Direct interaction between GRB2 and AGO2 regulates microRNA expression

Amy Kate Stainthorp

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Molecular and Cellular Biology

April, 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

Firstly I thank my supervisors, Prof. John Ladbury and Prof. Adrian Whitehouse, for their consistent support. You always helped me find a way forwards and motivated me when I felt things were not going well.

Additionally, I express my gratitude to Prof. Richard Bayliss, for continuing to support my academic development through assessing my progress reviews, and Dr Niamh Forde, for providing advice on my RNA work.

I would like to say a huge thank you to Dr Chi-Chuan Lin. Your help during my PhD has been invaluable, from teaching me techniques, sharing reagents, advising me on problems and presentations, answering my endless questions and taking me for some very necessary coffees. Thank you for always being there for me and always believing in me.

Thanks to all the other members of the Ladbury lab. I loved working with you all and feel privileged to have got to know you all during this time. Thank you also to all members of the Whitehouse lab who taught me new techniques.

I express my appreciation to the Wellcome Trust for funding this PhD and introducing me to an incredible community of researchers. Particular thanks to Asif Fazal, David Nicholson, James Murphy, Jonathan Foster and Sophie Moul; I'm so grateful to have shared this experience with you.

I wish to thank all the facilities managers for their assistance; SOT for keeping the LIGHT building running smoothly and doing a fantastic job to get us back in after lockdown; and the purchasing team for resolving any problems efficiently.

Finally, thank you to my parents for taking care of me from afar and listening to all my troubles. You helped more than you know.

Abstract

microRNAs (miRNAs) post-transcriptionally regulate gene expression to control vital cellular processes. Consequently, miRNAs are differentially expressed in diseases such as cancer, and such changes are often attributed to transcriptional control. However, miRNAs are produced by a multistep pathway and increasing studies show that miRNAs are subject to extensive post-transcriptional regulation. Biogenesis culminates in loading of a miRNA onto Argonaute 1-4 by DICER1, which can be disrupted by Argonaute 2 (AGO2) phosphorylation at tyrosine 393.

In this study, the role of GRB2 in regulation of miRNA-loading was investigated. FGFR2-dependent and independent mechanisms were explored using a range of cell-based assays and biophysical techniques. GRB2-controlled miRNAs were identified by small RNA sequencing of GRB2-deficient cells. Perturbation of these miRNAs in cancer was explored by a mixture of quantitative PCR and bioinformatics approaches. Finally, reciprocal regulation of GRB2 expression by the GRB2-dependent miRNAs was studied using luciferase reporter assay.

GRB2 was found to regulate miRNA-loading independently of FGFR2. GRB2 associated with the AGO2-DICER1 complex through an interaction mediated by GRB2 N-terminal SH3 domain and a proline-rich motif in AGO2 PAZ domain. GRB2 depletion dysregulated miRNA expression by two different mechanisms: 1) loading of let-7 family miRNAs was enhanced, and 2) transcription of other miRNAs, including the 17~92 cluster and mir-221 family, was diminished. Preliminary data suggested let-7 loading may be similarly deregulated in the A498 kidney carcinoma cell line. Finally, a miR-221-5p mimic targeted an 8mer site in the *GRB2* 3'UTR and downregulated GRB2 expression.

Overall, this study identifies a previously unknown role of GRB2 in regulation of miRNA biogenesis. GRB2-mediated inhibition of let-7 loading onto AGO2 may control oncogene expression and have implications in tumorigenesis. Additionally, miR-221-5p may silence GRB2.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

microRNAs (miRNAs) are small RNAs which post-transcriptionally regulate gene expression. *In silico* tools have predicted that they bind over half of the human transcriptome (Friedman et al., 2009) and consequently they help to control most cellular processes. Accordingly, aberrant miRNA expression is seen in many human diseases (Lu et al., 2005; Carè et al., 2007; Chen et al., 2008; Asrih and Steffens, 2013; Swarbrick et al., 2019). In particular, microarray analyses have demonstrated that miRNAs are dysregulated in most human cancers (Volinia et al., 2006). Many miRNAs have been assigned roles as oncogenic miRNAs (oncomirs) or tumour suppressor miRNAs, due to their respective targeting of tumour suppressor genes and oncogenes (Calin and Croce, 2006). Changes in miRNA expression are majorly attributed to a difference in transcriptional regulation (Lin and Gregory, 2015; Ali Syeda et al., 2020). However, the multi-step pathway of miRNA biogenesis allows for regulation at various stages and numerous mechanisms of post-transcriptional regulation have been described (Treiber et al., 2019). The final step in this pathway is ‘loading’, a process by which the miRNA is transferred onto the core miRNA-induced-silencing protein, Argonaute 1-4, by the miRNA-processing enzyme, DICER1 (1.3 miRNA biogenesis). Thus, the active miRNA-induced silencing complex (miRISC) is generated. Although most biogenesis steps occur at the rate of transcription, time-resolved small RNA sequencing experiments have shown that miRNA-loading is slower and therefore forms a kinetic bottleneck in the pathway (Reichholf et al., 2019). miRNA-loading therefore presents a potential opportunity whereby miRNA biogenesis may be post-transcriptionally controlled.

1.1 miRNA function

miRNAs were initially discovered in *Caenorhabditis elegans* through the complementarity between sequences in the *lin-14* 3' untranslated region (UTR) and the 22 nucleotide RNA *lin-4* (Lee et al., 1993; Wightman et al., 1993). Two similar classes of 20-25 nucleotide RNAs were described: miRNAs, which were understood to regulate endogenous genes, and small interfering RNAs (siRNAs), which had been identified to act as a defence mechanism in response to viral infection (Hamilton and Baulcombe, 1999). In the subsequent years, northern blotting experiments were used to identify miRNAs in various other species, including both human and plant cells, and some were found to be conserved between vertebrate and invertebrate species (Pasquinelli et al., 2000; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Reinhart et al., 2002). miRNAs are now known to have diverse and vital functions in most species. Disruption of their expression is consequently associated with cardiovascular disease, immune disorders, neuromuscular disease and cancer (Bartel, 2018).

Both siRNAs are miRNAs are incorporated into the RISC to induce specific post-translational regulation of gene expression. The minimal RISC is comprised of a small RNA bound to an AGO protein; the small RNA acts as a guide and binds complementary sequences in the messenger RNA (mRNA) of its target gene. siRNAs have exact complementarity with their targets and thus trigger mRNA cleavage (Hamilton and Baulcombe, 1999; Hammond et al., 2000; Zamore et al., 2000; Yang et al., 2000). In contrast, miRNAs have some sequence mismatch and consequently mediate gene silencing through translational repression (Hutvágner and Zamore, 2002) and mRNA degradation

(Bagga et al., 2005). However, these specific functions appear to be entirely sequence dependent: *in vitro* assays demonstrated that targeting of a synthetic miRNA with perfect complementarity to endogenous mRNA leads to target cleavage (Hutvágner and Zamore, 2002).

As miRNA target sites are not perfectly complementary, miRNAs are able to target multiple genes. Many factors influence the ability of a miRNA to bind its target. The miRNA seed sequence spans nucleotides (nt) 2-8 and, as a minimum, base pairing between six consecutive nt to the target RNA is required for silencing (Figure 1.1.1) (Lewis et al., 2003; Poy et al., 2004). In addition, factors, such as secondary structure, evolutionary conservation, AU-rich flanking sequences, location within the target RNA, propensity for binding of multiple miRNAs and base pairing towards the 3’ end of the miRNA, may influence miRNA targeting (Lewis et al., 2005; Long et al., 2007; Grimson et al., 2007; Broughton et al., 2016). Non-canonical binding of miRNAs to target sites without a perfect seed match has also been identified (Yekta et al., 2004; Chi et al., 2012; Loeb et al., 2012). Additionally, while miRNAs usually bind their targets in the 3’ UTR, examples of targeting the 5’ UTR and coding regions have been described (Lewis et al., 2005; Forman et al., 2008; Xu et al., 2014).

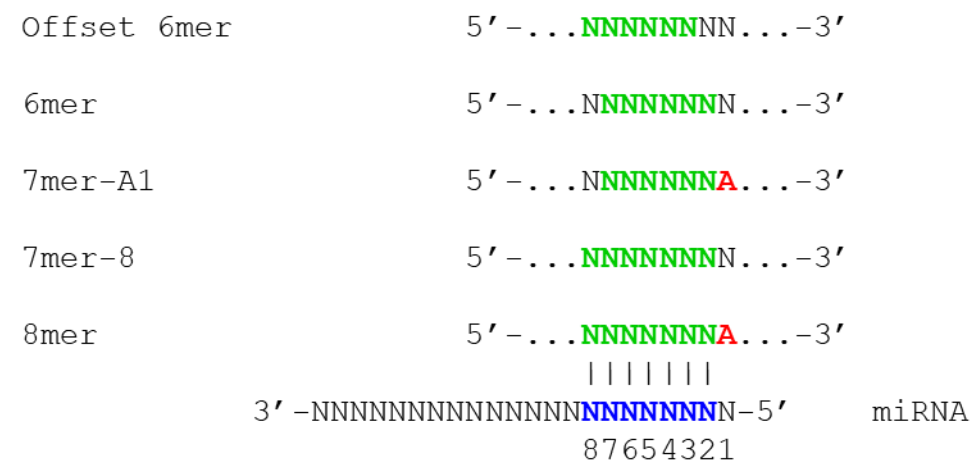


Figure 1.1.1 Canonical seed pairing of a miRNA to its target

miRNA seed (nucleotides 2-8) highlighted in blue. Complementary nucleotides in the mRNA target site shown in green. 3' adenosine required for 8mer and 7mer-A1 targeting sites in red. (Figure adapted from Friedman et al., 2009).

1.2 AGO2

Of the four AGO proteins present in the human genome (AGO1-4), cleavage assays using purified RISC components originally identified AGO2 as the only one which has catalytic slicer activity, allowing it to cleave target RNAs (Meister et al., 2004; Liu et al., 2004). Catalytic activity has subsequently been demonstrated for AGO3, but the enzyme has specific substrate requirements (Park et al., 2017). Mutational analyses revealed that a ‘catalytic tetrad’ is required for slicing activity, which is present only in these two proteins (Figure 1.2.1) (Nakanishi et al., 2012). However, human AGO1-4 have high overall sequence similarity (Figure 1.2.1) and as target slicing is not normally needed for miRNA-mediated gene silencing, all human AGO proteins are thought to function similarly in this pathway (Su et al., 2009). However, unlike the other AGO proteins, knockout of AGO2 in mice is embryonic lethal, suggesting this protein has a non-redundant function which is vital for development (Morita et al., 2007).

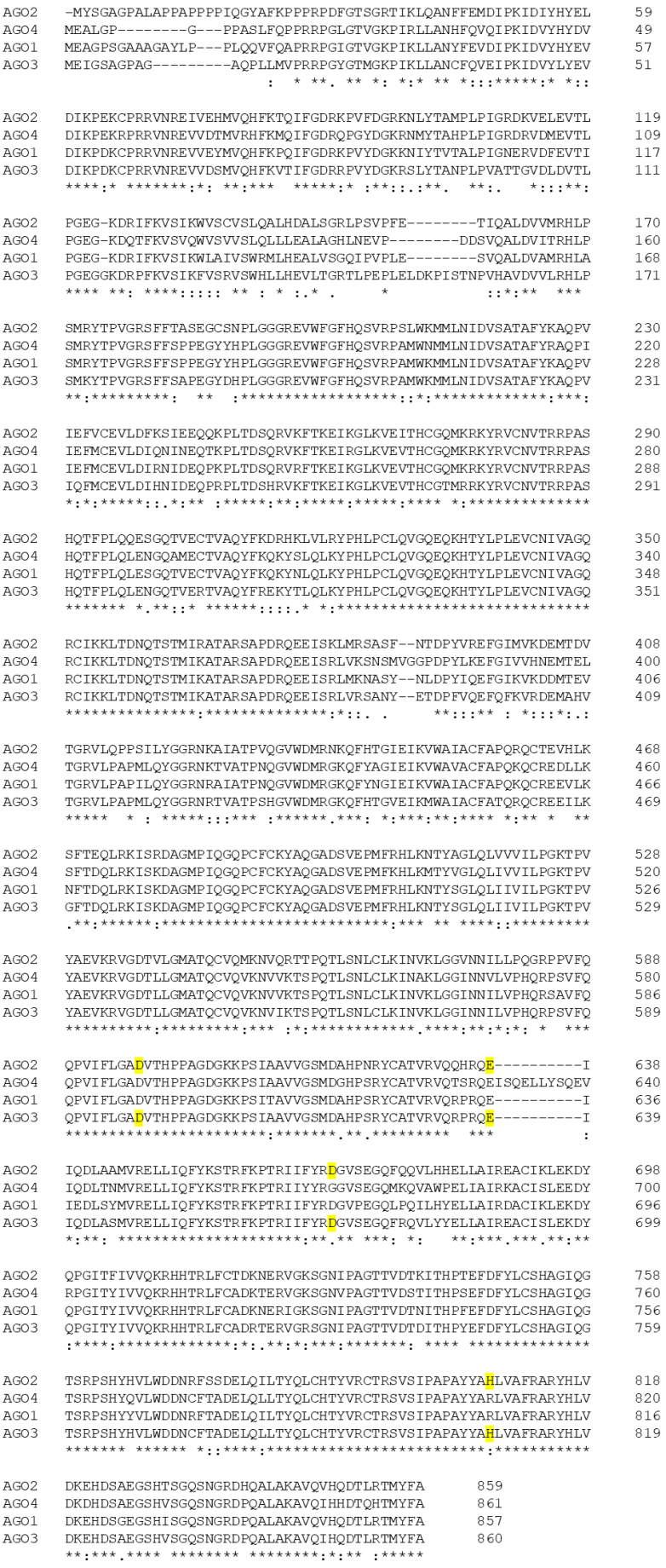


Figure 1.2.1 Alignments of human AGO1-4

Alignments of human proteins AGO1-4. Catalytic residues present in AGO2 and AGO3 are highlighted in yellow. Generated using Clustal Omega (Sievers et al., 2011; EMBL-EBI, 2021)

AGO proteins have four main domains: the amino-terminal (N-terminal) domain, the Piwi Argonaut and Zwille (PAZ) domain, the middle (MID) domain and the P-body-induced wimpy testes (PIWI) domain (Figure 1.2.2 A). The N-terminal-PAZ domains and MID-PIWI domains form two lobes, across which the miRNA is bound (Figure 1.2.2 B). Crystal structures of the PIWI domain have revealed that its fold is similar to that of RNase H enzymes and thus this domain confers AGO2 its slicer activity (Song et al., 2004). Additionally, PIWI is responsible for many protein interactions with AGO2. For example, NMR experiments have demonstrated that PIWI recruits the glycine-tryptophan repeat (GW)-containing effector proteins, trinucleotide repeat containing 6 (TRNC6) (Schirle and MacRae, 2012; Pfaff et al., 2013). The PAZ domain forms an oligonucleotide/oligosaccharide-binding (OB) fold (Yan et al., 2003) and anchors the 3’ end of the miRNA (Song et al., 2004). Similarly, the MID domain binds the 5’ phosphate and adopts a Rossman fold (Boland et al., 2010). Finally, native gel analyses of siRNA structure suggested that the N-terminal domain is required for unwinding of the miRNA duplex (Kwak and Tomari, 2012).

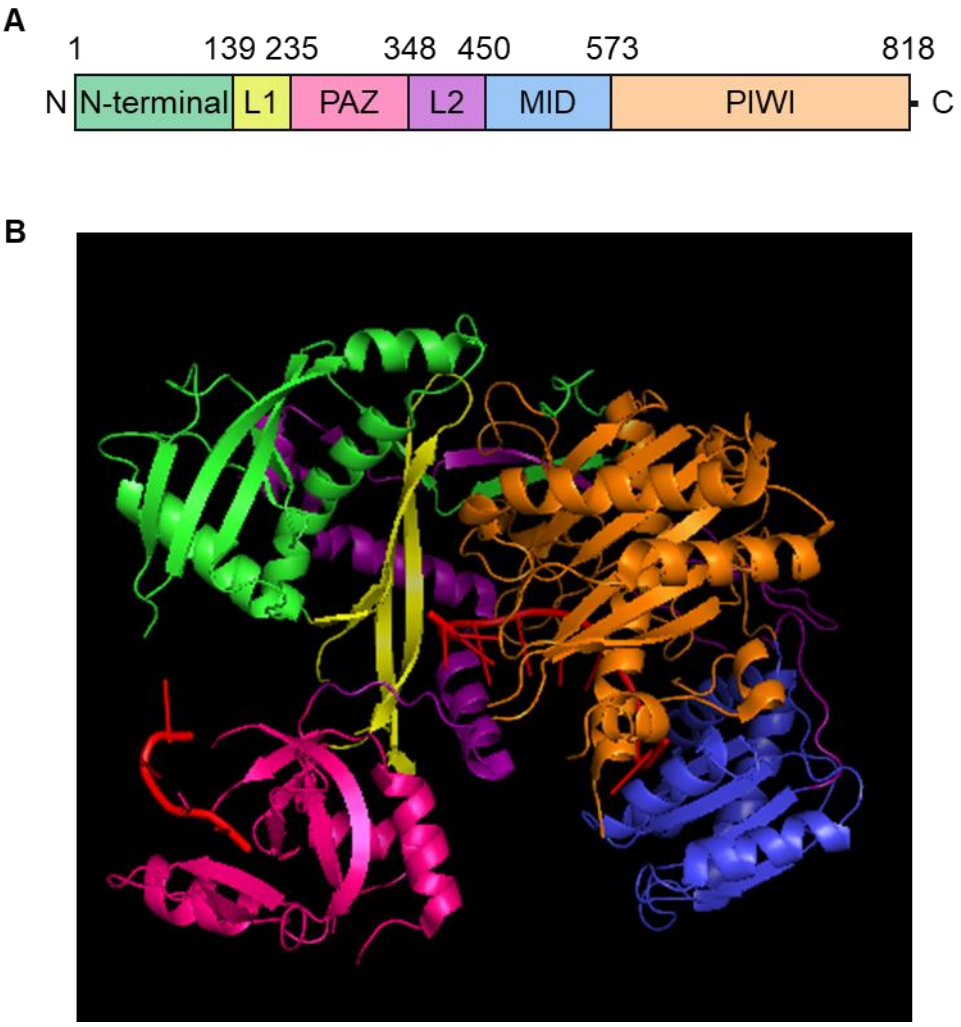


Figure 1.2.2 AGO2 structure

(A) Domain and **(B)** crystal structure of AGO2 bound to a miRNA. N-terminal domain in green, PAZ domain in pink, MID domain in blue, PIWI domain in orange, linker 1 (L1) in yellow, linker 2 (L2) in purple and miRNA in red. (Adapted from Elkayam et al., 2012, PDB: 4F3T.)

Upon binding of AGO2-miRNA to the target mRNA, AGO2 associates with TNRC6 proteins. Mass spectrometric analyses have demonstrated that TNRC6 proteins act as scaffolds and recruit mediators of translational repression and mRNA degradation (Landthaler et al., 2008). The poly(A)-nuclease (PAN2-PAN3) and carbon catabolite repression 4-like-negative regulator of transcription (CCR4-NOT) deadenylation complexes, the dead box protein 6 (DDX6) decapping activator and mRNA decapping enzymes 1 and 2 (DCP1/2) are all recruited to form a huge miRISC (Behm-Ansmant et al., 2006; Chen et al., 2009; Nishihara et al., 2013; Chen et al., 2014). Fluorescence microscopy of mammalian cells has suggested that TNRC6 also localises AGO2 to processing bodies (P bodies), which are the sites of mRNA silencing (Sen and Blau, 2005; Liu et al., 2005).

1.3 miRNA biogenesis

Most miRNAs are produced through a canonical biogenesis pathway. In this pathway multiple enzymes are required to process the initial miRNA transcript and thus generate the mature miRNA. These enzymes are highly conserved in vertebrates and invertebrates. For example, *Drosophila melanogaster* and *Caenorhabditis elegans* express AGO2 homologues DmAGO1 and CeALG-1/2 (Grishok et al., 2001; Okamura et al., 2004). Consequently, miRNA biogenesis is frequently studied in these systems.

The miRNA biogenesis pathway begins in the nucleus. Chromatin immunoprecipitation has revealed that RNA polymerase II transcribes the primary miRNA (pri-miRNA) (Figure 1.3.1) from genes which may be located both in the introns of coding genes or at distinct loci (Lee et al., 2004). miRNA genes may code for an individual miRNA or for a cluster of miRNAs, leading to

transcription of a polycistronic transcript (Lee et al., 2002). The pri-miRNA is hundreds of nt long with a central hairpin loop structure and is both capped and polyadenylated (Cai et al., 2004). Cell-based assays have suggested that polycistronic miRNAs are processed and miRNAs encoded in introns are removed co-transcriptionally, prior to splicing (Morlando et al., 2008). The DROSHA/DiGeorge syndrome critical region 8 (DGCR8) microprocessor complex precisely cleaves both strands of the pri-miRNA in the hairpin loop region (Figure 1.3.1), as shown by *in vitro* processing assays (Lee et al., 2003; Han et al., 2004). A precursor miRNA (pre-miRNA) with a 2 nt 3' overhang is thus generated. Non-canonical, microprocessor-independent mechanisms of miRNA biogenesis have also been described. These miRNAs may be generated from introns (termed mirtrons) (Ruby et al., 2007), from tRNAs, from snoRNAs (Babiarz et al., 2008) and from capped hairpin transcripts (Xie et al., 2013).

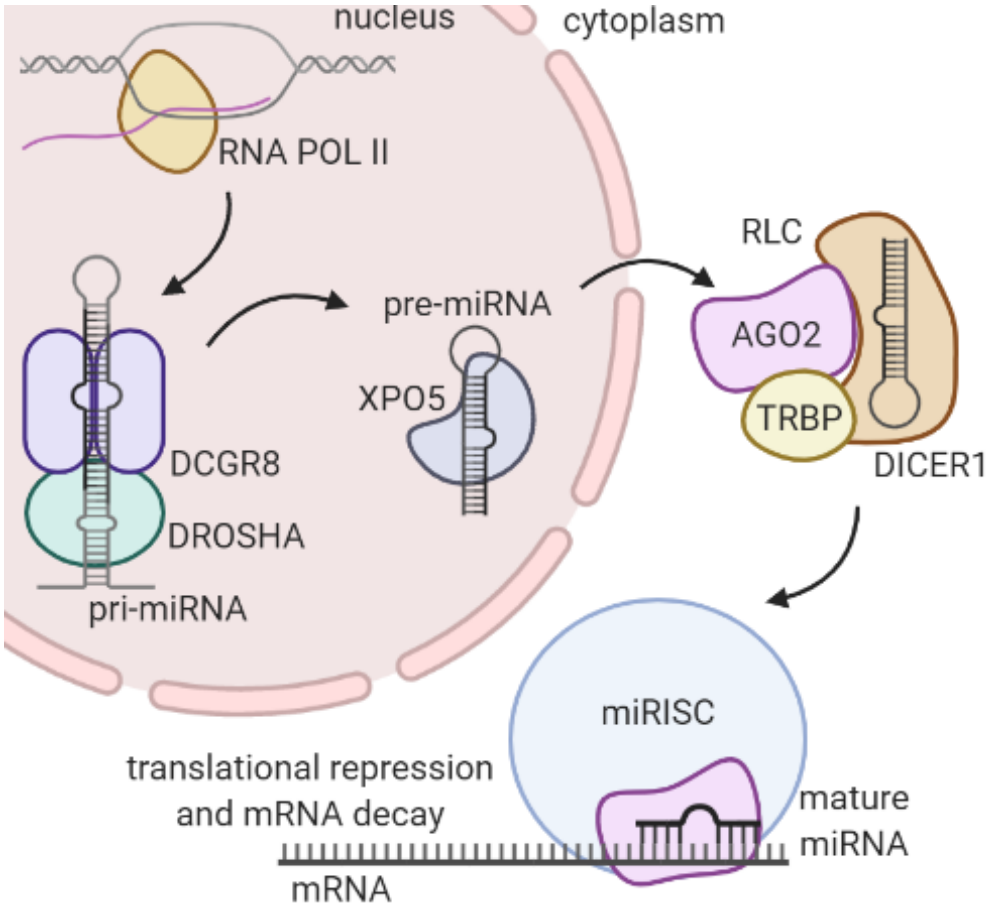


Figure 1.3.1 miRNA biogenesis pathway

miRNA biogenesis starts in the nucleus with transcription of primary miRNA (pri-miRNA) by RNA polymerase II (RNA POL II). The DROSHA/DGCR8 complex then processes the pri-miRNA, which is exported from the nucleus to the cytoplasm by XPO5. DICER1 cleaves the pre-miRNA to a mature miRNA duplex and binds AGO2 and TRBP to form the RISC loading complex (RLC). Upon loading of the miRNA duplex AGO2 ejects the passenger strand and the bound miRNA acts as a guide, leading AGO2 to target mRNAs. AGO2 recruits other proteins to form the miRNA induced silencing complex (miRISC) which facilitates translational repression and mRNA degradation. Created with BioRender.com.

Following cleavage, correctly processed pre-miRNAs are exported to the cell cytoplasm. Translocation occurs via a direct, and Ran-guanosine triphosphate (GTP) dependent, interaction with Exportin 5 (XPO5) (Figure 1.3.1) (Yi et al., 2003; Lund et al., 2004). Interestingly, although other canonical export routes are not known, knockout of *XPO5* only slightly reduces expression of canonical miRNAs, suggesting other export routes exist (Kim et al., 2016).

In the cytoplasm, pre-miRNAs are processed by the enzyme DICER1 to produce a mature miRNA duplex (Figure 1.3.1) (Grishok et al., 2001). DICER1 is a 220 kilodalton (kDa) enzyme, which possesses an RNA helicase domain, a domain of unknown function (DUF283), platform domain, PAZ domain, two RNase III domains and a double stranded RNA-binding domain (dsRBD) (Figure 1.3.2). A crystal structure of full length DICER1 revealed that the PAZ domain recognises the 3’ 2 nt overhang of the miRNA and acts as a molecular ruler to produce a duplex of ~22 nt (MacRae et al., 2006). Simultaneously, the 5’ phosphate of the opposite strand is secured by the platform domain (Tian et al., 2014). Comparison of EM structures demonstrated that upon pre-miRNA binding, DICER1 adopts an active state with the RNase III domains close to the RNA duplex (Taylor et al., 2013). The RNase III domains form a dimer and each catalyse the cleavage of an RNA strand to remove the apical loop (Zhang et al., 2004). The resulting mature miRNA duplex is then moved away from DICER1 (Tian et al., 2014), allowing AGO2 to bind DICER1 RNase III domain via its PIWI domain (Tahbaz et al., 2004), and the duplex is transferred to AGO2.

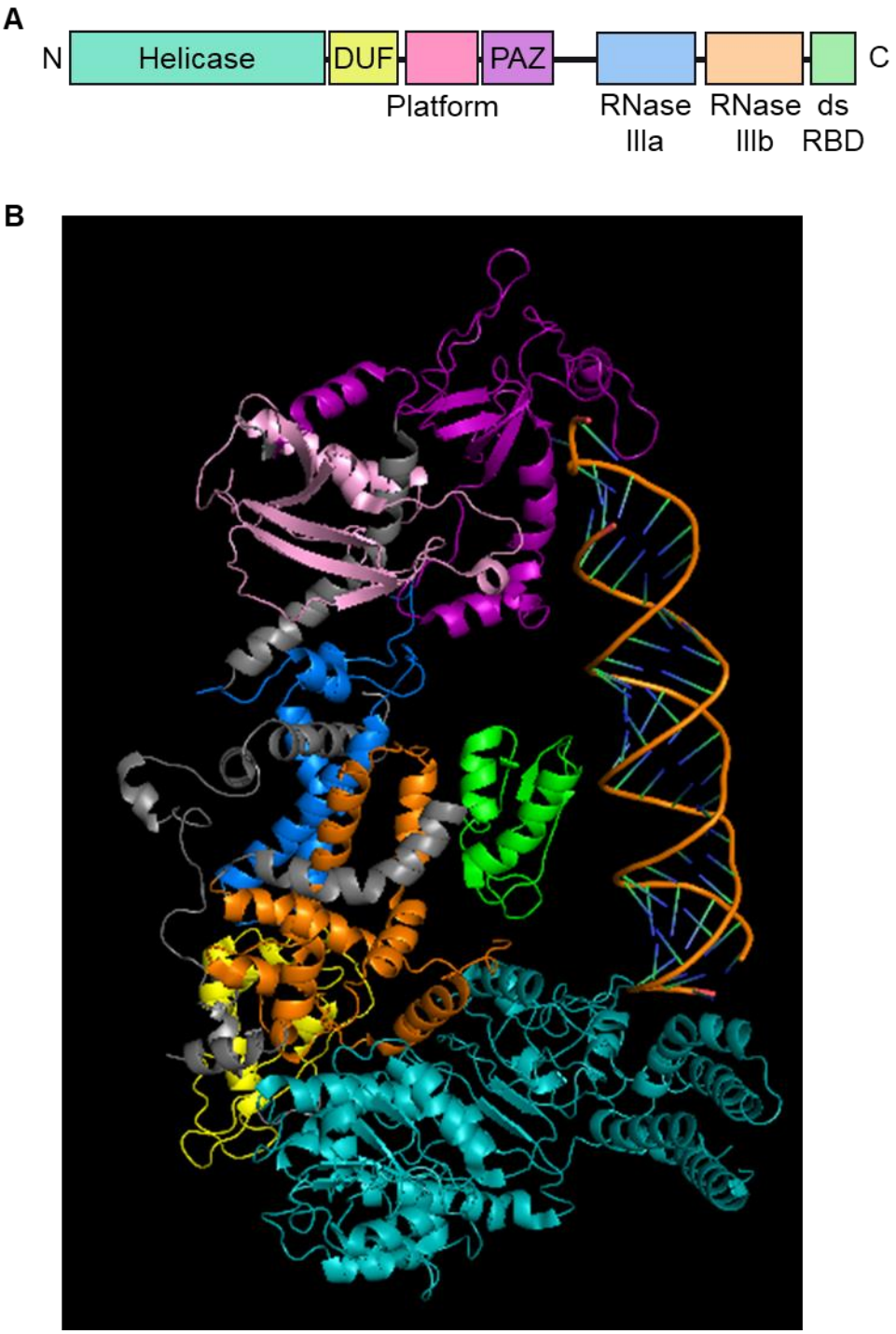


Figure 1.3.2 Structure of human DICER1

(A) Domain and **(B)** EM structures of human DICER1. Helicase domain in teal. Domain of unknown function (DUF283) domain in yellow. Platform domain in pink. PAZ domain in purple. RNase IIIa and b domains in blue and orange. Double stranded RNA-binding domain (dsRBD) in green. Pre-let-7 in orange. (PDB: 5ZAL, Z. Liu *et al.*, 2018)

Passing of the miRNA duplex from DICER1 to AGO2 triggers unwinding of the duplex and ejection of the passenger miRNA strand (Figure 1.3.1). Single molecule Förster resonance energy transfer (FRET) experiments have revealed conformational changes in *D. melanogaster* AGO2 during small RNA-loading: prior to loading, AGO2 is in an apo conformation, but upon duplex binding it transitions to a pre-miRISC state. After passenger strand release it forms an active conformation which is capable of gene silencing (Tsuboyama et al., 2018). The mechanism by which AGO2 unwinds the miRNA is not fully known but is thought to be through the N-terminal domain acting as a wedge to drive separation of the two strands (Kwak and Tomari, 2012). As AGO2 does not require ATP for miRNA unwinding, it is hypothesised that the energy for unwinding may be provided by changes in conformation bringing AGO2 from a high energy state to a low energy state (Treiber et al., 2019). miRNA strand selection is usually determined by the 5' base pairing, as NMR titration experiments indicated that the MID domain preferentially binds less thermodynamically stable 5' ends (Frank et al., 2010). Considering this 'rule', it was originally thought only one strand was loaded onto AGO2. Indeed, one strand is usually selected as the guide and forms the dominant species in the cell. However, northern blotting of mature miRNA species *D.melanogaster* and RNA sequencing in human cells suggested that either strand may be loaded and that the dominant species may change via 'arm switching' (Okamura et al., 2008; Guo et al., 2014).

Release of the passenger strand into the cytoplasm is slicer-independent (Gu et al., 2011; Park and Shin, 2015). It is thought that the endoribonuclease component 3 promoter of RISC (C3PO) is responsible for degradation of the ejected miRNA strand (Liu et al., 2009).

Loading of miRNAs onto AGO2 is facilitated by multiple other proteins. Firstly, western blotting and *in vitro* assays have demonstrated that trans-activation-responsive RNA-binding protein (TRBP) is an integral part of the miRISC-loading machinery (Gregory et al., 2005). DICER1, TRBP and AGO2 form a stable miRISC-loading complex (RLC) which is sufficient for correct loading of miRNAs onto AGO2 *in vitro* (Naruse et al., 2018). TRBP binds DICER1 directly to both promote proper processing of the pre-miRNA and assist miRNA handover to AGO2 (Chendrimada et al., 2005; Wang et al., 2009; Lee and Doudna, 2012). The role of a similar DICER1-binding protein, protein activator of PKR (PACT), is not yet fully characterised (Lee et al., 2006). However, assays using reconstituted RISC suggested that TRBP and PACT may help produce differently sized miRNAs, termed isomiRs (H.Y. Lee et al., 2013).

The heat shock protein 90 (HSP90) and heat shock cognate protein 70 (HSC70) chaperone systems are also highly involved in miRNA-loading. GST-pulldown experiments have demonstrated that HSC70, HSP90, heat shock protein 40 (HSP40) and the co-chaperones HSC70/HSP90 organising protein homologue (HOP) and P23 all associate with the N-terminus of AGO2. HSC70/HSP40 also bind the MID-PIWI domains (Tahbaz et al., 2001). Western blotting cell lysates has revealed that HSP90 protects the apo form of AGO2 from proteasomal degradation (Johnston et al., 2010). Additionally, it has been shown that the two chaperone systems promote ATP-dependent miRNA-loading (Kawamata et al., 2009; Iwasaki et al., 2010). In mammals, HSC70/HSP90 can even facilitate *in vitro* loading of siRNA duplexes independently of DICER1 and TRBP (Naruse et al., 2018). Prior to miRNA binding, AGO2 adopts multiple closed conformations. Binding of HSC70 partially opens AGO2; HSP90 then stabilises the open conformation, which is

able to accommodate miRNA duplexes (Tsuboyama et al., 2018). In *D. melanogaster*, chaperones extend the dwell time of DICER1 on AGO2, thus facilitating miRNA transfer (Iwasaki et al., 2015).

Knockdown studies have revealed that DICER1-independent miRNA-loading occurs for only one miRNA: miR-451. Mature miR-451 is generated through slicing by AGO2 PIWI domain (Yang et al., 2010). Similarly, although loading of miR-486 is via DICER1, AGO2 slicer activity is required to eliminate the passenger strand. Consequently, knockout of these two miRNAs mirrors the phenotype seen upon expression of catalytically dead AGO2 (Jee et al., 2018).

1.4 miRNA nomenclature

Specific nomenclature is required to differentiate between the different miRNA transcripts, but may differ slightly between different publications. Here, the miRNA gene is referred to with a lower case ‘r’ and in italics (e.g. *mir-221*). miRNA families are denoted in the same way but not italicised: mir-221. The primary transcript from this gene is represented as pri-mir-221, while the precursor transcript is pre-mir-221 (Lee et al., 2002). In this study, where both primary and precursor transcripts are described, the ‘pre’ prefix is used. The mature miRNA is denoted with a capitalised ‘R’: miR-221. 5p and 3p suffixes (miR-221-5p and miR-221-3p) refer to miRNAs which are produced from the 5p and 3p arms of the hairpin respectively (Griffiths-Jones et al., 2006). Letters indicate closely related sequences, such as with miR-18a and miR-18b. Numbers after the suffix (e.g. miR-19b-1 and miR-19b-2) are used if two different loci produce identical mature miRNAs (Lagos-Quintana et al., 2001).

1.5 Regulation of miRNA-loading independent of AGO2 modification

In addition to chaperones and TRBP, proteins help to regulate loading of specific miRNAs onto AGO2 (Figure 1.5.1). A large class of these proteins are RNA-binding proteins (RBPs), many of which bind the pre-miRNA loop (Piskounova et al., 2008; Newman et al., 2008; Trabucchi et al., 2009; Rau et al., 2011; Kawahara and Mieda-Sato, 2012; Choudhury et al., 2014; Wu et al., 2015). Probably the best characterised RBP regulator of miRNA-loading is LIN28A (Rybak et al., 2008; Heo et al., 2008). RNA affinity purification experiments have shown that LIN28A binds pre-let-7 in the cytoplasm and recruits terminal uridylation transferases (TUTs). TUT4 and TUT7 polyuridylate the pre-miRNAs (Heo et al., 2009; Thornton et al., 2012). Polyuridylation inhibits cleavage by DICER1 and ultimately leads to their degradation (Heo et al., 2008). When LIN28A is not expressed, RNA sequencing and *in vitro* assays have revealed that TUT2/4/7 add a single uridine to the 3’ end of class II let-7 miRNAs, promoting cleavage by DICER1 (Heo et al., 2012). LIN28B is also suggested to mediate polyuridylation of let-7 (Heo et al., 2008), although this is disputed as let-7 suppression does not require TUT4 (Piskounova et al., 2011). Inhibition may instead be via sequestration of pri-miRNA in the nucleus to preclude DROSHA-processing, as demonstrated by RNA immunoprecipitation (RIP) and immunofluorescence studies (Viswanathan et al., 2008; Newman et al., 2008; Piskounova et al., 2011). Two other RBPs, TRIM25 and muscleblind-like protein 1 (MBNL1), also bind pre-miRNAs to positively and negatively regulate LIN28-mediated processing (Rau et al., 2011; Choudhury et al., 2014).

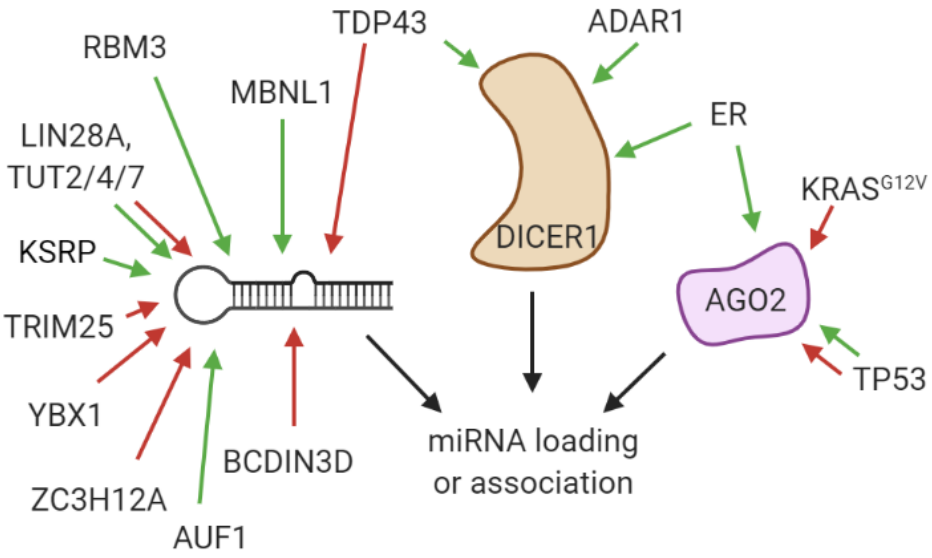


Figure 1.5.1 Schematic of proteins which regulate loading or association of miRNAs with AGO2 in humans

Proteins which bind pre-miRNAs, DICER1 or AGO2 to promote (green) or inhibit (red) miRNA-loading or association with AGO2. Created with BioRender.com.

RBPs may sequester pre-miRNAs from binding DICER1, such as TAR-DNA-binding protein 43 (TDP43) (King et al., 2014) and Y-box binding protein 1 (YBX1). YBX1 binding pre-mir-29b-2 may contribute to cell proliferation in glioblastoma multiforme (Wu et al., 2015). Immunoprecipitation (IP) experiments demonstrated that TDP43 binds the mir-1 family to block recruitment to DICER1 (King et al., 2014). Additionally, TDP43 associates with DICER1 in an RNA-dependent manner to facilitate cleavage of pre-mir-143 and pre-mir-547 (Kawahara and Mieda-Sato, 2012). Similarly, RNA-binding motif 3 (RBM3) (Pilotte et al., 2011) and KH type-splicing regulatory protein (KSRP) (Trabucchi et al., 2009) interact with pre-miRNAs to promote their association with DICER1 and DICER1-mediated processing respectively. Alternatively, single-molecule binding assays have suggested that loading of already processed, mature let-7 miRNA onto AGO2 is mediated by AU-rich-binding factor 1 (AUF1) (Yoon et al., 2015). In contrast to most other RBPs, adenosine deaminase acting on RNA (ADAR1) binds DICER1 directly to promote miRNA processing and loading onto AGO2 in a manner similar to TRBP (Ota et al., 2013).

RBPs also modify pre-miRNAs to inhibit DICER1 processing. The endoribonuclease zinc-finger-CCCH-domain-containing protein 12A (ZC3H12A) cleaves several pre-miRNAs in their apical loop. Analyses of gene expression databases indicated that high ZC3H12A combined with low DICER1 expression may be associated with worse prognosis in various cancers (Suzuki et al., 2011). Similarly, cell-based assays demonstrate BCDIN3-domain-containing protein (BCDIN3D) methylates pre-mir-145 on the 5’ phosphophate to block DICER1 binding, enhancing the tumour phenotype of breast cancer cells (Xhemalce et al., 2012). DICER1-mediated miRNA-loading onto AGO2 is thus mediated majorly by miRNA-binding proteins, rather than proteins which directly

interact with DICER1 or AGO2. These proteins may both enhance or diminish loading.

Few non-RNA binding proteins have been described in the regulation of miRNA-loading. In breast cancer cells, the oestrogen receptor β (ESR2) associates with AGO2, DICER1 and TRBP. RIP-sequencing experiments indicated that formation of this complex enhances loading of specific miRNAs, which may contribute to the less aggressive phenotype of ESR2 positive breast cancers (Tarallo et al., 2017). TP53 associates with AGO2 to both promote and inhibit association of specific miRNAs in response to DNA damage. However, comparison of RIP- and whole-cell-sequencing datasets suggested that this is uncoupled from miRNA biogenesis and does not affect overall miRNA expression. Interestingly, oncogenic mutation of TP53 reduces binding of tumour suppressor let-7 miRNAs (Krell et al., 2016). IP studies have demonstrated that the small GTPase KRAS also binds AGO2, through a direct interaction with AGO2 N-terminal domain. When KRAS harbours the oncogenic G12V mutation, binding inhibits AGO2 unwinding miRNAs. The authors suggest that this mechanism may contribute to an overall suppression of miRNA expression, which has been suggested as a general feature of tumour cells (Lu et al., 2005; Kumar et al., 2007; Shankar et al., 2016). However, the frequent upregulation of oncomirs in human cancer may complicate this interpretation (Hayashita et al., 2005; Calin and Croce, 2006). Furthermore, complete loss of *DICER1* inhibited tumorigenesis in a mouse model of retinoblastoma (Lambertz et al., 2010), demonstrating a vital role for miRNAs in cancer pathogenesis.

1.6 Post-translational modification of AGO2

AGO2 is regulated by a variety of post-translational modifications (PTMs), which modify both its expression and activity (Figure 1.6.1). Argonaute proteins are highly conserved in eukaryotic organisms (Figure 1.6.2). In this way, some PTMs are conserved between species (Golden et al., 2017; Quévillon Huberdeau et al., 2017).

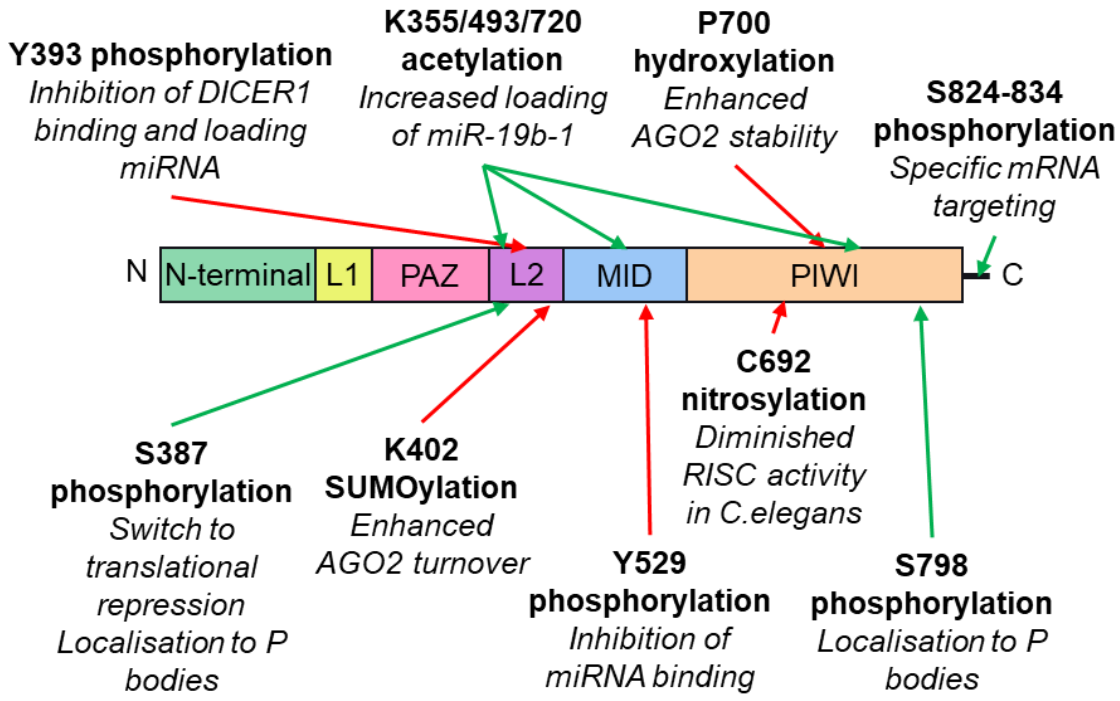


Figure 1.6.1 Post-translational modifications of AGO2

Schematic of various sites which are post-translationally modified on AGO2. Modifications are in bold and the consequences of each are in italics. Sites which enhance AGO2 function have green arrows, whereas sites which impede function have red arrows. All sites have been identified in human cells unless stated otherwise.

```
ALG-2_CAEEL   MFPLPVH---------------NGPRLGKLSIFEMPGDSLTSSS-------FMPDGGA--    36
AGO1_DROME    MYPVGQQSQWTPSPTRPQSPSQAQTSFDTLTSPPAPGSSVNPTAVTSPSAQNVAAGGATV    60
Ago2_DANRE    ------------------------------------------------------MYPIGA--     6
AGO2_XENLA    ------------------------------------------------------------      0
AGO2_CHICK    ---------------------------------------------------------MYPG--     4
AGO2_CHIMP    ------------------------------------------SHHPSQEQLRSLPT--    14
AGO2_MOUSE    ------------------------------------------------------------      0
AGO2_HUMAN    ------------------------------------------------------------      0
AGO2_HORSE    ------------------------------------------------------------      0

ALG-2_CAEEL   ----ETS--SSSQLG-------GSAHGAIGTKPDAGVQFQCPVRPNHGVEGRSILLRANH    83
AGO1_DROME    AGAAATAAQVASALGATTGSVTPAIATATPATQPDMPVFTCPRRPNLGREGRPIVLRANH   120
Ago2_DANRE    AGA----TELFQGRP-SSGSD---VSAPASPPAPQEYVFKPPQRPDFGTMGRTIKLQANF    58
AGO2_XENLA    ---------MYSGAG-P-VLV---PPTTTPPLPMPAYTFKPPPRPDFGTSGRTIKLQANV    46
AGO2_CHICK    A----------STAG-P-VLA---PPP-PLPPPVQGYAFKPPPRPDFGTSGRTIKLQANF    48
AGO2_CHIMP    TGT-----PYFIALH-K-ESL---APP-APPPPIQGYAFKPPPRPDFGTSGRTIKLQANF    63
AGO2_MOUSE    ---------MYSGAG-P-VLA---SPA-PTTSPIPGYAFKPPPRPDFGTTGRTIKLQANF    45
AGO2_HUMAN    ---------MYSGAG-P-ALA---PPA-P-PPPIQGYAFKPPPRPDFGTSGRTIKLQANF    44
AGO2_HORSE    ---------MYSGAG-P-VLA---PPA-PPPPPIQGYAFKPPPRPDFGTSGRTIKLQANF    45
                                          *  * **: *  ** * *:**

ALG-2_CAEEL   FAVRIPGGSVQHYQIDVFPDKCPRRVNREVIGCLISSF-SKYFTNIRPVYDGKRNMYTRE   142
AGO1_DROME    FQVTMPRGYVHHYDINIQPDKCPRKVNREIIETMVHAY-SKIFGVLKPVFDGRNNLYTRD   179
Ago2_DANRE    FEMEIPKLEVYHYEIDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVYDGRKNLYTAM   118
AGO2_XENLA    FEMDIPKIEIYHYDIDIKPEKCPRRVNREIVEHMVQHFKAQIFGDRKPVFDGRKNLYTAM   106
AGO2_CHICK    FEMDIPKIDIYHYELDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVFDGRKNLYTAM   108
AGO2_CHIMP    FEMDIPKIDIYHYELDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVFDGRKNLYTAM   123
AGO2_MOUSE    FEMDIPKIDIYHYELDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVFDGRKNLYTAM   105
AGO2_HUMAN    FEMDIPKIDIYHYELDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVFDGRKNLYTAM   104
AGO2_HORSE    FEMDIPKIDIYHYELDIKPEKCPRRVNREIVEHMVQHFKTQIFGDRKPVFDGRKNLYTAM   105
              * : :*   : **:::: *:****:****::  ::  : :: *   :**:**:.*:**

ALG-2_CAEEL   PLPIGTEP-MNFEVTLPGDSAVERKFSVTMKWIGQVCLSALDDAMEGRVRQVPHEAVQSI   201
AGO1_DROME    PLPIGNER-LELEVTLPGEG-KDRIFRVTIKWQAQVSLFNLEEALEGRTRQIPYDAILAL   237
Ago2_DANRE    PLPIGRDK-VELEVTIPGEG-KDRSFKVAIKWMSCVSLQALHEALSGRLPNIPFETIQAL   176
AGO2_XENLA    PLPIARDKQVELEVTLPGEG-KDRIFKVAIKWMACVSLQALHDALSGRLPNVPFETVQAL   165
AGO2_CHICK    PLPIGRDKQVELEVTLPGEG-KDRIFKVAIKWMSCVSLQALHDALSGRLPSVPFETIQAL   167
AGO2_CHIMP    PLPIGRDK-VELEVTLPGEG-KDRIFKVSIKWVSCVSLQALHDALSGRLPSVPFETIQAL   181
AGO2_MOUSE    PLPIGRDK-VELEVTLPGEG-KDRIFKVSIKWVSCVSLQALHDALSGRLPSVPFETIQAL   163
AGO2_HUMAN    PLPIGRDK-VELEVTLPGEG-KDRIFKVSIKWVSCVSLQALHDALSGRLPSVPFETIQAL   162
AGO2_HORSE    PLPIGRDK-VELEVTLPGEG-KDRIFKVAIKWVSCVSLQALHDALSGRLPSVPFETIQAL   163
              ****. :  ::::***:**:.  :* * *::** . *.*  *.:*:.**  .:*.::: ::

ALG-2_CAEEL   DVILRHLPSLKYTPVGRSFFTPPGVMKPGMQMHQESKLGGGREVWFGFHQSVRPSQWKMM   261
AGO1_DROME    DVVMRHLPSMTYTPVGRSFFSSPEGY--------YHPLGGGREVWFGFHQSVRPSQWKMM   289
Ago2_DANRE    DVVMRHLPSMRYTPVGRSFFTPSEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   228
AGO2_XENLA    DVVMRHLPSMRYTPVGRSFFTASEGC--------ANPLGGGREVWFGFHQSVRPSLWKMM   217
AGO2_CHICK    DVVMRHLPSMRYTPVGRSFFTASEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   219
AGO2_CHIMP    DVVMRHLPSMRYTPVGRSFFTASEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   233
AGO2_MOUSE    DVVMRHLPSMRYTPVGRSFFTASEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   215
AGO2_HUMAN    DVVMRHLPSMRYTPVGRSFFTASEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   214
AGO2_HORSE    DVVMRHLPSMRYTPVGRSFFTASEGC--------SNPLGGGREVWFGFHQSVRPSLWKMM   215
              **::*****: *********:                ****************** ****

ALG-2_CAEEL   LNIDVSATAFYRAMPVIEFVAEVLELPVQALAERRALSDAQRVKFTKEIRGLKIEITHCG   321
AGO1_DROME    LNIDVSATAFYKAQPVIDFMCEVLDIRDI-NEQRKPLTDSQRVKFTKEIKGLKIEITHCG   348
Ago2_DANRE    LNIDVSATAFYKAQPVIEFMCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   287
AGO2_XENLA    LNIDVSATAFYKAQPVIEFMCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   276
AGO2_CHICK    LNIDVSATAFYKAQPVIEFVCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   278
AGO2_CHIMP    LNIDVSATAFYKAQPVIEFVCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   292
AGO2_MOUSE    LNIDVSATAFYKAQPVIEFVCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   274
AGO2_HUMAN    LNIDVSATAFYKAQPVIEFVCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   273
AGO2_HORSE    LNIDVSATAFYKAQPVIEFVCEVLDFKSI-EEQQKPLTDSQRVKFTKEIKGLKVEITHCG   274
              ***********:* ***:*:.***::      ::: *:*:*********:***:******

ALG-2_CAEEL   AVRRKYRVCNVTRRPAQTQTFPLQLETGQTIECTVAKYFFDKYRIQLKYPHLPCLQVGQE   381
AGO1_DROME    QMRRKYRVCNVTRRPAQMQSFPLQLENGQTVECTVAKYFLDKYRMKLRYPHLPCLQVGQE   408
Ago2_DANRE    QMKRKYRVCNVTRRPASHQTFPLQQENGQTIECTVAQYFKDKYKLVLRYPHLPCLQVGQE   347
AGO2_XENLA    QMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   336
AGO2_CHICK    QMKRKYRVCNVTRRPASHQTFPLQQENGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   338
AGO2_CHIMP    QMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   352
AGO2_MOUSE    QMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   334
AGO2_HUMAN    QMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   333
AGO2_HORSE    QMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVAQYFKDRHKLVLRYPHLPCLQVGQE   334
               ::**************. *:**** *.***:*****:** *:::: *:************
```

```
ALG-2_CAEEL   QKHTYLPPEVCDIVPGQRCLKKLTDVQTSTMIKATARSAPERE-REICKLVSKAELSADPF   441
AGO1_DROME    HKHTYLPLEVCNIVAGQRCIKKLTDMQTSTMIKATARSAPDRE-REINNLVKRADFNNDSY   468
Ago2_DANRE    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQDEISKLMRSANFNTDPY   407
AGO2_XENLA    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPF   396
AGO2_CHICK    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPY   398
AGO2_CHIMP    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPY   412
AGO2_MOUSE    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPY   394
AGO2_HUMAN    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPY   393
AGO2_HORSE    QKHTYLPLEVCNIVAGQRCIKKLTDNQTSTMIRATARSAPDRQEEISKLMRSASFNTDPY   394
              :****** ***:** ****:***** ******:*******:*: ** :*:  *.:. * :

ALG-2_CAEEL   AHEFGITINPAMTEVKGRVLSAPKLLYGGRHRA--------------TTALPNQGVWDMR   487
AGO1_DROME    VQEFGLTISNSMMEVRGRVLPPPKLQYGGRVSTGLTGQQLFPPQNKVSLASPNQGVWDMR   528
Ago2_DANRE    VREFGVMVRDDMTEVNGRVLQAPSILYGGRNK---------------AIATPVQGVWDMR   452
AGO2_XENLA    VREFGIMVKDDMTDVTGRVLQPPSILYGGRSK---------------AIATPVQGVWDMR   441
AGO2_CHICK    VREFGIMVKDEMTDVTGRVLQPPSILYGGRNK---------------AIATPVQGVWDMR   443
AGO2_CHIMP    VREFGIMVKDEMTDVTGRVLQPPSILYGGRNK---------------AIATPVQGVWDMR   457
AGO2_MOUSE    VREFGIMVKDEMTDVTGRVLQPPSILYGGRNK---------------AIATPVQGVWDMR   439
AGO2_HUMAN    VREFGIMVKDEMTDVTGRVLQPPSILYGGRNK---------------AIATPVQGVWDMR   438
AGO2_HORSE    VREFGIMVKDEMTDVTGRVLQPPSILYGGRNK---------------AIATPVQGVWDMR   439
              .:***: :   * :* ****  *.: ****                 : * * *******

ALG-2_CAEEL   GKQFHTGMEVRTWAIACFAQQSHVKENDLRMFTTQLQRISTDAGMPIIGTPMFCKYASGV   547
AGO1_DROME    GKQFFTGVEIRIWAIACFAPQRTVREDALRNFTQQLQKISNDAGMPIIGQPCFCKYATGP   588
Ago2_DANRE    NKQFHTGIEIKVWAIACFAPQRQCTELLLKAFTDQLRKISRDAGMPIQGQPCFCKYAQGA   512
AGO2_XENLA    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKTFTEQLRKISRDAGMPIQGQPCFCKYAQGA   501
AGO2_CHICK    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKTFTEQLRKISRDAGMPIQGQPCFCKYAQGA   503
AGO2_CHIMP    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKSFTEQLRKISRDAGMPIQGQPCFCKYAQGA   517
AGO2_MOUSE    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKSFTEQLRKISRDAGMPIQGQPCFCKYAQGA   499
AGO2_HUMAN    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKSFTEQLRKISRDAGMPIQGQPCFCKYAQGA   498
AGO2_HORSE    NKQFHTGIEIKVWAIACFAPQRQCTEVHLKSFTEQLRKISRDAGMPIQGQPCFCKYAQGA   499
              .***.**:*:: ******* *    *  *: ** **::** ****** * * ***** *

ALG-2_CAEEL   EQVEPMFKYLKQTYSAIQLIVVVLPGKTPIYAEVKRVGDTVLGIATQCVQAKNAIRTTPQ   607
AGO1_DROME    DQVEPMFRYLKITFPGLQLVVVVLPGKTPVYAEVKRVGDTVLGMATQCVQAKNVNKTSPQ   648
Ago2_DANRE    DSVEPMFKHLKYTYQGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQVKNVQKTTPQ   572
AGO2_XENLA    DSVEPMFRHLKNTYTGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   561
AGO2_CHICK    DSVEPMFRHLKNTYTGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   563
AGO2_CHIMP    DSVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   577
AGO2_MOUSE    DSVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   559
AGO2_HUMAN    DSVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   558
AGO2_HORSE    DSVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTTPQ   559
              :.*****::** *: .:**:**:******:*************:****** **. :*:**

ALG-2_CAEEL   TLSNLCLKMNVKLGGVNSILLPNVRPRIFNEPVIFLGCDITHPAAGDTRKPSIAAVVGSM   667
AGO1_DROME    TLSNLCLKINVKLGGINSILVPSIRPKVFNEPVIFLGADVTHPPAGDNKKPSIAAVVGSM   708
Ago2_DANRE    TLSNLCLKINVKLGGVNNILLPQGRPLVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   632
AGO2_XENLA    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   621
AGO2_CHICK    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   623
AGO2_CHIMP    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   637
AGO2_MOUSE    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   619
AGO2_HUMAN    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   618
AGO2_HORSE    TLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSM   619
              ********:******:*.**:*. ** :*::******.*:*** *** :***********

ALG-2_CAEEL   DAHPSRYAATVRVQQHRQEIITDLTYMVRELLVQFYRNT-RFKPARIVVYRDGVSEGQLF   726
AGO1_DROME    DAHPSRYAATVRVQQHRQEIIQELSSMVRELLIMFYKSTGGYKPHRIILYRDGVSEGQFP   768
Ago2_DANRE    DAHPSRYCATVRVQQHRQDIIQDLATMVRELLIQFYKST-RFKPTRIIYYRDGISEGQFN   691
AGO2_XENLA    DAHPNRYCATVRVQQHRQEIIQDLSAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   680
AGO2_CHICK    DAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   682
AGO2_CHIMP    DAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   696
AGO2_MOUSE    DAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   678
AGO2_HUMAN    DAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   677
AGO2_HORSE    DAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKST-RFKPTRIIFYRDGVSEGQFQ   678
              ****.**.**********:** :*: ******: **:.*  :** **: ****:****:

ALG-2_CAEEL   NVLQYELRAIREACVMLESGYQPGITFIAVQKRHHTRLFAADKADQVGKAFNIPPGTTVD   786
AGO1_DROME    HVLQHELTAIREACIKLEPEYRPGITFIVVQKRHHTRLFCAEKKEQSGKSGNIPAGTTVD   828
Ago2_DANRE    QVLQHELLAIREACIKLEKDYQPGITFVVVQKRHHTRLFCMDRNERVGKSGNIPAGTTVD   751
AGO2_XENLA    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDRNERVGKSGNIPAGTTVD   740
AGO2_CHICK    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVD   742
AGO2_CHIMP    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVD   756
AGO2_MOUSE    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVD   738
AGO2_HUMAN    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVD   737
AGO2_HORSE    QVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVD   738
              :**::** ******: **  *:*****:.**********. :: :: **: *** *****
```

```
ALG-2_CAEEL   VGITHPTEFDFFLCSHAGIQGTSRPSHYHVLWDDNDLTADELQQLTYQMCHTYVRCTRSV   846
AGO1_DROME    VGITHPTEFDFYLCSHQGIQGTSRPSHYHVLWDDNHFDSDELQCLTYQLCHTYVRCTRSV   888
Ago2_DANRE    TKITHPFEFDFYLCSHAGIQGTSRPSHYHVLWDDNHFTSDELQVLTYQLCHTYVRCTRSV   811
AGO2_XENLA    TKITHPSEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   800
AGO2_CHICK    TKITHPSEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   802
AGO2_CHIMP    TKITHPTEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   816
AGO2_MOUSE    TKITHPTEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   798
AGO2_HUMAN    TKITHPTEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   797
AGO2_HORSE    TKITHPTEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRSV   798
              . **** ****:**** ****************** : :**** ****:***********

ALG-2_CAEEL   SIPAPAYYAHLVAFRARYHLVDRDHGSGEEGSQPSGTSSEDTTLSSMAKAVQVHPDSNNV   906
AGO1_DROME    SIPAPAYYAHLVAFRARYHLVEKEHDSGEGSH-QSGC-SEDRTPGAMARAITVHADTKKV   946
Ago2_DANRE    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDQQALAKAVQIHQDTLRT   869
AGO2_XENLA    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDQQALAKAVQVHQDTLRT   858
AGO2_CHICK    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDHQALAKAVQVHQDTLRT   860
AGO2_CHIMP    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDHQALAKAVQVHQDTLRT   874
AGO2_MOUSE    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDHQALAKAVQVHQDTLRT   856
AGO2_HUMAN    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDHQALAKAVQVHQDTLRT   855
AGO2_HORSE    SIPAPAYYAHLVAFRARYHLVDKEHDSAEGSH-TSGQ-SNGRDHQALAKAVQVHQDTLRT   856
              ********************:::*.*.* .   **  *:.    ::*:*: :* *: ..

ALG-2_CAEEL   MYFA      910
AGO1_DROME    MYFA      950
Ago2_DANRE    MYFA      873
AGO2_XENLA    MYFA      862
AGO2_CHICK    MYFA      864
AGO2_CHIMP    MYFA      878
AGO2_MOUSE    MYFA      860
AGO2_HUMAN    MYFA      859
AGO2_HORSE    MYFA      860
              ****
```

Figure 1.6.2 Alignment of AGO2 proteins in eukaryotes

Alignments of AGO1/2 proteins which participate in the miRNA pathway in various species. CAEEL (*Caenorhabditis elegans*), DROME (*Drosophila melanogaster*), DANRE (*Danio rerio*, zebrafish); XENLA (*Xenopus laevis*, African clawed frog); CHICK (chicken). Generated using Clustal Omega (Sievers et al., 2011; EMBL-EBI, 2021).

Ubiquitination of AGO2 promotes its degradation under specific conditions. For example, ubiquitination in response to T cell activation leads to cell differentiation (Bronevetsky et al., 2013). Ubiquitin control of AGO2 expression has also been linked to breast cancer, with reduced proteasomal degradation of AGO2 enhancing cell proliferation and migration. In this case, studies in mammalian cells suggested that the elevated AGO2 expression was downstream of epidermal growth factor receptor (EGFR) and mitogen activated protein kinase (MAPK) signalling (Adams et al., 2009). Additionally, proteasomal and autophagic degradation both appear to modulate AGO2 expression in response to miRNA levels (Smibert et al., 2013; Martinez and Gregory, 2013). Control of AGO2 levels in this way potentially serves to prevent aberrant AGO2 function by binding other RNAs.

Non-ubiquitin modifications may also regulate AGO2 expression. For example, *in vitro* and cell-based assays have demonstrated that small ubiquitin-like modifier (SUMO)-conjugating enzyme ubiquitin carrier protein 9 (UBC9) and SUMO-ligase Ran binding protein 2 (RANBP2) SUMOylate AGO2 at lysine 402 to promote turnover (Sahin et al., 2014). Conversely, hydroxylation of AGO2 at proline 700 by collagen prolyl-4-hydroxylase both increases stability (Qi et al., 2008) and promotes HSC70/HSP90 binding to enhance AGO2 activation (Wu et al., 2011).

AGO2 PTM may specifically alter AGO2 activity. Under conditions of cell stress and viral infection, AGO2 is poly-ADP-ribosylated to inhibit translational repression and increase expression of interferon genes (Leung et al., 2011; Seo et al., 2013). Similarly, mass spectrometric analyses have identified that mRNA degradation and translational repression are prevented by S-nitrosylation of *C. elegans* ALG-1 at cysteine 692, which is conserved in mammalian AGO2. S-

nitrosylation occurs in response to nitrogen released by host bacteria (Seth et al., 2019). Another recent observation is acetylation of AGO2 at lysine 720, 493 and 355, by P300/CREB-binding protein (CBP). Acetylation allows AGO2 to recruit pre-mir-19b-1; formation of this pre-miRNA deposit complex enhances processing by DICER1. Consequently, increased miR-19b expression, along with high AGO2 acetylation, is correlated with poor prognosis in lung cancer (H. Zhang et al., 2019).

Phosphorylation of AGO2 has significant influence over its function. Probably the best characterised phosphorylation site (phospho-site) in AGO2 is serine 387 (S387), which was initially found as a downstream target of the p38 MAPK pathway (Zeng et al., 2008). AKT3 was the first direct kinase to be identified for this site. Cell-based cleavage assays and immunofluorescence experiments demonstrated that phosphorylation by AKT3 initiates a shift in AGO2 activity from mRNA cleavage to translational repression (Horman et al., 2013). LIM domain-containing protein 1 (LIMD1) binds phosphorylated S387 and recruits TNRC6 proteins to facilitate mRNA silencing and localise AGO2 to P bodies (Bridge et al., 2017). S387 phosphorylation is both positively and negatively regulated by N-methyl-D-aspartate receptor (NMDAR) activation in neuronal cells: downstream AKT3-mediated phosphorylation permits translational repression of genes involved in synaptic plasticity (Rajgor et al., 2018). Conversely, NMDAR-inhibition of phosphorylation leads to AGO2 degradation and leads to spinal maturation (Paradis-Isler and Boehm, 2018). Finally, western blot analyses have suggested that direct phosphorylation of S387 by MEK excludes AGO2 from being sorted into exosomes. Consequently, increased S387 phosphorylation in KRAS transformed cells correlates with reductions in secreted miRNA (McKenzie et al., 2016).

Serine/threonine phosphorylation has also been recorded on AGO2 carboxy-terminus (C-terminus). RIP and luciferase assays have revealed that phosphorylation of a cluster of residues inhibits mRNA binding but is required for efficient gene silencing in both *C. elegans* ALG-1 (Quévillon Huberdeau et al., 2017) and human AGO2 (Golden et al., 2017). A cycle of phosphorylation of human AGO2 serine 824-834 is mediated by casein kinase I isoform α (CSNK1A1) and ankyrin repeat subunit C (ANKRD52)/protein phosphatase 6 catalytic subunit (PP6C). Dynamic phosphorylation of these residues may serve to ensure specific targeting of mRNAs (Golden et al., 2017). Similarly, phosphorylation of AGO2 serine 798 by an unknown kinase targets AGO2 to P bodies (Lopez-Orozco et al., 2015).

Several tyrosine phospho-sites on AGO2 have also been described. The first characterised was tyrosine 529 (Y529). Mutational analyses demonstrated that Y529 phosphorylation reduces miRNA binding due to the repulsive force between the phosphate of the modified residue and the 5’ end of the miRNA (Rüdel et al., 2010). Phosphorylation of Y529 is important for macrophage activation through released suppression of mRNAs encoding cytokines (Mazumder et al., 2013). Proto-oncogene tyrosine-protein kinase SRC (c-SRC) is the only known kinase for Y529. Phosphorylation of this residue, along with tyrosine 393 (Y393) and 749, promotes tumorigenesis by increasing cell proliferation and migration in various cancer cell lines including the A549 lung cancer cell line (T. Liu et al., 2020).

Y393 phosphorylation inhibits DICER1 binding AGO2 and loading miRNAs. Interestingly, small-RNA sequencing experiments suggested that, under hypoxia, phosphorylation of Y393 by EGFR suppresses loading of long-loop miRNAs. These miRNAs are enriched in tumour suppressors, therefore

increased Y393 phosphorylation correlates with higher cell migration and poor prognosis in breast cancer (Shen et al., 2013). Y393 is dephosphorylated by protein tyrosine phosphatase 1B (PTP1B) and leads to premature senescence in $HRAS^{G12V}$-transformed cells (Yang et al., 2014). In *Theileria annulate*-infected macrophages, increased expression of the adaptor protein growth factor receptor-bound protein 2 (GRB2) is proposed to facilitate GRB2-mediated recruitment of PTP1B to AGO2: GST-pulldown experiments suggested association of both proteins with GRB2, but the mechanism of binding was not elucidated nor shown to be direct (Haidar et al., 2018). Decreased AGO2 phosphorylation leads to suppression of JNK-interacting protein-2 (JIP2) and consequently macrophage dissemination (Haidar et al., 2018). Finally, the non-receptor tyrosine kinase FYN phosphorylates AGO2 at an unknown site and thus releases suppression of myelin basic protein in oligodendrocytes (Müller et al., 2015).

1.7 MAPK regulation of miRNA biogenesis

The MAPK pathway has been frequently implicated in regulation both of AGO2 and of miRNA biogenesis via modification of other proteins. MAPK pathways are widely conserved and have diverse but vital roles in a cell's response to external stimuli. Growth factors, inflammatory cytokines and environmental stresses may each initiate a different MAPK pathway. The intracellular pathway is characterised by a phosphorylation cascade which amplifies the signal and culminates in gene transcription to generate a biological outcome (Figure 1.7.1). External stimuli may be detected by a transmembrane receptor, such as the receptor tyrosine kinases (RTKs) which bind growth factors. IP experiments have revealed that RTKs transduce the signal to the MAPKs via recruitment of

GRB2 and the guanine nucleotide exchange factor son of sevenless (SOS) (Lowenstein et al., 1992; Buday and Downward, 1993). SOS triggers guanine nucleotide exchange on the membrane-bound small GTPase RAS. GTP-binding allows RAS to recruit the proto-oncogene serine/threonine-protein kinase RAF, a MAPK kinase kinase (MAPKKK) (Stokoe et al., 1994). Localisation of RAF to the membrane results in its phosphorylation and activation. RAF then phosphorylates and activates the MAPK kinase (MAPKK), MEK1/2 (Kyriakis et al., 1992). MEK1/2 is a serine/threonine and tyrosine kinase; it phosphorylates and activates the MAPK extracellular-regulated kinase 1/2 (ERK1/2) (Crews et al., 1992). Upon phosphorylation, the serine/threonine kinase ERK1/2 translocates to the nucleus and phosphorylates many transcription factors (TF) (Chen et al., 1992). Consequently, the cell transcriptome is altered to induce cell proliferation. The p38 and JUN N-terminal kinase (JNK) pathways involve different proteins but are governed by the same principles.

Stimuli	Growth Factor	Cytokines, stress	Environmental Stress
Receptor	**RTK**	e.g. TNFR	-
Adaptors/ Activators	**GRB2**, SOS, **RAS**	TRAF2/3/6, RHO	RAC1
MAPKKK	RAF	MEKK1-4, MLK, ASK1, TAK1	MEKK1/4, MLK
MAPKK	**MEK1/2**	MAPKK3/6, MAPKK4	MAPKK4/7
MAPK	**ERK1/2**	**p38**	JNK1/2/3
Substrates	ETS, ELK1, MYC, STAT1/3, **ESR2**, MNK1/2	HSP27, PLA2, MNK1/2, **MAPKAPK2**, MYC, ELK1, ATF2, STAT1	JUN, ATF2, ELK1, DCP4, **TP53**, NFAT
Outcome	Cell proliferation, differentiation	Apoptosis, inflammation, stress response, cell differentiation, survival, proliferation	

Figure 1.7.1 Proteins of the MAPK signalling pathways and their involvement in miRNA biogenesis

Outcomes and proteins at each stage of the ERK1/2, p38 and JNK1/2/3 mitogen-activated protein kinase (MAPK) pathways. Proteins involved in regulation of miRNA biogenesis or AGO2 function are in bold. MAPK kinase (MAPKK), MAPKK kinase (MAPKKK), receptor tyrosine kinase (RTK). (Adapted from: Soares-Silva, Diniz, Gomes, & Bahia, 2016; Wei and Liu, 2002).

