Dissertation
submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
M.Res. Felix Preston Williams
born in: Levallois-Perret (France)
Oral examination: 08.03.2021

Structure and RNA-binding preferences of the C-terminal domains of TRIM2 and TRIM3

Referees: Dr Orsolya Barabas
Prof. Claudio Joazeiro

Abstract

TRIM2 and TRIM3 are two highly homologous members of the tripartite motif (TRIM) protein family. They are primarily expressed in neurons and have roles in cell differentiation and polarisation. Although several roles have been proposed linked to their E3 ligase activity, which is mediated by their N-terminal domains, their proposed role in RNA binding, mediated by their C-terminal NHL domains, has not yet been studied. Here I present an initial characterisation of the RNA binding motif of the TRIM2 and TRIM3 NHL domains, demonstrating that they preferentially bind cytosine rich sequences, put forward a hypothesis as to the likely orientation of the RNA along the top surface of the NHL domain, as well as present structures for the TRIM2 NHL and TRIM3 filamin domains. Additionally, I show a difference in stability between the TRIM2 and TRIM3 NHL domains in solution, which I assessed by biophysical measurements and molecular dynamics simulations. I then explore possible interactions between the NHL and filamin domains and how these might differ between TRIM2 and TRIM3. The results presented show differences in the motif and mode of RNA binding as well as in inter-domain interactions when compared to the well studied homologues of these proteins, Brat and TRIM71.

Zusammenfassung

TRIM2 und TRIM3 sind zwei homologe Vertreter der Familie der tripartite motif (TRIM)-Proteinfamilie. Sie werden vor allem in Neuronen exprimiert und spielen eine Rolle in der Differenzierung von Zellen und der Ausbildung der Zellpolarität. Obwohl mehrere Funktionen im Zusammenhang mit der E3-Ligaseaktivität der N-terminalen Domänen dieser Proteine vorgeschlagen wurden, wurde die Fähigkeit der C-terminalen Domänen RNA zu binden bislang nicht genauer untersucht. In dieser Arbeit präsentiere ich eine erste Beschreibung des RNA-Motivs der NHL-Domänen von TRIM2 und TRIM3 und zeige, dass diese bevorzugt Cytosin-reiche Sequenzen binden. Ich schlage ein Modell der Orientierung der RNA auf der Oberseite der NHL-Domäne vor und beschreibe die Kristallstrukturen der TRIM2 NHL und TRIM3 Filamin-Domänen. Außerdem zeige ich anhand von biophysikalischen Daten und Molekulardynamiksimulationen, dass sich die NHL-Domänen von TRIM2 und TRIM3 in ihrer Stabilität unterscheiden. Ich beschreibe auch mögliche Interaktionen zwischen der NHL- und Filamin-Domäne und wie sich diese Interaktionen bei TRIM2 und TRIM3 unterscheiden. Zusammen zeigen die Ergebnisse, dass sich TRIM2 und TRIM3 sowohl im Hinblick auf das bevorzugte RNA-Motiv und die Bindungsart, als auch in ihren Interdomän-Interaktionen von den besser untersuchten Homologen TRIM71 und Brat unterscheiden.

Publications

F. P. Williams, K. Haubrich, C. Perez-Borrajero, and J. Hennig, "Emerging RNA-binding roles in the TRIM family of ubiquitin ligases", *Biological Chemistry*, vol. 400, p1443-1464, 2019.

Contents

Abbreviations

Bis-Tris	Bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane
cDNA	Complementary DNA
CLIP-seq	Cross-linking immunoprecipitation and sequencing
CSP	Chemical shift perturbation
CV	Column volumes
DSF	Differential scanning fluorimetry
dsRNA	Double stranded RNA
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
HECT	Homologous to E6-AP C-terminus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSQC	Heteronuclear single quantum coherence spectroscopy
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal calorimetry
MD	Molecular dynamics
miRNA	Micro RNA
mRNA	Messenger RNA
NF-L	Neurofilamin light chain
NHL	NCL-1, HT2A, Lin41
NMR	Nuclear magnetic resonance spectroscopy
NOE	Nuclear Overhauser effect
OD	Optical density
PCR	Polymerase chain reaction
ppm	Parts per million
RBCC	RING, B-box, coiled coil
RBD	RNA binding domain
RBR	RING between RING
RING	Really interesting new gene
RIP-seq	RNA immunoprecipitation and sequencing
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
SANS	Small angle neutron scattering
SAXS	Small angle X-ray scattering
SEC	Size exclusion chromatography
SIA	Scaffold independent analysis
SIRPA	Signal-regulatory protein alpha
snRNP	Small nuclear ribonucleoproteins

SPRY	Spore lysis A and ryanodine receptors
ssRNA	Single stranded RNA
TCEP	Tris(2-carboxyethyl)phosphine
TRIM	Tripartite motif
Tris	Tris(hydroxymethyl)aminomethane
UTR	Untranslated region

1 Introduction

TRIM proteins are a family of proteins characterised by the possession of a ubiquitin E3 ligase domain. I will therefore, to provide context, give a very brief introduction to ubiquitination before discussing TRIM proteins specifically.

1.1 Ubiquitination

As E3 ligases, TRIM proteins catalyse the last step of the ubiquitination cascade, a series of reactions by which proteins are post-translationally modified by additions of the 76 residue protein ubiquitin, leading to a variety of downstream fates (Hershko and Ciechanover, 1998; Komander and Rape, 2012).

Ubiquitination can consist of the addition of a single ubiquitin to the target protein (monoubiquitination) or addition of a single ubiquitin to multiple sites on the target protein (multi-monoubiquitination) but often consists of additions of chains of ubiquitin molecules to the target protein (polyubiquitination). In polyubiquitination, ubiquitins are added to pre-existing ubiquitin chains by aminolysis reactions between the added ubiquitin's C terminus and the amino groups of the already present ubquitin's lysine sidechains or N-terminus. As there are seven lysines in ubiquitin's sequence (K6, K11, K27, K29, K33, K48 and K63) and one N terminal residue (M1) this gives 8 different possible amide bonds that a second ubiquitin can form to a ubiquitin already attached to the target protein. Ubiquitin chains can be made up of multiple ubiquitins connected by the same linkage (homotypic) or by varying linkages (heterotypic). As each ubiquitin has multiple lysines, chains can even be branched. The different ubiquitin chain linkages can be conceptualised as a ubiquitin code which may be read by a variety of proteins possessing ubiquitin binding domains and triggers specific outcomes for the tagged protein depending on the ubiquitin chain linkage it has acquired (Akutsu et al., 2016; Komander and Rape, 2012).

Common outcomes and the linkages to which they are associated are as follows. The most well know role of ubiquitination is to target proteins for degradation by the 26S proteasome. A variety of linkages can specify this fate, most classically K48 linked polyubiquitin chains but also K11 polyubiquitination which triggers degradation of the targets of the anaphase promoting complex, thereby participating in cell cycle regulation (Chau et al., 1989; Matsumoto et al., 2010). Ubiquitinated proteins can also be targeted to non-proteasomal degradation through at variety of mechanisms including recruiment to autophagosomes and lysosomes, both of which are linked to K63 polyubiquitination (Huang et al., 2006; Kirkin et al., 2009). Moreover, ubiquitination can affect protein localisation, as in the case of the monoubiquitination of p53 which causes it to be trafficked

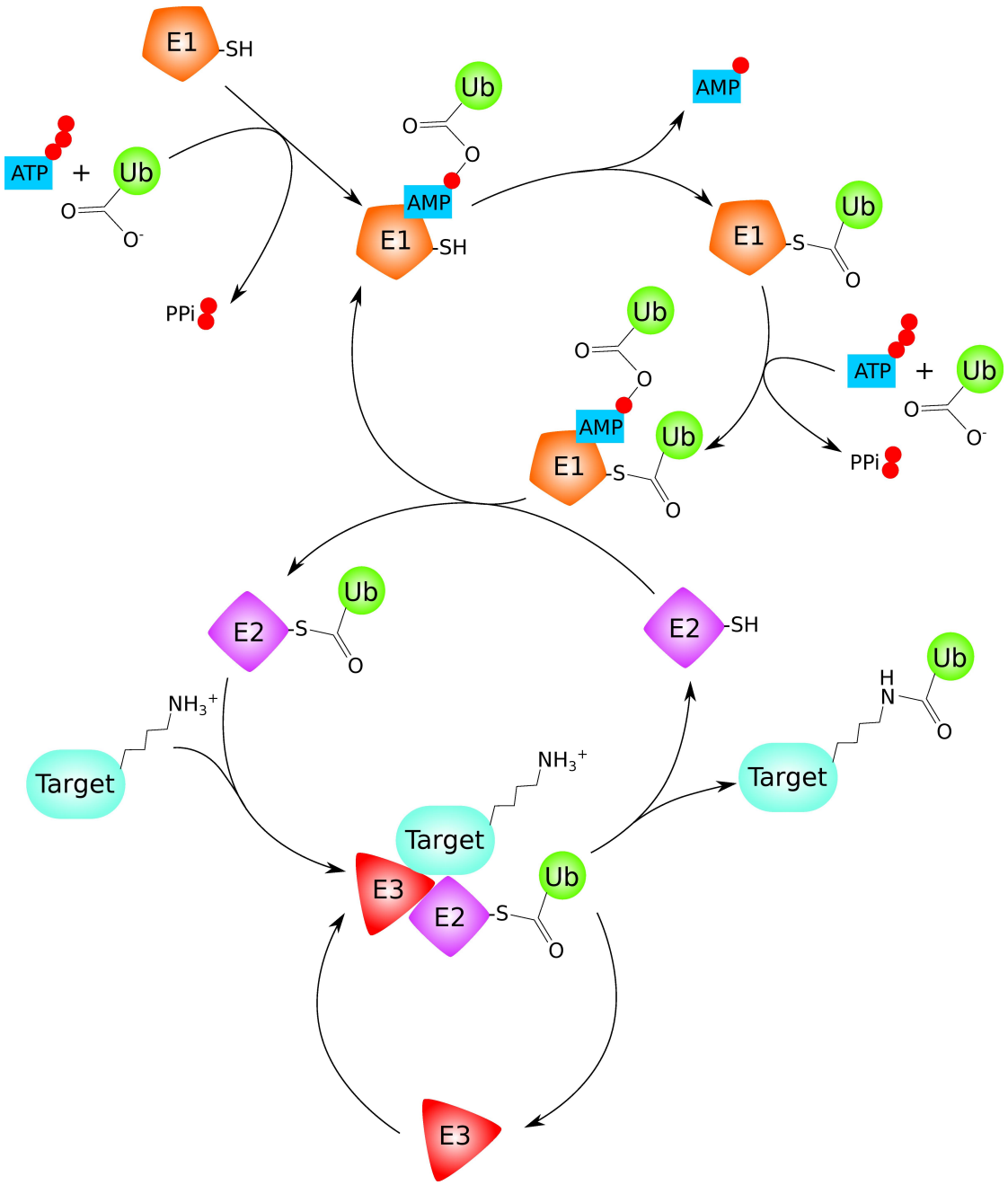


Figure 1: Illustration of the ubiquitination cascade in the case of a RING E3 ligase catalysing mono-ubiquitination of a target protein at a lysine sidechain. Note that in the case of a HECT or RBR E3 ligase an additional intermediate step exists whereby a thioester bond would be formed between a catalytic cysteine on the E3's surface and ubiquitin.

out of the cytoplasm, or trigger downstream signalling, as in the case of the regulation of replication error bypass by PCNA monoubiquitination (Hoege et al., 2002; Li et al., 2003).

The ubiquitination cascade, which catalyses ubiquitination, is composed of three types of enzyme operating in sequence (See Figure 1, page 9). E1 ubiquitin ligases hydrolyse ATP to activate ubiquitin by forming a thioester bond between an E1 cysteine and the C-terminus of the ubiquitin (Haas et al., 1982). E2 ligases catalyse the transfer of the ubiquitin from the E1 ligase to their own catalytic cysteine (Stewart et al., 2016). Finally, E3 ligases catalyse the transfer of the ubiquitin from the E2 catalytic cysteine, either directly or through their own catalytic cysteine, to a residue (usually a lysine) on the target protein or on a pre-existing ubiquitin chain (Pickart and Eddins, 2004).

The number of these enzymes increases the further down the cascade they are: in humans two E1 ligases are known, compared to at least 38 E2 ligases and over 600 E3 ligases (Dou et al., 2013; Jin et al., 2007; Ye and Rape, 2009). The E3 ligases can be split into three groups based on their catalytic domain, these are: HECT (homologous to E6-AP C-terminus) E3 ligases, RBR (RING between RING) E3 ligases, and RING (really interesting new gene) E3 ligases. The latter group includes TRIM proteins and constitutes the vast majority of E3 ligases (Deshaies and Joazeiro, 2009). The increase in number of ligases corresponds to increasing layers of specificity: the two E1 ligases show different ubiquitin charging efficiencies for different E2 ligases (Jin et al., 2007). The different E2 ligases show preferences for different types of bond formation, E3 ligases, or polyubiquitin chain linkages. Examples of bond formation specificity include Ube2W and Ube2J2 which catalyse aminolysis to the α-amino group and hydroxyl attachment respectively as opposed to the more common attachment to lysine sidechains. E3 ligase selectivity is exemplified by Ube2L3 which is active in conjuction with HECT and RBR E3 ligases but not RING E3 ligases. Finally polyubiquitin chain specificity is illustrated by Ube2S, Ube2K and Ube2N–Ube2V1 which catalyse the formation of K11, K48 and K63 polyubiquitin chains respectively (Stewart et al., 2016; Ye and Rape, 2009). As the final enzymes in the cascade, E3s match E2s with the target protein and, in the case of HECT E3s, also determine chain linkages (Wang and Pickart, 2005). The catalytic mechanism depends on the type of E3 ligase: in HECT and RBR E3 ligases catalysis of ubiquitin-target bond formation passes through an intermediate step whereby a catalytic cysteine on the E3 forms a thioester bond with the ubiquitin prior to catalysing attachment to the target. Conversely, in the case of RING domains the ubiquitin goes from being covalently bound to the E2 to being covalently bound to the substrate without any covalent binding to the E3.

High resolution structures of RING E3 ligases bound to both substrate and

an E2 are relatively rare (Brown et al., 2015; McGinty et al., 2014). However, a common theme with these structures is that of multisite interactions involving multiple domains (Deshaies and Joazeiro, 2009). In this context, understanding of RING E3 ligases relies not only on a structural understanding of the RING domain but also of distinct substrate binding domains as well as how these domains interact as a whole. Several projects carried out in our lab, including this one, therefore focus on characterising the substrate binding domains of the TRIM family of E3-ligases and combining that information using an integrative structural biology approach.

1.2 TRIM proteins

1.2.1 The tripartite motif

The TRIM family of proteins are an ancient, highly diverse, group of proteins that constitutes one of the largest families of E3 ligases (Crawford et al., 2018). The family is defined by the presence of the following three domains in the following order: a RING domain, one or two B-box domains and a coiled-coil, hence these three domains are sometimes collectively called the RBCC (Reddy et al., 1992; Reymond et al., 2001).

The RING domain provides E3 ligase activity, it is characterised by the presence of seven cysteines and one histidine that together bind two structural zinc ions in a cross-brace manner, thereby stabilising the domain's $\beta\beta\alpha$-fold (Freemont, 1993; Joazeiro and Weissman, 2000; Meroni and Diez-Roux, 2005). The RING domain may also, in the case of some TRIM proteins but not others (for example TRIM32 and TRIM25 respectively), mediate multimerisation (Koliopoulos et al., 2016). However, even when the RING domain itself does not drive dimerisation or multimerisation there is reason to believe that some higher order assembly is necessary for the function of many RING domains (Streich et al., 2013).

The B-box, which may be present in either one or two copies, is another zinc binding domain (Borden et al., 1993), indeed it is called B-box by analogy to the RING domain which was initially designated as the A-box (Reddy and Etkin, 1991). In cases where two B-boxes are present the first is a so-called type 1 B-box and the second is a type 2 B-box, in the cases where only one is present it is invariably a type 2 B-box. Much like the RING domain it maintains its $\beta\beta\alpha$ structure through a cross-brace coordination of two zincs although these are coordinated by three histidines and five cysteines, one of which can, in the case of B-box type 2, be substituted for an aspartic acid. The high degree of structural similarity between B-boxes and RING domains as well as their frequent association has lead to some authors hypothesising that they may be

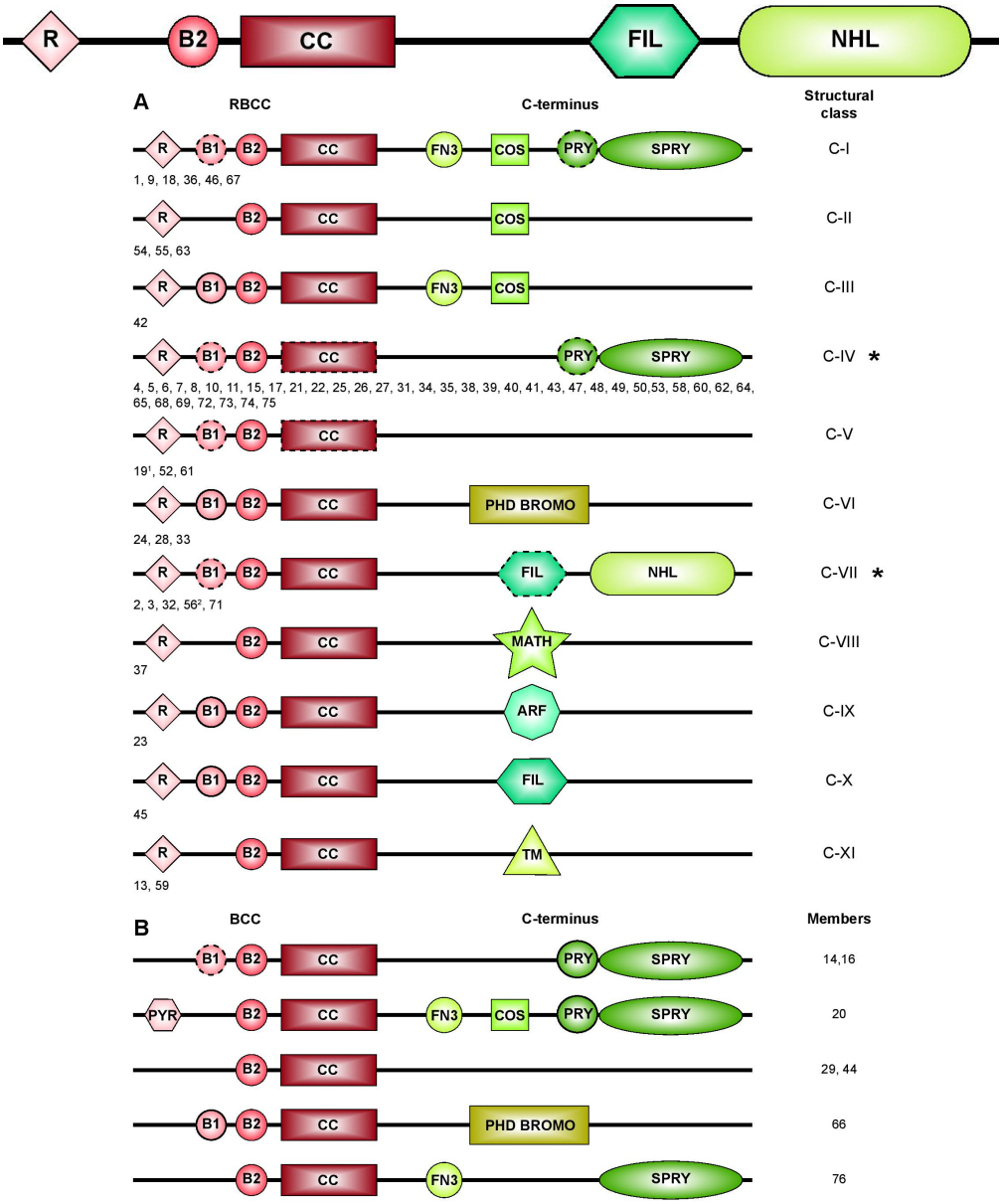


Figure 2: Adapted from Williams et al. (2019), domain arrangement of TRIM2 and TRIM3 (**top**) as well as all other human TRIM proteins (**bottom**). TRIMs are classified as in Ozato et al. (2008), domains with dashed lines are not present in every instance of the class, classes with well established RNA binding roles are marked with an asterisk. Members without a RING domain were not classified by Ozato et al. (2008) but are included here for completion.

evolutionarily related through duplication (Massiah et al., 2007). The B-box's

function is less well understood than that of the RING domain. It may play a role in dimerisation or even higher order association but this role does not appear to apply to all TRIM proteins (Diaz-Griffero et al., 2006; Koliopoulos et al., 2016; Li et al., 2011). It also appears that in some contexts the domain can play a regulatory role by acting as an auto-inhibitor (Dickson et al., 2018).

The coiled-coil mediates homo-dimerisation (Sanchez et al., 2014) and can play a role in recruitment of protein or RNA ligands or provide a binding surface for domains, ensuring their correct positioning relative to one another for ubiquitination (Haubrich et al., 2020; Koliopoulos et al., 2018; Stoll et al., 2019). Some authors claim the domain also mediates hetero-dimerisation but the reported results could merely reflect higher order hetero-multimerisation (Hatakeyama, 2017; Reymond et al., 2001). It also appears to play a role in ensuring TRIM proteins are trafficked to the correct compartment (Reymond et al., 2001).

Finally the C-terminal domains are variable with 11 main domain arrangements recognised (including some variability within each group, see Figure 2 on page 12), underlining the versatility of this arrangement (Ozato et al., 2008).

The TRIM protein family is ancient, it is present in at least some representatives of all eukaryotic supergroups but has undergone extensive diversification in animals (Marín, 2012) and especially vertebrates which posses a highly diversified set of SPRY domain (spore lysis A and ryanodine receptors) containing TRIM proteins (Sardiello et al., 2008).

1.2.2 A family with diverse biological functions

TRIM proteins are involved in a variety of roles including cellular differentiation, autophagy, apoptosis, DNA repair, tumour suppression and innate immunity (van Gent et al., 2018; Hatakeyama, 2017). This broad range of roles leads to involvement in diverse pathologies including schizophrenia, axonal neuropathy, hepatocellular and squamous cell carcinomas as well as sarcotubular myopathy and limb girdle muscular distrophy (2H type) (Tocchini and Ciosk, 2015). The importance of the TRIM family in cancer was recognised early. When the family was defined it already contained the transforming oncogenes PML (TRIM19) and RET (TRIM27) as well as an oncogenic fusion protein of TRIM24 and the kinase B-raf known as T18 (Reddy et al., 1992; Zhong et al., 1999).

Many TRIM proteins play a central function in innate immunity with a variety of roles and behaviours being evident (van Gent et al., 2018; Rajsbaum et al., 2014). One of the most remarkable being the case of TRIM5α which recognises the HIV capsid and forms a hexagonal lattice on its surface (Li et al., 2016). Other innate immunity functions include TRIM21 acting as a receptor of intracellular antibody-opsonized viruses and TRIM56's regulation of the TLR3 signalling

system (Shen et al., 2012; Vaysburd et al., 2013). Their important role in innate immunity is also reflected in their high number in fish, which have particularly developed innate immune systems and where they may be specifically expressed in particularly sensitive areas (Langevin et al., 2019; Rajsbaum et al., 2014).

A large number of TRIM proteins are involved in neuronal development with roles in processes such as axon specification and guidance, control of cell differentiation and adult neuron maintenance (van Beuningen et al., 2015; Marchetti et al., 2014; Menon et al., 2015; Mitschka et al., 2015).

1.2.3 TRIM proteins as E3 ligases

Almost all TRIM proteins possess a catalytic RING domain. In RING E3 ligases, unlike HECT and RBR E3 ligases, the ubiquitin is never covalently bound by the RING domain but addition of the E3 ligase fixes the previously flexible E2-ubiquitin in an orientation optimal for nucleophilic attack of the thioester by a target protein lysine (Buetow and Huang, 2016; Pruneda et al., 2011). The target may, of course, be another protein but auto-ubiquitination is also a functionally relevant role for several TRIM proteins (Choudhury et al., 2017; Diaz-Griffero et al., 2006).

The structural details of ubiquitination activity by the RING domain are best studied in the case of TRIM5α where it is believed that ubiquitination actively involves all three RING domains in a "two plus one" model wherein ubiquitin chain extension is carried out by RING domains acting as a dimer and undergoing partner exchange to extend K63 ubiquitin chains (Fletcher et al., 2018). Kinetic studies suggest that this type of cooperative mechanism may be present more generally across TRIM proteins with the ubiquitination activity for TRIM5α, TRIM25 and TRIM32 exhibiting kinetics matching allosteric cooperativity (Streich et al., 2013). This dependence on higher order assembly would be consistent with the observation that the RING domain's catalytic activity is enhanced by the presence of other members of the RING-B-box-coiled coil motif (to varying degrees depending on the domains present and the TRIM studied), even in E2 ubiquitin discharge assays where a natural target is absent. As well as with the observation that certain RING domains dimerise without involvement from other domains in a manner critical to their catalytic activity (Koliopoulos et al., 2016). Indeed, involvement of at least two domains in catalysis may be a conserved feature of RING domains in general (Dou et al., 2013).

TRIM proteins are able to form a variety of different types of ubiquitination linkages with reported linkages including the relatively common K63 and K48 linkages (Lu et al., 2019; Zhang et al., 2017) but also the more uncommon K6, K27 and K33 (Sparrer et al., 2017; Ye et al., 2017; Zanchetta et al., 2017). Ubiquitination

can take the form of polyubiquitination or monoubiquitination with some TRIM proteins able to switch between the two (Fletcher et al., 2018; Schreiber et al., 2015). This diversity of ubiquitination results in diverse outcomes for the targeted proteins including stabilisation, localisation changes, signalling or degradation (Lu et al., 2019; Peisley et al., 2014; Zanchetta et al., 2017; Zhang et al., 2017). Additionally, the RING domain can act as an E3 in the addition of ubiquitin-like proteins ISG-15 and SUMO to target proteins (Chu and Yang, 2010; Zou and Zhang, 2006). TRIM proteins with these reported functions intriguingly include TRIM32, a protein closely related to TRIM2 and TRIM3, and TRIM25, another known RNA-interacting protein. Both these examples are also able to act as a ubiquitin E3 ligase (Koliopoulos et al., 2016).

1.3 TRIM-NHL proteins

1.3.1 A branch of the TRIM family

The TRIM-NHL subfamily (class VII in the Ozato et al. (2008) classification) is defined by the presence of an NHL domain, a six bladed β-propeller, at the C-terminus of the protein. The NHL domain was first defined by Slack and Ruvkun (1998) based on an alignment of NCL-1 (a *C. elegans* TRIM protein), HT2A (a *H. sapiens* protein now usually refered to as TRIM32) and Lin41 (the *C. elegans* ortholog to the human protein TRIM71). With each NHL domain containing six NHL repeats.

Two exceptions to this rule exist which I choose to nevertheless include in the TRIM-NHL subfamily: first there is the case of TRIM32 where the fourth NHL repeat is replaced with a different sequence which, although known to be compatible with β-strand formation, has an unknown impact on the structure (Slack and Ruvkun, 1998). Recent modelling work by Bawa et al. (2020) based on the orthologous *D. melanogaster* protein Thin, suggests the structure of TRIM32 matches that of other NHL domains closely but shows extended IV c-d and Vd-VIa loops (see Section 1.4.1, page 20). Second, the sequence of TRIM56 does not fit that of canonical NHL repeats and is instead composed of homologous NHL-like repeats (Liu et al., 2016). Beyond homology, another strand of evidence suggesting that these repeats form a single domain are the experiments carried out by Shen et al. (2012) which showed equal disruption to protein activity in mutants in which all of repeat IV and part of repeat V were deleted and those where part of repeat V and all of repeat VI were deleted.

Within the TRIM-NHL family a phylogenetic classification of NHL domains carried out by Kumari et al. (2018) shows a clustering of TRIM2, TRIM3 and TRIM71 NHL domains together. Conversely, TRIM56 and TRIM32 are part of the outgroup along with several non-TRIM NHL containing proteins not reported

Domain	TRIM3	Lin-41	Brat	TRIM56	TRIM32
Whole sequence	66.9%	33.2%	26.2%	22.4%	19.4%
RING	90.5%	42.9%	N/A	50.0%	40.0%
B-box	71.4%	38.1%	29.3%	40.5%	23.8%
Coiled-coil	50.0%	21.8%	21.2%	18.4%	19.2%
Filamin	56.3%	38.8%	N/A	N/A	N/A
NHL	81.5%	37.3%	37.5%	26.7%	23.7%

Table 1: Table showing domain-wise percent identity to TRIM2 for all human TRIM-NHLs as well as the *D.melanogaster* protein Brat. Note that the coiled coil domain boundaries are based on those for TRIM2 where everything between the B-box and the filamin domain was included.

to bind RNA such as the unknown function protein, NHLRC-2, the peptidyl lyase, PAL, and the bacterial receptor protein, PknD (Biterova et al., 2018; Chufán et al., 2009; Good et al., 2004). This is somewhat unsurprising given the non-standard NHL repeats present in both TRIM56 and TRIM32. On the other hand, a phylogenetic classification of human TRIM proteins based on the tripartite motif sequence alone shows that TRIM2, TRIM3, TRIM56 and TRIM71 all cluster together, along with TRIM45, which has a C-terminal Filamin domain but no NHL, while TRIM32 clusters separately with PML/TRIM19, a non-NHL protein whose isoform dependent C-terminal domain remains poorly understood (Li et al., 2017; Williams et al., 2019). This suggests that, with the possible exception of TRIM32, these proteins are related to each other through a process of whole gene duplication and divergence as opposed to single domain duplication. This may have implications as to the evolution of the roles of these proteins as we might expect this process would be more compatible with a slow process of role specialisation and diversification rather than the abrupt acquisition of radically new functions.

1.3.2 The known roles of TRIM2 and TRIM3

TRIM2 and TRIM3 are two highly homologous proteins (67% identity throughout their sequence, see Table 1, page 16) in the TRIM-NHL family. Their domain arrangement is as follows: an N-terminal RING domain, a single B-box, a coiled coil, a filamin domain and finally a C-terminal NHL domain (see Figure 2, page 12). Similarly to their better studied *Drosophila* orthologue Brat (brain tumour), their expression is mainly neuronal.

Although Brat's domain arrangement differs through the absence of the RING and filamin domains and the presence of two rather than one B-boxes, their roles are, similarly to Brat and other TRIM-NHL proteins (Marchetti et al.,

2014; Mukherjee et al., 2016; Reichardt et al., 2018), related to cell differentiation and survival, neuronal development and adult cell maintenance and regulation of oncogenesis, along with an emerging role in innate immunity.

TRIM2 has been shown to play a crucial role in the differentiation of neuronal precursor cells and the development of brain, eye and spinal cord tissue in the development of *Xenopus laevis* embryos. Its down-regulation leads to increased apoptosis and sensory and motor function loss (Lokapally et al., 2020).

TRIM3 is also thought to be critical to brain development as, like Brat, it suppresses c-Myc expression and acts on Notch signalling to promote switching from symmetric to asymmetric division and promote differentiation (Chen et al., 2014; Mukherjee et al., 2016). It is worth noting that in Brat these roles have been linked to RNA binding (Betschinger et al., 2006; Laver et al., 2015; Lee et al., 2006; Shi et al., 2013). Likewise, TRIM2 regulates neuron polarization by ubiquitinating neurofilamin light chain (NF-L) leading to tubulin (which is bound by NF-L) being made available for axon formation (Khazaei et al., 2011). This is consistent with the finding that TRIM2 knockout mice exhibit a progressive neurodegeneration phenotype, attributed to an accumulation of NF-L in axons (Balastik et al., 2008). In line with the ataxia phenotype observed in mice, two TRIM2 mutations have been linked to Charcot-Marie-Tooth disease in humans (Pehlivan et al., 2015; Ylikallio et al., 2013). Interestingly, the mutation found by Pehlivan et al. (2015) is located in the NHL domain, this was interpreted by the authors as potentially destabilising the entire protein.

The hypothesis that ataxia is caused by NF-L accumulation has, however, been called into question by the recent work of Li et al. (2020), who produced a strain of mice with a defective RING domain. These mice did not display an ataxia phenotype despite their inability to directly ubiquitinate NF-L. Additionally, a comparison of brain lysates from wild type mice, mice lacking a TRIM2-RING domain, and mice expressing only a TRIM2-RING domain showed no difference in NF-L levels, suggesting the pathological focal accumulation of NF-L observed by Balastik et al. (2008) is not linked to an overall increase in NF-L levels. The exact mechanism of the neuropathy observed in TRIM2 deficient mice and patients is therefore an open question.

TRIM2 also has a role in ischemic tolerance by ubiquitinating and down-regulating the pro-apoptotic protein Bim (Thompson et al., 2011). In accordance with this role, a mutation inactivating the TRIM2-RING domain was linked to patients suffering severe responses to minor or trivial head-injuries (Ibrahim et al., 2020). This anti-apoptotic interaction with Bim can also have a tumourigenic effect (Chen et al., 2015).

TRIM2 also limits the entry of New World adenoviruses through its interactions with the protein SIRPA (signal-regulatory protein alpha) which is a known

regulator of phagocytosis (Sarute et al., 2019).

Both TRIM2 and TRIM3 are believed to be involved in long-term potentiation regulation (Cheung et al., 2010; Ohkawa et al., 2001). TRIM3 also regulates memory consolidation, hippocampal synaptic plasticity, and spine density through its poly-ubiquitination of γ-actin (Hung et al., 2010; Schreiber et al., 2015). Similarly, TRIM3 has been linked to addiction pathways as it is reported to K48 polyubiquitinate the chromatin remodeller INO80, in the nucleus accumbens, targeting it for degradation. In rats trained to self administer cocaine which had then been deprived of the drug for 30 days, TRIM3 expression was found to be decreased, leading to increased INO80 levels associated with more self administration of the drug upon its return (Werner et al., 2019).

TRIM3 has also been found to inhibit autophagy through K48 linked polyubiquitination of Beclin1, a key regulator of autophagy initiation (Lu et al., 2019). Both proteins have also been reported to bind Myosin V although they do not appear to ubquitinate it (Balastik et al., 2008; Ohkawa et al., 2001; Schreiber et al., 2015).

Finally, these diverse regulatory roles have resulted in multiple examples of involvement in both tumour suppression and upregulation for each protein (Chen et al., 2014; Lin et al., 2020; Lu et al., 2019; Rao et al., 2020). TRIM2 and TRIM3 have also been linked to Alzheimer's disease, Huntingdon's disease, schizophrenia and non-acquired focal epilepsy (Collaborative, 2019; Jin et al., 2012; Schonrock et al., 2012; Martins-de Souza et al., 2009).

In conclusion, TRIM2 and TRIM3 have been associated with a variety of functions, however precise information of the biological pathways involved is often sparse and structural information completely absent. As such, a more detailed characterisation of these proteins has the potential to advance understanding in multiple research areas.

1.3.3 The biological role of RNA binding among TRIM-NHLs

The role of RNA binding amongst TRIM-NHL proteins emerged over the course of several discoveries: early research indicated that TRIM32 and Brat both interacted with RNA interacting proteins (Fridell et al., 1995; Sonoda and Wharton, 2001) or had RNA related roles without any identified interactors (Frank and Roth, 1998; Slack et al., 2000) but direct RNA binding of the TRIM-NHL proteins remained hypothetical (Slack et al., 2000). It was also progressively established that the NHL domain played a crucial role in these interactions (Arama et al., 2000; Loedige et al., 2013; Sonoda and Wharton, 2001). In fact, an experiment by Loedige et al. (2013) demonstrated that, at least within the TRIM-NHL family, the NHL domain is the primary determinant of RNA binding as swapping the NHL domains of TRIM32 and TRIM71 swapped the mRNA targets of the proteins.

Upon the development of mRNA interactome capture methodology TRIM71 was found to be a direct RNA binder (Kwon et al., 2013). This was soon followed by the discovery that Brat was also a direct RNA binding protein, with it being hypothesised that most, if not all, TRIM-NHL proteins shared this function based on the conserved charge of the TRIM-NHL top surface (Loedige et al., 2014).

New discoveries place a new emphasis on RNA binding as a common and crucial role of TRIM proteins in general (Williams et al., 2019). Notable functions include: the regulation of micro-RNA (miRNA) by TRIM25, TRIM71 and NHL-2 (Choudhury et al., 2014; Hammell et al., 2009; Rybak et al., 2009), the regulation of innate immune response by TRIM25 (Choudhury et al., 2017), as well as translational regulation in the context of cell differentiation by TRIM71 and Brat (Harris et al., 2011; Worringer et al., 2014). Moreover, Brat also has a well studied role in *Drosophila* embryogenesis, during which it suppresses translation of *hunchback* mRNA (Arvola et al., 2017). It has traditionally been interpreted as acting in concert with Pum and Nanos to help establish an anterior-posterior gradient of Hunchback protein(Loedige et al., 2015). However, recent results suggest that Brat binds RNA independently of Pum and Nanos, leading to the hypothesis that its main role is instead to direct clearance of maternal *hunchback* mRNA during the maternal to zygotic transition (Laver et al., 2015; Macošek et al., 2020).

1.3.4 The link between RNA binding and ubiquitination

Several strands of evidence lead us to believe that the ubiquitination and RNA binding functions are linked in many RNA binding TRIMs.

It is believed that TRIM25's ubiquitination activity is regulated by RNA as RNase I treatment decreases TRIM25's ubiquitination of ZAP (Choudhury et al., 2017). A recently proposed model for how this can occur suggests that RNA binds to both the PRY-SPRY domain and the coiled-coil bringing PRY-SPRY bound substrate into close proximity to the RING domain, allowing ubiquitination (Haubrich et al., 2020).

The ability of RNA binding to modulate ubiquitination activity would not be unprecedented. For example, the Roquins, two highly homologous RING E3 ligases which regulate immunological tolerance (Heissmeyer and Vogel, 2013), have been reported to display complex cross-talk between RNA binding and ubiquitination, with ubiquitination activity increasing when RNA is bound for certain E2s but decreasing for others (Zhang et al., 2015). Likewise, recruitment of E3 ligases to their ubiquitination targets via RNA binding has a precedent in the case of listerin, a member of the ribosome-associated protein quality control process. Listerin is recruited to the prematurely terminated polypeptide chains produced by stalled ribosomes by binding to the 60S ribosomal subunit (Lyumkis

et al., 2014). It then polyubiquitinates these polypeptide chains marking them for degradation (Bengtson and Joazeiro, 2010; Joazeiro, 2019).

Moreover, TRIMs such TRIM71 often segregate to P-bodies where they both regulate proteins through ubiquitination, such as ALG-1(Zou et al., 2013) and bind and regulate locally present miRNAs such as pri-miRNA29a (Treiber et al., 2017). TRIM71, Mei-P26 and NHL-2 all form part of the miRISC complex where they regulate ALG-1 and AGO1 with downstream effects on mRNAs *let-7* and *lsy-6* (Davis et al., 2018; Hammell et al., 2009; Neumuller et al., 2008). The high specificity of the regulation to a limited number of miRNAs strongly suggests that direct TRIM-miRNA interaction may be involved, possibly in tandem with ubiquitination activity (Williams et al., 2019).

This is reminiscent of the function of the protein MEX-3C which binds the 3'UTR of its target mRNA which it then transports to the cytosol. There, it promotes mRNA degradation by Argonaute proteins using a mechanism dependent on ubiquitination activity (Cano et al., 2012).

In another example it was demonstrated that TRIM56's function as a restrictor of the Zika virus depends on the presence of both its RNA binding C-terminal domain and its RING domain (Yang et al., 2019), suggesting this is yet another case in which RNA binding and ubiquitination are linked.

Reymond et al. (2001) argued that, allowing for some domain deletions, the rigid conservation of the order of the domains in the TRIM protein family throughout evolution suggests that the arrangement is part of "an integrated functional structure, rather than a collection of separate modules". If this is indeed the case we should expect that RNA binding and ubiquitination are directly coupled in at least some cases and do not constitute two independent functions of TRIM2 and TRIM3.

1.4 The NHL and Filamin domains

1.4.1 The structure of NHL domains

The NHL domain is defined by the presence of NHL repeats (usually 6) as defined by Slack and Ruvkun (1998), these fold into a six bladed β-propeller where every blade is made up of four β strands. Each NHL repeat contributes the outer β-strand (strand d) of one blade and then the three inner β-strands of the next blade (Edwards et al., 2003), as illustrated in Figure 3, page 21.

This roughly cylindrical shape produces a large surface at each base where interactions with other macromolecules may occur, the top surface being defined as the surface on which the b to c loops are present (Edwards et al., 2003). In the case of protein-protein interactions, binding may occur at the top or bottom surface (Cho et al., 2006; Lee et al., 2006) but so far all well characterised protein-

RNA interactions occur at the top surface (Kumari et al., 2018; Loedige et al., 2015)

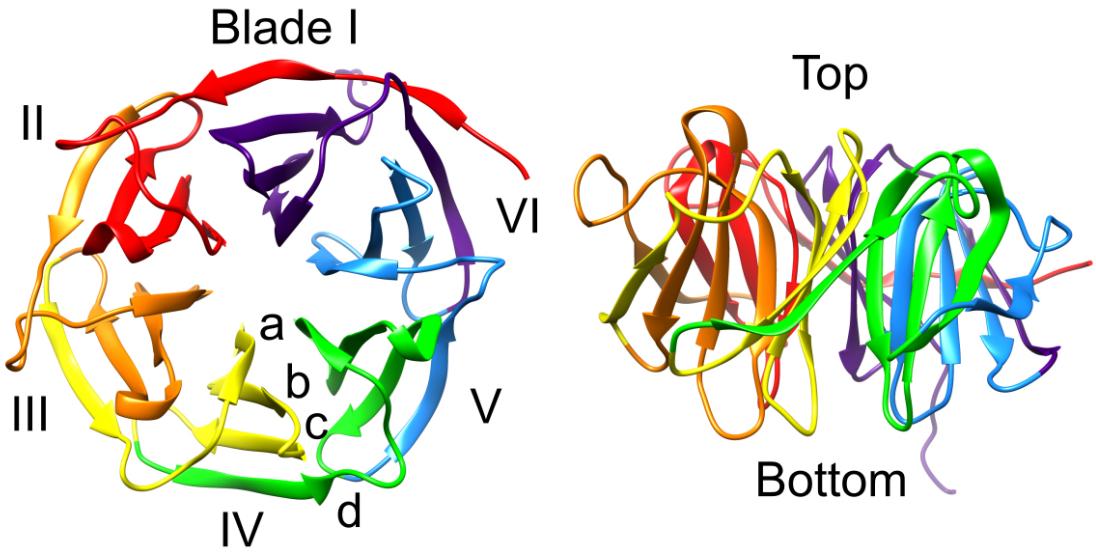


Figure 3: Adapted from Williams et al. (2019); structure of the Brat NHL domain as described in Edwards et al. (2003), blades and β-strands with conventional numbering and NHL repeats coloured in sequence order from red to purple, left image shows the NHL domain viewed from the top while that on the right shows it viewed from the side.

The NHL domain itself appears more widespread than TRIM proteins (although it was defined based on these), with examples also occurring in prokaryotes (Good et al., 2004). These non-TRIM NHL containing proteins have diverse roles such as forming part of transmembrane Serine/Threonine protein kinase receptors, acting as catalytic domains in the α-amidation of the carboxy terminus of neuropeptides and peptide hormones, or having unknown function (Biterova et al., 2018; Chufán et al., 2009; Good et al., 2004). There does not appear to have been any attempt to review the evolutionary or functional relationships of these various non-TRIM NHL domains in the literature to date, however these diverse functions underline the versatility of NHL domains

The "modular" design of the NHL domain is of some interest to the field of synthetic biology and protein evolution. Voet et al. (2014) demonstrated that an NHL domain may be assembled from three identical double NHL repeat peptides or two identical triple NHL repeat peptides. They then hypothesised that the evolutionary origin of the domain may reside in the duplication of smaller modules such as these. Another protein domain with a similar modular design is the WD40 domain, which also forms β-propellers composed of 4-stranded β-sheets with each repeat contributing three strands to one blade and

one to the preceding as is the case for NHL domains. Contrary to NHL domains, WD40 domains more often form seven bladed β-propellers but also have roles in RNA binding as discussed below (Stirnimann et al., 2010).

1.4.2 The NHL domain as an RNA binder

So far two RNA binding NHL domains, those of Brat and TRIM71 have been crystallised with RNA bound (Kumari et al., 2018; Loedige et al., 2015), while the binding motif of Brat paralogue Mei-P26 and its *C. elegans* orthologue NHL-2 have also been characterised.

The homologous proteins Brat, Mei-P26 and NHL-2 have similar binding motifs with a preference for UU[G/A]UU[G/A], UUUACA and poly-uridine respectively (Davis et al., 2018; Loedige et al., 2015). The binding of RNA by Brat was characterised at atomic resolution by Loedige et al. (2015), the RNA was found to lie flat along a positively charge band on the top surface of the NHL domain with different bases fitting into specific grooves.

