Structural insights into mimicry of TGF-β signaling by the murine parasite H. polygyrus

by

Ananya Mukundan

Bachelor of Science, University of Michigan, 2016

Submitted to the Graduate Faculty of the

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Ananya Mukundan

It was defended on

June 1, 2023

and approved by

Gordon Rule, Professor, Department of Biological Sciences

Marijn Ford, Associate Professor, Department of Cellular Biology

Greg Delgoffe, Associate Professor, Department of Immunology

Dissertation Director: Andrew Hinck Professor, Department of Structural Biology

Structural insights into mimicry of TGF-β signaling by the murine parasite *H. polygyrus*

Ananya Mukundan, PhD

University of Pittsburgh, 2023

TGF-β is a key regulator of the adaptive immune system, acting to convert naïve CD4+ T-cells into Foxp3+ T_{regs}. The mouse intestinal hookworm *Heligmosomoides polygyrus* takes advantage of this pathway, surviving in its host by using a five-domain secreted TGF-β mimic protein, TGM or TGM-1 to expand the population of T_{regs} in the host. TGM-1 binds to the TGF-β receptors directly and is part of a larger family of 10 secreted parasitic proteins (TGM-1:TGM-10), all of which lack sequence and structural homology to TGF-β, and are instead composed of five CCP domains (D1-D5). In this dissertation, biophysical methods were used to study TGM-1 and TGM-4, how they bind the TGF-β receptors and how to leverage these structural details towards the development of TGF-β-targeting therapeutics. SPR, ITC, and NMR showed that binding of TGM-1 to the TGF-β receptors is modular, with D1 and D2 binding to TβRI and D3 binding to TβRII. The domains of TGM-1 compete with TGF-β for binding to TβRI and TβRII, showing that TGM-1 mimics TGF-β both functionally and structurally. The solution structure of TGM-1 D3 and the TGM-1 D3:TβRII complex show that the TGM-1 D3 CCP domain is adapted to allow for expansion of C-terminal segment, partially exposing several hydrophobic residues that interact with the edged β-strand that TβRII also uses to bind TGF-β. Further studies revealed that domains 4 and 5 of TGM-1 bind a co-receptor CD44 that potentiates TGM-1 signaling *in vitro*. This co-receptor binds to TGM-1 and TGM-4 but not TGM-6, indicating that different TGM isoforms bind to distinct co-receptors depending on the domain 4-5 composition. This led to the

hypothesis that cell-targeted TGF-β inhibitors could be developed by fusing known TGF-β inhibitors to domains 4 and 5 of TGM-1. Initial tests into these fused proteins indicate inhibition of TGF-β signaling, with studies into cell-specific inhibition and potentiation to be performed in the future. This dissertation demonstrates how *H. polygyrus* has adapted a family of proteins to mimic TGF-β signaling and opens up unique avenues for inhibition of TGF-β signaling.

Table of Contents

List of Tables

List of Figures

1.0 Introduction: TGF-β signaling in physiologic and pathologic contexts

1.1 Overview of TGF-β family and signaling

1.1.1 Discovery of TGF-β family proteins

In 1980, Sporn and Todaro proposed "autocrine secretion", which hypothesized that malignant cells secrete polypeptides that act on the cells' own external receptors to self-stimulate growth[1]. Such polypeptides were designated as Transforming Growth Factors or TGFs, the first of which, sarcoma growth factor (SGF), was isolated from murine fibroblasts infected with an RNA virus[1,2]. TGFs were noted to both transform non-malignant cells into the neoplastic phenotype, thus allowing for growth in soft agar, and also to bind to epidermal growth factor (EGF) receptors[2,3]. The Sporn/Roberts lab in the late 1970s isolated TGFs from non-neoplastic adult murine tissues; murine tissue was frozen and the protein was extracted using acetic acid/ethanol and purified via a size-exclusion column in 1M acetic acid[3]. This purification led to the discovery that TGFs could be divided into two distinct families based on their function, with TGF-α binding to EGF receptors and TGF-β acting to induce phenotypic transformation of non-neoplastic cells[2]. At the same time, the Moses lab independently isolated TGF-β from murine fibroblast cell lines, using a combination of soft-agar growth assays and ion-exchange chromatography, the latter step separating TGF-α from TGF-β[2].

During the 1980s, a total of three isoforms of TGF-β were discovered, TGF-β1,TGF-β2, and TGF-β3. Rik Derynck was given large quantities of human platelet-produced TGF-β from the Sporn/Roberts lab. Using clostripain, an arginine-targeting protease, digestion he was able to

identify a continuous 55 amino acid sequence, along with several smaller sequences[2,4]. From this sequence, two complementary 44 residue probes, along with 16 smaller 14 residue probes, were generated and tested against a human genomic library[4]. Though extensive screening of multiple cell types was required, eventually a cDNA sequence corresponding to the mRNA of the TGF-β precursor was generated revealing a 112 amino acid biologically active protein from a 390 amino acid TGF-β precursor[4]. Following the discovery of TGF-β1, R&D systems attempted to purify TGF-β1 from porcine platelets. The final step of purification up to that point involved a reverse-phase chromatography step with a C18 column; however, when a C4 column was utilized, three protein peaks were noted over an acetonitrile gradient of 24%-40%, all with the expected molecular weight of biologically active TGF-β at ca. 25 kDa[5]. These three peaks corresponded to the TGF-β1 homodimer, a TGF-β1:TGF-β2 heterodimer, and a TGF-β2 homodimer. The discovery of TGF-β3 soon followed with cDNA cloning identifying the final TGF-β isoform, TGF-β3[6,7]. cDNA analysis by other labs revealed other proteins with high sequence homology to the TGF-βs, leading to what is now known as the 33-member TGF-β family.

1.1.2 Evolution of TGF-β family proteins

TGF-βs are functionally diverse and are essential for multiorgan (palate, lung, heart, etc.) development[8-11], regulation of both the pro and anti-inflammatory arms of the adaptive immune system[12-17], along with wound healing and angiogenesis through promotion of vascular fibrosis[18-20]. TGF-βs belong to a larger family of proteins including BMPs (bone morphogenic proteins) and GDFs (growth and differentiation factors) which are involved in embryonic development[21], skeletal development and repair, and several skeletal pathologic conditions[22]. Other proteins in the

family include activins and inhibins which activate and repress follicular-stimulating hormone (FSH) production respectively[23], as well as acting in erythropoiesis[24], adipose metabolism, and neuronal cell differentiation[25]. Lesser known members of the family include myostatin, which regulates skeletal muscle growth[26], and nodal, which is involved in left-right axis determination[27].

While the TGF-βs are only present in vertebrates, members of the TGF-β family have been found in all metazoans[28]. Evolutionarily, the BMPs emerged first, followed by the activins and inhibins, and finally the TGF-βs[29,30]. This is evidenced by the presence of the BMPs in both worm and fly species[31], the presence of activins and inhibins in flies only[32], and the presence of the TGF-βs in vertebrates[33]. Phylogeny of the TGF-β family proteins show that the BMPs and GDFs are the proteins from which other members of the family diverge, with the TGF-βs differentiating themselves relatively late, and the Lefty proteins being the most divergent (Fig. 1).

1.1.3 Structure of TGF-β family proteins

TGF-βs are homodimeric signaling proteins with an interchain disulfide bond connecting the two monomers. They signal through a heterotetrameric signaling complex, binding to two pairs of serine/threonine kinases, designated as the type I and type II receptors respectively[34,35] (Fig. 2). TGF-β first binds its type II receptor (TβRII), two receptors per TGF-β homodimer, and then only can two type I receptors (TβRI) bind the binary TGF-β:$T\beta RII_2$ complex[36]. Formation of the heterotetramer leads to autophosphorylation of the type II receptor, which then phosphorylates the type I receptor[37]. Signals are then propagated either by the canonical Smad proteins[38], or a non-Smad pathway: MAPK[39], PI3K[40], etc. (Fig. 2).

