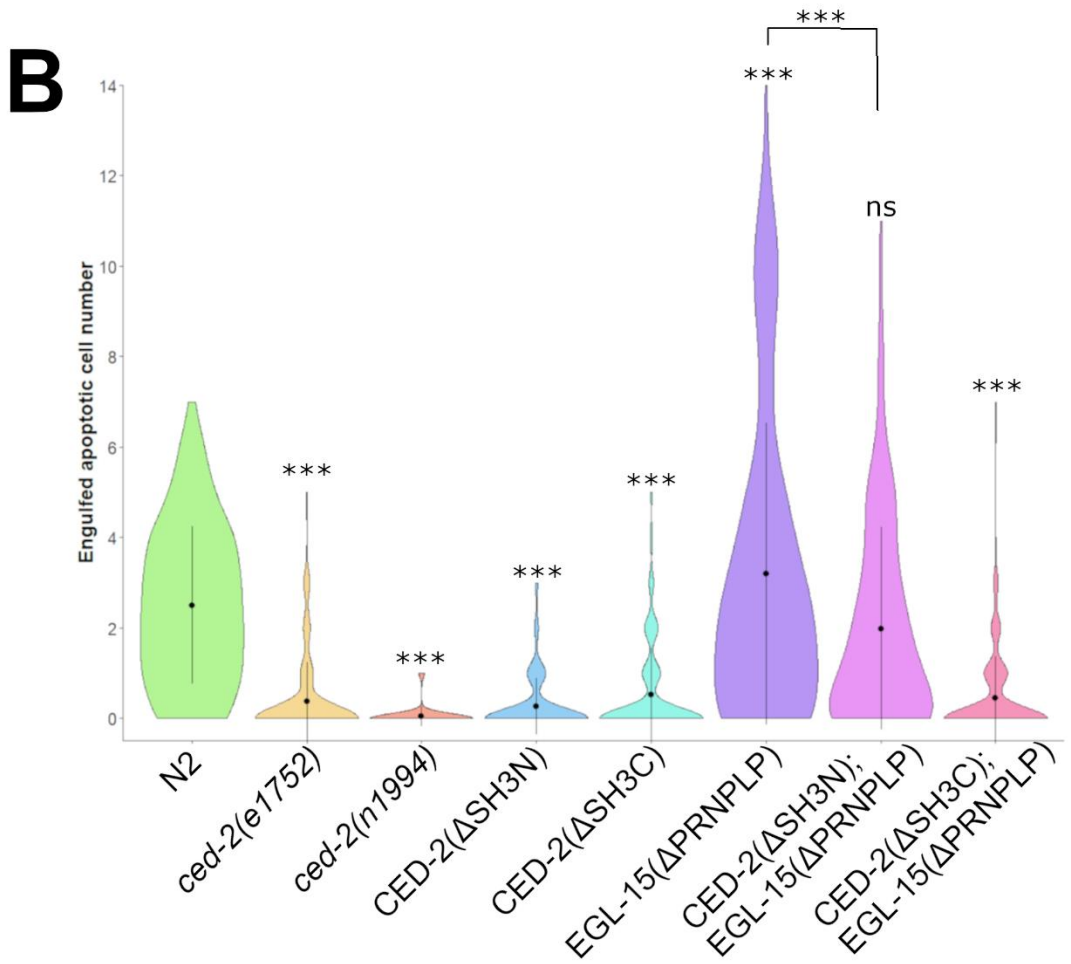
zone near the gonad bend, as observed for wild-type N2 nematodes (Figure 4-11A). Representative gonad images are displayed in Figure 4-11A. c*ed-2(e1752)* and CED-2(ΔSH3C) nematodes also display a DTC pathfinding defect, which will be discussed later (section 4.4.4). The distribution of values is shown on a violin plot (Figure 4-11B). Wild-type (N2) animals establish a normal range, where there are between 0-7 engulfed cells. As expected, two *ced-2* loss-of-function alleles (*ced-2(e1752)* and *ced-2(n1994)*) show a severe reduction in engulfed germ cell numbers that is statistically significant compared to wild-type. CED-2(ΔSH3N) and CED-2(ΔSH3C) single mutants also show severe reduction in engulfment, establishing these alleles as *ced-2* loss-of-function. The EGL-15(ΔPRNPLP) single mutant seems to show an unusual engulfment phenotype. While 86.18% of the animals show an engulfed corpse number that falls within the wild-type range, there are several nematodes in the population showing an increase in engulfment numbers. 21 out of 152 scored EGL-15(ΔPRNPLP) animals had a high (>7) number of engulfed corpses. The EGL-15(ΔPRNPLP) single mutant and CED-2(ΔSH2N);EGL-15(ΔPRNPLP) double mutant animals have two representative images, with the top one showing no engulfment and the bottom one displaying very high engulfed corpse numbers. To see whether loss of EGL-15 PRNPLP motif had a significant effect on engulfment, an F-test was used. This statistical test determines whether two populations show the same variance (null hypothesis). Indeed, there was significant difference in variance of engulfed apoptotic cell number between the wild-type and EGL-15(ΔPRNPLP) animals (Figure 4-11B). This suggests that the EGL-15 PRNPLP motif could be required for normal, tightly controlled cell engulfment in the gonad. CED-2(ΔSH3N);EGL-15(ΔPRNPLP) double mutant shows a loss-of-function rescue phenotype. The

F-test determined that the variance in engulfed cell numbers was not significantly different from wild-type ($p = 0.09$), but it was significantly different ($p = 1.47 * 10^{-7}$) compared to EGL-15(ΔPRNPLP) animals (Figure 4-11B). However, the CED-2(ΔSH3C);EGL-15(ΔPRNPLP) double mutant still appears to show a severe impairment in engulfment, which is almost identical to CED-2(ΔSH3C) single mutant (Figure 4-11B). The results show that in CED-2, both SH3N and SH3C domains are required for functional engulfment (Figure 4-11). It was also established that the PRNPLP proline-rich motif of EGL-15 is required for maintaining correct engulfment function (Figure 4-11B). Furthermore, there is a genetic interaction between the SH3N domain of CED-2 and the PRNPLP motif of EGL-15, as combined loss of these protein regions restored apoptotic cell engulfment to wild-type range (Figure 4-11).

It is important to note that the width of the violin plot is not a proportional representation of how many animals had a certain number of engulfed cell corpses, as the results for engulfment loss-of-function are heavily skewed towards zero. To separately evaluate the number of hermaphrodites with fully non-functional engulfment function, a bar chart was produced, which shows the percentage of the animals showing no cell engulfment (Figure 4-11C). In wild-type (N2), 12.08% of the animals show no engulfment. CED-2 loss of function alleles *ced-2(e1752)* and *ced-2(n1994)* show a severe dysfunction of engulfment, with 79.61% and 95.00% of nematodes showing no engulfment respectively. 80.41% CED-2(ΔSH3N) hermaphrodites show no engulfment, while in CED-2(ΔSH3C) strain engulfment is impaired in 71.43% of animals. EGL-15(ΔPRNPLP) strain shows no engulfment in 25.00% of scored animals. In double mutant strains, 34.57% of CED-2(ΔSH3N);EGL-15(ΔPRNPLP) animals showed engulfment dysfunction, while the number in CED-

2(ΔSH3C);EGL-15(ΔPRNPLP) is 72.67%. The results show that both SH3 domains of CED-2 are required for engulfment to take place. EGL-15(ΔPRNPLP) animals do not show a large increase in the percentage of nematodes showing impaired engulfment compared to N2, suggesting that this strain is similar to wild-type when it comes to the “no engulfment” phenotype. CED-2(ΔSH3N);EGL-15(ΔPRNPLP) double mutant showed a slight rescue of engulfment compared to *ced-2* loss of function strains, indicating a potential genetic interaction between these two genes. However, removing both the CED-2 SH3C domain and the EGL-15 PRNPLP motif did not rescue engulfment impairment, indicating that there is no genetic interaction between these alleles.


A
N2
ced-2 (e1752)
ced-2(n1994)
CED-2(ΔSH3N)
CED-2(ΔSH3C)
EGL-15(ΔPRNPLP)
CED-2(ΔSH3N);
EGL-15(ΔPRNPLP)
CED-2(ΔSH3C);EGL-15(ΔPRNPLP)


B
Engulfed apoptotic cell number
0
2
4
6
8
10
12
14
***
***
***
***
***
***
ns
***
N2
ced-2(e1752)
ced-2(n1994)
CED-2(ΔSH3N)
CED-2(ΔSH3C)
EGL-15(ΔPRNPLP)
CED-2(ΔSH3N);
EGL-15(ΔPRNPLP)
CED-2(ΔSH3C);
EGL-15(ΔPRNPLP)

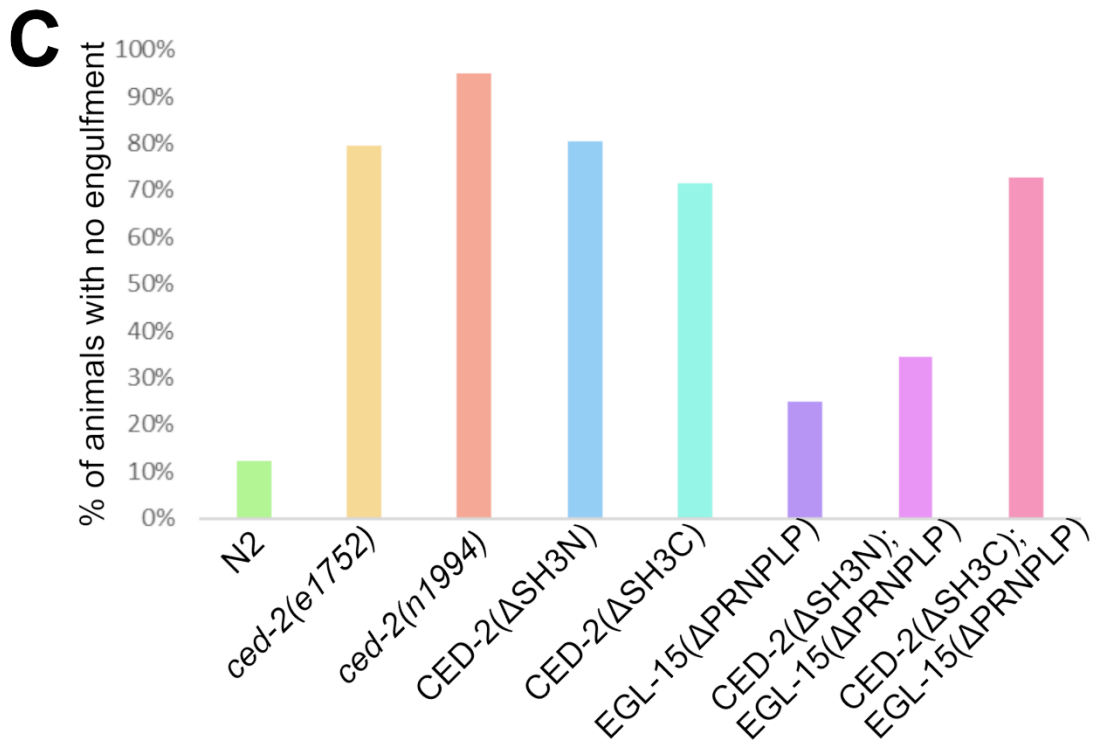


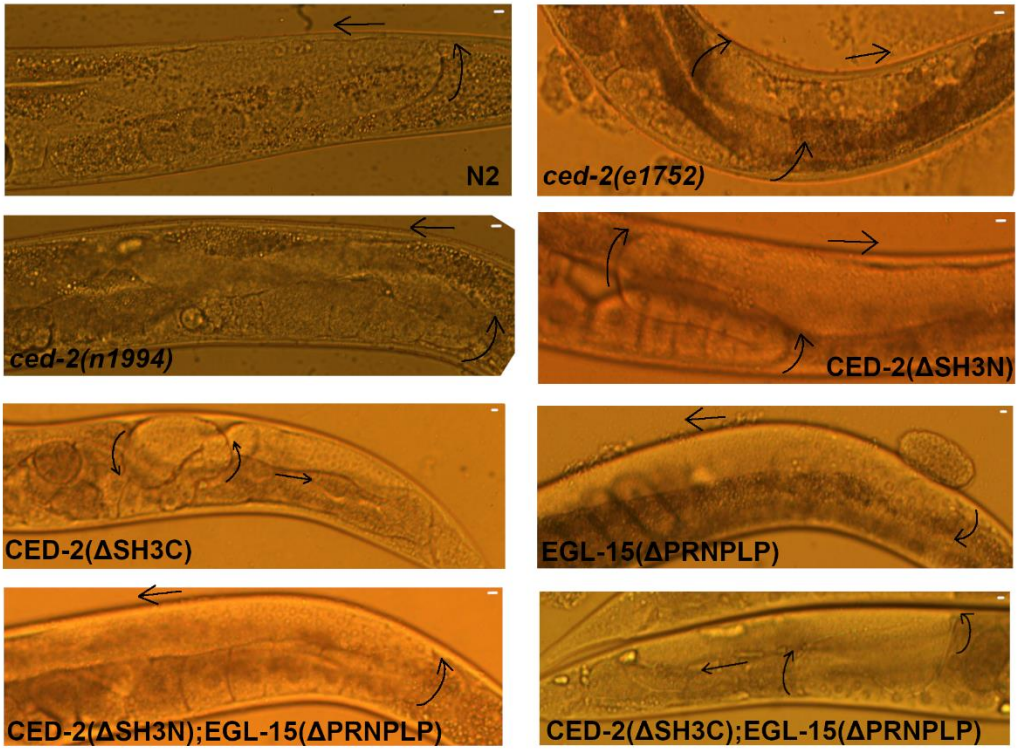
Figure 4-11. CED-2 SH3N domain and EGL-15 PRNPLP motif are required for CED-2 loss of function phenotype. (A) Representative images of AO-stained *C. elegans* gonad arms. EGL-15(ΔPRNPLP) and CED-2(ΔSH3N);EGL-15(ΔPRNPLP) strains have two examples – one for low engulfment (top), and one for high engulfment (bottom). Scale bar – 5 µm. **(B)** Violin plots representing distribution (with mean and standard deviation) of AO staining results. 150 animals were scored for each strain (except 20 for *ced-2(e1994)* and 97 for CED-2(ΔSH3N)) over at least 3 different experiments. F-test was used as a statistical test. Ns – not significant, *** - $p \leq 0.001$ **(C)** Bar chart showing the percentage of animals in each strain which had no apoptotic germline cell engulfment**.**

4.4.4 Aberrant gonad morphology of *ced-2* mutants

While recording the numbers of engulfed apoptotic germline corpses, gonad morphology was also noted (Figure 4-12). As expected, all gonads in wild-type animals were normal. In *ced-2(e1752)* mutants, an extra gonad arm turn occurs at a frequency of 0.66%. *ced-2(n1994)* mutants were not scored due to their scrawny morphology, where the gonad arms were difficult to visualise. 4.12% of CED-2(ΔSH3N) mutants and 14.29% of CED-2(ΔSH3C)

animals showed an extra gonad turn. Gonad morphology was normal in EGL-15(ΔPRNPLP) nematodes. Interestingly, CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) double mutants also showed normal gonad morphology. While CED-2(ΔSH3C)*;*EGL-15(ΔPRNPLP) double mutants had a low frequency of the gonad morphology defect phenotype (5.33%; Figure 4-12). The results suggest that CED-2 SH3N and SH3C domains are involved in correct DTC pathfinding. Furthermore, CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) double mutants showed a rescue phenotype, suggesting a genetic interaction between *ced-2* and *egl-15.*

A



B

Strain	Total animals	Morphology defect	% with morphology defect
N2	149	0	0.00%
ced-2(e1752)	152	1	0.66%
ced-2(n1994)	20	N/A	N/A
CED-2(ΔSH3N)	97	4	4.12%
CED-2(ΔSH3C)	189	27	14.29%
EGL-15(ΔPRNPLP)	152	0	0.00%
CED-2(ΔSH3N); EGL-15(ΔPRNPLP)	188	0	0.00%
CED-2(ΔSH3C); EGL-15(ΔPRNPLP)	150	8	5.33%

Figure 4-12. Effect of *ced-2* and *egl-15* mutations on normal gonad morphology. (A) Nomarski images showing the gonad morphology in wild-type (N2), *ced-2* and *egl-15* mutant strains. Gonad turns are indicated by curved arrows. Scale bar – 5 µm. **(B)** Quantification of gonad morphology defect phenotype. The defect observed was an extra gonad arm turn. The phenotype was not scored in *ced-2(n1994)* mutants due to difficulties visualising the gonad. The animals were scored over at least 3 different experiments.

4.4.5 Brood size in mutant animals

Physiological apoptosis of germ cells is thought to act as quality control, as well as provide additional cytoplasm and organelles to the developing

oocytes (Huelgas Morales and Greenstein, 2018). However, no clear connection between dysfunction of apoptotic germline cell engulfment and brood size has been established. In order to investigate whether perturbation of engulfment can affect brood size, progeny number was counted (Figure 4-13). Wild-type (N2) nematodes showed an average brood size of 285.0 ± 55.5. Loss of function *ced-2(e1752)* animals show a significantly reduced brood size (162.9 ± 65.5). CED-2(ΔSH3N) mutants had an average brood size of 227.7 ± 55.8, which is significantly different from wild-type, while CED-2(ΔSH3C) mutants showed an average brood size of 254.7 ± 40.7, which was not significantly different from wild-type animals. EGL-15(ΔPRNPLP) mutants had an average brood size of 276.3 ± 36.9. CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) animals showed an average brood size of 280.7 ± 29.6, while CED-2(ΔSH3C)*;*EGL-15(ΔPRNPLP) double mutants had an average brood size of 252.7 ± 54.2; neither of which is significantly different from wild-type average brood. The results show that functional CED-2 is needed for maintenance of brood size. The SH3N domain appears to be required for this function. Furthermore, loss of the EGL-15 PRNPLP sequence in the CED-2(ΔSH3N) mutant background restores normal brood size numbers, suggesting that *egl-15* and *ced-2* genetically interact in brood size maintenance.