MAPK signalling proteins are highly involved in the regulation of AGO2 (Figure 1.7.1). Examples have already been described with EGFR, GRB2, RAS, MEK, ESR2, p38 and TP53 (Zeng et al., 2008; Shen et al., 2013; Shankar et al., 2016; Krell et al., 2016; McKenzie et al., 2016; Tarallo et al., 2017; Haidar et al., 2018). Kinases in these pathways also phosphorylate other miRNA biogenesis proteins, demonstrating the importance of MAPK signalling in regulation of miRNA expression. Western blot analyses have identified that p38 phosphorylates DROSHA to inhibit the interaction with DGCR8. Reduced miRNA biogenesis consequently mediates stress-induced cell death (Yang et al., 2015). Additionally, a downstream target of p38, MAPK-activated protein kinase 2 (MAPKAP2), phosphorylates deadbox helicase 5 (DDX5). Phosphorylation promotes nuclear localisation of DDX5, where it enhances microprocessor processing of several pri-miRNAs (Hong et al., 2013). Similarly, mutational studies have demonstrated that ERK1/2 phosphorylates both DGCR8 and TRBP to facilitate their stabilisation. A pro-growth profile of miRNA expression is thus generated (Paroo et al., 2009; Herbert et al., 2013). A downstream target of ERK, ribosomal protein S6 kinase (S6K), also phosphorylates TRBP to the same effect (Warner et al., 2016). ERK phosphorylation of XPO5 leads to PIN1 isomerisation and a reduction in pre-miRNA nuclear export. Suppression of miR-122 by this mechanism may promote tumour development in liver cancer (Sun et al., 2016).

1.8 miRNAs in cancer

Although miRNAs may be either over or under-expressed in human cancers, miRNA profiling of tissue samples has suggested that an overall reduction in miRNA biogenesis is characteristic of many cancers (Lu et al., 2005).

Concurrently, the miRNA biogenesis proteins DICER1 and DROSHA are frequently downregulated or may harbour inactivating mutations (Figure 1.8.1) (Lin and Gregory, 2015). However, abrogation of DICER1 expression is inhibitory to tumour formation (Lambertz et al., 2010). Interestingly, expression of AGO2 is rarely altered in cancer. Exceptions to this include upregulation in serous ovarian carcinoma and myeloma (Y. Zhou et al., 2010; Vaksman et al., 2012) and downregulation in melanoma (Völler et al., 2013). Furthermore, in one study, staining tissue microarrays indicated that AGO2 expression may help to classify different breast cancer subtypes. However, no correlation was found between AGO2 expression and prognosis (Casey et al., 2019).

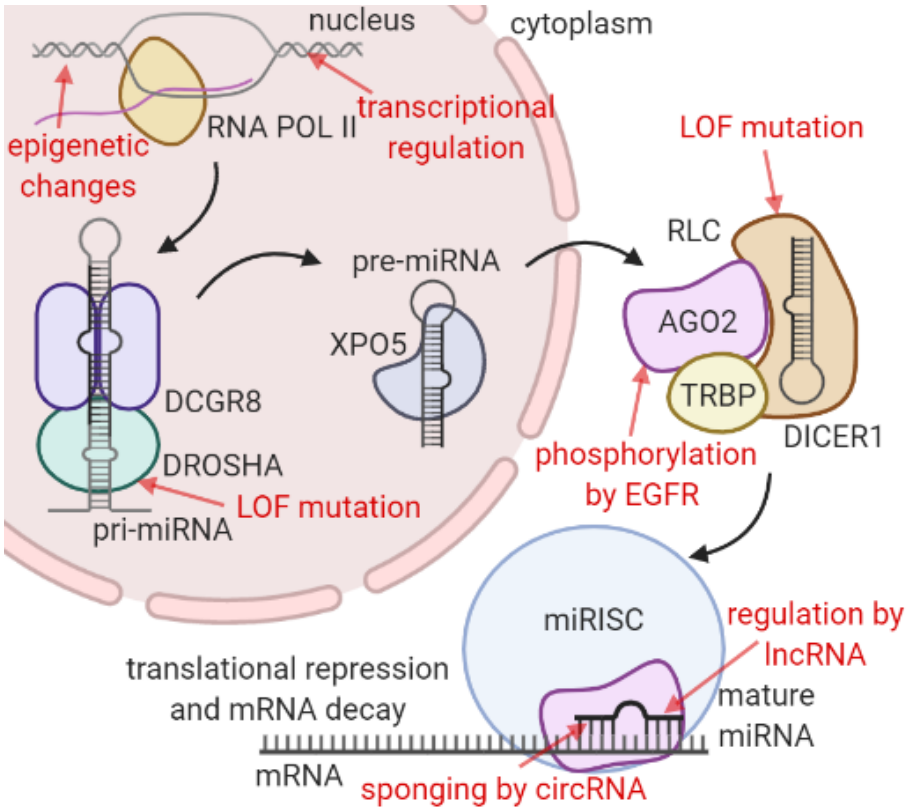


Figure 1.8.1 Mechanisms of altered regulation of miRNAs in cancer

Expression and activity of miRNAs may be altered in cancer via genetic mutation of miRNAs and miRNA biogenesis proteins. Loss of function (LOF) mutations are reported for both DROSHA and DICER1. Additionally, epigenetic changes and interactions with other RNAs may regulate miRNA expression. Alternatively, AGO2 phosphorylation and interactions with other proteins may affect miRNA biogenesis or activity. Created using BioRender.com.

Dysregulation of specific miRNAs may be easier to characterise. Almost all human cancers have a deregulated miRNA expression signature (Lu et al., 2005). Furthermore, some cancers show over and under expression of the same miRNAs (Volinia et al., 2006). These miRNAs may target well known tumour suppressor genes or proto-oncogenes, suggesting the miRNAs themselves act as oncomirs or tumour suppressors. For example, luciferase assays have demonstrated that the let-7 family of miRNAs targets *RAS.* Accordingly, decreased let-7 expression is proposed to mediate increased levels of RAS protein in lung cancer (Johnson et al., 2005). Different expression of many miRNAs in cancer is frequently a result of genetic alteration (Figure 1.8.1) (Zhang et al., 2006). One reason for this is that miRNA genes are located in regions of the genome which are particularly susceptible to mutation (Calin et al., 2004). For example, microarray experiments have shown that the oncogenic *mir-17~92* cluster of miRNA genes is often amplified in B cell lymphoma leading to elevated miR-17~92 expression (L. He et al., 2005). Conversely, tumour suppressor genes *mir-15* and *mir-16* are deleted in B cell chronic lymphocytic leukaemia (Calin et al., 2002). Point mutations may also modify miRNA expression, through a reduction in transcription (Raveche et al., 2007), pri-miRNA processing (Calin et al., 2005; Auyeung et al., 2013) and miRNA-loading (Trissal et al., 2018). Less often, mutations in the miRNA seed sequence may inhibit targeting of mRNAs (Trissal et al., 2018).

Epigenetic changes also regulate miRNA expression in cancer (Figure 1.8.1). In many cancers, changes in miRNA promotor methylation alters their transcription. Methylation-specific polymerase chain reaction (PCR) initially identified hypermethylation of *mir-124a* in colon cancer (Ueda et al., 2014).

Consequent transcriptional inactivation leads to increased levels of the miR-124a target cyclin D kinase 6.

Changes in miRNA biogenesis may also account for dysregulated miRNA expression (see 1.5, 1.7). One of the best characterised is phosphorylation of AGO2 Y393 by EGFR and its significance in breast cancer prognosis (Figure 1.8.1) (Shen et al., 2013).

More recently, it has emerged that circular RNAs (circRNAs) may act as sponges to inhibit specific miRNAs (Figure 1.8.1). Cell-based assays and analysis of tumour samples have suggested that circNRIP1 sponges miR-149-5p leading to increased expression of AKT1 and cell proliferation, migration and invasion in gastric cancer (X. Zhang et al., 2019). Similarly, long non-coding RNAs (lncRNAs) may alter miRNA expression (Figure 1.8.1). In hepatocellular carcinoma, increased levels of lncRNA MIR4435 enhance expression of miR-147a to promote cell proliferation (Kong et al., 2019). Finally, feedback mechanisms are a frequent characteristic of miRNA regulation (Tsang et al., 2007) and frequently contribute to cancer progression (Zhao et al., 2016; Shen et al., 2017; J. Liu et al., 2018). For example, quantitative PCR (qPCR) analysis has demonstrated that transforming growth factor β (TGFB1) activates expression of miR-1269. miR-1269 then targets negative regulators of TGFB1 signalling and promotes colorectal cancer progression (Bu et al., 2015).

miRNAs have been suggested as a prognostic and diagnostic tool to aid treatment of human cancers. Initially, microarray experiments found that the miRNA fingerprint of metastases of unknown origin may be better at predicting poorly differentiated cancers than mRNA expression (Rosenfeld et al., 2008). The discovery that exosomes containing miRNAs are released into the circulation then suggested that circulating miRNAs may provide a non-invasive

method for cancer diagnosis (Valadi et al., 2007). miRNAs in the blood are protected from degradation by the vesicles or carrier proteins to which they are attached (Valadi et al., 2007; Arroyo et al., 2011). Furthermore, miRNAs have been detected in 12 different body fluids (Weber et al., 2010). Serum miRNAs were originally found to show distinct expression patterns in lung and colorectal cancer, while maintaining consistent levels between healthy individuals (Chen et al., 2008). A more recent review found 279 circulating miRNAs consistently deregulated in gastric, colorectal, hepatocellular, pancreatic, lung, breast, cervical and prostate cancers (He et al., 2015).

Simultaneously, miRNAs may be used in therapeutics to treat cancer. Use of synthetic miRNA mimics may be particularly effective, due to their ability to target multiple oncogenes. One of the most successful miRNA mimics is for miR-34 (MRX34) (Wiggins et al., 2010), which was used in the first phase I clinical trial for a miRNA mimic. Although the trial was closed early due to severe immune toxicity, liposomal administration of MRX34 reduced expression of many miR-34 targets with a dose response and tumour burden was reduced in 4% of patients (Hong et al., 2020). Equally, synthetic siRNAs may be designed to target oncogene expression. While no siRNA-based therapeutic has been approved for cancer treatment, the hereditary transthyretin-mediated amyloidosis and acute hepatic porphyria treatments, Patisiran and Givosiran, have both been brought to the market (Hoy, 2018; Scott, 2020).

Alternatively, reducing the expression of oncomirs may provide a useful avenue for cancer therapy. Oncomirs may be targeted by antisense oligomirs called antimiRs, such as the antimiR for miR-155 (Babar et al., 2012). In a mouse model of lymphoma, attachment of the antimiR to a peptide, which forms an α-helix under acidic pH, targeted the drug to cancer cells. Additionally, this

method allowed non-endocytic uptake and reduced clearance by the liver (Cheng et al., 2015).

1.9 GRB2

GRB2 is an adaptor protein which classically recruits signalling proteins to stimulated receptors. It is composed of three domains: a SRC homology 2 (SH2) domain flanked by two SRC homology 3 (SH3) domains (Figure 1.9.1) (Lowenstein et al., 1992). Thus, GRB2 is able to bind multiple proteins at any one time to form protein complexes. While SH3 domains bind proline-rich motifs (PRMs) (Cicchetti et al., 1992; Booker et al., 1993), SH2 domains interact with phosphorylated tyrosine residues, such as those presented on active RTKs (Moran et al., 1990). GST-pulldown experiments have shown that binding of GRB2 to stimulated RTKs is via an interaction between its SH2 domain and a phospho-tyrosine (pY) (Lowenstein et al., 1992). The pY may be presented directly on the receptor or another adaptor protein which is bound to the receptor. With both its C and N-terminal SH3 domains (CSH3, NSH3) it then recruits SOS (Rozakis-Adcock et al., 1993; Li et al., 1993; Chardin et al., 1993). GRB2 is highly conserved in eukaryotes (Figure 1.9.2) and GRB2-dependent signalling is seen in both *C. elegans* and *D. melanogaster* (Downward, 1994).

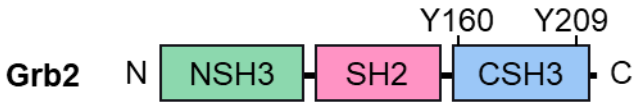


Figure 1.9.1 Domain structure of GRB2

GRB2 possesses N and C-terminal SH3 domains, which flank an SH2 domain. Tyrosines 160 and 209 can be phosphorylated to induce dimer dissociation and release from FGFR2 C-terminal tail.

```
SEM5_CAEEL    MEAVAEHDFQAGSPDELSFKRGNTLKVLNKDEDPHWYKAELDGNEGFIPSNYIRMTECNW 60
GRB2_DROME    MEAIAKHDFSATADDELSFRKTQILKILNMEDDSNWYRAELDGKEGLIPSNYIEMKNHDW 60
GRB2A_XENLA   MEAIAKYDFKATADDELSFKRGDVLKVLNEECDQNWYKAELNGKDGFIPKNYIEMKAHPW 60
GRB2_DANRE    MEAIAKYDFKATADDELSFKRGEILKVLNEECDQNWYKAELNGKDGFIPKNYIEMKAHPW 60
GRB2_CHICK    MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKGGFIPKNYIEMKPHPW 60
GRB2_HUMAN    MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPKNYIEMKPHPW 60
GRB2_HORSE    MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPKNYIEMKPHPW 60
GRB2_CHIMP    MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPKNYIEMKPHPW 60
GRB2_MOUSE    MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPKNYIEMKPHPW 60
              ***:*::**.* : *****:: : **:** : * :**:***:*: *:**.***.*.   *

SEM5_CAEEL    YLGKITRNDAEVLLKKPTVRDGHFLVRQCESSPGEFSISVRFQDSVQHFKVLRDQNGKYY 120
GRB2_DROME    YYGRITRADAEKLLS-N-KHEGAFLIRISESSPGDFSLSVKCPDGVQHFKVLRDAQSKFF 118
GRB2A_XENLA   FFGKIPRAKAEEMLGKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
GRB2_DANRE    FFGRIPRARAEEILNKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
GRB2_CHICK    FFGKIPRAKAEEMLGKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQQFKVLRDGAGKYL 119
GRB2_HUMAN    FFGKIPRAKAEEMLSKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
GRB2_HORSE    FFGKIPRAKAEEMLSKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
GRB2_CHIMP    FFGKIPRAKAEEMLSKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
GRB2_MOUSE    FFGKIPRAKAEEMLSKQ-RHDGAFLIRESESAPGDFSLSVKFGNDVQHFKVLRDGAGKYF 119
              : *:* *  ** :*     ::* **:* .**:**:**:**:  :.**:****** .*:

SEM5_CAEEL    LWAVKFNSLNELVAYHRTASVSRTHTILLSDMN---------------VETKFVQALFDF 165
GRB2_DROME    LWVVKFNSLNELVEYHRTASVSRSQDVKLRDMI---------------PEEMLVQALYDF 163
GRB2A_XENLA   LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVPQVHGGDRATSLPQQPTYVQALFDF 179
GRB2_DANRE    LWVVKFNSLNELVDYHRTTSVSRNQQIFLRDIEQVP------------QHPTYVQALFDF 167
GRB2_CHICK    LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVP------------QQPTYVQALFDF 167
GRB2_HUMAN    LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVP------------QQPTYVQALFDF 167
GRB2_HORSE    LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVP------------QQPTYVQALFDF 167
GRB2_CHIMP    LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVP------------QQPTYVQALFDF 167
GRB2_MOUSE    LWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQMP------------QQPTYVQALFDF 167
              **.********** ***::****.: : * *:                .   ****:**

SEM5_CAEEL    NPQESGELAFKRGDVITLINKDDPNWWEGQLNNRRGIFPSNYVCPYNSNKSNSNVAPGFN 225
GRB2_DROME    VPQESGELDFRRGDVITVTDRSDENWWNGEIGNRKGIFPATYVTPYHS------------ 211
GRB2A_XENLA   DPQEDGELGFRRGDFIQVVDNSDPNWWKGTCLSQTGMFPRNYVTPVNRNM---------- 229
GRB2_DANRE    DPQEDGELGFRRGDFIQVLDNSDPNWWKGACHGQTGMFPRNYVTPVNRNM---------- 217
GRB2_CHICK    DPQEEGELGFRRGDFIQVLDNSDPNWWKGACHGQTGMFPRNYVTPVNRNI---------- 217
GRB2_HUMAN    DPQEDGELGFRRGDFIHVMDNSDPNWWKGACHGQTGMFPRNYVTPVNRNV---------- 217
GRB2_HORSE    DPQEDGELGFRRGDFIHVMDNSDPNWWKGACHGQTGMFPRNYVTPVNRNV---------- 217
GRB2_CHIMP    DPQEDGELGFRRGDFIHVMDNSDPNWWKGACHGQTGMFPRNYVTPVNRNV---------- 217
GRB2_MOUSE    DPQEDGELGFRRGDFIHVMDNSDPNWWKGACHGQTGMFPRNYVTPVNRNV---------- 217
               ***.*** *:***.* : :..* ***:*   .: *:** .** * :

SEM5_CAEEL    FGN 228
GRB2_DROME    --- 211
GRB2A_XENLA   --- 229
GRB2_DANRE    --- 217
GRB2_CHICK    --- 217
GRB2_HUMAN    --- 217
GRB2_HORSE    --- 217
GRB2_CHIMP    --- 217
GRB2_MOUSE    --- 217
```

Figure 1.9.2 Alignment of GRB2 proteins in eukaryotes

Alignments of GRB2 proteins in various species. CAEEL (*Caenorhabditis elegans*), DROME (*Drosophila melanogaster*), DANRE (*Danio rerio*, zebrafish); XENLA (*Xenopus laevis*, African clawed frog); CHICK (chicken). Generated using Clustal Omega (Sievers et al., 2011; EMBL-EBI, 2021).

Unphosphorylated GRB2 exists in a monomer-dimer equilibrium with K_D ~0.7 μM, as determined by microscale thermophoresis (MST) (Lin et al., 2012). Single molecule fluorescence imaging microscopy has revealed that both pY-binding and tyrosine 160 (Y160) phosphorylation disrupts dimer formation (Figure 1.9.1, Ahmed et al., 2015). Y160 is phosphorylated by c-SRC, BCR/ABL and FGFR2 in response to growth factor stimulation (Jones and Benjamin, 1997; Li et al., 2001; Ahmed et al., 2013). Additionally, in non-stimulated cells (i.e. cells that have been deprived of growth factor), cycling of GRB2 phosphorylation is observed due to the phosphatase and kinase activities of SHP2 and FGFR2 (Ahmed et al., 2013).

In addition to its role linking transmembrane receptors to intracellular signalling pathways, GRB2 has been implicated in several other mechanisms. For example, knockdown-rescue experiments have indicated that GRB2 hinders actin polymerisation to inhibit myogenesis, through an interaction involving CSH3 (Mitra and Thanabalu, 2017). Additionally, GRB2 facilitates the nuclear localisation of phosphatase and tensin homologue (PTEN) and thus regulates the DNA damage response (Hou et al., 2019). Receptor-independent functions of GRB2 may be unsurprising considering its localisation to both the cytoplasm and nucleus, as determined by immunohistochemistry of both breast cancer and normal cells (Verbeek et al., 1997). Upon stimulation of cells with growth factor, GRB2 is trafficked to the plasma membrane and to endosomes (Watanabe et al., 2000; Jiang and Sorkin, 2002; Fortian and Sorkin, 2014).

GRB2 may be dysregulated in cancer. For example, upregulation of GRB2 is seen in gastric, colorectal, breast, liver, kidney and oesophageal carcinomas (Yip et al., 2000; Yu et al., 2008; Yu et al., 2009; Li et al., 2014; Zhang et al., 2014; Zhang and Zhang, 2017). In contrast, tissue microarray experiments have

suggested that downregulation of GRB2 in conjunction with upregulation of fibroblast growth factor receptor 2 (FGFR2) and phospholipase C-γ (PLCγ) is associated with poor prognosis in lung adenocarcinoma and ovarian cancer patients (Timsah et al., 2015; Timsah et al., 2016). Regulation of GRB2 by miRNAs has also been implicated in cancer. One example of this is reduced targeting of GRB2 by miR-329 in pancreatic cancer. Decreased miR-329 expression releases suppression of GRB2 and consequently enhances cell proliferation (Wang et al., 2016).

Various attempts have been made to develop cancer therapeutics which target GRB2. For example, a peptide which binds GRB2 SH3 was shown to have cytotoxic effects in an erythroleukemic cell line expressing the breakpoint cluster region protein (BCR)-Abelson murine leukemia viral oncogene homologue (ABL) fusion. The peptide bound both SH3 domains with nanomolar affinity (Ye et al., 2008). However, drugs targeting the SH3 domains have not yet reached the clinic. The difficulty in targeting GRB2 SH3 may be inherent in the promiscuous nature of these interactions. Similarly, therapeutic inhibition of GRB2 SH2 domain has shown equally little promise. Peptides have been developed which bind the SH2 domain and diminish proliferation in breast cancer cell lines (Chen et al., 2010). Nevertheless, peptides against GRB2 SH2 have rarely been tested in animal models.

1.10 SRC homology domains

SH3 domains are formed of around 60 residues, which can have high sequence variability. Despite this, their structure is well conserved. SH3 domains fold into an anti-parallel β-barrel, with the n-SRC and RT loops forming a large peptide-binding surface (Musacchio et al., 1992; Booker et al., 1993). Mutational

analyses have demonstrated that the consensus SH3-binding site consists of two proline residues separated by any other two residues (PxxP) (Ren et al., 1993). This PxxP motif forms a polyproline II helix, allowing the two prolines to be bound by two shallow pockets in the SH3 domain. The interaction is mediated by multiple hydrophobic contacts (Yu et al., 1994). A third binding site is designated the ‘specificity pocket’. The specificity pocket is negatively charged and interacts with a positive residue adjacent to the PxxP motif. In this way, two binding orientations of proline-rich peptides may be observed: class I, ‘+xxPxxP’, or class II, ‘xPxxPx+’ (Feng et al., 1994). Characteristically, SH3 domains are highly promiscuous and have typically low (micromolar) binding affinities, as measured by fluorescence titration assays (Chen et al., 1993; Yu et al., 1994). Non-consensus SH3-domain binding motifs have also been identified (Lewitzky et al., 2001; Liu et al., 2003; Hoelz et al., 2006; Aitio et al., 2008; Santiveri et al., 2009). These ligands make multiple interactions with the specificity pocket and consequently bind with unusually high affinity.

SH2 domains are protein modules formed of around 100 residues. Like SH3 domains, they are well conserved (Kawata et al., 1997) and function to interact with other proteins. However, their mode of binding differs significantly to that of SH3 domains. Crystal structures have revealed that SH2 domains bind phosphorylated tyrosine residues with high affinity (Matsuda et al., 1990; Moran et al., 1990; Eck et al., 1993; Waksman et al., 1993). Specificity is inferred through recognition of residues C-terminal to pY (Zhou et al., 1993).

1.11 FGFR2

The fibroblast growth factor receptor (FGFR) family of RTKs are defined by their high affinity binding to fibroblast growth factor (FGF) ligands (Lee et al., 1989).

Humans express FGFR1-4, with the best studied being FGFR1, 2 and 3 (Givol and Yayon, 1992; Turner and Grose, 2010). FGFRs consist of three extracellular immunoglobulin-like domains (Ig), which are separated from the intracellular tyrosine kinase domain (KD) by a transmembrane helix (TM) (Lee et al., 1989) (Figure 1.11.1). The tyrosine kinase domain is flanked by a juxtamembrane region (JM) and C-terminal tail (CT). Two isoforms of FGFR2, FGFR2IIIb and FGFR2IIIc, are generated by alternative splicing of the second Ig domain. Consequently, the isoforms have different ligand-binding specificity: cell-based binding assays have revealed that FGFR2IIIb binds FGF7 and FGF10, whereas the IIIc isoform binds FGF2, FGF9 and FGF18 (Miki et al., 1992). Additionally, *in situ* hybridisation experiments in mouse embryos have identified that FGFR2IIIb and IIIc are expressed in epithelial and mesenchymal cells respectively (Orr-Urtreger et al., 1993).

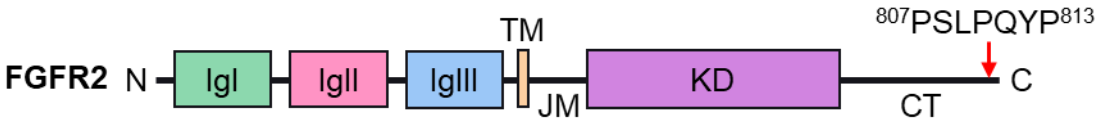


Figure 1.11.1 Domain structure of FGFR2

FGFR2 possesses three ligand-binding Ig domains, a transmembrane helix (TM), juxtamembrane region (JM), kinase domain (KD) and C-terminal tail (CT). In the CT, residues 807-813 form a proline-rich motif.

FGFR2 signalling in response to growth factor stimulation has been extensively studied. FGFs bind between Ig2 and Ig3 (Plotnikov et al., 1999), causing the receptors to homodimerise. Heparin or heparin sulphate proteoglycans (HSPGs) are required as cofactors for growth factor binding (Rapraeger et al., 1991; Ornitz et al., 1992). Dimerisation allows trans-phosphorylation of the kinase domains, and crystal studies have shown that this releases the kinase from its auto-inhibited conformation (Mohammadi et al., 1996). Time-resolved mass spectrometry experiments have demonstrated that phosphorylation occurs in a sequential manner. Initial phosphorylation of a single tyrosine in FGFR activation loop (A-loop, Y656 in FGFR2) increases activity of the kinase domain 50-100 fold. Phosphorylation of a further four tyrosine residues creates binding sites for intracellular signalling proteins. Finally, phosphorylation of the second A-loop tyrosine (Y657 in FGFR2) leads to a 500-1000 fold increase in kinase activity (Furdui et al., 2006).

Activation of FGFR, and other RTKs, is followed by docking of intracellular signalling proteins via their SH2 domains (Figure 1.11.2). Various signalling pathways are consequently initiated. For example, the FGFR-bound adaptor FGFR substrate 2 (FRS2) is phosphorylated, allowing GRB2 to dock onto phospho-FRS2 (Kouhara et al., 1997). GRB2 then recruits SOS to activate RAS-MAPK signalling (Chardin et al., 1993). Additionally, knockdown experiments have identified that binding of GRB2-associated binder 1 (GAB1) to GRB2 initiates the phosphatidylinositol 3-kinase (PI3K)/AKT pathway (Lamothe et al., 2004). PLCγ binds FGFR directly to stimulate signalling via inositol triphosphate (IP_3) and diacylgycerol (DAG) (Mohammadi et al., 1991). Finally, signal transducers and activators of transcription (STATs) are also activated by direct recruitment to FGFR (Su et al., 1997). Through these various

signalling pathways, FGFR2 regulates cell proliferation, apoptosis, survival, differentiation, angiogenesis, migration and invasion (Ornitz and Itoh, 2015).

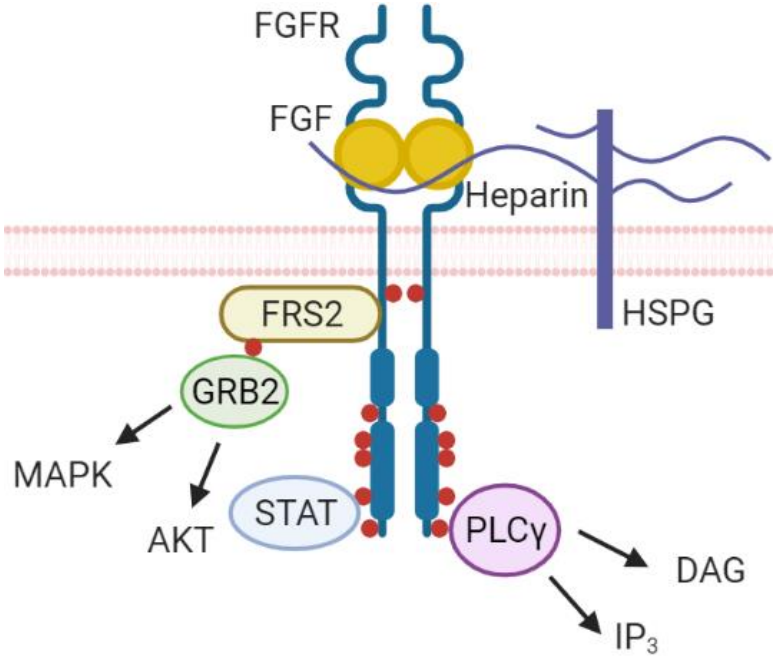


Figure 1.11.2 Signalling via activated FGFRs

FGFR is activated by binding of FGF and the cofactor Heparin/HSPG. Stimulation is followed by trans-phosphorylation of the intracellular region. Phosphorylation increases kinase activity and creates phosphotyrosine binding sites for SH2 domain-containing proteins. FRS2 and GRB2 couple FGFR to MAPK and AKT pathways. Binding of PLCγ links the receptor to IP_3 and DAG signalling. STAT is also recruited and activated. Phosphotyrosines are shown in red. Created with BioRender.com.

Under basal conditions, a mixture of isothermal titration calorimetry (ITC) and cell-based assays have revealed that GRB2 CSH3 binds the PRM ^{807}PSLPQYP^{813} in the CT of FGFR2 (Ahmed et al., 2010). Binding of a GRB2 dimer leads to formation of a tetrameric complex. In this state, Y656 is phosphorylated and FGFR2 is kinase active. However, GRB2 sterically hinders further phosphorylation of FGFR2. Hence, FGFR2 is primed ready for growth factor mediated activation but unable to activate downstream signalling pathways (Figure 1.11.3). Upon ligand stimulation of FGFR2, GRB2 is phosphorylated on Y160 and tyrosine 209 (Y209, Figure 1.9.1), causing it to monomerise and dissociate from FGFR2 C-terminus, allowing FGFR2 to become fully phosphorylated (Lin et al., 2012).

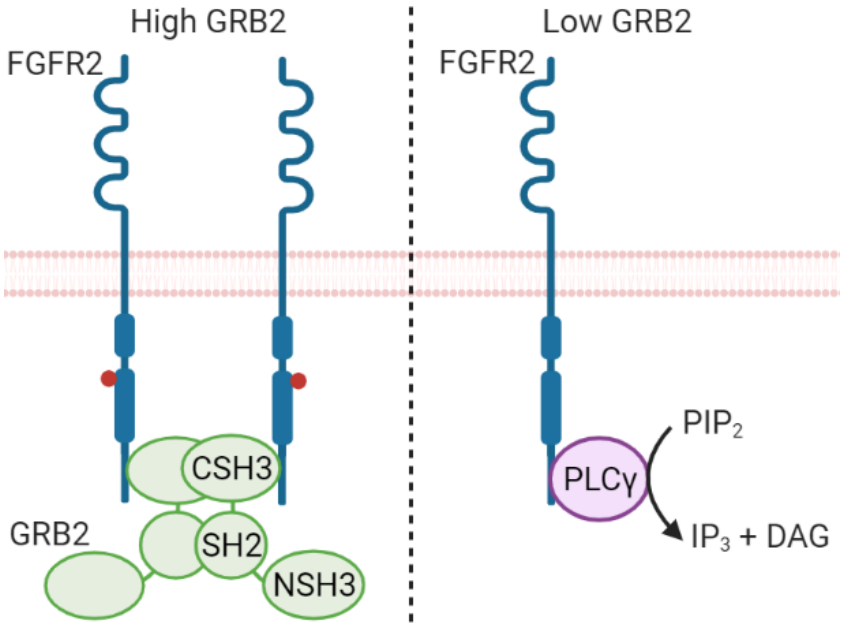


Figure 1.11.3 FGFR2 signalling in absence of growth factor

Under non-stimulated conditions, the C-terminal SH3 domain of GRB2 binds FGFR2 proline-rich motif to form a tetrameric complex. FGFR2 Y656 becomes phosphorylated (red circle). When GRB2 expression is low, PLCγ binds the same motif. PLCγ is activated and converts PIP_2 to IP_3 and DAG. Created with BioRender.com.

Interestingly, under non-stimulated conditions, binding of signalling proteins to FGFR2 is dynamic. A combination of *in vitro* kinase assays and western blotting experiments have shown that in the GRB2-FGFR2 complex, FGFR2 can phosphorylate GRB2 Y209, causing GRB2 to dissociate from the receptor. Dissociation of GRB2 allows SH2 domain-containing protein tyrosine phosphatase 2 (SHP2) to bind FGFR2. Consequently, FGFR2 phosphorylates SHP2 tyrosine 542 to upregulate its activity. SHP2 then dephosphorylates both FGFR2 and GRB2, allowing GRB2 to again bind FGFR2 CT. Y656 again is phosphorylated and the cycle repeats (Ahmed et al., 2013).

When GRB2 expression is low, PLCγ binds the same PRM on FGFR2. Binding upregulates PLCγ activity and the phospholipase hydrolyses phosphatidylinositol 4,5-bisphosphate to produce IP_3 and DAG (Figure 1.11.3) (Timsah et al., 2014). Signalling via these molecules promotes cell invasion (Timsah et al., 2015).

In a similar mechanism, NMR analysis has indicated that the the SH3 domain of SRC-family kinase, FYN, binds a PRM in the CT of RTK ERBB2. This interaction is hypothesised to contribute to oncogenesis in breast cancers, which demonstrate SRC-family kinase interaction with ERBB2 (Bornet et al., 2014).

Amplification and activating mutations of FGFR2 are both seen in cancer. Gene fusions also occur, although less commonly (Helsten et al., 2016). In endometrial cancer, gain of function point mutations in the Ig domains increase ligand binding affinity and release autoinhibition (Dutt et al., 2008; Helsten et al., 2016). Similarly, activating point mutations of both the ligand-binding and kinase domains have been detected in squamous cell lung carcinoma (Liao et al.,

2013). In stomach cancer, reverse transcription coupled with PCR has identified FGFR2 isoforms in which the CT is truncated (Itoh et al., 1994). The loss of specific residues in the CT is thought to inhibit receptor internalisation and promote FRS2 binding (Cha et al., 2009). Additionally, mutation of serine 780 in the CT in bladder cancer disrupts a negative feedback loop mediated by ERK1/2. Cells stably expressing this mutation show increased FGFR2 signalling and motility (Szybowska et al., 2019). Therefore, it appears the cause of aberrant FGFR2 signalling in cancer is not solely restricted to mutation of the ligand-binding and kinase domains.

Multiple drugs have been developed to target FGFR2 in cancer. Erdafitinib is a pan-FGFR inhibitor which is approved for the treatment of FGFR2 mutated urothelial carcinoma (Markham, 2019). It is a small molecule which is taken up by cells to inhibit FGFR signalling (Perera et al., 2017). Recently, in a phase II clinical trial, the small-molecule pemigatinib has been shown to have anti-tumour effects in advanced cholangiocarcinoma with FGFR2 fusion (Abou-Alfa et al., 2020). Pemigatinib is a potent inhibitor of FGFR1, 2 and 3 and weaker inhibition of FGFR4 (P.C.C. Liu et al., 2020).

1.12 Aims and hypotheses

Previous studies implicate GRB2 and other MAPK proteins in regulation of miRNA-loading (Shen et al., 2013; Shankar et al., 2017; Haidar et al., 2018). This thesis will investigate a direct role for GRB2 in control of miRNA-loading. FGFR2-dependent and independent mechanisms of regulation will be explored.

1. To determine if FGFR2 phosphorylates AGO2 and if this is regulated by GRB2

Hypothesis: GRB2 and FGFR2 cooperate to regulate RLC formation via FGFR2-mediated phosphorylation of AGO2 Y393.

RTK-mediated phosphorylation of AGO2 Y393 precludes DICER1-binding AGO2 to inhibit miRNA-loading (Shen et al., 2013). The potential for FGFR2 to phosphorylate AGO2 Y393 and consequently control RLC assembly will be evaluated. GRB2 mediates and inhibits FGFR2 substrate phosphorylation under both basal and starved conditions (Lowenstein et al., 1992; Lin et al., 2012). Therefore, GRB2 control of FGFR2-regulation of AGO2 will be studied.

2. To identify if GRB2 binds AGO2 directly and if this modulates the AGO2-DICER1 interaction independently of RTK

Hypothesis: GRB2 binds AGO2 directly to regulate RLC formation independently of RTK-mediated AGO2 phosphorylation.

GRB2 has been suggested to associate with AGO2 (Haidar et al., 2018). The interaction between GRB2 and AGO2 will be assessed and mechanism of binding elucidated. The effect of GRB2 on AGO2-binding DICER1 in absence of FGFR2 will be probed.

3. To identify miRNAs which are regulated by GRB2 and the stage at which GRB2 controls their expression

Hypothesis: GRB2-binding AGO2 regulates miRNA-loading. Consequent changes in miRNA expression have relevance in tumorigenesis and mediate feedback loops to control expression of GRB2.

Disruption of miRNA biogenesis may produce a miRNA signature with specific properties, such as being pro- or anti-tumorigenic (Volinia et al., 2006). miRNAs which are dysregulated upon depletion of GRB2 will be identified and the stage at which miRNA biogenesis is disrupted will be determined. Concurrent changes in expression of miRNA target genes will be examined. The potential for GRB2-regulation of miRNA to be disrupted in cancer will be considered. Finally, the potential for feedback suppression of GRB2 by the GRB2-dependent miRNAs will be explored.

This work will reveal a novel mechanism in the regulation of miRNA biogenesis and a new cellular function of GRB2. GRB2-mediated control of miRNA expression may have implications in cancer therapy.

Chapter 2: Materials and Methods

2.1 Antibodies

Table 2.1.1 Antibodies

Antibodies and clones, where relevant, are identified. Details are provided of their applications, dilutions, the antigen against which they were raised, the host species in which they were raised, and their supplier. Western blotting (immunoblotting, IB), immunoprecipitation (IP), RNA immunoprecipitation (RIP), carboxy-terminus (C-terminus), Cell Signalling Technologies (CST).

Antibody	Application	Dilution	Antigen	Host	Supplier, catalogue number
AGO2 (EPR10411)	IB IP, RIP	1:1000 1:100	Human AGO2 350-450	Rabbit	Abcam (ab186733)
AGO2 (2E12-1C9)	IP	1:50	Human AGO2 483-859	Mouse	Abcam (ab57113)
AGO2 pY393	IB	1:1000	A peptide corresponding to human AGO2 pY393	Rabbit	Abcam (ab215746)
AGO2 pY393 #1	IB IP	1:1000 1:50	Peptide: NTDP-pY-VREFGC	Rabbit	Genscript (custom)
AGO2 pY393 #2	IB IP	1:1000 1:50	Peptide: NTDP-pY-VREFGC	Rabbit	Genscript (custom)
GRB2	IB	1:1000	Human GRB2 C-terminus	Rabbit	CST (3972)
GRB2 C7	IB	1:1000	Mouse GRB2 54-164	Mouse	Santa Cruz (sc-8034)
FGFR2 (Bek C17)	IB	1:1000	Human FGFR2 C-terminus	Rabbit	Santa Cruz (sc-122)
FGFR2 (Bek C8)	IB	1:1000	Human FGFR2 805-821	Mouse	Santa Cruz (sc-6930)
pFGFR2 Y656/7 (55H2)	IB	1:1000	A peptide corresponding to human FGFR1 pY653/4	Mouse	CST (3476)
ERK1/2 (p44/42)	IB	1:1000	A peptide corresponding to rat p44 C-terminus	Rabbit	CST (4695)

pERK1/2	IB	1:1000	A peptide corresponding to human ERK1 pY202/204	Rabbit	CST (4370)
pY99	IB	1:1000	pY	Mouse	Santa Cruz (sc-7020)
DICER1	IB	1:1000	Human DICER1	Rabbit	CST (3363)
TUT4 (9F6.1)	IB	1:250	An epitope in the N-terminal region of human TUT4	Mouse	Sigma-Aldrich (MABE1526)
HMGA2 (D1A7)	IB	1:1000	A peptide corresponding to the N-terminal region of human HMGA2	Rabbit	CST (8179)
LIN28B (D4H1)	IB	1:1000	A peptide corresponding to the C-terminal region of human LIN28B	Rabbit	CST (11965)
GAPDH (14C10)	IB	1:500 1:1000	A peptide corresponding to a region near the C-terminus of human GAPDH	Rabbit	CST (2118)
α-Tubulin	IB	1:500 1:1000	A peptide corresponding to human α-Tubulin	Rabbit	CST (2144)
β-Actin (13E5)	IB	1:500 1:1000	A synthetic protein corresponding to the N-terminus of human β-Actin	Rabbit	CST (4970)
GFP (D5.1)	IB	1:1000	A peptide corresponding to the C-terminal region of GFP	Rabbit	CST (2956)
RFP	IB	1:1000	A peptide derived from residues 22-36	Rabbit	Genscript (A00682)

FLAG (M2)	IB	1:1000	A peptide with sequence DYKDDDDK	Mouse	Sigma-Aldrich (F1804)
HRP-conjugated Rabbit IgG	IB	1:1000	Rabbit heavy and light chain IgG	Goat	CST (7074)
HRP-conjugated Mouse IgG	IB	1:1000	Mouse heavy and light chain IgG	Horse	CST (7076)
Normal Rabbit IgG	IP	1:100 1:200	None	Rabbit	CST (2729)
Normal Mouse IgG	IP	1:40	None	Mouse	Santa Cruz (sc-2025)

2.2 Buffers

Lysis Buffer

50 mM Hepes-NaOH pH 7.5, 1% (v/v) Igepal-CA630 (Sigma, 18896) 10 mM NaF, 1 mM Na_3VO_4, 10% (v/v) glycerol, and 150 mM NaCl.

Lysis Buffer + Protease Inhibitor

1x lysis buffer, and 1x Calbiochem protease inhibitor (PI) cocktail III (added fresh, Fisher Scientific, 12818395).

SDS Loading Buffer

0.25% (w/v) bromophenol blue, 50% (v/v) glycerol, 10% (w/v) sodium dodecyl sulphate (SDS), 0.25 M Tris-HCl pH 6.8, and 10% (v/v) β-mercaptoethanol.

10X SDS Running Buffer

250 mM Tris pH 8.3, 1.92 M glycine, and 1% (w/v) SDS.

10X Transfer Buffer

250 mM Tris pH 8.3 and 1.92 M glycine.

1X Transfer Buffer

1X transfer buffer and 20% (v/v) methanol.

Wash Buffer

50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), and 0.1% (v/v) Tween-20.

Blocking Buffer

Wash buffer and 3% (w/v) bovine serum albumin (BSA).

RNA Immunoprecipitation Lysis Buffer

25 mM Tris-HCl pH 7.4, 150 mM KCl, 5 mM EDTA, 0.5 mM DTT, 0.5% NP-40, 100 U/ml RNase inhibitor (added fresh, New England Biolabs (NEB), M0314), and 1X Calbiochem protease inhibitor cocktail III (added fresh). Filtered using 0.22 μm filter.

RNA Immunoprecipitation Immunoprecipitation Buffer

10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, and 100 U/ml RNase inhibitor (added fresh). Filtered using 0.22 μm filter.

RNA Immunoprecipitation Wash Buffer

20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.4% NP-40 (v/v), and 100 U/ml RNase inhibitor (added fresh). Filtered using 0.22 μm filter.

Proteinase K Buffer

7.8 mM Tris-HCl pH 7.5, 117 mM NaCl, 0.39 mM EDTA, 1% (w/v) SDS, and 100 U/ml Proteinase K (add last, NEB, P8107). Filtered using 0.22 μm filter.

Annealing Buffer

10 mM Tris-HCl pH 7.5, 100 mM NaCl, and 1 mM EDTA.

Wash Buffer A

20 mM Tris-HCl pH 8, 150 mM NaCl, and 1 mM β-mercaptoethanol.

Wash Buffer B

20 mM Tris-HCl pH 8, 150 mM NaCl, 10% (v/v) glycerol, and 1 mM β-mercaptoethanol.

Glutathione-S-Transferase Elution Buffer

20 mM Tris-HCl pH 8, 150 mM NaCl, 20 mM reduced glutathione, and 1 mM β-mercaptoethanol.

Coomassie Protein Stain

0.1% (w/v) Coomassie Brilliant Blue, 40% (v/v) methanol, and 10% (v/v) glacial acetic acid (10%).

Gel Wash Buffer

40% (v/v) methanol and 10% (v/v) glacial acetic acid (10%).

Maltose-Binding Protein Elution Buffer

20 mM Tris-HCl pH 8, 150 mM NaCl, 10% (v/v) glycerol, 10 mM maltose, 1 mM β-mercaptoethanol.

Imidazole Elution Buffer A

20 mM Tris-HCl pH 8, 150 mM NaCl, 200 mM imidazole, and 1 mM β-mercaptoethanol.

Imidazole Elution Buffer B

20 mM Tris-HCl pH 8, 150 mM NaCl, 10% (v/v) glycerol, 200 mM imidazole and 1 mM β-mercaptoethanol.

Microscale Thermophoresis Buffer

20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% glycerol and 1 mM β-mercaptoethanol.

2.3 Plasmids

Table 2.3.1 Plasmids

Plasmids are listed along with their accession number, where applicable. Details are given of the application for which each plasmid was used, along with their major feature, their source, and, where applicable, their catalogue number. Amino-terminal (N-terminal), carboxy-terminal (C-terminal), New England Biolabs (NEB).

Plasmid	Application	Features	Source, Catalogue number
pEGFP-C1-AGO2	Mammalian Cell Expression	Full length AGO2 with N-terminal GFP-tag	Addgene (21981)
pEGFP-C1-AGO2-Y393F	Mammalian Cell Expression	Full length AGO2 with Y393F mutation and N-terminal GFP-tag	This study
pEGFP-C1-AGO2-Y322/529F	Mammalian Cell Expression	Full length AGO2 with Y322F and Y529F mutations and N-terminal GFP-tag	This study
pEGFP-N2-EGFR	Mammalian Cell Expression	Full length EGFR with C-terminal GFP-tag	Kin Man-Suen (University of Leeds, UK)
pEGFP-N2	Mammalian Cell Expression	GFP	Clontech
FLAG-FGFR2	Mammalian Cell Expression	Full length FGFR2 with N-terminal FLAG-tag	Natalie Stephenson (University of Leeds, UK)
pcDNA-3.1-RFP-N-GRB2	Mammalian Cell Expression	Full length GRB2 with an N-terminal RFP-tag	Chi-Chuan Lin (University of Leeds, UK)
pcDNA-3.1-RFP-N-GRB2-W36K	Mammalian Cell Expression	Full length GRB2 with a W36K mutation and N-terminal RFP-tag	Chi-Chuan Lin (University of Leeds, UK)

pcDNA-3.1-RFP-N-GRB2-R86A	Mammalian Cell Expression	Full length GRB2 with R86A mutation and N-terminal RFP-tag	Chi-Chuan Lin (University of Leeds, UK)
pcDNA-3.1-RFP-N-GRB2-W193K	Mammalian Cell Expression	Full length GRB2 with W193K mutation and N-terminal RFP-tag	Chi-Chuan Lin (University of Leeds, UK)
pcDNA-3.1-RFP-N-AGO2	Mammalian Cell Expression	Full length AGO2 with N-terminal RFP-tag	This study
pcDNA-3.1-RFP-N	Mammalian Cell Expression	Full length RFP	Tony Ng (King's College London, UK)
pcDNA-3.1	Mammalian Cell Expression	No insert	Tony Ng (King's College London, UK)
pDNA6/Myc-His B	Mammalian Cell Expression	Full length GRB2 with N-terminal Myc tag	Chi-Chuan Lin (University of Leeds, UK)
MSCV	Mammalian Cell Expression	No insert	Lin He (Addgene # 24828) (Olive et al., 2009)
MSCV-let-7-sponge	Mammalian Cell Expression	Sequence complementary to let-7 family miRNA	Phil Sharp (Addgene #29766) (Kumar et al., 2008)
pET28a-GRB2	Bacterial Cell Expression	Full length GRB2 with N and C-terminal His tags	(Ahmed et al., 2010)
pET28a-GRB2-W36K	Bacterial Cell Expression	Full length GRB2 with W36K mutation N and C-terminal His tags	Chi-Chuan Lin (University of Leeds, UK)
pET28a-GRB2-W193K	Bacterial Cell Expression	Full length GRB2 with W193K mutation N and C-terminal His tags	Chi-Chuan Lin (University of Leeds, UK)
pET28a-GRB2-W36/193K	Bacterial Cell Expression	Full length GRB2 with W36/193K mutation N and C-terminal His tags	Chi-Chuan Lin (University of Leeds, UK)
pET28a-PAZ	Bacterial Cell Expression	AGO2 PAZ domain (residues 228-348) N and C-terminal His tags	Chi-Chuan Lin (University of Leeds, UK)
pGEX2T-GRB2	Bacterial Cell Expression	Full length GRB2 with N-terminal GST-tag	(Ahmed et al., 2010)

pGEX2T-NSH3	Bacterial Cell Expression	GRB2 N-terminal SH3 domain (residues 1-58) with N-terminal GST-tag	(Ahmed et al., 2010)
pGEX2T-SH2	Bacterial Cell Expression	GRB2 SH2 (residues 60-152) domain with N-terminal GST-tag	(Ahmed et al., 2010)
pGEX2T-CSH3	Bacterial Cell Expression	GRB2 C-terminal SH3 domain (residues 156-215) with N-terminal GST-tag	(Ahmed et al., 2010)
pGEX-2T	Bacterial Cell Expression	GST alone	GE Healthcare
pMAL-c5x-N-terminal	Bacterial Cell Expression	AGO2 N-terminal domain (residues 1-139) with N-terminal MBP-tag	This study
pMAL-c5x-PAZ	Bacterial Cell Expression	AGO2 PAZ domain (residues 235-348) with N-terminal MBP-tag	This study
pMAL-c5x-MID	Bacterial Cell Expression	AGO2 MID domain (residues 450-573) with N-terminal MBP-tag	This study
pMAL-c5x-PIWI	Bacterial Cell Expression	AGO2 PIWI domain (residues 517-818) with N-terminal MBP-tag	This study
pMAL-c5x-PAZ-4A	Bacterial Cell Expression	AGO2 PAZ domain (residues 235-348) with ^{323}PHLP^{326}/AAAA mutation and N-terminal MBP-tag	This study
pMAL-c5x	Bacterial Cell Expression	MBP alone	NEB (N8108)
pLenti6/V5-DEST-FGFR2	Lentiviral Transduction	Full length FGFR2 with P810A/P813A mutation	Natalie Stephenson (University of Leeds, UK)
pCas9-Guide-gRNA1	GRB2 knockdown	Cas9, target sequence for gRNA against *GRB2*	OriGene (KN200469G2)
pCas9-Guide-gRNA2	GRB2 knockdown	Cas9, target sequence for gRNA against *GRB2*	OriGene (KN200469G1)

pCas9-Guide-scmbld	GRB2 knockdown control	Cas9, Scrmbld gRNA target sequence	OriGene (GE100003)
pUC	GRB2 knockdown	*GFP* gene without promoter and puromycin resistance gene under control of PGK promoter, both flanked by homologous arms targeted against *GRB2*	OriGene (KN200469-D)
lentiCRISPR-V2	GRB2 knockdown	Cas9, puromycin selectable marker, gRNA targeting GRB2	GenScript
pRL-TK (AF025846)	Luciferase Assay	Renilla luciferase gene under control of a thymidine kinase promoter	Promega (E2241)
pGL3-control-mir-17^{WT}	Luciferase Assay	pGL3-control containing mir-17 targeting site in *GRB2*	This study
pGL3-control-mir-17^{mut}	Luciferase Assay	pGL3-control containing mutated mir-17 targeting site in *GRB2*	This study
pGL3-control-mir-221-5p^{WT}-8mer	Luciferase Assay	pGL3-control containing miR-221-5p 8mer targeting site in *GRB2*	This study
pGL3-control-mir-221-5p^{mut}-8mer	Luciferase Assay	pGL3-control containing mutated miR-221-5p 8mer targeting site in *GRB2*	This study
pGL3-control-mir-221-5p^{WT}-7mer-A1_1	Luciferase Assay	pGL3-control containing miR-221-5p 7mer-A1 targeting site in *GRB2* 5'UTR	This study
pGL3-control-mir-221-5p^{mut}-7mer-A1_1	Luciferase Assay	pGL3-control containing mutated miR-221-5p 7mer-A1 targeting site in *GRB2* 5'UTR	This study
pGL3-control-mir-221-5p^{WT}-7mer-A1_2	Luciferase Assay	pGL3-control containing miR-221-5p 7mer-A1 targeting site in *GRB2* 3'UTR	This study

pGL3-control-mir-221-5p^{mut}-7mer-A1_2	Luciferase Assay	pGL3-control containing mutated miR-221-5p 7mer-A1 targeting site in *GRB2* 3'UTR	This study
pGL3-control-mir-221-3p^{WT}-7mer-A1	Luciferase Assay	pGL3-control containing miR-221-3p 7mer-A1 targeting site	This study
pGL3-control-mir-221-3p^{mut}-7mer-A1	Luciferase Assay	pGL3-control containing mutated miR-221-3p 7mer-A1 targeting site	This study
pGL3-control-GRB2	Luciferase Assay	pGL3-control containing *GRB2* complementary DNA	NovoPro, custom
pGL3-control (U47296)	Luciferase Assay	Firefly luciferase gene under control of SV40 promoter	Promega (E1741)

2.4 Primers

Table 2.4.1 Primer sequences for various applications

Unless otherwise stated, primers were custom ordered from Integrated DNA Technologies. Site-directed mutagenesis (SDM), quantitative polymerase chain reaction (qPCR). Lowercase 'p' indicates 5' phosphorylation.

Primer Name	Application	Sequence/Reference
AGO2 N-terminal domain FP	Molecular cloning	CATCATCATATGATGTACTCGGGAGCCGGCCCC
AGO2 N-terminal domain RP	Molecular cloning	CATCATGTCGACTTAGCTCACGCAGGACACCCACTT
AGO2 PAZ domain FP	Molecular cloning	CATCATCATATGTGTGAAGTTTTGGATTTTAAA
AGO2 PAZ domain RP	Molecular cloning	CATCATGTCGACTTATGCCACAATGTTACAGACCTC
AGO2 MID domain FP	Molecular cloning	CATCATCATATGGTGTGGGCCATTGCGTGCTTC

AGO2 MID domain RP	Molecular cloning	CATCATGTCGACTTAGCCTCCCAGCTTGACGTTGAT
AGO2 PIWI domain FP	Molecular cloning	CATCATCATATGCTGGTGGTGGTCATCCTGCCC
AGO2 PIWI domain RP	Molecular cloning	CATCATGTCGACTTACACCAGGTGGTACCTGGCCCG
AGO2 Y393F FP	SDM	CAACACAGATCCATTCGTCCGTGAATTTGG
AGO2 Y393F RP	SDM	CCAAATTCACGGACGAATGGATCTGTGTTG
AGO2 Y322F FP	SDM	GTTCTGCGCTTCCCCCACCTC
AGO2 Y322F RP	SDM	CAACTTGTGCCTGTCCTTGAAATAC
AGO2 Y529F FP	SDM	ACGCCCGTGTTCGCCGAGGTC
AGO2 Y529F RP	SDM	CTTGCCGGGCAGGATGAC
AGO2 ^{323}PHLP^{326}>AAAA FP	SDM	GCTGCAGCCGCATGTTTACAAGTCGGACAGGAG
AGO2 ^{323}PHLP^{326}>AAAA RP	SDM	GTAGCGCAGAACCAACTTGTG
mir-17 7mer-m8 WT *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACAAGGTCAGGAGT
mir-17 7mer-m8 WT *GRB2* TS antisense	Cloning luciferase vectors	pCTAGACTCCTGACCTTGTGATCCGCCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCT
mir-17 7mer-m8 mutant *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGCTCACGCCTGTAATCCCAAAAAAAAAGGAGGCCAAGGTGGGCGGATCACAAGGTCAGGAGT
mir-17 7mer-m8 mutant *GRB2* TS antisense	Cloning luciferase vectors	pCTAGACTCCTGACCTTGTGATCCGCCCACCTTGGCCTCCTTTTTTTTTGGGATTACAGGCGTGAGCCT
mir-221 (3p) 7mer-A1 WT *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGAAGTCATGGGAACGCAGCCTCCAGTTGTAGCAGGTTTCACTATTCCTATGCTGGGGTACAT

mir-221 (3p) 7mer-A1 WT *GRB2* TS antisense	Cloning luciferase vectors	pCTAGATGTACCCCAGCATAGGAATAGTGA AACCTGCTACAACTGGAGGCTGCGTTCCC ATGACTTCCT
mir-221 (3p) 7mer-A1 mutant *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGAAGTCATGGGAACGCAGCCTC CAGTTTTTTTTGGTTTCACTATTCCTATGCT GGGGTACAT
mir-221 (3p) 7mer-A1 mutant *GRB2* TS antisense	Cloning luciferase vectors	pCTAGATGTACCCCAGCATAGGAATAGTGA AACCAAAAAAAAACTGGAGGCTGCGTTCCCA TGACTTCCT
miR-221-5p 8mer WT *GRB2* TS sense	Cloning luciferase vectors	pCTAGATCCAGATATGAACAGGGCCCAGG CCTGGAGCGTTTGCTGTGCCAGGAGGCGG CAGCTCTTCTT
miR-221-5p 8mer WT *GRB2* TS antisense	Cloning luciferase vectors	pCTAGAAGAAGAGCTGCCGCCTCCTGGCA CAGCAAACGCTCCAGGCCTGGGCCCTGTT CATATCTGGAT
miR-221-5p 8mer mutant *GRB2* TS sense	Cloning luciferase vectors	pCTAGATCCAGATATGAACAGGGCCCAGG CCTGGAGCGTTTGCTGAAAAAAAAGGCGG CAGCTCTTCTT
miR-221-5p 8mer mutant *GRB2* TS antisense	Cloning luciferase vectors	pCTAGAAGAAGAGCTGCCGCCTTTTTTTTC AGCAAACGCTCCAGGCCTGGGCCCTGTTC ATATCTGGAT
miR-221-5p 7mer-A1_1 WT *GRB2* TS sense	Cloning luciferase vectors	pCTAGAAGCCAGGAGGTATTGCTGCTTCG GCGACCGGGCGGCGGCAGCGGCGGCGG CGGCTGTGGCAGT
miR-221-5p 7mer-A1_1 WT *GRB2* TS antisense	Cloning luciferase vectors	pCTAGACTGCCACAGCCGCCGCCGCCGCT GCCGCCGCCCGGTCGCCGAAGCAGCAATA CCTCCTGGCTT
miR-221-5p 7mer-A1_1 mutant *GRB2* TS sense	Cloning luciferase vectors	pCTAGAAAAAAAAAGGTATTGCTGCTTCGG CGACCGGGCGGCGGCAGCGGCGGCGGC GGCTGTGGCAGT
miR-221-5p 7mer-A1_1 mutant *GRB2* TS antisense	Cloning luciferase vectors	pCTAGACTGCCACAGCCGCCGCCGCCGCT GCCGCCGCCCGGTCGCCGAAGCAGCAATA CCTTTTTTTTT

miR-221-5p 7mer-A1_2 WT *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGCGTTCTCCTGCTTGTACTAAGC CAGGAGGCTTTAAGCTCCAGCTTTAAGGGT TGTGAGCCT
miR-221-5p 7mer-A1_2 WT *GRB2* TS antisense	Cloning luciferase vectors	pCTAGAGGCTCACAACCCTTAAAGCTGGAG CTTAAAGCCTCCTGGCTTAGTACAAGCAGG AGAACGCCT
miR-221-5p 7mer-A1_2 mutant *GRB2* TS sense	Cloning luciferase vectors	pCTAGAGGCGTTCTCCTGCTTGTACTAAAA AAAAAGGCTTTAAGCTCCAGCTTTAAGGGT TGTGAGCCT
miR-221-5p 7mer-A1_2 mutant *GRB2* TS antisense	Cloning luciferase vectors	pCTAGAGGCTCACAACCCTTAAAGCTGGAG CTTAAAGCCTTTTTTTTTAGTACAAGCAGGA GAACGCCT
GRB2 LHA FP	Amplification of genomic DNA	GTGGGGCTCAAAATAAGCTGC
GRB2 LHA RP	Amplification of genomic DNA	TTAAGGGCCAGCTCATTCCTC
GRB2 RHA FP	Amplification of genomic DNA	CATCTTGTTGGTCATGCGGC
GRB2 RHA RP	Amplification of genomic DNA	ATGGAAGTGCAGGGGGAAAG
GRB2 75278 FP	Amplification of genomic DNA	ACACCACACATCACACCAACT
GRB2 75726 RP	Amplification of genomic DNA	TGGCCACTGCTCTTAATCGT
GRB2 75278 NdeI FP	Amplification of genomic DNA	CATCATCATATGACACCACACATCACACCA ACT
GRB2 75726 SalI RP	Amplification of genomic DNA	CATCATGTCGACTGGCCACTGCTCTTAATC GT

let-7a-5p	qPCR	Hs_let-7a_2 miScript Primer Assay (Qiagen, MS00031220)
let-7d-5p	qPCR	Hs_let-7d_1 miScript Primer Assay (Qiagen, MS00003136)
let-7g-5p	qPCR	Hs_let-7g_2 miScript Primer Assay (Qiagen, MS00008337)
miR-17-5p	qPCR	Hs_miR-17_2 miScript Primer Assay (Qiagen, MS00029274)
miR-181c-5p	qPCR	Hs_miR-181c_2 miScript Primer Assay (Qiagen, MS00008841)
miR-19b-3p	qPCR	Hs_miR-19b_2 miScript Primer Assay (Qiagen, MS00031584)
miR-20a-5p	qPCR	Hs_miR-20a_1 miScript Primer Assay (Qiagen, MS00003199)
miR-221-5p	qPCR	Hs_miR-221*_1 miScript Primer Assay (Qiagen, MS00009170)
pre-let-7a-1	qPCR	Hs_let-7a-1_PR_1 miScript Precursor Assay (Qiagen, MP00000007)
pre-let-7g	qPCR	Hs_let-7g_1_PR miScript Precursor Assay (Qiagen, MP00000070)
pre-miR-17	qPCR	Hs_mir-17_1_PR miScript Precursor Assay (Qiagen, MP00001064)
pre-miR-181c	qPCR	Hs_mir-181c_1_PR miScript Precursor Assay (Qiagen, MP00001099)
pre-miR-19b-1	qPCR	Hs_mir-19b-1_PR_1 miScript Precursor Assay (Qiagen, MP00001386)
pre-miR-20a	qPCR	Hs_mir-20a_1_PR miScript Precursor Assay (Qiagen, MP00001484)
pre-miR-221	qPCR	Hs_mir-221_1_PR miScript Precursor Assay (Qiagen, MP00001617)
RNU6-6P	qPCR	TACAGAGAAGATTAGCATGG
GAPDH FP	qPCR	GGTGTGAACCATGAGAAGTATGA

GAPDH RP	qPCR	GAGTCCTTCCACGATACCAAAG
GRB2 FP	qPCR	TTCACTGCTGCTCCTCTTTC
GRB2 RP	qPCR	CTAGGACTATACGTGGCCTTAAAC
NRAS FP	qPCR	CAGTGCCATGAGAGACCAATAC
NRAS RP	qPCR	TCTGCTCCCTGTAGAGGTTAAT
HMGA1 FP	qPCR	GCCTGGGATCTGAGTACATATTG
HMGA1 RP	qPCR	CGGAAGCAAAGTAGGGTTAGG
HMGA2 FP	qPCR	ATGGTGCACCGAGTGATTT
HMGA2 RP	qPCR	TGGGACTTAAGAGGTAGTAGTGT
DICER1 FP	qPCR	CCCTGTTTGGTGCTATGTTTATG
DICER1 RP	qPCR	CTACTGGTATGTTGATGGGAG
LIN28B FP	qPCR	GGCAGCAGGTCTGTGAATTA
LIN28B RP	qPCR	GATCAGCTTACAGGGCTAAGAG

2.5 Molecular cloning

2.5.1 Transformation of bacterial cells with plasmid DNA

For all DNA production and manipulation, competent DH5α cells (NEB, C2987) were transformed with plasmid DNA. DH5α cells have *lacZ*ΔM15, *rec*A1 and *end*A mutations; they are therefore optimised for molecular cloning techniques and have high, good quality DNA yields. 20 µl cells were thawed on ice before 1 µl ice-cold plasmid was added. The cells were flicked to mix and incubated in ice for 10 mins. Cells were heat shocked at 42°C for 45 seconds, then immediately transferred to ice for another 5 mins. 950 µl lysogeny broth (LB)

was added and cells incubated at 37°C, 200 rotations per minute (rpm) for 1 h. Cells were then pelleted by centrifugation at 3380 xg for 1 min. The pellet was resuspended in 50 μl LB and spread on agar plates containing 100 μg/ml ampicillin or 50 μg/ml kanamycin as appropriate. Plates were incubated at 37°C for 16 h.

2.5.2 Making glycerol stocks

Following transformation of bacterial cells, glycerol stocks were made for long term storage of bacterial cells expressing plasmid. A single colony was expanded into 10 ml LB containing the appropriate antibiotic (2.5.1). Cultures were incubated at 37°C for 16 h at 200 rpm. 500 μl of the culture was mixed with 500 μl sterile 50% glycerol and stored at -80°C.

2.5.3 Plasmid preparation from bacterial cells

To prepare plasmid DNA for molecular cloning or mammalian cell transfection, DNA was purified from an overnight culture. Cells from either a glycerol stock or a single colony from agar plate were expanded into 10 ml LB, supplemented with the appropriate antibiotic (2.5.1), and incubated for 16 h at 37°C, 200 rpm. Cells were pelleted by centrifugation at 2200 xg for 10 mins. Plasmid DNA was purified from cell pellets using the QIAprep Spin Miniprep Kit (Qiagen, 27106) according to manufacturer's instructions. DNA concentration was then quantified at 260 nm, using a NanoDrop 2000 (Thermo Scientific, ND-2000).

2.5.4 Restriction digestion

Plasmid DNA and PCR products were cut in either a double or single restriction digest. A 20 μl reaction was set up using 1 μg DNA, 10 U enzyme (NEB, Table 2.5.1, 2.5.10), 1x buffer and nuclease-free water (Invitrogen, 10977023). Buffer

was CutSmart, 3.1, 2.1 or 1.1 (NEB, various) as dictated by the NEBcloner v1.4.2 double digest finder (NEB, 2020b). Reactions were incubated at 37°C for 30 min. DNA was purified by agarose gel electrophoresis (2.5.6) and gel extraction (2.5.7).

Table 2.5.1 Enzymes used in molecular cloning

Complementary DNA (cDNA)

Vector	Application	Enzyme(s)
pcDNA3.1-RFP-N	Insertion of *AGO2* cDNA	BamHI and NotI
pMALc5x	Insertion of AGO2 domain cDNA	NdeI and SalI
pGL3-control	Insertion of miRNA targeting sites in *GRB2*	XbaI
pMALc5x	Insertion of G1 genome around the start of *GRB2*	NdeI and SalI
Parental DNA	Site-directed mutagenesis	DpnI

2.5.5 Dephosphorylation of cut vectors

During restriction digestion, 5’ phosphates were removed from vector ends by addition of 1 U calf intestinal phosphatase (CIP, NEB, M0290) to the reaction.

2.5.6 Agarose gel electrophoresis

DNA of different sizes was separated by agarose gel electrophoresis. 1% (for DNA fragments >500 bp) or 2% (for DNA fragments <500 bp) agarose gel was made by dissolving agarose to the desired percentage in 1x Tris-Acetate-EDTA (TAE, National Diagnostics, EC-872). 1x Sybr safe (Invitrogen, S33102) was added, before the mixture was poured into a mould and left to set. The gel was set up in a DNA running tank and covered in 1x TAE. Purple loading dye (NEB, B7024S) was diluted to 1x with DNA and loaded into the wells of the gel. DNA

was separated by electrophoresis at 100 V until purple loading dye could be seen to have migrated 1/3 the length of the gel. 1 kb (NEB, N3232) and 100 bp (NEB, N3231) DNA ladders were used to estimate the size of DNA fragments.

2.5.7 Gel extraction

DNA fragments were purified from agarose gels with the QIAquick gel extraction kit (Qiagen, 28706), according to manufacturer's instructions.

2.5.8 Ligation

To ligate digested insert DNA into digested vector backbone, DNA was incubated with 1 μl T4 DNA ligase (NEB, M0202) and 1x T4 DNA ligase buffer. The reaction was incubated in a 20 μl volume, for 16 h at room temperature. A 3:1 insert:backbone ratio was used. Prior to transformation into cells, DNA was purified using the QIAquick purification kit (Qiagen, 28106), according to manufacturer's instructions.

2.5.9 Polymerase chain reaction of plasmid DNA

PCR amplification of a region of plasmid DNA was used to produce insert DNA with specific restriction sites, which could then be spliced into another vector. 10 ng DNA was amplified with 1 U Vent (NEB, M0254) and 1x ThermoPol reaction buffer (NEB, B9004) in a 50 μl reaction, with 1 μM forward and reverse primers and 400 μM dNTP mix (NEB, N0447). The reactions were placed in a thermocycler and DNA amplified under the conditions in Table 2.5.2. Annealing temperature was calculated using NEB Tm Calculator v1.12.0 (NEB, 2020c). After PCR, DNA was purified by agarose gel electrophoresis (2.5.6) and gel extraction (2.5.7).

Table 2.5.2 Cycling conditions for amplification of insert DNA from plasmid

Stage	Temperature	Time	Repeat
Initial Denaturation	95°C	5 min	-
Denaturation	95°C	30 sec	
Primer Annealing	55-65°C	30 sec	x30
Extension	72°C	1 min per kb	
Final Extension	72°C	10 min	-

2.5.10 Site directed mutagenesis

Mutations were introduced into plasmid DNA by site directed mutagenesis (SDM), using either the PrimeSTAR GXL polymerase (Takara Bio, R050, Table 2.5.3) or Q5 kit (NEB, E0554, Table 2.5.4). For both kits, reactions were made up according to manufacturer's instructions. After PCR, DNA amplified by the Q5 polymerase was incubated with the KLD enzyme mix from the same kit, as instructed in the handbook, before bacterial cells were transformed with 5 μl of the mix (2.5.1). DNA amplified with the PrimeSTAR polymerase was purified using the QIAquick purification kit (2.5.8) and parental plasmid DNA digested by incubating with DpnI (NEB, R0176, 2.5.4). Mutant plasmid DNA was again purified using the QIAquick purification kit, before being transformed into DH5α cells (2.5.1).

Table 2.5.3 Cycling conditions used for site-directed mutagenesis with PrimeSTAR polymerase

Stage	Temperature	Time	Repeat
Initial Denaturation	95°C	3 min	-
Denaturation	95°C	10 sec	
Primer Annealing	55°C, 60°C	15 sec	x34
Extension	68°C	1 min per kb	
Final Extension	68°C	10 min	-

Table 2.5.4 Cycling conditions use for site-directed mutagenesis with the Q5 kit

Primer annealing temperature was calculated using the NEBaseChanger tool v1.3.0 (NEB, 2020a).

Stage	Temperature	Time	Repeat
Initial Denaturation	95°C	30 sec	-
Denaturation	95°C	10 sec	
Primer Annealing	67°C, 70°C	30 sec	x34
Extension	72°C	30 sec per kb	
Final Extension	72°C	2 min	-

2.5.11 Extraction of genomic DNA from mammalian cells

Genomic DNA was purified from human embryonic kidney 293 SV40 T-antigen-expressing (HEK293T) cells using the DNeasy Blood and Tissue kit (Qiagen, 69504), following the manufacturer's protocol.

2.5.12 Polymerase chain reaction of genomic DNA

The start of the *GRB2* gene was amplified to genotype G1 clones. Reactions were set up in 50 μl, containing 500 ng genomic DNA, 1 U Phusion polymerase

(NEB, M0530), 1x Phusion HF buffer (NEB, B0518), 400 μM dNTP and 1 μM each primer. Cycling conditions were as stated in Table 2.5.5.

Table 2.5.5 Cycling conditions used to amplify a fragment of genomic DNA

Stage	Temperature	Time	Repeat
Initial Denaturation	95°C	5 min	-
Denaturation	95°C	30 sec	
Primer Annealing	56°C	1 min	x34
Extension	72°C	3 min	
Final Extension	72°C	10 min	-

2.5.13 Sequencing

Plasmids and PCR fragments were sequenced using either GENEWIZ or Source Bioscience.

2.6 Mammalian cell culture

2.6.1 Culture of mammalian cell lines

Cell lines were cultured in medium (for type see Table 2.6.1) supplemented with 50 μg/ml gentamicin (Gibco, 15710064) and fetal bovine serum (FBS, Sigma-Aldrich, F7524, for percentage see Table 2.6.1). Cells were maintained at 37°C with 5% CO_2 in T75 flasks (Corning, 10492371). When cells reached confluency, they were washed three times with phosphate buffered saline (PBS, Gibco, 10010056), followed by incubation in 5 ml 1x trypsin (Gibco, 15400054) for 5 mins at 37°C and 5% CO_2 to dissociate cells from the culture vessel. 0.5

ml of the dissociated cells was pelleted by centrifugation at 120 xg for 5 min, resuspended in culture media and transferred to a new T75 flask.

