Recognition of precursor microRNAs by the Dicer cofactor TRBP

Banushan Balansethupathy

PhD

University of York
Biology

September 2018

Abstract

MicroRNAs (miRNAs) are small non-coding RNAs that enable post-transcriptional gene regulation through RNA interference (RNAi). The biogenesis of miRNAs involves distinct enzymatic processing steps, each of which requires a specific set of proteins. The processing of precursor miRNAs (pre-miRNAs) is performed by the endoribonuclease Dicer. Dicer enzymes are conserved in all organisms that utilise RNAi and they always associate with a double-stranded (ds) RNA-binding protein. In mammals, Dicer associates with TAR element binding-protein (TRBP) but the precise role of TRBP in miRNA biogenesis remains unclear. Research over the last 15 years has suggested that TRBP aids Dicer substrate recognition, facilitates diffusion of Dicer along dsRNA substrates, influences the length of Dicer products, and regulates Dicer activity or stability. Human TRBP is a 366-residue protein that comprises three dsRNA-binding domains (dsRBDs) separated by long unstructured linkers. The two N-terminal Type-A dsRBDs bind dsRNA while the C-terminal Type-B domain mediates protein-protein interactions. The goal of this thesis is to explore the structure-function relationship and RNA-binding properties of TRBP. DsRBDs recognise dsRNA via three interaction surfaces: Regions 1-3. Previous research has highlighted the contribution of Region 2 and suggested that the sequence and molecular dynamics of this loop may tune the dsRNA-binding affinity of the domain. Here, nuclear magnetic resonance (NMR) spectroscopy was used to evaluate the structure and ps-ns backbone dynamics of the two Type-A dsRBDs of TRBP, with particular focus on RNA-binding region 2. Analysis of ^{15}N NMR relaxation parameters and residue composition of this region revealed differences between the two dsRBDs. A series of single and double mutant variants were made in single and tandem dsRBD constructs, and their dsRNA-binding properties were tested using several pre-miRNAs. A previously reported difference in binding affinity between dsRBD-1 and 2 was shown to be dependent on pH, which was also shown to affect the foldedness of dsRBD-1. Single-site mutations were shown to alter the dynamics of Region 2 and to affect RNA-binding affinity. These analyses suggested that Region 2 may be required for sensing the presence of non-Watson-Crick features in a pre-miRNA. The implications of these results on Dicer activity are unclear, but the characterisation of binding properties and backbone dynamics reported here provide further insight into pre-miRNA recognition by TRBP and proteins with multiple dsRBD more generally.

Table of Contents

List of Tables

List of figures

Acknowledgments

I would firstly like to thank my supervisor's Dr Michael Plevin and Professor Mike Williamson for selecting me as the PhD candidate for this project and supporting me over the past four years. In particular, Michael Plevin for his frequent advice and help, which he doled out with patience throughout and to the very end. Without his guidance I would not be the researcher I am today.

I am also thankful for my TAP panel, Professor Jennifer Potts and Dr Dimitris Lagos for their suggestions in regard to this project. In addition, I am grateful for the support from everyone on L1, especially Dr Clément Dégut for his wisdom in and out of the lab. I would also like to show my gratitude to Dr Alex Heyam, who has helped me many times throughout these four years. Thank you to Leanne Archer who shared with me the ups and downs of the PhD journey.

Finally, I am thankful for my wonderful family and friends who have supported me in many ways and helped me get through. Especially my wonderful partner Amy Creighton for her support.

Authors declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1- Introduction

1.1 MicroRNAs; Introduction, discovery and role

MicroRNAs (miRNAs) are small non-coding RNAs that can regulate the expression of genes at the post-transcriptional level. In humans, miRNAs are the principal mediators of RNA interference (RNAi), a process in which small regulatory single-stranded (ss) RNAs base pair with complementary sequences on target mRNA, which in turn leads to the degradation and/or translation repression of the mRNA (Bartel 2004; 2009; Friedman et al. 2009).

The discovery of miRNAs began with an investigation carried out on the gene *lin-4*. This specific gene was shown to control the development of *C. elegans* larvae. It did not code for a specific protein but still carried a functional role (Lee et al. 1993). Two small non-coding RNAs derived from the *lin-4* gene. The sRNAs varied in size and one was seen to fold into a stem loop and proposed to be the precursor of the other shorter RNA. Later, it was determined that the lin-4 RNAs had an antisense complementarity to various sites in the 3' untranslated region (UTR) of the lin-14 mRNA (Wightman et al. 1993). This region was previously identified to be required for the repression of *lin-14*. The sites of interaction found on both lin-4 and lin-14 were probed by a series of molecular and biochemical experiments. This highlighted that the direct but imprecise base pairing of the sRNA and mRNA is important for the regulation of Lin-14 protein expression (Wightman et al. 1993).

The discovery and investigation into the lin-4 RNA hypothesised a model where by *lin-14* gene expression is regulated by the RNA as part of regulatory pathway which triggers the transformation in cell divisions of the first larval stage to those of the second. Which is instigated by the downregulation of the nuclear protein Lin-14 (Lee et al. 1993; Wightman et al. 1993). This information hinted at a new mechanisms of gene regulation,

but it was only seven years later when a second RNA was identified to control genes in this manner, let-7 (Reinhart et al. 2000). It was shown that *let-7* and *lin-41* were evolutionarily conserved throughout metazoans and homologues were found in molluscs, sea urchins, flies and humans (Pasquinelli et al. 2000). This group of sRNAs involved in gene regulation through RNA interference have since been termed microRNAs. Computational predictions of miRNA target sites indicate that miRNAs may be involved in a broader network of gene regulation (Lewis et al. 2003).

As described, the first miRNAs discovered in *C. elegans* lin-4 and let-7 carry functional roles in controlling development, mutated *lin-4* cells were unable to make the transition from the first to the second larval stage (Lee et al. 1993). In adult *C. elegans*, *lin-4* has been shown to be required for the down regulation of *lin-14* and the life-span regulation of the worm (Boehm and Slack 2005). Let-7 has also been shown to be important for the larval to adult transition as it regulates the expression of several developmental genes (Reinhart et al. 2000). In vertebrates, miR-124 has been shown to play a role in regulating neural development, it was shown through *in vitro* assays where miR-124 overexpression triggered, and knockdown assays prevented neural development (Gao 2010).

The expression levels of miRNAs have been studied in various tumours and identified to be significantly altered (Lu et al. 2005; Jansson and Lund 2012; Peng and Croce 2016). Lu et al (2005) identified that the global miRNA expression was lower in cancer cell tissues compared with normal tissue. The expression of 217 mammalian miRNAs was investigated using bead-based flow cytometric expression profiling (Lu et al. 2005). It was concluded that the downregulation of some miRNAs might play a crucial role in the generation or maintenance of tumour cells. The loss of miRNAs in tumour cells suggests that certain miRNAs play a role in tumour repression. However, it has been shown that miRNAs from the miR17-92 cluster act with oncogenic potential (He et al. 2005; Olive et al. 2010). MiR-155 has been shown to be sufficient to induce tumorigenesis, experiments in transgenic mice where miR-155 had been overexpressed showed pre-

leukemic pre-B cells which proliferate to B cell leukaemia (Costinean et al. 2006).

Let-7 one of the first miRNAs discovered has also been shown to have a role in cancer as the expression of this miRNA was found to be decreased in lung tumours compared to normal lung tissue and overexpression of let-7 was able to suppress cancer cell growth, shown in both *in vitro* and *in vivo* studies (Johnson et al. 2007; Kumar et al. 2008)*.* The production of miRNAs is under heavy control which will be discussed further below. The dysregulation of miRNA production has been associated with disease progression and cancers (Melo et al. 2009; 2010; Gurtner et al. 2016).

1.2 MiRNA biogenesis

The production of miRNAs relies on several processing steps within cells (Lee et al. 2002). The primary transcript of miRNAs produced by RNA polymerase II can contain multiple hairpins, each of which can give rise to a different miRNA. MiRNAs are generated from these primary RNA transcripts in a two-step process (Figure 1.1). The initial primary miRNA (pri-miRNA) is a 1-2 kilobase nuclear transcript, which possesses a short imperfect stem loop structure. This pri-miRNA is recognized by the nuclear "Microprocessor" complex, which is made up of the nuclear protein Di George Syndrome Critical region 8 (DGCR8) and the RNase III enzyme Drosha (Han et al. 2004). The Microprocessor excises the imperfect stem loop from the pri-miRNA (Lee et al. 2003; Denli et al. 2004). The cleavage product, which is termed the precursor miRNA (pre-miRNA), has a hairpin structure with a two nucleotide (nt) 3'-overhang (Bartel 2009). The pre-miRNA is assembled into a complex with the nucleocytoplasmic transporter factor Exportin-5 and RanGTP, which facilitates the translocation of the pre-miRNA into the cytoplasm (Bohnsack et al. 2004; Lund et al. 2004). To produce the final miRNA for RNAi one more processing step is required where the 60-70 nt pre-miRNA is processed in the cytoplasm and is cleaved by a second RNase III enzyme called Dicer (Bernstein et al. 2001; Feng et al. 2012). The structure of Dicer has been

elucidated to give further insight on the mechanisms of pre-miRNA processing (MacRae et al. 2006; Lau et al. 2009; 2012). The processing step by Dicer has been revealed to be aided by a class of proteins known as double-stranded RNA binding proteins (dsRBPs) (Koscianska et al. 2011; Ha and Kim 2014). The product of Dicer cleavage is a 20-25 bp duplex, which contains the functional miRNA guide sequence (Kim et al. 2009). One strand of the miRNA/miRNA* duplex is transferred to the RNA-induced silencing complex (RISC), which is responsible for miRNA-directed RNAi in humans (Pratt and MacRae 2009). The intermediate protein complex made up of TRBP, Dicer and Argonaute, part of the RISC loading complex (RLC) is potentially responsible for assisting the transfer of the correct guide strand miRNA from Dicer to RISC (MacRae et al. 2008).
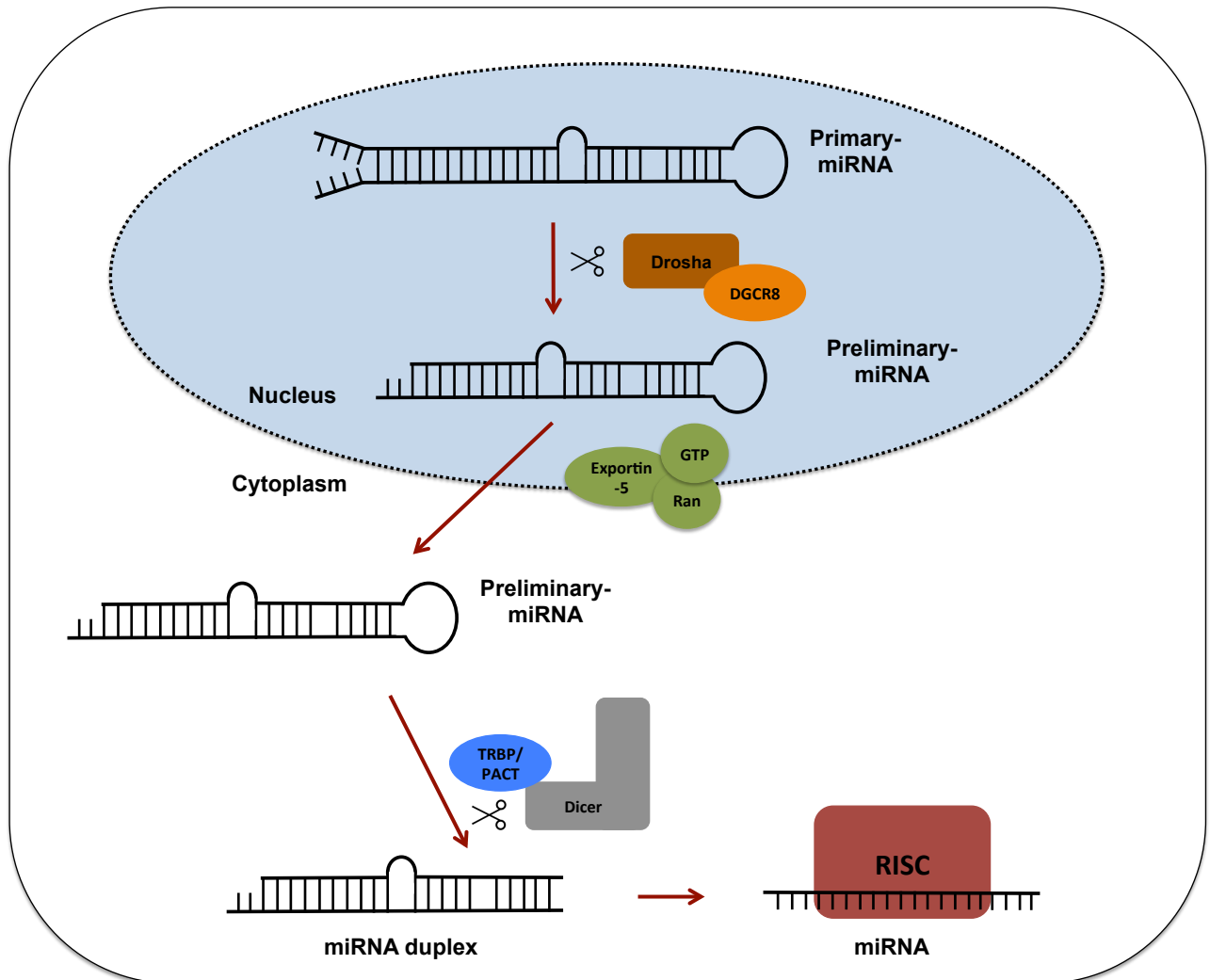
Figure 1.1 Overview of the miRNA biogenesis pathway.

Pre-miRNAs are RNA hairpins produced in the nucleus from primary miRNA and exported to the cytoplasm. The endoribonuclease Dicer then processes pre-miRNAs to remove the terminal loop to produce a RNA duplex, this duplex contains the miRNA sequence which gets loaded onto the RISC.

1.3 The role TRBP and PACT play in facilitating Dicer processing of dsRNA

As described, miRNAs are produced from long RNA transcripts, and require several regulated enzymatic processing steps during their maturation (Bartel 2009). Two multi-domain dsRBPs, TAR element binding protein (TRBP) and the Protein activator of PKR (PACT) have been identified to interact and aid Dicer in the processing of pre-miRNAs (Chendrimada et al. 2005; Haase et al. 2005; Lee et al. 2006; Lee and Doudna 2012; Lee et al. 2013). Biochemical and structural studies have

shown that TRBP and PACT associate with Dicer (Kok et al. 2007; Laraki et al. 2008; niels et al. 2009; Wilson et al. 2015).

In vivo studies of TRBP have identified that presence of this dsRBP can be crucial for RNAi activity, with knockout studies showing a loss of RNAi activity (Chendrimada et al. 2005; Haase et al. 2005; Lee and Doudna 2012; Lee et al. 2013). This was tested by investigating mRNA silencing by short hairpin RNA (shRNA), when TRBP and PACT were depleted. The shRNA requires processing by Dicer and this identified that both proteins contributed to Dicer processing as its mRNA silencing was affected by the depletion of TRBP and PACT. Another study contradicted these claims and identified that siRNA-mediate silencing was not affected by the depletion of TRBP and PACT (Kok et al. 2007). However, the effectiveness of the knockdown was not measured and knockdown of TRBP and PACT were carried out using siRNAs. As effectiveness of the siRNA knockdown technique can depend on the presence and activity of TRBP and PACT, this is not the best technique to knockdown these dsRBPs.

In vitro binding studies have revealed that Dicer in complex with either TRBP or PACT can bind to dsRNA with a higher affinity relative to Dicer alone (nivas Chakravarthy et al. 2010; Lee and Doudna 2012; Lee et al. 2013). This dsRNA binding affinity enhancement suggests that these dsRBPs may be required for the initial recruitment of dsRNA to Dicer.

1.3.1 Altering the processing rate of Dicer

Existing explorations into TRBP and PACT have underlined that the two dsRBPs impose different effects on Dicer for the processing of dsRNA (Lee et al. 2013). One role that has been proposed for TRBP is that it can affect the Dicer processing rate of dsRNA. Several *in vitro* studies have shown TRBP can either increase the processing rate of dsRNA by Dicer or have a conflicting function and decrease the rate of Dicer processing, dependent on the dsRNA and its respective structure (Lee and Doudna 2012; Lee et al. 2013). *In vitro* dicing experiments identified that the rate change TRBP

imposes on Dicer processing can be affected by structural elements of the dsRNA, such as an unpaired nucleotide at the Dicer cleavage site, which was shown to inhibit processing rate of Dicer bound TRBP. In contrast to TRBP, PACT in complex with Dicer has not been shown to increase the rate of processing of Dicer. However, where TRBP has been shown to increase the rate of Dicer processing for pre-siRNAs, PACT has been identified to severely decrease the rate of pre-siRNA processing compared to a Dicer-TRBP complex or Dicer alone (Lee et al. 2013).

1.3.2 Facilitation of iso-miR production by Dicer cleavage of pre-miRNAs

Variations in the length of the miRNA product (iso-miRNAs), which are derived from the processing of a single pre-miRNA by Dicer, can play a role in finding downstream miRNA targets (Cloonan et al. 2011). Thus, the mechanism behind the production of iso-miRs is important. TRBP and PACT have been shown to have differing effects on Dicer production of iso-miRs with product lengths differing in one nucleotide (Figure 1.2). TRBP has been shown to stimulate iso-miR production and alter the number of iso-miRs produced by Dicer. On the other hand, a Dicer-PACT complex produces the same number of iso-miRs as Dicer would alone suggesting PACT does not influence iso-miR production of Dicer (Fukunaga et al. 2012; Wilson et al. 2015). Knockout (KO) studies involving TRBP and PACT showed a similar trend in iso-miR production. TRBP KO studies in HeLa cells revealed altered levels of iso-miR expression, which was not seen for PACT KO studies. The effect that TRBP has on Dicer cleavage site selection and therefore iso-miR production suggests that this dsRBP may facilitate the positioning of the dsRNA on Dicer (Ha and Kim 2014; Kim et al. 2014). The PACT-Dicer complex not being able to produce iso-miRs could suggest PACT may interact with Dicer for the processing of certain pre-miRNAs to hinder the production of iso-miRs.

In a recent study Dicer partner proteins TRBP in humans and Loquacious-PB (Loqs-PB) in *Drosophila* were shown to be required for tuning the length

of miRNAs produced by Dicer (Zhu et al. 2018). They identified that base-mis-matches in the pre-miRNA stem can alter the length of miRNAs produced by Dicer with its partner proteins TRBP and Loqs-PB.

1.3.3 Substrate gatekeeper for Dicer in an RNA-rich environment

A recent study described that TRBP acts as a "gatekeeper" for Dicer in finding correct dsRNA substrates (Fareh et al. 2016) (Figure 1.2). It was determined that TRBP was able to ensure the efficient processing of pre-miRNAs by Dicer in an RNA crowded environment. Using tRNA as a competitor molecule revealed that the rate of processing of pre-miRNA was severely inhibited when Dicer was unaccompanied by TRBP. However, a Dicer-TRBP complex remained unaffected to an RNA crowded environment. Competition RNA binding assays showed that the binding affinity of Dicer to the pre-miRNA was severely reduced with a 1000-fold excess of tRNA present but with Dicer in complex with TRBP the presence of tRNA did not present any pre-miRNA binding inhibition. This newly identified role of TRBP further highlights the significance of TRBP in the miRNA biogenesis pathway, whether this phenomenon is present within PACT is yet to be investigated.

1.3.4 Facilitating diffusion of Dicer along dsRNAs

Both TRBP and PACT have been shown to have an ATP-independent diffusion property on dsRNA (Koh et al. 2013; 2017) (Figure 1.2). The exact role this diffusion property plays on dsRNA processing has yet to be elucidated but Dicer-TRBP diffusion activity does correlate with dsRNA cleavage. A possible motivation for this diffusion mechanism to exist can be to allow the dsRBPs to scan the overall structure of the dsRNA and in turn find the correct cleavage site for Dicer. The diffusion mechanism could

also be required for selection of appropriate substrates by scanning the dsRNA substrate.
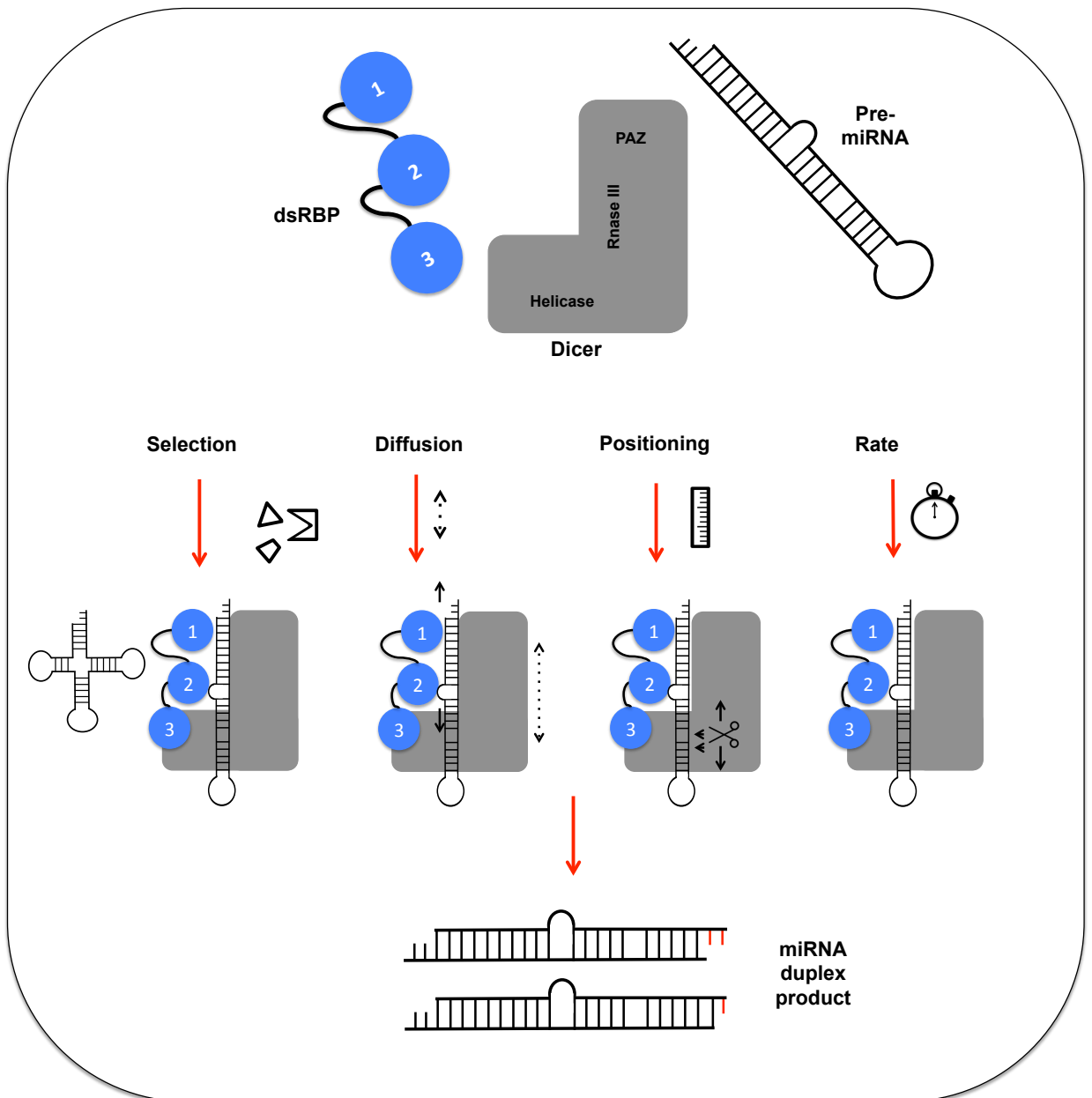
Figure 1.2 Impact of TRBP and PACT in Dicer processing of pre-miRNA.

Model of TRBP and PACT function in miRNA biogeneis. TRBP and PACT facilitate the "selection" specificity of Dicer for correct dsRNA substrates. "Diffusion" by Dicer on dsRNA is independent of ATP and is facilitated by TRBP. "Positioning" of Dicer on the dsRNA to produce iso-miRs with differing lengths is facilitated by TRBP and PACT. TRBP and PACT both affect the "rate" of dsRNA cleavage by Dicer.

1.3.5 Post-transcriptional regulation of Dicer cofactors and its subsequent consequence on miRNA processing

It has been reported that both TRBP and PACT undergo phosphorylation at several sites, but the effect post-translational modification has on regulating miRNA biogenesis has only been studied for TRBP (Paroo et al. 2009; Warner et al. 2016). Phosphorylation of TRBP has been shown to increase stability of the protein but also the stabilization of the miRNA-processing complex (Paroo et al. 2009). Phosphomimic TRBP transfection into human cell lines showed an increase in miRNA production relative to the controls (Paroo et al. 2009). This was credited towards the increased stability of the TRBP-Dicer complex as *in vitro* studies showed no differences in Dicer processing when complexed with wild-type or phosphor-mimic TRBP (Chakravarthy et al. 2010). A variety of kinases have been shown to modulate TRBP phosphorylation. Increasing the stability of the Dicer-TRBP complex through phosphorylation of TRBP, showed an increase in miRNA production and in turn the downstream miRNA silencing function.

1.4 The RNA Induced Silencing Complex (RISC)

The RNA induced silencing complex is made up of by a family of ribonucleoproteins that target genes for silencing, RNA interference is a powerful mechanism of gene silencing required by many organisms. It was identified when the mechanisms involved in the RNA interference pathway were investigated (Hammond et al. 2001; Pratt and MacRae 2009).

The RNA induced silencing pathway requires one of two types of RNAs to target specific genes, miRNAs and small interfering RNAs (siRNAs). SiRNAs are a class of double-stranded RNA molecules, which like miRNAs can regulate the expression of genes using RNAi. Unlike miRNAs, siRNAs are processed from long, fully complementary dsRNAs (Tomari and Zamore 2005). The mechanism of RNAi regulation requires various proteins in the complex to form the RNA induced silencing complex (RISC), which uses the guide RNAs to control the regulation of genes. Argonaute

was identified to be part of the RISC and the first biochemical evidence for it was shown through a purification of pre-loaded RISC from Drosophila S2 cells (Hammond et al. 2001). Biochemical and structural studies have shown Argonaute proteins to be the catalytic core of the RNA effector complex (Liu et al. 2004). Therefore, Argonaute proteins are abundant in nature and found in plants, animals, fungi and protist (Carmell et al. 2002). The number of Argonaute genes present vary in different organisms for example *C.elegans* carry 27 whereas humans only contain 4 (Grishok et al. 2001).

Argonaute proteins form the core of the RISC and contain a PIWI domain which has been implicated to be the required for mRNA cleavage. In some Argonaute proteins this domain has been shown to have a fold similar to that of RNAse H (Song et al. 2004; Yuan et al. 2005). Insight into the function of Argonaute has shown the PAZ domain found in Argonaute has a fold sufficient for oligonucleotide binding. This domain has the ability to anchor the single stranded 3' end of the ssRNAs (Lingel et al. 2004). Further structural studies into archaeal Argonaute proteins identified that the PIWI domain has a conserved binding region for the 5'phosphate of sRNAs. This hypothesizes a model where the sRNA is lodged between the PAZ and PIWI domains of the protein and therefore positions the target mRNA near the catalytic centre (Peters and Meister 2007). All human Ago proteins have been shown to bind to microRNAs but only complexes containing Ago-2 support mRNA cleavage. The catalytic activity of Argonaute was identified by mutations carried out on three amino acids (DDH) found in the PIWI domain (Liu et al. 2004).

The loading of the RISC complex in humans depends on key proteins, such as Dicer and dsRNA binding proteins TRBP in humans, which makes up the RISC loading complex (RLC). These three proteins are highly conserved across eukaryotes and represent the minimal components required for a functioning RISC. TRBP has been shown to form a complex with Ago2 in the presence of Dicer identified by immunoprecipitation experiments (Chendrimada et al. 2005; Haase et al. 2005; Lee and Doudna

2012; Lee et al. 2013). *In vitro* constitution of the RISC with TRBP, DICER and Ago2 has shown that it is capable of processing pre-miRNAs and cleaving target mRNAs (MacRae et al. 2008). For precise mRNA target inhibition, the correct complementary miRNA must be loaded onto RISC. Reconstituted RLC with DICER, TRBP and Ago2 displayed loading activity by unwinding of the miRNA/miRNA* duplex and selection of the correct guide strand dependent on the thermodynamic asymmetry property of processed pre-miRNAs (Khvorova et al. 2003). In *Drosophila*, selection of the miRNA guide strand has been shown to be mediated by the dsRNA binding protein R2D2, which was shown to sense the relative stability between the two ends of the siRNA. TRBP has been shown to bind siRNA *in vitro* and sense the asymmetry of dsRNA siRNAs by binding preferentially to one end of asymmetric siRNAs. This could potentially control the loading of the guide strand into Ago2 (Gredell et al. 2010).

1.5 Double stranded RNA binding domains (dsRBDs)

The double-stranded (ds) RNA binding domain (dsRBD) is a protein domain between 65-70 amino acids in length. One of the first identified role of the domain is to specifically bind to double stranded RNA (dsRNA) (Krovat and Jantsch 1996; Ramos et al. 2000). However, this is not a necessary property required for the definition of this domain as they have also been shown to interact with other protein domains (Krovat and Jantsch 1996; Gleghorn and Maquat 2014). DsRBDs can be found in eukaryotes, prokaryotes and have been shown to be present in the last common ancestor of metazoans (Kerner et al. 2011). DsRBD harbouring proteins have been implicated in many roles ranging from antiviral response to RNA transport and RNA silencing (Masliah et al. 2013). MiRNA biogenesis across most organisms has a minimum of one RNaseIII enzyme and usually multiple dsRBD containing proteins to aid with the recognition of dsRNA.

1.5.1 Discovery of dsRBDs

The first discovery of dsRBDs started with the identification of this domain in two specific proteins: Staufen found in Drosophila required for the localization of maternal RNAs; and Xlrbpa in Xenopus which has been identified as a homologue of the TRBP (Johnston et al. 1992). A 76-amino acid region in Xlrbpa was identified to be sufficient in binding to dsRNA. Through sequence searches it was identified that Xlrbpa shares high similarity with regions found within TRBP and shares sequence similarity to three regions found in the Drosophila protein Staufen, which was not known to be a dsRNA binding protein. Using Xlrpba and a region of Staufen with high sequence conservation to the RNA binding domain of Xlrpba, binding assays were carried out with several RNA probes and identified protein constructs bound strongly to only RNAs with a double-stranded secondary structure (Johnston et al. 1992).

1.5.2 The basic fold of a dsRBD

The first three-dimensional structures of dsRBDs were solved by solution NMR and identified a conserved protein fold in both *E. coli* RNase III dsRBD and the third dsRBD of *Drosophila* Staufen (Bycroft et al. 1995; Kharrat et al. 1995). An α-β-β-β-α topology was uncovered where two α helices are packed against a three stranded anti-parallel β sheet (Figure 1.3). To date, X-ray crystallography and NMR studies have determined structures of more than 30 dsRBDs (Masliah et al. 2013).
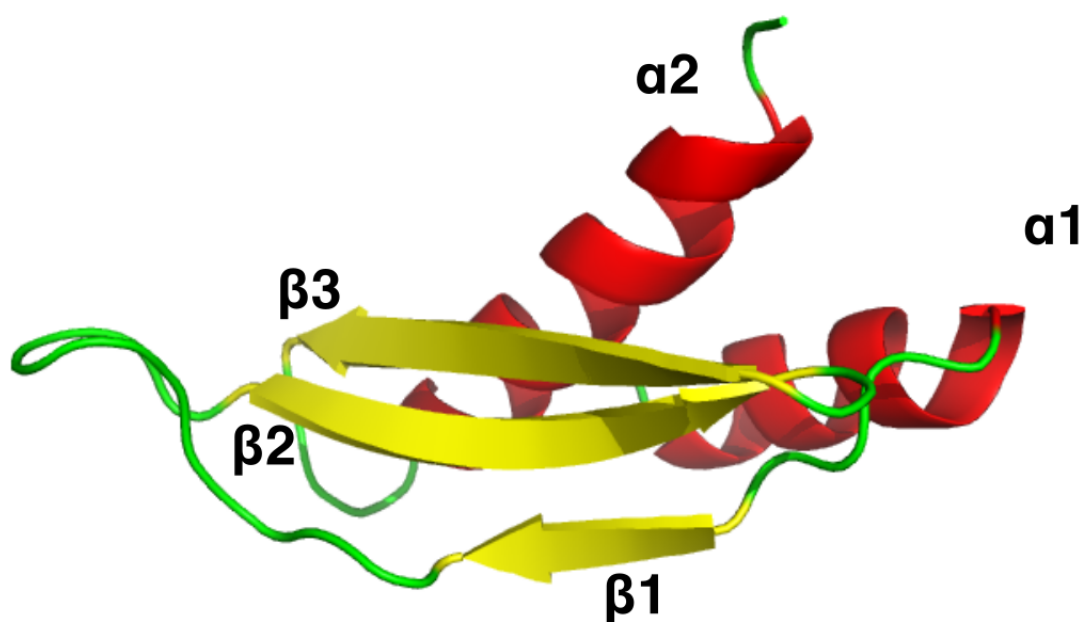
Figure 1.3 Secondary structure of a typical dsRBD.

Taken from a dsRBD found in Staufen (PDB accession code 1STU). The 3D structure holds the typical α-β-β-β-α fold. Alpha helix in red and beta sheet in yellow. The protein structure was visualised using Pymol.

The conserved residues across dsRBDs have been shown to be individually important for the specific conformation of the dsRBD. Hydrophobic residues present in α1, β1, β2 and α2 of most dsRBDs help give the canonical α-β**-**β-β-α fold of dsRBDs. In addition to conserved hydrophobic residues there are two aromatic side chains, which are nearly absolutely conserved between most dsRBDs, a tyrosine residue found in the β1 strand (19th residue of most dsRBDs) and a phenylalanine residue in the β2 strand (residue position 34 of most dsRBDs) (Figure 1.4A). The importance of their conservation can stem from their role of being indirectly involved in the binding of dsRNA by maintaining the correct orientation of the positively charged residues on the RNA binding regions. Furthermore, the position of these aromatic rings plays a role in the structure of the domain as they lie on the edge of the dsRBD to contribute to the overall stability of the domain (Ryter and Schultz 1998).

1.5.3 Double stranded RNA recognition regions of Type-A dsRBDs

Mutational analysis of dsRBDs identified three regions required and conserved for the interaction with dsRNA (Krovat and Jantsch 1996; Erard et al. 1998; Ryter and Schultz 1998; Ramos et al. 2000; Chang and Ramos 2005). Region 1 is located in the first alpha helix with the Region 2 in the loop linking strand β1 and β2 and the Region 3 at the N-terminus of helix α2. Each of these regions contains high conservation across most dsRBDs, specifically the conserved glutamic acid residue found at the N-terminal end in helix α1, the GPxH motif conserved in β1- β2 loop and the positively charged residues in the KKxAK motif at the N-terminal end of helix α2, conserved with only slight variations between dsRBDs (Figure 1.4A).

DsRNA recognition by dsRBDs is highly specific due to the dsRBDs ability to discriminate between dsRNA from other nucleic acid polymers (dsDNA, ssRNA and DNA/RNA hybrids) (Masliah et al. 2013).The conformation adopted by dsRNA is characterized as A-form helix, which is significant for dsRBD recognition. The A-form dsRNA double helix is described by a deep and narrow major groove, and a wide and shallow minor grove (Saenger and Saenger 1984). Another characteristic that should be highlighted is the presence of 2'-OH functional groups on the ribose sugar not present in dsDNA, 2'-OH groups line the entire minor groove, and this may favour the binding of dsRBDs to dsRNA over dsDNA. In summary dsRBDs have the ability to distinguish dsRNA from other molecules by probing the presence of 2'-OH on the dsRNA and the unique structure of the A-form helix and bind across successive minor grooves and the intervening major groove on the dsRNA helix.

NMR spectroscopy and X-ray crystallography have given high-resolution and atomic resolution structures of dsRBDs in complex with various types of RNA molecules (Ramos et al. 2000; Wang et al. 2011; Yamashita et al. 2011). The high-resolution structures have highlighted the RNA binding interface of dsRBDs and exposed the various parts of the domain which

combine to present a surface with the ability to bind to the A-form helix structure that the dsRNA hold (Chang and Ramos 2005; Masliah et al. 2013). The recognition is carried out by the several regions of the domain contacting the bases located at the minor grooves and the phosphate backbone of the major groove (Figure 1.4B)

(A)

```
                              1        10        20        30        40        50        60
HsTRBP-D1|Q15633|30-97     ....TPISLLQEYGTRIGK.TP.V..YDLLKAEGQAHQPNFTFRVTVGDTSC.TG.Q.GPSKKAAKHKAAEVALKHLKG
HsTRBP-D2|Q15633|159-227   ....NPVGALQELVVQKGWRLP.E..YTVTQESGPAHRKEFTMTCRVERFI.EIG.S.GTSKKLAKRNAAAKMLLRVHT
HsPACT-D1|O75569|34-101    ....TPIQVLHEYGMKTKN.IP.V..YECERSDVQIHVPTFTFRVTVGDITC.TG.E.GTSKKLAKHRAAEAAINILKA
HsPACT-D2|O75569|126-194   ....NPIGSLQELAIHHGWRLP.E..YTLSQEGGPAHKREYTTICRLESFM.ETG.K.GASKKQAKRNAAEKFLAKFSN
HsDGCR8-D1|Q8WYQ5|511-578  ....SEVCILHEYMQRVLKVRP.V..YNFFECENPSE..PFGASVTIDGVTYGSG.T.ASSKKLAKNKAARATLEILIP
HsDGCR8-D2|Q8WYQ5|620-685  .....PYQILHECLKRN.H.GMGDTSIKFEVVPGKNQKSEYVM..ACGKHT.VRG.W.CKNKRVGKQLASQKILQLL.H
XlRBPA-D1|Q91836|20-87     ....TPIQLLHEFGTKTGN.HP.V..YTLEKAEGQAHNPSFTFRLVIGDITS.LG.E.GPSKKTPKQKAAEFALNILRG
XlRBPA-D2|Q91836|112-180   ....NPVGGSLQELAVQKGWRLP.E..YTVAQESGPPHKREFTITCRVETFV.ETG.S.GTSKQVARVAAEKLLTKFKT
HsPKR-D1|P19525|9-77       ....FFMEELNTYRQKQGV.VL.K..YQELPNSGPPHDRRFTFQVIIDGREFPEG.E.GRSKKEAKNAAAKLAVEILNK
HsPKR-D2|P19525|100-167    ....NYIGLINRIAQKKRL.TV.N..YEQC.ASGVHGPEGFHYKCKMGQKEYSIG.T.GSTKQEAKQLAAKLAYLQILS
AtHYL1-D1|O04492|15-84     ....VFKSRLQEYAQKYKLPTP.V..YEIVK.EGPSHKSLFQSTVILDGVRYNSLPG.FFNRKAAEQSAAEVALRELAK
AtHYL1-D2|O04492|101-170   ....LCKNLLQEYAQKMNYAIP.L..YQCQKVETLGRVTQFTCTVEIGGIK.YTG.AATRTKKDAEISAGRTALLAIQS
HsADAR1-D1|P55265|503-571  ....NPISGLLEYAQFASQ.TC.E..FNMIEQSGPPHEPRFKFQVVINGREFPPA.E.AGSKKVAKQDAAMKAMTILLE
HsADAR1-D2|P55265|614-682  ....SPVTTLLECMHKLGN.SC.E..FRLLSKEGPAHEPKFQYCVAVGAQTFPSV.S.APSKKVAKQMAAEEAMKALHG
HsADAR2-D1|P78563|78-144   ....LPKNALMQLNEIK.P.GL.Q..YTLLSQTGPVHAPLFVMSVEVNGQV.FEG.S.GPTKKKAKLHAAEKALRSFVQ
HsADAR2-D2|P78563|231-298  PSGKNPVMILNELR...P.GL.K..YDFLSESGESHAKSFVMSVVVDGQF.FEG.S.GRNKKLAKARAAQSALAAIFN
Aa.RNaseIII|O67082|151-219 ....DYKTILQEITQKRWKERP.E..YRLISVEGPHHKKKFIVEAKIKEYR.TLG.E.GKSKKEAEQRAAEELIKLLEE
consensus>50               .....pi..lqe..........p.e..ye.....gp.h...%...v.vd......g...g.skk.akq.Aae.al..l..
```

Region 1   Region 2   Region 3

(B)

Region 1   **Region 3**   **Region 2**

minor   major   minor

Figure 1.4 DsRNA interaction by a type-A dsRBD

(A) Sequence alignment of type-A dsRBDs found in various organisms and dsRBPs *RNA binding regions highlighted. Sequences were aligned using T-coffee (Notredame et al. 2000) and rendered using ESPript (Gouet et al. 2003).(B) Structure of TRBP-D1 with RNA binding regions highlighted (RED) interacting with dsRNA . Edited using Pymol (PDP accession code 5N8L)*

In region 1 of most dsRBDs there are around 4 residues, which produce the interaction surface, one of which is a highly conserved glutamate (Figure 1.4A). The other 3 residues lack conservation between dsRBDs of different proteins, which could be an explanation for the difference in RNA binding properties of various dsRBDs. The residues required for dsRNA interaction in region 1, range from bulky and hydrophobic amino acids (e.g. valine or isoleucine) to amino acids whose side chains are capable of hydrogen bond interactions. They have roles in either interacting with the ribose and bases of the dsRNA or interacting with other amino acids in the dsRBD for correct folding leading to successful RNA binding interaction. The highly conserved RNA binding residue glutamate found in region 1 has been shown to be required for RNA binding as mutating this residue has shown to abolish RNA binding (Ramos et al. 2000). The carboxylic group of this glutamate makes a hydrogen bond with the 2'-OH group of a ribose sugar of the RNA.

RNA binding region 2 found in the loop between β1- β2, has been shown to interact with the minor groove of the dsRNA, one helix turn away from the minor groove interacting with RNA binding region 1 (Figure 1.5). The conserved amino acids in this region match the trend of GPxHxx across most dsRBDs, this region relies on interactions conserved and seen in structures of most dsRBDs. The carbonyl group of the peptide backbone of the histidine residue makes a hydrogen bond to a 2'OH group of one strand of the dsRNA with the imidazole ring of the residue making a hydrogen bond to the 2'OH of the previous ribose on the other strand. The histidine residue has been shown in various structures to be central in RNA binding of this region and the importance of this residue has been shown with mutations of this residue knocking out RNA binding (Krovat and Jantsch 1996; Yamashita et al. 2011). There are slight sequence variations in RNA binding region 2 between dsRBDs. For example, the presence of two histidine residues in RNase III or two proline residues in Xlrpba preceding the RNA binding histidine. Nonetheless, in all three structures the side chain imidazole ring of a histidine makes contact with the 2'OH of

the dsRNA (Figure 1.5). Comparing region 2 of various dsRBDs has also shown a disparity in amino acid composition of this loop region between dsRBDs of different proteins and dsRBDs within the same multi-domain protein.

(A) TRBP dsRBD2

(B) Xlrbpa dsRBD2

(C) RNase III
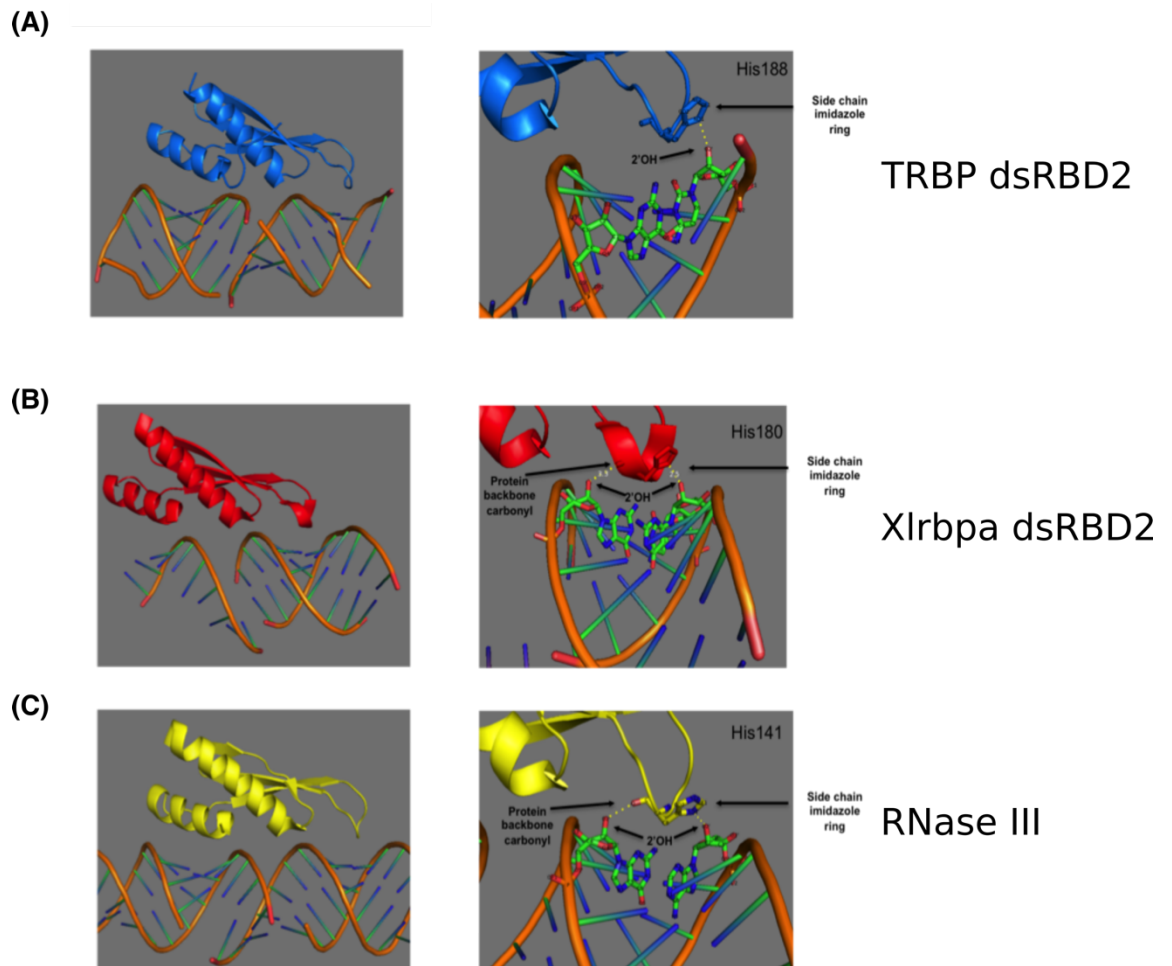
Figure 1.5 High-resolution X-ray structures of the RNA binding region 2 interacting with dsRNA in various dsRBDs.

Crystal structures of dsRBDs interacting with dsRNA with RNA binding region 2 highlighted (A) TRBP-dsRBD2 interaction with a dsRNA (PDB:3ADL). Interaction of His188 in RNA binding region 2 of TRBP dsRBD2 with dsRNA. (B) Xlrbpa dsRBD2 interaction with dsRNA (PDB:1DI2). Interaction of His180 in RNA binding region 2 of Xlrbpa dsRBD2 with dsRNA. (C) RNase III dsRBD interaction with dsRNA (PDB:4m2z). Interaction of His 141 in RNA binding region 2 interacting with dsRNA. Pymol was used to visualize the structures.

RNA binding region 3 found in the N-terminal tip of helix α2, consists of a well conserved amino acid motif of KKxAK, with this surface interacting with the width of the major groove. This region has been shown to contact the negatively charged phosphodiester backbone of both RNA strands. The side chains of the first and the third lysine of the motif interact with one strand with the second lysine interacting with the other strand forming an arch shaped structure. A dsRBD found in Drosha has been shown to not interact with dsRNA as it lacks amino acids found in the three RNA binding regions. However, Kranick et al (2017) were able to engineer dsRNA binding activity into this dsRBD through site directed mutagenesis at the canonical RNA binding region 3. This region alone was able to facilitate the interaction of the dsRBD with dsRNA, which highlights the importance of this RNA binding region (Kranick et al. 2017).