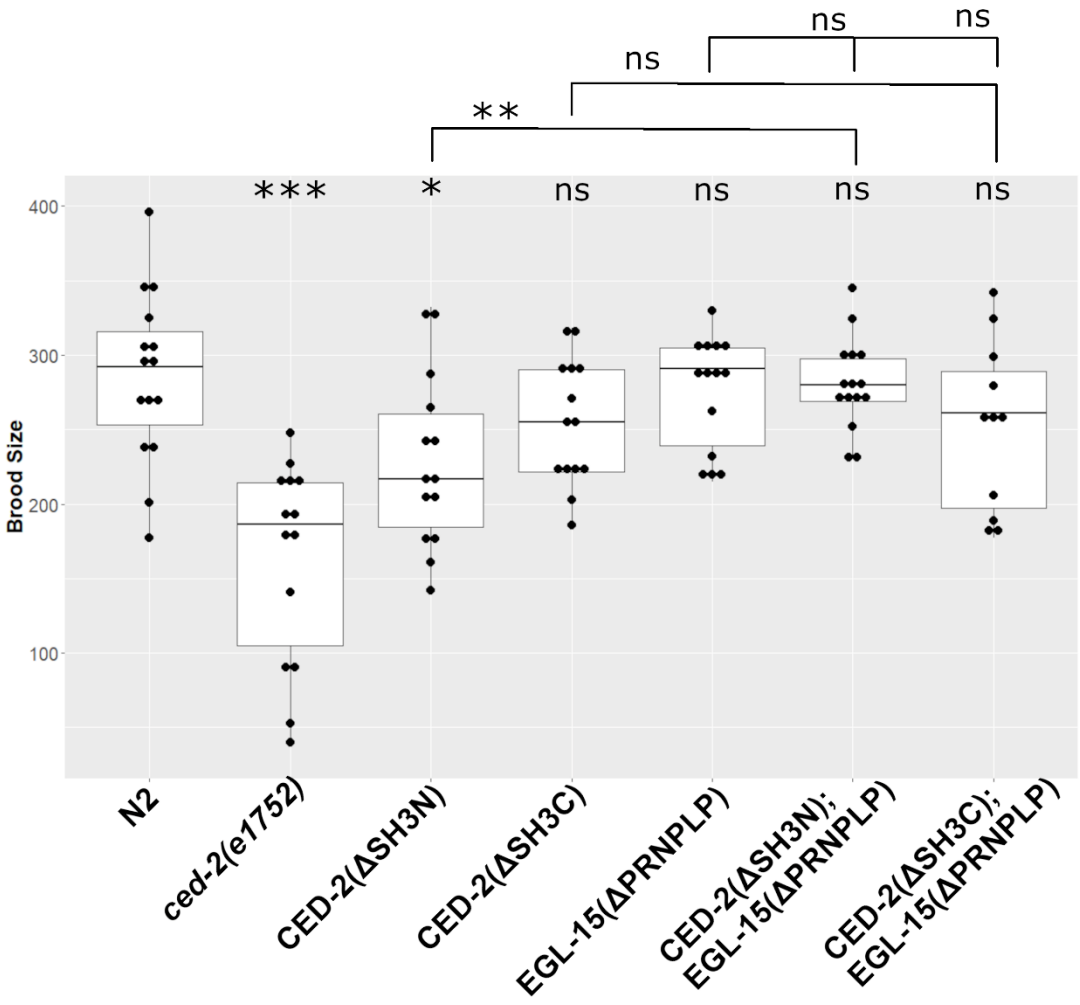


Figure 4-13. Quantification of C. elegans brood size. N = 15 for each strain (including outliers). Ns – not significant, * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001.

4.4.6 RNAi screen

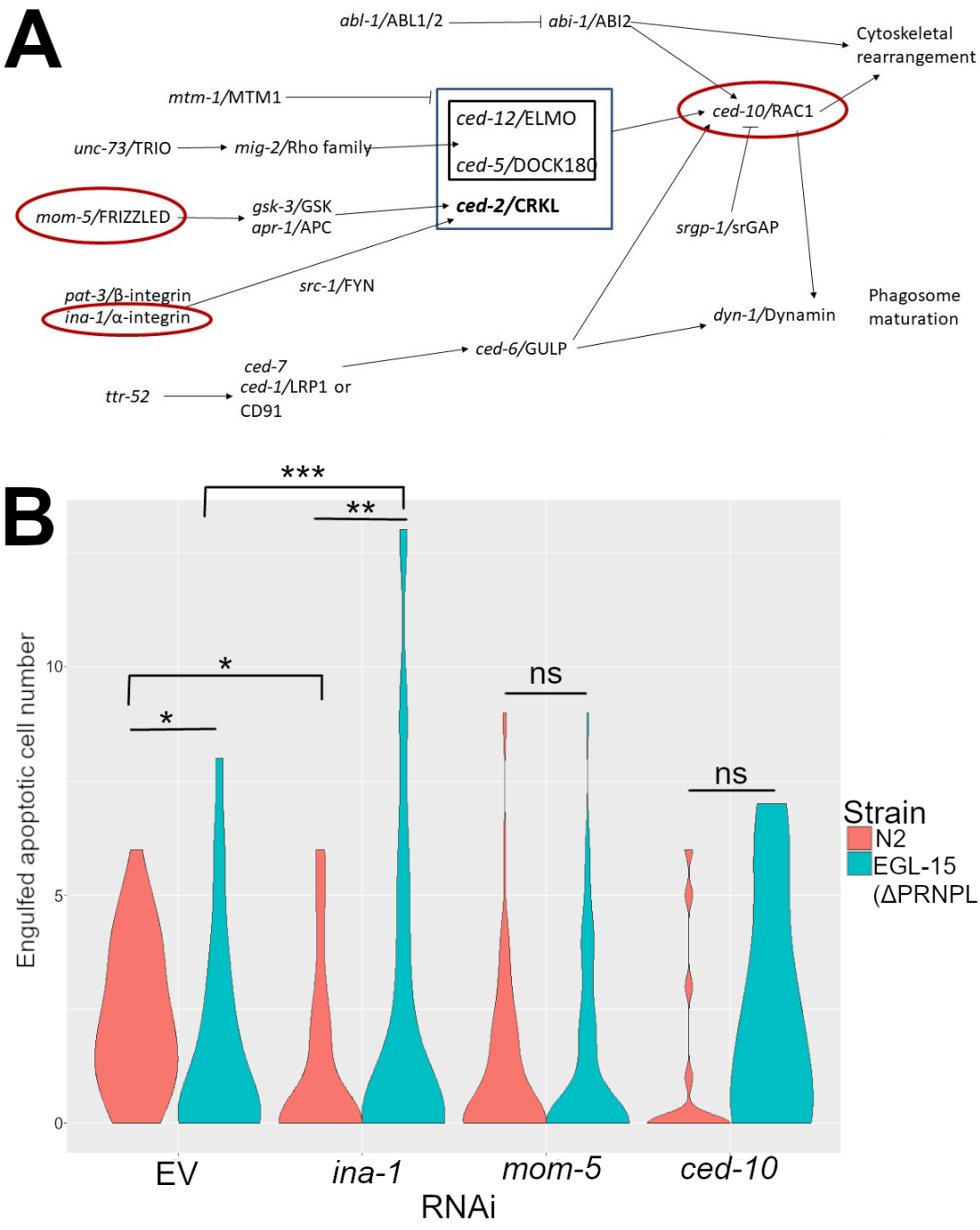
As AO staining and brood size results hinted at a potential genetic interaction between *ced-2* and *egl-15*, an RNAi experiment was conducted to see whether *egl-15* is involved in the engulfment signalling pathway (Figure 4-14). *C. elegans* were fed with RNAi bacteria from L1 stage onwards and stained with AO at day 1 of adulthood. The N2 strain was used as a wild-type control, with EGL-15(ΔPRNPLP) as the experimental strain. RNAi against several genes in the apoptotic cell engulfment pathway was used. *mom-5* and *ina-1* are two different “inputs” upstream of *ced-2*, while *ced-10* is the central

effector essential for apoptotic cell engulfment by neighbouring cells (Figure 4-14A).

The distribution of engulfed corpse number is shown as a violin plot in Figure 4-14B. In the wild-type control animals fed with bacteria expressing the empty RNAi vector (EV), the number of engulfed corpses ranged between 0 and 6, similar to previous AO staining results, while EGL-15(ΔPRNPLP) nematodes fed with the same bacteria showed a range of 0-8 engulfed cell corpses. The variance between these phenotypes was significant. Knocking down *ina-1* using RNAi in N2 did not affect the engulfment range (0-6), while in EGL-15(ΔPRNPLP) animals grown on *ina-1* RNAi bacteria the range was 0-13. The variance between these populations under *ina-1* RNAi was also significant. RNAi-mediated reduction of *mom-5* expression in both strains resulted in similar engulfment range: 0-9 in both N2 and EGL-15(ΔPRNPLP) animals, and the variance shows no statistical significance between the strains. In wild-type nematodes fed with bacteria expressing *ced-10* RNAi, the engulfment range varied between 0 and 6. In EGL-15(ΔPRNPLP) animals fed with the same bacteria, the engulfment range was 0-7, with no statistically significant variance.

To separately evaluate the effect of RNAi-mediated gene knockdown on engulfment function, a bar chart was produced which shows the percentage of hermaphrodites with non-functional engulfment (Figure 4-14C). 13.04% of control animals fed with EV RNAi bacteria showed no engulfment, while in EGL-15(ΔPRNPLP) mutants fed with EV RNAi, 46.00% of animals showed no engulfment. When *ina-1* expression was knocked down using RNAi, 62.50% and 51.35% of N2 and EGL-15(ΔPRNPLP) animals (respectively) showed no engulfment. RNAi-mediated knockdown of *mom-5* resulted in 54.43% of N2 nematodes without engulfment of apoptotic germline cells, while in EGL-

15(ΔPRNPLP) 62.50% of animals lacked engulfment. 73.33% of N2 animals fed with *ced-10* RNAi bacteria showed no engulfment, while in EGL-15(ΔPRNPLP) it was 42.86% of nematodes. Please note that the screen could not be fully completed due to time constraints – while the previous AO staining up to 150 animals were screened per strain, in the RNAi screen, there were up to 50 animals scored per strain. Overall, the results suggest that *egl-15* could have roles in the engulfment pathway, with *egl-15* and *mom-5* potentially being in the same pathway due to lack of additive effects, but the experimental method requires refinement to obtain a better answer (see section 4.5.8 for discussion).


A
abl-1/ABL1/2
abi-1/ABI2
Cytoskeletal rearrangement
mtm-1/MTM1
ced-12/ELMO
ced-5/DOCK180
ced-10/RAC1
unc-73/TRIO
mig-2/Rho family
mom-5/FRIZZLED
gsk-3/GSK
apr-1/APC
ced-2/CRKL
srgp-1/srGAP
pat-3/β-integrin
ina-1/α-integrin
src-1/FYN
dyn-1/Dynamin
Phagosome maturation
ced-6/GULP
ced-7
ttr-52
ced-1/LRP1 or CD91
B
***
**
*
*
ns
ns
Engulfed apoptotic cell number
10
5
0
EV
ina-1
mom-5
ced-10
RNAi
Strain
N2
EGL-15 (ΔPRNPLP)

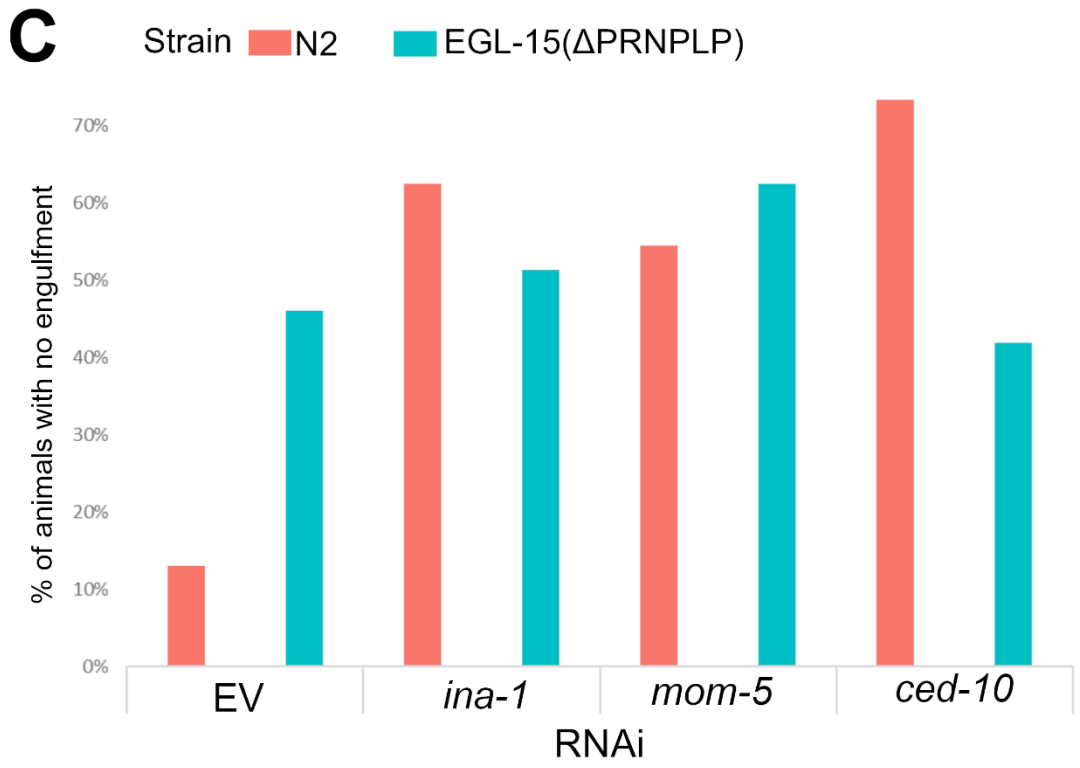


Figure 4-14. EGL-15 has a potential role in canonical apoptotic cell engulfment pathway. A) A part of the apoptotic germline cell engulfment pathway. CED-2/CED-5/CED-12 complex formation is mediated by multiple upstream activators, leading to activation of CED-10 and eventual cell engulfment. **B)** AO staining results from wild-type and EGL-15(ΔPRNPLP) animals treated with RNAi. The nematodes were fed with RNAi bacteria from hatching and AO-stained at young adulthood. Ns – not significant, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. **C)** Bar chart representing the percentage of animals in which engulfment was non-functional.

4.4.7 CED-2 dimerisation

CRKL, a human orthologue of CED-2, can dimerise, and in the previous Chapter it was shown that dimerisation might play a part in enhanced receptor activation (Figure 3-9). CED-2 dimerisation has not yet been investigated in published work. To see whether CED-2 can dimerise, GST-tagged recombinant CED-2 protein was expressed in *E. coli* and purified. The GST tag was cleaved off using thrombin. Some protein was fluorescently labelled using Atto^{488}. 100 nM of Atto^{488}-labelled protein, as well as varying concentrations of unlabelled CED-2 were then used to measure dimerisation dissociation constant by MST

($K_{d\ dimer}$ = 11.3 ± 0.4 µM)(Figure 4-15). The results confirm that CED-2 can dimerise under MST conditions, although with much lower $K_{d\ dimer}$ than CRKL (Figure 3-8, K_d = 0.093 µM ± 0.008 µM). Furthermore, $K_{d\ dimer}$ for CED-2 is larger than that for EGL-15 proline-rich motif binding (K_d = 3.5 ± 0.8 µM, Figure 3-4), suggesting that CED-2 is unlikely to bind EGL-15 as a dimer.

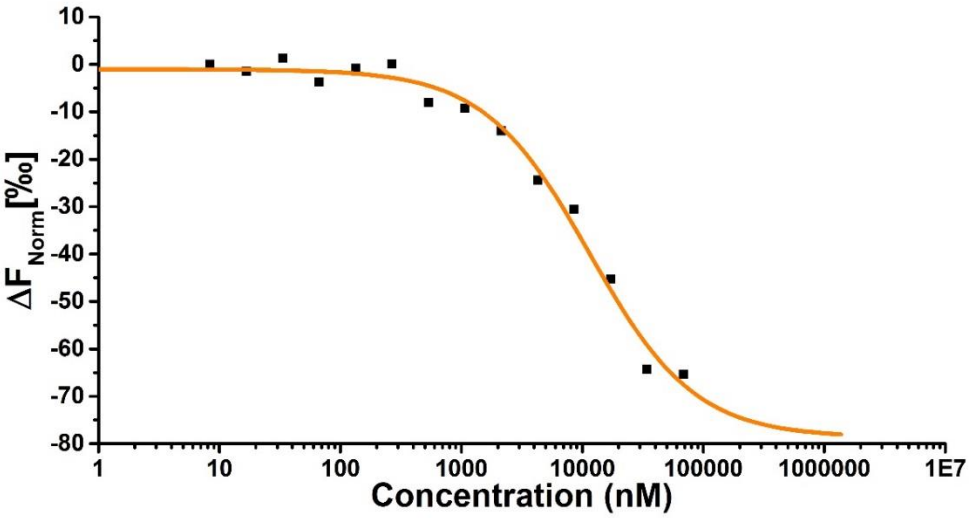


Figure 4-15. CED-2 can form dimers in solution. Dimerisation of CED-2 was measured using MST. $K_{d\ dimer}$ = 11.3 ± 0.4 µM.

4.5 Discussion

The functions of EGL-15 and CED-2 proteins have been previously reported (DeVore et al., 1995; Reddien and Horvitz, 2000; Goodman et al., 2003; Chai et al., 2012). However, the information gained on these functions often relies on EMS-generated mutants with premature stop codons which leave a truncated protein that is partially functioning or non-functioning (Reddien and Horvitz, 2000; Goodman et al., 2003). Therefore, there is a need to investigate *C. elegans* protein function using more precise mutations, as even a small part of the protein can have a big impact on protein function. For example, it has been shown in humans that a proline-rich motif in FGFR2 is important in normal cell signalling, as well as cancer development (Ahmed et al., 2010; Lin et al., 2012; Timsah et al., 2014). In CRKL, SH3 domains have differing functions: SH3N is the main interactor with proline-rich motif-containing proteins, while the SH3 domain acts as a dimerisation module (Harkiolaki et al., 2006). So far in *C. elegans* research, most commonly used *egl*-15 and *ced*-2 strains contain premature stop codons which delete the whole C-terminal tail of EGL-15 and disable both SH3 domains of CED-2. Usage of CRISPR/Cas9 in *C. elegans* has enabled the generation of precisely edited mutants. This allowed investigation of the importance of the proline-rich motifs in EGL-15 in *C. elegans*, as well as the functional importance of the SH3 domains in CED-2.

4.5.1 *egl-15* proline-rich motif mutants have no apparent phenotype

The EGL-15 C-terminal tail undergoes alternative splicing, and many different variants have been detected (WormBase, release WS280). The tail can have three potential proline-rich motifs: PVNLPSEP, PPNDP and PRNPLP (Figure 4-8). Deletion of either of these motifs did not produce any obviously

debilitating phenotypes (e.g. “bag of worms”, lethality) as seen in *egl-15* loss-of-function mutants (Goodman et al., 2003), suggesting that these motifs are not involved in canonical EGL-15 signalling. *egl-15* is also known to be involved in several other processes. However, the phenotypes of these are not clearly apparent. This includes CAN neuron migration (Fleming et al., 2005), muscle arm extension (Dixon et al., 2006), and fluid homeostasis (Sasson and Stern, 2004). So far, *ced-2* is not known to be involved in these functions, and as the focus of this project was to look into the interaction between EGL-15 and CED-2, these *egl-15*-related phenotypes were not investigated in detail. It would be interesting to see whether animals lacking any of the proline-rich motifs would show phenotypes related to these other EGL-15 roles.