Initial efforts to characterise the binding motif for TRIM71 faced more difficulties, however, as the only pattern that could be extracted from the RNA-compete experiments carried out by Loedige et al. (2015) was the presence of a single highly conserved adenine. Subsequent analysis by Kumari et al. (2018) resolved this problem by determining that the binding of TRIM71 was structure rather than sequence dependent: TRIM71 preferentially binds stem loops having a three nucleotide loop whose third base is a purine and where the final base pair of the stem is U-A. This binding occurs via an end-on interaction of the loop with a positively charged cavity located in the middle of the top surface.

It thus appears that, while most well characterised NHL domains from the TRIM-NHL family display some level of affinity for RNA, with the expectation that nearly all of them are RNA binders (Loedige et al., 2014), this RNA binding role can be achieved through various RNA binding modes. It should however be noted that the analysis carried out by Kumari et al. (2018) did not detect any preference for structured RNA in NHL domains not homologous to TRIM71, suggesting that it may be unique when it comes to its binding mode.

The role of NHL domains in binding RNA appears to be part of a wider trend of RNA binding by β-propeller domains as these domains on the whole display extreme versatility in their roles and interactions (Stirnimann et al., 2010) with WD40 β-propeller domains being enriched in mRNA interactome studies (Kwon et al., 2013). A prominent example is Gemin5 which binds RNA as part of the SMN (survival of motor neurons) complex to produce the snRNPs that later form part of the spliceosome Lau et al. (2009).

1.4.3 The filamin domain

Between the coiled coil and the NHL domain of TRIM2 and TRIM3 lies the filamin domain (so called because of its filamin-like immunoglobulin fold), which is also present at this position in TRIM71 and forms the only C-terminal domain of TRIM45. It is composed of seven β-strands that form two antiparallel β-sheets (Tocchini et al., 2014).

Its role is poorly understood, early work by Loedige et al. (2013) suggest it may be involved in the translational repression role of mammalian TRIM71, while Tocchini et al. (2014) reported that a 16 amino acid insertion likely to disrupt the fold of the filamin domain caused defects in the somatic, but not germline, functions of *C. elegans* TRIM71. More recently it was reported that phosphorylation on position S443 on the TRIM2 filamin domain allows AXL binding, ubiquitination and down-regulation by TRIM2 (Rao et al., 2020). The filamin domain also appears to be necessary and sufficient for TRIM2's role in internalising New World adenoviruses suggesting it may be the domain that interacts with the phagocytosis regulating protein SIRPA (Sarute et al., 2019). Its association with NHL domains appears to be ancient as it is present in the mollusc orthologue to TRIM2/3, L-TRIM, and in the *C. elegans* TRIM71 orthologue Lin-41, despite being absent in Brat (van Diepen et al., 2005; Tocchini et al., 2014). This suggests that the association may have been present in the common ancestor of Chordata, Ecdysozoa (containing arthropods and nematodes) and Lophotrochozoa (containing molluscs), although the exact phylogenetic position of such an ancestor is unclear as the phylogeny of early bilateria remains an area of active study (Philippe et al., 2019).

In conclusion, TRIM2 and TRIM3 are two proteins which are part of an extremely versatile and ancient family of E3 ligases. They, like many of their homologues, are responsible for a diverse set of regulatory roles with impacts on a variety of pathways of scientific or medicinal interest. However, their roles remain poorly understood and almost all research that does exist focuses on their ubiquitination activities despite it being expected that they perform RNA binding activities through their C-terminal domains. For this reason this thesis will focus on exploring the ligands and structures of these C-terminal domains.

2 Aims of Thesis

Many of TRIM2 and TRIM3's reported roles are expected to involve RNA binding, given the ubiquity of such interactions in other TRIM-NHL proteins and involvement of RNA binding in the same roles when carried out by orthologues. Moreover, RNA binding could both modulate the ubiquitination activity of these proteins as well as recruit ubiquitination targets making it crucial to our overall understanding of these proteins as a whole. This thesis therefore aims to characterise the RNAs bound by TRIM2 and TRIM3. Given the expectation that such RNA binding would be dependent on the proteins' C-terminal domains, this also entails characterising these domains as well as their interactions between each other.

The information sought during this project includes the structures of these C-terminal domains, their consensus RNA binding motif and the strength of protein-RNA interactions as well as their location with residue and base level resolution. In order to achieve this I studied domain-RNA interactions *in vitro* using nuclear magnetic resonance spectroscopy (NMR), characterised the structure and dynamics of the domains using a combination of NMR and X-ray crystallography and investigated inter-domain interactions using small angle X-ray scattering (SAXS). I then compared the results of our investigation to previously characterised TRIM-NHL proteins as well as compared the domains for TRIM2 and TRIM3 to see how the C-terminal domains of these proteins might differ from each other and closely related homologues.

In short I aimed to:

- Solve the structure of C-terminal domains
- Find the RNA binding motifs
- Characterise the RNA binding sites
- Compare these to previously known TRIM-NHL C-terminal domains
- Characterise the interdomain interactions

3 Structural biology methodologies

As the main method employed in this thesis is nuclear magnetic resonance spectroscopy (NMR), including certain more specialised applications with which the reader may not have experience, I have written a short introduction to the method as well as the applications used over the course of this project. I will also provide a short introduction to the two other structural biology methodologies employed in this project, X-ray crystallography and small angle X-ray scattering (SAXS), with a focus on the particular applications I used. If the reader is looking for a more in depth introduction to the topics of NMR or SAXS I would recommend Cavanagh (2007) and Keeler (2005), and Schnablegger and Singh (2011), Svergun and Koch (2003), and Putnam et al. (2007), respectively.

3.1 Spin, shielding and the basics of the NMR experiment

NMR is based on the measurement of nuclear spin, a quantum number associated, in our case, with the nucleus of an atom. When studying biological macromolecules we usually focus on the nuclei ^{1}H, ^{15}N and ^{13}C. These are all characterised by having spin 1/2. This means that, when measured with respect to a given axis (say for example an external applied magnetic field), they may orient themselves either "up" or "down" corresponding to a secondary spin number m=+1/2 or m=-1/2. When placed in an external magnetic field certain orientations will become slightly more energetically favoured in accordance with the equation:

$$E_m = -\hbar\gamma m B_0$$

Where $\hbar$ is Plank's constant over 2π, B_0 is the strength of the applied magnetic field in Tesla and γ is a constant that depends on the nucleus studied and is approximately $26.8{\cdot}10^7$rad$\cdot$ s$^{-1}{\cdot}$T^{-1} for ^{1}H, $6.7{\cdot}10^7$rad$\cdot$s$^{-1}{\cdot}$T^{-1} for ^{13}C and $-2.7{\cdot}10^7$rad$\cdot$s$^{-1}{\cdot}$T^{-1} for ^{15}N (Harris et al., 2001). This leads to a slight excess of one population over the other in accordance with the Boltzmann distribution and a measurable macroscopic magnetic moment.

When taken in bulk these spins behave in a manner analogous to angular momentum and can in some cases be treated "classically", although it should be emphasised that this is due to the averaging of a great many spins, with any individual spin behaving in a quantum manner (Levitt, 1988). Notably, the vector associated with the imbalance of spins will precess around the axis of the applied magnetic field, conventionally denoted as the z-axis, with a characteristic frequency, called the Larmor frequency, equal to $\gamma B_0/2\pi$.

The core steps of an NMR experiment are:

- the creation of a magnetic moment along the z-axis, achieved through the application of an external magnetic field,
- a displacement of the magnetic moment from the z-axis using a precisely timed series of one or more radio frequency pulses (including the possibility of magnetisation being transferred from one set of nuclei to another),
- and the observation of the precession of the magnetic moment around z using a coil. The resulting signal is called the free induction decay, or FID, and consists of overlapped sinusoidal signals (with some loss of signal over the course of the acquisition) from which frequencies may be extracted by Fourier transformation.

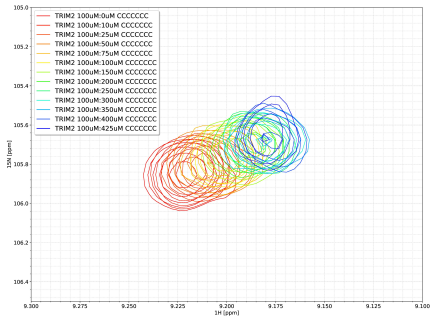
The reader may have noticed at this point that the methodology I have described only allows us, for the moment, to distinguish between different nuclear compositions but not between identical nuclei present in a different chemical context. We can, in fact, distinguish these thanks to an effect called shielding: when a molecule is placed within a magnetic field the electrons will produce a small induced magnetic field which "shields" the nuclei from experiencing the full effect of the external magnetic field. The magnitude of this effect depends on the local electronic environment of the nucleus, that is to say its chemical environment, the resulting shift in frequency is therefore called chemical shift. Chemical shift is usually denoted δ and expressed in a parts per million (ppm) scale determined with reference to a given chemical standard. This scale has the advantage of, first of all, making the units more convenient to work with and, most importantly, of giving measurements independent of the field strength of the magnet used.

As every nucleus in a given macromolecule exists in a slightly different chemical environment it will give a slightly different chemical shift. Assuming we can assign a chemical shift to every one of the nuclei of interest (usually the ^{1}H and ^{15}N in the protein backbone) we can track the chemical environment of each of these nuclei.

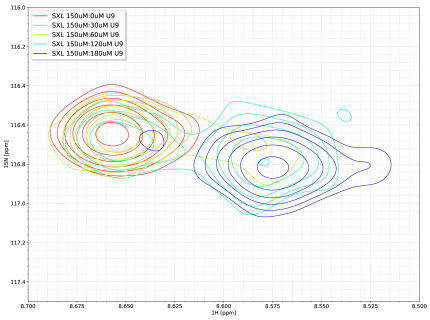
3.2 NMR titrations

NMR is highly sensitive to the chemical environment of the nuclei composing the molecule of interest. Binding of a ligand to our macromolecule of interest will modify the local electronic environment of nuclei near the binding site, changing the degree of shielding experienced by these nuclei and therefore their chemical shift. This change in the chemical shift (δ) is called a chemical shift perturbation (CSP) and can be denoted $\Delta\delta$ (Williamson, 2013).

For practical purposes three different binding cases are often distinguished based on the frequency with which the ligand binds and dissociates from the macromolecule k_{ex} (for exchange), relative to the difference in frequency of the signal given by the bound and unbound state (which we will call $\Delta_\omega = \omega_{bound} - \omega_{unbound}$). k_{ex} is dependent on the strength of the interaction (approximate K_d stated here assume that the ligand binds at diffusion controlled rates and therefore that k_{ex} depends entirely on k_{off}). This gives rise to three exchange regimes: fast exchange in which $\Delta_\omega << k_{ex}$, which is usually associated with weak binding (roughly $K_d > 10\mu$M), slow exchange in which $\Delta_\omega >> k_{ex}$, which is usually associated with strong binding (roughly $K_d < 1\mu$M) and intermediate exchange in which $\Delta_\omega \approx k_{ex}$ which is associated with intermediate binding affinities (Teilum et al., 2017).



(a) A titration of C-7 RNA against TRIM2-NHL in the fast exchange regime.



(b) A titration of U-9 RNA against SXL in the slow exchange regime (see Hennig et al. (2013)).

Figure 4: Comparison of protein-RNA titrations carried out in the fast and slow exchange regime. Note how in fast exchange the position of the peaks shifts with the addition of RNA while in the slow exchange peaks slowly disappear and then reappear.

In our titrations all binding is in the fast exchange regime. In this case the ligand binds and dissociates from the protein many times over the course of the NMR experiment, this means the observed chemical shift (δ_{obs}) will be an average of the unbound chemical shift (δ_{free}) and the bound chemical shift (δ_{bound}) weighted according to the proportion of bound (LP) and free (P) protein according to the equation:

$$\delta_{obs} = \frac{\delta_{bound}[LP] + \delta_{free}[P]}{[LP] + [P]}$$

From this, assuming a high enough ligand to protein fraction can be reached so that δ_{bound} can be estimated, the bound and free fraction of protein can be

assessed at different ligand concentrations and a K_d determined (see Figure 4a, page 27 for an illustration and Section 4.9, page 44 for the equation used).

In the case where binding is very tight all available ligand will be bound, therefore the protein will not bind or dissociate from the ligand over the course of the NMR experiment, giving rise to two distinct population of bound and unbound protein. We will observe two different chemical shifts δ_{bound} and δ_{free} with intensities proportional to the relative sizes of these populations (see Figure 4b, page 27).

The final case is intermediate exchange, in this case the ligand will bind and dissociate from the protein a small number of times over the course of the NMR experiment. This gives rise to a large number of populations having been bound to the ligand for different fractions of the total experiment time which results in peak broadening and signal loss over the course of the titration (the total area under the curve should remain constant but, in practice, signal often decreases to below noise levels). Signal should, however, recover upon saturation of the protein with ligand as we return to a single uniform population.

3.3 NMR relaxation experiments

If we observe a magnetic moment precessing around the z-axis (so in the x-y plane) we will notice that the signal does not persist indefinitely but decays over time, this is termed relaxation and measuring it can give us insight into the dynamics of our system.

We usually distinguish two main components to relaxation, the longitudinal relaxation (measured by T_1 or its inverse R_1) and the transverse relaxation (measured by T_2 or its inverse R_2). The longitudinal relaxation is caused by our magnetic moment slowly returning to the z-axis and is dominated by interactions of the excited nuclei with surrounding molecules (called the lattice) while transverse relaxation is caused by spins losing coherence with one another over time (we may think of this as the magnetic moment being "smeared out" in the x-y plane over the course of the experiment and therefore decreasing in magnitude) and is dominated by the interaction of excited nuclei with each other. Both of these parameters can be measured using appropriate pulse sequences.

One component of relaxation which we can measure easily is the nuclear Overhauser effect (NOE) due to through space dipole-dipole interactions. If we take two nuclei I and S which are close in space and perturb S this perturbation will transfer from spin S to spin I. If we denoted I_0^I and I^I the integral of the peak associated with spin I in the unperturbed and perturbed state respectively then we can define the heteronuclear NOE enhancement $\eta_{I\{S\}}$ according to the following equation:

$$\eta_{I\{S\}} = \frac{I^I - I_0^I}{I_0^I}$$

The values taken by this equation depend on the tumbling rate of the I-S bond studied and therefore provides us with information as to the dynamics of the protein in that location (Gil and Navarro-Vázquez, 2017).

Data from NMR relaxation parameters can then be fitted to a variety of models to relate quantities such as T_1, T_2 and NOE enhancement to internal and external molecular motions. These are related by known equations to the spectral density function J(ω) which is itself equal to the Fourier transform of the correlation function $C(t)$ for a given nucleus. These equations are (Fushman et al., 2001):

$$R_1 = 3(d^2 + c^2)J(\omega_N) + d^2\left[3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)\right]$$

$$R_2 = \frac{1}{2}(d^2 + c^2)\left[4J(0) + 3J(\omega_N)\right] + \frac{1}{2}d^2\left[J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + R_{ex}$$

$$NOE = 1 + \frac{\frac{\gamma_H}{\gamma_N}d^2\left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right]}{R_1}$$

With c and d being terms representing contributions to relaxation from chemical shift anisotropy and dipolar coupling and R_{ex} being the chemical exchange contribution to R_2. c and d can be calculated based on default values based on previous measurements and known physical quantities. Assuming one can determine R_{ex}, (which can be done by repeating measurements, R_1, R_2 and NOE at different field strengths or, as was done in our case, by only selecting residues believed to be in a well folded part of the protein not subject to chemical exchange to determine an initial model), one can therefore fit a model of J(ω) with a sufficiently low number of parameters.

One of the most popular ways of doing this is the "model free" (so called because initial assumptions are not explicitly made about the molecular motions studied but rather about the correlation function) approach put forward by Lipari and Szabo (1982a). In this model, overall macromolecular tumbling and internal motions are uncorrelated, the correlation function is therefore a product of an overall correlation function $C_O(t)$ and an internal correlation function $C_I(t)$ (i.e. $C(t) = C_O(t)C_I(t)$). The $C_I(t)$ function is a single-exponential approximation of the form:

$$C_I(t) = S^2 + (1 - S^2)e^{-\frac{t}{\tau_e}}$$

Where S^2 is an order parameter taking a value of 0 when the internal motion is completely isotropic and a value of 1 when the internal motion is completely

restricted. S^2 is therefore related to whether a residue is part of a folded part of a protein (as unfolded regions will have very low S^2 values) and to what degree it undergoes substantial internal motion, but holds no information as to the speed of such motions. τ_e is an effective correlation time which contains information on both the rate of motion and the spatial nature of this motion, making it harder to interpret directly (Lipari and Szabo, 1982a).

If we assume the overall motion of the macromolecule is isotropic, we get $C_O(t) = e^{-\frac{t}{\tau_R}}$, where τ_R is the correlation time of the macromolecule as a whole. This means J(ω), the Fourier transform of C(t) will fit the equation (Korzhnev et al., 2001; Lipari and Szabo, 1982a):

$$J(\omega) = \frac{S^2\tau_R}{1+(\omega\tau_R)^2} + \frac{(1-S^2)\tau_e^{'}}{1+(\omega\tau_e^{'})^2}$$

Where S^2 and τ_R are as defined above and $1/\tau_e^{'}=1/\tau_R+1/\tau_e$.

This equation was extended by Clore et al. (1990b) who noticed that it was insufficient to explain ^{15}N relaxation data in the case of the backbone amide groups of certain proteins. This is because the previous model makes two assumptions: first that the internal motions are much faster than the overall tumbling ($\tau_e << \tau_R$), and second that the internal motions are fast relative to the frequency of the nucleus studied ($\omega\tau_e << 1$), which is not necessarily true in the case of protein motions of larger amplitude (in the case of ^{15}N, at the field strengths usually used this requires $\tau_e < 50-100ps$) They therefore slightly extended the model put forward by Lipari and Szabo (1982a) by adding a second exponential component to $C_I(t)$ to give:

$$C_I(t) = S^2 + (1-S_f^2)e^{-\frac{t}{\tau_f}} + S_f^2(1-S_s^2)e^{-\frac{t}{\tau_s}}$$

Where $S^2 = S_f^2 S_s^2$. With S_f^2 and S_s^2 being order parameters analogous to S^2 and relating to motions occurring on fast and slow, respectively, time-scales τ_f and τ_s different by at least one order of magnitude. This gives a spectral density function J(ω) as follows:

$$J(\omega) = \frac{S^2\tau_R}{1+(\omega\tau_R)^2} + \frac{(1-S_f^2)\tau_f^{'}}{1+(\omega\tau_f^{'})^2} + \frac{S_f^2(1-S_s^2)\tau_s^{'}}{1+(\omega\tau_s^{'})^2}$$

Where S^2, S_f^2, S_s^2 and τ_R are defined as above while $1/\tau_f^{'}=1/\tau_R+1/\tau_f$ and $1/\tau_s^{'}=1/\tau_R+1/\tau_s$. Note that when only fast or slow motions are present (i.e. $S_f^2 = S^2$ or $S_f^2 = 1$), this simplifies to the previous model (Clore et al., 1990b).

So far all the cases we have looked at have dealt with the case in which the overall macromolecular motion is isotropic. However, for less globular proteins tumbling is often anisotropic or axially symmetric. The characteristics of this motion can be represented by a rotational diffusion tensor **D**, this is a 3x3 matrix whose trace is $1/2\tau_c$, where τ_c is the overall rotational correlation time (Berlin et al., 2013). This tensor can then be used to calculate a version of Lipari and Szabo with anisotropic tumbling:

$$J(\omega) = \sum_{\eta=-2}^{2} c_\eta \left[\frac{S^2\tau_\eta}{1+(\omega\tau_\eta)^2} + \frac{(1-S^2)\tau_{e,\eta}}{1+(\omega\tau_{e,\eta})^2} \right]$$

Where $1/\tau_{e,\eta}=1/\tau_\eta+1/\tau_e$ and τ_η and c_η can be calculated from eigenfunctions and associated eigenvalues of **D** (see Korzhnev et al. (2001) for full expansions of these values). The two preceding equations can be combined to give the equation our model uses (with the caveat that this describes a fully anisotropic molecule, of which our axially symmetric molecule is a special case):

$$J(\omega) = \sum_{\eta=-2}^{2} c_\eta \left[\frac{S^2\tau_\eta}{1+(\omega\tau_\eta)^2} + \frac{(1-S_f^2)\tau_{f,\eta}}{1+(\omega\tau_{f,\eta})^2} + \frac{S_f^2(1-S_s^2)\tau_{s,\eta}}{1+(\omega\tau_{s,\eta})^2} \right]$$

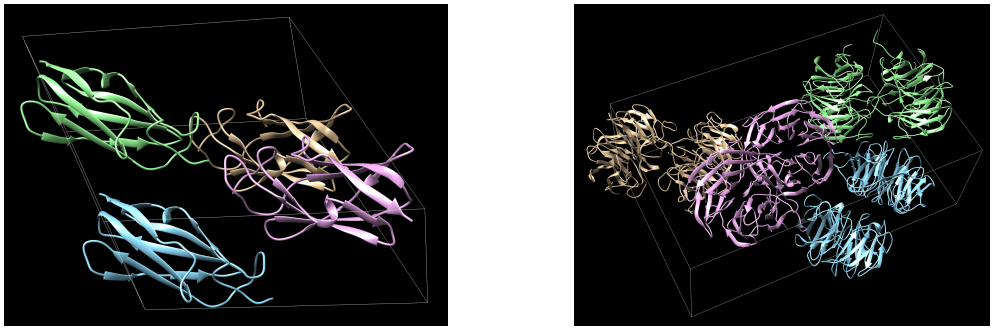
3.4 X-ray crystallography

X-ray crystallography is the most well established of structural biology methods and remains the method through which the vast majority of atomic resolution protein structures are determined to this day.

The first, and often most laborious, step is the growth of a crystal, that is to say a large, well ordered lattice of atoms related by translation relationships. These are obtained by preparing a solution of pure, concentrated protein and making it slowly come out of solution to form the crystals. This is achieved by placing the protein under metastable conditions, that is to say that the protein remaining in solution is thermodynamically unfavoured but not to such an extent to cause immediate and disordered precipitation. Under these conditions protein crystals may spontaneously nucleate and grow, or form around impurities or purposefully introduced nucleants. Crystallisation screens may be set up such that the protein conditions either already are metastable, such as in the batch crystallisation method, or the buffer conditions will change over time such that the protein becomes metastable, as is the case in the vapour diffusion method (Luft et al., 2014).

The crystal is characterised by the unit cell, that is to say the smallest unit from which the whole rest of the crystal may be constructed purely by repeated

translations (the crystal is in fact often modelled as an infinitely repeated unit cell). The unit cell can then be identical to, or be subdivided into multiple copies of, the asymmetric unit. The asymmetric unit is the smallest unit from which the unit cell may be constructed by symmetry relations that are not pure translation (for example rotation). Finally the asymmetric unit can contain a single copy of the protein or multiple copies of the protein in different conformations. An example of this can be seen in Figure 5, page 32, which shows the unit cells of the two crystal structures presented in this thesis: the unit cell for filamin contains four asymmetric units each composed of one copy of the protein, while the NHL unit cell also contains four asymmetric units but each of these is composed of two protein chains.



(a) TRIM3-Filamin (b) TRIM2-NHL

Figure 5: Illustrations of the unit cells the two crystal structures presented in this thesis (see the Results section for more details). One unit cell is shown in each case with limits shown in white with the chains making up each asymmetric unit displayed in a different colour.

Once we have a crystal of sufficient quality we can then acquire X-ray scattering data from it. The ordered nature of the crystal allows for constructive interference to build up from multiple unit cells, we will therefore observe strong signal at certain specific points whereas at all other points the interference will be destructive and no signal will be observed. To determine the location of these points it is useful to picture the scattering as the reflection of an X-ray wave hitting a plane of the crystal with an angle θ and then being reflected by the same angle θ. We can now imagine a second parallel wave passing through this plane but being reflected by an equivalent plane a distance d from the first (the regularities of the crystal allow many such planes). The extra distance travelled by the second wave will be $2d\sin(\theta)$, if this is a multiple of the wavelength (λ) the two waves will be in phase and the interference will be constructive, if not the interference will be destructive. From this we get Bragg's law which determines the conditions under which constructive interference will occur:

$$n\lambda = 2dsin(\theta)$$

If we apply this to a three dimensional crystal with a unit cell defined by vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ we can define the so called Miller indices (h,k,l) for each observed scattering vector $\vec{S}$ (defined by the incident and outgoing vectors such that $\vec{S} = \vec{k_{in}} - \vec{k_{out}}$) such that:

$$\vec{a} * \vec{S} = h$$
$$\vec{b} * \vec{S} = k$$
$$\vec{c} * \vec{S} = l$$

where h, k and l are integers (Putnam et al., 2007).

The crystal is then rotated in the X-ray beam in order to obtain scattering data in all necessary orientations (in crystals with higher symmetry a lower range of angles is generally needed to acquire a full dataset (Dauter, 1999)). The electron density (ρ) at a particular position, specified by coordinates x,y, and z, of a unit cell of volume V is related to scattering by the equation:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h,k,l} |F_{hkl}| e^{i\alpha_{hkl}} e^{-2i\pi(hx+ky+lz)}$$

where α_{hkl} is the phase associated with the structure factor amplitude $|F_{hkl}|$ for Miller indices h, k and l. $|F_{hkl}|$ is proportional to the square root of the intensity measured by the detector but the information encoded in α_{hkl} is lost during our experiment, this gives rise to the phase problem (Taylor, 2003).

The phase problem may be solved in a variety of ways including isomorphous replacement, multiwave anomalous diffraction and single wavelength anomalous diffraction. However, the method we chose to solve the phase problem was molecular replacement, in this methodology a previously solved highly homologous structure is used (often with more variable regions such as loops and sidechains removed) with the assumption that such a structure will closely resemble the structure of our macromolecule of interest. Different translations and rotations of the homologous structure are tested such that it provides the best possible prediction of the observed data and this orientation is then used to provide an initial estimate of the phase information, thereby constructing a plausible electron density map. Once this has been achieved, modelling of our actual macromolecule of interest can begin with gradual improvements of the model in theory translating to improved estimates for the phase information until a satisfactory model has been produced.

3.5 SAXS

Small angle X-ray scattering is in many ways conceptually similar to X-ray crystallography except that the protein molecules are acquired in solution and we therefore acquire data corresponding to molecules whose orientation both relative to each other and the detection apparatus has been averaged. This presents the advantage that the macromolecule or complex need not be crystallised to be studied and is measured under conditions usually more biologically relevant than those prevalent in a crystal lattice. The main disadvantage is that the averaging causes rapid loss of signal with increasing resolution, as high resolution contributions average out limiting the technique to maximum resolutions in the range of 10Å-50Å and therefore only allows us to probe the rough shape of the macromolecule.

When conducting a SAXS experiment we first of all measure scattering (we limit ourselves to elastic scattering throughout this discussion) in the presence of our molecule of interest ($\rho(\mathbf{r})$) and compare it to scattering by the solvent alone (ρs). This gives us the scattering due to the biological macromolecule of interest alone ($\Delta\rho(\mathbf{r})$) according to the equation:

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho s$$

With contrast arising from the difference in electron density between the solvent ($0.33e^-/Å^3$ for pure water) and the macromolecule ($0.44e^-/Å^3$ for protein, $0.55e^-/Å^3$ for DNA), meaning that signal is higher for electron dense macromolecules such as nucleic acids. It is interesting to note that this property can be used to separate contributions from nucleic acids and protein in systems containing both components by using a buffer with an electron density matching that of protein (Tokuda et al., 2016). These "contrast matching" experiments play a major role in the applications of the analogous small angle neutron scattering (SANS) method.

Having corrected for the scattering for the buffer we may also wish to establish a unit to describe our scattering pattern. As the macromolecules are oriented randomly relative to the measurement device our scattering pattern will be centrally symmetric, therefore the only measurement we need to describe the position of a scattering intensity signal on our detector is the angle 2θ from the center of our beam (we use 2θ as a convention stemming from the practice of modelling the scattering as a reflection by the sample with incident and reflected angles of θ giving a total displacement from the incident beam angle of 2θ). The scattering intensity at a given position will depend on the wavelength (λ) of the X-ray beam used. We will therefore define a unit called s (q is also often used) independent of the X-ray wavelength as follows:

$$s = \frac{4\pi sin(\theta)}{\lambda}$$

s (or q) usually given in units of nm^{-1} or $Å^{-1}$. Expressing our intensities this way allows us to express the scattering angle independently of the set-up of any given SAXS beamline allowing us to easily compare SAXS curves from different sources.

Having isolated the contribution of the macromolecule alone and established a unit independent of the X-ray wavelength used we can then start to extract structure information. If our scattering particles are randomly oriented and randomly distributed throughout our mixture, corresponding to the case of dilute non-interacting proteins, their scattering does not interfere with each other and their scattering intensities are summed (Svergun and Koch, 2003). We can therefore consider each particle in isolation. The scattering of a single particle, itself made up of many atoms is the interference pattern produced by the incoming X-rays from every atom inside the particle (Schnablegger and Singh, 2011). This is expressed by the Debye formula:

$$I(s) = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i(s) f_j(s) \frac{sin(sr_{ij})}{sr_{ij}}$$

where I(s) is the scattering intensity at s, N is the total number of atoms within the particle, f_i and f_j are the scattering factors of atoms i and j respectively and r_{ij} is the distance between atoms i and j. It is worth noting that the expression $\sin(sr_{ij})/sr_{ij}$ comes from the rotational averaging of the vector $\vec{r}_{ij}$. If one considers the particle to be homogeneous we can express this in terms of the function P(r) which represents the frequency of pairs of points of distance r within the particle (Koch et al., 2003). The scattering is then related to P(r) by the formula:

$$I(s) = 4\pi \int_0^\infty P(r) \frac{sin(sr_{ij})}{sr_{ij}} dr$$

As is clear from the previous two equations each distance r_{ij} contributes a term of the form $\sin(x)/x$ to the intensity. The P(r) can therefore in theory be computed from the intensity via the inverse Fourier transform:

$$P(r) = \frac{r^2}{2\pi^2} \int_0^\infty s^2 I(s) \frac{sin(sr)}{sr} ds$$

In practice as we do not have I(s) measurements from 0 to ∞ we have to make some reasonable assumptions about our distance distributions to produce

a P(r) (such as the value being 0 for r larger than the size of our particle and r=0, the value being positive for all other values, the curve being smooth, etc. see Svergun (1992) for an in depth discussion).

Another quantity that is often measured independently is the radius of gyration R(g) of a molecule which corresponds to the root-mean-square distance from the center of the particle of each scatterer. This can be calculated from the P(r) but also can be estimated, based on the assumption that the particle is roughly spherical (Guinier et al., 1955), from the Guinier plot (ln[I(s)] vs s^2) using the following equation, which only holds true at low s:

$$ln[I(s)] = ln[I(0)] - \frac{s^2 R_G^2}{3}$$

Likewise, one can also determine folding state and other general properties of the protein. These, along with the P(r), can then be compared to a model to determine whether they are consistent with a given hypothesis.

4 Materials and methods

4.1 Protein buffer compositions

Buffer(s)	Component	Concentration
NMR assignment	Sodium phosphate pH 6.0	50mM
	NaCl	100mM
	TCEP	0.5mM
TRIM2 NHL RNA titration	BisTris pH 5.8	50mM
	NaCl	50mM
	TCEP	0.5mM
TRIM3 Filamin-NHL lysis TRIM3 NHL lysis	HEPES pH 7 (pHed at 4°C)	100mM
	NaCl	500mM
	Glycerol	10% (v/v)
	β-mercaptoethanol	2.5mM
	RNase A	0.05μg/mL
	DNase I	0.25μg/mL
	EDTA-free protease inhibitor	120μg/mL
TRIM3 Filamin-NHL IMAC TRIM3 NHL IMAC	HEPES pH 7 (pHed at 4°C)	100mM
	NaCl	500mM
	Glycerol	10% (v/v)
	β-mercaptoethanol	2.5mM
TRIM3 Filamin-NHL SEC TRIM3 NHL SEC	HEPES pH 7 (pHed at 4°C)	20mM
	NaCl	200mM
	Glycerol	5% (v/v)
	TCEP	1mM
	Sodium azide	0.01% (w/v)
TRIM2 NHL lysis TRIM2 Filamin lysis TRIM3 Filamin lysis	Tris pH 7.8 (pHed at 4°C)	50mM
	NaCl	500mM
	β-mercaptoethanol	1mM
	RNase A	0.05μg/mL
	DNase I	0.25μg/mL
	EDTA-free protease inhibitor	120μg/mL
TRIM2 NHL IMAC TRIM2 Filamin IMAC TRIM3-Filamin IMAC	Tris pH 7.8 (pHed at 4°C)	50mM
	NaCl	500mM
	β-mercaptoethanol	1mM
TRIM2 Fil-NHL lysis	Tris pH 7.8 (pHed at 4°C)	50mM
	NaCl	100mM
	β-mercaptoethanol	1mM
	RNase A	0.05μg/mL
	DNase I	0.25μg/mL
	EDTA-free protease inhibitor	120μg/mL

Buffer(s)	Component	Concentration
TRIM2 Fil-NHL IMAC	Tris pH 7.8 (pHed at 4°C)	50mM
	NaCl	100mM
	β-mercaptoethanol	1mM
SAXS	Tris pH 7.8 (pHed at room temperature)	100mM
	NaCl	200mM
	TCEP	0.5mM
TRIM3 Filamin Crystallisation	Tris pH 7.8 (pHed at room temperature)	50mM
	NaCl	200mM
	TCEP	0.5mM

Table 2: Composition of protein buffers used in this project

4.2 Media and media components

Medium	Component	Concentration
LB (prepared by EMBL kitchen)	Tryptone	10g/L
	Yeast extract	5g/L
	NaCl	10g/L
	pH to 7.4 and autoclave	
Minimal medium (using ^1H_2O or ^2H_2O)	10x M9 medium	10%(v/v)
	100x Trace element solution	1%(v/v)
	20%(w/v) unlabelled glucose	2%(v/v)
	or ^{13}C glucose	**or** 2g/L
	$MgSO_4$	2mM
	$CaCl_2$	0.1mM
	Thiamin	6mg/L
	NH_4Cl	0.5g/L
	filter sterilise	
10x M9 salts (using ^1H_2O or ^2H_2O)	NaH_2PO_4	60g/L
	K_2HPO_4	30g/L
	NaCl	5g/L
	pH to 7.4 and filter sterilise	
100x Trace elements (using ^1H_2O or ^2H_2O)	EDTA	5g/L
	$FeCl_3$ x $7H_2O$	6g/L
	$MnCl_2$ x $4H_2O$	1.15g/L
	$CoCl_2$ x $4H_2O$	0.8g/L
	$ZnSO_4$ x $7H_2O$	0.7g/L
	$CuCl_2$ x $2H_2O$	0.3g/L
	H_3BO_3	0.02g/L
	$(NH_4)6Mo_7O_{24}$ x $4H_2O$	0.25g/L
	EDTA	5g/L
	stir until golden and filter sterilise	

Table 3: Composition of media (and buffers used in media) used in this project

4.3 Cloning

Two expression plasmids were used for the constructs produced in this project pETM22, a modified version of pET-24d, which places construct expression under the control of a T7/lac promoter, carries a kanamycin resistance gene and is designed to produce protein with a 3C cleavable N-terminal Thioredoxin-6His solubility and affinity tag, and pET28-MHL which also places construct

expression under the control of a T7/lac promoter, carries a kanamycin resistance gene and is designed to produce protein with a TEV cleavable N-terminal 6His affinity tag.

All cloning was carried out using restriction free cloning with the strategy being, briefly, as follows: the sequence of interest was isolated and amplified by polymerase chain reaction (PCR), the PCR product was separated from the template and primers by agarose gel electrophoresis and the relevant band was extracted from the gel using a Qiagen MinElute gel extraction kit as per manufacturer's instructions, then the eluted PCR product was used as a super-primer for amplification into the target vector (van den Ent and Löwe, 2006).

After cloning, *E. coli* cells of the DH5α strain were used for large scale plasmid production, cells were transformed by heat shock, plated on agar with the appropriate antibiotic then grown at 37°C overnight. Colonies were then picked, used to seed LB for overnight growth at 37°C, and then the plasmid was extracted using a Qiagen Miniprep kit as per manufacturer's instructions. Plasmids obtained were sequenced by a commercial service (GATC or eurofins).

All constructs used in this project along with the plasmid and *E. coli* strain in which they were expressed are listed in Table 4, page 40.

Construct name	Boundaries	Plasmid	Expression system
TRIM2-Filamin	318-427	pETM22	BL21(DE3)
TRIM2-NHL	468-744	pETM22	BL21(DE3)
TRIM2-Filamin-NHL	318-744	pETM22	BL21(DE3)
TRIM3-Filamin	318-420	pETM22	BL21(DE3)
TRIM3-NHL	464-744	pETM22	Strain 4
TRIM3-Filamin-NHL	301-744	pET28-MHL	BL21(DE3)

Table 4: Constructs used in this project. All boundaries are given relative to the canonical human sequences (TRIM2:Q9C040-1, TRIM3:O75382-1) given by UniProt (Consortium, 2018). Strain 4 is from de Marco et al. (2007), see text for details.

4.4 Protein purification

Protein constructs (i.e. TRIM3-NHL:464-744, TRIM3-Filamin:318-420, TRIM3-Filamin-NHL:301-744, TRIM2-NHL:468-744, TRIM2-Filamin:310-420 were TRIM2-Filamin-NHL:318-744) were purified as follows: proteins were expressed using the *E. coli* strain BL21(DE3), in the cases of all constructs other than TRIM3-NHL or the *E. coli* strain BL21(DE3) co-expressing the bacterial chaperone proteins KJE, ClpB and GroELS (Strain 4) (de Marco et al., 2007) in the case of TRIM3-NHL. Cells were grown in LB, ^{15}N minimal medium or ^{2}H-^{15}N-^{13}C minimal

medium as appropriate at 37°C to an OD_{600} of 0.8. Following induction with 0.2mM IPTG, proteins were expressed overnight at 18°C. Cells were pelleted, re-suspended in the relevant lysis buffers (see Table 2, page 38) and stored at -80°C until needed for purification. Cells were thawed and lysed using a French press in the case of TRIM3-NHL, TRIM3-Filamin and TRIM3-Filamin-NHL, or by sonication in the case of TRIM2-NHL, TRIM2-Filamin and TRIM2-Filamin-NHL. The lysates were cleared by centrifugation at 28 800g for 30 minutes at 4°C and filtered. The protein was purified using two rounds of Nickel affinity chromatography followed by size exclusion chromatography: first the eluate was applied to a Nickel-Sepharose column (GE) before being washed with 10CV of the relevant IMAC buffer + 25mM imidazole and eluted with the relevant IMAC buffer + 500mM imidazole in the case of TRIM3-NHL, TRIM3-Filamin and TRIM3-Filamin-NHL, or eluted with the relevant IMAC buffer + 1M imidazole in the case of TRIM2-NHL, TRIM2-Filamin and TRIM2-Filamin-NHL. The tagged protein thus obtained was then cleaved using 3C protease in all cases other than TRIM3-Filamin-NHL and with TEV protease in the case of TRIM3-Filamin-NHL, this was done during an overnight dialysis at 4°C against the relevant IMAC buffer. A reverse Nickel affinity chromatography step was then carried out by applying the cleaved protein to the Nickel-Sepharose column, washing with the relevant IMAC buffer + 25mM imidazole and collecting the flow through. The flow through was then subjected to a final size exclusion chromatography purification step through application to an S75 gel filtration column into the appropriate buffer.

In the case of our ^{2}H-^{1}5N-^{13}C TRIM2-NHL construct (468-744), in order to be able to acquire NMR experiments involving backbone nuclei I needed to exchange backbone ^{2}H for ^{1}H from the solvent (as these experiments are acquired on ^{1}H nuclei), due to the extreme stability of the NHL domain core this exchange was very slow, I therefore stored the protein at a concentration of 5μM at 37°C in TRIM2 NHL IMAC buffer+0.5M urea for 40 days. Following this time the protein was concentrated and buffer exchanged into NMR assignment buffer.

4.5 Crystallisation of constructs and acquisition of X-ray crystallography data

TRIM2-NHL crystals were obtained at room temperature by hanging drop vapour diffusion using a starting protein sample of 10mg/mL in 20mM Tris pH 7.5, 150mM NaCl, 20% glycerol and 1mM DTT mixed at a 1:1 ratio with a reservoir mix of 0.2M potassium chloride and 24% (w/v) polyethylene glycol 3350 pH 7.5 as first described by Guan et al. (2014). A first dataset was acquired

on a rotating-anode diffractometer at EMBL Heidelberg (generator: RIGAKU MircroMax-007, detector: image plate detector Mar345 dtb, goniostats: Mardtb from Marresearch) at a 1.87Å resolution and an initial model was prepared (R_{free}=0.26061 and R_{work}=0.21076) phasing was carried out using molecular replacement based on the NHL domain of Brat (PDB accession code: 1Q7F, Edwards et al. (2003)). A slightly higher resolution dataset containing useful data to 1.80Å was acquired at the ESRF on beamline ID23 (Gabadinho et al., 2010; Nurizzo et al., 2006), using this data and the previously prepared model I prepared an improved model (R_{free}=0.2228 and R_{work}=0.1909, full statistics in Table S4 page 118). All TRIM2-NHL crystallisation and data acquisition was carried out by Jaelle Foot, initial model preparation was carried out by Jaelle Foot and Brice Murciano, final model preparation was carried out by me.

TRIM3-Filamin crystals were obtained by sitting drop vapour diffusion at 8°C using a starting protein sample of 10mg/mL in TRIM3-Filamin crystallisation buffer mixed at a 1:1 ratio with a reservoir mix of 0.1 sodium cacodylate pH 6.5 plus 1M sodium citrate, cryo-protection was with 25% glycerol. Data was acquired to a resolution of 2.134 Å at ESRF on beamline ID30 (McCarthy et al., 2018). Phasing was carried out using molecular replacement based on the filamin domain of TRIM71 (PDB accession code: 4UMG, Tocchini et al. (2014)), model quality indicators were R_{free}=0.2575 and R_{work}=0.2270 (full statistics in Table S3, page 115).

Intensities were integrated using XDS (Kabsch, 2010a,b), subsequent analysis and refinement was carried out in Phenix versions 1.18.2 and 1.17.1 (Liebschner et al., 2019) and manual model building was carried out using COOT (Emsley and Cowtan, 2004).

4.6 NMR experiments

All RNA titrations, TRIM2-NHL assignment experiments and TRIM2-NHL ^{1}H-^{15}N-HSQC experiments were acquired on a 18.8T (800MHz) Bruker Avance III instrument equipped with a cryogenic triple resonance HCN probe. TRIM2-Filamin-NHL, TRIM3-Filamin-NHL, TRIM3-NHL, TRIM2-Filamin and TRIM3-Filamin ^{1}H-^{15}N-HSQC experiments were acquired on a 16.4T (700MHz) Bruker Avance III instrument equipped with a triple resonance HCN probe. TRIM2-Filamin and TRIM3-Filamin assignment experiments were acquired on a 14.1T (600MHz) Bruker Avance III instrument equipped with a cryogenic triple resonance HCN probe.

TRIM2-NHL was assigned with spectra acquired using the TROSY approach applied to ^{1}H-^{15}N-HSQC, HNCACB, HN(CO)CACB (Nietlispach, 2005; Yang and Kay, 1999), HNCA (Eletsky et al., 2001; Salzmann et al., 1999; Schulte-

Herbrüggen and Sørensen, 2000), HNCO (Salzmann et al., 1998) and HN(CA)CO (Salzmann et al., 1999), these were acquired at 310K with 200μM protein in NMR assignment buffer (see Table 2, page 38).

^{1}H-^{15}N-HSQC titration experiments were acquired using the pulse sequences as described in Mori et al. (1995), later experiments were acquired using an apodization weighted acquisition scheme to improve signal to noise (Simon and Kostler, 2019). All titration experiments were carried out at 298K in TRIM2-NHL RNA titration buffer (see Table 2, page 38).

Relaxation experiments were acquired using the topspin pulse sequences hsqcnoef3gpsi, hsqct2etf3gpsitc3d.2 and hsqct1etf3gpsi3d.2 (Kay et al., 1989) for NOE, T_2 and T_1 measurements respectively. Delays for T_1 measurements were: 20ms (duplicated), 150ms (duplicated), 500ms, 800ms, 1000ms, 1600ms, 2400ms and 3200ms, while those for T_2 measurements were: 20ms (duplicated), 50ms, 100ms, 150ms (duplicated), 400ms, 500ms, 650ms, 800ms, 1000ms and 1600ms. Experiments were acquired at 298K in TRIM2 NHL RNA titration buffer.