Although there is not much evidence in the literature that dsRBDs have a sequence specific interaction with dsRNA, solution structure and mutational studies of ADAR2 have highlighted the sequence specific interaction of some dsRBDs (Stefl et al. 2010). However, most structural and biochemical information suggests that dsRBDs do not favour specific dsRNA sequences and dsRBDs mainly interact with non-sequence dependent parts of the dsRNA, such as the phosphate backbone and the 2'OH groups of the ribose.

1.5.4 Type-B dsRBDs

As described previously the canonical dsRBD is structured with an α-β-β-β-α fold which a gives a specific secondary structure with the ability to recognize and bind dsRNA. However, further research into dsRBDs identified several domains which still hold the canonical fold but lack the key property of binding dsRNA. These domains termed type-B dsRBDs, have been discovered to form homodimers or even heterodimers with other dsRBDs containing proteins. This feature adds to the function of dsRBPs past the obvious dsRNA binding property.

Human Staufen 1 contains one dsRBD, which follows the type-B rule (dsRBD5) (Gleghorn et al. 2013). There are a variation of ways a dsRBD can have the classic α-β- β-β-α fold without the ability to bind to dsRNA. For example, in the (DUF)283 domain of Dicer and Dicer-like proteins the extension of the loop between α1 and β1 causes α1 to cross at α2 in type-B dsRBDs, this hinders the alignment of α1 with the dsRNA minor groove for the interaction of RNA binding region 1 to dsRNA (Qin et al. 2010). Positioning of amino acids in the B-type dsRBDs can also play a role in knocking out dsRNA binding as can be seen in the two isoforms of STAU1 where they differ by an additional 6 amino acids in the β1 sheet, the addition of these 6 residues was shown to inhibit RNA binding by repositioning the key RNA binding motif found in the loop between β1 and β2 (Bycroft et al. 1995).

Variations in sequence within the actual RNA binding regions could also occur in type-B dsRBDs. In the X-ray crystal structure of human Stau1 dsRBD5 it was shown that it lacked the conserved glutamate required for RNA binding of region1, region 2 was shown to be shorter and therefore lacks the GPxH motif and region 3 lacks the KKxAK motif which is replaced by negatively charged residues which repel dsRNA (Gleghorn and Maquat 2014). TRBP and PACT also contain one type-B dsRBDs that rather than containing dsRNA binding properties, is required for the interaction with protein cofactors (Laraki et al. 2008; Heyam et al. 2015).

1.6 Multi-domain significance of double stranded RNA binding proteins

Although structural biology has highlighted atomic resolution detail of individual dsRBDs, the interplay of dsRBDs with each other in a multi domain protein hasn't been studied in detail (Lunde et al. 2007). RNA recognition by dsRBPs can range in various ways using single or multiple dsRBDs; the domains can independently bind to dsRNA and have little or no contact between them. Alternatively, they can interact with each other to produce a structural conformation required for dsRNA recognition and binding (Tian et al. 2004).

The need for a multi-domain dsRBP can be due to individual dsRBD demonstrating a weak specificity and affinity for the dsRNA and the combination of multiple domains can give high affinity and specificity for dsRNAs. For example, in the interaction of TRBP and protein kinase R (PKR) with dsRNA, each dsRBD domain binds with lower affinity to the dsRNAs compared to the multi-domain protein (Ucci et al. 2007; Yamashita et al. 2011; Benoit et al. 2013).

The interplay of separate domains on a dsRBP can be dependent on the linker region connecting the dsRBDs and therefore the linker may also a play a role in the dsRNA interaction of the protein. The linker can be important in the formation of dsRBP/dsRNA complex as in some cases the linker plays a role in forming the binding surface of the dsRBD by specifically folding to form inter-domain contacts (Teplova et al. 2011; Mackereth and Sattler 2012). However, in other cases the linker can be intrinsically disordered and flexible and in this case the length of the linker may be important rather than the conserved sequence. The flexible long linker can also act as a tether by increasing the local concentration of one dsRNA-binding domain upon the binding of the other, however this is dependent on the length of the linker. In the case of TRBP and ADAR2 they present examples where the two domains interact with their target dsRNAs independently and no physical interaction can be seen between the two domains, this can be due to the flexibility and the length of the linker connecting the domains (Stefl et al. 2010; Benoit et al. 2013).

1.7 Structure of precursor microRNAs

The role of the structure of pre-miRNAs in recognition by dsRBP still remains unclear. Pre-miRNAs have an A-form helical structure and dsRBD interaction has been identified to be non-sequence specific. This proposes that structural features within pre-miRNAs such as bulges, internal loops and non-Watson-crick base pairing may play a role in dsRNA recognition by dsRBPs (Figure 1.6) (Krol et al. 2004). In a study on pre-microRNA

33

structure and its effect on binding various dsRBPs it was shown that ADAR2 and Staufen1 showed a higher affinity for structured RNAs (pre-let7 and TAR RNA) (Bevilacqua et al. 1998). However, pre-let-7A and TAR RNA only contain some structural defects and not the whole spectrum of structural features that pre-miRNAs can exhibit. The two dsRBDs of PKR have also been shown to bind to dsRNAs which contain bulges and internal loops (Bevilacqua et al. 1998). PKR dsRBD-1 is capable of binding to structurally different dsRNAs much more favourably than PKR-dsRBD2 (Bevilacqua et al. 1998; Heinicke et al. 2011).

(A)

(B)
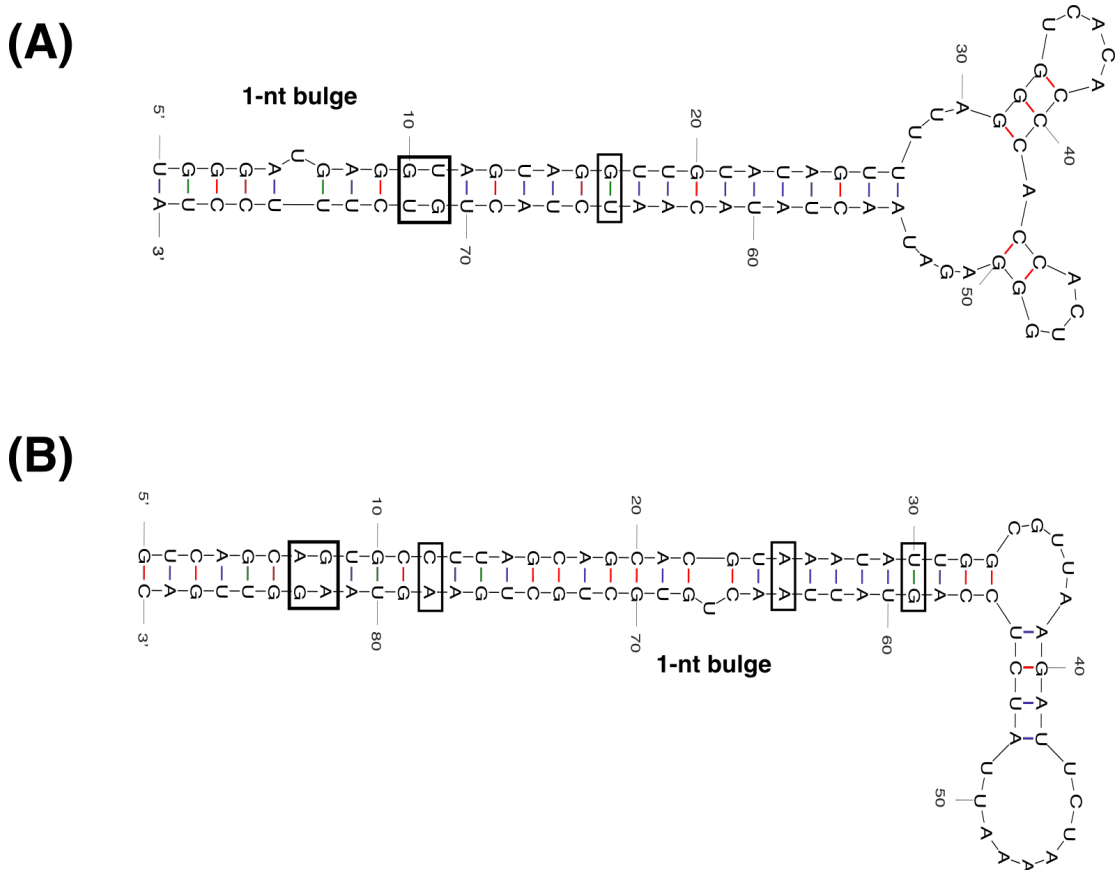
Figure 1.6 Structural features found in pre-miRNAs.

Structures of pre-miRNAs were created using mfold (Zuker, 2003) . Non Watson-Crick base pairs are annotated in black boxes. (A) Structure of Pre-let-7a-1 (B) Structure of Pre-miR-16-1.

1.8 Structure of TRBP and PACT

TRBP and PACT are multi-domain proteins with three dsRBDs, two N-terminal type-A dsRBDs, which bind to dsRNA and one type-B dsRBD found at the C-terminus, which mediates protein-protein interaction (Laraki et al. 2008; Yamashita et al. 2011; Benoit et al. 2013; Wilson et al. 2015). Comparison of the linker regions between TRBP and PACT show that the linker regions are not as highly conserved as the dsRBDs and the length of the linker region differs substantially between TRBP and PACT (Figure 1.7B).
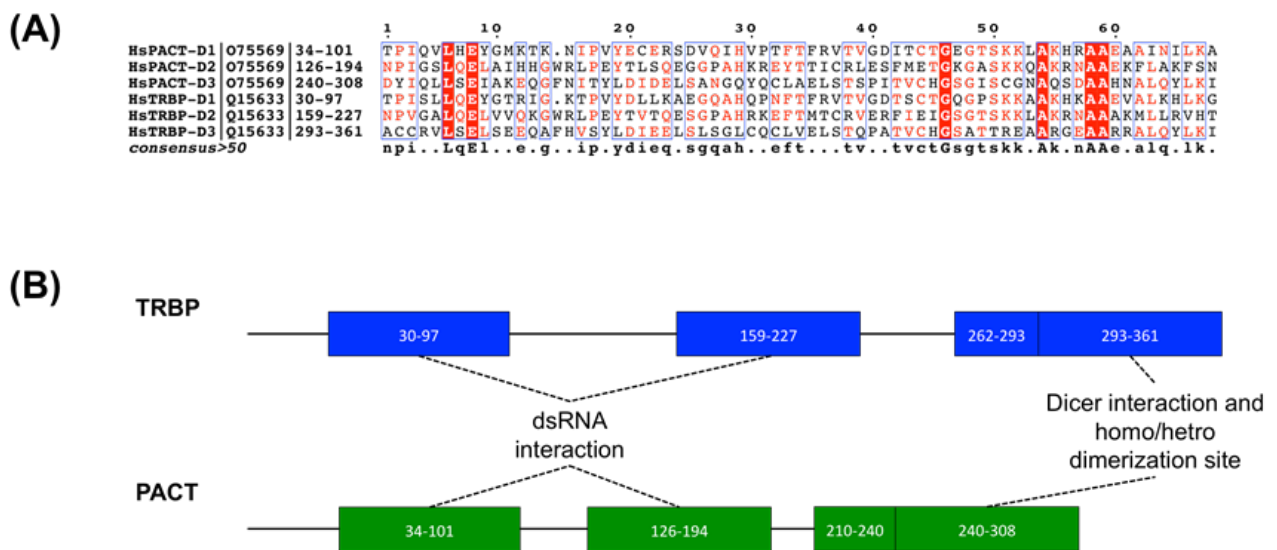
Figure 1.7 Sequence conservation and domain structure and interaction regions of TRBP and PACT
(A)Sequence conservation of the dsRBDs found in TRBP and PACT. Aligned using T-Coffee and rendered using ESPript.

(B)Schematic of the structure of both dsRBPs which contain 3 dsRBDs. Two type-A dsRBDs at the N-terminal with one type-B dsRBD at the C-terminal. The third dsRBD has a N-terminal extension for both TRBP and PACT.

In both TRBP and PACT, domains 1 and 2 are type-A dsRBDs and have been shown to have the ability to bind to dsRNA non-specifically (Yamashita et al. 2011; Benoit et al. 2013). When TRBP was first discovered it was thought to specifically recognise structured RNAs such

as HIV-1 trans-activation response (TAR) RNA (Gatignol et al. 1991; 1993). But since then both TRBP and PACT have been shown to interact with a broad range of dsRNA targets. A linker region in both TRBP and PACT separates the two N-terminal dsRBDs. The linker region between TRBP-D1 and TRBP-D2 has been shown to be highly flexible and the dsRBDs have been shown to not interact with each other in complex with dsRNA or in a free conformation, which could be due to the length and flexibility of the linker region (Benoit et al. 2013). The effect of the short linker between the two dsRBDs in PACT in dsRNA binding remains to be elucidated.

The first and second dsRBD structures of TRBP solved by X-ray crystallography have shown only minimal differences between the dsRBDs with a small shift of the first α-helix (Figure 1.8) (Yamashita et al. 2011). Yamashita et al, identified a difference in the domains at the loop between α1 helix and β2 strand, which in dsRBD2 possesses an extra tryptophan residue (Trp152) (Yamashita et al. 2011). The side chain of this tryptophan protrudes into the hydrophobic pocket made up of Lys150 and other residues. The interaction between the aromatic ring of Trp152 and the positively charged lysine side chain on the opposite side of the dsRNA binding surface has been said to contribute to the further stabilization of the domain structure.

Structures of TRBP and its binding partners have been studied and elucidated. The RNA binding portion of TRBP, TRBP-D1 and TRBP-D2 have been shown to confer the canonical α-β-β-β-α fold. Interestingly TRBP-D1 contains an additional small alpha helix at the N-terminal end (α0) (Masliah et al. 2018; Paithankar et al. 2018). This α-helix has been shown to fold back on the cleft between α1 and α2 and is stabilized by several side chain hydrophobic interactions between α0 and α1 (Figure 1.8F). The only structure at the time of writing for PACT is the N-terminal type A dsRBD, PACT-D1 (Figure 1.8C). The solution structure highlights this dsRBD does not digress from the canonical dsRBD fold. Structures of the type-A dsRBDs of TRBP have been elucidated in complex with

dsRNAs, a crystal structure of TRBP-D2 with a dsRNA and solution structures of TRBP-D1 and TRBP-D2 in complex with a 19 bp siRNA (Yang et al. 2010; Masliah et al. 2018). These structures identify that the type-A dsRBDs of TRBP interact with dsRNA with the three canonical dsRBD interaction regions (Figure 1.8E &F).

The C-terminal domains of TRBP and PACT (TRBP-D3 and PACT-D3) do not bind to dsRNA but are required for the interaction with other proteins, the presence of a type-B dsRBD is common among dsRNA binding proteins (Gleghorn and Maquat 2014). Both TRBP-D3 and PACT-D3 have been shown to bind to Dicer and PACT-D3 has also been shown to interact with Protein kinase R (PKR) (Park et al. 1994; Heyam et al. 2015) . There is also evidence that both TRBP-D3 and PACT-D3 can homodimerize and heterodimerize (Laraki et al. 2008; Singh et al. 2011; Heyam et al. 2018). 3.2 Å crystal structure of TRBP-D3 binding with the N-terminal helicase of Dicer, a four-helix bundle of 97 residues, highlights an interface, which comprises a hydrophobic core surrounded by electrostatic interactions (Wilson et al. 2015).
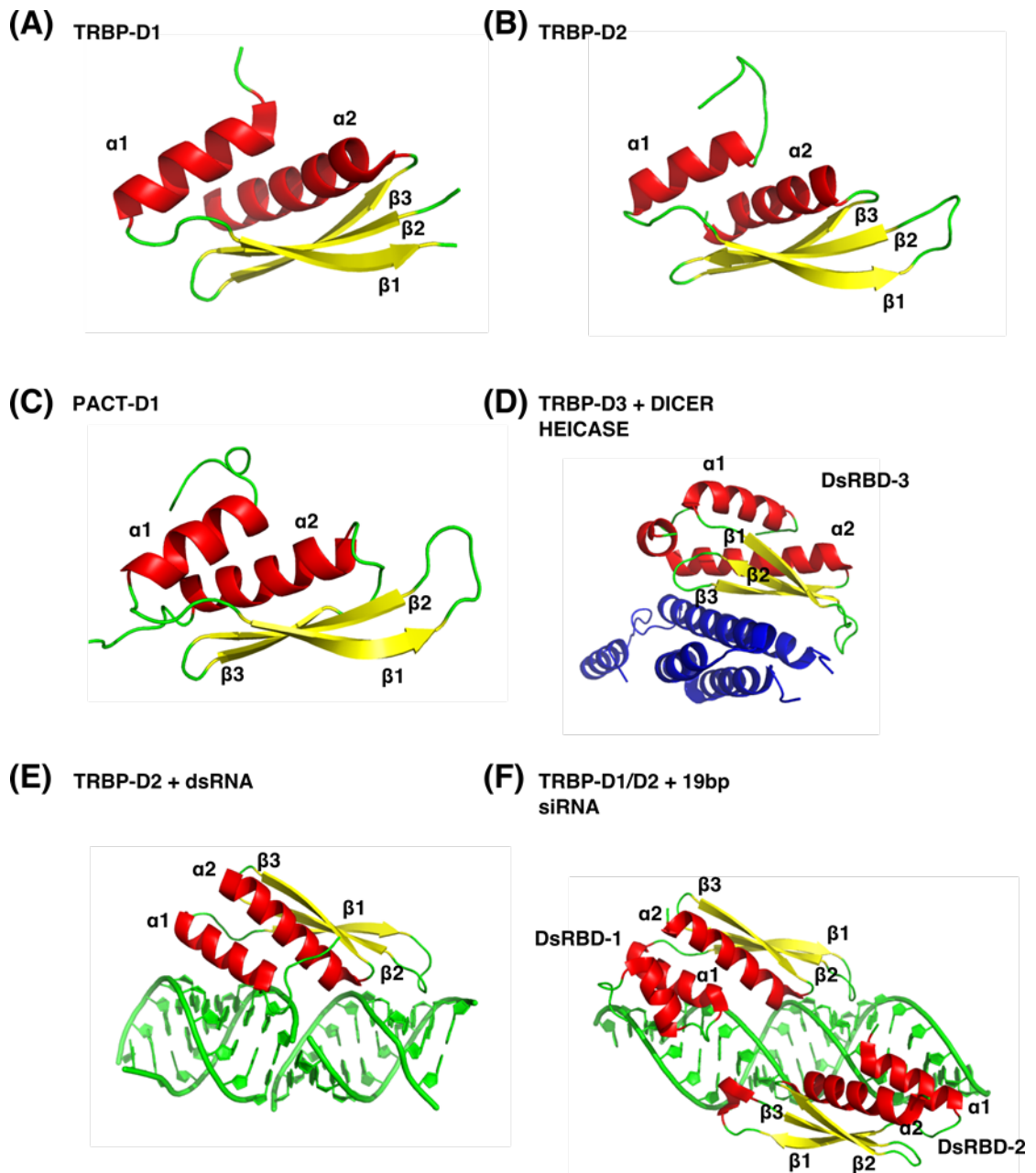
Figure 1.8 3D structures of TRBP and PACT dsRBDs individually and interacting with their respective binding partners.

(A) Solution structure of TRBP-D1 (PDB: 3LLH) (B) Solution structure of TRBP-D2 (PBB: 2CPN) (C) Solution structure of PACT-D1 (PDB: 1DIX) (D) Crystal structure of TRBP-D3 interacting with part of the Dicer helicase (PDB:4WYQ) (E) Crystal structure of TRBP-D2 interacting with a nicked dsRNA (PDB: 3ADL). (F) Solution structure of TRBP-D1 and TRBP-D2 interacting with a 19bp siRNA. Structures were visualized using Pymol.

1.9 Interaction of TRBP with dsRNA

The RNA binding properties of TRBP have been studied and the two N-terminal dsRBDs have been shown to interact with dsRNA independently. RNA binding affinity studies have suggested that the two dsRBDs bind to dsRNA with differing affinities (Table 1.1). However, the difference in affinity ranges from low to high depending on the experimental technique and the dsRNA used (Table 1.1). Two methods have been used to study the interaction of TRBP dsRBDs with dsRNAs, isothermal calorimetry (ITC) and electrophoretic mobility shift assays (EMSAs). A difference in binding affinity between dsRBDs in multi-domain proteins has been elucidated for other proteins such as PKR, HYL1 and ADAR1 (Gabel et al. 2006; Rasia et al. 2010). Structural, dynamic and sequence differences have been proposed to explain the difference in affinities seen in dsRBDs in proteins.

Table 1.1 Binding affinities of TRBP-D1 and TRBP-D2 with various dsRNAs

RNA	Domain	K_D (μM)	Error (μM)	Technique	Reference
Pre-miR-155	D1	15.4	0.8	ITC	Benoit et
	D2	3.7	0.5		al., 2013
21 bp duplex	D1	0.22		ITC	Yamashita
	D2	0.11			et al., 2011
TAR RNA	D1	8		EMSA	Daviet et
	D2	0.06			al., 2000
22 bp duplex	D1	3.5	0.3	EMSA	Acevedo
	D2	1.7	0.1		et al., 2015
33 bp duplex	D1	0.9	0.3	EMSA	Acevedo
	D2	1	0.08		et al., 2015

As mentioned previously, the dsRBDs of TRBP have been shown to bind dsRNA non-sequence specifically, however there is evidence for TRBP in recognising structural imperfections in pre-miRNAs to the extent of it being unable to bind efficiently to imperfections in the dsRNA stem (Takahashi et al. 2013). TRBP was seen binding to a dsRNA with a 1x1 nt A.A mismatch in the stem with a weaker binding affinity than a perfect dsRNA duplex.

Full length TRBP and PACT have been shown to bind dsRNA with different modes of interaction. EMSA and size exclusion chromatography (SEC) showed that two molecules of TRBP or PACT bound to one siRNA. However, TRBP binds to the siRNA in two sequential steps (Takahashi et al. 2013). Whereas, PACT bound to siRNA in one step and the deletion of PACT-D3 disrupted this binding mode and the final protein complex was formed in two steps. This indicates that PACT dimerization affinity is stronger than the siRNA binding affinity and TRBP has a much weaker dimerization affinity compared to dsRNA binding affinity. This has been further suggested with NMR studies of the third domain of TRBP and PACT, PACT-D3 was shown to homodimerize much tighter than TRBP-D3 with a difference in amino acid composition between the two dsRBDs at the dimerization interface being the cause of this difference in homodimerization properties (Heyam et al. n.d.). These differences seen in the homodimerization and in the interaction of TRBP and PACT with dsRNA could allow for Dicer to discriminate between dsRNAs dependent on which cofactor of Dicer it is bound.

1.10 The properties of RNA binding region 2 in dsRNA recognition of dsRBDs

NMR analysis of the two dsRBDs found in HYL1 and their RNA binding properties showed that the first dsRBD (HYL1-1) bound to dsRNA with a much higher affinity compared HYL1-2. This difference in binding was attributed to RNA binding region 2 on both dsRBDs, as the second domain showed no interaction of this region with dsRNA in the NMR analysis (Rasia et al. 2010). The lack of a histidine residue in the β1-β2 loop of HYL1-2 could be a main contributor in the non-interaction of this region, however the structural properties of the loop region could also have been a contributing factor. The flexibility of loop 2 (RNA binding region 2) has been studied for several dsRBDs using NMR relaxation studies, in the apo and RNA bound state of dsRBDs (Wostenberg et al. 2010; 2012).

Interestingly analysis of dynamics through (^{1}H-^{15}N) Heteronuclear NOE experiments of the two N-terminal dsRBDs of TRBP identified a difference in flexibility in the β1-β2 loop region between the two dsRBDs. TRBP-D1 contained a more flexible loop region compared to TRBP-D2 (Benoit et al. 2013). The importance of this difference in dynamics for dsRBD function has yet to be elucidated; additionally, how this difference in dynamics exists between the two dsRBDs has not yet been identified.

Crystal structures of dsRBDs with dsRNA have highlighted that many of the dsRNA binding residues use both the backbone and side-chain groups to interact with dsRNA. Therefore, the protein backbone flexibility of these regions is likely important in playing a role in fine-tuning the dsRNA binding interaction of dsRBDs. The importance of amino acid composition and length of RNA binding region 2 in dsRBDs has been highlighted in the dsRBDs of RDE-4, the β1-β2 loop region is much shorter compared to other dsRBDs and lacks important dsRNA binding residues accounting for the proteins relatively weak dsRNA binding affinity.

1.11 Aim of this study

It has been identified that TRBP is important in the processing of miRNAs by Dicer and the loading of the RISC. However, the dsRNA interacting properties of TRBP have yet to be fully elucidated. The interaction of the individual dsRBDs of TRBP has been somewhat studied but the differences in the dsRNA binding properties, amino acid composition and motional dynamics in the important dsRNA interacting regions have not been as extensively studied compared to other dsRBDs.

Previous studies into the RNA binding regions of the type-A dsRBDs of TRBP have only targeted RNA binding region 3. Mutagenesis of this region identified its importance in RNA binding of the dsRBDs. Therefore, this study will focus on RNA binding region 2 in the β1-β2 loop of the two N-terminal dsRBDs and determine the role of this region in the specific dsRNA recognition of each dsRBD.

This study will focus on:
- The amino acid composition of RNA binding region 2 of TRBP-D1 and TRBP-D2 and how they play a role in dsRNA binding
    - Identification of the residues required for dsRNA binding.
    - Differences in this region between the two dsRBDs.
- The difference in dynamics of the β1-β2 loop between TRBP-D1 and TRBP-D2
    - Further analysis of the difference in loop dynamics
    - The cause of the difference in loop dynamics

This study should elucidate the importance of amino acid composition of RNA binding region 2 in the dsRBDs of TRBP. Additionally, investigating motional dynamics of dsRNA binding region 2 should highlight the importance of structural features of dsRNA binding regions in dsRNA interactions.

Investigating RNA binding properties of the dsRBDs in TRBP can help to further elucidate how TRBP facilitates Dicer in processing of pre-miRNAs,

such as iso-miR production, rate of dicer production, diffusion ability of TRBP domains and dsRNA specificity.

Chapter 2- Materials and Methods

2.1 Protein production methods

2.1.1 Vector preparations

Codon optimized DNA sequences (Eurofins Genomics) were cloned into a pETFPP expression vectors produced by the University of York Technology Facility (TF). This vector contained a gene for the resistance to kanamycin, an open reading frame coding for a maltose binding protein (MBP) solubility tag flanked by a 6xHis tag and a protease cleavage site (GLEVLFQ/GPAM) for human rhinovirus (HRV) 3C protease (produce by York TF). Due to the cleavage site of 3C protease present between the glutamate and glycine, all protein constructs of interest have a non-native sequence of GPAM at the N-terminus. List of plasmids can be found in the appendix.

The protein of interest used in this study was cloned after the 3C cleavage site, the cloning was performed by the University of York TF. A T7 promoter and a *lac* operator element required for protein expression is upstream of the ORF. This allowed protein expression to be induced by isopropyl β-D-1 thiogalactopyranoside (IPTG).

2.1.2 Transformation into E. coli

Transformation of the vector was carried out in the BL21 (DE3) strain of *Escherichia coli (E. coli).* An advantage of this strain is that the genes encoding proteases have been removed to prevent degradation of the overexpressed protein of interest.

Transformation was carried out with 50 µL of chemically competent *E. coli* cells incubated on ice for 30 minutes with 2 µL of plasmid in a 1.5 mL microcentrifuge tube. The mixture was then heat shocked with 45 seconds at 42°C followed by incubation on ice for 2 minutes. The mixture was then

added to 450 µL of Lysogeny broth (LB) media and incubated at 37°C for 1 hour with agitation. The cell mixture was then plated on LB/agar plates containing 50 µg/mL of kanamycin and incubated overnight at 37°C.

2.1.3 Site directed mutagenesis

Primers for site directed mutagenesis of protein constructs were designed using PrimerX (http://www.bioinformatics.org/primerx/). Primers were designed with specific properties such as length of primer (25-45 nt), melting temperatures of < 78°C and GC content was required to be between 40%-60%. Mutagenesis was carried out on freshly prepared DNA plasmid (100 ng) using the QuickChange lightning mutagenesis kit, the reaction was carried out according to the manufactures protocol.

2.1.4 Sequencing of mutant constructs

5 single colonies were picked from LB agar plates for each mutant protein construct and used to inoculate 5 mL of LB for 16 hours at 37°C. The 5 mL cultures were centrifuged and the extraction of the plasmid was carried out using a Miniprep kit (QIAprep Spin Miniprep Kit) according to manufactures protocol. Plasmids were then sent off for sequencing (Eurofins genomics). The sequenced plasmids were then compared with the WT protein DNA sequence and effectiveness of protein mutation was checked.

2.1.5 Protein expression for NMR studies

The same conditions and protocol were used for the expression of all labelled protein constructs in this study. 10 mL of LB with 50 µg/mL of kanamycin was inoculated with a single colony and grown for 6 hours at 37° C with shaking at 120 rpm. The optical density at 600 nM (OD_{600}) was measured and a volume (V) calculated to give a cell OD_{600} of 0.05 for the new media was extracted and centrifuged. The cells were then resuspended in 50 mL M9/kanamycin minimal medium.

$$V = \frac{OD_{600} \times New\ volume}{Required\ OD_{600}}$$

The 50 mL M9 preculture was grown for 16 hours at 37°C with agitation at 120 rpm. The cells were then resuspended in 1 L of M9 minimal medium again using the above equation to determine a starting OD_{600} of 0.05 for the 1 L culture. The culture was then grown at 37°C for 6-8 hours or until it reached an OD_{600} between 0.6-0.8. Protein expression was then induced by supplementing the culture with IPTG (final concentration in culture 1 mM). After induction the culture was grown at 20°C for 16 hours. The culture was then centrifuged at 5000 RCF for 20 minutes to pellet the cells. Cell pellets were either stored at -20°C or lysed immediately.

2.1.6 Protein expression in LB

When protein constructs were not required to be labelled for NMR studies, protein expression was carried out in LB. A pre-culture of 20 mL of LB/kanamycin was inoculated and grown for 16 hours at 37°C with agitation at 120 rpm. The optical density was measured, and the volume required for a starting $OD_{600\ of}$ 0.05 for 1 L of LB was centrifuged and cell pellet was resuspended and transferred to the 1 L LB culture. The culture was grown for 4 hours or until it reached an OD_{600} of 0.6-0.8 and induced with IPTG, after induction the culture was grown at 20°C for 16 hours. The cell culture was then centrifuged at 5000 RCF for 20 minutes to pellet the cells. Cell pellets were used as described above.

2.1.7 Assessment of protein expression and solubility

To investigate protein expression and solubility, 50 mL cultures of protein constructs were grown as above. Cell pellets were resuspended in lysis buffer. Dependent on the method of lysis (chemical or sonication) the lysis buffer for chemical lysis was also supplemented with 1 mg/mL of hen egg lysozyme (Sigma) and Triton X-100. Gel samples were made of total protein before centrifugation, the supernatant and centrifuged cell pellet

resuspended in the same volume of lysis buffer. Samples were then examined by SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) for protein expression and solubility.

2.1.8 Large scale protein purification

1L cultures of His-MBP tagged protein constructs were grown as above, cell pellets were resuspended in lysis buffer as described above and lysed using sonication. Sonication of samples were carried out on ice with 3 minutes total sonication (3s on, 3s off, power level 5.5) carried out twice with a short incubation on ice in between (Sonicator 3000 Misonix). The lysate was then centrifuged for 30 minutes at 50,000 RCF, the supernatant was stored on ice prior to loading over a 5 mL nickel affinity column (HisTrap HP GE) equilibrated with Ni-IMAC binding buffer. Bound His-MBP tagged protein was then eluted with a Ni-IMAC elution buffer. The eluted protein was then buffer exchanged into Ni-IMAC binding buffer using a desalting column (HiPrep 26/10 GE). The concentration of the buffer exchanged eluted protein was then calculated by UV spectrophotometry using the hypothetical extinction coefficients identified by ProtParam calculated by the number of tryptophan's, tyrosine's and phenylalanine's present in the protein (https://web.expasy.org/protparam/).

Table 2.1 UV extinction coefficients of TRBP constructs

Protein construct	Molecular weight (kDa)	ε_{280} (cm^{-1}M^{-1})
TRBP-D1	8.9	2980
TRBP-D2	8.4	6990
TRBP-D12	22.9	9970

The His-MBP solubility tag was cleaved from the eluted protein using His-tagged HRV 3C protease (York TF) added in a ratio of 1 mg of protease per 50 mg of eluted protein at 4°C for 16 hours. Dithiothreitol (DTT) was added to the eluted protein to give a final concentration of 1 mM prior to the addition of the 3C protease. After cleavage a second nickel affinity purification was carried out to separate the protein from the His tagged solubility tag.

The eluted protein of interest was then concentrated to a volume below 1 mL using a spin column with a 5 kDa MWCO (VivaSpin 20), which was them passed over a S75 Superdex 16/600 column (GE healthcare) equilibrated in size exclusion buffer at a flow rate of 1 mL/min. Eluted fractions were investigated by SDS-PAGE and fractions containing the pure protein of interest were pooled and concentrated (VivaSpin 6). Protein samples were then dialysed in the buffer required for the different experiments.

2.1.9 Media and buffers for protein production and purification

LB medium

Component	Concentration (g/L)
Tryptone	10
Yeast Extract	5
NaCl	10
Agar (for plates)	15

M9 Base 1 L

Component	Concentration (g/L)
Na_2HPO_4 (anhydrous)	6
KH_2PO_4	3
NaCl	0.5
NH_4Cl or $^{15}NH_4Cl$	1

1 L M9 medium

Component	Volume
M9 base	1 L
20% D-Glucose or 99% [U-^{13}C] D-glucose	10 mL
1 M MgSO$_4$	1 mL
100 mM CaCl$_2$	1 mL
100 mM MnCl$_2$	1 mL
50 mM ZnSO$_4$	1 mL
100 mM FeCl$_3$	0.5 mL
100x Vitamin cocktail (Gibco)	10 mL
50 mg/ mL Kanamycin	1 mL

Lysis buffer (pH 7.5) used for sonication

Component	Concentration
HEPES	50 mM
KCl	100 mM
MgCl$_2$	10 mM
CaCl$_2$	10 mM
Imidazole	10 mM
Glycerol	10 %
DNase	20 µg/mL
Leupeptin	1 µg/mL
Pepstatin	1 µg/mL

Ni-IMAC binding buffer pH 7.5

Component	Concentration
Tris-HCl pH 7.5	20 mM
NaCl	500 mM
Imidazole	20 mM

Ni-IMAC elution buffer pH 7.5

Component	Concentration
Tris-HCl pH 7.5	20 mM
NaCl	500 mM
Imidazole	500 mM

Size exclusion buffer

Component	Concentration
Tris-HCl	23.5 mM
KCl	100 mM
MgCl$_2$	10 mM
DTT	5 mM

2.2 RNA production methods

Two pre-miRNAs were used in this study. Pre-miR 155 which was based on hsa-miR-155 stem loop (miRbase accession code:MI0000681). Pre-let-7a sequence was based on has-let-7a-1 (miRbase accession code:MI0000060). The 5' and 3' ends of the RNAs were modified for the inclusion of 2 guanines on the 5' end to aid transcription initiation and the 3' end to base pair with the modified 5' end.

2.2.1 DNA template production

Double stranded DNA templates for RNA production were designed with a T7 promoter sequence and created by annealing two oligonucleotides (Appendix). The two oligonucleotides, a forward and reverse primer, required for the dsDNA template were annealed in DNA annealing buffer by heating a mixture of the two primers at 95 °C for 5 minutes and then letting the mixture cool at room temperature for 20 minutes. DNA templates were then investigated by gel electrophoresis. 100 ng of DNA in 2x loading buffer was loaded on an agarose gel and run for 100 Volts for 50 minutes.

2.2.2 RNA production

RNAs of interest were produced through an *in vitro* transcription (IVT) reaction. DNA templates were incubated with components required for IVT for 4 hours at 37°C, the reaction was stopped by the incubation of the reaction at -20°C.

50 µL IVT reactions

Components	Volume (µL)
rATP (100 mM)	2
rUTP (100 mM)	2
rGTP (100 mM)	2
rCTP (100 mM)	2
GMP (100 mM)	2
$MgCl_2$ (100 mM)	*
T7 RNA polymerase (40ug/ul)	1.3
10x T7 buffer	5
Nuclease free H_2O	*
DNA template (500 ng/ µL)	3.5

*Volume depends on the specific $MgCl_2$ concentration required for the dsRNA

10X T7 buffer

Components	Concentration
Tris-HCl pH 8	400 mM
Spermidine	10 mM
DTT	50 mM
Triton X-100	1%

2.2.3 Large scale RNA purification

RNA IVT reactions were purified by anion exchange chromatography. 1 mL of the reaction was passed over a 1 mL MonoQ HR 5/50 GL column (GE Healthcare) equilibrated with Anion exchange Buffer A. Bound RNA was then eluted using a gradient of Anion exchange Buffer B. An initial elution gradient from 0-100% per mL was carried out to determine the elution volume of the RNA of interest. The elution gradient was then optimized according to the RNA purified. Contents of eluted fractions were then analysed by gel electrophoresis using a 15% urea denaturing gel to check for pure RNA fractions.

2.2.4 3' RNA labelling

RNAs were labelled at the 3' end with fluorescein 5-thiosemicarbazide (FTSC). 0.5 nmoles of RNA was labelled in a 50 µL reaction volume. The labelling strategy uses two steps (Zearfoss and Ryder 2012). First the 3' terminal ribose sugar is oxidized forming a reactive aldehyde. This oxidized sugar is then conjugated to an aldehyde-reactive chemical tag (FTSC). Oxidation is carried out by sodium periodate and as this requires vicinal hydroxyls, the reaction is specific to RNAs and only modifies the 3' terminal ribose.

Step 1- Oxidation of the terminal ribose sugar
The reaction was assembled as described below and reacted for 90 minutes at room temperature. The RNA was then precipitated by adding 2.5 µL of 5 M NaCl, 1 µL of 20 mg/ml glycogen and 100 µL of 100% ethanol. Incubated at -20°C for 20 minutes and spun at 16,000x g for 25 minutes.

Reagent	Volume (µL)
Sodium acetate pH 5.2 (0.5 M)	10
RNA (100 µM)	5
$NaIO_4$	10
H_2O	25

Step 2 -Labelling of the reactive aldehyde

The pelleted RNA from the previous reaction was resuspended in FTSC labelling solution and incubate at 4°C overnight in the dark.

FTSC labelling solution

Reagent	Volume (µL)
Sodium acetate pH 5.2 (0.5 M)	80
FTSC (200 mM)	3
H_2O	317

Step 3-RNA clean up

The labelled RNA was precipitated as mentioned above and resuspended in 20 mM Tris-HCl pH 7.5. The RNA was then further purified using a centrifuge spin column (Bio-Spin 6,Bio-rad) equilibrated in 20 mM Tris-HCl pH 7.5.

Successfulness of labelling was determined by running labelled product on a 15% urea denaturing gel and visualising the gel on an Amersham Typhoon (GE) based on the excitation and emission of the dye (Ex492: Em:516).

2.2.5 Buffers for RNA production

DNA annealing buffer

Component	Concentration
Tris-HCl pH 8.0	10 mM
NaCl	50 mM
EDTA	1 mM

Anion exchange Buffer A

Component	concentrations
Tris-HCl pH 7.0	20 mM
NaCl	100 mM
$MgCl_2$	10 mM

Anion exchange Buffer B

Components	concentrations
Tris-HCl pH 7.0	20 mM
NaCl	1000 mM
MgCl$_2$	10 mM

2.3 Gels

2.3.1 SDS PAGE

Gel components

Component	15% resolving gel	4% stacking gel
30% Acrylamide (40:1)	4.9 mL	650 µL
1.5 M Tris-HCl, pH 8.8	2.5 mL	-
0.5 M Tris-HCl pH 6.8	-	1.25 mL
10% SDS	100 µL	50 µL
Distilled H$_2$0	2.5 mL	3.05 mL
10% APS	100 µL	50 µL
TEMED	10 µL	5 µL

Running buffer

Component	Concentration
Tris	3 g/L
Glycine	14 g/L
SDS	1 g/L

Sample buffer

Component	
1M Tris-HCl pH 7.2	1 mL
Distilled H$_2$0	3 mL
10% SDS	10 mL
Bromophenol blue	0.06 g
Glycerol	12 g

Staining

1 L Coomassie stain

Component	
Ethanol	450 mL
Acetic Acid	100 mL
Distilled H_2O	450 mL
Brilliant Blue R	2.5 g

1 L Destain

Component	
Ethanol	100 mL
Acetic Acid	100 mL
Distilled H_2O	800 mL

2.3.2 Agarose gel

Components for 2% gel

Component	
Agarose	2 g
1XTBE	100 mL
Sybr™ Safe DNA gel stain	10 µL

*Running buffer 1XTBE

2.3.3 Urea denaturing gel

Components for 15% gel

Component	
40% Acrylamide	9.4 mL
10X TBE	2.5 mL
Urea	10.5 g
APS 10%	100 µL
TEMED	10 µL
H_2O	5.6 mL

2.3.4 Native Gel for EMSA

10% Native gel

Component	Volume
40% Acrylamide	3 mL
10X TBE	1 mL
APS 10%	100 µL
TEMED	10 µL
H_2O	9 mL

*Running buffer 1XTBE

Running buffer 1X TBE made from 10x TBE used for both native and urea denaturing gels

10XTBE components	
Tris base	108 g
Boric acid	55 g
EDTA, sodium salt	7.5 g

Staining of RNA for EMSAs and urea denaturing gels were carried out using Sybr™ Gold. Gels were incubated in 40 mL of 1xTBE with 4 µL of stain.

Molecular weight ladders used

Ladder	Gel	Company
Precision Plus Protein™	SDS page gels	Bio-Rad
Low range ssRNA ladder	Urea denaturing gels	BioLabs Inc
dsRNA ladder	Urea denaturing gels	BioLabs Inc
Low molecular weight ladder	Agarose gel	BioLabs Inc

2.4 Biophysical methods

2.4.1 SEC-MALS

SEC-MALS is the combination of size exclusion chromatography with multi-angle light scattering analysis. This technique enables characterisation of protein properties such as size and oligomerisation state. The measurement of light scattered by a molecule into a plurality of angles is used to determine the absolute molar mass and average size of a molecule in solution. The light scattered by a molecule such as a protein is proportional to both the mass concentration and molecular weight of the protein (Wyatt 1993).

100 µL of purified protein sample at 3 mg/ml was passed over a Superdex 75 10/30 analytical gel filtration column (GE Healthcare), pre-equilibrated in SEC-MALLS buffer at a flow rate of 0.5 mL/min. The light scattering, and refractive index measurements were recorded using in-line Wyatt Dawn HELEOS-II and Wyatt rEX Optilab detectors. The measurements were then analysed using ASTRA software version 5.3.4.12 (Wyatt Technology).

SEC-MALLS buffer pH 6.5

Component	Concentration
Phosphate buffer	23.5 mM
KCl	100 mM
MgCl$_2$	10 mM
DTT	5 mM

2.4.2 Circular dichroism

Circular dichroism (CD) uses the calculated difference between the absorption of left-handed circularly polarised light and right handed circularly polarised light dependent on a molecule that contains one or more chiral chromophores. Measurements of molecules are carried out in

57

the visible and ultra-violet (UV) region. The molecule of interest containing chiral chromophores will absorb one circular polarised light (CPL) state more than the other and the CD signal over the corresponding wavelength will be non-zero. The CD signal is either positive or negative depending on which CPL state is absorbed to a greater extent. A positive CD signal is given if the left CPL is absorbed greater than the right CPL and a negative signal is given if the left CPL is absorbed less than the right CPL (Greenfield 2007).

The absorption of L-CPL and R-CPL can be used in structural biology to evaluate the secondary structure, folding and binding properties of proteins (Greenfield 2007). The chromophores in the amides of the polypeptide backbone of proteins are used in CD spectroscopy and different structural elements in proteins have characteristic CD spectra. Proteins with α-helical structures have negative bands at 222 nm and 208 nm and a positive band at 193 nm. Proteins with β-pleated sheets have negative bands at 218 nm and positive bands at 195 nm (Greenfield 2007). Unstructured proteins have low ellipticity above 210 nm and negative bands at 195 nm. Therefore, CD spectroscopy is a useful tool to determine the structural characteristics of proteins.

CD of the protein of interest was carried out on a JASCO J-810. A CD curve of the buffer alone was taken before a CD curve of the protein to normalise the CD signal of the protein. CD spectroscopy was carried out on a 400 µL sample of protein at a concentration of 0.1 mg/mL in a 1 mm pathlength cuvette.

CD data in millidegrees for each protein construct were extracted and the molar ellipticity for the protein constructs was calculated from the original CD data given in millidegrees using the equation below.

$$(\theta) = \frac{m^{\circ} \times M}{10 \times L \times C}$$

θ = molar ellipticity (deg*cm^{2}dmol)

m$^{\circ}$ = millidegrees

M= average molecular weight (g/mol)

L= path length of cell (cm)

C= concentration (mg/mL)

CD buffer

Component	Concentration
Bis-Tris propane	23.5 mM
KCl	100 mM
TECEP	1 mM

2.5 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) is a specific spectroscopy technique on nuclei used in far reaching applications throughout the biological sciences (Marion 2013). NMR utilizes a large superconductive magnet to probe the intrinsic spin properties of atomic nuclei and uses electromagnetic waves in the radio frequency range to manipulate spins alignment.

The principle of NMR is based on the quantum-mechanical property of a nuclei also known as spin, which depends on the organisation of the protons and neutrons of the nuclei. Atomic nuclei with an even number of protons and neutrons have zero spin and atomic nuclei which contain an odd number of protons and neutrons have a non-zero spin and have a magnetic dipole moment. Nuclei such as ^{1}H, ^{13}C, ^{15}N, ^{31}P and ^{19}F are useful in NMR as their spin of value ½ make their behaviour in a magnetic field easy to predict, in contrast with other non-zero spin value and are almost exclusively used. The magnetic moment of these nuclei forces the spin to behave as tiny bar magnets (dipole) (Levit. 2008).

As described above NMR takes advantage of nuclei with ½ spin as they can only exist in one of two possible states in an applied magnetic field. The spin will naturally line up either parallel (low energy state) or antiparallel (high energy state) to the direction of an external magnetic field. This alignment can be perturbated by applying a secondary oscillatory magnetic field in the form of a radio wave of frequency specific to the observed nucleus and the static magnetic field (Larmor frequency). As the nuclei spin returns to its low energy state, energy is emitted by the nuclei in the form of a RF radiation at a frequency close to the excitation frequency depending both on the static magnetic field and the local magnetic field perceived by the atom. The relaxation of the nucleus to its equilibrium state in a magnetic field can be measured using coil probes placed in close proximity to the sample to give an NMR signal also known as the free induction decay (FID) which can be processed by Fourier transform to yield an NMR spectrum of the nuclei.

The relaxation of a nucleus to its equilibrium state is related to its resonant frequency. The resonant frequency of the relaxation is dependent on the local magnetic field perceived by the nucleus, this field is affected by the chemical environment and neighbouring nuclei. Therefore, the chemical environment of a specific nucleus can be derived from its resonant frequency. This can allow for different nuclei in a molecule to be distinguished and create distinguished peaks in the NMR spectrum. The difference between the radiation frequency of a reference nuclei (methyl protons of DSS in my experiment), and the observed frequency of a nuclei in its environment is the chemical shift of this specific nucleus and is measured in parts per million (ppm).

NMR spectroscopy can be a powerful tool in obtaining structural and dynamic information of proteins. In proteins NMR spectroscopy the most commonly probed nuclei are ^{1}H, ^{15}N and ^{13}C, these nuclei can be found in the backbone of proteins or the side chain groups of amino acids.