4.5.2 EGL-15 PRNPLP proline-rich motif is the CED-2 binding site

A physical interaction between EGL-15 and CED-2 has been shown before using the yeast two-hybrid method (Lo et al., 2010). The Type-IV EGL-15 C-terminal tail, to which CED-2 binds, contained a unique PRNPLP proline-rich motif, while PVNLPSEP and PPNDP motifs were present in all four tail types used in the study (Lo et al., 2010). As CED-2 is comprised of both SH2 and SH3 domains, the protein could bind the EGL-15 C-terminal tail on phosphorylated tyrosine residues or proline-rich motifs. The PRNPLP proline-rich motif of EGL-15 was found to be the binding site for full-length CED-2 protein, and this interaction is likely to be mediated by the SH3 domains, because the peptides used in MST were not phosphorylated (Figure 4-9). The interaction is moderately strong (K_d = 3.5 ± 0.8 µM) for a SH3-proline-rich motif interaction, on par with the established binding range (Pawson, 1995), suggesting that it is possible for this interaction to occur *in vivo*. Interestingly, the binding affinity between the *C. elegans* proteins is very similar to the binding

affinity between human FGFR2 C-terminal tail and CRKL^{2xSH3} construct (Figure 3-4, K_d = 2.510 µM ± 0.264 µM). A GST pulldown was also attempted to show the interaction by pulling down CED-2 present in animal lysate with GST-Pro3 peptide (Figure 4-10). This was, however, unsuccessful due to non-specific binding of CED-2 to the GST tag.

4.5.3 EGL-15 PRNPLP motif has potential roles in germ cell engulfment

After establishing a direct interaction between the EGL-15 PRNPLP motif and CED-2, the functional consequences were investigated. CED-2-related phenotypes were investigated in the EGL-15(ΔPRNPLP) mutant. When engulfment of apoptotic germ cells was measured, the EGL-15(ΔPRNPLP) mutant strain showed a significantly different variance in engulfed apoptotic cell number compared to wild-type hermaphrodites (Figure 4-11). The results appear to indicate that EGL-15(ΔPRNPLP) animals are unable to tightly control the number of engulfed cells. The PRNPLP motif could be involved in regulating engulfment, and when it is deleted, gonadal sheath cells could become overactive and engulf an excessive number of germ cells, perhaps by losing their ability to distinguish between apoptotic and non-apoptotic cells. Wild-type *egl-15* could be selectively expressed in gonad sheath cells in EGL-15(ΔPRNPLP) mutant background to see whether correct EGL-15 function is indeed required in sheath cells to regulate engulfment. Alternatively, the PRNPLP proline-rich motif could be required for physiological apoptosis, and loss of it has induced excessive cell death. EGL-15 involvement in physiological apoptosis would not be surprising, as active RAS/MAPK signalling, which the RTK can activate, is known to be required for its induction (Gumienny et al., 1999). Wild-type *egl-15* could be selectively expressed in the germline of EGL-

15(ΔPRNPLP) animals to see if it can rescue the “loss of tight engulfment control” phenotype. Furthermore, wild-type and EGL-15(ΔPRNPLP) could be used in an RNAi screen, utilising knockdowns of various activators (*pax-2, egl-38)* and inhibitors (*lin-35)* of *C. elegans* physiological apoptosis to establish a potential role of EGL-15 in physiological apoptosis signalling pathway.

4.5.4 CED-2 SH3 domains are required for engulfment

In human cells, the two CRKL SH3 domains have differing functions. The SH3N domain binds to downstream effectors DOCK1 and C3G (Birge et al., 2009), and it is important for oncogenic function of CRKL in NSCLC (Cheung et al., 2011). While SH3C is lacking structural features to enable proline-rich motif binding (Jankowski et al., 2012), it can act as a dimerisation point for CRKL (Harkiolaki et al., 2006). This CRKL dimerisation can also enhance activation of FGFR2 (Figure 3-9). However, for *C. elegans* CED-2, only loss-of-function mutants containing premature stop codons have been used for research previously (Reddien and Horvitz, 2000). It was worthwhile to see whether the two CED-2 SH3 domains had separate roles in CED-2 function. Deletion of individual SH3 domains resulted in similar CED-2 loss-of-function phenotypes: 80.41% of CED-2(ΔSH3N) and 71.43% of CED-2(ΔSH3C) animals showed no engulfment of apoptotic germline cells, which was very similar to established *ced-2(e1752)* loss-of-function allele mutants (79.61%) (Figure 4-11C). This established both CED-2 SH3 domains as important regions of CED-2 for its function in engulfment.

4.5.5 CED-2 SH3 domains are involved in DTC pathfinding

The two SH3 domains are also important for CED-2 function in DTC pathfinding during gonad development (Figure 4-12). It has been previously

reported that loss of CED-2 function leads to abnormal cell morphology, with a frequency of 39% in *ced-2(e1752)* and 38% in *ced-2(n1994)* mutants (Reddien and Horvitz, 2000). Unfortunately, the same frequency was not observed in *ced-2(e1752)* strain, in which only 0.66% of the animals exhibited this phenotype, while *ced-2(1994)* were not scored. SH3 domain deletion mutants also did not reach the previously reported frequency of gonad morphology defects. The aberrant gonad phenotyped showed an extra gonad arm turn in 4.12% of CED-2(ΔSH3N) mutants, and in 14.29% of CED-2(ΔSH3C) mutants (Figure 4-12). These results establish CED-2 SH3 domain mutants as loss-of-function alleles with a minor role in DTC pathfinding. During DTC migration, CED-5 present at focal adhesions recruits CED-2 and CED-12 to form a GEF complex, which allows activation of CED-10 and in turn downstream signalling, enabling correct migration (Gumienny et al., 2001). The weak effect of loss of CED-2 function can be due to multiple factors. First, CED-2/CED-5/CED-12 is not the only complex which activates CED-10 – it can also be activated by integrins, as well as UNC-73 (DeBakker et al., 2004). Integrins can also activate SRC-1 and MIG-2 without any involvement of CED-10 (Wong and Schwarzbauer, 2012). Overall DTC migration is regulated by several different pathways, making CED-2 partially dispensable for this function. It could also be that one SH3 domain can partially compensate for loss of the other. Previous work implied that CED-2 could be interacting with CED-5 via its SH2 domain, rather than the SH3 (Kang et al., 2011). If that is the case, then the remaining SH3 domain (either N or C) could be able to interact with CED-12, still forming the GEF complex. The next step in determining the precise function of CED-2 in engulfment and DTC migration would involve biochemical studies. Interactions between recombinant CED-2 protein constructs (including either point mutants which would disable

functions of each domains, or constructs lacking whole domains) and potential CED-2 interactors such as CED-12, CED-5, GSK-3, APR-1 and SRC-1 (Figure 4-3; (Neukomm et al., 2011)) could be investigated utilising either MST or pulldown assays. This would allow to fine-tune already existing knowledge of the CED-2 signalling pathway which was built primarily from genetic interactions.

4.5.6 Functional CED-2 is needed for brood size maintenance

Physiological apoptosis in the *C. elegans* gonad is thought to act as a quality control for embryos, as well as provide cytoplasm and organelles such as mitochondria to the developing oocytes (Huelgas Morales and Greenstein, 2018). However, so far it is unknown whether the presence of unengulfed cell corpses would affect the quality of oocytes. Brood size was measured in *ced-2(e1752)* loss of function animals (Figure 4-13). Indeed, loss of *ced-2* function significantly reduced brood size to an average of 162.9 ± 65.5 progeny (compared to 285.0 ± 55.5 in wild-type hermpahrodites). There is however a possibility that *ced-2(e1752)* contains some other background mutations that can affect brood size. To have more confidence in the hypothesis, the brood size of *ced-2* SH3 domain mutants generated by CRISPR/Cas9 was also measured. It appears that the SH3N domain of CED-2 is indeed required to maintain a brood size – CED-2(ΔSH3N) mutant animals had an average brood size of 227.7 ± 55.8, which is significantly different from wild-type. However, the reduction of brood size is not a direct consequence of impaired apoptotic germ cell engulfment. Although CED-2(ΔSH3C) hermaphrodites showed a severe engulfment impairment phenotype, the average brood size in these animals (254.7 ± 40.7) was not significantly different from wild-type, rendering the initial hypothesis incorrect. Therefore CED-2 could be acting in the germline or

oocytes to allow correct embryo development. This could be tested by germline-specific expression of full-length CED-2 in CED-2(ΔSH3N) animals. Alternatively, CED-2 could be needed for survival in developing egg or during any of the larval stages. To see if that is the case, the number of hatched CED-2(ΔSH3N) mutant eggs could be measured, while a development assay would be employed to see whether the animals tend to encounter lethality at certain larval stages.

4.5.7 EGL-15 PRNPLP deletion rescues CED-2(ΔSH3N) loss of function phenotypes

To establish if there is a genetic interaction between the EGL-15 PRNPLP proline-rich motif and the CED-2 SH3 domains, *ced-2;egl-15* double mutants were generated. CED-2 related phenotypes, as well as brood size, were measured in these animals. In single mutant CED-2(ΔSH3N) hermaphrodites, engulfment of apoptotic germ cells is severely impaired. EGL-15(ΔPRNPLP) single mutant animals show an aberrant engulfment of germ cells, where the distribution of engulfed cell values is significantly different from that of wild-type *C. elegans*. However, in CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) double mutant, the distribution of engulfed cell values is not statistically significant compared to wild-type, indicating that this strain shows a rescue phenotype (Figure 4-11B). Furthermore, if only the total impairment of engulfment (zero corpses engulfed) is considered, the double mutant shows a rescue phenotype as well: in CED-2(ΔSH3N) mutants, 80.41% of nematodes show no engulfment, in EGL-15(ΔPRNPLP) 25.00% lack cell corpse engulfment, and in the CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) double mutant 34.57% show no engulfment (Figure 4-11C). The CED-2(ΔSH3N)*;*EGL-15(ΔPRNPLP) animals also showed no gonad morphology defect (0%,

compared to 4.12% in CED-2(ΔSH3N)), further reinforcing that loss of EGL-15 PRNPLP rescued engulfment and DTC pathfinding defects of CED-2(ΔSH3N) hermaphrodites (Figure 4-12B). A similar result was also observed using the brood size assay (Figure 4-13). CED-2(ΔSH3N);EGL-15(ΔPRNPLP) double mutants had an average brood size that is similar to wild-type, and significantly different to CED-2(ΔSH3N) single mutant, suggesting that loss of the EGL-15 PRNPLP in the CED-2 SH3N deletion background rescued the CED-2 loss of function phenotype.

The next step would be to see whether CED-2 and EGL-15 interaction and rescue occurs within the same tissue, which would imply that the proteins participate in a signalling pathway together. Although the physical interaction between EGL-15 and CED-2 has been established (Lo et al., 2010), (Figure 4-9), it is unclear whether this interaction can occur *in vivo*. So far, EGL-15 expression has been visualised in the hypodermis, intestine, vulva muscles and several neurons (Mounsey et al., 2002; Huang and Stern, 2004; Bülow et al., 2004; Lo et al., 2008; WormBase release WS280). But presence of EGL-15 in tissues harbouring CED-2 cannot be fully ruled out. *egl-15* transcript has been detected in the somatic gonad precursor cells (Mathies et al., 2019; WormBase release WS280), which give rise to somatic gonad cells such as the DTC and sheath cells (Kimble and Hirsh, 1979). Thus, it is important to see whether EGL-15 spice variant containing the PRNPLP proline-rich motif is present in the somatic gonad tissue within adult *C. elegans*. For this, CRISPR/Cas9 could be used to add a fluorescent tag to EGL-15, and confocal microscopy could be used to visualise the tissues where this isoform is expressed. Alternatively, full-length EGL-15 isoform containing the PRNPLP motif could be expressed in specific tissues (gonadal sheath for engulfment, and DTC for gonad

morphology) in a CED-2(ΔSH3N);EGL-15(ΔPRNPLP) double mutant background to see if CED-2 loss-of-function phenotype could be induced again. Furthermore, it would be interesting to see if EGL-15 ligands EGL-17 and LET-756 are required for apoptotic cell engulfment. As the ligands are important for development, auxin-inducible degron system could be used to deplete EGL-17 and/or LET-756 in L4 or young adult animals, followed by recording engulfed apoptotic cell number in day-1 adult nematodes.

In contrast to the SH3N deletion strain results, loss of the EGL-15 PRNPLP in the CED-2(ΔSH3C) mutant background did not rescue *ced-2* loss-of-function phenotypes of this allele. AO staining results in the double mutant showed that the CED-2(ΔSH3C);EGL-15(ΔPRNPLP) strain still acts as a CED-2 loss-of-function (Figure 4-11B, C). Gonad morphology in the double mutant was also defective in 5.33% of animals (Figure 4-12). However, an interaction between the two proteins cannot be ruled out completely. Several CED-2(ΔSH3C);EGL-15(ΔPRNPLP) double mutants showed a strange phenotype (Appendix E), where the animal did not have a functional reproductive system and could not produce any progeny. To investigate this further, nematodes DAPI stained to visualise the aberrant phenotypes of the reproductive system. The development of the gonad throughout life could also be tracked, to gain insight on how the development of the reproductive system goes wrong in those animals, which could then hint at potential pathways and genes involved in this case. Furthermore, wild-type *egl-15* and *ced-2* could be selectively expressed in germ line and somatic gonad precursors of CED-2(ΔSH3C);EGL-15(ΔPRNPLP) animals to see whether the reproductive system development dysfunction could be rescued.

The loss of EGL-15 PRNPLP proline-rich motif rescued multiple CED-2 loss of function phenotypes when combined with the loss of CED-2 SH3N domain – namely apoptotic germline cell engulfment, DTC migration and brood size defects. This indicates that the PRNPLP motif could be required for CED-2 loss-of-functions phenotypes to occur, suggesting that it could be a binding site for direct negative regulators of CED-2 signalling, or alternatively for proteins participating in a parallel pathway which indirectly negatively regulates CED-2 signalling. However, this negative regulation only occurs when the SH3N domain of CED-2 is deleted, which indicates that the two SH3 domains of CED-2 have different roles in the biological function of the protein.

4.5.8 EGL-15 might be involved in engulfment pathway

As a genetic interaction was established between the EGL-15 PRNPLP motif and the CED-2 SH3N domain, it is possible that EGL-15 can be involved in apoptotic cell engulfment pathway. The RNAi focused-screen in which *ina-1, mom-5* or *ced-10* were knocked down, was not very conclusive, likely due to the complexity of the engulfment pathway. The variance between wild-type and EGL-15(ΔPRNPLP) animals fed with control RNAi (EV) bacteria is significantly different, corroborating the previous AO staining results (Figure 4-11B). Strangely, EGL-15(ΔPRNPLP) hermaphrodites fed with control RNAi bacteria showed a higher dysfunction in engulfment (46.00%) than wild-type animals (13.04%), or even EGL-15(ΔPRNPLP) hermaphrodites fed with OP50 bacteria (25.00%, Figure 4-11C). This could be due to several reasons. Firstly, it could be due to the fact that a smaller number of animals was scored in this RNAi screen compared to previous AO staining assay (up to 72 per strain). It could also be that the EGL-15(ΔPRNPLP) strain is more sensitive to RNAi bacteria in general, because these mutants showed a more severe engulfment defect

phenotype under control RNAi. RNAi-mediated knockdown of *ina-1* in wild-type hermaphrodites changed the variance of engulfed cell corpses significantly, as well as increased the number of animals with dysfunctional engulfment. During *ina-1* RNAi, EGL-15(ΔPRNPLP) showed a significantly different variance in engulfed number of cells compared to both wild-type animals grown under same RNAi, and EGL-15(ΔPRNPLP) fed with control RNAi bacteria, while total number of hermaphrodites showing no engulfment was similar. These results show that reduction of *ina-1* function in both wild-type and EGL-15(ΔPRNPLP) animals can affect the range of the number of engulfed germline cell corpses, perhaps indicating that EGL-15 could signal parallel to integrins. When *mom-5* is knocked down, the variance in engulfed corpses between wild-type and EGL-15(ΔPRNPLP) does not differ significantly. Interestingly, for both strains, the variance is not significant compared to controls either. These results suggest that perhaps *mom-5* and *egl-15* could act in the same pathway, as there are seemingly no additive phenotypic effects. Furthermore, the percentage of animals with dysfunctional engulfment during *mom-5* RNAi is similar between wild-type and EGL-15(ΔPRNPLP) animals (54.43% and 62.50%, respectively). Results for *ced-10* RNAi are perplexing. There is no significant difference in variance of the engulfed cell number between strains grown on *ced-10* RNAi and same strains fed with control RNAi bacteria. However, the percentage of animals with no engulfment is very high for wild-type (73.33% compared to 13.04% control), while for EGL-15(ΔPRNPLP) animals it is similar to control (42.86% under *ced-10* RNAi vs 46.00% control). Results for *ced-10* RNAi are very inconclusive, most likely due to very low number of animals that were scored (15 N2 and 14 EGL-15(ΔPRNPLP)). These findings need to be reinforced with a higher number of nematodes. But the result still suggests that

EGL-15 could be involved in apoptotic cell engulfment pathway through CED-2 and CED-10, which has never been investigated before, establishing a new role for EGL-15 protein. To fully establish whether EGL-15 has a role in engulfment signalling, RNAi against more genes and double RNAi against combination of genes within the pathway should be tested, with more animals scored for each condition.

This work did not touch upon the function of CED-2 in engulfment of somatic cells. In the embryo, 113 somatic cells undergo programmed cell death, and 18 additional cells do so during the L2 stage. CED-2/5/12 pathway activation allows neighbouring cells to engulf apoptotic somatic cells (Wang and Yang, 2016). In the future, the importance of separate CED-2 SH3 domains, as well as the involvement of EGL-15 (and PRNPLP motif specifically) for somatic cell engulfment should be investigated.