All NMR experiments were processed using NMRpipe (Delaglio et al., 1995) and assignments were carried out in CcpNmr Analysis version 2.4.2 (Vranken et al., 2005).

4.7 Relaxation experiments analysis

Relaxation experiments were analysed using PINT version 2.1.0 (Ahlner et al., 2013). Once relaxation rates for individual residues had been extracted an overall rotational diffusional tensor as well as individual residues' order parameter were determined using ROTDIF 3.1 (Berlin et al., 2013; Walker et al., 2004). In this software T_1, T_2 and NOE values were fitted to an extended Lipari-Szabo model (Clore et al., 1990b; Lipari and Szabo, 1982a,b). Giving rise, for an anisotropic model, to the following equation:

$$J(\omega) = \sum_{\eta=-2}^{2} c_\eta \left[\frac{S^2 \tau_\eta}{1 + (\omega\tau_\eta)^2} + \frac{(1 - S_f^2)\tau_{f,\eta}}{1 + (\omega\tau_{f,\eta})^2} + \frac{S_f^2(1 - S_s^2)\tau_{s,\eta}}{1 + (\omega\tau_{s,\eta})^2} \right]$$

see Section 3.3, page 28 for a full explanation of these variables.

However, as relaxation data was only be acquired at one field strength, there are an insufficient number of parameters to fit this model. Thus the rotational diffusion tensor (**D**) is initially determined by excluding flexible residues (Berlin et al., 2013; Walker et al., 2004). This was done on the basis of NOE measurements, wherein residues having a ratio of unsaturated signal to saturated signal < 0.75 were flagged as flexible and removed from initial calculations of **D**.

4.8 RNA titrations

Lyophilized RNAs were purchased from IBA and Metabion and resuspended in water. Each RNA was progressively added to a single 100μM protein sample. In the case of the TRIM2-NHL construct, titrations were carried out in TRIM2 NHL RNA titration buffer, with the exception of the titration against UUG UUG which was carried out in NMR assignment buffer (see Table 2 page 38). As protein samples in this buffer showed no signs of aggregation or any changes in the ^{1}H-^{15}N-HSQC spectrum following flash freezing in liquid nitrogen and storage at -80°C (see Figure S1, page 112), large scale protein preparations were produced and the resulting samples were stored as above and aliquots were used as needed.

4.9 Analysis of RNA titrations

^{1}H-^{15}N-HSQC peaks were tracked manually in CcpNmr Analysis version 2.4.2 (Vranken et al., 2005). Chemical shift perturbations were then expressed as weighted distances(Williamson, 2013; Williamson et al., 1997) according to the equation:

$$\Delta\delta = \sqrt{\frac{1}{2}\left((\delta_{H1} - \delta_{H2})^2 + 0.14\,(\delta_{N1} - \delta_{N2})^2\right)}$$

Chemical shift perturbations were then used to determine the K_d (assuming a single binding site) using the equation:

$$\Delta\delta_{obs} = \Delta\delta_{max}\frac{[P]_t + [L]_t + K_d - \sqrt{([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t}}{2[P]_t}$$

Where $\Delta\delta_{obs}$, $\Delta\delta_{max}$, $[P]_t$ and $[L]_t$ represent the observed weighted distance, the weighted distance at full binding of the protein by the ligand, and total protein and ligand concentrations, respectively. This procedure was carried out by tracking multiple peaks (the 20 peaks showing the largest CSPs) while fitting different $\Delta\delta_{max}$ for each peak but fitting the same K_d as recommended in the literature (Arai et al., 2012; Williamson, 2013). Fitting was carried out using R, in which parameters were estimated using a nonlinear least squares approach.

Within a single titration point, weighted chemical shift perturbations were classified as significant if they exceeded the average CSP under those conditions plus one standard deviation (Schumann et al., 2007).

For CSP comparison representations, weighted distances were calculated as above. Each vector representing a single CSP was then scaled to the longest vector, in all residues and for all titrations, at the titration point used.

4.10 Small angle X-ray scattering and associated modelling

SAXS data was collected at the P12 beamline at the PETRA III storage ring at DESY (Hamburg) operated by EMBL (Blanchet et al., 2015). Measurements were carried out at 20°C in SAXS buffer (see Table 2, page 38) in flow cell mode. For each sample 20 to 45 frames of 0.05 to 0.145s were acquired using wavelengths between 0.122798 nm and 0.124026 nm and a pilatus6M detector. Primus was used to manually check frames for radiation damage and other abnormalities then selected frames were averaged and buffer subtracted (Konarev et al., 2003). The pairwise distribution function was calculated using GNOM (Svergun, 1992).

The modelling of individual TRIM2-Filamin-NHL structures with different linker conformations was carried out by restricted molecular dynamics simulations in CNS(1.2) using an ARIA framework as previously described (Brünger, 2007; Brünger et al., 1998; Linge et al., 2003; Rieping et al., 2006; Simon et al., 2010). Structures of the TRIM2 NHL(469-744) and filamin (321-427), obtained by homology modelling based on the TRIM3 filamin domains were kept fixed and connected by a 42 amino acid linker with randomised Φ and Ψ angles. These structures were then energy minimised using Cartesian dynamics simulated annealing, the simulated annealing was composed of three phases: one at 20 000K with 40 000 steps, and two cooling phases from 2000K to 1000K and from 1000K to 50K each composed of 40 000 steps. Each structure was then fitted against the SAXS data using CRYSOL (Svergun et al., 1995). This was compared to data generated using the ensemble optimisation method, EOM (Bernadó et al., 2007; Tria et al., 2015), using the same structures and default parameters.

4.11 Protein melting temperature determination by differential scanning fluorimetry

Protein melting temperatures were determined by the differential scanning fluorimetry (DSF) method which is based on measuring changes in the fluorescence of tryptophans at 330nm and 350nm which is sensitive to their chemical environment and therefore varies over the course of protein unfolding (Alexander et al., 2014). Proteins were soaked into a capillary tube and heated at a rate of 1°C/min, over the course of this time fluorescence at 330nm, 350nm, the ratio of these two values and scattering were measured. The software provided by the manufacturer was then used to find maxima of the derivative of these values from which the melting temperature was calculated. All DSF experiments were carried out in TRIM3-NHL SEC buffer (see Table 2, page 38) using a Nanotemper Prometheus NT.48.

5 Results

5.1 TRIM3 binds RNA with likely NHL domain involvement

Prior to my starting the project an initial RNA immunoprecipitation and sequencing (RIP-seq) experiment (Zhao et al., 2010) was carried out on TRIM3. This methodology finds RNAs bound by a given protein. The methods consists in immunoprecipitating the protein of interest, thereby capturing bound RNA, removing non-specific low affinity binders using a low salt wash and then eluting bound RNAs with a high salt wash, these RNAs are then used to produce a cDNA library which is then sequenced using high throughput methods. It should be noted that this method, unlike the popular alternative of cross-linking and immunoprecipitation (CLIP-seq) may also detect indirectly bound RNAs if the protein forms part of a complex and does not provide an RNA binding site (Wessels et al., 2016). In our case the procedure was carried out using mouse brain lysate (in keeping with the neuronal expression pattern of TRIM3) and three different antibodies. The vast majority (96%) of RNAs detected were mRNAs. We decided to pool enriched RNA reads from all three antibodies and focus on the mRNA hits. This experiment was not repeated for TRIM2 due to a lack of suitable antibodies at the time.

Given that most studied post-transcriptional regulation of mRNA translation occurs at the 5′ UTR (Wilkie et al., 2003) we decided to focus on these in our analysis. We analysed the data using DREME, a piece of software that discovers short motifs (up to 8 base pairs) in a paired set of sequences (Bailey, 2011). Using this we compared the 5′UTRs of enriched and non-enriched RNA sequences detected in our RIP-seq experiment (there were 17563 and 40523 sequences, respectively, in the enriched and non-enriched RNA pools). This analysis showed an enrichment of CG rich sequences in the 5′UTRs of RNAs enriched in our RIP-seq experiment, suggesting our protein may preferentially bind CG rich sequences (see Figure 6 for the top hits and Table S1, page 114 for full results). As the closely related TRIM71 protein has been shown to interact with the 3′UTRs of mRNAs (Aeschimann et al., 2017) we also analysed these (we retrieved 15484 enriched and 36586 non-enriched sequences in this case). The motifs discovered were more heterogeneous with no clear overall pattern (see Table S2, page 114).

These experiments demonstrated that TRIM3 binds RNA (although potentially indirectly), as the NHL domain is the locus of RNA binding for TRIM-NHLs known to bind RNA my project focused on characterising the NHL domains of TRIM2 and TRIM3.

	Motif	Logo	E-value	Unerased E-value
1.	CGSCBC		8.6e-088	4.4e-063
2.	GCBGCG		4.5e-068	6.0e-059
3.	GCCG		6.0e-053	4.3e-061
4.	CGMGS		5.9e-050	1.6e-060
5.	GHSCG		3.1e-037	1.5e-046
6.	GCG		3.3e-035	1.3e-033
7.	CYYCG		5.1e-028	1.5e-039
8.	CYCG		1.3e-023	2.5e-031
9.	CGS		1.2e-014	1.7e-006
10.	GGGRG		5.4e-009	1.9e-019

Figure 6: Top ten sequences enriched in the 5'UTRs of RNAs bound by TRIM3, as determined by RIP-seq followed by DREME analysis. Note that E-values are the enrichment p-value times the number of candidate motif tested, these can be either erased or unerased, meaning either excluding or including, respectively, sites that match a previously found motif.

5.2 Structure of the TRIM2-NHL domain

It was impossible to directly study the structure and RNA binding of the TRIM3-NHL domain as it was insufficiently stable to easily purify or carry out lengthy experiments (covered in more detail in section 5.7, page 66). Instead I focused on the TRIM2-NHL domain. This was successfully purified as described in section 4.4, page 40, following which resonances from the ^{1}H-^{15}N-HSQC spectrum of TRIM2-NHL were recorded (see section 4.6, page 42). The observation that peaks show good dispersion confirms that the protein is well folded with each residue occupying a distinct chemical environment, in contrast to the tight clustering of peaks around 8ppm observed for unfolded proteins where each residue occupies a similar chemical environment (Page et al., 2005). The next task was to assign each peak to a specific residue in order to allow us to determine residue level information from titrations. The backbone NH groups of 265 out of 270 (98.1%) non-proline residues were successfully assigned using

standard 3-dimensionnal NMR experiments for protein backbone assignment (Sattler et al., 1999). In order to improve signal and resolution, this assignment was carried out using a transverse relaxation optimized spectroscopy (TROSY) approach (Pervushin et al., 1997; Salzmann et al., 1998). Moreover a higher temperature (310K) than subsequent experiments was used to increase protein tumbling thereby decreasing relaxation. Both of these factors resulted in small changes to the chemical shifts relative to those measured in our RNA titration experiments. Given the crowded nature of the spectrum, this introduced some ambiguities in the assignment of the spectrum obtained under RNA titration conditions. This was resolved using temperature titrations and some additional backbone assignment experiments. The assignment was successfully transferred between the two conditions for 256 out of 270 (94.8%) non-proline residues (see Figure 7, page 49). However, it should be noted that, due to overlap and some of these peaks being extremely weak only around 242 out of 270 (89.5%), on average, non-proline residue backbone peaks were successfully tracked in our RNA titration assays (see Table S6, page 119).

The larger size of the NHL domain (leading to faster relaxation and more crowded NMR spectra) would have made structure determination by NMR difficult, we therefore solved the structure of the TRIM2-NHL domain by X-ray crystallography. The NHL domain of TRIM2 was crystallised by Jaelle Foot in accordance with the protocol developed by Guan et al. (2014). Guan et al. (2014) obtained high quality crystals and good diffraction but were unable to solve the structure, presumably due to their use of the distantly related PKnD NHL domain for molecular replacement. I then solved the structure using molecular replacement based on the Brat NHL structure solved by Edwards et al. (2003) as described in more detail in section 4.5 (page 41). Resolution was 1.8Å and R_{work} and R_{free} were 0.1909 and 0.228 respectively, full statistics may be found in Table S4 (page 118). The TRIM2 NHL domain is similar in structure to previously studied NHL domains. It consists of a six bladed β-propeller roughly 40Å in diameter and 27Å in height, each blade of the β-propeller is made up of four anti-parallel β-strands, labelled a to d starting from the center of the propeller (see Figure 8, page 50). Each NHL repeat forms the d-sheet of one blade, a long loop running along the top surface of the NHL domain, and the a, b and c sheets of the next blade. The first NHL repeat is shared between blades I and II and the last between blades VI and I, this is identical to the layout for the Brat NHL domain (see Figure 3, page 21).

5.3 Dynamics of the TRIM2-NHL domain

I also decided to carry out NMR relaxation experiments to characterise the dynamics of the NHL domain. This had two main aims, first as a comparison

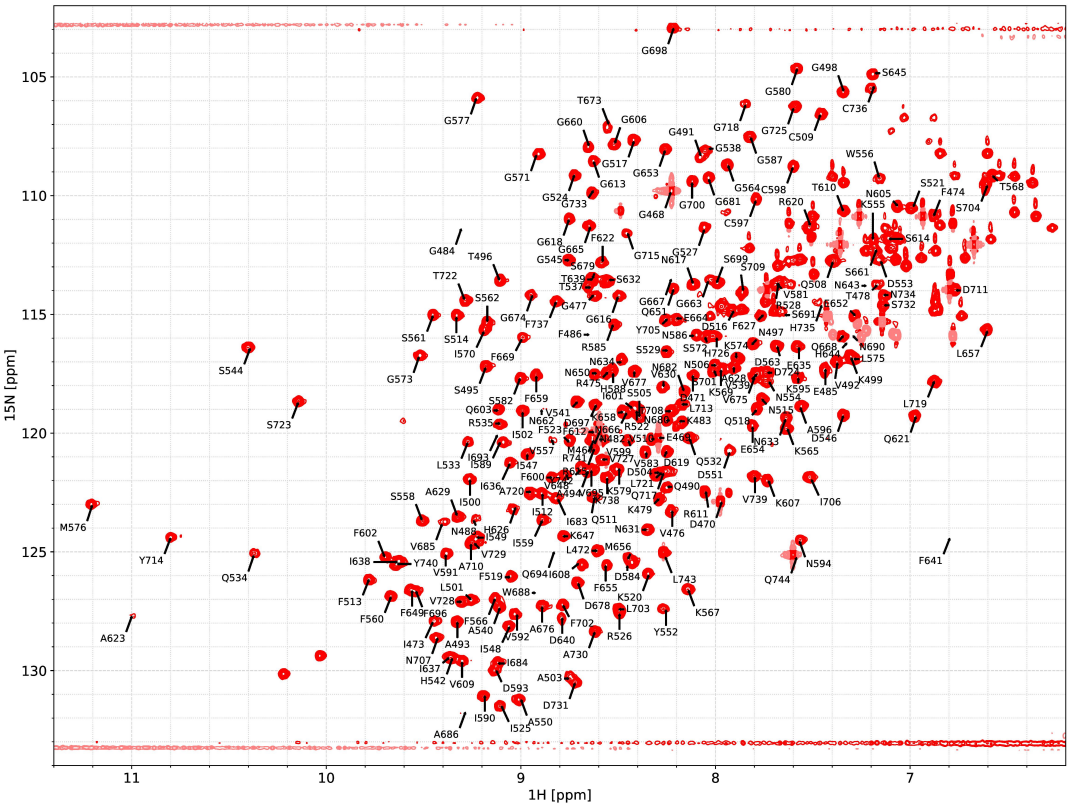
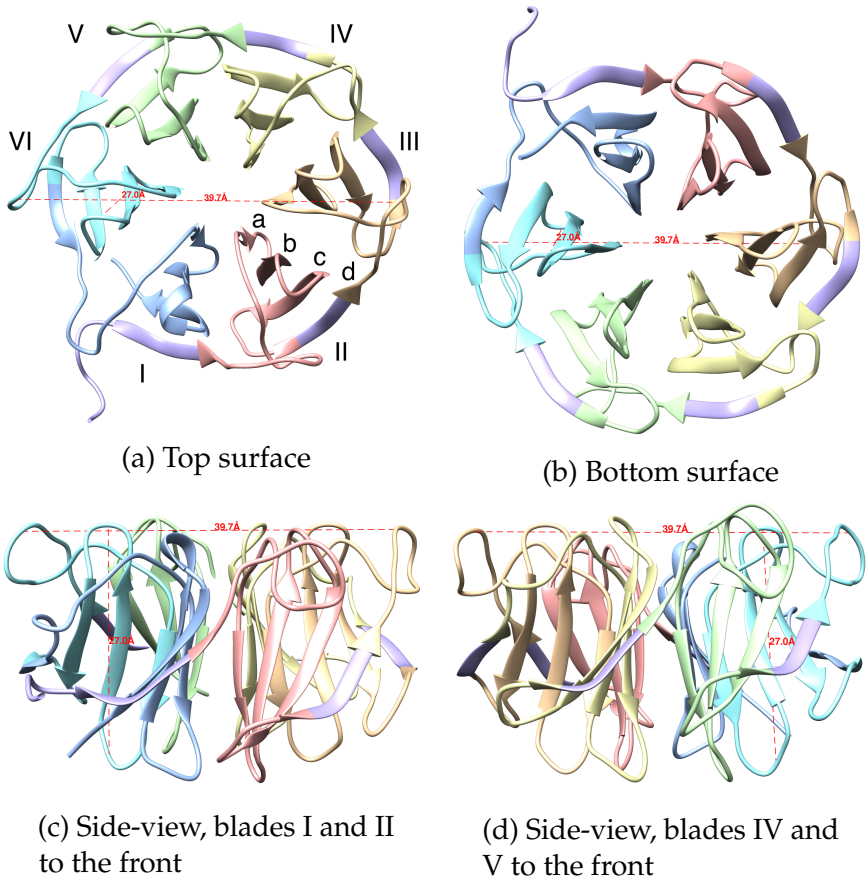


Figure 7: Assigned TRIM2-NHL ^{1}H-^{15}N-HSQC acquired at 298K in TRIM2 NHL RNA titration buffer (see Table 2, page 38). Note that some assigned peaks are too weak to be seen at the displayed signal to noise levels, these were generally not successfully tracked during RNA titrations. Also of interest is the strong downfield shift of residues M576, Y714, Q534 and A623 as discussed in Section 5.4.

to see if this could shed light on the relatively unstable TRIM3-NHL domain (as discussed in section 5.7, page 66) and second to assess the likelihood of substantial conformational changes upon RNA binding. Such conformational change is more plausible if residues are already flexible.

The relaxation parameters T_1, T_2 and NOE enhancements were measured for TRIM2-NHL residues, the results are displayed in Figure 9, page 51. Under conditions where internal motions are fast, the overall rotational correlation time τ_c is in the nanosecond range or longer, and there is no chemical exchange, R_2/R_1 depends purely on the overall rotational correlation time of the protein. Deviations from this are indicative of either conformational exchange or slow motions on the nanosecond timescale (Clore et al., 1990a; Kay et al., 1989; Korzhnev et al., 2001). As can be seen from the R_2/R_1 ratios in Figure 9C, the majority



(a) Top surface

(b) Bottom surface

(c) Side-view, blades I and II to the front

(d) Side-view, blades IV and V to the front

Figure 8: Structure of the TRIM2-NHL. NHL repeats as described in Slack and Ruvkun (1998) are coloured from red to blue with amino acids not assigned to a given repeat left in purple. Blades and β-strands are numbered according to convention.

of the residues in TRIM2-NHL have a similar R_2/R_1, moreover most have high NOE enhancement values, suggesting that their internal motions occur on a fast time-scale. All this suggests that most residues apart from those at the termini (which appear flexible, as expected from the crystal structure) are part of a single folded domain with fast, short range internal motions. An initial estimate for τ_c of 15ns was calculated from the median R_2/R_1.

In order to interpret these results at the level of individual residues, I modelled our data with input from our crystal structure using an extended Lipari-Szabo model (Clore et al., 1990b) implemented by the program ROTDIF (Berlin et al., 2013; Walker et al., 2004) (see Sections 3.3 and 4.7 on pages 28 and 43 respectively for more details). This modelling suggested that protein tumbling was not isotropic but best represented by an axially symmetric slightly oblate ($D_{\parallel}/D_{\perp}$=0.94) diffusion tensor with a τ_c of 15.19±0.04ns. As part of this model,

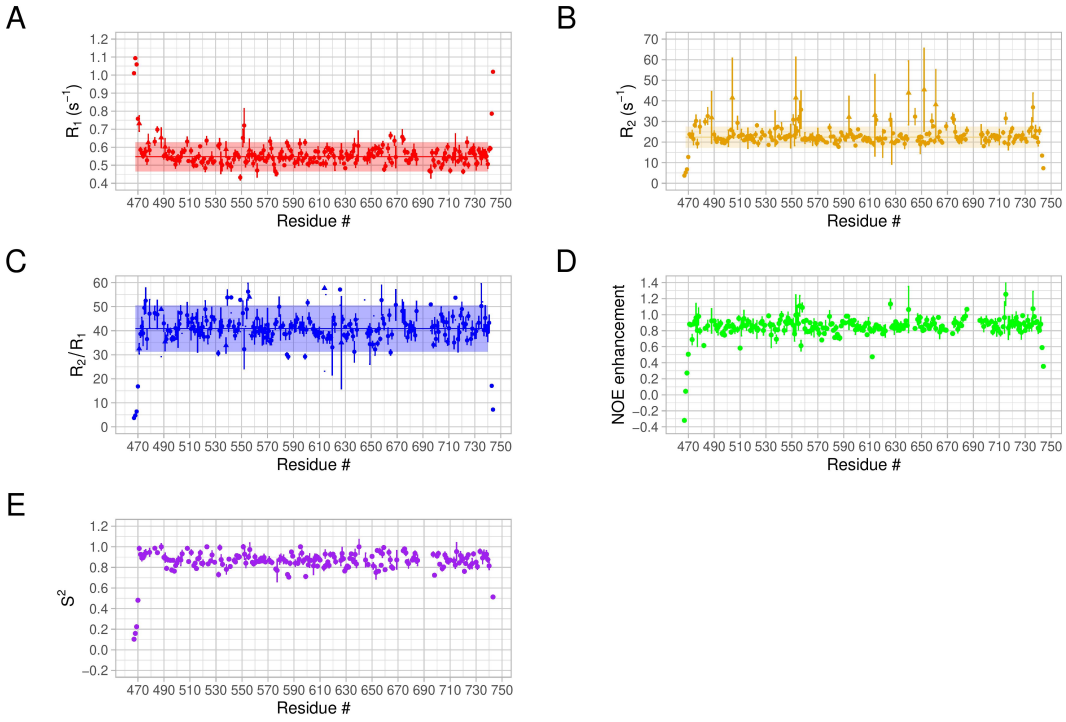


Figure 9: Per residue ^{1}H-^{15}N relaxation data for TRIM2-NHL acquired at 298K. R_1 (**A**),R_2 (**B**) and NOE enhancement (**D**) results along with calculated R_2/R_1 ratios (**C**) and S^2 ratios determined by our extended Lipari-Szabo model (**E**) are represented. Lines represent median value, rectangles are median $\pm$ one standard deviation, peaks with a worse but acceptable fit to exponential decay in relaxation data acquisition are highlighted with triangles.

S^2 values were calculated for each residue (see Figure 9E, page 51). The high S^2 values for most residues again indicate that most non-terminal residues form part of a stable domain that doesn't undergo substantial internal motions. This model appears to be plausible, the axis determined by the model matches the main axis predicted by the ab initio diffusion tensor predictor, ELM (Ryabov et al., 2006), relatively well (11° offset). The model was also somewhat robust to small changes in the data used. When 10% of residues were removed randomly and ROTDIF was rerun the average offset of the axis of symmetry from the ELM prediction was 20° over 10 runs with two of those runs giving large offsets (42° and 52°) but eight matching previous predictions relatively closely (7° to 18° offset)(see Figure S11, page 128). τ_c predictions were stable with all values being between 15.15 and 15.22ns. The lack of stability of the predicted axis may be due to the model being almost isotropic.

5.4 The top surface of the TRIM2-NHL domain

The top surface of the NHL domain is the locus of RNA binding for other studied cases such as Brat and TRIM71. The top surface of the TRIM2 NHL domain is positively charged, much like those of Brat, TRIM71 and the RNA binding TRIM32 homologue Thin (Loedige et al., 2013, 2014; Schwamborn et al., 2009). This is in line with the homology modelling and hypothesised RNA binding roles discussed by Kumari et al. (2018) and Loedige et al. (2014), and contrasts with the lack of positive charge seen in non-RNA binding NHL domains such as those of PAL, PknD and NHLRC2 (see Figure 10, page 52).

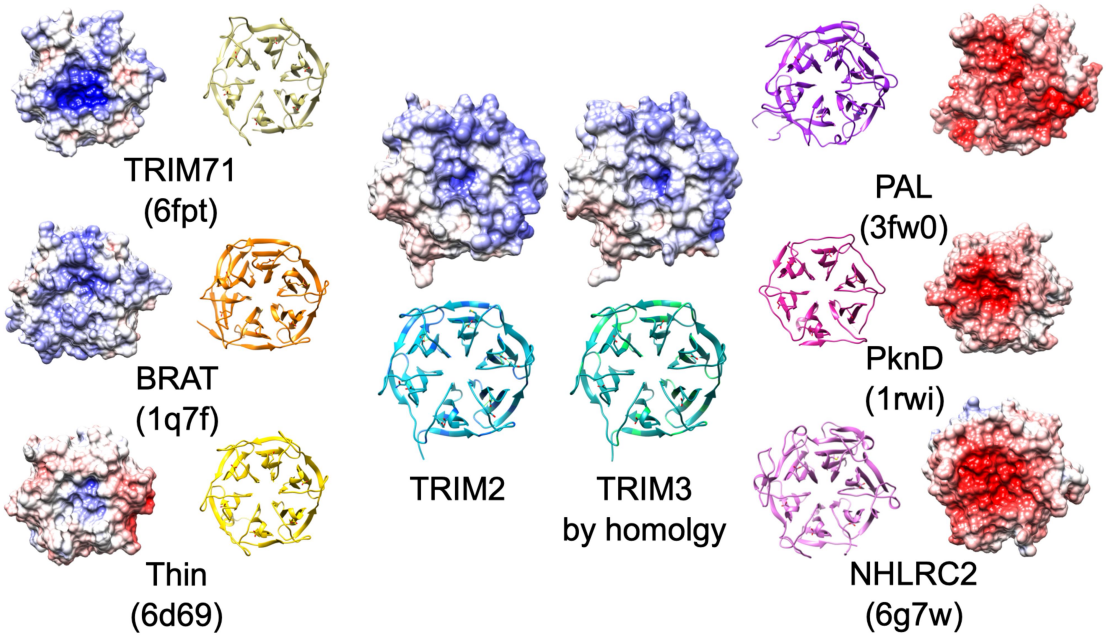


Figure 10: Top surfaces of all NHL domains present in the PDB (excluding artificial NHL domains designed by Voet et al. (2014)) along with those of TRIM2-NHL and a homology model of TRIM3-NHL. Surfaces are coloured blue (positive) to red (negative) depending on the electrostatic potential of the top surface of the protein, these were calculated using the APBS server (Baker et al., 2001). Note the difference between NHL domains known or believed to bind RNA (left) and those for which this is not a known role (right).

The top surface is characterised by the presence of long loops (10-15 amino acids for TRIM2-NHL with an average length of 13.8), these loops are anchored in place by a network of hydrogen bonds (see Figure 11, page 53). Especially notable are those formed by the ring of six aspartates D504, D551, D593, D640, D687 and D731 as these are highly conserved throughout NHL domains (including that of TRIM3), with glutamate sometimes tolerated (Slack and Ruvkun, 1998).

Mutation of these aspartates has been linked to Charcot-Marie Tooth disease in the case of TRIM2 (Pehlivan et al., 2015) and Limb-Girdle muscular distrophy in the case of TRIM32 (Frosk et al., 2002; Schoser et al., 2005), reportedly due to decreased stability of the mutant form (Kudryashova et al., 2011). Their sidechains form hydrogen bonds to the backbone of residues Q534, M576, A623 and Y714 as can be observed from the characteristic downfield chemical shift experienced by these residues' backbone NHs (see Figure 7, page 49). These hydrogen bonds are also apparent in the crystal structure for the equivalently positioned N670 and T487, although as these residues are unassigned they could not be confirmed by NMR.

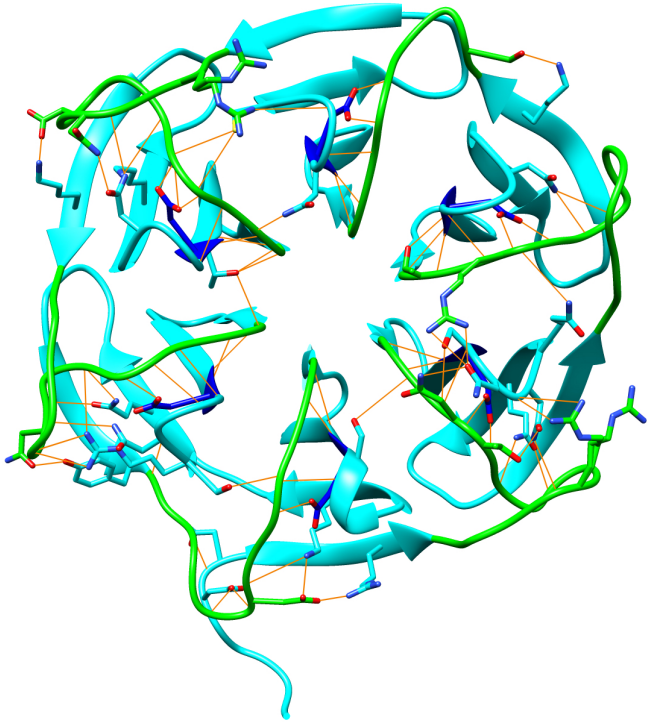


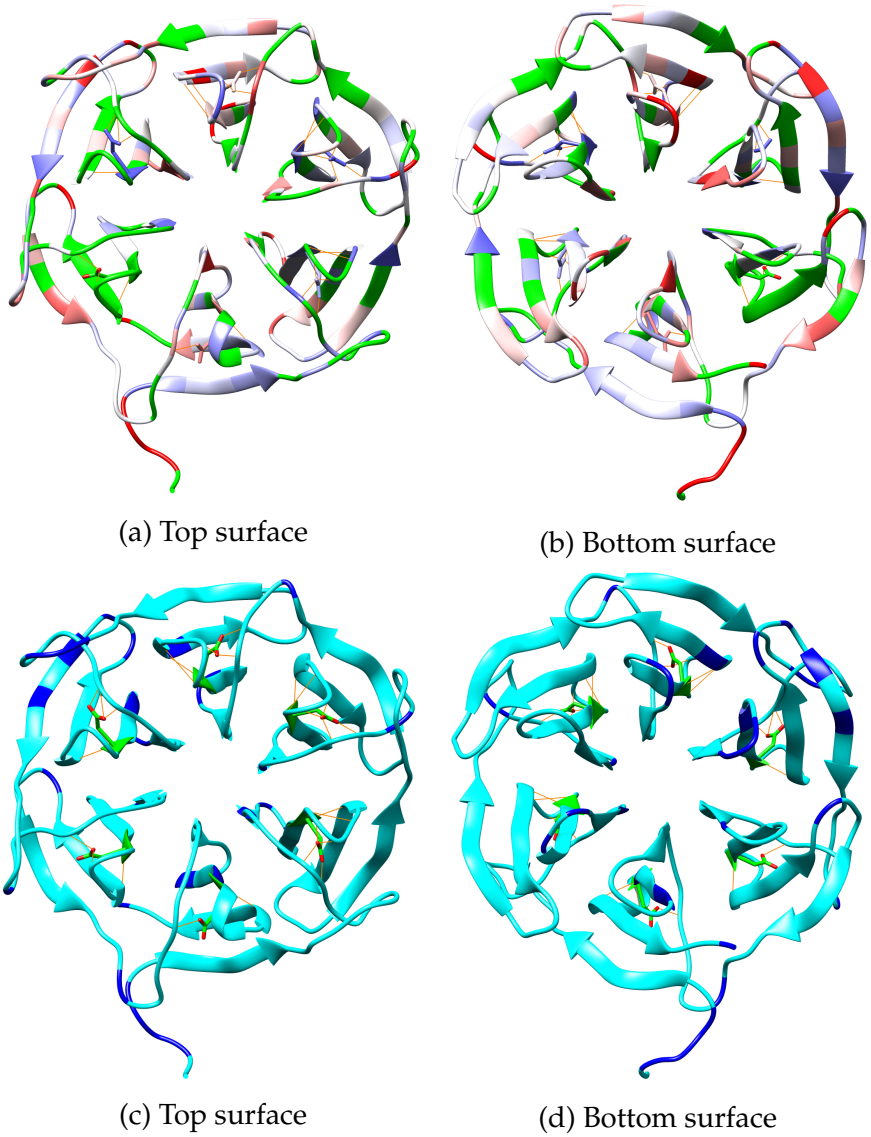
Figure 11: Illustration of the dense network of hydrogen bonds anchoring the loops on the top surface of TRIM2-NHL. Potential hydrogen bonds formed by the top surface loops (highlighted in green) were detected using a relaxed (deviations of up to 0.4Å and up to 20° are tolerated) version of the criteria set out in Mills and Dean (1996) and displayed in orange. Highlighted in dark blue are the conserved ring of aspartates D504, D551, D593, D640, D687 and D731.

This network appears to give greater stability to the backbone of the top surface residues compared to those of the bottom surface despite the presence of considerably longer loops. The average S^2 of residues on or near the top surface

(top surface loops plus the first residue of connecting β-strand) is 0.888 compared to an average S^2 of 0.872 over the whole structure, excluding unstructured termini, both of which lie near the bottom surface (average is 0.848 including them) and an average S^2 of 0.854 for the bottom surface excluding termini (see Figure 12a, page 55).

Additionally, of the 31 residues where our model fits both an S^2_{fast} and an S^2_{slow} (Clore et al., 1990b), indicating the potential presence of slower (on average for these loops τ_e=1.9ns), larger scale dynamics, 8 are at or near the top surface of which 6 are distant from the center of the top surface, while 14 are at or near the bottom surface. Of the remaining, 5 are in the termini and 4 are in β-strands not at the top or bottom of the structure (see Figure 12d, page 55). One should, however, be cautious when interpreting the S^2_{fast} data given that it implies fitting a relatively complex model based on relaxation rates that have only been determined at one field strength. However, when taken together with the S^2 it appears clear that the top surface is as rigid if not more so than the rest of the domain.

These results line up relatively well with molecular dynamics (MD) simulations carried out by our collaborator Thomas Hoffmann. As part of our investigation of why TRIM3-NHL is less stable than TRIM2-NHL, three 1μs MD simulations of the NHL domain of TRIM2 were carried out. For each of these the first 500ns were discarded as during this time results may be more reflective of the crystal structure derived model adapting to the forcefield used in the simulation, the results for the three runs were then combined giving us a total 1.5μs simulation. Examining the root mean square fluctuation (RMSF) of C_α atoms and all atoms over the course of these simulations likewise tend to show a relatively rigid backbone at the top surface (see Figures 13a and 13b, page 56), although some flexible side-chains are present.



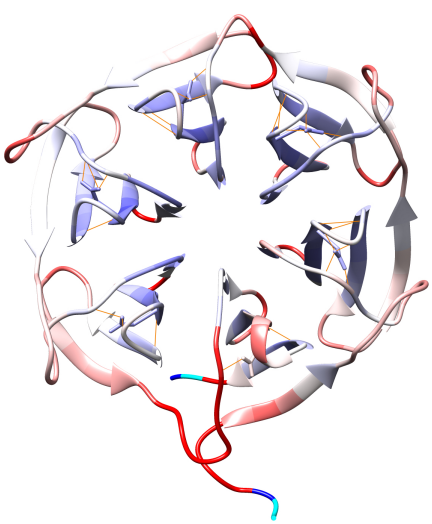
(a) Top surface

(b) Bottom surface

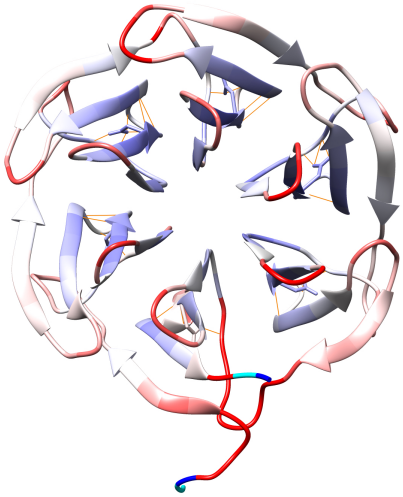
(c) Top surface

(d) Bottom surface

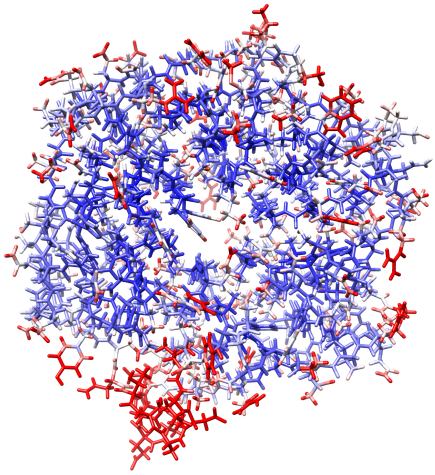
Figure 12: View of the top and bottom surfaces of TRIM2-NHL. **Top:** Residues are coloured blue (high) to red (low) according to their calculated S^2. In each case colour scale maximum is at average plus/minus one standard deviation and average is white. Peaks where accurate measurements were lacking for either T_1, T_2 or NOE are coloured in green. **Bottom:** Residues where our relaxation data model suggest the presence of slower timescale dynamics (for which both an S^2_{slow} and an S^2_{fast} were fitted) are highlighted in dark blue, the structurally important residues D504, D551, D593, D640, D687 and D731 are highlighted in green and the hydrogen bonds formed by their sidechains are displayed.



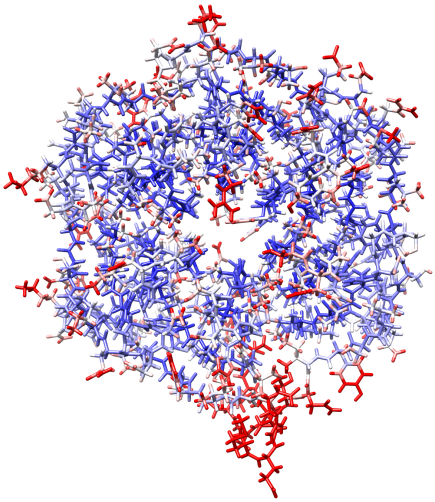
(a) Top surface, C_α RMSF



(b) Bottom surface, C_α RMSF



(c) Top surface, RMSF by atom



(d) Bottom surface, RMSF by atom

Figure 13: **Top:** RMSF of the C_α of each residue over the course of 1.5μs of simulation time. **Bottom:** RMSF of each atom over the course of the same simulation. Residues and atoms are coloured blue (low) to red (high) according to their calculated RMSF. In each case colour scale maximum is at average plus/minus one standard deviation and average is white.

5.5 TRIM2-NHL preferentially binds C rich RNA sequences through its top surface

Having characterised the structure and dynamics of the TRIM2-NHL domain and assigned its NMR spectrum we can now look at the RNA binding of the domain.

Through NMR titrations of a variety of RNA sequences against TRIM2-NHL, I was able to confirm that it does, as expected, bind RNA, as evidenced by large CSPs visible throughout the course of these titrations (see Figure 14, page 58). Using the crystal structure we obtained of TRIM2-NHL as well as the backbone assignment for our HSQC spectra I could confirm that binding is mediated by the top surface of the NHL domain, similarly to other known NHL-RNA interactions (see Figures 15a and 15b on page 59). The location of chemical shift perturbations is consistent with the RNA binding along a curve running from between blades II and III through the rough center of the top surface to between blades IV and V. This can be contrasted to Brat-NHL where the RNA instead runs from between blades II and III to between blades VI and I (see Figures 15c and 15d, page 59). This change in orientation is consistent, however, with the charge distribution on the top surface of the NHL domain.

In total twenty-seven RNA sequences of a length of 3 to 6 bases were titrated against TRIM2-NHL of which 26 showed binding in the fast exchange regime allowing an estimation of the K_d of the interaction for 24 of those, the two remaining having insufficiently strong binding to estimate K_d at the RNA to protein ratios used (Table 5 and Table S6). If we restrict ourselves to 6-mers we can see a wide variety of binding affinities with K_d values ranging from 89μM to 1 300μM, suggesting binding occurs with a relatively high degree of selectivity. This is in line with our preliminary experiments carried out to study the TRIM2-NHL using scaffold independent analysis (SIA) methodology (Beuth et al., 2007; Collins et al., 2015). In order to establish the suitability of our system for SIA, a method for determining strong RNA binding sequences based on titrations against pools with one position fixed to a particular base while all other positions are randomised, I titrated a randomised 6-N RNA pool against our protein and no large CSPs could be observed (see Figure S9, page 126). Given the large chemical shift perturbations seen for some of our high affinity 6-mers this result suggests that a large portion of the pool binds our protein with very low affinity, suggesting that RNA binding to TRIM2-NHL is highly sequence specific. Although it should be noted that this titration was carried out in a different buffer to most of our other titrations (assignment buffer rather than titration buffer, see Table 2, page 38), which may slightly reduce binding. This combination of weak and specific binding made the SIA approach unsuitable to our system, which is why it was not pursued further.

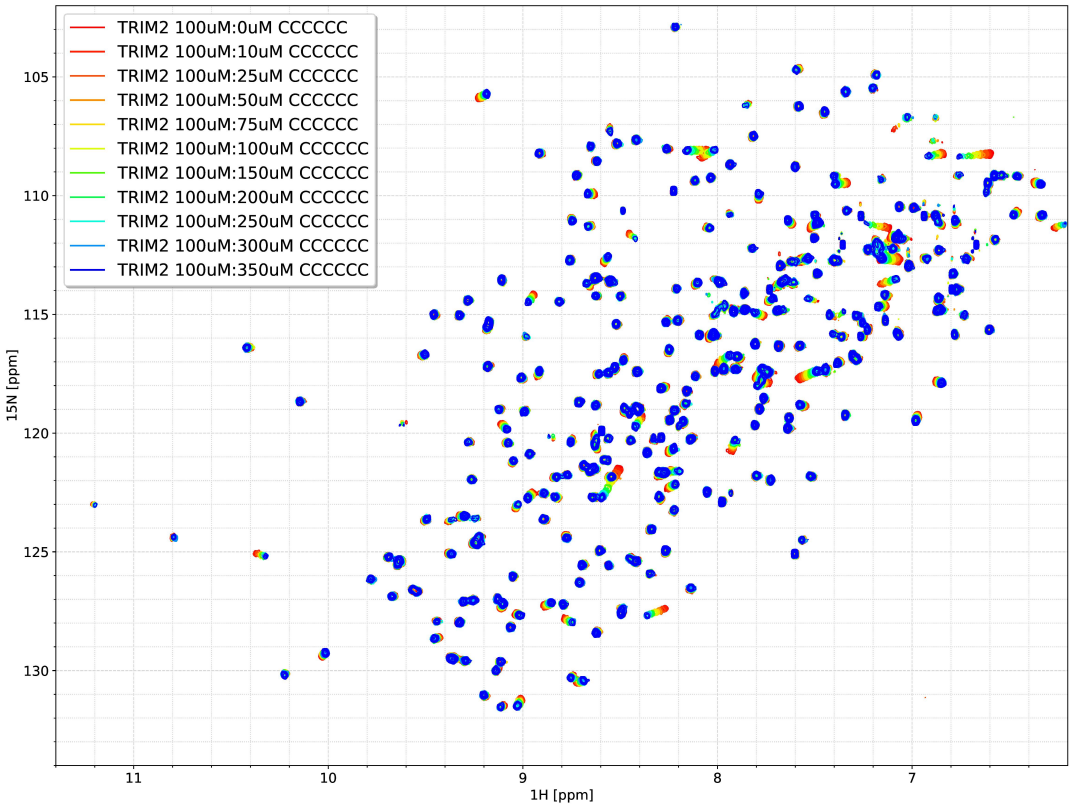
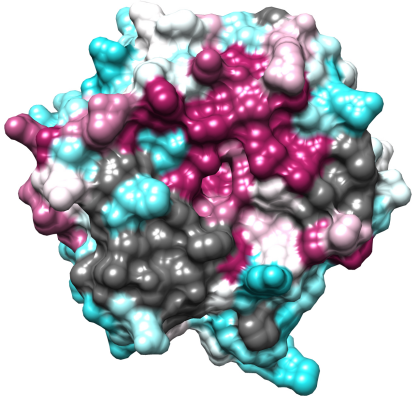


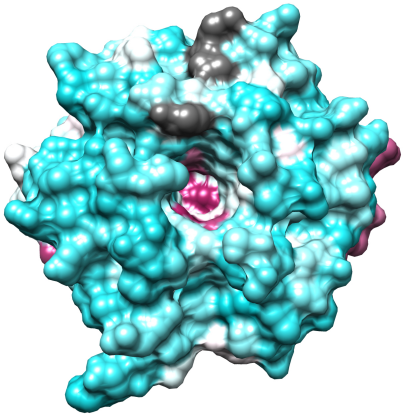
Figure 14: Titration of the RNA 6-mer, CCC CCC against TRIM2-NHL carried out at 298K in TRIM2 NHL RNA titration buffer.