Fast time-scale local motions of proteins can be obtained by NMR spectroscopy by investing specific spin relaxation times. Spin relaxation methods of proteins most commonly probe the amide ^{15}N spin which will elucidate the motion of the protein backbone (Jarymowycz and Stone 2006; Kleckner and Foster 2011).

2.5.1 2D Nuclear magnetic resonance spectroscopy

(^{1}H-^{15}N) HSQC is one of the most common experiments for the study of proteins by NMR as it displays the proton and nitrogen chemical shifts of the backbone amide groups of proteins. A HSQC spectrum gives a 2D spectrum with one axis for the proton and one axis for the nitrogen and the spectrum contains a peak for each unique proton attached to a nitrogen. Therefore, an amide in the protein backbone will yield a peak in the HSQC spectra and displays a fingerprint of the protein backbone.

(^{1}H-^{15}N) Transverse relaxation-optimized spectroscopy (TROSY) is 2D NMR technique most commonly used to study larger proteins. (^{1}H-^{15}N) TROSY experiments were used for analysis of the double domain protein construct, TRBP-D12.

2.5.2 3D Triple resonance backbone assignment

The assignment of a protein amide cross peak in a (^{1}H-^{15}N) HSQC/TROSY experiment can be carried out by using 3D triple resonance experiments where a third nucleus is probed (^{13}C) (Ikura et al. 1990). 3D triple resonance experiments can probe the side chain carbons or carbonyls directly linked to an amide in an amino acid or the carbons of the preceding amino acid. For example, a 3D HNCO experiment reports on the carbonyl group of the preceding amino acid. A 3D HNCACO experiment reports on the carbonyl group of the amino acid probed and the carbonyl group of the preceding amino acid. Using these experiments, the NH cross peaks of amino acids can be sequentially linked helping with the backbone assignment. Additional experiments such as a HNCA and HNCOCA

experiment can be used in the same manner which probe the Cα, to sequentially link (^{1}H-^{15}N) cross peaks. Once sequentially linked the type of amino acid the NH cross peak represents can be elucidate by the side chain carbon chemical shifts elucidated from HNCACB or HNCOCACB experiments. These give chemical shifts for the Cα and Cβ linked to the NH group. The sequentially linked NH cross peaks can then be assigned with the protein sequence.

Spectra of triple resonance assignment experiments were analysed, and backbone assignment of proteins were carried out using CcpNMR Analysis version 2.4 (Vranken et al. 2005).

2.5.3 ^{15}N relaxation experiments

T_1 relaxation experiments determine the rate at which magnetization returns to the thermodynamic equilibrium along the z-axis. $T_{1ρ}$ is characterized by the decay of magnetisation spin-locked by a radio frequency field. $T_{1ρ}$ can be converted into giving spin-spin relaxation time (T_2), which is the decay constant for the component of magnetization that is perpendicular to the external field. (^{1}H-^{15}N) heteronuclear NOE experiments provide information about the motion of individual N-H bonds residues which have a faster motion than the overall tumbling of the molecule, show a decreased NOE intensity compared to the residues in structured regions. All three relaxation experiments are affected by the local motion of a protein and can shed light on the dynamics of a protein (Palmer 2004; Kleckner and Foster 2011).

2.6 NMR experimental set up

For all NMR samples purified protein were prepared with 10% D_2O and 0.05 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and loaded either in a 5 mm thin wall NMR tube (Wilmad) or a 5 mm thin Shigemi tube (Shigemi Inc). NMR data collection was carried out on two spectrometers (Table 2.2).

Table 2.2 NMR spectrometer details

Spectrometer	Field (MHz)	Probe	Abbreviation
Bruker Avance II	700	Triple resonance cryoprobe	700_York
Bruker Avance I	800	Triple resonance cryoprobe	800_Shef

2.6.1 (^{1}H-^{15}N) HSQC

HSQC experiments were carried out in Tris-HCl pH 6-7.5 20 mM, 100 mM KCl,10 mM MgCl$_2$ and 5 mM DTT all carried out on the 700_York using the pulse sequence hsqcetf3gpsi. Nitrogen and proton offsets used were 118 ppm and 4.7 ppm respectively, the spectral widths used were 28 ppm for nitrogen and 16 ppm for proton. Most (^{1}H-^{15}N) HSQC experiments were recorded with 2048 real and imaginary points in the proton dimension and 128 real and imaginary points in the nitrogen dimension.

HSQC spectra were processed using TopSpin version 3.5 pl 7. An apodization function, zero filling, Fourier transformation and phase correction (if required) were carried out. A gaussian function was applied to in the ^{1}H and the indirect dimension.

2.6.2 Assignment of WT TRBP-D12 and TRBP-D12$_{Mut}$

Assignment of WT-TRBP D12 was carried out on 0.5 mM sample of [^{15}N-^{13}C] labelled sample in 20 mM Tris-HCL pH 6.5, 100 mM KCl, 10 mM MgCl2, 0.5% Glycerol, 5 mM TECEP, 10% D$_2$0, 0.05 mM of DSS at pH 7.0. Experiments were recorded on the 800_Shef at 4°C. Experimental parameters are described below.

Table 2.3 Experimental parameters for assignment of WT-TRBP-D12 and TRBP-D12$_{Mut}$

Pulse sequence	Nucleus	Offset (ppm)	Spectral width (ppm)	No of points (real & imaginary)
Trosyetf3psi.2	^{1}H	4.7	12.5	4096
	^{15}N	117	36	256
Trhncogp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	60
	^{13}C	173	14	80
Trnhncacogp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	48
	^{13}C	173	14	80
Trnhncacbgp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	48
	^{13}C	41	70	180
Trnhncocacbgp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	48
	^{13}C	41	70	180
Trnhncaetgp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	48
	^{13}C	52	28	180
Trnhncocaetgp3d	^{1}H	4.7	12.5	2048
	^{15}N	117.5	30	48
	^{13}C	52	28	180

2.6.3 Backbone assignment of TRBP-D1 Q56P and TRBP-D2 P186Q

Assignment of TRBP-D1 Q56P and TRBP-D2 P186Q was carried out on ~1-1.5 mM sample of [^{15}N-^{13}C] labelled sample in 20 mM Tris-HCl pH 6.5, 100 mM KCl, 10 mM MgCl2, 0.5% Glycerol, 5 mM TECEP, 10% D$_2$0, 50 uM of DSS at pH 7.0. Experiments were recorded on the 700_York.

Table 2.4 Experimental parameters for assignment of TRBP-D1 Q56P and TRBP-D2 P186Q

Pulse sequence		Offset (ppm)	Spectral width (ppm)	No of points (real & imaginary)
B_trosyetf3psi.2	^{1}H	4.7	12.5	4096
	^{15}N	117	36	256
B_trhncagp3d.2	^{1}H	4.7	13.7	1024
	^{15}N	118	26	128
	^{13}C	53.2	30	200
B_trhncacbgp3d.2	^{1}H	4.7	13.7	1024
	^{15}N	118	26	160
	^{13}C	42.5	65	180
B_trhncocacbgp3d.2	^{1}H	4.7	13.7	1024
	^{15}N	118	26	160
	^{13}C	42.5	65	180
B_trhncacogp3d.2	^{1}H	4.7	13.7	1024
	^{15}N	118	35	160
	^{13}C	173.5	14	180

2.6.4 ^{15}N relaxation experiments of WT TRBP-D12 and TRBP-D12$_{Mut}$

^{15}N T_1, $T_{1\rho}$ and (^1H-^{15}N) Heteronuclear NOE relaxation experiments were recorded using the Sheffield_800 spectrometer at 4°C on a ~0.35 mM [^{15}N] labelled sample of WT TRBP-D12 and a ~0.35 mM [^{15}N] labelled sample of TRBP-D12$_{Mut}$. The buffer used was 20 mM Tris-HCl pH 7.0, 100 mM KCl, 10 mM MgCl$_2$ and 5mM TECEP at pH 7.0. T_1 experiment were recorded with randomly ordered relaxation delays of 0,80,240,400,640,800,1200,1760,2400 (ms) with a repeated relaxation delay of 400ms. $T_{1\rho}$ experiment were recorded with randomly ordered relaxation delays of 1,20,30,40,60,90,110,150,200 (ms) with a repeated relaxation delay of 20ms. (^1H-^{15}N) Heteronuclear experiments were carried out using a recycle delay of 10 seconds.

The exact same buffer conditions and experimental parameters were used for calculating ^{15}N relaxation rates for WT TRBP-D2 and TRBP-D2 P186Q. All relaxation experiments were carried out on the 800_Shef.

Relaxation data was processed using Felix NMR, the first dimension ^1H and the indirect dimension (^{15}N) were processed with a sine-bell function and phase shifted. Zero filling and baseline corrections were applied. The data was then referenced to DSS, which has a peak at 0 ppm in the one-dimensional ^1H spectrum.

Table 2.5 Experimental parameters for ^{15}N relaxation experiments

Pulse sequence	Nucleus	Offset (ppm)	Spectral width (ppm)	No of points (real & imaginary)
Trosyetf3gpsi.2	^1H ^{15}N	4.7 117	12.5 36	4096 256
T1N15_tr.nl	^1H ^{15}N	4.7 117.5	12.5 30	2048 260
T1rhoN15_tr.nl	^1H ^{15}N	4.7 117.5	12.5 30	2048 260
N15HetNoe_tr.nl	^1H ^{15}N	4.7 117.5	12.5 30	2048 260

2.6.5 ^{15}N Relaxation data analysis

CCPN analysis version 2.4 was used to transfer cross peak assignments to the relaxation data. T_1 and $T_{1\rho}$ values were obtained from a two-parameter exponential decay to zero model (Equation below). The errors in the relaxation rates calculated using the bootstrap error method on CCPN analysis 2.4.

$$I_{(t)} = I(0) \exp\left(-\frac{t}{T_{1,1\rho}}\right)$$

$I_{(t)}$ =Intensity of peak at time point (t)

$I_{(0)}$ =Intensity of peak at 0

(^{1}H-^{15}N) Heteronuclear-NOE values were calculated from the ratio of peak intensities from the proton saturated (Sat) and non-proton saturated experiment (ref).

$$NOE\ value = \frac{Sat\ Intensity}{Ref\ Intensity}$$

$$Error = abs\ (value) \times \sqrt{\left(\left(\frac{Sat\ noise}{Sat\ intensity}\right)^2 + \left(\frac{Ref\ noise}{Ref\ intensity}\right)^2\right)}$$

2.6.6 $T_{1\rho}$ conversion to T_2

$T_{1\rho}$ rates calculated were converted to T_2 values using the equation below (Lakomek et al. 2012). It uses T_1 and $T_{1\rho}$ values to calculate T_2 value. T_1 and $T_{1\rho}$ were first converted to R_1 and $R_{1\rho}$, using these values and the equation below $R_{1\rho}$ rates were converted to R_2.

$$R_2 = \frac{((R_{1\rho} - R_1)cos^2\theta)}{sin^2\theta}$$

θ = W1/ΔW

ΔW = spin lock frequency offset (9528.99) – ^{15}N shift in Hz

W1= spin lock field strength = 1866

2.7 RNA protein interaction studies

2.7.1 Electrophoretic mobility shift assay

Electrophoretic mobility shift assay (EMSA) is a commonly used electrophoresis technique used to study the interaction of protein with nucleic acids (Fried 1989). It can identify if a protein binds to a nucleic acid or even the binding affinity of a protein for a specific nucleic acid. It can also determine if one or more protein molecules are involved in binding. EMSA depends on the migration of the complex compared to the free molecule in a native acrylamide gel. Complex formation dependent on protein concentration can be used to determine the binding affinity of a protein.

A constant concentration of RNA was incubated with an increasing concentration of protein in EMSA binding buffer (20 mM Tris-HCl pH 6.5, 200 mM NaCl, 5% Glycerol) and run on an 10% native gel. The gel was run at 100 volts for 2 hours at 4°C. The free RNA and RNA-protein complex was then visualised by Sybr™ Gold staining, a fluorescent stain that detects nucleic acids. The gels were then imaged using BioRad Gel Doc™ EZ gel documentation system. The estimated affinity of the protein construct for the dsRNA was calculated by analysis the disappearance of the free RNA band.

2.7.2 Fluorescence anisotropy to measure RNA: protein binding affinities

Fluorescence anisotropy is a principle where light emitted by a fluorophore has unequal intensities across different axes of polarization. The anisotropy can describe the tumbling rate of a molecule the fluorophore is attached to. This principle can be used to measure binding constants of complexes by measuring the change in anisotropy caused by the change in rotational time when a molecule is bound to a partner molecule. Binding is calculated by using the difference in anisotropy of the free, partially

bound and fully bound molecule. It can be used to monitor protein-nucleic acid interactions (Pagano et al. 2011).

Titrations were performed using increasing concentrations of protein with 50 mM 3' fluorescin labelled RNAs in 200 uL volumes. The reaction buffer contained 20 mM Tris pH 7.5, 200 mM KCl, 5 mM TECEP and 5% glycerol. Fluorescence readings were taken in a black flat-bottom 96-well plate (NUNC) at 25°C in a BMG POLARstar Optima plate reader with the detector gain set on 50 mM of labelled RNA.

2.7.3 Determination of binding model for the interaction of RNA: protein

When fitting a model to the binding isotherm obtained from the fluorescence anisotropy, to give an assessment of the binding behaviour of the RNA-protein interaction, there are few things that have to be considered. A sensible binding model would have to consider parameters such as stoichiometry of the reacting molecules. In previous research on TRBP interacting with pre-miRNAs, it has been identified that two double domain constructs can bind to one pre-miRNA, which also allows for four single dsRBDs to bind one pre-miRNA (Benoit et al., 2013). The RNA: protein stoichiometry of 1:4 for our reactions makes the fitting of the saturation curve obtained from fluorescence anisotropy complicated. Although we suspect that each site is a single independent binding site, each domain binding to the RNA will have a different macroscopic binding affinity. Due to when each site on the pre-miRNA is continuously filled by a dsRBD it reduces the number of free sites available on the pre-miRNA for the subsequent binding domain. The macroscopic binding constants or the stepwise equilibrium constant of each domain binding to the RNA cannot be identified from this RNA: protein interaction study as the distinctive binding steps of the RNA being saturated (1:1, 1:2, 1:3 or 1:4 RNA: protein complexes) cannot be extracted from the fluorescence anisotropy data. Fluorescence anisotropy lacks the precision for this. The binding isotherm produced cannot be reliably fitted to a model with four

different binding constants. Therefore, the most appropriate approach to fit the binding isotherms is to fit the data to an equation of single identical independent binding sites, which will estimate the overall binding constant. Although this model cannot precisely define the stepwise binding constants of each single domain it will provide us with a global affinity of the four domains binding to one pre-miRNA. This model can still aid in comparing the binding properties of the various protein constructs and discover the most interesting constructs for further informative investigation.

The dissociation constants for the RNA protein complexes were determined by fitting the binding curves with the equation below in Prism (GraphPad). The equation is based on a one site binding model.

$$Y = \frac{Bmax \times X}{K_D + X}$$

Y= Change in anisotropy

Bmax= maximum change in anisotropy

X= protein concentration

K_D= is the dissociation constant.

Chapter 3- Analysis of protein and RNA constructs produced in this study

3.1 Production of recombinant TRBP domains

3.1.1 Introduction

Recombinant proteins are used throughout the biological sciences and the production of proteins for scientific studies is a biotechnological process. It can be accomplished by the manipulation of gene expression in organisms or using cell free expression systems. Chemical synthesis of proteins is not viable for most proteins due to their size and their complexity; therefore, cells and their cellular machinery are used to produce proteins from modified genetic templates. The most commonly used cellular expression systems are bacterial cells, insect cells and mammalian cells. Each system has its own advantages and disadvantages, ranging from yield, protein size capabilities and post-translational modification of proteins. Bacterial protein expression systems contain advantages such as effortlessness of culture growth and the production of high yields of recombinant protein.

Studies required milligram quantities of highly purified protein, occasionally specific isotope labelled, therefore expression of TRBP constructs were carried out in *E. coli*. TRBP is a multi-domain protein with three dsRBDs separated by long flexible linkers. The two N-terminal dsRBDs interact with dsRNA whereas the C-terminal dsRBD interacts with protein cofactors of TRBP (Gatignol et al. 1991; Laraki et al. 2008; niels et al. 2009). Interested in the RNA binding properties of TRBP, the third domain (TRBP-D3) was disregarded and protein constructs from the RNA binding portion of TRBP were produced (Table 3.1). Each protein construct was flanked by a His-MBP affinity and solubility tag to aid with the solubility and purification of the protein.

3.1.2 Solubility test of protein constructs using different lysis methods

To determine the best method of cell lysis for protein solubility, two methods were tested with protein constructs TRBP-D2 and TRBP-D12. 50 mL cultures of *E. coli* BL21 (D23) cells expressing the protein of interest were grown, split into two and centrifuged. Cell pellets were either lysed by using a buffer with lysozyme or by sonication. Gel samples were made of the: total cell content after lysis, supernatant after centrifugation and the cell pellets (Figure 3.1).

	Residue number	No of amino acids	Predicted molecular weight (kDa)	Molecular weight with tag (kDa)
TRBP-D1	19-99	80	8.9	51.9
TRBP-D2	157-228	71	8.4	51.4
TRBP-D12	19-228	209	22.9	65.9
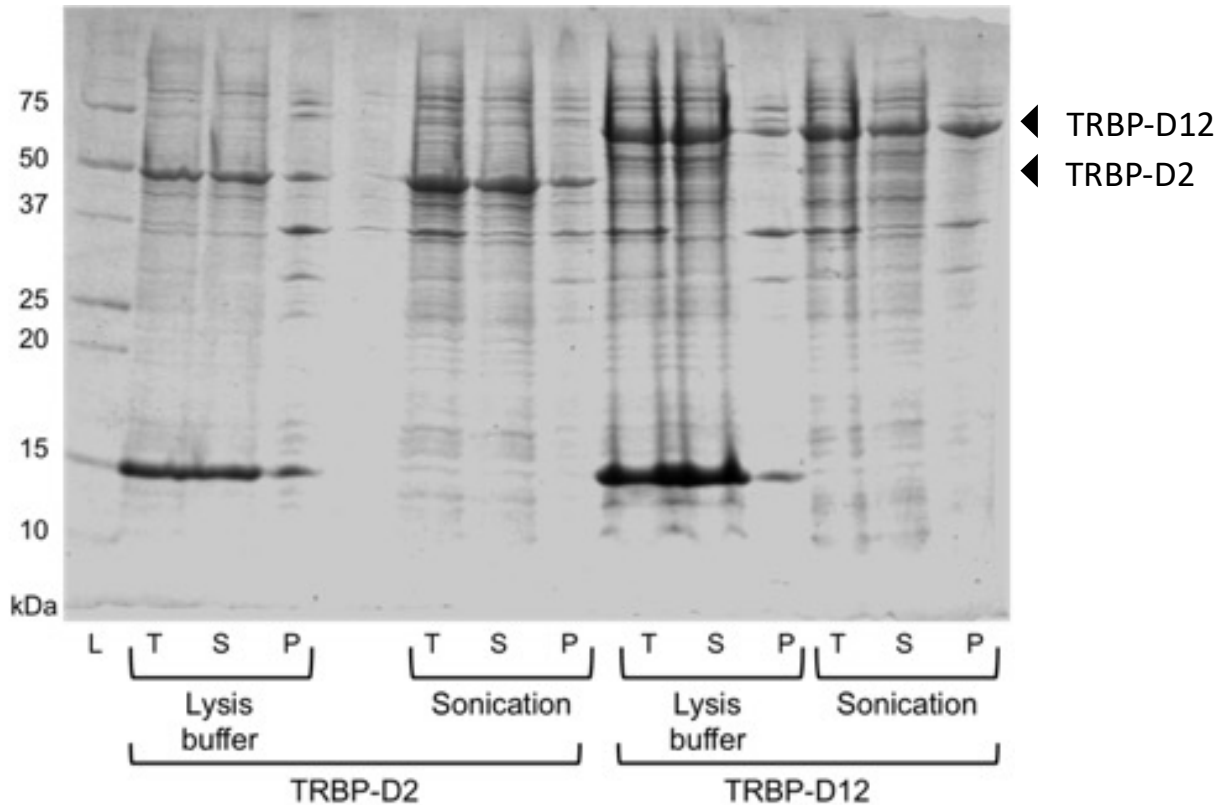
Figure 3.1 Lysis test of TRBP constructs to check for solubility.

15% SDS-PAGE gel of the lysis tests and solubility of TRBP-D2 and TRBP-D12. For each test; 3 lanes were run: Total protein before centrifugation (T); Supernatant after centrifugation (S); and pellet resuspended after centrifugation (P). Lane (L) is the protein ladder, with masses in kDa. The band seen around 15 kDa in the lysis buffer lanes is lysozyme. TRBP-D2 is the overexpressed band at around 50 kDa. TRBP-D12 protein band is around 75 kDa.

The lysis test results show that both methods can extract soluble protein from cells. TRBP-D1 was not tested but due to its similarity in domain structure would be anticipated to behave similarly to the TRBP-D2 construct. Sonication was taken forward for future purifications of proteins due to the ease of this method with large-scale cultures with smaller cell suspension volumes. Sonication is also advantageous in the breakdown of nucleic acids without the need of DNase and RNase which can affect downstream assays.

3.1.3 Scale up of protein expression and purification of TRBP constructs

1 L cultures of the His-MBP tagged TRBP constructs were grown in either M9 minimal medium containing (^{15}N) labelled ammonium chloride or both (^{15}N) labelled ammonium chloride and (^{13}C) labelled glucose for NMR related studies or in LB for higher yields and non-NMR related studies.

Following lysis of the cell pellet by sonication, several steps were carried out to purify the protein of interest and obtain pure protein. The protein of interest was purified form the endogenous *E. coli* proteins using nickel immobilized metal affinity chromatography (Ni-IMAC). Once purified, the protein was cleaved away from its affinity tag (His-MBP) by the protease 3C and a second nickel IMAC purification step was carried out to remove the cleaved protein construct from the affinity tag (Figure 3.2).

(A)

L PH FT E AC PH FT E AC PH FT E AC

◀ His-MBP
◀ TRBP-D12
◀ TRBP-D1 &D2

TRBP-D1   TRBP-D2   TRBP-D12

(B)

◀ His-MBP
◀ TRBP-D1

L Input 250 255 260 265 270 275 280 285 325 330

Elution volume (mL)
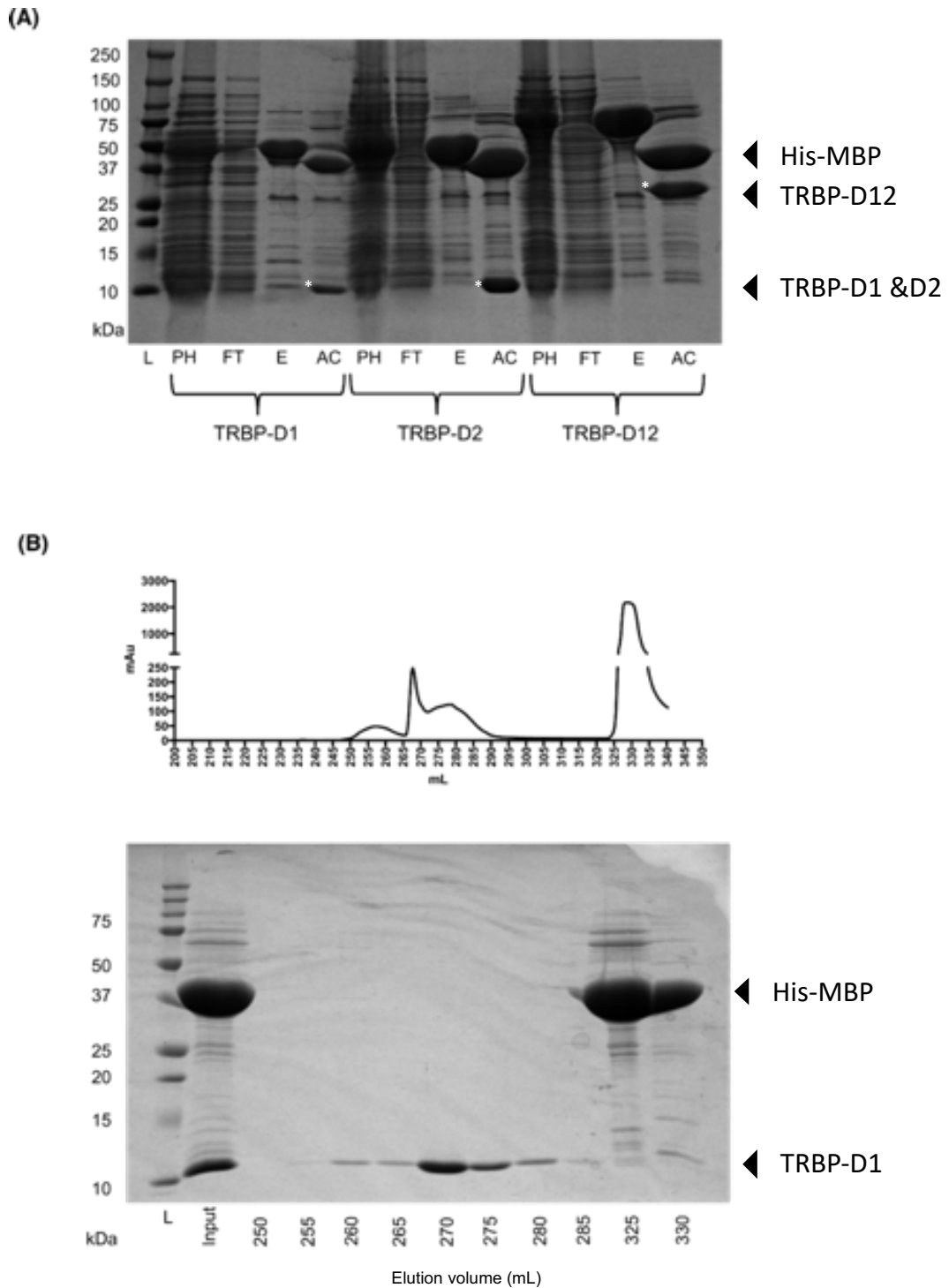
Figure 3.2 Purification of TRBP constructs by Ni-IMAC

15% SDS-PAGE gel analysis of Nickel IMAC purification of TRBP constructs. A) SDS-PAGE of TRBP constructs after the 1st Nickel IMAC purification and cleavage. For each TRBP construct four lanes were run: Protein sample pre Ni-IMAC (PH); Flow-through from Ni-IMAC purification (FT); Elution of bound protein from Ni-IMAC (E); protein construct after dialysis and 3C protease cleavage (AC); White astericks are used to highlight the protein of interest. B) Example of 2nd Ni-IMAC purification of TRBP-D1 after cleavage of the affinity tag. 15% SDS page gel accompanied by the 2nd Ni-IMAC purification profile.

Following the 2nd Ni-IMAC purification, fractions containing the protein of interest were first concentrated and then purified using SEC (Figure 3.3).

(A)
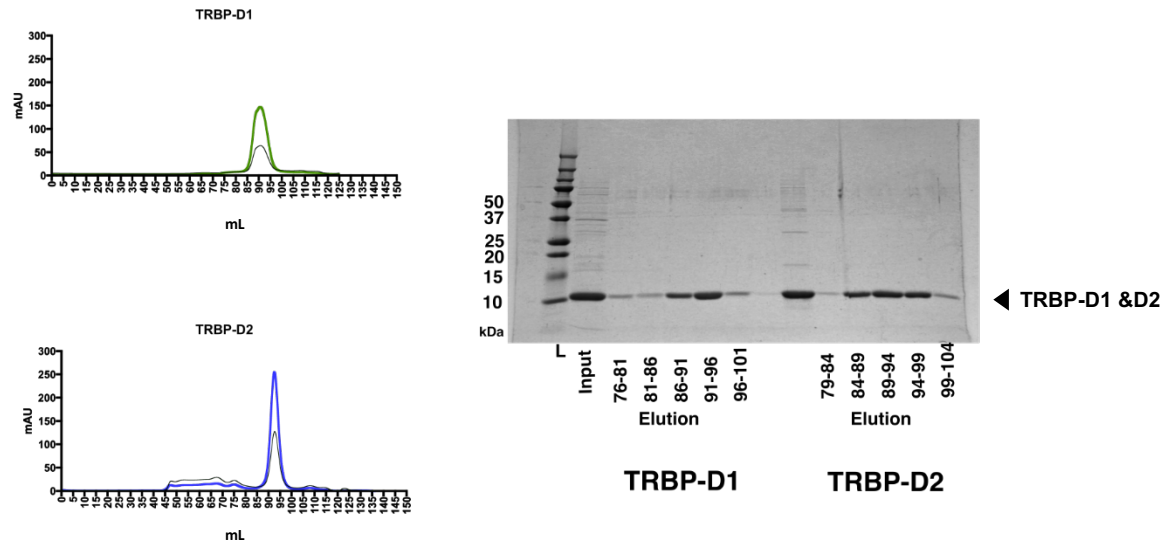
(B)
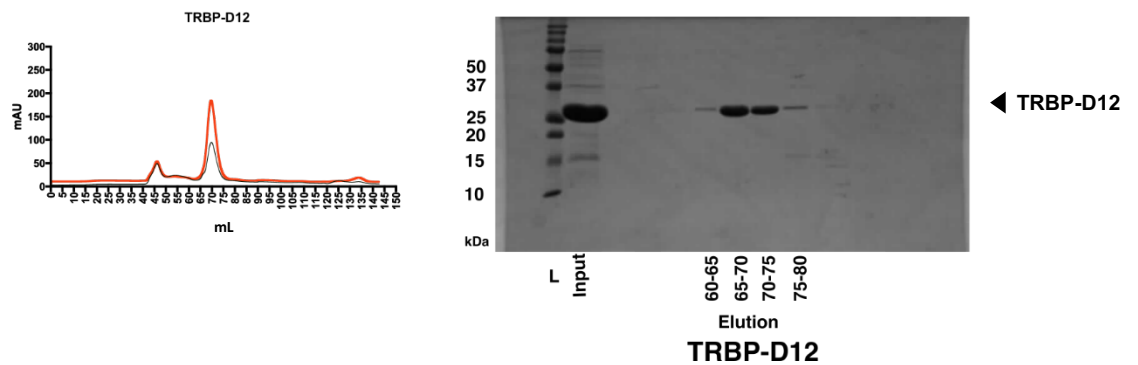
Figure 3.3 Size Exclusion chromatography of TRBP constructs.

A) TRBP-D1 and TRBP-D2 SEC profiles (B) TRBP-D12 SEC profile. 5 mL fractions of the single dominant peaks seen for each protein construct were run on a 15% SDS-PAGE gel for analysis, displayed next to the SEC profiles. Protein ladder (L) Black lines on the SEC profiles represent the absorbance at A260. SE was carried out on Superdex S75 (16/600)

Size exclusion chromatography was carried out on all protein constructs to remove the contaminants that can be seen in the input lane of the SDS-PAGE gels (Figure 3.3). All protein constructs show single dominant peaks in their respective chromatogram.

3.1.4 Summary

The purity of TRBP constructs seen after the final SEC purification step was determined to be sufficient for downstream studies. Single purified bands were seen on SDS-PAGE gels at the correct size for TRBP-D1 and TRBP-D2 (Figure 3.3A). TRBP-D12 can be seen to run at a slightly higher molecular weight than expected on the SDS-PAGE gel. This phenomenon has been seen for this protein construct before, regardless TRBP-D12 appears to be pure with no contaminant proteins after size exclusion chromatography (Figure 3.3B).

3.2 Characterisation of the RNA binding portion of TRBP

To determine whether the protein constructs purified were folded and to test their oligomerization characteristics, various biophysical techniques were conducted. Previous research in the literature has highlighted certain characteristics of these dsRBDs (Benoit et al., 2013; Yamashita et al., 2011).Therefore, it is important to determine whether the purified constructs to be used here have the same properties, so downstream investigations can align with results obtained in the literature.

3.2.1 SEC-MALS shows that the dsRBDs of TRBP are monomeric

TRBP-D1 and TRBP-D2 were examined by SEC-MALS, which is used to determine the absolute molar mass and average size of molecules in solution. A size exclusion step is carried out prior to the multi-angle light scattering procedure to separate the protein of interest from protein aggregates.

(A)

(B)
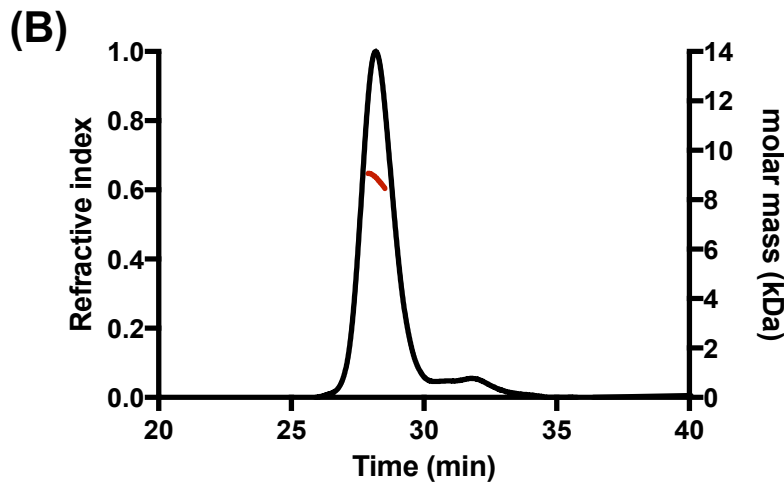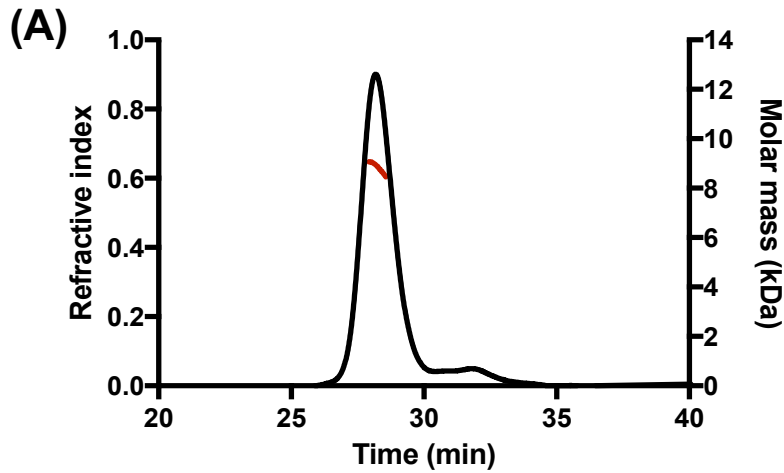
Figure 3.4 SEC-MALS of TRBP-D1 and TRBP-D2 constructs.

The refractive index is shown as a black line with the calculated molecular mass in red. A)SEC-MALS of TRBP-D1 B) SEC-MALS of TRBP-D2.

SEC-MALS of the two N-terminal dsRBDs of TRBP confirms that both dsRBDs are monomeric in solution (Figure 3.4A & 3.4B). Both constructs eluted at a similar volume and the calculated molecular weight by SEC-MALS of TRBP-D1 and TRBP-D2 are 8.9 kDa and 8.8 kDa respectively. These values are consistent with the monomeric molecular mass of each construct (Table 3.1).

3.2.2 NMR spectroscopy of TRBP constructs shows that the proteins are folded

All three TRBP constructs were grown in minimal media containing (^{15}N) labelled ammonium chloride, which will allow for two-dimensional nuclear magnetic resonance spectroscopy (2D NMR) on the proteins to characterise the protein fold. The 2D NMR technique carried out was heteronuclear single-quantum correlation spectroscopy (HSQC), which detects correlations between nuclei of two different types connected by a single covalent bond. In this study, the two bonded nuclei probed are (^{1}H-^{15}N), as each residue within a protein apart from proline has a proton attached to nitrogen in the peptide bond. Therefore, each amide produces a cross peak in the HSQC spectrum representing a specific residue (Marion 2013).

Analysis of the three TRBP constructs by 2D (^{1}H, ^{15}N) HSQC suggests that all constructs are folded as evidenced by the high dispersion of cross peaks in the proton dimension. Moreover, the majority of cross peaks can be individually distinguished (Figure 3.5). In the TRBP-D12 2D HSQC spectrum, overlapping peaks can be seen around the middle of the spectrum. This is most likely due to the long unstructured linker region found between the two dsRBDs (TRBP-D1 and TRBP-D2) (Figure 3.5C).

Figure 3.5 2D (^{1}H,^{15}N) HSQC of TRBP constructs.

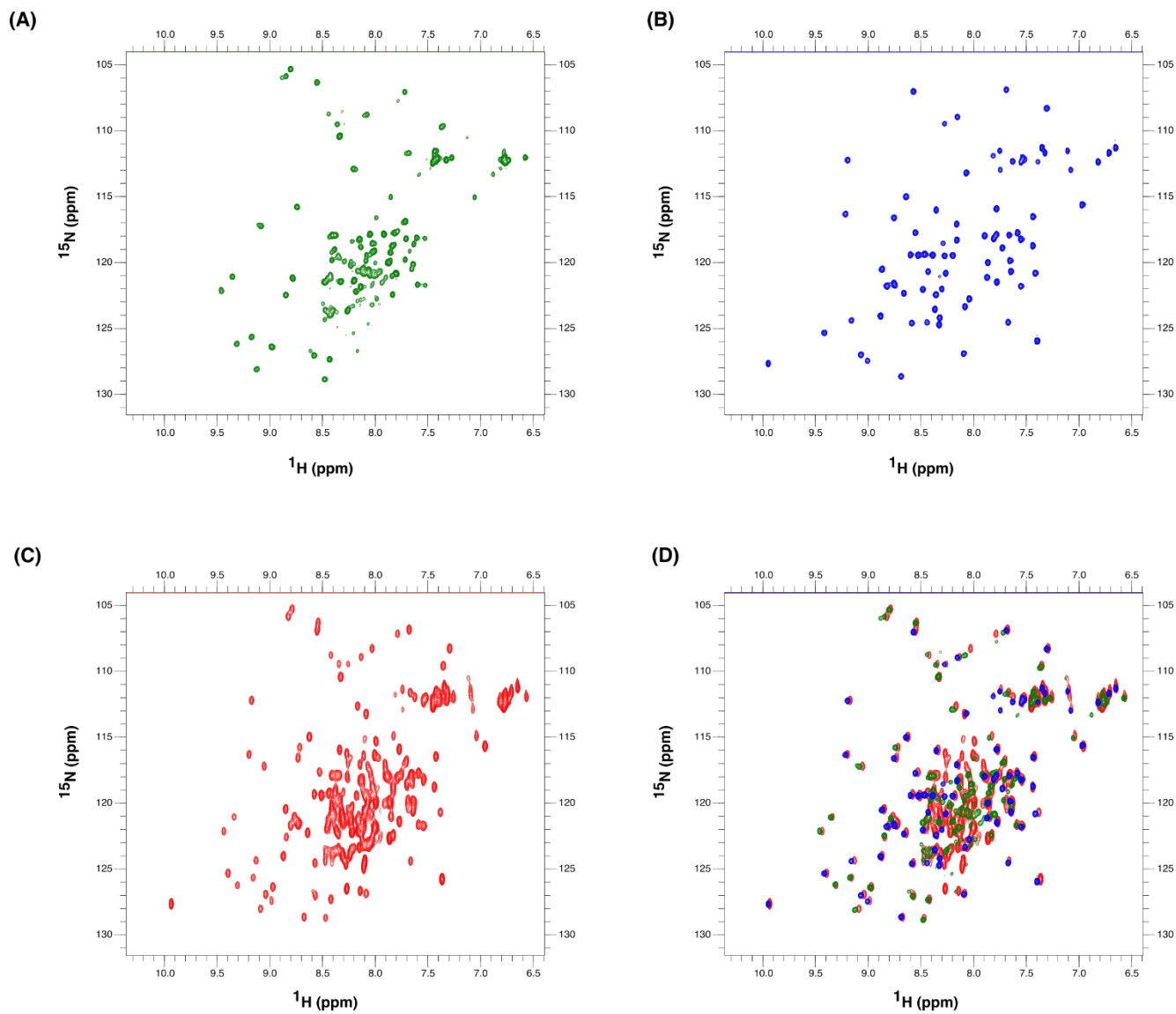2D HSQC spectra of A) TRBP-D1. B) TRBP-D2.C) TRBP-D12. (D) Overlay of TRBP-D1, TRBP-D2 and TRBP-D12. All protein samples were in NMR buffer with a pH of 7.0 (Tris 23.5 mM, KCl 100 mM, MgCl$_2$ 10 mM, DTT 5 mM, D$_2$O 10% and DSS 50 uM).

3.2.3 Summary

SEC-MALS and NMR studies have shown that the RNA binding portions of TRBP (D1 and D2) form stable monomeric proteins. SEC-MALS data is not available for the double domain construct (TRBP-D12) but interpreting the SEC data one can assume the protein is a monomer in solution from its elution volume (Figure 3.3B). NMR data shows all three protein constructs are folded in solution, in the buffer conditions tested. Consistent with previous NMR analysis of these dsRBDs, the NMR data suggests the two dsRBDs do not interact with each other in the TRBP-D12 construct as the cross peaks from the individual dsRBDs overlay nicely with TRBP-D12 with no large shifts in cross peaks (Figure 3.5D) (Benoit et al. 2013; Benoit and Plevin 2013).

These results are consistent with previous biophysical work carried out on TRBP and its dsRBDs.

3.3 Biophysical approach to analysing TRBP-D1 folded/unfolded equilibrium

3.3.1 Introduction

Previous NMR studies of TRBP-D1 have shown that the dsRBD exists in a state of equilibrium between a folded and unfolded protein (Benoit. 2013). This phenomenon was characterised by the observation that the number of non-side chain cross peaks observed was almost two times the number of peaks expected. Two sets of peaks could be distinguished based on their intensities; one minor set of cross peaks with low intensities clustered around the middle of the spectrum suggesting an unfolded species of protein being present and a major set of cross peaks with high intensities dispersed across the proton dimension belonging to the folded species. Previous work into this phenomenon with EXSY experiments detected an exchange between the two sets of peaks, demonstrating that TRBP-D1 exists in equilibrium of folded and unfolded protein.

The phenomenon could not be explained and a solution to get rid of this state of equilibrium was never determined. It is important to address this issue since downstream characterisation of this dsRBD requires the correct concentration of folded and active protein for accurate analysis as proteins require a specific structure and conformation for their function.

3.3.2 NMR suggests the folded/unfolded equilibrium of TRBP-D1 is affected by pH

A set of 2D (^{1}H-^{15}N) HSQC spectra of TRBP-D1 were collected across a pH range from 6.0 to 7.5. At pH 6.5 the approximate number of non-side chain cross peaks seen in the spectrum was significantly higher than the number of cross peaks expected (79 cross peaks) (Table 3.2). With an increase in pH, the number of cross peaks in the spectrum decreases, the set of cross peaks that disappear with an increase in pH are mostly located in the middle of the spectrum suggesting the concentration of the unfolded species of TRBP-D1 diminishing (Figure 3.6). The reduction in the number of peaks in the middle of the spectrum seen with an increase in pH suggests a move into only a single folded species of the protein being present.

Table 3.2 TRBP-D1 – Number of peaks seen in the 2D (^{1}H-^{15}N) HSQC spectra at different pH.

pH of buffer	Number of peaks
6.0	140
6.5	135
7.0	100
7.5	85

(A) TRBP-D1 pH 6.0

(B) TRBP-D1 pH 6.5

(C) TRBP-D1 pH 7.0
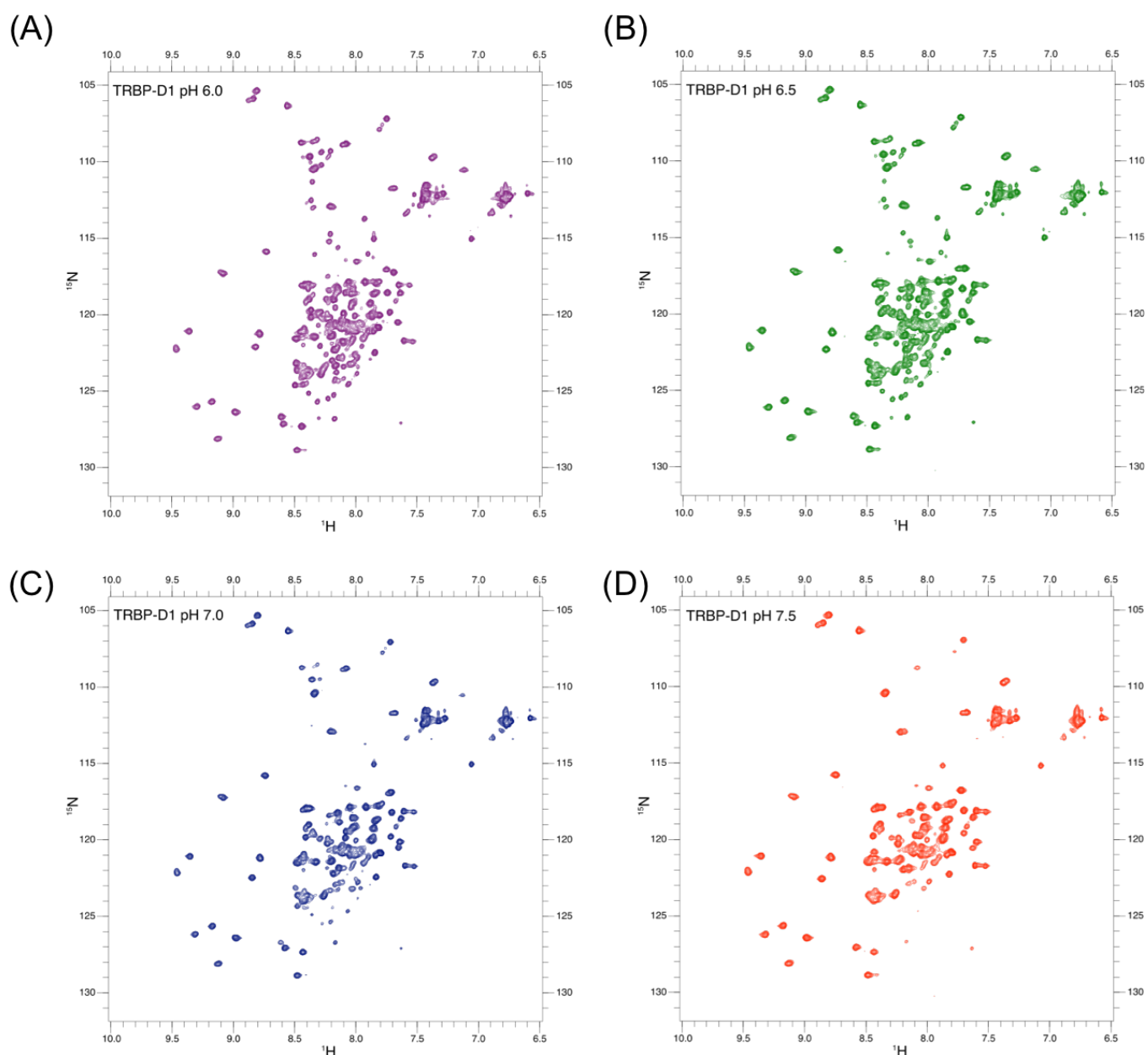
(D) TRBP-D1 pH 7.5

Figure 3.6 2D (^{1}H,^{15}N) HSQC experiments of TRBP- D1 to analyse the effect of pH on the protein stability.