4.5.9 CED-2 can dimerise

In human cells, CRKL has been shown to dimerise via its SH3C domains (Harkiolaki et al., 2006). In previous Chapter it was also shown that CRKL dimer can interact with FGFR2 and enhance its phosphorylation (Figure 3-9). Determining whether CED-2 can dimerise is imperative to be able to interpret potential interaction mechanism with EGL-15. In this Chapter, the formation of CED-2 dimer was shown (Figure 4-15), with the $K_{d\ dimer}$ of 11.3 ± 0.4 μM. It would be interesting to see if CED-2 also uses SH3C domains to dimerise, and whether dimerisation occurs *in vivo*. Although the dimerisation dissociation constant for CED-2 is weaker than that for CRKL ($K_{d\ dimer}$ = 0.093 μM ± 0.008 μM), dimerisation would be dependent on the cellular context. Furthermore, $K_{d\ dimer}$ for CED-2 is also weaker than K_d of CED-2 interaction with EGL-15 peptide

containing PRNPLP motif (K_d = 3.5 ± 0.8 μM), suggesting that CED-2 is unlikely to bind EGL-15 as a dimer. However, even if CED-2 dimerisation is not relevant when it comes to CED-2 interaction with EGL-15, dimerisation could still be an important part of CED-2 signalling in any of its roles.

4.6 Conclusions

In this Chapter, several new findings on *C. elegans* EGL-15 and CED-2 proteins were described. The EGL-15 PRNPLP proline-rich motif was established as a CED-2 binding site, although this direct interaction has not yet been shown *in vivo* biochemically. However, a genetic interaction has been shown between the EGL-15 PRNPLP proline-rich motif and the CED-2 SH3N domain that affects engulfment of apoptotic germline cell corpses, DTC pathfinding and brood size maintenance. It was also discovered that both SH3N and SH3C domains of CED-2 are important for normal function, but in a newly established role of CED-2 in brood size maintenance, only an intact SH3N domain is required. While these findings suggest that EGL-15 proline-rich motif could be important for signalling, it is unclear how this signalling is initiated. The roles of EGL-15 ligands EGL-17 and LET-756 are unclear. Thus, it is also unclear whether the PRNPLP proline-rich motif is important “tier 2” signalling in *C. elegans*. As “tier 2” signalling in human cells is described as occurring without receptor stimulation, it could be difficult to investigate whether it happens in the animals – loss of receptor-stimulating ligand genes can have severe consequences to nematode development and health. Thus, the next step would be to first see whether EGL-17 and LET-756 are involved in EGL-15 PRNPLP motif function. Loss of EGL-17 can cause a “bag-of-worms” phenotype (Burdine et al., 1998), while loss of LET-756 can be lethal (Goodman et al.,

2003). To overcome these severe phenotypes, auxin-inducible degron system could be used to induce degradation of EGL-15 ligands during adulthood when germline cell engulfment occurs and when embryos are produced to investigate the engulfment and brood size phenotypes, as well as specifically during the second turn of the gonad to investigate ligand involvement in the DTC migration function. Alternatively, once EGL-15 presence is established in the gonadal sheath, germ cells or DTC, EGL-15 lacking the extracellular domain could be selectively expressed in these tissues to see whether ligand-induced signalling is required for EGL-15 phenotypes in cell engulfment, DTC migration and brood size maintenance.

Chapter 5 Conclusions and future directions

This work focused on “tier 2” RTK signalling that is mediated by interactions between proline-rich C-terminal tails of RTKs and SH3 domains of binding partner proteins, in absence of extracellular stimulus. These interactions were found to be involved in some cancers (Gartside et al., 2009; Timsah et al., 2016). It is crucial to find more proteins that could participate in this “tier 2” signalling in order to further establish its importance in normal cell signalling, as well as cancer. CRKL, an adaptor protein containing SH2 and SH3 domains, was chosen for this work. This protein is an established oncogene, but knowledge of its upstream binding partners is lacking. Thus, it was a good candidate for this project. Chapter 3 focused on the FGFR-CRKL interaction, while the purpose of Chapter 4 was to determine whether this interaction is evolutionarily conserved in *C. elegans* model organism.

5.1 Novel interactions between CRKL SH3 domains and RTK C-terminal tails

A dot-blot screen showed that CRKL can bind proline-rich C-terminal tails of various RTKs. Many of these interactions are novel, thus this assay can act as a starting point in expanding our knowledge of CRKL involvement in various RTK signalling pathways in development and disease.

5.1.1 Interaction between FGFR2 and CRKL

Chapter 3 was focused on establishing the interaction between FGFR2 and CRKL, as well as the impact of this interaction on intracellular signalling. The results uncovered various aspects of this interaction. CRKL was shown bind FGFR2 via its SH2 and SH3N domains, suggesting that binding can occur to

both active (phosphorylated) and inactive (non-phosphorylated) receptor. Dimeric CRKL enhanced FGFR2 autotransphosphorylation, thus potentially its activation. In turn, active FGFR2 was shown to phosphorylate CRKL on multiple sites. The next big question is how do these findings connect? There are several potential scenarios. For example, dimeric CRKL could constitutively bind two molecules of FGFR2, forming a heterotetramer and hindering full FGFR2 activation, as it occurs within the FGFR/GRB2 heterotetramer (Lin et al., 2012). It could also be that CRKL-induced FGFR2 dimerisation leads to aberrant receptor activation even in absence of activating FGF ligand. In that case, CRKL could act similar to PLCγ1 (Timsah et al., 2014). Although CRKL does not seem to bind to the same proline-rich motif of FGFR2 as GRB2 (Ahmed et al., 2010), the motifs are within close proximity, which makes steric hindrance a possibility. Thus, CRKL and GRB2 could compete for binding of the FGFR C-terminal tail, and if CRKL outcompetes GRB2, it could lead to uncontrolled receptor activation and downstream signalling, potentially even in the absence of extracellular stimulus.

5.1.2 Downstream signalling pathways

Establishing the effect of FGFR2 and CRKL interaction on downstream signalling pathways was complicated. Activation of RAP1 and RAC1 pathways was investigated first, as they are well known downstream effectors of CRKL (Gotoh et al., 1995; Birge et al., 2009), but the results were highly variable and inconclusive. However, RPPA revealed a global change in cell signalling in cells with lowered CRKL expression, where in a broad sense, cell signalling occurs through the AKT pathway, while CRKL knockdown causes the switch towards MAPK signalling pathway. This dataset provides an overview on how CRKL

presence can shape cell signalling which can act as a guidance tool for more in-depth research of CRKL signalling.

5.2 Interaction between EGL-15 and CED-2

The results in Chapter 4 have shown that the interaction between EGL-15 and CED-2 is evolutionarily conserved in *Caenorhabditis elegans*. Physical interaction between EGL-15 and CED-2 proteins was shown using a biophysical assay, while phenotypic assays showed a genetic interaction between *egl-15* and *ced-2*, suggesting that *in vivo* these genes could participate in same or similar pathways. Surprisingly, the results show that EGL-15 PRNPLP motif could act as a negative regulator of CED-2 signalling (when SH3N domain is deleted), which is opposite from what was found in human cells, where CRKL was found to increase FGFR2 phosphorylation. It was discovered that the EGL-15 PRNPLP proline-rich motif is physiologically relevant in these animals, showing that it has a novel role in regulating apoptotic germ cell engulfment, DTC migration and normal brood size maintenance.

In conclusion, this work shows that the interaction between FGFR2 and CRKL is evolutionarily conserved. Furthermore, these interactions in both human cells and *C. elegans* nematode animals are mediated by the proline-rich motif within the RTK C-terminal tail and the SH3 domains of the adapter protein. However, it is yet unclear whether the “tier 2” signalling could be evolutionarily conserved, as the involvement of EGL-15 ligands in apoptotic cell engulfment, DTC migration and brood size has not yet been investigated. Thus further work is required to establish the presence of “tier 2” signalling in other organisms.

References

Acquaviva, J., Wong, R. and Charest, A. 2009. The multifaceted roles of the receptor tyrosine kinase ROS in development and cancer. *Biochimica et Biophysica Acta - Reviews on Cancer*. **1795**(1), pp.37–52.

Ahmed, Z., George, R., Lin, C.C., Suen, K.M., Levitt, J.A., Suhling, K. and Ladbury, J.E. 2010. Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cellular Signalling*. **22**(1), pp.23–33.

Alessi, D.R., James, S.R., Downes, C.P., Holmes, A.B., Gaffney, P.R.J., Reese, C.B. and Cohen, P. 1997. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Bα. *Current Biology*. **7**(4), pp.261–269.

An, R., Wang, Y., Voeller, D., Gower, A., Kim, I.K., Zhang, Y.W. and Giaccone, G. 2016. CRKL mediates EML4-ALK signaling and is a potential therapeutic target for ALK-rearranged lung adenocarcinoma. *Oncotarget*. **7**(20), pp.29199–29210.

Anderson, D., Koch, C.A., Grey, L., Ellis, C., Foran, M.F. and Pawson, T. 1990. Binding of SH2 Domains of Phospholipase C. *Science*. **250**(4983), pp.979–982.

Arribere, J.A., Bell, R.T., Fu, B.X.H., Artiles, K.L., Hartman, P.S. and Fire, A.Z. 2014. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in caenorhabditis elegans. *Genetics*. **198**(3), pp.837–846.

Atfi, A., Abécassis, L. and Bourgeade, M.F. 2005. Bcr-Abl activates the AKT/FoxO3 signalling pathway to restrict transforming growth factor-β-mediated cytostatic signals. *EMBO Reports*. **6**(10), pp.985–991.

Azoury, S.C., Reddy, S., Shukla, V. and Deng, C.X. 2017. Fibroblast growth factor receptor 2 (FGFR2) mutation related syndromic craniosynostosis. *International Journal of Biological Sciences*. **13**(12), pp.1479–1488.

Ballif, B.A., Arnaud, L., Arthur, W.T., Guris, D., Imamoto, A. and Cooper, J.A. 2004. Activation of a Dab1/CrkL/C3G/Rap1 Pathway in Reelin-Stimulated Neurons. *Current Biology*. **14**, pp.606–610.

Basu, S., Totty, N.F., Irwin, M.S., Sudol, M. and Downward, J. 2003. Akt Phosphorylates the Yes-Associated Protein, YAP, to Induce Interaction with 14-3-3 and Attenuation of p73-Mediated Apoptosis some of the metabolic effects of Akt, whereas BAD can account for some of the survival-promoting effects. The most convincingly. *Molecular Cell*. **11**, pp.11–23.

Bauer, N.N., Chen, Y.W., Samant, R.S., Shevde, L.A. and Fodstad, O. 2007. Rac1 activity regulates proliferation of aggressive metastatic melanoma. *Experimental Cell Research*. **313**(18), pp.3832–3839.

Beurel, E., Grieco, S.F. and Jope, R.S. 2015. Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. *Pharmacol Ther*. **0**, pp.114–131.

Bian, X., Liang, Z., Feng, A., Salgado, E. and Shim, H. 2018. HDAC inhibitor suppresses proliferation and invasion of breast cancer cells through regulation of miR-200c targeting CRKL. *Biochemical Pharmacology*. **147**, pp.30–37.

Birge, R.B., Kalodimos, C., Inagaki, F. and Tanaka, S. 2009. Crk and CrkL adaptor proteins: Networks for physiological and pathological signaling. *Cell Communication and Signaling*. **7**, pp.1–23.

Bock, H.H. and May, P. 2016. Canonical and non-canonical reelin signaling. *Frontiers in Cellular Neuroscience*. **10**(Jun), pp.1–20.

Bong, S. and Rhee, S.C. 1995. Significance of PIP2 hydrolysis and regulation of phospholipase C isozymes. *Current Opinion in Cell Biology*. **7**, pp.183–189.

Borland, C.Z., Schutzman, J.L. and Stern, M.J. 2001. Fibroblast growth factor signaling in Caenorhabditis elegans. *BioEssays*. **23**(12), pp.1120–1130.

Branda, C.S. and Stern, M.J. 2000. Mechanisms controlling sex myoblast migration in Caenorhabditis elegans hermaphrodites. *Developmental Biology*. **226**(1), pp.137–151.

Brasher, B.B. and Van Etten, R.A. 2000. c-Abl has high intrinsic tyrosine kinase activity that is stimulated by mutation of the Src homology 3 domain and by autophosphorylation at two distinct regulatory tyrosines. *Journal of Biological Chemistry*. **275**(45), pp.35631–35637.

Brenner, S. 1974. The genetics of Caenorhabditis elegans. *Genetics*. (77), pp.71–94.

Bülow, H.E., Boulin, T. and Hobert, O. 2004. Differential functions of the C. elegans FGF receptor in axon outgrowth and maintenance of axon position. *Neuron*. **42**(3), pp.367–374.

Burdine, R.D., Branda, C.S. and Stern, M.J. 1998. EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in C. elegans. *Development*. **125**(6), pp.1083–1093.

Byerly, L., Cassada, R.C. and Russell, R.L. 1976. The life cycle of the nematode Caenorhabditis elegans. *Developmental Biology*. **51**(1), pp.23–33.

Cabello, J., Neukomm, L.J., Günesdogan, U., Burkart, K., Charette, S.J., Lochnit, G., Hengartner, M.O. and Schnabel, R. 2010. The Wnt pathway controls cell death engulfment, spindle orientation, and migration through CED-10/Rac. *PLoS Biology*. **8**(2).

Cai, L., Wang, H. and Yang, Q. 2017. CRKL overexpression promotes cell proliferation and inhibits apoptosis in endometrial carcinoma. *Oncology Letters*. **13**(1), pp.51–56.

Cecchetelli, A.D. and Cram, E.J. 2017. Regulating distal tip cell migration in space and time. *Mechanisms of Development*. **148**, pp.11–17.

Cha, J.Y., Lambert, Q.T., Reuther, G.W. and Der, C.J. 2008. Involvement of fibroblast growth factor receptor 2 isoform switching in mammary oncogenesis. *Molecular Cancer Research*. **6**(3), pp.435–445.

Chai, Y., Tian, D., Yang, Y., Feng, G., Cheng, Z., Li, W. and Ou, G. 2012. Apoptotic regulators promote cytokinetic midbody degradation in C. Elegans. *Journal of Cell Biology*. **199**(7), pp.1047–1055.

Chardin, P., Camonis, J.H., Gale, N.W., Van Aelst, L., Schlessinger, J., Wigler, M.H. and Bar-Sagi, D. 1993. Human Sos1: A guanine nucleotide exchange factor for Ras that binds to GRB2. *Science*. **260**(5112), pp.1338–1343.

Chellaiah, A.T., McEwen, D.G., Werner, S., Xu, J. and Ornitz, D.M. 1994. Fibroblast Growth Factor Receptor (FGFR) 3. Alternative splicing in immunoglobulin-like domain III creates a receptor highly specific for acidic FGF/FGF-1. *The Journal of Biological Chemistry*. **269**, pp.11620–11627.

Chen, E.B. and Stern, M.J. 1998. Understanding cell migration guidance: lessons from sex myoblast migration in C. elegans. *Trends in Genetics*. **14**(8), pp.322–327.

Cheung, H.W., Du, J., Boehm, J.S., He, F., Weir, B.A., Wang, X., Butaney, M., Sequist, L. V., Luo, B., Engelman, J.A., Root, D.E., Meyerson, M., Golub, T.R., Jänne, P.A. and Hahn, W.C. 2011. Amplification of CRKL induces transformation and epidermal growth factor receptor inhibitor resistance in human non-small cell lung cancers. *Cancer Discovery*. **1**(7), pp.608–625.

Chong, H., Lee, J. and Guan, K.L. 2001. Positive and negative regulation of Raf kinase activity and function by phosphorylation. *EMBO Journal*. **20**(14), pp.3716–3727.

Clark, S.G., Stern, M.J. and Horvitz, H.R. 1992. C. elegans cell-signalling gene sem-5 encodes a protein with SH2 and SH3 domains. *Letters to Nature*. **356**(March), pp.340–344.

Cram, E.J., Shang, H. and Schwarzbauer, J.E. 2006. A systematic RNA interference screen reveals a cell migration gene network in C. elegans. *Journal of Cell Science*. **119**(23), pp.4811–4818.

Crittenden, S.L., Eckmann, C.R., Wang, L., Bernstein, D.S., Wickens, M. and Kimble, J. 2003. Regulation of the mitosis/meiosis decision in the Caenorhabditis elegans germline. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **358**(1436), pp.1359–1362.

Cunningham, D.L., Sarhan, A.R., Creese, A.J., Larkins, K.P.B., Zhao, H., Ferguson, H.R., Brookes, K., Marusiak, A.A., Cooper, H.J. and Heath, J.K. 2020. Differential responses to kinase inhibition in FGFR2-addicted triple negative breast cancer cells: a quantitative phosphoproteomics study. *Scientific reports*. **10**(1), p.7950.

DeBakker, C.D., Haney, L.B., Kinchen, J.M., Grimsley, C., Lu, M., Klingele, D., Hsu, P., Chou, B., Cheng, L., Blangy, A., Sondek, J., Hengartner, M.O., Wu, Y., Ravichandran, K.S., Carolina, N., Hill, C. and Hill, C. 2004. Phagocytosis of Apoptotic Cells Is Regulated by a UNC-73 / TRIO-MIG-2 / RhoG Signaling Module and Armadillo Repeats of CED-12 / ELMO. *Current Biology*. **14**, pp.2208–2216.

Defilippi, P., Di Stefano, P. and Cabodi, S. 2006. p130Cas: a versatile scaffold in signaling networks. *Trends in Cell Biology*. **16**(5), pp.257–263.

Dell, K.R. and Williams, L.T. 1992. A novel form of fibroblast growth factor receptor 2. Alternative splicing of the third immunoglobulin-like domain confers ligand binding specificity. *Journal of Biological Chemistry*. **267**(29), pp.21225–21229.

Deng, N., Goh, L.K., Wang, H., Das, K., Tao, J., Tan, I.B., Zhang, S., Lee, M., Wu, J., Lim, K.H., Lei, Z., Goh, G., Lim, Q.Y., Tan, A.L.K., Poh, D.Y.S., Riahi, S., Bell, S., Shi, M.M., Linnartz, R., Zhu, F., Yeoh, K.G., Toh, H.C., Yong, W.P., Cheong, H.C., Rha, S.Y., Boussioutas, A., Grabsch, H., Rozen, S. and Tan, P. 2012. A comprehensive survey of genomic

alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. *Gut*. **61**(5), pp.673–684.

Dent, P., Haser, W., Haystead, T.A.J., Vincent, L.A., Roberts, T.M. and Sturgill, T.W. 1992. Activation of mitogen-activated protein kinase kinase by v-raf in NIH 3T3 cells and in vitro. *Science*. **257**(5075), pp.1404–1407.

DeVore, D.L., Horvitz, H.R. and J.Stern, M. 1995. An FGF receptor signaling pathway is required for the normal cell migrations of the sex myoblasts in C. elegans hermaphrodites. *Cell*. **83**(4), pp.611–620.

Dixon, S.J., Alexander, M., Fernandes, R., Ricker, N. and Roy, P.J. 2006. FGF negatively regulates muscle mebrane extension in Caenorhabditis elegans. *Development*. **133**(7), pp.1263–1275.