Table 2.6.1 Culture conditions for cell lines used in this study

Cells were maintained in media supplemented with fetal bovine serum (FBS) and gentemicin, at 37°C and varying levels of CO_2. Puromycin was added to the media to select stable cell lines. Dulbecco′s Modified Eagle′s Medium (DMEM), Minimum Essential Medium (MEM), Roswell Park Memorial Institute (RPMI), Iscove's Modified Dulbecco's Medium (IMDM).

Cell Line	Medium	FBS	CO_2	Puromycin
HEK293T	DMEM (Gibco, 11520416)	10%	5%	0.8 μg/ml
NIH3T3	DMEM	10%	5%	-
PC3	DMEM	10%	5%	-
MCF-7	DMEM	10%	5%	-
MDA-MB-231	DMEM	10%	5%	0.5 μg/ml
A498	MEM (Gibco, 11095080)	10%	5%	4 μg/ml
H520	RPMI 1640 (Gibco, 11835030)	10%	5%	0.5 μg/ml
Caco2	IMDM (Gibco, 12440061)	20%	5%	20 μg/ml

2.6.2 Freezing cell line stocks

Stocks of cell lines were made for long term storage. When a T75 flask of cells reached confluency, cells were dissociated using trypsin digestion (2.6.1). 2.5 ml of cells were pelleted by centrifugation at 180 xg for 5 min and resuspended in 500 μl Recovery cell culture freezing media (Gibco, 12648010). The cells were stored at -80°C for 24 h then transferred to liquid nitrogen.

2.6.3 Generation of a GRB2 knockout cell line

HEK293T cells with GRB2 knockout were generated using the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated endonuclease 9 (Cas9) GRB2 knockout kit (OriGene, KN200469), according to manufacturer's instructions. The kit relies on homologous recombination to insert episomal DNA into the start of the *GRB2* gene, both knocking out *GRB2* and knocking in GFP and puromycin resistance ($Puro^R$) genes.

HEK293T cells were transfected in a 6-well plate (Corning, CLS3516) at 60% confluency. Three transfections were performed according to the kit protocol; each of the three pCas9-guide plasmids (Table 2.3.1) was transfected simultaneously alongside pUC (Table 2.3.1). pCas9-guide encoded the guide RNA (gRNA) targeted to *GRB2* or the scrambled control, and the Cas9 protein (Table 2.6.2). pUC contained the GFP gene, with no promoter, and $Puro^R$, under control of the phosphoglycerate kinase 1 (PGK) promoter, (GFP-PGK-$Puro^R$) flanked by GRB2-directed 5' and 3' homology arms. The transfection reagent was TurboFectin (OriGene, TF81001). Cells were passaged after 48 h (2.6.1) and cultured until they reached P7, to allow both integration of the GFP-PGK-$Puro^R$ cassette and to dilute out cells which expressed extra-chromosomal $Puro^R$. As $Puro^R$ had its own promoter, cells which expressed this gene as the episomal form were still puromycin resistant.

Table 2.6.2 Sequences of guide RNAs used in Cas9-mediated knockdown of GRB2 in HEK293T cells

Guide RNA	Sequence
gRNA1	TTTGAAGTCATATTTGGCGA
gRNA2	TTTTGAAGCTCAGCTCGTCG

Cells were split into 10 cm dishes (Corning, CLS430591) in media containing 0.8 μg/ml puromycin to select for cells with $Puro^R$. From this point on, all cell lines generated using these cells were always cultured in media containing 0.8 μg/ml puromycin. Cells were then genotyped for the correct insertion. Genomic DNA was purified (2.5.11), before left and right homologous arms were amplified (2.5.12) and sent for sequencing (2.5.13).

Monoclonal cell lines were generated by seeding ~20 cells in a 10 cm dish and incubating until colonies could be seen by eye. Clones were selected by removing culture media, placing cloning disks (Sigma-Aldrich, Z374431-100EA) soaked in trypsin over the colonies and incubating the plate for 5 min at 37°C. The disks were added to wells of a 96-well plate (Corning, 10695951) containing culture media (Table 2.6.1), to allow attachment of cells. Once cells reached confluency, they were expanded into a larger culture vessel. The process was continued until enough cells were produced to freeze down stocks (2.6.2), analyse protein expression (2.7) and genotype.

To genotype G1 clones, genomic DNA was purified (2.5.11). A region at the start of the *GRB2* gene, containing the sequence against which the gRNA was targeted, was amplified by PCR (2.5.12). These PCR products were cloned into the pMAL-c5x vector (2.5.4 – 2.5.8) and transformed into DH5α cells (2.5.1).

DNA was purified from 10 colonies (2.5.3) and this was sent for sequencing (2.5.13).

2.6.4 Lentiviral transduction of FGFR2 expression

HEK293T cells with FGFR2 expression were generated by lentiviral transduction. To generate lentiviral stock, HEK293FT cells were initially split into a 6-well plate. When they reached 60-80% confluency, cells were transfected with 3 μl ViraPower™ Lentiviral Packaging Mix (Invitrogen, K4975–00), containing plasmids expressing virus envelope and packaging proteins, and 1 μg pLenti plasmid expressing the FGFR2 gene. The transfection was completed as described (2.6.7), using 10 μl Lipofectamine 2000 (Invitrogen, 11668019) as the transfection reagent and in a total volume of 500 μl.

After 24 h, cell media was replaced with fresh culture media and, after a further 24 h, media containing viral stock was collected. Media was centrifuged at 2000 xg, 4°C for 15 min, to pellet cell debris, and supernatant stored at -80°C.

HEK293T cells which were to be transduced with FGFR2 expression were split into a 6-well plate. When they reached 50-70% confluency, culture media was replaced with the thawed viral titre. 6 μg/ml polybrene was added to improve the efficiency of transduction. After 24 h the media was replaced with normal culture media containing 10 μg/ml blasticidin to select for cells with FGFR2 transduction. After a further 48 h, cells were split into 10 cm dishes. Cells were cultured for ten days, with fresh media being applied every three days. Clones were then selected by removing culture media, placing cloning disks soaked in trypsin over the colonies and incubating the plate for 5 min at 37°C. The disks were added to wells of a 96-well plate containing culture media, to allow attachment of cells. Once cells reached confluency, they were expanded into a

larger culture vessel. The process was continued until enough cells were produced to freeze down stocks (2.6.2) and analyse protein expression (2.7).

2.6.5 Knockdown of GRB2 in cancer cell lines

GRB2 was knocked down in various cell lines by lentiviral-mediated delivery of lentiCRISPR-V2 (GenScript) (Sanjana et al., 2014). The lentiCRISPR-V2 plasmid contains sequences encoding Cas9 and gRNA, as well as a selectable puromycin marker. gRNA sequences had been designed to target *GRB2* and have minimal binding to other regions of the genome (Table 2.6.3, GenScript) (Sanjana et al., 2014). Cells at 50% confluency were infected by application of a 1:1 mixture of normal culture media (Table 2.6.1) and lentiviral titre. Various mixtures of virus containing different gRNA sequences were used to enhance knockdown (Table 2.6.3). After 3 days, the media was removed and replaced with media supplemented with puromycin (for concentrations see Table 2.6.1) to select modified cells. Infected and parental control cells were passaged in media containing puromycin until all control cells were dead. GRB2 expression was then examined by western blotting (2.8, 2.7).

Table 2.6.3 Sequences of guide RNAs used in Cas9-mediated knockdown of GRB2 in cancer cell lines

Guide RNA	Sequence
1	GCAGCAGCCGACATACGTCC
2	ATTATGTCACCCCCGTGAAC
3	GCACCTGTTCTATGTCCGC

2.6.6 Kill curves

The minimum concentration of puromycin needed to select GRB2 knockout and knockdown (kd) cells was determined using a kill curve. Cells were seeded in 96-well plates to reach 50% confluency the next day. Media was then supplemented with media containing puromycin at increasing concentrations spanning 0-100 μg/ml. Fresh media was applied every 2-3 days for a total of 10 days. The minimum puromycin concentration required to kill all cells at this time was determined by cell counting.

2.6.7 Transfection

Mammalian cells were manipulated to express exogenous protein by transfection of plasmid DNA. HEK293T cells were split into 10 cm dishes or 6-well plates and transfected at 50% confluency. When 10 cm dishes were used, 5 μg plasmid DNA and 25 μl Metafectene (Biontex, T020) were each incubated in 500 μl Opti-MEM (Gibco, 31985062) for 5 min, room temperature. 2 μg plasmid DNA and 7 μl Metafectene was used for cells in 6-well plates. The plasmid DNA was added to the Metafectene and mixed by slowly pipetting up and down once, to prevent shearing of the DNA. The mixture was incubated for 15 min, at room temperature, and then added dropwise to the cells. Unless stated otherwise, protein expression was analysed 48 h post-transfection.

Breast cancer cell lines were also transfected in 10 cm dishes using the above protocol, but with 15 μg plasmid DNA and 30 μl *Trans*IT-BrCa (Cambridge Bioscience, MIR 5500) as the transfection reagent. MCF-7 cells were transfected at 50% confluency and MDA-MB-231 cells at 70% confluency.

HEK293T cells were transfected with a miRNA mimic (Sigma-Aldrich, Table 2.6.3) using the same protocol, with the following alterations. Cells were

transfected at 20% confluency in 6-well plates. The miRNA mimic was added to give final concentration 125 nM upon addition of the transfection mixture to the cells. As before, transfections were complete 48 h before analysis of protein and mRNA expression.

Table 2.6.3 miRNA mimics

Target	Catalogue Number
miR-17-5p	HMI0264
miR-20a-5p	HMI0367
miR-106a-5p	HMI0037
miR-221-5p	HMI0397
miR-221-3p	HMI0398
miR-222-3p	HMI0400
Negative control	HMC0002

2.6.8 Serum starvation

Cells were starved of serum by removal of media containing 10% FBS. They were washed three times with 10 ml PBS and then cultured in DMEM, supplemented only with 50 μg/ml geneticin and any appropriate antibiotics. HEK293T cells were starved for 16 h, whereas MDA-MB-231 and MCF-7 cells were starved for 18 h. 4 h prior to lysis, starvation media was removed from the breast cancer cell lines and they were washed as before, before fresh starvation media was added.

2.6.9 Stimulation

Prior to stimulation, cells were starved (2.6.8). FGF-stimulated cells were stimulated with 20 ng/ml FGF2, 7, 9, 10 (R&D Systems, 233-FB, 251-KG, 273-F9, 345-FG) or a cocktail of FGF2 and 9 for 15 min, at 37°C, 5% CO_2. EGF-

stimulated cells were stimulated with 10 ng/ml EGF (R&D Systems, 236-EG) for 5 min, at 37°C, 5% CO_2.

2.6.10 Kinase inhibition

FGFR2 was inhibited by addition of 30 nM SU5402 (Cambridge Bioscience, SM99) for 16 h prior to lysis. To inhibit PTP1B, media was supplemented with 10 μM CinnGEL 2-methylester (Insight Biotechnology, sc-205633) 2 h before lysis. Stock solutions of both inhibitors were made by dissolving in dimethyl sulphoxide (DMSO) to 1000x their working concentration, which was then diluted into starvation media. An equal volume of DMSO was added to control cells.

2.7 Western blotting

2.7.1 SDS-PAGE

Proteins studied by *in vitro* techniques utilising mammalian cell lysates (2.8, 2.9, 2.10, 2.19) and bacterial protein expression (2.18, 2.19, 2.23) were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). SDS loading buffer was added to lysates to 1x concentration and incubated at 95°C for 15 min to denature and add negative charge to the proteins. Samples were briefly centrifuged at 21,130 xg to collect condensate. Supernatant was then loaded into wells of a 4-20% polyacrylamide precast gel (Bio-Rad, 4561094), along with prestained protein standards (Bio-Rad, 1610374). Proteins were separated by electrophoresis at a constant voltage of 120 V for 50 min in SDS running buffer.

2.7.2 Transfer

After SDS-PAGE, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane using either semi-dry or dry transfer. Semi-dry transfer was completed using extra thick filter paper (Bio-Rad, 1703965) and membrane (GE Healthcare, 10600038). Two sheets of filter paper were soaked for 10 min in transfer buffer and the membrane was dipped in methanol for 1 min before use. The components were layered in a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad, 1703940) in the following order: filter paper, membrane, gel, filter paper. Proteins were transferred at 20 V for 55 min. Alternatively, proteins were transferred using the iBlot 2 Gel Transfer Device (Invitrogen, IB21001) and PVDF transfer stacks (Invitrogen, IB24001), according to manufacturer's instructions.

2.7.3 Ponceau staining

To detect high levels of protein, PVDF membranes were rinsed briefly in in distilled water (dH_2O) and then ponceau stained. Membranes were incubated in ponceau stain (Sigma-Aldrich, P3504) for 10 min with rocking. All following membrane incubations were completed with rocking. They were then washed by incubating in dH_2O, until background stain was removed. Membranes were photographed on a G:Box Chemi XT4 (Syngene) using manual settings, white light and automatic exposure. Stain was removed by washing with wash buffer.

2.7.4 Immunoblotting

Following transfer, membranes were rinsed briefly in dH_2O and blocked for at least 30 min in blocking buffer. Primary antibody was added at the appropriate dilution (Table 2.1.1) and the membrane incubated at 4°C for 16 h. The membrane was washed in wash buffer for 1 h at room temperature, with six

buffer changes. Next, the membrane was incubated in blocking buffer and 1:1000 diluted horse-radish peroxidase (HRP)-conjugated secondary antibody (Table 2.1.1), at room temperature, for 1 h. Membranes were then washed as before.

Membranes were imaged by enhanced chemiluminescence; the membrane was incubated in equal volumes of peroxide and luminol western ECL substrates (Bio-Rad, 1705061) for 1 min. Light emissions were then visualised on a G:Box Chemi XT4 using manual settings. If bands were faint, the process was repeated using strong western blotting substrate (Thermo Scientific, 34094) diluted 10x in western ECL substrates.

2.8 Mammalian cell lysis

Mammalian cells were lysed by removal of culture media, washing in PBS and incubation in lysis buffer +PI. 400 μl buffer was added to 10 cm dishes and 50 μl to 6-well plates. Cells were incubated on ice for 5 min, before cells were scraped from the base of the culture vessel and transferred to 1.5 ml tubes. Cells were incubated on ice for a further 15 min, then centrifuged at 21,130 xg, 4°C for 15 min to pellet debris. Cell lysate was transferred to a fresh tube.

Protein concentration was then measured by Bradford assay. 1 μl lysate (or lysis buffer +PI as a blank) was added to 1 ml Pierce Coomassie plus Bradford assay reagent (Thermo Scientific, 23238) and absorbance measured at 600 nm using a spectrophotometer (6320D spectrophotometer, Jenway, 632 501).

2.9 Immunoprecipitation

Protein complexes were immunoprecipitated from mammalian cell lysate following lysis and quantification of protein concentration (2.8). 1 mg of cell lysate was used, and reactions set up in 500 μl, containing the appropriate dilution of antibody (Table 2.1.1). Normal IgG was used as a control. Equal masses of test and normal IgG were added. Reactions were incubated with rotation at 4°C for 16 h. 20 μl protein A/G-plus agarose beads (Santa Cruz, sc-2003) were washed once in lysis buffer and added to the reaction. Reactions were incubated for 1 h at room temperature, with rotation. Lysate was then removed by centrifuging at 100 xg to pellet beads and taking off the supernatant. Beads were washed three times in 1 ml lysis buffer, centrifuging as before. Bound proteins were then released by boiling beads in 50 μl 2x SDS loading buffer and analysed by SDS-PAGE and western blotting (2.7).

2.10 GFP/RFP Trap

GFP and RFP-tagged proteins were precipitated from mammalian cell lysates overexpressing the appropriate protein (2.6.7). Following cell lysis (2.8), reactions were set up containing 1 mg cell lysate in 500 μl and 5 μl GFP/RFP-Trap magnetic agarose (ChromoTek, gtma-20/rtma-20). Prior to use, magnetic beads were washed once in 1 ml lysis buffer by separating the beads from supernatant using a DynaMag-2 magnet (Invitrogen, 12321D). Reactions were incubated either for 1 h at room temperature or 16 h at 4°C, with rotation. Supernatant was removed as before, and beads washed three times in 1 ml lysis buffer, before bound proteins were released by boiling in 50 μl 2x SDS loading buffer and analysed by SDS-PAGE and western blotting (2.7).

2.11 Mass spectrometry

Transiently overexpressed GFP-AGO2 (2.6.7) was purified from HEK293T cells by GFP-Trap (2.10). After SDS-PAGE (2.7.1), the protein gel was stained in InstantBlue Coomassie protein stain (Expedeon, ISB1L) for 15 min. The band corresponding to GFP-AGO2 was excised and mass spectrometry was completed by the Proteomics Facility at the University of Bristol, UK. Briefly, the protein was subject to in-gel trypsin digestion by the DigestPro automated digestion system (CEM). Peptides were then analysed on a Bruker Daltonics UltrafleXtreme 2 mass spectrometer (Bruker) by matrix-assisted laser desorption/ionization-time of flight-tandem mass spectrometry (MALDI-ToF-MS/MS). Reads were searched for peptide modifications against the human Uniprot database and the GFP-AGO2 sequence using SEQUEST (Eng et al., 1994). In addition to serine, threonine and tyrosine phosphorylation, oxidation and carbamidomethylation were considered. Data was filtered using a 5% false discovery rate.

2.12 Förster resonance energy transfer

HEK293T cells were seeded onto glass coverslips (Fischer Scientific, 12313138), in a 12-well plate (Corning, CLS3513). Prior to seeding, coverslips were rinsed in 70% ethanol and left to dry before they were washed three times in PBS. 24 h later, cells were transfected with RFP and GFP-tagged proteins (2.6.7) and after a further 24 h were serum-starved (2.6.8). Following this, cells were washed three times in PBS and fixed by incubating in 4% paraformaldehyde (PFA, Boster Biological Technology, AR1068) in PBS for 20 min at room temperature. Coverslips were then adhered to microscope slides

(Thermo Fischer, 10756991) using mounting media (Sigma-Aldrich, F6057) and left to dry.

Fluorescence signals were recorded on LSM880 + airyscan inverted or upright microscope (Zeiss) at the University of Leeds Bioimaging and Flow Cytometry facilities. RFP fluorescence was excited at 561 nm and detected between 607-633 nm; GFP fluorescence excited by 488 nm and detected between 499-544 nm; and FRET signal excited with 488 nm and detected between 610-693. The detector gain was 1.0.

2.13 RNA extraction and complementary DNA synthesis

2.13.1 RNA extraction

Total cell RNA was isolated from HEK293T cells in either a 6-well plate or 10 cm dish. Cells were homogenised by pipetting up and down several times in 750 μl TRIzol Reagent (Invitrogen, 15596026) and incubating for 5 min at room temperature. Chloroform was then added at a dilution of 1 in 5 and shaken vigorously for 15 sec. Homogenates were incubated for a further 2 min at room temperature, then centrifuged at 12,000 xg for 15 min to separate the organic phase, containing DNA and protein, interphase and aqueous phase, containing RNA. 3/4 the volume of the aqueous phase was taken, and the chloroform extraction was repeated. The aqueous phase was then mixed with an equal volume of isopropanol and 1 μl glycogen (Thermo Scientific, R0551), and incubated for 16 h at -20°C to precipitate RNA.

Precipitated RNA was pelleted by centrifugation at 12,000 xg for 10 min at 4°C. The pellet was washed twice in 1 ml 80% ethanol, vortexing the pellet each time and centrifuging at 7500 xg for 5 min at 4°C. The pellet was air dried for 5-10

min and then dissolved in 20 μl nuclease-free water. RNA concentration was determined at 260 nm on a NanoDrop 2000.

2.13.2 DNase treatment

Contaminating DNA was removed from purified RNA using the DNA-free DNA removal kit (Invitrogen, AM1906). Reactions were completed in a 20 μl volume containing 1x DNase I buffer. Where RNA concentration was under 200 ng/μl, 1 μl DNase was added to 20 μl RNA and the reaction was incubated for 30 min at 37°C. Higher concentrations of RNA were diluted to 200-500 ng/μl, before reactions were initially set up containing 1.5 μl DNase and the reaction incubated as before. Another 1.5 μl DNase was then added and the reaction incubated as previously. DNase treatment was stopped by addition of 2-4 μl DNase inactivation reagent, dependent on the total volume of DNase added, and incubating reactions for 2 min at room temperature. RNA was separated from the inactivation reagent by centrifugation at 10,000 xg and 4°C for 1.5 min and transferring to a new tube.

2.13.3 Reverse transcription

Complementary DNA (cDNA) was synthesised from total cell RNA using the miScript II reverse transcriptase kit (Qiagen, 218161), according to manufacturer's instructions. 1 ug of RNA was used in combination with the HiFlex buffer, which contains both oligo-dT and random primers to allow subsequent amplification of miRNA alongside precursor miRNA and mRNA. All RNA species are first polyadenylated. In the same reaction, mature miRNA is then converted to cDNA using the oligo-dT primers, whereas precursor miRNA and mRNA is reverse transcribed using both oligo-dT and random priming. The

oligo-dT primers contain the universal tag which is used to amplify mature miRNA only during qPCR (see 2.15).

2.14 Small RNA sequencing

Total cell RNA was extracted (2.13.1) and DNase-treated (2.13.2). Library synthesis and sequencing was carried out by Ragini Medhi (University of Cambridge, UK). Library preparation and sequencing were done using the NEXTFLEX small RNA-seq kit (Perkin Elmer, NOVA-5132). Briefly, 3' and 5' adapters were ligated, before the first strand of cDNA was synthesised by reverse transcription. cDNA was amplified by PCR, then cDNA ~150 bp (corresponding to RNA ~20 nt) was selected by PAGE. Sequencing was performed on the HiSeq 1500 sequencer (Illumina). Sequence analysis was done by Dapeng Wang (University of Leeds, UK). Adaptor sequences were trimmed using Cutadapt (Martin, 2011) and read quality assessed with FastQC (Babraham_Institute, 2019). Sequencing reads were aligned to the human genome assembly GRCh38 (GENCODE, Release 31, Frankish et al., 2021) using Bowtie (Langmead et al., 2009). A count matrix was produced using featureCounts (Liao et al., 2014) which was based on the microRNA annotation file from miRbase (v22, Kozomara et al., 2019), before differential expression was analysed using the DEseq2 package (Love et al., 2014).

2.15 Quantitative polymerase chain reaction

qPCR was done using the miScript SYBR green PCR kit (Qiagen, 218073). Reactions were made up as stated in the manufacturer's handbook in a total volume of 20 μl and using 3-20 ng cDNA per reaction. Cycling conditions were as per the manufacturer's instructions, in a Rotor-Gene-Q cycler (Qiagen) and

using a rotor-disk 72 rotor (Qiagen, 9018899). Precursor and mature miRNA expression was normalised to the small nuclear RNA RNU6-6P, whereas the housekeeping gene GAPDH was used for mRNAs. Fold change was calculated using the delta-delta-ct method.

2.16 RNA immunoprecipitation

In RIP experiments, AGO2 was immunoprecipitated and co-precipitated RNA was analysed. Mammalian cells were lysed (2.8) using RIP lysis buffer. For each reaction, 25 μl Magna ChIP protein A/G magnetic beads (Sigma-Aldrich, 16-663) were prepared by washing twice in 0.5 ml RIP wash buffer. Antibody was bound by incubating in 100 μl RIP wash buffer with the appropriate dilution of antibody (Table 2.1.1) for 40 min at room temperature, on rotation. Normal IgG was used as a control. The beads were washed as before. Reactions were set up containing 100 μl cell lysate and 900 μl RIP IP buffer, then incubated at 4°C with rotation, for 16 h. The beads were washed six times in RIP wash buffer.

RNA was then purified from each RIP and 10 μl whole cell lysate. Reactions were made up to 150 μl with proteinase K buffer and incubated at 55°C for 30 min to digest protein. 100 μl RIP wash buffer and 750 μl TRIzol were added and RNA was extracted (2.13.1). Whole cell RNA was DNase treated (2.13.2), then cDNA synthesised from all samples (2.13.3). RNA was quantified by qPCR (2.15), normalising to GAPDH and quantifying fold change by the delta-delta-ct method.

2.17 Luciferase assay

miRNA targeting sites in the *GRB2* transcript were identified by luciferase assay, as described previously (Jin et al., 2013).

2.17.1 Synthesis of plasmids containing potential targeting sites in GRB2

miRNA targeting sites were predicted by both using TargetScan 7.2 (Lewis et al., 2005; TargetScan, 2020) and manual searching of the *GRB2* transcript. 5' phosphorylated oligos containing the predicted sites (Table 2.4.1) were annealed to their complementary strand by incubating at 10 μM in annealing buffer at 95°C for 6 min, then at room temperature for 30 min. Annealed oligos were spliced into the XbaI-digested, CIP-dephosphorylated pGL3-control vector and confirmed by sequencing (2.5.4, 2.5.1, 2.5.3, 2.5.13)

2.17.2 Assay for luciferase activity

HEK293T cells were transfected in 12-well plates with 1 μg pGL3 plasmid, 5 ng pRL-TK plasmid and 100 nM miRNA mimic (Sigma-Aldrich, Table 2.6.3, 2.6.7). After 48 h, luciferase activity was quantified on a POLARstar Omega microplate reader (BMG LABTECH, Centre for Biomolecular Interactions, University of Leeds) using the Dual-Luciferase reporter assay system (Promega, E1910), according to manufacturer's instructions.

2.18 Protein production

2.18.1 Bacterial cell transformation for protein expression

To prepare a stock of bacterial cells expressing a plasmid, the same protocol was used as for molecular cloning (2.5.1). However, if the stock was for protein expression, competent BL21(DE3) cells (NEB, C2527) were used. BL21(DE3)

cells do not express Lom and OmpT proteases to maximise protein expression. They do express the T7 polymerase, which binds T7 promoters to induce protein expression upon addition of isopropyl β-d-1-thiogalactopyranoside (IPTG, Calbiochem, 5820).

2.18.2 Protein Expression

Recombinant protein was expressed by incubating cells at 37°C, 200 rpm in LB supplemented with the appropriate antibiotic (2.5.1) until optical density (OD) 0.8-1.0 was reached. Protein expression was induced with 0.1% IPTG at 20°C for 16 h. Cells were then harvested by centrifuging at 5500g using the Fiberlite F8S-6 x 1000y rotor (Thermo Scientific, 096-061137) and Sorvall RC BIOS 6 centrifuge system (Thermo Scientific, 75007044-7), 20°C, for 10 min.

2.18.3 Purification of GST-tag proteins

Cell pellets were resuspended in 5 ml wash buffer A per 1 L LB and lysed by sonication (Fisherbrand Model 120 Sonic Dismembrator, Fisher Scientific, 12337338). During sonication cells were kept on ice. Cell debris was pelleted by centrifugation at 48000 xg using a SS-34 rotor (Thermo Scientific, 28020TS) for 1 h. This and all subsequent steps were performed at 20°C, for GRB2 proteins, or room temperature, for all other proteins. Glutathione superflow resin (Clontech, 635608) was washed once in wash buffer A and incubated with lysates for 1 h on rotation, to allow glutathione-S-transferase (GST)-tagged protein to bind. 0.5 ml beads were used per 1 L LB. The beads were washed with wash buffer A in a gravity flow column until no colour change was observed upon mixing 10 μl eluted wash buffer with 200 μl Bradford reagent. GST-tagged protein was eluted with 5 ml glutathione elution buffer. Protein purity was analysed by SDS-PAGE (2.7.1) and staining with Coomassie dye for 20

minutes, followed by washing in protein gel wash buffer until bands were visible. Stained gels were imaged using visible light and automatic exposure on a G:box Chemi XT4. Concentration was determined by its absorbance at 280 nm on a NanoDrop 2000.

2.18.4 Purification of MBP-tag proteins

Cell pellets were resuspended in 5 ml wash buffer B per 1 L LB and lysed by sonication (2.18.3). Cell debris was pelleted by centrifugation at 48000 xg using a SS-34 rotor (Thermo Scientific, 28020TS) for 1 h. This and all subsequent steps were performed at 4°C. Maltose binding protein (MBP)-tag proteins were purified using a 5 ml MBPTrap HP column (GE Healthcare, 28918778). The column was equilibrated with 10 ml wash buffer B before lysate was applied. Non-specifically bound protein was removed by washing with 1 L wash buffer B. MBP-protein was eluted by application of 50 ml maltose elution buffer and collected in 5 ml fractions. The flow rate for all steps was 2.5 ml/min, except for sample application which was 0.5 ml/min. Protein concentration was determined by its absorbance at 280 nm on a NanoDrop 2000 and purity was analysed by SDS-PAGE (2.7.1) and Coomassie staining (2.18.3).

Protein was purified further by size exclusion chromatography (SEC). A Sephadex G-200 column (271233, Sigma Aldrich) was equilibrated in 200 ml wash buffer B, then protein was applied. Proteins were eluted by application of wash buffer B and collected in 2 ml fractions. All steps were completed at flow rate 0.9 ml/min. Purity was analysed by SDS-PAGE (2.7.1) and Coomassie staining (2.18.3).

2.18.5 Purification of His-tag proteins

Cell pellets were resuspended in wash buffer A, for GRB2 proteins, or wash buffer B, for PAZ. 5 ml was used per 1 L LB. Cells were lysed by sonication (2.18.3). Cell debris was pelleted by centrifugation at 48000 xg using a SS-34 rotor (Thermo Scientific, 28020TS) for 1 h. For this and all subsequent steps GRB2 protein was kept at 20°C and PAZ at 4°C. TALON metal-affinity resin (Takara Bio, 635504) was washed once in wash buffer A and incubated with lysates for 1 h on rotation, to allow His-tagged protein to bind. 0.5 ml beads were used per 1 L LB. The beads were washed with wash buffer A in a gravity flow column until no colour change was observed upon mixing 10 μl wash buffer with 200 μl Bradford reagent. His-tagged GRB2 proteins were eluted with 5 ml imidazole elution buffer A and His-PAZ with imidazole elution buffer B. Protein purity was analysed by SDS-PAGE (2.7.1) and Coomassie staining (2.18.3), and concentration was determined by its absorbance at 280 nm on a NanoDrop 2000.

2.18.6 RNase treatment

If stated, contaminating RNA was removed from preparations of recombinant AGO2 PAZ domain by incubating cell lysates with 20 μg/ml RNase H for 16 h at 4°C.

2.19 Pulldown using GST-tagged protein

GST-GRB2 domains were purified (2.18.3), but not eluted from beads. The amount of bound protein was quantified by SDS-PAGE (2.7.1) and Coomassie staining (2.18.3). The volume of beads for each protein was adjusted to ensure equal amounts of each, then volumes were made equal with unbound

glutathione superflow resin. A maximum volume of 50 μl of beads was used. Beads were washed once in lysis buffer and used to pull down endogenous protein from mammalian lysate, following the same protocol as for GFP/RFP Trap (2.10). Bound protein was analysed by SDS-PAGE, ponceau staining of GST-protein, and western blotting (2.7.1-2.7.4)

2.20 Isothermal titration calorimetry

Prior to ITC, all proteins were dialysed into matched wash buffer A. Peptides were dissolved into matched wash buffer A. Concentration was estimated using absorbance at A280 on a NanoDrop 2000, due to presence of a single tyrosine residue in each peptide.

1.2 mM peptide was titrated in a solution of 80 μM His-tagged GRB2, with an initial injection of 0.5 μl followed by 13 3 μl injections and spacing of 90 seconds. The experiment was completed at 25°C. Alternatively, 500 μM MBP-tagged protein was titrated into 20 μM His-tagged GRB2. An initial 0.5 μl injection was followed by 19 2 μl injections, spaced every 120 seconds, at 25°C. Experiments were performed on a Microcal iTC200 (Malvern Analytical, Centre for Biomolecular Interactions, University of Leeds). Excluding the initial peak, peaks were integrated to give a Wiseman curve, which was fitted with Microcal Origin using a 1:1 binding model to yield the binding affinity, K_D, and stoichiometry, n.

2.21 Microscale thermophoresis

2.21.1 Protein labelling

Prior to protein labelling and MST, proteins were dialysed into MST buffer. 100 μM GRB2 protein was labelled by incubating with 200 μM Atto 488 NHS ester (Sigma-Aldrich, 41698) in a 100 μl volume, in 0.1 μM sodium bicarbonate pH 8.3. The reaction was completed for 1 h at room temperature, in the dark. Labelled GRB2 was separated from unconjugated dye on a PD10 desalting column (GE, 17-0851-01) and eluted in MST buffer.

2.21.2 AGO2 PAZ with GRB2

Prior to MST, His-tagged PAZ domain dialysed into MST buffer. A two-fold dilution series of unlabelled PAZ was set up and mixed with labelled GRB2, yielding final concentrations of 100 nM GRB2 and 20-653,000 nM PAZ. Proteins were transferred to capillaries (Monolith, MO-AK005) and fluorescence was recorded using MST instrument (Monolith, NT.115, Centre for Biomolecular Interactions, University of Leeds). Atto 488 was excited with a blue LED set to a power level that achieved fluorescence between 800 and 1200 count. MST power was used at 20% and reactions were performed in triplicate.

2.21.3 GRB2 with DICER1

Protein preparation and MST experiments done by Chi-Chuan Lin (University of Leeds, UK). A dilution series of 3.66 to 120,000 nM unlabelled GRB2 was set up an combined with 100 nM labelled DICER1. A DICER1 construct was used which contained residues 1276-end (RNase III domains and dsRBD, $DICER1^{CT}$).

2.21.4 AGO2 with DICER1

Protein preparation and MST experiments done by Chi-Chuan Lin (University of Leeds, UK). AGO2 was preincubated with 0, 25 or 50 μM GRB2. AGO2 was then labelled and mixed with unlabelled $DICER1^{CT}$. $DICER1^{CT}$ was used in a dilution series from 3.66 to 120,000 nM.

2.22 Circular dichroism

Circular dichroism (CD) studies were completed with the assistance of G Nasir Khan (University of Leeds, UK). Experiments were completed using 200 μl 0.2 mg/ml protein in a 1 mm path-length curvette, on a Chirascan CD spectropolarimeter (Applied Photophysics). Far ultraviolet (UV) absorbance was measured over wavelengths 190-260 nm.

2.23 Protein crosslinking

Chemical crosslinking of His-PAZ and GRB2 was completed using disuccinimidyl tartrate (DST, Thermo Scientific, 20589). DST was dissolved to 25 mM in DMSO. DST was added dropwise to a solution of PAZ and GRB2 in wash buffer B and mixed by inverting. The final concentrations were 2.5 mM DST and 250 μM each protein. The reaction was incubated for 20 min at 21°C, then quenched by the addition of Tris pH 8 to a final concentration of 20 mM.

After crosslinking, protein complexes of different sizes were separated by SEC using a Superdex Increase 75 column (Sigma-Aldrich, GE29-1487-21). The column was equilibrated with 200 ml wash buffer B. Sample was applied, then eluted by washing with 200 ml wash buffer B. Protein was fractionated in 2 ml.

All steps were completed at a flow rate of 1 ml/min. Protein complexes were analysed SDS-PAGE (2.7.1) and staining with Coomassie dye (2.18.3).

2.24 *In silico* docking

Docking experiments were performed by Chi-Chuan Lin (University of Leeds, UK). Docking of GRB2 (PDB: 1GRI, Maignan *et al.*, 1995) to AGO2 PAZ domain (PDB: 6RA4, Greenidge *et al.*, 2019), was performed using the HADDOCK2.2 web server (van Zundert et al., 2016). To facilitate the SH3-PxxP-mediated complex, the ^{323}PHLP^{326} motif on AGO2 PAZ and hydrophobic binding pocket residues in GRB2 CSH3 were defined as the 'active residues (directly involved in binding)'. GRB2 residues tyrosine 7, phenylalanine 9, tryptophan 36, phenylalanine 47, proline 49, and tyrosine 52 were selected (Vidal et al., 1999). Passive residues around the active residues were automatically included in the docking process. Both defined structures were prepared using pymol.

2.25 Analysis of miRNA expression in human cancer

2.25.1 OnocmiR miRNA expression database

Expression of miRNAs identified in our screen was analysed across different human cancers using OncomiR (Wong et al., 2017). miRNAs were filtered for $\log_2$(fold change in mean expression) = 0.5-25.0.

2.25.2 The Cancer Genome Atlas colorectal carcinoma cohort

Freely available data from colorectal carcinoma patients was downloaded from The Cancer Genome Atlas (TCGA) (TCGA_Research_Network, 2020) by Henry Wood (University of Leeds, UK). Expression of specific miRNAs, mRNAs and

clinical information was selected. All patient data had been anonymised prior to downloading.

2.26 Statistics

For experiments where N ≥ 3 (see figure legend), significance of differences were calculated using a two-tailed Student's t-test. During DESeq2 analysis of miRNA expression (2.14), p values were adjusted using the Benjamini-Hochberg method to account for multiple significance testing (Love et al., 2014). For qPCR experiments, all statistics were performed on log2 transformed fold changes. Correlations between *GRB2* and miRNA expression data from the TCGA were calculated using the Pearson correlation coefficient and p values were transformed according to the Bonferroni correction for multiple comparisons. Paired t-tests were used to calculate the significance of difference in expression between matched tumour and normal samples from the TCGA.

Chapter 3: Generation of GRB2-knockdown and FGFR2-expressing cell lines

Increasing evidence presents AGO2 as a target of signalling-mediated regulation (Jee and Lai, 2014; Müller et al., 2020). GRB2 has previously been suggested to bind AGO2, although this has not been shown to be direct (Haidar et al., 2018). Additionally, GRB2 expression promoted phosphorylation of AGO2 by EGFR, although again the role of the adaptor was not characterised (Shen et al., 2013). It is possible that FGFR2 also phosphorylates AGO2, as there is a high level of redundancy in RTK signalling (Lemmon and Schlessinger, 2010). This chapter therefore describes the generation of GRB2-knockdown and FGFR2-expressing cell lines, which will later be used to characterise FGFR2- and GRB2-regulation of AGO2.

Initially, CRISPR/Cas9 gene editing was used to knockdown GRB2 in HEK293T cells. FGFR2 was then expressed in both GRB2 kd and unmodified HEK293T by lentiviral transduction. Stable cell lines with varying GRB2 and FGFR2 expression later served as useful tools to identify the roles of both proteins in AGO2 regulation.

3.1 CRISPR/Cas9 editing to knockdown GRB2

Four HEK293T cell lines were generated to study the roles of GRB2 and FGFR2 on AGO2 regulation: 1) GRB2 expressed, FGFR2 not expressed (parental HEK293T); 2) GRB2 kd, FGFR2 not expressed; 3) GRB2 expressed, FGFR2 expressed; 4) GRB2 kd, FGFR2 expressed (Figure 3.1.1). HEK293T cells were used because of the ease with which this cell line can be maintained and transfected.

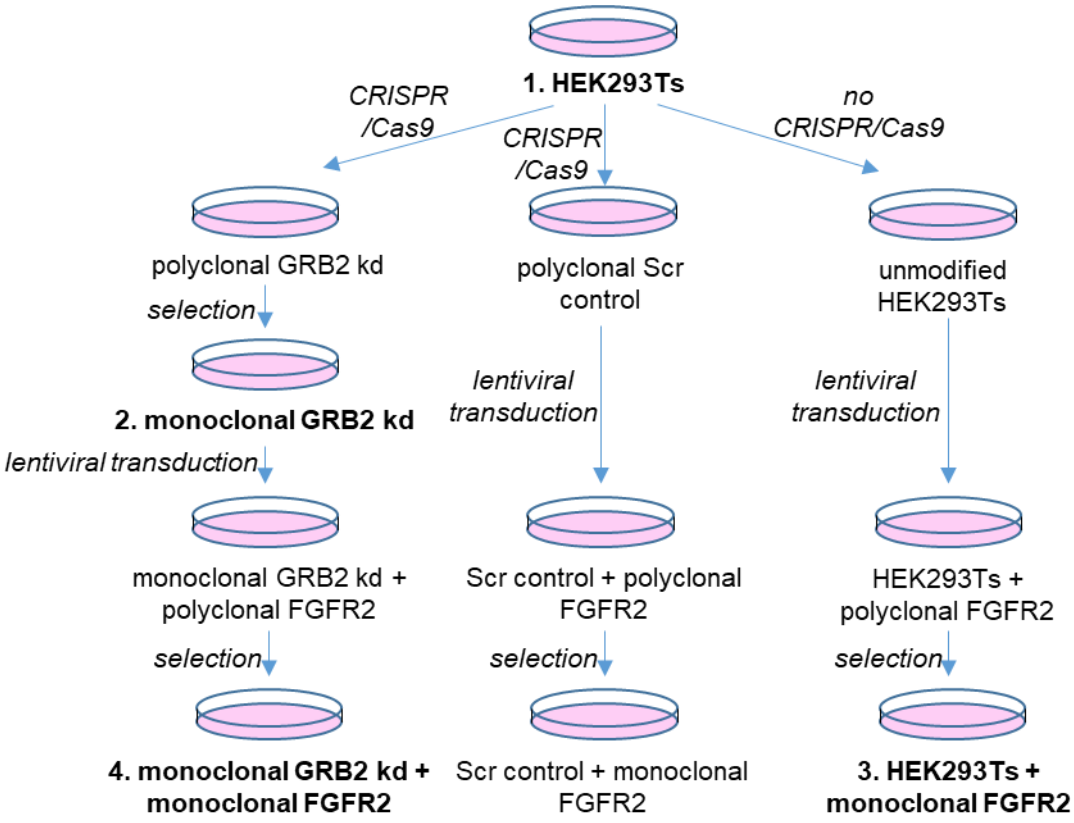


Figure 3.1.1 Scheme for making HEK293T cell lines with and without both GRB2 and FGFR2 expression

HEK293T cells were subject to CRISPR/Cas9 gene editing using a guide RNA targeted against either *GRB2* or with a scrambled sequence (Scr control). GRB2 knockdown (kd) cells were generated. FGFR2 expression was then transduced in GRB2 kd, scr and unmodified HEK293T cells.

GRB2 expression was originally knocked down by CRISPR/Cas9 gene editing. Two plasmids were simultaneously transfected: a pCas9-guide, which encoded a guide RNA (gRNA) targeted to *GRB2* and the Cas9 protein (Figure 3.1.2 A), and pUC, which contained both GFP and puromycin resistance genes (GFP-Puro^R) flanked by GRB2-directed 5' and 3' homology arms (Figure 3.1.2 B). Three pCas9-guide plasmids were used: two with different gRNA sequences targeted against *GRB2* (gRNA1 and gRNA2, Figure 3.1.2 C) and one with a scrambled sequence. Expression of these plasmids lead to *GRB2* knockout by homologous recombination (HR): the gRNA-Cas9 complex cut the genome at the start of *GRB2*, producing a double strand break which was repaired using the pUC template. *GRB2* was consequently disrupted and GFP-Puro^R inserted (Figure 3.1.2 D). Cells with incorporation of Puro^R were then selected with 0.8 μg/ml puromycin. The minimum concentration of puromycin required to kill unmodified HEK293T cells was determined in a kill curve.

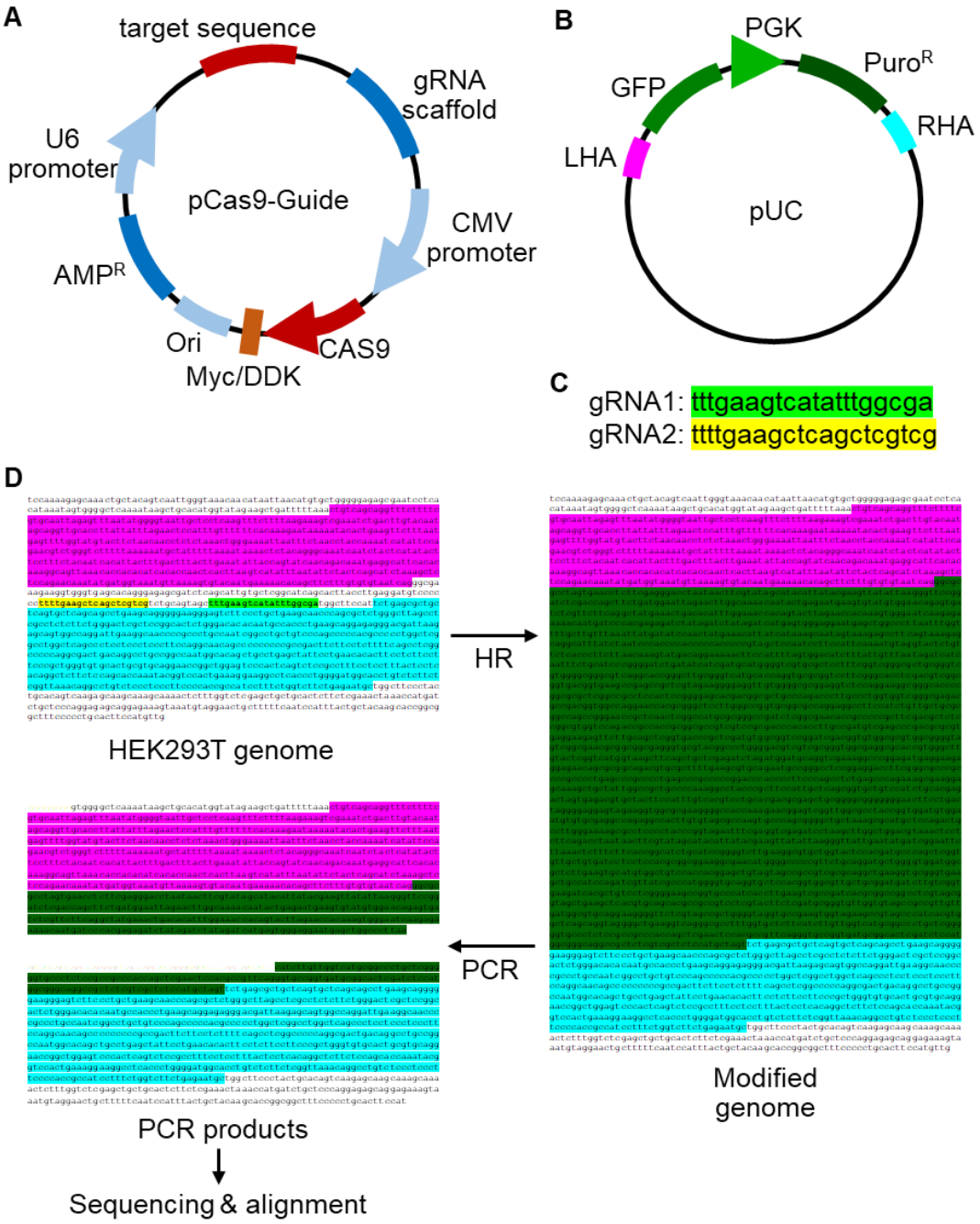


Figure 3.1.2 CRISPR/Cas9 system used to knockdown GRB2

(A) pCas9-Guide plasmid map. OriGene, KN200469. **(B)** pUC plasmid map. OriGene, KN200469. LHA = left hand homologous arm. RHA = right hand homologous arm. PGK = phosphoglycerate kinase 1 promoter. **(C)** guide RNA (gRNA) sequences used to target the start of *GRB2*. **(D)** Gene editing of *GRB2* produced by correct cutting and homologous recombination. Purple highlight = LHA. Cyan highlight = RHA. Light green highlight = sequence targeted by gRNA1. Yellow highlight = sequence targeted by gRNA2. Dark green highlight = inserted sequence containing GFP and puromycin resistance genes.

gRNA1 and gRNA2-modified polyclonal cell lines (pG1 and pG2) showed reduced GRB2 expression at passage 1 when compared to scrambled-modified cell line (Scr) and passaging in media supplemented with puromycin further reduced GRB2 (Figure 3.1.3 A, B). This was presumably due to the selection of CRISPR-edited cells. Initially, to ensure the cells contained the GFP-Puro^R insert at the start of *GRB2*, cells were genotyped by sequencing of the right and left hand homologous arms (RHA and LHA, Figure 3.1.2 D). Genomic DNA was extracted and both homologous arms amplified by PCR. PCR products were purified and analysed by sequencing. The sequencing reads generated from both pG1 arms aligned to *GRB2,* suggesting that GFP-Puro^R had inserted correctly into the *GRB2* gene (Appendix A.1, Figure A.1.1). Sequencing from the RHA of pG2 also matched *GRB2.* However, the sequencing read from the LHA of pG2 mapped to *PNMA5* (Appendix A.1, Figure A.1.2). Neither of the fragments from Scr mapped to *GRB2*. Instead, the LHA fragment aligned with *PNMA5* and the RHA to chromosome 5 (Appendix A.1, Figure A.1.3). Primers used for PCR and sequencing were designed using PrimerBlast (Ye et al., 2012), which generates target-specific primer sequences. However, identification of other regions of the genome suggests that there was non-specific binding and amplification at these loci. It is likely this occurred in absence of the correct insertion of GFP-Puro^R. pG1 was therefore selected to generate monoclonal cell lines and to transduce FGFR2 expression.

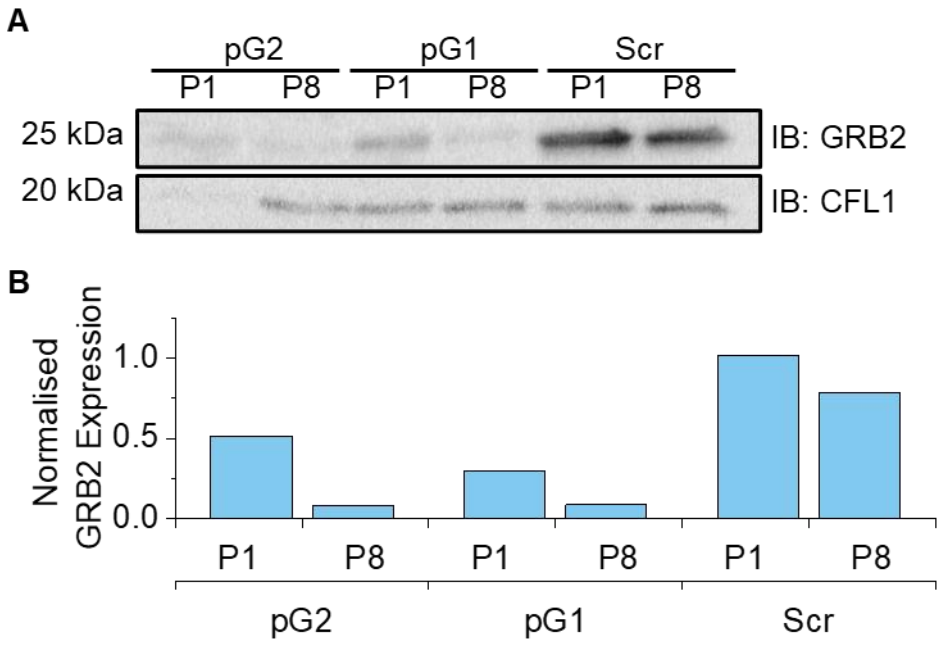


Figure 3.1.3 pG1 and pG2 show reduced expression of GRB2

(A, B) Western blot and quantification of GRB2 expression in polyclonal CRISPR-edited cell lines at passage 1 (P1) and 8 (P8). GRB2 was normalised to the housekeeping gene cofilin 1 (CFL1). pG1 and pG2 = polyclonal GRB2 knockdown. Scr = scrambled control.

3.2 Validation of GRB2 knockdown cell lines

As GFP-Puro^R was correctly inserted into the genome of pG1, monoclonal cell lines (G1) were generated from these cells. Correct incorporation of GFP-Puro^R into *GRB2* was again identified by PCR amplification of the LHA and RHA used in homologous recombination (Figure 3.1.2 D). Sanger sequencing of the PCR products demonstrated that both homologous arms mapped to *GRB2*, suggesting that GFP-Puro^R was appropriately inserted (Appendix A.1, Figures A.1.4-5). Clones were further genotyped to determine if any had complete GRB2 knockout. HEK293T cells have on average between 1 and 3 copies of chromosome 17, on which GRB2 is located (Lin et al., 2014). However, not all copies would need to be knocked out by insert of GFP-Puro^R to produce a complete knockout cell line; some may be disrupted by Cas9-cutting and repair by non-homologous end joining (NHEJ). NHEJ is more frequently used to repair double strand breaks than HR, but is less accurate so can result in small insertions and deletions. If these cause a frame shift in the coding region this will knockout *GRB2*.

G1 clones with the lowest GRB2 expression were selected by western blotting (Figure 3.2.1). They were then genotyped by PCR amplification over the region of *GRB2* which was targeted by gRNA1. Correctly repaired breaks produced a fragment of around 449 bp. All three cell lines produced fragments corresponding to this size, and some also gave products of other sizes (Figure 3.2.2 A). Strangely, a band of 2897 bp, corresponding to a copy of *GRB2* with the GFP-Puro^R insert, was not observed. Possibly the PCR conditions used were not sufficient to amplify a fragment of this size.

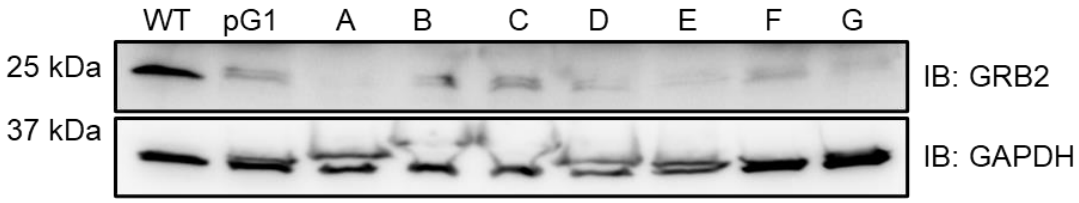


Figure 3.2.1 Selection of GRB2 knockdown clones for genotyping

Western blot and quantification of GRB2 expression in wild type (WT) and gRNA1 modified polyclonal (pG1) and various monoclonal cell lines.

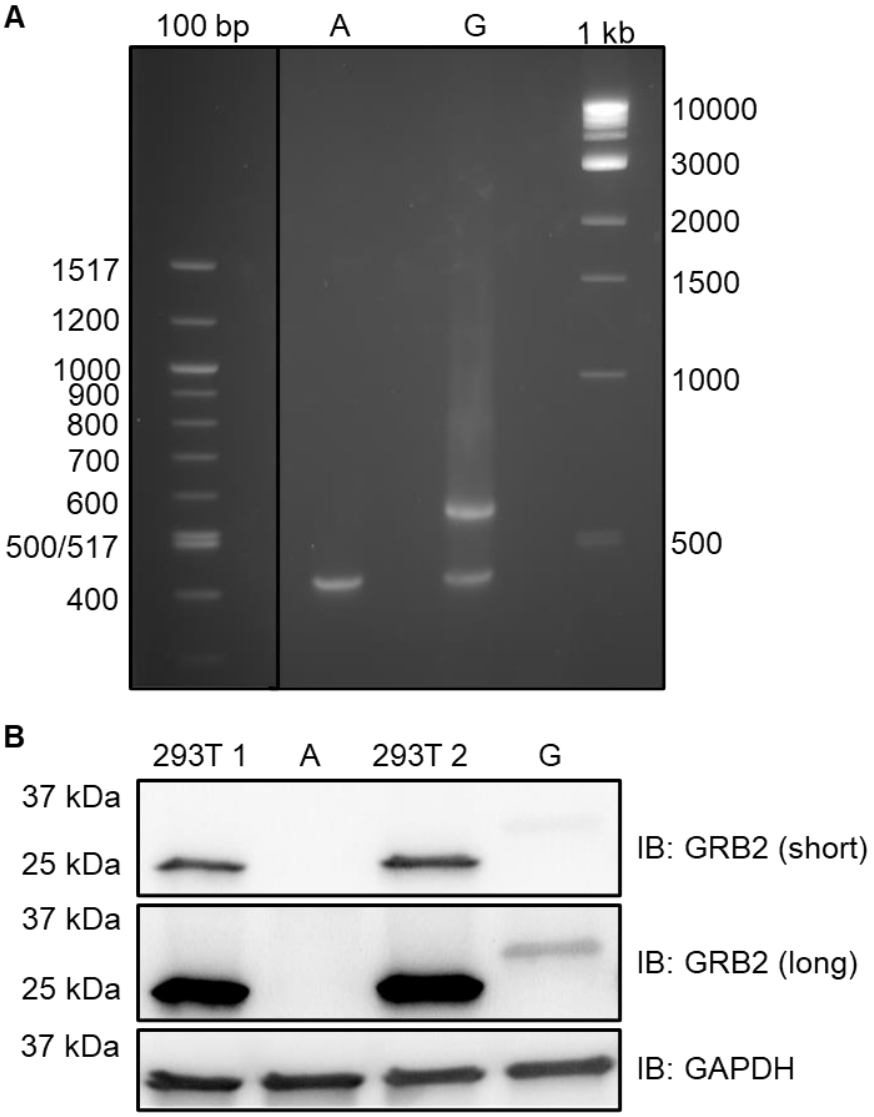


Figure 3.2.2 GRB2 is knocked out and knocked down in two clones

(A) DNA gel of PCR fragments corresponding to the region of *GRB2* which is targeted by gRNA2. **(B)** Western blot of GRB2 expression in HEK293T (293T), the GRB2 knockout clone (A) and GRB2 knockdown clone (G). Long and short refer to exposure time.

Sequencing was used to determine if the PCR fragments contained any mutations. The single product for clone A had a 5 nt deletion causing a frame shift in *GRB2* (Figure 3.2.2 B; Appendix A.1, Figure A.1.6). Clone A therefore had full GRB2 knockout. Clone G had two PCR products: one had a 4 nt deletion, knocking out *GRB2* (Figure 3.2.2 B; Appendix A.1, Figure A.1.7). The larger PCR product had insertions of 139 and 1 nt as well as a deletion of 19 nucleotides, producing an overall insertion of 141 nt (Figure 3.2.2 B; Appendix A.1, Figure A.1.7). Consequently, the mutation maintained the open reading frame of this *GRB2* gene. A 30 kDa GRB2 protein was encoded which contained a truncated NSH3 domain and a 53 residue insert at its very N-terminus (Appendix A.2, Figure A.2.1). Alignment of the insert revealed it mapped to GFP. Western blotting using an antibody targeted at GRB2 CSH3 revealed that the mutant protein was expressed at very low levels (Figure 3.2.2 B). Blotting with the CSH3 antibody also eliminated the band at 25 kDa (Figure 3.2.2 B, Figure 3.2.1 A), suggesting that this band may indicate binding of that antibody (NSH3-targeted) to a non-specific product. The absence of any WT *GRB2* in both clones was further confirmed by cloning the PCR products into a vector, transforming into bacteria and sequencing plasmids from ≥10 colonies (Table 3.2.1).

Table 3.2.1 Genotyping of GRB2 knockdown clones

PCR fragments spanning the start of *GRB2* were generated from three clones. Products were analysed on a DNA gel to determine their size and mutations identified by sequencing. PCR fragments were also inserted into a plasmid and multiple replicates sequenced.

Clone	Product / bp	Mutation	Further sequencing
A	~420	5 nt deletion	11/11 correct mutant
G	~600	141 nt insertion	10/10 correct mutant
G	~420	4 nt deletion	11/11 correct mutant

3.3 Transduction of FGFR2 expression in HEK293T cells

HEK293T cells express no endogenous FGFR2 as shown by western blotting (Figure 3.3.1). FGFR2 expression was therefore transduced into both monoclonal GRB2 knockdown cells (G3F) and scrambled cells (ScrF) using a lentiviral titre expressing FGFR2 (FGFR2IIIc isoform). Importantly, the 'G3' GRB2 clone was isolated from pG1, but had not been genotyped to confirm the extent of GRB2 knockdown; this was verified by western blotting only (Figure 3.3.1 A). Both wild type FGFR2 ($FGFR2^{WT}$) and FGFR2 with GRB2-binding $^{810}PQYP^{813}$ mutated to $^{810}AQYA^{813}$ ($FGFR2^{PxxP}$) were used. Varying levels of FGFR2 expression were observed due to differing transduction efficiencies and for $ScrF^{PxxP}$ expression increased with passaging (Figure 3.3.1 A).

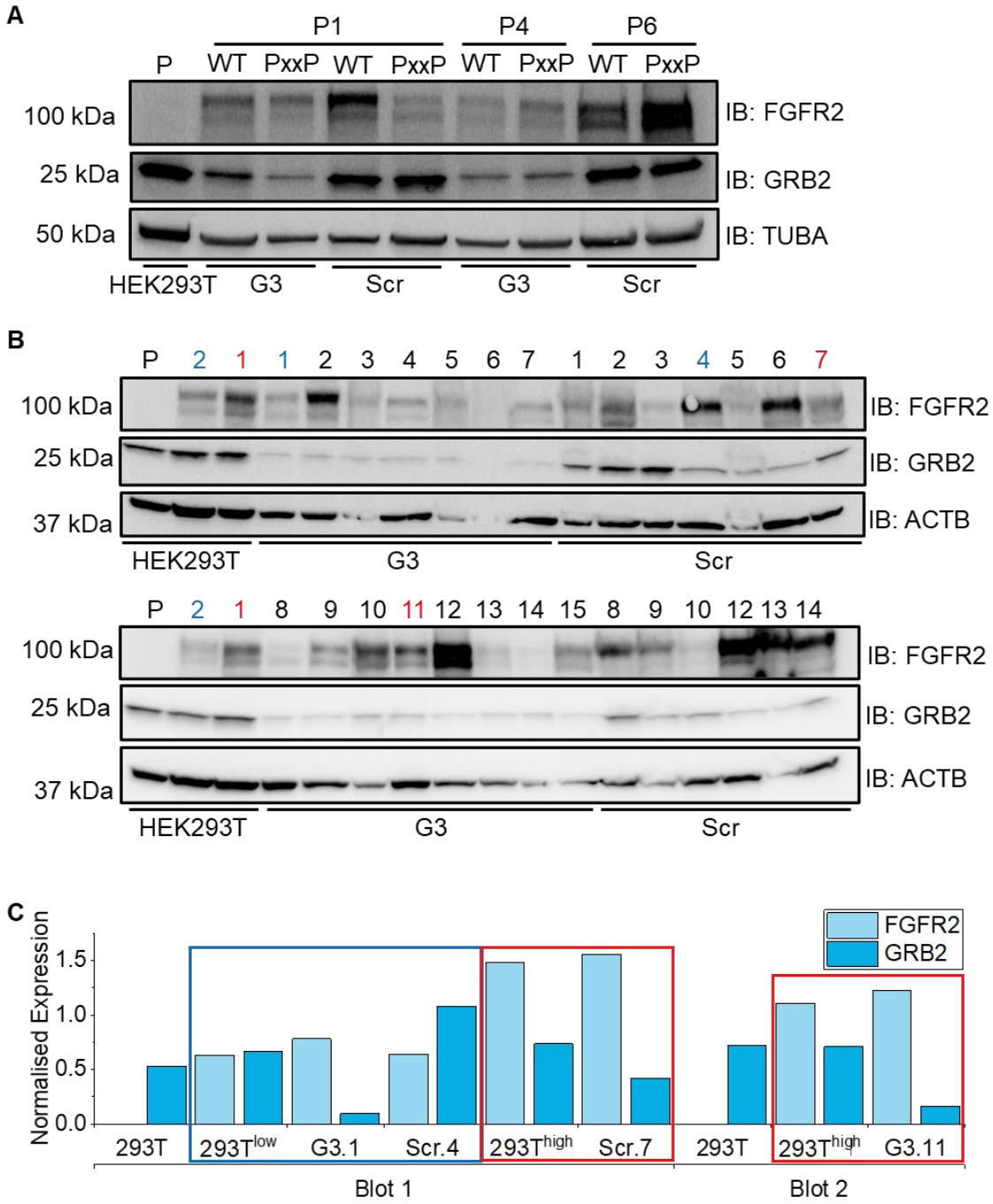


Figure 3.3.1 Transduction of FGFR2 expression into GRB2 knockdown and control HEK293T cells

(A) FGFR2 expression at passage 1 (P1) and later passages after transduction into GRB2 knockdown (G3) and scrambled control (Scr) cells. Both wild type FGFR2 (WT) and FGFR2 with the ^{810}PQYP^{813} motif mutated to AQYA (PxxP) were transduced. GRB2 and FGFR2 expression analysed by western blotting (IB) with α-Tubulin (TUBA) loading control. **(B, C)** Western blot and quantification of FGFR2 and GRB2 expression in multiple clones. Blue text/boxes indicates clones with FGFR2 expression which is comparable to the HEK293T clone stably expressing low FGFR2 (F^{low}). Red text/boxes indicates clones with higher FGFR2 expression, similar to the FGFR2-expressing HEK293T clone (F^{high}). Protein expression normalised to β-Actin (ACTB).

G3F^{WT} and ScrF^{WT} were used to make monoclonal cell lines. The level of FGFR2 expression in multiple clones was compared to that of two existing HEK293T cell lines with wild type FGFR2 expression (F^{low} and F^{high}) (Figure 3.3.1 B). F^{low} and F^{high} had undergone lentiviral transduction to express FGFR2^{WT}, but not CRISPR/Cas9 gene editing, and therefore served as a GRB2/FGFR2^{WT}-expressing cell line, similar to ScrF^{WT}. F^{low} had lower FGFR2 expression and F^{high} expressed higher levels of FGFR2, as determined by western blot (Figure 3.3.1 B). G3F^{WT} clone 1 and ScrF^{WT} clone 4 were selected to be used in experiments with F^{low}, while G3F^{WT} clone 11 and ScrF^{WT} clone 7 with F^{high} (Figure 3.3.1 C). Clones were not selected for the FGFR2^{PxxP}-expressing cell lines. Ultimately, control cell lines generated from HEK293T cells were used in experiments rather than Scr cell lines. In Scr cells, GFP-Puro^{R} had inserted randomly into the genome which could change gene expression and cell behaviour. Unless otherwise stated, the FGFR2-expressing clones used in experiments were F^{high} and G3F^{WT} 11, donated F1 and G3F1 from hereon in.

Chapter 4: Investigation into the phosphorylation of AGO2 by FGFR2

AGO2 is regulated by tyrosine phosphorylation at several sites. One of the best characterised is Y393. Y393 is phosphorylated by EGFR under both stimulated and hypoxic conditions. Phosphorylation of Y393 inhibits DICER1 binding and loading long-loop miRNAs onto AGO2, and therefore correlates with increased cell survival and invasiveness (Shen et al., 2013). The same site is phosphorylated by c-SRC under normoxia and this promotes cell migration (T. Liu et al., 2020).

GRB2 is also implicated in the regulation of Y393 phosphorylation. Knocking down GRB2 diminished AGO2 phosphorylation by EGFR (Shen et al., 2013). Conversely, the adaptor has also been proposed to recruit PTP1B to AGO2 to facilitate dephosphorylation of Y393 (Yang et al., 2014; Haidar et al., 2018).

Mass spectrometry screening data from the Ladbury laboratory previously identified that AGO2 associated with FGFR2 CT (either through direct interaction or as part of a complex) in the absence of growth factor stimulation (Table 4.1.1). Under these conditions phosphotyrosyl binding sites in the C-terminus of FGFR2 for downstream effector proteins are absent. The observed interaction suggests that FGFR2 may bind and phosphorylate AGO2 under these conditions.

Table 4.1.1 Peptide sequences and proteins bound in mass spectrometry analysis of receptor tyrosine kinase interactions under non-stimulated conditions

Prolines which form PxxP motifs are indicated in red in the peptide sequence. AGO2 in red. Data contributed by Arndt Rohwedder and Zahra Timsah.

FGFR1	FGFR2	INSR	IGFR	ALK	PDGFR	HER2
EPLPEEPCLPRHPAQLANGGLKRR	EPSLPQFPHINGSVKT	GKKNGRILTLPRSNPS	ESVPLDPSAS	ERSPAAPPPLPTTSSGKAAK	CDSPLEPQDEPEPEPQL	RPQPPSPRE
YBX1	YBX1	RPSA	HSC70	HSPB1	HSPB1	HSPB1
RPSA	RPSA	YBX1	IF-B1	ILF2	YBX1	RPSA
SFPQ	SFPQ	G3BP1	GRP78	RPSA	RPSA	ILF2
HNRNPK	HNRNPL	RNPS1	TP63	HNRNPK	HNRNPK	
HNRNPU	WIBG	HNRNPK	OAS1	RMSB	RMSB	
G3BP1	HSC70	HNRNPDL	HNRNPU	TCP1	ILF2	
CAPRIN1	CAPRIN1			DNAJB6	DNAJB6	
TCP1	HNRNPU			EIF3F	EIF3F	
G3BP2	ACL					
STAU1	DDX1					
YBX3	DBN1					
SND1	**AGO2**					
	RNPS1					

Protein binding to FGFR2 CT under non-stimulated conditions is mediated by SH3 domains binding to the proline-rich motif ^{807}PSLPQYP^{813} (Ahmed et al., 2010; Timsah et al., 2014). AGO2 does not possess an SH3 domain, and therefore is unlikely to bind FGFR2 CT directly. Instead, it is possible that GRB2 recruits AGO2 to FGFR2. One way in which this could occur is through the binding of AGO2 to GRB2 NSH3 domain which is not involved in the complex with FGFR2 (Chapter 1, Figure 1.11.3). In this Chapter, the potential roles of GRB2 and FGFR2 in regulating both AGO2 phosphorylation and the AGO2-DICER1 interaction will be investigated.

4.2 GRB2 associates with AGO2 and FGFR2 under non-stimulated conditions

Initially, to reflect the conditions under which AGO2 associated with FGFR2, regulation of AGO2 by FGFR2 and GRB2 was studied in non-stimulated HEK293T cells. Serum-starved conditions, i.e. cells which had been starved in media in the absence of FBS, were adopted to represent the non-stimulated, basal state. Under these conditions the receptor should not be unable to initiate a downstream signalling response. To identify whether this was the case, the phosphorylation of ERK1/2 (pERK1/2) was measured, which provides a marker for upregulation of the MAPK pathway. pERK1/2 was assessed in both parental HEK293T cells, which do not express FGFR2, and HEK293T cells stably expressing FGFR2 (F1), under both starved and stimulated conditions. Cells were depleted of FBS overnight, followed by stimulation with FGF9 for 15 min. Phosphorylation was assessed by western blotting. The serum-starved cells showed no evidence of pERK1/2 (Figure 4.2.1 A), suggesting that the serum-starvation conditions corresponded to a basal state. FGF9-stimulated HEK293T

cells expressed less pERK1/2 than F1. The low level of pERK1/2 in stimulated HEK293T cells may have been a result of FGF9-activation of FGFR3, which is expressed at low levels (Uhlén et al., 2015).

Under non-stimulated conditions, it was hypothesised that GRB2 may recruit AGO2 to FGFR2. To investigate the potential interaction between AGO2 and GRB2, GST-tagged GRB2 was immobilised on agarose beads and incubated with lysates from serum-starved HEK293T and F1 cells. Bound proteins were washed and analysed by western blotting. The blot clearly showed bands corresponding to both AGO2 and FGFR2 (Figure 4.2.1 B). AGO2 was observed to be associated with GRB2 in both the presence and absence of FGFR2 expression.

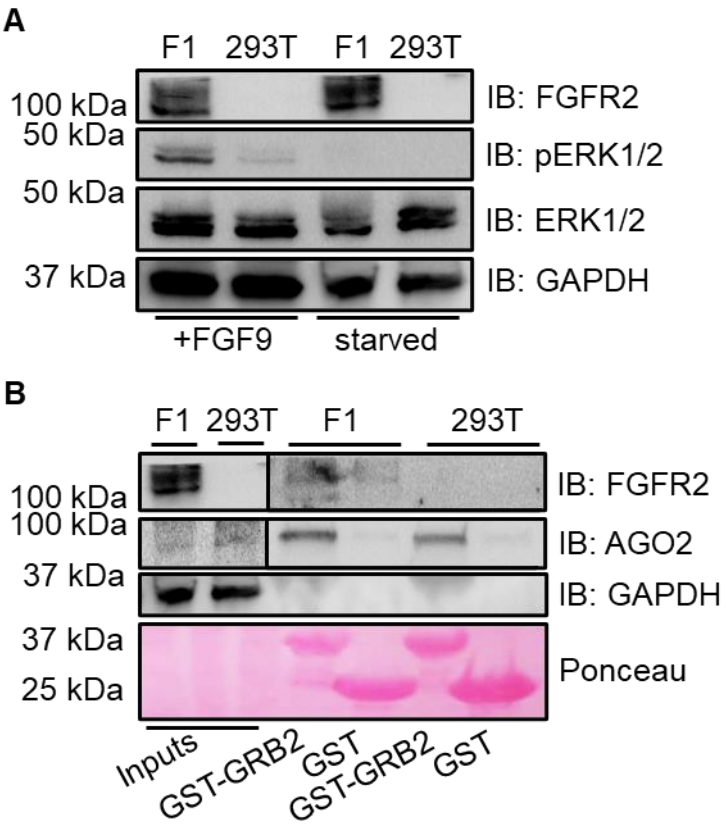


Figure 4.2.1 GRB2 binds AGO2 and FGFR2 in serum-starved HEK293T cells

(A) HEK293T (293T) and HEK293T with stable FGFR2 expression (F1) cells were incubated in starvation media for 16 hours (16 h) before stimulation with FGF9. Cells were lysed and pERK1/2 expression analysed by western blotting (IB). **(B)** Lysates from serum-starved HEK293T and F1 cells were incubated with purified GST proteins before washing and analysis of bound proteins by western blotting.