2D (^{1}H,^{15}N) HSQC of A) TRB-D1 at pH 6.0, B) TRBP-D1 at pH 6.5. C) TRBP-D1 at pH 7.0. D) TRBP-D1 at pH 7.5. 2D (^{1}H,^{15}N) HSQC experiments were carried out on a single sample with altering the pH of the sample after the 2D (^{1}H,^{15}N) HSQC was recorded at each pH. Starting concentration of TRBP-D1 was 0.65 mM; the protein concentration changed only minimally as buffering of the sample was achieved with highly concentrated acid and base.
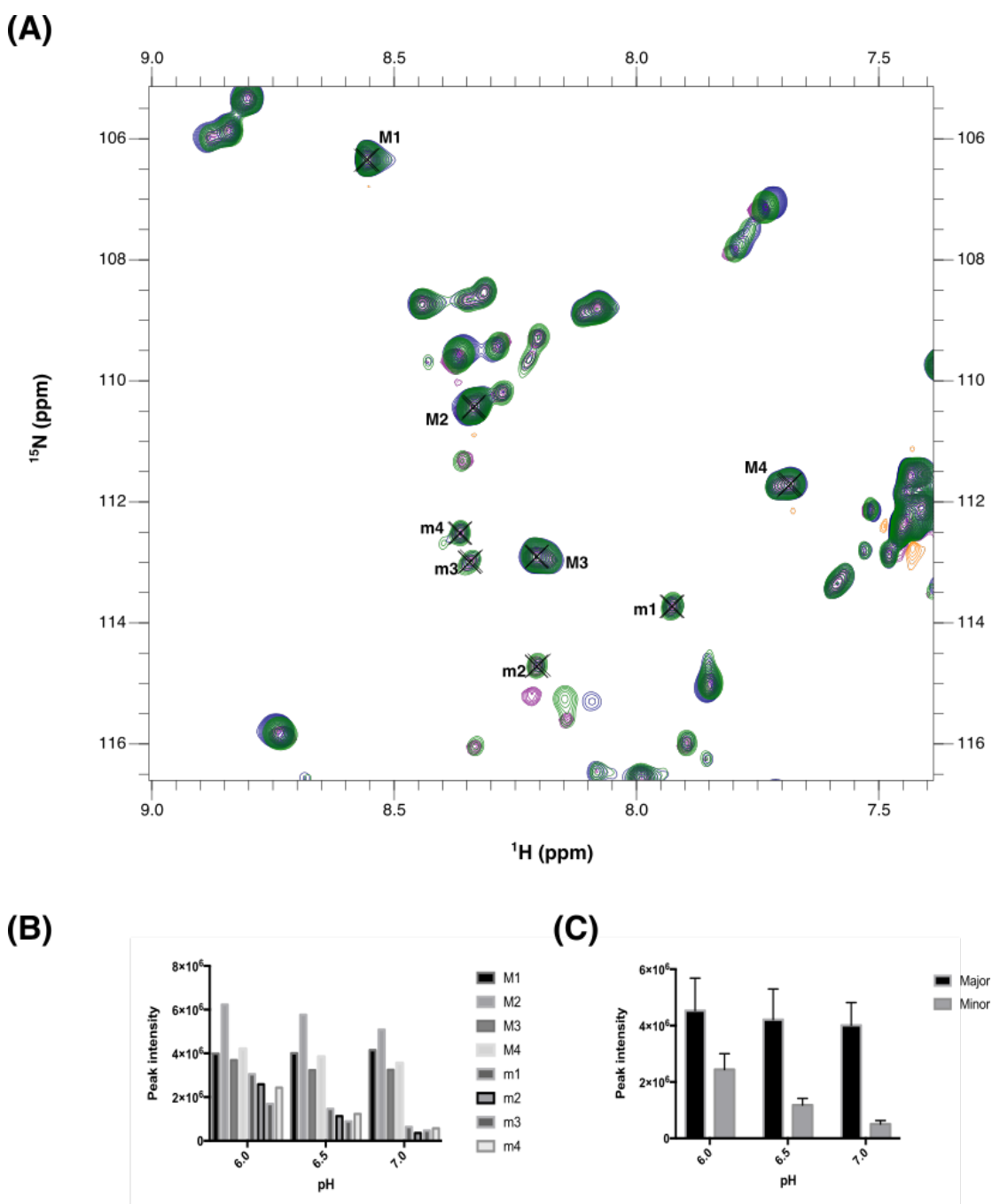
Figure 3.7 Analysis of the folded/unfolded equilibrium of TRBP-D1 using NMR spectroscopy.

A) Cross peaks were picked from the major species (M) and minor species (m) based on the cross-peak intensity. B) Intensity of cross peaks from both the major and minor set species C) Average peak intensity of all the cross peaks in both the major and minor set.

The cross peaks in the 2D (^{1}H,^{15}N) HSQC experiments that disappear with an increase in pH can be distinguished from other cross peaks in the spectrum as they have lower intensities across the whole pH range (Figure 3.7B). The low intensity peaks in the spectrum are denoted as a set of minor species of cross peaks, which represent the residues in the unfolded state of the protein. The high intensity peaks are denoted as the major species of cross peaks representing the protein in the folded state (Figure 3.7A).

To semi-quantitatively investigate the effects of buffer pH on the unfolded/folded equilibrium phenomenon seen for TRBP-D1, a set of cross peaks in both major and minor set were picked, which represent amino acids in both the folded and unfolded state (Figure 3.7A). NMR spectroscopy of TRBP-D1 suggests that an increase in the pH of the NMR buffer reduces the concentration of the unfolded state of the protein being present, seen in the 2D (^{1}H-^{15}N) HSQC experiments. The intensity of the cross peaks in the minor species decreases with an increase in pH and at pH 7.0 the minor set of cross peaks picked as residues in the unfolded state have much lower intensities compared to the cross peaks in the folded state (major set of cross peaks). At pH 7.5 the NMR spectrum displayed no peaks that correspond to residues in the minor species of cross peaks, suggesting that at this pH the protein may exist fully in the folded state. This data suggests that TRBP-D1 can be shifted into the folded state of the folded/unfolded equilibrium with an increase in the pH of the buffer. A caveat to this analysis is that only a few cross peaks were picked to represent the folded/unfolded state of the whole protein. Another problem that the NMR analysis of the TRBP-D1 phenomenon encounters is the NH groups in the unfolded state will exchange with the buffer with an increase in pH more readily than the NH groups in the folded state. What may be observed is the protein not being shifted into the folded state of equilibrium but rather not being able to visualise the minor cross peaks in the spectrum due to exchange. Therefore, further analysis of this phenomenon is required by another approach.

3.3.3 Characterisation of the unfolded/folded equilibrium of TRBP-D1 by CD spectroscopy

Previous NMR studies and the results here indicated that TRBP-D1 exists in a folded/unfolded equilibrium. NMR analysis suggested that at pH 7.5 TRBP-D1 is found in its folded state. To check if this phenomenon and result can be seen with another technique, circular dichroism spectroscopy was carried out.

CD has many applications but most interestingly for this study it can be used to estimate the secondary structure of proteins and the folding properties (Greenfield, 2008). The near UV CD spectrum of proteins provides information on the tertiary structure; the different types of regular secondary structures found in proteins give rise to characteristic CD spectra. Circular dichroism (CD) uses circularly polarized light and depends on the absorption of left and right-handed light by structures within proteins. Different structural elements within proteins have characteristic CD spectra, α-helical proteins have negative bands at 222 nm and 208 nm and a positive band at 195 nm. β helices have negative bands at 218nm and positive bands at 195 nm. Therefore, CD spectroscopy can be used to determine the protein fold by analysing the ellipticity. Disordered proteins have low molar ellipticity above 210 nm, using this property the folded/unfolded equilibrium of TRBP-D1 will be investigated by CD spectroscopy.

Circular Dichroism spectroscopy was conducted on TRBP-D1 with a pH ranging from 5.6 to 7.6. CD spectra were recorded for TRBP-D1 at each pH (Figure 3.8). CD spectroscopy was also conducted on TRBP-D2 as a control. The molar ellipticity of the proteins at 223 nM was calculated, which will give an indication of the fold of the protein. TRBP-D1 can be seen to have a lower molar ellipticity at 223 nM at pH 5.6 and as the pH of the buffer increases the molar ellipticity of the protein increases, suggesting a shift into a more ordered protein conformation (folded state) (Figure 3.8E). The molar ellipticity at 223 nm plateaus after pH 7.0, this plateau suggests

no further changes in the ratio of folded/unfolded protein. The change in molar ellipticity for TRBP-D2 with pH is not as considerable as TRBP-D1.

(A)

(B)

(C)

(D)

(E)

(F)
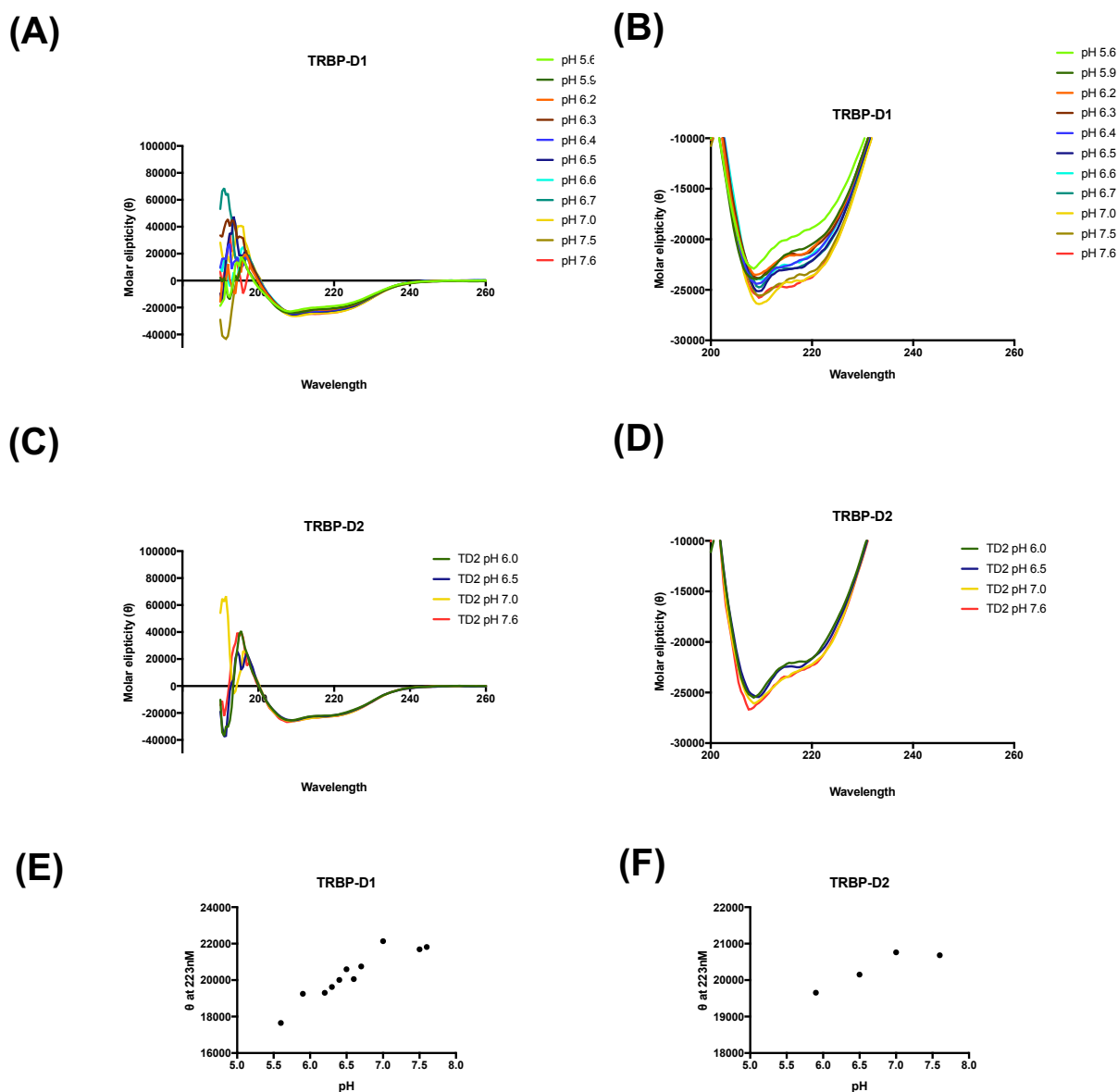
Figure 3.8 Characterisation of the folded/unfolded equilibrium of TRBP-D1 using circular dichroism.

A) Full CD spectra collected of TRBP-D1 in buffer with a pH ranging from 5.6 to 7.6 B) Zoomed in CD spectra of TRBP-D1 to highlight the change in CD wavelength at each pH. C) Full CD spectra collected of TRBP-D2 at various pHs. D) Zoomed in CD spectra ofTRBP-D2 to highlight the change in CD wavelength at each pH. E) Molar elipticity of TRBP-D1 at 223 nM to distinguish protein stability at each pH range. F) Molar elipticity of TRBP-D2 at 223 nM to distinguish protein stability at each pH range.

3.3.4 Summary

TRBP-D1 has been shown to exist in equilibrium between folded and unfolded protein. To compare the binding activity and study the properties of TRBP-D1 compared to other dsRBDs such as TRBP-D2. TRBP-D1 is required to be in its folded and most ordered protein conformation. Through NMR and CD spectroscopy, it has been shown that the equilibrium between folded and unfolded protein is affected by the pH. With a pH greater than 7 in the buffers tested TRBP-D1 exists as a folded species. Therefore, RNA binding studies of TRBP-D1 will be carried out at pH 7.5 to be certain that work will be carried out with an active and folded protein, with no contribution of the unfolded state on RNA binding.

3.4 Characterisation of TRBP mutant constructs

3.4.1 Introduction

To study the RNA binding properties of the dsRBDs found in TRBP, mutant TRBP constructs were created to understand the importance of amino acid sequence in RNA binding regions (Figure 3.9). The residues chosen to study this are found in RNA binding region 2, the research focused on trying to elucidate how specific residues in this region played a role in the structure of the region and its RNA binding property of both dsRBDs.

To check whether point mutations targeted the correct amino acids in a mutant protein construct and to check if the mutation has not disrupted the overall fold and structure of the protein 2D (^{1}H,^{15}N) HSQC experiments were carried out.
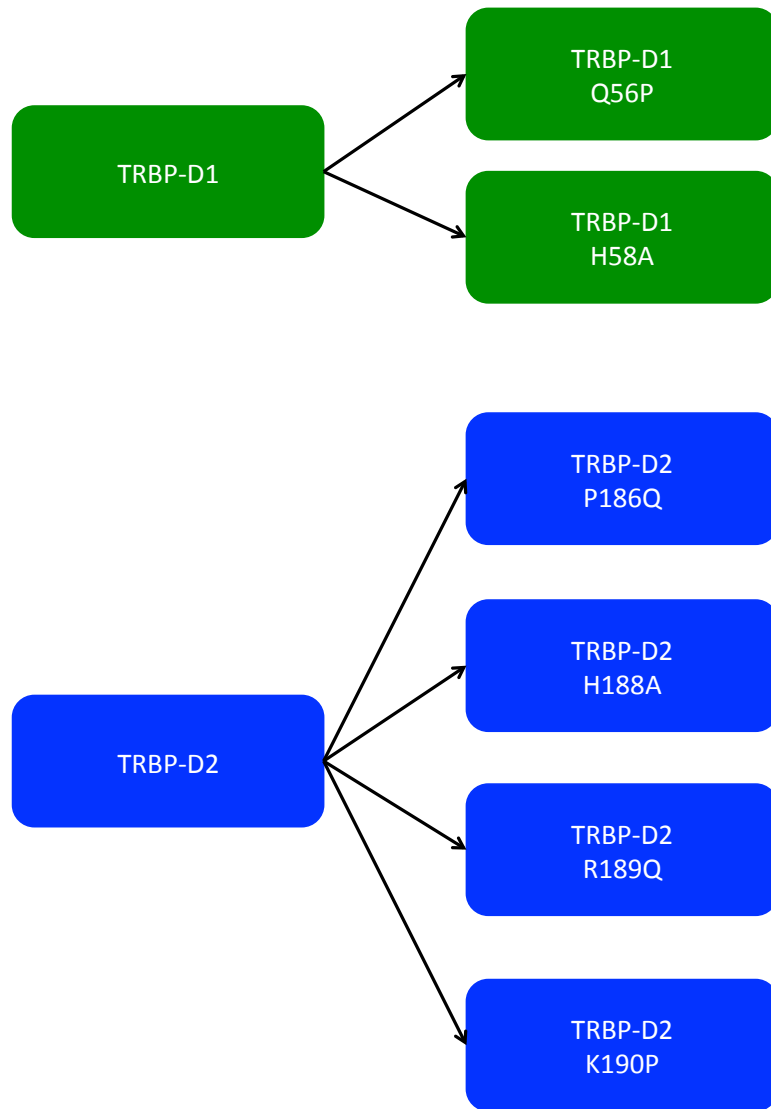
Figure 3.9 Schematics of mutant TRBP dsRBD constructs

Diagram highlighting the point mutations introduced in TRBP-D1 and TRBP-D2 for characterisation of RNA binding region 2.

3.4.2 NMR analysis suggests mutations in RNA binding region 2 do not disrupt the fold of the protein

^{15}N labelled TRBP wild type (WT) and mutant constructs were produced and purified at an appropriate concentration for NMR spectroscopy (Fig 3.10).

2D ($^{1}H,^{15}N$) HSQC experiments of TRBP- D1 WT and the mutant constructs show that the protein constructs are folded due to the dispersion of peaks in the protein dimension and mutations have not disrupted the overall fold of the protein domain (Figure 3.10). The experiments were collected with a buffer at pH 7.0 to reduce the effects of unfolded state seen for TRBP-D1, which was still seen in the mutant constructs. The specific pH of 7.0 was used to reduce chemical exchange of the amide groups of flexible regions of the folded protein specifically at the site of the mutations.
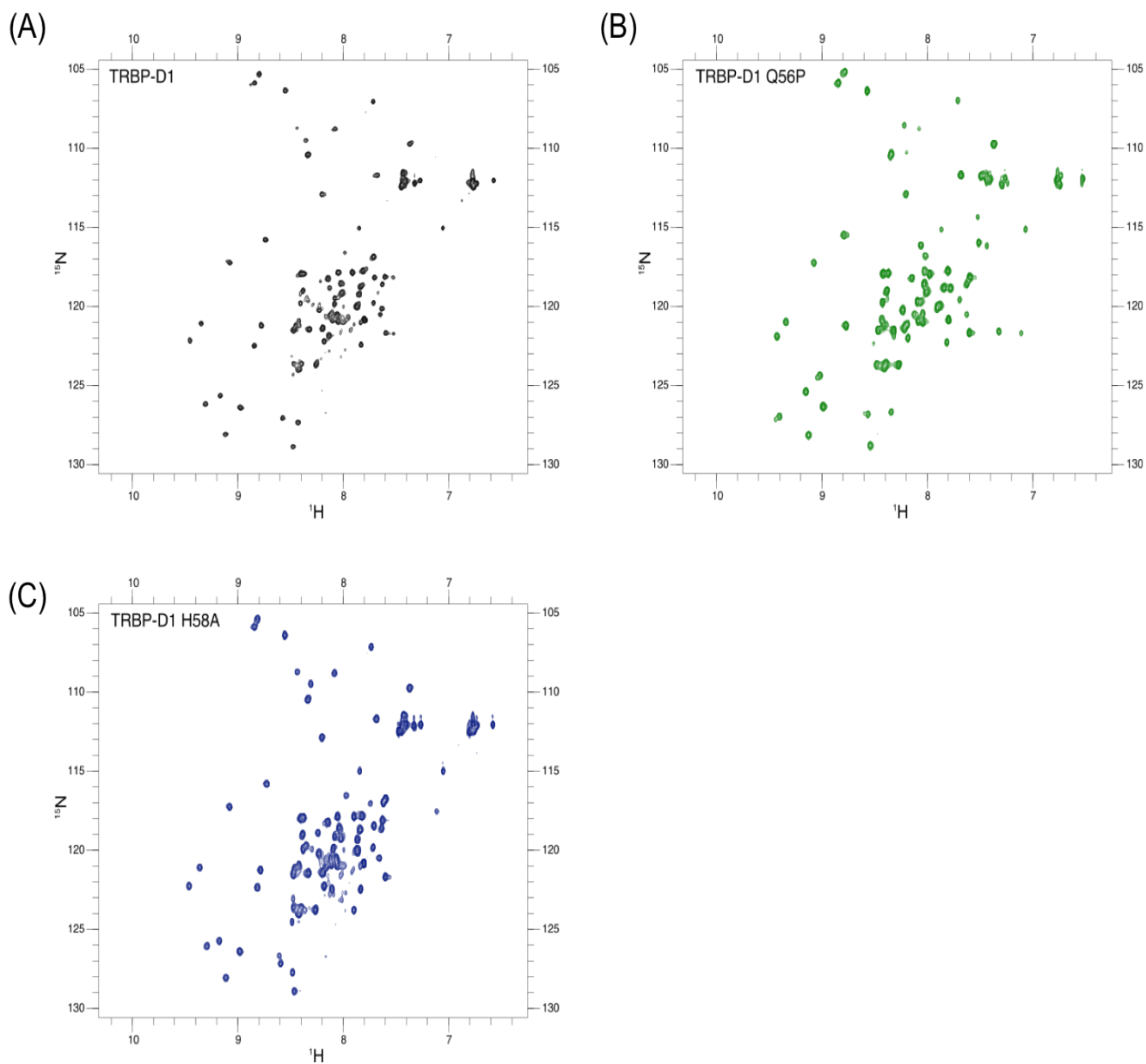
Figure 3.10 2D (^{1}H,^{15}N) HSQC spectra of TRBP-D1 WT and mutant constructs.

2D (^{1}H,^{15}N) HSQC spectra of A) TRBP-D1 WT B) TRBP-D1 Q56P C) TRBP-D1 H58A.

2D (^{1}H,^{15}N) HSQC experiments of TRBP- D2 wild type and the mutant constructs show the protein is folded which can be seen by the dispersion of peaks in the proton dimension (Figure 3.11). Difference in cross peak positions can be seen for the mutants compared to the wild type protein in only a few positions indicating the mutation has affected the residue of interest and the surrounding residues, with not much effect on the whole protein itself, this is to be expected as the mutations carried out are in a flexible loop between two β−strands.
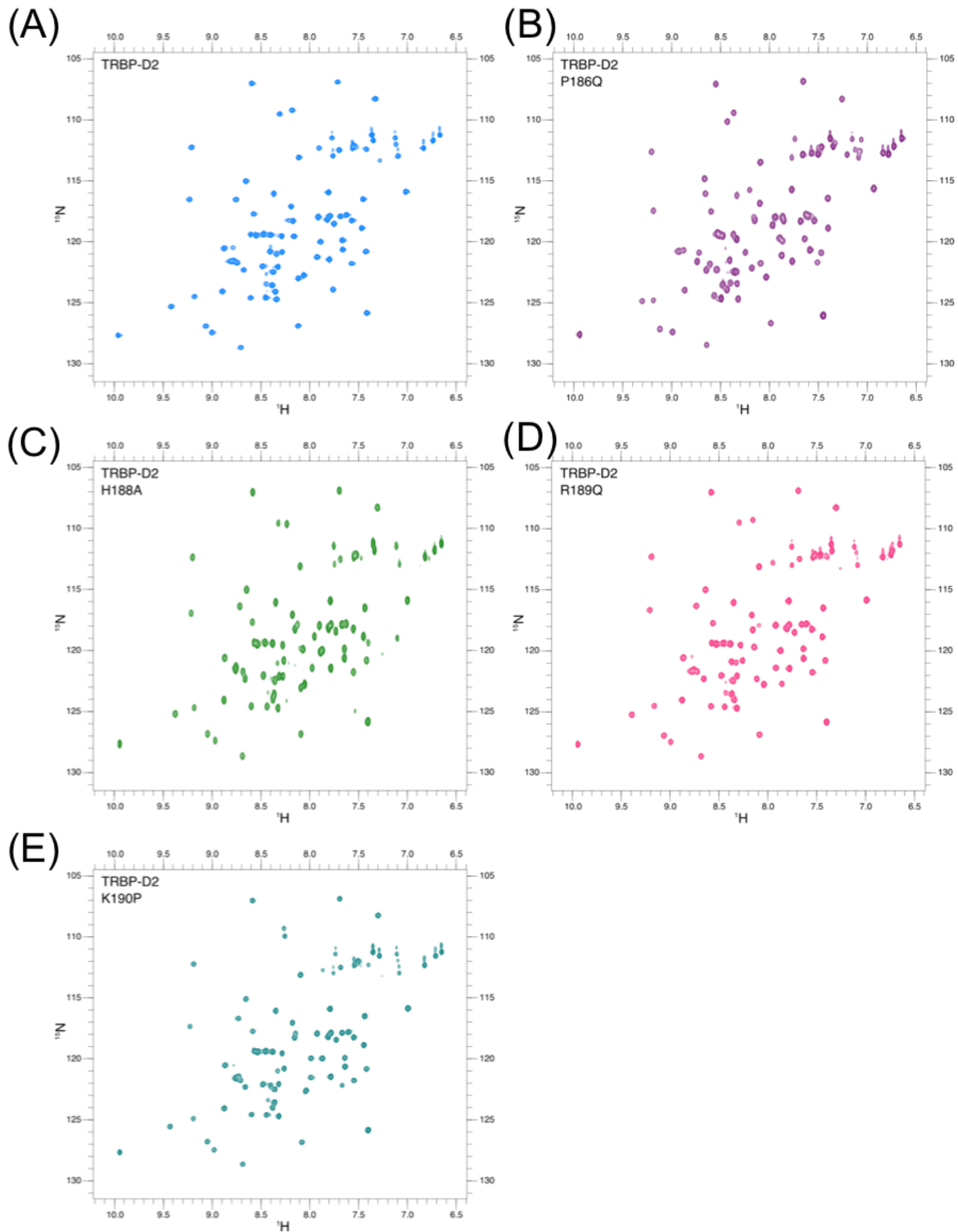
Figure 3.11 2D (^{1}H,^{15}N) HSQC spectra of TRBP-D2 wt and mutant constructs.

2D (^{1}H,^{15}N) HSQC spectra of A) TRBP-D2 WT B) TRBP-D2 P186Q C) TRBP-D2 H188A. D) TRBP-D2 R189Q E) TRBP-D2 K190P.

3.4.3 Summary

NMR analysis of each construct indicates that although mutations have been carried out, the protein maintains its overall fold. All mutant constructs show well dispersed cross peaks in the 2D (^{1}H,^{15}N) HSQC spectrum, which is ideal for further in-depth NMR studies of these protein constructs. Further examination on the impact of the mutations on the protein requires assignment of the cross peaks to precisely identify the structural regions affected.

3.5 Production of RNAs for RNA binding studies

3.5.1 Introduction

There are various methods for producing RNA, using organic chemistry reactions or enzymatic reactions to produce RNA. The most commonly used method is an enzymatic reaction that employs the use of an RNA polymerase to transcribe RNA strands from a template DNA strand. This method is cheap and efficient in producing high yields of RNA. Three commonly used polymerases with high specificity for their 23 base bair promoter are SP6, T3 and T7 RNA polymerases. After production of RNAs *in vitro*, purification of the RNA away from contaminants such as free nucleotides, the DNA template and the polymerase is required. Which can be carried out using simple RNA clean up kits or more advanced techniques such as anion exchange for higher purity.

3.5.2 DNA template production

In vitro transcription requires pure linear DNA template containing a promoter region for the polymerase enzyme. Consequently, DNA templates were prepared from two oligonucleotides annealed together to form a dsDNA template, each DNA template was flanked with the T7 polymerase initiation sequence. The two oligonucleotides were mixed in DNA annealing buffer and heated for 5 min at 95°C and then incubated at RT for 30 minutes for hybridization of the oligonucleotides to occur (Figure 3.12)
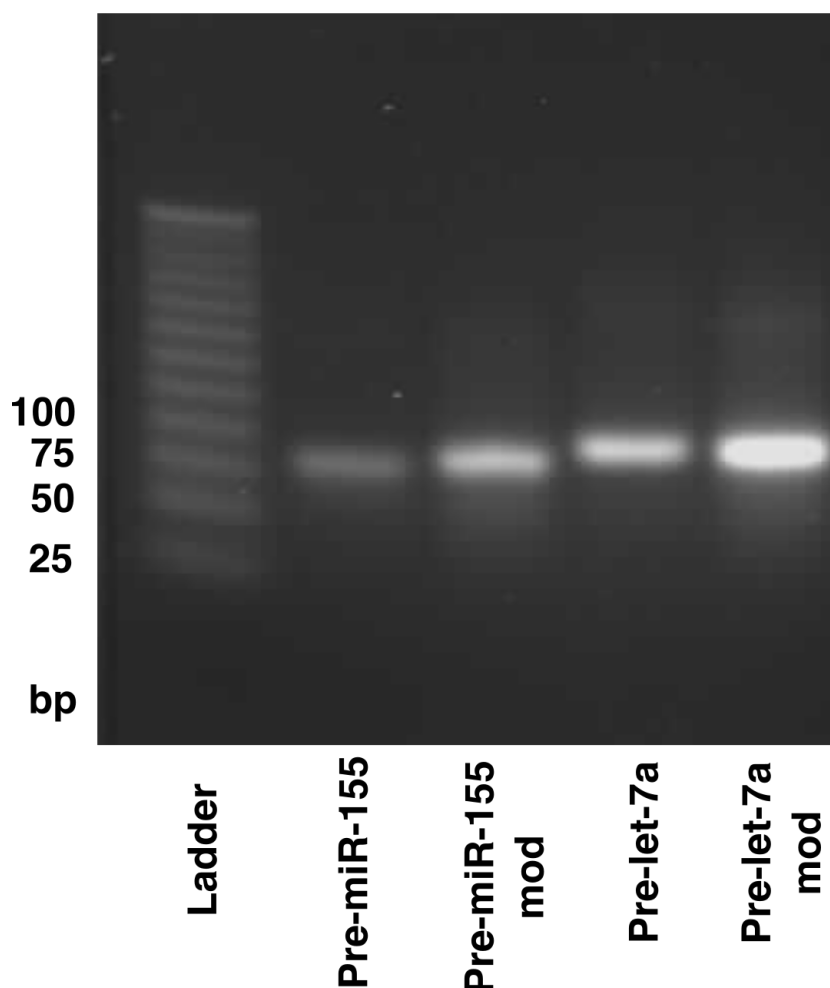
Figure 3.12 DNA template production

2% Agarose gel of DNA templates for in vitro transcription produced by annealing two oligonucleotides.

3.5.3 In vitro transcription of RNAs

To determine the best conditions for RNA production through *in vitro* transcription using T7 polymerase, the concentration of $MgCl_2$ required for a reaction was tested. Small-scale reactions were set up with varying concentrations of $MgCl_2$ to determine the best concentration required for increased yields of RNA production.

The RNA band of interest can be seen as the over stained band around the 50 nt marker for 155 WT and 155 modified (155 mod) and 80nt for let7a WT and let7a modified (let7a mod) (Figure 3.13). The best $MgCl_2$ concentration for each RNA construct was determined by identifying the highest yield of RNA production from the intensity of the RNA band on the denaturing gel.
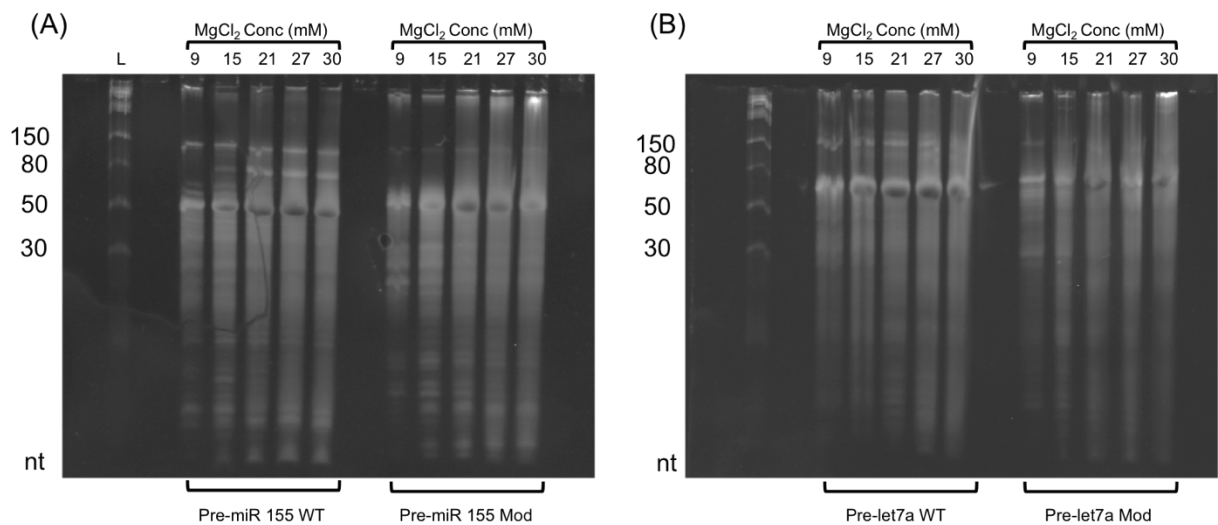
Figure 3.13 MgCl₂ concentration test for IVT reactions.

15% urea denaturing gel analysis of RNA yield dependent on MgCl₂ concentration. A) Gel of Pre-miR 155 and Pre-miR 155 mod small scale in vitro transcription reactions B) Gel of Pre-let7a and Pre-let7a mod RNAs small scale in vitro transcription reactions. The concentration of MgCl₂ used is noted on top for each lane. (L) is the RNA ladder used to check the RNA produced is the correct size.

Following determination of appropriate concentration of MgCl$_2$, 5 mL reactions of each RNA construct were carried out. IVT reactions were incubated in a water bath at 37°C for 4 hours. To increase RNA yield, extra T7 polymerase and MgCl$_2$ were added at two hours into the reaction. Reactions were stopped by freezing the reaction at -20 °C. Large scale purification of dsRNAs were then carried out using anion exchange chromatography (Figure 3.14).

(A)

(B)

Elution Volume (ml)
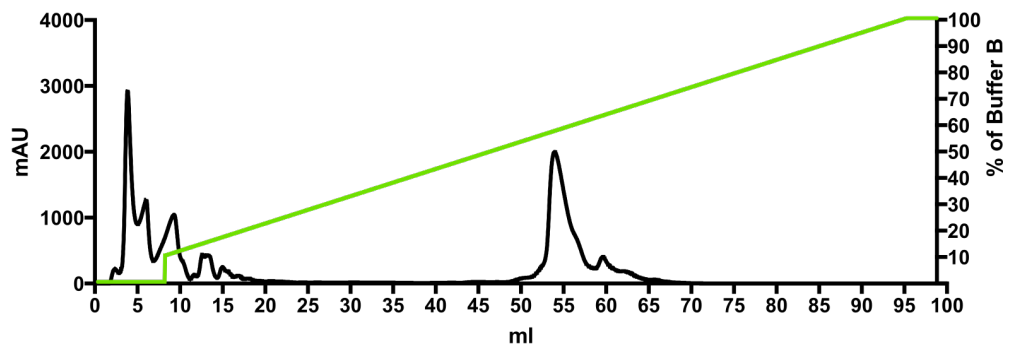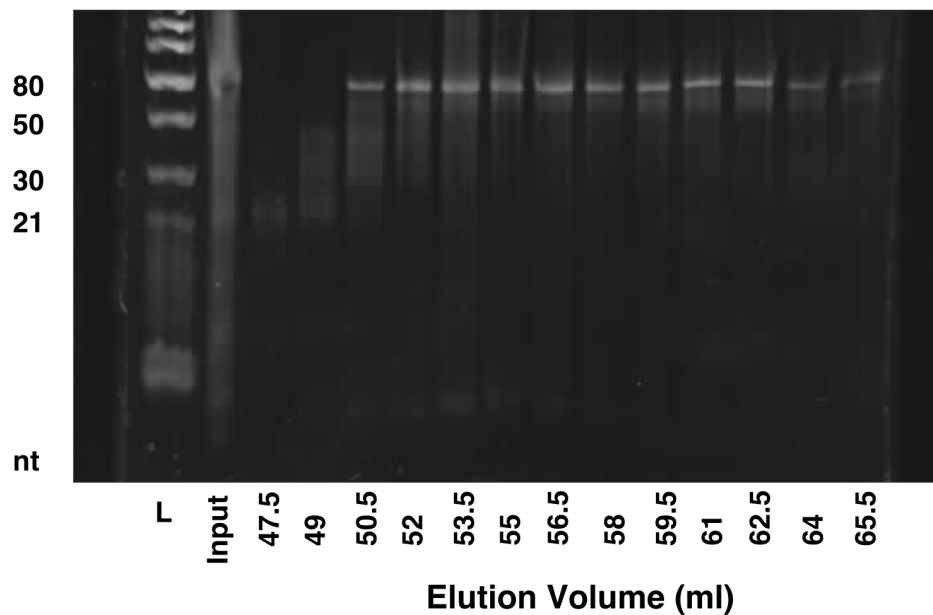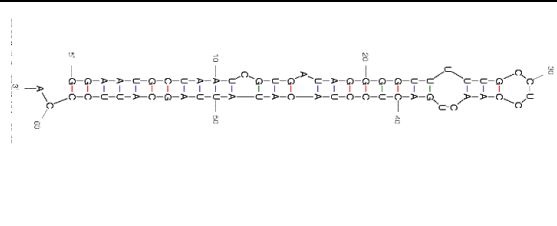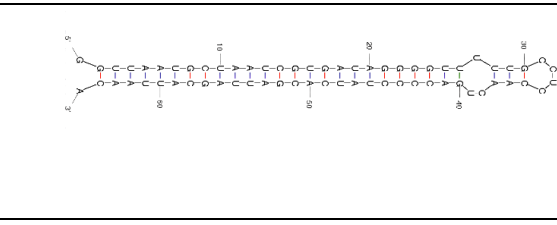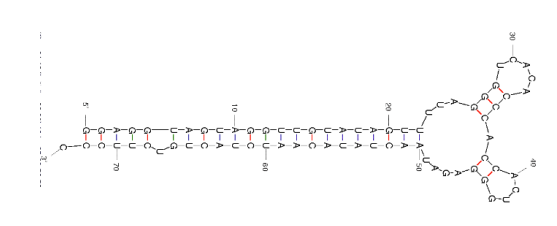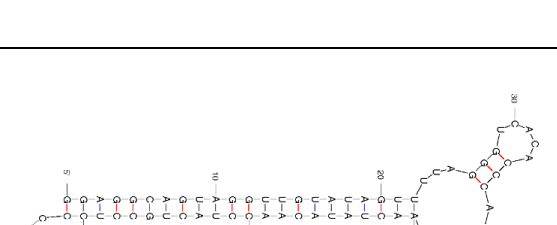
Figure 3.14 Large scale purification of in vitro produced Pre-let7a.

A) Anion exchange profile for Pre—let7a B) 15 % urea denaturing gel of 1.5 mL fractions eluted from the anion exchange purification. Anion exchange was carried out on a Mono Q 5/50 GL column. The black line represents the absorbance at A260 nm. The green line represents the concentration of Buffer B (Tris-HCl pH 7.0 20 mM, MgCl$_2$ 10 mM and 1 M NaCl).

Table 3.3 DsRNA constructs

DsRNA structures were predicted using Mfold.

dsRNA	Number of nucleotides	MgCl$_2$ conc for IVT (mM)	Structure
Pre-miR-155	61	21	
Pre-miR-155 mod	65	21	
Pre-let-7a	73	21	
Pre-let-7a mod	73	21	

3.5.4 End labelling of dsRNAs

Scientific studies of nucleic acids to determine the function and to study their properties often require labelling of the nucleic acid with chemical tags or radioisotopes. Labelling of oligonucleotides can be carried out during synthesis or post synthesis using an enzyme-catalysed reaction. They can be end labelled or nucleotides can be internally modified. For downstream studies to investigate the RNA binding properties of the dsRBDs found in TRBP, dsRNAs investigated were labelled on the 3' end through post-synthetic conjugation of a chemical tag (Zearfoss and Ryder, 2012). Figure 3.15 shows successful labelling of Pre-let7a and Pre-miR 155 constructs that were used for downstream RNA binding assays.

(A)   **(B)**
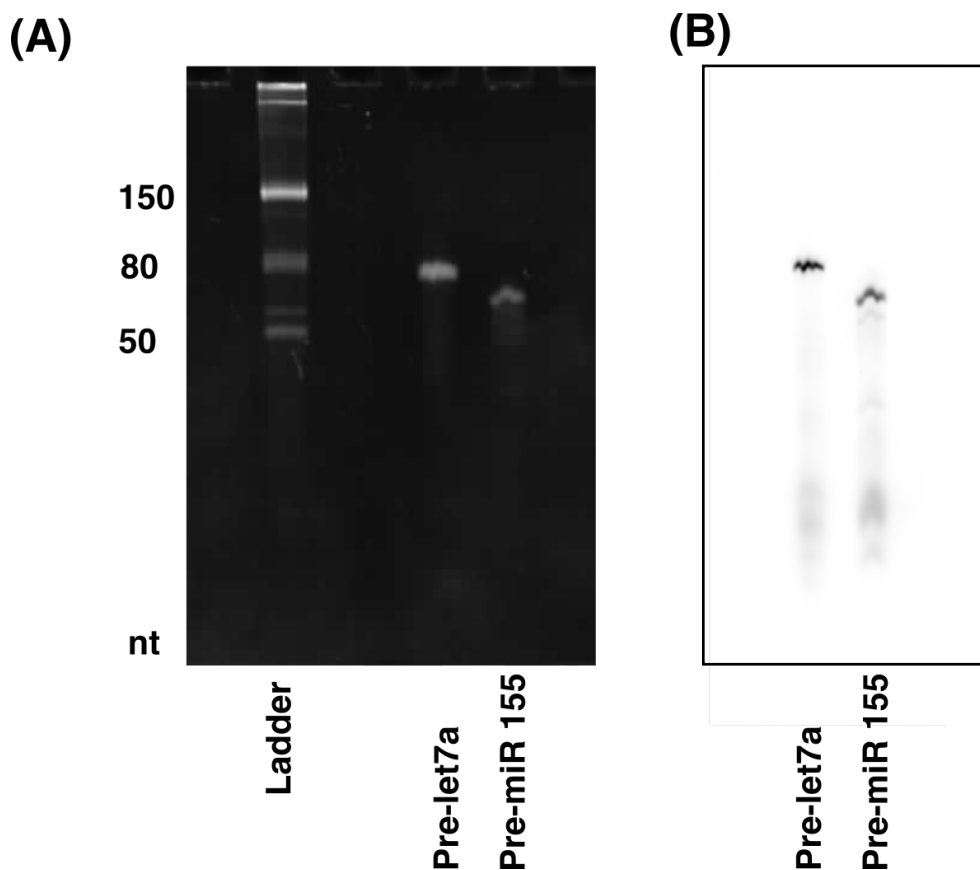
Figure 3.15 3' end labelling of dsRNAs with fluorescein 5-thiosemicarbazide.

15% urea denaturing gel of dsRNA constructs visualised two ways. A) Detection of dsRNA using SYBR Gold B) Detection of the fluorescent label FTSC.

Chapter 4- Characterisation of the type-A dsRBDs of TRBP by NMR spectroscopy

4.1 Introduction

The unique structure of a protein is determined by its amino acid sequence; however, proteins carry the ability to occupy ensembles of conformations and are not strictly static entities. Local motion within a protein may be essential for its function and investigations into protein dynamics can be crucial to understanding the mechanism of protein function (Yang et al. 2014). This chapter will explore the protein dynamics and local motions of the type-A dsRBDs of TRBP.

NMR spectroscopy has been widely used to study dsRNA-binding proteins, from the singular dsRBDs of a protein to the whole multi-domain protein. NMR spectroscopy has been used to elucidate solution structures, molecular dynamics and molecular interaction regions of dsRBDs and linker regions found in dsRBPs (Stefl et al. 2010a) Bycroft et al., 1995; Lu and Hall, 1997; Gabel et al., 2006; Nanduri et al., 2000; Masliah et al., 2018; Kharrat et al., 1995; Mueller et al., 2010; Benoit et al., 2013; Benoit and Plevin, 2013; Chiliveri and Deshmukh, 2014). Motional dynamics of dsRBDs have been shown to play an important role in their function (Agarwal and Barnwal 2009). For example, the interferon-inducible antiviral kinase PKR (PKR) has been shown through NMR spectroscopy to contain two dsRBDs that exhibit a difference in motional flexibility on a millisecond timescale, which has been suggested to be the reason for a difference in RNA binding activity seen between the two dsRBDs. The motionally more flexible PKR-dsRBD1 has a higher RNA binding activity compared to PKR-dsRBD2 (Yang et al. 2014).

To date only minimal dynamic data has been elucidated for TRBP especially for the dsRNA binding regions. Previous work by Benoit et al

highlighted through (^{1}H-^{15}N) heteronuclear NOE experiments that the type-A dsRBDs of TRBP confer similar rigidity across each dsRBD. However, further analysis of the data displayed a difference in flexibility in the loop region between β1-β2 in the dsRBDs (Benoit et al. 2013). Interestingly this region is vital for RNA binding of the dsRBD and what was seen is that TRBP-D2 has a more rigid β1-β2 loop compared to TRBP-D1. This led to questioning how this difference between these two dsRBDs exists and what role does this phenomenon play in the binding of dsRNAs. To find answers for these questions and further study the dynamics of the dsRBDs in TRBP, mutagenesis of this loop region and in-depth NMR dynamics experiments were carried out.

NMR spectroscopy is a powerful tool for studying proteins; compared to other structural techniques such as X-ray crystallography and electron microscopy, NMR provides information on the protein structure in the solution state therefore it can be employed to obtain vital information about the dynamics of a protein. NMR spectroscopy permits the analysis of specific residues and atoms (Palmer 2004) .

^{15}N NMR relaxation studies can provide information relating to the dynamics and motions within a protein on a ps to ms timescale. NMR relaxation experiments are influenced by the global tumbling of a molecule but also by local motional dynamics within the protein structure (Palmer 2001; Jarymowycz and Stone 2006).To study dynamic motions within a protein, accurate measurement of various relaxation parameters is required. NMR relaxation experiments were collected at a single magnetic field strength of 800Mhz to analyse the motional properties of the type-A dsRBDs found in TRBP. Spin-lattice relaxation (T_1), Spin-spin relaxation time in the rotating frame ($T_1\rho$) and (^{1}H-^{15}N) Heteronuclear-NOE experiments were collected.

4.2 Backbone assignment of TRBP-D12

To examine the structure and dynamics of proteins using NMR spectroscopy it is first important to acquire resonance assignment of proteins, to distinguish which chemical shift in a spectrum corresponds to which atom within a protein sequence. Therefore, this chapter will begin by discussing the attempt to assign the NMR spectrum of the double domain protein construct TRBP-D12, the difficulties that were encountered and how they were solved to produce a partially assigned NMR spectrum of TRBP-D12.

The backbone assignment of proteins is commonly carried out using 3D triple resonance experiments where three nuclei are probed ^{1}H, ^{13}C, ^{15}N. Triple resonance experiments have the ability to transfer the magnetization over the peptide bond to probe the amide, carboxyl group and the side chain carbons of an amino acid but also the carbonyl groups preceding that specific amino acid (Ikura et al. 1990). Hence, triple resonance experiments can be carried out to sequential link each (^{1}H-^{15}N) cross peaks. Triple resonance experiments such as HNCACB can then be used to identify corresponding cross peaks to a specific amino acid in the protein sequence by analysing the chemical shifts of Cα and Cβ cross peaks linked to a (^{1}H-^{15}N) cross peak.

Table 4.1 Triple resonance experiments carried out for backbone assignment of TRBP-D12

Experiments	Nuclei probed	
	i :nuclei's connected to NH group of amino acid	i-1 :nuclei's connected to NH group of previous amino acid
HNCO	X	CO
HN(CA)CO	CO	CO
HNCACB	CA,CB	CA,CB
HN(CO)CACB	X	CA CB
HNCA	CA	CA
HN(CO)CA	X	CA

Following initial protein stability tests of TRBP-D12 in different buffer conditions and temperatures, [^{13}C, ^{15}N] labelled TRBP-D12 was prepared at ~0.5 mM and a standard set of triple-resonance backbone assignment experiments was recorded (Table 4.1).