The estimated K_d of these RNAs show a clear pattern of stronger binding for sequences containing cytosines in central and 3′ positions. Some examples illustrating this being the sequence AAC CCG and the swapped sequence CCG AAC which have affinities of 135μM and 660μM respectively or the sequence AAC AAC and the similar sequence AAU AAU which have affinities of 140μM and 490μM respectively. A final clear example of the preference for cytosine at the center of the sequence are the four RNAs with varying third positions: CCC CCC, CCG CCC, CCU CCC and CCA CCC which have K_d 98μM, 230μM, 250μM and 610μM, respectively. Purines appear to be tolerated or even be slightly stabilising in the first two positions of the sequence while sequences containing uracil were generally poor binders (although fewer of these were tested, precisely because of initial experiments suggesting their low affinity). This aspect can again be contrasted to the case of Brat for which the preferred binding motif is UUG UUG (Loedige et al., 2015).

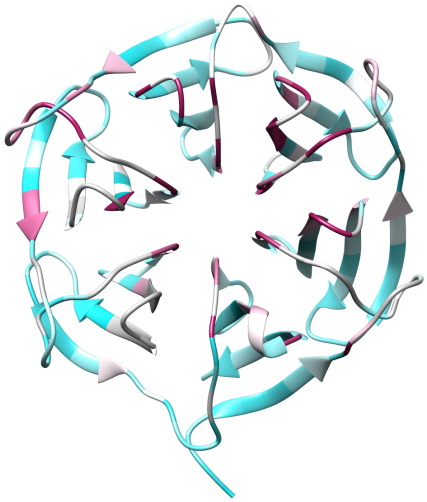
Some differences in affinity between closely related sequences remain hard



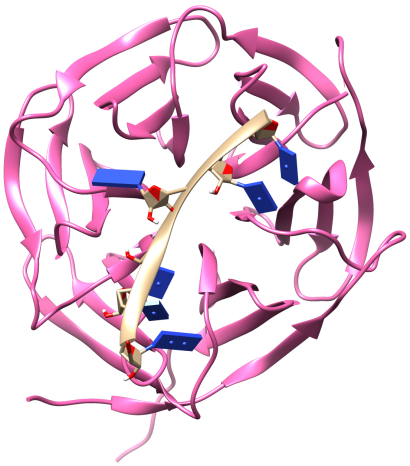
(a) Top view, TRIM2-NHL CSPs upon C-6 titration



(b) Bottom view, TRIM2-NHL CSPs upon C-6 titration



(c) Top view, TRIM2-NHL CSPs upon C-6 titration



(d) Top view, Brat-NHL with UUGUUG bound

Figure 15: Top panels and bottom left panel represent CSPs upon titration of C6 against TRIM2-NHL. Residues are coloured cyan (low) to purple (high) according to CSPs at the 1:3 protein to RNA ratio. In each case colour scale maximum is at average plus/minus one standard deviation and average is white, peaks that could not be tracked are coloured grey. Bottom right panel is the structure of Brat-NHL, coloured pink, bound to an RNA with sequence UUG UUG, coloured tan (Loedige et al., 2015)

to explain based on a simple model of six nucleotides bound with each base independently providing an additive binding penalty or gain depending on its position in the sequence. For example, when one compares the calculated

RNA	**$K_d \pm$ Standard error**
CCG CCG	Not estimated due to signal loss
CCC CCC C	$74\mu M \pm 4\mu M$
CAC CCC	$89\mu M \pm 9\mu M$
GCG CCC	$89\mu M \pm 9\mu M$
GGC CCC	$90\mu M \pm 10\mu M$
CCC CCC	$98\mu M \pm 6\mu M$
AAC CCG	$135\mu M \pm 5\mu M$
ACC CCC	$139\mu M \pm 8\mu M$
AAC AAC	$140\mu M \pm 10\mu M$
ACA CCC	$190\mu M \pm 20\mu M$
GCG CCG	$190\mu M \pm 30\mu M$
GAC CCG	$200\mu M \pm 10\mu M$
GAC CCC	$220\mu M \pm 20\mu M$
GAC CGC	$220\mu M \pm 20\mu M$
CCG CCC	$230\mu M \pm 20\mu M$
ACA CCG	$230\mu M \pm 40\mu M$
CCU CCC	$250\mu M \pm 20\mu M$
UUU UUU	$330\mu M \pm 40\mu M$
ACA ACA	$490\mu M \pm 10\mu M$
AAU AAU	$490\mu M \pm 50\mu M$
CCA CCC	$610\mu M \pm 50\mu M$
CCG AAC	$660\mu M \pm 30\mu M$
AAA AAA	$800\mu M \pm 500\mu M$
CCG	$900\mu M \pm 200\mu M$
UUG UUG*	$1\,300\mu M \pm 300\mu M$
CCC CC	Not estimated due to contamination
CCCC	Not estimated due to weak binding
CCC	Not estimated due to weak binding

Table 5: Estimated K_d for different RNAs titrated against TRIM2-NHL. *titration performed in a different buffer to other titrations (see Section 4.8, page 44).

affinities for sequences GCG CCC and GCG CCG, $89\mu M$ and $190\mu M$ respectively, it would appear that the 3′ terminal G is highly destabilising while in the context of the sequences GAC CCC and GAC CCG with K_d estimates of $220\mu M$ and $200\mu M$ the difference is within measurement error. One simple way to explain such discrepancies would be to suppose that, unlike Brat-NHL, TRIM2-NHL forms interactions with fewer than 6 nucleotides (say n nucleotides where $n < 6$) and the measured K_d of one of our 6-mers depends primarily on the most favoured sub-sequence of length n that exists within our 6-mer (with potentially a

small affinity bonus or penalty associated with RNA sliding). I therefore decided to try to establish whether binding affinity dropped off under a certain length of RNA by testing different lengths of poly-C RNAs. These tests of different RNA lengths demonstrate that TRIM2-NHL forms stabilising interactions with 5 to 6 RNA nucleotides as the binding of C-7mer is not significantly stronger than that of C-6mer (74μM versus 98μM). I would hypothesise that the slight improvement in affinity is attributable to the ability of C-7mer to slide on the binding site without detaching, giving a lower k_{off}. The affinity of C-4mer and C-3mer is so weak that it cannot be measured in our assays (due to the CSPs increasing linearly with RNA addition in the range studied, see Table S6, page 119, making fitting the binding equation described in Section 4.9 impossible). C-5mer was tested but results could not be satisfactorily interpreted due to triethylamine contamination (see Figure S8, page 125). It should also be noted that the RNA CCG had a measurable, though low, affinity while that of CCC could not be measured in our assays, this suggests that the presence of the 3' terminal G significantly increases the RNA's affinity.

One RNA, with sequence CCG CCG showed dramatic signal loss (see Figure S10, page 127). The observation that this signal loss affects nearly all peaks in a relatively uniform fashion as well as the fact that signal does not return even at high RNA to protein ratios leads us to conclude that this is not due to the binding of the RNA being in the intermediate exchange regime but to the formation of a large, slow tumbling, complex. The fact that this signal loss occurred at relatively low RNA concentrations, as well as the high binding affinities of other G rich sequences, suggests that this interaction is relatively strong, however I was unable to produce an estimate for binding affinity.

Only linear RNAs have so far been tested, the possibility that TRIM2-NHL preferentially binds structured RNA is therefore still open. Although the analysis of the topic by Kumari et al. (2018) may suggest this is unlikely as described in more detail in Section 6.3. Individual titrations along with estimated K_d and chemical shift perturbation plots can be found in Supplementary Figure S13 and Supplementary figure S14, starting on pages 130 and 134 respectively.

5.6 Analysis of chemical shift perturbations gives indications concerning direction of RNA binding

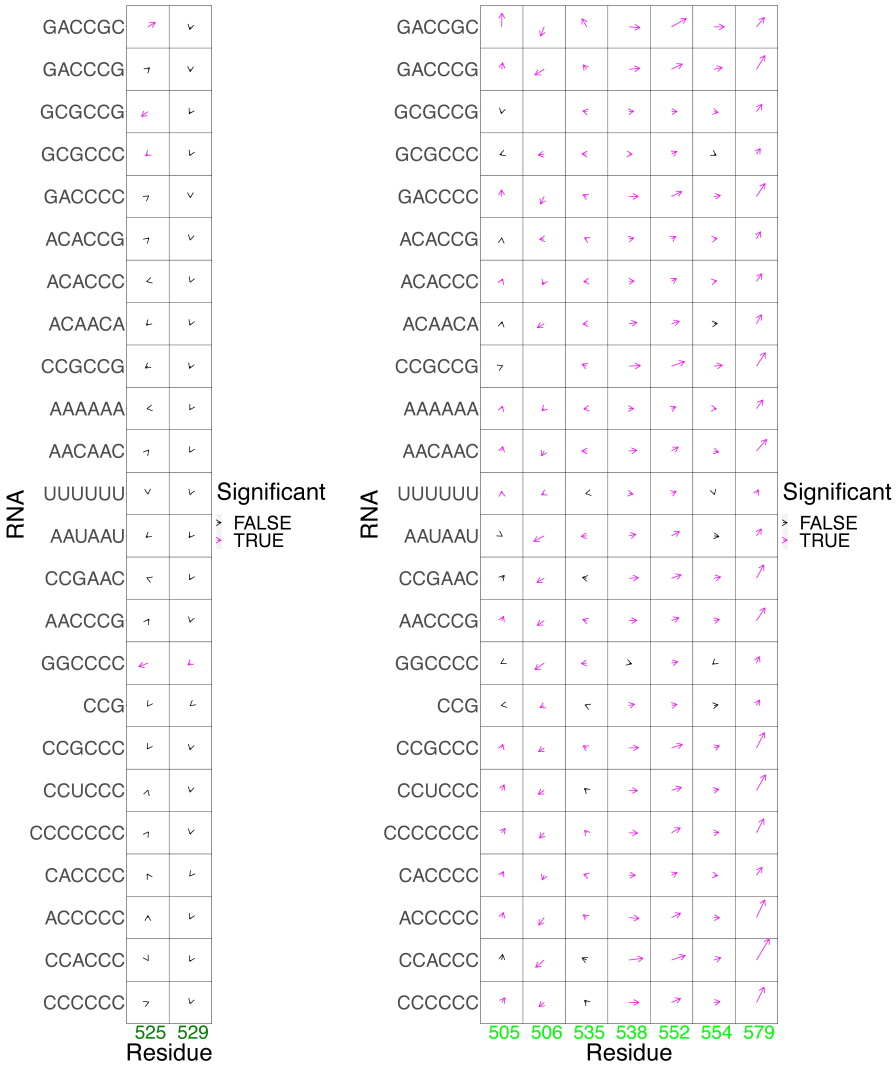
Given that chemical shift perturbations are sensitive to the chemical environment of the residue studied, binding of different ligands results in a different pattern of chemical shift perturbations, even once binding affinity has been corrected for. This implies that when we compare the CSPs of two RNAs, at similar percentages of protein in the bound state, differences will be due to differences in the local

chemical environment of the residues probed. Therefore, by comparing the CSPs experienced by a residue at the binding surface we gain information about the RNA nearby, I attempted to use this to determine the likely direction of RNA binding. I did this by comparing chemical shift perturbations at the titration point closest to 50% binding (this choice is discussed in detail in Section 6.1, page 75).

As can be seen in Figure 16a, two residues have CSPs that appear to depend almost exclusively on the first nucleotide of the RNA sequence. The first is I525 which shows a pattern of only giving significant CSPs when bound by an RNA with a guanine at the 5' terminus. The effect for S529 is much more subtle with smaller, mostly non significant, CSPs but these do seem correlated to those observed for I525.

There exists a wider range of residues for which chemical shift perturbations appear to be correlated to properties of the first half of the RNA sequence although the rules relating sequence to CSP are sometimes less clear-cut (see Figure 16b, page 63). The first example is N506 for which most sequences giving a large chemical shift perturbation are characterised by the presence of a purine at position 2 or 3. Conversely, most of the sequences giving a non-significant chemical shift perturbation for residue I505 have a purine in the 3rd position. In the case of N554 most of the larger CSPs appear to be associated with the specific sequence GAC in the first half of the RNA sequence. For residue Y552 larger chemical shift perturbations seem to be associated with the leading sequences GAC and CCN. For residue R535 there appears to be a pattern of little to no ^{15}N component to the CSP for sequences starting with AA or GG while most (but not all) other sequences tend to produce some shift in the ^{15}N dimension. The size of the chemical shift perturbation experienced by K579 appears to depend on the entire leading triplet with certain triplets, such as GAC consistently giving large chemical shift perturbations, and others such as ACA or GCG giving much smaller ones. Finally the pattern for G538 is far from clear but the presence of cytosine in the 3rd position seems to make larger CSPs somewhat more likely.

Five residues' chemical shift perturbations appear to correlate with properties of the center of the RNA sequence (see Figure 17a, page 64). The first is T537 where large CSPs appear to be associated with the presence of non-cytosine nucleotides, especially purines, in the 3^{rd} and 4^{th} position. In the case of N551 some sequences give CSPs that occur almost entirely in the ^{1}H dimension while others have a ^{15}N component, those for which this is most pronounced appear upon titration with sequences containing a purine in the 5^{th} position. The pattern for A676 is not particularly clear but seems to show larger CSPs for some sequences that have purines in positions 3 and 4. In the case of Q717 two different types of large chemical shift perturbations can be observed, a more or



(a) 5′ dependent (b) 5′half of sequence dependent

Figure 16: Plot of selected chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation. Residue numbers are coloured according to their classification in Figure 18.

less pure downfield shift in the ^{1}H dimension which appears to be associated with some sequences containing a G in the third position while the sequence

(a) Middle dependent

(b) 3′ dependent

(c) Unclear relationship

Figure 17: Plot of selected chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation. Residue numbers are coloured according to their classification in Figure 18.

GAC CGC shows a pure upfield shift in the ^{15}N dimension, this sequence is the only one tested to contain a G in the 5^{th} position which may explain this particular effect. Finally the size of the CSPs exhibited by A720, although weak, appear to be somewhat correlated to those shown by Q717.

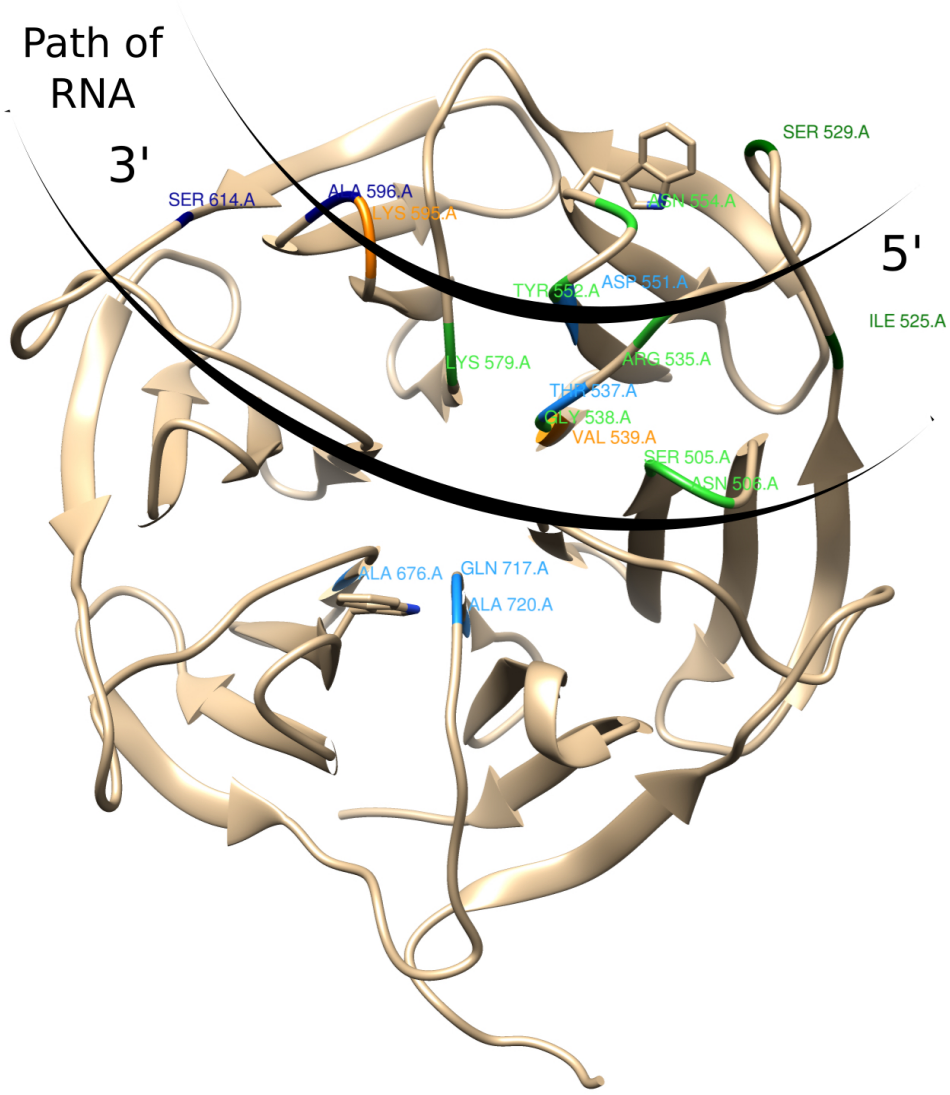


Figure 18: Structure of the TRIM2-NHL domain with proposed path followed by bound RNA. Residues of interest are coloured according to which nucleotide of the RNA their chemical shift perturbations depend on. Residues whose CSPs appear to depend on the first nucleotide of the RNA sequence are coloured in dark green, residues whose CSPs depend on nucleotides toward the 5′ end of the sequence are coloured in lime green, residues whose CSPs depend on nucleotides close to the middle of the sequence are coloured in vivid blue and residues whose CSPs depend on nucleotides at the 3′ end of the sequence are coloured in navy blue. For context, some residues showing an unclear relationship between RNA sequence bound and CSP are displayed in orange.

Two residues' chemical shift perturbations correlate with the 3′ nucleotide

of the RNA sequence bound (see Figure 17b, page 64). The residue showing the clearest pattern is displayed by A596 which only shows significant chemical shift perturbations upon binding of sequences with a 3' terminal uracil. Another noticeable pattern is that displayed by S614 which appears to only show a simultaneous downfield shift in both the ^{1}H and ^{15}N dimension when one of the last two residues is a pyrimidine.

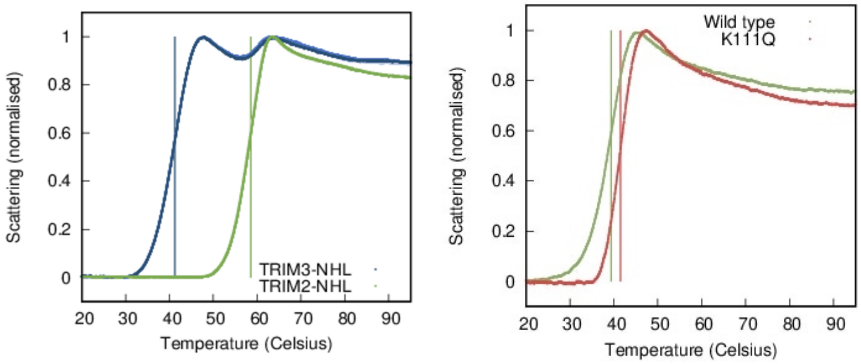
Finally, two residues in this area also show significant CSPs that do not show a clear link to the bound sequence (see Figure 17c, page 64). V539 appears to show consistent small CSPs in both the ^{1}H and ^{15}N dimensions, suggesting the residue's chemical environment may not be affected by the sequence bound. Similarly, all CSPs of residue K595 follow a similar direction but there is not a clear pattern as to which residues show stronger chemical shift perturbations.

Plots analogous to those presented for Figures 16 and 17 but including all residues displaying a significant chemical shift in at least one titration are presented in Figure S15 on page 138.

A lot of these patterns are not particularly robust but when taken together they suggest that the 5' end of the RNA lies somewhere between blades II and III and the 3' end lies between blades IV and V with the center passing close to the center of the NHL domain (see Figure 18, page 65). Moreover, the orientation of the RNA matches that observed for Brat-NHL (although the binding site differs for the 3' end of the RNA).

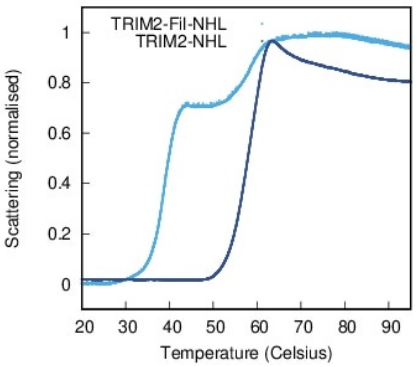
5.7 TRIM2 and TRIM3 NHL domains have differing stabilities possibly linked to interactions with Filamin domains

Despite testing multiple constructs for TRIM3-NHL (constructs tested all ended at the end of the protein sequence, 744, and started at positions 444, 455, 461 , 464, 465, 466, 467, 468 and 469 with a start position of 464 finally being selected), I was unable to find one with sufficient stability to carry out similar experiments to those performed on TRIM2-NHL. The melting temperature of TRIM3-NHL is considerably lower than that of TRIM2-NHL at around 40°C compared to around 60°C (there is batch to batch variability of up to 2°C in our measurements, see Figure 19a, page 67). This is not a particularly low melting temperature, being relatively close to the 45°C limit past which a majority of proteins will successfully crystallise (Dupeux et al., 2011). However, even when stored on ice the protein had a tendency to precipitate over the course of a few hours, I hypothesised that this is due to slow unfolding. This was surprising as the two domains have 81.4% sequence identity (and 94.3% similarity using a BLOSUM80 matrix), which lead us to assume their behaviours would be relatively similar (see Figure 25, page 77).

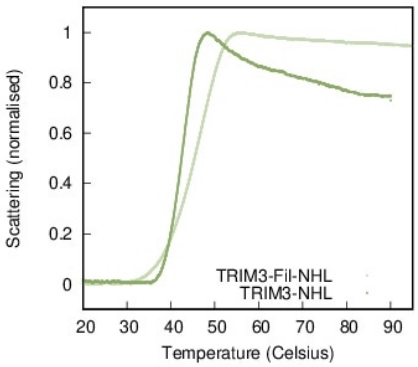


(a) Scattering at different temperatures TRIM2 and TRIM3

(b) Scattering at different temperatures TRIM3-NHL wild type and mutant



(c) Scattering at different temperatures TRIM2-NHL and TRIM2-Filamin-NHL



(d) Scattering at different temperatures TRIM3-NHL and TRIM3-Filamin-NHL

Figure 19: Scattering measurements during NanoDSF experiments for a variety of constructs.

In order to attempt to explain this disparity we partnered with Thomas Hoffman to run MD simulations of the TRIM2 and TRIM3 NHL domains. As outlined above these were carried out using our TRIM2-NHL crystal structure and a TRIM3-NHL homology model based on the crystal structure and consisted of three MD simulations of 1μs for each domain. The first half of each simulation was then discarded and the three simulations were combined to give an equivalent of 1.5μs simulated. In the case of the wild type domains simulations were carried out at 298K, 323K and 373K in order to capture less frequent events which would not occur on time-scales of the order of 1.5μs at 298K. We then analysed differences in residue interactions (using a residue interaction network analysis approach) and RMSF. At 298K differences in RMSF between the two

simulations were relatively subtle but at 373K there appeared to be a substantial increase in flexibility in TRIM3-NHL relative to TRIM2-NHL around the c and d strands of blade I (see Figures 20 and 21 on pages 68 and 69 respectively). This suggests that there may exist a mechanism whereby lower stability of blade I leads to an overall lower stability of the domain.

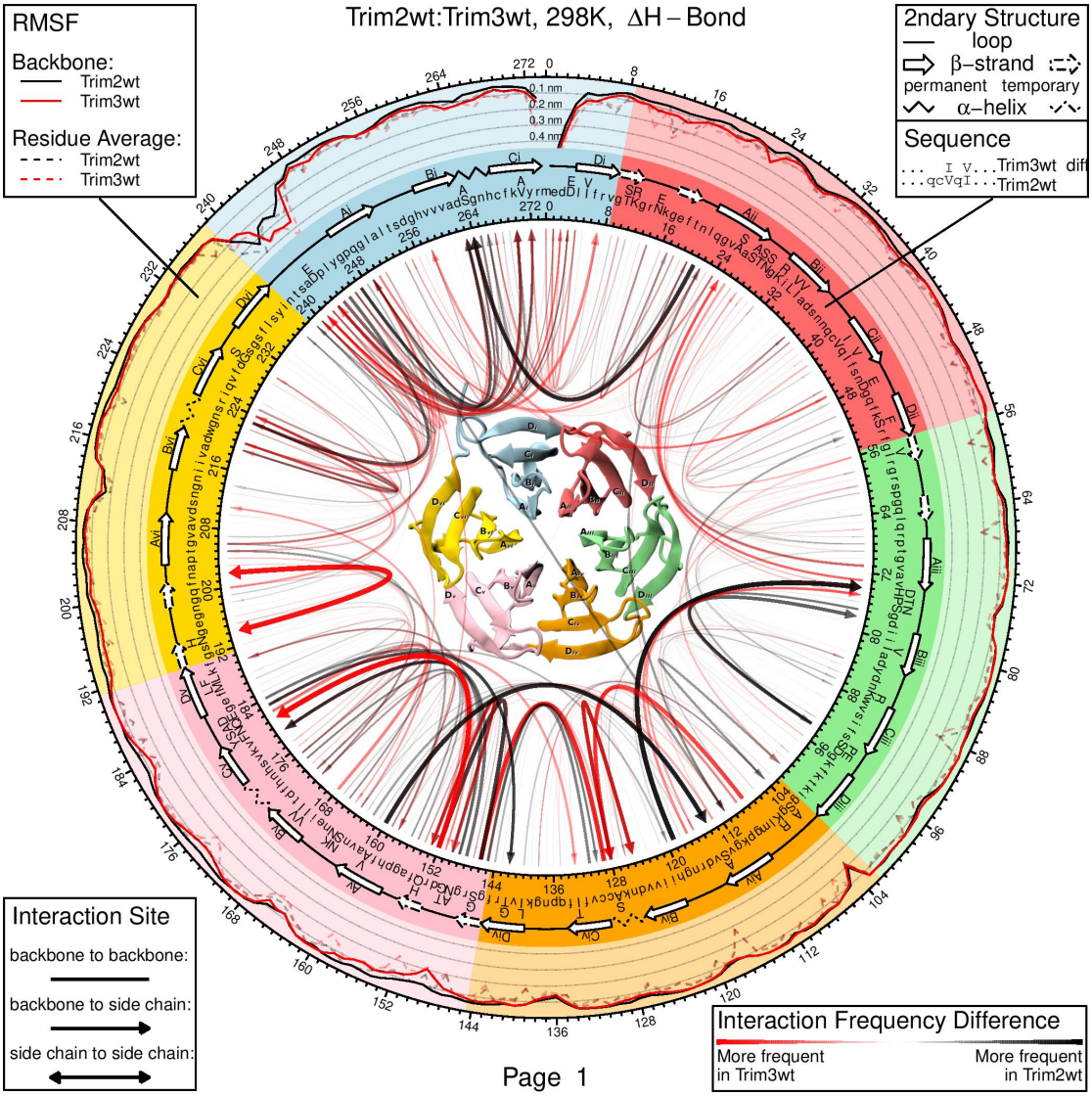


Figure 20: Comparison of MD runs for TRIM3-NHL wild type and TRIM2-NHL wild type at 298K. Structure in the center is represented from the bottom. Sequence numbering matches the domain, not the full length sequence. The inner sequence corresponds to TRIM2-NHL. Arrows represent differences in hydrogen bond frequencies.

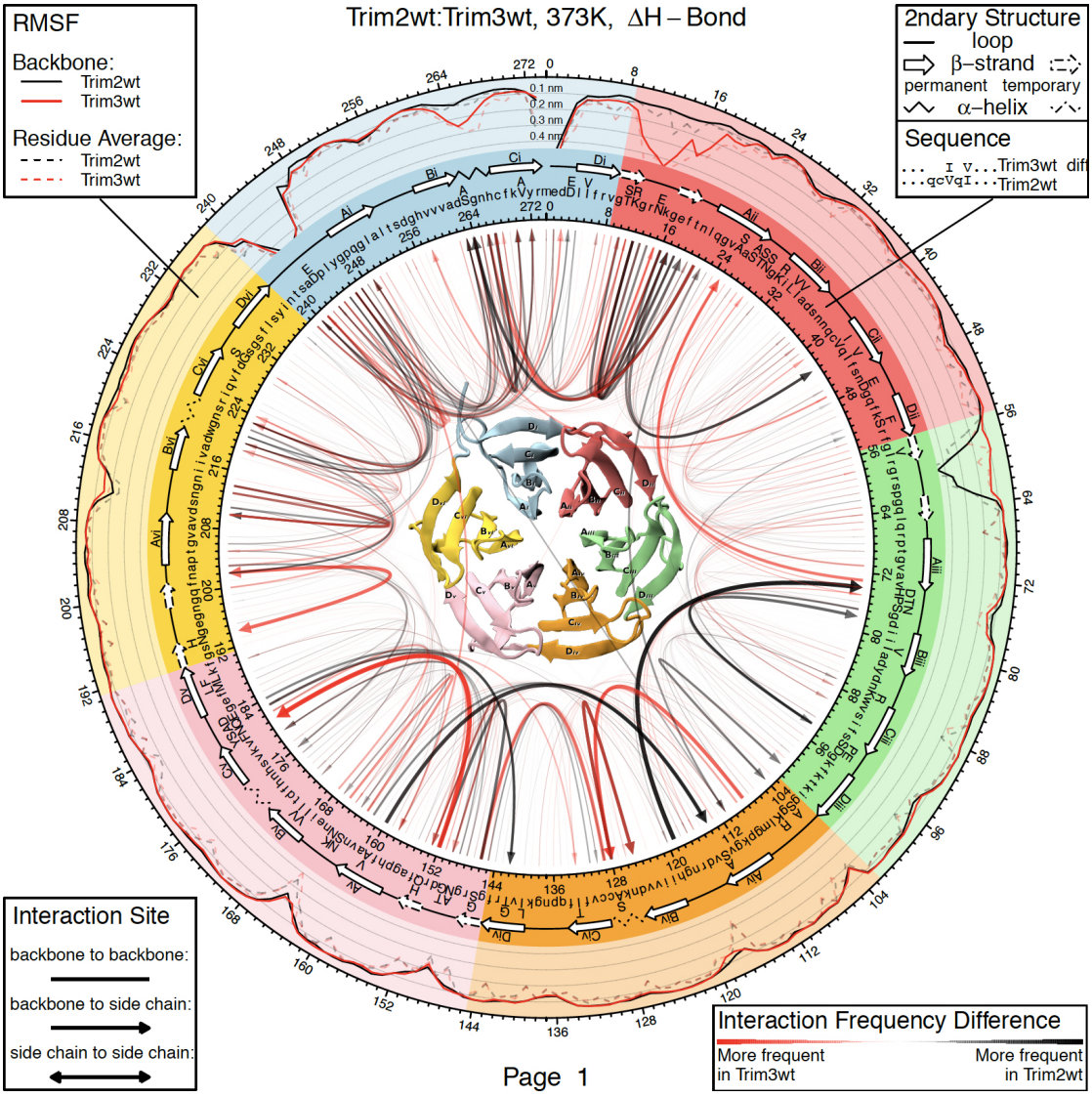


Figure 21: Comparison of MD runs for TRIM3-NHL wild type and TRIM2-NHL wild type at 373K. Structure in the center is represented from the bottom. Sequence numbering matches the domain, not the full length sequence. The inner sequence corresponds to TRIM2-NHL. Arrows represent differences in hydrogen bond frequencies.

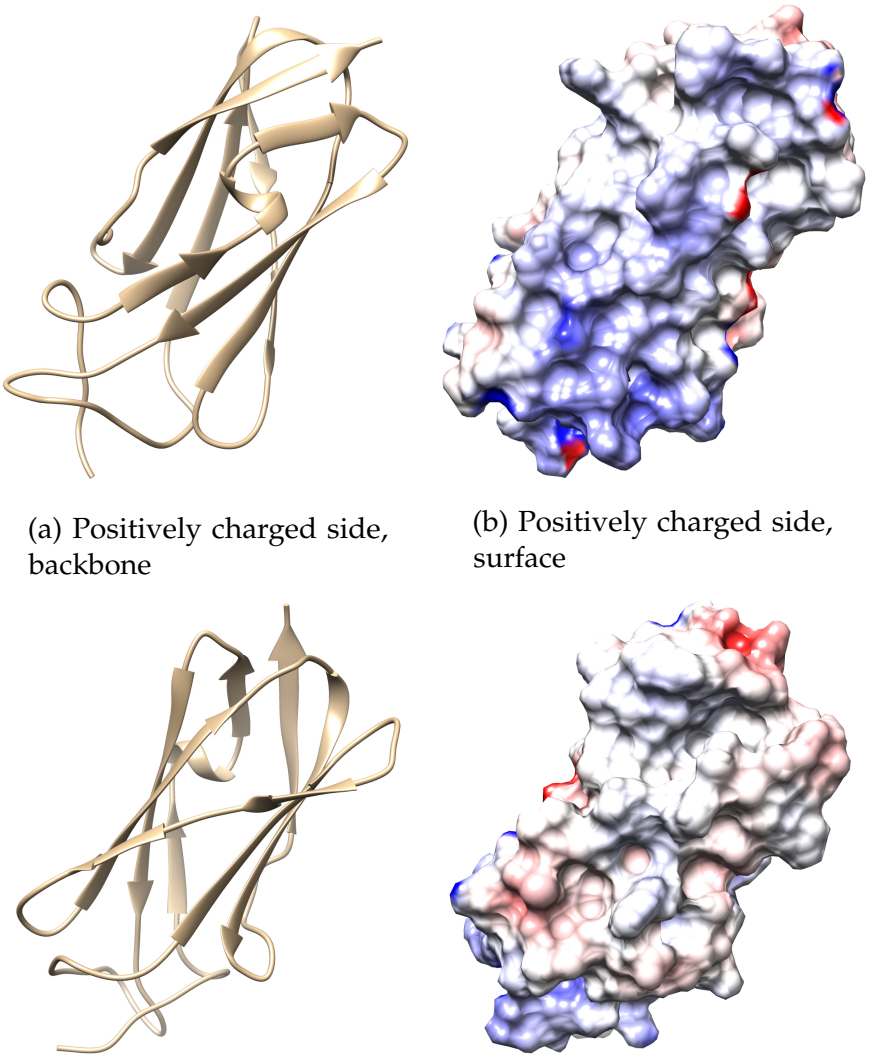
Based on this data as well as a look at which non-conservative mutations were likely to be disruptive, I attempted to stabilise the TRIM3-NHL domain by mutating selected residues towards the TRIM2-NHL sequence, towards sequences present in other organisms, or based on an analysis of residue interactions. Mutations tested were Q490E, K579Q, K579R, A582S, V628A, A732S, and combinations thereof. These produced small improvements to the stability on the order of 2°C

but did not produce sufficient improvements to solve the precipitation issue (see Figure 19b). Most notably mutations A732S and K111Q produced a within-batch stabilisation of 1.5 °C and 2°C respectively, MD simulations were carried out for these interesting mutants in order rationalise their effect which is discussed in more detail in Section 6.5, page 81. Unfortunately, the A732S-K111Q double mutant only increased stability by 2°C, indicating that the stabilisation effects displayed by the two mutants are not cumulative.

I also attempted to make several circularisation constructs of TRIM3-NHL using the split intein methodology described in Waldhauer et al. (2015). This was thought to be a promising methodology as, first of all, the NHL domain is characterised by a close proximity of the N and C-termini requiring only short linker insertions. And second of all, given that blade I appears slightly more flexible than the other blades, an "unravelling" mechanism wherein blade I unfolds destabilising other blades seems like a credible unfolding pathway for the NHL domain. We would therefore expect that stabilising blade I by linking its c and d strands would stabilise the NHL domain overall. Unfortunately, a mass-spectrometry analysis of the protein produced showed that the circularisation failed.

I then observed that the melting curve of our TRIM2-Filamin-NHL construct appeared to contain two distinct melting events while our TRIM3-Filamin-NHL construct showed only one melting event with the midpoint of the curve occurring around 5°C higher than that for TRIM3-NHL alone (see Figures 19c and 19d, page 67). This would be consistent with the TRIM3 NHL and Filamin domains forming a stabilising interaction with each other. The TRIM3-Filamin-NHL construct also seemed to show better behaviour during expression and purification of the protein. I do not believe that the presence of the linker alone is responsible for the stabilisation of TRIM3-NHL as constructs of different lengths (up to inclusion of around half the linker) were tested during our initial construct optimisation without showing a pattern of increasing stability with increasing length.

Given the interesting properties of the TRIM3 Filamin-NHL construct I attempted to crystallise various filamin-NHL constructs for TRIM2 and TRIM3 as well as the filamin domains of TRIM2. However crystallisation of these constructs was unsuccessful. I successfully crystallised the filamin domain of TRIM3 from which I obtained a dataset to a resolution of 2.13Å and built a model of the structure (R_{work}/R_{free}=0.2270/0.2575). Full statistics are presented in Table S3 on page 115. The structure matches known structures relatively closely with a root mean square deviation (RMSD) of 1.47Å between our structure and that determined by Tocchini et al. (2014) for the filamin domain of TRIM71 (see Figure S4, page 121). The structure is, as expected, constituted of two antiparallel β-sheets,



(a) Positively charged side, backbone

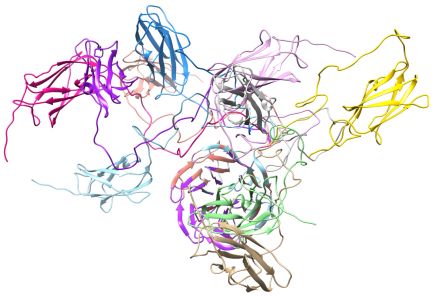
(b) Positively charged side, surface

(c) Opposite side, backbone

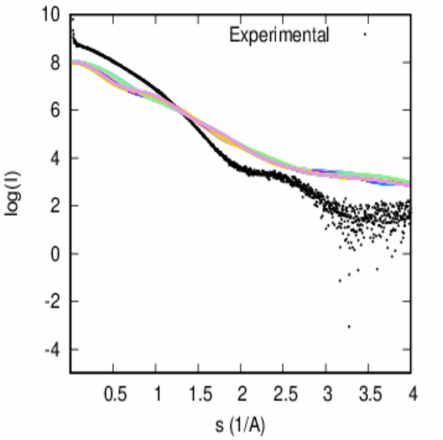
(d) Opposite side, surface

Figure 22: Structure of the TRIM3 filamin domain with calculated electrostatics.

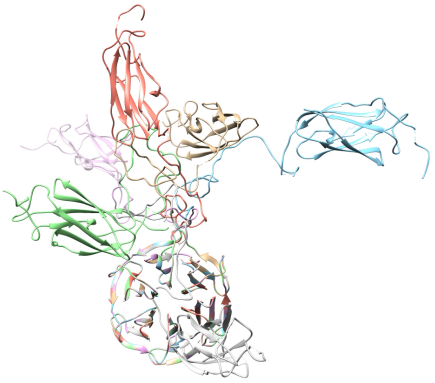
one side of the structure is slightly positively charged (see Figure 22, page 71) which could mediate interactions with the negatively charged bottom surface of the NHL domain or with RNA. In order to further characterise the interaction between the filamin and NHL domains in solution, I also acquired SAXS data for all NHL, filamin and filamin-NHL constructs as well as ^{1}H-^{15}N-HSQC data for each of these. These were of somewhat variable quality owing to the differing stabilities of the domains studied (see Figures 23 and 24, on pages 72 and 74 respectively).



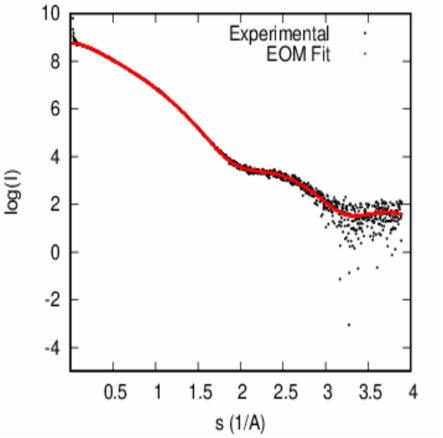
(a) Structures of best 10 individual fits to SAXS data



(b) Best 10 individual fits to SAXS data



(c) Structures making up the EOM fit of the data



(d) EOM fit of mixture of structures to SAXS data

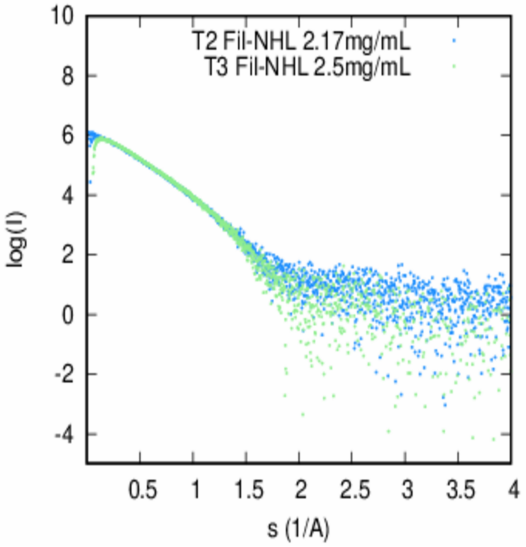
Figure 23: Structures and fits to TRIM2-Filamin-NHL SAXS data for individual structures (**top**) and a mix of structures assembled by EOM (**bottom**).

In line with the NanoDSF results, our SAXS data also seems to indicate a lack of interaction between the filamin and NHL domains of TRIM2. Based on our structure of the TRIM2-NHL domain, as well as a homology model based on our crystal structure of TRIM3-filamin we generated 200 structures by randomising the linker between the NHL and filamin domains. None of these structures fit the data well, with the best χ^2 being equal to 70 (see Figures 23a and 23b, page 72), whereas a mixture of 6 structures generated using EOM (Tria et al., 2015) provided a good fit with a χ^2 of 5.2, indicating that a mixture of conformations exists in solution and therefore no strong interaction between the

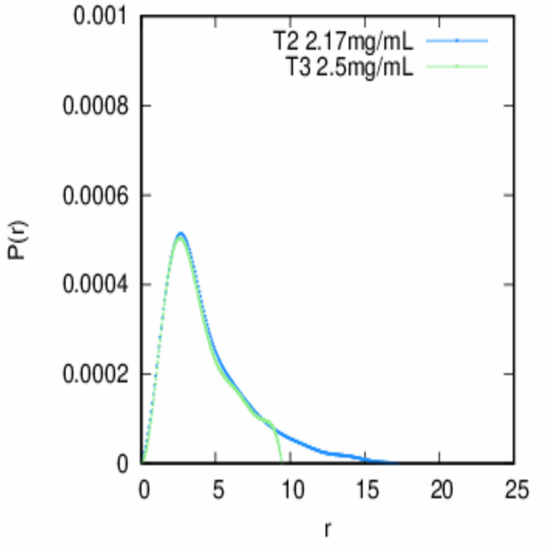
two domains exists (see Figures 23c and 23d, page 72). Weak interactions may be still be present given that the average R_g of the single structures generated (37.6 Å) is slightly larger than the R_g we can calculate from our SAXS curves (33.4 to 34.9 Å). This indicates that the structure may be slightly more compact than we would expect given independent domains. See Table S5, page 116 for the SAXS measurement statistics, reported according to guidelines (Trewhella et al., 2017).

However, recently acquired SAXS data from the TRIM3-Filamin-NHL construct indicates that it does not substantially differ from the TRIM2-Filamin-NHL in the level of its domain interaction with both the SAXS curves and P(r) curves having a very similar shape over a range of concentrations (see Figures 24a and 24b). This suggests that any difference in the interactions between the NHL domain and the Filamin domain must be linked to short term interactions rather than the formation of a stable domain-domain complex. These interactions, if they exist, may be detectable in ^{1}H-^{15}N-HSQC spectra of the different constructs.