4.3 Overexpression of FGFR2 results in tyrosine phosphorylation of AGO2

To test if Y393 was phosphorylated in lysates from HEK293T cells expressing FGFR2, a mass spectrometry approach was adopted, which provides a sensitive detection of post-translational modification through measuring molecular mass change. Three conditions were analysed. First was EGF-stimulated parental HEK293T cells transiently overexpressing GFP-EGFR. EGF acts as a positive control because stimulation of the receptor is known to result in phosphorylated Y393 (pY393) (Shen et al., 2013). Second was HEK293T with stable FGFR2 expression (F1), stimulated with FGF2/9 (cocktail of FGF2 and FGF9), and third, FGF2/9-stimulated GRB2-depleted HEK293T (G1; Chapter 3, 3.2) transiently overexpressing FGFR2. In all cell lines, GFP-AGO2 was overexpressed. Cells were serum-starved and then stimulated with growth factor (10 ng/ml EGF; or 20 ng/ml FGF2/9). FGF2 and FGF9 preferentially activate the FGFR2IIIc isoform which was used in this study. Both growth factors were selected since it was not known which was more efficient in stimulating AGO2 phosphorylation. GFP-AGO2 was purified using GFP-Trap and run on an SDS PAGE gel. The band corresponding to GFP-AGO2 was excised and trypsin digested before peptides were subjected to MALDI-ToF-MS/MS analysis (see Chapter 2, 2.11). Successful cell stimulation and protein expression was confirmed by western blotting (Figure 4.3.1).

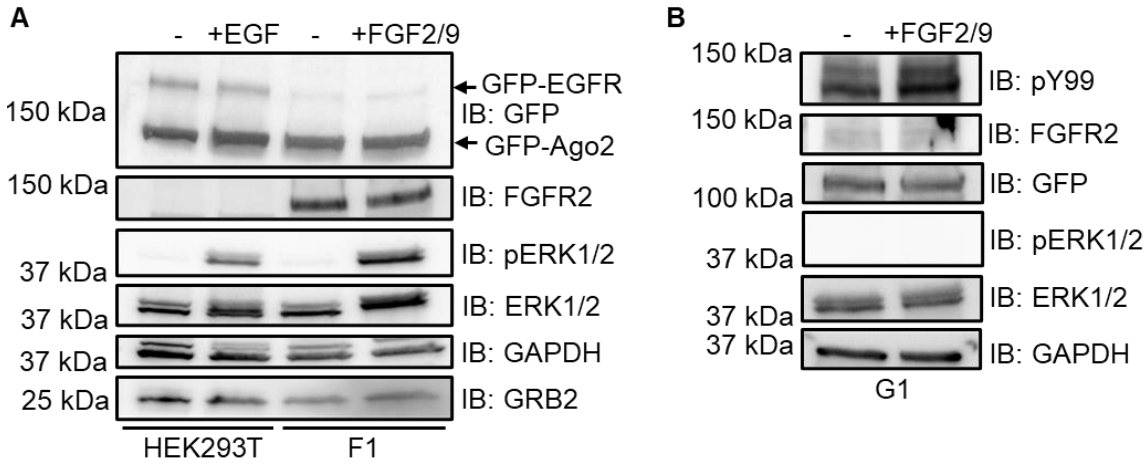


Figure 4.3.1 Lysates used in mass spectrometry analysis of GFP-AGO2

(A) HEK293T cells overexpressing GFP-EGFR, HEK293T cells stably expressing FGFR2 (F1) cells and **(B)** GRB2-depleted HEK293T (G1) cells overexpressing FGFR2, all overexpressing GFP-AGO2, were stimulated with the indicated growth factor. Protein expression was analysed by western blotting (IB).

Four peptides containing sequences including Y393 were identified, but none showed phosphorylation of this residue under any condition (Table 4.3.1). The number of times each peptide was detected is referred to as a count and represents the abundance of that peptide in each sample. Serine 387 (S387) was phosphorylated in a small number of peptides in all three conditions, as has been previously observed (Zeng et al., 2008). Coverage rates of 92.88%, 87.57% and 82.70% were obtained for GFP-AGO2 from EGF-stimulated cells, FGF-stimulated F1 cells and FGF-stimulated G1 cells respectively. Since no evidence of pY393 was observed in the EGF-stimulated positive control, it is possible that Y393 was phosphorylated but that it was not detected.

Table 4.3.1 Sequences and counts of peptides spanning Y393 identified in mass spectrometry analysis of GFP-AGO2

GFP-AGO2 was purified from EGF-stimulated HEK293T cells overexpressing GFP-EGFR (293T+EGF); FGF2/9-stimulated HEK293T stably expressing FGFR2 (F1+FGF); and FGF2/9-stimulated, GRB2-depleted HEK293T overexpressing FGFR2 (G1+FGF). Bound protein was analysed by SDS PAGE and the band corresponding to GFP-AGO2 trypsin digested before being subject to MALDI-ToF-MS/MS. Capital letters indicate non-phosphorylated residues whereas lowercase letters indicate phosphorylated residues. Y393 is highlighted.

Sequence	293T+EGF	F1+FGF	G1+FGF
SAsFNTDPYVR	3	2	5
SASFNTDPYVR	14	18	18
SASFNTDPYVREFGIMVK	1	0	0
SASFNTDPYVREFGIMVKDEMTDVTGR	2	0	0

Interestingly, phosphorylation of additional tyrosine residues (Y322, Y529, Y804/805) was observed in GFP-AGO2 purified from FGF-stimulated G1 cells compared to FGF-stimulated parental HEK293T cells (Table 4.3.2). However, the G1 cells transiently overexpressed FGFR2, from a plasmid, whereas the HEK293T cells had stable FGFR2 expression, from a gene inserted into the cell genome. Therefore, it is not clear whether the increase in tyrosine phosphorylation in G1 cells was due to the lack of GRB2 or the higher levels of FGFR2 resulting from transient expression as opposed to stable expression.

Table 4.3.2 AGO2 phosphorylation sites and peptide counts identified in mass spectrometry analysis of GFP-AGO2

GFP-AGO2 was purified from EGF-stimulated HEK293T cells overexpressing GFP-EGFR; FGF2/9-stimulated HEK293T with stable FGFR2 (F1); and FGF2/9-stimulated, GRB2-depleted HEK293T (G1) overexpressing FGFR2. MALDI-ToF-MS/MS analysis of GFP-AGO2 identified peptides with phosphorylated residues. The number of peptides spanning each site, either phosphorylated or unmodified, is listed. Bracketed numbers indicate the probability that the correct phosphorylated residue was identified. Only sites with probabilities over 50% have been included. If the probability varied with different peptides this is presented as a range.

Phospho site	HEK293T+EGF		F1+FGF		G1+FGF	
	Phosphorylated	Unmodified	Phosphorylated	Unmodified	Phosphorylated	Unmodified
Y393	0	20	0	20	0	23
S189	0	38	1 (100%)	45	1 (97.2%)	50
S253	1 (99.9%)	24	0	22	0	50
T307	1 (99.2%)	34	1 (95.5%)	18	0	30
Y322	0	34	0	24	4 (100%)	38
T338	0	56	0	55	5 (87.1%)	68
S387	3 (100%)	17	2 (100%)	18	5 (96.8%)	18
Y529	0	22	0	13	3 (100%)	11
S727	0	17	1 (100%)	18	0	24
Y804 or Y805	0	40	0	35	6 (50:50)	40
S824	5 (100%)	31	9 (97.1-100%)	20	7 (89.2-100%)	67
T830	2 (54.5-70%)	34	3 (80.5%)	30	2 (54.6-70%)	72
S831	4 (83.2-87.9%)	32	0	33	1 (54.6-70%)	73
S834	4 (97.6-99.9%)	32	4 (90.2-99.8%)	29	2 (90.1-96.7%)	72

To test if the increase in AGO2 phosphorylation of tyrosines Y322, Y529, Y804/805 seen in G1 cells was due to the knockdown of GRB2 or transient overexpression of FGFR2, AGO2 phosphorylation status was analysed using a general pY antibody (pY99). FLAG-FGFR2 was transfected into both parental HEK293T and G1 cells. Cells were then serum-starved overnight, before being stimulated with FGF2/9. Endogenous AGO2 was immunoprecipitated using a high stringency buffer (1% Tween-20) and general tyrosine phosphorylation analysed by western blotting with pY99. In both cell lines, overexpression of FGFR2 resulted in tyrosine phosphorylation of AGO2 under both starved and stimulated conditions (black arrow, Figure 4.3.2). AGO2 showed higher tyrosine phosphorylation in unmodified HEK293T cells than the G1 cells. Thus, it appeared that the increase in tyrosine phosphorylation observed by mass spectrometry was due to the transient overexpression of FGFR2. Furthermore, in HEK293T cells, stimulation enhanced AGO2 phosphorylation. Together, these observations suggested that GRB2 may regulate AGO2 phosphorylation both independently and cooperatively with FGFR2-stimulation.

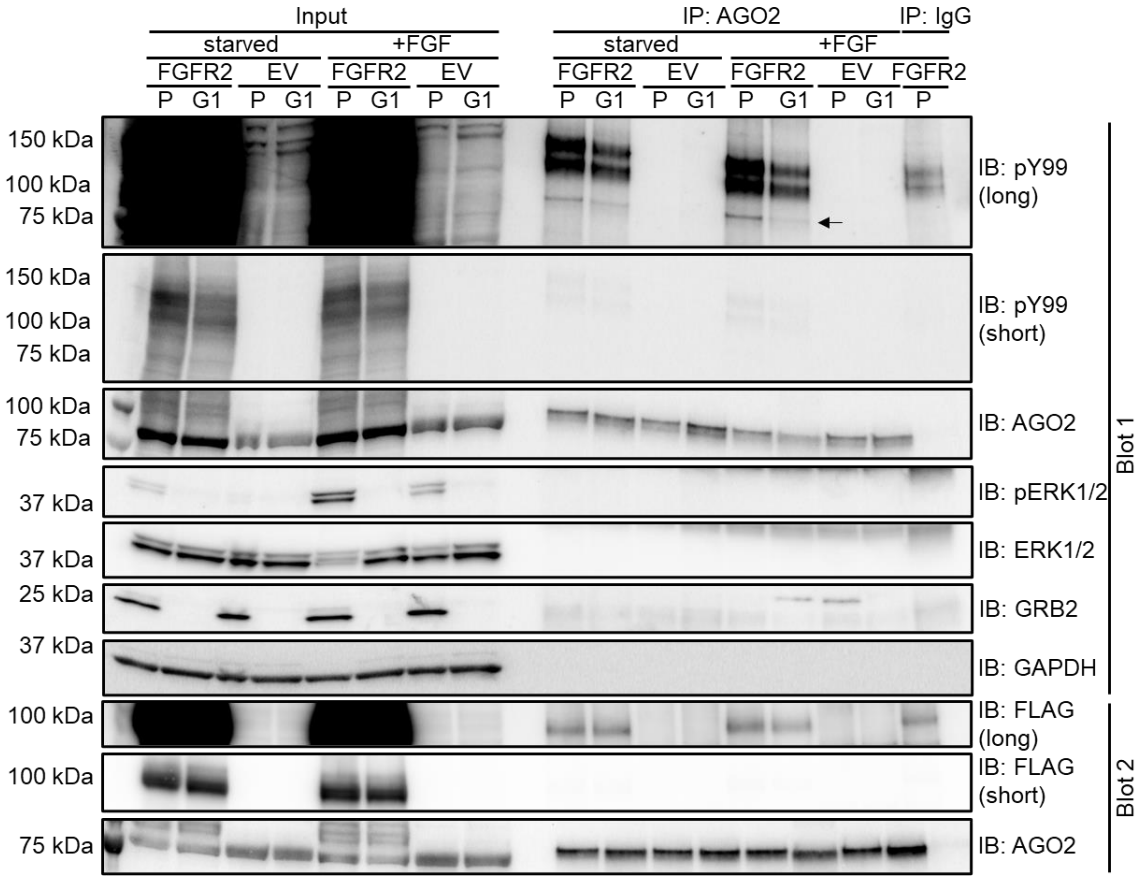


Figure 4.3.2 FGFR2 overexpression mediates tyrosine phosphorylation of AGO2

FLAG-FGFR2 or the empty vector (EV) was transfected into parental HEK293T cells (P) or G1 cells. Cells were either starved or stimulated with 20 ng/ml FGF2/9 and AGO2 immunoprecipitated. Tyrosine phosphorylation (pY99) and presence of FLAG-FGFR2 in immunoprecipitates (IPs) was analysed by western blotting (IB). Arrow indicates AGO2. Long and short refer to length of exposure time.

Interestingly, the pY99 blot also revealed bands at ~120 and ~140 kDa in the AGO2 IPs (Figure 4.3.2). The molecular weight of 120 kDa corresponds to FGFR2. A new blot was probed with FLAG antibody, to minimise bleed through of old antibody signal. However, similar levels of FLAG-FGFR2 were pulled down in both AGO2 and IgG IPs (Figure 4.3.2). The pY99 signal must therefore have been due to pulldown of another tyrosine phosphorylated protein of this size.

The data in Section 1.2 suggested FGFR2 may phosphorylate AGO2 in HEK293T cells transiently overexpressing FGFR2. AGO2 phosphorylation was enhanced by stimulation, but phosphorylation of AGO2 also was seen under basal conditions. GRB2 augmented AGO2 phosphorylation (Figure 4.3.2). However, pY393 could not be detected by mass spectrometry, as demonstrated by the absence of pY393 identified in the positive control, EGF-stimulated cells (Table 4.3.1). It therefore remained to be determined if FGFR2 could phosphorylate Y393 in cells.

4.4 Interaction of DICER1 with AGO2 is not controlled by FGFR2

Prior to this point, the data revealed that FGFR2 overexpression may enhance tyrosine phosphorylation of AGO2. Previous reported data shows that phosphorylation of Y393 results in inhibition of DICER1 binding to AGO2 (Shen et al., 2013). Thus, the presence, or absence, of the AGO2-DICER1 complex can give an indirect measure of pY393 on AGO2.

To assess whether GRB2 would mediate, or inhibit, FGFR2-dependent Y393 phosphorylation, the DICER1-AGO2 interaction was examined in both

HEK293T stably expressing FGFR2 (F1) and GRB2 knockdown cells stably expressing FGFR2 (G3F1). Cells with stable FGFR2 expression were used to simplify the experiment. The cells were starved overnight, before being stimulated with FGF2/9. Endogenous AGO2 was then immunoprecipitated and blotted for DICER1 pulldown. pY393 and pY99 were not probed as both antibodies produced strong nonspecific bands in the IgG control at the expected molecular weight of AGO2 (Figure 4.4.1). The appearance of a non-specific band when blotting F1 IPs with the pY99 antibody, but not when blotting IPs from HEK293T cells overexpressing FGFR2 (Figure 4.3.2), was likely to be a result of the longer exposure required in this experiment.

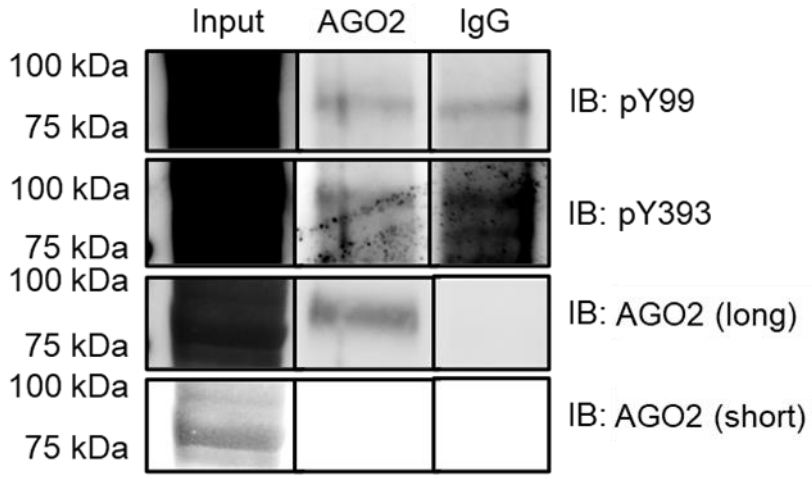


Figure 4.4.1 The pY393-specific and pY99 antibodies produce bands for samples immunoprecipitated with IgG

Lysates from HEK293T cells with stable FGFR2 expression (F1) were taken, then proteins were immunoprecipitated using AGO2 or IgG antibodies and analysed by western blotting (IB) for phosphorylated tyrosine (pY99) and phosphorylated AGO2 Y393 (pY393). Outlines indicate cropping of blot to remove other lanes; all bands are taken from the same blot and at the same length of exposure time. Long and short refer to exposure time.

The level of DICER1 found in AGO2 immunoprecipitates was not significantly changed upon stimulation with FGF2/9, in both G3F1 and F1 cells (Figure 4.4.2). G3F1 cells do not have GRB2 knockout but have reduced GRB2 compared to F1 (Figure 4.4.2, Figure 3.3.1). Additionally, non-specific binding of this GRB2 antibody to a 25 kDa protein (Figure 3.2.1 A, Figure 3.2.2 B) and over-exposure of the blot obscured the apparent extent of knockdown. Stimulation of G3F1 did not result in upregulation of pERK1/2. Additionally, this cell line had diminished FGFR2 phosphorylation compared to F1, under both serum-starved and stimulated conditions. It therefore appeared that knockdown of GRB2 hindered FGFR2 activation. GRB2 could not be observed in AGO2 immunoprecipitates, potentially due binding of the secondary antibody to IgG upon the high exposure. Overall, these results indicated that stably expressed FGFR2 did not regulate the AGO2-DICER1 interaction in HEK293T cells. Knockdown of GRB2 expression did not affect this.

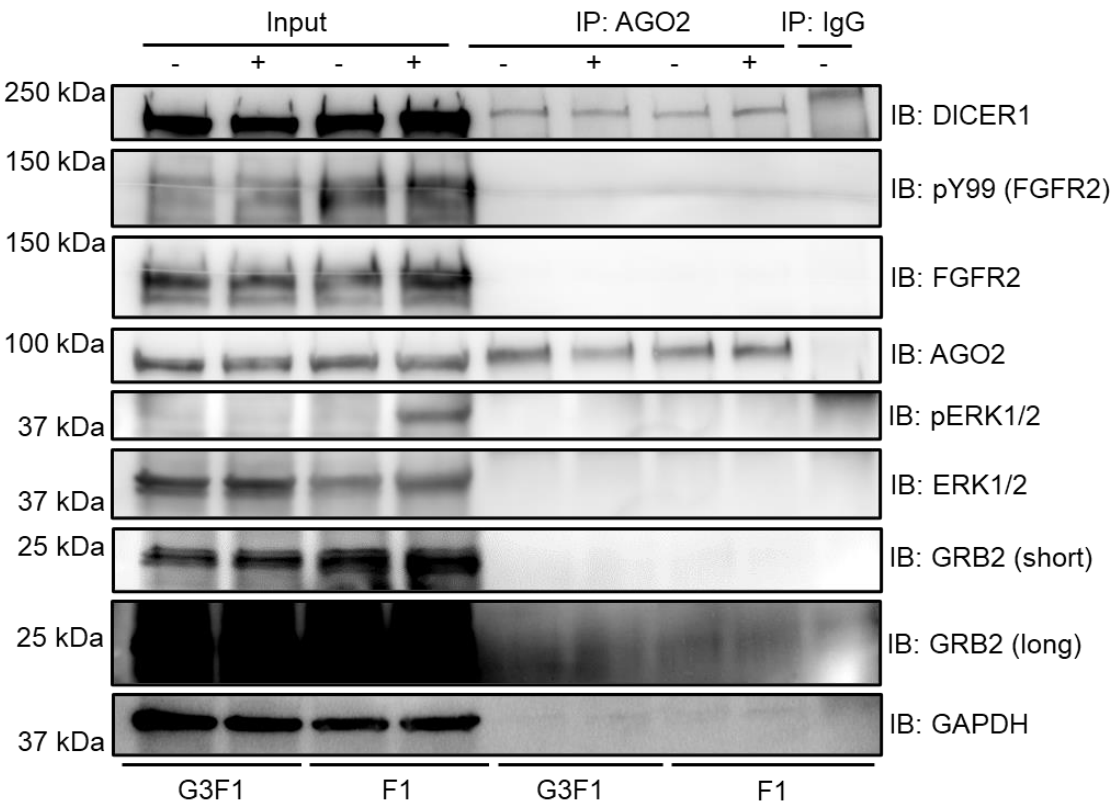


Figure 4.4.2 FGF-stimulation of GRB2-expressing and GRB2-knockdown HEK293T cells stably expressing FGFR2 does not reduce AGO2 binding DICER1

Western blot (IB) of DICER1 associated with endogenous AGO2 which has been immunoprecipitated (IP) from G3F1 (HEK293T with GRB2 knockdown and FGFR2 expression) and F1 (HEK293T with FGFR2 expression). Cells were starved (-) or stimulated with FGF2/9 for 15 minutes (+). N = 2. Long and short refer to exposure time.

To evaluate whether high, as opposed to endogenous, expression of GRB2 resulted in FGFR2 regulation of the interaction between AGO2 and DICER1, the appearance of the AGO2-DICER1 complex was examined in cells overexpressing GRB2. The empty vector (EV) was used as a control. GRB2 or EV were transfected into F1 cells, which were then serum-starved before stimulation with FGF2/9. In this experiment, serum starvation was completed in the presence or absence of an FGFR2 kinase inhibitor (SU5402, 30 nM), to allow comparison of basally activated, monophosphorylated FGFR2 and inactive FGFR2. Thus, FGFR2-inhibited cells act as a negative control for phosphorylation of AGO2 by FGFR2 in both serum-starved and FGF-stimulated cells. Endogenous AGO2 was immunoprecipitated and analysed by western blotting. As before, the level of DICER1 bound to AGO2 was not changed upon FGFR2 stimulation, nor upon FGFR inhibition (Figure 4.4.3 A, B). No changes were observed in both EV and GRB2-transfected cells. Therefore, it seemed that elevated GRB2 expression did not regulate AGO2 binding to DICER1 in cells in the presence of basally activated FGFR2 kinase, nor in the presence of FGF2/9-stimulated FGFR2 kinase.

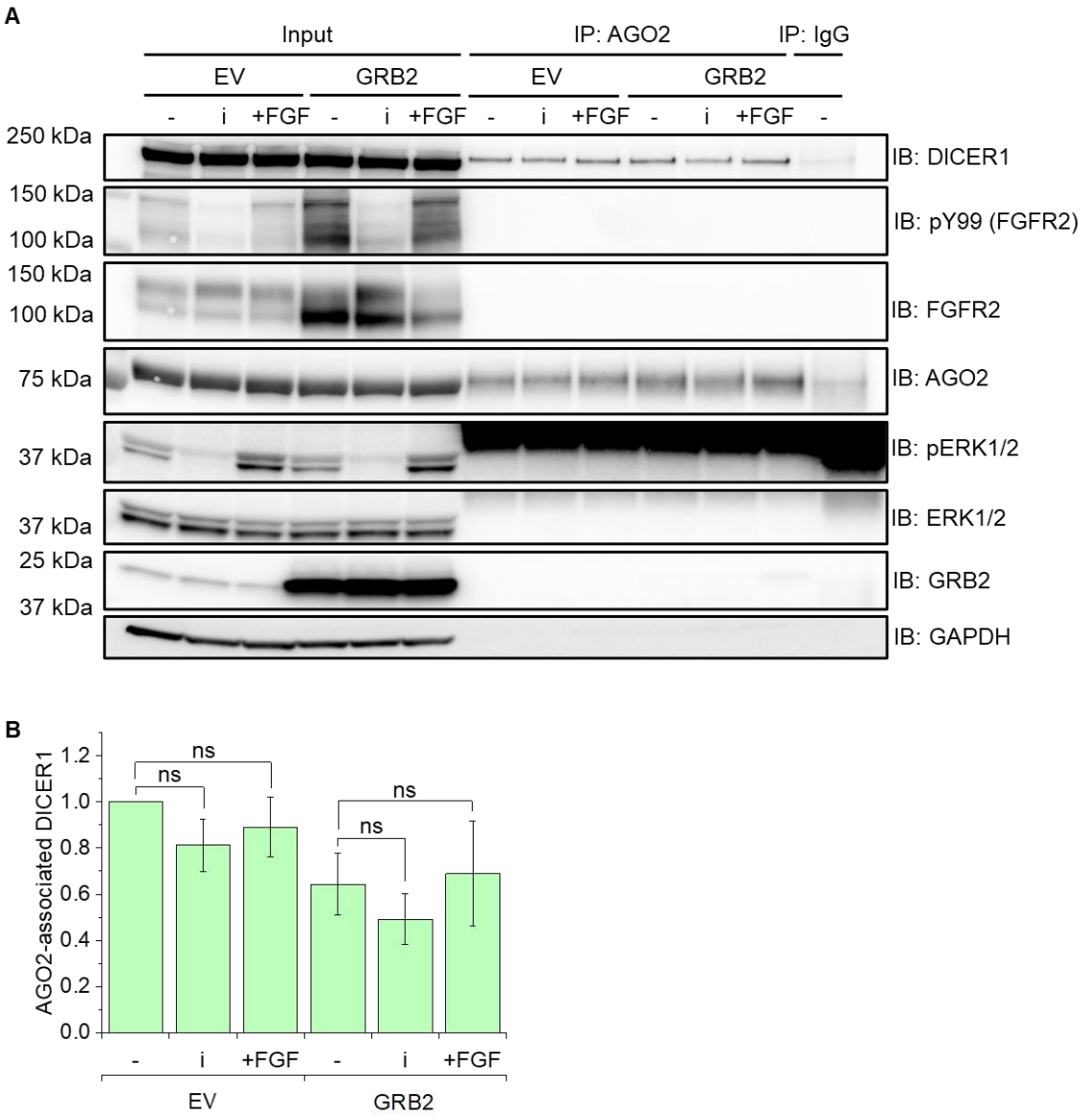


Figure 4.4.3 AGO2 binding DICER1 is not regulated by FGFR2 in HEK293T cells stably expressing FGFR2 and with GRB2-overexpression

(A) Western blot (IB) and **(B)** quantification of DICER1 associated with endogenous AGO2 in HEK293T stably expressing FGFR2 (F1) with and without GRB2 overexpression. DICER1 was normalised against AGO2. Cells were cultured in starvation media containing either DMSO or FGFR2 inhibitor (i) before being stimulated with FGF2/9 and used to immunoprecipitate (IP) AGO2. N = 3. Error bars show standard error of mean. EV = empty vector. ns = not significant by Student's t-test.

pY393 is dephosphorylated by PTP1B (Yang et al., 2014). Therefore, it was possible that PTP1B dephosphorylation of Y393 masked changes in AGO2 binding DICER1. To determine this, AGO2 was immunoprecipitated in cells which had been cultured in presence and absence of PTP1B inhibitor (CinnGEL-2me, 10 μM). F1 cells were incubated under serum-starved conditions in the presence, and absence, of FGFR2 and PTP1B inhibitors, before being stimulated with FGF2/9. Endogenous AGO2 was immunoprecipitated and analysed by western blotting. However, the level of DICER1 associated with AGO2 was not significantly reduced upon FGF stimulation in PTP1B-inhibited cells, nor in DMSO-treated control cells (Figure 4.4.4). Consequently, it seemed AGO2 was not regulated by FGFR2, nor PTP1B, in serum-starved and FGF-stimulated HEK293T cells stably expressing FGFR2. It should be noted that, although the PTP1B inhibitor was used as previously described (Vota et al., 2013; Chamorro et al., 2015), PTP1B inhibition by CinnGEL-2me should be confirmed by western blotting of EGF-induced signalling, so could not be analysed in these cells.

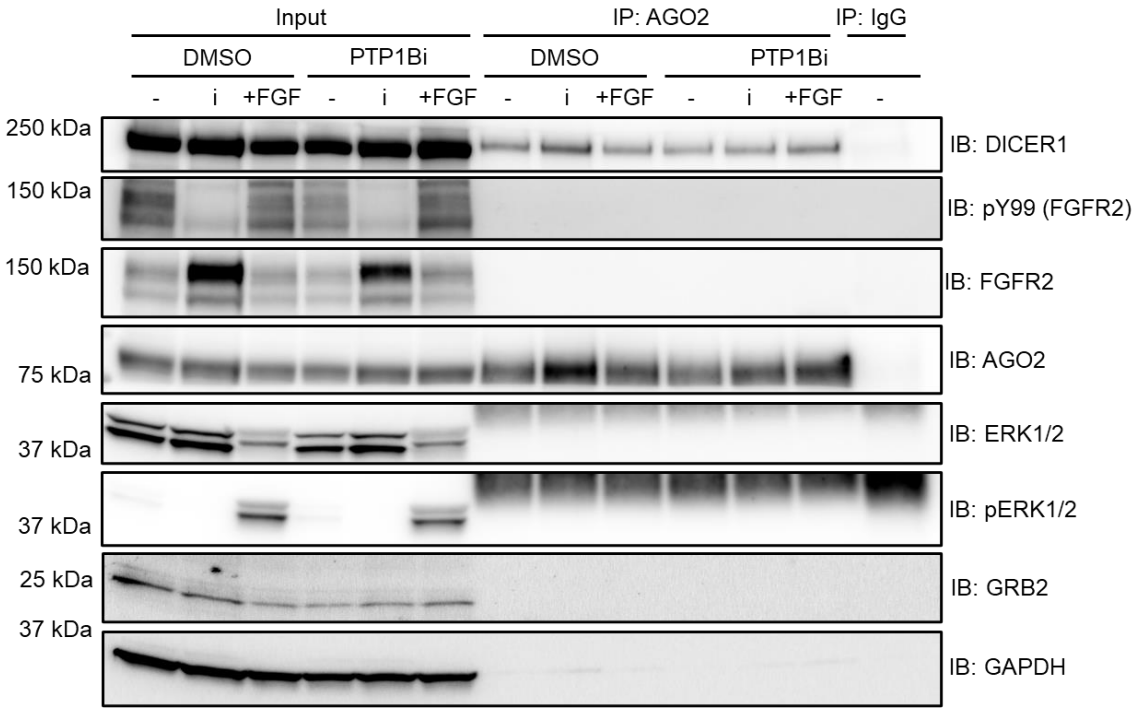


Figure 4.4.4 FGFR2 does not regulate the interaction between AGO2 and DICER1 in PTP1B-inhibited HEK293T cells stably expressing FGFR2

Western blot (IB) of DICER1 associated with endogenous AGO2 in F1 (HEK293T stably expressing FGFR2) with and without PTP1B inhibition. Cells were cultured in starvation media supplemented with FGFR2 inhibitor (i), PTP1B inhibitor (PTP1Bi) or DMSO, before being stimulated with FGF2/9 and used to immunoprecipitate (IP) AGO2. N = 2.

Overall, these experiments indicated that stably expressed activated FGFR2 did not regulate binding of AGO2 to DICER1 in HEK293T cells, regardless of GRB2 expression. Therefore, these results do not provide evidence for FGFR2-phoshorylation of AGO2 Y393.

4.5 FGFR2 does not phosphorylate AGO2 in breast cancer cell lines

The lack of FGFR2-dependent changes in DICER1-binding AGO2 in HEK293T cells stably expressing FGFR2 (F1) suggested this cell line may be inappropriate for studying FGFR2 regulation of AGO2. To explore if FGFR2 may phosphorylate Y393 in another system, AGO2 phosphorylation was examined in MCF-7 and MDA-MB-231 breast cancer cell lines. Both express the two main isoforms of FGFR2: FGFR2IIIb and FGFR2IIIc (Nurcombe et al., 2000). FGFR2IIIb is preferentially activated by FGF7 and 10, while activation of FGFR2IIIc is favoured with FGF2 and 9 (Miki et al., 1992).

MDA-MB-231 and MCF-7 cells were serum-starved, in the presence or absence of 30 nM SU5402. Cells were then stimulated with a panel of FGFs (2, 7, 9 and 10). Endogenous AGO2 was immunoprecipitated and blotted with pY393. The AGO2 IP had been optimised to allow blotting with the pY393 antibody. It was found that IP with IgG from a rabbit host produced a strong band at 90 kDa upon blotting with pY393 but using IgG from a mouse host did not. The reactions from each cell line were analysed on separate gels. It was previously observed that running the reactions on a larger midi gel gave blurred AGO2 bands, which may make it harder to detect pY393 bands. Blots were probed initially for pY393, then for pY99 to assess FGFR2 stimulation. In both cell lines, FGFR2 expression was low, making it hard to detect phosphorylation of FGFR2

activation loop (Y656/7; Figure 4.5.1 A, B; Figure 4.5.2 B). However, an FGFR2-specific pY99 band was difficult to distinguish from the background due to presentation of pY residues on other proteins in the cell lysate (Figure 4.5.1 A, B). Furthermore, pY393 demonstrated nonspecific binding in the input, which may indicate binding to other pYs. Stripping these sites can result in reduced antibody binding upon re-probing. Consequently, FGFR2 stimulation was inferred from the downstream activation of ERK. In both cell lines, pERK was reduced upon incubation with SU5402 and enhanced upon stimulation. The increase in ERK phosphorylation was particularly apparent with FGF2, 7 and 10 in MCF-7 cells (Figure 4.5.1 A), and with FGF2 and 10 in MDA-MB-231 cells (Figure 4.5.1 B). Thus, FGFR2 appeared to be properly inhibited and stimulated under these conditions.

AGO2 phosphorylation was assessed by blotting for pY393. Bands in the IP samples at ~90 kDa could be observed upon a very long exposure of pY393 blots. However, a band also appeared in the IgG control (Figure 4.5.1). It was therefore concluded that phosphorylation of Y393 could not be detected on endogenous AGO2 under these conditions.

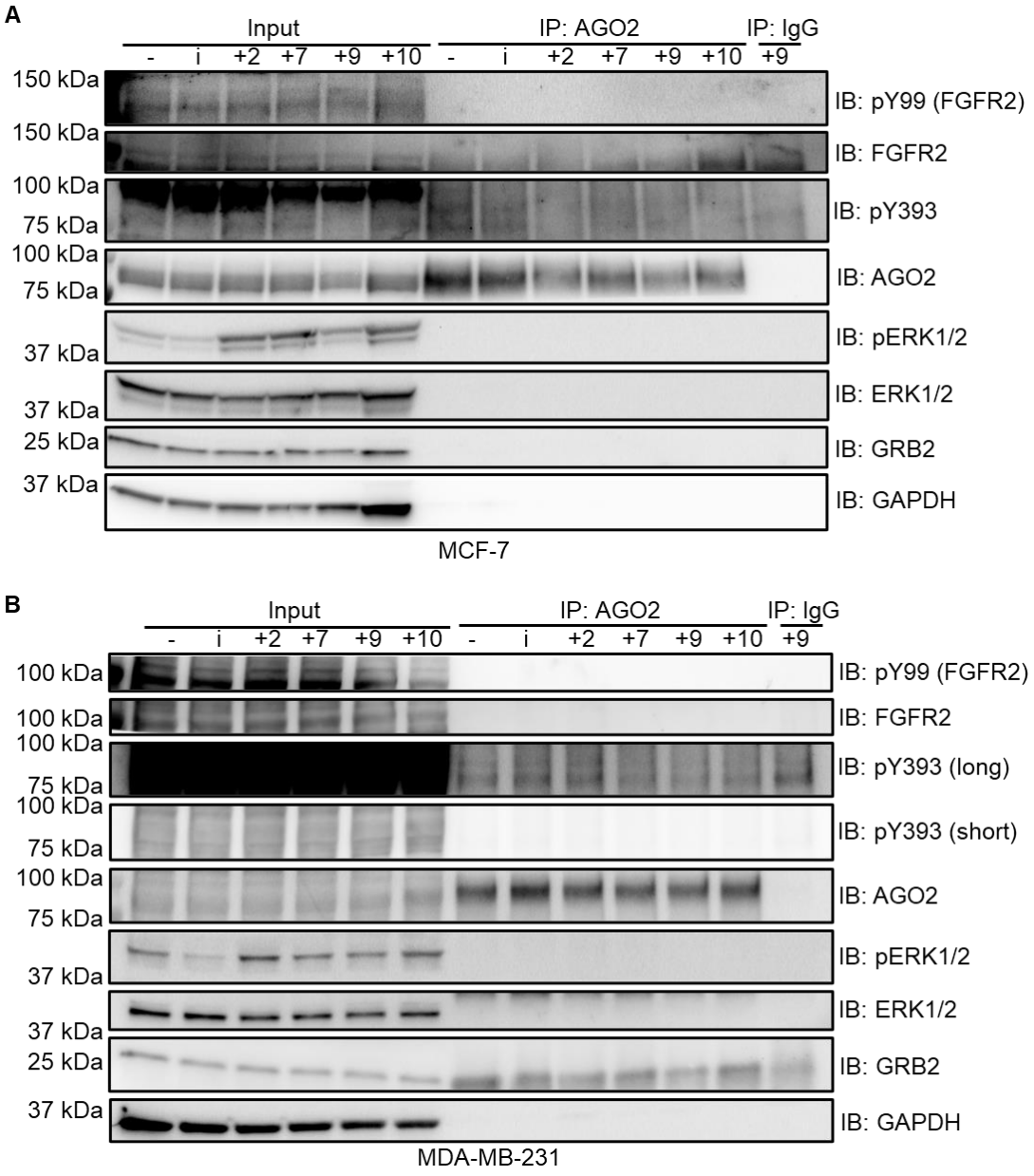


Figure 4.5.1 Y393 is not phosphorylated on endogenous AGO2 in breast cancer cells expressing endogenous FGFR2IIIb and FGFR2IIIc

(A) MCF-7 and **(B)** MDA-MB-231 cells were incubated in serum-free media containing FGFR2 inhibitor (i) or DMSO 18 h before stimulation with indicated FGF (+2, +7, +9, +10) or cells were left unstimulated (-). Lysates were used to immunoprecipitate (IP) AGO2 and Y393 phosphorylation (pY393) was analysed by western blotting (IB). Long and short refer to exposure time.

Previous studies have identified AGO2 phosphorylation (Shen et al., 2013; T. Liu et al., 2020). Therefore, to test if FGFR2 may phosphorylate Y393 when AGO2 is overexpressed, cells were transfected with GFP-AGO2. As the pY393 antibody appeared to not be specific for AGO2 pY393 in cell lysates (Figure 4.5.1 A, B), it was possible it also bound other sites on AGO2. Therefore, a mutant of AGO2 was used with the two previously described phosphotyrosyl sites on AGO2, Y529 and Y322, mutated to phenylalanine (GFP-AGO2 2YF). The mutant AGO2 was transfected into MDA-MB-231 and MCF-7 cells.

Cells were serum-starved in the presence, or absence, of FGFR2 inhibitor. They were then stimulated with a panel of FGFs. GFP proteins were pulled down and blotted for pY393. As with the previous experiment, FGFR2 activation could not be assessed by blotting pY99 and pY656/7 (the activation loop tyrosine), so instead ERK phosphorylation was examined (Figure 4.5.2 A, B). pERK was only slightly reduced upon incubation of MCF-7 cells with SU5402, when compared with serum-starved cells (Figure 4.5.2 B). Therefore, the FGFR2 inhibitor therefore may not have completely abrogated FGFR2 kinase activity in this cell line. However, a decrease in pERK was observed upon incubation of the inhibitor with MDA-MB-231 cells, suggesting that FGFR2 was inhibited in this cell line (Figure 4.5.2 A). Additionally, both cell lines showed increased ERK activation in response to stimulation with growth factor when compared to serum-starved cells (Figure 4.5.2 A, B), suggesting that FGFR2 was appropriately activated under these conditions. pERK was consistently higher in MCF-7 cells stimulated with FGF2, 7 and 10, whereas MDA-MB-231 cells displayed increased pERK upon stimulation with only FGF2 and 10 (Figure 4.5.1, Figure 4.5.2). Potentially these growth factors preferentially activate MAPK signalling, rather than other downstream pathways.

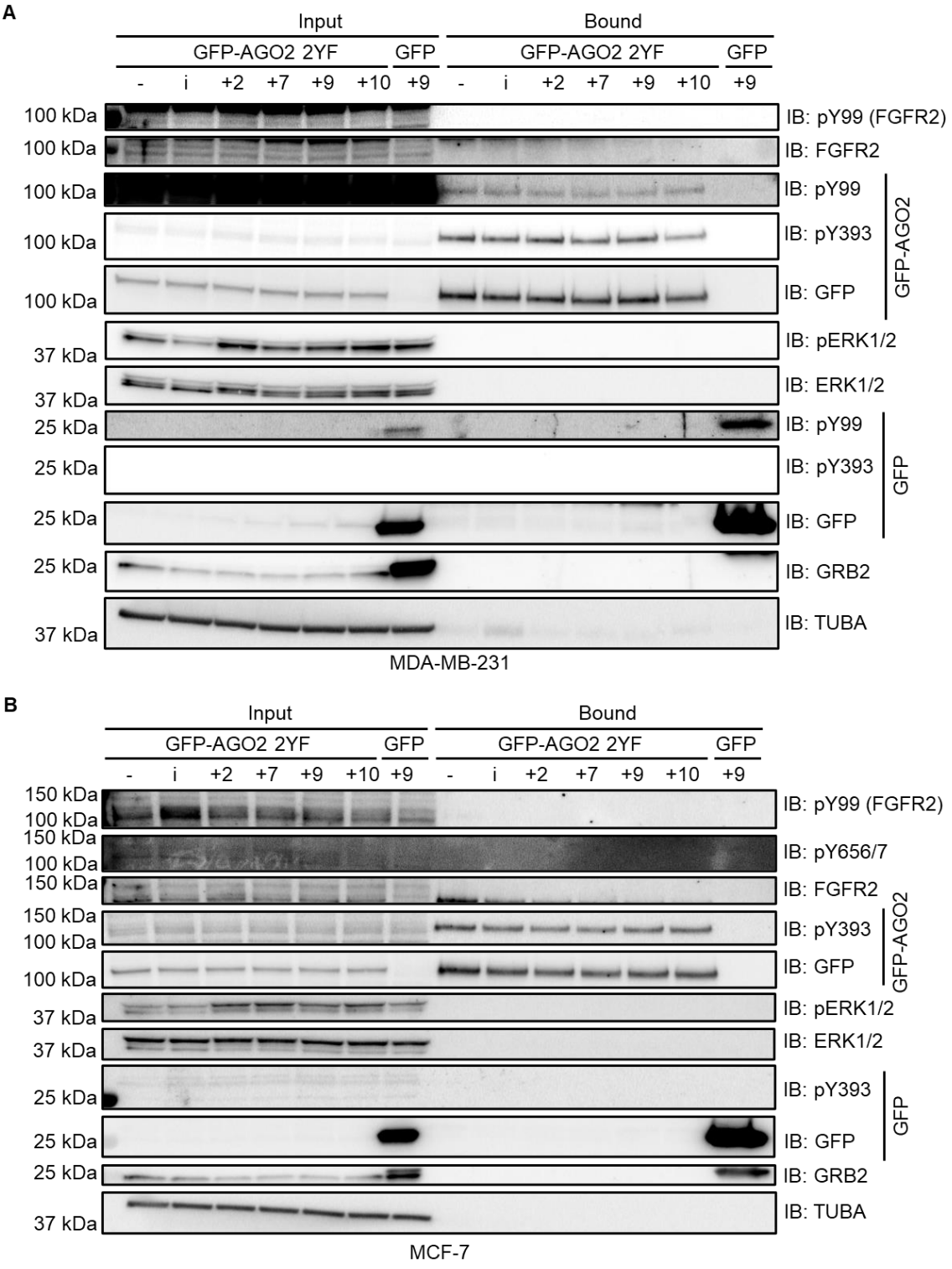


Figure 4.5.2 Y393 is phosphorylated on GFP-AGO2 but is not regulated by FGFR2 in cells with endogenous and stable FGFR2 expression

(A) MDA-MB-231 and **(B)** MCF-7 cells were transfected with GFP-AGO2 or GFP plasmids. Cells were incubated in serum-free media containing FGFR2 inhibitor (i) or DMSO 18 h before stimulation with indicated FGF (+2, +7, +9, +10) or were left unstimulated (-). Lysates were used to pulldown GFP proteins and Y393 phosphorylation (pY393) was analysed by western blotting (IB). α-Tubulin (TUBA) was used as the loading control.

Y393 phosphorylation could be observed on GFP-AGO2 2YF at similar levels in every condition, in both cell lines (Figure 4.5.2 A, B). The pY393 antibody did not pick up phosphorylation on GFP, suggesting it was specific for AGO2. Similarly, blotting with pY99 revealed phosphorylation of GFP-AGO2 2YF in MDA-MB-231 cells (Figure 4.5.2 A). However, unlike the pY393 antibody, the antibody for pY99 also detected phosphorylation on GFP, so only the pY393 antibody was valid in this experiment. It therefore appeared that Y393 was phosphorylated on GFP-AGO2 2YF in both breast cancer cell lines. However, there did not appear to be any FGFR2-dependent changes in Y393 phosphorylation; consequently, it was unlikely that the kinase responsible for this was FGFR2. Overall, both breast cancer cell lines demonstrate that FGFR2 stimulation does not regulate AGO2 pY393. MDA-MB-231 cells additionally show that Y393 phosphorylation is not mediated by basal activation of FGFR2.

To rule out the possibility that the lack of Y393 phosphorylation observed on endogenous AGO2 was due to poor binding of the antibody, new antibodies were generated (pY393 #1 and pY393 #2). The antibodies were tested using a range of *in vitro* and cell-based methods. Data suggested they were able to differentiate between phosphorylated and non-phosphorylated Y393 (data not shown). However, the antibodies also bound many other phospho-sites, including sites on FGFR2 (data not shown). Re-purification of the antibodies resulted in an extremely low yield (<4%, data not shown).

Crucially, the antibodies needed to be able to differentiate between non-phosphorylated and phosphorylated Y393 in the context of the full-length protein. To test this, an AGO2 construct was made with Y393 mutated to phenylalanine (GFP-AGO2 Y393F). GFP-AGO2 Y393F was transfected into HEK293T cells, which express EGFR. The cells were stimulated with EGF, which has previously been shown to result in phosphorylation of Y393 (Shen et al., 2013). A band at the expected molecular weight of GFP-AGO2 was observed for both wild type (WT) and Y393F mutant AGO2 (Figure 4.5.3). However, for both pY393 #1 and 2, this band was stronger for the mutant. It is therefore likely that the antibodies recognised phosphorylation on another protein of a similar molecular weight. Alternatively, they may have bound another phosphorylated residue on AGO2.

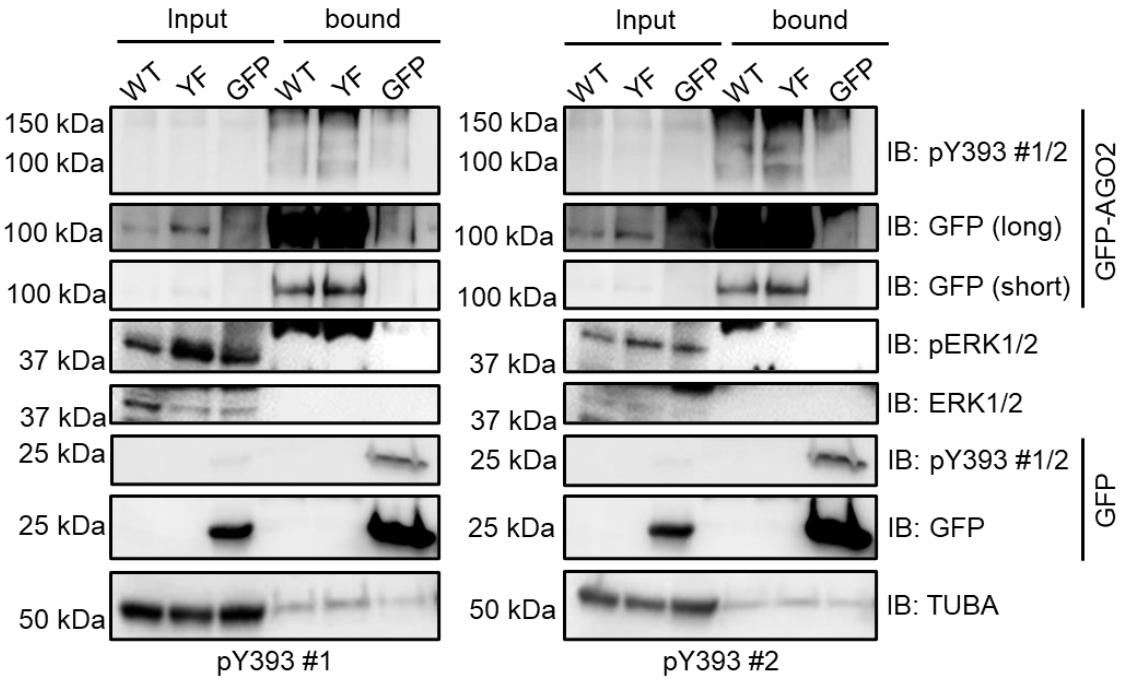


Figure 4.5.3 Custom generated antibodies are not specific for pY393

pY393 #1 and pY393 #2 were generated by GenScript. HEK293T cells overexpressing wild type GFP-AGO2 (WT), GFP-AGO2 Y393F (YF) and GFP were stimulated with EGF and GFP proteins were pulled down. pY393 #1 and 2 were then used to western blot (IB) both inputs and bound proteins. α-Tubulin (TUBA) was used as a loading control. Long and short refer to exposure time.

4.6 Discussion

4.6.1 Phosphorylation of Y393

GRB2 associated with both FGFR2 and AGO2 in unstimulated HEK293T cells (Figure 4.2.1 B). It was therefore proposed that GRB2 may recruit AGO2 to FGFR2 to allow FGFR2 to phosphorylate AGO2 Y393 under basal conditions. However, FGFR2-phosphorylation of AGO2 Y393 was not identified. Attempts to detect pY393 on GFP-AGO2 purified from stimulated HEK293T cells using mass spectrometry failed. Peptides containing pY393 were not identified for the EGF-stimulated positive control (Table 4.3.1), so the experimental procedure was likely to be at fault. To mimic the conditions in which EGFR-phosphorylated AGO2 Y393 was originally detected, both EGFR and AGO2 were overexpressed in HEK293T cells (Shen et al., 2013). Potentially the ratios of plasmids transfected needed optimising to ensure the corrected protein concentrations for phosphorylation. Additionally, several steps which have previously been described to assist visualisation of AGO2 phospho-sites were not completed (Rüdel et al., 2011). Although phosphatase inhibitors were included in the lysis buffer, cells were not treated with pervanadate prior to lysis. Additionally, phosphopeptides were not enriched by TiO_2 affinity purification.

The lack of FGFR2-mediated phosphorylation of AGO2 Y393 detected by western blotting may be more convincing. Y393 could not be detected by IP of endogenous AGO2 from breast cancer cell lines expressing endogenous FGFR2 (Figure 4.5.1 A, B). Similarly, no FGFR2-dependent changes in AGO2 Y393 phosphorylation were seen by pulldown of GFP-AGO2 2YF from the breast cancer cells (Figure 4.5.2 A, B). However, a background level of Y393 phosphorylation was observed in this experiment, suggesting these cells

express another kinase for Y393. MDA-MB-231 cells and MCF-7 cells both express EGFR (Osborne et al., 1981; Fitzpatrick et al., 1984) and c-SRC (Biscardi et al., 1998; Uhlén et al., 2015). Although EGFR and c-SRC require activation by growth factor, it is likely that some molecules remained active even in serum-starved cells. As a demonstration of this, pERK1/2 in serum-starved cells was reduced but not abrogated (Figure 4.5.1, Figure 4.5.2). Therefore, it is likely that EGFR or c-SRC was responsible for the phosphorylation of Y393 observed here (Shen et al., 2013; T. Liu et al., 2020).

Several possible explanations may exist for why the FGFR2-independent Y393 phosphorylation was seen by GFP pulldown but not endogenous AGO2 IP. Possibly the commercial pY393 antibody used here was not sensitive enough to detect low levels of pY393. Alternatively, AGO2 Y393 may not normally be phosphorylated in these cells, but overexpression of AGO2 may allow it to become phosphorylated. Y393 phosphorylation may therefore be more relevant in tumours where AGO2 is overexpressed, as has been shown for head and neck, colon, ovarian, bladder and gastric cancers (Chang et al., 2010; Li et al., 2010; Papachristou et al., 2011; Vaksman et al., 2012; Zhang et al., 2013; Yang et al., 2013). This is consistent with the previously identified roles for pY393 in promoting cancer pathogenesis (Shen et al., 2013; T. Liu et al., 2020).

On the other hand, the presence of Y393 phosphorylation on GFP-AGO2 2YF, where pY322 was inhibited, suggests phosphorylation of AGO2 Y322 may inhibit phosphorylation of Y393. In support of this, Y322 was phosphorylated by mass spectrometry (Table 4.3.2). Y322 is surface exposed (Figure 4.6.6 A), so protein binding the phosphorylated residue may restrict kinase access to Y393. It may therefore be interesting to 1) repeat the mass spectrometry experiment

with GFP-AGO2 2YF, and, 2) simultaneously blot pY393 on GFP-AGO2 WT and 2YF.

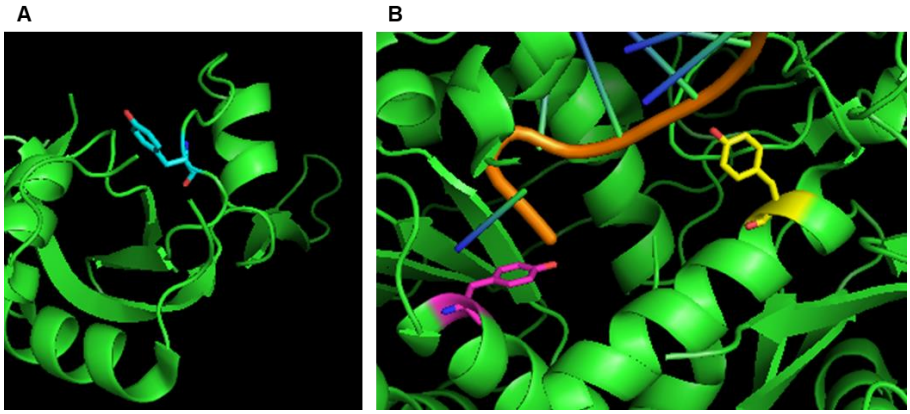


Figure 4.6.6 Positioning of AGO2 Y322, Y529 and Y804

Crystal structure of AGO2 (green) bound to a guide RNA (orange), with residues highlighted: **(A)** Y322 (blue), **(B)** Y529 (purple) and Y804 (yellow) (4W5N, Schirle & MacRae, 2012).

Y322 has previously been identified by mass spectrometry (Quévillon Huberdeau et al., 2017), although a role for it has not been characterised. Of the other two pY sites detected by mass spectrometry, Y529 has widely been reported to inhibit miRNA binding AGO2 through electrostatic repulsion (Rüdel et al., 2011; Mazumder et al., 2013; Patranabis and Bhattacharyya, 2016). Y804, as far as we know, has not before been identified. Both Y529 and Y804 point towards the miRNA (Figure 4.6.6 B). Therefore, Y804 may also preclude miRNA binding via steric hinderance.

4.6.2 DICER1 binding AGO2

Phosphorylation of Y393 inhibits DICER1 binding AGO2 (Shen et al., 2013). Consistent with the lack of FGFR2-dependent phosphorylation of Y393, no changes in the association of DICER1 with AGO2 were seen upon FGFR2-inhibition nor stimulation in HEK293T cells stably expressing FGFR2 (Figure 4.4.2, Figure 4.4.3, Figure 4.4.4). GRB2 been suggested to recruit PTP1B to AGO2 in *T. annulata*-infected macrophages, which express high levels of GRB2 (Haidar et al., 2018). If GRB2-overexpression is required to mediate FGFR2-phosphorylation of AGO2, simultaneous dephosphorylation of Y393 by PTP1B may explain the lack of changes in AGO2 binding DICER1 in these cells (Figure 4.4.3). Therefore, it may be interesting to assess the AGO2-DICER1 interaction in cells with both GRB2-overexpression and PTP1B-inhibition.

Chapter 5: GRB2 binds AGO2 directly to complex AGO2 and DICER1

In the previous Chapter it was demonstrated that GRB2 was associated with AGO2 in HEK293T cells (Figure 4.2.1 B). Therefore, GRB2 may bind AGO2 directly to regulate miRNA biogenesis.

This Chapter focuses on characterising the interaction mechanism between GRB2 and AGO2. The consequent effect of GRB2 on RLC assembly, independent of AGO2 phosphorylation, is also investigated. In the canonical pathway of miRNA biogenesis, the miRISC is formed via loading of a miRNA onto AGO2 by the enzyme DICER1 (Grishok et al., 2001). DICER1 processes the pre-miRNA to remove the hairpin loop, resulting in a 21-nucleotide miRNA duplex. DICER1 binds AGO2 directly through an interaction between DICER1 RNase III domain and AGO2 PIWI domain (Bernstein et al., 2001; Tahbaz et al., 2004). In this complex, DICER1 loads the miRNA onto AGO2 (Chapter 1, Figure 1.3.1).

While AGO2 and DICER1 form the core machinery for RISC-loading, other proteins bind the RLC to guide this process. Examples of such proteins include the DICER1 cofactors, TRBP and PACT (Chendrimada et al., 2005; Lee et al., 2006), as well as the DICER1-binding proteins ADAR1 and TDP43 (Kawahara and Mieda-Sato, 2012; Ota et al., 2013). Additionally, the HSC70/HSP90 chaperone systems stabilise AGO2 prior to miRNA-binding (Iwasaki et al., 2010; Iwasaki et al., 2015; Naruse et al., 2018; Tsuboyama et al., 2018; Park et al., 2019). ESR2 has also been demonstrated to associate with the RLC (Tarallo et al., 2017) Aside from these examples, few binding partners of AGO2 in the RLC have been identified. It is likely that other proteins which associated

with the RLC also regulate loading, such as via stabilisation of the complex or by protein recruitment. This Chapter reports an investigation into whether GRB2 can perform such a role.

5.1 GRB2 binds AGO2 under non-stimulated conditions

Preliminary work using GST-pulldown suggested GRB2 binds endogenous AGO2 in both unstimulated HEK293T cells and unstimulated HEK293T with stable FGFR2 expression (F1) (Figure 4.2.1). To assess whether this interaction occurs in multiple cell types, a GST-pulldown experiment was completed using HEK293T, NIH3T3 (mouse embryonic fibroblast) and PC3 (human prostate) cell lines. Cells were serum-starved before lysates were taken and incubated with GST-GRB2. Consistent with the results obtained using HEK293T cells, in both NIH3T3 and PC3 cell lines GST-GRB2 bound AGO2 while GST alone did not, or bound only a negligible amount (Figure 5.1.1 A). It is therefore likely that the GRB2-AGO2 interaction occurs ubiquitously and is not cell-type specific. As relatively little AGO2 was pulled down compared to the amount of GST-GRB2 it appears that the interaction is weak.

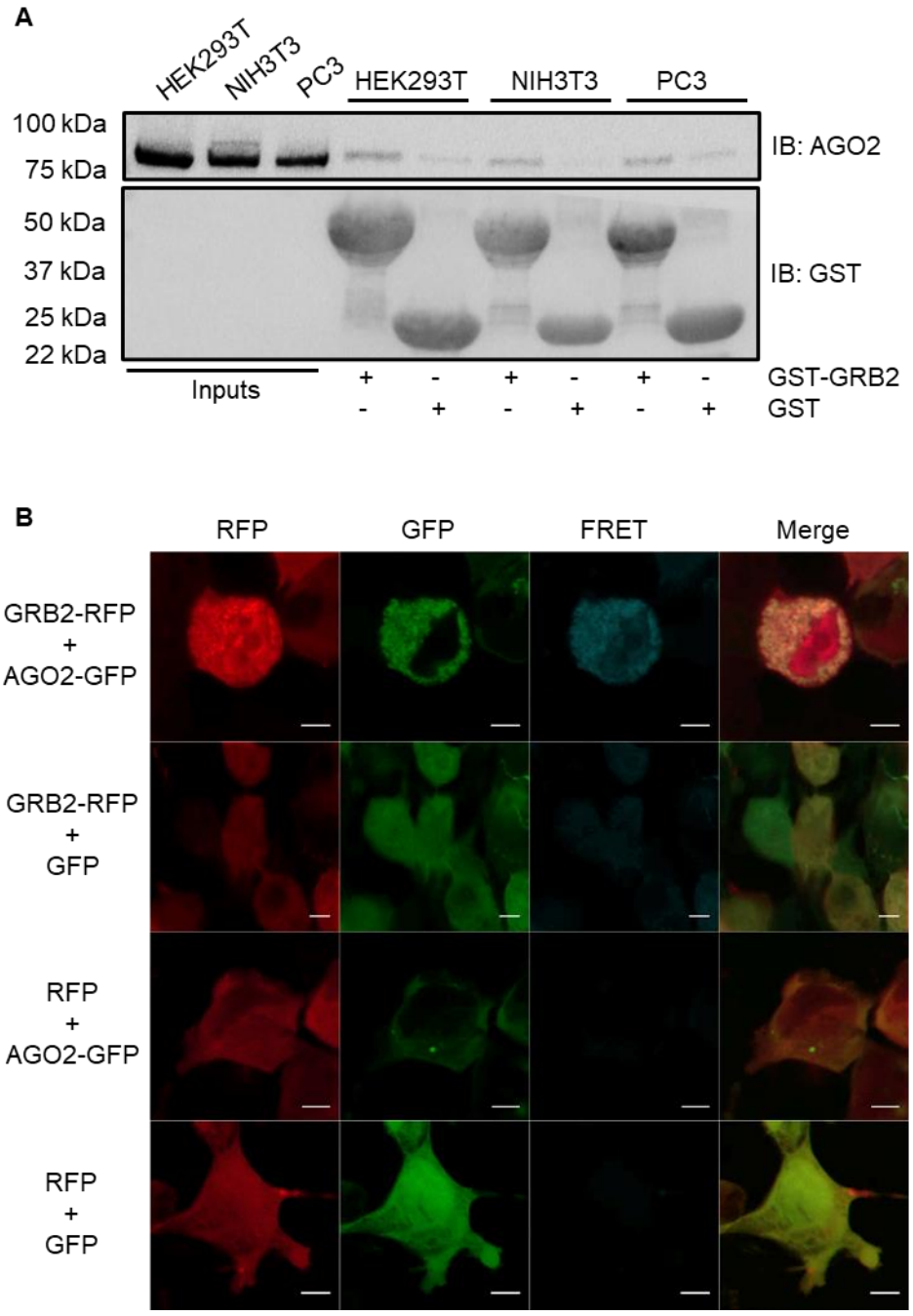


Figure 5.1.1 GRB2 complexes with AGO2 under basal conditions

(A) GST-GRB2 pulldown of AGO2 in three cell lines under non-stimulated conditions. Purified GST proteins were immobilised on beads and incubated with lysate from serum-starved cells. Bound proteins were analysed by western blotting (IB). **(B)** Fluorescence and Förster resonance energy transfer signals of RFP-tagged GRB2 and GFP-tagged AGO2. HEK293T cells overexpressing fluorescent proteins were serum-starved before imaging. N = 3. Scale bars are 5 µm.

Next, to demonstrate that GRB2 and AGO2 associate in a cellular environment, a FRET experiment was used. HEK293T cells were transfected with plasmids expressing GFP-tagged AGO2 and RFP-tagged GRB2 (or GFP and RFP as controls). The cells were then serum-starved and fixed onto coverslips. Finally, cells were imaged for GFP, RFP and FRET signal. As expected, GFP-AGO2 could be seen throughout the cell cytoplasm but not in the nucleus (Figure 5.1.1 B). RFP-GRB2, GFP and RFP were expressed in both the cytoplasm and the nucleus. Cells expressing both GFP-AGO2 and RFP-GRB2 had a greater FRET signal than the controls (Figure 5.1.1 B). Thus, these proteins are within close proximity (~10 nm) of each other and are likely to form part of the same complex. The HEK293T cell line used did not express FGFR2, confirming that GRB2 and AGO2 associated independently of this receptor. In fact, the FRET signal was observed throughout the cell cytoplasm, rather than only at the cell membrane. Therefore, it seemed the two proteins were able to form a complex which did not contain any membrane-bound receptor.

5.2 GRB2 directly binds AGO2 ^{323}PHLP^{326}

To ascertain if binding of GRB2 to AGO2 is direct, ITC was used. GRB2 bound AGO2 under basal conditions (Figure 5.1.1 A, B), when there are only low levels of tyrosine phosphorylation. Consequently, it is probable the interaction is via GRB2 SH3 domain binding AGO2 proline-rich motif. AGO2 has 3 regions containing canonical PxxP motifs (Figure 5.2.1 A). Therefore, three peptides spanning each motif were used (AGO2 #1, #2 and #3, see Table 5.2.1 for sequences). 1.2 mM peptide was titrated into a solution of recombinant purified 80 μM GRB2. Only peptide AGO2 #3, spanning residues 317-333 and located in the PAZ domain of AGO2, showed binding to GRB2 (Figure 5.2.1 B-D, Table

5.2.1). A symmetrical binding curve was produced, indicating binding of the peptide to a single site on GRB2. The curve was therefore fitted with a 1:1 binding model. However, a stoichiometry of 1.48 ±0.04 sites was produced. It is likely that this stoichiometry reflected a lower real concentration of peptide than was input in the analysis software. Degradation of the peptide may have occurred through repeated freeze-thawing, exposure to oxygen and storage in solutions above pH 8. Consequently, the binding curve was refitted with manual adjustments in peptide concentration until a stoichiometry of n = 1 was reached. The final peptide concentration was 800 μM. A K_D of 4.27 ± 1.17 μM was produced, which is characteristic of SH3 domain interactions (Yu et al., 1994). It is important to note that at 80 μM GRB2 was a dimer, so a complex was formed consisting of two GRB2 proteins each bound to a peptide on the same site (2:2 binding). From the crystal structure of AGO2 it is clear that the ^{323}PHLP^{326} motif is surface exposed (Figure 5.2.1 E) (Schirle, Sheu-Gruttadauria, & MacRae, 2014) . Therefore, in the context of the entire protein, this motif is available to bind GRB2.

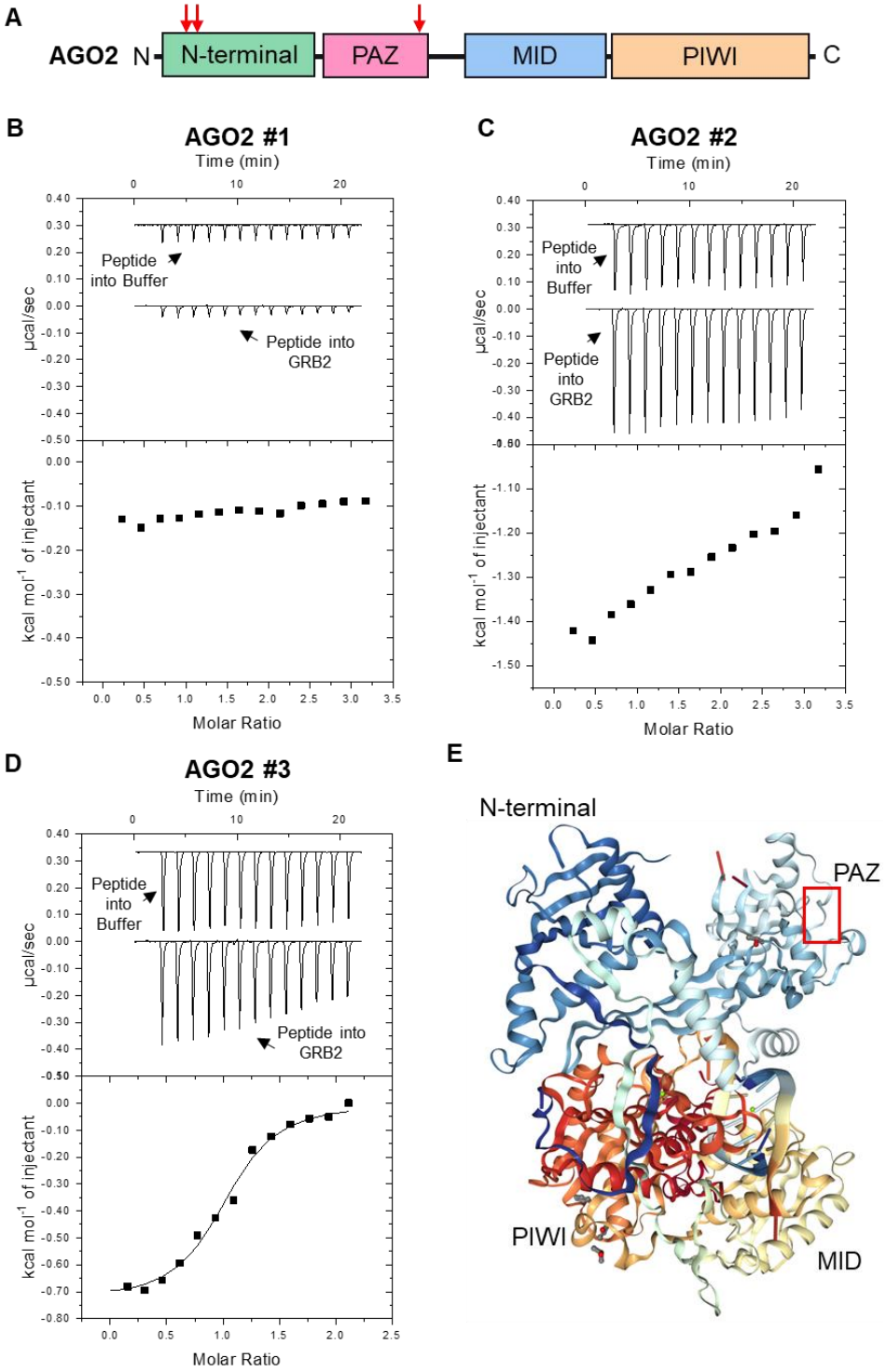


Figure 5.2.1 A peptide spanning a proline-rich region in the PAZ domain of AGO2 binds GRB2 by ITC

(A) Domain structure of AGO2. Proline rich regions containing canonical PxxP motifs are indicated by red arrows. **(B-D)** Isothermal titration calorimetry isotherms of GRB2 with three peptides spanning each of AGO2 proline-rich regions. Titration of peptide AGO2 #3 into GRB2 gives a measured affinity of 4.27 ± 1.17 µM. A solution containing 1.2 mM peptide was titrated into an 80 µM solution of GRB2 or buffer alone as a control. Peaks were integrated to give a Wiseman curve which was fitted using a 1:1 binding model to yield the binding affinity, K_D, and stoichiometry, n. **(E)** The crystal structure of AGO2 (PDB: 4W5N, Schirle et al., 2014). ^{323}PHLP^{323} is indicated inside the red box.

Table 5.2.1 AGO2 peptides tested for binding to GRB2

Sequences, locations in AGO2 and measured constants of peptides used in isothermal titration calorimetry with GRB2. Prolines with the PxxP consensus are highlighted in red.

Peptide	Residues	Domain	Sequence	Affinity	Stochiometry
AGO2 #1	1-20	N-terminal	MYSGAGPALAPPAPPPPIQG	-	-
AGO2 #2	21-40	N-terminal	YAFKPPPRPDFGTSGRTIKL	-	-
AGO2 #3	317-333	PAZ	KLVLRYPHLPCLQVGQE	4.27 µM ± 1.17 µM	1.00 ±0.03 sites

To confirm that GRB2 binding AGO2 is mediated by the PAZ domain, ITC was used to probe the interaction of GRB2 with isolated MBP-tagged AGO2 domains. 500 μM MBP-tagged AGO2 domains were titrated into a solution of 20 μM GRB2, or buffer alone as a control. MBP alone, the N-terminal domain, the MID domain and the PIWI domain, did not produce any evidence of binding at 25°C (Figure 5.2.2 A-D). Conversely, a binding curve was observed upon titration of AGO2 PAZ domain into GRB2 (Figure 5.2.2 E). The higher heat produced by titration of MBP-PIWI into GRB2 than of any other domain may have indicated binding of this domain as well. However, a control titration of MBP-PIWI into buffer also produced peaks of this magnitude (Figure 5.2.2 D), indicating that the heat did not represent a GRB2-PIWI interaction. The binding curve produced upon titration of the PAZ domain was symmetrical, suggesting it bound one site on GRB2. The curve was therefore fitted with a 1:1 binding model. A stoichiometry of 2.39 ±0.02 was produced, indicating that two PAZ domains bound each GRB2 protein. However, it was unlikely that there were two binding events as the binding curve was not biphasic. As the PAZ protein preparation was not nuclease-treated, it is possible that the presence of RNA affected the concentration of free PAZ domain. For example, two PAZ domains could be bound to the same piece of RNA or DNA. The curve was therefore refitted with 212 μM PAZ at which n = 1. The resulting K_D of 585 nM ±61 nM was similar to that measured with the PAZ domain peptide. Consequently, this provided stronger evidence that it is this domain which binds, and that binding is via an SH3 domain interacting with a proline-rich motif.