The 2D (^{1}H-^{15}N) TROSY experiment at pH 6.5 of TRBP-D12 contained approximately 290 cross peaks, which is much higher than the expected number of cross peaks of 194 (Figure 4.1). The extra peaks are due to the previously seen effect of TRBP-D1 existing in equilibrium between two states in chapter 3. The presence of TRBP-D1 in two states complicated the interpretation of the NMR spectra for TRBP-D12 but does not hinder the assignment procedure as the two states could be distinguished as a major and minor species of peaks as seen previously for TRBP-D1. NMR assignment experiments were carried out at pH 6.5 to help identify and assign the maximum number of chemical shifts possible for the protein.
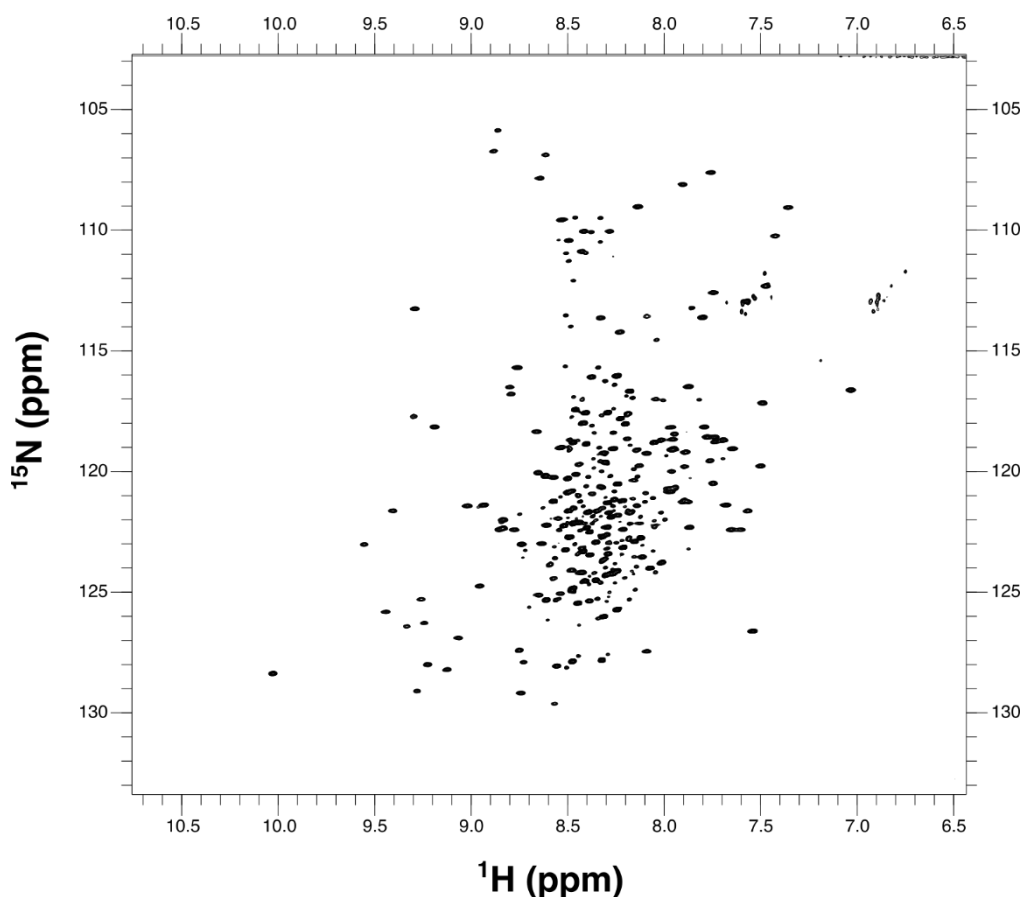
Figure 4.1 2D (^{1}H, ^{15}N) TROSY of WT TRBP-D12.

Assignment of specific nuclei probed in the triple resonance backbone assignment experiments, HNCA, HNCACB, HNCO, HCOCAB were attempted for TRBP-D12. The set of minor cross peaks seen (unfolded state of TRBP-D1) in the 2D (^{1}H,^{15}N) TROSY experiments did not possess corresponding cross peaks in all triple resonance experiments carried out. Interestingly in the major set of cross peaks not all contained corresponding cross peaks in the HNCACB or HNCOCACB experiments. Most of them contained cross peaks for the Cα in the HNCA experiments but additional magnetization transfer via the Cβ led to a significant loss of detectable signal in the HNCACB experiments. The cross peaks in the 2D (^{1}H,^{15}N) TROSY expected to be from the structured dsRBDs elucidated from the overlap of the single dsRBDs with TRBP-D12, displayed no NMR signals in the 3D HNCACB triple resonance experiments. A potential reason for not being able to see cross peaks in the 3D triple resonance experiments

could be due to the poor signal to noise ratio of the experiments. The poor signal to noise observed for these experiments is most likely due to the fast decay of NMR signals due to the slow tumbling of the large protein in solution (~22 kD). Several approaches could have been attempted to overcome the problems with the signal to noise seen for TRBP-D12. For example, deuterium labelling of the protein, which improves the signal-to-noise by decreasing the relaxation rate of the ^{13}C and ^{15}N spins (Sattler and Fesik 1996). However, another approach was conducted, which was to carry out assignment experiments on the single domains and then assign the linker region using the original TRBP-D12 triple resonance experiments. To assist with the backbone assignment of the double domain protein TRBP-D12, single dsRBDs proteins with single point mutations between the β1-β2 loops were produced.

4.2.1 Backbone assignment of TRBP-D1 Q56P

[^{15}N,^{13}C] labelled TRBP-D1 Q56P was prepared at a concentration of ~1.1 mM and 3D HNCOCACB, HNCACB and HNCA triple-resonance backbone assignment experiments were recorded. TRBP-D1 Q56P is a mutation carried out in RNA binding region 2 of the dsRBD to mimic the amino acid composition of TRBP-D2 for downstream studies of this region.

Figure 4.2 presents an assigned (^{1}H,^{15}N) TROSY of TRBP-D1 Q56P. It contains 110-120 peaks including side chain amides, marginally higher than the number of peaks predicted from the sequence, which is due to the two-state phenomenon seen for TRBP-D1. The peaks in the unfolded state could be distinguished from the peaks in the folded state by comparing the intensities of the peaks. 83% of backbone amide resonances in the folded state were assigned, with 83% of Cα spins and 82% of Cβ spins.

Figure 4.2 Assigned 2D (1H,15N) TROSY spectrum of TRBP-D1 Q56P

Assigned cross peaks are labeled with the residue number and the amino acid type.

The secondary structure of TRBP-D1 Q56P can be predicted from the NMR chemical shifts of H,N,CO, Cα and Cβ using TALOS-N(Shen and Bax, 2013). TALOS-N can predict dihedral angles from the chemical shifts it is presented using a database of measured or predicted chemical shifts from X-ray structures. The TALOS-N secondary structure prediction (Figure 4.3) is in good agreement with secondary structures of TRBP-D1 based on crystallographic and NMR data, which suggests that it adopts the same fold and that the assignment of TRBP-D1Q56P was successful (Yamashita et al. 2011; Masliah et al. 2018).

Figure 4.3 TALOS-N prediction of TRBP-D1 Q56P

Secondary structure prediction by TALOS-N using HN, Cα and Cβ chemical shifts. Coloured bars show the prediction confidence for each structure type. Above the graph is the secondary structure from a known NMR structure (PDB accession 5N8M)

The N-terminal region of TRBP has recently been shown to contain a small alpha helix preceding the canonical αββα fold of dsRBD (Masliah et al. 2018; Paithankar et al. 2018). The TALOS-N prediction has also highlighted this structure in the protein construct of TRBP-D1 used this study (Figure 4.3).

4.2.2 Backbone assignment of TRBP-D2 P186Q

[^{15}N,^{13}C] labelled TRBP-D2 P186Q was prepared at a concentration of ~1.2mM and a standard set of triple-resonance backbone assignment experiments were recorded, 3D HNCOCACB, HNCACB and HNCA experiments. TRBP-D2 P186Q is a mutation carried out in RNA binding region 2 of the dsRBD to mimic the amino acid composition of TRBP-D1 for downstream studies of this region.

Figure 4.4 presents an assigned 2D (^{1}H,^{15}N) TROSY spectrum of TRBP-D2 P186Q. It contains ~90 peaks including side chain amides, close to the number of peaks predicted from its sequence. 93% of backbone amide resonances were assigned, with 92% of Cα spins and 91% of Cβ spins.
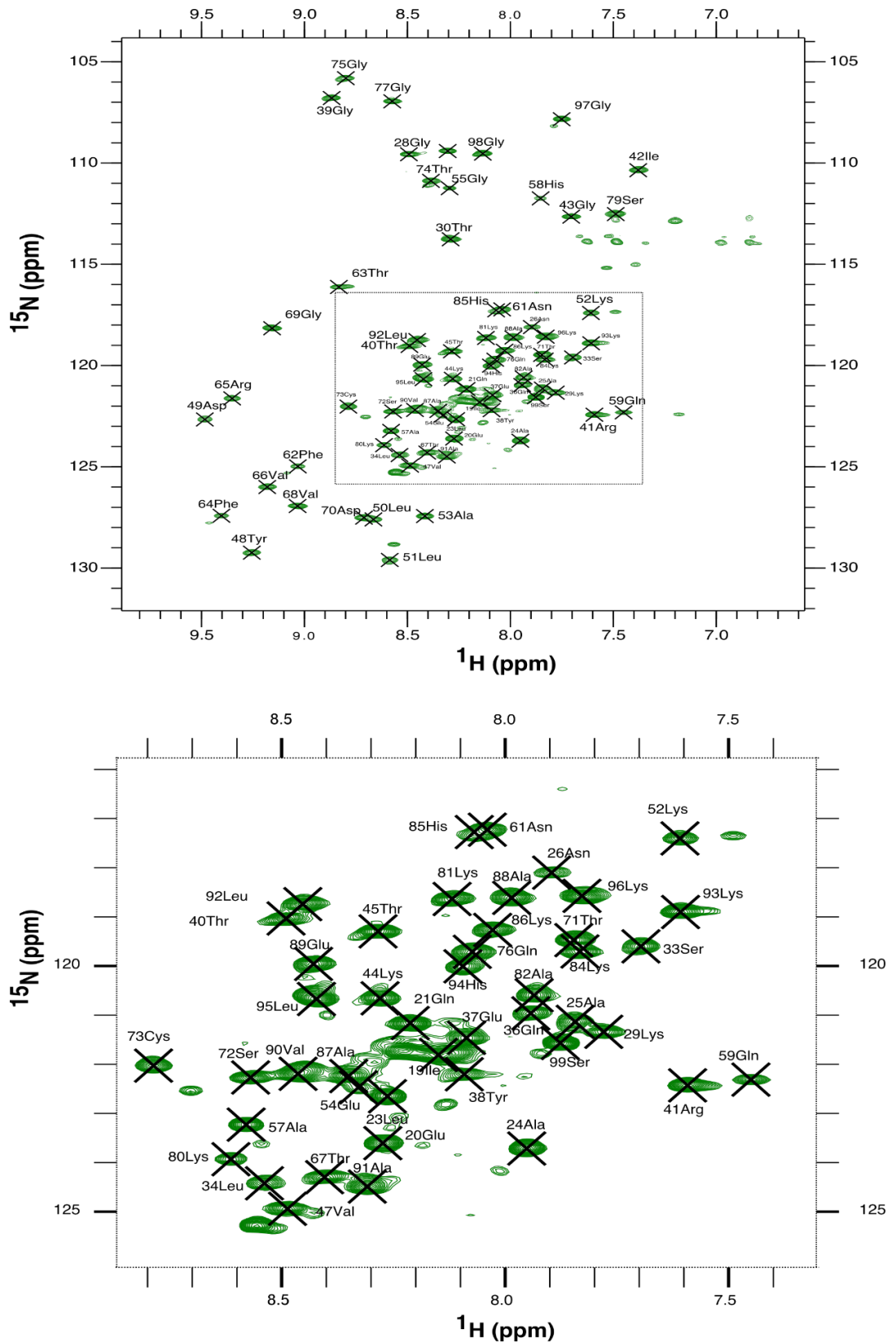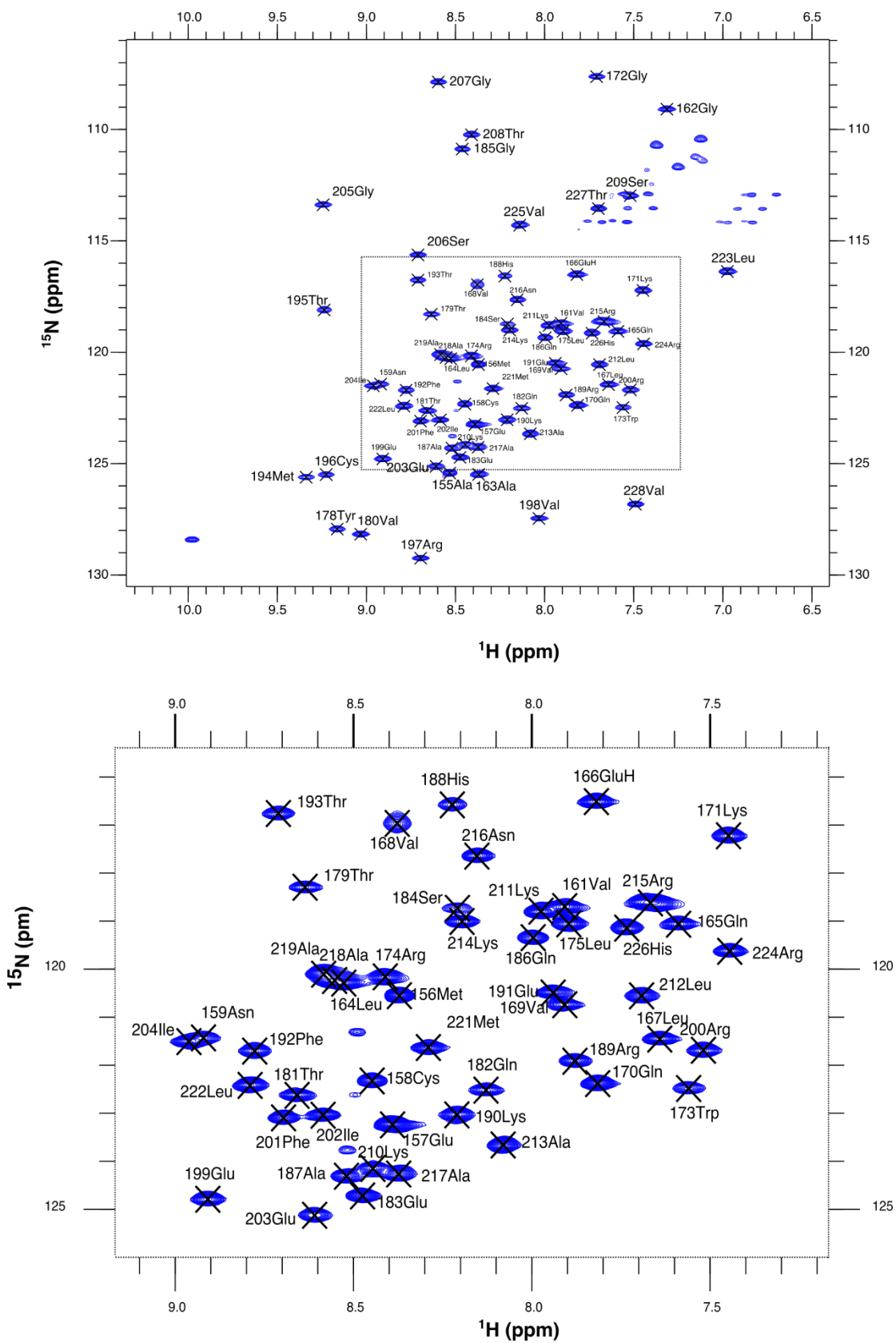
Figure 4.4 Assigned 2D (^{1}H,^{15}N) TROSY spectrum of TRBP-D2 P186Q

Assigned cross peaks are labeled with the residue number and the amino acid type.

Using TALOS-N the secondary structure of TRBP-D2 P186Q was predicted from the Cα and Cβ chemical shifts collected. The TALOS-N secondary structure prediction (Figure 4.5) is in good agreement with the secondary structure based on crystallographic and NMR data for TRBP-D2, which suggests that the assignment of TRBP-D2 P186Q was successful (Yamashita et al. 2011; Masliah et al. 2018).
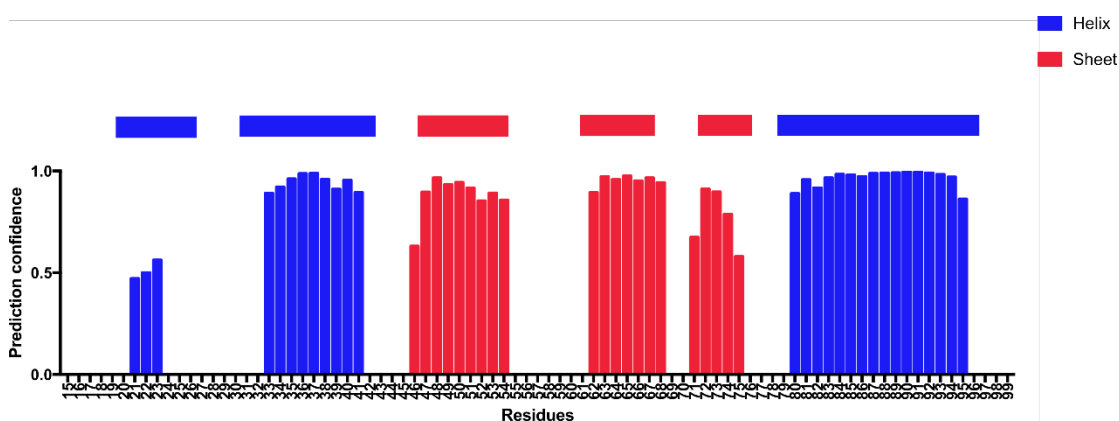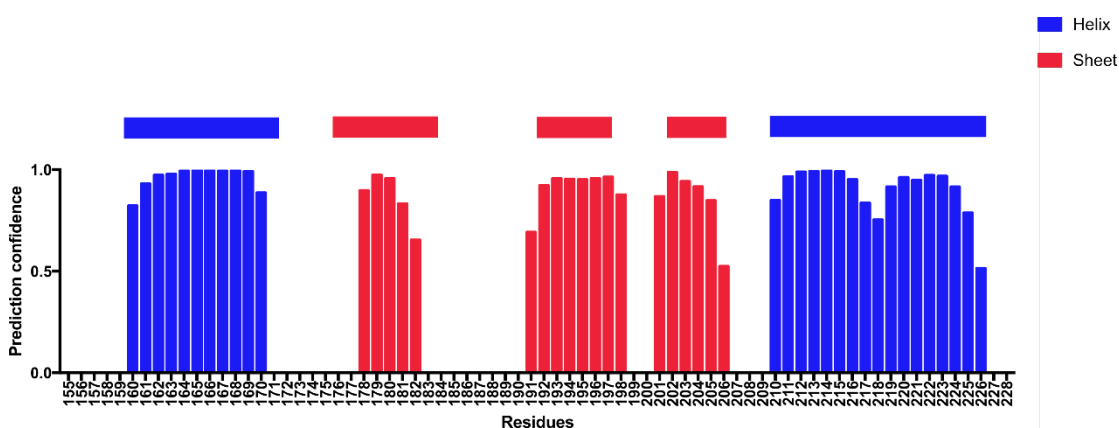

Figure 4.5 TALOS-N prediction of TRBP-D2 P186Q

Secondary structure prediction by TALOS-N using HN, Cα and Cβ chemical shifts. Coloured bars show the prediction confidence for each structure type. Above the graph is the secondary structure from a known NMR structure of TRBP-D2 (PDB accession 5N8M).

4.2.3 Backbone assignment of TRBP-D12 WT

With the aid of cross peak assignments from the single dsRBDs, most of the chemical shifts of the dsRBDs in the TRBP-D12 spectrum were assigned (Figure 4.7). However, due to peak overlap in the TRBP-D12 spectrum transfers of peak assignments were difficult and not all the peaks in each dsRBD were assigned in TRBP-D12. The initial triple resonance experiments of TRBP-D12 helped to identify and assign some of the linker peaks but due to the bad signal to noise not all peaks were assigned. The percentage of peaks assigned for D1, D2 and the linker were 67%,90%,26% respectively. Figure 4.6 identifies the peaks that have been assigned for TRBP-D12.
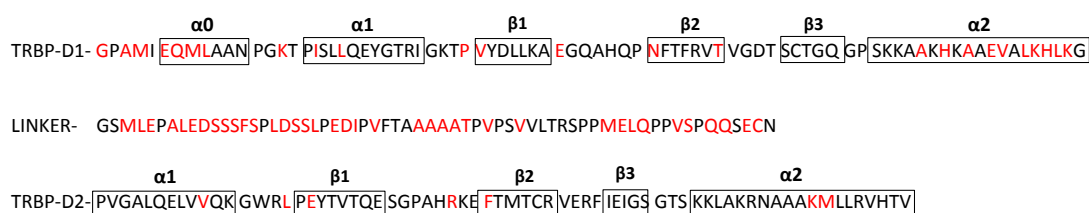
Figure 4.6 Cross peaks assigned for TRBP-D12

Representation of TRBP-D12 peaks that were assigned using TRBP-D12 assignment experiments and transfer of assignments from the single dsRBDs, TRBP-D1 Q56P and TRBP-D2 P186Q. In red are residues which have not been assigned.
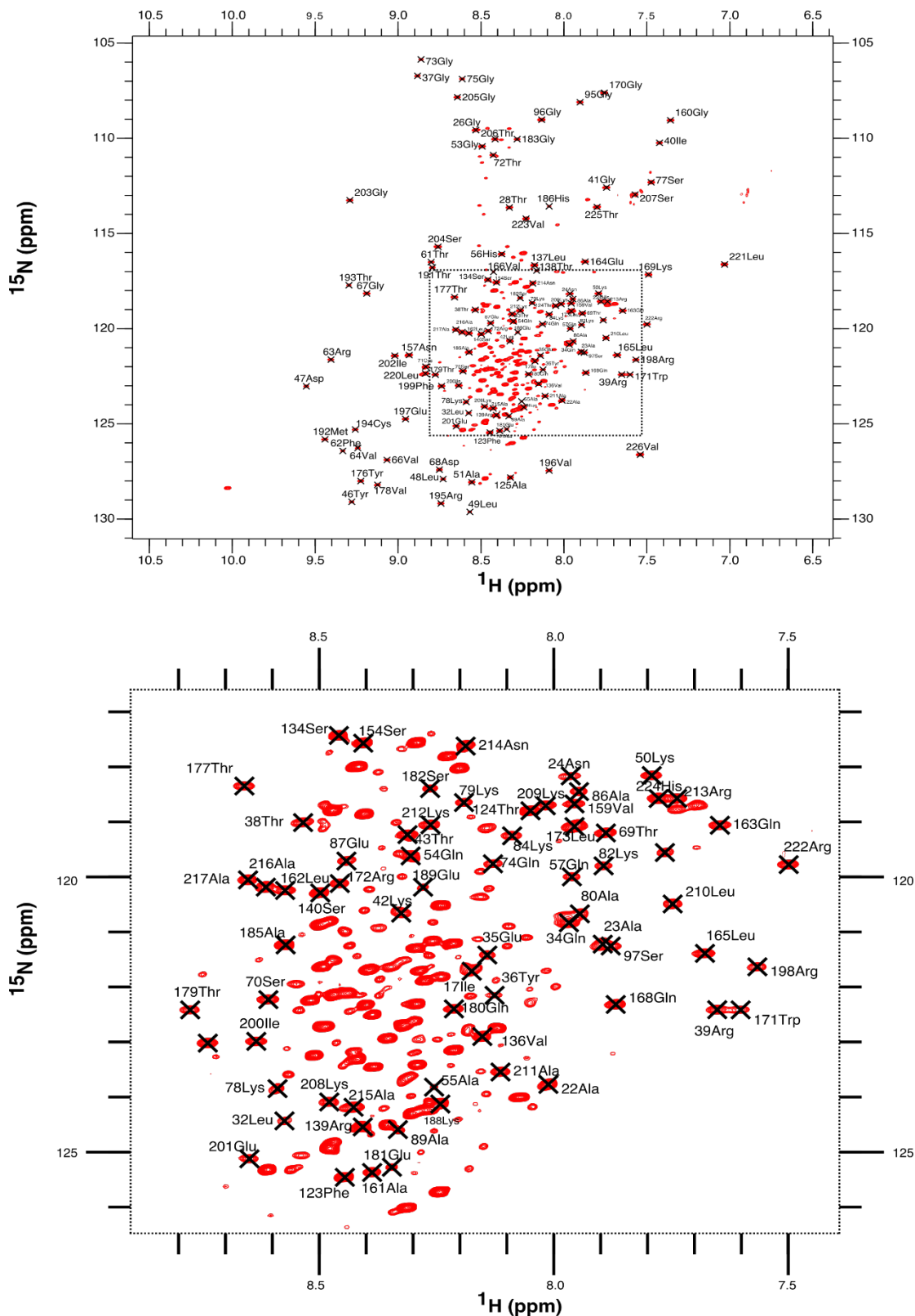
Figure 4.7 Assigned 2D (^{1}H,^{15}N) TROSY spectrum of TRBP-D12 WT

Assigned peaks are labeled with the residue number and the amino acid type. Bottom panel is the centre of the spectrum enlarged. The minor set of cross peaks belonging to the unfolded state of TRBP-D1 were not assigned.

4.3 Characterisation of TRBP mutant dsRBDs by NMR spectroscopy

2D (^{1}H,^{15}N) HSQC experiments were carried out on mutant dsRBD constructs of TRBP to examine whether point mutations in the dsRBD has affected the overall fold of the protein construct. Compound chemical shift differences between the amide groups of the WT and mutant constructs were calculated, which will identify the amino acids affected by each single point mutation. Assignments were transferred from the wild type protein to the mutant constructs where possible. Compound chemical shift differences were calculated using the Equation below (Williamson 2013).

$$(\Delta\delta = [\Delta\delta H^2 + (\Delta\delta N/6.5)^2]^{1/2}$$

4.3.1 Characterisation of TRBP-D1 mutants

Compound chemical shift differences were calculated between TRBP-D1 WT and the mutant constructs TRBP-D1 Q56P and TRBP-D1 H58A (Figure 4.8). Chemical shift values where the matching residue in the mutant spectra could not be assigned because of an unidentifiable shift of the peak were given an arbitrary value of below zero. Chemical shift differences seen between the WT TRBP-D1 and TRBP-D1 Q56P highlighted major chemical shift differences concentrated around the β1-β2 loop, around the site where the glutamine residue was mutated for a proline (Q56P). This can be perceived through a chemical shift difference value of 0.3 and higher. Major chemical shift differences could be seen for TRBP-D1 H58A in the β1- β2 loop, perceived through not being able to link matching peaks in the mutant construct without assigning TRBP-D1 H58A.
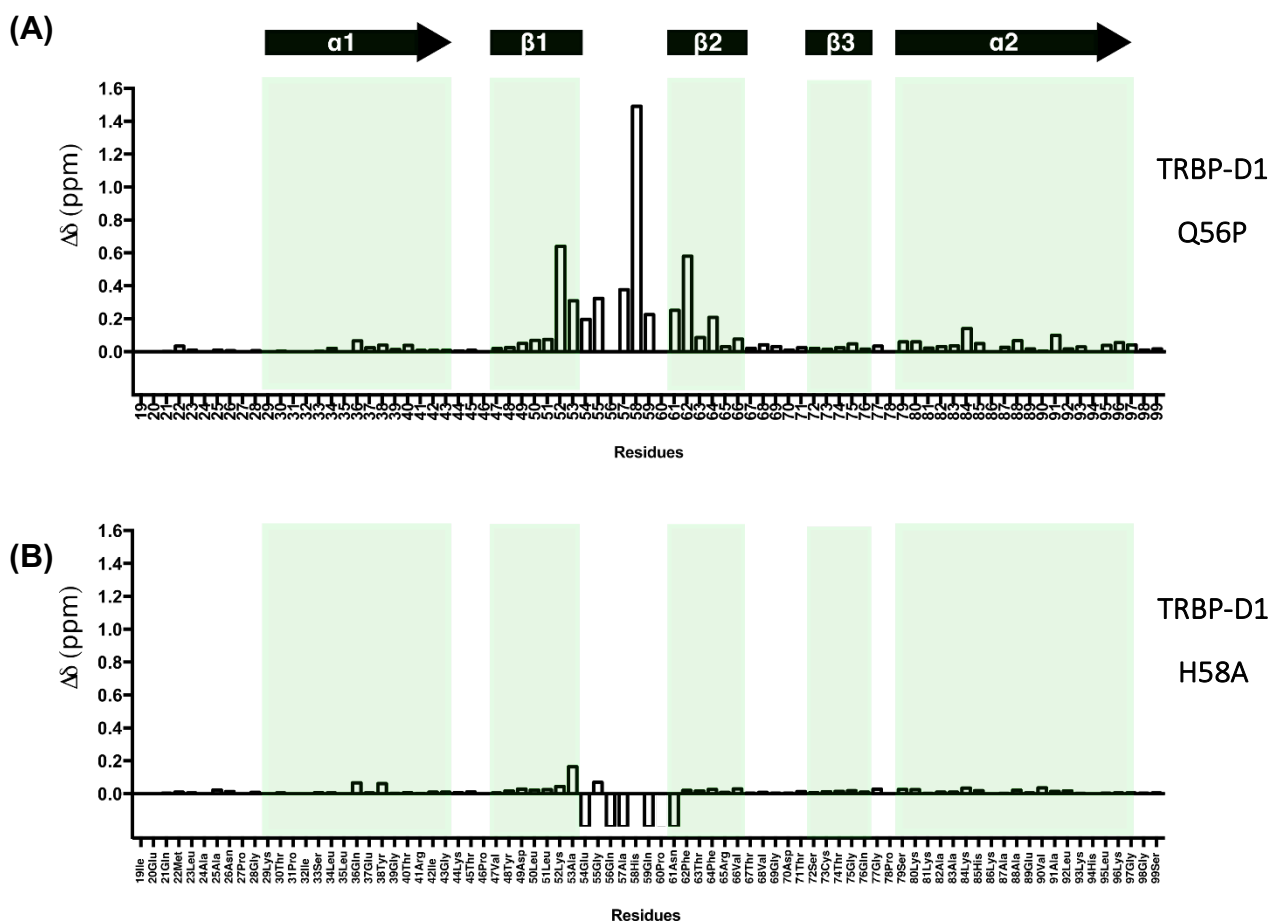
Figure 4.8 Chemical shift differences between TRBP WT and mutant constructs.

(A) Compound chemical shift differences of TRBP-D1 Q56P. (B) Compound chemical shift differences of TRBP-D1 H58A. Shift in cross peaks that were not able to be distinguished were given an arbitrary value of below 0.

4.3.2 Characterisation of TRBP-D2 mutants

Major chemical shift differences in the TRBP-D2 mutant constructs can be observed in the residues surrounding the single point mutation for each construct as expected. Large chemical shift differences can be seen for mutants TRBP-D2 P186Q (Figure 4.8A) and TRBP-D2 K190P (Figure 4.8D) in the β1- β2 loop compared to the other mutant constructs, which could be due to the removal of a proline residue in the P186Q mutation and an addition of a proline residue in the K190P mutant, changing the backbone structure or dynamics and therefore the chemical environment of the residues in this loop region.
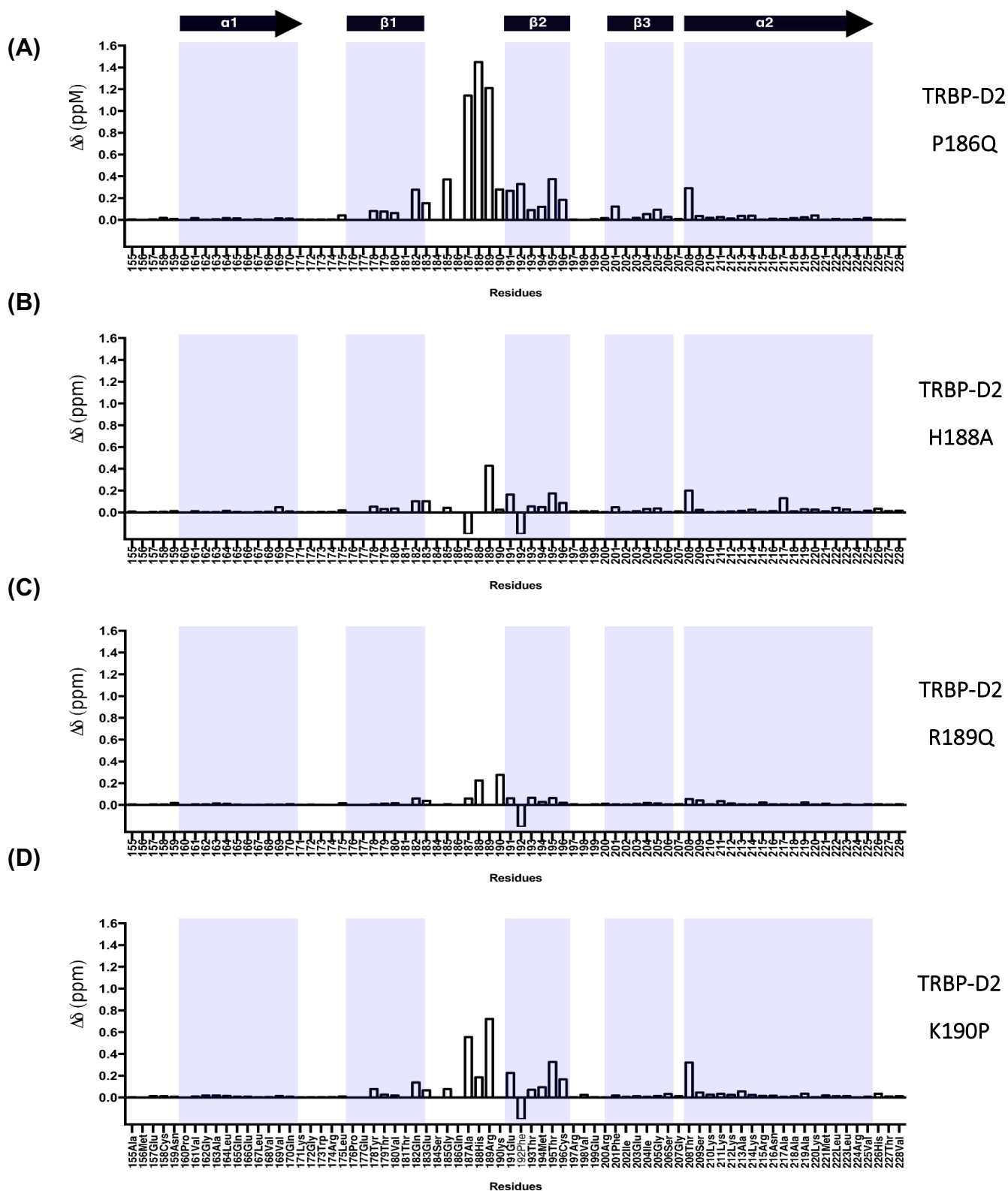
Figure 4.9 Chemical shift differences between TRBP-D2 WT and mutant constructs

(A) Compound chemical shift differences of TRBP-D2 P186Q. (B) Compound chemical shift differences of TRBP-D2 H188A (C) TRBP-D2 R189Q (D) TRBP-D2 K190P. Shift in cross peaks not distinguishable were given an arbitrary value of below 0.

Interestingly a large chemical shift difference can be visible for a residue not close in the amino acid sequence for example threonine 208. From inspection of the structure of TRBP-D2, this residue may not be close in sequence but is close in proximity to the loop region in the structure (Figure 4.10). There could be different possibilities for why Thr-208 may be affected by the changing of the β1- β2 loop, one could be that the modification in loop residues changes the structure of the loop and therefore decreases the distance between the residues in the loop region and 208Thr, affecting the chemical environment of the NH group of 208Thr. Another explanation could be that the residues found in the β1- β2 loop could be interacting with 208Thr and therefore changing the amino acid composition of the loop by mutagenesis would affect this interaction.

(A)   **(B)**
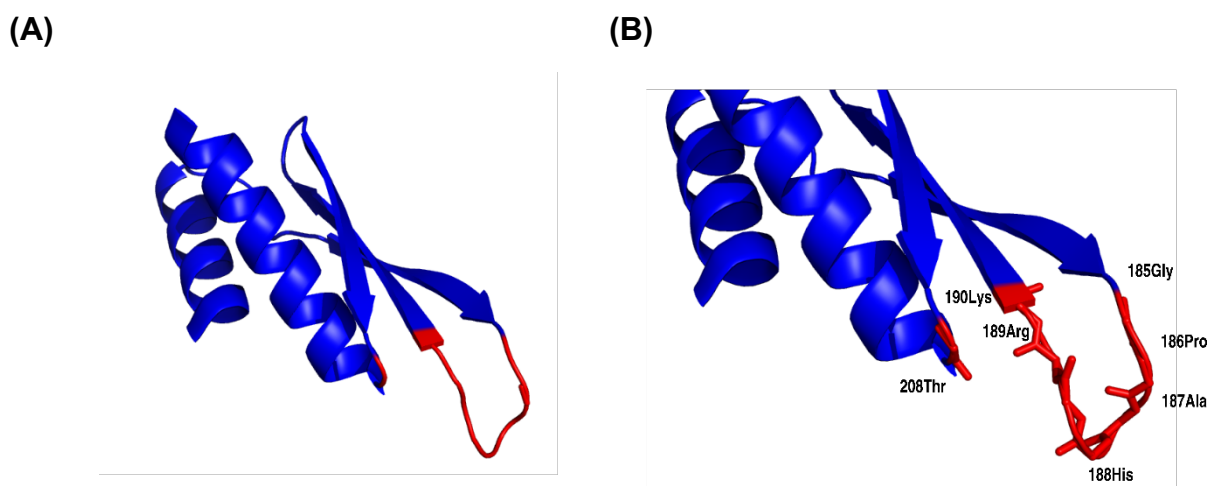
Figure 4.10 Structure of TRBP-D2 highlighting chemical shift changes due to mutagenesis.

A) Structure of TRBP- D2, residues affected largely by mutagenesis in the β1- β2 loop are coloured red B) Zoomed image orientated on the β1- β2 loop region highlighting residues affected largely by mutagenesis of TRBP-D2. Images created in Pymol using PDB accession 3ADL.

4.4 ^{15}N relaxation analysis of the type-A dsRBDs in TRBP

4.4.1 NMR ^{15}N relaxation analysis of TRBP-D12 shows variations in motional dynamics between the two dsRBDs

^{15}N NMR relaxation analysis has previously shown that the two type-A dsRBDs of TRBP contain similar dynamics between the structural features of the dsRBDs such as the alpha helices and beta strand but interestingly (^{1}H-^{15}N) Heteronuclear NOE data identified that the loop region between β1-β2 in the two dsRBDs displayed a difference in flexibility (Benoit et al. 2013). This loop region was shown to be more flexible in TRBP-D1 compared to TRBP-D2. To further study this phenomenon the analysis of this loop region was enhanced with additional ^{15}N NMR relaxation data, the longitudinal relaxation time T_1 and the transverse relaxation time T_2 calculated from the relaxation in the rotating frame $T_1\rho$.

^{15}N T_1 times were calculated for WT TRBP-D12 (Figure 4.11A). Regions with high flexibility have lower T_1 values, for example the values measured in the linker region connecting the two dsRBDs. Analysis of the T_1 relaxation times of the residues within the dsRBD shows that alpha helices and beta strands are rigid in structure and the regions connecting the alpha helices and beta strands also display similar rigidity, excluding the loop joining β1-β2, which expresses the highest flexibility within the dsRBDs.

Regions within proteins with high T_2 values exhibit higher flexibility, which is visible within the linker regions connecting the two dsRBDs (Figure 4.11B). Similar to the T_1 data the T_2 relaxation times indicate that the alpha helices and beta strands confer particular rigidity. Similar to the T_1 relaxation times for the amino acids in the β1-β2 loop, the T_2 relaxation times indicate a much more rigid loop in TRBP-D2 compared to TRBP-D1. Interestingly the β1-β2 loop in TRBP-D2 has lower T_2 values compared to the structured regions, which is not indicative of a flexible loop, this phenomenon will be discussed further later.

The (^{1}H-^{15}N) heteronuclear NOE values calculated for TRBP-D12 further prove the rigidity seen with the structured regions (high noe values) of the protein and the flexibility seen within the loop regions (low noe values) (Figure 4.11C). Following the trend seen for the T_1 and T_2 relaxation experiments, the β1-β2 loop region between the two dsRBDs again displays a difference in rigidity in the (^{1}H-^{15}N) Heteronuclear noe experiment.

The three relaxation parameters calculated for WT TRBP-D12 have highlighted some interesting variation between the two dsRBDs in the multi-domain protein. Most of the structured regions between the two dsRBDs show similar rigidity but the β1-β2 loops show different dynamic properties. The ^{15}N relaxation experiments indicate that the β1-β2 loop in TRBP-D2 is more rigid than the loop of TRBP-D1. This will be further explored by investigating the difference in amino acid composition of the β1-β2 loop between the two dsRBDs.
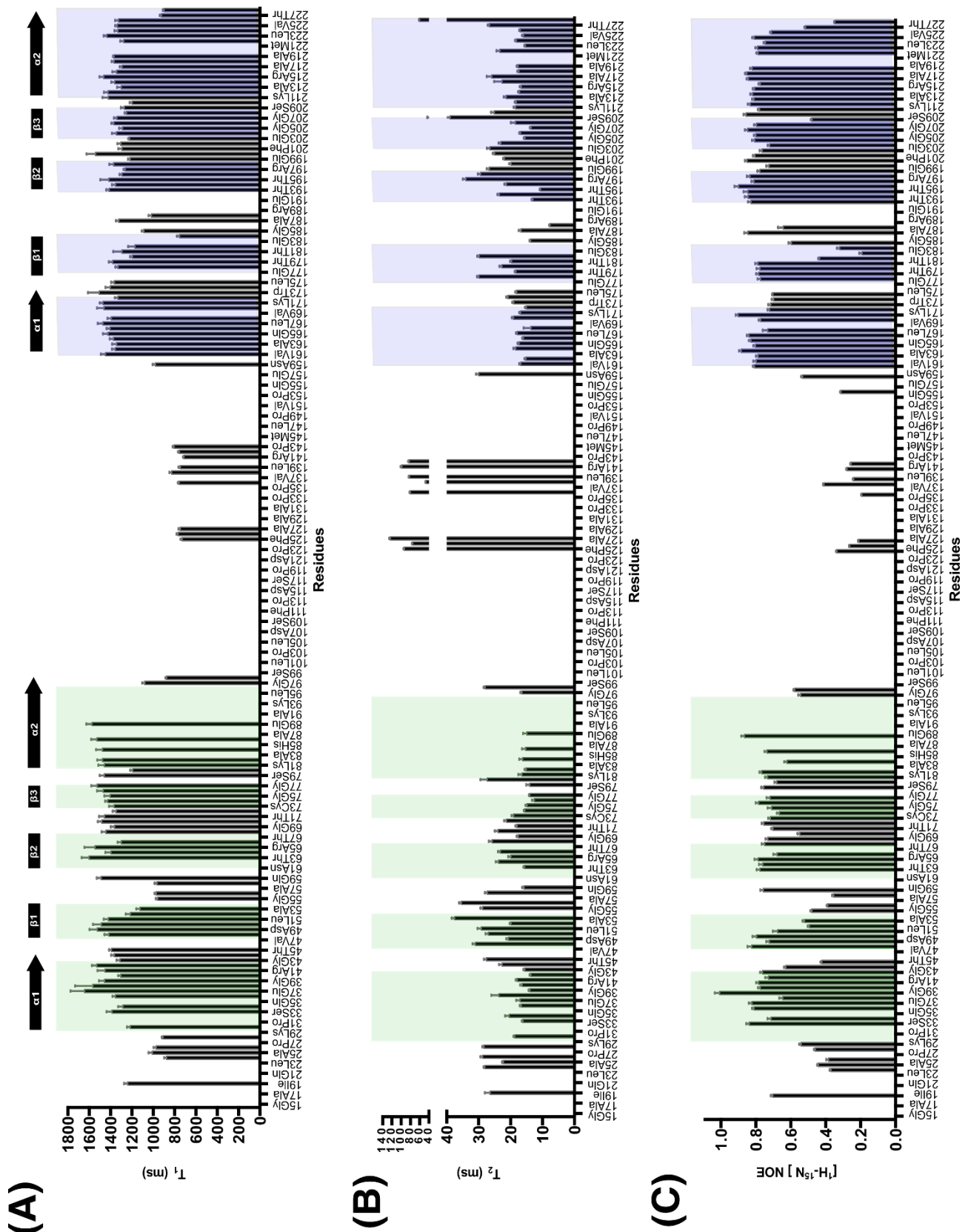
Figure 4.11 ^{15}N relaxation parameters of WT TRBP-D12

(A) T_1 relaxation values (B) T_2 relaxation values (C) (^1H-^{15}N) HET-NOE values for TRBP-D12 WT. Coloured regions correspond to the structured regions of the protein (Green-TRBP-D1, Blue TRBP-D2)

4.4.2 Characterisation of the dsRNA binding regions of the type-A dsRBDs by NMR spectroscopy

Like many type-A dsRBDs, the N-terminal dsRBDs of TRBP each contain three canonical dsRNA binding regions. Region 1 is located in the first α helix, with the second region in the loop adjoining β1 and β2 strands, and the third region on the N-terminal tip of the α2 helix (Figure 4.12).
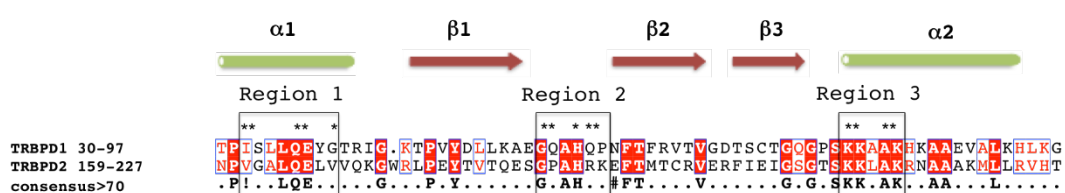
Figure 4.12 Amino acid composition the two N-terminal dsRBDs of TRBP.

Sequence of TRBP dsRBDs with the three RNA binding regions annotated. Essential RNA binding amino acid residues are labeled by an * over the residue.

Comparison of the ^{15}N relaxation times between TRBP-D1 and TRBP-D2 within TRBP-D12, revealed differences of the two domains within the β1-β2 loop. Interestingly this is one of the three dsRNA-binding regions of canonical dsRBDs (Figure 4.12). DsRNA binding region 2 in TRBP-dsRBD1 shows increased flexibility compared to TRBP-dsRBD2. This phenomenon is most pronounced in the T_2 relaxation parameter. The other two RNA binding regions have more similar relaxation rate profiles between the two dsRBDs (Table 4.2 & 4.3).