Structurally the TGF-β family proteins evolved from the cystine knot growth factor domain[33], which are characterized by two β-sheets composed of antiparallel β-strands (β-ribbons)

and a cystine knot, a motif which contains three disulfide bonds, with two disulfides linking parallel β-strands (Cys II – Cys V, Cys III – Cys VI) with the remaining disulfide bond passing through the ring formed to two other parallel β-strands (Cys I – Cys IV) (Fig. 3A)[33,41]. Members of the TGF-β family are secreted with a 250 amino acid prodomain, which acts to aid in folding and dimerization of the homodimer but also prevents functional activity for some but not all of the proteins in the family. All of the TGF-β isoforms are latent when associated with the prodomain, as are the activins and some of the BMP proteins[42]. Once the proteins are secreted through the endoplasmic reticulum, they are cleaved between the prodomain and growth factor homodimer in the Golgi by the membrane-bound serine proprotein convertases, including but not limited to furin[33]. Further processing is required for removal of the growth factor from the prodomain, for TGF-β this requires association with latent TGF-β binding protein (LTBP) or with glycoprotein A repetitions dominant (GARP) in the extracellular matrix and force transmitted by integrins[43], or through further protease cleavage[44] (Fig. 4).

The prodomain of TGF-β is composed of a bowtie knot, two arm regions, two straitjacket regions with latency lassos, and two association regions (Fig. 3B). The arm domains both contain two antiparallel β-sheets with four strands each that cross at the bowtie region via β-strands 8 and 9, linking the two arm regions, creating a closed arm conformation[43]. The arm domains also contain the RGD sequence which bind integrin αvβ6 for cleavage of the prodomain from the growth factor domain[43]. The bowtie knot links the two arm domains via two disulfide bonds for TGF-β1 and TGF-β3, and three disulfide bonds for TGF-β2. The straitjacket region is N-terminal to the arm region and surrounds the growth factor domain (Fig. 3B). The straitjacket is composed of a long alpha helix (α1) which contacts both monomers of the growth factor and stabilizes binding to the growth factor through contacts with hydrophobic residues in three turns of the alpha helix[33,43],

along with a latency lasso. The latency lasso connects the α1 helix of the straitjacket with the α2 helix of the arm domain and displaces the α3 helix of the growth factor domain, thus preventing receptor binding to the TGF-β growth factor (Fig. 3B). Finally, the prodomain includes ten residues at the opposite end of the bowtie knot which are structurally disordered without the presence of LTBP or GARP that associates with either protein via cysteine residues at this region[45]. The association region binds to the TGF-β-like binding region in LTBP and the LRRC region in GARP; these regions do not share structural similarity and so the association region is capable of existing in multiple conformations[33].

The mature TGF-β homodimer is formed by two 112-residue monomers connected by an interchain disulfide bond. The homodimer forms a 'hand over hand' structure, with each monomer resembling a left hand curled on itself (Fig. 3C). The 'heel' of one monomer presses against the 'palm' of the second monomer. The hands consist of four β-strand fingers, and the thumb is formed by an α-helix[33]. In some of the TGF-β members (TGF-βs, AMH, activins), the thumb is tethered to β-strand 1 by a disulfide bond which displaces the α-helix from the palm. While the growth factor structure that has been described has been termed a 'closed' conformation, an 'open' conformation for family members has been shown to exist, particularly with TGF-β3, Activin A, and GDF-8[46,47]. The three TGF-β isoforms are similar in structure, as are the BMP homodimers and the activin homodimers. Differences in the homodimer structures between the three subfamilies are most evident in the heel α-helix, and at the fingertip regions. These slightly different conformations reflect their ability to bind their respective type I and type II receptors.

1.1.4 Structure of TGF-β family receptors and binding paradigms

The TGF-β family members bind to type I and type II receptors, each of which consist of an ectodomain, a transmembrane domain, and a kinase domain. There are seven type I receptors and five type II receptors, allowing for promiscuity in receptor-ligand interactions. The three TGF-β isoforms are unique in that they are the only TGF-β family members to bind the type II receptor TβRII (Table 1). TGF-β family ligands can be distinguished by the receptors they bind and also by the Smads they signal through, with the TGF-βs and activins signaling through Smads 2 and 3 and the BMPs and GDFs signaling through Smads 1, 5, and 8[38].

The ectodomains of the type I (Alk1 – Alk7) and type II receptors (ActRII, ActRIIB, BMPRII, MisRII, TβRII), ca. 100 amino acids, contain a three-finger toxin fold characterized by a three-strand antiparallel β-sheet (β3- β5) surrounded by two antiparallel β-strands (β1- β2) and an extended loop connecting β4 and β5 (Fig. 5A-B). All of the TGF-β family receptors share four disulfide bonds in common, with ActRII, ActRIIB, and BMPRII containing an additional disulfide bond that displaces the β4-β5 loop away from the antiparallel β-sheet, and TβRII containing two distinct disulfide bonds from the other type II receptors. Alk1, Alk3, Alk5, and Alk6 also contain an additional disulfide bond that performs a similar function (Fig. 5A-B).

TβRI, or Alk5, the type I receptor that binds TGF-β, Activins, and GDF-8, is distinct from other type I receptors of the TGF-β family as it contains a prehelix extension[48], which is a five-amino acid PRDRP extension between β4 and the β4-β5 loop (Fig. 3D). This prehelix sequence is not present in other type I receptors, with the exception of Alk4[33], which has a PAGKP prehelix extension. This forces TβRI to bind closer to the fingertips of TGF-β and thus make contact with both TGF-β and TβRII[36,48,49] (Fig. 5A). Usually, the type I receptors bind at the wrist of the growth factor homodimer[49,50], contacting both monomers but not the type II receptor, whereas Alk5 and

Alk1 bind differently, with Alk1 still binding at the wrist but in a rotated conformation, due to a deletion of two residues in the prehelix extension (PHE) as opposed to a five-amino acid extension[33,51].

TβRII is unique as it only binds TGF-β. While most type II receptors bind to their growth factors through their exposed β5- β7 sheet and the β6- β7 loop, TβRII uses an edged β-strand, β4 on the β3-β4 sheet, along with β1 and β3 (Fig. 5B) to bind TGF-β[36,49]. This is possible due to the eight-residue extension of the β6-β7 loop and the absence of the disulfide connecting the β4-β5 loop to the β5 strand which is present in ActRII, ActRIIB, and BMPRII. This missing disulfide bond allows for the extended β6-β7 loop to block the surface of the β5-β7 sheet from binding the growth factor and exposes the β1-β2 sheet for binding. TβRII also contains a ten-residue extension on its β1-β2 loop; this loop is stabilized by two disulfide bonds, including one unique to TβRII and braces against the β3-β5 sheet, allowing for presentation of the β4 strand to TGF-β[49].

There are currently three paradigms for how the main TGF- β family members bind to their receptors, one for the TGF-β isoforms, one for the BMPs, and one for the activins. For the TGF-βs, the type I and type II receptor binding is cooperative, with the type II receptor binding first and the type I receptor binding the joint interface created[36,52]. This is not uniformly the case with the BMPs; some receptors bind a stepwise manner depending on the isoforms, but others do not[53,54]. For the TGF-β binding paradigm, TβRII inserts its edged β-strand (β4) in between the fingers of TGF-β, which is a site distinct from all the other growth factor type II receptors, which bind closer to the knuckles of the homodimer rather than at the fingertips (Fig. 5C). Ile76 of TβRII packs against a hydrophobic pocket of TGF-β, containing Trp332, Tyr390, and Val392. This is stabilized by hydrogen-bonded ion pairs of Asp55 on TβRII and Arg394 on TGF- β and Glu141 on TβRII and Arg325 of TGF-β[55]. These arginine residues are present in TGF-β1 and TGF-β3 and absent

on TGF-β2, which contributes for the low affinity binding of TGF-β2 to TβRII. These arginine residues have a unique position in TGF-β, with the loop lengths and sequences containing Arg325 and Arg394 being different for TGF-β in comparison to other growth factors. TβRI then binds to the interface formed by TGF-β and TβRII, allowing for additional contacts that allow for high affinity binding to the complex whereas binding of TβRI alone to TGF-β is considerably weaker by nearly 500-fold[36,49].