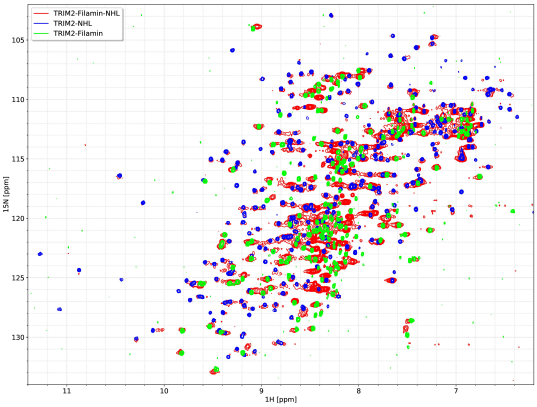
As these interactions would constitute a change to the chemical environment of each domain they should therefore change the chemical shift of participating residues. Conversely, in the case of a perfect lack of interaction we would expect the Filamin-NHL construct spectrum to be a simple sum of the NHL spectrum and the Filamin spectrum. I have acquired spectra for each of these constructs for both TRIM2 and TRIM3, as can be seen in Figures 24c and 24d on page 74. There is no obvious difference between the TRIM2 and TRIM3 constructs in the number of peaks whose chemical shift has changed significantly. However, a more detailed analysis may help us find relationships once we have accounted for issues such as the presence of a linker affecting residues in close proximity and the differing quality of the spectra used. To this end assignment of the TRIM2 and TRIM3 filamin domains was undertaken, with assignment of TRIM2 Filamin being complete at 88.8% of non-proline residues assigned while that of TRIM3-Filamin is ongoing with only 58.9% of non-proline residues assigned.



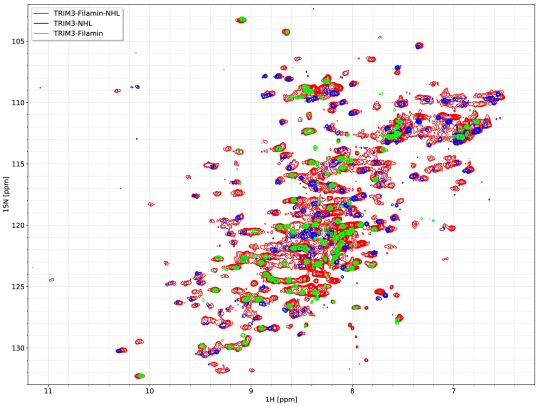
(a) Buffer subtracted SAXS data for TRIM2 Filamin-NHL (**blue**) and TRIM3 Filamin-NHL (**green**)



(b) P(r) plots for TRIM2 Filamin-NHL (**blue**) and TRIM3 Filamin-NHL (**green**)



(c) HSQC spectra for TRIM2 NHL (**blue**), Filamin (**green**) and Filamin-NHL (**red**) constructs



(d) HSQC spectra for TRIM3 NHL (**blue**), Filamin (**green**) and Filamin-NHL (**red**) constructs

Figure 24: Comparison of the TRIM2-Filamin-NHL construct and the TRIM3-Filamin-NHL construct by SAXS and NMR. Larger versions of the NMR spectra may be found in Figures S5 and S6, on pages 122 and 123 respectively.

6 Discussion

6.1 Choices relating to the analysis of chemical shift perturbation differences

The analysis of chemical shift perturbations that I carried out in order to try to determine the orientation of the RNA is not a standard method when analysing the interaction of two macromolecules, although similar methods are often used in the case of fragment screening during fragment based drug design (Shuker et al., 1996). It therefore seems useful to go over some of the assumptions inherent in the approach and some of the practical choices made when analysing the data.

The basic idea behind the method is that, as different RNAs are different ligands and therefore constitute a different chemical environment when bound, they will result in different chemical shift perturbations even in the fully bound state. In order to extract an orientation from this observation we do, however, have to assume that two types of effect remain local. First, we assume that changes to a certain base of the bound RNA do not significantly change the binding conformation of distant bases. If this were not the case, even if we could pinpoint a specific change in RNA sequence as causing specific differences in CSPs, there would be no way to associate this with the position of the mutated nucleotide on the surface, as large changes to the binding conformation would occur for all nucleotides. Second, changes in binding near a specific residue must produce changes in CSPs mostly at or near that residue. If, for example, binding of a specific RNA sequence were to cause significant conformational change this would translate to chemical shift perturbations at many locations throughout the protein. Fortunately, in our case both these assumptions are likely to mostly hold true as single stranded RNA is a quite flexible macromolecule (Isaksson et al., 2004) and our domain is quite rigid. One advantage of our approach is that we don't seek to explain changes in CSPs or interpret them in isolation. For example, if we were to change a base at a certain location and see an associated increase in the CSP of a given residue one could interpret that in multiple ways: the new base could form stronger interactions with the residue in question meaning these interactions are more frequent giving a larger CSP, conversely the base may fit the protein surface less well and may need to displace the residue further in order to bind. Our analysis avoids the need to distinguish cases such as these.

The next question is the practical aspects of how this analysis was carried out: first of all I decided to attempt to analyse the chemical shift perturbations in an approximately 50% bound state rather than attempt to calculate the $\Delta\delta_{max}$ for each residue, which would have yielded larger chemical shift perturbations which may, in turn, have been easier to analyse and would be less sensitive

to small variations in the peak position in a given spectrum. I did this due to concerns that calculation of the $\Delta\delta_{max}$ would not be sufficiently accurate, especially for residues with small chemical shift perturbations, it may however be interesting to attempt this approach in future. Second, I decided to use the closest point in my titration curves to 50% binding without attempting to correct for the slightly different percentage binding actually measured (see Table S6, page 119). This approach was chosen because any correction calculation would present the same drawbacks as attempting to calculate the $\Delta\delta_{max}$ and would be relatively difficult to implement in my titration analysis programs. It should be noted that although for most sequences the point closest to 50% binding falls within a range from 45% to 53% some of the weak binders do not reach these ratios. However, given that the differences in CSP magnitude discussed are mostly large relative to the differences in percentage bound RNA, all RNAs fall well within an acceptable range, with the notable exception of AAA AAA, which is a weak binder delivered in low quantities by the manufacturer. Finally I decided to manually interpret the differences in chemical shift perturbations. This approach is, of course, very open to biases in interpretation especially given the subtle differences often being observed. The low number of sequences tested and their bias towards certain types of higher affinity sequence probably makes any inferential statistical analysis impossible. However, although direct manual comparison of the CSPs has the advantage of allowing one to consider both magnitude and direction information it may allow for a more quantitative analysis if this information were collapsed into a single statistic, in future I would be interested in experimenting with the dot product of CSPs to attempt to address this.

6.2 Comparison of RNA binding in TRIM2-NHL and other NHL domains

The preference for a C rich RNA binding sequence would set TRIM2 apart from the most closely related NHL domains which all have U rich binding motifs: Brat, Mei-P26 and NHL have binding motifs UU[G/A]UU[G/A], UUUACA, and poly-U respectively (Davis et al., 2018; Loedige et al., 2015). Moreover, as discussed in Section 5.5 (page 57) the RNA binding site of Brat differs from that I propose for TRIM2-NHL inasmuch as the RNA runs from between blades II and III to between blades VI and I for Brat and from between blades II and III to between blades IV and V for TRIM2.

The very different RNA binding, both in terms of favoured sequences and binding site, of TRIM2-NHL and Brat-NHL was relatively unexpected as these proteins are usually considered orthologues. On the other hand, this is consistent

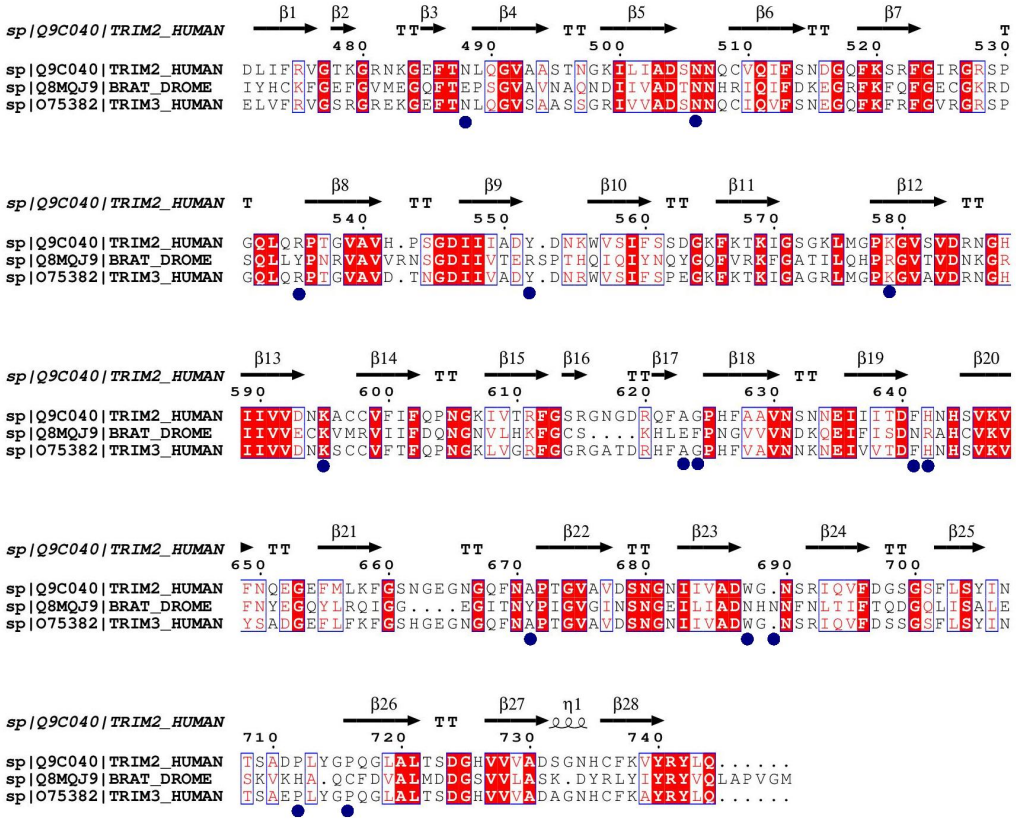


Figure 25: Alignment of TRIM2, TRIM3 and Brat NHL domains with secondary structure and residues determined to be RNA binding by Loedige et al. (2015) labelled with a blue dot.

with the very low degree of conservation of the RNA binding residues found by Loedige et al. (2015). Of those 15 residues only two, N800 and K891, corresponding to N506 and K595 in our sequence, are conserved in the TRIM2-NHL sequence. While in two cases the changes in sequence are conservative, these being R875 and K934 which correspond to K579 and H642 in our sequence (see Figure 25, page 77). It should be noted that of these four conserved and semi-conserved residues, three are in the binding site hypothesised in this work and undergo significant chemical shift perturbations upon RNA binding (see Figure 18, page 65). Kumari et al. (2018) already commented upon the surprising lack of conservation of the top surface of NHL domains in the subfamily of proteins containing Brat, TRIM2 and TRIM3. Their explanation for this is a hypothesised divergence in function but Brat, TRIM2 and TRIM3 actually have

several specific functions in common, for example both TRIM3 and Brat suppress Myc expression (Chen et al., 2014; Harris et al., 2011) and in the case of Brat this happens through post-transcriptional regulation involving the NHL domain (Ferreira et al., 2014; Harris et al., 2011). Given this, as well as the conservation of a few RNA binding residues, it seems possible that both the binding site and the sequences recognised have undergone a considerable degree of change over evolutionary time while maintaining the same underlying functions. In order to disprove or bring supporting information to this hypothesis a more detailed understanding of the exact molecular mechanisms underlying these roles would be necessary (most notably whether direct mRNA binding occurs in both cases), as well as residue level information concerning the binding site for several more Brat orthologues.

6.3 RNA binding: remaining questions and alternative hypotheses

It is unclear at this stage whether there exist other single stranded RNA sequences with significantly higher affinities to TRIM2-NHL than those we tested. Our sequences are weak binders compared to the literature on NHL domains where affinities of the order of 40-400nM are reported (Davis et al., 2018; Loedige et al., 2015), although these values are for 15 and 17-mer RNAs and were estimated by different methods (EMSA and fluorescence anisotropy) so they may not be directly comparable. A better direct comparison would be the isothermal calorimetry (ITC) measurements reported by Macošek et al. (2020) which showed a K_d of 1.4 μM for the interaction between a UUGUUG 6-mer and Brat. On the one hand, the strongest binding sequences I have so far tested (characterised by a C rich center and 3' end and one or two purines at the two positions nearest the 5' end) all have relatively similar K_d in the 90μM range, bind approximately 10x more tightly than some less preferred sequences, and resemble the sequences most enriched in our RIP-seq data, suggesting they may be close to the ideal binding motif. This would be consistent with a consensus sequence of the form [G/A/C][G/A/C]CCNC (the 5^{th} position has not been tested thoroughly enough to truly determine a sequence preference). On the other hand, I have been unable to establish a K_d for very CG rich sequences as they either cause unexplained signal loss or are unsuitable for titrations due to their propensity for dimerisation and G quadruplex formation. It should also be noted that although the binding affinities given for our best binding RNA 6-mer are relatively low, TRIM2 and TRIM3 form dimers and synergistic RNA binding of both NHL domains simultaneously could greatly increase both affinity and specificity (as high affinity binding would require the presence of two binding sites instead

of one). This is a mechanism of growing importance when interpreting RNA-protein interactions (Hennig et al., 2014; Hollmann et al., 2020; Jankowsky and Harris, 2015). It should be noted that none of our analyses so far would allow us to detect the presence of a double binding site as DREME can only find short non-gapped motifs, it may therefore be interesting to rerun our analysis with a tool capable of detecting gapped motifs (Bailey et al., 2015). Alternatively, affinity could be increased through involvement of additional domains as seen in Haubrich et al. (2020) and discussed more at length below.

The high degree of rigidity of the NHL domain top surface was an unexpected observation and could possibly be linked to the strong disfavouring of non cytosine residues at certain positions. It may be interesting to compare this to the RNA binding surface of more permissive RNA binding domains (RBDs).

I have so far restricted our analysis to RNAs without secondary structure as I expected TRIM2 and TRIM3 to bind similar sequences to Brat but structure could play a role in RNA binding preference, especially given the recent work by Kumari et al. (2018) on TRIM71. Although it should be noted that none of the TRIM2 or TRIM3 orthologues tested as part of that study showed a preference for structured RNA (see Figure S12, page 129). In order to eliminate these possibilities I would propose a shift to ITC experiments given that the RNAs tested would by necessity be larger, increasing the likelihood of signal loss through a variety of mechanisms (slower tumbling, binding of multiple proteins to a single RNA, etc.). If TRIM2-NHL did show a preference for stem loops the preferred stem loop sequence would probably be at least somewhat correlated to the preferred single stranded sequence so CG rich stem loops and dsRNA should be tested.

In order to confirm our RNA binding site several mutants at and around the binding site have already been designed (R535A, N554A, K555A, K579A N594A and K595), testing these should confirm our hypothesis concerning the binding site. The next obvious step is to carry out crystallisation trials with one of our better binding RNAs. Finally, we can get a more detailed understanding of how exactly our RNA binds by using our CSPs as docking restraints in a program such as HADDOCK (van Zundert et al., 2016) or, if we wish to keep our analysis as independent as possible, by applying a specialised ssRNA binding computational modelling tool such as that described by Pal and Levy (2019).

6.4 Interaction of the NHL domain with other domains

So far our data indicates that interactions between the NHL and filamin domains of TRIM2 and TRIM3 are either absent or transient. This can be compared to the findings of Kumari et al. (2018) when studying TRIM71 in which the NHL and filamin domain appear to interact, albeit weakly, and with some flexibility in

the relative orientation of the domains. This difference may, however, be caused either by crystal packing or the difference in linker length between TRIM71 which has a 10 residue long linker and TRIM2 and TRIM3 which have linker lengths of around 45 to 50 residues (there are 3 extra residues in the TRIM3 linker).

Given that these results suggest the NHL and filamin domains remain flexible relative to each other we can ask ourselves whether the C-terminus is flexible relative to the rest of the protein. So far we do not have much evidence in either direction but looking at the known interactions of our domains as well as other TRIM proteins can give us clues as to how C-terminal flexibility (internal as well as relative to the rest of the protein) could be modulated.

It is currently unknown if RNA binding would cause interactions between the NHL and filamin domains to form. It is possible that the filamin domain forms interactions with RNA although so far no such interaction has been reported. Now that we have a good idea of what sequences bind TRIM2-NHL it will be relatively easy to extend such RNAs and titrate them against constructs containing both filamin and NHL domains. The behaviour of TRIM2 and TRIM3 under these conditions may also differ as the first has an overall negatively charged filamin domain (pI 5.04 at domain boundaries equivalent to those of TRIM3) while the second has a positively charged filamin domain (pI 9.29), suggesting that TRIM3's filamin domain may be more likely to form interactions with RNA than that of TRIM2.

Similarly, by analogy with the work of Haubrich et al. (2020), we may see interactions between the C-terminal domains and the coiled-coil, especially upon RNA binding. Given the long flexible linker present between the NHL domain and the filamin domain, simultaneous RNA binding by the C-terminal domains and coiled-coil regions of the protein, as observed by Haubrich et al. (2020) in the case of TRIM25, would likely cause the structure to become significantly more compact, thus bringing any NHL bound ligand closer to the RING and B-box domains and thereby providing a clear mechanism for ubiquitination modulation. Indeed, we have preliminary data suggesting this may be the case in the form of a titration of the coiled-coil domain against TRIM2-NHL where addition of the coiled-coil to the NHL domain resulted in signal loss, suggesting that there may be an interaction between these two domains (see Figure S2, page 113). Unfortunately, the construct tested showed poor stability making it difficult to interpret this data as it is unclear if the interaction occurs with the correctly folded coiled-coil domain or with unfolded aggregates. Different constructs are currently being tested.

RNA binding could, of course, also bring RNA bound substrates into the vicinity of the RING domain, thereby targeting them for ubiquitination. Alterna-

tively, RNA could recruit TRIM2 and TRIM3 to the correct environment for them to form interactions with their ubiquitination targets, much like the reported segregation of TRIM71 to P-bodies. It is of course possible that the ubiquitination and RNA binding roles of these proteins are completely separate. However, as laid out in the introduction, the high degree of conservation within the NHL family of RNA binding and ubiquitination modules within the same protein strongly suggest that these are functionally linked (although it should be noted that Brat presents a counter-example to this).

Finally, TRIM2 and TRIM3 are known targets for post-translational modifications which could modulate inter-domain interactions. First of all there is the reported phosphorylation of the filamin domain which causes TRIM2 to bind and ubiquitinate AXL (Rao et al., 2020). The exact structural mechanism through which filamin domain phosphorylation at the serine 443 position leads to increased AXL binding, as well as whether additional mechanisms need to come into play in order for AXL to be brought into proximity with the RING domain, remains unknown. However we are in a position to relatively easily produce and characterise phosphomimetic mutants of TRIM2 filamin and filamin-NHL constructs. Second, auto-ubiquitination is a common feature of TRIM proteins (Choudhury et al., 2017; Diaz-Griffero et al., 2006), which preliminary experiments conducted in our lab suggest is also a feature of TRIM2 and TRIM3, which could affect the domain interactions of these proteins through some, as of yet, unknown mechanism.

6.5 TRIM3 NHL domain stability

As discussed in Section 5.7, the NHL domain of TRIM3 is considerably less stable than that of TRIM2. This is unfortunate as a comparison of the RNA binding behaviour of TRIM3-NHL with TRIM2-NHL could have been instructive as to whether differences in the two TRIM proteins' function could be caused by differences in their RNA interactions. It is still unclear what exactly causes this difference in stability given the high level of sequence similarity between the two proteins (81.4% sequence identity and 94.3% similarity) various approaches having failed to produce stabilising effects anywhere near the difference in stability observed. The MD simulations for TRIM3-NHL suggest an "unravelling" mechanism, wherein an unstable blade I unfolds and destabilises the entire domain, may be responsible. It is unfortunate that our circularisation methods did not succeed as this should in theory be a good way to stop unfolding of this type. Mass-spectrometry analysis of our products using this method confirmed that they were not the circularised protein expected but did not allow us to confirm what was produced, making it hard to devise an optimisation strategy.

The mutation strategy produced slight stabilisation effects which can be

explained using our MD data. I shall now briefly describe our proposed stabilisation mechanism for the two best mutations we have found so far: K579Q and A732S.

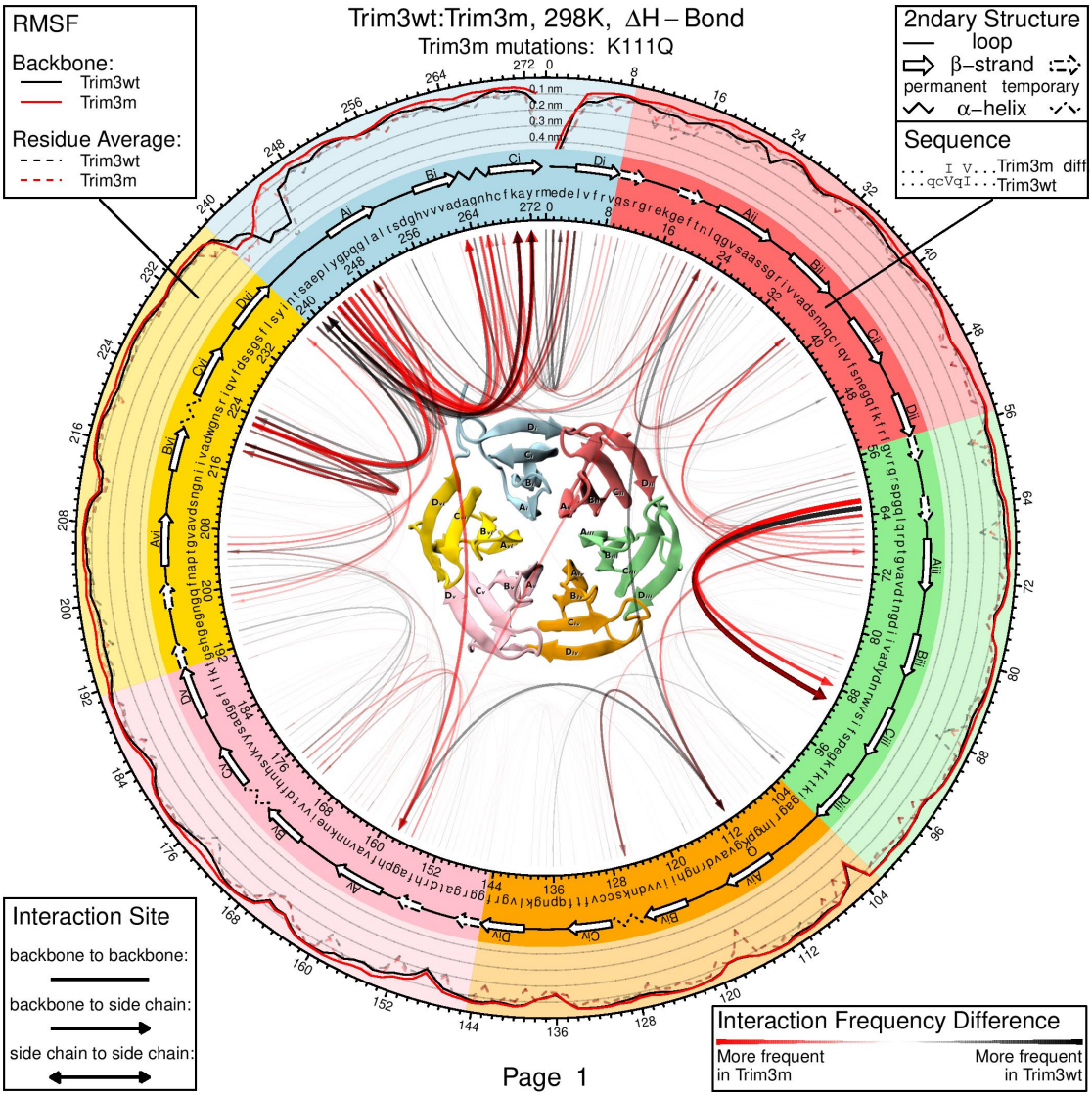


Figure 26: Comparison of MD runs for TRIM3-NHL wild type and the K579Q mutant. Structure in the center is represented from the bottom. Sequence numbering matches the domain, not the full length sequence.

We can rationalise the stabilising effects of K579Q (referred to by its domain numbering K111Q in Figures 26 and 19b) using the MD data. This mutation affects the hydrogen bonding network on the top surface surrounding the centre of the NHL domain as illustrated in Figures 26 and 27 on pages 82 and 27

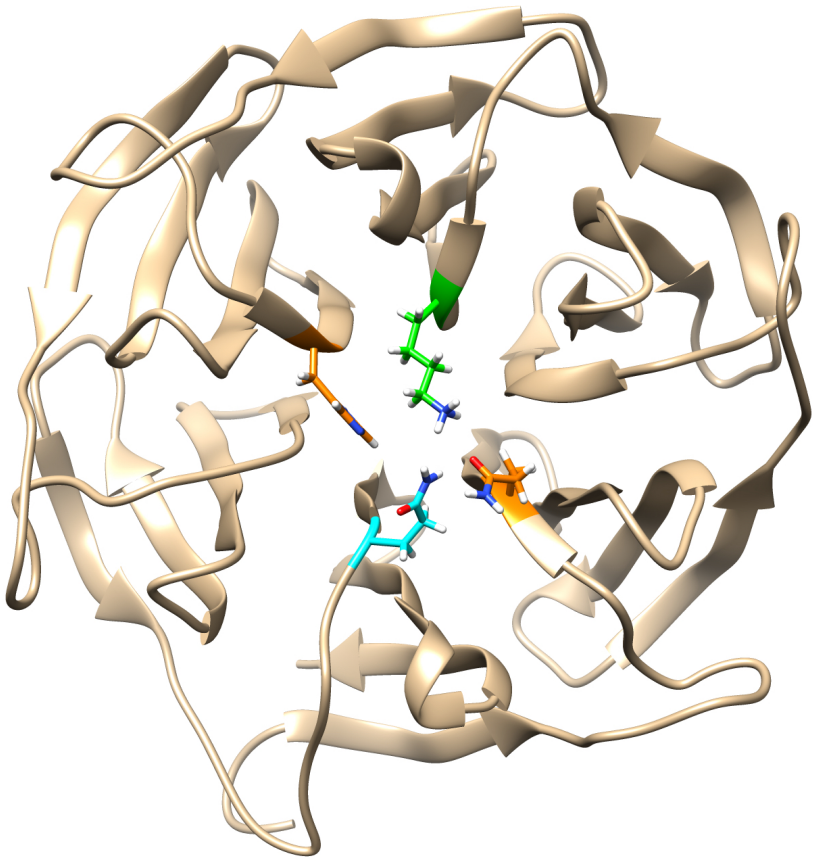


Figure 27: TRIM2-NHL structure with relevant sidechains highlighted (they are conserved with TRIM3-NHL) including residue K579 (numbered 111 in Figure 26 and coloured green) and those of Q490 and H626 (numbered 22 and 158, respectively, in the previously mentioned figure and coloured orange) and Q717 (numbered 249 in the figure and coloured light blue).

respectively. If we look at a comparison of the frequencies of hydrogen bond interactions in our simulations we see that the mutation appears to reduce interactions between the side-chain of residue 579 (numbered 111 in the figure and coloured green) and those of Q490 and H626 (numbered 22 and 158, respectively, in the figure and coloured orange), freeing the latter up for more frequent interactions with Q717 (numbered 249 in the figure and coloured light blue) which appears to have a stabilising effect on blade I as a whole, decreasing the RMSF of the blade VI-blade I linker and affecting the entire hydrogen bonding network of the region.

Similarly, the mutation A732S (numbered A264S in Figure 28, page 84) in-

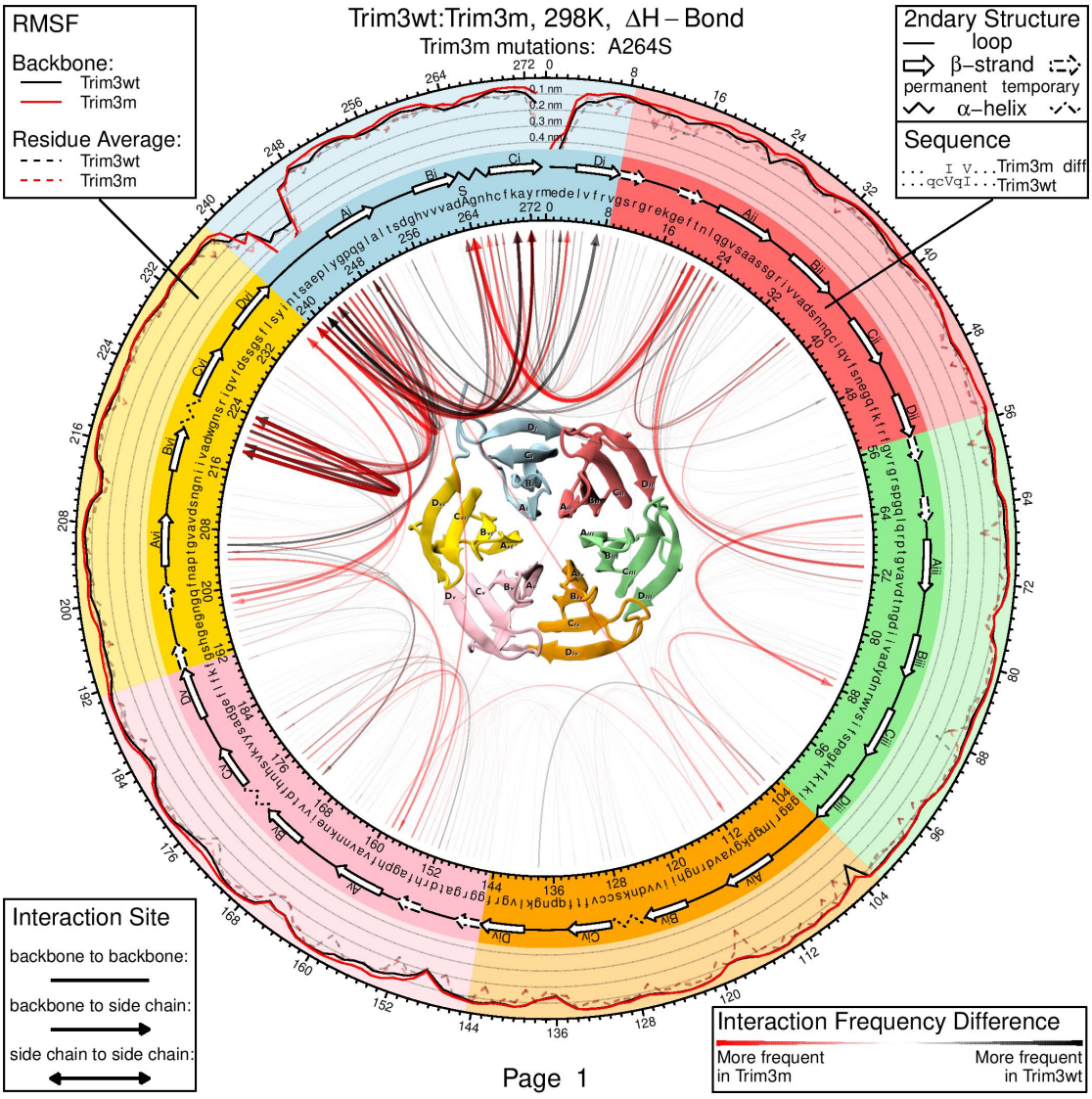


Figure 28: Comparison of MD runs for TRIM3-NHL wild type and the A732S mutant. Structure in the center is represented from the bottom. Sequence numbering matches the domain, not the full length sequence.

creased the protein's melting temperature by 1.5°C. This is particularly interesting given that this sequence variation is present in the TRIM3-NHL of some species (in fact, non-mammalian vertebrates mostly have a serine at position 732), suggesting that it may not have much of an effect on sequence binding. We supposed that this is due to it's ability to form a hydrogen bond to the backbone of residue L489 (numbered L21 in the figure) as is the case in TRIM2-NHL (see Figure 29, page 85). This corresponds to the output of our MD simulations

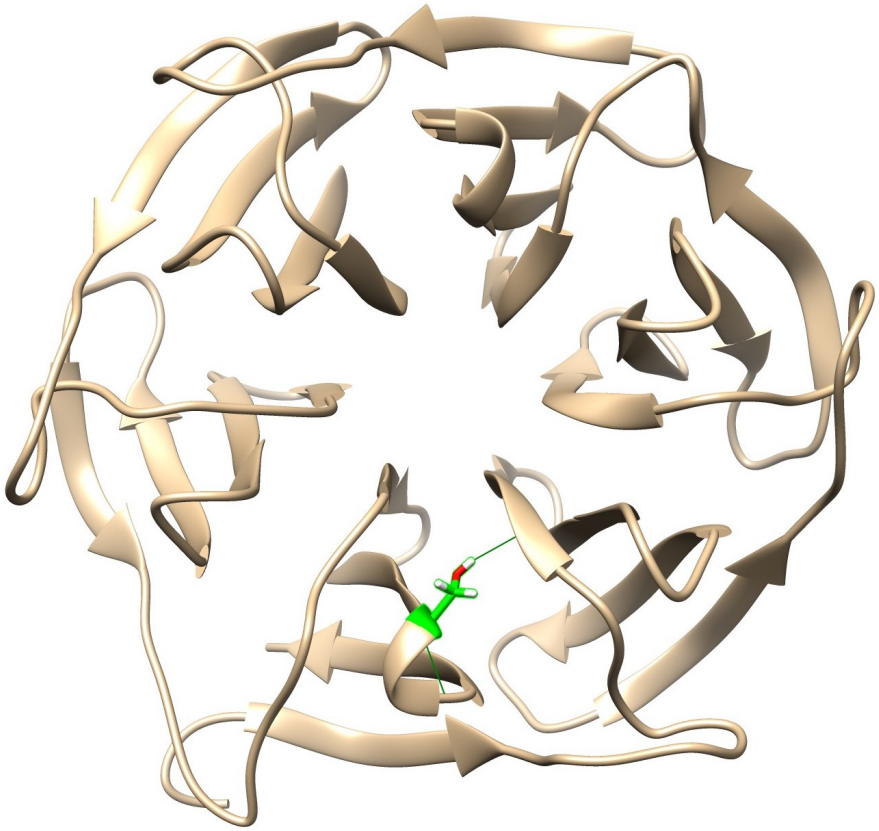


Figure 29: TRIM2-NHL structure with the sidechain of A732S highlighted in green and hydrogen bond formed to L489 displayed.

that suggest that in the mutant the hydrogen bond is formed and has a small stabilising effect on blades I and II (as evidenced by slightly lower RMSF values across this region) although it may destabilise the blade VI-blade I linker by some mechanism.

All of these interpretations rest on molecular dynamics simulations but could perhaps be confirmed if the stabilising effects of the addition of the filamin domain and these promising mutations are cumulative (unfortunately, the stabilising mutations so far tested are not cumulative with one another).

The stabilising effects of the filamin domain appear small when considering improvements in melting temperature but did qualitatively improve our protein's behaviour. Although it is clear from our SAXS data that this cannot be mediated by stable interactions between the NHL and filamin domain, our data is compatible with less frequent weak interactions which may have a stabilising effect. This is based on the observation that the R_g we would predict in the

absence of any interaction (37.6Å) is slightly larger than that calculated from our TRIM2-filamin-NHL data (33.4 to 34.9Å) which is in turn larger than that calculated from our TRIM3-filamin-NHL data (31.0 to 31.4Å). This suggests that TRIM3-filamin-NHL is more frequently in a more compact conformation than TRIM2-filamin-NHL which is in turn more compact than we would expect in the absence of inter-domain interactions. For this reason, a full analysis of the filamin-NHL NMR data may yield interesting results. It could also be interesting to check whether, in the absence of mutants that produce cumulative stabilising effects, a combination of filamin domain addition and NHL domain mutations may yield better results.

7 Conclusions and Outlook

In this work I have presented the first characterisation of the RNA binding of TRIM2 and TRIM3. Our data suggests a significant divergence from the sequences bound by studied orthologues as well as a possible shift of the location of the RNA binding site. This could have interesting implications when studying the evolution of TRIM-NHL proteins' RNA binding and how it relates to their biological roles.

I have also found a significant difference in the stability of the NHL domains of TRIM2 and TRIM3, explored some stabilisation strategies, and presented some initial data as to the potential mechanisms underlying this difference in stability. Although significant effort was expended in attempts to stabilise the TRIM3-NHL domain I have so far had only limited success. However, should stabilisation of this domain prove to be important to future projects some basic improvements are now available that may be expanded upon.

Additionally, the first atomic resolution structures for the C-terminal domains of these proteins was presented (filamin domain of TRIM3 and NHL domain of TRIM2) along with a characterisation of the dynamics of TRIM2-NHL with implications as to the role of flexibility in RNA binding in this case. NMR assignment of all stable C-terminal constructs is either complete or well under-way. This places us in a position to move forward rapidly with many open questions concerning these proteins. This includes inter-domain interactions which remains an important aspect of our understanding of TRIM protein function.

8 Acknowledgements

I would like to thank Janosch Hennig for giving me the opportunity to work on this project, for creating a work environment in which I had the opportunity to learn many new techniques from talented people and for giving me the freedom to focus on the elements that I found most interesting.

I would also like to thank the members of my TAC committee: Orsolya Barabas, Claudio Joazeiro and Kyung-Min Noh for their time and regular thoughtful input into this project. I would also like to thank Rob Russell for taking the time to read my thesis and participate in my defence committee.

This project would not have been possible without the participation of many collaborators and lab members, some of the initial work on this project was carried out by Inga Loedige and Jaelle Foot as well two members of the Noh group, Maja Gehre and Weronika Stachera. Later on, discussion, feedback and advice from Cecilia Perez-Borrajero concerning the project was extremely helpful especially concerning reanalysis of the RIP-seq data.

I would also like to thank Pravin Jagtap for his advice on processing X-ray data and model building as well as Lyudmila Dimitrova-Paternoga for advice and help concerning data acquisition. I am also grateful to Bernd Simon for his teaching and advice about NMR and SAXS. Po-Chia Chen was also helpful in his explanations of SAXS.

I would also like to acknowledge the hard work put in by the students, Aysenur Yikilmaszoy and Nikolay Silakov, who participated in this project and would like to thank them for their perseverance in the face of an often difficult project, as well as Brice Murciano who kindly set up crystallisation screens for the sometimes rather numerous constructs they and I produced. Although they only joined the project for a week it was also extremely rewarding to supervise Aoife Hopkins and Junna Lino.

I am very grateful to Sandra Augsten and Pawel Masiewicz for their constant help, support and advice, as well as to my fellow PhD students Jakub Macosek, Nele Hollmann, Kevin Haubrich and Andrea Lomoschitz for their many scientific discussions and keeping the lab a fun place to work.

Finally I would like to thank my friends and family for the emotional and practical support (thanks especially to Matt Tarnowski on that front) they provided over the course of this PhD, most of all my loving wife, Olivia.

References

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., and Grosshans, H. (2017). Lin41 post-transcriptionally silences mRNAs by two distinct and position-dependent mechanisms. Mol Cell *65*, 476–489.

Ahlner, A., Carlsson, M., Jonsson, B.H., and Lundström, P. (2013). PINT: a software for integration of peak volumes and extraction of relaxation rates. J Biomol NMR *56*, 191–202.

Akutsu, M., Dikic, I., and Bremm, A. (2016). Ubiquitin chain diversity at a glance. Journal of Cell Science *129*, 875.

Alexander, C.G., Wanner, R., Johnson, C.M., Breitsprecher, D., Winter, G., Duhr, S., Baaske, P., and Ferguson, N. (2014). Novel microscale approaches for easy, rapid determination of protein stability in academic and commercial settings. Biochim Biophys Acta *1844*, 2241–2250.

Arai, M., Ferreon, J.C., and Wright, P.E. (2012). Quantitative analysis of multisite protein-ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. J Am Chem Soc *134*, 3792–3803.

Arama, E., Dickman, D., Kimchie, Z., Shearn, A., and Lev, Z. (2000). Mutations in the beta-propeller domain of the *Drosophila* brain tumor (Brat) protein induce neoplasm in the larval brain. Oncogene *19*, 3706–3716.

Arvola, R.M., Weidmann, C.A., Tanaka Hall, T.M., and Goldstrohm, A.C. (2017). Combinatorial control of messenger RNAs by Pumilio, Nanos and Brain Tumor proteins. RNA Biol *14*, 1445–1456.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653–1659.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic Acids Res *43*, W39–49.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences *98*, 10037–10041.

Balastik, M., Ferraguti, F., Pires-da Silva, A., Lee, T.H., Alvarez-Bolado, G., Lu, K.P., and Gruss, P. (2008). Deficiency in ubiquitin ligase TRIM2 causes accumulation of neurofilament light chain and neurodegeneration. Proc Natl Acad Sci U S A *105*, 12016–12021.

Bawa, S., Brooks, D.S., Neville, K.E., Tipping, M., Sagar, M.A., Kollhoff, J.A., Chawla, G., Geisbrecht, B.V., Tennessen, J.M., Eliceiri, K.W., and Geisbrecht, E.R. (2020). *Drosophila* TRIM32 cooperates with glycolytic enzymes to promote cell growth. eLife *9*.

Bengtson, M.H. and Joazeiro, C.A.P. (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature *467*, 470–473.

Berlin, K., Longhini, A., Dayie, T.K., and Fushman, D. (2013). Deriving quantitative dynamics information for proteins and RNAs using ROTDIF with a graphical user interface. J Biomol NMR *57*, 333–352.

Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using Small-Angle X-ray Scattering. Journal of the American Chemical Society *129*, 5656–5664.

Betschinger, J., Mechtler, K., and Knoblich, J.A. (2006). Asymmetric segregation of the tumor suppressor Brat regulates self-renewal in *Drosophila* neural stem cells. Cell *124*, 1241 – 1253.

van Beuningen, S.F., Will, L., Harterink, M., Chazeau, A., van Battum, E.Y., Frias, C.P., Franker, M.A., Katrukha, E.A., Stucchi, R., Vocking, K., Antunes, A.T., Slenders, L., Doulkeridou, S., Smitt, P.S., Altelaar, A.M., Post, J.A., Akhmanova, A., Pasterkamp, R.J., Kapitein, L.C., de Graaff, E., and Hoogenraad, C.C. (2015). TRIM46 controls neuronal polarity and axon specification by driving the formation of parallel microtubule arrays. Neuron *88*, 1208 – 1226.

Beuth, B., García-Mayoral, M.F., Taylor, I.A., and Ramos, A. (2007). Scaffold-independent analysis of RNA-protein interactions: the Nova-1 KH3-RNA complex. J Am Chem Soc *129*, 10205–10210.

Biterova, E., Ignatyev, A., Uusimaa, J., Hinttala, R., and Ruddock, L.W. (2018). Structural analysis of human NHLRC2, mutations of which are associated with FINCA disease. PLoS One *13*, e0202391.

Blanchet, C.E., Spilotros, A., Schwemmer, F., Graewert, M.A., Kikhney, A., Jeffries, C.M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., Fiedler, S., Roessle, M., and Svergun, D.I. (2015). Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). Journal of Applied Crystallography *48*, 431–443.

Borden, K.L., Martin, S.R., O'Reilly, N.J., Lally, J.M., Reddy, B.A., Etkin, L.D., and Freemont, P.S. (1993). Characterisation of a novel cysteine/histidine-rich metal binding domain from *Xenopus* nuclear factor XNF7. FEBS Lett *335*, 255–260.

Brown, N.G., VanderLinden, R., Watson, E.R., Qiao, R., Grace, C.R.R., Yamaguchi, M., Weissmann, F., Frye, J.J., Dube, P., Ei Cho, S., Actis, M.L., Rodrigues, P., Fujii, N., Peters, J.M., Stark, H., and Schulman, B.A. (2015). RING E3 mechanism for ubiquitin ligation to a disordered substrate visualized for human anaphase-promoting complex. Proceedings of the National Academy of Sciences *112*, 5272–5279.

Brünger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nat Protoc *2*, 2728–2733.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998). *Crystallography and NMR System*: A New Software Suite for Macromolecular Structure Determination. Acta Crystallographica Section D *54*, 905–921.

Buetow, L. and Huang, D.T. (2016). Structural insights into the catalysis and regulation of E3 ubiquitin ligases. Nat Rev Mol Cell Biol *17*, 626–642.

Cano, F., Bye, H., Duncan, L.M., Buchet-Poyau, K., Billaud, M., Wills, M.R., and Lehner, P.J. (2012). The RNA-binding E3 ubiquitin ligase MEX-3C links ubiquitination with MHC-I mRNA degradation. The EMBO journal *31*, 3596–3606.

Cavanagh, J. (2007). Protein NMR spectroscopy : principles and practice (Amsterdam Boston: Academic Press).

Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science *243*, 1576–1583.

Chen, G., Kong, J., Tucker-Burden, C., Anand, M., Rong, Y., Rahman, F., Moreno, C.S., Van Meir, E.G., Hadjipanayis, C.G., and Brat, D.J. (2014). Human Brat ortholog TRIM3 is a tumor suppressor that regulates asymmetric cell division in glioblastoma. Cancer Res *74*, 4536–4548.