Table 4.2 Average ^{15}N relaxation values of the RNA binding regions in TRBP-D1

	TRBP-D1		
	R1	R2	R3
	33S, 36Q,37E	55G, 57A,58H	80K,81K,84K
T1 (ms)	1463.4 ±158.2	1098.8 ±261.4	1378.4 ±159
T2 (ms)	16.7 ±0.5	26.9 ±8.2	20.1 ±6.3
^{15}N{^{1}H} NOE	0.83 ±0.01	0.50 ±0.18	0.69 ±0.07

Table 4.3 Average ^{15}N relaxation values of the RNA binding regions in TRBP-D2

	TRBP-D2		
	R1	R2	R3
	161V,162G, 165Q 166E	185G, 187A, 188H, 189R	210K,211K, 213A,214K
T1 (ms)	1410.8 ±38.6	1114.4 ±159.8	1322.6 ±94.7
T2 (ms)	17.6 ±1.3	12.8 ±4.6	20.6 ±3.5
^{15}N{^{1}H} NOE	0.81 ±0.03	0.71 ±0.18	0.82 ±0.04

Sequence comparison between the two dsRBDs highlighted differences in RNA binding region 2 (Figure 4.12). For instance, two positively charged residues (R189,K190) found in TRBP-D2 and not present in TRBP-D1. Additionally, dsRNA-binding region 2 within TRBP-D2 contains a proline residue (P186) preceding the RNA interacting histidine (H188) whereas this is a glutamine residue (Q56) in RNA binding region 2 of TRBP-D1. To study the contribution of these residue variations may play in dynamics and motion of RNA binding region 2, single point mutations probing the proline residue difference were carried out. In TRBP-D1, Q56 was mutated to a proline residue and in TRBP-D2 P186 was mutated to a glutamine residue. The mutation of each dsRBD was applied to the multi-domain protein (TRBP-D12) to produce TRBP-D12$_{Mut}$. Relaxation experiments were carried out on both the WT TRBP-D12 and TRBP-D12$_{Mut}$. Backbone assignments of TRBP-D12$_{Mut}$ were obtained by transferring assignments from mutant variants of single dsRBDs and WT TRBP-D12 protein.
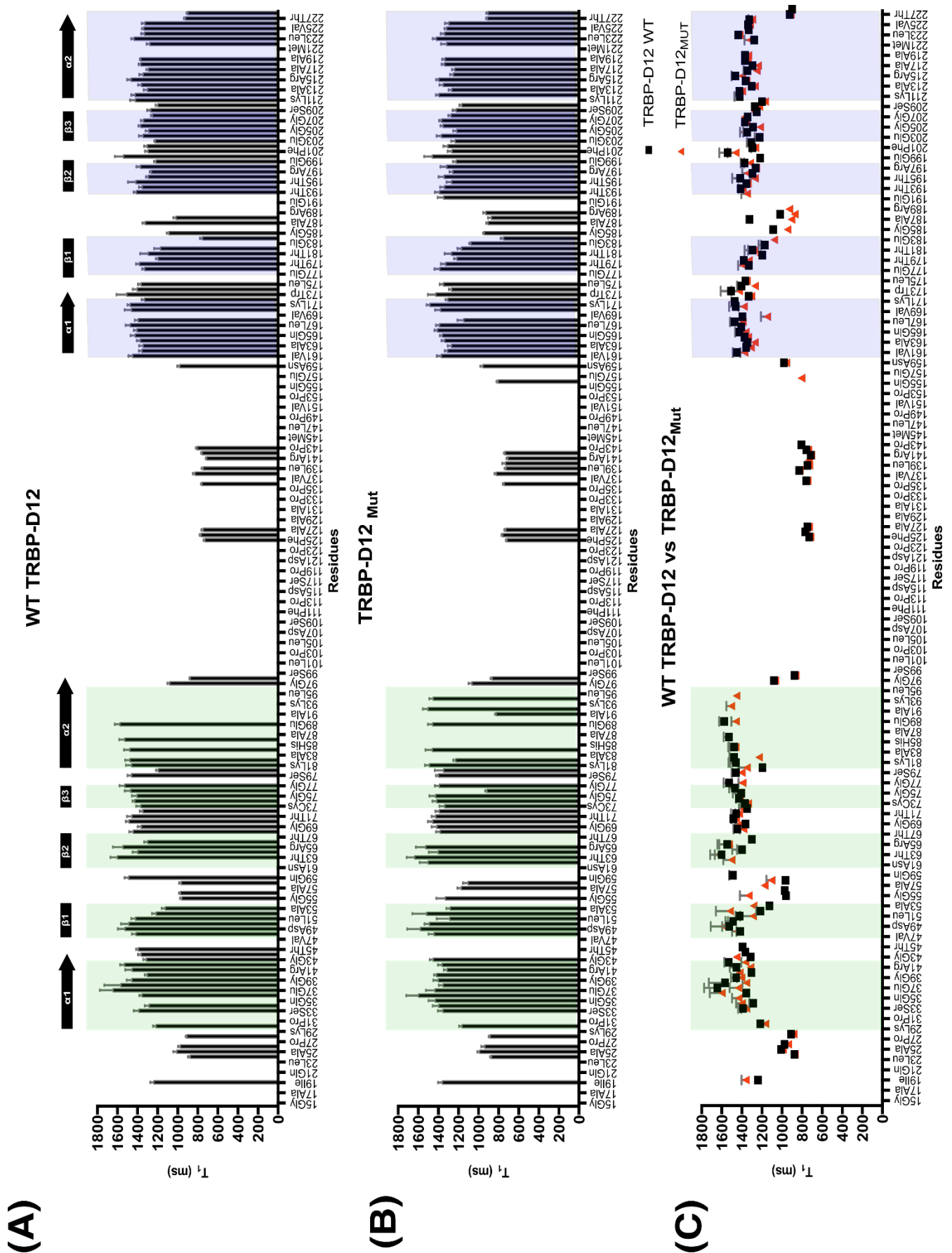
Figure 4.13 T₁ relaxation values of TRBP-D12 WT/Mut

T₁ relaxation values of A) TRBP-D12 (B) TRBP-D12$_{Mut}$ (C) Comparison of T₁ relaxation between WT TRBP-D12 (black squares) and TRBP-D12$_{Mut}$ (red triangles).
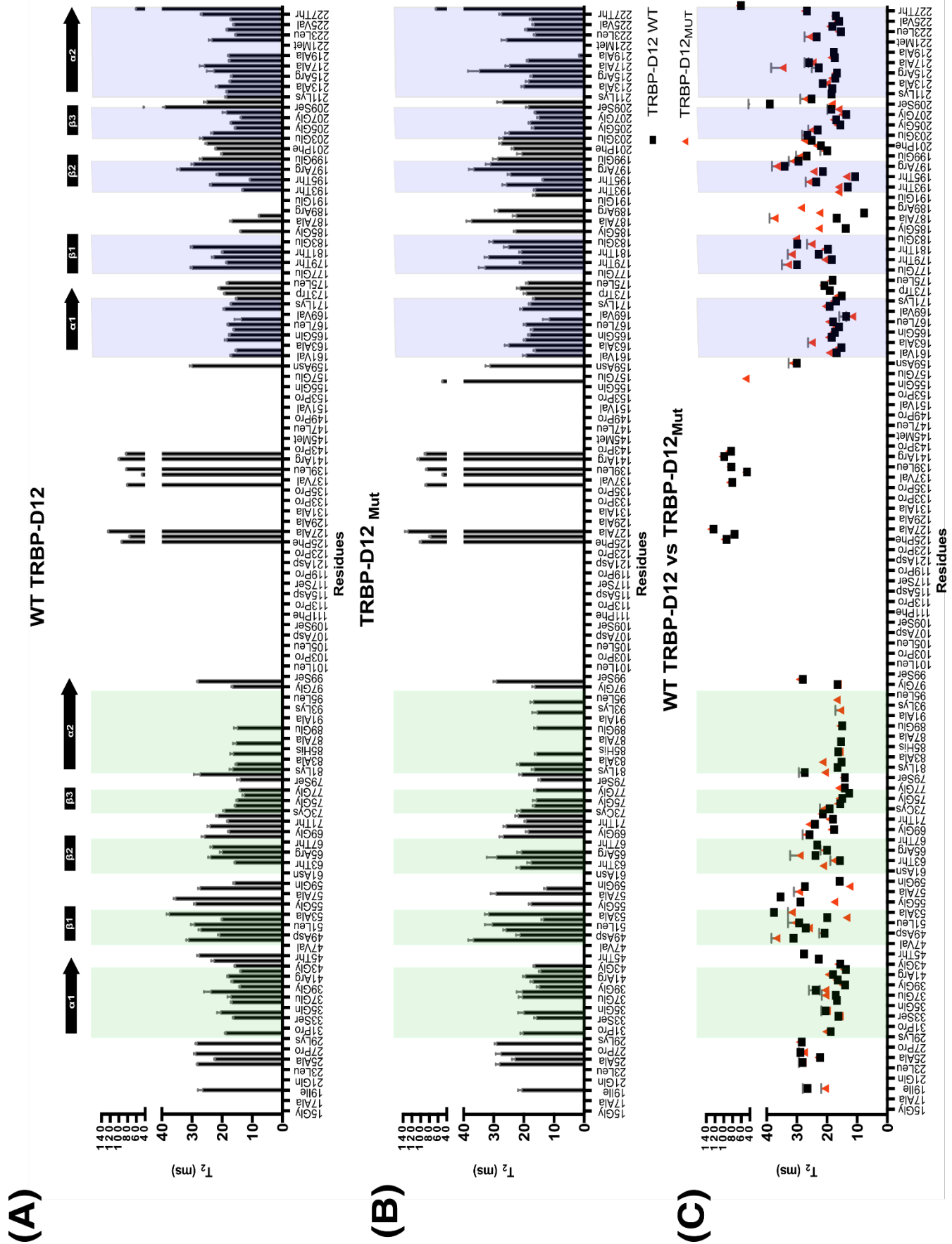
Figure 4.14 T_2 relaxation values of TRBP-D12 WT/Mut

T_2 relaxation values of A) TRBP-D12 (B) TRBP-D12$_{Mut}$ (C) Comparison of T_2 relaxation between WT TRBP-D12 (black squares) and TRBP-D12$_{Mut}$ (red triangles).
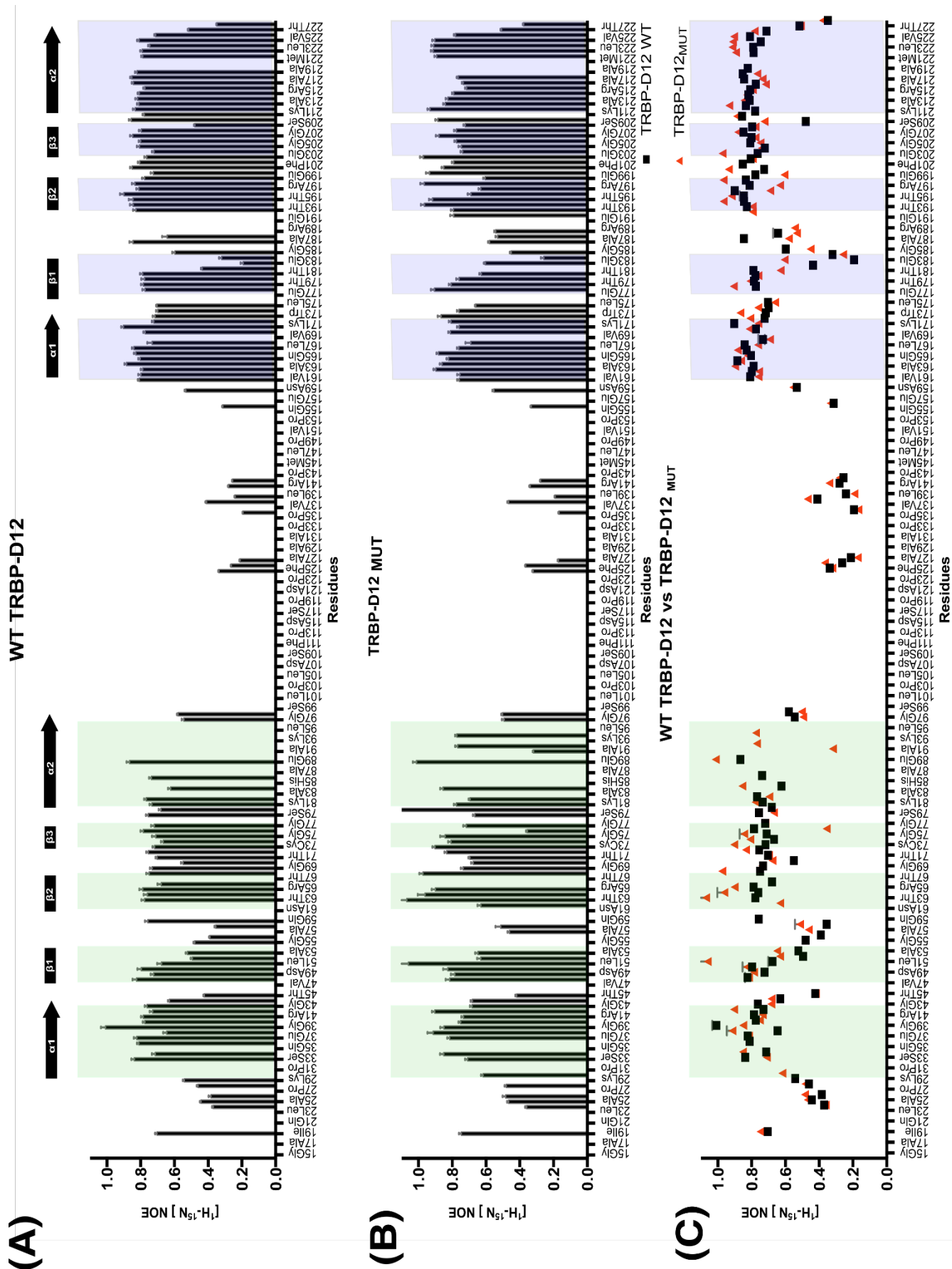
Figure 4.15 (^{1}H-^{15}N) Heteronuclear-NOE relaxation values of TRBP-D12 WT/Mut

(^{1}H-^{15}N) HET-NOE relaxation values of A) WT TRBP-D12 (B) TRBP-D12$_{Mut}$ C) Comparison of ^{15}N{^{1}H} NOE relaxation between WT TRBP-D12 (black squares) and TRBP-D12$_{Mut}$ (red triangles).

Comparison of the relaxation data between the WT TRBP-D12 and TRBP-D12$_{Mut}$ show that the mutations of the loop region have not altered the dynamics of the structured regions of the dsRBDs, but primarily affected the residues within the β1-β2 loop (Figure 4.13-15). In TRBP-D1 the substitution of the glutamine residue (Q56) for a proline residue has increased the rigidity of the β1-β2 loop region, which can be seen across the T_1 and T_2 relaxation experiments. The replacement of the proline residue in the β1-β2 loop of TRBP-D2 for a glutamine residue has been shown to affect the dynamics of the loop by decreasing the rigidity of this region.

Table 4.4 ^{15}N relaxation values of the β1-β2 loop in TRBP-D1 of WT TRBP-D12 and TRBP-D12$_{Mut}$

	WT TRBP-D12	TRBP-D12MUT
	55G, 56Q, 58H, 59E	55G,57A,58H
T_1	1098.8 ±261.4	1204.6 ±117.3
T_2	26.9 ±8.2	19.8 ± 8.6
^{15}N{^{1}H} NOE	0.50 ± 0.18	0.48 ± 0.036

Table 4.5 ^{15}N relaxation times of the β1-β2 loop in TRBP-D2 of WT TRBP-D12 and TRBP-D12$_{Mut}$

	WT TRBP-D12	TRBP-D12MUT
	185G,187A,188H, 189R	185G, 187A,188H,189R
T_1	1114.4 ±159.8	914.6 ± 30.5
T_2	12.8 ±4.6	27.8 ± 7.0
^{15}N{^{1}H} NOE	0.71 ±0.18	0.53 ±0.05

^{15}N relaxation experiments of the dsRBDs in TRBP-D12 exposed a disparity in structural dynamics between the two dsRBDs, in the β1-β2 loop, where RNA binding region 2 lies. With mutations in the β1-β2 loop of each dsRBD to mimic the amino acid composition of the other domain, the relaxation values were

essentially exchanged between the two dsRBDs. This suggests the difference in a single amino acid between the two loop regions could be the reason for the variation in motional dynamics of this loop.

4.4.3 ^{15}N relaxation analysis of TRBP-D2 WT & P186Q

To further investigate the structural dynamics of the dsRBDs of TRBP and to confirm the results observed in the double domain construct, relaxation analysis was carried out on TRBP-D2. Analysis on the single dsRBD should be able to solve some of the problems encountered with the double domain construct such as the overlapping peaks and extraction of relaxation data on the residues in the β1-β2 loop. In addition, it would be interesting to verify the T_2 relaxation rate phenomenon seen in RNA binding region 2 for TRBP-D2 in the double domain construct within the single dsRBD as well. Focusing on RNA binding region 2, relaxation experiments were also conducted on the TRBP-D2 P186Q mutant.

The NMR experiments were carried out with the same parameters and conditions used for TRBP-D12 WT and three relaxation parameters were calculated, T_1, T_2, (^{1}H-^{15}N) HET-NOE. The advantage of assessing the single dsRBD dynamics is that relaxation data were extracted from more residues in the β1-β2 loop than in the double- domain construct.
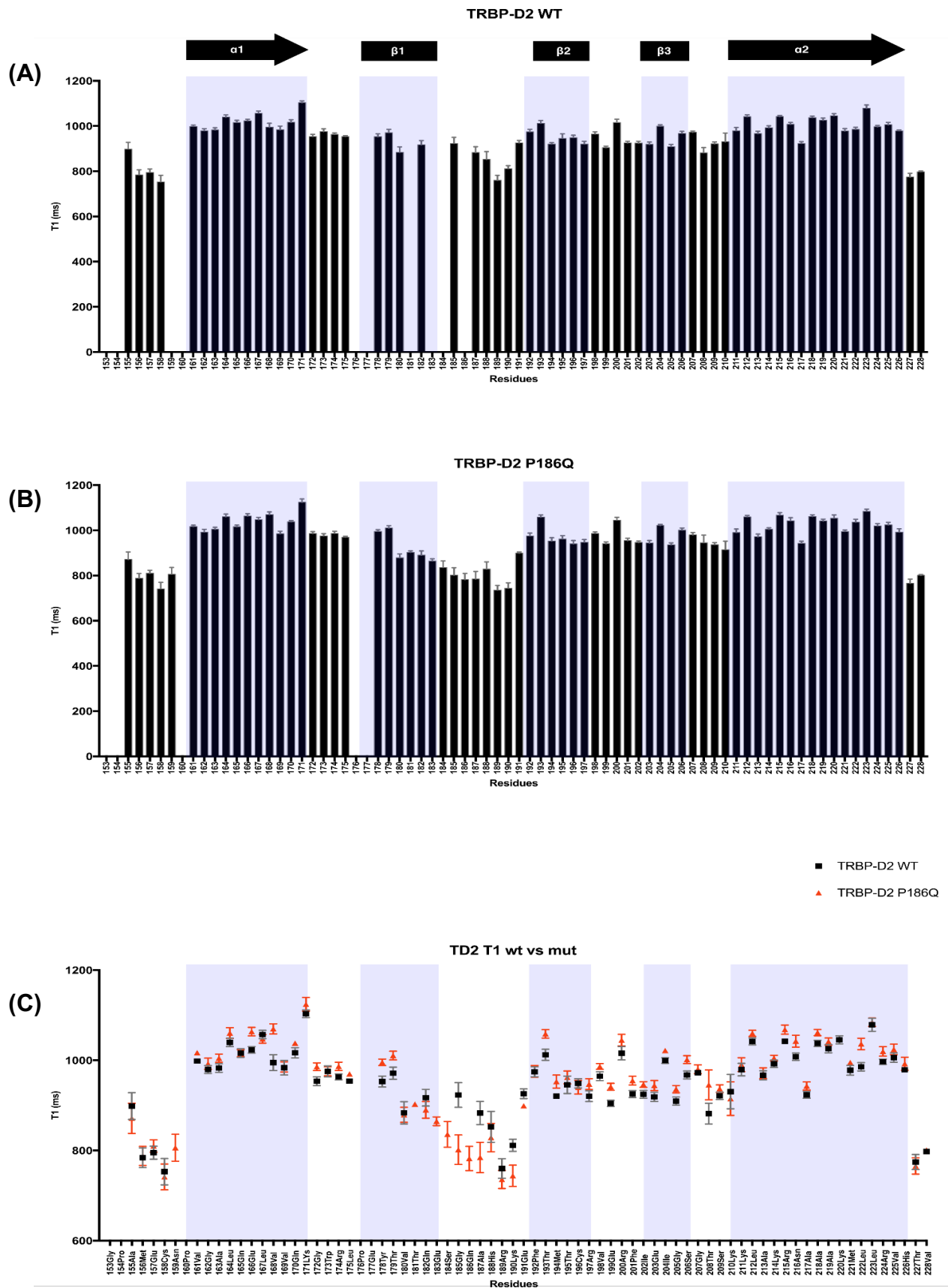
Figure 4.16 T₁ relaxation values of TRBP-D2 WT/P186Q

(A) T₁ relaxation values of TRBP-D2 (B) T₁ relaxation values of TRBP-D2 P186Q (C) Comparison of T₁ relaxation between TRBP-D2 WT (black squares) and P186Q (red triangles).
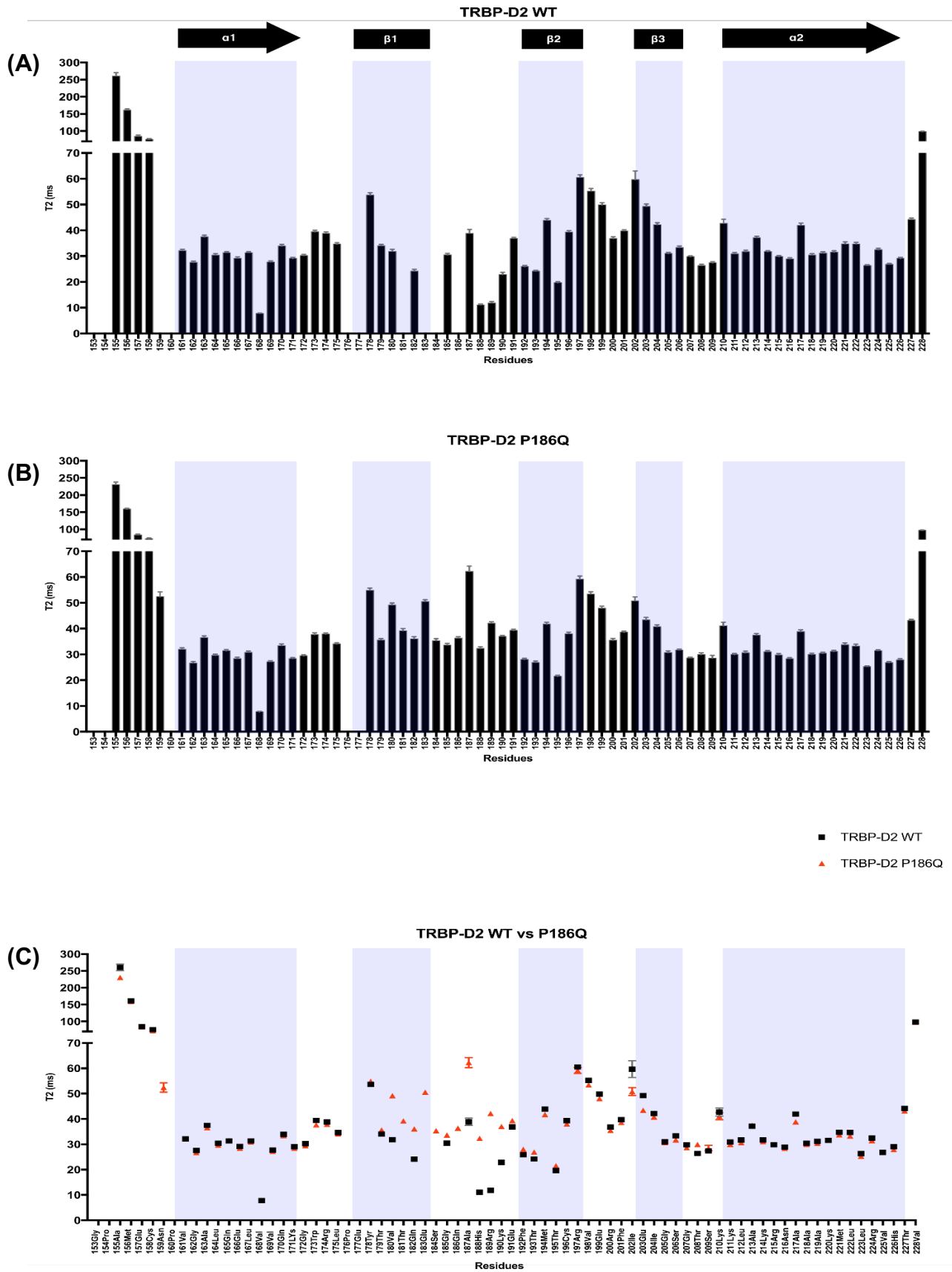
Figure 4.17 T₂ relaxation values of TRBP-D2 WT/P186Q

(A) T₂ relaxation values of TRBP-D2 (B) T₂ relaxation values of TRBP-D2 P186Q Comparison of T₁ relaxation between TRBP-D2 WT (black squares) and P186Q (red triangles).
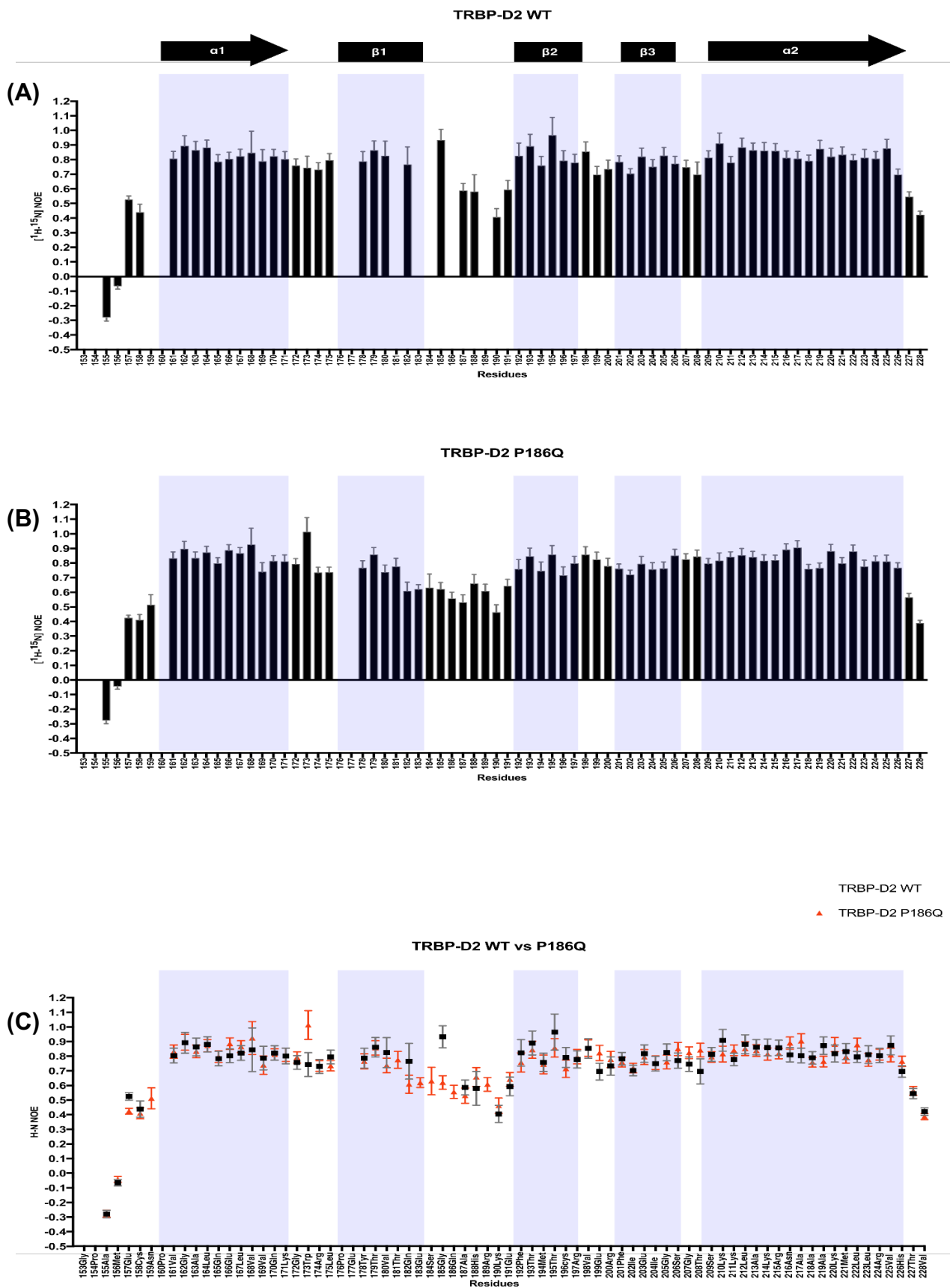
Figure 4.18 (^{1}H-^{15}N) HET-NOE relaxation values of TRBP-D2 WT/P186Q

(A) (^{1}H-^{15}N) HET-NOE relaxation values of TRBP-D2 (B) (^{1}H-^{15}N) HET-NOE relaxation values of TRBP-D2 P186Q Comparison of T_1 relaxation between TRBP-D2 WT (black squares) and P186Q (red triangles).

The ^{15}N relaxation experiments displayed rigidity in the structured regions with flexibility in the loop regions of the protein and the N-terminal residues. The T_1 and T_2 relaxation experiments show that the dynamics of the β1-β2 loop is altered with the removal of the proline residue from TRBP-D2. The T_1 and T_2 times indicate that the β1-β2 loop region of TRBP-D2 P186Q is more flexible than WT TRBP-D2. However, this difference in dynamics is not clearly displayed in the (^{1}H-^{15}N) Heteronuclear NOE studies of the two protein constructs.

Table 4.6 ^{15}N relaxation times of the β1-β2 loop in WT TRBP-D2 and TRBP-D2 P186Q

	WT TRBP-D2	TRBP-D2 P186Q
	185G,187A,188H,190K	185G,186Q,187A,188H,189R,190K
T_1 (ms)	846.1 ± 63.1	779.5 ± 35.1
T_2 (ms)	22.9 ± 11.9	40.6 ± 11.11
^{15}N{^{1}H} NOE	0.63 ± 0.22	0.57 ± 0.07

T_2 relaxation rates of TRBP-D2 WT and P186Q highlight there is rigidity in the structured regions of the dsRBDs and high flexibility seen at the N-terminal end of the dsRBD. Interestingly the T_2 relaxation rate phenomenon seen in RNA binding region 2 of TRBP-D12 can also be seen in the single dsRBD. The T_2 relaxation rates for the residues in the β1-β2 loop are significantly lower than for residues found in the structured regions not indicative of a flexible loop. With the substitution of the proline residue (P186) the T_2 relaxation rates of the loop region increase indicating an increase in flexibility. Interestingly Val-168 also has a low T_2 relaxation time compared to the rest of the domain.

A possible explanation for seeing this phenomenon of a low T_2 relaxation time in a flexible loop could be due to a conformational exchange process occurring with the residues in the loop structure, if the spins of the residues are switching between two or more states with different chemical shifts, this affects the T_2 relaxation rate by signifying a shorter T_2 value. Conformational exchange motion has previously been identified for loop regions of dsRBDs (Lu and Hall,

1997; Nanduri et al., 2000). Interestingly the removal of the proline residue increased the T_2 values of RNA binding region 2 specifically the histidine residue back to an expected value. This could suggest that the proline residue maybe important in the potential conformational exchange of this loop region, but further NMR analysis must be carried out to prove this.

4.4.4 Discussion

Investigating dynamics and local motions within a protein could help to further understand how a protein functions and carries out its specific role. [15]N relaxation experiments have been carried out to investigate the dynamics of the two type-A RNA binding domains of TRBP to inspect the similarities and differences in local motion between the two dsRBDs. Interestingly, differences in dynamics were identified between RNA binding region 2 between both domains. Therefore, it was important to first investigate how this difference in dynamics exists between two highly conserved dsRBDs and then to understand how this difference impacts the singular dsRBDs and within the multi-domain protein. This led to investigating the amino acid composition of this region and carrying out point mutations to find the cause for the difference in dynamics. The results indicate that the proline residue (P186Q) found in TRBP-D2 which is a glutamine residue in TRBP-D1 may be the cause of the difference in dynamics between the two domains in RNA binding region 2.

The investigations suggest that there might be some interesting phenomenon of RNA binding region 2 in TRBP-D2. The T_2 relaxation experiment of WT-TRBP-D2 showed residues in the β1-β2 loop region has noticeably lower T_2 values compared to the average T_2 value of the dsRBD not indicative of a flexible loop region. Which has been seen in other dsRBDs where residues required an exchange term for T_2 relaxation to be fitted. This requires the model-free approach proposed by Lipari and Szabo which calculates the exchange contribution to T_2 (Rex)((Lipari and Szabo, 1982). Therefore, some form of conformational exchange could also be occurring with the residues within this loop region, additionally point mutations with the removal of P186 suggests the ability to eliminate this phenomenon. It was also shown that an

insertion of a proline residue in the β1-β2 loop of TRBP-D1 had the ability to induce this T_2 phenomenon. Although the T_2 data for the $β1-β2$ loop in TRBP-D2 may be affected by residues experiencing exchange, the other two relaxation parameters highlight that this region is more rigid than the β1- β2 loop found in TRBP-D1.

The flexibility of the β1- β2 loop in dsRBDs has been hypothesized to be critical for binding, as the loop is required to rotate toward the dsRNA to interact with the 2'hydroxyl groups of the minor groove (Ryter and Schultz 1998; Doyle and Jantsch 2002; Masliah et al. 2013). A difference in dynamics between dsRBDs within a multi-domain dsRBP has been previously reported for the interferon-inducible antiviral kinase PKR. Using NMR spectroscopy, it was shown that the two structurally similar N-terminal dsRBDs exhibited a difference in motional flexibility on a millisecond timescale. It was suggested that the difference in backbone dynamics correlates with their differential dsRNA binding activity (Yang et al. 2014). Therefore, could the difference in backbone dynamics seen for TRBP-D1 and TRBP-D2 in β1- β2 loop region play a role in RNA binding of these dsRBDs? The dsRBDs of PKR have been shown to bind to structurally variable dsRNAs and PKR-dsRBD1 binds more favourably than PKR-dsRBD2. It has been hypothesised that the more dynamic PKR-dsRBD1 may confer the flexibility required for structurally distorted dsRNA. Therefore, could the difference in dynamics between TRBP-D1 and TRBP-D2 play a similar role in the binding of pre-miRNAs.

This chapter has identified a difference in dynamics between two dsRBDs within the same multi-domain protein. The data suggests that the difference in dynamics is due to amino acid composition. The next logical step would be to try and understand why this difference in dynamics exists between TRBP-D1 and TRBP-D2 and how changing the dynamics of the dsRBD can change its functions such as dsRNA binding. In the case of a dsRBD interacting with dsRNA, most of the conserved RNA binding amino acids found in dsRBDs use their backbones for interaction. Therefore, backbone flexibility can play a critical role in dsRNA binding activity of a dsRBD (Ryter and Schultz 1998; Ramos et al. 2000; Wang et al. 2011).

134

Chapter 5- Investigating the dsRNA binding properties of TRBP

5.1 Introduction

This chapter will focus on the RNA binding properties of the type-A dsRBDs of TRBP and more specifically the characteristics of the second RNA binding region of both dsRBDs. The amino acid sequence encompasses the structural dynamics and chemical properties of protein structures therefore each specific residue in region 2 could play a role in dsRNA recognition. Subsequently this will be explored through mutagenesis and investigating the RNA binding affinities of WT and mutant dsRBD constructs.

5.2 The RNA binding relationship of the dsRBDs of TRBP

The two type-A dsRBDs of TRBP that bind to dsRNA are connected by an unstructured long flexible linker, previous studies have shown that two dsRBDs do not interact with each other within the multi-domain protein (Benoit et al. 2013). The presence of a linker between dsRBDs can play a role in the interaction of the dsRBDs with dsRNA, some proteins require the linker region to form an integral part of the recognition surface or tether the dsRBDs to restrict local diffusion by defining a maximum distance between dsRBDs in a length dependent manner (Mackereth and Sattler 2012). The tethering of multiple dsRBDs can yield to an increase in local concentration of one dsRBD upon binding of another dsRBD within a multi-domain protein (Shamoo et al. 1995; Tian et al. 2004). In case of an infinitely long linker the affinity of two dsRBDs within a multi-domain protein is additive and in the case of a short linker it is multiplicative (Shamoo et al. 1995). Linker length dictating and restricting the local concentration of dsRBDs in a multi-domain protein can contribute to the interaction of a weak interacting dsRBDs. A high-affinity dsRBD can first bind to a substrate and restrict the local concentration of the weaker binding domain dependent on the linker length and enhance substrate binding.

Investigations were carried out to analyse the RNA binding affinities of the singular domains (TRBP-D1 & TRBP-D2) and the double domain protein (TRBP-D12) with various dsRNAs. The method of choice for analysing binding affinities was fluorescence anisotropy; each dsRNA investigated was labelled with a fluorophore on the 3' end. Multiple single dsRBDs and multiple TRBP-D12 are capable of binding a pre-miRNA. However, fluorescence anisotropy binding data were fit to a single site-binding isotherm, with the assumption that each binding even is independent.

The linker between the dsRBDs in TRBP is long and flexible separating the dsRBDs therefore the expected binding affinity of TRBP-D12 is to be the sum of the dsRBD affinities. However, analysing the RNA binding affinity of TRBP-D12 we can see that for the dsRNAs tested TRBP-D12 binds to dsRNA stronger than the individual TRBP dsRBDs (Figure 5.1). A 2-4-fold increase in RNA binding affinity for the double domain construct was seen compared to the single dsRBD (Table 5.1). This indicates that although the linker is long and flexible it plays a role in RNA binding by increasing the affinity of the single dsRBDs by tethering them to each other, possibly by increasing the RNA binding surface by positioning of the dsRBDs or by increasing the local concentration of one dsRBD when one is already bound.

Table 5.1 RNA binding affinities of TRBP-D1, TRBP-D2 and TRBP-D12 for various dsRNAs

	Pre-miR -155				Pre-Let-7a			
	Native		modified		Native		modified	
	K_D	error	K_D	error	K_D	error	K_D	error
TRBP-D1	946	190	519	60	777	88	1015	142
TRBP-D2	914	103	571	59	630	57	824	78
TRBP-D12	381	58	247	36	292	45	285	38
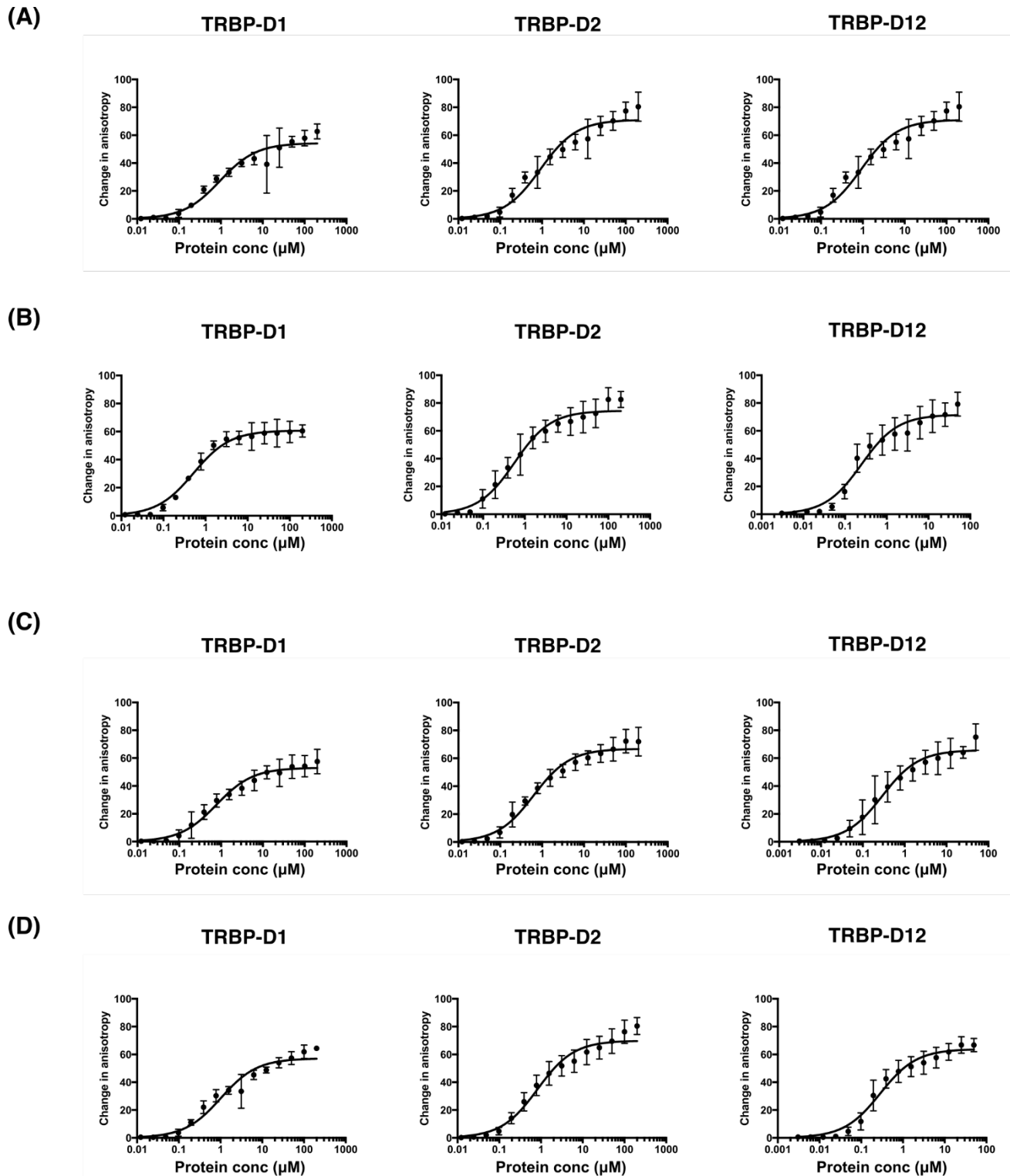
Figure 5.1 Binding curves for TRBP-D1, TRBP-D2 and TRBP-D12 for various dsRNAs

(A) Interaction of Pre-miR 155 WT vs. protein constructs (B Interaction of Pre-miR 155 mod vs. protein constructs (C) Interaction of Pre-let7a vs. protein constructs (D) Interaction of Pre-let7a mod vs. protein constructs. Experimental data are averaged from three independent fluorescent anisotropy experiments.

5.3 The role of dsRNA structure on the dsRNA binding affinity of the dsRBDs found in TRBP

The binding of most dsRBDs to precursor microRNAs has been identified to be sequence independent (Masliah et al. 2013). Therefore, it has been suggested that recognition of dsRNAs by dsRBDs depends on the structural features of the dsRNA such as bulges and internal loops (Bevilacqua et al. 1998; Krol et al. 2004; Heinicke et al. 2011; Liu et al. 2018). To investigate how structural features of pre-miRNAs may be recognised by TRBP, modified versions of two pre-miRNAs were produced (Pre-miRNA-155 & Pre-let-7a). Structural features have been removed to produce a perfect dsRNA duplex of the pre-miRNA (Figure 5.2). Pre-miR-155 contains two 1-nt bulges in the middle of dsRNA stem and two G.U wobble base pairs in the dsRNA stem. Pre-let7a contains only 1-nt bulge close to the 5' & 3' end of the duplex and two G/U wobble base pairs.

(A)

(B)

(C)

(D)
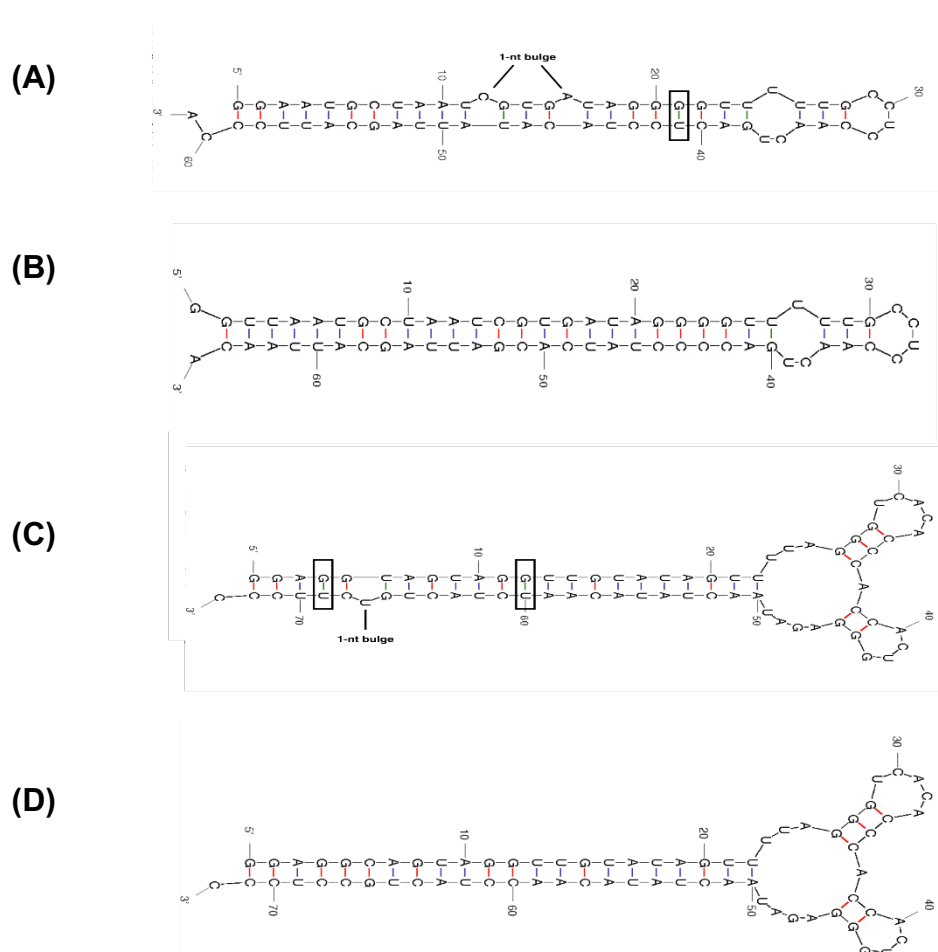
Figure 5.2 Predicted secondary structure of dsRNAs used in RNA binding studies

(A) Pre-miR 155 WT (B) Pre-miR 155 mod (C) Pre-let7a WT (D) Pre-let7a mod. Secondary structures of the dsRNAs were predicted using Mfold.

138

The RNA binding analysis of TRBP-D12 with the wild type and modified versions of pre-miRNAs showed that for TRBP-D12 a large difference in RNA binding affinities cannot be seen (Table 5.1). However, the single domains of TRBP can been seen to interact with a higher affinity (2-fold) for pre-miR-155 mod compared to pre-miR-155 wt. Interestingly with the removal of structural imperfections in pre-let-7a the RNA binding affinity by the single dsRBDs increased marginally. Other structured dsRNAs need to be investigated to further understand the role of structural imperfections in the dsRNA stem of pre-miRNAs. However, the data presented here suggests that the RNA binding affinities of the single dsRBDs of TRBP can indeed be affected by 1-nt bulges and non- Watson-Crick base pairing. The RNA binding data also implies that structural features of dsRNAs do not only impede RNA binding but can also enhance the affinity as seen by pre-let-7a and its modified version. Hence, the position of structural features on the dsRNA could possible play a role in dsRNA recognition.

It was expected for the dsRBDs of TRBP to bind to both modified version of the pre-miRNAs with similar binding affinities, but this was not the case as pre-let-7a mod has ~1.5-2-fold weaker binding affinity for the single dsRBDs compared to pre-miR-155 mod. A potential explanation for this could be due to the length of the dsRNA stem, the Mfold structures reveal that Pre-miR-155 mod has a 25 bp stem whereas Pre-let-7a mod has a 21 bp stem (Figure 5.2). This is in agreement with data indicating that the binding affinity of TRBP increases with RNA duplex length (Acevedo et al. 2015).