In contrast to this model, the BMP type I and type II receptors do not contact each other at all. The BMP type I receptors bind to the 'wrist' region of the homodimer, i.e., the dimer interface, while the type II receptors bind to the knuckle region (Fig. 5D). The BMP type II receptor uses its β5-β7 sheet to bind its homodimer, as due to the shorter β6-β7 loop it is available to act as a binding interface, mainly through hydrophobic interactions as opposed to the electrostatic interactions, hydrogen pairs, and hydrophobic interactions of the TGF-β:TβRII interface (Fig 5C). The type I receptors that bind to BMP receptors do not contact the type II receptor and more thoroughly contact both growth factors monomers. This binding is mostly hydrophobic with a larger binding interface in comparison to TβRI binding to the TGF- β growth factors[30,56]. The difference in type I receptor specificity is mediated by the β4-β5 loop, which in BMPRIa is disordered when unbound but ordered as a helix when bound (Fig. 5D). The β4-β5 loop in TβRI contains the hydrophobic prehelix extension that clashes with the wrist region of TGF-β due to steric overlap and thus TβRI if shifted towards the fingers of TGF-β, contacting both TGF-β and TβRII.

The activin paradigm has more in common with the BMP paradigm, with a lack of cooperativity[56]. In both cases the type I receptors utilize a phenylalanine residue to project into hydrophobic pockets of either GDF11 or BMP2/9. Hydrophobic interactions are formed between the type I receptors and conserved tryptophan residues of both ligands (Fig. 5D-E). However, the

BMP ligand utilizes a much higher surface area to bind the type I receptor in comparison to the activins, which is reflected in a relative difference in binding affinities. More interactions between the prehelix loop of BMP2 and Alk3 occur in comparison to the single residue interaction of GDF11 with TβRI, which is compensated by a greater degree of interaction between GDF11 and the prehelix extension of TβRI. This prehelix extension is not present in BMP type I receptors, accounting for this difference[56].

1.1.5 Intracellular TGF-β signaling

Once the ligand is bound to the extracellular domains of TGF-β receptors, signaling occurs through the intracellular kinase domains. The kinase domains of the type II receptors are constitutively active and consist of an amino-terminal lobe and a carboxy-terminal lobe, the former with a five strand β-sheet and an α-helix, and the latter with helices and a catalytic loop. The two lobes of the type II receptor kinase domains create an ATP-binding pocket, with the β1-β2 strands of the amino-terminal contacting the phosphate of the ATP and the regions connecting the two lobes forming a hydrogen bond with the adenosine[56,57]. The type I receptors, including TβRI, are structurally similar but also contain a glycine-serine (GS) rich domain N-terminal to the amino-lobe. TβRII phosphorylates the TβRI GS loop at the sequence: 185-TTSGSGSGLP-194; the phosphorylation of the GS loop then allows the L3 loop of Smad2 to bind the L45 loop of TβRI[57].

Smad proteins contain two globular domains, MH1 and MH2 (Mad homology 1 and 2), linked with a proline-rich set of residues. The MH1 domain acts to bind DNA while the MH2 domain is recognized by the β4-β5 loop of the type I receptor at the Ser-X-Ser motif at the L3 loop of the MH2 domain[57]; these two motifs determine specificities to either Smad 2/3 or Smad 1/5/8[58].

Upon phosphorylation of either Smad 2/3 or Smad 1/5/8, Smad4 is recruited to the complex and translocated to the nucleus where it either activates or represses target genes[59] (Fig. 2). Outside of phosphorylation, there are also other post-translational modifications that can impact the stability, activity, and subcellular localization of TβRI and TβRII[57]. Other proteins interacting with the kinase domains of TβRI and TβRII include inhibitory Smads such as Smad7 acting to negatively regulate TGF-β activity[60], and 14-3-3ε acting to promote TGF-β activity[61].

Outside of Smad signaling, there are also non-canonical non-Smad TGF-β ligand signaling pathways, including branches of the mitogen-activated protein kinases (MAPK) pathway: extracellular signal regulated kinases 1/2 (Erk 1/2), Jun amino terminal kinase (JNK), p38 kinase, IkB kinase (IKK), phosphatidylinositol-3 kinase, along with the Akt, and the Rho GTPases[40,62,63]. These alternate pathways can either be Smad-dependent or Smad-independent, highlighting that TGF-β signaling is part of a larger interconnected set of signaling pathways than the canonical Smad pathway model would suggest.

Due to the variety and crosstalk between these signaling pathways, the effect of the TGF-β ligands is highly pleiotropic. TGF-β ligands can have opposing effects based on the cellular context. Smads can associate with a variety of DNA-binding transcription factors which can greatly increase the breath of genes that are transcribed, and even then, the Smad proteins are themselves regulated by other phosphorylation or ubiquitylation modifications. Taking into account non-Smad signaling pathways, which may or may not interact with Smad signaling pathways, leads to even further variety in signaling output in which the cell type and thus the cell-specific transcription factors and cofactors present, along with the cellular environment, can lead to highly context-dependent responses to TGF-β ligands[64]. Finally, TGF-β has been implicated in

regulation of non-coding RNAs, microRNAs, long noncoding RNAs, etc., which act as posttranslational regulators of mRNA transcripts, leading to even further pleiotropy.

1.2 TGF-β in physiological contexts

1.2.1 TGF-β in cellular proliferation and differentiation

TGF-β was initially discovered as a stimulator of malignant cell growth, but has been shown to have antiproliferative activities in a variety of cell types[65,66] including epithelial cells and keratinocytes[67]. The mechanism for this growth inhibition is through transcriptional upregulation of cyclin-dependent kinase (CDK) inhibitor proteins[68-70]. Cdks control cell cycle progression, and TGF-β inhibits these Cdks through the p15 and p21 proteins which combined inhibit Cdks 1, 2, 4, and 6[68-70]. Another strategy TGF-β utilizes is to inhibit the expression and production of proliferative proteins including Cdc25A[64,71]. TGF-β also has the potential to indirectly impact cellular proliferation by antagonizing known proliferators, including EGF[72]. However, TGF-β also has the ability to act as a stimulator of proliferation for malignant cells, chondrocytes, mesenchymal stem cells, and fibroblasts and endothelial cells under certain conditions[64]. TGF-β is highly implicated in the differentiation of multiple cell types, including neuronal cells[73], several different immune cell types[12], embryonic and somatic stem cells[74], and epithelial cells[75]. Epithelial cells proliferate and differentiate as they build higher-order tissues; these cycles of proliferation and differentiation are controlled by neighboring layers of mesenchymal cells[75,76]. They can differentiate into multiple cell types, and when they differentiate into mesenchymal cells the process is called epithelial – mesenchymal transition (EMT).

TGF-β2 has been shown to induce EMT in endothelial cells that line heart valve septa[77], while TGF-β1 has been shown to induce EMT in mouse mammary epithelium[78], and TGF-β3 in palatal cells[79]. TGF-β-invoked Smad signaling leads to an increase in expression and activity in EMT transcription factors[80,81]. These Smads also act to directly express mesenchymal genes[82]. Non-Smad signaling pathways activated by TGF-β also lead to EMT, including the RHO-like GTPase, PI3K, ERK, JNK, and MAPK pathways[81]. Activation of these pathways aids in filopodia formation, dissolution of tight junctions, and contributes more to changes in gene expression that promote a mesenchymal over an epithelial phenotype[80,81]. Inhibition of these pathways has also been shown to prevent TGF-β induced EMT[83].

1.2.2 TGF-β in embryonic development

TGF-β is also heavily involved in embryonic development, TGF-β2 in cardiac development, and TGF-β3 in palatal development[84]. TGF-β2 deficient mice show cardiac defects[11], as TGF-β2 is detectable in the myocardium by day eight in mice[84]. TGF-β receptors are also found in precardiac tissue, and thus the growth factors are likely to serve as early signals for formation of endocardial and myocardial precursors. Additionally, TGF-β2 and TGF-β3 have been implicated in endocardial EMT, leading to the establishment of mesenchymal cells in the cardiac tissue[84]. Similarly TGF-β3 mutant mice are associated with cleft lip and cleft palate[85], as TGF-β3 is expressed in medial edge epithelial cells prior to and during palatal fusion[9]. While homozygous TGF-β3 null mice have palatal defects, heterozygous mice are unaffected, and the induced palatal defects can be mitigated by Smad2 overexpression[85]. TGF-β1 is less clearly tied to a specific organ system, but homozygous null TGF-β1 mice die soon after birth due to multifocal inflammatory organ disease[10].