Feller, S.M. 2001. Crk family adaptors - signalling complex formation and biological roles. *Oncogene*. **20**, pp.6348–6371.

Felschow, D.M., McVeigh, M.L., Hoehn, G.T., Civin, C.I. and Fackler, M.J. 2001. The adapter protein CrkL associates with CD34. *Blood*. **97**(12), pp.3768–3775.

Feng, S., Chen, J.K., Yu, H., Simon, J.A. and Schreiber, S.L. 1994. Two binding orientations for peptides to the Src SH3 domain: Development of a general model for SH3-ligand interactions. *Science*. **266**(5188), pp.1241–1247.

Filippakopoulos, P., Müller, S. and Knapp, S. 2009. SH2 domains: modulators of nonreceptor tyrosine kinase activity. *Current Opinion in Structural Biology*. **19**(6), pp.643–649.

Fleming, T.C., Wolf, F.W. and Garriga, G. 2005. Sensitized genetic backgrounds reveal a role for C. elegans FGF EGL-17 as a repellent for migrating CAN neurons. *Development*. **132**(21), pp.4857–4867.

Franke, F.C., Müller, J., Abal, M., Medina, E.D., Nitsche, U., Weidmann, H., Chardonnet, S., Ninio, E. and Janssen, K.P. 2019. The Tumor Suppressor SASH1 Interacts With the Signal Adaptor CRKL to Inhibit Epithelial–Mesenchymal Transition and Metastasis in Colorectal Cancer. *Cmgh*. **7**(1), pp.33–53.

Franke, T.F., Kaplan, D.R. and Cantley, L.C. 1997. PI3K: Downstream AKTion blocks apoptosis. *Cell*. **88**(4), pp.435–437.

Fu, L., Dong, Q., Xie, C., Wang, Y. and Li, Q. 2015. CRKL protein overexpression enhances cell proliferation and invasion in pancreatic cancer. *Tumor Biology*. **36**(2), pp.1015–1022.

Fullard, J.F., Kale, A. and Baker, N.E. 2009. Clearance of apoptotic corpses. *Apoptosis*. **14**(8), pp.1029–1037.

Gartside, M.G., Chen, H., Ibrahimi, O.A., Byron, S.A., Curtis, A. V., Wellens, C.L., Bengston, A., Yudt, L.M., Eliseenkova, A. V., Ma, J., Curtin, J.A., Hyder, P., Harper, U.L., Riedesel, E., Mann, G.J., Trent, J.M., Bastian, B.C., Meltzer, P.S., Mohammadi, M. and Pollock, P.M. 2009. Loss-of-function fibroblast growth factor receptor-2 mutations in melanoma. *Molecular Cancer Research*. **7**(1), pp.41–54.

Goodman, S.J., Branda, C.S., Robinson, M.K., Burdine, R.D. and Stern, M.J. 2003. Alternative splicing affecting a novel domain in the C. elegans EGL-

15 FGF receptor confers functional specificity. *Development*. **130**(16), pp.3757–3766.

Gorry, M.C., Preston, R.A., White, G.J., Zhang, Y., Singhal, V.K., Losken, H.W., Parker, M.G., Nwokoro, N.A., Post, J.C. and Ehrlich, G.D. 1995. Crouzon syndrome: Mutations in two spliceoforms of FGFR2 and a common point mutation shared with jackson - weiss syndrome. *Human Molecular Genetics*. **4**(8), pp.1387–1390.

Gotoh, T., Hattori, S., Nakamura, S., Kitayama, H., Noda, M., Takai, Y., Kaibuchi, K., Matsui, H., Hatase, O. and Takahashi, H. 1995. Identification of Rap1 as a target for the Crk SH3 domain-binding guanine nucleotide-releasing factor C3G. *Molecular and Cellular Biology*. **15**(12), pp.6746–6753.

Gumienny, T.L., Brugnera, E., Tosello-Trampont, A.C., Kinchen, J.M., Haney, L.B., Nishiwaki, K., Walk, S.F., Nemergut, M.E., Macara, I.G., Francis, R., Schedl, T., Qin, Y., Van Aelst, L., Hengartner, M.O. and Ravichandran, K.S. 2001. CED-12/ELMO, a novel member of the CrkII/Dock180/Rac pathway, is required for phagocytosis and cell migration. *Cell*. **107**(1), pp.27–41.

Gumienny, T.L., Lambie, E., Hartwieg, E., Horvitz, H.R. and Hengartner, M.O. 1999. Genetic control of programmed cell death in the Caenorhabditis elegans hermaphrodite germline. *Development*. **126**(5), pp.1011–1022.

Guris, D.L., Fantes, J., Tara, D., Druker, B.J. and Imamoto, A. 2001. Mice lacking the homologue of the human 22q11.2 gene CRLK phenocopy neurocristopathies of DiGeorge syndrome. *Nature Genetics*. **27**(3), pp.293–298.

Gustafson, T.A., He, W., Craparo, A., Schaub, C.D. and O'Neill, T.J. 1995. Phosphotyrosine-dependent interaction of SHC and insulin receptor substrate 1 with the NPEY motif of the insulin receptor via a novel non-SH2 domain. *Molecular and Cellular Biology*. **15**(5), pp.2500–2508.

Hadari, Y.R., Kouhara, H., Lax, I. and Schlessinger, J. 1998. Binding of Shp2 Tyrosine Phosphatase to FRS2 Is Essential for Fibroblast Growth Factor-Induced PC12 Cell Differentiation. *Molecular and Cellular Biology*. **18**(7), pp.3966–3973.

Hall, D.H., Winfrey, V.P., Blaeuer, G., Hoffman, L.H., Furuta, T., Rose, K.L., Hobert, O. and Greenstein, D. 1999. Ultrastructural features of the adult hermaphrodite gonad of Caenorhabditis elegans: Relations between the germ line and soma. *Developmental Biology*. **212**, pp.101–123.

Harkiolaki, M., Gilbert, R.J.C., Jones, E.Y. and Feller, S.M. 2006. The C-Terminal SH3 Domain of CRKL as a Dynamic Dimerization Module Transiently Exposing a Nuclear Export Signal. *Structure*. **14**(12), pp.1741–1753.

Hart, K.C., Robertson, S.C., Kanemitsu, M.Y., Meyer, A.N., Tynan, J.A. and Donoghue, D.J. 2000. Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4. *Oncogene*. **19**(29), pp.3309–3320.

Hasegawa, H., Kiyokawa, E., Tanaka, S., Nagashima, K., Gotoh, N., Shibuya, M., Kurata, T. and Matsuda, M. 1996. DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane.

Molecular and Cellular Biology. **16**(4), pp.1770–1776.

Hattori, Y., Itoh, H., Shinya, U., Hosokawa, K., Ochiai, A., Ino, Y., Ishii, H., Sakamoto, H., Yamaguchi, N., Yanagihara, K., Hirohashi, S., Sugimura, T. and Terada, M. 1996. Immunohistochemical Detection of K-sam Protein in Stomach Cancer. *Clinical Cancer Research*. **2**(August), pp.1373–1381.

Hedgecock, E.M., Sulston, J.E. and Thomson, J.N. 1983. Mutations affecting programmed cell deaths in the nematode Caenorhabditis elegans. *Science*. **220**(4603), pp.1277–1279.

Hirsh, D., Oppenheim, D. and Klass, M. 1976. Development of the reproductive system of Caenorhabditis elegans. *Developmental Biology*. **49**(1), pp.200–219.

Ten Hoeve, J., Arlinghaus, R.B., Guo, J.Q., Heisterkamp, N. and Groffen, J. 1994. Tyrosine phosphorylation of CRKL in Philadelphia+ leukemia. *Blood*. **84**(6), pp.1731–1736.

ten Hoeve, J., Morris, C., Heisterkamp, N. and Groffen, J. 1993. Isolation and chromosomal localization of CRKL, a human crk-like gene. *Oncogene*. **8**(9), pp.2469–2474.

Honda, H. 1925. Experimental and Cytological Studies on Bisexual and Hermaphrodite Free-Living Nematodes, With Special Reference to Problems of Sex. *Journal of Morphology and Physiology*. **40**(2), pp.192–233.

Horvitz, H.R. 1999. Genetic control of programmed cell death in the nematode Caenorhabditis elegans. *Cancer Research*. **59**(7 SUPPL.), pp.1011–1022.

Hsu, T.Y. and Wu, Y.C. 2010. Engulfment of Apoptotic Cells in C. elegans Is Mediated by Integrin α/SRC Signaling. *Current Biology*. **20**(6), pp.477–486.

Huang, H., Li, L., Wu, C., Schibli, D., Colwill, K., Ma, S., Li, C., Roy, P., Ho, K., Songyang, Z., Pawson, T., Gao, Y. and Li, S.S.C. 2008. Defining the specificity space of the human Src homology 2 domain. *Molecular and Cellular Proteomics*. **7**(4), pp.768–784.

Huang, P. and Stern, M.J. 2004. FGF signaling functions in the hypodermis to regulate fluid balance in C. elegans. *Development*. **131**(11), pp.2595–2604.

Hubbard, S.R. 1999. Structural analysis of receptor tyrosine kinases. *Progress in Biophysics and Molecular Biology*. **71**(3–4), pp.343–358.

Hubbard, S.R., Mohammadi, M. and Schlessinger, J. 1998. Autoregulatory mechanisms in protein-tyrosine kinases. *Journal of Biological Chemistry*. **273**(20), pp.11987–11990.

Hubbard, S.R. and Till, J.H. 2000. Protein Tyrosine Kinase Structure and Function. *Ann. Rev. Biochem*. **69**, pp.373–398.

Hubbard, S.R., Wei, L., Ellis, L. and Hendrickson, W.A. 1994. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature (London)*. **372**(6508), pp.746–754.

Huelgas Morales, G. and Greenstein, D. 2018. C. elegans germline cell death, live! *PLoS Genetics*. **14**(7), pp.1–3.

Hunter, D.J., Kraft, P., Jacobs, K.B., Cox, D.G., Yeager, M., Hankinson, S.E., Wacholder, S., Wang, Z., Welch, R., Yu, K., Chatterjee, N., Orr, N., Willett,

W.C., Graham, A., Ziegler, R.G., Berg, C.D., Buys, S.S., Catherine, A., Feigelson, H.S., Calle, E.E., Thun, M.J., Richard, B., Tucker, M., Gerhard, D.S., Jr, J.F.F., Robert, N., Thomas, G. and Chanock, S.J. 2007. GWAS study identifies alleles in FGFR2 associated with risk of breast cancer. *Nature genetics*. **39**(7), pp.870–874.

Hurwitz, M.E., Vanderzalm, P.J., Bloom, L., Goldman, J., Garriga, G. and Horvitz, H.R. 2009. Abl kinase inhibits the engulfment of apopotic cells in Caenorhabditis elegans. *PLoS Biology*. **7**(4), pp.0876–0891.

Itoh, H., Hattori, Y., Sakamoto, H., Ishii, H., Kishi, T., Sasaki, H., Yoshida, T., Koono, M., Sugimura, T. and Terada, M. 1994. Preferential Alternative Splicing in Cancer Generates a K-sam Messenger RNA with Higher Transforming Activity. *Cancer Research*. **54**, pp.3237–3242.

Jabs, E.W., Lp, X., Scott, A.F., Meyers, G., Chen, W., Eccles, M., Mao, J., Charnas, L.R., Jackson, C.E. and Jaye, M. 1994. Jackson-Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor č. *Nature Genetics*. **8**(november), pp.275–279.

Jankowski, W., Saleh, T., Pai, M.T., Sriram, G., Birge, R.B. and Kalodimos, C.G. 2012. Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nature Chemical Biology*. **8**(6), pp.590–596.

Johnson, D.E., Lu, J., Chen, H., Werner, S. and Williams, L.T. 1991. The human fibroblast growth factor receptor genes: a common structural arrangement underlies the mechanisms for generating receptor forms that differ in their third immunoglobulin domain. *Molecular and Cellular Biology*. **11**(9), pp.4627–4634.

De Jong, R., Ten Hoeve, J., Heisterkamp, N. and Groffen, J. 1997. Tyrosine 207 in CRKL is the BCR/ABL phosphorylation site. *Oncogene*. **14**(5), pp.507–513.

Kanazawa, S., Fujiwara, T., Matsuzaki, S., Shingaki, K., Taniguchi, M., Miyata, S., Tohyama, M., Sakai, Y., Yano, K., Hosokawa, K. and Kubo, T. 2010. bFGF regulates PI3-Kinase-Rac1-JNK pathway and promotes fibroblast migration in wound healing. *PLoS ONE*. **5**(8), pp.1–12.

Kandel, E.S., Skeen, J., Majewski, N., Di Cristofano, A., Pandolfi, P.P., Feliciano, C.S., Gartel, A. and Hay, N. 2002. Activation of Akt/Protein Kinase B Overcomes a G2/M Cell Cycle Checkpoint Induced by DNA Damage. *Molecular and Cellular Biology*. **22**(22), pp.7831–7841.

Kane, L.P., Shapiro, V.S., Stokoe, D. and Weiss, A. 1999. Induction of NF-κB by the Akt/PKB kinase. *Current Biology*. **9**(11), pp.601–604.

Kang, S., Elf, S., Dong, S., Hitosugi, T., Lythgoe, K., Guo, A., Ruan, H., Lonial, S., Khoury, H.J., Williams, I.R., Lee, B.H., Roesel, J.L., Karsenty, G., Hanauer, A., Taunton, J., Boggon, T.J., Gu, T.-L. and Chen, J. 2009. Fibroblast Growth Factor Receptor 3 Associates with and Tyrosine Phosphorylates p90 RSK2, Leading to RSK2 Activation That Mediates Hematopoietic Transformation. *Molecular and Cellular Biology*. **29**(8), pp.2105–2117.

Kang, Y., Xu, J., Liu, Yong, Sun, J., Sun, D., Hu, Y. and Liu, Yingfang 2011. Crystal structure of the cell corpse engulfment protein CED-2 in Caenorhabditis elegans. *Biochemical and Biophysical Research*

Communications. **410**(2), pp.189–194.

Katoh, M. 2008. Cancer genomics and genetics of FGFR2 (Review). *International Journal of Oncology*. **2**(33), pp.233–237.

Kay, B.K., Williamson, M.P. and Sudol, M. 2000. The importance of being proline : the interaction of proline-rich motifs in signaling proteins with their cognate domains. *The FASEB Journal*. **14**, pp.231–241.

Kim, Y.H., Kwei, K.A., Girard, L., Salari, K., Kao, J., Wang, P., Hernandez-boussard, T., Gazdar, A.F., Minna, J.D. and Pollack, J.R. 2010. Genomic and functional analysis identified CRKL as an oncogene amplified in lung cancer. *Oncogene*. **29**(10), pp.1421–1430.

Kimble, J. and Hirsh, D. 1979. The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Developmental Biology*. **70**(2), pp.396–417.

Kimble, J.E. and White, J.G. 1981. On the Control of Germ Cell Development in Caenorhabditis elegans. *Developmental Biology*. **219**, pp.208–219.

Kitayama, H., Sugimoto, Y., Matsuzaki, T., Ikawa, Y. and Noda, M. 1989. A ras-related gene with transformation suppressor activity. *Cell*. **56**(1), pp.77–84.

Klint, P. and Claesson-Welsh, L. 1999. Signal transduction by fibroblast growth factor receptors. *Frontiers in bioscience : a journal and virtual library*. **4**(7), pp.165–177.

Knudsen, B.S., Feller, S.M. and Hanafusa, H. 1994. Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk. *Journal of Biological Chemistry*. **269**(52), pp.32781–32787.

Kobashigawa, Y. and Inagaki, F. 2012. CrkL is not Crk-like. *Nature Chemical Biology*. **8**(6), pp.504–505.

Kokel, M., Borland, C.Z., DeLong, L., Horvitz, H.R. and Stern, M.J. 1998. clr-1 encodes a receptor tyrosine phosphatase that negatively regulates an FGF receptor signaling pathway in Caenorhabditis elegans. *Genes and Development*. **12**(10), pp.1425–1437.

Kostrzewska-Poczekaj, M., Bednarek, K., Jarmuz-Szymczak, M., Bodnar, M., Filas, V., Marszalek, A., Bartochowska, A., Grenman, R., Kiwerska, K., Szyfter, K. and Giefing, M. 2020. Copy number gains of the putative CRKL oncogene in laryngeal squamous cell carcinoma result in strong nuclear expression of the protein and influence cell proliferation and migration. *Scientific Reports*. **10**(1), pp.1–9.

Kouhara, H., Hadari, Y.R., Spivak-Kroizman, T., Schilling, J., Bar-Sagi, D., Lax, I. and Schlessinger, J. 1997. A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell*. **89**(5), pp.693–702.

Kuroyanagi, H., Kobayashi, T., Mitani, S. and Hagiwara, M. 2006. Transgenic alternative-splicing reporters reveal tissue-specific expression profiles and regulation mechanisms in vivo. *Nature Methods*. **3**(11), pp.909–915.

Lemmon, M.A. and Schlessinger, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell*. **141**(7), pp.1117–1134.

Li, Y., Dowbenko, D. and Lasky, L.A. 2002. AKT/PKB phosphorylation of p21Cip/WAF1 enhances protein stability of p21Cip/WAF1 and promotes cell survival. *Journal of Biological Chemistry*. **277**(13), pp.11352–11361.

Light, Y., Paterson, H. and Marais, R. 2002. 14-3-3 Antagonizes Ras-Mediated Raf-1 Recruitment to the Plasma Membrane To Maintain Signaling Fidelity. *Molecular and Cellular Biology*. **22**(14), pp.4984–4996.

Lim, W.A., Richards, F.M. and Fox, R.O. 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Letters to Nature*. **372**(November), pp.375–379.

Lin, C.C., Melo, F.A., Ghosh, R., Suen, K.M., Stagg, L.J., Kirkpatrick, J., Arold, S.T., Ahmed, Z. and Ladbury, J.E. 2012. Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell*. **149**(7), pp.1514–1524.

Lin, F., Chengyao, X., Qingchang, L., Qianze, D., Enhua, W. and Yan, W. 2015. CRKL promotes lung cancer cell invasion through ERK-MMP9 pathway. *Molecular Carcinogenesis*. **54**(S1), pp.E35–E44.

Ling, L., Zhu, T. and Lobie, P.E. 2003. Src-CrkII-C3G-dependent activation of Rap1 switches growth hormone-stimulated p44/42 MAP kinase and JNK/SAPK activities. *Journal of Biological Chemistry*. **278**(29), pp.27301–27311.

Liu, D. 2014. The adaptor protein Crk in immune response. *Immunology and Cell Biology*. **92**(1), pp.80–89.