Chen, V.B., Arendall, W.B.r., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12–21.

Chen, X., Dong, C., Law, P.T.Y., Chan, M.T.V., Su, Z., Wang, S., Wu, W.K.K., and Xu, H. (2015). MicroRNA-145 targets TRIM2 and exerts tumor-suppressing functions in epithelial ovarian cancer. Gynecol Oncol *139*, 513–519.

Cheung, C.C., Yang, C., Berger, T., Zaugg, K., Reilly, P., Elia, A.J., Wakeham, A., You-Ten, A., Chang, N., Li, L., Wan, Q., and Mak, T.W. (2010). Identification of BERP (brain-expressed RING finger protein) as a p53 target gene that modulates seizure susceptibility through interacting with GABA(A) receptors. Proc Natl Acad Sci U S A *107*, 11883–11888.

Cho, P.F., Gamberi, C., Cho-Park, Y.A., Cho-Park, I.B., Lasko, P., and Sonenberg, N. (2006). Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos. Curr Biol *16*, 2035–2041.

Choudhury, N.R., Heikel, G., Trubitsyna, M., Kubik, P., Nowak, J.S., Webb, S., Granneman, S., Spanos, C., Rappsilber, J., Castello, A., and Michlewski, G. (2017). RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. BMC Biology *15*, 105.

Choudhury, N.R., Nowak, J.S., Zuo, J., Rappsilber, J., Spoel, S.H., and Michlewski, G. (2014). TRIM25 is an RNA-specific activator of Lin28a/TuT4-mediated uridylation. Cell Reports *9*, 1265–1272.

Chu, Y. and Yang, X. (2010). SUMO E3 ligase activity of TRIM proteins. Oncogene *30*, 1108 EP –.

Chufán, E.E., De, M., Eipper, B.A., Mains, R.E., and Amzel, L.M. (2009). Amidation of bioactive peptides: the structure of the lyase domain of the amidating enzyme. Structure (London, England : 1993) *17*, 965–973.

Clore, G.M., Driscoll, P.C., Wingfield, P.T., and Gronenborn, A.M. (1990a). Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear 15N-1H NMR spectroscopy. Biochemistry *29*, 7387–7401.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C., and Gronenborn, A.M. (1990b). Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. Journal of the American Chemical Society *112*, 4989–4991.

Collaborative, E. (2019). Ultra-rare genetic variation in the epilepsies: A whole-exome sequencing study of 17,606 individuals. Am J Hum Genet *105*, 267–282.

Collins, K.M., Oregioni, A., Robertson, L.E., Kelly, G., and Ramos, A. (2015). Protein-RNA specificity by high-throughput principal component analysis of NMR spectra. Nucleic acids research *43*, e41–e41.

Consortium, T.U. (2018). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research *47*, D506–D515.

Crawford, L.J., Johnston, C.K., and Irvine, A.E. (2018). TRIM proteins in blood cancers. J Cell Commun Signal *12*, 21–29.

Dauter, Z. (1999). Data-collection strategies. Acta Crystallographica Section D *55*, 1703–1717.

Davis, G.M., Tu, S., Anderson, J.W.T., Colson, R.N., Gunzburg, M.J., Francisco, M.A., Ray, D., Shrubsole, S.P., Sobotka, J.A., Seroussi, U., Lao, R.X., Maity, T., Wu, M.Z., McJunkin, K., Morris, Q.D., Hughes, T.R., Wilce, J.A., Claycomb, J.M., Weng, Z., and Boag, P.R. (2018). The TRIM-NHL protein NHL-2 is a co-factor in the nuclear and somatic RNAi pathways in *C. elegans*. eLife *7*, e35478.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277–293.

Deshaies, R.J. and Joazeiro, C.A.P. (2009). RING domain E3 ubiquitin ligases. Annu Rev Biochem *78*, 399–434.

Diaz-Griffero, F., Li, X., Javanbakht, H., Song, B., Welikala, S., Stremlau, M., and Sodroski, J. (2006). Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. Virology *349*, 300–315.

Dickson, C., Fletcher, A.J., Vaysburd, M., Yang, J.C., Mallery, D.L., Zeng, J., Johnson, C.M., McLaughlin, S.H., Skehel, M., Maslen, S., Cruickshank, J., Huguenin-Dezot, N., Chin, J.W., Neuhaus, D., and James, L.C. (2018). Intracellular antibody signalling is regulated by phosphorylation of the Fc receptor TRIM21. eLife *7*, e32660.

van Diepen, M.T., Spencer, G.E., van Minnen, J., Gouwenberg, Y., Bouwman, J., Smit, A.B., and van Kesteren, R.E. (2005). The molluscan RING-finger protein L-TRIM is essential for neuronal outgrowth. Mol Cell Neurosci *29*, 74–81.

Dou, H., Buetow, L., Sibbet, G.J., Cameron, K., and Huang, D.T. (2013). Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. Nature Structural & Molecular Biology *20*, 982–986.

Dupeux, F., Röwer, M., Seroul, G., Blot, D., and Márquez, J. (2011). A thermal stability assay can help to estimate the crystallization likelihood of biological samples. Acta Crystallogr D Biol Crystallogr *67*, 915–919.

Edwards, T.A., Wilkinson, B.D., Wharton, R.P., and Aggarwal, A.K. (2003). Model of the brain tumor-Pumilio translation repressor complex. Genes Dev *17*, 2508–2513.

Eletsky, A., Kienhöfer, A., and Pervushin, K. (2001). TROSY NMR with partially deuterated proteins. Journal of Biomolecular NMR *20*, 177–180.

Emsley, P. and Cowtan, K. (2004). *Coot*: model-building tools for molecular graphics. Acta Crystallographica Section D *60*, 2126–2132.

van den Ent, F. and Löwe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods *67*, 67–74.

Ferreira, A., Boulan, L., Perez, L., and Milán, M. (2014). Mei-P26 mediates tissue-specific responses to the Brat tumor suppressor and the dMyc proto-oncogene in *Drosophila*. Genetics *198*, 249–258.

Fletcher, A.J., Vaysburd, M., Maslen, S., Zeng, J., Skehel, J.M., Towers, G.J., and James, L.C. (2018). Trivalent RING assembly on retroviral capsids activates TRIM5 ubiquitination and innate immune signaling. Cell Host & Microbe *24*, 761–775.e6.

Frank, D.J. and Roth, M.B. (1998). *ncl-1* is required for the regulation of cell size and ribosomal RNA synthesis in *Caenorhabditis elegans*. J Cell Biol *140*, 1321–1329.

Freemont, P.S. (1993). The RING finger. A novel protein sequence motif related to the zinc finger. Ann N Y Acad Sci *684*, 174–192.

Fridell, R.A., Harding, L.S., Bogerd, H.P., and Cullen, B.R. (1995). Identification of a novel human zinc finger protein that specifically interacts with the activation domain of lentiviral Tat proteins. Virology *209*, 347–357.

Frosk, P., Weiler, T., Nylen, E., Sudha, T., Greenberg, C.R., Morgan, K., Fujiwara, T.M., and Wrogemann, K. (2002). Limb-girdle muscular dystrophy type 2H associated with mutation in TRIM32, a putative E3-ubiquitin-ligase gene. Am J Hum Genet *70*, 663–672.

Fushman, D., Cowburn, D., James, T.L., Dötsch, V., and Schmitz, U. (2001). [7] - Nuclear Magnetic Resonance Relaxation in Determination of Residue-Specific 15N Chemical Shift Tensors in Proteins in Solution: Protein Dynamics, Structure, and Applications of Transverse Relaxation Optimized Spectroscopy (Academic Press), vol. 339. 109–126.

Gabadinho, J., Beteva, A., Guijarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M.W., Brockhauser, S., Flot, D., Gordon, E.J., Hall, D.R., Lavault, B., McCarthy, A.A., McCarthy, J., Mitchell, E., Monaco, S., Mueller-Dieckmann, C., Nurizzo, D., Ravelli, R.B.G., Thibault, X., Walsh, M.A., Leonard, G.A., and McSweeney, S.M. (2010). *MxCuBE*: a synchrotron beamline control environment customized for macromolecular crystallography experiments. Journal of Synchrotron Radiation *17*, 700–707.

van Gent, M., Sparrer, K.M.J., and Gack, M.U. (2018). TRIM proteins and their roles in antiviral host defenses. Annu Rev Virol *5*, 385–405.

Gil, R.R. and Navarro-Vázquez, A. (2017). Chapter 1 application of the nuclear Overhauser effect to the structural elucidation of natural products. In: Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 2: Data Acquisition and Applications to Compound Classes, vol. 2 (The Royal Society of Chemistry), 1–38.

Good, M.C., Greenstein, A.E., Young, T.A., Ng, H.L., and Alber, T. (2004). Sensor domain of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinase, PknD, forms a highly symmetric beta-propeller. Journal of Molecular Biology *339*, 459–469.

Guan, X., Li, J., Lu, X., Dong, Y., Chen, W., and Li, X. (2014). Expression, purification, crystallization and preliminary X-ray diffraction analysis of the C-terminal NHL domain of human TRIM2. Acta Crystallogr F Struct Biol Commun *70*, 673–675.

Guinier, A., Fournet, G., Walker, C., Yudowitch, K., Sons, J.W.., Chapman, and (Londyn)., H. (1955). Small-angle Scattering of X-rays. Structure of matter series (Wiley).

Haas, A.L., Warms, J.V., Hershko, A., and Rose, I.A. (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. J Biol Chem *257*, 2543–2548.

Hammell, C.M., Lubin, I., Boag, P.R., Blackwell, T.K., and Ambros, V. (2009). *nhl-2* modulates microRNA activity in *Caenorhabditis elegans*. Cell *136*, 926–938.

Harris, R.E., Pargett, M., Sutcliffe, C., Umulis, D., and Ashe, H.L. (2011). Brat promotes stem cell differentiation via control of a bistable switch that restricts BMP signaling. Dev Cell *20*, 72–83.

Harris, R.K., Becker, E.D., de Menezes, S.M.C., Goodfellow, R., and Granger, P. (2001). NMR nomenclature. Nuclear spin properties and conventions for

chemical shifts (IUPAC Recommendations 2001). Pure and Applied Chemistry *73*, 1795–1818.

Hatakeyama, S. (2017). TRIM family proteins: Roles in autophagy, immunity, and carcinogenesis. Trends Biochem Sci *42*, 297–311.

Haubrich, K., Augsten, S., Simon, B., Masiewicz, P., Perez, K., Lethier, M., Rittinger, K., Gabel, F., Cusack, S., and Hennig, J. (2020). RNA binding regulates TRIM25-mediated RIG-I ubiquitylation. bioRxiv .

Heissmeyer, V. and Vogel, K.U. (2013). Molecular control of Tfh-cell differentiation by Roquin family proteins. Immunological Reviews *253*, 273–289.

Hennig, J., Militti, C., Popowicz, G.M., Wang, I., Sonntag, M., Geerlof, A., Gabel, F., Gebauer, F., and Sattler, M. (2014). Structural basis for the assembly of the Sxl-Unr translation regulatory complex. Nature *515*, 287–290.

Hennig, J., Wang, I., Sonntag, M., Gabel, F., and Sattler, M. (2013). Combining NMR and small angle X-ray and neutron scattering in the structural analysis of a ternary protein-RNA complex. J Biomol NMR *56*, 17–30.

Hershko, A. and Ciechanover, A. (1998). The ubiquitin system. Annu Rev Biochem *67*, 425–479.

Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135–141.

Hollmann, N.M., Jagtap, P.K.A., Masiewicz, P., Guitart, T., Simon, B., Provaznik, J., Stein, F., Haberkant, P., Sweetapple, L.J., Villacorta, L., Mooijman, D., Benes, V., Savitski, M.M., Gebauer, F., and Hennig, J. (2020). Pseudo-RNA-binding domains mediate RNA structure specificity in Upstream of N-Ras. Cell Rep *32*, 107930.

Huang, F., Kirkpatrick, D., Jiang, X., Gygi, S., and Sorkin, A. (2006). Differential regulation of EGF receptor internalization and degradation by multiubiquitination within the kinase domain. Mol Cell *21*, 737–748.

Hung, A.Y., Sung, C.C., Brito, I.L., and Sheng, M. (2010). Degradation of postsynaptic scaffold GKAP and regulation of dendritic spine morphology by the TRIM3 ubiquitin ligase in rat hippocampal neurons. PLOS ONE *5*, 1–11.

Ibrahim, O., Sutherland, H.G., Maksemous, N., Smith, R., Haupt, L.M., and Griffiths, L.R. (2020). Exploring neuronal vulnerability to head trauma using a whole exome approach. J Neurotrauma .

Isaksson, J., Acharya, S., Barman, J., Cheruku, P., and Chattopadhyaya, J. (2004). Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. Biochemistry *43*, 15996–16010.

Jankowsky, E. and Harris, M.E. (2015). Specificity and nonspecificity in RNA-protein interactions. Nature reviews. Molecular cell biology *16*, 533–544.

Jin, J., Cheng, Y., Zhang, Y., Wood, W., Peng, Q., Hutchison, E., Mattson, M.P., Becker, K.G., and Duan, W. (2012). Interrogation of brain miRNA and mRNA expression profiles reveals a molecular regulatory network that is perturbed by mutant huntingtin. J Neurochem *123*, 477–490.

Jin, J., Li, X., Gygi, S.P., and Harper, J.W. (2007). Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature *447*, 1135–1138.

Joazeiro, C.A. and Weissman, A.M. (2000). RING finger proteins: mediators of ubiquitin ligase activity. Cell *102*, 549–552.

Joazeiro, C.A.P. (2019). Mechanisms and functions of ribosome-associated protein quality control. Nature Reviews Molecular Cell Biology *20*, 368–383.

Kabsch, W. (2010a). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr *66*, 133–144.

Kabsch, W. (2010b). *XDS*. Acta Crystallographica Section D *66*, 125–132.

Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry *28*, 8972–8979.

Keeler, J. (2005). Understanding NMR spectroscopy (Chichester, England Hoboken, NJ: Wiley).

Khazaei, M.R., Bunk, E.C., Hillje, A.L., Jahn, H.M., Riegler, E.M., Knoblich, J.A., Young, P., and Schwamborn, J.C. (2011). The E3-ubiquitin ligase TRIM2 regulates neuronal polarization. J Neurochem *117*, 29–37.

Kirkin, V., McEwan, D.G., Novak, I., and Dikic, I. (2009). A role for ubiquitin in selective autophagy. Mol Cell *34*, 259–269.

Koch, M.H.J., Vachette, P., and Svergun, D.I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution *36*, 147–227.

Koliopoulos, M.G., Esposito, D., Christodoulou, E., Taylor, I.A., and Rittinger, K. (2016). Functional role of TRIM E3 ligase oligomerization and regulation of catalytic activity. EMBO J *35*, 1204–1218.

Koliopoulos, M.G., Lethier, M., van der Veen, A.G., Haubrich, K., Hennig, J., Kowalinski, E., Stevens, R.V., Martin, S.R., Reis e Sousa, C., Cusack, S., and Rittinger, K. (2018). Molecular mechanism of influenza A NS1-mediated TRIM25 recognition and inhibition. Nature Communications *9*, 1820.

Komander, D. and Rape, M. (2012). The ubiquitin code. Annual Review of Biochemistry *81*, 203–229.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography *36*, 1277–1282.

Korzhnev, D.M., Billeter, M., Arseniev, A.S., and Orekhov, V.Y. (2001). NMR studies of Brownian tumbling and internal motions in proteins. Progress in Nuclear Magnetic Resonance Spectroscopy *38*, 197–266.

Kudryashova, E., Struyk, A., Mokhonova, E., Cannon, S.C., and Spencer, M.J. (2011). The common missense mutation D489N in TRIM32 causing limb girdle muscular dystrophy 2H leads to loss of the mutated protein in knock-in mice resulting in a TRIM32-null phenotype. Hum Mol Genet *20*, 3925–3932.

Kumari, P., Aeschimann, F., Gaidatzis, D., Keusch, J.J., Ghosh, P., Neagu, A., Pachulska-Wieczorek, K., Bujnicki, J.M., Gut, H., Grosshans, H., and Ciosk, R. (2018). Evolutionary plasticity of the NHL domain underlies distinct solutions to RNA recognition. Nat Commun *9*, 1549.

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nature Structural &Amp; Molecular Biology *20*, 1122 EP –.

Langevin, C., Levraud, J.P., and Boudinot, P. (2019). Fish antiviral tripartite motif (TRIM) proteins. Fish and Shellfish Immunology *86*, 724 – 733.

Lau, C.k., Bachorik, J.L., and Dreyfuss, G. (2009). Gemin5-snRNA interaction reveals an RNA binding function for WD repeat domains. Nat Struct Mol Biol *16*, 486–491.

Laver, J.D., Li, X., Ray, D., Cook, K.B., Hahn, N.A., Nabeel-Shah, S., Kekis, M., Luo, H., Marsolais, A.J., Fung, K.Y., Hughes, T.R., Westwood, J.T., Sidhu, S.S.,

Morris, Q., Lipshitz, H.D., and Smibert, C.A. (2015). Brain tumor is a sequence-specific RNA-binding protein that directs maternal mRNA clearance during the *Drosophila* maternal-to-zygotic transition. Genome Biol *16*, 94.

Lee, C.Y., Wilkinson, B.D., Siegrist, S.E., Wharton, R.P., and Doe, C.Q. (2006). Brat is a Miranda cargo protein that promotes neuronal differentiation and inhibits neuroblast self-renewal. Developmental Cell *10*, 441 – 449.

Levitt, M. (1988). 2: Density-operator theory of pulses and precession. In: Pulse methods in 1D and 2D liquid-phase NMR.

Li, C., Peng, Q., Wan, X., Sun, H., and Tang, J. (2017). C-terminal motifs in promyelocytic leukemia protein isoforms critically regulate PML nuclear body formation. J Cell Sci *130*, 3496–3506.

Li, J.J., Sarute, N., Lancaster, E., Otkiran-Clare, G., Fagla, B.M., Ross, S.R., and Scherer, S.S. (2020). A recessive *Trim2* mutation causes an axonal neuropathy in mice. Neurobiol Dis *140*, 104845.

Li, M., Brooks, C.L., Wu-Baer, F., Chen, D., Baer, R., and Gu, W. (2003). Mono- versus polyubiquitination: Differential control of p53 fate by Mdm2. Science *302*, 1972–1975.

Li, X., Yeung, D.F., Fiegen, A.M., and Sodroski, J. (2011). Determinants of the higher order association of the restriction factor TRIM5alpha and other tripartite motif (TRIM) proteins. J Biol Chem *286*, 27959–27970.

Li, Y.L., Chandrasekaran, V., Carter, S.D., Woodward, C.L., Christensen, D.E., Dryden, K.A., Pornillos, O., Yeager, M., Ganser-Pornillos, B.K., Jensen, G.J., and Sundquist, W.I. (2016). Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. eLife *5*, e16269.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Poon, B.K., Prisant, M.G., Read, R.J., Richardson, J.S., Richardson, D.C., Sammito, M.D., Sobolev, O.V., Stockwell, D.H., Terwilliger, T.C., Urzhumtsev, A.G., Videau, L.L., Williams, C.J., and Adams, P.D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861–877.

Lin, Z., Lin, X., Zhu, L., Huang, J., and Huang, Y. (2020). TRIM2 directly deubiquitinates and stabilizes Snail1 protein, mediating proliferation and metastasis of lung adenocarcinoma. Cancer Cell Int *20*, 228.

Linge, J.P., Habeck, M., Rieping, W., and Nilges, M. (2003). ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics *19*, 315–316.

Lipari, G. and Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. Journal of the American Chemical Society *104*, 4546–4559.

Lipari, G. and Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. analysis of experimental results. Journal of the American Chemical Society *104*, 4559–4570.

Liu, B., Li, N.L., Shen, Y., Bao, X., Fabrizio, T., Elbahesh, H., Webby, R.J., and Li, K. (2016). The C-terminal tail of TRIM56 dictates antiviral restriction of influenza A and B viruses by impeding viral RNA synthesis. J Virol *90*, 4369–4382.

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., and Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic acids research *41*, 518–532.

Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R., Engelmann, J.C., Krahn, M.P., and Meister, G. (2015). The crystal structure of the NHL domain in complex with RNA reveals the molecular basis of *Drosophila* brain-tumor-mediated gene regulation. Cell Rep *13*, 1206–1220.

Loedige, I., Stotz, M., Qamar, S., Kramer, K., Hennig, J., Schubert, T., Loffler, P., Langst, G., Merkl, R., Urlaub, H., and Meister, G. (2014). The NHL domain of BRAT is an RNA-binding domain that directly contacts the *hunchback* mRNA for regulation. Genes Dev *28*, 749–764.

Lokapally, A., Neuhaus, H., Herfurth, J., and Hollemann, T. (2020). Interplay of TRIM2 E3 ubiquitin ligase and ALIX/ESCRT complex: Control of developmental plasticity during early neurogenesis. Cells *9*.

Lu, Q., Zhang, Y., Ma, L., Li, D., Li, M., Liu, P., and Li, J. (2019). TRIM3 negatively regulates autophagy through promoting degradation of Beclin1 in Ewing sarcoma cells. Onco Targets Ther *12*, 11587–11595.

Luft, J.R., Newman, J., and Snell, E.H. (2014). Crystallization screening: the influence of history on current practice. Acta Crystallogr F Struct Biol Commun *70*, 835–853.

Lyumkis, D., Oliveira dos Passos, D., Tahara, E.B., Webb, K., Bennett, E.J., Vinterbo, S., Potter, C.S., Carragher, B., and Joazeiro, C.A.P. (2014). Structural basis

for translational surveillance by the large ribosomal subunit-associated protein quality control complex. Proceedings of the National Academy of Sciences of the United States of America *111*, 15981–15986.

Macošek, J., Simon, B., Linse, J.B., Winter, S., Foot, J., Perez, K., Rettel, M., Ivanović, M., Masiewicz, P., Murciano, B., Savitski, M.M., Hub, J.S., Gabel, F., and Hennig, J. (2020). Structure and dynamics of the quaternary *hunchback* mRNA translation repression complex. bioRxiv , 2020.09.08.287060.

Marchetti, G., Reichardt, I., Knoblich, J.A., and Besse, F. (2014). The TRIM-NHL protein Brat promotes axon maintenance by repressing *src64B* expression. J Neurosci *34*, 13855–13864.

de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., and Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. BMC Biotechnology *7*, 32.

Marín, I. (2012). Origin and diversification of TRIM ubiquitin ligases. PLOS ONE *7*, e50030–.

Massiah, M.A., Matts, J.A.B., Short, K.M., Simmons, B.N., Singireddy, S., Yi, Z., and Cox, T.C. (2007). Solution structure of the MID1 B-box2 CHC(D/C)C2H2 zinc-binding domain: Insights into an evolutionarily conserved RING fold. Journal of Molecular Biology *369*, 1–10.

Matsumoto, M.L., Wickliffe, K.E., Dong, K.C., Yu, C., Bosanac, I., Bustos, D., Phu, L., Kirkpatrick, D.S., Hymowitz, S.G., Rape, M., Kelley, R.F., and Dixit, V.M. (2010). K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. Mol Cell *39*, 477–484.

McCarthy, A.A., Barrett, R., Beteva, A., Caserotto, H., Dobias, F., Felisaz, F., Giraud, T., Guijarro, M., Janocha, R., Khadrouche, A., Lentini, M., Leonard, G.A., Lopez Marrero, M., Malbet-Monaco, S., McSweeney, S., Nurizzo, D., Papp, G., Rossi, C., Sinoir, J., Sorez, C., Surr, J., Svensson, O., Zander, U., Cipriani, F., Theveneau, P., and Mueller-Dieckmann, C. (2018). ID30B – a versatile beamline for macromolecular crystallography experiments at the ESRF. Journal of Synchrotron Radiation *25*, 1249–1260.

McGinty, R.K., Henrici, R.C., and Tan, S. (2014). Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. Nature *514*, 591–596.

Menon, S., Boyer, N.P., Winkle, C.C., McClain, L.M., Hanlin, C.C., Pandey, D., Rothenfußer, S., Taylor, A.M., and Gupton, S.L. (2015). The E3 ubiquitin ligase

TRIM9 is a filopodia off switch required for netrin-dependent axon guidance. Developmental Cell *35*, 698 – 712.

Meroni, G. and Diez-Roux, G. (2005). TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. Bioessays *27*, 1147–1157.

Mills, J.E. and Dean, P.M. (1996). Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. J Comput Aided Mol Des *10*, 607–622.

Mitschka, S., Ulas, T., Goller, T., Schneider, K., Egert, A., Mertens, J., Brustle, O., Schorle, H., Beyer, M., Klee, K., Xue, J., Gunther, P., Bassler, K., Schultze, J.L., and Kolanus, W. (2015). Co-existence of intact stemness and priming of neural differentiation programs in mES cells lacking TRIM71. Sci Rep *5*, 11126.

Mori, S., Abeygunawardana, C., Johnson, M.O., and van Zijl, P.C. (1995). Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. J Magn Reson B *108*, 94–98.

Mukherjee, S., Tucker-Burden, C., Zhang, C., Moberg, K., Read, R., Hadjipanayis, C., and Brat, D.J. (2016). *Drosophila* Brat and human ortholog TRIM3 maintain stem cell equilibrium and suppress brain tumorigenesis by attenuating Notch nuclear transport. Cancer Res *76*, 2443–2452.

Neumuller, R.A., Betschinger, J., Fischer, A., Bushati, N., Poernbacher, I., Mechtler, K., Cohen, S.M., and Knoblich, J.A. (2008). Mei-P26 regulates microRNAs and cell growth in the *Drosophila* ovarian stem cell lineage. Nature *454*, 241–245.

Nietlispach, D. (2005). Suppression of anti-TROSY lines in a sensitivity enhanced gradient selection TROSY scheme. Journal of Biomolecular NMR *31*, 161–166.

Nurizzo, D., Mairs, T., Guijarro, M., Rey, V., Meyer, J., Fajardo, P., Chavanne, J., Biasci, J.C., McSweeney, S., and Mitchell, E. (2006). The ID23-1 structural biology beamline at the ESRF. Journal of synchrotron radiation *13*, 227–238.

Ohkawa, N., Kokura, K., Matsu-Ura, T., Obinata, T., Konishi, Y., and Tamura, T.A. (2001). Molecular cloning and characterization of neural activity-related RING finger protein (NARF): a new member of the RBCC family is a candidate for the partner of myosin V. J Neurochem *78*, 75–87.

Ozato, K., Shin, D.M., Chang, T.H., and Morse, Herbert C, r. (2008). TRIM family proteins and their emerging roles in innate immunity. Nature reviews. Immunology *8*, 849–860.

Page, R., Peti, W., Wilson, I.A., Stevens, R.C., and Wüthrich, K. (2005). NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. Proceedings of the National Academy of Sciences *102*, 1901–1905.

Pal, A. and Levy, Y. (2019). Structure, stability and specificity of the binding of ssDNA and ssRNA with proteins. PLoS Comput Biol *15*, e1006768.

Pehlivan, D., Coban Akdemir, Z., Karaca, E., Bayram, Y., Jhangiani, S., Yildiz, E.P., Muzny, D., Uluc, K., Gibbs, R.A., Elcioglu, N., Lupski, J.R., and Harel, T. (2015). Exome sequencing reveals homozygous TRIM2 mutation in a patient with early onset CMT and bilateral vocal cord paralysis. Hum Genet *134*, 671–673.

Peisley, A., Wu, B., Xu, H., Chen, Z.J., and Hur, S. (2014). Structural basis for ubiquitin-mediated antiviral signal activation by rig-i. Nature *509*, 110–114.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated *T2* relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proceedings of the National Academy of Sciences *94*, 12366.

Philippe, H., Poustka, A.J., Chiodin, M., Hoff, K.J., Dessimoz, C., Tomiczek, B., Schiffer, P.H., Müller, S., Domman, D., Horn, M., Kuhl, H., Timmermann, B., Satoh, N., Hikosaka-Katayama, T., Nakano, H., Rowe, M.L., Elphick, M.R., Thomas-Chollier, M., Hankeln, T., Mertes, F., Wallberg, A., Rast, J.P., Copley, R.R., Martinez, P., and Telford, M.J. (2019). Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria. Curr Biol *29*, 1818–1826.

Pickart, C.M. and Eddins, M.J. (2004). Ubiquitin: structures, functions, mechanisms. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1695*, 55 – 72, the Ubiquitin-Proteasome System.

Pruneda, J.N., Stoll, K.E., Bolton, L.J., Brzovic, P.S., and Klevit, R.E. (2011). Ubiquitin in motion: structural studies of the ubiquitin-conjugating enzyme-ubiquitin conjugate. Biochemistry *50*, 1624–1633.

Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys *40*, 191–285.

Rajsbaum, R., García-Sastre, A., and Versteeg, G.A. (2014). TRIMmunity: the roles of the TRIM E3-ubiquitin ligase family in innate antiviral immunity. J Mol Biol *426*, 1265–1284.

Rao, L., Mak, V.C.Y., Zhou, Y., Zhang, D., Li, X., Fung, C.C.Y., Sharma, R., Gu, C., Lu, Y., Tipoe, G.L., Cheung, A.N.Y., Mills, G.B., and Cheung, L.W.T. (2020). p85beta regulates autophagic degradation of AXL to activate oncogenic signaling. Nat Commun *11*, 2291.

Reddy, B.A. and Etkin, L.D. (1991). A unique bipartite cysteine-histidine motif defines a subfamily of potential zinc-finger proteins. Nucleic acids research *19*, 6330–6330.

Reddy, B.A., Etkin, L.D., and Freemont, P.S. (1992). A novel zinc finger coiled-coil domain in a family of nuclear proteins. Trends Biochem Sci *17*, 344–345.

Reichardt, I., Bonnay, F., Steinmann, V., Loedige, I., Burkard, T.R., Meister, G., and Knoblich, J.A. (2018). The tumor suppressor Brat controls neuronal stem cell lineages by inhibiting Deadpan and Zelda. EMBO Rep *19*, 102–117.

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., Guffanti, A., Minucci, S., Pelicci, P.G., and Ballabio, A. (2001). The tripartite motif family identifies cell compartments. The EMBO journal *20*, 2140–2151.

Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T.E., and Nilges, M. (2006). ARIA2: Automated NOE assignment and data integration in NMR structure calculation. Bioinformatics *23*, 381–382.

Ryabov, Y.E., Geraghty, C., Varshney, A., and Fushman, D. (2006). An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. Journal of the American Chemical Society *128*, 15432–15444.

Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., and Wulczyn, F.G. (2009). The let-7 target gene mouse *lin-41* is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nat Cell Biol *11*, 1411–1420.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998). TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. Proceedings of the National Academy of Sciences *95*, 13585–13590.

Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999). TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. Journal of the American Chemical Society *121*, 844–848.

Sanchez, J.G., Okreglicka, K., Chandrasekaran, V., Welker, J.M., Sundquist, W.I., and Pornillos, O. (2014). The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. Proc Natl Acad Sci U S A *111*, 2494–2499.

Sardiello, M., Cairo, S., Fontanella, B., Ballabio, A., and Meroni, G. (2008). Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. BMC Evolutionary Biology *8*, 225.

Sarute, N., Ibrahim, N., Medegan Fagla, B., Lavanya, M., Cuevas, C., Stavrou, S., Otkiran-Clare, G., Tyynismaa, H., Henao-Mejia, J., and Ross, S.R. (2019). TRIM2, a novel member of the antiviral family, limits New World arenavirus entry. PLoS Biol *17*, e3000137.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress in Nuclear Magnetic Resonance Spectroscopy *34*, 93–158.

Schnablegger, H. and Singh, Y. (2011). The SAXS guide: getting acquainted with the principles (Austria: Anton Paar GmbH).

Schonrock, N., Humphreys, D.T., Preiss, T., and Götz, J. (2012). Target gene repression mediated by miRNAs miR-181c and miR-9 both of which are down-regulated by amyloid-beta. J Mol Neurosci *46*, 324–335.

Schoser, B.G.H., Frosk, P., Engel, A.G., Klutzny, U., Lochmuller, H., and Wrogemann, K. (2005). Commonality of TRIM32 mutation in causing sarcotubular myopathy and LGMD2H. Ann Neurol *57*, 591–595.

Schreiber, J., Vegh, M.J., Dawitz, J., Kroon, T., Loos, M., Labonte, D., Li, K.W., Van Nierop, P., Van Diepen, M.T., De Zeeuw, C.I., Kneussel, M., Meredith, R.M., Smit, A.B., and Van Kesteren, R.E. (2015). Ubiquitin ligase TRIM3 controls hippocampal plasticity and learning by regulating synaptic gamma-actin levels. J Cell Biol *211*, 569–586.

Schulte-Herbrüggen, T. and Sørensen, O.W. (2000). Clean TROSY: Compensation for relaxation-induced artifacts. Journal of Magnetic Resonance *144*, 123 – 128.

Schumann, F.H., Riepl, H., Maurer, T., Gronwald, W., Neidig, K.P., and Kalbitzer, H.R. (2007). Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. J Biomol NMR *39*, 275–289.

Schwamborn, J.C., Berezikov, E., and Knoblich, J.A. (2009). The TRIM-NHL protein TRIM32 activates microRNAs and prevents self-renewal in mouse neural progenitors. Cell *136*, 913–925.

Shen, Y., Li, N.L., Wang, J., Liu, B., Lester, S., and Li, K. (2012). TRIM56 is an essential component of the TLR3 antiviral signaling pathway. J Biol Chem *287*, 36404–36413.

Shi, W., Chen, Y., Gan, G., Wang, D., Ren, J., Wang, Q., Xu, Z., Xie, W., and Zhang, Y.Q. (2013). Brain tumor regulates neuromuscular synapse growth and endocytosis in *Drosophila* by suppressing Mad expression. J Neurosci *33*, 12352–12363.

Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.

Simon, B. and Kostler, H. (2019). Improving the sensitivity of FT-NMR spectroscopy by apodization weighted sampling. J Biomol NMR *73*, 155–165.

Simon, B., Madl, T., Mackereth, C.D., Nilges, M., and Sattler, M. (2010). An efficient protocol for NMR-spectroscopy-based structure determination of protein complexes in solution. Angewandte Chemie International Edition *49*, 1967–1970.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. Mol Cell *5*, 659–669.

Slack, F.J. and Ruvkun, G. (1998). A novel repeat domain that is often associated with RING finger and B-box motifs. Trends Biochem Sci *23*, 474–475.

Sonoda, J. and Wharton, R.P. (2001). *Drosophila* Brain Tumor is a translational repressor. Genes Dev *15*, 762–773.

Martins-de Souza, D., Gattaz, W.F., Schmitt, A., Rewerts, C., Maccarrone, G., Dias-Neto, E., and Turck, C.W. (2009). Prefrontal cortex shotgun proteome analysis reveals altered calcium homeostasis and immune system imbalance in schizophrenia. Eur Arch Psychiatry Clin Neurosci *259*, 151–163.

Sparrer, K.M.J., Gableske, S., Zurenski, M.A., Parker, Z.M., Full, F., Baumgart, G.J., Kato, J., Pacheco-Rodriguez, G., Liang, C., Pornillos, O., Moss, J., Vaughan, M., and Gack, M.U. (2017). TRIM23 mediates virus-induced autophagy via activation of TBK1. Nat Microbiol *2*, 1543–1557.

Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Research *26*, 423–440.

Stirnimann, C.U., Petsalaki, E., Russell, R.B., and Muller, C.W. (2010). WD40 proteins propel cellular networks. Trends Biochem Sci *35*, 565–574.

Stoll, G.A., Oda, S.i., Chong, Z.S., Yu, M., McLaughlin, S.H., and Modis, Y. (2019). Structure of KAP1 tripartite motif identifies molecular interfaces required for retroelement silencing. Proceedings of the National Academy of Sciences *116*, 15042.

Streich, Frederick C, J., Ronchi, V.P., Connick, J.P., and Haas, A.L. (2013). Tripartite motif ligases catalyze polyubiquitin chain formation through a cooperative allosteric mechanism. The Journal of biological chemistry *288*, 8209–8221.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). *CRYSOL* – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. Journal of Applied Crystallography *28*, 768–773.

Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. Journal of Applied Crystallography *25*, 495–503.

Svergun, D.I. and Koch, M.H. (2003). Small-angle scattering studies of biological macromolecules in solution. Reports on Progress in Physics *66*, 1735.

Taylor, G. (2003). The phase problem. Acta Crystallographica Section D *59*, 1881–1890.

Teilum, K., Kunze, M.B.A., Erlendsson, S., and Kragelund, B.B. (2017). (S)Pinning down protein interactions by NMR. Protein science : a publication of the Protein Society *26*, 436–451.

Thompson, S., Pearson, A.N., Ashley, M.D., Jessick, V., Murphy, B.M., Gafken, P., Henshall, D.C., Morris, K.T., Simon, R.P., and Meller, R. (2011). Identification of a novel Bcl-2-interacting mediator of cell death (Bim) E3 ligase, tripartite motif-containing protein 2 (TRIM2), and its role in rapid ischemic tolerance-induced neuroprotection. J Biol Chem *286*, 19331–19339.

Tocchini, C. and Ciosk, R. (2015). TRIM-NHL proteins in development and disease. Semin Cell Dev Biol *47-48*, 52–59.

Tocchini, C., Keusch, J.J., Miller, S.B., Finger, S., Gut, H., Stadler, M.B., and Ciosk, R. (2014). The TRIM-NHL protein LIN-41 controls the onset of developmental plasticity in *Caenorhabditis elegans*. PLOS Genetics *10*, 1–13.

Tokuda, J.M., Pabit, S.A., and Pollack, L. (2016). Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS. Biophys Rev *8*, 139–149.

Treiber, T., Treiber, N., Plessmann, U., Harlander, S., Daiß, J.L., Eichner, N., Lehmann, G., Schall, K., Urlaub, H., and Meister, G. (2017). A compendium of RNA-binding proteins that regulate microRNA biogenesis. Molecular Cell *66*, 270–284.e13.

Trewhella, J., Duff, A.P., Durand, D., Gabel, F., Guss, J.M., Hendrickson, W.A., Hura, G.L., Jacques, D.A., Kirby, N.M., Kwan, A.H., Pérez, J., Pollack, L., Ryan, T.M., Sali, A., Schneidman-Duhovny, D., Schwede, T., Svergun, D.I., Sugiyama, M., Tainer, J.A., Vachette, P., Westbrook, J., and Whitten, A.E. (2017). 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. Acta Crystallogr D Struct Biol *73*, 710–728.

Tria, G., Mertens, H.D.T., Kachala, M., and Svergun, D.I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ *2*, 207–217.

Vaysburd, M., Watkinson, R.E., Cooper, H., Reed, M., O'Connell, K., Smith, J., Cruickshanks, J., and James, L.C. (2013). Intracellular antibody receptor TRIM21 prevents fatal viral infection. Proceedings of the National Academy of Sciences *110*, 12397–12401.

Voet, A.R.D., Noguchi, H., Addy, C., Simoncini, D., Terada, D., Unzai, S., Park, S.Y., Zhang, K.Y.J., and Tame, J.R.H. (2014). Computational design of a self-assembling symmetrical beta-propeller protein. Proceedings of the National Academy of Sciences of the United States of America *111*, 15102–15107.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696.

Waldhauer, M.C., Schmitz, S.N., Ahlmann-Eltze, C., Gleixner, J.G., Schmelas, C.C., Huhn, A.G., Bunne, C., Büscher, M., Horn, M., Klughammer, N., Kreft, J., Schäfer, E., Bayer, P.A., Krämer, S.G., Neugebauer, J., Wehler, P., Mayer, M.P., Eils, R., and Di Ventura, B. (2015). Backbone circularization of *Bacillus subtilis* family 11 xylanase increases its thermostability and its resistance against aggregation. Mol Biosyst *11*, 3231–3243.

Walker, O., Varadan, R., and Fushman, D. (2004). Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from (15)N relaxation data using computer program ROTDIF. J Magn Reson *168*, 336–345.

Wang, M. and Pickart, C.M. (2005). Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. EMBO J *24*, 4324–4333.

Werner, C.T., Mitra, S., Martin, J.A., Stewart, A.F., Lepack, A.E., Ramakrishnan, A., Gobira, P.H., Wang, Z.J., Neve, R.L., Gancarz, A.M., Shen, L., Maze, I., and Dietz, D.M. (2019). Ubiquitin-proteasomal regulation of chromatin remodeler INO80 in the nucleus accumbens mediates persistent cocaine craving. Sci Adv *5*, eaay0351.

Wessels, H.H., Hirsekorn, A., Ohler, U., and Mukherjee, N. (2016). Identifying RBP Targets with RIP-seq (New York, NY: Springer New York). 141–152.

Wilkie, G.S., Dickson, K.S., and Gray, N.K. (2003). Regulation of mRNA translation by 5′- and 3′-UTR-binding factors. Trends Biochem Sci *28*, 182–188.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall, W.B.r., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S., and Richardson, D.C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci *27*, 293–315.

Williams, F.P., Haubrich, K., Perez-Borrajero, C., and Hennig, J. (2019). Emerging RNA-binding roles in the TRIM family of ubiquitin ligases. Biol Chem *400*, 1443–1464.

Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc *73*, 1–16.

Williamson, R.A., Carr, M.D., Frenkiel, T.A., Feeney, J., and Freedman, R.B. (1997). Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation. Biochemistry *36*, 13882–13889.

Worringer, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The *let-7*/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. Cell Stem Cell *14*, 40–52.

Yang, D. and Kay, L.E. (1999). Improved 1HN-detected triple resonance TROSY-based experiments. Journal of Biomolecular NMR *13*, 3–10.

Yang, D., Li, N.L., Wei, D., Liu, B., Guo, F., Elbahesh, H., Zhang, Y., Zhou, Z., Chen, G.Y., and Li, K. (2019). The E3 ligase TRIM56 is a host restriction factor of Zika virus and depends on its RNA-binding activity but not miRNA regulation, for antiviral function. PLoS Negl Trop Dis *13*, e0007537.

Ye, W., Hu, M.M., Lei, C.Q., Zhou, Q., Lin, H., Sun, M.S., and Shu, H.B. (2017). TRIM8 negatively regulates TLR3/4-mediated innate immune response by blocking TRIF-TBK1 interaction. J Immunol *199*, 1856–1864.

Ye, Y. and Rape, M. (2009). Building ubiquitin chains: E2 enzymes at work. Nature Reviews Molecular Cell Biology *10*, 755–764.

Ylikallio, E., Poyhonen, R., Zimon, M., De Vriendt, E., Hilander, T., Paetau, A., Jordanova, A., Lonnqvist, T., and Tyynismaa, H. (2013). Deficiency of the E3 ubiquitin ligase TRIM2 in early-onset axonal neuropathy. Hum Mol Genet *22*, 2975–2983.

Zanchetta, M.E., Napolitano, L.M.R., Maddalo, D., and Meroni, G. (2017). The E3 ubiquitin ligase MID1/TRIM18 promotes atypical ubiquitination of the BRCA2-associated factor 35, BRAF35. Biochim Biophys Acta Mol Cell Res *1864*, 1844–1854.

Zhang, J., Zhang, C., Cui, J., Ou, J., Han, J., Qin, Y., Zhi, F., and Wang, R.F. (2017). TRIM45 functions as a tumor suppressor in the brain via its E3 ligase activity by stabilizing p53 through K63-linked ubiquitination. Cell Death Dis *8*, e2831.

Zhang, Q., Fan, L., Hou, F., Dong, A., Wang, Y.X., and Tong, Y. (2015). New insights into the RNA-binding and E3 ubiquitin ligase activities of Roquins. Scientific Reports *5*, 15660 EP –.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of Polycomb-associated RNAs by RIP-seq. Molecular Cell *40*, 939–953.

Zhong, S., Delva, L., Rachez, C., Cenciarelli, C., Gandini, D., Zhang, H., Kalantry, S., Freedman, L.P., and Pandolfi, P.P. (1999). A RA-dependent, tumour-growth suppressive transcription complex is the target of the PML-RARalpha and T18 oncoproteins. Nat Genet *23*, 287–295.

Zou, W. and Zhang, D.E. (2006). The interferon-inducible ubiquitin-protein isopeptide ligase (E3) EFP also functions as an ISG15 E3 ligase. J Biol Chem *281*, 3989–3994.

Zou, Y., Chiu, H., Zinovyeva, A., Ambros, V., Chuang, C.F., and Chang, C. (2013). Developmental decline in neuronal regeneration by the progressive change of two intrinsic timers. Science *340*, 372–376.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. Journal of Molecular Biology *428*, 720–725.