1.2.3 TGF-β in wound healing and adaptive immunity

TGF-β1 was shown as early as 1983 to be able to increase the speed of wound healing in vivo[86]. Platelets store TGF-β1 and release them at the site of the wound[87]. TGF-β1 then acts to attract monocytes and fibroblasts leading to fibroblast proliferation, and the formation of granulation tissue[64]. TGF-β1 induces production of ECM proteins, including fibronectin and collagen, which promotes ECM deposition[87], while also preventing degradation of the ECM through inhibition of metalloproteinases[88]. Different isoforms of TGF-β lead to different effects on wound healing, with TGF-β1/2 promoting ECM deposition early on but not impacting final scarring, whereas addition of TGF-β3 reduces scarring[89,90], which has made TGF-β3 an attractive molecular scaffold for development of anti-scar wound healing therapeutics[91].

TGF-β is also hugely important as a regulator of the immune system, particularly through its impacts on immune cell differentiation and proliferation. TGF-β acts in a highly pleiotropic manner in the immune system, with effects on T-cell, B-cell, macrophage, and dendritic cell populations, though it is most classically known for its impact on T-cells. TGF-β acts to promote receptors that commit the CD8+ T-cells to their differentiation state[92]. TGF-β performs similar actions for other T-cell subtypes, such as the CD4+CD25+Foxp3+ thymus-derived regulatory T-cells (T_{regs}), invariant natural killer T-cells, and intraepithelial lymphocytes, by promoting the survival of their precursors[12].

T_{regs} act to downregulate the immune systems and can originate in the peripheral immune system and in the thymus (pT_{regs} and tT_{regs} respectively). In the thymus, TGF-β plays an important role in induction of tT_{regs} by promoting the survival of precursor cell populations[93]. Mice lacking TβRI were shown to have lesser amounts of Foxp3+ thymic lymphocytes[94]. In the peripheral immune system, TGF-β has been shown to act as an immunosuppressive cytokine, inhibiting T-

cell proliferation through minimizing interleukin-2 (IL-2) gene transcription and protein production[95,96]. However, this is cell-context specific, as TGF-β is not able to inhibit the proliferation of activated T-cells due to the decreased expression of TβRII at the cell surface[97]. Loss of TGF-β signaling also reduces some populations of activated T-cells, particularly CD4+ and CD8+ T-cells[92]. Broadly, TGF-β can either aid in the activation of T-cell subtypes (pT_{regs}, Th17, Th9, follicular T-cells) or inhibit the activation of T-cell subtypes (cytotoxic T lymphocytes, Th1, Th2 cells) through regulating the expression of T-cell lineage-defining transcription factors[12].

TGF-β inhibits Cyclin A and Cdk2, which both impact the growth cycle of B-cells[98], while also inducing apoptosis in murine B-cells[99], and in B-cell progenitor cells[100]. Mice deficient in TGF-β show an expanded B-cell population, and spleen cells absent of TβRII show higher levels of bromodeoxyuridine (BrdU) incorporation[101]. TGF-β also impacts B-cell differentiation through control of immunoglobulins, inhibiting IgG synthesis[101] and promoting production of IgA antibodies[102]. Mice deficient in TGF-β in B-cells specifically show significant losses of IgA, both in the serum and mucosa[101]. Additionally, TGF-β inhibits IgE, which is a mast cell activator, while also inhibiting proliferation, and mast cell degranulation[103]. However, TGF-β also acts to promote the expression of some mast cell inflammatory mediators, including IL-6[104]. Similarly, stimulation with TGF-β inhibits macrophage inflammatory responses when in combination with other cytokines, but stimulation with TGF-β alone promotes inflammatory cytokine production[105,106]. Depending on the cell context, physical location, and the presence of inflammatory cytokines, TGF-β can act to stimulate or inhibit the proliferation and differentiation a variety of immune cell subtypes, making TGF-β not only vital for maintaining immune homeostasis, but also difficult to study due to its pleiotropy.

1.3 TGF-β in pathological contexts

1.3.1 TGF-β in autoimmune disorders

Dysregulation of TGF-β is the cause of many pathologies: connective tissue and skeletal disorders, autoimmune conditions, several cancer types, fibrotic diseases, etc. For example, deficiencies in T-cell specific TGF-β signaling can lead to autoimmunity in mice[107,108], though this is dependent on both lymphopenia and loss of TGF-β signaling. Issues with TGF-β signaling have been implicated in inflammatory bowel disease, due to the role of TGF-β and IL-10 in intestinal homeostasis[109,110]. IL-10 acts to promote TGF-β secretion in intestinal T-cells[111], and also to differentiate naïve T-cells into T_{regs}[112]. Mutations in TGF-β and IL-10 have been correlated with inflammatory bowel disease (IBD)[113,114], and mice with deletions in both of these pathways end up with ulcerative colitis (UC)[115].

In rheumatoid arthritis (RA) TGF-β has been known to act in both a protective and a pathologic manner. Higher levels of TβRII have been found in RA synovial fibroblasts with respect to osteoarthritis (OA) patients[116], and intraperitoneal administration of TGF-β into collagen-induced arthritis mouse models were shown to reduce the severity of arthritic symptoms[117]. This is supported in other rodent models, with *Streptococcus*-induced arthritic rat models also showing symptomatic improvement upon administration of TGF-β[117]. On the other hand, TGF-β seems to have the opposite function in healthy animal models. Injection of TGF-β1 into healthy mice induces joint inflammation and lymphocyte recruitment, and treatment with TGF-β antibodies represses synovial inflammation and production of pro-inflammatory cytokines[117].

Multiple sclerosis (MS) is studied in lab mice through the experimental autoimmune encephalitis (EAE) murine model. TGF-β has both a protective and disease-provoking impact on

this mouse model, though the former has more evidence to support it. Administration of TGF-β1 reduces disease incidence, inflammation, and demyelination[118] and improves clinical course, including mitigation of central nervous system damage[119]. MS penetrance is highly variable amongst patients, and T-cells from patients with a more stable disease course expressed more TGF-β than patients with a more severe and active disease course[120]. Yet, TGF-β also has the ability to differentiate Th17 cells, which provoke the EAE phenotype, and mice with T-cell deletion of TGF-β1 are resistant to EAE[121].

1.3.2 TGF-β in organ fibrosis

TGF-β also has a key role in pathogenic fibrotic diseases, with dysregulated fibrotic responses often resulting from aberrant wound healing. TGF-β induces pulmonary fibrosis, with TGF-β acting to promote EMT and the production of smooth muscle actin and collagen in the idiopathic pulmonary fibrosis (IPF)[122], occurring due to both canonical Smad signaling and non-canonical pathways including the JNK, MAPK, and ERK pathways[122]. TGF-β promotes the expression of connective tissue growth factor (CTGF) and follistatin-like protein 1 (FSTl1), and inhibits long noncoding RNAs that may act to inhibit fibrogenesis[123]. Additionally, cancer patients treated with bleomycin as a chemotherapeutic have an up to 10% chance of acquiring TGF-β-driven bleomycin-induced pulmonary fibrosis, and lung epithelial TβRII knockout mice were more likely to have an increase in alveolar epithelial cell lifetime[124].

TGF-β has also been implicated in both cardiac fibrosis and nephrotic fibrosis. In patients with ischemic cardiomyopathy (ICM), TGF-β and phosphorylated Smad2 (pSmad2) were shown to be increased[125]. Additionally in nephritic glomeruli cells, TGF-β inhibits degradation of matrix protein components and reduces the activity of the serine protease plasminogen activator (PA)[126].

TGF-β also alters the cell surface protein landscape by adjusting the proportions of α1β1 and α5β1 integrins in nephritic glomerular cells[127] to promote cell-matrix adhesion. Additionally, in patients with more severe nephrotic disease such as diabetic nephropathy, expression of multiple isoforms of TGF-β in the glomerulus and the interstitium were shown to be increased[128].