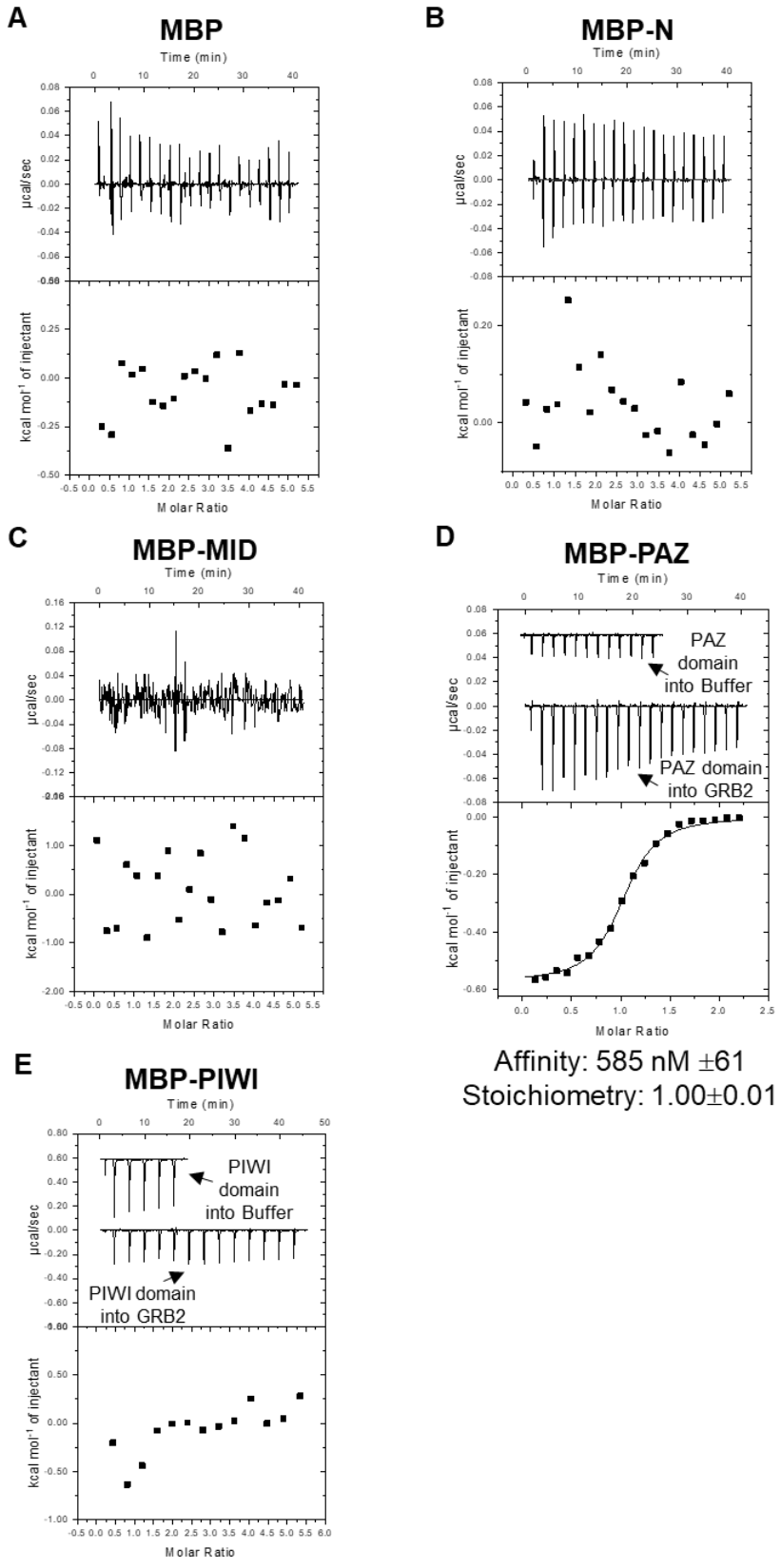


Figure 5.2.2 Only the PAZ domain of AGO2 interacts with GRB2

(A-F) Isothermal titration calorimetry isotherms of MBP-tagged AGO2 domains titrated into GRB2. An affinity of 585 ±61 nM was measured for MBP-PAZ. A solution containing 500 µM AGO2 domain was titrated into a 20 µM solution of GRB2 or buffer alone as a control. Peaks were integrated to give a Wiseman curve which was fitted using a 1:1 binding model to yield the binding affinity, K_D, and stoichiometry, n.

In an attempt to analyse the make-up of PAZ-GRB2 complexes, GRB2 and PAZ were crosslinked. His-tagged PAZ was RNase-treated to eliminate any effect of RNA on PAZ-binding GRB2. RNase catalyses the hydrolysis of phosphodiester bonds in the RNA backbone, thus removing RNA from the protein preparation. Purified GRB2 and His-PAZ, both at 250 μM, were incubated in a 10-fold excess of disuccinimidyl tartrate (DST) and resulting complexes analysed by SDS-PAGE. Two crosslinked products were immediately visible, corresponding to: 1) a complex ~40 kDa, probably one GRB2 bound to one PAZ domain, and 2) a complex ~55 kDa, either GRB2 dimer (GRB2-GRB2) or one GRB2 bound to two PAZ domains (PAZ-GRB2-PAZ, Table 5.2.2, Figure 5.2.3 A). However, the ~55 kDa complex was unlikely to be PAZ-GRB2-PAZ, as a single binding event between PAZ and GRB2 was observed by ITC (Figure 5.2.1 D, Figure 5.2.2 D). Separation of the complexes by SEC and analysis by SDS-PAGE also revealed higher order complexes. In addition to the complexes at ~40 kDa and ~55 kDa, a complex at ~80 kDa was observed, corresponding to GRB2 dimer bound to two PAZ domains in an overall ratio of 1:1 (PAZ-GRB2-GRB2-PAZ, Figure 5.2.3 B). Overall, this data suggested that PAZ and GRB2 form a GRB2-PAZ complex via a single binding surface, but further work is required to conclusively determine if, in vitro, PAZ can also bind another site on GRB2, allowing GRB2 to interact with two PAZ domains at once.

Table 5.2.2 Predicted molecular weights of possible complexes containing GRB2 and PAZ

Molecular weights were calculated from the sequence of each recombinant protein.

Complex	Molecular Weight (kDa)
PAZ	16.0
GRB2	26.2
GRB2-PAZ	42.2
GRB2-GRB2	52.5
PAZ-GRB2-PAZ	58.2
GRB2-GRB2-PAZ	68.5
PAZ-GRB2-GRB2-PAZ	84.5

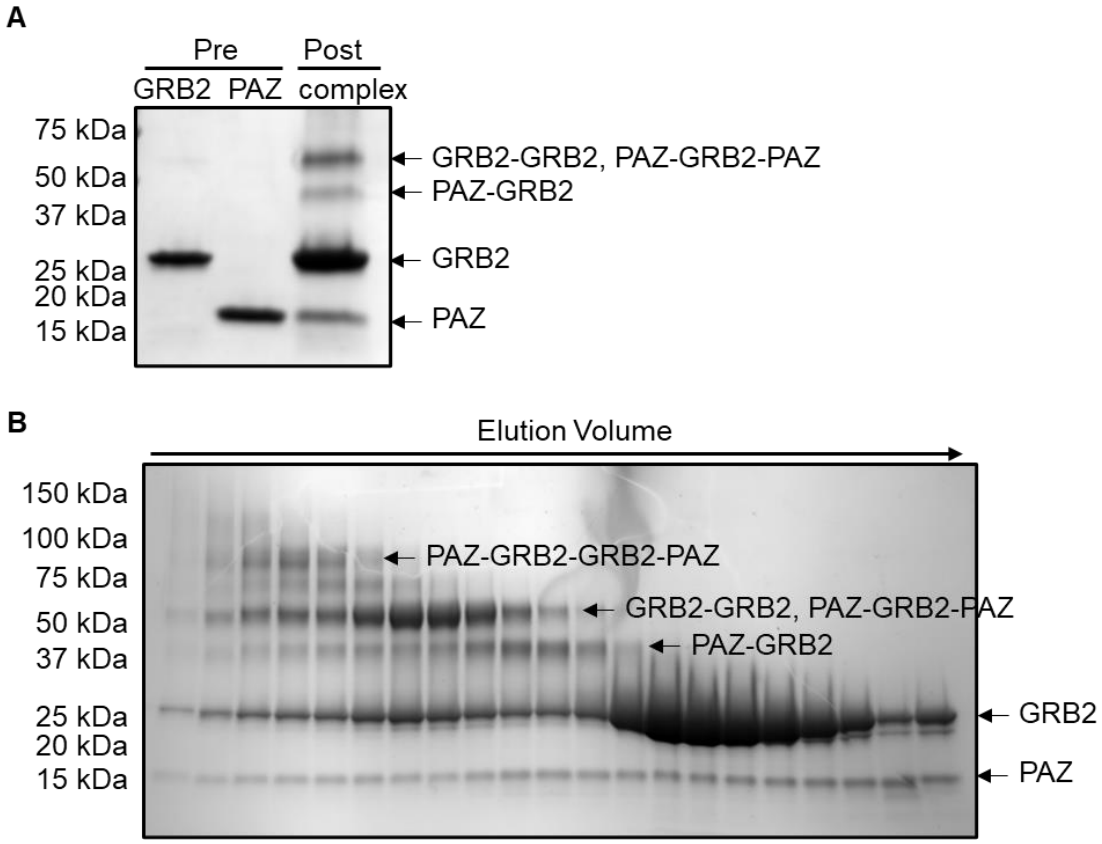


Figure 5.2.3 Crosslinking of GRB2 and PAZ suggest multiple complexes may form *in vitro*

(A) Pure GRB2 and PAZ proteins before and after crosslinking and **(B)** after size exclusion chromatography to separate crosslinked proteins. 250 μM GRB2/His-PAZ was crosslinked using 2.5 nM DST before analysis by SDS PAGE. Uncrosslinked PAZ and GRB2 and crosslinked complexes are indicated.

To determine if binding of the PAZ domain to GRB2 was dependent on the ^{323}PHLP^{326} motif, ITC was again used. Residues 323-326 of MBP-tagged AGO2 were mutated to four alanines (MBP-PAZ 4A). 500 μM MBP-PAZ 4A was titrated into a solution of 20 μM GRB2. Mutation abrogated PAZ binding GRB2 (Figure 5.2.4 A), confirming that these residues were essential for the interaction.

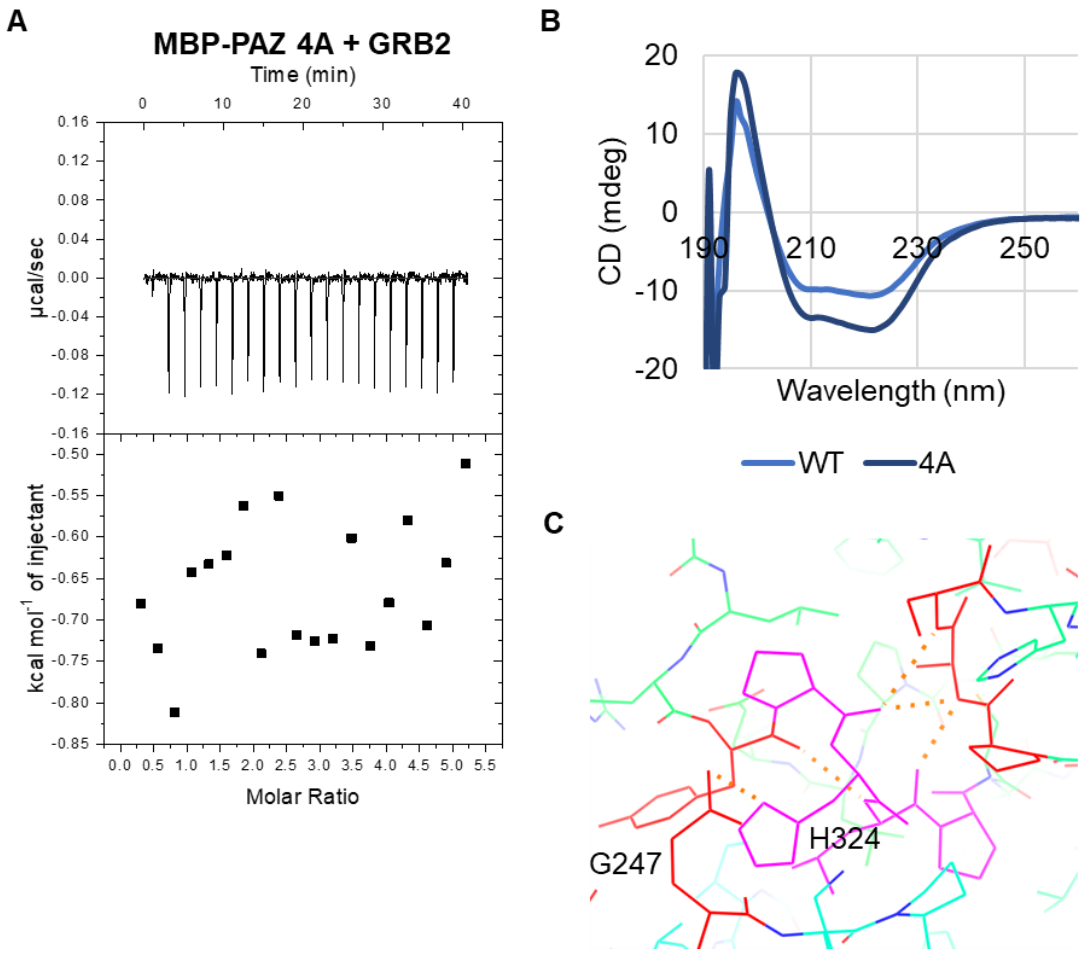


Figure 5.2.4 Mutation of ^{323}PHLP^{326} abrogates PAZ binding GRB2

(A) Isothermal titration calorimetry isotherm of MBP-PAZ with ^{323}PHLP^{326} mutated to four alanine residues (4A) with GRB2. A solution containing 500 µM MBP-protein was titrated into a 20 µM solution of GRB2. **(B)** Far-UV circular dichroism spectra of wild type (WT) and 4A MBP-PAZ secondary structure. Measurements were made at pH 7.5 and 20°C. **(C)** Crystal structure of AGO2 containing ^{323}PHLP^{326} (magenta). Hydrogen bonds with interacting residues (red) are indicated. (PDB: 4W5N, Schirle et al., 2014).

The lack of binding to MBP-PAZ 4A may have been due to mutation of ^{323}PHLP^{326}, or alternatively could have been a result of destabilisation of the overall protein structure. To test this, the circular dichroism (CD) spectra of wild type (WT) MBP-PAZ and MBP-PAZ 4A was measured. Spectra were recorded at 190-260 nm to investigate secondary structure. Changes in protein structure can be identified by a shift in the CD spectra. However, the spectra were almost identical (Figure 5.2.4 B). It should be noted that the presence of the MBP-tag would have contributed to the observed spectra. As both proteins had this tag, the chance of seeing smaller changes in PAZ domain structure may have been reduced. Both spectra had a high alpha helical content, as shown by minima at 210 and 221 nm. There was a very small shift in the point at which each ellipticity = 0 mdeg (WT = 202.0, 4A = 202.3). However, this may have been expected as the replacement of histidine 324 with an alanine residue may disrupt hydrogen bonding with glutamine 247 to slightly destabilise the surrounding region (Figure 5.2.4 C). The changes in peak size between the two spectra were the result of small differences in protein concentration. Additionally, the noise observed before 200 nm was due to presence of glycerol in the buffer, which absorbs light below this wavelength. Overall, this data suggested that the 4A mutation did not result in PAZ unfolding. This was consistent with the location of ^{323}PHLP^{326} in a loop region (Figure 5.2.1 E).

As all ITC experiments had been completed with PAZ domain which had not been RNase treated, it was possible that residual RNA was bound to the domain in these experiments. To assess whether RNA binding to PAZ could have had an impact on the GRB2-PAZ interaction, AGO2 PAZ domain was pre-treated with RNase prior to binding analysis with MST. A dilution series of unlabelled His-PAZ was set up and mixed 1:1 with labelled GRB2 to final

concentrations 0.02-653 μM PAZ and 100 nM GRB2. Importantly, GRB2 is a monomer at 100 nM. Binding curves were produced by both PAZ preparations. Curves were fitted according to the K_D fit model, which assumes one to one binding. Use of PAZ domain which had not been treated with RNase gave a dissociation constant of 17.80 ±0.95 μM (Figure 5.2.5 A) and RNase-treated PAZ domain bound GRB2 with a similar affinity: K_D 6.14 ±0.31 μM (Figure 5.2.5 B).

Interestingly, the change in normalised fluorescence for RNase-treated PAZ domain was half that of the untreated protein. A smaller change in fluorescence is indicative of formation of a smaller complex. Consequently, this data was consistent with the hypothesis that presence of RNA dimerised PAZ and allowed two PAZ domains to simultaneously bind to GRB2, thus helping to explain the initial stoichiometry of n ~2 by ITC (Figure 5.2.2 D). RNase treatment destroyed the PAZ dimers and reduced the size of the final GRB2-PAZ complex. However, the similar K_Ds in both experiments indicated that the presence of RNA did not significantly affect the affinity of GRB2 for AGO2 PAZ domain.

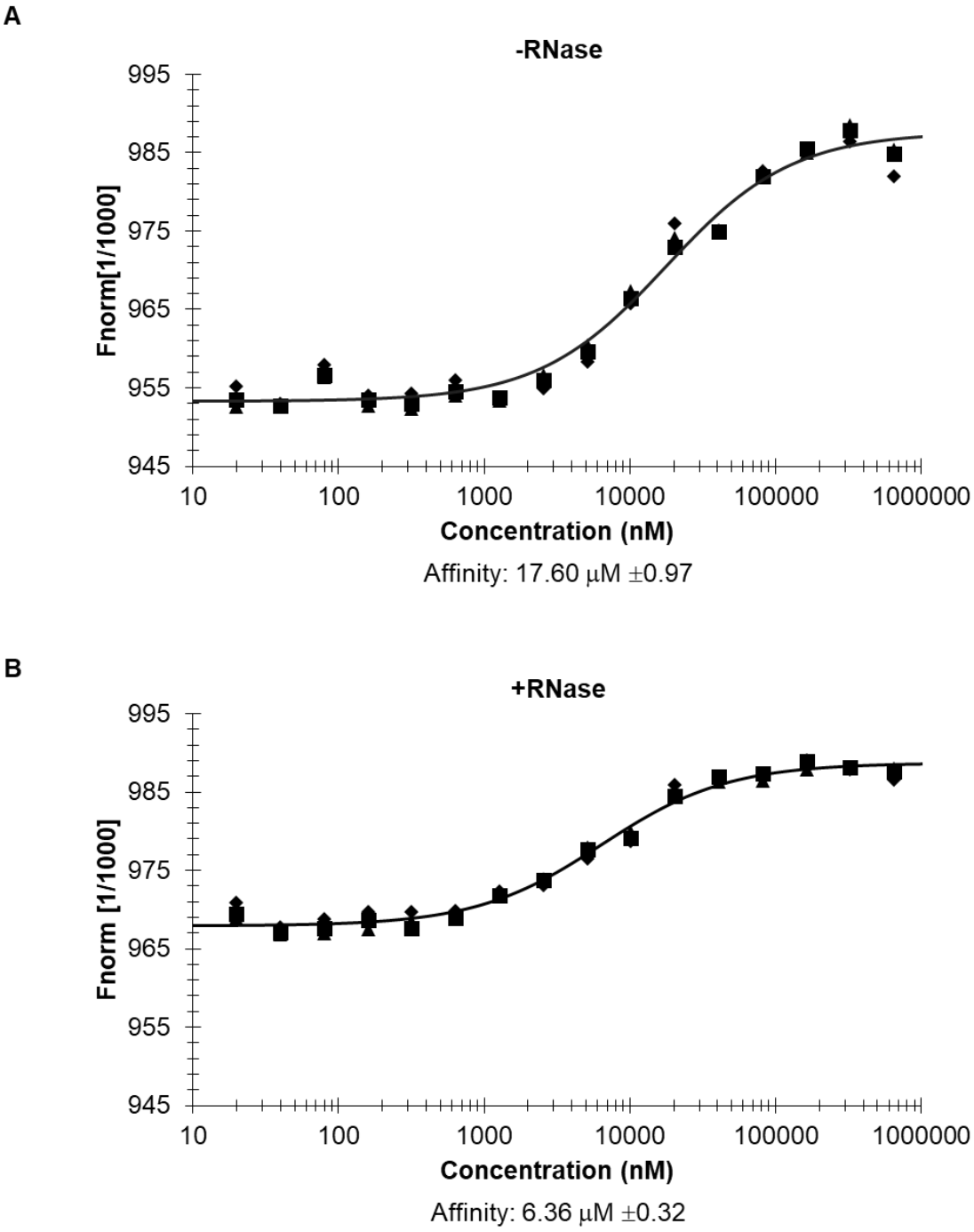


Figure 5.2.5 Treatment of His-PAZ with RNase does not affect GRB2-binding

(A, B) Microscale thermophoresis of GRB2 with His-tagged PAZ domain which had or had not been treated with RNase. A dilution series of 20-653,000 nM unlabelled PAZ and 100 nM labelled GRB2 was used. Reactions were performed in triplicate.

5.3 GRB2 NSH3 domain binds AGO2 under non-stimulated conditions

Biophysical studies indicated that AGO2 PAZ domain binds a single site on GRB2 via a PxxP motif. GRB2 has two SH3 domains (Figure 5.3.1 A). Therefore, to ascertain which domain AGO2 bound a GST-pulldown experiment was used to determine which GRB2 domain bound AGO2. GST-tagged GRB2 domains were purified and immobilised on beads. HEK293T cells were serum-starved and lysates incubated with the GST-beads, before bound proteins were analysed by western blotting. Binding of AGO2 to the N-terminal SH3 domain (NSH3) was apparent, suggesting this domain binds AGO2 (Figure 5.3.1 B, C). In some replicates of the experiment limited binding was observed for the C-terminal SH3 domain (CSH3). Negligible binding was observed to the SRC homology 2 (SH2) domain (Figure 5.3.1 B, C). It therefore appeared that AGO2 primarily binds NSH3. However, potentially through use of high protein concentrations and due to the promiscuous nature of SH3 domain interactions, AGO2 may also weakly bind CSH3 in these experiments.

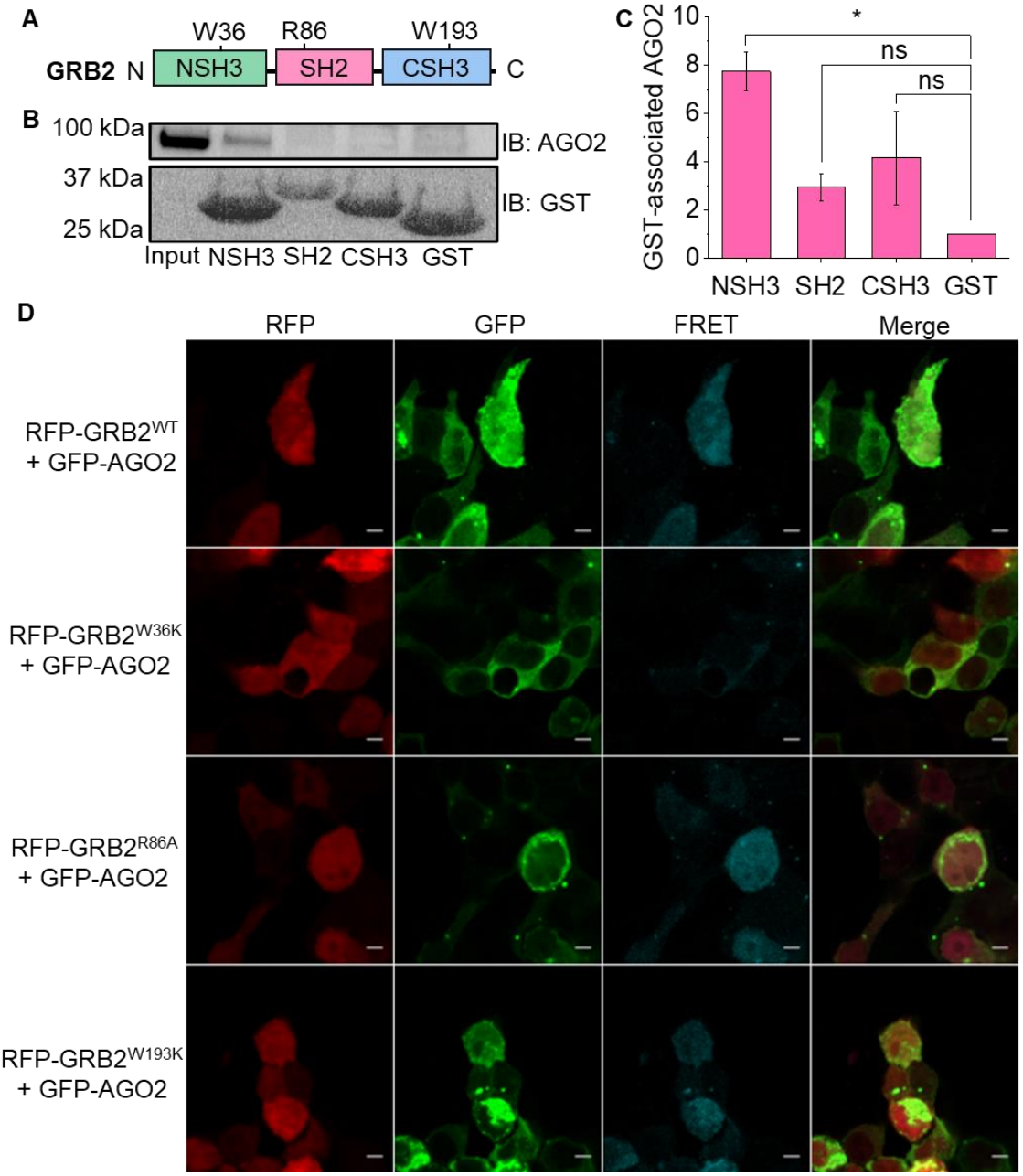


Figure 5.3.1 GRB2 N-terminal SH3 domain associates with AGO2 under non-stimulated conditions

(A) Domain structure of GRB2. Indicated residues can be mutated to knock out function of each domain. **(B)** Western blot (IB) and **(C)** quantification of AGO2 associated with GST-tagged GRB2 domains, normalised to GST. Purified GST-proteins were immobilised on beads and incubated with lysate from serum-starved HEK293T cells. N = 3. *p < 0.05. ns = not significant by Student's t-test. **(D)** Förster resonance energy transfer between wild type (WT) and mutant RFP-tagged GRB2 and GFP-tagged AGO2 proteins expressed in serum-starved HEK293T cells. Scale bars are 5 µm.

To confirm the domain(s) of GRB2 which bind AGO2, FRET was used. Three GRB2 constructs were made with a single point mutation in each, to knockout the function of each domain individually. Mutation only of the domain(s) which bound AGO2 would result in a decrease in FRET, compared to WT GRB2. The pY-binding deficient SH2 domain mutant (GRB2 R86A, Huang & Sorkin, 2005) and the PxxP-binding incompetent NSH3 and CSH3 domain mutants (GRB2 W36K and GRB2 W193K respectively, Bisson et al., 2011) were selected (Figure 5.3.1 A). Cells were transfected with the appropriate RFP-GRB2 plasmid along with GFP-AGO2. After serum starvation, the cells were fixed and imaged for FRET. FRET signal was reduced with the W36K mutant but not with the R86A or W193K mutants. (Figure 5.3.1 D). Therefore, mutation of the SH2 and CSH3 domains did not appear to inhibit GRB2 binding AGO2. In contrast, mutation of NSH3 diminished the GRB2-AGO2 interaction. Therefore, this data supported the conclusion that it was only this domain which bound GRB2 in cells.

MST was performed to further elucidate which domain of GRB2 binds AGO2. A dilution of unlabelled His-PAZ was set up and mixed 1:1 with labelled GRB2, resulting in final concentrations 0.02-653 μM PAZ and 100 nM GRB2. Mutations were introduced into GRB2 to knockout the binding capacities of: 1) NSH3 only (W36K), 2) CSH3 only (W193K), and, 3) both N and CSH3 (W36/193K). MST of $GRB2^{W193K}$ with PAZ produced a binding curve. Fitting the curve gave a measured K_D of 13.10 ±1.61 μM (Figure 5.3.2 B). The K_D is similar to that reported for wild type GRB2 (Figure 5.2.5), suggesting that CSH3 is dispensable for binding.

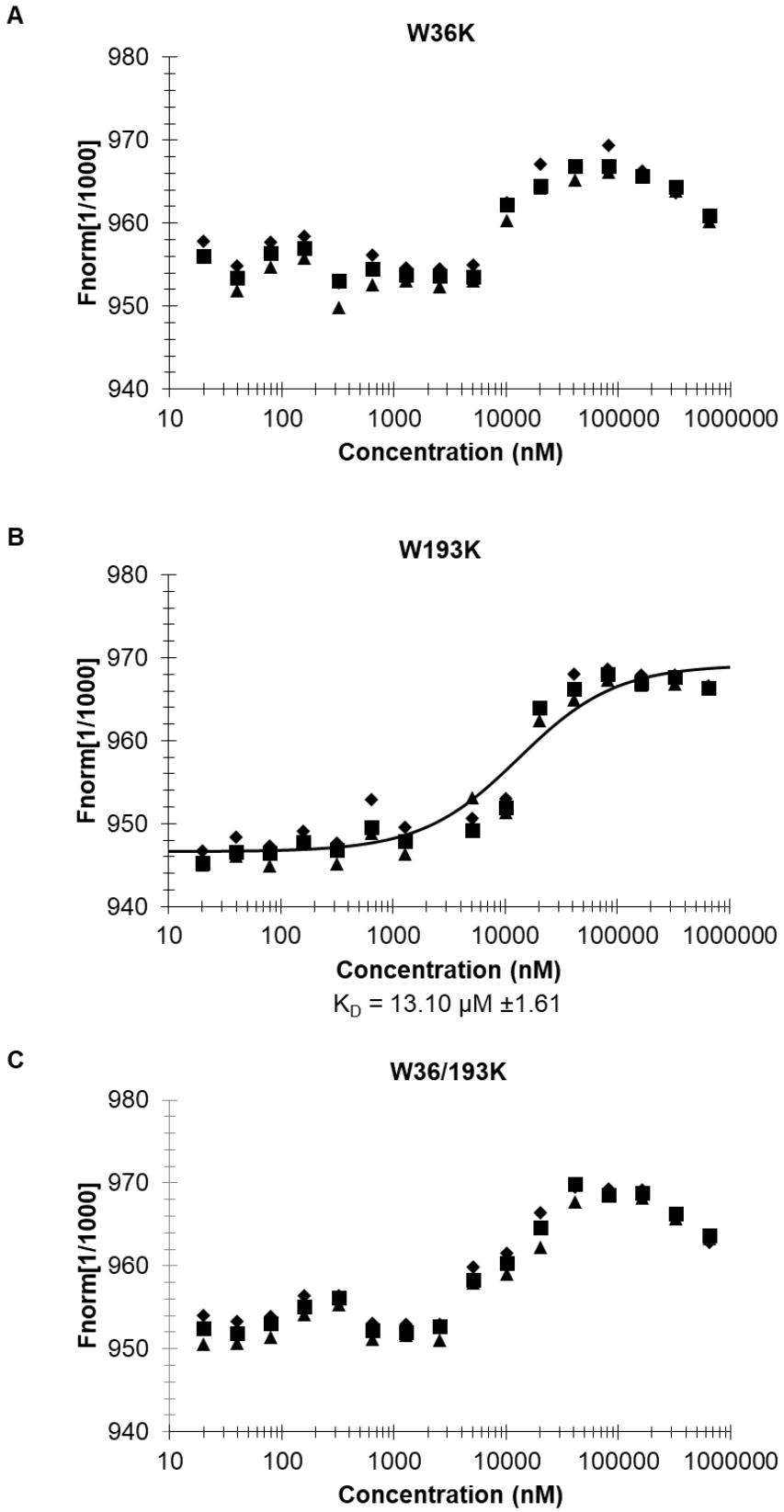


Figure 5.3.2 Mutation of GRB2 NSH3, but not CSH3, reduces binding to PAZ

Microscale thermophoresis of **(A)** GRB2 W36K (inhibits NSH3 binding), **(B)** W193K (inhibits CSH3 binding) and **(C)** W36/193K with His-tagged PAZ domain which had not been treated with RNase. A dilution series of 20-653,000 nM unlabelled PAZ and 100 nM labelled GRB2 was used. Reactions were performed in triplicate.

For the W36K and W36/193K mutant GRB2 proteins an initial increase in normalised fluorescence indicated binding to PAZ. At high concentrations of PAZ fluorescence decreased (Figure 5.3.2 A, C). Therefore, it is likely that in these reactions, self-association of PAZ disrupted the GRB2-PAZ complex. As this occurred only with the W36K-containing mutants, it appeared that mutation of NSH3 reduced the affinity of GRB2 for PAZ.

Having demonstrated that AGO2-PAZ binds to the NSH3 of GRB2, the structural basis of the interaction was investigated. To model how GRB2 may bind AGO2 PAZ domain, an *in silico* docking experiment was performed using the HADDOCK2.2 web server (van Zundert et al., 2016). HADDOCK uses an information-driven approach to determine protein-protein interactions, employing experimental and predicted data to determine the structure of the protein complex. It consistently performs well in Critical Assessment of PRedicted Interactions (CAPRI) experiments (Janin, 2005; Lensink and Wodak, 2013). However, it should be acknowledged that some structures may be hard to predict in a docking experiment, for example if the two interacting proteins exist as part of a multicomponent assembly.

To facilitate docking of GRB2 NSH3 to AGO2 PAZ, specific residues were selected as the active residues involved in binding. The hydrophilic binding pocket of GRB2 was defined by residues Y7, F9, W36, F47, P49, and Y52 (Vidal et al., 1999). The ^{323}PHLP^{326} motif was selected for PAZ. A model of GRB2 bound to PAZ was successfully produced (Figure 5.3.3). A large binding interface across GRB2 NSH3 could be seen, with W36 clearly interacting with ^{323}PHLP^{326} (Figure 5.3.3). In addition to the residues used to define GRB2 NSH3, F9, T12, D15, Q34, N35, F47, N51 and Y52 potentially contributed to the interaction (Figure 5.3.3).

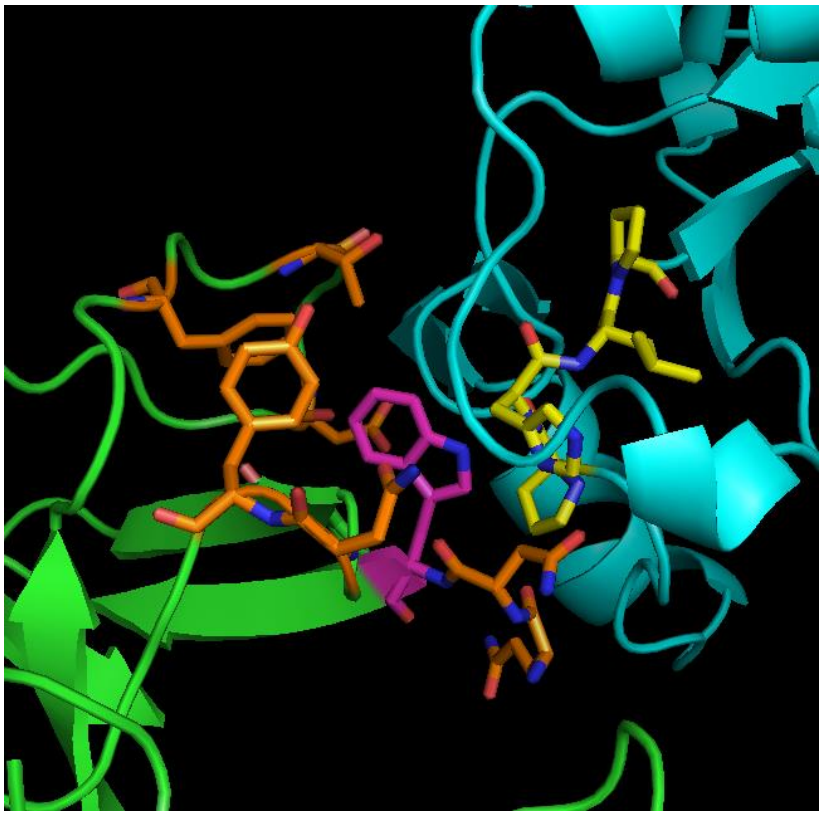


Figure 5.3.3 *In silico* modelling of AGO2 PAZ domain bound to GRB2 NSH3

Crystal structure of AGO2 PAZ domain (PDB: 6RA4; Greenidge et al., 2019; blue)) docked to GRB2 (PDB: 1GRI; Maignan et al., 1995; green) via an interaction between GRB2 W36 (purple) and ^{323}PHLP^{326} (yellow). Other residues which may contribute towards the interaction are shown in orange. Docking was performed using HADDOCK2.2 (van Zundert et al., 2016). Data contributed by Chi-Chuan Lin, University of Leeds.

As it appears that many residues may mediate association of GRB2 to AGO2, mutation of W36 alone may have been insufficient to completely inhibit the interaction in MST experiments (Figure 5.3.2). Therefore, mutation of W36 may have reduced the affinity, but could not abrogate binding. Overall, both the biophysical and cell-based experiments indicated that only the N-terminal SH3 domain of GRB2 binds AGO2 PAZ domain.

5.4 GRB2 binds AGO2-DICER1 to form a trimeric complex

The data presented thus far suggest that under non-stimulated conditions GRB2 NSH3 binds AGO2 ^{323}PHLP^{326}. To determine if GRB2 forms a larger complex containing both AGO2 and DICER1, GST-pulldown was used. HEK293T cells were deprived of growth factor and lysates were taken. Lysates were incubated with recombinant GST-tagged GRB2 and bound proteins were analysed by western blotting. GRB2 associated with both AGO2 and DICER1 (Figure 5.4.1 A). The presence of GRB2 in a complex with the RNA-processing proteins AGO2 and DICER1 might suggest a miRNA-loading regulatory role for the adaptor protein.

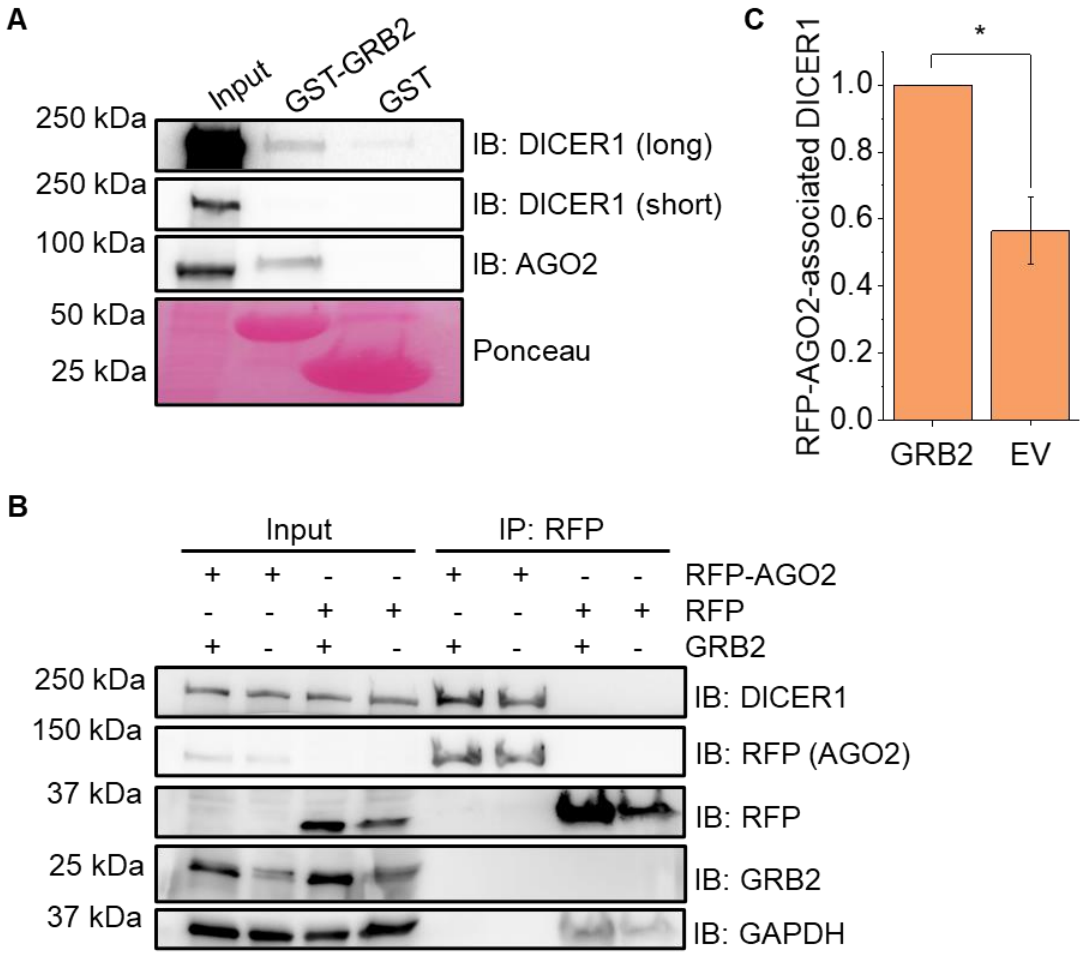


Figure 5.4.1 GRB2 may slightly stabilise AGO2 binding DICER1 under non-stimulated conditions

(A) Western blot of AGO2 and DICER1 pulldown by GST-GRB2 in HEK293T cells. HEK293T cells were starved before lysis. **(B)** Western blot (IB) and **(C)** quantification of DICER1 pulled down by RFP-AGO2 in cells with and without GRB2 expression. G1 cells were transfected with RFP proteins and either GRB2 or the empty vector (EV), then serum-starved before lysis. DICER1 was normalised to RFP. N = 3. Error bars show standard error of mean. *p < 0.05. ns = not significant by Student's t-test. Long and short refer to exposure time.

To investigate if GRB2 regulated the interaction between AGO2 and DICER1, an RFP-pulldown experiment was utilised. RFP-AGO2, or RFP alone as a control, was transfected into GRB2-depleted G1 cells (Chapter 3, 3.2), along with GRB2, or the empty vector (EV) as a control. Cells were serum-starved before lysates were taken. Lysates were incubated with RFP-trap beads and bound proteins analysed by western blotting. RFP-AGO2 bound more DICER1 in G1 cells overexpressing GRB2 (Figure 5.4.1 B, C). Quantification of the level of DICER1 bound showed an increase of ~40% from G1 cells to GRB2-overexpressing cells (Figure 5.4.1 C). Therefore, presence of GRB2 appeared to promote AGO2 binding DICER1. It should be noted that a band can be observed at the expected molecular weight of GRB2 in the input for G1 cells expressing EV (Figure 5.4.1 B). A different GRB2 antibody did not produce this band (Figure 3.2.2 B); therefore, it appeared to be a non-specific band produced by this GRB2 antibody.

GRB2 may promote the AGO2-DICER1 interaction in two ways; 1) it may bind both AGO2 and DICER1 directly and recruit DICER1 to AGO2, or 2) it may bind AGO2 and induce a conformational change that enhances the ability of AGO2 to form a stable tertiary complex, strengthening the association of DICER1 with AGO2 (Figure 5.4.2 A). While experiments presented here reveal GRB2 has a direct interaction with AGO2, it is not known if GRB2 also binds DICER1 directly. To investigate this, MST was performed using GRB2 and DICER1. Recombinant DICER1^{CT} was expressed and purified from bacterial cells. A truncated DICER1 protein was used as expression and purification of this protein gives much higher yields than the full-length protein. DICER1^{CT} contains both RNase III domains, which bind AGO2, as well as the dsRBD. The protein preparation was incubated with RNase to remove contaminating RNA which

may be bound to DICER1^{CT}. To perform MST experiments, 100 nM labelled DICER1 was mixed with a dilution series of 6.33 to 120,000 nM unlabelled GRB2. A binding curve was not produced, suggesting that the two proteins did not interact (Figure 5.4.2 B). Therefore, it appeared that GRB2 did not directly bind DICER1 to recruit it to AGO2 (Figure 5.4.2 A).

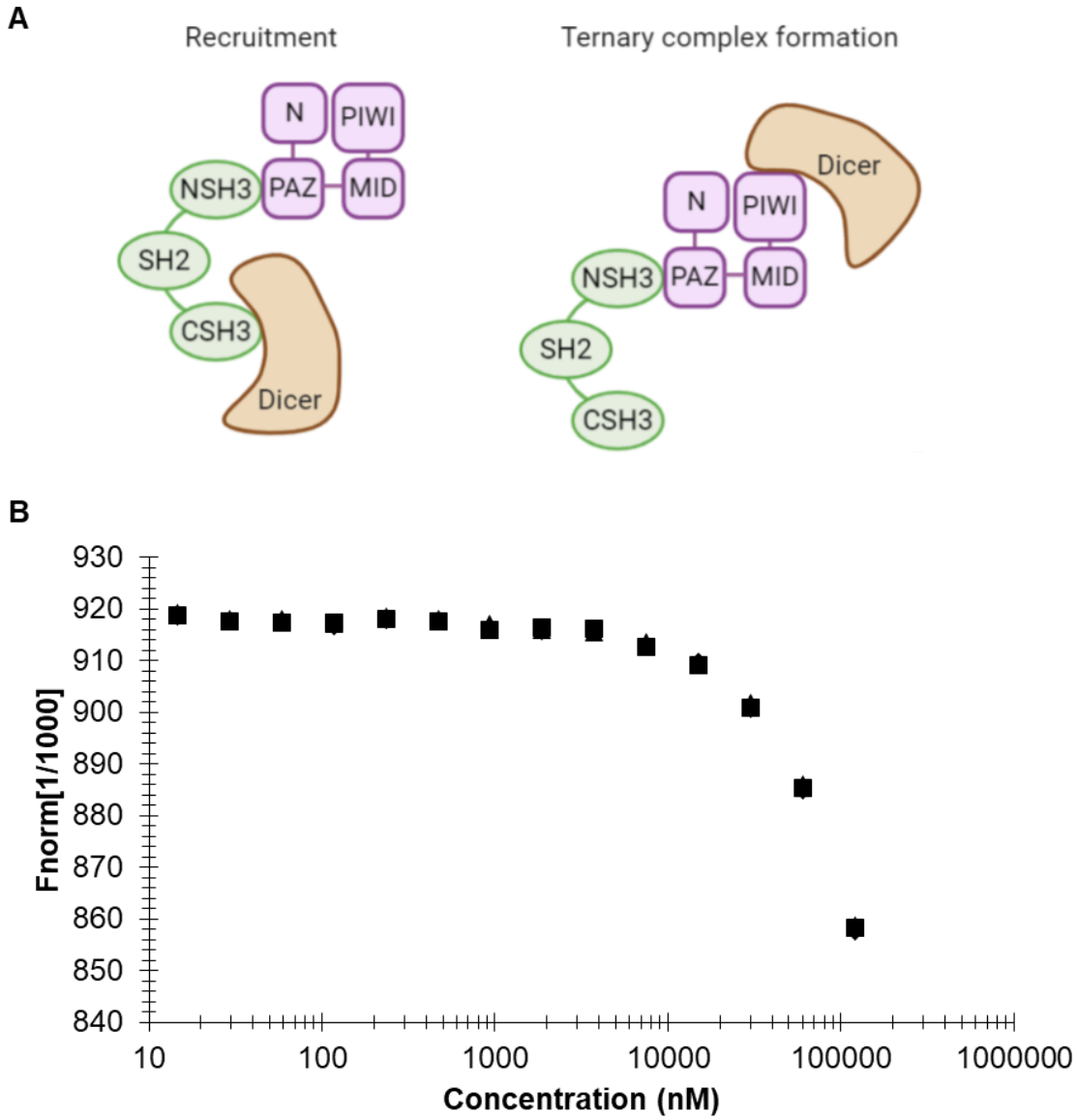


Figure 5.4.2 GRB2 does not bind DICER1 by MST

(A) Schematic of possible mechanisms by which GRB2 may interact with both AGO2 and DICER1. The data reported herein suggest that the likely combination is the ternary complex. Created with BioRender.com. **(B)** A dilution series of 3.66 to 120,000 nM unlabelled GRB2 was set up with 100 nM labelled DICER1^{CT}. Data contributed by Chi-Chuan Lin.

To investigate if complexing of GRB2 and AGO2 promoted AGO2 binding DICER1, the AGO2-DICER1 interaction was analysed by MST. Full-length AGO2 and DICER1^{CT} were expressed in *E. coli*. During purification both proteins were treated with RNase to eliminate RNA. To set up MST experiments, labelled AGO2 was combined with a dilution series of DICER1 from 3.66 to 120,000 nM. MST of full length AGO2 and DICER1 produced a binding curve that could be fitted. A measured dissociation constant of 17.60 ±1.67 μM was produced (Figure 5.4.3, Table 5.4.1). Upon preincubation of AGO2 with increasing concentrations of GRB2, a slight dose increase in affinity of AGO2 for DICER1 was recorded. At 50 μM GRB2, AGO2 bound DICER1 with K_D 4.58 ±0.63 μM (Figure 5.4.3, Table 5.4.1). While it appears that GRB2 only very slightly promotes AGO2 binding DICER1, the change in affinity may indicate that GRB2, AGO2 and DICER1 form a complex with AGO2 bridging the other two proteins (Figure 5.4.2 A). However, this limited change in affinity is unlikely to have physiological relevance.

Table 5.4.1 Measured dissociation constants for DICER1 binding AGO2 with preincubation in increasing concentrations of GRB2

100 nM labelled AGO2 was preincubated with GRB2, before a dilution series of 3.66 to 120,000 nM unlabelled DICER1 was set up and used in microscale thermophoresis with AGO2. Data contributed by Chi-Chuan-Lin.

GRB2 concentration	Affinity
0 μM	17.60 +/-1.67 μM
25 μM	7.86 +/-1.32 μM
50 μM	4.58 +/-0.63 μM

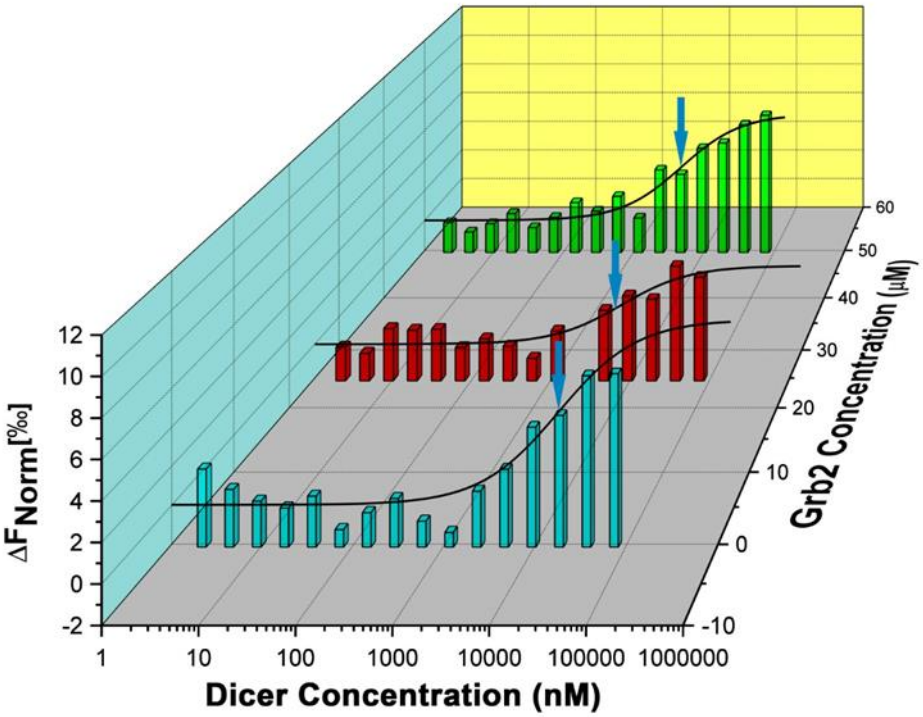


Figure 5.4.3 GRB2 may slightly stabilise AGO2-DICER1 via formation of a tertiary GRB2-AGO2-DICER1 complex

Microscale thermophoresis of DICER1 with AGO2 which has been preincubated with increasing concentrations of GRB2. A dilution series of 3.66 to 120,000 nM unlabelled DICER1 was set up with 100 nM labelled AGO2, which had been preincubated with 0, 25 and 50 μM GRB2. Arrows indicate K_D. Data contributed by Chi-Chuan Lin.

5.5 Discussion

The work reported in this Chapter represents the first demonstration of: 1) direct interaction between GRB2 and AGO2, and 2) direct binding of any protein to AGO2 PAZ domain. GRB2 has previously been suggested to recruit PTP1B to AGO2 (Haidar et al., 2018), although a direct interaction was not shown. Mapping of the GRB2 interactome has revealed association with both miRNA biogenesis factors and mediators of gene silencing, under both starved and EGF-stimulated conditions (Duggal et al., 2019). GRB2 may therefore play a larger role in miRNA biogenesis and function than has previously been identified.

The major function of PAZ domains is to bind small RNAs (Yan et al., 2003). The only reported protein-protein interaction mediated by a PAZ domain is between DICER1 and Flock house virus B2 (FHVB2) protein (Singh et al., 2009). However, this interaction involved FHVB2 protein C-terminus, which has no sequence similarity to that of an SH3 domain. Furthermore, DICER1 PAZ domain does not contain a PxxP motif. AGO2 PAZ domain has been demonstrated to recruit RAD51 to DNA double-strand breaks, although the interaction was not confirmed to be direct (Wang and Goldstein, 2016). RAD51 does not possess an SH3 domain capable of binding ^{323}PHLP^{326}. Originally, it was thought that PAZ-PAZ domain interactions mediated DICER1 binding PAZ-PIWI domain (PPD) proteins, such as AGO2. However it has now been shown binding occurs via the RNase III domain of DICER1 and the PIWI domain of the PPD protein (Tahbaz et al., 2004). In fact, many AGO2 interactors bind this domain, such as the GW182 proteins which recruit effectors of mRNA repression to the miRISC (Behm-Ansmant et al., 2006).

Slightly different binding constants were produced when the affinity of GRB2 for AGO2 was measured in different experiments. Use of AGO2 PAZ domain demonstrated approximately 5-fold higher affinity than that to a peptide spanning ^{323}PHLP^{326} (Figure 5.2.1 D, Figure 5.2.2 D). The change in K_D may be expected due to stabilisation of the interaction by flanking sequences in the full domain. However, the 20-fold weaker binding of PAZ domain to GRB2 by MST is more surprising (Figure 5.2.5 A). Potentially, addition of the MBP-tag stabilised folding of PAZ used in ITC. Alternatively, GRB2 dimer may bind PAZ with a higher affinity. Dimerisation of GRB2 involves a large change in conformation (Schiering et al., 2000; McDonald et al., 2008) and affects binding of NSH3 to SOS (Ahmed et al., 2015).

One limitation of the data presented thus far is the lack of binding demonstrated for endogenous GRB2 to endogenous AGO2. Pulldown of GRB2 was not observed in an IP of endogenous AGO2 (Figure 4.4.2). Antibody binding AGO2 may disrupt the GRB2-AGO2 interaction; the AGO2 antibody used in these experiments binds a sequence between 350 - 450 and therefore adjacent to the ^{323}PHLP^{326} motif. IP of endogenous GRB2, using an antibody which binds GRB2 SH2 or CSH3, and blotting for endogenous AGO2 may be more successful.

Another limitation the work presented here may be that further experiments are required to fully elucidate how a trimeric GRB2-AGO2-DICER1 complex is formed. GST-GRB2 pulldown suggests the three proteins form a complex together (Figure 5.4.1 A). However, it does not clarify which interactions are indirect and what other proteins may be involved in a functional complex. While MST experiments indicate GRB2 does not bind DICER1 RNase III and dsRBD domains (Figure 5.4.3, Table 5.4.1), the helicase, DUF283, platform and PAZ

domains were not tested. Similarly, although MST suggests preincubation of AGO2 with GRB2 may slightly promote DICER1-binding AGO2 (Figure 5.4.3), it does not directly show that stabilisation is through formation of a complex involving all three proteins. Further investigation is required to confirm that the three proteins form a complex with AGO2 bridging GRB2 and DICER1. *In vitro* experiments such as SPR, or in cell techniques, e.g. three molecule FRET, may determine this.

Chapter 6: Specific microRNAs are regulated by GRB2

In the previous Chapter it was demonstrated that GRB2 binds AGO2 directly, via an interaction between GRB2 N-terminal SH3 domain and AGO2 PAZ domain (Figure 5.2.2, Figure 5.3.1). GRB2 was also found to be included in a heterotrimeric complex with AGO2 and DICER1 (Figure 5.4.1), and therefore appeared to be able to integrate into the RLC. However, the interaction with DICER1 did not appear to be direct: instead, it appeared that a complex was formed with AGO2 bridging GRB2 and DICER1.

This Chapter will examine whether the formation of a GRB2-AGO2-DICER1 complex leads to regulation of DICER1-loading miRNAs onto AGO2 by GRB2. It will also investigate the role of this mechanism in cancer. DICER1-mediated processing of pre-miRNA and subsequent loading of mature miRNA onto AGO2 is the last stage in the canonical pathway of miRNA biogenesis (Chapter 1, Figure 1.3.1). Consequently, this process directly produces the active miRISC and must be tightly controlled. miRNAs may act as oncomiRs and tumour suppressor miRNAs. Therefore, dysregulated loading of these miRNAs can regulate cancer pathogenesis.

A well-known example of how loading is perturbed in cancer is through phosphorylation of AGO2 Y393 by EGFR and c-SRC, which inhibits DICER1 loading pro-tumorigenic miRNAs onto AGO2 (Shen et al., 2013; T. Liu et al., 2020). Additionally, ESR2 associates with the RLC to regulate loading of specific miRNAs and may mediate the better prognosis in ESR2+ breast cancers. However, the detailed molecular mechanism for this was not elucidated (Tarallo et al., 2017). Finally, in response to DNA damage, TP53 binds AGO2 to modulate its association with miRNA, but this mechanism is

thought to be uncoupled from classical loading as it does not affect overall miRNA biogenesis. TP53 mutants differentially regulate binding of tumour suppressor miRNAs to AGO2 (Krell et al., 2016). In this Chapter, it is proposed that GRB2 specifically regulates loading of let-7 family miRNAs and additionally controls expression of the mir-221 and mir-17~92 clusters by a different mechanism. The potential for GRB2-regulation of miRNA to contribute to tumorigenesis is explored.

6.1 GRB2 regulates two groups of miRNAs by different mechanisms

Direct binding of GRB2 to AGO2 and complexing with DICER1 suggested that GRB2 may moderate miRNA-loading (Figure 5.4.1 A). To determine if the presence of GRB2 affected miRNA expression, small-RNA sequencing was performed in G1 (Chapter 3, 3.2) and parental HEK293T cells. Two G1 clones were used to distinguish GRB2-dependent changes in miRNA expression and random effects of CRISPR/Cas9 mutagenesis; of note, while clone 1 had complete knockout of GRB2, clone 2 had low expression of a NSH3-mutated GRB2 protein. G1 and HEK293T cells were serum-starved and total RNA was purified. Libraries were prepared, using size selection for RNA ~20 nt, and sequenced, before reads were aligned to the human genome and differential expression analysed (Chapter 2, 2.14). Three comparisons of differentially expressed miRNAs were then performed: 1) clone 1 vs HEK293T, 2) clone 2 vs HEK293T, and, 3) clone 1 vs clone 2. In the third comparison, miRNAs which were significantly dysregulated in the same direction were retained. Thus, miRNAs were identified which were deregulated in both clones compared to

HEK293T, but not to each other. A cohort of miRNAs were identified with both enhanced and diminished expression (Figure 6.1.1).

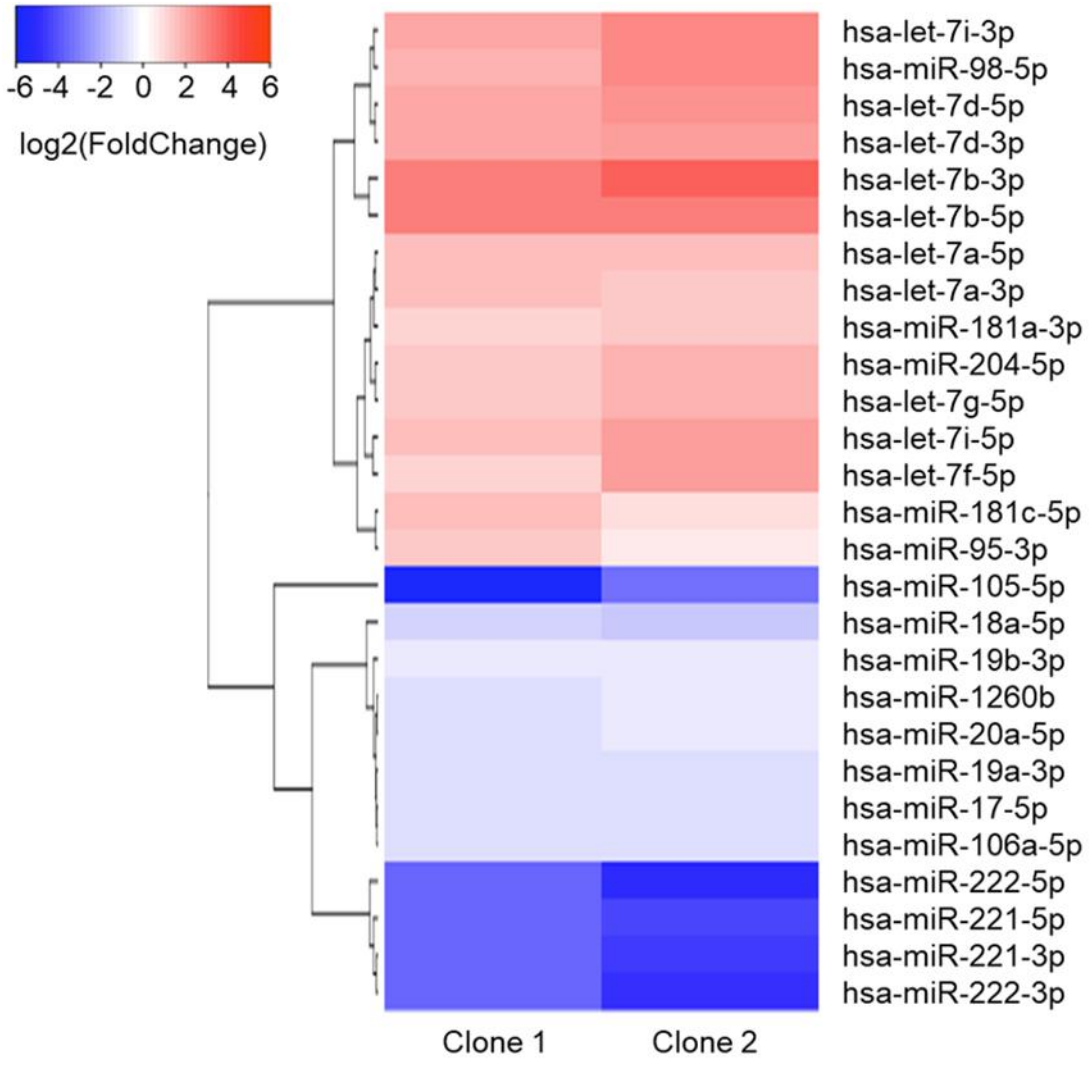


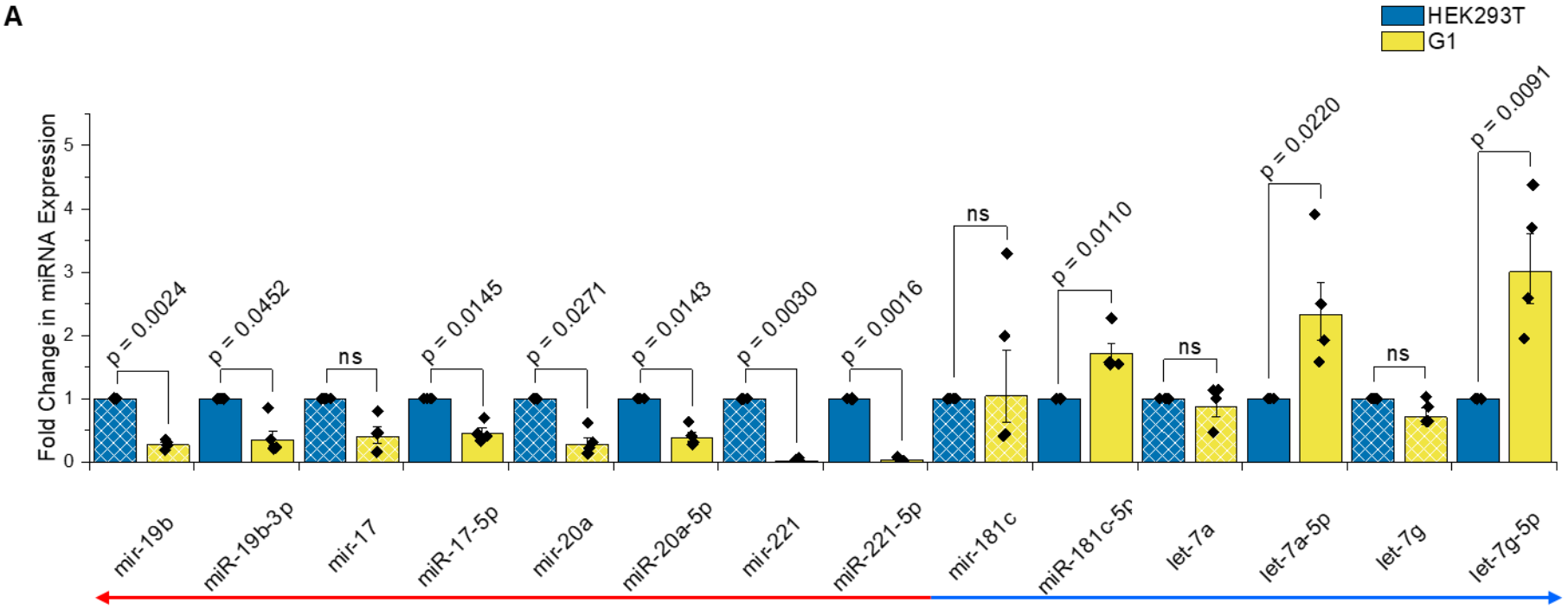
Figure 6.1.1 miRNAs are up and downregulated in G1 cells

Small RNA sequencing of two G1 clones revealed increased (blue) and decreased (red) miRNA expression. N = 2. $p < 0.05$ (Benjamini-Hochberg-adjusted). Library preparation and sequencing was performed by Ragini Medhi (University of Cambridge, UK). Sequencing analysis was done by Dapeng Wang (University of Leeds, UK) and heat map generation by Kin Man-Suen (University of Leeds, UK).

It was immediately apparent that miRNAs of the same family were deregulated in a similar manner. Members of the mir-17 (miR-18a-3p, miR-20a-5p, miR-17-5p, miR-106a-5p); mir-19 (miR-19a-3p, miR-19b-3p); the mir-221 (miR-222-3p, miR-222-5p, miR-221-3p, miR-221-5p) families were all significantly diminished, whereas the let-7 (let-7i-3p, let-7d-5p, let-7d-3p, let-7b-3p, let-7b-5p, miR-98-5p, let-7i-5p, let-7f-5p, let-7a-5p, let-7a-3p and let-7g-5p) and the mir-181 (miR-181a-3p, miR-181c-5p) families were significantly enhanced in the absence of GRB2 (Figure 6.1.1, $p < 0.05$). The most strongly downregulated were the mir-221 family. For example, expression of miR-221-5p was reduced 12.7- and 22.6-fold in clones 1 and 2 respectively. In contrast, let-7b-5p and 3p had the greatest increase in expression, with levels of let-7b-5p enhanced 8.7- and 8.9-fold in each clone.

As GRB2 both positively and negatively regulated miRNA expression it was unlikely that expression of both groups of miRNAs were controlled by the same mechanism. To investigate if any miRNAs were regulated by GRB2 at the stage of miRNA-loading onto AGO2, expression of both precursor and mature miRNAs was measured. Parental HEK293T and G1 cells were serum-starved, before total RNA was isolated. As similar dysregulation of miRNA expression was seen in each of the clones, both were used in this and all following experiments. Expression of several GRB2-regulated miRNAs, and their relative precursors, were then quantified by qPCR. The miRNAs downregulated in G1 (mDGs): miR-19b-3p, miR-17-5p, miR-20a-5p, miR-221-5p; and the miRNAs upregulated in G1 (mUGs): miR-181c-5p, let-7a-5p and let-7g-5p were selected. For all miRNAs, mature transcripts were significantly differentially expressed in a manner that matched results from the small RNA sequencing (Figure 6.1.2 A, $p < 0.05$). It should be noted that these p values were not adjusted to account

for multiple significance testing as only miRNAs which had already been identified as dysregulated in HEK293T cells were analysed (Figure 6.1.1).



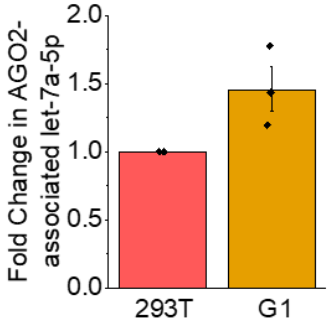


Figure 6.1.2 Upregulation of let-7 family miRNAs occurs via loading whereas miRNA downregulation happens prior to this

(A) Fold change in expression of all precursor miRNA (hatched, precursor and primary) and mature miRNA (plain) in G1 cells (yellow) compared to HEK293T cells (blue). Precursor miRNAs are produced prior to loading, whereas mature miRNAs have been loaded onto AGO2. Red and blue arrows indicate miRNAs which were downregulated and upregulated respectively in GRB2-depleted HEK293T by small-RNA seq. Total cell RNA was extracted from serum-starved cells then miRNA expression was quantified by qPCR against RNU6-6P control. N = 4. **(B)** Fold change in AGO2-associated let-7a-5p in G1 and HEK293T cells. AGO2 was immunoprecipitated from serum-starved cells. RNA was extracted from immunoprecipitates and miRNA quantified by qPCR against GAPDH control. N = 3. Error bars indicate standard error of mean. p values are shown if $p < 0.05$. ns = not significant, by Student's t-test.

Like their mature counterparts, mDG precursor transcripts were consistently diminished in G1 cells. The decrease in precursor expression was significant for all except pre-mir-17 due to variation in the measured levels of this transcript (Figure 6.1.2 A, $p < 0.05$). It is possible that the lack of significance for pre-mir-17 was due to less accurate measurement of precursor miRNAs, which are expressed at lower levels than mature miRNA. For all miRNAs, the reduction in pre-mir was similar or greater than for mature. For example, pre-mir-221 was reduced 53.2-fold whereas miR-221-5p was downregulated 31.1-fold. Both primary and precursor miRNA transcripts were bound by the ‘precursor’ primer (Chapter 1, Figure 1.3.1). Consequently, it was likely that both the pri- and pre-miRNA were diminished and that regulation occurred before DROSHA-mediated processing of pri-miRNA.

In contrast to the mDG, for all tested mUG, precursor transcripts were not significantly deregulated (Figure 6.1.2 A, $p < 0.05$). Pre-mir-181c, pre-let-7a-1 and pre-let-7g displayed fold changes of 1.0, 0.9 and 0.7 respectively in G1 cells compared to HEK293T. However, it should be noted that pre-mir-181c showed fluctuations in the measured levels of this transcript. miR-181c-5p is expressed at much lower levels than the let-7 miRNAs, so measurements of its precursor are less likely to be accurate. An increase in mature but not precursor miRNA may have indicated changes in loading or turnover.

To determine if increased cellular mUG expression correlated with an increase in the level of miRNA which was loaded onto AGO2, expression of AGO2-associated let-7 was quantified. AGO2 was immunoprecipitated from serum-starved HEK293T and G1 cells. Co-precipitated RNA was extracted and pulldown of mature let-7a-5p was quantified by qPCR. AGO2-associated let-7a-5p was consistently increased upon GRB2 depletion (1.5-fold, Figure 6.1.2 B).

However, high variation in the measured levels of this transcript meant it did not reach significance. It was likely that further optimising of the washing steps was required to ensure more consistent pulldown of RNA. Overall, these results suggested that GRB2 inhibited loading of let-7 family miRNAs onto AGO2.

Expression levels of precursor and mature transcripts suggested that GRB2 regulated miRNA biogenesis by two mechanisms: 1) GRB2 hindered loading of let-7 onto AGO2, and 2) GRB2 promoted expression of mir-17, mir-19 and mir-221 prior to DROSHA-mediated pri-miRNA processing.

Of the entire let-7 family, only group II let-7 miRNAs were significantly overexpressed in both G1 cell lines (Figure 6.1.1). Group II pre-let-7 require 3' mono-uridylation by a terminal uridylyl transferase (TUT4, TUT7 or TUT2) in order to be bound and processed by DICER1 (Heo et al., 2012). Therefore, it could be argued that increased miRNA loading onto AGO2 is likely to be due to more efficient TUT processing, rather than through GRB2 binding AGO2. To determine if TUT is dysregulated in G1 cells, TUT4 expression was blotted in both G1 and HEK293T cells. TUT4 was selected as it has the highest expression in HEK293T cells (Uhlén et al., 2015; Protein Atlas, 2020). As with previous experiments, cells were serum-starved before lysis. TUT4 expression was unchanged between G1 and HEK293T cells (Figure 6.1.3). Therefore, it is unlikely that GRB2 regulates TUT protein levels to mediate changes in miRNA-loading.