Supplementary figures and tables

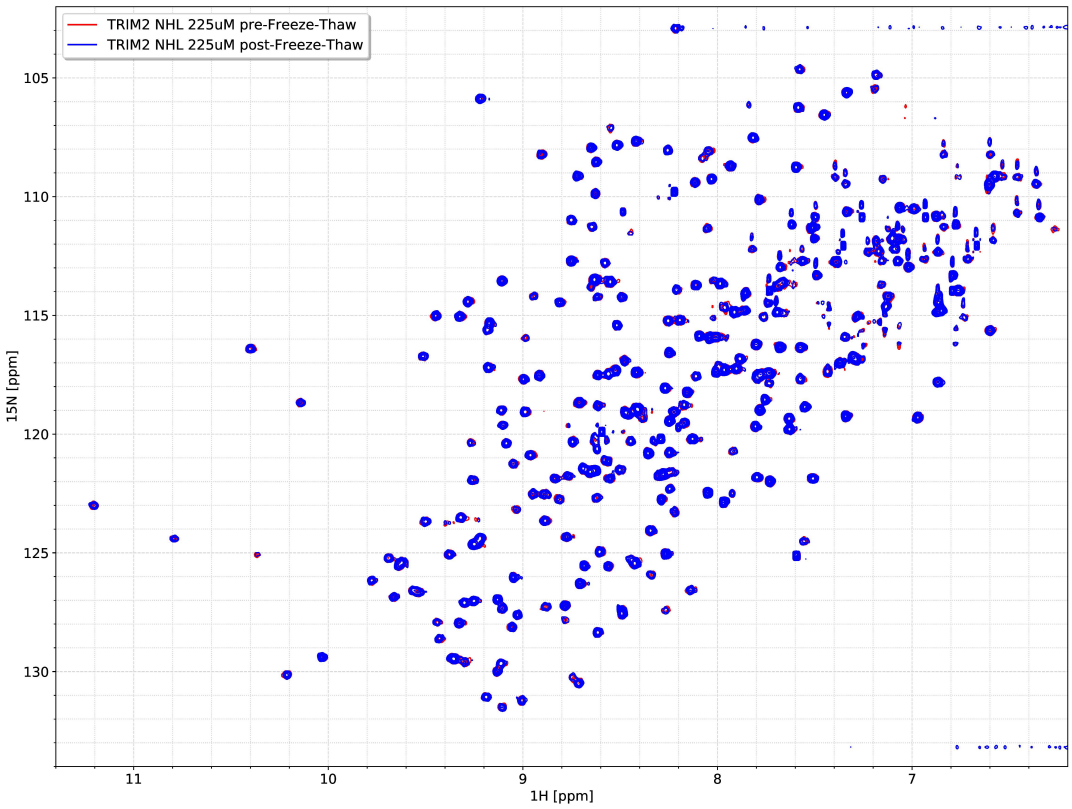


Figure S1: HSQCs of TRIM2-NHL before and after a flash-freeze thaw cycle carried out at 225μM protein concentration in TRIM2 NHL RNA titration buffer (see Table 2 page 2)

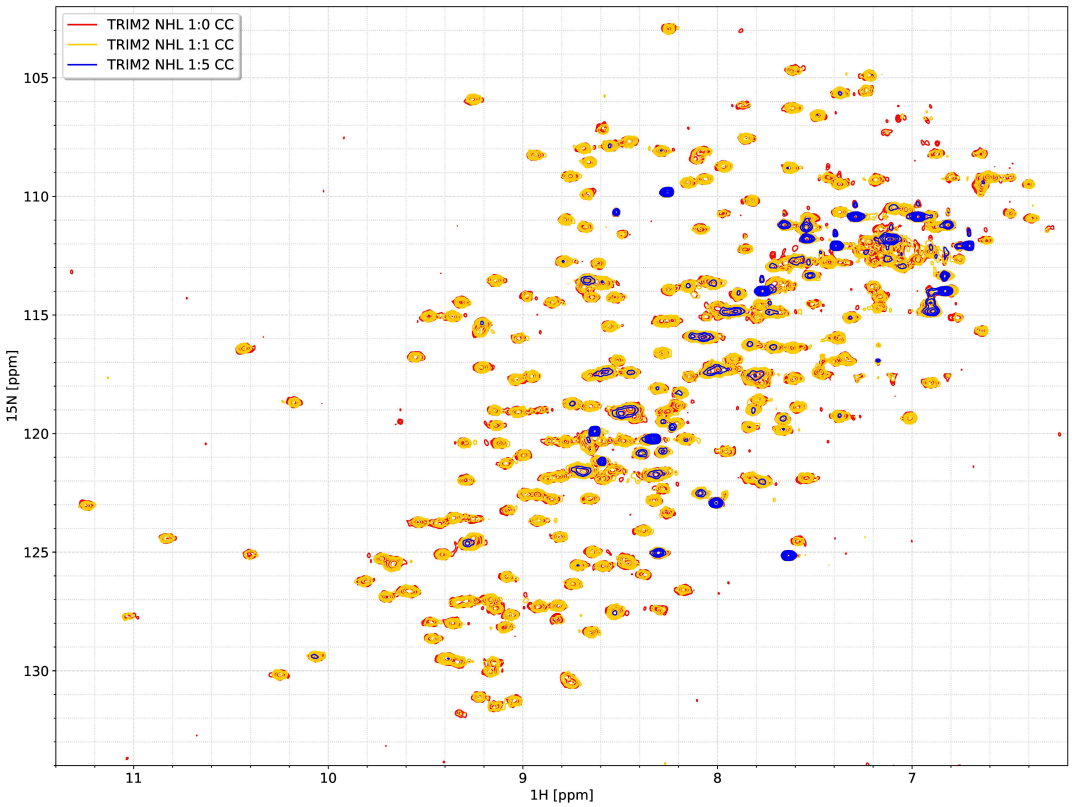


Figure S2: Titration of a TRIM2-coiled-coil construct (155-290), prepared by a collaborator against 100μM TRIM2-NHL concentration in TRIM2 NHL RNA titration buffer (see Table 2 page 2).

Motif	Fraction present in enriched sequences	Fraction present in non-enriched sequences	p-value	E-value
CGSCBC	13786/17563	28534/40523	9.4e-093	8.6e-088
GCBGCG	8621/17563	16610/40523	5.0e-073	4.5e-068
GCCG	11564/17563	23837/40523	6.7e-058	6.0e-053
CGMGS	11122/17563	22864/40523	6.7e-055	5.9e-050
GHSCG	11572/17563	24297/40523	3.6e-042	3.1e-037
GCG	14467/17563	31433/40523	3.9e-040	3.3e-035
CYYCG	10649/17563	22418/40523	6.5e-033	5.1e-028
CYCG	10588/17563	22436/40523	1.7e-028	1.3e-023
CGS	14063/17563	31091/40523	1.6e-019	1.2e-014
GGGRG	8938/17563	19271/40523	8.1e-014	5.4e-009
CYCCUC	4457/17563	9220/40523	5.1e-012	3.3e-007
HCG	12228/17563	27139/40523	1.6e-010	1.0e-005
GAGSAG	4881/17563	10478/40523	6.8e-007	3.8e-002
CYAUUUU	698/17563	1285/40523	7.6e-007	4.2e-002

Table S1: Table of motifs found by DREME in the 5'UTRs of RNAs enriched in a RIP-seq experiment against TRIM3. Non standard letter codes are as follows: B=CGU, D=AGU, H=ACU, K=GU, M=AC, R=AG, S=CG and Y=CU. E-values are the enrichment p-value times the number of candidate motifs tested, this calculation excludes sites that match a previously found motif.

Motif	Fraction present in enriched sequences	Fraction present in non-enriched sequences	p-value	E-value
UUUKUAA	7993/15484	15968/36586	1.1e-062	1.1e-057
CGHC	13447/15484	29683/36586	9.9e-059	9.1e-054
CCGDG	10505/15484	22100/36586	8.0e-059	7.0e-054
ADCG	13043/15484	28600/36586	2.0e-058	1.7e-053
UGNACA	12493/15484	27156/36586	4.5e-058	3.7e-053

Table S2: Table of motifs found by DREME in the 3'UTRs of RNAs enriched in a RIP-seq experiment against TRIM3. Non standard letter codes are as follows: B=CGU, D=AGU, H=ACU, K=GU, M=AC, R=AG, S=CG and Y=CU. E-values are the enrichment p-value times the number of candidate motifs tested, this calculation excludes sites that match a previously found motif.

	TRIM3-Filamin 318-420
Data collection	
Wavelength (Å)	0.97625
Resolution (Å)	27.8-2.132 (2.208-2.132)
Space group	C 1 2 1
Cell dimensions	
a, *b*, *c* (Å)	56.02, 33.50, 58.84
α, β, γ (°)	90, 117.25, 90
Reflections	
Total	16042 (1189)
Unique	5379 (428)
Multiplicity	3.0 (2.7)
Completeness (%)	96.67 (77.72)
I/σ(I)	3.82 (1.15)
Wilson B factor (Å^2)	24.5
Data quality	
R-merge	0.2174 (0.9505)
R-meas	0.265 (1.178)
R-pim	0.1495 (0.6867)
CC 1/2	0.966 (0.482)
CC*	0.991 (0.806)
Refinement	
Reflections used in refinement	5129 (393)
Reflections used for R-free	247 (36)
R_{work}/R_{free}	0.2270/0.2575 (0.2670/0.3277)
CC_{work}/CC_{free}	0.936/0.924 (0.708/0.711)
Number of non-hydrogen atoms	
Total	851
Protein	740
Ligand/ion	0
Water	111
Protein residues	104
R.M.S deviations from ideality	
Bond lengths (Å)	0.001
Bond angles (°)	0.450
Ramachandran statistics	
Favoured (%)	97.06
Allowed (%)	2.94
Outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	1.36
Average B-factor (Å^2)	28.55
Macromolecule	28.05
Solvent	31.89

Table S3: Statistics for crystal data and model of TRIM3-Filamin. Solved by molecular replacement using the PDB structure 4UMG (Tocchini et al., 2014) all parentheses refer to highest resolution shell. Clashscore is the number of all-atom steric clash overlaps $\geq$0.4Å per thousand atoms as defined in Chen et al. (2010).

	TRIM2-Fil-NHL	TRIM3-Fil-NHL
Sample details		
Organism	*Homo Sapiens*	*Homo Sapiens*
Source	*E. coli* BL21(DE3)	*E. coli* BL21(DE3)
Uniprot sequence ID (boundaries)	Q9C040-1 (318-744)	O75382-1 (301-744)
Description	TRIM2 Filamin and NHL domains (318-744)	TRIM3 Filamin and NHL domains (301-744)
Molecular mass from chemical composition (Da)	46 099	48 047
Loading concentration (mg/mL)	0.14-17.34	1.25-10
Concentration (μM)	3-376	26-208
Injection volume (μL)	30	30
Solvent composition	50mM Tris pH7.8, 200mM NaCl, 0.5mM TCEP	50mM Tris pH7.8, 200mM NaCl, 0.5mM TCEP
Collection parameters		
Source and instrument	Hamburg PETRA-III P12 with Dectris Pilatus 6M	Hamburg PETRA-III P12 with Dectris Pilatus 6M
Wavelength (Å)	1.23	1.24
Sample-detector distance (m)	3	3
q-measurement range	0.002260-0.7405	0.002974-0.7267
Radiation damage monitoring	frame by frame comparison	
Exposure time (s)	0.045x20	0.145x45
Sample configuration	sample changer with flow through capillary measurement	
Sample temperature (°C)	20	20
Software employed		
SAXS data processing	Data reduction using BsxCube, solvent subtraction and merging using PRIMUS/qt from ATSAS 2.8.4	
Basic analyses	Guinier, P(r), V_P calculated with PRIMUS/qt from ATSAS 2.8.4	
Atomic structure modelling	CRYSOL 2.8.3 and EOM 2.1	
Structural parameters		
Guinier analysis		
I(0) (raw)	39473.8 $\pm$ 117.75	0.070 $\pm$ 0.00029
R_g (Å)	33.4 $\pm$ 2.8	31.4 $\pm$ 0.2
qR_g max	1.26	1.17
Coefficient of correlation R^2	0.79	0.76
P(r) analysis from AUTOGNOM		
I(0) (raw)	39590	0.06857 $\pm$ 0.0001664
R_g (Å)	34.9	31.0 $\pm$ 0.7
D_{max} (Å)	122.1	94.4
q range ($Å^{-1}$)	0.007606-0.2394	0.01239-0.2588
Quality estimate from GNOM	0.6704	0.7947
Porod volume ($Å^{-3}$)	71 820	75 310

Table S5: Details of SAXS samples

	TRIM2-NHL 469-744
Data collection	
Wavelength (Å)	0.8729
Resolution (Å)	39.7-1.8 (1.864-1.8)
Space group	P 2_1 2_1 2_1
Cell dimensions	
a, *b*, *c* (Å)	44.177, 83.197, 132.758
α, β, γ (°)	90, 90, 90
Reflections	
Total	209482 (19807)
Unique	46038 (4526)
Multiplicity	4.6 (4.4)
Completeness (%)	99.49 (99.56)
I/σ(I)	6.72 (1.24)
Wilson B factor (Å^2)	27.06
Data quality	
R-merge	0.1375 (1.092)
R-meas	0.1552 (1.238)
R-pim	0.07004 (0.5696)
CC 1/2	0.99 (0.488)
CC*	0.998 (0.81)
Refinement	
Reflections used in refinement	44752 (4400)
Reflections used for R-free	1286 (126)
R_{work}/R_{free}	0.1909/0.2228 (0.3261/0.3361)
CC_{work}/CC_{free}	0.967/939 (0.614/0.635)
Number of non-hydrogen atoms	
Total	4605
Protein	4256
Ligand/ion	0
Water	349
Protein residues	560
R.M.S deviations from ideality	
Bond lengths (Å)	0.014
Bond angles (°)	1.20
Ramachandran statistics	
Favoured (%)	94.42
Allowed (%)	4.68
Outliers (%)	0.90
Rotamer outliers (%)	0.00
Clashscore	4.08
Average B-factor (Å^2)	37.08
Macromolecule	36.75
Solvent	41.07
Number of TLS groups	11

Table S4: Statistics for crystal data and model of TRIM2-NHL. Solved by molecular replacement using the PDB structure 1Q7F (Edwards et al., 2003) all parentheses refer to highest resolution shell. Clashscore is the number of all-atom steric clash overlaps $\geq$0.4Å per thousand atoms as defined in Chen et al. (2010). Note that the high proportion of non-favoured Ramachandran angles is due to the geometry of the loops on the top surface (Figure S3).

RNA	Number of peaks tracked	Of which backbone peaks	Top protein:RNA ratio reached	Percent protein bound at point closest to 50%
CCC CCC	244	242	1:3.5	50%
CCA CCC	246	244	1:4.5	41%
ACC CCC	244	242	1:4.5	52%
CAC CCC	245	243	1:4.0	52%
CCC CCC C	244	242	1:4.25	43%
CCU CCC	243	241	1:4.0	50%
CCG CCC	243	241	1:4.5	47%
CCG CCG	209	208	1:3.5	Not estimated
CCG	247	245	1:5.5	36%
GGC CCC	244	242	1:4.75	52%
AAC CCG	246	244	1:5.5	45%
CCG AAC	245	243	1:9.0	50%
AAU AAU	244	242	1:4.5	49%
UUU UUU	246	244	1:6.0	48%
AAC AAC	243	241	1:3.5	51%
AAA AAA	245	243	1:2.0	15%
ACA ACA	244	242	1:4.0	42%
ACA CCC	244	242	1:4.0	51%
ACA CCG	245	243	1:4.0	47%
GAC CCC	244	242	1:4.0	48%
GCG CCC	244	242	1:5.0	52%
GCG CCG	242	240	1:5.0	45%
GAC CCG	244	242	1:5.0	50%
GAC CGC	242	240	1:5.0	53%
CCC C	preliminary:27	preliminary:27	1:3.5	Not estimated
CCC	preliminary:27	preliminary:27	1:4.0	Not estimated

Table S6: Number of peaks tracked in RNA titration assays against TRIM2-NHL note that the RNA:protein ratio reached in assays varied quite widely due to variability in the amounts of RNA provided by suppliers

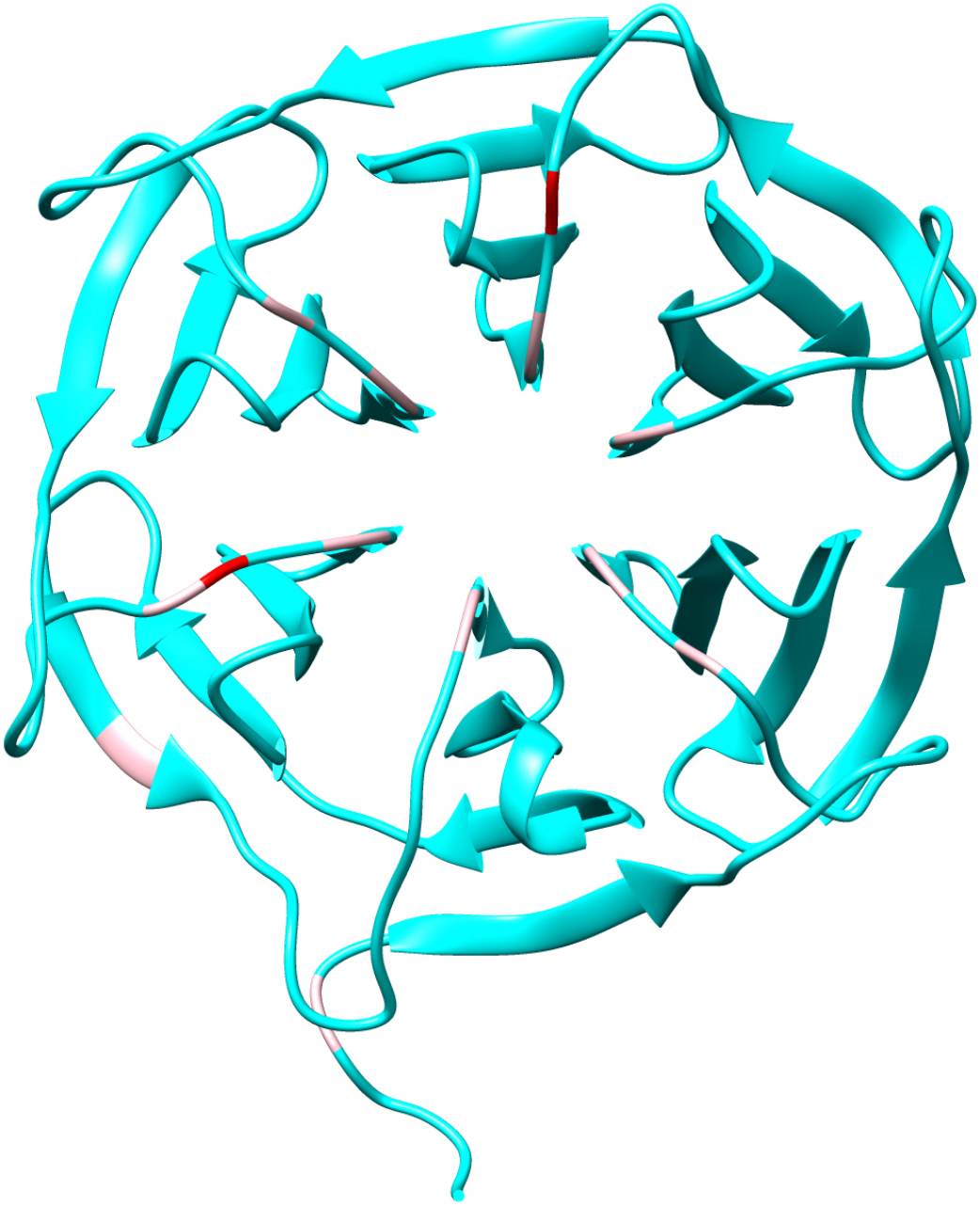


Figure S3: Location of residues in chain A of the TRIM2-NHL model whose Φ and Ψ angles fall in the outlier (red) or allowed (pink) region of the Ramachandran plot as defined by the Molprobity server (Williams et al., 2018)

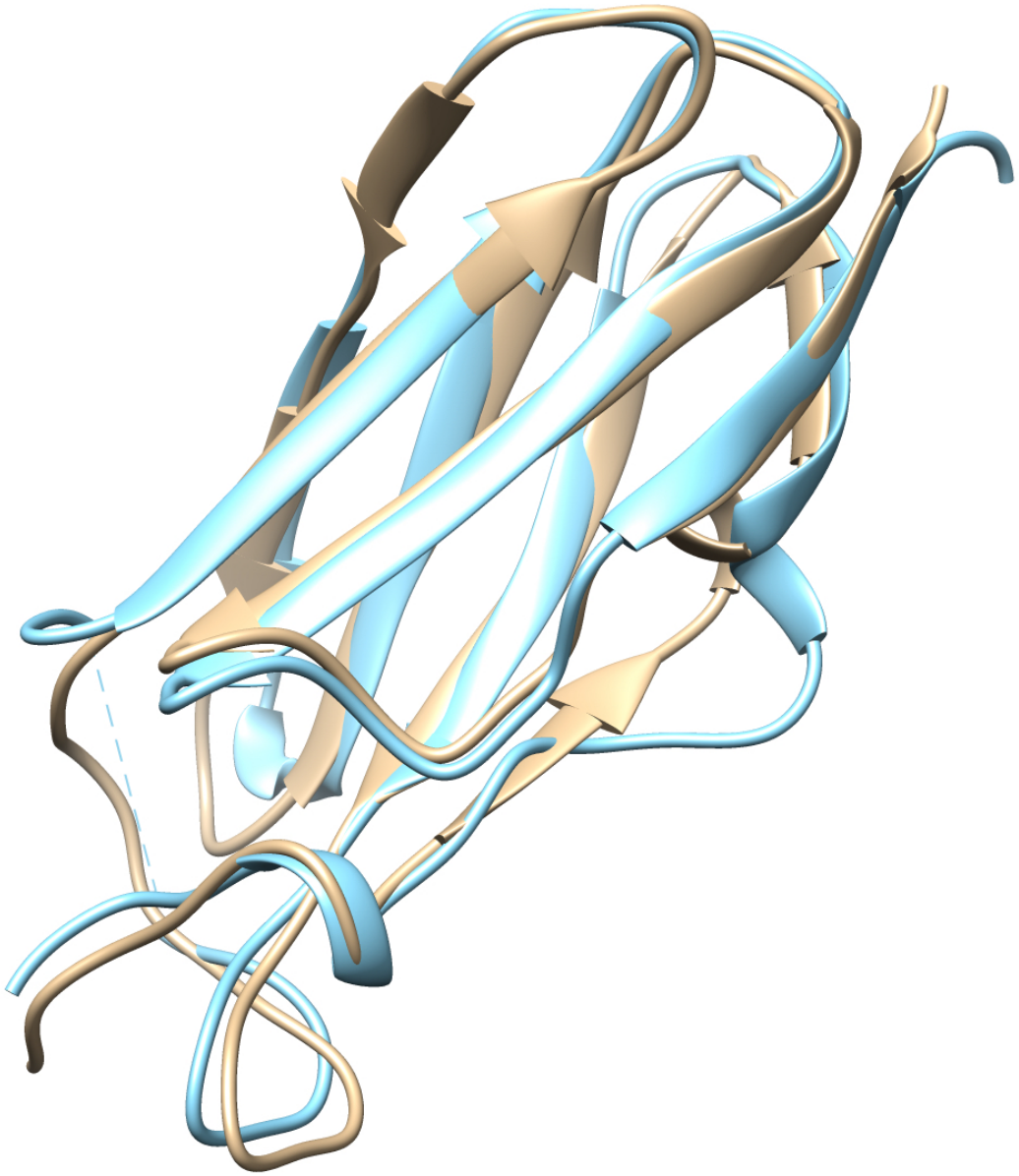


Figure S4: Comparison of the structures of TRIM3-Filamin as determined by us (**tan**) and that of the TRIM71-Filamin domain as determined by Tocchini et al. (2014) (PDB:4umg,**blue**).

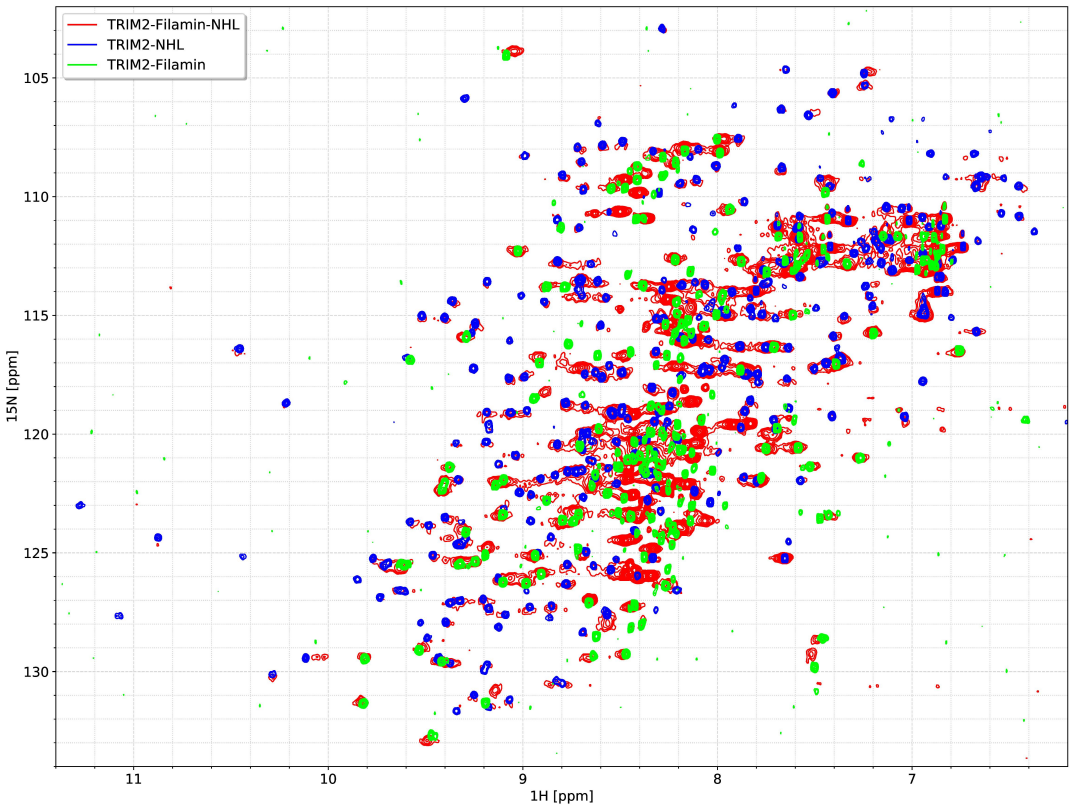


Figure S5: HSQC spectra for TRIM2 NHL (**blue**), Filamin (**green**) and Filamin-NHL (**red**) constructs

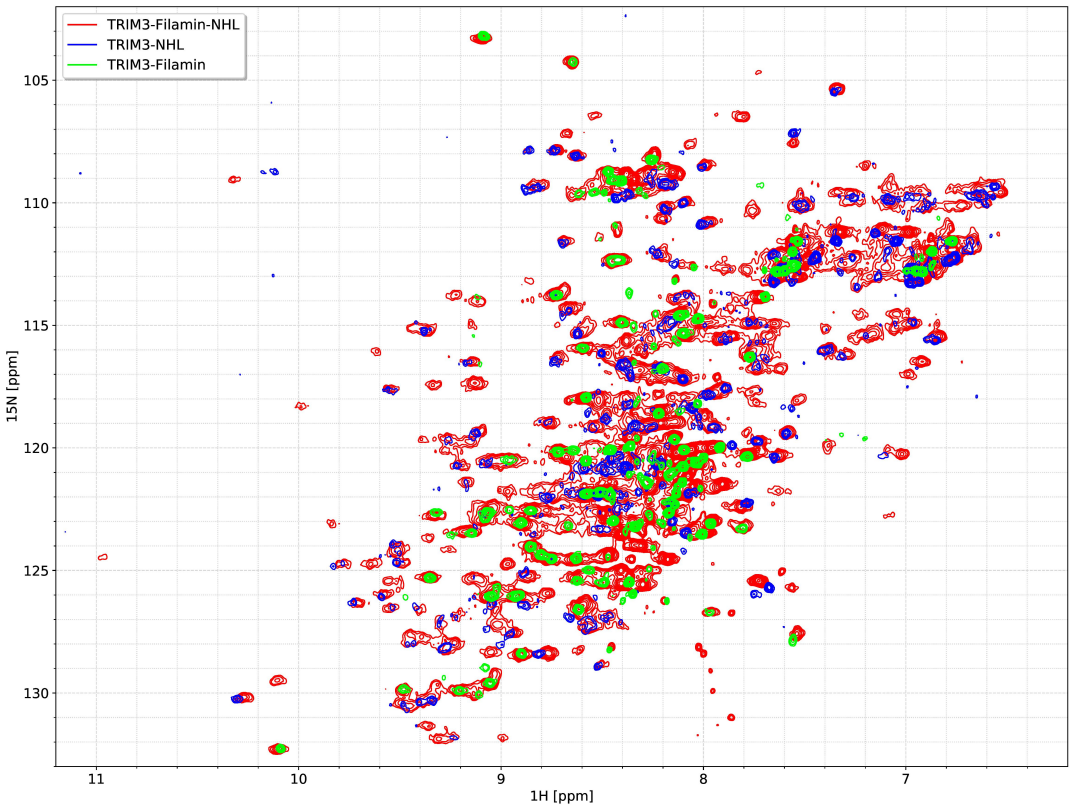


Figure S6: HSQC spectra for TRIM3 NHL (**blue**), Filamin (**green**) and Filamin-NHL (**red**) constructs

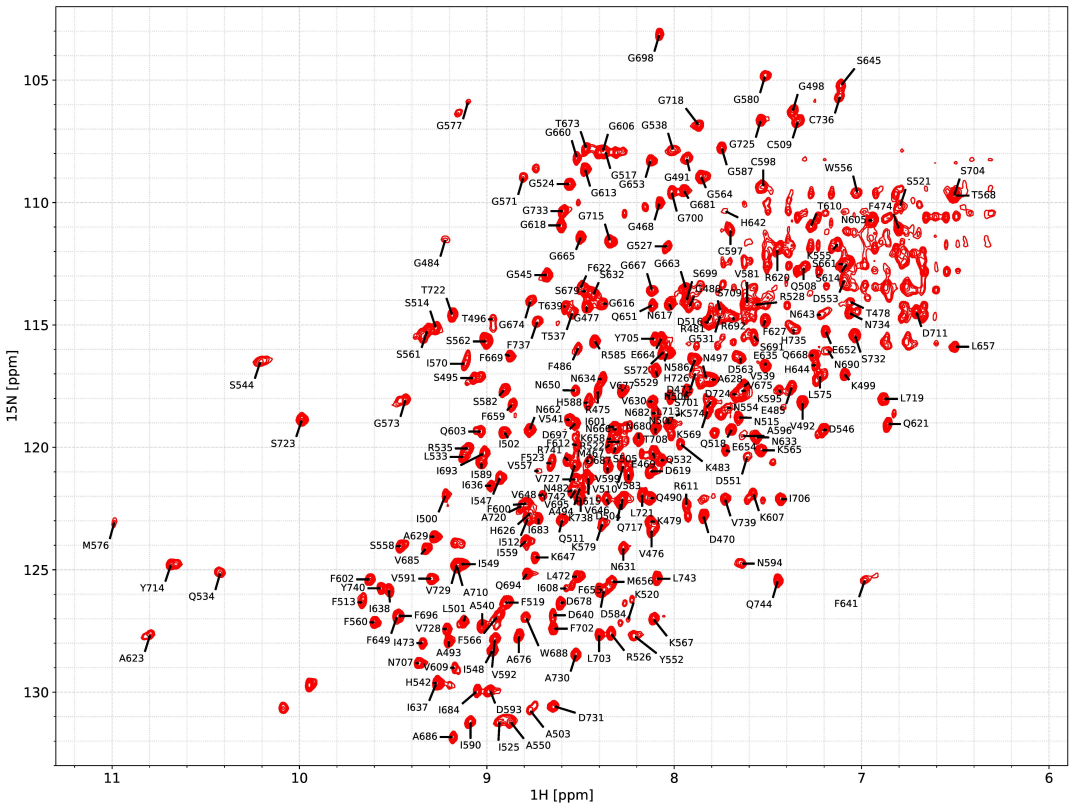


Figure S7: Assigned TRIM2-NHL ^{1}H-^{15}N-TROSY-HSQC with deuterated sidechains, acquired at 310K in NMR assignment buffer (see Table 2, page 38).

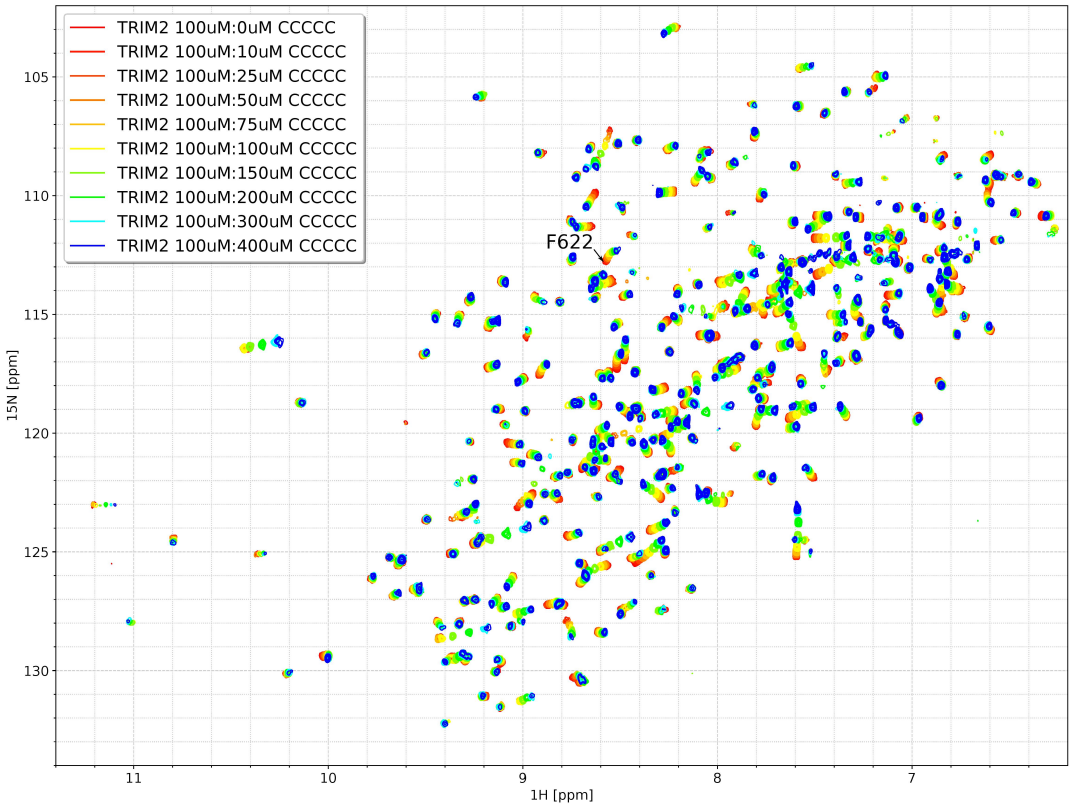


Figure S8: Titration of RNA C-5mer against TRIM2-NHL in TRIM2 NHL RNA titration buffer carried out at 298K. Note the curvature evident in the chemical shift perturbations of certain peaks (for example the peak associated with residue F622 which has a starting at position of roughly 8.6-113ppm) this is evidence of two separate binding events occurring simultaneously: the resultant CSP at each titration point is the addition of the CSP associated with the first binding event and the CSP associated with the second binding event, each of which is associated with a different K_d, hence the curvature observed if we track the peaks through the titrations. In this case we attribute this to contamination with triethylamine detected in the 1D spectra associated with this titration.

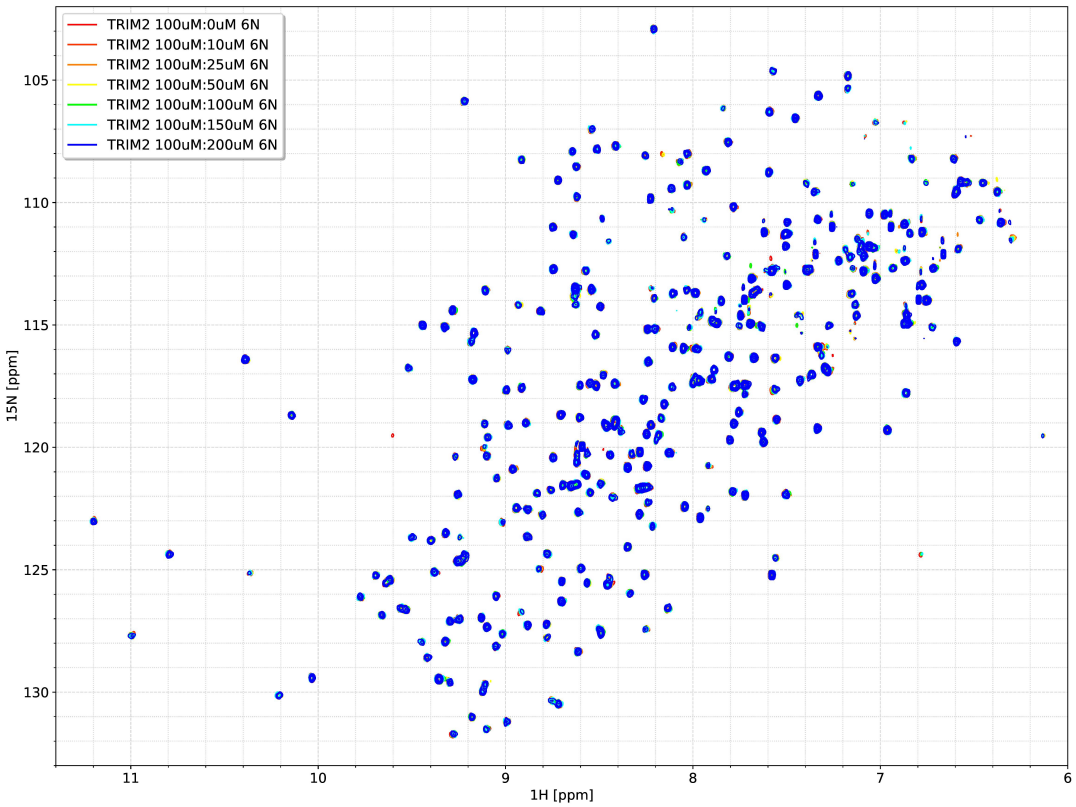


Figure S9: Titration of a randomised RNA N-6mer against TRIM2-NHL carried out at 298K. Note that this experiment was not carried out in the usual TRIM2-NHL RNA titration buffer but in TRIM2-NHL assignment buffer (see Table 2, page 38).

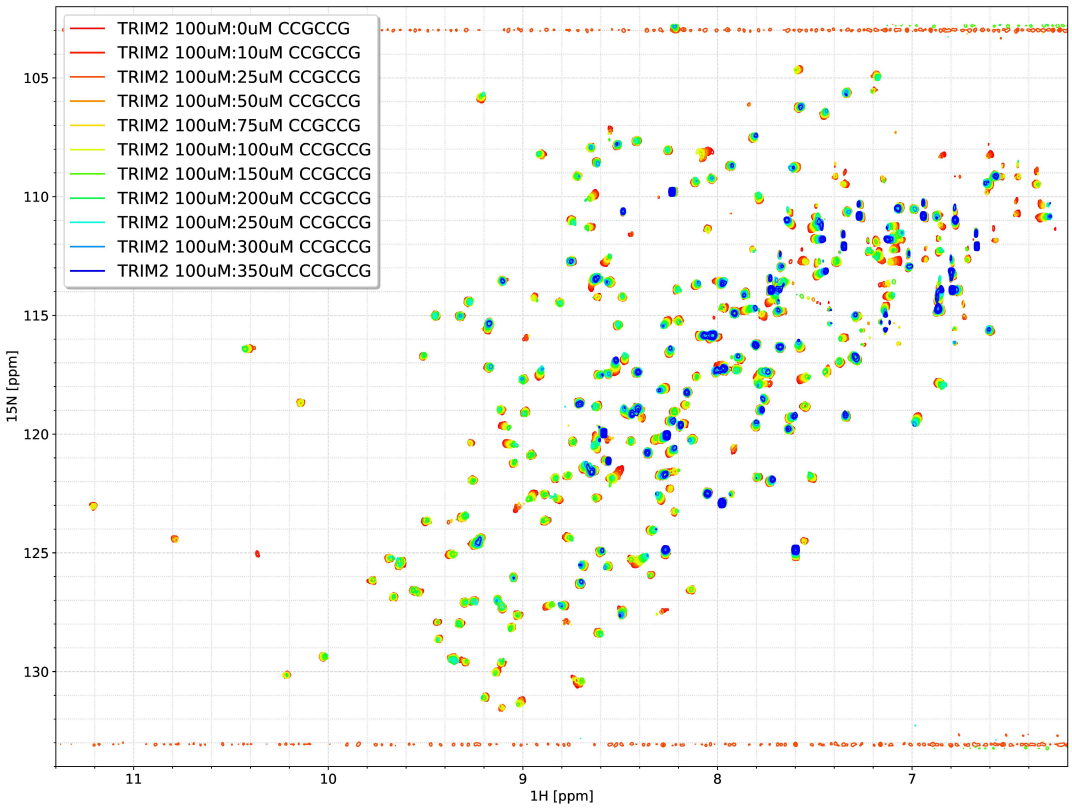
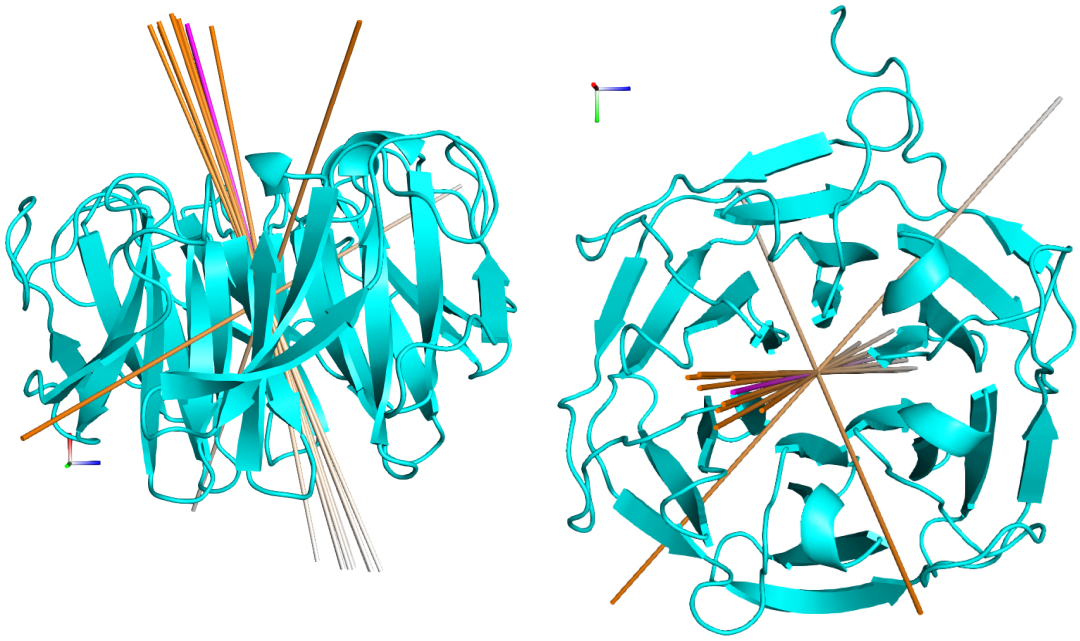


Figure S10: Titration of an RNA with sequence CCG CCG against TRIM2-NHL carried out at 298K. Note the loss of signal affecting almost all peaks.