Lo, T.W., Bennett, D.C., Goodman, S.J. and Stern, M.J. 2010. Caenorhabditis elegans fibroblast growth factor receptor signaling can occur independently of the multi-substrate adaptor FRS2. *Genetics*. **185**(2), pp.537–547.

Lo, T.W., Branda, C.S., Huang, P., Sasson, I.E., Goodman, S.J. and Stern, M.J. 2008. Different isoforms of the C. elegans FGF receptor are required for attraction and repulsion of the migrating sex myoblasts. *Developmental Biology*. **318**(2), pp.268–275.

Luo, B., Hiu, W.C., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., Hinkle, G., Boehm, J.S., Beroukhim, R., Weir, B.A., Mermel, C., Barbie, D.A., Awad, T., Zhou, X., Nguyen, T., Piqani, B., Li, C., Golub, T.R., Meyerson, M., Hacohen, N., Hahn, W.C., Lander, E.S., Sabatini, D.M. and Root, D.E. 2008. Highly parallel identification of essential genes in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*. **105**(51), pp.20380–20385.

Marais, R., Light, Y., Paterson, H.F. and Marshall, C.J. 1995. Ras recruits Raf-1 to the plasma membrane far activation by tyrosine phosphorylation. *EMBO Journal*. **14**(13), pp.3136–3145.

Markaki, M. and Tavernarakis, N. 2010. Modeling human diseases in Caenorhabditis elegans. *Biotechnology Journal*. **5**(12), pp.1261–1276.

Mathies, L.D., Ray, S., Lopez-Alvillar, K., Arbeitman, M.N., Davies, A.G. and Bettinger, J.C. 2019. MRNA profiling reveals significant transcriptional differences between a multipotent progenitor and its differentiated sister. *BMC Genomics*. **20**(1), pp.1–15.

Matsuda, M., Mayer, B.J., Fukui, Y. and Hanafusa, H. 1990. Binding of transforming protein, P47gag-crk, to a broad range of phosphotyrosine-

containing proteins. *Science*. **248**(4962), pp.1537–1539.

Matsuda, M., Tanaka, S., Nagata, S., Kojima, A., Kurata, T. and Shibuya, M. 1992. Two species of human CRK cDNA encode proteins with distinct biological activities. *Molecular and Cellular Biology*. **12**(8), pp.3482–3489.

Matsuki, T., Pramatarova, A. and Howell, B.W. 2008. Reduction of Crk and CrkL expression blocks reelin-induced dendritogenesis. *Journal of Cell Science*. **121**(11), pp.1869–1875.

Mayer, B.J., Hamaguchi, M. and Hanafusa, H. 1988. A novel viral oncogene with structural similarity to phospholipase C. *Nature*. **332**(6161), pp.272–275.

McCarter, J., Bartlett, B., Dang, T. and Schedl, T. 1997. Soma-germ cell interactions in Caenorhabditis elegans: Multiple events of hermaphrodite germline development require the somatic sheath and spermathecal lineages. *Developmental Biology*. **181**(2), pp.121–143.

Miki, T., Bottaro, D.P., Fleming, T.P., Smith, C.L., Burgess, W.H., Chan, A.M. and Aaronson, S.A. 1992. Determination of ligand-binding specificity by alternative splicing : Two distinct growth factor receptors encoded by a single gene. *Proceedings of the National Academy of Sciences of the United States of America*. **89**(January), pp.246–250.

Mohammadi, M., Honegger, A.M., Rotin, D., Fischer, R., Bellot, F., Li, W., Dionne, C.A., Jaye, M., Rubinstein, M. and Schlessinger, J. 1991. A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. *Molecular and Cellular Biology*. **11**(10), pp.5068–5078.

Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B.K., Hubbard, S.R. and Schlessinger, J. 1997. Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science*. **276**(5314), pp.955–960.

Moon, A.M., Guris, D.L., Seo, J.H., Li, L., Hammond, J., Talbot, A. and Imamoto, A. 2006. Crkl deficiency disrupts Fgf8 signaling in a mouse model of 22q11 deletion syndromes. *Developmental Cell*. **10**(1), pp.71–80.

del Moral, P.M., De Langhe, S.P., Sala, F.G., Veltmaat, J.M., Tefft, D., Wang, K., Warburton, D. and Bellusci, S. 2006. Differential role of FGF9 on epithelium and mesenchyme in mouse embryonic lung. *Developmental Biology*. **293**(1), pp.77–89.

Mounsey, A., Bauer, P. and Hope, I.A. 2002. Evidence suggesting that a fifth of annotated Caenorhabditis elegans genes may be pseudogenes. *Genome Research*. **12**(5), pp.770–775.

Muralidharan, V., Dutta, K., Cho, J., Vila-Perello, M., Raleigh, D.P., Cowburn, D. and Muir, T.W. 2006. Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II. *Biochemistry*. **45**(29), pp.8874–8884.

Neukomm, L.J., Nicot, A.S., Kinchen, J.M., Almendinger, J., Pinto, S.M., Zeng, S., Doukoumetzidis, K., Tronchère, H., Payrastre, B., Laporte, J.F. and Hengartner, M.O. 2011. The phosphoinositide phosphatase MTM-1 regulates apoptotic cell corpse clearance through CED-5-CED-12 in C. elegans. *Development*. **138**(10), pp.2003–2014.

Nishiwaki, K. 1999. Mutations affecting symmetrical migration of distal tip cells in Caenorhabditis elegans. *Genetics*. **152**(3), pp.985–997.

Nosaka, Y., Arai, A., Miyasaka, N. and Miura, O. 1999. CrkL mediates Ras-dependent activation of the Raf/ERK pathway through the guanine nucleotide exchange factor C3G in hematopoietic cells stimulated with erythropoietin or interleukin-3. *Journal of Biological Chemistry*. **274**(42), pp.30154–30162.

Oda, T., Heaney, C., Hagopian, J.R., Okuda, K., Griffin, J.D. and Druker, B.J. 1994. Crkl is the major tyrosine-phosphorylated protein in neutrophils from patients with chronic myelogenous leukemia. *Journal of Biological Chemistry*. **269**(37), pp.22925–22928.

Ohtsuka, T., Shimizu, K., Yamamori, B., Kuroda, S. and Takai, Y. 1996. Activation of brain B-Raf protein kinase by Rap1B small GTP-binding protein. *Journal of Biological Chemistry*. **271**(3), pp.1258–1261.

Ong, S.H., Guy, G.R., Hadari, Y.R., Laks, S., Gotoh, N., Schlessinger, J. and Lax, I. 2000. FRS2 Proteins Recruit Intracellular Signaling Pathways by Binding to Diverse Targets on Fibroblast Growth Factor and Nerve Growth Factor Receptors. *Molecular and Cellular Biology*. **20**(3), pp.979–989.

Ong, S.H., Hadari, Y.R., Gotoh, N., Guy, G.R., Schlessinger, J. and Lax, I. 2001. Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins. *Proceedings of the National Academy of Sciences of the United States of America*. **98**(11), pp.6074–6079.

Ornitz, D.M. and Itoh, N. 2001. Protein family review Fibroblast growth factors Gene organization and evolutionary history. *Genome*. **2**(3), pp.1–12.

Ornitz, D.M. and Itoh, N. 2015. The fibroblast growth factor signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology*. **4**(3), pp.215–266.

Ornitz, D.M., Xu, J., Colvin, J.S., McEwen, D.G., MacArthur, C.A., Coulier, F., Gao, G. and Goldfarb, M. 1996. Receptor specificity of the fibroblast growth factor family. *Journal of Biological Chemistry*. **271**(25), pp.15292–15297.

Orr-Urtreger, A., Bedford, M.T., Burakova, T., Arman, E., Zimmer, Y., Yayon, A., Givol, D. and Lonai, P. 1993. Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR2). *Developmental Biology*. **158**(2), pp.475–486.

Paix, A., Folkmann, A., Rasoloson, D. and Seydoux, G. 2015. High efficiency, homology-directed genome editing in Caenorhabditis elegans using CRISPR-Cas9ribonucleoprotein complexes. *Genetics*. **201**(1), pp.47–54.

Park, T.-J., Boyd, K. and Curran, T. 2006. Cardiovascular and Craniofacial Defects in Crk-Null Mice. *Molecular and Cellular Biology*. **26**(16), pp.6272–6282.

Park, T.J. and Curran, T. 2008. Crk and Crk-like play essential overlapping roles downstream of disabled-1 in the reelin pathway. *Journal of Neuroscience*. **28**(50), pp.13551–13562.

Park, W.J., Meyers, G.A., Li, X., Theda, C., Day, D., Oriow, S.J., Jones, M.C. and Jabs, E.W. 1995. Novel FGFR2 mutations in crouzon and jackson-weiss syndromes show allelic heterogeneity and phenotypic variability.

Human Molecular Genetics. **4**(7), pp.1229–1233.

Pawson, T. 1995. Protein modules and signalling networks. *Nature*. **373**(6515), pp.573–580.

Paznekas, W.A., Cunningham, M.L., Howard, T.D., Korf, B.R., Lipson, M.H., Grix, A.W., Feingold, M., Goldberg, R., Borochowitz, Z., Aleck, K., Mulliken, J., Yin, M. and Jabs, E.W. 1998. Genetic heterogeneity of Saethre-Chotzen syndrome, due to TWIST and FGFR mutations. *American Journal of Human Genetics*. **62**(6), pp.1370–1380.

Pelech, S.L. and Sanghera, J.S. 1992. MAP kinases: Charting the regulatory pathways. *Science*. **257**(5075), pp.1355–1356.

Plotnikov, A.N., Schlessinger, J., Hubbard, S.R. and Mohammadi, M. 1999. Structural Basis for FGF Receptor Dimerization and Activation et al Dimerization of the extracellular domains leads to juxtaposition of the cytoplasmic domains and. *Cell*. **98**, pp.641–650.

Powers, C.J., McLeskey, S.W. and Wellstein, A. 2000. Fibroblast growth factors, their receptors and signaling. *Endocrine-Related Cancer*. **7**(3), pp.165–197.

Przylepa, K.A., Paznekas, W., Zhang, M., Golabi, M., Bias, W., Bamshad, M.J., Carey, J.C., Hall, B.D., Stevenson, R., Orlow, S.J., Cohen Jr, M.M. and Jabs, E.W. 1996. Fibroblast growth factor receptor 2 mutations in Beare-Stevenson cutis gyrata syndrome. *Nature Genetics*. **13**(August), pp.492–494.

Raghunath, M., Patti, R., Bannerman, P., Lee, C.M., Baker, S., Sutton, L.N., Phillips, P.C. and Damodar Reddy, C. 2000. A novel kinase, AATYK induces and promotes neuronal differentiation in a human neuroblastoma (SH-SY5Y) cell line. *Molecular Brain Research*. **77**(2), pp.151–162.

Reddien, P.W. and Horvitz, H.R. 2000. CED-2/CrkII and CED-10/Rac control phagocytosis and cell migration in Caenorhabditis elegans. *Nature Cell Biology*. **2**(3), pp.131–136.

Reichman, C.T., Mayer, B.J., Keshav, S. and Hanafusa, H. 1992. The product of the cellular crk gene consists primarily of SH2 and SH3 regions. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*. **3**(7), pp.451–460.

Reintjes, N., Li, Y., Becker, A., Rohmann, E., Schmutzler, R. and Wollnik, B. 2013. Activating Somatic FGFR2 Mutations in Breast Cancer. *PLoS ONE*. **8**(3).

Rhodes, J., York, R.D., Tara, D., Tajinda, K. and Druker, B.J. 2000. CrkL functions as a nuclear adaptor and transcriptional activator in Bcr-Abl-expressing cells. *Experimental Hematology*. **28**(3), pp.305–310.

Richter, M., Murai, K.K., Bourgin, C., Pak, D.T. and Pasquale, E.B. 2007. The EphA4 receptor regulates neuronal morphology through SPAR-mediated inactivation of rap GTPases. *Journal of Neuroscience*. **27**(51), pp.14205–14215.

Robinson, D.R., Wu, Y.M. and Lin, S.F. 2000. The protein tyrosine kinase family of the human genome. *Oncogene*. **19**(49), pp.5548–5557.

Ronan, T., Macdonald-Obermann, J.L., Huelsmann, L., Bessman, N.J., Naegle,

K.M. and Pike, L.J. 2016. Different Epidermal Growth Factor Receptor (EGFR) agonists produce unique signatures for the recruitment of downstream signaling proteins. *Journal of Biological Chemistry*. **291**(11), pp.5528–5540.

La Rosée, P., Holm-Eriksen, S., Konig, H., Härtel, N., Ernst, T., Debatin, J., Mueller, M.C., Erben, P., Binckebanck, A., Wunderle, L., Shou, Y., Dugan, M., Hehlmann, R., Ottmann, O.G. and Hochhaus, A. 2008. Phospho-CRKL monitoring for the assessment of BCR-ABL activity in imatinib-resistant chronic myeloid leukemia or Ph+ acute lymphoblastic leukemia patients treated with nilotinib. *Haematologica*. **93**(5), pp.765–769.

Rosen, M.K., Yamazaki, T., Gisht, G.D., Kayi, C.M., Pawsont, T. and Kay, L.E. 1995. Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. *Letters to Nature*. **374**, pp.477–479.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T. and Bowtell, D. 1993. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Letters to Nature*. **363**(May), pp.83–85.

Rutland, P., Louise, J., Reardon, W., Baraitser, M., Hayward, R., Jones, B., Malcolm, S., Winter, R.M., Oldridge, M., Slaney-, S.F., Poole, M.D. and Wilkie, A.O.M. 1995. Identical mutations in the FGFR2 gene cause both Pfeiffer and Couzon syndrome phenotypes. *Nature Genetics*. **9**(february), pp.173–176.

Sadowski, I., Stone, J.C. and Pawson, T. 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. *Molecular and Cellular Biology*. **6**(12), pp.4396–4408.

Sakkab, D., Lewitzky, M., Posern, G., Schaeper, U., Sachs, M., Birchmeier, W. and Feller, S.M. 2000. Signaling of hepatocyte growth factor/scatter factor (HGF) to the small GTPase Rap1 via the large docking protein Gab1 and the adapter protein CRKL. *Journal of Biological Chemistry*. **275**(15), pp.10772–10778.

Salgia, R., Uemura, N., Okuda, K., Li, J.L., Pisick, E., Sattler, M., De Jong, R., Druker, B., Heisterkamp, N., Lan Bo Chen, Groffen, J. and Griffin, J.D. 1995. CRKL links p210(BCR/ABL) with paxillin in chronic myelogenous leukemia cells. *Journal of Biological Chemistry*. **270**(49), pp.29145–29150.

Sasson, I.E. and Stern, M.J. 2004. FGF and PI3 kinase signaling pathways antagonistically modulate sex muscle diferrentiation in C. elegans. *Development*. **131**(21), pp.5381–5392.

Schlessinger, J. 2000. Cell Signaling by Receptor Tyrosine Kinases. *Cell*. **103**, pp.211–225.

Schlessinger, J. 2014. Receptor tyrosine kinases: Legacy of the first two decades. *Cold Spring Harbor Perspectives in Biology*. **6**(3), pp.1–13.

Schlessinger, J. 1988. Signal transduction by allosteric receptor oligomerization. *Trends in Biochemical Sciences*. **13**(11), pp.443–447.

Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A. V, Yeh, B.K., Yayon, A., Linhardt, R.J. and Mohammadi, M. 2000. Crystal Structure of a Ternary FGF-FGFR-Heparin Complex Reveals a Dual Role for Heparin in FGFR Binding and Dimerization. *Molecular Cell*. **6**, pp.743–750.

Schmidt, J.W., Wehde, B.L., Sakamoto, K., Triplett, A.A., Anderson, S.M., Tsichlis, P.N., Leone, G. and Wagner, K.-U. 2014. Stat5 Regulates the Phosphatidylinositol 3-Kinase/Akt1 Pathway during Mammary Gland Development and Tumorigenesis. *Molecular and Cellular Biology*. **34**(7), pp.1363–1377.

Schönherr, C., Yang, H.L., Vigny, M., Palmer, R.H. and Hallberg, B. 2010. Anaplastic lymphoma kinase activates the small GTPase Rap1 via the Rap1-specific GEF C3G in both neuroblastoma and PC12 cells. *Oncogene*. **29**(19), pp.2817–2830.

Schutzman, J.L., Borland, C.Z., Newman, J.C., Robinson, M.K., Kokel, M. and Stern, M.J. 2001. The Caenorhabditis elegans EGL-15 Signaling Pathway Implicates a DOS-Like Multisubstrate Adaptor Protein in Fibroblast Growth Factor Signal Transduction. *Molecular and Cellular Biology*. **21**(23), pp.8104–8116.

Ségaliny, A.I., Tellez-Gabriel, M., Heymann, M.F. and Heymann, D. 2015. Receptor tyrosine kinases: Characterisation, mechanism of action and therapeutic interests for bone cancers. *Journal of Bone Oncology*. **4**(1), pp.1–12.

Senechal, K., Halpern, J. and Sawyers, C.L. 1996. The CRKL adaptor protein transforms fibroblasts and functions in transformation by the BCR-ABL oncogene. *Journal of Biological Chemistry*. **271**(38), pp.23255–23261.

Senechal, K., Heaney, C., Druker, B. and Sawyers, C.L. 1998. Structural Requirements for Function of the Crkl Adapter Protein in Fibroblasts and Hematopoietic Cells. *Molecular and Cellular Biology*. **18**(9), pp.5082–5090.

Seo, J.-H., Suenaga, A., Hatakeyama, M., Taiji, M. and Imamoto, A. 2009. Structural and Functional Basis of a Role for CRKL in a Fibroblast Growth Factor 8-Induced Feed-Forward Loop. *Molecular and Cellular Biology*. **29**(11), pp.3076–3087.

Shigeno-Nakazawa, Y., Kasai, T., Ki, S., Kostyanovskaya, E., Pawlak, J., Yamagishi, J., Okimoto, N., Taiji, M., Okada, M., Westbrook, J., Satta, Y., Kigawa, T. and Imamoto, A. 2016. A pre-metazoan origin of the CRK gene family and co-opted signaling network. *Scientific Reports*. **6**(January), pp.1–14.

Smith, J.J., Richardson, D.A., Kopf, J., Yoshida, M., Hollingsworth, R.E. and Kornbluth, S. 2002. Apoptotic Regulation by the Crk Adapter Protein Mediated by Interactions with Wee1 and Crm1/Exportin. *Molecular and Cellular Biology*. **22**(5), pp.1412–1423.

Su, N., Jin, M. and Chen, L. 2014. Role of FGF/FGFR signaling in skeletal development and homeostasis: Learning from mouse models. *Bone Research*. **2**(December 2013), pp.1–24.