5.4 Investigating the amino acid composition of RNA binding region 2 in TRBP-D1 & TRBP-D2

As discussed in Chapter 1 RNA binding region 2 of dsRBDs is highly conserved and important for the interaction between dsRBDs and dsRNA. Therefore, this part of the chapter will focus on understanding the importance of the residues in this region in the dsRBDs of TRBP. Interestingly the two type-A dsRBDs of TRBP have a different amino acid composition in this

specific region. Analysing the difference in amino acid composition of RNA binding region 2 of TRBP-D2 compared to TRBP-D1 will expose how each residue functions in RNA binding of this region and help to further elucidate how this region interacts with dsRNA. NMR analysis of RNA binding region 2 of the dsRBDs in TRBP in the previous chapter has highlighted differences in dynamics between the two dsRBDs. It was shown that a single point mutation Q56P and P186Q in TRBP-D1 and TRBP-D2 respectively can alter the flexibility. This will be further investigated to understand how dynamics of this region plays a role in dsRNA binding of each dsRBD.
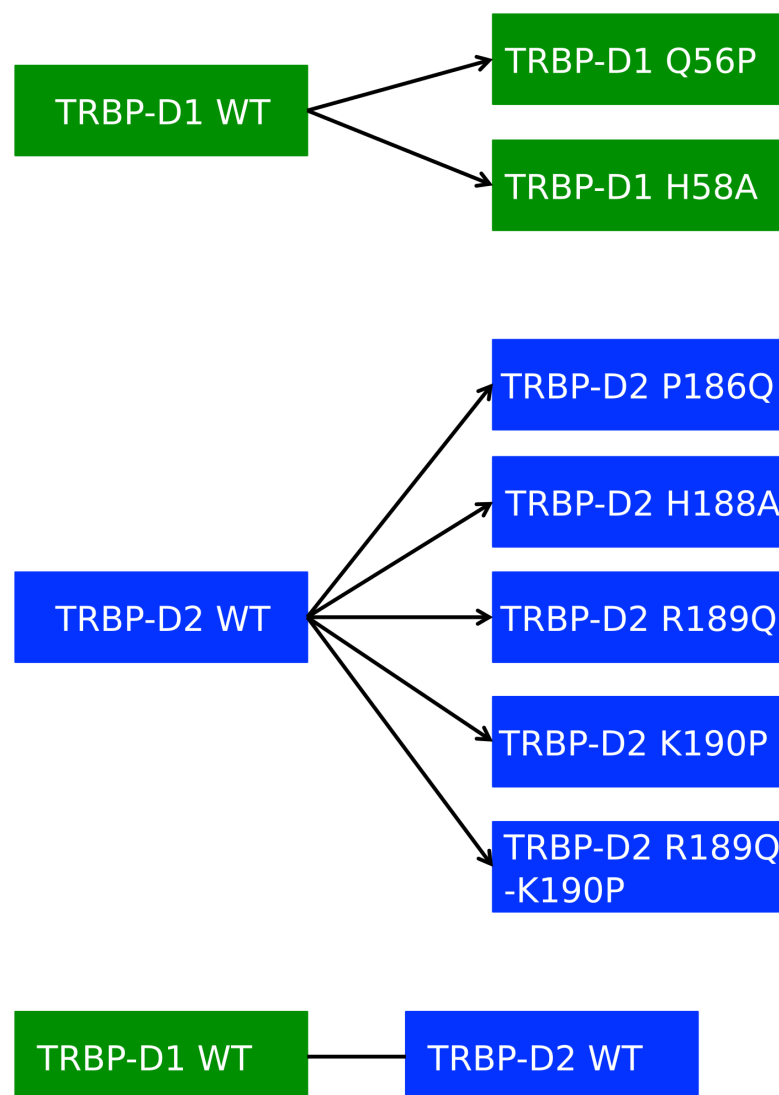
Figure 5.3 Schematic of TRBP constructs created by mutagenesis for RNA binding studies.

5.4.1 Investigating RNA binding region 2 in TRBP-D2 by EMSA

Individual residues in the β1- β2 loop of TRBP-D2 were mutated to mimic the amino acid composition of TRBP-D1 and tested to determine whether mutations alter the RNA binding (Figure 5.3). The sequence alignment highlighted differences in residues between the regions of the TRBP domains. The first difference that was investigated is the presence of a Proline residue preceding the RNA interacting histidine (G**P**AHRK) in TRBP-D2, TRBP-D1 has a glutamine at the equivalent position. The second difference that was investigated in this region is the presence of two positively charged residues (GPAH**RK**) in TRBP-D2 that are either glutamine or a proline residue in TRBP-D1. The role these differences in amino acid composition play in RNA binding was initially investigated through EMSA.

Gels were analysed, and the estimated dsRNA binding affinities were examined by determining the disappearance of the free RNA band at various protein concentrations for each protein construct. EMSA results show that a single amino acid substitution in RNA binding region 2 can reduce the affinity of TRBP-D2 for pre-miRNA (Figure 5.4). This can be seen by the disappearance of the free RNA band at a higher protein concentration compared to WT TRBP-D2. Mutation of the positively charged residues found in TRBP-D2 had little to no effect on the RNA binding affinity suggesting that these residues do not interact with the negatively charged backbone of the dsRNA. However, the P186Q mutation reduced binding affinity of TRBP-D2 for both pre-miRNAs tested.

Although the mutation K190P has introduced another proline residue into the β1- β2 loop, this mutation had minimal effect on the binding affinity, which suggests the positioning of the proline in TRBP-D2 may be important rather than the number of proline residues present.

The results obtained imply that differences in the residue composition of the β1- β2 loop of these dsRBDs may be required for fine-tuning the affinity of the two RNA binding dsRBDs in TRBP. These findings suggest a hypothesis that the proline residue at the position preceding the histidine may be play a role in

interacting with the dsRNA or facilitating other residues such as the histidine for RNA binding.

(A)
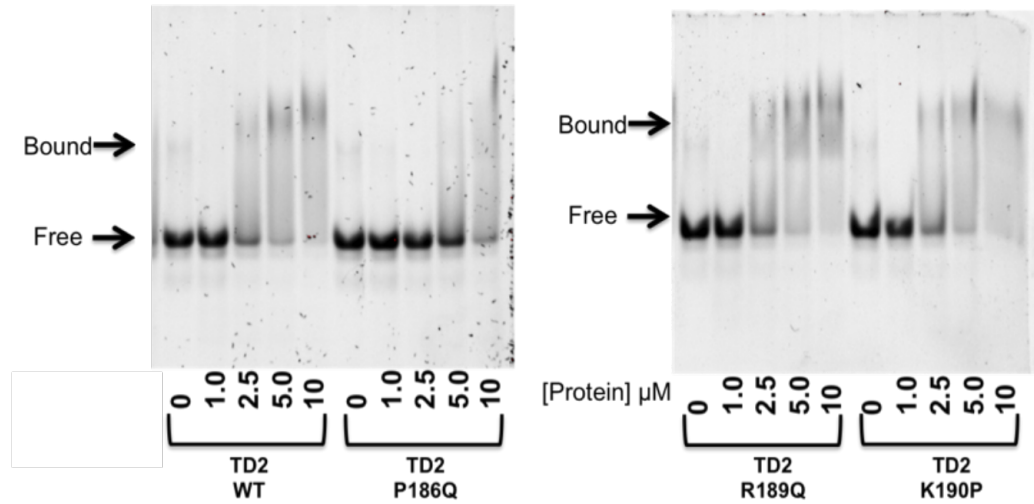
(B)
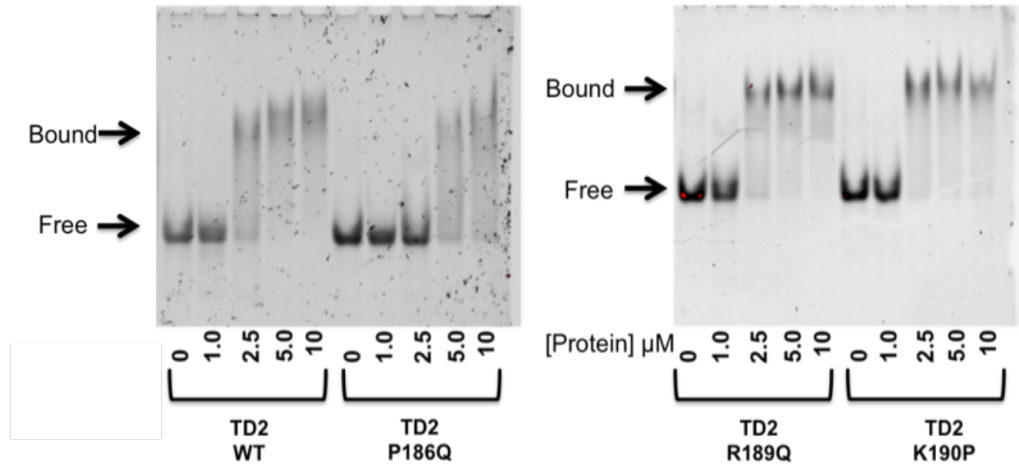
Figure 5.4 Electrophoretic mobility shift assay of two pre-miRNAs with various TRBP wt and mutant constructs for binding analysis.

(A) 200 nM of Pre-miRNA 155 incubated with an increasing concentration of TRBP-D1, TRBP-D2, TRBP-D2 P186Q, TRBP-D2 R189Q and TRBP-D2 K190P protein constructs. (B) 200 nM of Pre-let 7a incubated with an increasing concentration of TRBP-D1, TRBP-D2, TRBP-D2P186Q, TRBP-D2 R189Q and TRBP-D2K190P protein constructs. EMSAs were carried out on 10% native polyacrylamide gels in triplicates

5.4.2 Quantitative analysis of the dsRNA binding affinity of TRBP-D2 and mutations affecting RNA binding region 2

To further analyse RNA binding properties region 2 of TRBP-D2, a more quantitative analyse of determining the RNA binding affinity was carried out. Fluorescence anisotropy was used, and protein binding affinities were measured with 3' FTSC labelled dsRNAs. To further our understanding of RNA binding region 2, the histidine residue was mutated to an alanine to analyse its importance in RNA binding (Figure 5.3). Similar experiments have been conducted on dsRBDs for other proteins, which revealed that mutation resulted in a lower affinity or no binding to dsRNA (Bycroft et al. 1995; Krovat and Jantsch 1996). The EMSA analyses of the two positively charged amino acids in RNA binding region 2 showed that single point mutations at these positions did not alter the RNA binding affinity. Therefore, a double mutant construct was created to understand how this may affect RNA binding of this region (TRBP-D2 RK189QP).

RNA binding curves obtained through fluorescence anisotropy were fitted with a single site-specific binding equation (Figure 5.5). Although it has been shown that multiple dsRBDs bind to pre-miRNAs such as pre-miR-155, a single site binding equation was used to calculate a global affinity of the multiple binding sites, assuming all sites are independent.

(A)

	Pre-miR -155				Pre-Let-7a			
	Native		modified		Native		modified	
	K_D (nM)	error	K_D (nM)	error	K_D (nM)	error	K_D (nM)	error
TRBP-D2	914	103	571	59	630	57	824	78
TRBP-D2 P186Q	1272	250	1265	202	1276	162	1047	149
TRBP-D2 H188A	1798	316	1285	210	1527	261	2277	395
TRBP-D2-R189Q-K190P	1643	280	1410	271	1954	371	2433	388

(B)
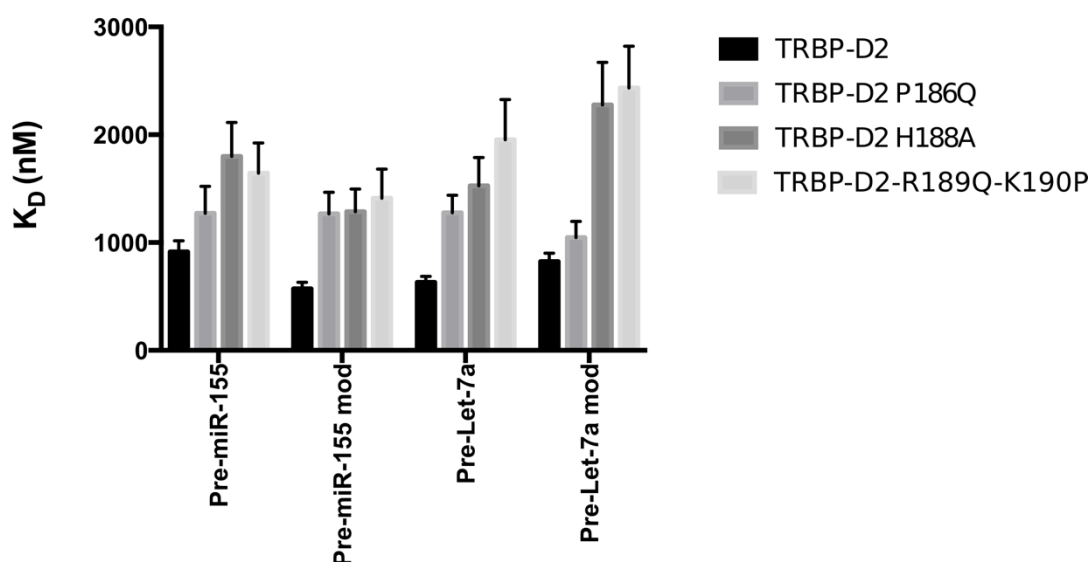
Figure 5.5 dsRNA binding affinities of WT TRBP-D2 and mutant constructs

(A)Table of binding affinities of WT and mutant constructs of TRBP-D2 for various pre-miRNAs produced (B) Bar chart format of binding affinities. Binding affinities were calculated from three independent fluorescence anisotropy experiments.

The fluoresce anisotropy analysis of RNA binding region 2 of TRBP-D2 has revealed that mutations to specific residues within this region has an effect on the RNA binding affinity of the dsRBD (Figure 5.5).

For all pre-miRNAs tested, the mutation of the histidine residue for an alanine residue did not knock out RNA binding but merely decreased the RNA binding affinity of TRBP-D2. This result was unexpected as others have shown that the removal of the conserved histidine residue knocks out binding of dsRBDs (Krovat and Jantsch 1996; Yamashita et al. 2011). Approximately a 2-fold weaker binding affinity was seen for TRBP-D2 H188A compared to TRBP-D2 WT for all dsRNAs tested (Figure 5.5 & 5.6).

EMSA results suggested that mutation of the two positively charged residues (R189, K190) found in RNA binding region 2 of TRBP-D2 did not affect the RNA binding affinity of the dsRBD when mutated individually. However, the double mutation to mimic the amino acid composition of TRBP-D1 presents that the binding affinity is affected with the modification of both residues in TRBP-D2. TRBP-D2 R189Q-K190P shows a 2-3-fold weaker binding affinity for the dsRNAs tested compared to WT TRBP-D2 (Figure 5.5 & 5.6).

Structural analysis of RNA binding region 2 in TRBP-D2 highlighted that the proline residue (P186) is required for this region to attain certain rigidity and with the removal by substituting this residue for a glutamine residue decreased the flexibility of this region. To understand how this change in rigidity affects dsRNA binding of this region and the whole dsRBD, dsRNA binding tests were carried out with TRBP-D2 P186Q and various dsRNAs. The fluorescence anisotropy data highlighted that for all dsRNAs tested TRBP-D2 P186Q has a weaker binding affinity compared to WT TRBP-D2. A 2-fold difference in binding affinity can be seen for pre-let-7a and pre-miR-155 mod and slightly lower affinities for pre-miR-155 and pre-let-7a mod. The proline residue has not been seen to directly interact with dsRNA but the binding studies indicate it plays a role in the efficient binding of TRBP-D2 to dsRNAs. Possibly maintaining the correct positioning of the RNA interacting histidine residue.

(A)

(B)

(C)

(D)

Figure 5.6 Binding curves for WT TRBP-D2 and mutant constructs

(A) Interaction of Pre-miR 155 WT vs. protein constructs (B Interaction of Pre-miR 155 mod vs. protein constructs (C) Interaction of Pre-let7a vs. protein constructs (D) Interaction of Pre-let7a mod vs. protein constructs. Experimental data are averaged from three independent fluorescent anisotropy experiments.

5.4.3 Investigating RNA binding region 2 in TRBP-dsRBD1

Following investigations into RNA binding region 2 of TRBP-dsRBD2, this region was investigated in TRBP-D1. As described previously between the two dsRBDs there are major differences in amino acid composition in this region 2, to understand how this difference in amino acid composition affects the dsRNA binding capabilities of both dsRBDs, dsRNA binding studies were carried out on mutant protein constructs.

Single point mutations in RNA binding region 2 of TRBP-D1 were carried out to partially mimic the amino acid composition of TRBP-D2. To understand how the presence of a proline residue at the start of loop region can influence dsRNA binding. The conserved histidine residue within region 2 in TRBP-D1 was also mutated to an alanine to investigate this residues impact on dsRNA binding.

Mutagenesis of RNA binding region 2 in TRBP-D1 has shown that the substitution of a proline residue within this region (Q56P) has increased the RNA binding affinity of the dsRBD for Pre-miR-155 2-fold compared to WT TRBP-D1 (Figure 5.7). For the other dsRNAs tested a difference in binding affinities could not be observed. Pre-miR-155 features the most structurally non-perfect dsRNA duplex in this study so possibly mutating the glutamine to a proline residue in this region, has affected how this dsRBD recognises structural imperfection and could potentially help the dsRBD position itself around them more efficiently. This hypothesis requires further testing with other dsRNAs.

Interestingly mutating the histidine 58 to alanine did not reduce the dsRNA binding affinity, which was not seen for TRBP-D2 (Figure 5.7). TRBP-D1 H58A also reveals that with the removal of the histidine residue the dsRBD binds 2.5-fold stronger to pre-miR-155 similar to what was seen for TRBP-D1 Q56P. This suggests that by changing the amino acid composition of region 2 in TRBP-D1 could change how this dsRBD interacts with pre-miRNAs with structural imperfections in the dsRNA stem.

For pre-miR-155 mod, pre-let-7a and pre-let-7a mod these results indicate that the conserved histidine residue within this loop may not be required for the interaction of dsRNA for this dsRBD as removing the conserved histidine was shown to have no effect on the RNA binding affinity. However, for pre-miR-155, a more structurally imperfect dsRNA the removal of the histidine residue and changing the glutamine to a proline residue has increased the RNA binding affinity of TRBP-D1. This suggests the original amino acid composition of the loop region in TRBP-D1 could hinder the dsRBD from binding to dsRNAs with more structurally imperfect stems or have no effect on dsRNA binding.

(A)

| | Pre-miR -155 | | | | Pre-Let-7a | | | |
| | Native | | modified | | Native | | modified | |
	K_D (nM)	error	K_D (nm)	error	K_D (nm)	error	K_D (nm)	error
TRBP-D1	945	190	519	60	777	88	1015	142
TRBP-D1 Q56P	430	85	529	120	731	98	957	163
TRBP-D1 H58A	361	63	648	67	748	125	943	140

(B)
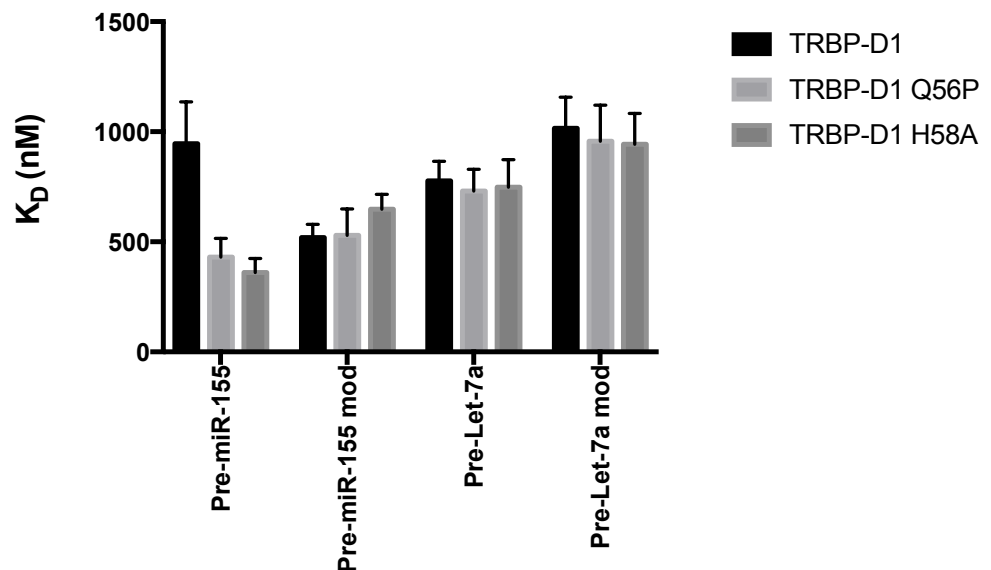
Figure 5.7 *dsRNA binding affinities of WT TRBP-D2 and mutant constructs*

(A)Table of binding affinities of WT and mutant constructs of TRBP-D2 (B) Bar chart format of binding affinities. Binding affinities were calculated from three independent fluorescent anisotropy experiments.
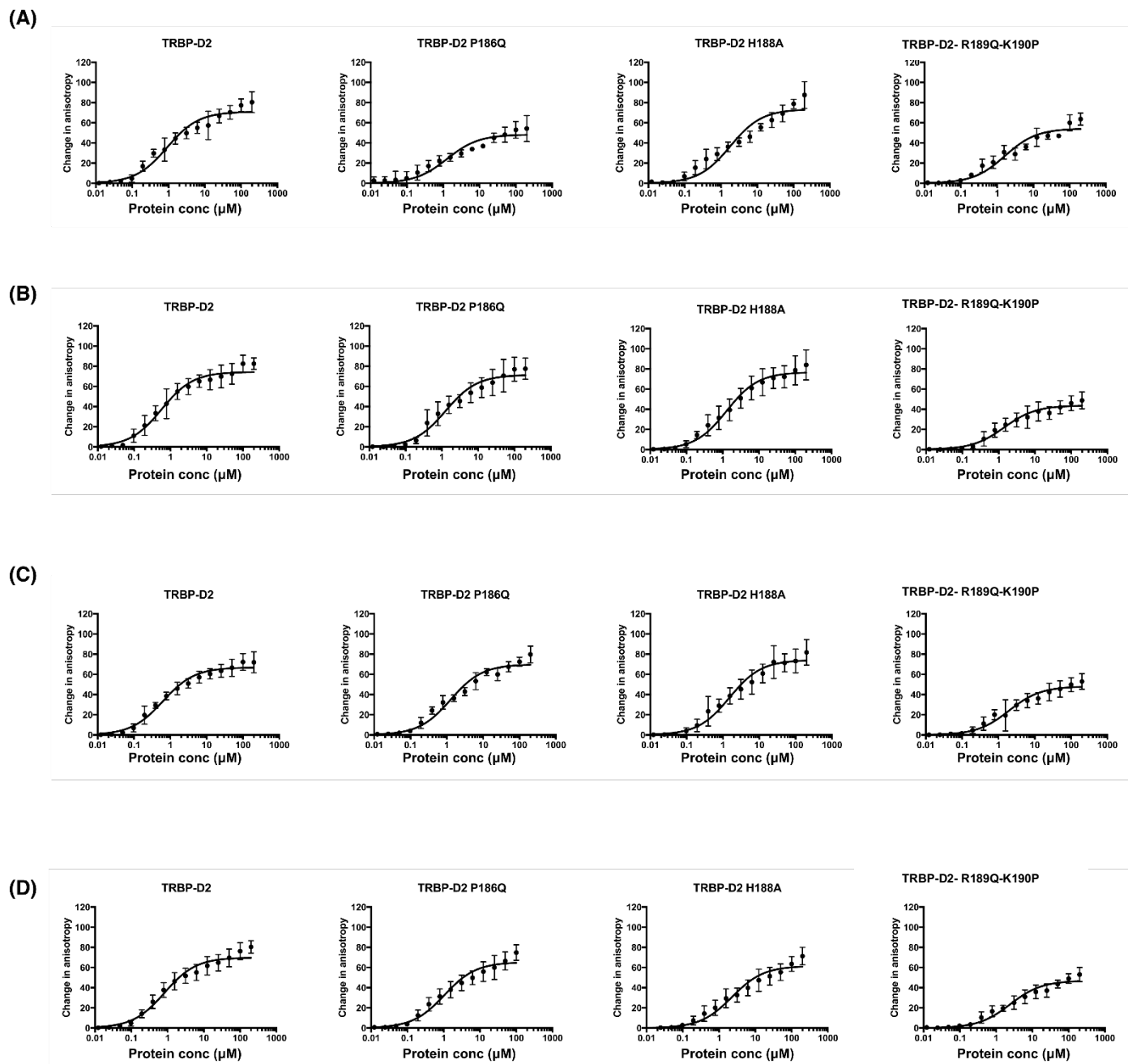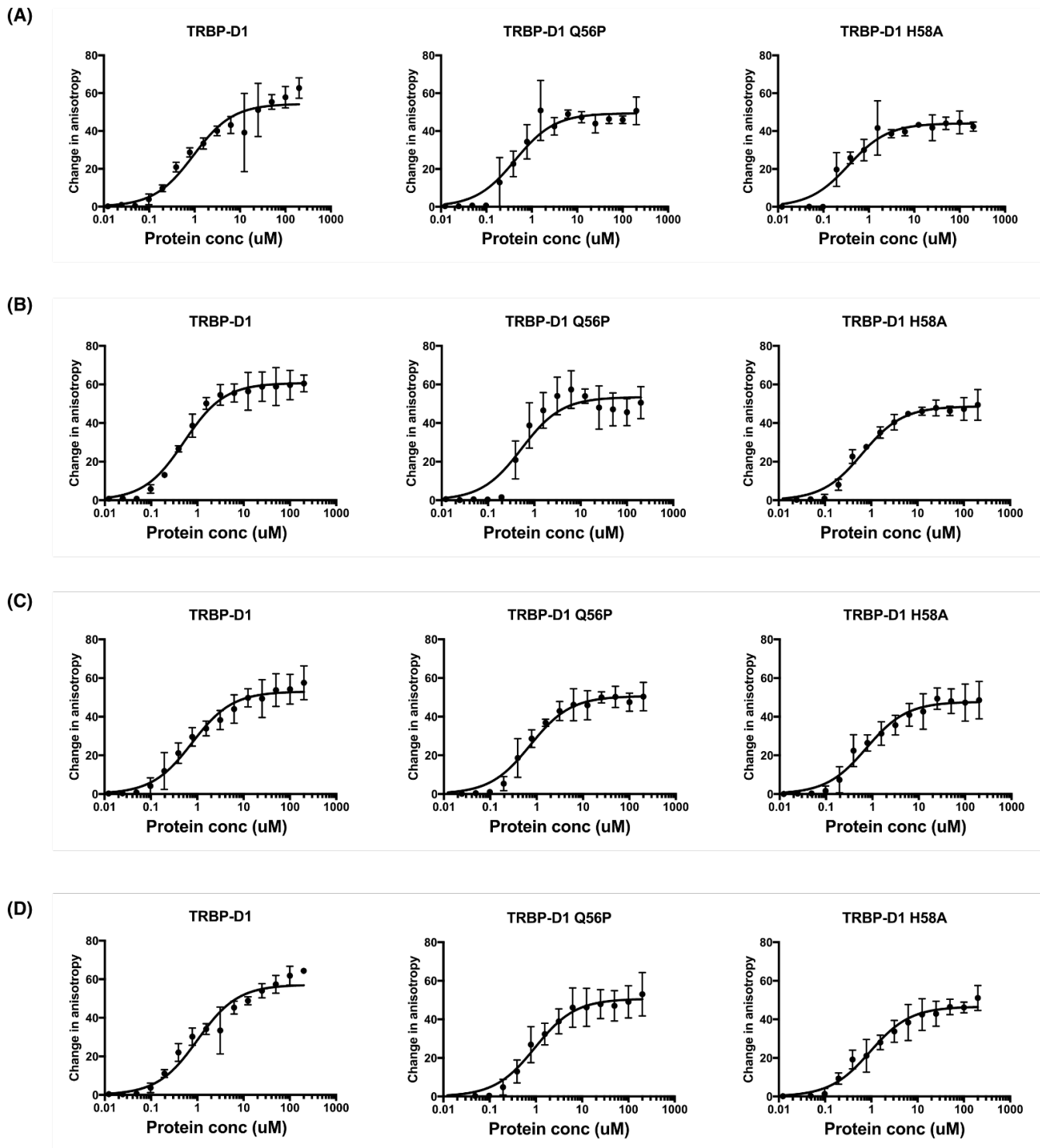
Figure 5.8 Binding curves for WT TRBP-D1 and mutant constructs

(A) Interaction of Pre-miR 155 WT vs. protein constructs (B Interaction of Pre-miR 155 mod vs. protein constructs (C) Interaction of Pre-let7a vs. protein constructs (D) Interaction of Pre-let7a mod vs. protein constructs. Experimental data are averaged from three independent fluorescent anisotropy experiments.

5.5 Discussion

DsRNA binding domains are often found in modular RNA binding proteins, in combination with other functional domains or as multiple copies (Mackereth and Sattler, 2012). TRBP contains three dsRBDs, two of which bind to dsRNA. The results have shown that the two N-terminal dsRBDs of TRBP do in fact bind to dsRNA and that the linker connecting the two dsRBD which has been shown not to interact with dsRNA, does enhance the interaction of the dsRBDs with dsRNA. The linker acts as a tether to potentially increase the RNA binding surface of the dsRBDs cooperatively or increasing the local concentration of one dsRBD when one is already bound. Interestingly, the RNA binding studies conducted with the separate dsRBDs of TRBP highlighted that both dsRBDs have similar RNA binding affinities for the dsRNAs tested, which is contrary to previous studies. An explanation for this could be due to the pH of the assay conducted, as it has been identified that TRBP-D1 exists in a state of equilibrium of folded and unfolded protein and investigating this has shown that an increase in pH disrupts this equilibrium between folded and unfolded protein. Studies here were carried out at pH 7.5 to ensure TRBP-D1 existed in a maximally folded state, which likely corresponds to its active state. Therefore, the contribution of the unfolded/folded state of equilibrium will not affect RNA binding and the investigation can focus on RNA binding region 2.

DsRBDs have been shown to bind to dsRNA independent of the dsRNA sequence (Masliah et al. 2013). As a result, it has been suggested that the dsRBDs found in TRBP recognise pre-miRNA through structural features within the dsRNA stem, pre-miRNA stems are highly enriched in bulges or non-canonical base pairs (Warf et al. 2011). This property of TRBP was investigated by analysing the binding affinity for native and modified pre-miRNAs sequences. The data showed that there was not a major difference in RNA binding between a structured dsRNA compared to a dsRNA with a perfect dsRNA stem with no helical defects for TRBP-D12, although a two-fold difference in binding affinity was observed for the single domains of TRBP for pre-miR-155-wt and its modified version. This suggests that two dsRBDs are required together to help facilitate TRBP bind to structurally imperfect dsRNAs.

Interestingly, the removal of structural features in pre-let-7a has reduced RNA binding affinity of the single dsRBDs of TRBP. This suggests that helical defects within a dsRNA stem may not only hinder the RNA binding affinity of dsRBDs but also aid. The positioning of the structural feature needs to also be taken into consideration. More dsRNAs need to be tested to further prove this but the investigation shows that the single dsRBDs of TRBP are affected by the presence of helical imperfections ubiquitous to pre-miRNAs. Yet again no difference in binding between pre-let-7a and its modified version was seen for TRBP-D12 suggesting the need for a multi-domain protein for the recognition of dsRNA.

RNA binding region 2 of dsRBDs has been well studied in various dsRNA binding proteins but the specific amino acid composition of this region and the differences seen within this region between dsRBDs has not been extensively investigated. How do the highly conserved and non-conserved residues in this region in TRBP play a role in dsRNA binding of the protein? Could the differences in amino acid composition in this region play a role in fine-tuning the dsRNA binding affinity of the dsRBD and therefore RNA binding properties of TRBP? Contrary to data surrounding other dsRBDs, the mutation of the conserved histidine residue in RNA binding region 2 did not knock out dsRNA binding in TRBP-D1 and TRBP-D2. Here it was shown that replacing the histidine residue with an alanine in TRBP-D1 had no significant effect on the RNA binding affinity for three out of the four dsRNAs tested. In fact, the RNA-binding affinity was enhanced 2-3 fold for of pre-miR-155. TRBP-D2 H188A on the other hand showed a reduction in RNA binding affinity for all dsRNAs tested (Figure 5.9). These data suggest that the conserved histidine residue in TRBP-D1 may not be involved in RNA binding of this region, as mutating this residue has not decreased the RNA binding affinity of the dsRBD, which was seen for TRBP-D2. The presence of a proline residue in region 2 of TRBP-D2 which is absent from TRBP-D1, was also examined. The data suggests that the introduction of a proline residue into region 2 of TRBP-D1 did not have a significant impact on the dsRNA binding affinity of this dsRBD for three out of the four dsRNAs tested. However, mutation of the proline residue in TRBP-D2 in region 2 (P186Q) did weaken binding affinity of this dsRBD for the dsRNAs investigated.

The mutational studies of RNA binding region 2 of TRBP-D2 have shown that each residue within this region is required for the efficient binding of not only this region but the whole dsRBD as point mutations have shown to decrease the RNA binding affinity.

(A)

		Pre-miR -155				Pre-Let-7a			
		Native		modified		Native		modified	
		K_D (nM)	error	K_D (nm)	error	K_D (nm)	error	K_D (nm)	error
TRBP-D1	WT	945	190	519	60	777	88	1015	142
	Q56P	430	85	529	120	731	98	957	163
	H58A	361	63	648	67	748	125	943	140
TRBP-D2	WT	914	103	571	59	630	57	824	78
	P186Q	1272	250	1265	202	1276	162	1047	149
	H188A	1798	316	1285	210	1527	261	2277	395
	R189Q-K190P	1643	280	1410	271	1954	371	2433	388
TRBP-D12	WT	381	58	247	36	292	45	285	38

(B) Pre-miR-155

(C) Pre-miR 155 mod

(D) Pre-let-7a

(E) Pre-let-7a mod
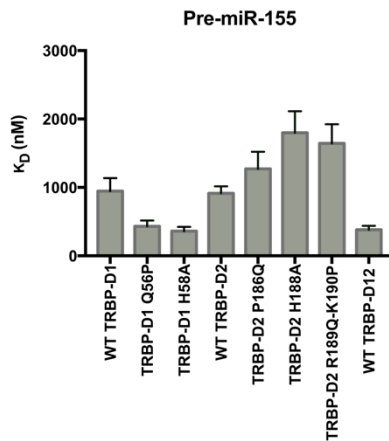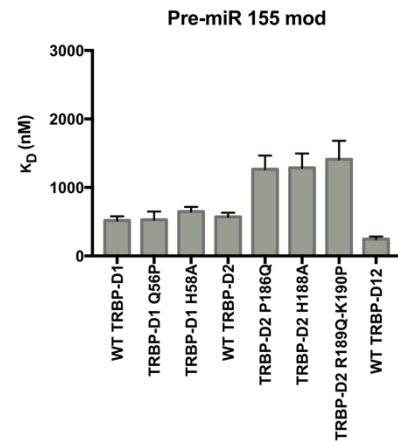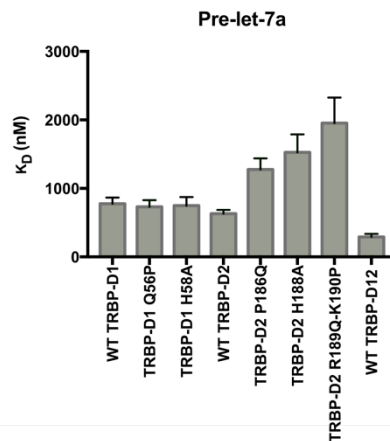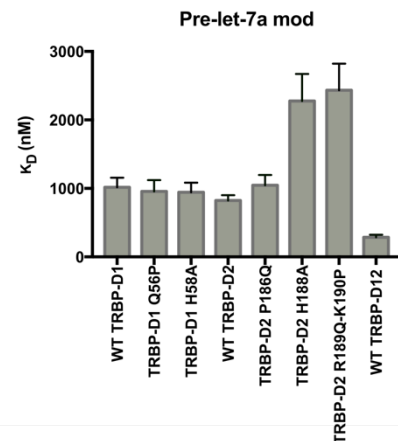
Figure 5.9 Summary of RNA binding affinities of TRBP constructs with dsRNAs

(A) Table of RNA binding affinities of TRBP constructs with various dsRNAs. Bar chart format of binding affinities for TRBP constructs for (B) Pre-miR-155 (C) Pre-miR-155 mod (D) Pre-let-7a (E) Pre-let-7a mod.

Chapter 6- Conclusions and Discussion

6.1 The folded/unfolded equilibrium of TRBP-D1 is affected by pH

TRBP-D1 has previously been shown to exist in an equilibrium between fully folded and unfolded state, whereas TRBP-D2 appeared to be stable (Benoit, 2013). Further investigation into this phenomenon showed that the equilibrium could be shifted towards the folded state by increasing the pH above 7.0. Trying to identify a condition whereby the protein is in its fully folded state was important, as dsRBDs have been shown to require the canonical fold to interact with dsRNA (Ryter and Schultz 1998). This also allowed for the investigation to focus on the RNA binding regions of TRBP-D1 with no contribution from the folded/unfolded equilibrium on RNA binding. A possible reason for pH effecting the protein conformation could be due to properties pH can have on ionizable amino acids which have a large influence on protein structure and stability (Grimsley et al. 2009).

The dsRBD (1732-1811) found in Dicer-1 (DCL1) of *Arabidopsis thaliana* was shown to be intrinsically disordered but acquired a folded structure when bound to dsRNA (Suarez et al. 2015). It was also seen to be disordered in tandem constructs with the other dsRBD found in DCL1. Although the dsRBD was intrinsically disordered it was seen to carry out its roles in RNA recognition. Disordered proteins and disordered regions within proteins are thought to enable multiple target recognition due their inherent flexibility (Dyson 2012). This suggests that the state of equilibrium between folded/unfolded in TRBP-D1 could be important for its function in RNA binding.

Like DCL1-dsRBD, TRBP-D1 also exists in this state of equilibrium in the double domain construct (TRBP-D12). RNA-binding studies of the dsRBDs of TRBP have shown that the two dsRBDs interacted with dsRNA with differing affinities, however the pH of the experiment was pH 6.5 (Benoit et al., 2013). In this study RNA binding studies performed at pH 7.5 (at which TRBP-D1 has been determined to be in a folded state) identified that the two dsRBDs bind

154

with similar affinities. This suggests a novel dsRBD feature found in TRBP whereby RNA binding of TRBP-D1 can be regulated by pH by altering the structure of TRBP-D1.

6.2 Pre-miRNA recognition by the modular protein TRBP

As presented in section 1.6, dsRBDs are commonly found in modular RNA binding proteins, both in multiple copies and in combination with other functional domains (Mackereth and Sattler, 2012; Tian et al., 2004). Mutations and truncations of single dsRBDs in a multi-domain protein have shown to diminish or completely abolish the ability of the protein to interact with dsRNA (Schmedt et al. 1995; Krovat and Jantsch 1996). By existing in multiple copies, the domains can allow the protein to bind with RNA with higher specificity and affinity (Stefl et al. 2006; Benoit et al. 2013). TRBP like other dsRBPs has been shown contain dsRBDs that fold independently with no intermolecular interaction between the domains (Benoit et al. 2013).

The RNA binding investigations in this study of the individual domains and the tandem domain construct showed that the modular design of the RNA binding portion of TRBP is required for a higher dsRNA binding affinity. The double domain construct was seen to bind 2-3-fold higher than the single dsRBDs. This suggests that tethering of the dsRBDs influences RNA binding of the dsRBDs by increasing their affinities for dsRNA.

A long and flexible linker such as the one found in TRBP between the two type-A dsRBDs has been suggested to allow modular proteins to recognise a diverse set of targets, whereas short linkers between dsRBDs has been suggested to predispose the domains to bind to a contiguous stretch of RNA (Tian et al. 2004). This could allow for structural features of pre-miRNAs to block RNA binding as shorter linkers could restrict the distance between dsRBDs required for binding around structural features. Interestingly, PACT contains a short linker (25aa) between the two type-A dsRBDs. Could this shorter linker play a role in the potential differential dsRNA binding between TRBP and PACT? The RNA binding data presented here identified that TRBP-D12 can bind to pre-miRNA with structural features tighter than the singular dsRBDs suggesting that linker is sufficiently long to allow the dsRBDs to bind

around the structural imperfections. However, for pre-miRNAs with a perfect duplex, a short linker such as the one found in PACT could be advantageous as linker length can be directly linked to the affinity of a modular protein. Assuming the free energy of individual dsRBDs binding is additive, the affinity of a protein with multiple dsRBDs could be as high as the product of the affinities of the individual dsRBDs. However, with a long flexible linker separating multiple dsRBDs, one dsRBD will bind to the dsRNA and the second domains sweeps a volume proportional to the length of the linker (Shamoo et al. 1995; Tian et al. 2004). In this volume, the concentration of the second dsRBD is different compared to in the free solution. As the linker length decreases the affinity of the second dsRBD binding increases due to the restricted sweeping volume which favours binding of the second dsRBD. Therefore, the affinity of the modular protein increases.

Linker length could therefore be advantageous or disadvantageous in modular proteins such as TRBP and PACT dependent on the substrate they are interacting with. Further work determining the RNA binding properties of PACT single and double domain constructs could elucidate how the difference in linker length contributes to dsRNA binding and why Dicer requires two dsRBPs for facilitating miRNA processing.

6.3 The first two dsRBDs of TRBP have variations in backbone dynamics between the two dsRBDs

^{15}N relaxation studies of the type-A dsRBDs of TRBP have previously and with this study have revealed subtle differences in backbone dynamics between the two dsRBDs on the ps-ns timescale. Interestingly the difference in dynamics localises to RNA binding region 2. Here, the dynamics of the two dsRBDs was further investigated and the difference was attributed to the amino acid composition of RNA binding region 2. A single amino acid difference between the two dsRBDs was able to modify the dynamics of this RNA binding region. The native sequence of TRBP-D2 has a proline residue at the N-terminus of the loop region where this is a glutamine in TRBP-D1. Proline residues within a polypeptide bond can impose conformational rigidity compared to other amino acids due to its distinctive cyclic structure.

Mutagenesis of the dsRBDs in TRBP-D12 to mimic the amino acid composition at region 2 of the other domain (TRBP-D1 Q56P & TRBP-D2 P186Q) were carried out. The results showed that the introduction of a proline residue into the β1-β2 loop in TRBP-D1 increased the rigidity of this loop region, while substituting out the proline residue in TRBP-D2 decreases the rigidity. The data suggests that dynamics of RNA binding region 2 region can be fine-tuned by a single amino acid.

Differences in the presence of proline residues in RNA binding region 2 seen in TRBP-D12 also exists between the type-A dsRBDs of PACT (Figure 6.1). Investigating whether a difference in dynamics also exists between the dsRBDs of PACT at RNA binding region 2 would determine whether this difference in dynamics is a conserved feature of multi-domain dsRBPs involved in miRNA biogenesis.

6.4 The contribution of amino acid composition of region 2 in RNA binding by the type-A dsRBDs of TRBP

Investigating the amino acid composition of RNA binding region 2 in both type-A dsRBDs of TRBP and the role this plays in RNA binding presented some interesting results. Mutagenesis of the conserved histidine residue in region 2 has shown to have no negative effect on RNA binding for TRBP-D1 and only weakened the RNA binding affinity of TRBP-D2 (2-3 fold). Interestingly, mutating the histidine residue to an alanine in TRBP-D1 enhanced the binding affinity of this dsRBD for pre-miR-155.

Investigations of region 2 in other dsRBDs has shown the importance of this region in RNA binding and its residues, with mutations of the histidine in the dsRBD of Xlrbpa knocking out RNA binding of the whole dsRBD (Krovat and Jantsch, 1996). Mutating the histidine residue in region 2 of TRBP-D2 showed to have only a small impact on the RNA binding affinity of this dsRBD. For all dsRNAs tested the RNA binding affinity decreased ~2 fold compared to WT-TRBP-D2. However, in this study the histidine residue in TRBP-D1 has been shown to be unnecessary for dsRNA binding.

Replacing the glutamine residue with a proline residue in region 2 of TRBP-D1 showed an increase in dsRNA binding affinity for pre-miR-155. Replacing the proline residue with a glutamine in RNA binding region 2 of TRBP-D2 had a negative effect on dsRNA binding. For all dsRNAs tested the TRBP-D2 P186Q mutant had a weaker binding affinity compared to WT TRBP-D2.

Region 2 has been reported to be important in RNA binding specifically the conserved histidine residue. However, the binding data in this study has suggested that the dsRBDs of TRBP can endure mutations in RNA binding region 2 without largely affecting the RNA binding of the dsRBDs. Mutagenesis of RNA binding region 2 in TRBP-D1 had no effect on the binding affinity of dsRNA with perfect (modified pre-miRNAs) or near perfect (pre-let-7a) dsRNA stem. Therefore, could the other two RNA binding regions be the main contributors to RNA binding with region 2 fine tuning the RNA binding of the dsRBD?

The differential binding modes seen for the dsRBDs of TRBP due to RNA binding region 2 could be important in fine tuning the affinity of the protein for dsRNA.

6.5 DsRNA binding by TRBP is affected by structural features of the pre-miRNA

Investigating the role of structural features on dsRNAs can play in dsRNA binding of TRBP showed that for pre-miR-155 a dsRNA with two 1-nt bulges in the dsRNA stem, the RNA binding affinities of the WT single dsRBD constructs were weaker than for pre-miR-155 (mod) with no structural imperfections in the dsRNA stem. The data presented here and by others suggest that the dsRBDs of TRBP can recognise structural features of pre-miRNAs (Acevedo et al. 2015). This in turn could help Dicer to discriminate between pre-miRNA substrates as Drosophila Dicer-1 has been shown to not distinguish or recognise structural features in the stem region of pre-miRNAs (Tsutsumi et al. 2011). TRBP has previously been shown to facilitate discrimination of pre-miRNA like species by Dicer. TRBP was shown to ensure the efficient processing of Dicer for a pre-miRNA in a tRNA crowded environment (Fareh et al. 2016). The processing of the pre-miRNA was severely hindered when Dicer was present alone. However, when complexed with TRBP the processing efficiency was not affected. This has not yet been linked to the dsRNA recognition property of TRBP, but one can hypothesise that the ability of TRBP in recognising structural features in pre-miRNAs could facilitate Dicer binding to pre-miRNAs in an RNA crowded environment.

TRBP recognition of dsRNA features could help facilitate the positioning of the protein on dsRNA and therefore help positioning of Dicer on the pre-miRNA. This can influence Dicer cleavage positioning, which would lead to a difference in iso-miR production (Lee and Doudna, 2012). TRBP has been shown to help tune miRNA length by Dicer dependent on base-mismatches in the pre-miRNA stem (Zhu et al. 2018). The data presented in this work supports this proposed model of TRBP dsRBDs recognising structural features to help Dicer positioning and processing.

6.6 The role of loop dynamics on dsRNA recognition

A difference in dynamics between dsRBDs in a multi-domain protein has been reported for the dsRBDs of PKR on a millisecond timescale (Yang et al. 2014). The study proposed a correlation between the difference in flexibility and dsRNA binding affinity and specificity. Here, to draw a link between dynamics and RNA binding of TRBP, RNA binding studies were carried out with WT and mutant TRBP constructs which altered the dynamics of RNA binding region 2. Altering the backbone dynamics of RNA binding region 2 in TRBP-D1 (Q56P) by increasing the rigidity has shown to have no substantial impact on the RNA binding of this dsRBD for dsRNAs that lack structural features in the dsRNA helix. However, for pre-miR-155 a dsRNA with structural features in the dsRNA stem, TRBP-D1 Q56P showed a 2-fold increase in RNA binding affinity compared to WT- TRBP-D1. Decreasing the rigidity of region 2 in TRBP-D2 showed that this dsRBD required the $\beta 1$-$\beta 2$ loop to attain certain rigidity for a higher RNA binding affinity of this dsRBD for the dsRNAs tested.

The NMR relaxation studies also suggested that residues in the $\beta 1$-$\beta 2$ loop in TRBP-D2 may be experiencing conformational exchange between states not seen in the loop region of TRBP-D1. This phenomeon has not previously been reported for the dsRBDs of TRBP. This could be important as multiple conformational sub-states of this region may allow binding to various dsRNAs substrates.