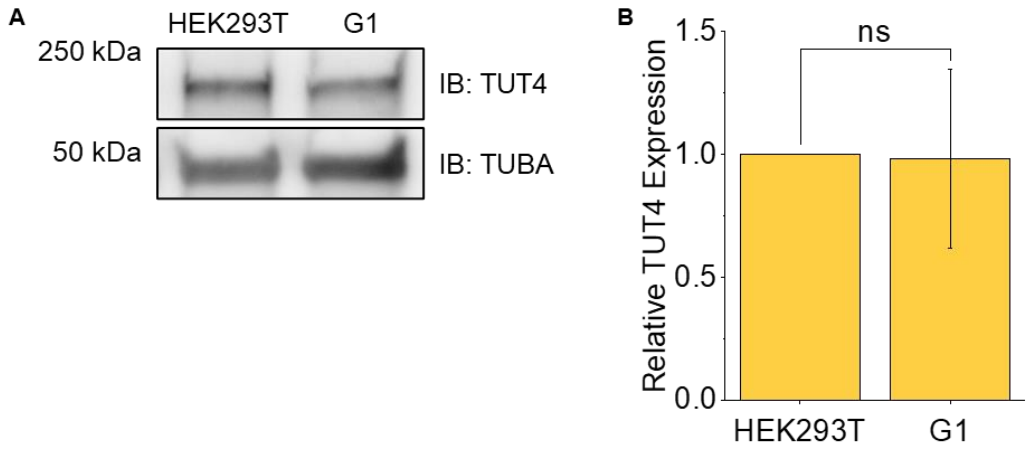


Figure 6.1.3 TUT4 expression is not dysregulated in G1 cells

(A) Western blot and **(B)** quantification of TUT4 expression in HEK293T cells. Cells were deprived of growth factor before lysis. Quantification of TUT4 was relative to α-Tubulin (TUBA). N = 3. Error bars show standard error of mean. ns = not significant at $p < 0.05$, Student's t-test.

Parental HEK293T cells were used as the control for miRNA expression in G1 cells instead of Scr cells (generated by CRISPR/Cas9 modification of HEK293T cells using a gRNA with a scrambled sequence, see Chapter 3, 3.1). HEK293T cells were selected as the Scr cell genome was modified by random insertion of a puromycin resistance gene. Consequently, expression of some miRNAs may have been altered in Scr cells, which would result as their being recorded as differentially regulated in G1 cells. However, the parental HEK293T cells express neither GFP, nor N-acetyl transferase which confers puromycin resistance. To confirm that the measured changes in miRNA were not mediated by either protein, miRNA expression in Scr cells was compared to HEK293T cells. Cells were serum-starved, before total cell RNA was isolated and miRNA expression measured by qPCR. Expression of one mUG, let-7a-5p, and one mDG, miR-221-5p, was examined. Neither miRNA was significantly or consistently differentially expressed in Scr compared to HEK293T (Figure 6.1.4). let-7a-5p and miR-221-5p displayed mean fold changes of 1.0 and 1.1 respectively. Therefore, these data demonstrate that the changes in miRNA expression seen in G1 cells were not an effect of puromycin treatment nor of GFP expression.

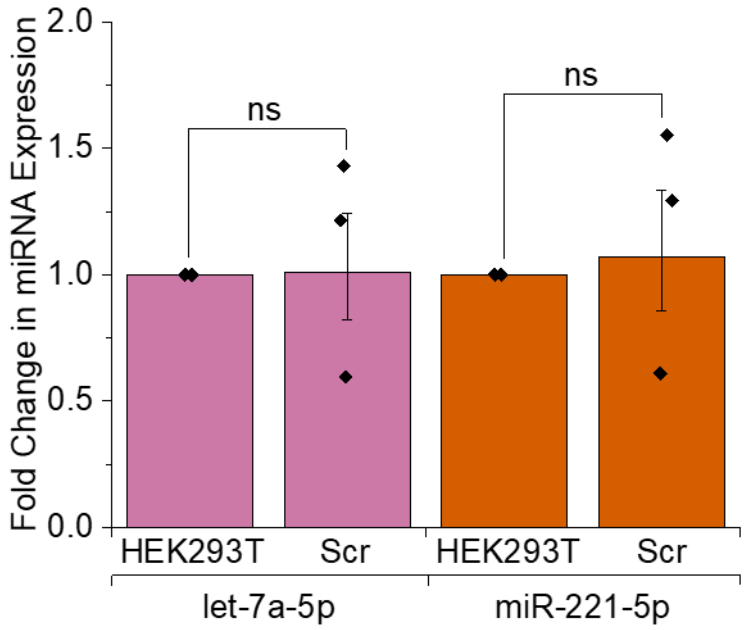


Figure 6.1.4 let-7a-5p and miR-221-5p expression is unchanged between HEK293T and Scr control cells

Cells were serum-starved and then RNA was extracted. Expression of let-7a-5p and miR-221-5p was quantified by qPCR. Fold change in miRNA expression from HEK293T to Scr cells was calculated. Error bars indicate standard error of mean. N = 3. ns = not significant at $p < 0.05$, by Student's t-test.

6.2 GRB2 moderation of let-7 controls expression of the miRNA targets

Increased let-7 loading onto AGO2 in the absence of GRB2 suggested augmented activity of these miRNAs. Consequently, cells depleted of GRB2 should have reduced expression of let-7 family target oncogenes. To test this, the expression of let-7 targets was quantified alongside let-7g-5p. HEK293T and G1 cells were serum-starved and total cell RNA extracted. RNA expression was then quantified by qPCR. As before, let-7g-5p expression was enhanced 3.0-fold in G1 cells compared to HEK293T (Figure 6.1.2 A, Figure 6.2.1 A). Expression of *NRAS*, *LIN28B, HMGA1* and *HMGA2* mRNAs were significantly decreased 1.4-, 5.0-, 1.9- and 6.1- fold respectively in the G1 cells (Figure 6.2.1 A, $p < 0.05$) (Bos et al., 1985; Tamimi et al., 1993; Cooper, 1996; Guo et al., 2006). As only these validated let-7 targets were tested it is possible that all could be dysregulated as a result of let-7 overexpression, so p values were not adjusted to account for multiple comparisons.

Protein expression of the let-7 targets was also examined. HEK293T and G1 cells were deprived of FBS, then cells were lysed and protein levels were examined by western blotting. LIN28B and HMGA2 protein expression were significantly diminished 4.6- and 6.5-fold in G1 cells (Figure 6.2.1 B, C; $p < 0.05$). Therefore, GRB2 appeared to inversely regulate expression of let-7 miRNAs and their target mRNAs.

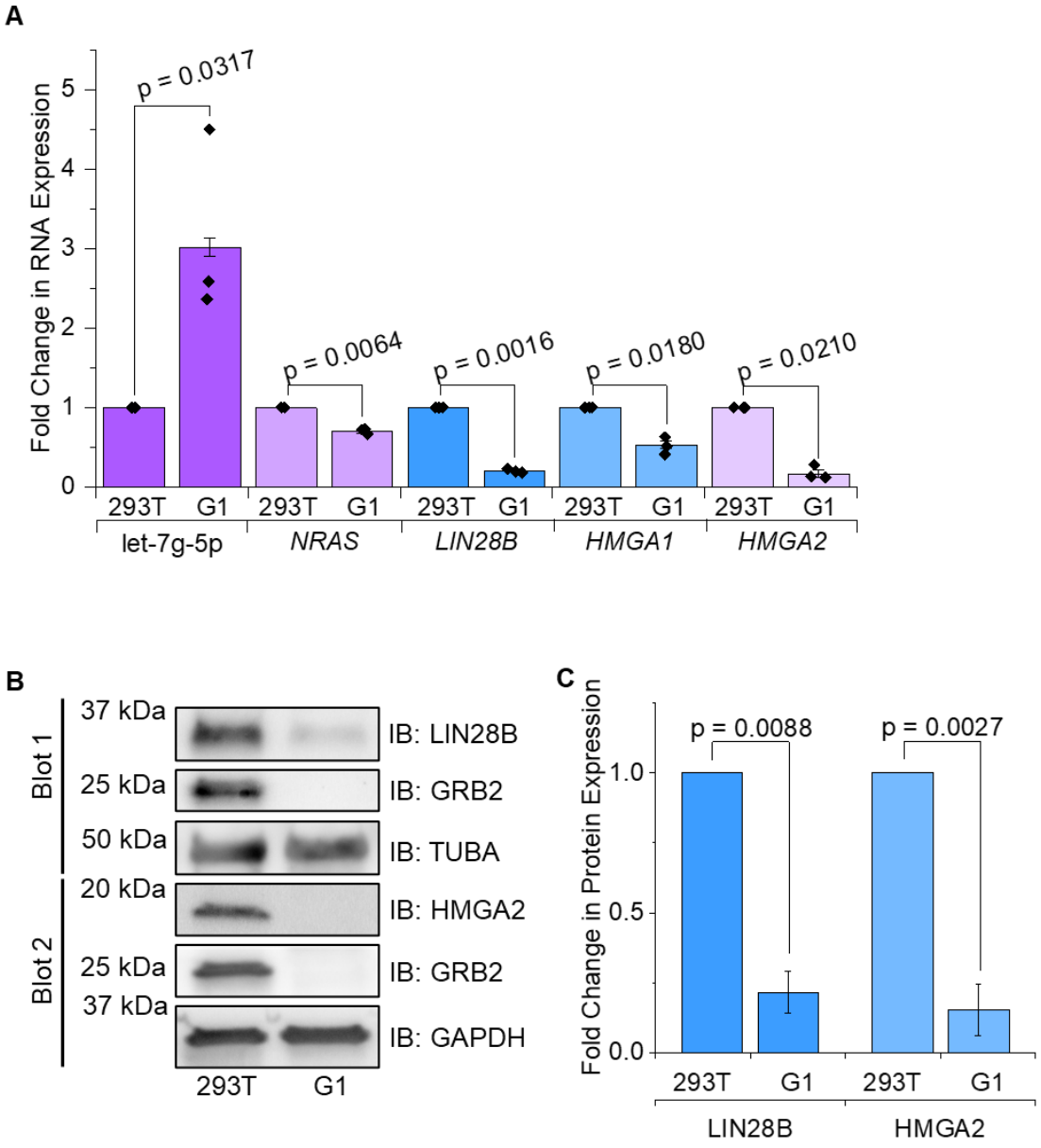


Figure 6.2.1 Upregulation of let-7 miRNA correlates with downregulation of let-7 target oncogenes in G1 cells

(A) Fold change in let-7 miRNA and target mRNAs in HEK293T and G1 cells. Total cell RNA was extracted from serum-starved cells and all RNAs quantified by qPCR. let-7g-5p was normalised against RNU6-6P whereas mRNAs were relative to GAPDH. N = 3. **(B)** Western blot (IB) and **(C)** quantification of let-7 target protein expression in HEK293T and G1 cells. Expression was normalised against the housekeeping gene GAPDH or α-Tubulin (TUBA). N = 3. Error bars indicate standard error of mean (SEM). p values are shown if $p < 0.05$, by Student's t-test.

To determine if the reduced expression of let-7 target genes in G1 was facilitated by increased let-7 activity, a rescue experiment was performed. If high let-7 silenced these genes, sponge-mediated inhibition of let-7 would release suppression. However, if another mechanism dictated expression of these targets in G1, then their expression would remain unchanged. G1 and HEK293T cells were transfected with a let-7 sponge or the empty vector as a negative control. The cells were serum-starved overnight before total cell RNA was extracted and quantified by qPCR. In both cell lines, expression of the sponge resulted in a slight increase of all three targets (Figure 6.2.2). *LIN28B* and *HMGA2* and were significantly increased in G1 cells but not HEK293T, with fold changes of 1.2 and 1.7. Conversely, *DICER1* showed significant upregulation in only the parental HEK293T cells (1.3-fold, $p < 0.05$). As before, only these three genes were assessed and raw p values were used. The small changes in gene expression in both cell lines suggested that the experiment needed optimising to completely rescue expression. However, as *LIN28B*, *HMGA2* and *DICER1* expression in G1 was consistently rescued by the sponge, it is likely that their downregulation in G1 was a result of let-7 activity.

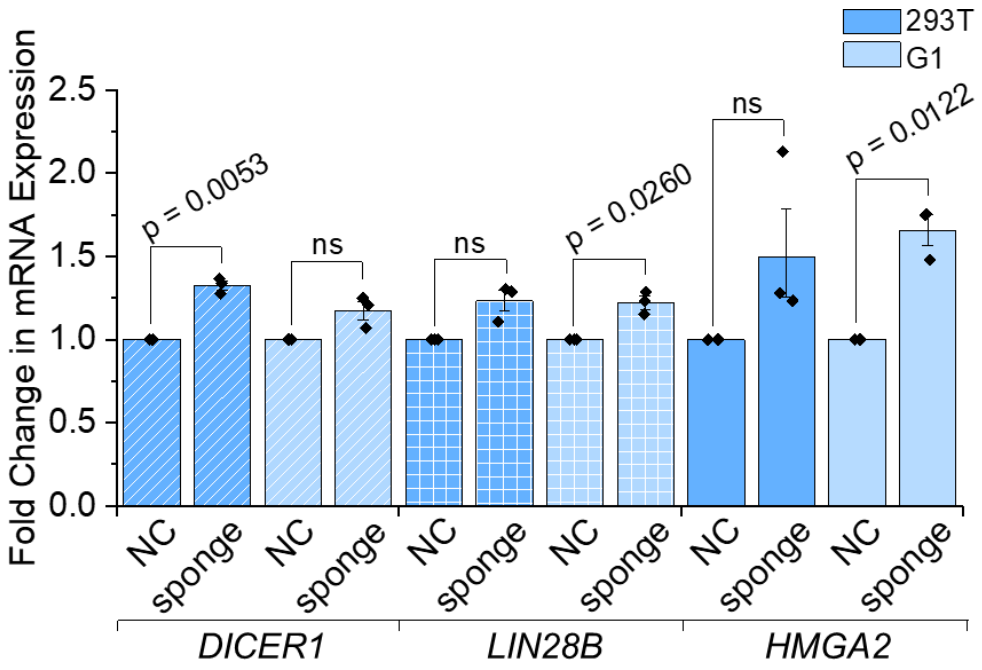


Figure 6.2.2 Transfection of a let-7 sponge rescues target expression in G1 cells

Fold change in mRNA expression in HEK293T and G1 cells expressing a let-7 sponge. Cells were transfected with a let-7 sponge or the empty vector as a negative control (NC) and serum-starved. Total cell RNA was extracted and all RNAs quantified by qPCR with normalisation to GAPDH. N = 3. Error bars indicate standard error of mean. p values are shown if $p < 0.05$. ns = not significant by Student's t-test.

6.3 Investigation of miRNA expression in colorectal cancer patient samples

Many of the mDGs are may function as oncomirs (i.e. miRNAs whose modified expression leads to oncogenesis), including the mir-17, mir-19 and mir-221 families (Ciafrè et al., 2005; Dews et al., 2006; Hayashita et al., 2005; H. He et al., 2005a; L. He et al., 2005b; E. J. Lee et al., 2007; Sylvestre et al., 2007). Conversely, the mUGs, predominantly the let-7 family, are normally reported as tumour suppressors (Takamizawa et al., 2004; Johnson et al., 2005; Akao et al., 2006; Mayr et al., 2007; Shell et al., 2007; Yu et al., 2007). Therefore, GRB2 may mediate a pro-tumorigenic miRNA profile in human cells.

To test this observation, expression data for the mDGs and mUGs in human cancers was extracted using the online database OncomiR (Wong et al., 2017). miRNAs were selected with log2(fold change in mean expression) = 0.5-25.0 and displayed in a heatmap with agglomerative hierarchical clustering. mDGs were more likely to be overexpressed (Figure 6.3.1 A) whereas mUGs were more likely to be downregulated (Figure 6.3.1 B). This was particularly evident for cancers derived from epithelial cells, such as colon, rectal, stomach and, to a lesser extent, lung adenocarcinomas. For example, colon adenocarcinoma showed enhanced expression of mDGs: miR-17-5p, miR-20a-5p, miR-19a-3p, miR-19b-3p, miR-18a-5p, miR-106a-5p, miR-222-3p, miR-222-5p and miR-221-5p. Equally, expression of mUGs was reduced: miR-181-3p, let-7i-5p, miR-204-5p, let-7g-5p, let-7f-5p, miR-98-5p, miR-181c-5p, let-7d-5p, let-7a-5p and let-7b-5p. These data demonstrated that GRB2 may facilitate expression of a pro-tumorigenic miRNA signature.

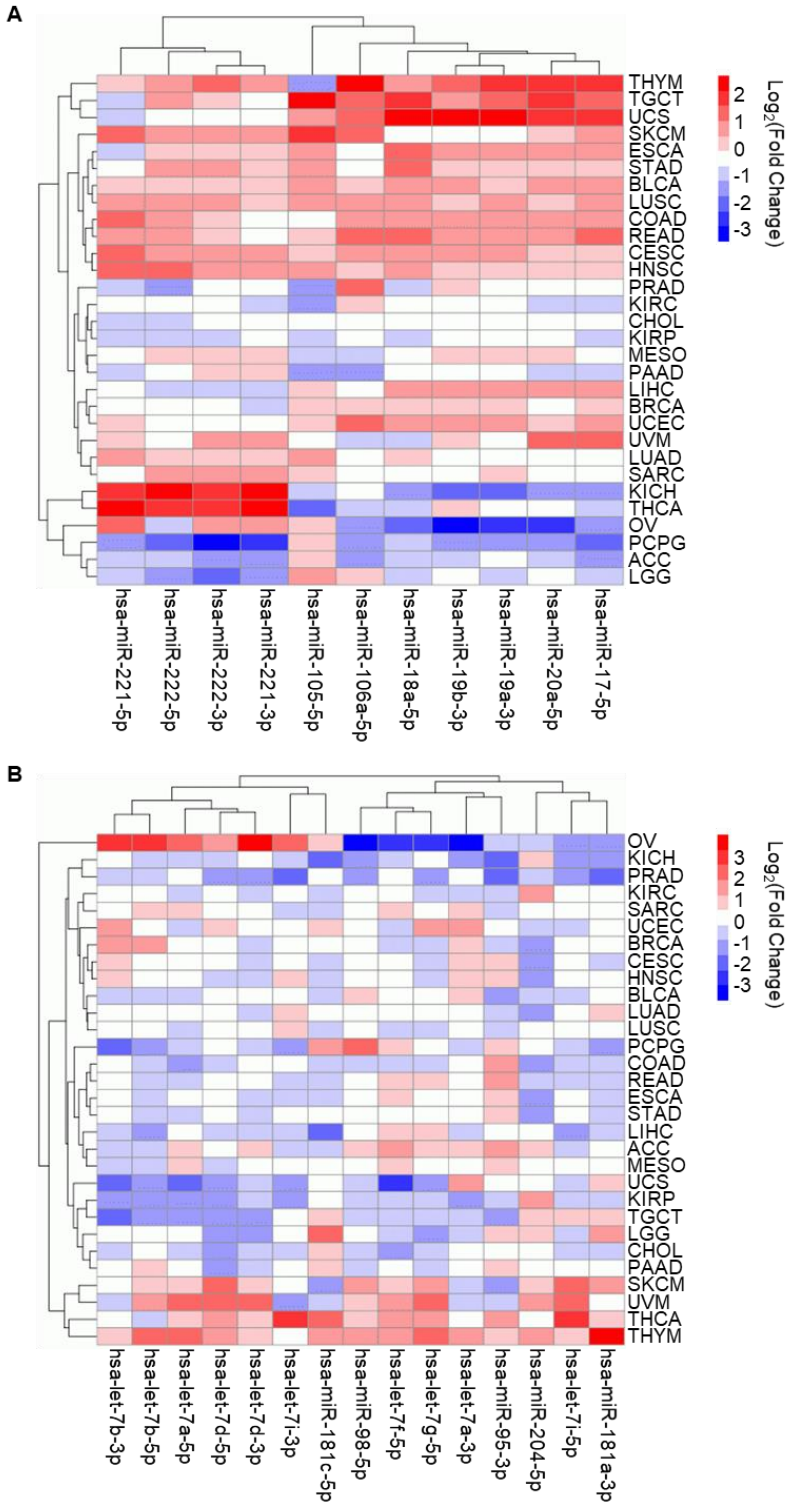


Figure 6.3.1 Differential expression of GRB2-regulated miRNAs in cancer

Heatmap **(A)** mDG and **(B)** mUG expression in various human cancers. Hierarchical clustering was completed using OncomiR (Wong et al., 2017). miRNAs were filtered for $\log_2$(mean expression) = 0.5-25.0. ACC: adrenocortical carcinoma; BLCA: bladder urothelial carcinoma; BRCA: breast invasive carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL: cholangiocarcinoma; COAD: colon adenocarcinoma; ESCA: esophageal carcinoma; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LGG: brain lower grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; MESO: mesothelioma; OV: ovarian serous cystadenocarcinoma; PAAD: pancreatic adenocarcinoma; PCPG: pheochromocytoma and paraganglioma; PRAD: prostate adenocarcinoma; READ: rectal adenocarcinoma; SARC: sarcoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UCS: uterine carcinosarcoma; UVM: uveal melanoma.

To explore if GRB2 did indeed regulate miRNA expression in colorectal cancer, a detailed analysis of patient samples was completed. Normalised expression data from the colorectal carcinoma cohort in The Cancer Genome Atlas (TCGA) was analysed (TCGA_Research_Network, 2020). The cohort contained 448 primary tumour samples with miRNA, mRNA and clinal data, eight of which also had data from matched normal tissue. *GRB2* expression (fragments per kilobase of transcript per million mapped reads upper quartile, FPKM-UQ) was correlated to miRNA level (reads per million) throughout the group of primary tumour samples. The Bonferroni correction was used to account for multiple significance testing. Significant negative correlations were found between let-7d, miR-106a, miR-181a-2, miR-19a, miR-19b (-1 and -2), and miR-95 with *GRB2* (Figure 6.3.2 A, C-G; $p < 0.05$). Conversely, let-7f (-1 and -2) was the only miRNA which was positively associated with *GRB2* (Figure 6.3.2 B).

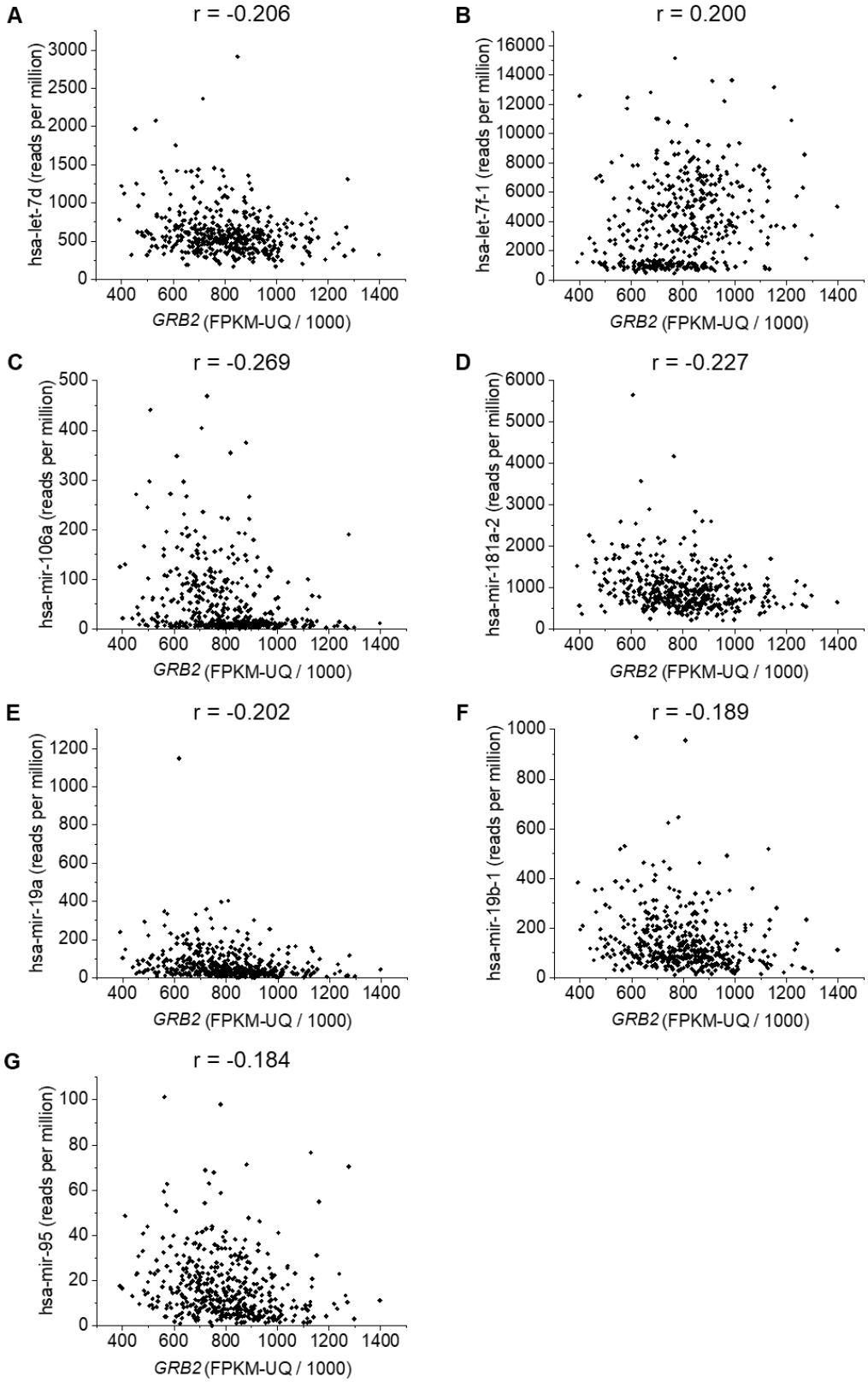


Figure 6.3.2 Associations of GRB2-regulated miRNAs with *GRB2* in colorectal cancer patients

Scatter diagrams showing significant correlations between expression of *GRB2* and indicated miRNA in primary colorectal tumour samples taken from The Cancer Genome Atlas. Pearson correlation coefficients (r) were calculated and considered significant at $p < 0.05$ with Bonferroni-adjustment. N = 448.

Filtering the data according to prior radiation and chemotherapy, as well as for tumour subtype, overall American Joint Committee for Cancer (AJCC) stage and lymph node (N) staging did not reveal any other correlations to *GRB2*. When only primary tumour samples at metastasis stage M0 were considered, miR-17, miR-18a and miR-181c showed significant negative associations with *GRB2*, in addition to all those reported for the entire primary tumour cohort (N = 323, data not shown). The M1 sample size was much smaller (N = 63), which may explain the lack of significant correlations in this group. miR-181c was also negatively correlated with *GRB2* in primary tumours at tumour stage T2, along with miR-181a-2 (N = 74). Selection for T3 tumours demonstrated negative correlations between both let-7g and let-7d and *GRB2*, as well the previously described associations with let-7f (1 and 2), miR-106a, miR-181a-2 and miR-19a (N = 304, data not shown). Again, sample sizes for T1 (N = 12) and T4 (N = 56) were smaller.

Overall, these data suggested that some miRNAs in this cohort of colorectal cancer patients may be associated with *GRB2*. However, they generally do not seem to be regulated by GRB2 in the same manner as in HEK293T cells. Two exceptions may be let-7d and let-7g.

It is possible that specific regulation of miRNA by GRB2 occurs only when GRB2 is particularly enhanced or diminished. Consequently, patient samples were sorted according to *GRB2* level and miRNA expression in the top and bottom 10% compared. let-7d, miR-106a and miR-181a-2 were significantly differentially expressed between the two groups (Figure 6.3.3, $p < 0.05$). All were reduced in the '*GRB2* high' samples and therefore matched the correlations found for all primary tumour samples (Figure 6.3.2).

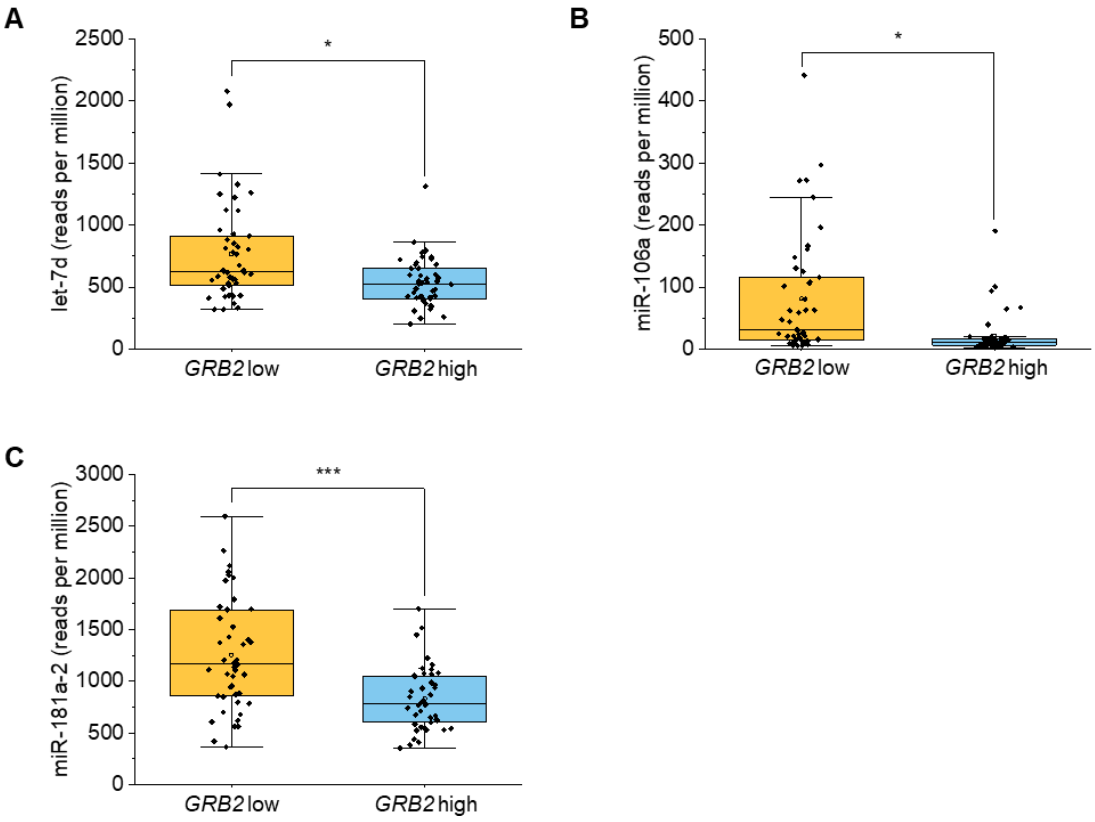


Figure 6.3.3 Expression of GRB2-regulated miRNAs in colorectal cancer patients with high and low *GRB2*

Colorectal primary tumour samples with the highest and lowest *GRB2* expression were compared in box plots. Data from The Human Cancer Genome Atlas. N = 45. *p < 0.05, ***p < 0.001, by Student's t-test (Bonferroni-corrected).

Overall, correlations between *GRB2* and miRNA in primary colorectal tumours did not reproduce those seen in G1 cells. However, GRB2 regulation of miRNA may occur early in the tumorigenesis process before other mechanisms take over. Thus, changes in GRB2 and miRNA may be observed between normal and tumour tissue, but not across tissue samples. Expression of *GRB2* and the miRNAs was compared in eight matched tumour and normal tissue samples. *GRB2* was significantly diminished in tumour cells (Figure 6.3.4 A, $p < 0.05$). let-7a (-1, -2 and -3), let-7b and let-7d were all decreased, while a fourth let-7 family member, miR-98, was enhanced (Figure 6.3.4 B, C, D, H). miR-181a-2 also had reduced expression in tumour tissue (Figure 6.3.4 E). Finally, miR-19a and miR-19b (-1 and -2) were upregulated (Figure 6.3.4 F, G). miR-19a was not detected in any of the normal samples but had low expression in tumour tissue.

The reduction in both *GRB2* and mUG (let-7a, let-7b, let-7d, miR-181a-2) from normal to tumour tissue suggests that GRB2 did not inhibit their expression in the early stages of colorectal tumorigenesis. Equally, the increase in mDG (miR-19a and miR-19b) indicates that GRB2 did not enhance their expression in this system. However, it is possible that early events leading to tumorigenesis were masked by continued development of the cancers.

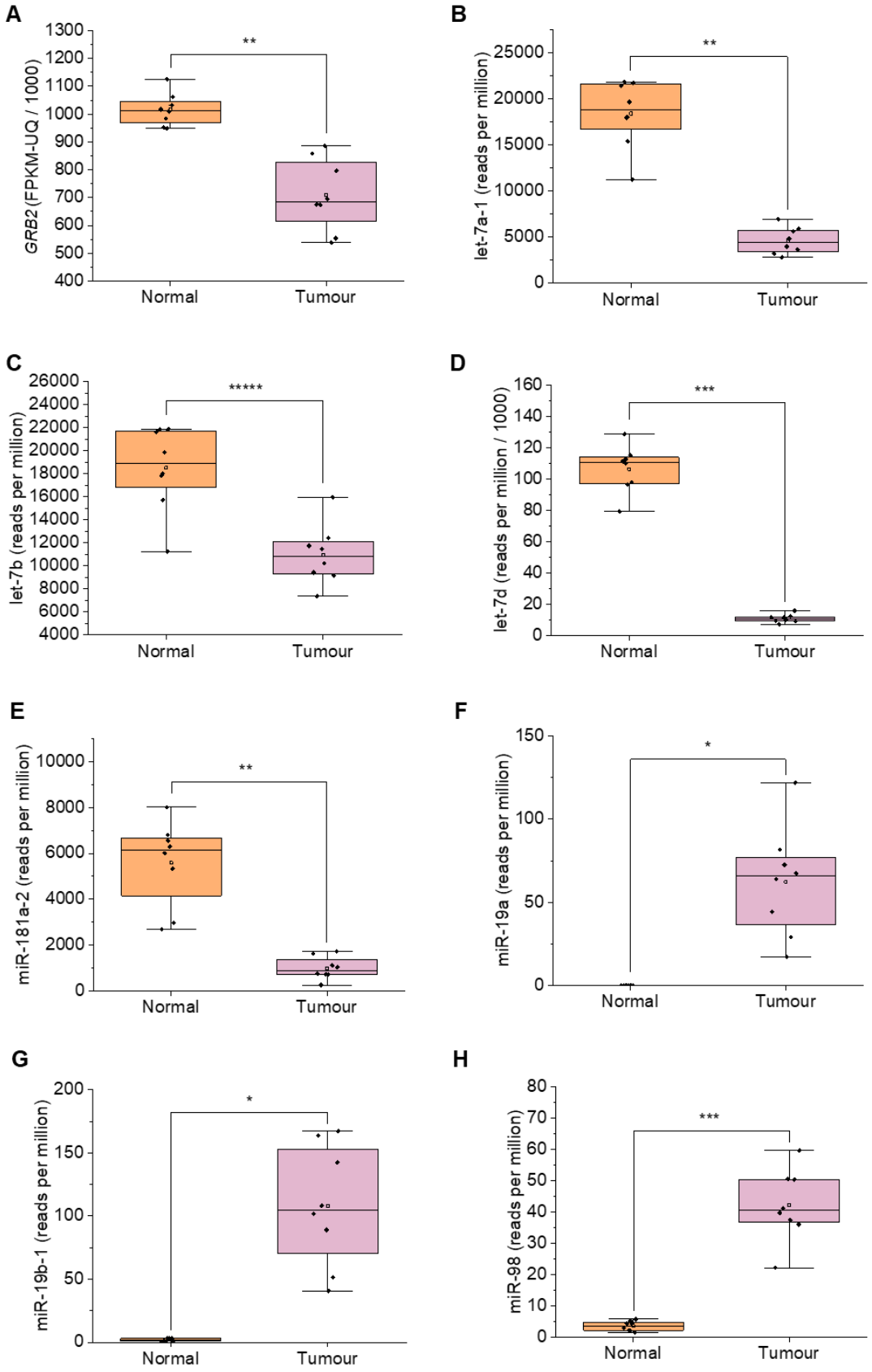


Figure 6.3.4 Expression of *GRB2* and the GRB2-regulated miRNAs in matched normal and colorectal cancer tissue

Box plots showing RNA expression in matched normal and primary colorectal tumour samples. Data is from The Cancer Genome Atlas. N = 8. *p < 0.05, **p < 0.01, ***p < 0.001, *****p < 0.00001 by Student's t-test (Bonferroni-adjusted).

6.4 Screen for GRB2-dependent miRNA expression in cancer cell lines

Up to this point, it appeared that the GRB2-regulated miRNAs may form a pro-tumorigenic signature which could be induced by high GRB2 (Figure 6.3.1). However, analysis of a cohort of colorectal cancer patient samples from the TCGA did not reveal the expected miRNA profile when correlated with *GRB2* expression. In an attempt to find cancers which mimicked dysregulation of miRNAs seen in G1 cells, a screen was completed. GRB2 was knocked down in four cancer cell lines and miRNA expression was examined. Stable cell lines were generated by CRISPR/Cas9-mediated genome editing. Cells were transduced with lentivirus expressing Cas9 and gRNAs targeted to GRB2 (Chapter 2, 2.6.5) and then selected with puromycin. Successful GRB2 knockdown was obtained in H520 lung squamous cell carcinoma, A498 kidney carcinoma, MDA-MB-231 breast adenocarcinoma and Caco2 colorectal adenocarcinoma cells (Figure 6.4.1). All knockdown cell lines expressed varying levels of GRB2 as determined by western blot. Two A498 knockdown cell lines were produced, with one showing a greater reduction in GRB2 (hereon in referred to as kd 1).

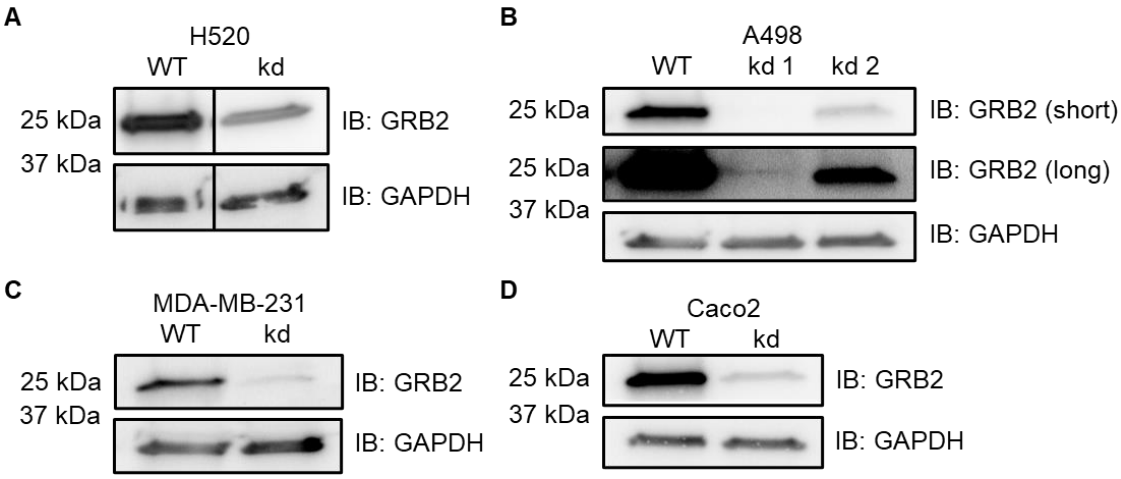


Figure 6.4.1 Stable knockdown of GRB2 in various cancer cell lines

Western blots (IB) showing GRB2 expression in **(A)** H520, **(B)** A498, **(C)** MDA-MB-231 and **(D)** Caco2 wild type (WT) and knockdown (kd) cell lines. Parental cells were transduced with lentivirus expressing Cas9 and gRNAs targeting GRB2. GRB2 kd cells were selected with puromycin.

To determine if the miRNA expression profiles of any GRB2 kd cancer cell lines reflected that of G1 cells, mUG and mDG levels were examined. Parental and knockdown cells were serum-starved and then total RNA was isolated. miRNA expression was quantified by qPCR against RNU6-6P. Importantly, only two repeats were completed for Caco2, H520 and A498 cells and therefore the data presented for these cell lines represent preliminary results only. For all cell lines, miRNAs from the same family were similarly dysregulated (Figure 6.4.2). H520 kd cells showed a general decrease in expression of the let-7, mir-19 and mir-17 families: for example, let-7a-5p was diminished 2.1-fold, miR-19-3p was reduced 1.6-fold and miR-17-5p was suppressed 1.7- fold (Figure 6.4.2 A). Expression of miR-181c-5p and miR-221-5p were similar in parental and knockdown H520 cells.

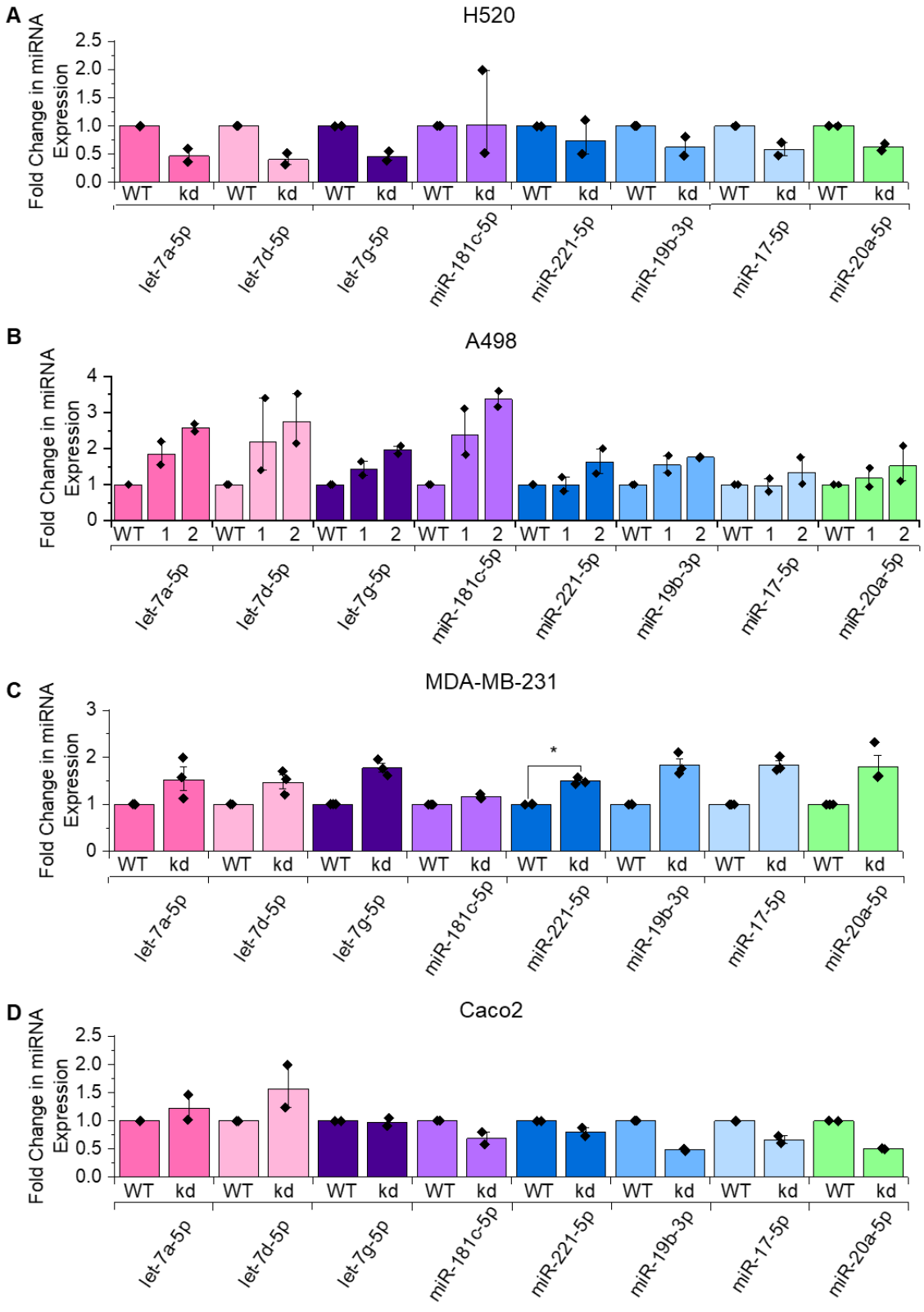


Figure 6.4.2 Expression of mUG and mDG in GRB2-knockdown cancer cell lines

Fold change in mUG and mDG in **(A)** H520, **(B)** A498, **(C)** MDA-MB-231 and **(D)** Caco2 wild type (WT) and knockdown (kd) cell lines. Cells were starved of FBS prior to RNA extraction. miRNA expression was quantified by qPCR with normalisation to RNU6-6P. For H520, A498 and Caco2 N = 2; for MDA-MB-231 N = 3. Error bars indicate standard error of mean. *p < 0.05 by Student's t-test (Bonferroni-corrected).

GRB2 knockdown in both A498 and MDA-MB-231 cells resulted in a general increase in miRNA expression (Figure 6.4.2 B, C). Significant upregulation of miR-221-5p was observed for the MDA-MB-231 cell line (1.5-fold, Figure 6.4.2 C, $p < 0.05$). The Bonferroni adjustment was applied to account for multiple significance testing of the GRB2-regulated miRNAs in new cell lines. miR-181c-5p was expressed similarly in both WT and kd MDA-MB-231. The other miRNAs tested were augmented to a similar extent as each other: for example, let-7a-5p was upregulated 1.6-fold (Figure 6.4.2 C).

miRNA expression was similar in both A498 knockdown cell lines (Figure 6.4.2 B). In general, let-7 family miRNAs showed the greatest amplification (let-7a-5p kd 1 = 1.9, kd 2 = 2.6), along with miR-181c-5p (kd 1 = 2.5, kd 2 = 3.4). miR-19b-3p was also enhanced (kd 1 = 1.5, kd 2 = 1.8), whereas miR-221-5p, miR-17-5p and miR-20a-5p were not consistently dysregulated.

Caco2 GRB2 knockdown cells exhibited a possible 1.6-fold increase in let-7d-5p compared to parental cells (Figure 6.4.2 D). However, let-7a-5p, let-7g-5p, miR-181c-5p and miR-221-5p did not appear to be differentially expressed. miR-19b-3p, miR-17-5p and miR-20a-5p were each diminished 2.0-, 1.5- and 2.0-fold.

Overall, these cell lines show different miRNA signatures in response to GRB2 knockdown, both to each other and to HEK293T cells. However, like G1, A498, MDA-MB-231 and to a lesser extent Caco2 kd cells all showed some upregulation of let-7. Caco2 kd cells also behaved similarly to G1 as they showed diminished mir-17 (miR-17-5p and miR-20a-5p) and miR-19b-3p.

To determine if the increase in let-7 seen in A498 and MDA-MB-231 kd cells was a result of enhanced loading on AGO2, let-7 precursor transcripts were

examined. Cells were incubated in serum-free media before total cell RNA was isolated. Precursor and mature let-7 transcripts were quantified by qPCR. While let-7a-5p and let-7g-5p were upregulated in A498 kd compared to parental, on average precursor let-7 transcripts were similar in all three cell lines (Figure 6.4.3 A). In both replicates, pre-let-7a-1 levels remained around 1 (kd 1 = 1.1; kd 2 = 1.3). Some variation was seen in the measured fold change was recorded for pre-let-7g (kd 1 = 1.2; kd 2 = 1.5). More replicates are required to determine conclusively if let-7 precursor transcripts are expressed at similar levels in both WT and GRB2 kd cells, while mature miRNA is augmented. However, the data available so far suggest that GRB2 knockdown in A498 cells may increase let-7 expression at the point of loading onto AGO2.

In MDA-MB-231 GRB2 kd cells, pre-let-7a-1 was significantly enhanced on average 1.5-fold and all repeats showed upregulation compared to WT cells (Figure 6.4.3 B, $p < 0.05$). Two out of three experimental replicates demonstrated similar increases in pre-let-7g. 1.5-fold amplification of precursor let-7 is similar to the increases of 1.6- and 1.8-fold of let-7a-5p and let-7g-5p. Overall, these data suggest that overexpression of let-7 in this MDA-MB-231 occurs prior to loading.

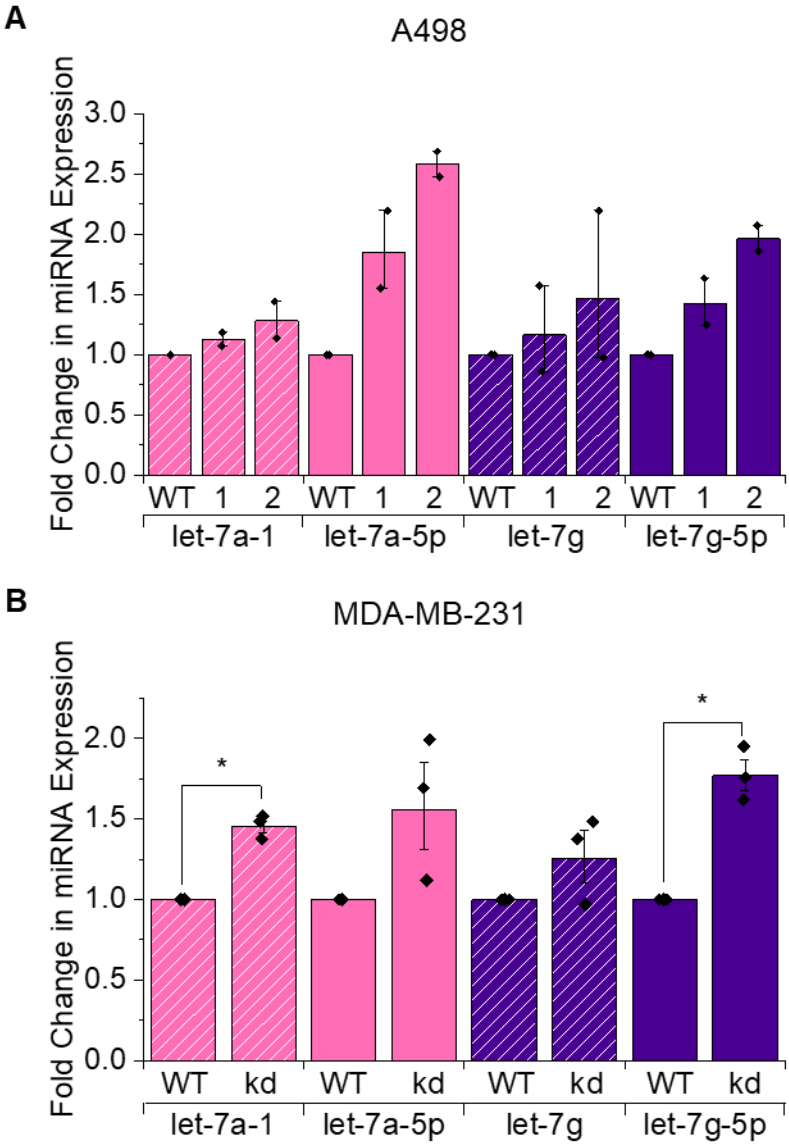


Figure 6.4.3 Expression of precursor and mature let-7 in GRB2-knockdown cancer cell lines

Fold change in precursor (lined) and mature (plain) let-7 miRNA in **(A)** A498 and **(B)** MDA-MB-231 wild type (WT) and knockdown (kd) cell lines. Cells were serum-starved, before total cell RNA was isolated. miRNA was quantified by qPCR with normalisation to RNU6-6P. For MDA-MB-231 N = 3, for A498 N = 2. *$p < 0.05$ (Bonferroni-corrected). Error bars indicate standard error of mean.

6.5 miR-221-5p targets GRB2

Feedback loops are a common feature of miRNA regulation and may mediate either amplification or downregulation of the miRNA or protein in question (Tsang et al., 2007). Therefore, to establish if this happened with GRB2, it was necessary to determine if any of the miRNAs dysregulated upon GRB2 depletion in turn targeted *GRB2*. Although miRNAs usually bind sequences in the 3' UTR, there are several accounts of miRNA-targeting of 3' UTR and coding sequences. Therefore, the full length *GRB2* transcript was searched for potential miRNA targeting sites (TS) using Targetscan 7.2 (Lewis et al., 2005; TargetScan, 2020). Of the GRB2-regulated miRNAs, miR-17-5p, miR-20a-5p, miR-106a-5p, miR-105-5p, miR-221-3p, miR-222-3p and miR-221-5p were predicted to target *GRB2*. An 8mer and two 7mer-A1 TS were identified for miR-221-5p (Figure 6.5.1 A). One 7mer-A1 TS (7mer-A1_1) was located at the start of the 5'UTR, whereas the second (7mer-A1_2) and the 8mer were in the 3'UTR.

A

```
GRB2_3'UTR_8mer      5'-...TGGAGCGUUUGCUGUGCCAGGA...-3'
                                 |  | |||||||
miR-221-5p             3'-UUUAGAUGUAACAUACGGUCCA-5'

GRB2_7mer-A1_1                        5'-AGCCAGGA...-3'
                                          ||||||
miR-221-5p             3'-UUUAGAUGUAACAUACGGUCCA-5

GRB2_7mer-A1_2       5'-...UCCUGCUUGUACUAAGCCAGGA...-3'
                                 |  || ||||||
miR-221-5p             3'-UUUAGAUGUAACAUACGGUCCA-5'
```

B

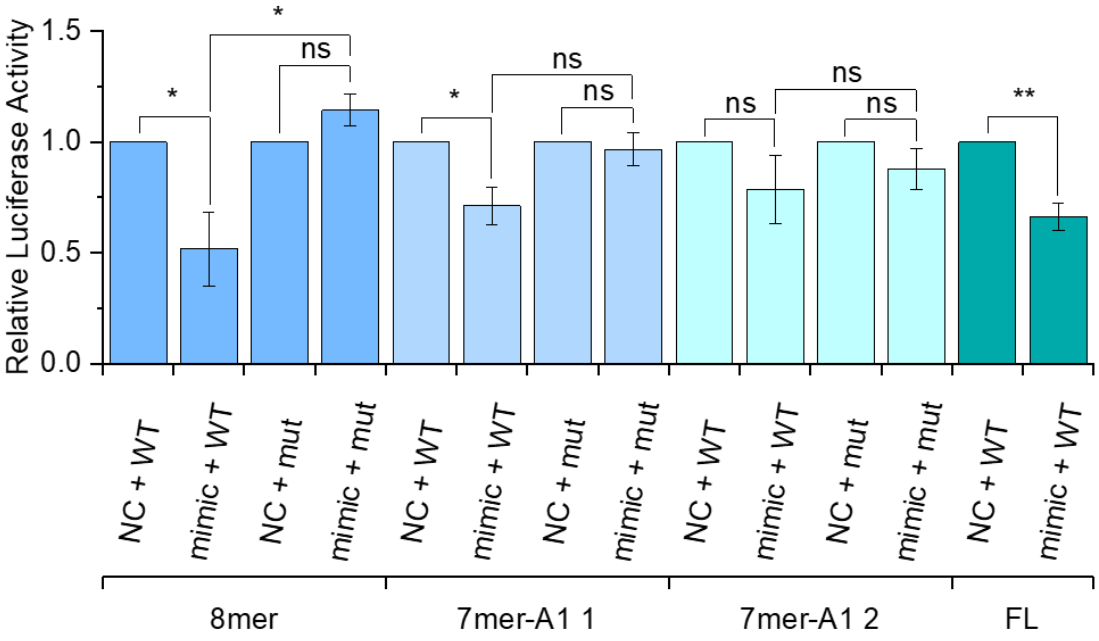


Figure 6.5.1 miR-221-5p targets *GRB2*

(A) Alignment of miR-221-5p and potential targeting sites in *GRB2*. **(B)** Luciferase assay of a miR-221-5p mimic with each of the three predicted targeting sites in *GRB2* mRNA and the full length *GRB2* transcript. Luciferase activity was measured in HEK293T cells expressing a miRNA mimics (miR-221-5p, mimic; negative control, NC) along with a control renilla luciferase plasmid and plasmids containing the firefly luciferase gene fused to wild type (WT) or mutant GRB2 targeting sites, or a sequence spanning the full length *GRB2* transcript (FL). N = 3.

Luciferase assays were used to determine if any of the miRNAs could repress *GRB2* expression. An approach was adopted where isolated 63 nt sequences spanning the predicted TS were fused to the 3' end of a luciferase reporter (*luc*). In these constructs, the seed sequence of each TS was either WT or mutated to a stretch of adenosines. A construct was also used which contained the full length *GRB2* transcript (*GRB2*^{FL}) downstream of *luc*. However, these experiments are limited as use of both incomplete 3'UTRs and a full transcript downstream of another coding sequence does not mimic physiological miRNA targeting of a 3'UTR. The luciferase reporter construct was transfected into HEK293T cells along with the appropriate miRNA mimic (Chapter 2, 2.17). A scrambled miRNA mimic was also used as a negative control (NC). For each condition, luciferase activity was measured.

Significantly reduced luciferase activity was recorded upon transfection of a miR-221-5p mimic along with the 8mer, 7mer-A1_1 and *GRB2*^{FL} constructs, when compared to the negative control mimic (Figure 6.5.1 B). Luciferase activity was not significantly diminished for any of the seed-mutated constructs. The miR-221-5p mimic suppressed luciferase activity by 50% when *luc* was fused to the 8mer TS, and 70% when *luc* was fused to both 7mer-A1_1 and *GRB2*^{FL}. As the reduction in luciferase activity was not any more pronounced with *GRB2*^{FL}, it was likely that the miR-221-5p mimic did not target multiple sites in the *GRB2* transcript. Therefore, the 8mer TS was probably the site of miR-221-5p targeting.

To investigate if miR-221-5p could regulate GRB2, GRB2 expression was quantified in cells treated with the miR-221-5p mimic. A mixture containing 125 nM mimic was used to transfect HEK293T cells. The cells were serum-starved, before GRB2 expression was analysed (48 h post-transfection). Expression of

GRB2 mRNA and protein levels were examined by qPCR and western blotting. Transfection of the miR-221-5p mimic reduced expression of *GRB2* mRNA by 16% (Figure 6.5.2 A) and GRB2 protein by 17% (Figure 6.5.2 B, C), compared to transfection of a negative control (NC). Consequently, expression of the miR-221-5p mimic resulted in a small decrease in GRB2.

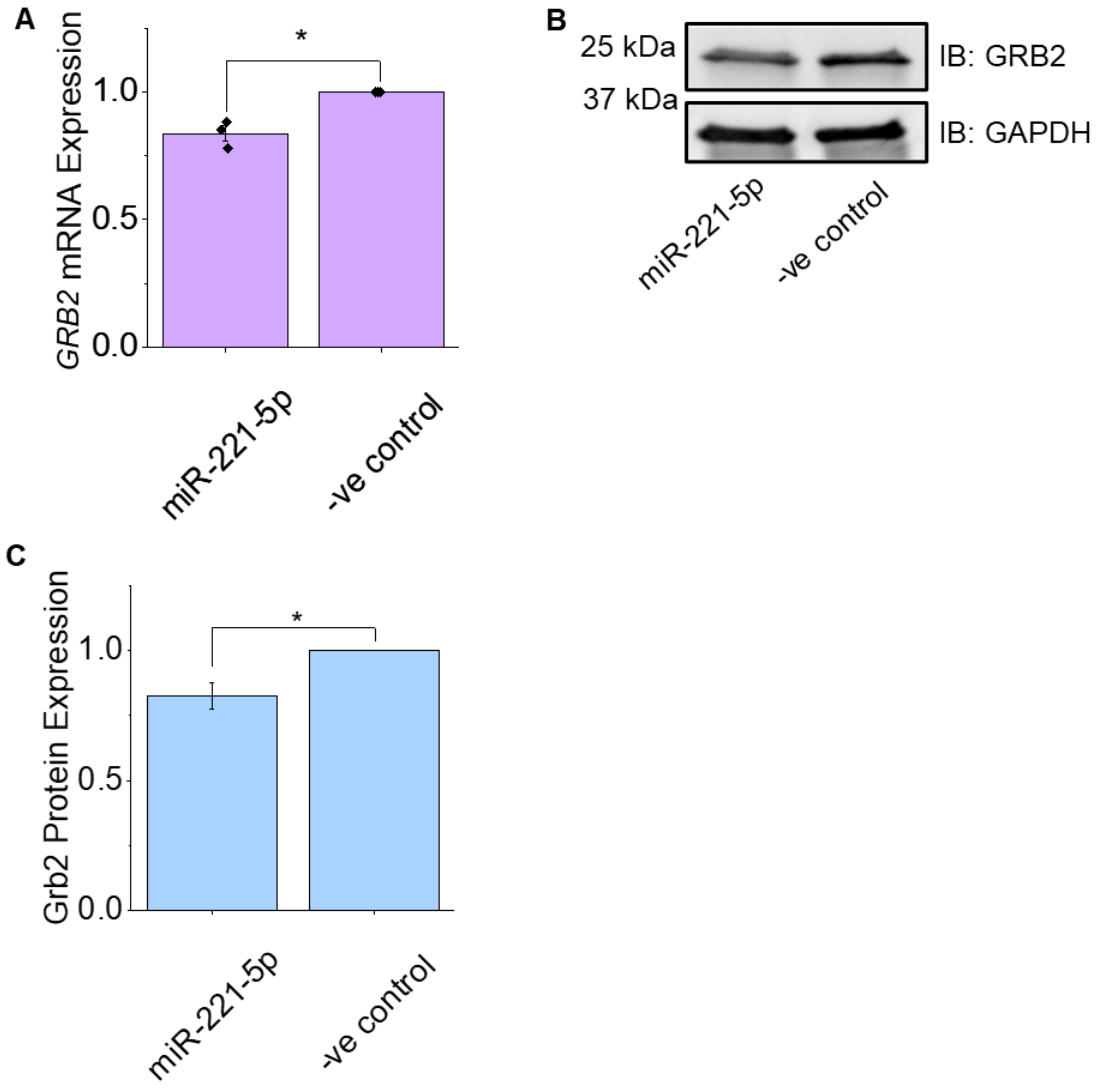


Figure 6.5.2 GRB2 expression is dysregulated by miR-221-5p overexpression

(A) Fold change in *GRB2* mRNA expression upon transfection of miR-221-5p. Total cell RNA was extracted from serum-starved HEK293T cells expressing miR-221-5p or NC mimics and *GRB2* expression quantified by qPCR against GAPDH. N = 3. **(B)** Western blot and **(C)** quantification of GRB2 protein expression upon transfection of miR-221-5p, normalised to GAPDH. HEK293T cells overexpressing miR-221-5p or NC mimics were serum-starved before lysis. N = 3.

6.6 Discussion

Data presented in this Chapter suggest a novel role for GRB2 in regulation of miRNA expression. GRB2 depletion in HEK293T cells promoted loading of let-7 family miRNAs onto AGO2 (Figure 6.1.2 A, B) and suppressed expression of let-7 targets such as HMGA2 (Figure 6.2.1). Therefore, a model may be proposed whereby GRB2 inhibits miRNA-loading and consequently releases miRNA-mediated translational silencing (Figure 6.1.1).

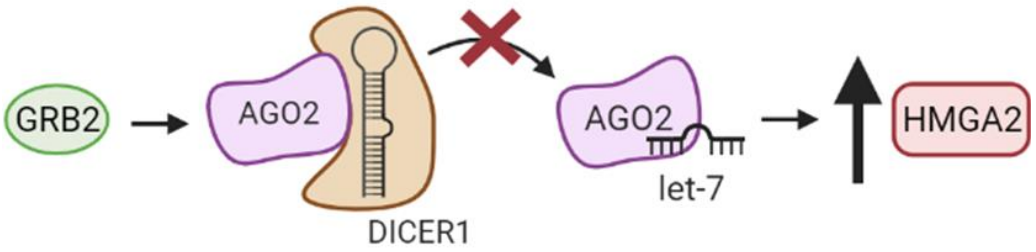


Figure 6.6.1 Model of GRB2-mediated let-7 regulation

GRB2 hinders loading of let-7 onto AGO2. Reduced let-7 activity releases repression of targets such as HMGA2. Created with BioRender.com.

While it was evident that upregulation of let-7 in G1 cells happened at the point loading, the stage at which miR-181c expression was augmented was less certain (Figure 6.1.2 A). Importantly, pre-mir-181c was expressed at much lower levels than either of the pre-let-7 transcripts, which made quantification less reliable. More repeats of this experiment may be required to reduce variation and thus ascertain if both the mir-181 family and the let-7 family are regulated by loading.

Analysis of let-7 target expression demonstrated that, of all those tested, LIN28B and HMGA2 were downregulated to the greatest extent in G1 cells (Figure 6.2.1). Sponge-mediated rescue of these targets was slightly more significant in G1 than in HEK293T (Figure 6.2.1) but did not match the level of repression observed upon GRB2 depletion. Generation of G1 cells which stably express the let-7 sponge may achieve better rescue (Y.T. Lee et al., 2013; Wang et al., 2015; Rowe et al., 2016).

HMGA2 is a transcription factor which is widely reported to promote cancer progression both independently (Cooper, 1996) and as a result of regulation by let-7 (Mayr et al., 2007). LIN28B is also frequently described as an oncogene, mainly attributed to its control of let-7 (Guo et al., 2006; Viswanathan et al., 2008; Newman et al., 2008; Heo et al., 2008). However, LIN28B is may additionally have pro-tumour effects independent of let-7 and has been suggested as a biomarker for chemosensitivity in colon cancer (Ma et al., 2018).

Specific regulation let-7 and their oncogenic targets intimated involvement of the GRB2-let-7 axis in tumour development (Figure 6.2.1). However, analysis of patient data from the TCGA suggested that this does not occur in colorectal cancer (Figure 6.3.2). Significant downregulation of *GRB2* in tumour samples may preclude initiation of this mechanism (Figure 6.3.4). Indeed, cancer cell

lines displayed differing miRNA profiles in response to GRB2 knockdown (Figure 6.4.2), demonstrating context-specific control of miRNA expression. Cell type may be an important determinant of this; the most unique expression signature was shown in H520 squamous carcinoma cells (Figure 6.4.2 A). By contrast, miRNA profiles of A498, MDA-MB-231 and Caco2 adenocarcinoma cells were more similar (Figure 6.4.2 B-D). For example, while H520 kd exhibited downregulation of let-7, all adenocarcinoma cell lines showed some enhancement of this family, albeit to different extents. Varying GRB2-dependent control on let-7 in different cancer cell lines may be unsurprising. let-7 expression may be dysregulated by numerous mechanisms (Viswanathan et al., 2008; Heo et al., 2008; Michlewski and Cáceres, 2010; Wang et al., 2011). Therefore, these may dominate over the GRB2-repression of let-7-loading. Furthermore, it is possible that GRB2 knockdown dysregulates miRNA expression in ways not observed in HEK293T, due to the involvement of GRB2 in multiple signalling pathways.

It should be noted that measurements of miRNA expression in cancer cell lines represent preliminary results only, with only two replicates done for H520, A498 and Caco2 cells. Furthermore, all miRNAs but one showed no statistical significance. However, application of the Bonferroni correction may cause significant results to be recorded as non-significant (Perneger, 1998). Instead, miRNAs which consistently showed dysregulation in the same way may represent true changes in expression and provide avenues for further investigation. Of all the cancer cell lines, only A498 GRB2 kd cells showed increased let-7 expression which could be mediated at the stage of miRNA-loading (Figure 6.4.3). Potentially, this feature is specific to the kidney, as it was also observed in G1 human embryonic kidney cells.

Analysis of miRNA expression in colorectal cancer patient samples and cell lines highlighted potential areas for further study. In the TGCA dataset let-7d was significantly correlated with *GRB2* (Figure 6.3.2, Figure 6.3.3). Furthermore, let-7d-5p was enhanced upon GRB2 knockdown in Caco2 cells (Figure 6.4.2 D), suggesting that the correlation shown in tissue samples may represent regulation of let-7d by GRB2. Neither analyses revealed similar changes of the other members of the let-7 family; thus it appears GRB2 may exert specific control over let-7d expression.

Of the miRNAs which were predicted to bind *GRB2*, in these luciferase assays, miR-221-5p targeted GRB2 directly through binding an 8mer TS in the GRB2 3' UTR (Figure 6.5.1). However, the *luc* reporter constructs used contained either 63 nt sequences corresponding to only a small part of the 3'UTR, or the entire *GRB2* transcript downstream of *luc*. Thus, neither mimicked a normal mRNA transcript. Structural features of the miRNA TS are integral to targeting efficiency (Long et al., 2007) and were unlikely to be properly replicated here. To obtain more meaningful results, the experiment should be completed using only the full 3'UTR.

Downregulation of GRB2 was demonstrated upon transfection of a miR-221-5p mimic (Figure 6.5.2). However, a small decrease in expression was recorded despite the high concentration of mimic (125 nM). Furthermore, miR-221-5p is the less dominant arm of miR-221 and is less likely to be expressed at high enough levels to target GRB2. Transfection of a miRNA inhibitor is required to determine if endogenous miR-221-5p targets GRB2. If endogenous miR-221-5p does target GRB2, it may be hypothesised that a negative feedback loop could be formed by which GRB2 maintains its own expression (Figure 6.6.2).

Figure 6.6.2 Potential model of feedback loop regulating GRB2 expression

High levels of GRB2 increase expression of pri-mir-221. If endogenous miR-221-5p targets GRB2, increased miR-221-5p could mediate translational silencing of GRB2. Created with BioRender.com.

Chapter 7: Discussion

In this study, two potential mechanisms of miRNA regulation were investigated: firstly, mediated by FGFR2, and secondly, by GRB2 and independent of FGFR2. Although FGFR2-control of miRNA biogenesis was not identified, GRB2 was found to control miRNA expression in two ways: 1) GRB2 inhibited loading of miRNAs onto AGO2, and 2) GRB2 augmented miRNA biogenesis before the stage of DROSHA-mediated pri-miRNA processing.

7.1 FGFR2 regulation of AGO2

FGFR2 does not appear to phosphorylate AGO2 Y393 and consequently does not regulate AGO2 binding DICER1 in non-stimulated and stimulated HEK293T cells stably expressing FGFR2 (Figure 4.4.2, Figure 4.4.3), nor in MDA-MB-231 and MCF-7 breast cancer cell lines (Figure 4.5.1, Figure 4.5.2). However, FGFR2 overexpression in HEK293T cells mediates phosphorylation of Y322, Y529 and Y804/805 upon stimulation with FGF2/9 (Table 4.3.2, Figure 4.3.2). Further work to examine phosphorylation of these residues may determine if this mechanism occurs in cancers with amplified FGFR2 expression. Additionally, *in vitro* phosphorylation assays may determine if FGFR2 is the direct kinase for these residues.

FGFR2-independent Y393 phosphorylation was detected upon exogenous expression of AGO2 in MCF-7 and MDA-MB-231 cells (Figure 4.5.2). AGO2 is upregulated in more severe breast cancer subtypes (Kulkarni et al., 2019) and at more advanced stages (Zhou et al., 2019). Therefore, it may be interesting to investigate other kinases which mediate changes miRNA-loading in these cancers.

7.2 GRB2 binding AGO2 hinders miRNA-loading

GRB2 directly binds AGO2 PAZ domain to incorporate itself in the RLC (Chapter 5, Figure 5.2.2, Figure 5.4.1 A). Additionally, GRB2 inhibits miRNA-loading onto AGO2 (Chapter 6, Figure 6.1.2). The GRB2-binding ^{323}PHLP^{326} motif forms the pocket in which the 3’ end of the miRNA is docked (Figure 7.2.1). Consequently, GRB2 binding may alter AGO2 structure in this region, to hinder loading of let-7 miRNAs. Reduced loading of let-7 results in a decrease in their activity, as seen by an increase in targets such as HMGA2 (Figure 7.2.2).

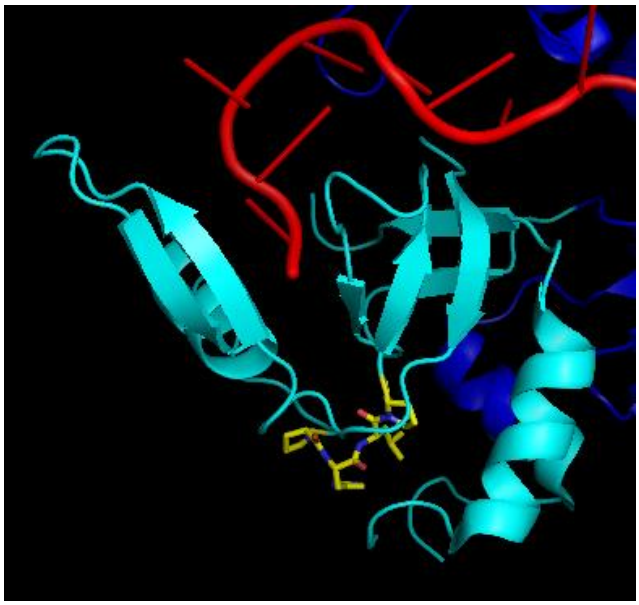


Figure 7.2.1 GRB2 binding AGO2 ^{323}PHLP^{326} may inhibit miRNA-loading

(A) Crystal structural model of AGO2 (dark blue) bound to a miRNA (red). AGO2 PAZ domain shown in cyan. ^{323}PHLP^{326} in yellow. (PDB: 4W5N. Schirle, Sheu-Gruttadauria, & MacRae, 2014).