(a) Viewed from side

(b) Viewed from top

Figure S11: Side and top view of the TRIM2-NHL domain with the axis of symmetry calculated by ROTDIF using either **purple**: all data or **orange**: data minus a randomly selected 10% of residues. The red, blue and green axes plotted in the corner correspond to those of the tensor predicted by ELM(Ryabov et al., 2006)

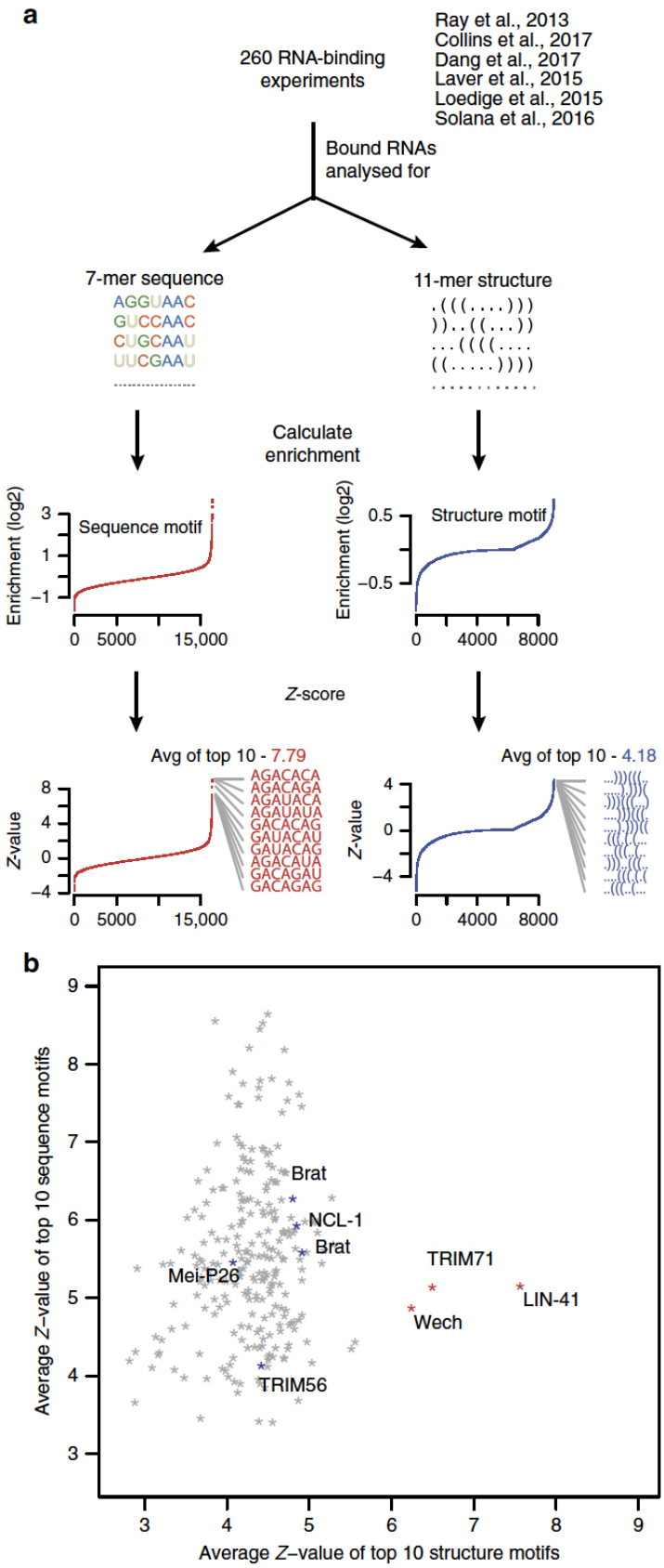


Figure S12: Figure from Kumari et al. (2018) showing their analysis to find NHL domains that preferentially bind structured RNA. Note that in this analysis, which does not include TRIM2 or TRIM3 but does include orthologues to TRIM2 and TRIM3, only TRIM71 and its orthologues show a strong preference for structured RNA.

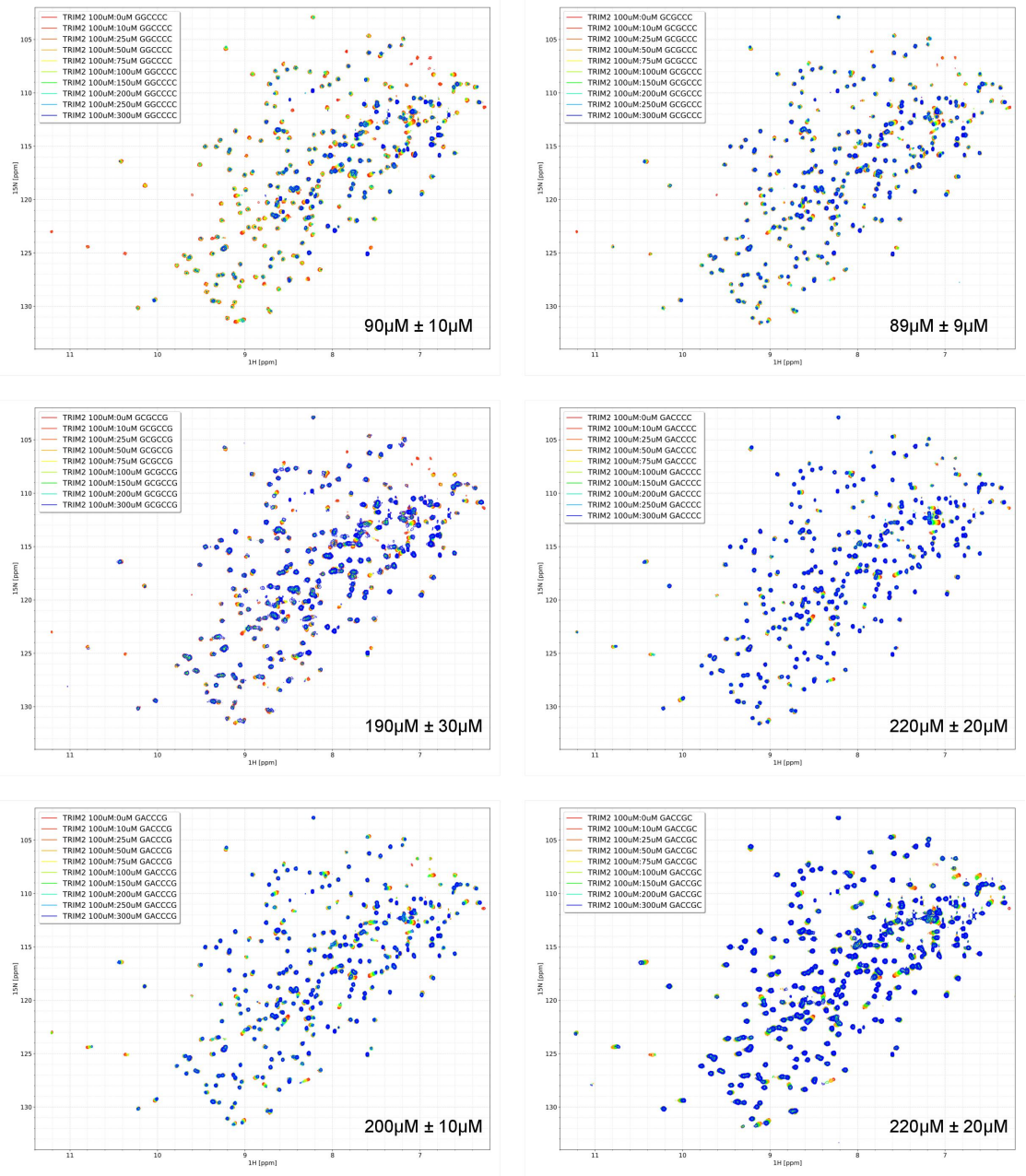


Figure S13: HSQC of titration of various RNAs with a leading G against TRIM2-NHL.

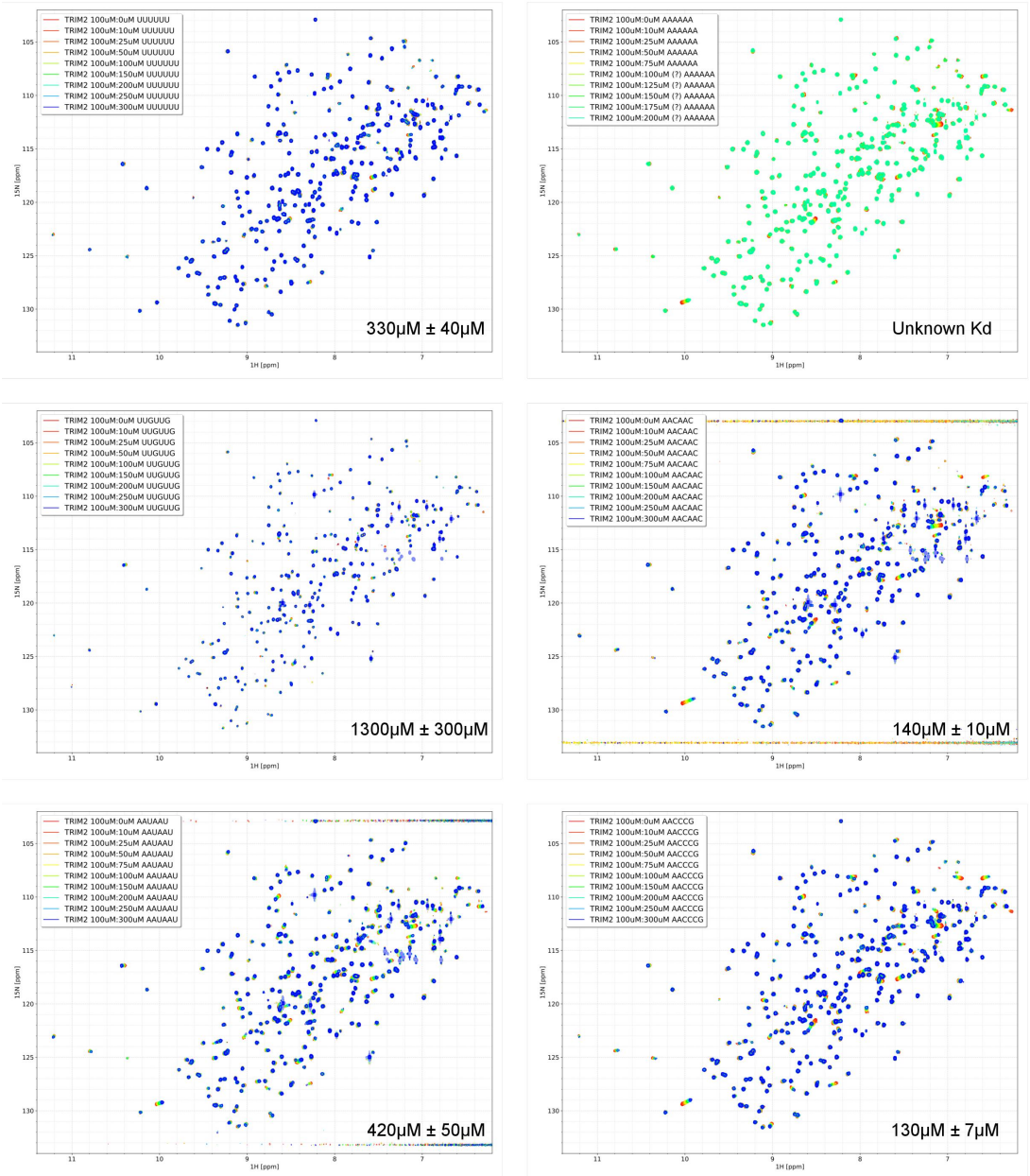


Figure S13: HSQC of titration of various A and U containing RNAs against TRIM2-NHL.

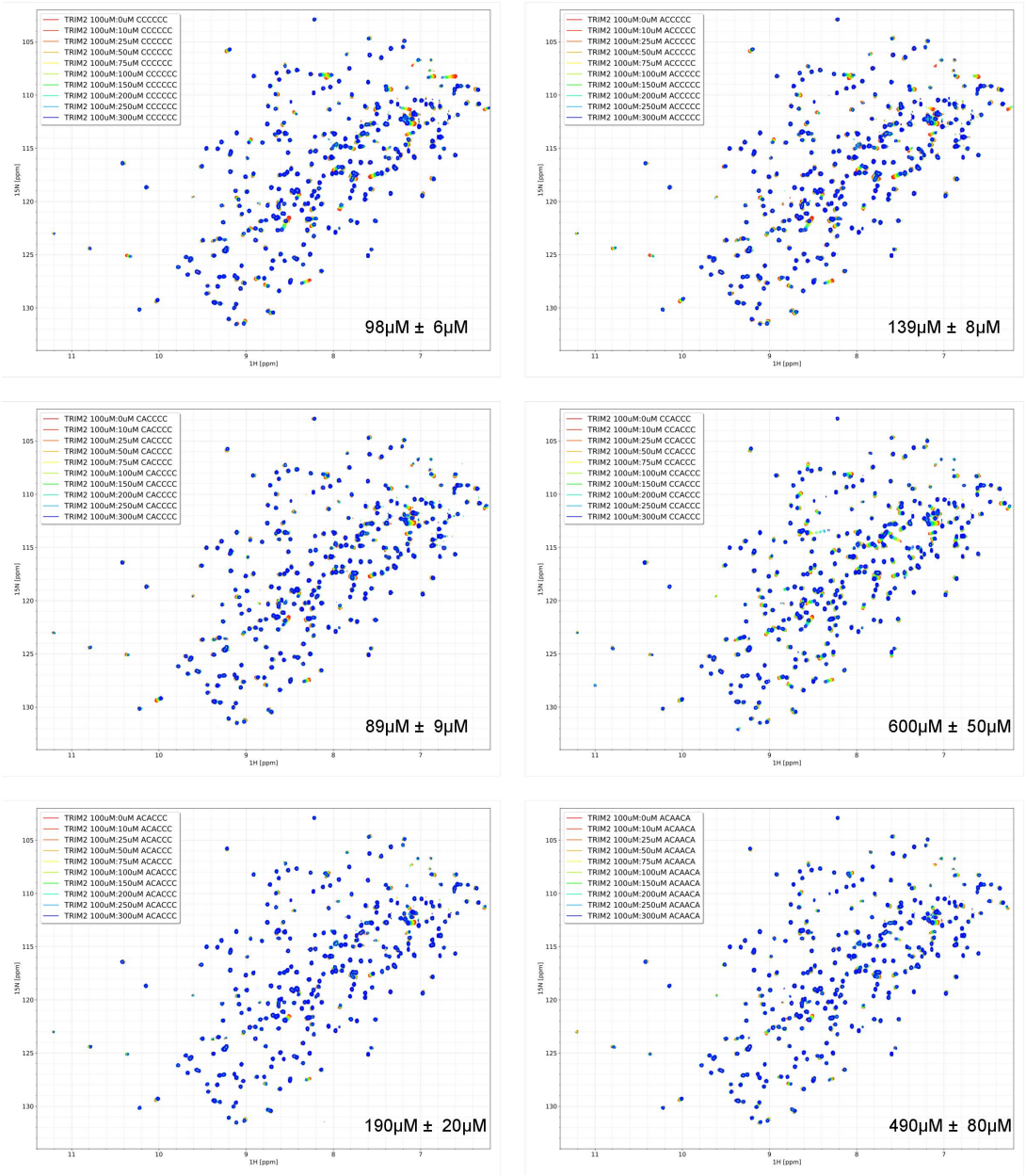


Figure S13: HSQC of titration of various A and C containing RNAs against TRIM2-NHL

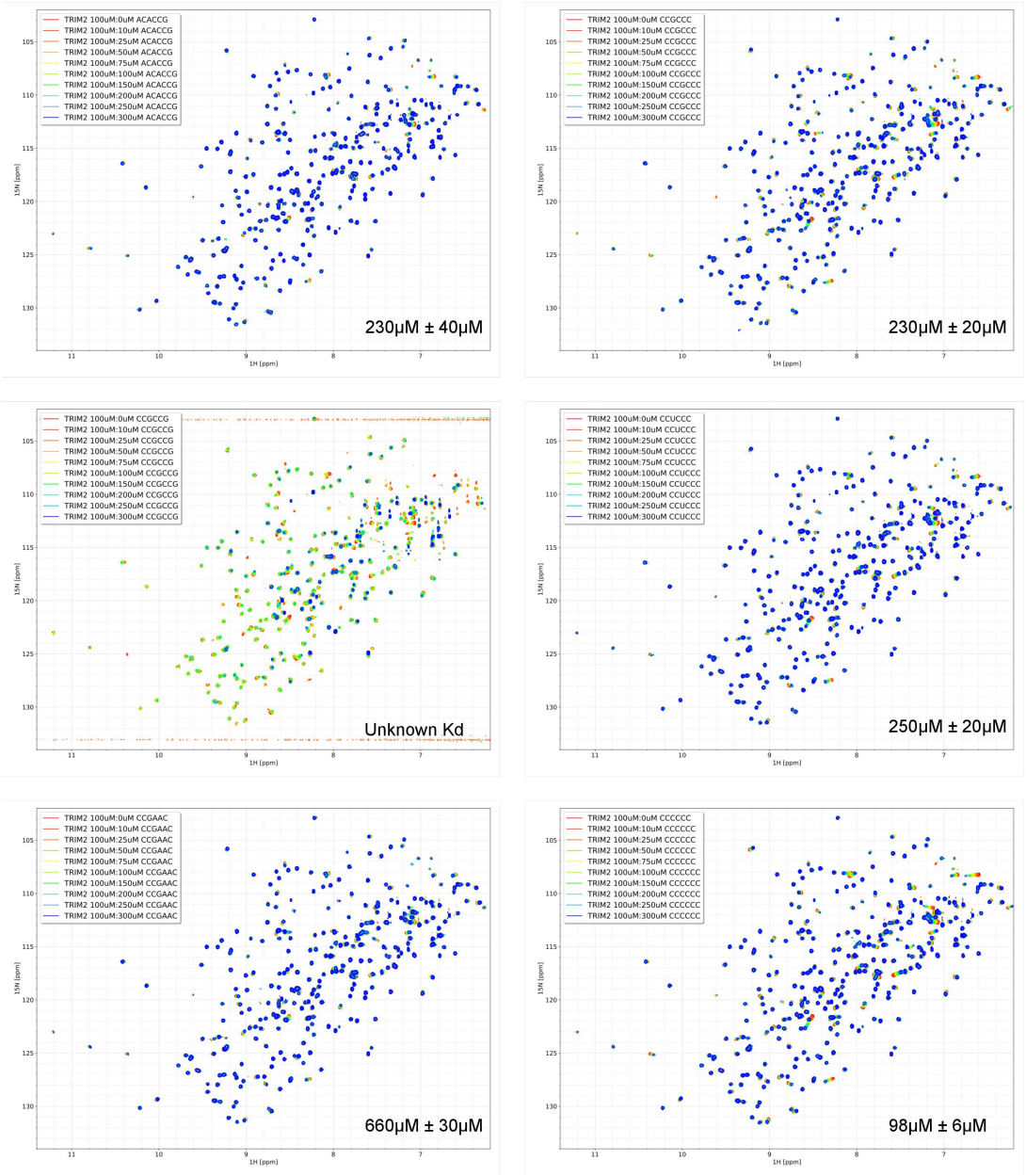


Figure S13: HSQC of titration of various RNAs containing the sequence CCG against TRIM2-NHL

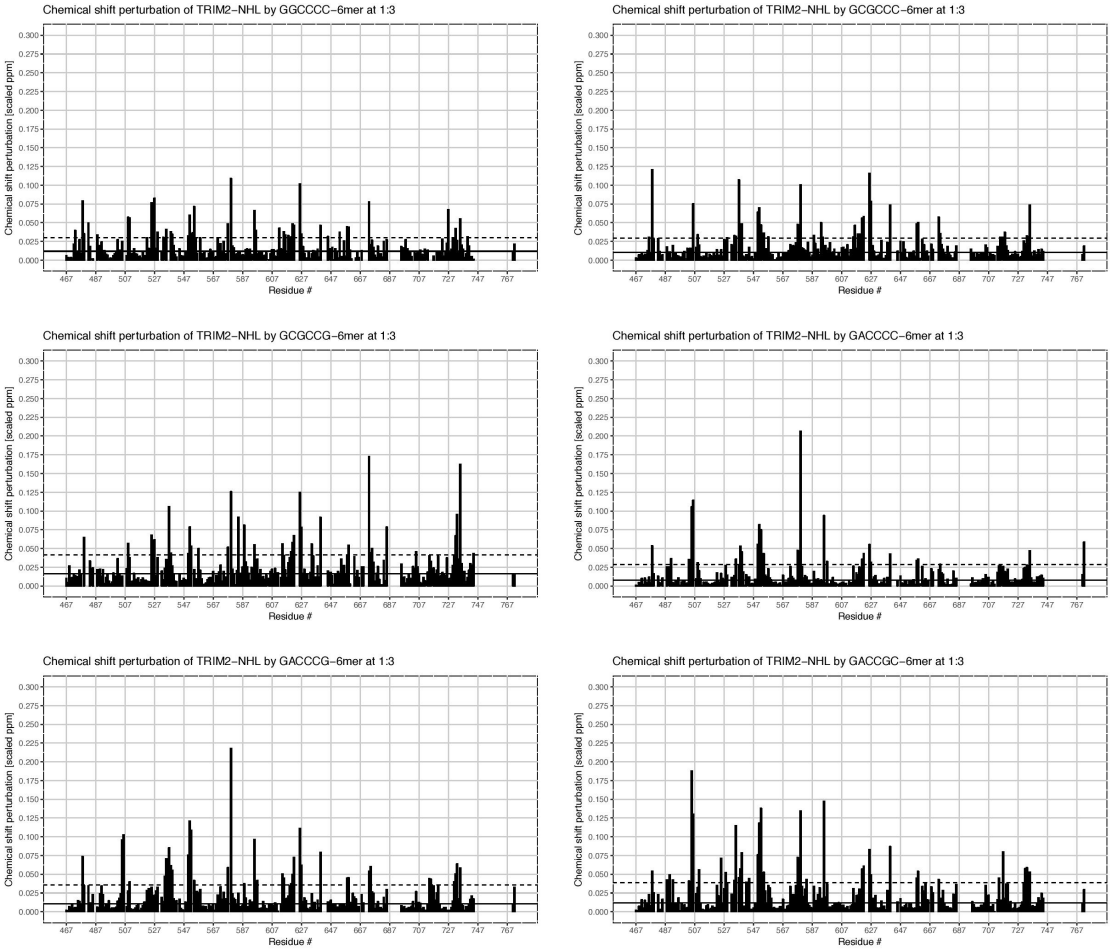


Figure S14: CSPs from titrations of various RNAs with a leading G against TRIM2-NHL

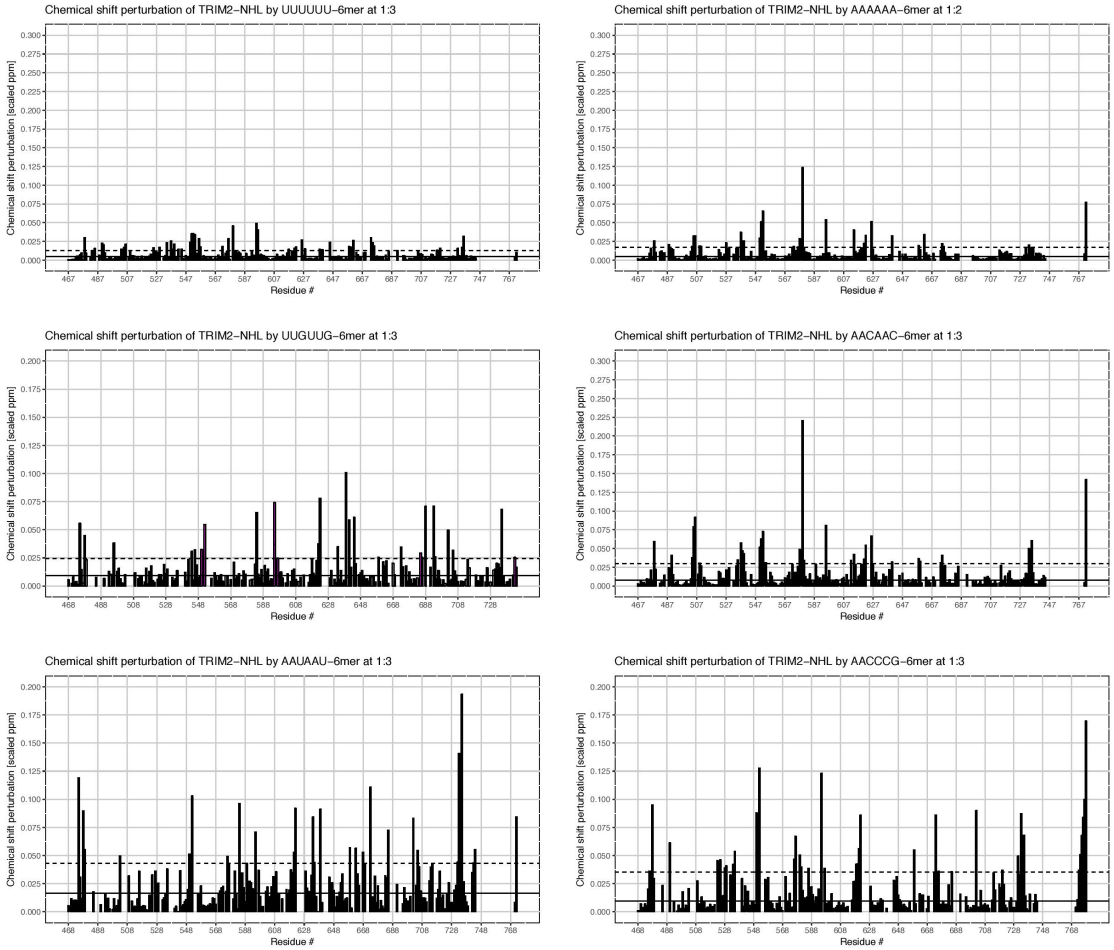


Figure S14: CSPs from titrations of various A and U containing RNAs against TRIM2-NHL

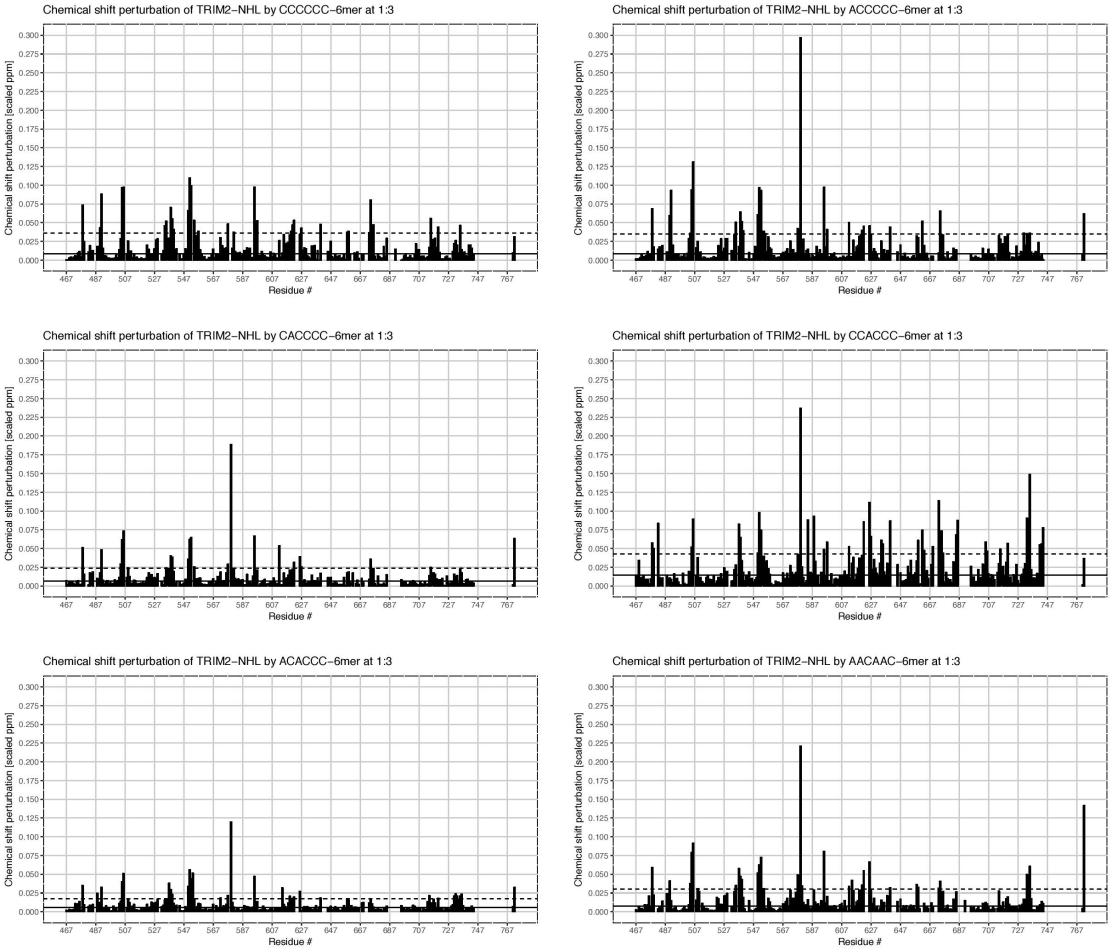


Figure S14: CSPs from titrations of various A and C containing RNAs against TRIM2-NHL

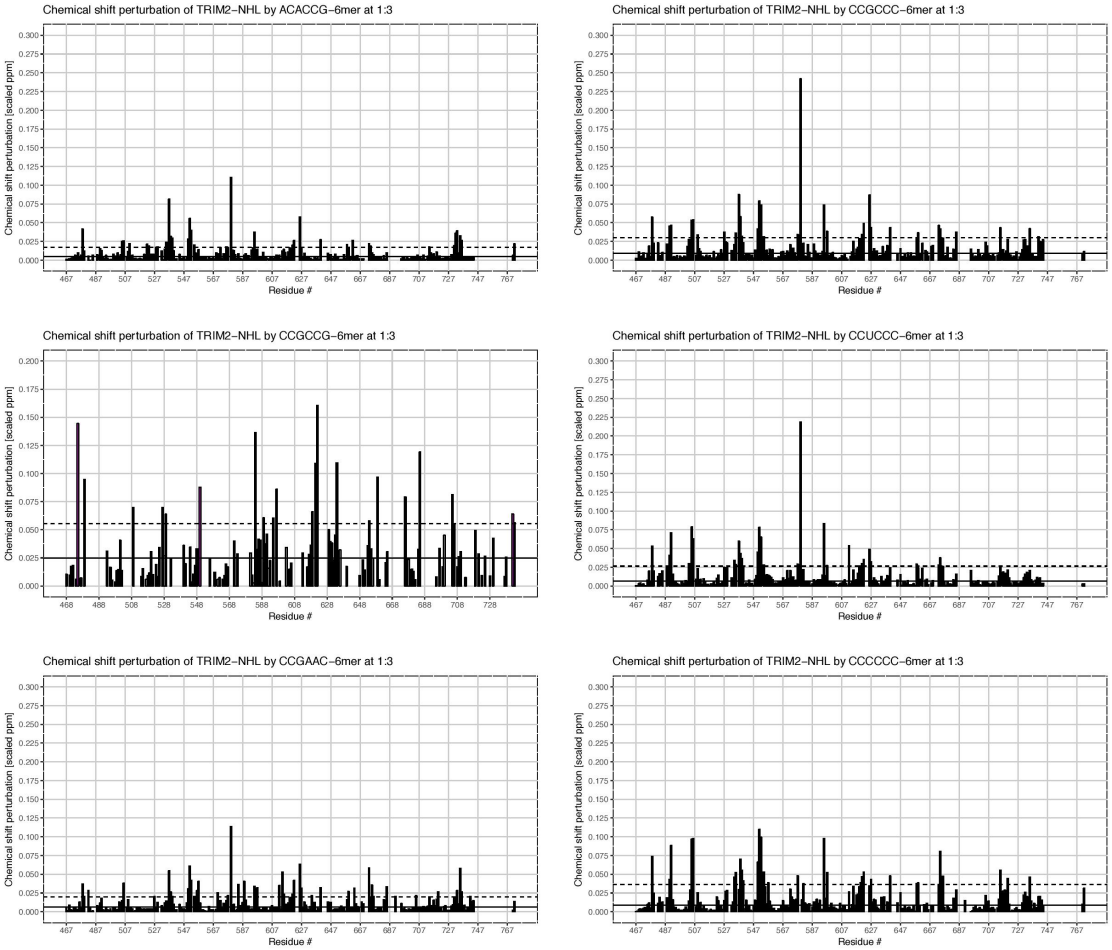


Figure S14: CSPs from titration of various RNAs containing the sequence CCG against TRIM2-NHL

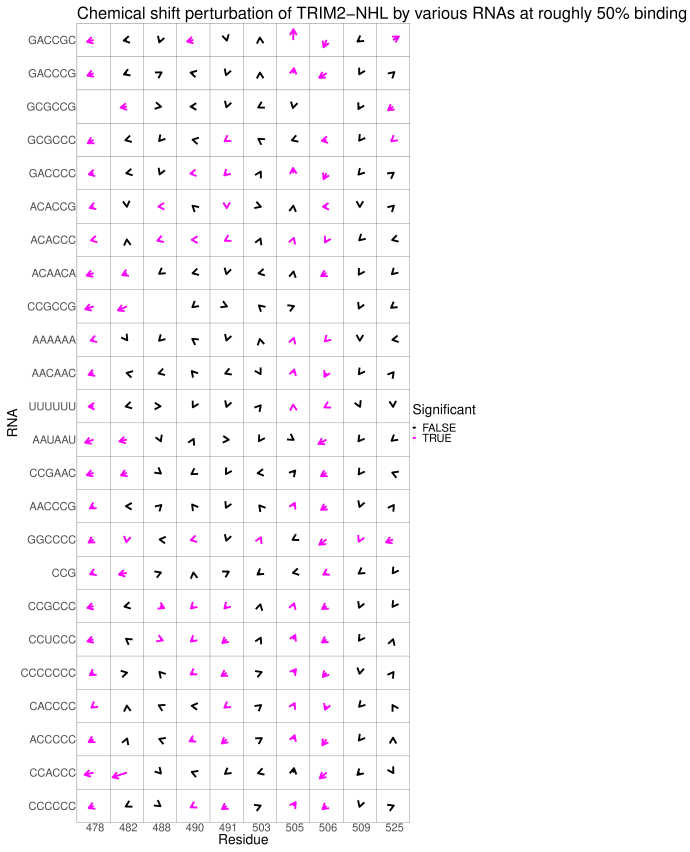


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

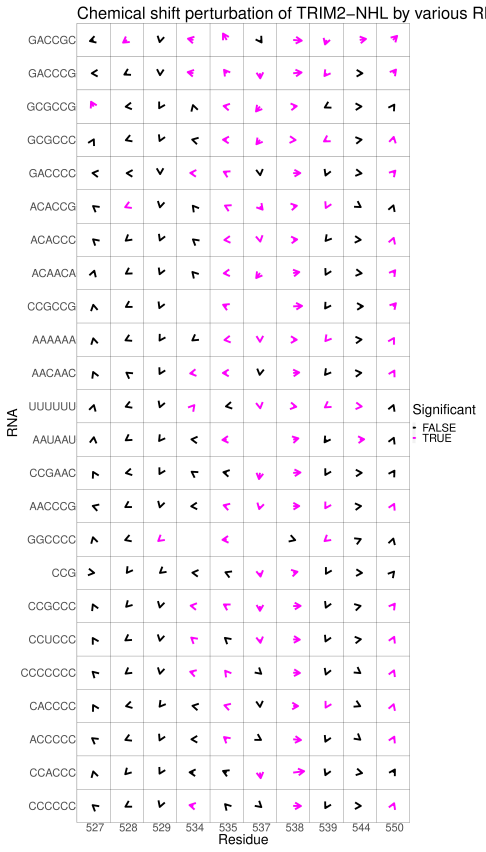


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

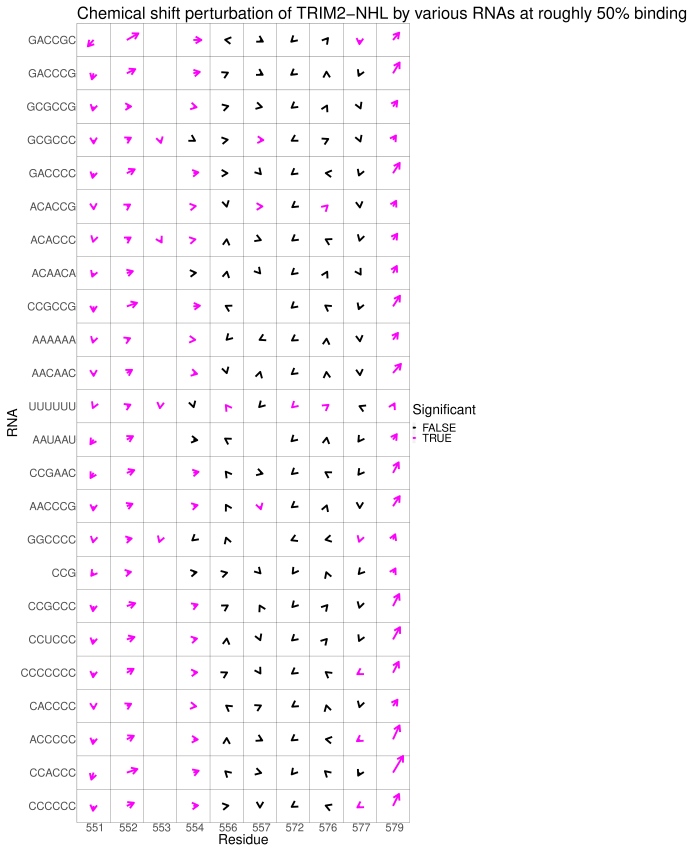


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

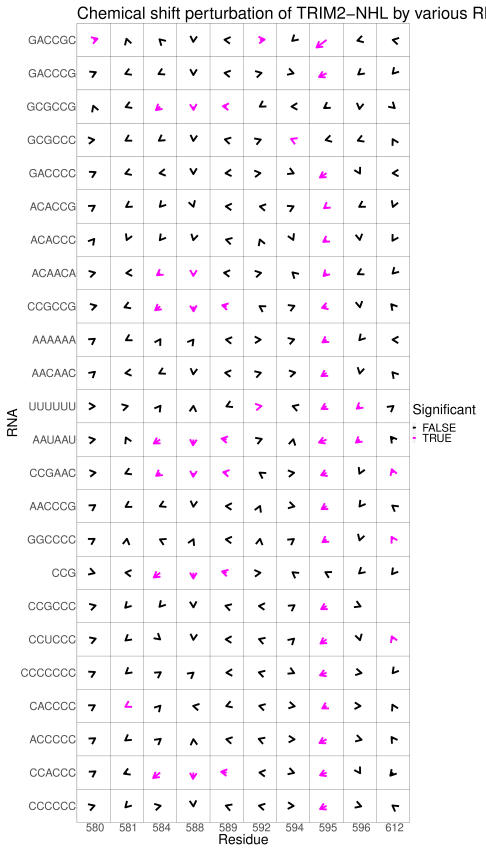


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

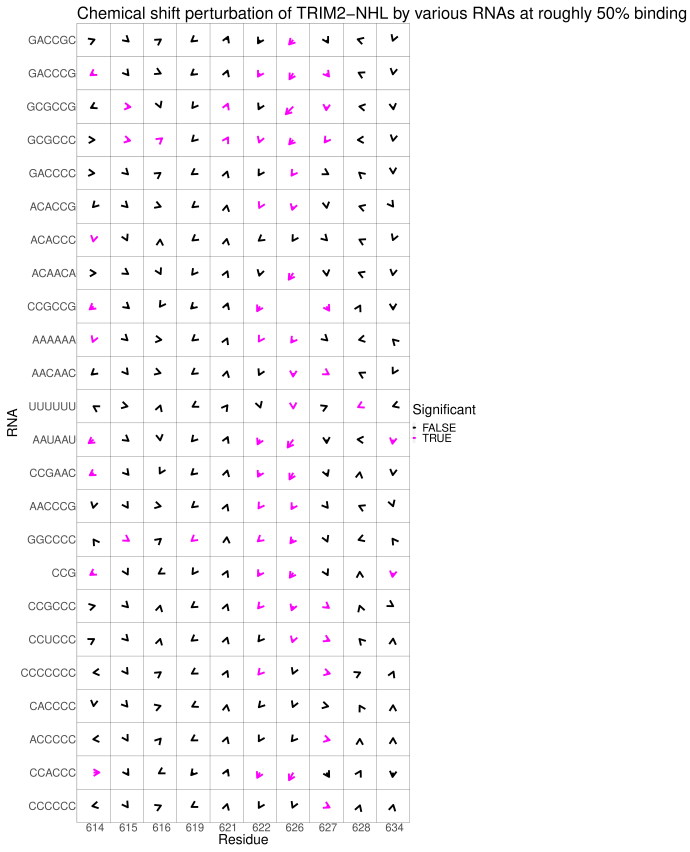


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

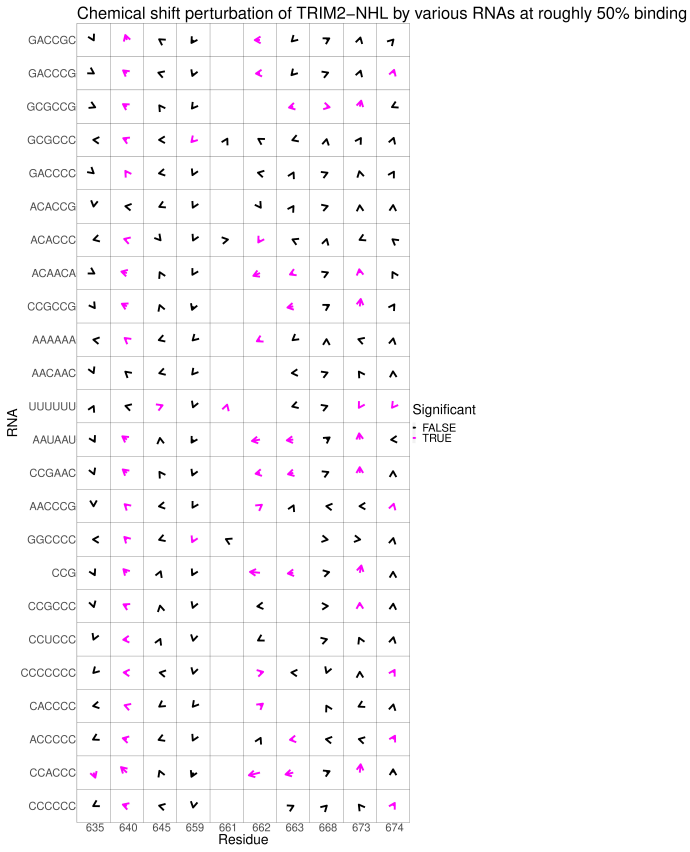


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

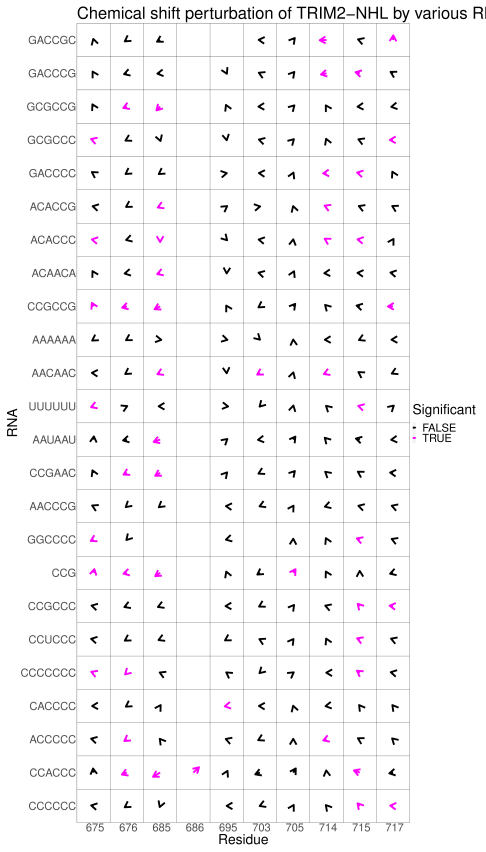


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.

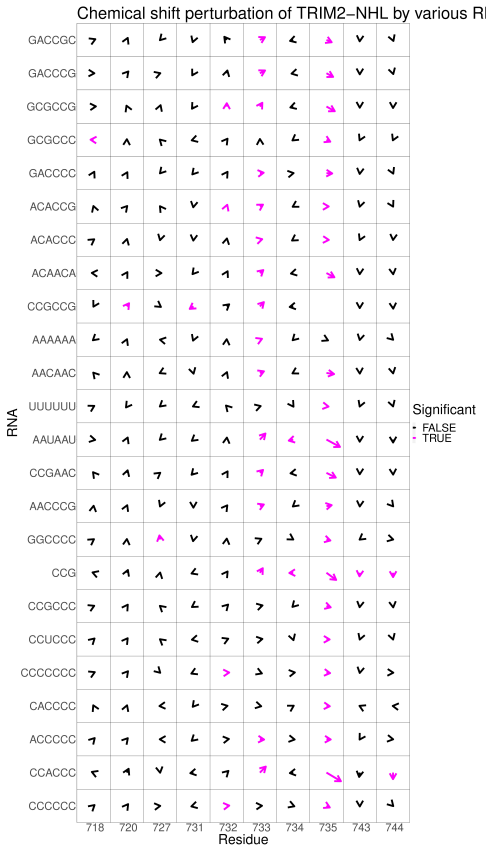


Figure S15: Plot of chemical shift perturbations upon titration of various RNAs against TRIM2-NHL. CSP is shown for the data point closest to that giving 50% bound protein. Dimensions were scaled as described in Section 4.9 then arrows were scaled such that the largest CSP of all RNAs studied has a length equal to half the side of one square. Arrows are coloured purple if they represent chemical shift perturbations significantly larger than the average CSP observed at the given data point for the given RNA; the cutoff being average plus one standard deviation.