1.3.3 TGF-β in tumors

In oncogenic settings, TGF-β acts pleiotropically, with both tumor promoting and suppressing activities dependent upon the type of cancer and the cancer stage. Cancer cells are known to evade the growth inhibitory and the immunosuppressive activities that TGF-β promotes, while taking advantage of its pro-inflammatory aspects. TGF-β acts as a tumor suppressor through its role in inhibiting cell growth via suppressing Cdk function and enhancing the expression of Cdk inhibitors[69,70]. Loss of function mutations of TGF-β signaling components: TβRI/TβRII/Smad2 lead to increased tumor progression and metastasis in several carcinomas[129,130]. TGF-β also plays a key role in the tumor microenvironment, specifically for signaling between stromal and epithelial cells. In a mouse model where selective deletion of TβRII in stromal cells was performed, nearby epithelial cells were more prone to malignant conversion, and other growth factors were upregulated[131]. TGF-β induces T_{regs} from naïve CD4+ T-cells. As such, TGF-β signaling can act to generate an immunosuppressive environment in which solid tumors can thrive. In many cancers including squamous cell carcinoma, colorectal cancer, and melanoma, TGF-β signaling can act to induce chemotherapeutic resistance[132].

Overall, TGF-β signaling enables tumor suppression via inhibition of oncogenic cell growth[133]; as cancers progress, TGF-β-mediated immunosuppression leads to a cytotoxic-deficient environment, with T_{regs} suppressing anti-tumor immunity through the secretion of

immunosuppressive cytokines[17,134] and the direct inhibition of effector T-cells[135]. Overexpression of TGF-β has been correlated with high tumor burden and poor clinical outcomes in melanoma, breast cancer, and renal cell carcinoma[136-139].

1.4 Current state of TGF-β-targeting therapeutics

There are currently several TGF-β targeting therapeutics in clinical trials; however, there have been mostly neutral or even mildly adverse outcomes. This has potentially been due to inappropriate animal models and selection of patients, or to the use of TGF-β inhibitory therapeutics as a singular treatment rather than pairing with other chemotherapeutic agents[140]. As an adjuvant to PD-1 and PD-L1 therapy, TGF-β inhibitors have been shown to surpass checkpoint monotherapy *in vivo*[141,142]. The major classes of therapeutics include TGF-β kinase inhibitors, TGF-β-inhibitory monoclonal antibodies, TGF-β ligand traps, integrin inhibitors, and antisense oligonucleotides[140,143].

TGF-β kinase inhibitors block the ability of the TGF-β receptor kinase domains to bind ATP. This class of inhibitors suppresses both canonical and non-canonical TGF-β signaling. The kinase inhibitor galunisertib acts on TβRI and was shown to be successful in phase I studies in patients with advanced solid tumors[144] and with malignant glioma[145], but in phase II studies studying clinical efficacy in combined therapy with temozolomide radiotherapy, no clinical benefit to galunisertib was found[145]. Another competitive inhibitor of the ATP binding domain of TβRI, LY3200882, was tolerated as monotherapy and with combination therapy with gemcitabine and nab-paclitaxel[146]. Out of 139 patients tested total, four glioma patients experienced partial response to treatment and three pancreatic cancer patients experienced a stable disease course, meaning that

even with multicenter studies and looking over a variety of cancer subtypes, the efficacy of TGF-β kinase inhibitors is still minimal and seems to be patient-specific[146].

Monoclonal antibodies block TGF-β receptor-ligand binding and also prevent activation of latent TGF-β. Scholar Rock is currently developing an inhibiting antibody specific to latent TGF-β1, SRK-181[147], with the thought that isoform specific TGF-β inhibition will lead to less side-effects, as pan-inhibition of TGF-β has demonstrated cardiotoxic side effects in the past[148]. SRK-181 has only been tested in mouse models, and while TGF-β antibodies have fared well in pre-clinical trials, during phase I trials, these therapeutics have faced various safety concerns[149,150].

Similarly, TGF-β ligand traps have been developed to sequester TGF-β and prevent it from binding its receptors by using soluble forms of the receptor ectodomains, i.e., soluble TβRI/TβRII. Most of these traps are still in the pre-clinical phases, with two exceptions: AVID200, a trap containing the extracellular domains of both receptors fused to an Fc domain[151], which is in the process of recruiting patients for trials, and Bintrafusp alfa, a trap which combines a monoclonal antibody against PD-L1 to the ectodomain of TβRII[152], which has been given accelerated approval by the FDA for PD-L1-expressing cervical cancers, but has yet to undergo phase II studies[153].

Integrin inhibitors have also been tested for inhibition of TGF-β signaling in both cancer models and pulmonary fibrosis. In the former case, none of the tested therapeutics have shown an increase in overall survival[140]. For the latter, Pliant has developed a dual anti-integrin $\alpha_v\beta_1$ integrin $\alpha_v\beta_6$ inhibitor for idiopathic pulmonary fibrosis. Though the trial is ongoing, the drug has currently met safety and tolerability requirements and is showing promising efficacy. While the field of TGF-β inhibitors spans a breadth of techniques and targets, there has not yet been a breakthrough towards targeting aberrant TGF-β signaling in cancer therapeutics, though some have shown promise in combination settings.

1.5 Tables and Figures

Table 1: Ligand-Receptor pairings for TGF-β superfamilyAdapted from Hinck et al[30].

Ligand	Type I Receptor	Type II Receptor	Smads
TGF-βs	TβRI	TβRII	2/3
	Alk1		
BMPs/GDFs	Alk1	ActRII	1/5/8
	Alk2	ActRIIB	
	Alk3	BMPRII	
	Alk6		
Activins	Alk4	ActRII	2/3
	TβRI	ActRIIB	
Inhibins		ActRII	
		ActRIIB	
Nodal	Alk4	ActRII	2/3
	Alk7	ActRIIB	
MIS	Alk3	MISRII	1/5/8
	Alk6		

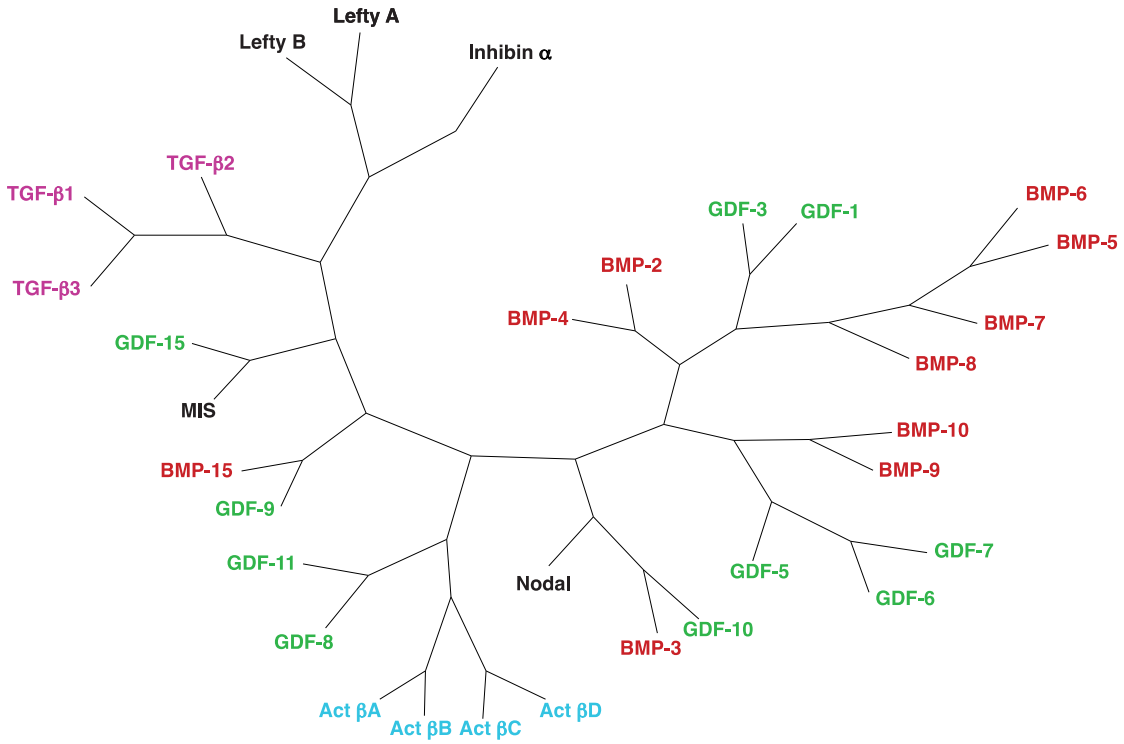

Figure 1: Phylogenetic tree of the TGF-β superfamily. Figure from Hinck et al[30]. Phylogenetic tree, colored to highlight groups of ligands within the TGF-β superfamily: red (BMPs), green (GDFs), purple (TGF-βs), blue (activins).