Su, X., Zhan, P., Gavine, P.R., Morgan, S., Womack, C., Ni, X., Shen, D., Bang, Y.J., Im, S.A., Ho Kim, W., Jung, E.J., Grabsch, H.I. and Kilgour, E. 2014. FGFR2 amplification has prognostic significance in gastric cancer: Results from a large international multicentre study. *British Journal of Cancer*. **110**(4), pp.967–975.

Sulston, J.E. and Horvitz, H.R. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology*. **56**(1), pp.110–

156.

Sulston, J.E., Schierenberg, E., White, J.G. and Thomson, J.N. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology*. **100**(1), pp.64–119.

Tanaka, K., Abe, M. and Sato, Y. 1999. Roles of extracellular signal-regulated kinase 1/2 and p38 mitogen-activated protein kinase in the signal transduction of basic fibroblast growth factor in endothelial cells during angiogenesis. *Japanese Journal of Cancer Research*. **90**(6), pp.647–654.

Tanaka, S., Morishita, T., Hashimoto, Y., Hattori, S., Nakamura, S., Shibuya, M., Matuoka, K., Takenawa, T., Kurata, T., Nagashima, K. and Matsuda, M. 1994. C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proceedings of the National Academy of Sciences of the United States of America*. **91**(8), pp.3443–3447.

Teyra, J., Huang, H., Jain, S., Guan, X., Dong, A., Liu, Y., Tempel, W., Min, J., Tong, Y., Kim, P.M., Bader, G.D. and Sidhu, S.S. 2017. Comprehensive Analysis of the Human SH3 Domain Family Reveals a Wide Variety of Non-canonical Specificities. *Structure*. **25**(10), pp.1598-1610.e3.

Timsah, Z., Ahmed, Z., Ivan, C., Berrout, J., Gagea, M., Zhou, Y., Pena, G.N.A., Hu, X., Vallien, C., Kingsley, C. V., Lu, Y., Hancock, J.F., Liu, J., Gladden, A.B., Mills, G.B., Lopez-Berestein, G., Hung, M.C., Sood, A.K., Bogdanov, M. and Ladbury, J.E. 2016. Grb2 depletion under non-stimulated conditions inhibits PTEN, promotes Akt-induced tumor formation and contributes to poor prognosis in ovarian cancer. *Oncogene*. **35**(17), pp.2186–2196.

Timsah, Z., Ahmed, Z., Lin, C.C., Melo, F.A., Stagg, L.J., Leonard, P.G., Jeyabal, P., Berrout, J., O'Neil, R.G., Bogdanov, M. and Ladbury, J.E. 2014. Competition between Grb2 and Plcγ1 for FGFR2 regulates basal phospholipase activity and invasion. *Nature Structural and Molecular Biology*. **21**(2), pp.180–188.

Tiong, K.H., Mah, L.Y. and Leong, C.O. 2013. Functional roles of fibroblast growth factor receptors (FGFRs) signaling in human cancers. *Apoptosis*. **18**(12), pp.1447–1468.

Tsygankova, O.M., Saavedra, A., Rebhun, J.F., Quilliam, L.A. and Meinkoth, J.L. 2001. Coordinated Regulation of Rap1 and Thyroid Differentiation by Cyclic AMP and Protein Kinase A. *Molecular and Cellular Biology*. **21**(6), pp.1921–1929.

Turner, N. and Grose, R. 2010. Fibroblast growth factor signalling: From development to cancer. *Nature Reviews Cancer*. **10**(2), pp.116–129.

Ueda, T., Sasaki, H., Kuwahara, Y., Nezu, M., Shibuya, T., Sakamoto, H., Ishii, H., Yanagihara, K., Mafune, K., Makuuchi, M. and Terada, M. 1999. Deletion of the Carboxyl-Terminal Exons of K-sam/FGFR2 by Short Homology-mediated Recombination, Generating Preferential Expression of Specific Messenger RNAs. *Cancer Research*. **59**, pp.6080–6086.

Uemura, N. and Griffin, J.D. 1999. The adapter protein Crkl links Cbl to C3G after integrin ligation and enhances cell migration. *Journal of Biological Chemistry*. **274**(53), pp.37525–37532.

Uemura, N., Salgia, R., Li, J.L., Pisick, E., Sattler, M. and Griffin, J.D. 1997. The

BCR/ABL oncogene alters interaction of the adapter proteins CRKL and CRK with cellular proteins. *Leukemia*. **11**(3), pp.376–385.

Vazquez, F., Ramaswamy, S., Nakamura, N. and Sellers, W.R. 2000. Phosphorylation of the PTEN Tail Regulates Protein Stability and Function. *Molecular and Cellular Biology*. **20**(14), pp.5010–5018.

Waksman, G. and Kuriyan, J. 2004. Structure and specificity of the SH2 domain. *Cell*. **116**(2 Suppl).

Wang, X., Wang, J., Gengyo-Ando, K., Gu, L., Sun, C.L., Yang, C., Shi, Yong, Kobayashi, T., Shi, Yigong, Mitani, S., Xie, X.S. and Xue, D. 2007. C. elegans mitochondrial factor WAH-1 promotes phosphatidylserine externalization in apoptotic cells through phospholipid scramblase SCRM-1. *Nature Cell Biology*. **9**(5), pp.541–549.

Wang, X. and Yang, C. 2016. Programmed cell death and clearance of cell corpses in Caenorhabditis elegans. *Cellular and Molecular Life Sciences*. **73**(11–12), pp.2221–2236.

Ward, S., Argon, Y. and Nelson, G.A. 1981. Sperm morphogenesis in wild-type and fertilization-defective mutants of Caenorhabditis elegans. *Journal of Cell Biology*. **91**(1), pp.26–44.

Werner, S., Duan, D.R., Vries, C.D.E., Peters, K.G., Johnson, D.E. and Williams, L.T. 1992. Differential Splicing in the Extracellular Region of Fibroblast Growth Factor Receptor 1 Generates Receptor Variants with Different Ligand-Binding Specificities. *Molecular and Cellular Biology*. **12**(1), pp.82–88.

Whitman, M., Downes, C.P., Keeler, M., Keller, T. and Cantley, L. 1988. Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate. *Nature*. **332**(6165), pp.644–646.

Wilkie, A.O.M. 1997. Craniosynostosis: Genes and mechanisms. *Human Molecular Genetics*. **6**(10 REV. ISS.), pp.1647–1656.

Williamson, M.P. 1994. The structure and function of proline-rich regions in proteins. . **260**, pp.249–260.

Wong, M.C. and Schwarzbauer, J.E. 2012. Gonad morphogenesis and distal tip cell migration in the Caenorhabditis elegans hermaphrodite. *Wiley Interdisciplinary Reviews: Developmental Biology*. **1**(4), pp.519–531.

Yarden, Y. and Axel, U. 1988. Growth factor receptor tyrosine kinases. *Ann. Rev. Biochem*. **57**, pp.443–78.

Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P. and Ornitz, D.M. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell*. **64**(4), pp.841–848.

Yeung, C.L., Ngo, V.N., Grohar, P.J., Arnaldez, F.I., Asante, A., Wan, X., Khan, J., Hewitt, S.M., Khanna, C., Staudt, L.M. and Helman, L.J. 2013. Loss-of-function screen in rhabdomyosarcoma identifies CRKL-YES as a critical signal for tumor growth. *Oncogene*. **32**(47), pp.5429–5438.

Yoon, S. and Seger, R. 2006. The extracellular signal-regulated kinase: Multiple substrates regulate diverse cellular functions. *Growth Factors*. **24**(1), pp.21–44.

Zhang, J., Gao, X., Schmit, F., Adelmant, G., Eck, M.J., Marto, J.A., Zhao, J.J. and Roberts, T.M. 2017. CRKL Mediates p110β-Dependent PI3K Signaling in PTEN-Deficient Cancer Cells. *Cell Reports*. **20**(3), pp.549–557.

Zhang, S., Li, F., Zhou, T., Wang, G. and Li, Z. 2020. Caenorhabditis elegans as a Useful Model for Studying Aging Mutations. *Frontiers in Endocrinology*. **11**(October), pp.1–9.

Zhang, Y.L., Wang, R.C., Cheng, K., Ring, B.Z. and Su, L. 2017. Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer Biology and Medicine*. **14**(1), pp.90–99.

Zhao, T., Miao, Z., Wang, Z., Xu, Y., Wu, J., Liu, X., You, Y. and Li, J. 2013. Overexpression of CRKL correlates with malignant cell proliferation in breast cancer. *Tumor Biology*. **34**(5), pp.2891–2897.

Zhao, X., Weir, B.A., LaFramboise, T., Lin, M., Beroukhim, R., Garraway, L., Beheshti, J., Lee, J.C., Naoki, K., Richards, W.G., Sugarbaker, D., Chen, F., Rubin, M.A., Jänne, P.A., Girard, L., Minna, J., Christiani, D., Li, C., Sellers, W.R. and Meyerson, M. 2005. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. *Cancer Research*. **65**(13), pp.5561–5570.

Zheng, C.F. and Guan, K.L. 1993. Properties of MEKs, the kinases that phosphorylate and activate the extracellular signal-regulated kinases. *Journal of Biological Chemistry*. **268**(32), pp.23933–23939.

Zhou, S., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., Birge, R.B., Fajardo, J.E., Chou, M.M., Hanafusa, H., Schaffhausen, B. and Cantley, L.C. 1993. SH2 domains recognize specific phosphopeptide sequences. *Cell*. **72**(5), pp.767–778.

Appendix A

PCR primers

Primer purpose	Primer sequence
egl-15(jez10) mutant detection	5'- TGA CAC TTG GTG GAA CTC CC -3'
	5'- AAT TCG GGT TTG CTC ATG CT -3'
egl-15(jez20) mutant detection	5'- TGC TCA ACG AAC CAA TCG GA -3'
	5'- ACT GAG CAT AAG TGT GGA GAA GG -3'
egl-15(jez8) detection	5'- AAC ACA CAA TTT GCG ACG AC -3'
	5'- TCT GCA ACC CAT TAC ACC AA -3'
egl-15 "pro3" oligo	5'- TAT ATC CAG TCA CAT CCA CTC TAC TCG AAA ATC ATA ATC AAA AAG AAC ATG ACG CCT CGC AAC CCG TTG CCG ACA AAG GAA ACT ATT GTC -3'
egl-15 pro3 cloning	5'- GCG GGA TCC TAT ATC CAG TCA -3'
	5'- TAA AGC GGC CGC GAC AAT AGT -3'
ced-2(jez45) detection	5'- CCG TGA AGA GGC TCA CAA GT -3'
	5'- AAC GAA CCA GGA TTG GTG GG -3'
	5'- GGC GGA AAA GGC AAC GAA AA -3'
ced-2(jez50) deletion detection	5'- TGC GCC TTT AAA CTT GTG TCG -3'
	5'- ACA CGA ATT GGG GTT CGG TG -3'
ced-2 amplifying from cDNA	5'- ATG ACG ACA AAC GGG TTC GA -3'
	5'- TTA TTC GCT GAC GGC GGT GA -3'
ced-2 cloning to pGEX	5'- CGC GGA TCC ATG ACG ACA AAC GGG TTC GA -3'
	5'- AAG GAA AAA AGC GGC CGC TTA TTC GCT GAC GGC GGT GA -3'
egl-15(jez10) repair	5'- ATT CCA GAA AAA GAG AAA ACA TCG CCC TCT TTC AGC TGT AGA TCA AAA CGT AAT TTC GTA CAT AAC GTA TTT A -3'
egl-15(jez8) repair	5'- TTT GCG ACG ACT ATG AGT CCA ATT TTT CTG TAG AAA ATC ATC TCT ACT GCA ATG ACA ATA TGC TGA AAA A -3'
egl-15(jez20) repair	5'- TCT ACT CGA AAA TCA TAA TCA AAA AGA ACA TGA CGA CAA AGG AAA CTA TTG TCT AAC TTC GTT TTT GAC T -3'
ced-2(jez45) repair	5'- TCA TTT TAA AAT GCG AGT TTT GAC TGA GGC AAG CCT GCT GGC CGC CTA CAA GAA GTT TCA CAA TGA CAG AAC CTC AAA AGG CGC CAG TCA ATC GTC GAT CGG AAG CAG CG -3'
ced-2(jez50) repair	5'- CGA TTT TTA TCA TGA AAA ACA TCT CCT AAA CCG AAA AAA AAA TTC AAT ATT TCA GAG CGA ATA AAA TCG TTT TTT TTT TCT AAA TTT TTA ACC TTG AAA TTT ATT GAT TC -3'
CRKL R21A mutation generation	5'- GGC CGG TGT CT**G CG**C AGG AGG CGC AG -3'
	5'- CTG CGC CTC CTG **CGC** AGA CAC CGG CC -3'
CRKL R39A mutation generation	5'- GGT ATG TTC CTC GTC **GCG** GAT TCT TCC ACC TGC -3'
	5'- GCA GGT GGA AGA ATC **CGC** GAC GAG GAA CAT ACC -3'
CRKL W160L mutation generation	5'- GAA GCC TGA AGA ACA G**CT** GTG GAG TGC CCG GAA C -3'
	5'- GTT CCG GGC ACT CCA C**AG** CTG TTC TTC AGG CTT C -3'
CRKL Y198F mutation generation	5'- GAA TTC CAA CAG TT**T** TGG GAT CCC AGA ACC -3'
	5'- GGT TCT GGG ATC CCA **A**AA CTG TTG GAA TTC -3'
CRKL Y207F mutation generation	5'- GAA CCT GCT CAT GCA T**TT** GCT CAA CCT CAG ACC -3'
	5'- GGT CTG AGG TTG AGC **AA**A TGC ATG AGC AGG TTC -3'
CRKL W275L mutation generation	5'- GAA TAT AAA TGG CCA G**CT** GGA AGG CGA AGT GAA CG -3'
	5'- CGT TCA CTT CGC CTT CC**A G**CT GGC CAT TTA TAT TC -3'

Thermocycler reaction programmes

Purpose	Programme
egl-15 pro-3 cloning	98°C (30s) -> [98°C (10s) -> 64°C (30s) -> 72°C (10s)]x25 -> 72°C(5min)
egl-15(jez10) and *egl-15(jez8)* detection	95°C (5min) -> [95°C (10s) -> 60°C (15s) -> 72°C (3min)]x32 -> 72°C (5min)
egl-15(jez20) detection	98°C (5min) -> [98°C(10s) -> 68°C (15s) -> 72°C (1min30s)]x35 -> 72°C (5min)
ced-2(jez45) detection	95°C (5min) -> [95°C (10s) -> 62°C (15s) -> 72°C (1min)]x32 -> 72°C (5min)
ced-2(jez50) detection	95°C (5min) -> [95°C (10s) -> 62.5°C (15s) -> 72°C (1min30s)]x32 -> 72°C (5min)
CRKL site-directed mutagenesis	92°C (30s) -> [92°C (30s) -> 65-68°C (1min) -> 68°C (8min)]x18 -> 68°C (10min)
Annealing temperatures: R21A – 68°C, R39A – 67°C, W160L – 68°C, Y207F – 65°C, W275L – 65°C.	

CRISPR/Cas9 guide RNA sequences

Purpose	Sequence
egl-15 PVNLPSE targeting	ACGTTTTGATCTACCTCTGA **TGG**
egl-15 PPNDP targeting	GTAGAGATGATTCGGATCAT**TGG**
egl-15 PRNPLP targeting	TTTGTCGGCAACGGGTTGCG **AGG**
ced-2 pre-SH3N targeting	CCGACGACAACCTCAATGAT **AGG**
ced-2 post-SH3N targeting	AAACTATGTACAAATTCAGA **TGG**
ced-2 pre-SH3C targeting	AACGTCACTTTGGCTTTTGC **TGG**
ced-2 post-SH3C targeting	ACGATTTTATTCGCTGACGG **CGG**

Proteins expressed and purified

Protein	Tag	Plasmid origin
CRKL 2xSH3	GST	Addgene (#36403)
FGFR2 C58	GST	Dr Chi-Chuan Lin, University of Leeds
FGFR2 C24	GST	Dr Chi-Chuan Lin, University of Leeds
FGFR2 C34	GST	Dr Chi-Chuan Lin, University of Leeds
CRKL	GST	Addgene (#36400)

Peptide sequences

FGFR2-C58 -

TTNEEYLDLSQPLEQYSPSYPDTRSSCSSGDDSVFSPDPMPYEPCLPQYPHINGSVKT

FGFR2-C24 - TTNEEYLDLSQPLEQYSPSYPDTR

FGFR2-C34 - SSCSSGDDSVFSPDPMPYEPCLPQYPHINGSVKT

EGL-15 pro1 - QKKRKHRPLSA**PVNLPSE**VDQNPQHTICDD

EGL-15 pro2 - TICDDYESNFSVE**PPNDP**NHLYCNDNMLKN

EGL-15 pro3 - YIQSHPLYSKIIIKKNMT**PRNPLP**TKETIV

Plasmids

Plasmid Name	Purpose	Source
pGEX-4T-2	EGL-15 Pro3 cloning; CED-2 cloning	
FGFR2-GFP	FGFR2-GFP expression	Ahmed et al, 2010

Antibodies

Antibody (Primary)	Manufacturer, catalogue number
CED-2	Novus Biologicals, 42090002
GST	CST, #2625
FGFR2 (Bek C-8)	Santa Cruz Bioctechnology, Sc-6930
CRKL (Mouse)	CST, #3182
CRKL (Rabbit)	CST, #38710
Rabbit IgG	CST, #2729
α-Tubulin	CST, #2144
p-Y (pY99)	Santa Cruz Biotechnology, sc-7020
His-tag	CST, #2365
CRKL pY207	CST, #3181
Phospho-FGF Receptor pY653/654	CST, #3471
CRK-II	CST, #3492
RAP1	CST, #8818 (Antibody included in the kit)
RAC1	CST, #8815 (Antibody included in the kit)
Antibody (secondary)	**Manufacturer, catalogue number**
Anti-Rabbit IgG (HRP conjugated)	CST, #7074
Anti-Mouse IgG (HRP conjugated)	CST, #7076

C. elegans mutant detection

Mutant	Detection method
egl-15(jez10)	PCR amplified. MboII restriction enzyme cut.
egl-15(jez8)	PCR amplified. DpnII restriction enzyme cut.
egl-15(jez20)	PCR amplified. BsaHI restriction enzyme cut.
ced-2(jez45)	PCR amplified. Size difference
ced-2(jez50)	PCR amplified. Size difference