The RNA binding and NMR relaxation studies presented here suggest that rigidity of RNA binding region 2 may be required in dsRNA recognition of pre-miRNAs with structurally imperfect dsRNA stems. This is contrary to what has been suggested for the dsRBDs of PKR, where one dsRBD with higher flexibility has been suggested to be able adapt to a variety of viral dsRNAs with imperfect A-form structures (Yang et al. 2014). A caveat for the hypothesis made for the dsRBDs of PKR and the dsRBDs of TRBP in this study is the low amount of dsRNAs tested against in their investigation.

6.7 Contribution of RNA binding region 2 in dsRBDs

RNA binding studies of WT and mutant TRBP-D2 constructs showed that amino acid composition of RNA binding region 2 is important for the RNA binding activity of this dsRBD but can endure mutations in this region. Investigation of the amino acid composition of RNA binding region 2 in TRBP-D1 showed that replacing the histidine residue with an alanine and altering the dynamics enhanced the binding affinity for pre-miR-155 and did not affect the affinity for the other dsRNAs tested. This suggests that the native sequence in region 2 of TRBP-D1 may hinder the dsRBD from binding to structurally imperfect dsRNAs. The results suggest that dynamics and amino acid composition of RNA binding region 2 may play a role in sensing imperfections in miRNA precursors. However, this assumption requires further investigation with RNA binding studies of more dsRNAs.

The RNA binding studies carried out on TRBP-D1 and TRBP-D2 indicate that the contribution of RNA binding region 2 may differ between the two dsRBDs dependent on the dsRNA it is binding to. This could be a feature of multi-domain proteins to be able to recognise and bind to various dsRNAs with higher specificity and affinity.

Although RNA binding region 2 has conserved amino acids between different dsRBDs (GPxHxx), the differences in residues, length and dynamics of this region in the β1-β2 loop could play a role in the differential RNA binding properties of the different dsRBDs. For example, the β1-β2 loop region in the dsRBD found in human Dicer has been shown to be three amino acids shorter than the canonical length. Dynamics studies into the loop region in Dicer-dsRBD showed there was only a minimal increase in dynamics compared with the rest of the domain and suggested to contribute to the high binding affinity of this dsRBD (Wostenberg et al. 2012). This study has shown that the dynamics between the dsRBDs of TRBP can also be fine-tuned by a single amino acid.

The dsRBD domain has been found in functionally diverse proteins and must recognise a vast amount of RNAs. The data presented in this thesis suggests that RNA binding region 2 in the dsRBDs of TRBP can influence the dsRNA

properties of a dsRBD. Out of the three RNA binding regions in dsRBDs, most of the sequence length variability exists in region 2 (Tian et al. 2004). Therefore, the amino acid composition and dynamics of this loop region could be different between various dsRBDs to accommodate for the vast amount of dsRNA targets.
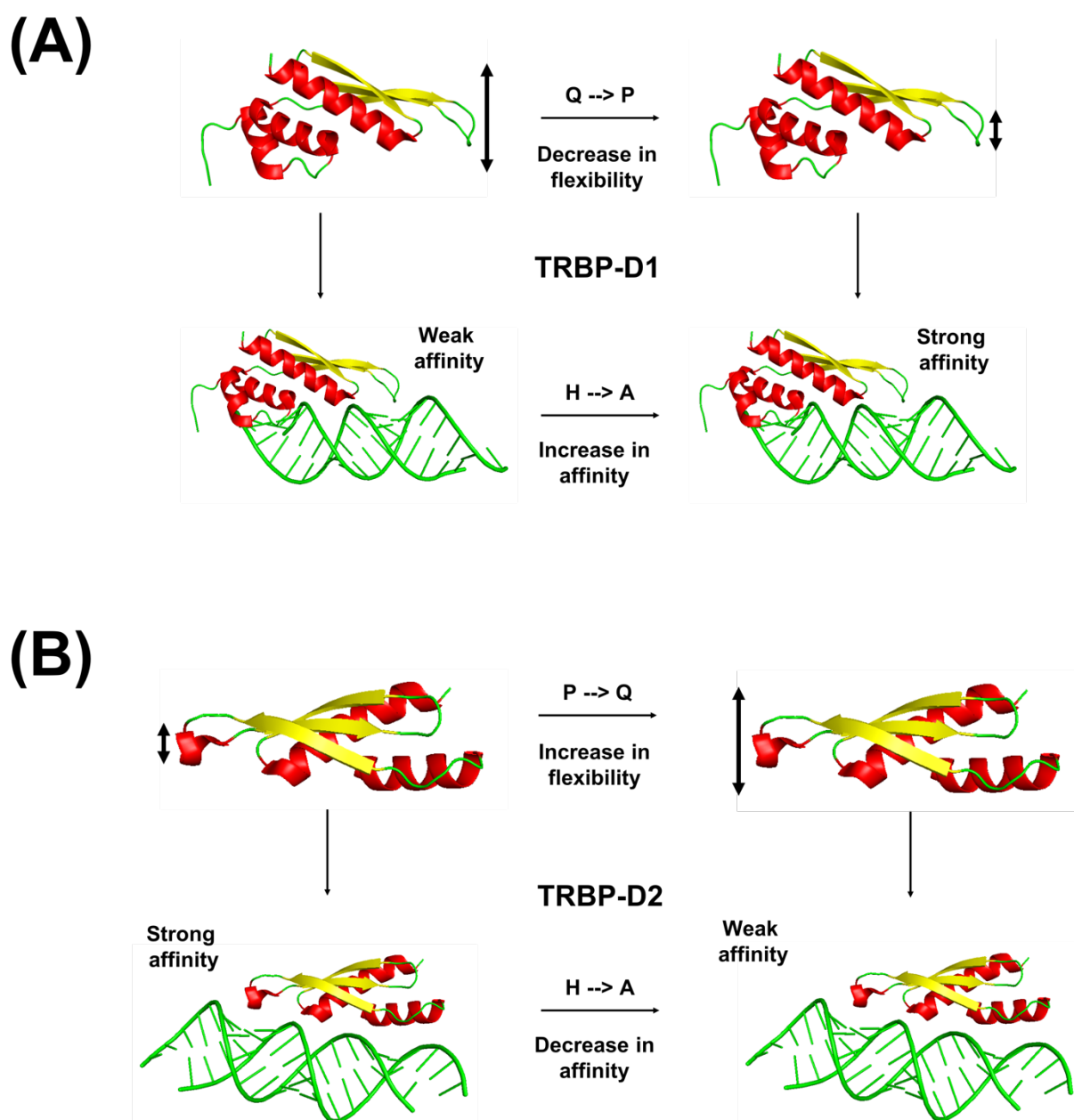
Figure 6.2 Contribution of the residues in the type-A dsRBDs of TRBP in structure and the dsRNA binding property of RNA binding region 2.

(A) Model of RNA binding region 2 in TRBP-D1 and the contribution of amino acid in RNA binding and flexibility of region 2 elucidated from the data. (B) Model of RNA binding region 2 in TRBP-D2 and the contribution of amino acid in RNA binding and flexibility of region 2 elucidated from the data. Created using Pymol. Arrows on the dsRBD structure indicate loop region flexibility.

Chapter 7- Suggestions for future research

This thesis sets out to further elucidate how TRBP interacts with dsRNAs, with the hope of understanding how TRBP interaction with dsRNA links to pre-miRNA processing by Dicer. This work has created a toolbox of mutants that change the RNA binding properties of the dsRBDs in TRBP, which can be used to study how dsRNA interaction properties of TRBP govern processing of pre-miRNAs by Dicer.

Changing the flexibility or amino acid composition of region 2 in one or both dsRBDs of TRBP and investigating this effect on Dicer processing such as iso-miR production and the processing rate of Dicer would elaborate on how the RNA binding properties of TRBP contribute to Dicer processing. As mentioned before a study highlighted that TRBP acts as a substrate gatekeeper for Dicer, how this is governed by the RNA binding by TRBP is yet to be elucidated. Would weakening the RNA binding affinity of one or both dsRBDs by mutagenesis affect the gate keeping property of TRBP? Would changing the dynamics of RNA binding regions affect recognition of the correct substrate for Dicer?

TRBP has been shown to diffuse along dsRNAs and a potential property of this diffusion is to scan the structure of dsRNAs to find the correct cleavage site for Dicer. Therefore, what role do the RNA binding regions in TRBP play in the diffusion mechanism and in turn the correct positioning of Dicer? Could the difference in dynamics observed for the two dsRBDs play a role in the diffusion of TRBP? The results from this thesis suggests dynamics in region 2 could play a role in recognition of structures within the dsRNA stem of pre-miRNAs, therefore this region could play a role in diffusion of the protein along pre-miRNAs with various structural features.

Investigations of the RNA binding properties of PACT could further help to understand how the RNA binding interaction of dsRBPs play a role in their function in miRNA biogenesis. Potential differences seen in the binding of dsRNA by TRBP and PACT could help to elucidate why Dicer requires two functional dsRBPs for miRNA biogenesis and how each protein carries out its differential role in Dicer processing of pre-miRNAs.

The studies here were conducted on the single type-A dsRBDs and a multi-domain construct of just the type-A dsRBD (TRBP-D12). However full length TRBP consists of three dsRBDs connected by long flexible linkers. Therefore, it would be interesting to investigate the RNA binding properties studied in this work in the context of the full-length protein. The data in this thesis highlighted that the two dsRBDs in TRBP have different dynamics in RNA binding region 2. How will switching the dynamics in RNA binding region 2 between TRBP-D1 and TRBP-D2 affect the dsRNA recognition by the full-length protein; furthermore, how will this affect the miRNA processing capabilities of TRBP in the miRNA biogenesis pathway?

Appendix

Oligonucleotides for DNA templates of dsRNAs

Pre-let-7a

Forward Primer

TAATACGACTCACTATAGGAGGTAGTAGGTTGTATAGTTTTAGGGTCAC
ACCCACCACTGGGAGATAACTATACAATCTACTGTCTTCCC

Reverse Primer

GGGAAGACAGTAGATTGTATAGTTATCTCCCAGTGGTGGGTGTGACCCT
AAAACTATACAACCTACTACCTCCTATAGTGAGTCGTATTA

Pre-miR-155

Forward Primer

TAATACGACTCACTATAGGAATGCTAATCGTGATAGGGGTTTTTGCCTC
CAACTGACTCCTACATATTAGCATTCCCA

Reverse Primer

TTTTGGGAATGCTAATATGTAGGAGTCAGTTGGAGGCAAAAACCCCTAT
CACGATTAGCATTCCTATAGTGAGTCGTATTA

Pre-let-7a mod

Forward Primer

TAATACGACTCACTATAGGAGGCAGTAGGTTGTATAGTTTTAGGGTCA
CACCCACCACTGGGAGATAACTATACAACCTACTGCCTCCC

Reverse Primer

GGGAGGCAGTAGGTTGTATAGTTATCTCCCAGTGGTGGGTGTGACCCT
AAAACTATACAACCTACTGCCTCCTATAGTGAGTCGTATTA

Pre-miR-155 mod

Forward Primer

TAATACGACTCACTATAGGTTAATGCTAATCGTGATAGGGGTTTTTGC
CTCCAACTGACCCCTATCACGATTAGCATTAAC

Reverse Primer

TGTTAATGCTAATCGTGATAGGGGTCAGTTGGAGGCAAAAACCCCTAT
CACGATTAGCATTAACCTATAGTGAGTCGTATTA

Primers for Mutagenesis of TRBP
TRBP-D1 Q56P

Forward Primer

GAA GGC CGA GGG CCC GGC CCA CCA ACC GAA C

Reverse Primer

GTT CGG TTG GTG GGC CGG GCC CTC GGC CTT C

TRBP-D1 H58A

Forward Primer

GAA GGC CGA GGG CCA AGC CGC GCA ACC GAA CTT CAC CTT TC

Reverse Primer

GAA AGG TGA AGT TCG GTT GCG CGG CTT GGC CCT CGG CCT TC

TRBP-D2 P186Q

Forward Primer

CCA AGA GAG CGG CCA GGC ACA TCG CAA GGA G

Reverse Primer

CTC CTT GCG ATG TGC CTG GCC GCT CTC TTG G

TRBP-D2 H188A

Forward Primer

CAA GAG AGC GGC CCT GCA GCG CGC AAG GAG TTC ACC ATG

Reverse Primer

CAT GGT GAA CTC CTT GCG CGC TGC AGG GCC GCT CTC TTG

TRBP-D2 R189Q

Forward Primer

GAG CGG CCC TGC ACA TCA GAA GGA GTT CAC CAT GAC

Reverse Primer

GTC ATG GTG AAC TCC TTC TGA TGT GCA GGG CCG CTC

TRBP-D2 K190P

Forward Primer

GCC CTG CAC ATC GCC CGG AGT TCA CCA TGA C

Reverse Primer

GTC ATG GTG AAC TCC GGG CGA TGT GCA GGG C

TRBP-D2 RK189QP

Forward Primer

CAG GTC ATG GTG AAC TCC GGC TGA TGT GCA GGG CCG CTC

Reverse Primer

GAG CGG CCC TGC ACA TCA GCC GGA GTT CAC CAT GAC CTG

TRBP Constructs

List of plasmids used for protein expression

Tag	Protein
His-MBP	TRBP-D1
His-MBP	TRBP-D2
His-MBP	TRBP-D12
His-MBP	TRBP-D1 Q56P
His-MBP	TRBP-D1 H58A
His-MBP	TRBP-D2 P186Q
His-MBP	TRBP-D2 H188A
His-MBP	TRBP-D2 R189Q
His-MBP	TRBP-D2 K190P
His-MBP	TRBP-D2 R189Q-K190P

DNA Sequence of protein constructs

6x His with MBP tag and 3C cleavage site

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCATGAAAATCGAAG
AAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCT
CGCTGAAGTCGGTAAGAAATTCGAGAAGATACCGGAATTAAAGTCACC
GTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCA
ACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGT
GGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCGGACAAAGCGT
TCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACG
GCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTA
TAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCC
GGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTT
CAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGG
GGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAAGACGTG
GGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGAC
CTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAG
AAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGT
GGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGG
TACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCT
GAGCGCAGGTATTAAC
GCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCCTCG
AAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAA
ACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAA
AGATCCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAAT
CATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTAC
TGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCT
GAAAGACGCGCAGACTCGTATCACCAAGGGCCTGGAAGT
TCTGTTCCAGGGACCAGCAATG

MGSSHHHHHHSSMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVE
HPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKL
YPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAK
GKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGL
TFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYG

VTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNK
DKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTA
VINAASGRQTVDEALKDAQTRITKGLEVLFQ^GPAM

TRBP-D1_19-99

ATCGAGCAGATGTTAGCCGCCAACCCGGGTAAAACCCCGATTAGCCTG
CTGCAAGAGTACGGTACACGCATCGGCAAAACACCGGTGTATGACTTA
CTGAAGGCCGAGGGCCAAGCCCACCAACCGAACTTCACCTTTCGCGTG
ACCGTTGGTGACACAAGCTGCACCGGCCAGGGCCCTAGTAAGAAGGC
CGCAAAGCACAAGGCCGCCGAAGTGGCCCTGAAACACCTGAAGGGCG
GTAGC

IEQMLAANPGKTPISLLQEYGTRIGKTPVYDLLKAEGQAHQPNFTFRVTVGD
TSCTGQGPSKKAAKHKAAEVALKHLKGGS

TRBP-D2_157-228

GAGTGTAATCCGGTGGGCGCACTGCAGGAACTGGTGGTGCAGAAGGG
CTGGCGCTTACCGGAGTACACCGTGACCCAAGAGAGCGGCCCTGCAC
ATCGCAAGGAGTTCACCATGACCTGCCGTGTGGAGCGCTTCATCGAGA
TCGGCAGCGGTACCAGCAAGAAGCTGGCCAAGCGCAATGCAGCCGCC
AAGATGCTGCTGCGCGTGCACACAGTG

ECNPVGALQELVVQKGWRLPEYTVTQESGPAHRKEFTMTCRVERFIEIGS
GTSKKLAKRNAAAKMLLRVHTV

TRBP-D12_19-228

ATCGAGCAGATGTTAGCCGCCAACCCGGGTAAAACCCCGATTAGCCTG
CTGCAAGAGTACGGTACACGCATCGGCAAAACACCGGTGTATGACTTA
CTGAAGGCCGAGGGCCAAGCCCACCAACCGAACTTCACCTTTCGCGTG
ACCGTTGGTGACACAAGCTGCACCGGCCAGGGCCCTAGTAAGAAGGC
CGCAAAGCACAAGGCCGCCGAAGTGGCCCTGAAACACCTGAAGGGCG
GTAGCATGCTGGAACCGGCCCTGGAGGACAGCAGCAGCTTCAGTCCG
CTGGATAGTAGCCTGCCGGAAGACATCCCGGTGTTCACCGCAGCAGCA
GCCGCAACCCCTGTTCCTAGTGTTGTTCTGACCCGCAGTCCGCCGATG
GAGTTACAACCGCCGGTTAGCCCTCAGCAGAGTGAGTGTAATCCGGTG

GGCGCACTGCAGGAACTGGTGGTGCAGAAGGGCTGGCGCTTACCGGA
GTACACCGTGACCCAAGAGAGCGGCCCTGCACATCGCAAGGAGTTCAC
CATGACCTGCCGTGTGGAGCGCTTCATCGAGATCGGCAGCGGTACCAG
CAAGAAGCTGGCCAAGCGCAATGCAGCCGCCAAGATGCTGCTGCGCG
TGCACACAGTG

IEQMLAANPGKTPISLLQEYGTRIGKTPVYDLLKAEGQAHQPNFTFRVTVGD
TSCTGQGPSKKAAKHKAAEVALKHLKGGSMLEPALEDSSSFSPLDSSLPE
DIPVFTAAAAATPVPSVVLTRSPPMELQPPVSPQQSECNPVGALQELVVQK
GWRLPEYTVTQESGPAHRKEFTMTCRVERFIEIGSGTSKKLAKRNAAAKML
LRVHTV

Abbreviations

Ago	Argonaute
CD	Circular dichroism
dsRNA	Double-stranded RNA
dsRBD	Double-stranded RNA binding domain
dsRBP	Double-stranded RNA binding protein
DNA	Deoxyribonucleic acid
DSS	4,4-dimethyl-4silapentane-1-sulphonic acid
DTT	Dithiothreitol
EDTA	Ethylenediamineteraacetic acid
EMSA	Electrophoretic mobility shift assay
EXSY	Exchange Spectroscopy
FA	Fluorescence anisotropy
FTSC	Fluorescein 5-thiosemicarbazide
HEPES	4-(2-hydroxyl)-1-piperazineethanesulfonic acid
HSQC	Heteronuclear Single Quantum Coherence
IVT	*In vitro transcription*
miRNA	microRNA
mRNA	Messenger RNA
NMR	Nuclear magnetic resonance
PDB	Protein Data Bank
Pre-miRNA	Precursor microRNA
Pri-miRNA	Primary microRNA
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel
SEC	Size Exclusion Chromatography
SEC-MALS	Size Exclusion Chromatography coupled to Multi-Angle Light Scattering
siRNA	Short interfering RNA
TCEP	Tris(2-carboxyethel) phosphine
TROSY	Transverse relaxation-optimized spectroscopy
UTR	Untranslated region

References

Agarwal, G.P., and Barnwal, R.P. 2009. Double-Stranded RNA Binding Domain (dsRBD) of PKR Shows Variable Dynamics in the Presence of Bacteriophage Pf1: An NMR Insight and its Possible Implications. The Open Structural Biology Journal **3**(1): 42–50. doi:10.2174/1874199100903010042 .

Bartel, D.P. 2004. MicroRNAs Genomics, Biogenesis, Mechanism, and Function. Cell **116**(2): 281–297. doi:10.1016/S0092-8674(04)00045-5 .

Bartel, D.P. 2009. MicroRNAs: Target Recognition and Regulatory Functions. Cell **136**(2): 215–233. doi:10.1016/j.cell.2009.01.002 .

Benoit, M., Imbert, L., Palencia, A., Pérard, J., Ebel, C., Boisbouvier, J., and Plevin, M.J. 2013. The RNA-binding region of human TRBP interacts with microRNA precursors through two independent domains. Nucleic Acids Research **41**(7): 4241–4252. doi:10.1093/nar/gkt086 .

Benoit, M.P., and Plevin, M.J. 2013. Backbone resonance assignments of the micro-RNA precursor binding region of human TRBP. Biomolecular NMR assignments **7**(2): 229–33. doi:10.1007/s12104-012-9416-8 .

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**(6818): 363. doi:10.1038/35053110 .

Bevilacqua, P., George, C., Samuel, C., and Cech, T. 1998. Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helixes. Biochemistry **37**(18): 6303–16. doi:10.1021/bi980113j .

Boehm, M., and Slack, F. 2005. A Developmental Timing MicroRNA and Its Target Regulate Life Span in C. elegans. Science: 1954–1957. doi:10.1126/science.1115596 .

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA (New York, N.Y.) **10**(2): 185–91.

Bycroft, M., Grünert, S., Murzin, A.G., Proctor, M., and Johnston, S.D. 1995. NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. The EMBO Journal **14**(14): 3563–3571. doi:10.1002/j.1460-2075.1995.tb07362.x .

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. 2002. The

Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes & Development: 2733–2742. doi:10.1101/gad.1026102 .

Chang, K.-Y.Y., and Ramos, A. 2005. The double-stranded RNA-binding motif, a versatile macromolecular docking platform. The FEBS journal **272**(9): 2109–17. doi:10.1111/j.1742-4658.2005.04652.x .

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**(7051): 740–4. doi:10.1038/nature03868 .

Chiliveri, S., and Deshmukh, M.V. 2014. Structure of RDE-4 dsRBDs and mutational studies provide insights into dsRNA recognition in the Caenorhabditis elegans RNAi pathway. The Biochemical journal **458**(1): 119–30. doi:10.1042/bj20131347 .

Cloonan, N., Wani, S., Xu, Q., Gu, J., Lea, K., Heater, S., Barbacioru, C., Steptoe, A.L., Martin, H.C., Nourbakhsh, E., Krishnan, K., Gardiner, B., Wang, X., Nones, K., Steen, J.A., Matigian, N.A., Wood, D.L., Kassahn, K.S., Waddell, N., Shepherd, J., Lee, C., Ichikawa, J., McKernan, K., Bramlett, K., Kuersten, S., and Grimmond, S.M. 2011. MicroRNAs and their isomiRs function cooperatively to target common biological pathways. Genome biology **12**(12): R126. doi:10.1186/gb-2011-12-12-r126 .

Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C.M. 2006. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proceedings of the National Academy of Sciences of the United States of America: 7024–9.

Doyle, M., and Jantsch, M.F. 2002. New and old roles of the double-stranded RNA-binding domain. Journal of structural biology **140**(1–3): 147–53. doi:10.1016/S1047-8477(02)00544-0 .

Erard, M., Barker, D., Amalric, F., Jeang, K., and Gatignol, A. 1998. An Arg/Lys-rich core peptide mimics TRBP binding to the HIV-1 TAR RNA upper-stem/loop. Journal of molecular biology **279**(5): 1085–99. doi:10.1006/jmbi.1998.1831 .

Fareh, M., Yeom, K.-H.H., Haagsma, A.C., Chauhan, S., Heo, I., and Joo, C. 2016. TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments. Nature communications **7**: 13694. doi:10.1038/ncomms13694 .

Feng, Y., Zhang, X., Graves, P., and Zeng, Y. 2012. A comprehensive analysis of precursor microRNA cleavage by human Dicer. RNA **18**(11):

2083–2092. doi:10.1261/rna.033688.112 .

Fried, M. 1989. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. Electrophoresis **10**(5–6): 366–76. doi:10.1002/elps.1150100515 .

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome research **19**(1): 92–105. doi:10.1101/gr.082701.108 .

Fukunaga, R., Han, B.W., Hung, J.-H.H., Xu, J., Weng, Z., and Zamore, P.D. 2012. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell **151**(3): 533–46. doi:10.1016/j.cell.2012.09.027 .

Gabel, F., Wang, D., Madern, D., Sadler, A., Dayie, K., Daryoush, M., Schwahn, D., Zaccai, G., Lee, X., and Williams, B. 2006. Dynamic Flexibility of Double-stranded RNA Activated PKR in Solution. Journal of Molecular Biology **359**(3): 610–623. doi:10.1016/j.jmb.2006.03.049 .

Gao, F.-B. 2010. Context-dependent functions of specific microRNAs in neuronal development. Neural Development: 25. doi:10.1186/1749-8104-5-25 .

Gatignol, A., Buckler, C., and and cellular biology, J.K. 1993. Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human P1/dsI kinase and Drosophila staufen. doi:10.1128/MCB.13.4.2193 .

Gatignol, A., Buckler-White, A., Berkhout, B., and Jeang, K. 1991. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. Science **251**(5001): 1597–1600. doi:10.1126/science.2011739 .

Gleghorn, M.L., Gong, C., Kielkopf, C.L., and Maquat, L.E. 2013. Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. Nature structural & molecular biology **20**(4): 515–24. doi:10.1038/nsmb.2528 .

Gleghorn, M.L., and Maquat, L.E. 2014. 'Black sheep' that don't leave the double-stranded RNA-binding domain fold. Trends in Biochemical Sciences **39**(7): 328–340. doi:10.1016/j.tibs.2014.05.003 .

Gouet, P., Robert, X., and Courcelle, E. 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic acids research **31**(13): 3320–3.

Gredell, J.A., Dittmer, M.J., Wu, M., Chan, C., and Walton, S. 2010. Recognition of siRNA asymmetry by TAR RNA binding protein. Biochemistry

49(14): 3148–55. doi:10.1021/bi902189s .

Greenfield, N.J. 2007. Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols **1**(6): nprot.2006.202. doi:10.1038/nprot.2006.202 .

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: 23–34. doi:10.1016/S0092-8674(01)00431-7 .

Gurtner, A., Falcone, E., Garibaldi, F., and Piaggio, G. 2016. Dysregulation of microRNA biogenesis in cancer: the impact of mutant p53 on Drosha complex activity. Journal of experimental & clinical cancer research : CR **35**: 45. doi:10.1186/s13046-016-0319-x .

Ha, M., and Kim, N.V. 2014. Regulation of microRNA biogenesis. Nature Reviews Molecular Cell Biology **15**(8): nrm3838. doi:10.1038/nrm3838 .

Haase, A.D., Jaskiewicz, L., Zhang, H., Lainé, S., Sack, R., Gatignol, A., and Filipowicz, W. 2005. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO reports **6**(10): 961–7. doi:10.1038/sj.embor.7400509 .

Hammond, S.M., Caudy, A.A., and Hannon, G.J. 2001. Post-transcriptional gene silencing by double-stranded RNA. Nature Reviews Genetics: 110. doi:10.1038/35052556 .

Han, J., Lee, Y., Yeom, K.-H.H., Kim, Y.-K.K., Jin, H., and Kim, V. 2004. The Drosha-DGCR8 complex in primary microRNA processing. Genes & development **18**(24): 3016–27. doi:10.1101/gad.1262504 .

He, L., Thomson, M.J., 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: 828. doi:10.1038/nature03552 .

Heinicke, L.A., Nallagatla, S., Hull, C.M., and Bevilacqua, P.C. 2011. RNA helical imperfections regulate activation of the protein kinase PKR: effects of bulge position, size, and geometry. Rna. doi:10.1261/rna.2636911 .

Heyam, A., Coupland, C.E., Dégut, C., Haley, R.A., Baxter, N.J., Jakob, L., Aguiar, P.M., Meister, G., Williamson, M.P., Lagos, D., and Plevin, M.J. (n.d.). Conserved asymmetry underpins homodimerization of Dicer-associated double-stranded RNA-binding proteins. Nucleic Acids Research: gkx928-. doi:10.1093/nar/gkx928 .

Heyam, A., Lagos, D., and Plevin, M. 2015. Dissecting the roles of TRBP and PACT in double-stranded RNA recognition and processing of noncoding RNAs. Wiley Interdisciplinary Reviews: RNA **6**(3): 271–289. doi:10.1002/wrna.1272 .

Ikura, M., Kay, L., and Bax, A. 1990. A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry **29**(19): 4659–67.

Jansson, M.D., and Lund, A.H. 2012. MicroRNA and cancer. Molecular oncology **6**(6): 590–610. doi:10.1016/j.molonc.2012.09.006 .

Jarymowycz, V.A., and Stone, M.J. 2006. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. Chem Rev **106**(5): 1624–1671. doi:10.1021/cr040421p .

Johnson, C.D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J., Chin, L., Brown, D., and Slack, F.J. 2007. The let-7 microRNA represses cell proliferation pathways in human cells. Cancer research: 7713–22. doi:10.1158/0008-5472.CAN-07-1083 .

Johnston, S.D., Brown, N., Gall, J., and Jantsch, M. 1992. A conserved double-stranded RNA-binding domain. Proceedings of the National Academy of Sciences **89**(22): 10979–10983. doi:10.1073/pnas.89.22.10979 .

Kerner, P., gnan, S., Marchand, L., gnan, B., and Vervoort, M. 2011. Evolution of RNA-Binding Proteins in Animals: Insights from Genome-Wide Analysis in the Sponge Amphimedon queenslandica. Molecular Biology and Evolution **28**(8): 2289–2303. doi:10.1093/molbev/msr046 .

Kharrat, A., Macias, M.J., Gibson, T.J., Nilges, M., and Pastore, A. 1995. Structure of the dsRNA binding domain of E. coli RNase III. The EMBO Journal **14**(14): 3572–3584. doi:10.1002/j.1460-2075.1995.tb07363.x .

Khvorova, A., Reynolds, A., and Jayasena, S.D. 2003. Functional siRNAs and miRNAs Exhibit Strand Bias. Cell: 209–216. doi:10.1016/S0092-8674(03)00801-8 .

Kim, N.V., Han, J., and Siomi, M.C. 2009. Biogenesis of small RNAs in animals. Nature Reviews Molecular Cell Biology: 126–139. doi:10.1038/nrm2632 .

Kim, Y., Yeo, J., Lee, J., Cho, J., Seo, D., Kim, J.-S., and Kim, N.V. 2014. Deletion of Human tarbp2 Reveals Cellular MicroRNA Targets and Cell-

Cycle Function of TRBP. Cell Reports **9**(3): 1061–1074. doi:10.1016/j.celrep.2014.09.039 .

Kleckner, I.R., and Foster, M.P. 2011. An introduction to NMR-based approaches for measuring protein dynamics. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1814**(8): 942–968. doi:10.1016/j.bbapap.2010.10.012 .

Koh, H.R., Kidwell, M.A., Doudna, J., and Myong, S. 2017. RNA Scanning of a Molecular Machine with a Built-in Ruler. Journal of the American Chemical Society **139**(1): 262–268. doi:10.1021/jacs.6b10387 .

Koh, H.R., Kidwell, M.A., Ragunathan, K., Doudna, J.A., and Myong, S. 2013. ATP-independent diffusion of double-stranded RNA binding proteins. Proceedings of the National Academy of Sciences of the United States of America **110**(1): 151–6. doi:10.1073/pnas.1212917110 .

Kok, K.H., Ng, M.-H.H., Ching, Y.-P.P., and Jin, D.-Y.Y. 2007. Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. The Journal of biological chemistry **282**(24): 17649–57. doi:10.1074/jbc.M611768200 .

Koscianska, E., Starega-Roslan, J., and Krzyzosiak, W.J. 2011. The role of Dicer protein partners in the processing of microRNA precursors. PloS one **6**(12): e28548. doi:10.1371/journal.pone.0028548 .

Kranick, J.C., Chadalavada, D.M., Sahu, D., and Showalter, S.A. 2017. Engineering double-stranded RNA binding activity into the Drosha double-stranded RNA binding domain results in a loss of microRNA processing function. PloS one **12**(8): e0182445. doi:10.1371/journal.pone.0182445 .

Krol, J., Sobczak, K., Wilczynska, U., Drath, M., Jasinska, A., Kaczynska, D., and Krzyzosiak, W.J. 2004. Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. The Journal of biological chemistry **279**(40): 42230–9. doi:10.1074/jbc.m404931200 .

Krovat, B., and Jantsch, M. 1996. Comparative mutational analysis of the double-stranded RNA binding domains of Xenopus laevis RNA-binding protein A. The Journal of biological chemistry **271**(45): 28112–9. doi:10.1074/jbc.271.45.28112 .

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: 3903–3908. doi:10.1073/pnas.0712321105 .

Lakomek, N.-A.A., Ying, J., and Bax, A. 2012. Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. Journal of biomolecular NMR **53**(3): 209–21. doi:10.1007/s10858-012-9626-5 .

Laraki, G., Clerzius, G., Daher, A., Melendez-Peña, C., Daniels, S., and Gatignol, A. 2008. Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions. RNA biology **5**(2): 92–103. doi:10.4161/rna.5.2.6069 .

Lau, P.-W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. 2012. The molecular architecture of human Dicer. Nat Struct Mol Biology **19**(4): 436. doi:10.1038/nsmb.2268 .

Lau, P.-W.W., Potter, C.S., Carragher, B., and MacRae, I.J. 2009. Structure of the human Dicer-TRBP complex by electron microscopy. Structure (London, England : 1993) **17**(10): 1326–32. doi:10.1016/j.str.2009.08.013 .

Lee, H.Y., and Doudna, J.A. 2012. TRBP alters human precursor microRNA processing in vitro. RNA (New York, N.Y.) **18**(11): 2012–9. doi:10.1261/rna.035501.112 .

Lee, H.Y., Zhou, K., ith, A., 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): 6568–76. doi:10.1093/nar/gkt361 .

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): 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. 2003. The nuclear RNase III Drosha initiates microRNA processing. Nature **425**(6956): 415–9. doi:10.1038/nature01957 .

Lee, Y., Hur, I., Park, S.-Y.Y., Kim, Y.-K.K., Suh, M.R., and Kim, V. 2006. The role of PACT in the RNA silencing pathway. The EMBO journal **25**(3): 522–32. doi:10.1038/sj.emboj.7600942 .

Lee, Y., Jeon, K., Lee, J.-T.T., Kim, S., and Kim, V. 2002. MicroRNA maturation: stepwise processing and subcellular localization. The EMBO journal **21**(17): 4663–70.

Levitt, M.H. (2008) Spin dynamics: Basics of nuclear magnetic resonance, 2nd edition. John Wiley & Sons,Ltd.

Lewis, B., Shih, I., Jones-Rhoades, M., and Cell, B.D. 2003. Prediction of

mammalian microRNA targets.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. 2004. Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. Nature Structural and Molecular Biology: nsmb777. doi:10.1038/nsmb777 .

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, M.J., 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): 1437–1441. doi:10.1126/science.1102513 .

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, R.H., and Golub, T.R. 2005. MicroRNA expression profiles classify human cancers. Nature **435**(7043): 834. doi:10.1038/nature03702 .

Lu, J., and Hall, K. 1997. Tertiary structure of RBD2 and backbone dynamics of RBD1 and RBD2 of the human U1A protein determined by NMR spectroscopy. Biochemistry **36**(34): 10393–405. doi:10.1021/bi9709811 .

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. 2004. Nuclear export of microRNA precursors. Science (New York, N.Y.) **303**(5654): 95–8. doi:10.1126/science.1090599 .

Lunde, B.M., Moore, C., and Varani, G. 2007. RNA-binding proteins: modular design for efficient function. Nature Reviews Molecular Cell Biology: 479–490. doi:10.1038/nrm2178 .

Mackereth, C.D., and Sattler, M. 2012. Dynamics in multi-domain protein recognition of RNA. Current Opinion in Structural Biology **22**(3): 287–296. doi:10.1016/j.sbi.2012.03.013 .

MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. 2008. In vitro reconstitution of the human RISC-loading complex. Proceedings of the National Academy of Sciences of the United States of America **105**(2): 512–7. doi:10.1073/pnas.0710869105 .

MacRae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, Z.W., Adams, P.D., and Doudna, J.A. 2006. Structural Basis for Double-Stranded RNA Processing by Dicer. Science **311**(5758): 195–198. doi:10.1126/science.1121638 .

Masliah, G., Barraud, P., and Allain, F.H. 2013. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cellular and molecular life sciences : CMLS **70**(11): 1875–95. doi:10.1007/s00018-012-1119-x .

Masliah, G., Maris, C., König, S.L., Yulikov, M., Aeschimann, F., Malinowska, A.L., Mabille, J., Weiler, J., Holla, A., Hunziker, J., Meisner-Kober, N., Schuler, B., Jeschke, G., and Allain, F. 2018. Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains. The EMBO Journal **37**(6): e97089. doi:10.15252/embj.201797089 .

Melo, S.A., Moutinho, C., Ropero, S., Calin, G.A., Rossi, S., Spizzo, R., Fernandez, A.F., Davalos, V., Villanueva, A., Montoya, G., Yamamoto, H., Schwartz, S., and Esteller, M. 2010. A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. Cancer cell **18**(4): 303–15. doi:10.1016/j.ccr.2010.09.007 .

Melo, S.A., Ropero, S., Moutinho, C., Aaltonen, L.A., Yamamoto, H., Calin, G.A., Rossi, S., Fernandez, A.F., Carneiro, F., Oliveira, C., Ferreira, B., Liu, C.-G.G., Villanueva, A., Capella, G., Schwartz, S., Shiekhattar, R., and Esteller, M. 2009. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. Nature genetics **41**(3): 365–70. doi:10.1038/ng.317 .

Mueller, G.A., Miller, M.T., DeRose, E.F., Ghosh, M., London, R.E., and Hall, T.M. 2010. Solution structure of the Drosha double-stranded RNA-binding domain. Silence **1**(1): 1–5. doi:10.1186/1758-907x-1-2 .

Nanduri, S., Rahman, F., Williams, B.R., and Qin, J. 2000. A dynamically tuned double-stranded RNA binding mechanism for the activation of antiviral kinase PKR. The EMBO Journal **19**(20): 5567–5574. doi:10.1093/emboj/19.20.5567 .

niels, S., Melendez-Peña, C.E., Scarborough, R.J., Daher, A., Christensen, H.S., Far, M., Purcell, D.F., Lainé, S., and Gatignol, A. 2009. Characterization of the TRBP domain required for Dicer interaction and function in RNA interference. BMC Molecular Biology **10**(1): 38. doi:10.1186/1471-2199-10-38 .

nivas Chakravarthy, Sternberg, S.H., Kellenberger, C.A., and Doudna, J.A. 2010. Substrate-specific kinetics of Dicer-catalyzed RNA processing. Journal of molecular biology **404**(3): 392–402. doi:10.1016/j.jmb.2010.09.030 .

nli, A., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. 2004. Processing of primary microRNAs by the Microprocessor complex. Nature **432**(7014): 231. doi:10.1038/nature03049 .

Notredame, C., Higgins, D., and Heringa, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology **302**(1): 205–17. doi:10.1006/jmbi.2000.4042 .

Olive, V., Jiang, I., and He, L. 2010. mir-17-92, a cluster of miRNAs in the

midst of the cancer network. The International Journal of Biochemistry & Cell Biology: 1348–1354. doi:10.1016/j.biocel.2010.03.004 .

Pagano, J.M., Clingman, C.C., and Ryder, S.P. 2011. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. RNA (New York, N.Y.) **17**(1): 14–20. doi:10.1261/rna.2428111 .

Paithankar, H., Jadhav, P.V., Naglekar, A.S., Sharma, S., and Chugh, J. 2018. 1H, 13C and 15N resonance assignment of domain 1 of trans-activation response element (TAR) RNA binding protein isoform 1 (TRBP2) and its comparison with that of isoform 2 (TRBP1). Biomolecular NMR Assignments **12**(1): 189–194. doi:10.1007/s12104-018-9807-6 .

Palmer, A. 2001. Nmr probes of molecular dynamics: overview and comparison with other techniques. Annual review of biophysics and biomolecular structure **30**: 129–55. doi:10.1146/annurev.biophys.30.1.129 .

Palmer, A.G. 2004. NMR Characterization of the Dynamics of Biomacromolecules. Chemical Reviews **104**(8): 3623–40. doi:10.1021/cr030413t .

Park, H., Davies, M., Langland, J., Chang, H., Nam, Y., Tartaglia, J., Paoletti, E., Jacobs, B., Kaufman, R., and Venkatesan, S. 1994. TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. Proceedings of the National Academy of Sciences of the United States of America **91**(11): 4713–7. doi:10.1073/pnas.91.11.4713 .

Paroo, Z., Ye, X., Chen, S., and Liu, Q. 2009. Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. Cell **139**(1): 112–22. doi:10.1016/j.cell.2009.06.044 .

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**: 86. doi:10.1038/35040556 .

Peng, Y., and Croce, C.M. 2016. The role of MicroRNAs in human cancer. Signal transduction and targeted therapy **1**: 15004. doi:10.1038/sigtrans.2015.4 .

Peters, L., and Meister, G. 2007. Argonaute Proteins: Mediators of RNA Silencing. Molecular Cell: 611–623. doi:10.1016/j.molcel.2007.05.001 .

Pratt, A.J., and MacRae, I.J. 2009. The RNA-induced Silencing Complex: A Versatile Gene-silencing Machine. Journal of Biological Chemistry **284**(27):

17897–17901. doi:10.1074/jbc.r900012200 .

Qin, H., Chen, F., Huan, X., Machida, S., Song, J., and Yuan, Y. 2010. Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. RNA (New York, N.Y.) **16**(3): 474–81. doi:10.1261/rna.1965310 .

Ramos, A., Grünert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., Johnston, D., and Varani, G. 2000. RNA recognition by a Staufen double-stranded RNA-binding domain. The EMBO Journal **19**(5): 997–1009. doi:10.1093/emboj/19.5.997 .

Rasia, R.M., Mateos, J., Bologna, N.G., Burdisso, P., Imbert, L., Palatnik, J.F., and Boisbouvier, J. 2010. Structure and RNA Interactions of the Plant MicroRNA Processing-Associated Protein HYL1. Biochemistry **49**(38): 8237–8239. doi:10.1021/bi100672x .

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, R.H., and Ruvkun, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature: 901. doi:10.1038/35002607 .

Ryter, J., and Schultz, S. 1998. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. The EMBO journal **17**(24): 7505–13. doi:10.1093/emboj/17.24.7505 .

Saenger, W., and Saenger, W. 1984. Principles of Nucleic Acid Structure. springer: 385–431. doi:10.1007/978-1-4612-5190-3_18 .

Schmedt, C., Green, S., Manche, L., Taylor, D., Ma, Y., and Mathews. 1995. Functional characterization of the RNA-binding domain and motif of the double-stranded RNA-dependent protein kinase DAI (PKR). Journal of molecular biology **249**(1): 29–44. doi:10.1006/jmbi.1995.0278 .

Shamoo, Y., Abdul-Manan, N., and Williams, K.R. 1995. Multiple RNA binding domains (RBDs) just don't add up. Nucleic Acids Research **23**(5): 725–728. doi:10.1093/nar/23.5.725 .

Singh, M., Castillo, D., Patel, C.V., and Patel, R.C. 2011. Stress-induced phosphorylation of PACT reduces its interaction with TRBP and leads to PKR activation. Biochemistry **50**(21): 4550–60. doi:10.1021/bi200104h .

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: 1434–1437. doi:10.1126/science.1102514 .

Stefl, R., Oberstrass, F., Hood, J., and Cell, J.M. 2010a. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove.

Stefl, R., Oberstrass, F.C., Hood, J.L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R.B., and Allain, F. 2010b. The Solution Structure of the ADAR2 dsRBM-RNA Complex Reveals a Sequence-Specific Readout of the Minor Groove. Cell **143**(2): 225–237. doi:10.1016/j.cell.2010.09.026 .

Takahashi, T., Miyakawa, T., Zenno, S., Nishi, K., Tanokura, M., and Ui-Tei, K. 2013. Distinguishable in vitro binding mode of monomeric TRBP and dimeric PACT with siRNA. PloS one **8**(5): e63434. doi:10.1371/journal.pone.0063434 .

Teplova, M., Wohlbold, L., Khin, N.W., Izaurralde, E., and Patel, D.J. 2011. Structure-function studies of nucleocytoplasmic transport of retroviral genomic RNA by mRNA export factor TAP. Nature Structural & Molecular Biology: 990. doi:10.1038/nsmb.2094 .

Tian, B., Bevilacqua, P.C., Diegelman-Parente, A., and Mathews, M.B. 2004. The double-stranded-RNA-binding motif: interference and much more. Nature reviews Molecular cell biology **5**(12): 1013.

Tomari, Y., and Zamore, P.D. 2005. Perspective: machines for RNAi. Genes & development **19**(5): 517–29. doi:10.1101/gad.1284105 .

Ucci, J.W., Kobayashi, Y., Choi, G., Alexandrescu, A.T., and Cole, J.L. 2007. Mechanism of Interaction of the Double-Stranded RNA (dsRNA) Binding Domain of Protein Kinase R with Short dsRNA Sequences†. Biochemistry **46**(1): 55–65. doi:10.1021/bi061531o .

Wang, Z., Hartman, E., Roy, K., Chanfreau, G., and Feigon, J. 2011. Structure of a Yeast RNase III dsRBD Complex with a Noncanonical RNA Substrate Provides New Insights into Binding Specificity of dsRBDs. Structure **19**(7): 999–1010. doi:10.1016/j.str.2011.03.022 .

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., 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): 9942–9955. doi:10.1093/nar/gkw631 .

Wightman, B., Ha, I., and Cell, R.G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.

Wilson, R.C., Tambe, A., Kidwell, M., Noland, C.L., Schneider, C.P., and Doudna, J.A. 2015. Dicer-TRBP Complex Formation Ensures Accurate Mammalian MicroRNA Biogenesis. Molecular Cell **57**(3): 397–407. doi:10.1016/j.molcel.2014.11.030 .

Wostenberg, C., Lary, J.W., Sahu, D., Acevedo, R., Quarles, K.A., Cole, J.L., and Showalter, S.A. 2012. The Role of Human Dicer-dsRBD in Processing Small Regulatory RNAs. PLoS ONE **7**(12): e51829. doi:10.1371/journal.pone.0051829 .

Wostenberg, C., Quarles, K.A., and Showalter, S.A. 2010. Dynamic origins of differential RNA binding function in two dsRBDs from the miRNA "microprocessor" complex. Biochemistry **49**(50): 10728–36. doi:10.1021/bi1015716 .

Yamashita, S., Nagata, T., Kawazoe, M., Takemoto, C., Kigawa, T., Güntert, P., Kobayashi, N., Terada, T., Shirouzu, M., Wakiyama, M., Muto, Y., and Yokoyama, S. 2011. Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. Protein Science **20**(1): 118–130. doi:10.1002/pro.543 .

Yang, L.-Q.Q., Sang, P., Tao, Y., Fu, Y.-X.X., Zhang, K.-Q.Q., Xie, Y.-H.H., and Liu, S.-Q.Q. 2014. Protein dynamics and motions in relation to their functions: several case studies and the underlying mechanisms. Journal of biomolecular structure & dynamics **32**(3): 372–93. doi:10.1080/07391102.2013.770372 .

Yang, S., Chen, H.-Y., Yang, J., Machida, S., Chua, N.-H., and Yuan, A.Y. 2010. Structure of Arabidopsis HYPONASTIC LEAVES1 and Its Molecular Implications for miRNA Processing. Structure **18**(5): 594–605. doi:10.1016/j.str.2010.02.006 .

Yuan, Y.-R., Pei, Y., Ma, J.-B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H.-Y., Dauter, Z., Tuschl, T., and Patel, D.J. 2005. Crystal Structure of A. aeolicus Argonaute, a Site-Specific DNA-Guided Endoribonuclease, Provides Insights into RISC-Mediated mRNA Cleavage. Molecular Cell: 405–419. doi:10.1016/j.molcel.2005.07.011 .

Zearfoss, R.N., and Ryder, S.P. 2012. End-Labeling Oligonucleotides with Chemical Tags After Synthesis. **941**: 181–193. doi:10.1007/978-1-62703-113-4_14 .

Zhu, L., Kandasamy, S.K., and Fukunaga, R. 2018. Dicer partner protein tunes the length of miRNAs using base-mismatch in the pre-miRNA stem. Nucleic acids research **46**(7): 3726–3741. doi:10.1093/nar/gky043 .

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization

prediction. Nucleic Acids Research **31**(13): 3406–3415. doi:10.1093/nar/gkg595 .