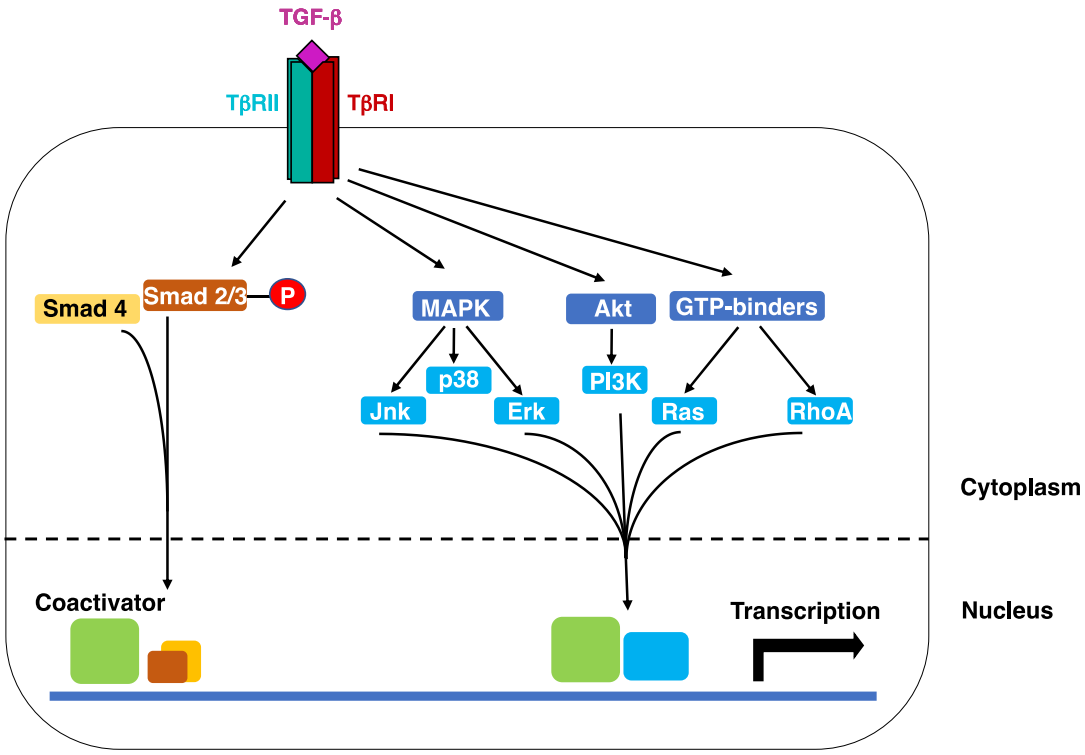

Figure 2: TGF-β signaling model. TGF-β forms a heterotetrameric complex by binding to the extracellular domains of two type I receptors (TβRI) and two type II receptors (TβRII). The kinase domains of these receptors are then phosphorylated. TGF-β either signals through the recruitment and phosphorylation of Smad2/3 which then complex with Smad4 and traverse to the nucleus for transcriptional regulation, or TGF-β signals through non-Smad pathways, including MAPK, Akt, GTP-binding proteins, etc. Adapted from "TGF-β signaling in context"[154], under CC BY 4.0.

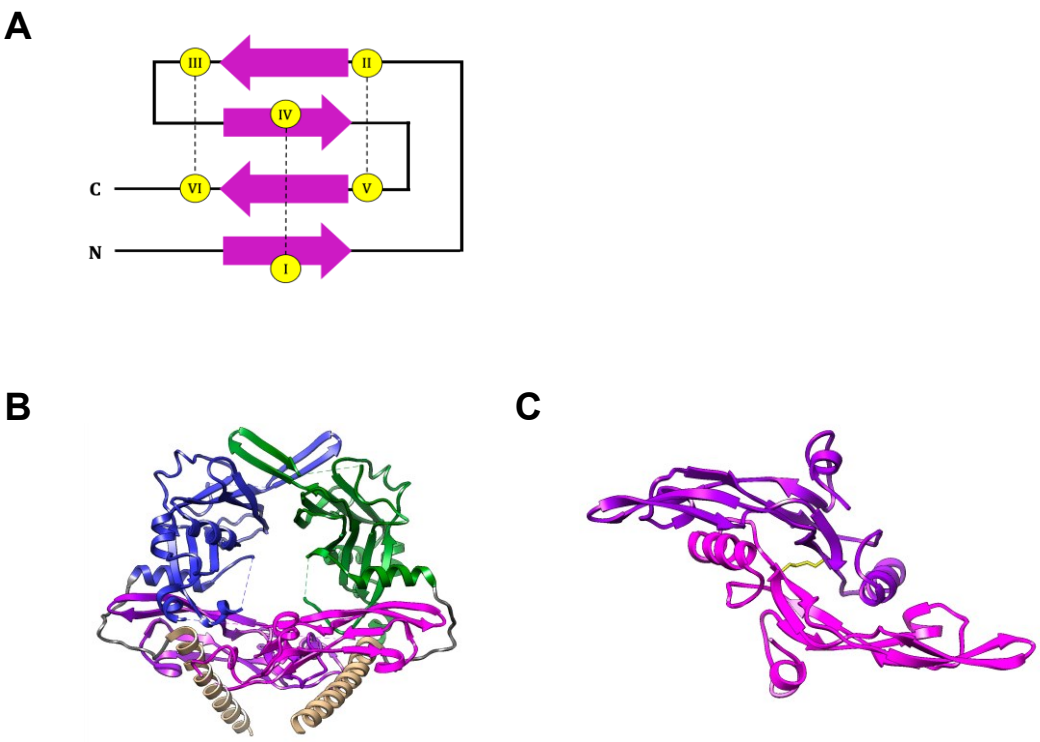

Figure 3: Structure of the TGF-β family prodomains and growth factors

A-E. A. Schema from Iyer et. al.[41] of the growth factor cystine knot structure, with beta strands (magenta) and cysteines (yellow) colored. B. Structure of the prodomain of TGF-β1, with arm domains (green and blue), latency lasso (gray),the association regions (tan), and the growth factor monomers (purple and pink), colored appropriately. PDB: 3RJR. C. Structure of the TGF-β growth factor homodimer in the closed confirmation, with interchain disulfide shown in yellow, PDB: 2PJY.