C. elegans strains

Strain Name	Genotype	Source
N2 (Bristol)		Caenorhabditis Genetics Center
LAD28	*egl-15(jez10)*	This study
LAD20	*egl-15(jez8)*	This study
LAD23	*egl-15(jez20)*	This study
LAD24	*ced-2(jez45)*	This study
LAD25	*ced-2(jez50)*	This study
LAD26	*ced-2(jez45) IV; egl-15(jez20) X*	This study
LAD27	*ced-2(jez50) IV; egl-15(jez20) X*	This study
CB3257	*ced-2(e1752) IV*	Caenorhabditis Genetics Center
MT4952	*ced-2(n1994) IV*	Caenorhabditis Genetics Center

Appendix B


A
1 MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN
81 MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD
161 VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSPGIH RDSSDRSAWY
241 MGPVSRQEAQ TRLQGQRHGM FLVRDSSTCP GDYVLSVSEN SRVSHYIINS LPNRRFKIGD QEFDHLPALL EFYKIHYLDT
Carbamidomethylation (+57.02)
Deamidation (NQ) (+0.98)
Oxidation (M) (+15.99)
Phosphorylation (STY) (+79.97)


390
397
321 TTLIEPAPRY PSPPMGSVSA PNLPTAEDNL EYVRTLYDFP GNDAEDLPFK KGEILVIIEK PEEQWWSARN KDGRVGMIPV
417 419 421
474
401 PYVEKLVRSS PHGKHGNRNS NSYGIPEPAH AYAQPQTTTP LPAVSGSPGA AITPLPSTQN GPVFAKAIQK RVPCAYDKTA
494 497
505
481 LALEVGDIVK VTRMNINGQW EGEVNGRKGL FPFTHVKIFD PQNPDENE

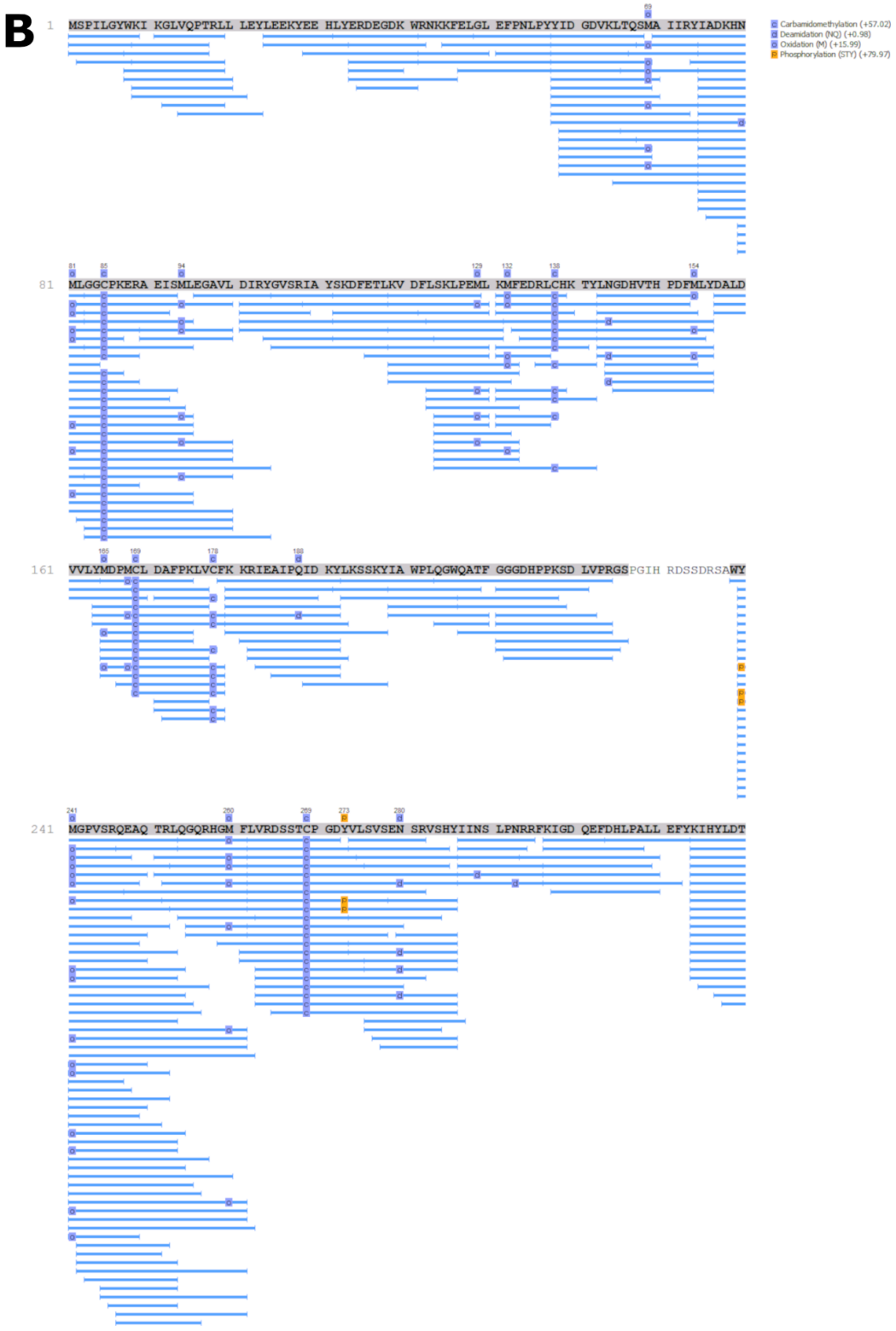

B
1 MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID GDVKLTQSMA IIRYIADKHN
81 MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD
161 VVLYMDPMCL DAFPKLVCFK KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSPGIH RDSSDRSAWY
241 MGPVSRQEAQ TRLQGQRHGM FLVRDSSTCP GDYVLSVSEN SRVSHYIINS LPNRRFKIGD QEFDHLPALL EFYKIHYLDT
Carbamidomethylation (+57.02)
Deamidation (NQ) (+0.98)
Oxidation (M) (+15.99)
Phosphorylation (STY) (+79.97)

Figure B-1. Peptide coverage of GST-CRKL construct. Non-phosphorylated (**A**) and phosphorylated (**B**) GST-CRKL was digested and undergone MS/MS to detect post-translational modification of tyrosine phosphorylation.

Appendix C

List of antibodies chosen from RPPA antibody list

Antibody	Type	Pathway, Function
FAK1	rabbit	Adhesion, VEGF Signalling
FAK1 P Y397	rabbit	Adhesion, VEGF Signalling
ATM/ATR Substrate P S/T	rabbit	Cell Cycle Control, DNA Repair
Aurora A/B/C P T288/T232/T198	rabbit	Cell Cycle
Bad P S112	rabbit	Apoptosis, Akt Signalling
CrkL P Y207	rabbit	Adaptor Proteins
HSP27 (HSPB1) P S78	rabbit	Chaperones, MAPK Signalling, Stress pathway
MEK1/2	rabbit	MAPK Signalling
MEK1/2 P S217/221	rabbit	MAPK Signalling
MNK1 (MKNK) P T197,T202	rabbit	MAPK Signalling, Translational Control
PI3 Kinase p110-alpha	rabbit	Phosphatidylinositol signalling system
PKC (pan) P S660 (beta-2)	rabbit	Calcium, cAMP, Lipid Signalling, PKC Signalling
Bim	rabbit	Apoptosis
Rap1	rabbit	Integrin Signalling, cAMP Signalling
Stat5	rabbit	Cytokine Signalling, Jak/Stat Signalling
Stat5 P Y694	rabbit	Cytokine Signalling, Jak/Stat Signalling
Akt P T308	rabbit	Akt Signalling, Metabolism
S6 Ribosomal protein P S235,S236	rabbit	Lipid Signalling, Metabolism, Translational Control
p44/42 MAPK (ERK1/2)	rabbit	MAPK Signalling
p44/42 MAPK (ERK1/2) P T202/T185,Y204/Y187	rabbit	MAPK Signalling
Src	rabbit	ErbB Signalling, VEGF Signalling, Adhesion
Akt	rabbit	Akt Signalling, Metabolism
Akt P S473	rabbit	Akt Signalling, Metabolism
beta-actin	rabbit	Housekeeping, Cytoskeleton
NFkB p65 S536	rabbit	inflammatory and immune responses
Chk1 P S345	rabbit	Cell Cycle Control
Chk2 P T68	rabbit	Cell Cycle Control
c-Jun P S73	rabbit	MAPK Signalling, Stress pathway
c-Myc	rabbit	MAPK Signalling, Transcription Factors
E-Cadherin	rabbit	Adhesion
Rb	rabbit	Apoptosis, Cell Cycle Control
4E-BP1 P T37,T46	rabbit	Metabolism, Translational Control, mTOR signalling
beta-Catenin	rabbit	Wnt Signalling
beta-Catenin P S33,S37,T41	rabbit	Wnt Signalling

PTEN	rabbit	Akt Signalling, Lipid Signalling, Metabolism
Ubiquitin (P4D1)	mouseIgG1	Protein Degradation, Housekeeping
PTEN P S380,T382,T383	rabbit	Tumor Suppressor, Phospho-Inositol Signalling, Akt Signalling
Cyclin D1	mouseIgG2a	Cell Cycle Control
SAPK/JNK (JNK2)	rabbit	MAPK Signalling, Stress pathway
p70 S6 Kinase P T421,S424	rabbit	Lipid Signalling, Metabolism, Translational Control
LKB1	rabbit	mTOR Signalling
GSK-3-alpha/beta P S21/S9	rabbit	Akt Signalling, Metabolism, Wnt Signalling, Hedgehog Signalling
Stat6 P Y641	rabbit	Cytokine Signalling, Jak/Stat Signalling
p53 P S15	rabbit	Apoptosis, Cell Cycle Control
p38 MAPK PT180,Y182	rabbit	MAPK Signalling, Stress pathway
p38 MAPK	rabbit	MAPK Signalling, Stress pathway
mTOR P S2448	rabbit	mTOR Signalling, Translational Control, Metabolism
mTOR	rabbit	mTOR Signalling, Translational Control, Metabolism
Stat1 P S727	rabbit	Cytokine Signalling, Jak/Stat Signalling
p21 CIP/WAF1	mouseIgG2a	Cell Cycle Control
Raf P S259	rabbit	MAPK Signalling, Ras Signalling
PLC-gamma1 P Y783	rabbit	Calcium, cAMP, Lipid Signalling
PDK-1 P S241	rabbit	Akt Signalling, Lipid Signalling, Metabolism
p90 S6 kinase (Rsk1-3) P T359,S363	rabbit	MAPK Signalling
PDK-1	rabbit	Akt Signalling, Lipid Signalling, Metabolism
p70 S6 Kinase	rabbit	Lipid Signalling, Metabolism, Translational Control
c-Myc P T58,S62	rabbit	MAPK Signalling, Transcription Factors
SAPK/JNK P T183,Y185	rabbit	MAPK Signalling, Stress pathway
Rb P S807,S811	rabbit	Apoptosis, Cell Cycle Control
Rb P S780	rabbit	Apoptosis, Cell Cycle Control
Raf P S338	rabbit	MAPK Signalling, Ras Signalling
Stat6	rabbit	Cytokine Signalling, Jak/Stat Signalling
Stat1 P Y701	rabbit	Cytokine Signalling, Jak/Stat Signalling
Src (family) P Y416	rabbit	ErbB Signalling, VEGF Signalling, Adhesion

Smad2/3 P S465/S423,S467/S425	rabbit	cell growth, apoptosis, morphogenesis, development and immune responses
Smad1/5 P S463/S465	rabbit	cell growth, apoptosis, morphogenesis, development and immune responses
Cyclin D1 P T286	rabbit	Cell Cycle Control
AMPK alpha	rabbit	Metabolism
AMPK alpha P T172	rabbit	Metabolism
Bcl-xl	rabbit	Apoptosis
Bid	rabbit	Apoptosis
Bim P S69	rabbit	Apoptosis
c-Jun N-term	rabbit	MAPK Signalling, Stress pathway
p53	rabbit	Apoptosis, Cell Cycle Control
GSK-3-beta P S9	rabbit	Akt Signalling, Metabolism, Wnt Signalling, Hedgehog Signalling
IRS-1	rabbit	Metabolism, Insulin Signalling
GSK-3-beta	rabbit	Akt Signalling, Metabolism, Wnt Signalling, Hedgehog Signalling
CrkL	mouseIgG1	Adaptor Proteins
HSP27 (HSPB1)	mouseIgG1	Chaperones, MAPK Signalling, Stress pathway
PKC-alpha	mouseIgG2b	Calcium, cAMP, Lipid Signalling, PKC Signalling
Ras	mouseIgG1	MAPK Signalling, Ras Pathway, GNBPs
Stat1	mouseIgG1	Cytokine Signalling, Jak/Stat Signalling
IRS-1 P S636/639	rabbit	Metabolism, Insulin Signalling
PLC-gamma1	rabbit	Calcium, cAMP, Lipid Signalling, PKC Signalling
SHP2 P Y542	rabbit	Tyrosine Phosphatases
Stat3	rabbit	Cytokine Signalling, Jak/Stat Signalling
XIAP	rabbit	Apoptosis
NFkB p105/p50	rabbit	inflammatory and immune responses
Calmodulin	mouseIgG1	inflammation, apoptosis, muscle contraction and intracellular movement
Profilin (C56B8)	rabbit	actin binding proteins, cell motility
ILK1 (4G9)	rabbit	cell adhesion, migration
Grb2	rabbit	adapter protein, Ras signalling
Cool1/betaPix	rabbit	focal adhesions
B-Raf [EP152Y]	rabbit	MAPK signalling
HSF1 [EP1711Y] P S303/307	rabbit	Heat shock factor, transcriptional activator
YAP1 [EP1674Y]	rabbit	transcriptional co-activator, Hippo signalling
MMP21 [EP1277Y]	rabbit	matrix metalloprotease
HSF1[EP1710Y]	rabbit	Heat shock factor, transcriptional activator

Histone H2A.X P S139 [EP854(2)Y]	rabbit	cell cycle, DNA Damage repair
Stat3 P Y705	rabbit	Cytokine Signalling, Jak/Stat Signalling
RhoA (67BC)	rabbit	cell migration, adhesion, proliferation and differentiation
Shc P Y239/40	rabbit	scaffold protein, MAPK Signalling
FOXO3a (75D8)	rabbit	apoptosis, Notch signalling
Histone H2A.X	rabbit	Chromatin Regulation
FOXO1 (C29H4)	rabbit	Cell Cycle regulation, (maybe apoptosis), regulates adipogenesis
Cortactin (H222)	rabbit	coordinate actin reorganization during cell movement
YAP P S127	rabbit	transcriptional co-activator, Hippo signalling
Stat3 P S727	rabbit	Cytokine Signalling, Jak/Stat Signalling
Akt substrate P (RXXS/T) (110B7E)	rabbit	Akt Signalling, Lipid Signalling, Metabolism
PKA substrate P (RRXS/T) (100G7E)	rabbit	cAMP Signalling
c-Abl P Y245	rabbit	cell proliferation, differentiation, apoptosis, cell adhesion, and stress responses
c-Abl P Y412 (247C7)	rabbit	cell proliferation, differentiation, apoptosis, cell adhesion, and stress responses
c-Abl	rabbit	Tyrosine Kinases
beta-Tubulin	rabbit	Housekeeping, Cytoskeleton
PKC-alpha P T638	rabbit	Calcium, cAMP, Lipid Signalling, PKC Signalling
GAPDH	mouseIgG2b	Housekeeping
Raf1 (C-12)	rabbit	MAPK Signalling, Ras Signalling
p21 CIP/WAF1 p T145	rabbit	Cell Cycle Control
p90 S6 kinase (Rsk1-3)	rabbit	MAPK Signalling
mTOR P S2481	rabbit	mTOR Signalling, Translational Control, Metabolism

Appendix D

Figure D-1. Alignment of FGFR2 and EGL-15 C-termini. Full-length FGFR2 and EGL-15 protein sequences were aligned, while the image was taken for the C-terminal portion. FGFR2 residues representing the “C58” C-terminal tail used in this work is shown by a black line above, with proline residues highlighted with a red star. EGL-15 proline-rich motifs used in this work are highlighted with a green line, while another proline-rich motif not used in this study is highlighted with a gray line.

Appendix E

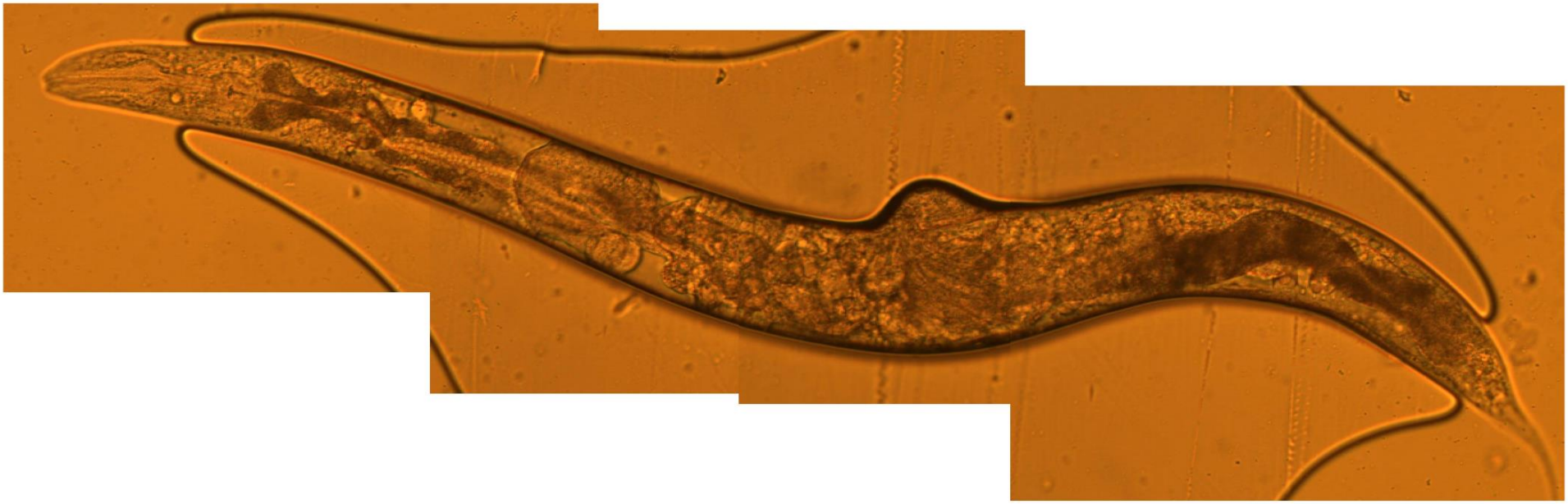


Figure E-1. Rare phenotype of CED-2(ΔSH3C);EGL-15(ΔPRNPLP). This animal was found during a brood size assay. It was capable of feeding, and moved normally, but did not lay any eggs.