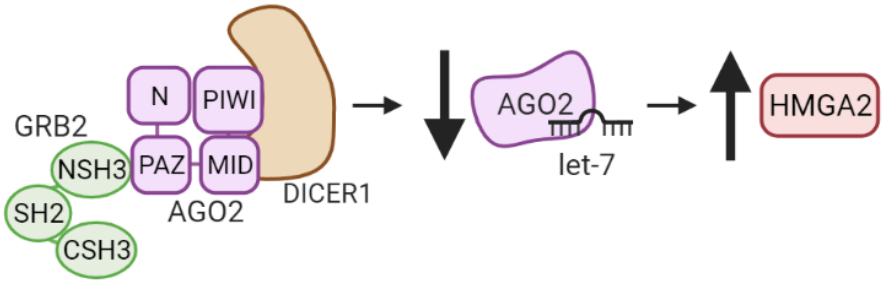


Figure 7.2.2 Model of GRB2 regulation of oncogene expression

Binding of GRB2 to AGO2 in the miRISC-loading complex (RLC) inhibits let-7 loading onto AGO2. let-7-bound AGO2 is reduced. The expression of oncogenic targets, such as HMGA2, is increased.

Typically, proteins which associate with the RLC promote loading of miRNAs (Chendrimada et al., 2005; Lee et al., 2006; Iwasaki et al., 2010). As GRB2 precludes loading, it appears the RLC may consist of many proteins with different functions and which may be required to fine-tune loading. While GRB2 inhibited miRNA-loading, it did not hinder AGO2 binding DICER1. In fact, it may have slightly stabilised the interaction (Figure 5.4.1, Figure 5.4.3). Consequently, GRB2 binding AGO2 may act to pause miRNA-loading at this point. Interestingly, GRB2 bound AGO2 PAZ domain when it was in both its monomer and dimer forms (Figure 5.2.2, Figure 5.2.5). It remains to be determined which of these complexes exist *in vivo*, raising the question of whether, through dimerization, GRB2 may bring together two RLCs. Formation of higher order RISC structures is observed in miRNA-mediated gene silencing: cooperative targeting by multiple AGO2-miRNA complexes results in enhanced targeting (Saetrom et al., 2007; Grimson et al., 2007) and can be mediated by TNRC6 (Elkayam et al., 2017; Briskin et al., 2020). Furthermore, AGO2-TNRC6 may form phase separated droplets with augmented deadenylation activity (Sheu-Gruttadauria and MacRae, 2018).

Other features of the GRB2-AGO2 interaction support the theory that miRNA-loading is the stage at which miRNA biogenesis is affected. Binding was observed in the cell cytoplasm (Figure 5.1.1 B), the same location as miRNA-loading (Gagnon et al., 2014). Furthermore, FRET was not observed in P-bodies and GRB2 did not localise to these regions, suggesting GRB2 does not form part of the mature miRISC.

GRB2 has previously been suggested to recruit PTP1B to AGO2 to allow AGO2 dephosphorylation (Haidar et al., 2018). Therefore, if GRB2 complexes AGO2, DICER1 and PTP1B, it appears the adaptor may play opposing roles in the

control of miRNA expression. However, the two mechanisms may not occur simultaneously; GRB2 binding AGO2-DICER1 is via AGO2 PxxP motif and occurs under growth factor deprivation. In contrast, GRB2 recruitment of PTP1B was recorded with cells in media supplemented with 10% FBS and therefore may be phosphorylation-dependent (Haidar et al., 2018).

An important question is why the loading of only some miRNAs was upregulated upon GRB2 knockdown. Phosphorylation of AGO2 Y393 suppresses loading of a subset of miRNAs with a long-loop structure in their precursor transcript (Shen et al., 2013). DICER1 has preference for binding these miRNAs (Tsutsumi et al., 2011). Thus, inhibition of DICER1 loading miRNAs onto AGO2 particularly affects these miRNAs. Interestingly, most of the miRNAs enhanced in G1 cells also have a long loop structure in their precursor transcript (eight or more nucleotides, $nt \geq 8$) (Table 7.2.1). Pre-mir-95 and pre-mir-181a-1 have medium length loops ($8 > nt \geq 6$) and pre-let-7g has a short loop ($6 > nt$) (Table 7.2.1). It should be noted that let-7a and let-7f have multiple precursors leading to the production of identical mature miRNAs (pre-let-7a-1, -2 and -3; pre-let-7f-1 and -2, respectively). However, the identical nature of the mature forms meant that the precursors responsible for upregulated let-7a-5p, let-7a-3p and let-7f-5p could not be differentiated in the small-RNA sequencing experiment or by qPCR.

Additionally, of the let-7 family, only group II miRNAs (let-7a-1, let-7a-3, let-7b, let-7d, let-7f-1, let-7f-2, let-7g, let-7i, and miR-98) were regulated by GRB2. Mono-uridylation at the 3’ end of these miRNAs generates the 2 nt overhang, required for processing by DICER1 (Heo et al., 2012). Consequently, while group I miRNAs have a 3’ overhang of two cytosines (CC), group II miRNAs have a cytosine followed by a uridine (CU). miRNAs with the CU overhang are

better substrates for DICER1 (Vermeulen et al., 2005). Thus, loading these group II let-7 miRNAs may be particularly affected by GRB2 knockdown, due to both their overhang sequence and long loop structure, which make them ideal substrates for DICER1.

Table 7.2.1 Loop length and 3’ overhang sequences of the precursors for mUGs and other let-7 family members

Pre-miRNA loop length in nucleotides. Mature 5p and 3p miRNAs which were enhanced in G1 cells are identified. For some mature miRNAs, several precursors exist (denoted with numbered suffixes). 3’ overhangs are taken from Heo et al., 2012 and miRbase (Kozomara and Griffiths-Jones, 2011).

miRNA	Loop length (nt)	3’ Overhang	Regulation by Grb2
let-7a-1	27	CU	5p & 3p
let-7a-2	5	CC	5p
let-7a-3	3	CU	5p & 3p
let-7b	30	CU	5p & 3p
let-7c	5	CC	-
let-7d	30	CU	5p & 3p
let-7e	4	CC	-
let-7f-1	5	CU	5p
let-7f-2	26	CU	5p
let-7g	4	CU	5p
let-7i	9	CU	5p & 3p
mir-98	11	CU	5p
mir-181c	11	C	5p
mir-181a-1	7	CC	3p
mir-181a-2	11	CC	3p
mir-204	15	GU	5p
mir-95	6	GCA	3p

Alternatively, the conformation adopted by GRB2-bound AGO2 may specifically preclude loading of let-7 family miRNAs. A similar mechanism was suggested to explain regulation of let-7 association with AGO2 by TP53 (Krell et al., 2016) and structural adjustments in AGO2 are central to the loading process (Tsuboyama et al., 2018). Crystallisation of the AGO2-GRB2 structure may reveal if this is the case.

While binding of GRB2 to AGO2 seems the most likely explanation for how let-7-loading is regulated, the specific regulation of all group II let-7 miRNAs may suggest the possibility that GRB2 may inhibit mono-uridylation of the pre-miRNAs by TUT. GRB2 does not regulate TUT4 expression (Figure 6.1.3). However, expression of TUT2 and TUT7 were not tested. Alternatively, it is possible GRB2 may inhibit uridylation by a different mechanism, such as by sequestering TUT from let-7. *In vitro* loading assays using purified AGO2, DICER1 and GRB2 may be required to demonstrate that GRB2 autonomously modulates let-7-loading.

Specific deregulation of the let-7 family may indicate GRB2-mediated control of embryonic development and differentiation of embryonic stem cells (ESCs) (Pasquinelli et al., 2000; Reinhart et al., 2000). GRB2 plays a vital role in early embryonic development and knocking it out in ESCs produces defects in differentiation (Cheng et al., 1998). Furthermore, suppression of let-7 in ESCs is attributed partly to extensive post-transcriptional control (Bracht et al., 2004; Thomson et al., 2006; Viswanathan et al., 2008; Heo et al., 2008). In mouse ESCs, it has been shown that a reduction in AGO2 expression specifically diminishes let-7 and let-7 silencing is enhanced by reciprocal feedback by LIN28A (Liu et al., 2021). It is possible that in this study reduced inhibition by LIN28B further augmented let-7 in G1 cells (Figure 6.2.1). Like let-7, loading of

the mir-181 family may be also regulated by GRB2 (Figure 6.1.2 A); mir-181 regulates both embryonic implantation (Chu et al., 2015) and later prenatal development (Garcia-Riart et al., 2017). Therefore, it may be interesting to examine miRNA expression upon GRB2 knockdown in ESCs.

GRB2-mediated inhibition of let-7 loading in AGO2 may also be of importance in kidney cancer (Figure 6.4.2 B, Figure 6.4.3 A). Further work is required to determine if correlations between GRB2 and let-7 exist in patients. GRB2 has been identified as a potential driver in renal clear cell carcinoma (RCC) (Zhang and Zhang, 2017). Furthermore, the let-7 family have been suggested as biomarkers for detection of patients with RCC (Fedorko et al., 2017) and reduced let-7b and c has been may mediate chemoresistance in RCC (Peng et al., 2015). Additionally, both let-7 and LIN28B are implicated in the development of Wilms tumour (Urbach et al., 2014).

If GRB2-regulation of miRNA-loading does contribute to tumorigenesis, targeting the GRB2-AGO2 interaction may not present a likely strategy for therapeutic intervention, due to the low affinity of SH3 domain interactions (Chen et al., 1993; Yu et al., 1994). Instead, it is possible that detection of changes in miRNA expression could aid cancer diagnosis or prognosis.

7.3 GRB2-mediated suppression of miRNA expression

While let-7 miRNAs appear to be regulated by loading onto AGO2, it is less clear how expression of the mDGs is controlled. As several members of the same cluster are affected (miR-17, miR-19a, miR-18a, miR-20a and miR-19b-1 of the mir-17~92 cluster; and miR-221 and miR-222 of the mir-221 cluster) it seems likely that transcription of these miRNAs is altered. Transcription of both

these clusters is activated by nuclear factor κB (NF-κB) and STAT3 (R. Zhou et al., 2010; Liu et al., 2014; Spaccarotella et al., 2014). Both TFs are activated by signalling pathways which are initiated by cytokines, growth factors and hormones. Consequently, it is difficult to postulate how NF-κB and STAT3 are regulated in the serum-starved cells. However, considering the involvement of GRB2 in these pathways (Zhang et al., 2003; Shostak and Chariot, 2015), it is possible that altered GRB2 expression could disrupt the mechanisms that regulate these TFs. Locus-specific chromatin IP followed by MS may identify changes TFs associated with the mir-17~19 and mir-221 clusters upon GRB2 depletion. Furthermore, mass spectrometry analysis and reverse phase protein array (RPPA) of protein expression and modification may help determine upstream changes in cell signalling pathways which regulate these TFs.

Alternatively, if GRB2 does not regulate mDG transcription, it is possible that it affects other biogenesis steps which occur before pri-miRNA cleavage by DROSHA. For example, DROSHA processing of mir-17~92 can only occur after removal of an inhibitor domain in the polycistronic pri-miRNA transcript (Du et al., 2015).

More work is required to determine the biological significance of GRB2-modulation of mDGs. Repression of mir-17 and mir-19 family mDGs was implicated in GRB2-diminished Caco2 cells (Figure 6.4.2 D), but not in analysis of the TCGA colorectal cancer cohort (Figure 6.3.2, Figure 6.3.3, Figure 6.3.4), nor in any of the other GRB2-depleted cancer cell lines (Figure 6.4.2 A-C). While GRB2 depletion diminished miR-221-5p in HEK293T (Figure 6.1.1, Figure 6.1.2 A), miR-221-5p was significantly upregulated upon GRB2 knockdown in MDA-MB-231 cells (Figure 6.4.2 C). It is therefore possible that upon tumorigenesis, other mechanisms prevail to control expression of mUG.

If miR-221-5p directly targets GRB2, suppression of GRB2 may reflect previously reported physiological roles of the miRNA. miR-221 functions in normal cells to control vascular processes, including angiogenesis and wound healing (Poliseno et al., 2006; Yajaira et al., 2007; Nam et al., 2008; Davis et al., 2009; Xiaojun et al., 2009; Mujahid et al., 2013). Although miR-221-5p is far less studied than the 3p arm, miR-221-5p has been suggested to protect against inflammation in colitis (Fang et al., 2015) and acute gouty arthritis (Li et al., 2021). Considering the role of GRB2 in transducing MAPK signalling downstream of cytokine receptors (Hibi and Hirano, 1998; Heinrich et al., 2003), as well as the RTKs vascular endothelial growth factor receptor (VEGFR) and platelet-derived growth factor receptor (PDGFR), miR-221-5p targeting GRB2 may help the miRNA to both mediate an anti-inflammatory response and regulate vascular processes.

While miR-221-3p is commonly cited as an oncomir (Tao et al., 2014; Shi et al., 2017; Wei et al., 2017; Li et al., 2018; Dong et al., 2019), the role of miR-221-5p is less clear. miR-221-5p is downregulated in prostate cancer (Taylor et al., 2010; Kiener et al., 2019) and colorectal cancer (Yuan et al., 2013), but upregulated in RCC (Liu et al., 2019). miR-221-5p targets the transcriptional regulator Smad family member 3 (SMAD3) in multiple myeloma (Fan et al., 2019) and suppressor of cytokine signalling 1 (SOCS1) in prostate cancer (Shao et al., 2018) to promote disease pathogenesis. However, miR-221-5p also silences disc domain receptor 1 (DDR1) (Jiang et al., 2020) and methyl-CpG-binding protein (MBD2) (Liu et al., 2014) to reduce cell proliferation and metastasis in gastric and colorectal carcinoma respectively. Consequently, if targeting of GRB2 is also considered, it is possible that miR-221-5p acts as a tumour suppressor, in contrast with the 3p arm. Opposing functions of -5p and -

3p arms has previously been described for miR-28 in nasopharyngeal cancer (Lv et al., 2019), miR-574 in gastric cancer (Z. Zhang et al., 2019) and miR-193a in both breast and gastric cancer (Tsai et al., 2016; Chou et al., 2018). It may therefore be interesting to explore if arm switching of miR-221 has clinical relevance.

7.4 Summary

Data presented in this thesis demonstrate a previously unreported role for GRB2 in the regulation of miRNA biogenesis. GRB2 binds AGO2 directly, via an interaction mediated by GRB2 NSH3 and AGO2 PAZ domain. In this way, GRB2 is incorporated into the RLC to hinder loading of let-7 miRNAs.

Bibliography

Abou-Alfa, G.K., Sahai, V., Hollebecque, A., Vaccaro, G., Melisi, D., Al-Rajabi, R., Paulson, A.S., Borad, M.J., Gallinson, D., Murphy, A.G., Oh, D.-Y., Dotan, E., Catenacci, D. V, Van Cutsem, E., Ji, T., Lihou, C.F., Zhen, H., Féliz, L. and Vogel, A. 2020. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *The Lancet Oncology*. **21**(5), pp.671–684.

Adams, B.D., Claffey, K.P. and White, B.A. 2009. Argonaute-2 expression is regulated by epidermal growth factor receptor and mitogen-activated protein kinase signaling and correlates with a transformed phenotype in breast cancer cells. *Endocrinology*. **150**(1), pp.14–23.

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.

Ahmed, Z., Lin, C.-C., Suen, K.M., Melo, F.A., Levitt, J.A., Suhling, K. and Ladbury, J.E. 2013. Grb2 controls phosphorylation of FGFR2 by inhibiting receptor kinase and Shp2 phosphatase activity. *The Journal of Cell Biology*. **200**(4), 493 LP – 504.

Ahmed, Z., Timsah, Z., Suen, K.M., Cook, N.P., Lee, G.R., Lin, C.-C., Gagea, M., Marti, A.A. and Ladbury, J.E. 2015. Grb2 monomer–dimer equilibrium determines normal versus oncogenic function. *Nature Communications*. **6**(1), p.7354.

Aitio, O., Hellman, M., Kesti, T., Kleino, I., Samuilova, O., Pääkkönen, K., Tossavainen, H., Saksela, K. and Permi, P. 2008. Structural Basis of PxxDY Motif Recognition in SH3 Binding. *Journal of Molecular Biology*.

382(1), pp.167–178.

Akao, Y., Nakagawa, Y. and Naoe, T. 2006. *let-7* MicroRNA Functions as a Potential Growth Suppressor in Human Colon Cancer Cells. *Biological and Pharmaceutical Bulletin*. **29**(5), pp.903–906.

Ali Syeda, Z., Langden, S.S.S., Munkhzul, C., Lee, M. and Song, S.J. 2020. Regulatory Mechanism of MicroRNA Expression in Cancer. *International journal of molecular sciences*. **21**(5), p.1723.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of molecular biology*. **215**(3), pp.403–410.

Arroyo, J.D., Chevillet, J.R., Kroh, E.M., Ruf, I.K., Pritchard, C.C., Gibson, D.F., Mitchell, P.S., Bennett, C.F., Pogosova-Agadjanyan, E.L., Stirewalt, D.L., Tait, J.F. and Tewari, M. 2011. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences*. **108**(12), 5003 LP – 5008.

Asrih, M. and Steffens, S. 2013. Emerging role of epigenetics and miRNA in diabetic cardiomyopathy. *Cardiovascular pathology : the official journal of the Society for Cardiovascular Pathology*. **22**(2), pp.117–125.

Auyeung, V.C., Ulitsky, I., McGeary, S.E. and Bartel, D.P. 2013. Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. *Cell*. **152**(4), pp.844–858.

Babar, I.A., Cheng, C.J., Booth, C.J., Liang, X., Weidhaas, J.B., Saltzman, W.M. and Slack, F.J. 2012. Nanoparticle-based therapy in an in vivo microRNA-155 (miR-155)-dependent mouse model of lymphoma.

Proceedings of the National Academy of Sciences of the United States of America. **109**(26), pp.E1695–E1704.

Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P. and Blelloch, R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes & Development* . **22**(20), pp.2773–2785.

Babraham_Institute 2019. FastQC. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R. and Pasquinelli, A.E. 2005. Regulation by let-7 and lin-4 miRNAs Results in Target mRNA Degradation. *Cell*. **122**(4), pp.553–563.

Bartel, D.P. 2018. Metazoan MicroRNAs. *Cell*. **173**(1), pp.20–51.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. and Izaurralde, E. 2006. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & development*. **20**(14), pp.1885–1898.

Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*. **409**, p.363.

Biscardi, J.S., Belsches, A.P. and Parsons, S.J. 1998. Characterization of human epidermal growth factor receptor and c-Src interactions in human breast tumor cells. *Molecular Carcinogenesis*. **21**(4), pp.261–272.

Bisson, N., James, D.A., Ivosev, G., Tate, S.A., Bonner, R., Taylor, L. and

Pawson, T. 2011. Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nature Biotechnology*. **29**, p.653.

Blast 2020. BLAST. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi.

Boland, A., Tritschler, F., Heimstädt, S., Izaurralde, E. and Weichenrieder, O. 2010. Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO reports*. **11**(7), pp.522–527.

Booker, G.W., Gout, I., Kristina^Downing, A., Driscoll, P.C., Boyd, J., Waterfield, M.D. and Campbell, I.D. 1993. Solution structure and ligand-binding site of the SH3 domain of the p85α subunit of phosphatidylinositol 3-kinase. *Cell*. **73**(4), pp.813–822.

Bornet, O., Nouailler, M., Feracci, M., Sebban-Kreuzer, C., Byrne, D., Halimi, H., Morelli, X., Badache, A. and Guerlesquin, F. 2014. Identification of a Src kinase SH3 binding site in the C-terminal domain of the human ErbB2 receptor tyrosine kinase. *FEBS Letters*. **588**(12), pp.2031–2036.

Bos, J.L., Toksoz, D., Marshall, C.J., Verlaan-de Vries, M., Veeneman, G.H., van der Eb, A.J., van Boom, J.H., Janssen, J.W. and Steenvoorden, A.C. 1985. Amino-acid substitutions at codon 13 of the N-ras oncogene in human acute myeloid leukaemia. *Nature*. **315**(6022), pp.726–730.

Bracht, J., Hunter, S., Eachus, R., Weeks, P. and Pasquinelli, A.E. 2004. Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *Rna*. **10**(10), pp.1586–1594.

Bridge, K.S., Shah, K.M., Li, Y., Foxler, D.E., Wong, S.C.K., Miller, D.C., Davidson, K.M., Foster, J.G., Rose, R., Hodgkinson, M.R., Ribeiro, P.S., Aboobaker, A.A., Yashiro, K., Wang, X., Graves, P.R., Plevin, M.J., Lagos,

D. and Sharp, T. V 2017. Argonaute Utilization for miRNA Silencing Is Determined by Phosphorylation-Dependent Recruitment of LIM-Domain-Containing Proteins. *Cell Reports*. **20**(1), pp.173–187.

Briskin, D., Wang, P.Y. and Bartel, D.P. 2020. The biochemical basis for the cooperative action of microRNAs. *Proceedings of the National Academy of Sciences*. **117**(30), 17764 LP – 17774.

Bronevetsky, Y., Villarino, A. V, Eisley, C.J., Barbeau, R., Barczak, A.J., Heinz, G.A., Kremmer, E., Heissmeyer, V., McManus, M.T., Erle, D.J., Rao, A. and Ansel, K.M. 2013. T cell activation induces proteasomal degradation of Argonaute and rapid remodeling of the microRNA repertoire. *The Journal of Experimental Medicine*. **210**(2), 417 LP – 432.

Broughton, J.P., Lovci, M.T., Huang, J.L., Yeo, G.W. and Pasquinelli, A.E. 2016. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Molecular Cell*. **64**(2), pp.320–333.

Bu, P., Wang, L., Chen, K.-Y., Rakhilin, N., Sun, J., Closa, A., Tung, K.-L., King, S., Kristine Varanko, A., Xu, Y., Huan Chen, J., Zessin, A.S., Shealy, J., Cummings, B., Hsu, D., Lipkin, S.M., Moreno, V., Gümüş, Z.H. and Shen, X. 2015. miR-1269 promotes metastasis and forms a positive feedback loop with TGF-β. *Nature communications*. **6**, p.6879.

Buday, L. and Downward, J. 1993. Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell*. **73**(3), pp.611–620.

Cai, X., Hagedorn, C.H. and Cullen, B.R. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*. **10**(12), pp.1957–1966.

Calin, G.A. and Croce, C.M. 2006. MicroRNA signatures in human cancers. *Nature Reviews Cancer*. **6**(11), pp.857–866.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F. and Croce, C.M. 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*. **99**(24), 15524 LP – 15529.

Calin, G.A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S.E., Iorio, M. V, Visone, R., Sever, N.I., Fabbri, M., Iuliano, R., Palumbo, T., Pichiorri, F., Roldo, C., Garzon, R., Sevignani, C., Rassenti, L., Alder, H., Volinia, S., Liu, C., Kipps, T.J., Negrini, M. and Croce, C.M. 2005. A MicroRNA Signature Associated with Prognosis and Progression in Chronic Lymphocytic Leukemia. *New England Journal of Medicine*. **353**(17), pp.1793–1801.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M. and Croce, C.M. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America*. **101**(9), 2999 LP – 3004.

Carè, A., Catalucci, D., Felicetti, F., Bonci, D., Addario, A., Gallo, P., Bang, M.-L., Segnalini, P., Gu, Y., Dalton, N.D., Elia, L., Latronico, M.V.G., Høydal, M., Autore, C., Russo, M.A., Dorn, G.W., Ellingsen, Ø., Ruiz-Lozano, P., Peterson, K.L., Croce, C.M., Peschle, C. and Condorelli, G. 2007.

MicroRNA-133 controls cardiac hypertrophy. *Nature Medicine*. **13**(5), pp.613–618.

Casey, M.C., Prakash, A., Holian, E., McGuire, A., Kalinina, O., Shalaby, A., Curran, C., Webber, M., Callagy, G., Bourke, E., Kerin, M.J. and Brown, J.A. 2019. Quantifying Argonaute 2 (Ago2) expression to stratify breast cancer. *BMC Cancer*. **19**(1), p.712.

Cha, J.Y., Maddileti, S., Mitin, N., Harden, T.K. and Der, C.J. 2009. Aberrant Receptor Internalization and Enhanced FRS2-dependent Signaling Contribute to the Transforming Activity of the Fibroblast Growth Factor Receptor 2 IIIb C3 Isoform. *The Journal of Biological Chemistry*. **284**(10), pp.6227–6240.

Chamorro, M.E., Maltaneri, R.E., Vittori, D.C. and Nesse, A.B. 2015. Protein tyrosine phosphatase 1B (PTP1B) is involved in the defective erythropoietic function of carbamylated erythropoietin. *The International Journal of Biochemistry & Cell Biology*. **61**, pp.63–71.

Chang, S.S., Smith, I., Glazer, C., Hennessey, P. and Califano, J.A. 2010. EIF2C is overexpressed and amplified in head and neck squamous cell carcinoma. *ORL; journal for oto-rhino-laryngology and its related specialties*. **72**(6), pp.337–343.

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), 1338 LP – 1343.

Chen, C.-H., Chen, M.-K., Jeng, K.-C.G. and Lung, F.-D.T. 2010. Effects of peptidic antagonists of Grb2-SH2 on human breast cancer cells. *Protein and peptide letters*. **17**(1), pp.44–53.

Chen, C.-Y.A., Zheng, D., Xia, Z. and Shyu, A.-B. 2009. Ago–TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nature Structural & Molecular Biology*. **16**(11), pp.1160–1166.

Chen, J.K., Lane, W.S., Brauer, A.W., Tanaka, A. and Schreiber, S.L. 1993. Biased combinatorial libraries: novel ligands for the SH3 domain of phosphatidylinositol 3-kinase. *Journal of the American Chemical Society*. **115**(26), pp.12591–12592.

Chen, R.H., Sarnecki, C. and Blenis, J. 1992. Nuclear localization and regulation of erk- and rsk-encoded protein kinases. *Molecular and Cellular Biology*. **12**(3), 915 LP – 927.

Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K., Guo, J., Zhang, Yujing, Chen, J., Guo, X., Li, Q., Li, X., Wang, W., Zhang, Yan, Wang, Jin, Jiang, X., Xiang, Y., Xu, C., Zheng, P., Zhang, Juanbin, Li, R., Zhang, H., Shang, X., Gong, T., Ning, G., Wang, Jun, Zen, K., Zhang, Junfeng and Zhang, C.-Y. 2008. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research*. **18**(10), pp.997–1006.

Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O. and Izaurralde, E. 2014. A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. *Molecular Cell*. **54**(5), pp.737–750.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K. and Shiekhattar, R. 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*. **436**, p.740.

Cheng, A.M., Saxton, T.M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W.,

Tortorice, C.G., Cardiff, R.D., Cross, J.C., Muller, W.J. and Pawson, T. 1998. Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. *Cell*. **95**(6), pp.793–803.

Cheng, C.J., Bahal, R., Babar, I.A., Pincus, Z., Barrera, F., Liu, C., Svoronos, A., Braddock, D.T., Glazer, P.M., Engelman, D.M., Saltzman, W.M. and Slack, F.J. 2015. MicroRNA silencing for cancer therapy targeted to the tumour microenvironment. *Nature*. **518**(7537), pp.107–110.

Chi, S.W., Hannon, G.J. and Darnell, R.B. 2012. An alternative mode of microRNA target recognition. *Nature structural & molecular biology*. **19**(3), pp.321–327.

Chou, N.-H., Lo, Y.-H., Wang, K.-C., Kang, C.-H., Tsai, C.-Y. and Tsai, K.-W. 2018. MiR-193a-5p and -3p Play a Distinct Role in Gastric Cancer: miR-193a-3p Suppresses Gastric Cancer Cell Growth by Targeting ETS1 and CCND1. *Anticancer research*. **38**(6), pp.3309–3318.

Choudhury, N.R., Nowak, J.S., Zuo, J., Rappsilber, J., Spoel, S.H. and Michlewski, G. 2014. Trim25 Is an RNA-Specific Activator of Lin28a/TuT4-Mediated Uridylation. *Cell reports*. **9**(4), pp.1265–1272.

Chu, B., Zhong, L., Dou, S., Wang, J., Li, J., Wang, M., Shi, Q., Mei, Y. and Wu, M. 2015. miRNA-181 regulates embryo implantation in mice through targeting leukemia inhibitory factor. *Journal of Molecular Cell Biology*. **7**(1), pp.12–22.

Ciafrè, S.A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C.-G., Sabatino, G., Negrini, M., Maira, G., Croce, C.M. and Farace, M.G. 2005. Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochemical and Biophysical Research Communications*. **334**(4), pp.1351–1358.

Cicchetti, P., Mayer, B.J., Thiel, G. and Baltimore, D. 1992. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science*. **257**(5071), 803 LP – 806.

Cooper, C.S. 1996. Translocations in solid tumours. *Current opinion in genetics & development*. **6**(1), pp.71–75.

Crews, C.M., Alessandrini, A. and Erikson, R.L. 1992. The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product. *Science*. **258**(5081), 478 LP – 480.

Davis, B.N., Hilyard, A.C., Nguyen, P.H., Lagna, G. and Hata, A. 2009. Induction of MicroRNA-221 by Platelet-derived Growth Factor Signaling Is Critical for Modulation of Vascular Smooth Muscle Phenotype * . *Journal of Biological Chemistry*. **284**(6), pp.3728–3738.

Dews, M., Homayouni, A., Yu, D., Murphy, D., Sevignani, C., Wentzel, E., Furth, E.E., Lee, W.M., Enders, G.H., Mendell, J.T. and Thomas-Tikhonenko, A. 2006. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nature Genetics*. **38**(9), pp.1060–1065.

Dong, Y., Zhang, N., Zhao, S., Chen, X., Li, F. and Tao, X. 2019. miR-221-3p and miR-15b-5p promote cell proliferation and invasion by targeting Axin2 in liver cancer. *Oncology letters*. **18**(6), pp.6491–6500.

Downward, J. 1994. The GRB2/Sem-5 adaptor protein. *FEBS letters*. **338**(2), pp.113–117.

Du, P., Wang, L., Sliz, P. and Gregory, R.I. 2015. A Biogenesis Step Upstream of Microprocessor Controls miR-17∼92 Expression. *Cell*. **162**(4), pp.885–899.

Duggal, S., Midha, M.K., Kumar, A. and Rao, K.V.S. 2019. Outlining the Grb2 interactome data and its interacting partners in HEK293 cells in absence and presence of epidermal growth factor. *Data in Brief*. **25**, p.104082.

Dutt, A., Salvesen, H.B., Chen, T.-H., Ramos, A.H., Onofrio, R.C., Hatton, C., Nicoletti, R., Winckler, W., Grewal, R., Hanna, M., Wyhs, N., Ziaugra, L., Richter, D.J., Trovik, J., Engelsen, I.B., Stefansson, I.M., Fennell, T., Cibulskis, K., Zody, M.C., Akslen, L.A., Gabriel, S., Wong, K.-K., Sellers, W.R., Meyerson, M. and Greulich, H. 2008. Drug-sensitive FGFR2 mutations in endometrial carcinoma. *Proceedings of the National Academy of Sciences*. **105**(25), 8713 LP – 8717.

Eck, M.J., Shoelson, S.E. and Harrison, S.C. 1993. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. *Nature*. **362**(6415), pp.87–91.

Elkayam, E., Faehnle, C.R., Morales, M., Sun, J., Li, H. and Joshua-Tor, L. 2017. Multivalent Recruitment of Human Argonaute by GW182. *Molecular cell*. **67**(4), pp.646-658.e3.

Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J. and Joshua-Tor, L. 2012. The Structure of Human Argonaute-2 in Complex with miR-20a. *Cell*. **150**(1), pp.100–110.

EMBL-EBI 2021. Clustal Omega. Available from: https://www.ebi.ac.uk/Tools/msa/clustalw2/.

Eng, J.K., McCormack, A.L. and Yates, J.R. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry*.

5(11), pp.976–989.

Fan, F.-Y., Deng, R., Lai, S.-H., Wen, Q., Zeng, Y., Gao, L., Liu, Y., Kong, P., Zhong, J., Su, Y. and Zhang, X. 2019. Inhibition of microRNA-221-5p induces osteogenic differentiation by directly targeting smad3 in myeloma bone disease mesenchymal stem cells. *Oncology letters*. **18**(6), pp.6536–6544.

Fang, K., Sideri, A., Law, I.K.M., Bakirtzi, K., Polytarchou, C., Iliopoulos, D. and Pothoulakis, C. 2015. Identification of a novel substance P (SP)-neurokinin-1 receptor (NK-1R) microRNA-221-5p inflammatory network in human colonic epithelial cells. *Cellular and molecular gastroenterology and hepatology*. **1**(5), pp.503–515.

Fedorko, M., Juracek, J., Stanik, M., Svoboda, M., Poprach, A., Buchler, T., Pacik, D., Dolezel, J. and Slaby, O. 2017. Detection of let-7 miRNAs in urine supernatant as potential diagnostic approach in non-metastatic clear-cell renal cell carcinoma. *Biochemia medica*. **27**(2), pp.411–417.

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), 1241 LP – 1247.

Fitzpatrick, S.L., LaChance, M.P. and Schultz, G.S. 1984. Characterization of Epidermal Growth Factor Receptor and Action on Human Breast Cancer Cells in Culture. *Cancer Research*. **44**(8), 3442 LP – 3447.

Forman, J.J., Legesse-Miller, A. and Coller, H.A. 2008. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proceedings of the National Academy of Sciences*. **105**(39), 14879 LP – 14884.

Fortian, A. and Sorkin, A. 2014. Live-cell fluorescence imaging reveals high stoichiometry of Grb2 binding to the EGF receptor sustained during endocytosis. *Journal of cell science*. **127**(Pt 2), pp.432–444.

Frank, F., Sonenberg, N. and Nagar, B. 2010. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*. **465**(7299), pp.818–822.

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Carbonell Sala, S., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I.T., García Girón, C., Gonzalez, J.M., Grego, T., Hardy, M., Hourlier, T., Howe, K.L., Hunt, T., Izuogu, O.G., Johnson, R., Martin, F.J., Martínez, L., Mohanan, S., Muir, P., Navarro, F.C.P., Parker, A., Pei, B., Pozo, F., Riera, F.C., Ruffier, M., Schmitt, B.M., Stapleton, E., Suner, M.-M., Sycheva, I., Uszczynska-Ratajczak, B., Wolf, M.Y., Xu, J., Yang, Y.T., Yates, A., Zerbino, D., Zhang, Y., Choudhary, J.S., Gerstein, M., Guigó, R., Hubbard, T.J.P., Kellis, M., Paten, B., Tress, M.L. and Flicek, P. 2021. GENCODE 2021. *Nucleic acids research*. **49**(D1), pp.D916–D923.

Friedman, R.C., Farh, K.K.-H., Burge, C.B. and Bartel, D.P. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research*. **19**(1), pp.92–105.

Furdui, C.M., Lew, E.D., Schlessinger, J. and Anderson, K.S. 2006. Autophosphorylation of FGFR1 Kinase Is Mediated by a Sequential and Precisely Ordered Reaction. *Molecular Cell*. **21**(5), pp.711–717.

Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A. and Corey, D.R. 2014. RNAi Factors Are Present and Active in Human Cell Nuclei. *Cell Reports*. **6**(1),

pp.211–221.

Garcia-Riart, B., Lorda-Diez, C.I., Marin-Llera, J.C., Garcia-Porrero, J.A., Hurle, J.M. and Montero, J.A. 2017. Interdigital tissue remodelling in the embryonic limb involves dynamic regulation of the miRNA profiles. *Journal of Anatomy*. **231**(2), pp.275–286.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. and Bairoch, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*. **31**(13), pp.3784–3788.

Givol, D. and Yayon, A. 1992. Complexity of FGF receptors: genetic basis for structural diversity and functional specificity. *The FASEB Journal*. **6**(15), pp.3362–3369.

Golden, R.J., Chen, B., Li, T., Braun, J., Manjunath, H., Chen, X., Wu, J., Schmid, V., Chang, T.-C., Kopp, F., Ramirez-Martinez, A., Tagliabracci, V.S., Chen, Z.J., Xie, Y. and Mendell, J.T. 2017. An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. *Nature*. **542**(7640), pp.197–202.

Greenidge, P.A., Blommers, M.J.J., Priestle, J.P. and Hunziker, J. 2019. How to Computationally Stack the Deck for Hit-to-Lead Generation: In Silico Molecular Interaction Energy Profiling for de Novo siRNA Guide Strand Surrogate Selection. *Journal of chemical information and modeling*. **59**(5), pp.1897–1908.

Gregory, R.I., Chendrimada, T.P., Cooch, N. and Shiekhattar, R. 2005. Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. *Cell*. **123**(4), pp.631–640.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright,

A.J. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*. **34**(suppl_1), pp.D140–D144.

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. 2007. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Molecular Cell*. **27**(1), pp.91–105.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G. and Mello, C.C. 2001. Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing. *Cell*. **106**(1), pp.23–34.

Gu, S., Jin, L., Zhang, F., Huang, Y., Grimm, D., Rossi, J.J. and Kay, M.A. 2011. Thermodynamic stability of small hairpin RNAs highly influences the loading process of different mammalian Argonautes. *Proceedings of the National Academy of Sciences of the United States of America*. **108**(22), pp.9208–9213.

Guo, L., Zhang, H., Zhao, Y., Yang, S. and Chen, F. 2014. Selected isomiR expression profiles via arm switching? *Gene*. **533**(1), pp.149–155.

Guo, Y., Chen, Y., Ito, H., Watanabe, A., Ge, X., Kodama, T. and Aburatani, H. 2006. Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. *Gene*. **384**, pp.51–61.

Haidar, M., Rchiad, Z., Ansari, H.R., Ben-Rached, F., Tajeri, S., Latre De Late, P., Langsley, G. and Pain, A. 2018. miR-126-5p by direct targeting of JNK-interacting protein-2 (JIP-2) plays a key role in Theileria-infected macrophage virulence. *PLOS Pathogens*. **14**(3), p.e1006942.

Hamilton, A.J. and Baulcombe, D.C. 1999. A Species of Small Antisense RNA

in Posttranscriptional Gene Silencing in Plants. *Science*. **286**(5441), 950 LP – 952.

Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*. **404**(6775), pp.293–296.

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H. and Kim, V.N. 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* . **18**(24), pp.3016–3027.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y. and Takahashi, T. 2005. A Polycistronic MicroRNA Cluster, miR-17-92, Is Overexpressed in Human Lung Cancers and Enhances Cell Proliferation. *Cancer Research*. **65**(21), 9628 LP – 9632.

He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G.A., Liu, C., Franssila, K., Suster, S., Kloos, R.T., Croce, C.M. and de la Chapelle, A. 2005. The role of microRNA genes in papillary thyroid carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*. **102**(52), 19075 LP – 19080.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J. and Hammond, S.M. 2005. A microRNA polycistron as a potential human oncogene. *Nature*. **435**(7043), pp.828–833.

He, Y., Lin, J., Kong, D., Huang, M., Xu, C., Kim, T.-K., Etheridge, A., Luo, Y., Ding, Y. and Wang, K. 2015. Current State of Circulating MicroRNAs as Cancer Biomarkers. *Clinical Chemistry*. **61**(9), pp.1138–1155.

Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Müller-Newen, G. and Schaper, F. 2003. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *The Biochemical journal*. **374**(Pt 1), pp.1–20.

Helsten, T., Elkin, S., Arthur, E., Tomson, B.N., Carter, J. and Kurzrock, R. 2016. The FGFR Landscape in Cancer: Analysis of 4,853 Tumors by Next-Generation Sequencing. *Clinical Cancer Research*. **22**(1), 259 LP – 267.

Heo, I., Ha, M., Lim, J., Yoon, M.-J., Park, J.-E., Kwon, S.C., Chang, H. and Kim, V.N. 2012. Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell*. **151**(3), pp.521–532.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J. and Kim, V.N. 2008. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell*. **32**(2), pp.276–284.

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., Yeom, K.-H., Han, J. and Kim, V.N. 2009. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell*. **138**(4), pp.696–708.

Herbert, K.M., Pimienta, G., DeGregorio, S.J., Alexandrov, A. and Steitz, J.A. 2013. Phosphorylation of DGCR8 Increases Its Intracellular Stability and Induces a Progrowth miRNA Profile. *Cell Reports*. **5**(4), pp.1070–1081.

Hibi, M. and Hirano, T. 1998. Signal transduction through cytokine receptors. *International reviews of immunology*. **17**(1–4), pp.75–102.

Hoelz, A., Janz, J.M., Lawrie, S.D., Corwin, B., Lee, A. and Sakmar, T.P. 2006. Crystal Structure of the SH3 Domain of βPIX in Complex with a High Affinity Peptide from PAK2. *Journal of Molecular Biology*. **358**(2), pp.509–522.

Hong, D.S., Kang, Y.-K., Borad, M., Sachdev, J., Ejadi, S., Lim, H.Y., Brenner, A.J., Park, K., Lee, J.-L., Kim, T.-Y., Shin, S., Becerra, C.R., Falchook, G., Stoudemire, J., Martin, D., Kelnar, K., Peltier, H., Bonato, V., Bader, A.G., Smith, S., Kim, S., O'Neill, V. and Beg, M.S. 2020. Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *British Journal of Cancer*. **122**(11), pp.1630–1637.

Hong, S., Noh, H., Chen, H., Padia, R., Pan, Z.K., Su, S.-B., Jing, Q., Ding, H.-F. and Huang, S. 2013. Signaling by p38 MAPK Stimulates Nuclear Localization of the Microprocessor Component p68 for Processing of Selected Primary MicroRNAs. *Science Signaling*. **6**(266), ra16 LP-ra16.

Horman, S.R., Janas, M.M., Litterst, C., Wang, B., MacRae, I.J., Sever, M.J., Morrissey, D. V, Graves, P., Luo, B., Umesalma, S., Qi, H.H., Miraglia, L.J., Novina, C.D. and Orth, A.P. 2013. Akt-mediated phosphorylation of Argonaute 2 down-regulates cleavage and up-regulates translational repression of microRNA targets. *Molecular cell*. **50**(3), pp.356–367.

Hou, B., Xu, S., Xu, Y., Gao, Q., Zhang, C., Liu, L., Yang, H., Jiang, X. and Che, Y. 2019. Grb2 binds to PTEN and regulates its nuclear translocation to maintain the genomic stability in DNA damage response. *Cell Death & Disease*. **10**(8), p.546.

Hoy, S.M. 2018. Patisiran: First Global Approval. *Drugs*. **78**(15), pp.1625–1631.

Huang, F. and Sorkin, A. 2005. Growth factor receptor binding protein 2-mediated recruitment of the RING domain of Cbl to the epidermal growth factor receptor is essential and sufficient to support receptor endocytosis. *Molecular biology of the cell*. **16**(3), pp.1268–1281.

Hutvágner, G. and Zamore, P.D. 2002. A microRNA in a Multiple-Turnover

RNAi Enzyme Complex. *Science*. **297**(5589), 2056 LP – 2060.

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**(12), 3237 LP – 3241.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T. and Tomari, Y. 2010. Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. *Molecular Cell*. **39**(2), pp.292–299.

Iwasaki, S., Sasaki, H.M., Sakaguchi, Y., Suzuki, T., Tadakuma, H. and Tomari, Y. 2015. Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. *Nature*. **521**(7553), pp.533–536.

Janin, J. 2005. Assessing predictions of protein–protein interaction: The CAPRI experiment. *Protein Science*. **14**(2), pp.278–283.

Jee, D. and Lai, E.C. 2014. Alteration of miRNA activity via context-specific modifications of Argonaute proteins. *Trends in cell biology*. **24**(9), pp.546–553.

Jee, D., Yang, J.-S., Park, S.-M., Farmer, D.T., Wen, J., Chou, T., Chow, A., McManus, M.T., Kharas, M.G. and Lai, E.C. 2018. Dual Strategies for Argonaute2-Mediated Biogenesis of Erythroid miRNAs Underlie Conserved Requirements for Slicing in Mammals. *Molecular Cell*. **69**(2), pp.265-278.e6.

Jiang, X., Jiang, M., Guo, S., Cai, P., Wang, W. and Li, Y. 2020. Promotion of miR-221-5p on the Sensitivity of Gastric Cancer Cells to Cisplatin and Its

Effects on Cell Proliferation and Apoptosis by Regulating DDR1. *OncoTargets and therapy*. **13**, pp.2333–2345.

Jiang, X. and Sorkin, A. 2002. Coordinated Traffic of Grb2 and Ras during Epidermal Growth Factor Receptor Endocytosis Visualized in Living Cells. *Molecular Biology of the Cell*. **13**(5), pp.1522–1535.

Jin, Y., Chen, Z., Liu, X. and Zhou, X. 2013. Evaluating the microRNA targeting sites by luciferase reporter gene assay. *Methods in molecular biology (Clifton, N.J.)*. **936**, pp.117–127.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D. and Slack, F.J. 2005. RAS Is Regulated by the let-7 MicroRNA Family. *Cell*. **120**(5), pp.635–647.

Johnston, M., Geoffroy, M.-C., Sobala, A., Hay, R. and Hutvagner, G. 2010. HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. *Molecular biology of the cell*. **21**(9), pp.1462–1469.

Jones, D.A. and Benjamin, C.W. 1997. Phosphorylation of Growth Factor Receptor Binding Protein-2 by pp60c-srcTyrosine Kinase. *Archives of Biochemistry and Biophysics*. **337**(2), pp.143–148.

Kawahara, Y. and Mieda-Sato, A. 2012. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proceedings of the National Academy of Sciences*. **109**(9), 3347 LP – 3352.

Kawamata, T., Seitz, H. and Tomari, Y. 2009. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nature*

Structural & Molecular Biology. **16**(9), pp.953–960.

Kawata, T., Shevchenko, A., Fukuzawa, M., Jermyn, K.A., Totty, N.F., Zhukovskaya, N. V, Sterling, A.E., Mann, M. and Williams, J.G. 1997. SH2 Signaling in a Lower Eukaryote: A STAT Protein That Regulates Stalk Cell Differentiation in Dictyostelium. *Cell*. **89**(6), pp.909–916.

Kiener, M., Chen, L., Krebs, M., Grosjean, J., Klima, I., Kalogirou, C., Riedmiller, H., Kneitz, B., Thalmann, G.N., Snaar-Jagalska, E., Spahn, M., Kruithof-de Julio, M. and Zoni, E. 2019. miR-221-5p regulates proliferation and migration in human prostate cancer cells and reduces tumor growth in vivo. *BMC Cancer*. **19**(1), p.627.

Kim, Y.-K., Kim, B. and Kim, V.N. 2016. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proceedings of the National Academy of Sciences*. **113**(13), E1881 LP-E1889.

King, I.N., Yartseva, V., Salas, D., Kumar, A., Heidersbach, A., Ando, D.M., Stallings, N.R., Elliott, J.L., Srivastava, D. and Ivey, K.N. 2014. The RNA-binding protein TDP-43 selectively disrupts microRNA-1/206 incorporation into the RNA-induced silencing complex. *The Journal of biological chemistry*. **289**(20), pp.14263–14271.

Kong, Qinglei, Liang, C., Jin, Y., Pan, Y., Tong, D., Kong, Qingcong and Zhou, J. 2019. The lncRNA MIR4435-2HG is upregulated in hepatocellular carcinoma and promotes cancer cell proliferation by upregulating miRNA-487a. *Cellular & Molecular Biology Letters*. **24**(1), p.26.

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.

Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. 2019. miRBase: from microRNA sequences to function. *Nucleic acids research*. **47**(D1), pp.D155–D162.

Kozomara, A. and Griffiths-Jones, S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research*. **39**(suppl_1), pp.D152–D157.

Krell, J., Stebbing, J., Carissimi, C., Dabrowska, A.F., de Giorgio, A., Frampton, A.E., Harding, V., Fulci, V., Macino, G., Colombo, T. and Castellano, L. 2016. TP53 regulates miRNA association with AGO2 to remodel the miRNA–mRNA interaction network. *Genome Research* . **26**(3), pp.331–341.

Kulkarni, R.P., Elmi, A., Alcantara-Adap, E., Hubrack, S., Nader, N., Yu, F., Dib, M., Ramachandran, V., Najafi Shoushtari, H. and Machaca, K. 2019. miRNA-dependent regulation of STIM1 expression in breast cancer. *Scientific reports*. **9**(1), p.13076.

Kumar, M.S., Erkeland, S.J., Pester, R.E., Chen, C.Y., Ebert, M.S., Sharp, P.A. and Jacks, T. 2008. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proceedings of the National Academy of Sciences of the United States of America*. **105**(10), pp.3903–3908.

Kumar, M.S., Lu, J., Mercer, K.L., Golub, T.R. and Jacks, T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genetics*. **39**(5), pp.673–677.

Kwak, P.B. and Tomari, Y. 2012. The N domain of Argonaute drives duplex

unwinding during RISC assembly. *Nature Structural &Amp; Molecular Biology*. **19**, p.145.

Kyriakis, J.M., App, H., Zhang, X., Banerjee, P., Brautigan, D.L., Rapp, U.R. and Avruch, J. 1992. Raf-1 activates MAP kinase-kinase. *Nature*. **358**(6385), pp.417–421.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. 2001. Identification of Novel Genes Coding for Small Expressed RNAs. *Science*. **294**(5543), 853 LP – 858.

Lambertz, I., Nittner, D., Mestdagh, P., Denecker, G., Vandesompele, J., Dyer, M.A. and Marine, J.-C. 2010. Monoallelic but not biallelic loss of Dicer1 promotes tumorigenesis in vivo. *Cell Death & Differentiation*. **17**(4), pp.633–641.

Lamothe, B., Yamada, M., Schaeper, U., Birchmeier, W., Lax, I. and Schlessinger, J. 2004. The Docking Protein Gab1 Is an Essential Component of an Indirect Mechanism for Fibroblast Growth Factor Stimulation of the Phosphatidylinositol 3-Kinase/Akt Antiapoptotic Pathway. *Molecular and Cellular Biology*. **24**(13), 5657 LP – 5666.

Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P.Y., Soll, S.J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M. and Tuschl, T. 2008. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* . **14**(12), pp.2580–2596.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*. **10**(3), p.R25.

Lau, N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P. 2001. An Abundant Class

of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans *Science*. **294**(5543), 858 LP – 862.

Lee, E.J., Gusev, Y., Jiang, J., Nuovo, G.J., Lerner, M.R., Frankel, W.L., Morgan, D.L., Postier, R.G., Brackett, D.J. and Schmittgen, T.D. 2007. Expression profiling identifies microRNA signature in pancreatic cancer. *International Journal of Cancer*. **120**(5), pp.1046–1054.

Lee, H.Y. and Doudna, J.A. 2012. TRBP alters human precursor microRNA processing in vitro. *RNA (New York, N.Y.)*. **18**(11), pp.2012–2019.

Lee, H.Y., Zhou, K., Smith, A.M., Noland, C.L. and Doudna, J.A. 2013. Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. *Nucleic acids research*. **41**(13), pp.6568–6576.

Lee, P.L., Johnson, D.E., Cousens, L.S., Fried, V.A. and Williams, L.T. 1989. Purification and complementary DNA cloning of a receptor for basic fibroblast growth factor. *Science*. **245**(4913), 57 LP – 60.

Lee, R.C. and Ambros, V. 2001. An Extensive Class of Small RNAs in Caenorhabditis elegans *Science*. **294**(5543), 862 LP – 864.

Lee, R.C., Feinbaum, R.L. and Ambros, V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*. **75**(5), pp.843–854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S. and Kim, V.N. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature*. **425**(6956), pp.415–419.

Lee, Y., Hur, I., Park, S.-Y., Kim, Y.-K., Suh, M.R. and Kim, V.N. 2006. The role

of PACT in the RNA silencing pathway. *The EMBO journal*. **25**(3), pp.522–532.

Lee, Y., Jeon, K., Lee, J.-T., Kim, S. and Kim, V.N. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*. **21**(17), pp.4663–4670.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H. and Kim, V.N. 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal*. **23**(20), pp.4051–4060.

Lee, Y.T., de Vasconcellos, J.F., Yuan, J., Byrnes, C., Noh, S.-J., Meier, E.R., Kim, K.S., Rabel, A., Kaushal, M., Muljo, S.A. and Miller, J.L. 2013. LIN28B-mediated expression of fetal hemoglobin and production of fetal-like erythrocytes from adult human erythroblasts ex vivo. *Blood*. **122**(6), pp.1034–1041.

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

Lensink, M.F. and Wodak, S.J. 2013. Docking, scoring, and affinity prediction in CAPRI. *Proteins: Structure, Function, and Bioinformatics*. **81**(12), pp.2082–2095.

Leung, A.K.L., Vyas, S., Rood, J.E., Bhutkar, A., Sharp, P.A. and Chang, P. 2011. Poly(ADP-ribose) Regulates Stress Responses and microRNA Activity in the Cytoplasm. *Molecular cell*. **42**(4), pp.489–499.

Lewis, B.P., Burge, C.B. and Bartel, D.P. 2005. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell*. **120**(1), pp.15–20.

Lewis, B.P., Shih, I., Jones-Rhoades, M.W., Bartel, D.P. and Burge, C.B. 2003. Prediction of Mammalian MicroRNA Targets. *Cell*. **115**(7), pp.787–798.

Lewitzky, M., Kardinal, C., Gehring, N.H., Schmidt, E.K., Konkol, B., Eulitz, M., Birchmeier, W., Schaeper, U. and Feller, S.M. 2001. The C-terminal SH3 domain of the adapter protein Grb2 binds with high affinity to sequences in Gab1 and SLP-76 which lack the SH3-typical P-x-x-P core motif. *Oncogene*. **20**(9), pp.1052–1062.

Li, F., Xu, J.-W., Wang, L., Liu, H., Yan, Y. and Hu, S.-Y. 2018. MicroRNA-221-3p is up-regulated and serves as a potential biomarker in pancreatic cancer. *Artificial cells, nanomedicine, and biotechnology*. **46**(3), pp.482–487.

Li, G., Zhang, H., Ma, H., Qu, S., Xing, Q. and Wang, G. 2021. MiR-221-5p is involved in the regulation of inflammatory responses in acute gouty arthritis by targeting IL-1β. *International Journal of Rheumatic Diseases*. **24**(3), pp.335–340.

Li, L.-Y., Li, E.-M., Wu, Z.-Y., Cao, H.-H., Shen, J.-H., Xu, X.-E., Chen, B., Wu, J.-Y. and Xu, L.-Y. 2014. Overexpression of GRB2 is correlated with lymph node metastasis and poor prognosis in esophageal squamous cell carcinoma. *International journal of clinical and experimental pathology*. **7**(6), pp.3132–3140.

Li, L., Yu, C., Gao, H. and Li, Y. 2010. Argonaute proteins: potential biomarkers for human colon cancer. *BMC cancer*. **10**, p.38.

Li, N., Batzer, A., Daly, R., Yajnik, V., Skolnik, E., Chardin, P., Bar-Sagi, D., Margolis, B. and Schlessinger, J. 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras

signalling. *Nature*. **363**(6424), pp.85–88.

Li, S., Couvillon, A.D., Brasher, B.B. and Van Etten, R.A. 2001. Tyrosine phosphorylation of Grb2 by Bcr/Abl and epidermal growth factor receptor: a novel regulatory mechanism for tyrosine kinase signaling. *The EMBO Journal*. **20**(23), pp.6793–6804.

Liao, R.G., Jung, J., Tchaicha, J., Wilkerson, M.D., Sivachenko, A., Beauchamp, E.M., Liu, Q., Pugh, T.J., Pedamallu, C.S., Hayes, D.N., Gray, N.S., Getz, G., Wong, K.-K., Haddad, R.I., Meyerson, M. and Hammerman, P.S. 2013. Inhibitor-sensitive FGFR2 and FGFR3 mutations in lung squamous cell carcinoma. *Cancer research*. **73**(16), pp.5195–5205.

Liao, Y., Smyth, G.K. and Shi, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*. **30**(7), pp.923–930.

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, S. and Gregory, R.I. 2015. MicroRNA biogenesis pathways in cancer. *Nature Reviews Cancer*. **15**(6), pp.321–333.

Lin, Y.C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van De Peer, Y., Tavernier, J. and Callewaert, N. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications*. **5**(1), pp.1–12.

Liu, J., Carmell, M.A., Rivas, F. V, Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L. and Hannon, G.J. 2004. Argonaute2 Is the

Catalytic Engine of Mammalian RNAi. *Science*. **305**(5689), 1437 LP – 1441.

Liu, J., Li, X., Wang, M., Xiao, G., Yang, G., Wang, H., Li, Y., Sun, X., Qin, S., Du, N., Ren, H. and Pang, Y. 2018. A miR-26a/E2F7 feedback loop contributes to tamoxifen resistance in ER-positive breast cancer. *International journal of oncology*. **53**(4), pp.1601–1612.

Liu, J., Rivas, F. V, Wohlschlegel, J., Yates, J.R., Parker, R. and Hannon, G.J. 2005. A role for the P-body component GW182 in microRNA function. *Nature Cell Biology*. **7**(12), pp.1261–1266.

Liu, P.C.C., Koblish, H., Wu, L., Bowman, K., Diamond, S., DiMatteo, D., Zhang, Y., Hansbury, M., Rupar, M., Wen, X., Collier, P., Feldman, P., Klabe, R., Burke, K.A., Soloviev, M., Gardiner, C., He, X., Volgina, A., Covington, M., Ruggeri, B., Wynn, R., Burn, T.C., Scherle, P., Yeleswaram, S., Yao, W., Huber, R. and Hollis, G. 2020. INCB054828 (pemigatinib), a potent and selective inhibitor of fibroblast growth factor receptors 1, 2, and 3, displays activity against genetically defined tumor models. *PloS one*. **15**(4), p.e0231877.

Liu, Q., Berry, D., Nash, P., Pawson, T., McGlade, C.J. and Li, S.S.-C. 2003. Structural Basis for Specific Binding of the Gads SH3 Domain to an RxxK Motif-Containing SLP-76 Peptide: A Novel Mode of Peptide Recognition. *Molecular Cell*. **11**(2), pp.471–481.

Liu, Q., Chen, X., Novak, M.K., Zhang, S. and Hu, W. 2021. Repressing Ago2 mRNA translation by Trim71 maintains pluripotency through inhibiting let-7 microRNAs T. W. Nilsen & J. L. Manley, eds. *eLife*. **10**, p.e66288.

Liu, S., Sun, X., Wang, M., Hou, Y., Zhan, Y., Jiang, Y., Liu, Zhanjie, Cao, X.,

Chen, P., Liu, Zhi, Chen, X., Tao, Y., Xu, C., Mao, J., Cheng, C., Li, C., Hu, Y., Wang, L., Chin, Y.E., Shi, Y., Siebenlist, U. and Zhang, X. 2014. A microRNA 221– and 222–Mediated Feedback Loop Maintains Constitutive Activation of NFκB and STAT3 in Colorectal Cancer Cells. *Gastroenterology*. **147**(4), pp.847-859.e11.

Liu, S., Wang, Y., Li, W., Yu, S., Wen, Z., Chen, Z. and Lin, F. 2019. miR-221-5p acts as an oncogene and predicts worse survival in patients of renal cell cancer. *Biomedicine & Pharmacotherapy*. **119**, p.109406.

Liu, T., Zhang, H., Fang, J., Yang, Z., Chen, R., Wang, Y., Zhao, X., Ge, S., Yu, J. and Huang, J. 2020. AGO2 phosphorylation by c-Src kinase promotes tumorigenesis. *Neoplasia (New York, N.Y.)*. **22**(3), pp.129–141.

Liu, Y., Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J., Kinch, L.N., Grishin, N. V and Liu, Q. 2009. C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science (New York, N.Y.)*. **325**(5941), pp.750–753.

Liu, Z., Wang, J., Cheng, H., Ke, X., Sun, L., Zhang, Q.C. and Wang, H.-W. 2018. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell*. **173**(5), pp.1191-1203.e12.

Loeb, G.B., Khan, A.A., Canner, D., Hiatt, J.B., Shendure, J., Darnell, R.B., Leslie, C.S. and Rudensky, A.Y. 2012. Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Molecular cell*. **48**(5), pp.760–770.

Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V. and Ding, Y. 2007. Potent effect of target structure on microRNA function. *Nature Structural & Molecular Biology*. **14**(4), pp.287–294.

Lopez-Orozco, J., Pare, J.M., Holme, A.L., Chaulk, S.G., Fahlman, R.P. and Hobman, T.C. 2015. Functional analyses of phosphorylation events in human Argonaute 2. *RNA (New York, N.Y.)*. **21**(12), pp.2030–2038.

Love, M.I., Huber, W. and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. **15**(12), p.550.

Lowenstein, E.J., Daly, R.J., Batzer, A.G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E.Y., Bar-Sagi, D. and Schlessinger, J. 1992. The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell*. **70**(3), pp.431–442.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., Downing, J.R., Jacks, T., Horvitz, H.R. and Golub, T.R. 2005. MicroRNA expression profiles classify human cancers. *Nature*. **435**(7043), pp.834–838.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E. and Kutay, U. 2004. Nuclear Export of MicroRNA Precursors. *Science*. **303**(5654), 95 LP – 98.

Lv, Y., Yang, H., Ma, X. and Wu, G. 2019. Strand-specific miR-28-3p and miR-28-5p have differential effects on nasopharyngeal cancer cells proliferation, apoptosis, migration and invasion. *Cancer Cell International*. **19**(1), p.187.

Ma, L., Zhao, Q., Chen, W. and Zhang, Y. 2018. Oncogene Lin28B increases chemosensitivity of colon cancer cells in a let-7-independent manner. *Oncology letters*. **15**(5), pp.6975–6981.

MacRae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D. and Doudna, J.A. 2006. Structural Basis for Double-Stranded RNA Processing by Dicer. *Science*. **311**(5758), 195 LP – 198.

Maignan, S., Guilloteau, J.P., Fromage, N., Arnoux, B., Becquart, J. and Ducruix, A. 1995. Crystal structure of the mammalian Grb2 adaptor. *Science*. **268**(5208), 291 LP – 293.

Markham, A. 2019. Erdafitinib: First Global Approval. *Drugs*. **79**(9), pp.1017–1021.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. **17**(1), pp.10–12.

Martinez, N.J. and Gregory, R.I. 2013. Argonaute2 expression is post-transcriptionally coupled to microRNA abundance. *RNA* . **19**(5), pp.605–612.

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), 1537 LP – 1539.

Mayr, C., Hemann, M.T. and Bartel, D.P. 2007. Disrupting the Pairing Between let-7 and Hmga2 Enhances Oncogenic Transformation. *Science*. **315**(5818), 1576 LP – 1579.

Mazumder, A., Bose, M., Chakraborty, A., Chakrabarti, S. and Bhattacharyya, S.N. 2013. A transient reversal of miRNA-mediated repression controls macrophage activation. *EMBO reports*. **14**(11), 1008 LP – 1016.

McDonald, C.B., Seldeen, K.L., Deegan, B.J., Lewis, M.S. and Farooq, A. 2008. Grb2 adaptor undergoes conformational change upon dimerization. *Archives of Biochemistry and Biophysics*. **475**(1), pp.25–35.

McKenzie, A.J., Hoshino, D., Hong, N.H., Cha, D.J., Franklin, J.L., Coffey, R.J., Patton, J.G. and Weaver, A.M. 2016. KRAS-MEK Signaling Controls Ago2

Sorting into Exosomes. *Cell reports*. **15**(5), pp.978–987.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. 2004. Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Molecular Cell*. **15**(2), pp.185–197.

Michlewski, G. and Cáceres, J.F. 2010. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nature structural & molecular biology*. **17**(8), pp.1011–1018.

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*. **89**(1), 246 LP – 250.

Mitra, P. and Thanabalu, T. 2017. Myogenic differentiation depends on the interplay of Grb2 and N-WASP. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. **1864**(3), pp.487–497.

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), 5068 LP – 5078.

Mohammadi, M., Schlessinger, J. and Hubbard, S.R. 1996. Structure of the FGF Receptor Tyrosine Kinase Domain Reveals a Novel Autoinhibitory Mechanism. *Cell*. **86**(4), pp.577–587.

Moran, M.F., Koch, C.A., Anderson, D., Ellis, C., England, L., Martin, G.S. and Pawson, T. 1990. Src homology region 2 domains direct protein-protein

interactions in signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*. **87**(21), pp.8622–8626.

Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T. and Hatada, I. 2007. One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. *Genomics*. **89**(6), pp.687–696.

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I. and Proudfoot, N.J. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nature Structural & Molecular Biology*. **15**(9), pp.902–909.

Mujahid, S., Nielsen, H.C. and Volpe, M. V 2013. MiR-221 and miR-130a Regulate Lung Airway and Vascular Development. *PLOS ONE*. **8**(2), p.e55911.

Müller, C., Schäfer, I., Luhmann, H.J. and White, R. 2015. Oligodendroglial Argonaute protein Ago2 associates with molecules of the Mbp mRNA localization machinery and is a downstream target of Fyn kinase . *Frontiers in Cellular Neuroscience* . **9**, p.328.

Müller, M., Fazi, F. and Ciaudo, C. 2020. Argonaute Proteins: From Structure to Function in Development and Pathological Cell Fate Determination . *Frontiers in Cell and Developmental Biology* . **7**, p.360.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. and Saraste, M. 1992. Crystal structure of a Src-homology 3 (SH3) domain. *Nature*. **359**(6398), pp.851–855.

Nakanishi, K., Weinberg, D.E., Bartel, D.P. and Patel, D.J. 2012. Structure of yeast Argonaute with guide RNA. *Nature*. **486**(7403), pp.368–374.

Nam, S., Kim, B., Shin, S. and Lee, S. 2008. miRGator: an integrated system for functional annotation of microRNAs. *Nucleic acids research*. **36**(Database issue), pp.D159–D164.

Naruse, K., Matsuura-Suzuki, E., Watanabe, M., Iwasaki, S. and Tomari, Y. 2018. In vitro reconstitution of chaperone-mediated human RISC assembly. *RNA (New York, N.Y.)*. **24**(1), pp.6–11.

NEB 2020a. NEBaseChanger. Available from: https://nebasechanger.neb.com.

NEB 2020b. NEBCloner. Available from: https://nebcloner.neb.com.

NEB 2020c. Tm Calculator. Available from: https://tmcalculator.neb.com.

Newman, M.A., Thomson, J.M. and Hammond, S.M. 2008. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* . **14**(8), pp.1539–1549.

Nishihara, T., Zekri, L., Braun, J.E. and Izaurralde, E. 2013. miRISC recruits decapping factors to miRNA targets to enhance their degradation. *Nucleic Acids Research*. **41**(18), pp.8692–8705.

Nurcombe, V., Smart, C.E., Chipperfield, H., Cool, S.M., Boilly, B. and Hondermarck, H. 2000. The Proliferative and Migratory Activities of Breast Cancer Cells Can Be Differentially Regulated by Heparan Sulfates *. *Journal of Biological Chemistry*. **275**(39), pp.30009–30018.

Okamura, K., Ishizuka, A., Siomi, H. and Siomi, M.C. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes & Development* . **18**(14), pp.1655–1666.

Okamura, K., Phillips, M.D., Tyler, D.M., Duan, H., Chou, Y. and Lai, E.C. 2008. The regulatory activity of microRNA* species has substantial influence on

microRNA and 3’ UTR evolution. *Nature structural & molecular biology*. **15**(4), pp.354–363.

Olive, V., Bennett, M.J., Walker, J.C., Ma, C., Jiang, I., Cordon-Cardo, C., Li, Q.-J., Lowe, S.W., Hannon, G.J. and He, L. 2009. miR-19 is a key oncogenic component of mir-17-92. *Genes & development*. **23**(24), pp.2839–2849.

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., Yayon, A., Flanagan, J.G., Svahn, C.M., Levi, E. and Leder, P. 1992. Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Molecular and Cellular Biology*. **12**(1), 240 LP – 247.

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.

Osborne, C.K., Hamilton, B., Nover, M. and Ziegler, J. 1981. Antagonism between epidermal growth factor and phorbol ester tumor promoters in human breast cancer cells. *The Journal of clinical investigation*. **67**(4), pp.943–951.

Ota, H., Sakurai, M., Gupta, R., Valente, L., Wulff, B.-E., Ariyoshi, K., Iizasa, H., Davuluri, R. V and Nishikura, K. 2013. ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing. *Cell*. **153**(3), pp.575-589Ota, H., Sakurai, M., Gupta, R., Valente, L.

Papachristou, D.J., Korpetinou, A., Giannopoulou, E., Antonacopoulou, A.G.,

Papadaki, H., Grivas, P., Scopa, C.D. and Kalofonos, H.P. 2011. Expression of the ribonucleases Drosha, Dicer, and Ago2 in colorectal carcinomas. *Virchows Archiv*. **459**(4), p.431.

Paradis-Isler, N. and Boehm, J. 2018. NMDA receptor-dependent dephosphorylation of serine 387 in Argonaute 2 increases its degradation and affects dendritic spine density and maturation. *Journal of Biological Chemistry* .

Park, J.H. and Shin, C. 2015. Slicer-independent mechanism drives small-RNA strand separation during human RISC assembly. *Nucleic acids research*. **43**(19), pp.9418–9433.

Park, M.S., Araya-Secchi, R., Brackbill, J.A., Phan, H.-D., Kehling, A.C., Abd El-Wahab, E.W., Dayeh, D.M., Sotomayor, M. and Nakanishi, K. 2019. Multidomain Convergence of Argonaute during RISC Assembly Correlates with the Formation of Internal Water Clusters. *Molecular Cell*. **75**(4), pp.725-740.e6.

Park, M.S., Phan, H.-D., Busch, F., Hinckley, S.H., Brackbill, J.A., Wysocki, V.H. and Nakanishi, K. 2017. Human Argonaute3 has slicer activity. *Nucleic Acids Research*. **45**(20), pp.11867–11877.

Paroo, Z., Ye, X., Chen, S. and Liu, Q. 2009. Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell*. **139**(1), pp.112–122.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E. and Ruvkun, G. 2000. Conservation of the sequence and

temporal expression of let-7 heterochronic regulatory RNA. *Nature*. **408**(6808), pp.86–89.

Patranabis, S. and Bhattacharyya, S.N. 2016. Phosphorylation of Ago2 and Subsequent Inactivation of let-7a RNP-Specific MicroRNAs Control Differentiation of Mammalian Sympathetic Neurons. *Molecular and Cellular Biology* . **36**(8), pp.1260–1271.

Peng, J., Mo, R., Ma, J. and Fan, J. 2015. let-7b and let-7c are determinants of intrinsic chemoresistance in renal cell carcinoma. *World journal of surgical oncology*. **13**, p.175.

Perera, T.P.S., Jovcheva, E., Mevellec, L., Vialard, J., De Lange, D., Verhulst, T., Paulussen, C., Van De Ven, K., King, P., Freyne, E., Rees, D.C., Squires, M., Saxty, G., Page, M., Murray, C.W., Gilissen, R., Ward, G., Thompson, N.T., Newell, D.R., Cheng, N., Xie, L., Yang, J., Platero, S.J., Karkera, J.D., Moy, C., Angibaud, P., Laquerre, S. and Lorenzi, M. V 2017. Discovery and Pharmacological Characterization of JNJ-42756493 (Erdafitinib), a Functionally Selective Small-Molecule FGFR Family Inhibitor. *Molecular Cancer Therapeutics*. **16**(6), 1010 LP – 1020.

Perneger, T. V 1998. What’s wrong with Bonferroni adjustments. *BMJ (Clinical research ed.)*. **316**(7139), pp.1236–1238.

Pfaff, J., Hennig, J., Herzog, F., Aebersold, R., Sattler, M., Niessing, D. and Meister, G. 2013. Structural features of Argonaute–GW182 protein interactions. *Proceedings of the National Academy of Sciences*. **110**(40), E3770 LP-E3779.

Pilotte, J., Dupont-Versteegden, E.E. and Vanderklish, P.W. 2011. Widespread regulation of miRNA biogenesis at the Dicer step by the cold-inducible

RNA-binding protein, RBM3. *PloS one*. **6**(12), pp.e28446–e28446.

Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D. and Gregory, R.I. 2011. Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell*. **147**(5), pp.1066–1079.

Piskounova, E., Viswanathan, S.R., Janas, M., LaPierre, R.J., Daley, G.Q., Sliz, P. and Gregory, R.I. 2008. Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. *The Journal of biological chemistry*. **283**(31), pp.21310–21314.

Plotnikov, A.N., Schlessinger, J., Hubbard, S.R. and Mohammadi, M. 1999. Structural Basis for FGF Receptor Dimerization and Activation. *Cell*. **98**(5), pp.641–650.

Poliseno, L., Tuccoli, A., Mariani, L., Evangelista, M., Citti, L., Woods, K., Mercatanti, A., Hammond, S. and Rainaldi, G. 2006. MicroRNAs modulate the angiogenic properties of HUVECs. *Blood*. **108**(9), pp.3068–3071.

Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., MacDonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P. and Stoffel, M. 2004. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature*. **432**(7014), pp.226–230.

Protein Atlas 2020. Human Protein Atlas. Available from: https://www.proteinatlas.org/about.

Qi, H.H., Ongusaha, P.P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Yujiang, Lee, S.W., Peng, J. and Shi, Yang 2008. Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature*. **455**, p.421.

Quévillon Huberdeau, M., Zeitler, D.M., Hauptmann, J., Bruckmann, A., Fressigné, L., Danner, J., Piquet, S., Strieder, N., Engelmann, J.C., Jannot, G., Deutzmann, R., Simard, M.J. and Meister, G. 2017. Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing in vivo *The EMBO Journal*. **36**(14), 2088 LP – 2106.

Rajgor, D., Sanderson, T.M., Amici, M., Collingridge, G.L. and Hanley, J.G. 2018. NMDAR-dependent Argonaute 2 phosphorylation regulates miRNA activity and dendritic spine plasticity. *The EMBO Journal*.

Rapraeger, A.C., Krufka, A. and Olwin, B.B. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science*. **252**(5013), 1705 LP – 1708.