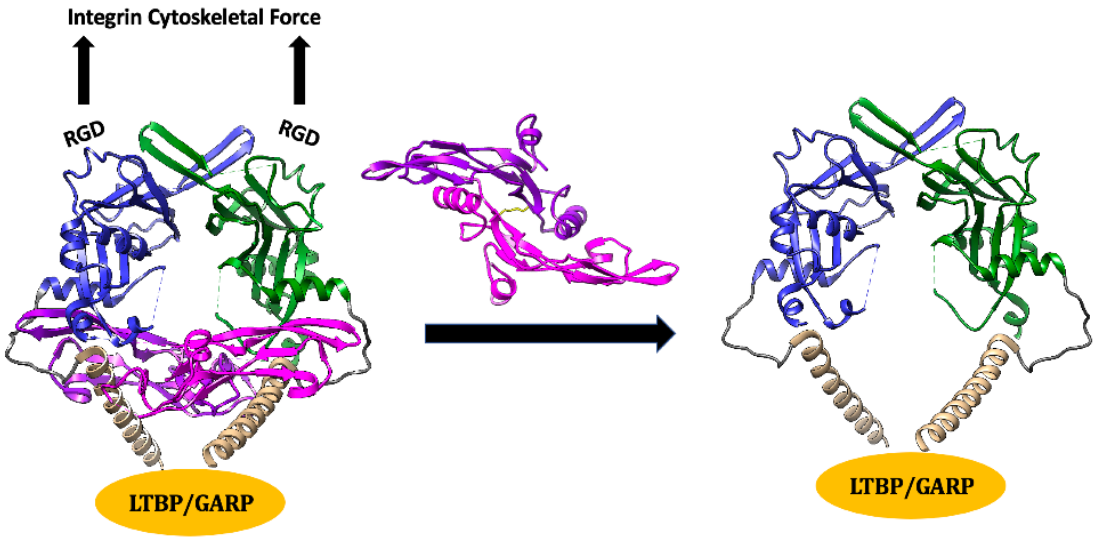

Figure 4: TGF-β is secreted in a latent form. Cartoon depiction of TGF-β bound to its prodomain, with key features indicated, arm domains (green and blue), latency lasso (gray),the association regions (tan), and the growth factor monomers (purple and pink), colored appropriately. Removal of the prodomain either by latent TGF-β binding protein (LTBP) or glycoprotein A repetition dominant (GARP), shown, PDB: 3RJR.

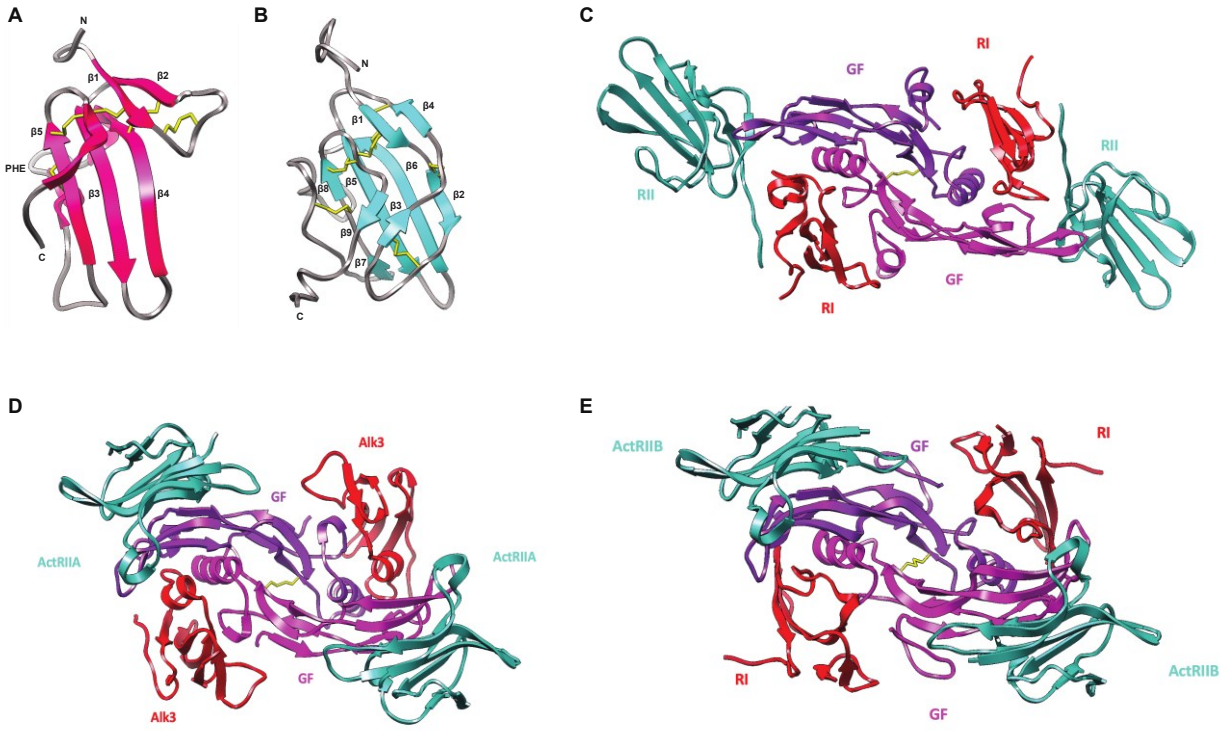

Figure 5: Structure of the TGF-β family receptors and receptor binding paradigms.

A-E. A. Structure of the extracellular domain type I TGF-β receptor TβRI, PDB: 2L5S. B. Structure of the type II TGF-β receptor TβRII, PDB: 1PLO. C. Structure of the ternary TGF-β:TβRI:TβRII complex, PDB: 2PJY. D. Structure of the ternary BMP2:Alk3:ActRIIA complex, PDB: 2GOO, 6OMN. E. Structure of the ternary GDF-11:TβRI:ActRIIB complex, PDB: 6MAC. All interchain disulfides shown in yellow.

2.0 Introduction: Helminth parasites and parasite-host immunoregulation

2.1 Overview of helminth-host immunomodulation

2.1.1 The hygiene hypothesis

In the 1960s Greenwood from the London School of Hygiene anecdotally noted the lesser prevalence of autoimmune diseases in areas with high rates of parasitic infection[155], and in a review assessing 10 years of admission to a Nigerian hospital, showed that out of near 100,000 admissions for connective tissue diseases, only 100 were autoimmune in nature[155]. This was accompanied by a perceived increased prevalence of autoimmune disease and allergy in 'developed' countries, with particularly high incidences of self-reported asthma and allergy in the UK, New Zealand, and Australia[156]. It was noted by Greenwood that mice from an autoimmune mouse model infected with malaria showed a delayed onset and reduction in symptom intensity[157]. It was further theorized that the difference in incidence of hay fever, allergy, and eczema amongst a cohort of British children born in 1958 and followed over 23 years was related to a difference in levels of microbial infection amongst the participants[158].

This data gave rise to the hygiene hypothesis and the old friends hypothesis[159]. The former hypothesizes an inverse correlation between childhood microbial infection and development of autoimmune disorders later in life. In childhood or early development, the immune system is deeply influenced by its environment, and so an individual that is more accustomed to infection is more prone towards a hyporesponsive immune state, due to immunomodulatory effector molecules secreted by parasites and other microbes[160]. The old friends' hypothesis theorizes that over

millennia parasites and other infectious organisms have co-evolved with their hosts, acting as a natural dampener on the mammalian immune system. Thus, the removal of these infectious agents has caused an unequal balancing of the reactive and tolerogenic arms of the immune system, thus leading to a higher prevalence of reactive immune conditions.

2.1.2 Helminth parasites and innate immunity

Helminth parasites are known for their ability to modulate their host's innate and adaptive immune systems. Helminth parasites interfere with responses to damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs). Damage to tissue structures from parasite entry and parasitic secretions lead to the induction of DAMPs and PAMPs, meaning that for successful infection, parasites must mitigate this initial innate response[161].

Glycan-rich helminth parasite excretory-secretory products (ES) have been linked to tolerogenic immunomodulation. *Schistosoma mansoni* (*S. mansoni*) soluble egg antigens (SEA) have been shown to suppress activation of dendritic cells via the Toll-like receptor (TLR), one of many pattern recognition receptors (PRRs), along with suppressing the production of inflammatory cytokines such as IL-6 and IL-12[162]. Similarly, infection with *S. mansoni* and *Brugia malayi (B. malayi)* have both been shown to alter the expression of TLR isoforms, while also guiding the intracellular signaling of the TLRs towards a more hyporesponsive immune phenotype[163].

Though the focus has been on TLRs, different ES products have the ability to impact other PRRs such as the C-type Lectins Receptors (CLRs), which are more directly involved in recognition of the glycan components of helminth-derived products[161]. The murine-infecting worm *Nippostrongylus brasiliensis (N. brasiliensis)* secretes a product that interacts with the host

surfactant protein D (SP-D), which is related to the C-type lectins[164]; without SP-D the host is unable to eject the parasite[164]. In *F. hepatica*, interaction of glycan-rich parasite ligands with CLRs leads to a more tolerogenic phenotype, with a higher level of IL-10 and TGF-β secretion[165].