Rau, F., Freyermuth, F., Fugier, C., Villemin, J.-P., Fischer, M.-C., Jost, B., Dembele, D., Gourdon, G., Nicole, A., Duboc, D., Wahbi, K., Day, J.W., Fujimura, H., Takahashi, M.P., Auboeuf, D., Dreumont, N., Furling, D. and Charlet-Berguerand, N. 2011. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nature Structural & Molecular Biology*. **18**(7), pp.840–845.

Raveche, E.S., Salerno, E., Scaglione, B.J., Manohar, V., Abbasi, F., Lin, Y.-C., Fredrickson, T., Landgraf, P., Ramachandra, S., Huppi, K., Toro, J.R., Zenger, V.E., Metcalf, R.A. and Marti, G.E. 2007. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood*. **109**(12), pp.5079–5086.

Reichholf, B., Herzog, V.A., Fasching, N., Manzenreither, R.A., Sowemimo, I. and Ameres, S.L. 2019. Time-Resolved Small RNA Sequencing Unravels

the Molecular Principles of MicroRNA Homeostasis. *Molecular Cell*. **75**(4), pp.756-768.e7.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. and Ruvkun, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*. **403**(6772), pp.901–906.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B. and Bartel, D.P. 2002. MicroRNAs in plants. *Genes & Development* . **16**(13), pp.1616–1626.

Ren, R., Mayer, B.J., Cicchetti, P. and Baltimore, D. 1993. Identification of a ten-amino acid proline-rich SH3 binding site. *Science*. **259**(5098), 1157 LP – 1161.

Rosenfeld, N., Aharonov, R., Meiri, E., Rosenwald, S., Spector, Y., Zepeniuk, M., Benjamin, H., Shabes, N., Tabak, S., Levy, A., Lebanony, D., Goren, Y., Silberschein, E., Targan, N., Ben-Ari, A., Gilad, S., Sion-Vardy, N., Tobar, A., Feinmesser, M., Kharenko, O., Nativ, O., Nass, D., Perelman, M., Yosepovich, A., Shalmon, B., Polak-Charcon, S., Fridman, E., Avniel, A., Bentwich, I., Bentwich, Z., Cohen, D., Chajut, A. and Barshack, I. 2008. MicroRNAs accurately identify cancer tissue origin. *Nature Biotechnology*. **26**(4), pp.462–469.

Rowe, R.G., Wang, L.D., Coma, S., Han, A., Mathieu, R., Pearson, D.S., Ross, S., Sousa, P., Nguyen, P.T., Rodriguez, A., Wagers, A.J. and Daley, G.Q. 2016. Developmental regulation of myeloerythroid progenitor function by the Lin28b–let-7–Hmga2 axis. *Journal of Experimental Medicine*. **213**(8), pp.1497–1512.

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. *Nature*. **363**(6424), pp.83–85.

Ruby, J.G., Jan, C.H. and Bartel, D.P. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature*. **448**(7149), pp.83–86.

Rüdel, S., Wang, Y., Lenobel, R., Körner, R., Hsiao, H.-H., Urlaub, H., Patel, D. and Meister, G. 2010. Phosphorylation of human Argonaute proteins affects small RNA binding. *Nucleic Acids Research*. **39**(6), pp.2330–2343.

Rüdel, S., Wang, Y., Lenobel, R., Körner, R., Hsiao, H.-H., Urlaub, H., Patel, D. and Meister, G. 2011. Phosphorylation of human Argonaute proteins affects small RNA binding. *Nucleic Acids Research*. **39**(6), pp.2330–2343.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R. and Wulczyn, F.G. 2008. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nature Cell Biology*. **10**(8), pp.987–993.

Saetrom, P., Heale, B.S.E., Snøve Jr, O., Aagaard, L., Alluin, J. and Rossi, J.J. 2007. Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic acids research*. **35**(7), pp.2333–2342.

Sahin, U., Lapaquette, P., Andrieux, A., Faure, G. and Dejean, A. 2014. Sumoylation of human argonaute 2 at lysine-402 regulates its stability. *PLoS ONE*.

Sanjana, N.E., Shalem, O. and Zhang, F. 2014. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods*. **11**(8), pp.783–784.

Santiveri, C.M., Borroto, A., Simón, L., Rico, M., Alarcón, B. and Jiménez, M.A. 2009. Interaction between the N-terminal SH3 domain of Nckα and CD3ε-

derived peptides: Non-canonical and canonical recognition motifs. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1794**(1), pp.110–117.

Schiering, N., Casale, E., Caccia, P., Giordano, P. and Battistini, C. 2000. Dimer Formation through Domain Swapping in the Crystal Structure of the Grb2-SH2−Ac-pYVNV Complex. *Biochemistry*. **39**(44), pp.13376–13382.

Schirle, N.T. and MacRae, I.J. 2012. The Crystal Structure of Human Argonaute2. *Science (New York, N.Y.)*. **336**(6084), pp.1037–1040.

Schirle, N.T., Sheu-Gruttadauria, J. and MacRae, I.J. 2014. Structural Basis for microRNA Targeting. *Science (New York, N.Y.)*. **346**(6209), pp.608–613.

Scott, L.J. 2020. Givosiran: First Approval. *Drugs*. **80**(3), pp.335–339.

Sen, G.L. and Blau, H.M. 2005. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nature Cell Biology*. **7**(6), pp.633–636.

Seo, G.J., Kincaid, R.P., Phanaksri, T., Burke, J.M., Pare, J.M., Cox, J.E., Hsiang, T.-Y., Krug, R.M. and Sullivan, C.S. 2013. Reciprocal Inhibition between Intracellular Antiviral Signaling and the RNAi Machinery in Mammalian Cells. *Cell Host & Microbe*. **14**(4), pp.435–445.

Seth, P., Hsieh, P.N., Jamal, S., Wang, L., Gygi, S.P., Jain, M.K., Coller, J. and Stamler, J.S. 2019. Regulation of MicroRNA Machinery and Development by Interspecies S-Nitrosylation. *Cell*. **176**(5), pp.1014-1025.e12.

Shankar, S., Pitchiaya, S., Malik, R., Kothari, V., Hosono, Y., Yocum, A.K., Gundlapalli, H., White, Y., Firestone, A., Cao, X., Dhanasekaran, S.M., Stuckey, J.A., Bollag, G., Shannon, K., Walter, N.G., Kumar-Sinha, C. and

Chinnaiyan, A.M. 2016. KRAS Engages AGO2 to Enhance Cellular Transformation. *Cell Reports*. **14**(6), pp.1448–1461.

Shankar, S., Tien, J.C.-Y., Siebenaler, R.F., Dommeti, V.L., Zelenka-Wang, S., Juckette, K.M., Xu, A., Mody, M., Goodrum, A., Tsaloff, G., Apel, I.J., Wang, L., Siddiqui, J., Shi, J., Kumar-Sinha, C. and Chinnaiyan, A.M. 2017. An Essential Role for Argonaute 2 in EGFR-KRAS Signaling in Pancreatic Cancer Development. *bioRxiv*., p.227264.

Shao, N., Ma, G., Zhang, J. and Zhu, W. 2018. miR-221-5p enhances cell proliferation and metastasis through post-transcriptional regulation of SOCS1 in human prostate cancer. *BMC Urology*. **18**(1), p.14.

Shell, S., Park, S.-M., Radjabi, A.R., Schickel, R., Kistner, E.O., Jewell, D.A., Feig, C., Lengyel, E. and Peter, M.E. 2007. Let-7 expression defines two differentiation stages of cancer. *Proceedings of the National Academy of Sciences*. **104**(27), 11400 LP – 11405.

Shen, J., Xia, W., Khotskaya, Y.B., Huo, L., Nakanishi, K., Lim, S.-O., Du, Y., Wang, Y., Chang, W.-C., Chen, C.-H., Hsu, J.L., Wu, Y., Lam, Y.C., James, B.P., Liu, X., Liu, C.-G., Patel, D.J. and Hung, M.-C. 2013. EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature*. **497**(7449), pp.383–387.

Shen, S., Huang, K., Wu, Y., Ma, Y., Wang, J., Qin, F. and Ma, J. 2017. A miR-135b-TAZ positive feedback loop promotes epithelial–mesenchymal transition (EMT) and tumorigenesis in osteosarcoma. *Cancer Letters*. **407**, pp.32–44.

Sheu-Gruttadauria, J. and MacRae, I.J. 2018. Phase Transitions in the Assembly and Function of Human miRISC. *Cell*. **173**(4), pp.946-957.e16.

Shi, J., Zhang, Y., Jin, N., Li, Y., Wu, S. and Xu, L. 2017. MicroRNA-221-3p Plays an Oncogenic Role in Gastric Carcinoma by Inhibiting PTEN Expression. *Oncology research*. **25**(4), pp.523–536.

Shostak, K. and Chariot, A. 2015. EGFR and NF-κB: partners in cancer. *Trends in Molecular Medicine*. **21**(6), pp.385–393.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*. **7**, p.539.

Singh, G., Popli, S., Hari, Y., Malhotra, P., Mukherjee, S. and Bhatnagar, R.K. 2009. Suppression of RNA silencing by Flock house virus B2 protein is mediated through its interaction with the PAZ domain of Dicer. *The FASEB Journal*. **23**(6), pp.1845–1857.

Smibert, P., Yang, J.-S., Azzam, G., Liu, J.-L. and Lai, E.C. 2013. Homeostatic control of Argonaute stability by microRNA availability. *Nature Structural & Molecular Biology*. **20**(7), pp.789–795.

Soares-Silva, M., Diniz, F.F., Gomes, G.N. and Bahia, D. 2016. The Mitogen-Activated Protein Kinase (MAPK) Pathway: Role in Immune Evasion by Trypanosomatids . *Frontiers in Microbiology* . **7**, p.183.

Song, J.-J., Smith, S.K., Hannon, G.J. and Joshua-Tor, L. 2004. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science*. **305**(5689), 1434 LP – 1437.

Spaccarotella, E., Pellegrino, E., Ferracin, M., Ferreri, C., Cuccuru, G., Liu, C., Iqbal, J., Cantarella, D., Taulli, R., Provero, P., Di Cunto, F., Medico, E., Negrini, M., Chan, W.C., Inghirami, G. and Piva, R. 2014. STAT3-mediated

activation of microRNA cluster 17~92 promotes proliferation and survival of ALK-positive anaplastic large cell lymphoma. *Haematologica*. **99**(1), pp.116–124.

Stokoe, D., Macdonald, S.G., Cadwallader, K., Symons, M. and Hancock, J.F. 1994. Activation of Raf as a result of recruitment to the plasma membrane. *Science*. **264**(5164), 1463 LP – 1467.

Su, H., Trombly, M.I., Chen, J. and Wang, X. 2009. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes & development*. **23**(3), pp.304–317.

Su, W.-C.S., Kitagawa, M., Xue, N., Xie, B., Garofalo, S., Cho, J., Deng, C., Horton, W.A. and Fu, X.-Y. 1997. Activation of Statl by mutant fibroblast growth-factor receptor in thanatophoric dysplasia type II dwarfism. *Nature*. **386**(6622), pp.288–292.

Sun, H.-L., Cui, R., Zhou, J., Teng, K., Hsiao, Y.-H., Nakanishi, K., Fassan, M., Luo, Z., Shi, G., Tili, E., Kutay, H., Lovat, F., Vicentini, C., Huang, H.-L., Wang, S.-W., Kim, T., Zanesi, N., Jeon, Y.-J., Lee, T.J., Guh, J.-H., Hung, M.-C., Ghoshal, K., Teng, C.-M., Peng, Y. and Croce, C.M. 2016. ERK Activation Globally Downregulates miRNAs through Phosphorylating Exportin-5. *Cancer Cell*. **30**(5), pp.723–736.

Suzuki, H.I., Arase, M., Matsuyama, H., Choi, Y.L., Ueno, T., Mano, H., Sugimoto, K. and Miyazono, K. 2011. MCPIP1 Ribonuclease Antagonizes Dicer and Terminates MicroRNA Biogenesis through Precursor MicroRNA Degradation. *Molecular Cell*. **44**(3), pp.424–436.

Swarbrick, S., Wragg, N., Ghosh, S. and Stolzing, A. 2019. Systematic Review of miRNA as Biomarkers in Alzheimer's Disease. *Molecular neurobiology*.

56(9), pp.6156–6167.

Sylvestre, Y., De Guire, V., Querido, E., Mukhopadhyay, U.K., Bourdeau, V., Major, F., Ferbeyre, G. and Chartrand, P. 2007. An E2F/miR-20a Autoregulatory Feedback Loop. *Journal of Biological Chemistry* . **282**(4), pp.2135–2143.

Szybowska, P., Kostas, M., Wesche, J., Wiedlocha, A. and Haugsten, E.M. 2019. Cancer Mutations in FGFR2 Prevent a Negative Feedback Loop Mediated by the ERK1/2 Pathway. *Cells*. **8**(6), p.518.

Tahbaz, N., Carmichael, J.B. and Hobman, T.C. 2001. GERp95 Belongs to a Family of Signal-transducing Proteins and Requires Hsp90 Activity for Stability and Golgi Localization. *Journal of Biological Chemistry* . **276**(46), pp.43294–43299.

Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W. and Hobman, T.C. 2004. Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Reports*. **5**(2), pp.189–194.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T. and Takahashi, T. 2004. Reduced Expression of the let-7 MicroRNAs in Human Lung Cancers in Association with Shortened Postoperative Survival. *Cancer Research*. **64**(11), 3753 LP – 3756.

Tamimi, Y., van der Poel, H.G., Denyn, M.M., Umbas, R., Karthaus, H.F., Debruyne, F.M. and Schalken, J.A. 1993. Increased expression of high mobility group protein I(Y) in high grade prostatic cancer determined by in situ hybridization. *Cancer research*. **53**(22), pp.5512–5516.

Tao, K., Yang, J., Guo, Z., Hu, Y., Sheng, H., Gao, H. and Yu, H. 2014. Prognostic value of miR-221-3p, miR-342-3p and miR-491-5p expression in colon cancer. *American journal of translational research*. **6**(4), pp.391–401.

Tarallo, R., Giurato, G., Bruno, G., Ravo, M., Rizzo, F., Salvati, A., Ricciardi, L., Marchese, G., Cordella, A., Rocco, T., Gigantino, V., Pierri, B., Cimmino, G., Milanesi, L., Ambrosino, C., Nyman, T., Nassa, G. and Weisz, A. 2017. The nuclear receptor ERβ engages AGO2 in regulation of gene transcription, RNA splicing and RISC loading. *Genome Biology*. **18**(189).

TargetScan 2020. TargetScan Human 7.2. Available from: http://www.targetscan.org/vert_72/.

Taylor, B.S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B.S., Arora, V.K., Kaushik, P., Cerami, E., Reva, B., Antipin, Y., Mitsiades, N., Landers, T., Dolgalev, I., Major, J.E., Wilson, M., Socci, N.D., Lash, A.E., Heguy, A., Eastham, J.A., Scher, H.I., Reuter, V.E., Scardino, P.T., Sander, C., Sawyers, C.L. and Gerald, W.L. 2010. Integrative Genomic Profiling of Human Prostate Cancer. *Cancer Cell*. **18**(1), pp.11–22.

Taylor, D.W., Ma, E., Shigematsu, H., Cianfrocco, M.A., Noland, C.L., Nagayama, K., Nogales, E., Doudna, J.A. and Wang, H.-W. 2013. Substrate-specific structural rearrangements of human Dicer. *Nature structural & molecular biology*. **20**(6), pp.662–670.

TCGA_Research_Network 2020. The Cancer Genome Atlas. Available from: https://www.cancer.gov/tcga.

Thomson, J.M., Newman, M., Parker, J.S., Morin-Kensicki, E.M., Wright, T. and Hammond, S.M. 2006. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development* . **20**(16),

pp.2202–2207.

Thornton, J.E., Chang, H.-M., Piskounova, E. and Gregory, R.I. 2012. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA (New York, N.Y.)*. **18**(10), pp.1875–1885.

Tian, Y., Simanshu, D.K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V.N. and Patel, D.J. 2014. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular cell*. **53**(4), pp.606–616.

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 &Amp; Molecular Biology*. **21**, p.180.

Timsah, Z., Berrout, J., Suraokar, M., Behrens, C., Song, J., Lee, J.J., Ivan, C., Gagea, M., Shires, M., Hu, X., Vallien, C., Kingsley, C. V, Wistuba, I. and Ladbury, J.E. 2015. Expression pattern of FGFR2, Grb2 and Plcγ1 acts as a novel prognostic marker of recurrence recurrence-free survival in lung adenocarcinoma. *American journal of cancer research*. **5**(10), pp.3135–3148.

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R. and Rosenfeld, M.G. 2009. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*. **459**(7249), pp.1010–1014.

Treiber, T., Treiber, N. and Meister, G. 2019. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature Reviews Molecular Cell Biology*. **20**(1), pp.5–20.

Trissal, M.C., Wong, T.N., Yao, J.-C., Ramaswamy, R., Kuo, I., Baty, J., Sun, Y., Jih, G., Parikh, N., Berrien-Elliott, M.M., Fehniger, T.A., Ley, T.J., Maillard, I., Reddy, P.R. and Link, D.C. 2018. MIR142 Loss-of-Function Mutations Derepress ASH1L to Increase HOXA Gene Expression and Promote Leukemogenesis. *Cancer Research*. **78**(13), 3510 LP – 3521.

Tsai, K.-W., Leung, C.-M., Lo, Y.-H., Chen, T.-W., Chan, W.-C., Yu, S.-Y., Tu, Y.-T., Lam, H.-C., Li, S.-C., Ger, L.-P., Liu, W.-S. and Chang, H.-T. 2016. Arm Selection Preference of MicroRNA-193a Varies in Breast Cancer. *Scientific Reports*. **6**(1), p.28176.

Tsang, J., Zhu, J. and van Oudenaarden, A. 2007. MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Molecular cell*. **26**(5), pp.753–767.

Tsuboyama, K., Tadakuma, H. and Tomari, Y. 2018. Conformational Activation of Argonaute by Distinct yet Coordinated Actions of the Hsp70 and Hsp90 Chaperone Systems. *Molecular Cell*. **70**(4), pp.722-729.e4.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H. and Tomari, Y. 2011. Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nature*

Structural & Molecular Biology. **18**(10), pp.1153–1158.

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

Ueda, Y., Ando, T., Nanjo, S., Ushijima, T. and Sugiyama, T. 2014. DNA methylation of microRNA-124a is a potential risk marker of colitis-associated cancer in patients with ulcerative colitis. *Digestive diseases and sciences*. **59**(10), pp.2444–2451.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A.-K., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J. and Pontén, F. 2015. Tissue-based map of the human proteome. *Science*. **347**(6220), p.1260419.

Urbach, A., Yermalovich, A., Zhang, J., Spina, C.S., Zhu, H., Perez-Atayde, A.R., Shukrun, R., Charlton, J., Sebire, N., Mifsud, W., Dekel, B., Pritchard-Jones, K. and Daley, G.Q. 2014. Lin28 sustains early renal progenitors and induces Wilms tumor. *Genes & development*. **28**(9), pp.971–982.

Vaksman, O., Hetland, T.E., Trope', C.G., Reich, R. and Davidson, B. 2012. Argonaute, Dicer, and Drosha are up-regulated along tumor progression in serous ovarian carcinoma. *Human Pathology*. **43**(11), pp.2062–2069.

Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J.J. and Lötvall, J.O. 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature Cell Biology*. **9**(6),

pp.654–659.

Verbeek, B.S., Adriaansen-slot, S.S., Rijksen, G. and Vroom, T.M. 1997. Grb2 overexpression in nuclei and cytoplasm of human breast cells: a histochemical and biochemical study of normal and neoplastic mammary tissue specimens. *The Journal of Pathology*. **183**(2), pp.195–203.

Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J. and Khvorova, A. 2005. The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA (New York, N.Y.)*. **11**(5), pp.674–682.

Vidal, M., Goudreau, N., Cornille, F., Cussac, D., Gincel, E. and Garbay, C. 1999. Molecular and cellular analysis of Grb2 SH3 domain mutants: interaction with Sos and dynamin11Edited by A. R. Fersht. *Journal of Molecular Biology*. **290**(3), pp.717–730.

Viswanathan, S.R., Daley, G.Q. and Gregory, R.I. 2008. Selective Blockade of MicroRNA Processing by Lin28. *Science*. **320**(5872), 97 LP – 100.

Volinia, S., Calin, G.A., Liu, C.-G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R.L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C.C. and Croce, C.M. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences*. **103**(7), 2257 LP – 2261.

Völler, D., Reinders, J., Meister, G. and Bosserhoff, A.-K. 2013. Strong reduction of AGO2 expression in melanoma and cellular consequences. *British Journal of Cancer*. **109**(12), pp.3116–3124.

Vota, D.M., Maltaneri, R.E., Wenker, S.D., Nesse, A.B. and Vittori, D.C. 2013. Differential Erythropoietin Action upon Cells Induced to Eryptosis by

Different Agents. *Cell Biochemistry and Biophysics*. **65**(2), pp.145–157.

Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. and Kuriyan, J. 1993. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: Crystal structures of the complexed and peptide-free forms. *Cell*. **72**(5), pp.779–790.

Wang, H.-W., Noland, C., Siridechadilok, B., Taylor, D.W., Ma, E., Felderer, K., Doudna, J.A. and Nogales, E. 2009. Structural insights into RNA processing by the human RISC-loading complex. *Nature structural & molecular biology*. **16**(11), pp.1148–1153.

Wang, Q. and Goldstein, M. 2016. Small RNAs Recruit Chromatin-Modifying Enzymes MMSET and Tip60 to Reconfigure Damaged DNA upon Double-Strand Break and Facilitate Repair. *Cancer Research*. **76**(7), 1904 LP – 1915.

Wang, X., Lu, X., Zhang, T., Wen, C., Shi, M., Tang, X., Chen, H., Peng, C., Li, H., Fang, Y., Deng, X. and Shen, B. 2016. mir-329 restricts tumor growth by targeting grb2 in pancreatic cancer. *Oncotarget*. **7**(16), pp.21441–21453.

Wang, Y., Chen, J., Yang, W., Mo, F., Senz, J., Yap, D., Anglesio, M.S., Gilks, B., Morin, G.B. and Huntsman, D.G. 2015. The Oncogenic Roles of DICER1 RNase IIIb Domain Mutations in Ovarian Sertoli-Leydig Cell Tumors. *Neoplasia*. **17**(8), pp.650–660.

Wang, Z., Lin, S., Li, J.J., Xu, Z., Yao, H., Zhu, X., Xie, D., Shen, Z., Sze, J., Li, K., Lu, G., Chan, D.T.-M., Poon, W.S., Kung, H. and Lin, M.C. 2011. MYC protein inhibits transcription of the microRNA cluster MC-let-7a-1~let-7d via noncanonical E-box. *The Journal of biological chemistry*. **286**(46),

pp.39703–39714.

Warner, M.J., Bridge, K.S., Hewitson, J.P., Hodgkinson, M.R., Heyam, A., Massa, B.C., Haslam, J.C., Chatzifrangkeskou, M., Evans, G.J.O., Plevin, M.J., Sharp, T. V and Lagos, D. 2016. S6K2-mediated regulation of TRBP as a determinant of miRNA expression in human primary lymphatic endothelial cells. *Nucleic Acids Research*. **44**(20), pp.9942–9955.

Watanabe, T., Shinohara, N., Moriya, K., Sazawa, A., Kobayashi, Y., Ogiso, Y., Takiguchi, M., Yasuda, J., Koyanagi, T., Kuzumaki, N. and Hashimoto, A. 2000. Significance of the Grb2 and Son of Sevenless (Sos) Proteins in Human Bladder Cancer Cell Lines. *IUBMB Life*. **49**(4), pp.317–320.

Weber, J.A., Baxter, D.H., Zhang, S., Huang, D.Y., How Huang, K., Jen Lee, M., Galas, D.J. and Wang, K. 2010. The MicroRNA Spectrum in 12 Body Fluids. *Clinical Chemistry*. **56**(11), pp.1733–1741.

Wei, W.-F., Zhou, C.-F., Wu, X.-G., He, L.-N., Wu, L.-F., Chen, X.-J., Yan, R.-M., Zhong, M., Yu, Y.-H., Liang, L. and Wang, W. 2017. MicroRNA-221-3p, a TWIST2 target, promotes cervical cancer metastasis by directly targeting THBS2. *Cell death & disease*. **8**(12), p.3220.

Wei, Z. and Liu, H.T. 2002. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research*. **12**(1), pp.9–18.

Wiggins, J.F., Ruffino, L., Kelnar, K., Omotola, M., Patrawala, L., Brown, D. and Bader, A.G. 2010. Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34. *Cancer research*. **70**(14), pp.5923–5930.

Wightman, B., Ha, I. and Ruvkun, G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*. **75**(5), pp.855–862.

Wong, N.W., Chen, Y., Chen, S. and Wang, X. 2017. OncomiR: an online resource for exploring pan-cancer microRNA dysregulation. *Bioinformatics*. **34**(4), pp.713–715.

Wu, C., So, J., Davis-Dusenbery, B.N., Qi, H.H., Bloch, D.B., Shi, Y., Lagna, G. and Hata, A. 2011. Hypoxia Potentiates MicroRNA-Mediated Gene Silencing through Posttranslational Modification of Argonaute2 . *Molecular and Cellular Biology*. **31**(23), pp.4760–4774.

Wu, S.-L., Fu, X., Huang, J., Jia, T.-T., Zong, F.-Y., Mu, S.-R., Zhu, H., Yan, Y., Qiu, S., Wu, Q., Yan, W., Peng, Y., Chen, J. and Hui, J. 2015. Genome-wide analysis of YB-1-RNA interactions reveals a novel role of YB-1 in miRNA processing in glioblastoma multiforme. *Nucleic acids research*. **43**(17), pp.8516–8528.

Xhemalce, B., Robson, S.C. and Kouzarides, T. 2012. Human RNA methyltransferase BCDIN3D regulates microRNA processing. *Cell*. **151**(2), pp.278–288.

Xiaojun, L., Yunhui, C., Shuo, Z., Ying, L., Jian, Y. and Chunxiang, Z. 2009. A Necessary Role of miR-221 and miR-222 in Vascular Smooth Muscle Cell Proliferation and Neointimal Hyperplasia. *Circulation Research*. **104**(4), pp.476–487.

Xie, M., Li, M., Vilborg, A., Lee, N., Shu, M.-D., Yartseva, V., Šestan, N. and Steitz, J.A. 2013. Mammalian 5′-Capped MicroRNA Precursors that Generate a Single MicroRNA. *Cell*. **155**(7), pp.1568–1580.

Xu, W., San Lucas, A., Wang, Z. and Liu, Y. 2014. Identifying microRNA targets in different gene regions. *BMC Bioinformatics*. **15**(7), p.S4.

Yajaira, S., Carlos, F.-H., S., P.J. and C., S.W. 2007. Dicer Dependent

MicroRNAs Regulate Gene Expression and Functions in Human Endothelial Cells. *Circulation Research*. **100**(8), pp.1164–1173.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L. and Zhou, M.-M. 2003. Structure and conserved RNA binding of the PAZ domain. *Nature*. **426**(6965), pp.469–474.

Yang, D., Lu, H. and Erickson, J.W. 2000. Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in Drosophila embryos. *Current Biology*. **10**(19), pp.1191–1200.

Yang, F.-Q., Huang, J.-H., Liu, M., Yang, F.-P., Li, W., Wang, G.-C., Che, J.-P. and Zheng, J.-H. 2013. Argonaute 2 is up-regulated in tissues of urothelial carcinoma of bladder. *International journal of clinical and experimental pathology*. **7**(1), pp.340–347.

Yang, J.-S., Maurin, T., Robine, N., Rasmussen, K.D., Jeffrey, K.L., Chandwani, R., Papapetrou, E.P., Sadelain, M., O'Carroll, D. and Lai, E.C. 2010. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences*. **107**(34), 15163 LP – 15168.

Yang, M., Haase, A.D., Huang, F.-K., Coulis, G., Rivera, K.D., Dickinson, B.C., Chang, C.J., Pappin, D.J., Neubert, T.A., Hannon, G.J., Boivin, B. and Tonks, N.K. 2014. Dephosphorylation of Tyrosine 393 in Argonaute 2 by Protein Tyrosine Phosphatase 1B Regulates Gene Silencing in Oncogenic RAS-Induced Senescence. *Molecular cell*. **55**(5), pp.782–790.

Yang, Q., Li, W., She, H., Dou, J., Duong, D.M., Du, Y., Yang, S.-H., Seyfried, N.T., Fu, H., Gao, G. and Mao, Z. 2015. Stress Induces p38 MAPK-

Mediated Phosphorylation and Inhibition of Drosha-Dependent Cell Survival. *Molecular Cell*. **57**(4), pp.721–734.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. and Madden, T.L. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. **13**(1), p.134.

Ye, Y.-B., Lin, J.-Y., Chen, Q., Liu, F., Chen, H.-J., Li, J.-Y., Liu, W.-Q., Garbay, C. and Vidal, M. 2008. The cytotoxicity of a Grb2-SH3 inhibitor in Bcr-Abl positive K562 cells. *Biochemical Pharmacology*. **75**(11), pp.2080–2091.

Yekta, S., Shih, I. and Bartel, D.P. 2004. MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science*. **304**(5670), 594 LP – 596.

Yi, R., Qin, Y., Macara, I.G. and Cullen, B.R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development* . **17**(24), pp.3011–3016.

Yip, S.S., Crew, A.J., Gee, J.M.W., Hui, R., Blamey, R.W., Robertson, J.F.R., Nicholson, R.I., Sutherland, R.L. and Daly, R.J. 2000. Up-regulation of the protein tyrosine phosphatase SHP-1 in human breast cancer and correlation with GRB2 expression. *International Journal of Cancer*. **88**(3), pp.363–368.

Yoon, J.-H., Jo, M.H., White, E.J.F., De, S., Hafner, M., Zucconi, B.E., Abdelmohsen, K., Martindale, J.L., Yang, X., Wood, W.H., Shin, Y.M., Song, J.-J., Tuschl, T., Becker, K.G., Wilson, G.M., Hohng, S. and Gorospe, M. 2015. AUF1 promotes let-7b loading on Argonaute 2. *Genes & Development* . **29**(15), pp.1599–1604.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J. and Song, E. 2007. let-7 Regulates Self

Renewal and Tumorigenicity of Breast Cancer Cells. *Cell*. **131**(6), pp.1109–1123.

Yu, G.-Z., Chen, Y., Long, Y.-Q., Dong, D., Mu, X.-L. and Wang, xi 2008. New insight into the key proteins and pathways involved in the metastasis of colorectal carcinoma. *Oncology reports*. **19**, pp.1191–1204.

Yu, G.Z., Chen, Y. and Wang, J.J. 2009. Overexpression of Grb2/HER2 signaling in Chinese gastric cancer: their relationship with clinicopathological parameters and prognostic significance. *Journal of cancer research and clinical oncology*. **135**(10), pp.1331–1339.

Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. and Schrelber, S.L. 1994. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell*. **76**(5), pp.933–945.

Yuan, K., Xie, K., Fox, J., Zeng, H., Gao, H., Huang, C. and Wu, M. 2013. Decreased Levels of miR-224 and the Passenger Strand of miR-221 Increase MBD2, Suppressing Maspin and Promoting Colorectal Tumor Growth and Metastasis in Mice. *Gastroenterology*. **145**(4), pp.853-864.e9.

Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. 2000. RNAi: Double-Stranded RNA Directs the ATP-Dependent Cleavage of mRNA at 21 to 23 Nucleotide Intervals. *Cell*. **101**(1), pp.25–33.

Zeng, Y., Sankala, H., Zhang, X. and Graves, P.R. 2008. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochemical Journal*. **413**(3), 429 LP – 436.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E. and Filipowicz, W. 2004. Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell*. **118**(1), pp.57–68.

Zhang, H., Wang, Y., Dou, J., Guo, Y., He, J., Li, L., Liu, X., Chen, R., Deng, R., Huang, J., Xie, R., Zhao, X. and Yu, J. 2019. Acetylation of AGO2 promotes cancer progression by increasing oncogenic miR-19b biogenesis. *Oncogene*. **38**(9), pp.1410–1431.

Zhang, J., Fan, X., Wang, C., Liu, B., Li, Q. and Zhou, X. 2013. Up-regulation of Ago2 expression in gastric carcinoma. *Medical Oncology*. **30**(3), p.628.

Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M.S., Giannakakis, A., Liang, S., Naylor, T.L., Barchetti, A., Ward, M.R., Yao, G., Medina, A., O'Brien-Jenkins, A., Katsaros, D., Hatzigeorgiou, A., Gimotty, P.A., Weber, B.L. and Coukos, G. 2006. microRNAs exhibit high frequency genomic alterations in human cancer. *Proceedings of the National Academy of Sciences*. **103**(24), 9136 LP – 9141.

Zhang, T., Ma, J. and Cao, X. 2003. Grb2 regulates Stat3 activation negatively in epidermal growth factor signalling. *The Biochemical journal*. **376**(Pt 2), pp.457–464.

Zhang, T. and Zhang, D. 2017. Integrating omics data and protein interaction networks to prioritize driver genes in cancer. *Oncotarget*. **8**(35), pp.58050–58060.

Zhang, X., Wang, S., Wang, H., Cao, J., Huang, X., Chen, Z., Xu, P., Sun, G., Xu, J., Lv, J. and Xu, Z. 2019. Circular RNA circNRIP1 acts as a microRNA-149-5p sponge to promote gastric cancer progression via the AKT1/mTOR pathway. *Molecular Cancer*. **18**(1), p.20.

Zhang, Y., Li, Z., Yang, M., Wang, D., Yu, L., Guo, C., Guo, X. and Lin, N. 2014. Identification of GRB2 and GAB1 Coexpression as an Unfavorable Prognostic Factor for Hepatocellular Carcinoma by a Combination of

Expression Profile and Network Analysis. *PLOS ONE*. **8**(12), p.e85170.

Zhang, Z., Pi, J., Zou, D., Wang, X., Xu, J., Yu, S., Zhang, T., Li, F., Zhang, X., Zhao, H., Wang, F., Wang, D., Ma, Y. and Yu, J. 2019. microRNA arm-imbalance in part from complementary targets mediated decay promotes gastric cancer progression. *Nature Communications*. **10**(1), p.4397.

Zhao, M., Luo, R., Liu, Y., Gao, L., Fu, Z., Fu, Q., Luo, X., Chen, Y., Deng, X., Liang, Z., Li, X., Cheng, C., Liu, Z. and Fang, W. 2016. miR-3188 regulates nasopharyngeal carcinoma proliferation and chemosensitivity through a FOXO1-modulated positive feedback loop with mTOR-p-PI3K/AKT-c-JUN. *Nature communications*. **7**, p.11309.

Zhou, Q., Ren, J., Hou, J., Wang, G., Ju, L., Xiao, Y. and Gong, Y. 2019. Co-expression network analysis identified candidate biomarkers in association with progression and prognosis of breast cancer. *Journal of Cancer Research and Clinical Oncology*. **145**(9), pp.2383–2396.

Zhou, R., Hu, G., Gong, A.-Y. and Chen, X.-M. 2010. Binding of NF-kappaB p65 subunit to the promoter elements is involved in LPS-induced transactivation of miRNA genes in human biliary epithelial cells. *Nucleic acids research*. **38**(10), pp.3222–3232.

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.

Zhou, Y., Chen, L., Barlogie, B., Stephens, O., Wu, X., Williams, D.R., Cartron, M.-A., van Rhee, F., Nair, B., Waheed, S., Pineda-Roman, M., Alsayed, Y.,

Anaissie, E. and Shaughnessy, J.D. 2010. High-risk myeloma is associated with global elevation of miRNAs and overexpression of EIF2C2/AGO2 *Proceedings of the National Academy of Sciences*. **107**(17), 7904 LP – 7909.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J. and Bonvin, A.M.J.J. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology*. **428**(4), pp.720–725.

List of Abbreviations

AGO	Argonaute
AGO2	Argonaute 2
A-loop	activation loop
antimiR	antisense oligomir
CD	circular dichroism
cDNA	complementary DNA
circRNA	circular RNA
CSH3	Grb2 C-terminal SH3 domain
c-SRC	proto-oncogene tyrosine-protein kinase Src
DAG	diacylglycerol
DGCR8	DiGeorge syndrome critical region 8
dsRBD	double stranded RNA-binding domain
EGFR	epidermal growth factor receptor
ESR2	oestrogen receptor β
ERK1/2	extracellular-regulated kinase 1/2
F1	monoclonal HEK293T expressing FGFR2
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FGFR2	fibroblast growth factor receptor 2
FRET	Förster resonance energy transfer
G1	monoclonal Grb2-depleted HEK293T
G3F1	monoclonal Grb2-depleted HEK293T expressing FGFR2
GFP-Puro^R	GFP and puromycin resistance insert
GRB2	growth factor receptor-bound protein 2
gRNA	guide RNA

GTP	guanosine triphosphate
HR	homologous recombination
HSC70	heat shock cognate protein 70
HSP90	heat shock protein 90
IP	immunoprecipitation
IP3	inositol triphosphate
ITC	isothermal titration calorimetry
kd	knockdown
K_D	equilibrium dissociation constant
kDa	kilodalton
LHA	left hand homologous arm
lncRNA	long non-coding RNA
MAPK	mitogen-activated protein kinase
MAPKK	MAPK kinase
MAPKKK	MAPK kinase kinase
mDG	microRNA downregulated in G1
MID	middle domain
miRISC	microRNA-induced silencing complex
miRNA	microRNA
mRNA	messenger RNA
MST	microscale thermophoresis
mUG	microRNA upregulated in G1
NC	negative control
NSH3	Grb2 N-terminal SH3 domain
nt	nucleotides
oncomir	oncogenic miRNA
P bodies	processing bodies

PACT	protein-activator of PKR
PAZ	Piwi Argonaute and Zwille
pG1	polyclonal Grb2-depleted HEK293T 1
pG2	polyclonal Grb2-depleted HEK293T 2
phospho-site	phosphorylation site
PI	protease inhibitor
PIWI	P-body-induced wimpy testes
PLCγ	phospholipase C
pre-miRNA	precursor microRNA
pri-miRNA	primary microRNA
PRM	proline-rich motif
PTEN	phosphatase and tensin homologue
PTM	post-translational modification
PTP1B	protein tyrosine phosphatase 1B
Puro^R	puromycin resistance gene
PxxP	consensus proline-rich motif
pY393	phosphorylated AGO2 tyrosine 393
pY99	phosphorylated tyrosine antibody
qPCR	quantitative polymerase chain reaction
RCC	renal clear cell carcinoma
RHA	right hand homologous arm
RIP	RNA-immunoprecipitation
RLC	miRISC-loading complex
RTK	receptor tyrosine kinase
Scr	HEK293T scramble control
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SEC	size exclusion chromatography
SH2	SRC homology 2 domain
SH3	SRC homology 3 domain
SHP2	SH2 domain-containing protein tyrosine phosphatase 2
siRNA	small interfering RNA
SOS	son of sevenless
STAT	signal transducer and activator of transcription
TCGA	The Cancer Genome Atlas
TF	transcription factor
TRBP	trans-activation responsive RNA-binding protein
TRNC6	trinucleotide repeat containing 6

Appendix A:

Genotyping of GRB2-knockdown cell lines

Cell lines were produced using CRISPR/Cas9 knockdown of GRB2 in HEK293T cells. A system was used whereby GFP and Puro^R genes were inserted into the start of the GRB2 gene, thus knocking out *GRB2*. Polyclonal cell lines, pG1 and pG2, were generated through use of gRNAs which targeted different sequences at the start of *GRB2* (Chapter 3, 3.1). Monoclonal cell lines, G1, were then selected (Chapter 3, 3.2). A Scr control cell line was made using a scrambled gRNA which should not target any region of the human genome.

A.1 Sequencing of GRB2 knockdown cell lines

pG1, pG2, Scr and G1 genomes were sequenced to check for a) correct insertion of GFP-Puro^R genes into the start of *GRB2*, and, b) mutations near the gRNA targeting site. To confirm insertion of GFP-Puro^R, PCR fragments spanning the RHA and LHA were generated. Sequences were then aligned to the human genome or to *GRB2* with the expected insert, using Blast (Altschul et al., 1990; Blast, 2020). Mutations near the gRNA target site were identified by sequencing of PCR products spanning this region and alignment to WT *GRB2*.

pG1 - GRB2 EU332843.1

```
G1RHA   51     TTGCTCTTGACTGTGCAGTATGGAAGCCAGCATTCTCAGAAGACCAGAAAGGATGGCGGT  110
               |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
GRB2    13452  TTGCTCTTGACTGTGCAGTAGGGAAGCCAGCATTCTCAGAAGACCAGAAAGGATGGCGGT  13511

G1RHA   111    GGGGAAGGGAGGGAGACAGGCCTGTTTAACCGAGAAGAGACAGGTGCCATCCCCAGGGTG  170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13512  GGGGAAGGGAGGGAGACAGGCCTGTTTAACCGAGAAGAGACAGGTGCCATCCCCAGGGTG  13571

G1RHA   171    AGGCCTTCCTTTCAGTGGACCGTATTTGGTGCTGGAGAAGAGCCTGTGAGGAGTAAAGGA  230
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13572  AGGCCTTCCTTTCAGTGGACCGTATTTGGTGCTGGAGAAGAGCCTGTGAGGAGTAAAGGA  13631

G1RHA   231    GGAAAGGCGGAGACTGAGTGGGACTCCAGCCGGTTCCTGCACGCAGTGCACACCCAGCGG  290
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13632  GGAAAGGCGGAGACTGAGTGGGACTCCAGCCGGTTCCTGCACGCAGTGCACACCCAGCGG  13691

G1RHA   291    GAAGGAAGAGGAAGTGTGTTCAGGAATAGCTCAGGCAGCTGTGCCATTGGCCGGCAGGCC  350
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13692  GAAGGAAGAGGAAGTGTGTTCAGGAATAGCTCAGGCAGCTGTGCCATTGGCCGGCAGGCC  13751

G1RHA   351    TGTCAGTCGCCTGGGGGCCGAGGCTGAAAAGAGGAAGAAGTCGGC-gggggggctgttg  409
               |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
GRB2    13752  TGTCAGTCGCCTGGGGGCCGAGGCTGAAAAGAGGAAGAAGTCGGCGGGGGGGGGCTGTTG  13811

G1RHA   410    cctggaagggagggaggagggctgagccaggccgagccaggggcgtggggCTGGGACA  469
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13812  CCTGGAAGGGAGGGAGGAGGGCTGAGCCAGGCCGAGCCAGGGGGCGTGGGGGCTGGGACA  13871

G1RHA   470    GCAGGCCGATTGGCAGGGCGGGGTTGCCTTCAATCCTGGCCACTGCTCTTAATCGTCCCT  529
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13872  GCAGGCCGATTGGCAGGGCGGGGTTGCCTTCAATCCTGGCCACTGCTCTTAATCGTCCCT  13931

G1RHA   530    CTCCTGCTTCAGGGTGGCATTGTGTGTCCCAGAGTGCCGGATCGAGTCCCAGAAGAGAGG  589
               |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
GRB2    13932  CTCCTGCTTCAGGGTGGCATTGTGTGTCCCAGAGTGCCGGAGCGAGTCCCAGAAGAGAGG  13991

G1RHA   590    CGAGGCTAAGCCCAGAGCGCTGGGTTGCTTCAGCAGGGAAGACTCCCTTCCCCCTGCTTC  649
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13992  CGAGGCTAAGCCCAGAGCGCTGGGTTGCTTCAGCAGGGAAGACTCCCTTCCCCCTGCTTC  14051

G1RHA   650    AGGCTGCTGAGCACTGAGCAGCGCTCAGAA  679
               ||||||||||||||||||||||||||||||
GRB2    14052  AGGCTGCTGAGCACTGAGCAGCGCTCAGAA  14081
```

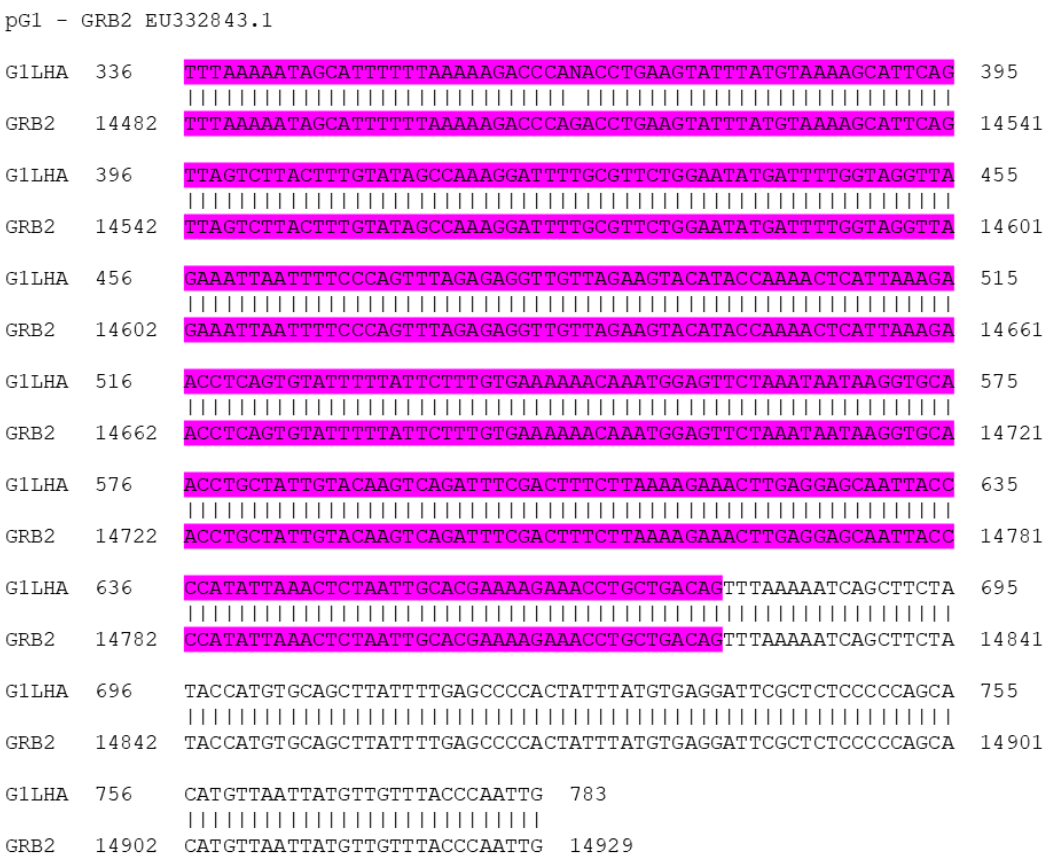


Figure A.1.1 Alignment of the right and left homologous arms of modified *GRB2* in pG1 cells with *GRB2* in the human genome

The pG1 genome was extracted and the homologous arms adjacent to the GFP-Puro^R insert were amplified by PCR. PCR fragments were sequenced and aligned to human nucleotide sequences using Blast (Altschul et al., 1990; Blast, 2020). Alignments were produced against *GRB2*. Cyan and purple highlights indicate RHA and LHA respectively. Green highlight denotes sequencing matching the GFP-Puro^R insert.

```
pG2 RHA - GRB2 EU332843.1

G2RHA  44     TGCTTTGNTTGCTCTTGACTGTGCNGTANGGAANCCNGCATTCTCNNANNACCANANNGG  103
              ||||||| |||||||||||||||| ||| |||| || ||||||||  |  |||| |  ||
GRB2   13444  TGCTTTGCTTGCTCTTGACTGTGCAGTAGGGAAGCCAGCATTCTCAGAAGACCAGAAAGG  13503

G2RHA  104    ATGGNGGTGGGGAAGGNAGGGAGACAGGCCTGTTTAACCGANAANANACAGGTGCCATCC  163
              |||| |||||||||| ||||||||||||||||||||||| || | ||||||||||||
GRB2   13504  ATGGCGGTGGGGAAGGGAGGGAGACAGGCCTGTTTAACCGAGAAGAGACAGGTGCCATCC  13563

G2RHA  164    CCAGGGTGAGGCCTTCCTTTCNNTGNACCGTATTTGGTGCTGGANAANANCCTGNGAGGA  223
              ||||||||||||||||||||  || |||||||||||||||||| || | |||| |||||
GRB2   13564  CCAGGGTGAGGCCTTCCTTTCAGTGGACCGTATTTGGTGCTGGAGAAGAGCCTGTGAGGA  13623

G2RHA  224    GTAAAGGANGAAAGGCNNANACTGANNGNGACTCCCGCCGGNTCCTGCNCNCNNTGNNCN  283
              |||||||| |||||||  | |||||  | |||||| ||||| |||||| | |  ||  |
GRB2   13624  GTAAAGGAGGAAAGGCGGAGACTGAGTGGGACTCCAGCCGGTTCCTGCACGCAGTGCACA  13683

G2RHA  284    CCCANCGGGAAGGAANAGGAAGTGTGTTCAGGAATAGCTCANGCACCTGTGCCATTGGCC  343
              |||| |||||||||| |||||||||||||||||||||||| ||| ||||||||||||||
GRB2   13684  CCCAGCGGGAAGGAAGAGGAAGTGTGTTCAGGAATAGCTCAGGCAGCTGTGCCATTGGCC  13743

G2RHA  344    GGCAGGCCTGTCAGTCNCCTGGGGGCCGAGGCTGANAAGAGGAANAANTCNGCGGNGGG  402
              ||||||||||||||| |||||||||||||||||| |||||||| || || |||| |||
GRB2   13744  GGCAGGCCTGTCAGTCGCCTGGGGGCCGAGGCTGAAAAGAGGAAGAAGTCGGCGGGGGG  13802


pG2 LHA - PNMA5 XM_017029253.1

G2LHA  49    GGGGGCCGGGCCCATAGCCCCTANCNGGTACAGGANAC-NNNNCGNNNANNCNCTANNCT  107
             |||||||||||||||||||| | ||||||||| ||     ||    |  | |||  ||
PNMA5  1944  GGGGGCCGGGCCCATAGCCCCTAACAGGTACAGGAGACACAAGCGGGGAACCACTACACT  2003

G2LHA  108   TNNNNCATGNGGNGCCNGGNCTTCNNATGGGGACCCCNNCNAANCNNGANCTGAANAANG  167
             |    |||| || ||| || ||||  ||||||||||  | || |  |  ||||| || |
PNMA5  2004  TAGCACATGGGGTGCCTGGGCTTCAGATGGGGACCCCAGCGAAGCAGGGGCTGAAGAAGG  2063

G2LHA  168   TGTGGCTGGGGGGCTCTGGCCTGGTCCTCTGGACNCCCCCNAATATCTACCCTGTGNACTT  227
             |||||||||||||||||||||||||||||||| ||||| ||||||||||||||| ||||
PNMA5  2064  TGTGGCTGGGGGGCTCTGGCCTGGTCCTCTGGACACCCCCAAATATCTACCCTGTGGACTT  2123

G2LHA  228   TGANCTGANCATGCCCNCTGGNACCCANGCCACNNGNGACCTGGAGGAGCCTGCCTGGNG  287
             ||| |||| ||||||| |||| ||||| |||||  | |||||||||||||||||||| |
PNMA5  2124  TGAGCTGAGCATGCCCACTGGCACCCAGGCCACGTGGGACCTGGAGGAGCCTGCCTGGGG  2183

G2LHA  288   CCTGCCTCTGCCNNAANGAGCCAGGGCTGCTGAGGAAGANCTGTGGGGCAGAANTGAANC  347
             ||||||||||||  || |||||||||||||||||||| |||||||||||||| |||| |
PNMA5  2184  CCTGCCTCTGCCAGAAGGAGCCAGGGCTGCTGAGGAAGAGCTGTGGGGCAGAAGTGAAGC  2243

G2LHA  348   CCCGCANGGGAGGCCNCNNGGTCCAGCCTGNCTTTANGATCATCTACACTGCATGAGGGG  407
             |||||| |||||||| |  |||||||||| ||||| ||||||||||||||||||||||
PNMA5  2244  CCCGCAGGGGAGGCCACAGGGTCCAGCCTGTCTTTAGGATCATCTACACTGCATGAGGGG  2303

G2LHA  408   AGCCCCANGAAGGCAGCATTCATANNCCCCTGCACCAGTGAGGAANANAAAGAGGANGAN  467
             ||||||| ||||||||||||||  |||||||||||||||||| | |||||||| ||
PNMA5  2304  AGCCCCAGGAAGGCAGCATTCATAAGCCCCTGCACCAGTGAGGAAGAGAAAGAGGAGGAG  2363

G1LHA  468   AATTAANAANANGATGATCNGNAGGANGAANAGGANAATAANGANGAGCANGA  520
             |||||| || | ||||||| | |||| ||| |||| ||||| || ||||| ||
PNMA5  2364  AATTAAGAAGAGGATGATCGGGAGGAGGAAGAGGAGAATAAGGAGGAGCAGGA  2416
```

Figure A.1.2 Alignment of the left and right homologous arms of pG2 cells to the human genome

The homologous arms flanking inserted DNA in pG2 cells were amplified by PCR and sequenced. Sequences were aligned against human nucleotide sequences using Blast (Altschul et al., 1990; Blast, 2020). The RHA (cyan) aligned to the expected region at the start of *GRB2*, whereas the product generated by the LHA primers matched *PNMA5*.

```
Scr RHA - Chromosome 5 AC237231.4

ScrRHA  40     GANGAANACACTGGCGGTTGGACTTCTAGAGGAAAGCACCGACAGGCACGGGCACACCAN  99
               || ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Chr.5   29996  GAGGAACACACAGGCGGTTGGACTTCTAGAGGAAAGCACCGACAGGCACGGGCACACCAG  30055

ScrRHA  100    CAGGCCACCAACTGGCAGAACGGCGTGGAGTT  131
               |||||||||||||||||||| |||||||||
Chr.5   30056  CAGGCCACCAACTGGCAGAACGACGTGGAGTT  30087


Scr LHA - PNMA5 XM_017029253.1

ScrLHA  569    GCACCCCATGTGCTAAGTGTANTGGTTCCCCGCTTGTGTCTCCTGTACCTGTTAGGGGCT  628
               |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
PNMA5   2018   GCACCCCATGTGCTAAGTGTAGTGGTTCCCCGCTTGTGTCTCCTGTACCTGTTAGGGGCT  1959

ScrLHA  629    ATGGGCCCGGCCCCCACTGAACCAGATGCTATTTGGTCTTGGTTGGTGAAGTCCCTTCCA  688
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PNMA5   1958   ATGGGCCCGGCCCCCACTGAACCAGATGCTATTTGGTCTTGGTTGGTGAAGTCCCTTCCA  1899

ScrLHA  689    GCTGGTGGAAAGTAAGGGGGCTTGATCTCGACC  721
               |||||||||||||||||||||||||||||||||
PNMA5   1898   GCTGGTGGAAAGTAAGGGGGCTTGATCTCGACC  1866
```

Figure A.1.3 Alignment of DNA from Scr cells against the human genome

Genomic DNA from Scr cells was amplified using primers designed to amplify the homologous arms flanking GFP-Puro^{R} inserted into the start of *GRB2*. PCR fragments were sequenced and aligned to human nucleotide sequences using Blast (Altschul et al., 1990; Blast, 2020). Alignments of the RHA were made against chromosome 5 and the LHA to *PNMA5.*

```
G1 - GRB2  EU332843.1

G1RHA   1      AGCAGCTCGAGANCAAAGAGTTTTGCTTTGCTTGCTCTTGACTGTGCAGTAGGGAAGCCA  60
               ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13421  AGCAGCTCGAGACCAAAGAGTTTTGCTTTGCTTGCTCTTGACTGTGCAGTAGGGAAGCCA  13480

G1RHA   61     GCATTCTCAGAAGACCAGAAAGGATGGCGGTGGGGAAGGGAGGGAGACAGGCCTGTTTAA  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13481  GCATTCTCAGAAGACCAGAAAGGATGGCGGTGGGGAAGGGAGGGAGACAGGCCTGTTTAA  13540

G1RHA   121    CCGAGAAGAGACAGGTGCCATCCCCAGGGTGAGGCCTTCCTTTCAGTGGACCGTATTTGG  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13541  CCGAGAAGAGACAGGTGCCATCCCCAGGGTGAGGCCTTCCTTTCAGTGGACCGTATTTGG  13600

G1RHA   181    TGCTGGAGAAGAGCCTGTGAGGAGTAAAGGAGGAAAGGCGGAGACTGAGTGGGACTCCAG  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13601  TGCTGGAGAAGAGCCTGTGAGGAGTAAAGGAGGAAAGGCGGAGACTGAGTGGGACTCCAG  13660

G1RHA   241    CCGGTTCCTGCACGCAGTGCACACCCAGCGGGAAGGAAGAGGAAGTGTGTTCAGGAATAG  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13661  CCGGTTCCTGCACGCAGTGCACACCCAGCGGGAAGGAAGAGGAAGTGTGTTCAGGAATAG  13720

G1RHA   301    CTCAGGCAGCTGTGCCATTGGCCGGCAGGCCTGTCAGTCGCCTGGGGGCCGAGGCTGAAA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13721  CTCAGGCAGCTGTGCCATTGGCCGGCAGGCCTGTCAGTCGCCTGGGGGCCGAGGCTGAAA  13780

G1RHA   361    AGAGGAAGAAGTCGGC-gggggggctgttgcctggaagggagggaggagggctgagcca  419
               ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
GRB2    13781  AGAGGAAGAAGTCGGCGGGGGGGGGCTGTTGCCTGGAAGGGAGGGAGGAGGGCTGAGCCA  13840

G1RHA   420    ggccgagccaggggggcgtgggggCTGGGACAGCAGGCCGATTGGCAGGGCGGGGTTGCCT  479
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13841  GGCCGAGCCAGGGGGCGTGGGGGCTGGGACAGCAGGCCGATTGGCAGGGCGGGGTTGCCT  13900

G1RHA   480    TCAATCCTGGCCACTGCTCTTAATCGTCCCTCTCCTGCTTCAGGGTGGCATTGTGTGTCC  539
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13901  TCAATCCTGGCCACTGCTCTTAATCGTCCCTCTCCTGCTTCAGGGTGGCATTGTGTGTCC  13960

G1RHA   540    CAGAGTGCCGGAGCGAGTCCCAGAAGAGAGGCGAGGCTAAGCCCAGAGCGCTGGGTTGCT  599
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    13961  CAGAGTGCCGGAGCGAGTCCCAGAAGAGAGGCGAGGCTAAGCCCAGAGCGCTGGGTTGCT  14020

G1RHA   600    TCAGCAGGGAAGACTCCCTTCCCCCTGCTTCAGGCTGCTGAGCACTGAGCAGCGCTCAGA  659
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14021  TCAGCAGGGAAGACTCCCTTCCCCCTGCTTCAGGCTGCTGAGCACTGAGCAGCGCTCAGA  14080

G1RHA   660    A  660
               |
GRB2    14081  A  14081


G1 - aligned to expected sequence

G1RHA   1      GCTGGGTTGCTTCAGCAGGGAAGACTCCCTTCCCCCTGCTTCAGGCTGCTGAGCACTGAG  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Expect  3332   GCTGGGTTGCTTCAGCAGGGAAGACTCCCTTCCCCCTGCTTCAGGCTGCTGAGCACTGAG  3273

G1RHA   61     CAGCGCTCAGAACTAGCATGGAGAGCGACGAGAGCGGCCTGCCCGCCATGGAGATCGAGT  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Expect  3272   CAGCGCTCAGAACTAGCATGGAGAGCGACGAGAGCGGCCTGCCCGCCATGGAGATCGAGT  3213

G1RHA   121    GCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCA  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Expect  3212   GCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCA  3153

G1RHA   181    CCCCCGAGCAGGGCCGC-TGACCAAACAAGATGA  213
               |||||||||||||||  |||||||  ||||||||
Expect  3152   CCCCCGAGCAGGGCCGCATGACCAA-CAAGATGA  3120
```

Figure A.1.4 Alignments of the RHA of modified *GRB2* in G1 cells

The RHA flanking the GFP-Puro^{R} insert in the G1 cell genome was amplified by PCR. Alignment of PCR fragments against human nucleotides sequences and the expected sequence were produced using Blast (Altschul et al., 1990; Blast, 2020). Cyan highlights indicates the RHA whereas green denotes inserted DNA.

```
G1 - GRB2  EU332843.1

G1LHA   61     CGCCCTGATTACACACAAAGAAGCTGTGTTTTTCATTGTACACTTTTAACATTTACCATC  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14220  CGCCCTGATTACACACAAAGAAGCTGTGTTTTTCATTGTACACTTTTAACATTTACCATC  14279

G1LHA   121    ATATTTGTTCTGGAGAGCTTTAGATGCTGAGTAGAATATTAAATATGACTTAAGTGAGTT  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14280  ATATTTGTTCTGGAGAGCTTTAGATGCTGAGTAGAATATTAAATATGACTTAAGTGAGTT  14339

G1LHA   181    GGTGCGATGTGTGGTGTTTAACTGCCTTTGTGTGAATGCCTCATTTGTCTGTTGATACTG  240
               ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14340  GGTGTGATGTGTGGTGTTTAACTGCCTTTGTGTGAATGCCTCATTTGTCTGTTGATACTG  14399

G1LHA   241    GTAATATTTCAAGTAAAGTCAAAGTAATGTGATTGTAGAAAGGAAGTATATGAGTAGATT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14400  GTAATATTTCAAGTAAAGTCAAAGTAATGTGATTGTAGAAAGGAAGTATATGAGTAGATT  14459

G1LHA   301    GATTTGCCCTGTAGAGTTTTATTTTAAAAATAGCATTTTTTAAAAAGACCCAGACCTGAA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14460  GATTTGCCCTGTAGAGTTTTATTTTAAAAATAGCATTTTTTAAAAAGACCCAGACCTGAA  14519

G1LHA   361    GTATTTATGTAAAAGCATTCAGTTAGTCTTACTTTGTATAGCCAAAGGATTTTGCGTTCT  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14520  GTATTTATGTAAAAGCATTCAGTTAGTCTTACTTTGTATAGCCAAAGGATTTTGCGTTCT  14579

G1LHA   421    GGAATATGATTTTGGTAGGTTAGAAATTAATTTTCCCAGTTTAGAGAGGTTGTTAGAAGT  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14580  GGAATATGATTTTGGTAGGTTAGAAATTAATTTTCCCAGTTTAGAGAGGTTGTTAGAAGT  14639

G1LHA   481    ACATACCAAAACTCATTAAAGAACCTCAGTGTATTTTTATTCTTTGTGAAAAAACAAATG  540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14640  ACATACCAAAACTCATTAAAGAACCTCAGTGTATTTTTATTCTTTGTGAAAAAACAAATG  14699

G1LHA   541    GAGTTCTAAATAATAAGGTGCAACCTGCTATTGTACAAGTCAGATTTCGACTTTCTTAAA  600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14700  GAGTTCTAAATAATAAGGTGCAACCTGCTATTGTACAAGTCAGATTTCGACTTTCTTAAA  14759

G1LHA   601    AGAAACTTGAGGAGCAATTACCCCATATTAAACTCTAATTGCACGAAAAGAAACCTGCTG  660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GRB2    14760  AGAAACTTGAGGAGCAATTACCCCATATTAAACTCTAATTGCACGAAAAGAAACCTGCTG  14819

G1LHA   661    ACAGTTTAAAAATCAGCTTCTATACCATGTGCAGCTTATT  700
               ||||||||||||||||||||||||||||||||||||||||
GRB2    14820  ACAGTTTAAAAATCAGCTTCTATACCATGTGCAGCTTATT  14859


G1 - aligned to expected sequence

G1LHA   3      ACTTCGTAT-ATGTATGCTATACGAAGTTATTAGGTCCCTCGAAGAGGTTCACTAGGCGC  61
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Expect  733    ACTTCGTATAATGTATGCTATACGAAGTTATTAGGTCCCTCGAAGAGGTTCACTAGGCGC  674

G1LHA   62     GCCCTGATTACACACAAAGAAGCTGTGTTTTTCATTGTACACTTTTAACATTTACCATCA  121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Expect  673    GCCCTGATTACACACAAAGAAGCTGTGTTTTTCATTGTACACTTTTAACATTTACCATCA  614

G1LHA   122    TATTTGTTCTGGAGAGCTTT  141
               ||||||||||||||||||||
Expect  613    TATTTGTTCTGGAGAGCTTT  594
```

Figure A.1.5 Alignments of the LHA of modified *GRB2* in G1 cells

The LHA adjacent to inserted GFP-Puro^R DNA was amplified from the G1 genome by PCR. DNA fragments were sequenced and aligned both to human nucleotide sequences and to the expected sequence using Blast (Altschul et al., 1990; Blast, 2020). The LHA is highlighted in purple and inserted GFP-Puro^R in green.

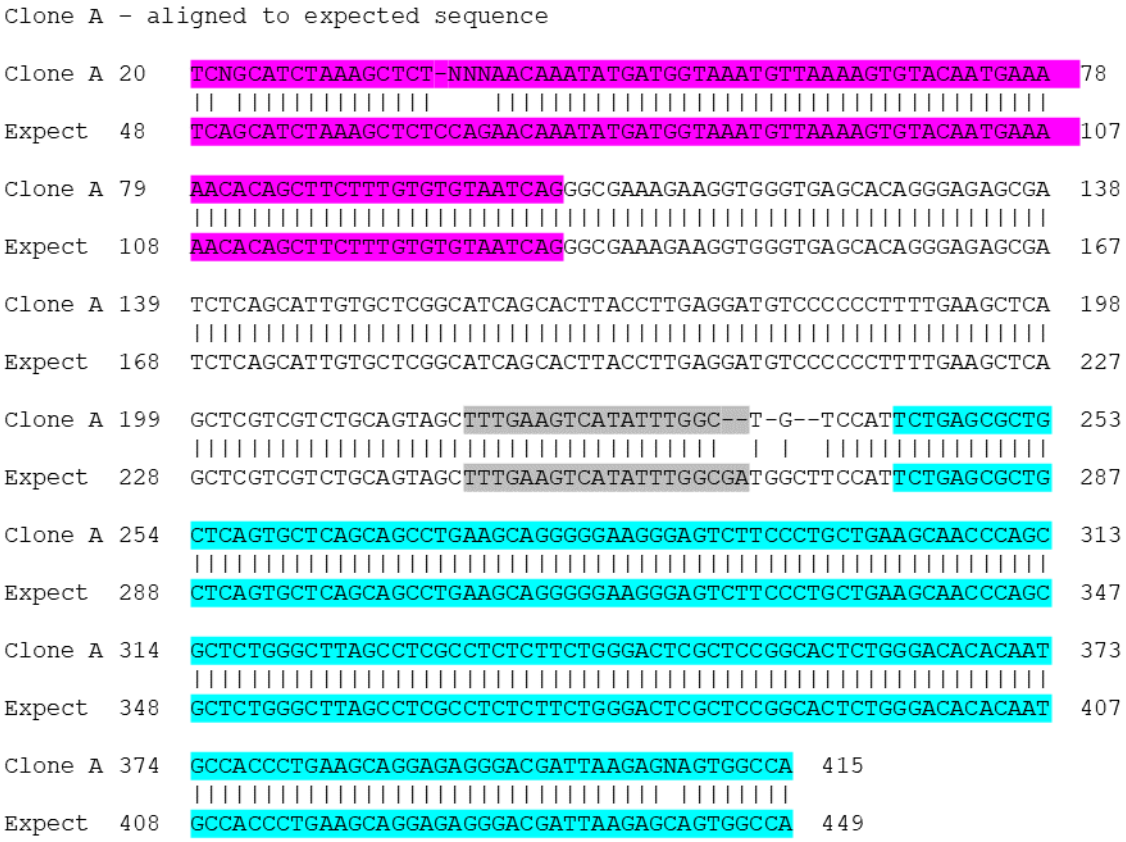


Figure A.1.6 Alignment of *GRB2* without GFP-Puro^R insert in the G1 clone A against wild type *GRB2*

Clone A was generated from pG1 cells. Genomic DNA was extracted and the region at the start of *GRB2* was amplified by PCR using two primers targeted against wild type *GRB2*. The PCR fragment was sequenced and Blast was used to align it against wild type *GRB2* (Altschul et al., 1990; Blast, 2020). Purple and cyan highlights indicate the left and right homologous arms. Grey highlights the sequence targeted by gRNA1.

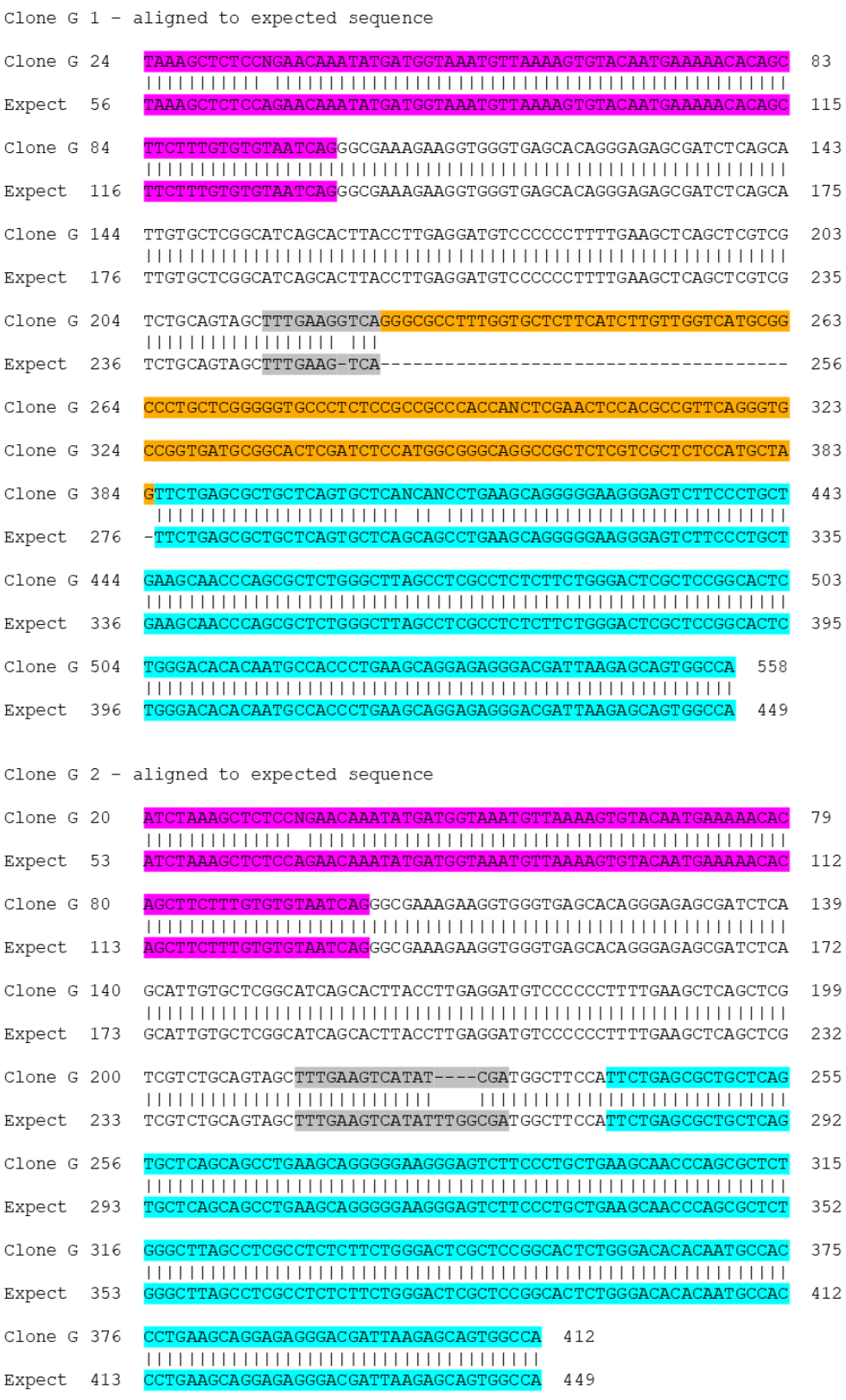


Figure A.1.7 Alignment of the start of *GRB2* without GFP-Puro^{R} insert from the G1 clone G

Clone G was selected from pG1 cells. Genomic DNA was extracted and the region at the start of *GRB2* was amplified by PCR. DNA fragments were sequenced and sequences aligned against expected wild type *GRB2* DNA using Blast (Altschul et al., 1990; Blast, 2020). RHA and LHA are each highlighted in cyan and purple, the gRNA1 sequence in grey, and a large DNA insert in orange.

A.2 Expected amino acid sequence of the GRB2 mutant expressed by G1 clone G

Clone G of the G1 cell line had low expression of a mutant GRB2 protein (Chapter 3, 3.2).

MESDESGLPAMEIECRITGTLNGVEFELVGGGEGTPEQGRMTNKMKSTKG
ALTFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPKNYIEM
KPHPWFFGKIPRAKAEEMLSKQRHDGAFLIRESESAPGDFSLSVKFGNDV
QHFKVLRDGAGKYFLWVVKFNSLNELVDYHRSTSVSRNQQIFLRDIEQVP
QQPTYVQALFDFDPQEDGELGFRRGDFIHVMDNSDPNWWKGACHGQTGMF
PRNYVTPVNRNV

Figure A.2.1 Expected sequence of the mutant GRB2 protein expressed in the G1 clone G

DNA which had been inserted into the start of *GRB2* in clone G was translated along with the rest of the gene using ExPASy (Gasteiger et al., 2003). Underlined residues are encoded by inserted DNA.