Along with their impacts on TLRs and dendritic cells (DCs), ES from different helminth species have been shown to interfere with macrophage activation[165]. The *F. hepatica* proteins Fh15 and Fh12 suppress LPS-induced TLR4 activation[166], inflammatory cytokines IL-12 and TNF-α[166,167], and induce the alternatively activated macrophage phenotype[167], which promotes wound repair and healing over inflammation[168]. In the same parasite, the protein FhHDM-1 acts to prevent macrophage activation, through mimicry of a host antimicrobial peptide that suppresses a pro-inflammatory response in the case of bacterial infection[169]. Thus, parasitic products not only interact with existing host machinery to enact a state of immune hyporesponsiveness, but parasites will often do so through mimicry of existing host proteins and peptides.

2.1.3 Helminth parasites and adaptive immunity

Broadly, helminth parasites shift the immune phenotype to a tolerogenic one, inducing T_{regs} and regulatory B-cells (B_{regs}). T_{regs} can be induced by TGF-β, and as such several parasites take advantage of the TGF-β signaling pathway to induce T_{regs} and generate a downregulated immune phenotype in their hosts. *Litosomoides sigmodontis (L. sigmodontis)*, a rat nematode, is able to decrease the function of anti-parasitic host immunity via upregulation of CD4+ CD25+ Foxp3+ T_{regs}[170,171] and of hyporesponsive CTLA4+ CD4+ Foxp3- Th2 cells[171]. These induced T_{regs} and Th2 cells in conjunction are so vital for the function of this parasite that only upon antibody-mediated depletion of these cells were the hosts able to gain infection resistance[170,171]. B_{regs} are B-

cells that act to create a tolerogenic environment in the organism, producing anti-inflammatory cytokines including IL-10, TGF-β, and IL-35[172]. In *S. mansoni* infection, an IL-10 dependent immunosuppressive phenotype is maintained[173], and B_{regs} act not only to produce this IL-10 but also to recruit T_{regs} to the site of inflammation in an allergic airway mouse model[174]. Some helminth parasites can also act on T_{regs} indirectly. *Toxocaris leonina (T. leonina)* produces a galectin-9 homologue, with galectin-9 having been shown to interact with the TGF-β signaling pathway to support the conversion of naïve CD4+ T-cells into CD4+ CD25+ Foxp3+ T_{regs}[175,176]. Preliminary experiments show this homologue can mitigate the symptoms of DSS-induced colitis, and that mice treated with this homologue show increased production of TGF-β, indicating that these parasitic products have the potential to be used towards developing immunosuppressive therapeutics[176].

Helminth ES products can also act on other effector immune cells. The cystatin protein Bm-CPI-2 in *B. malayi* acts to reduce T-cell priming by interfering with MHC Class II antigen presentation[177]. The parasite responsible for river blindness, *Onchocerca volvulus (O. volvulus)*, also performs a similar function with its cysteine protease inhibitor onchocystatin[178]. Parasites can direct the host towards a Th2 immune phenotype over a Th1 phenotype, preventing Th1 and Th17-mediated inflammation[165], shown with the *S. mansoni* glutathione transferase protein, which induces an eosinophil-rich immune response and mitigates inflammation in mouse models of colitis[179]. Other parasite products from *Ancylostoma caninum (A. caninum), Ancylostoma ceylanicum (A. ceylanicum), and B. malayi* have been shown to be homologous to the ShK protein, a sea anemone protein that acts to prevent voltage-gated potassium channel functioning[180] and whose promotion halts effector T-cell proliferation[181]. Helminth parasites can also act on effector

B-cells, with secreted proteins from *Ancylostoma* and *Acanthocheilonema viteae* acting on the B-cell receptor complex to prevent B-cell response to presented antigens[165].

2.1.4 Helminth parasites and allergy

In a mouse asthma model, infection with *L. sigmodontis* was shown to inhibit the asthma phenotype; production of immunoglobulins, sensitivity to methacholine, and eosinophilia were all reduced[182]. TGF-β expression on splenic CD4+ T-cells increased upon infection, and blocking of either TGF-β or T_{regs} by monoclonal antibodies partially restored sensitivity to methacholine, but did not change eosinophilia or immunoglobulin production[170,182]. This indicated that TGF-β signaling is one of the pathways that *L. sigmodontis* uses to minimize allergy, but not the only one.

Infection with *S. mansoni* and injection of *S. mansoni* eggs has been shown to reduce symptoms of allergy in a murine model. Mice treated with *S. mansoni* showed lower levels of reactive cytokines such as IL-4, IL-5, and IgE, reduced eosinophilia, and lower amounts of airway inflammation[183]. Mice treated with *S. mansoni* eggs had higher levels of CD4+ CD25+ T_{regs}, and depletion of these T_{regs} with a CD25 antibody was able to restore eosinophilia[183]. However, these results were dependent on the density of the infecting parasites and the length of infection. The usual markers of asthma phenotype: bronchoalveolar lavage (BAL) eosinophilia, bronchial inflammation, airway hyperresponsiveness were shown to be increased during acute stages of infection (8 weeks), and only start decreasing over time (12 weeks plus), indicating an initial immune response mounted against the parasite that only subsides during later stages of infection[173]. Additionally, BAL eosinophilia was shown to be dependent on parasite density, while bronchial inflammation as measured by Th2 cytokine production was shown to be dependent on length of infection[173].

However, in contrast to all of this, a recently performed meta-analysis of over 80 studies and nearly 100,000 patients showed no real association between helminth infection and allergic disease[184], indicating that more research needs to be done to characterize the molecules of these parasitic secretions that are leading to potential allergy minimization.

2.1.5 Impact of helminth parasites on autoimmunity

Autoimmune disorders in the context of childhood helminth infection are more difficult to study as they are relatively infrequent in younger patients; however, multiple studies have noted the lower prevalence of autoimmune disorders in countries with higher parasitic infection rates[155,159,165]. Some studies have utilized a positive titer of circulating anti-nuclear antibodies (ANA) as a measurement of autoimmune disease, though this measurement can also be positive in other pathological cases, including chronic infection, cancers, etc. and also in a physiologically normal individual[185]. In a study with over 600 Zimbabwean patients over a variety of age ranges (2-68) who had been exposed to *Schistosoma haematobium*, ANA levels were measured before and after treatment with praziquantel, an anti-parasitic drug[186]. Patients who were exposed to Schistosomes showed lower ANA levels in comparison to control patients, and this effect was partially reversed by administration of praziquantel, showing that there is at least some link between ANA level and parasitic infection[186].

More studies have shown that infection with helminths have mitigated autoimmune symptoms. In a mouse model of collagen-induced arthritis (CIA), arthritic symptoms were reduced upon parasite administration[187]. Levels of circulating type II collagen IgG and IgG2a antibodies had decreased, along with the expression of pro-inflammatory cytokines such as TNFα, IFNγ, and IL-17A[187]. However, this beneficial impact was dependent on the stage of infection with mice

exposed to the early-stage infection (pre-egg-laying stage) having no change in arthritic symptoms[188]. Mice exposed to the egg-laying stages of infection (7 weeks minimum) showed reduced levels of circulating antibodies, lower arthritic scores, higher productions of Th2 cytokines (IL-4, IL-5) and lower production of Th1 cytokines, though the level of Foxp3+ T_{regs} did not differ between control and infected mice, no matter the stage[188].

Infection with helminth parasites has shown to be beneficial in mouse models of non-obese diabetes (NOD), experimental autoimmune encephalitis (EAE), and other autoimmune diseases. In cases of *L. sigmodontis* infection, mice do not develop signs of diabetes, whereas control mice developed diabetes 80% of the time. This protection against disease is correlated with an increase in IL-4 and IL-5 production, and an increase in CD4+ CD25+ Foxp3+ T_{regs}, dependent upon TGF-β[189,190]. In general, studies have shown administration of helminth parasites to minimize or prevent autoimmune symptoms in mouse models.

2.1.6 The current state of helminth therapeutics

Due to their *in vitro* effects on allergy and autoimmune disorders, therapeutics utilizing live parasites and parasite products are in the clinical trial pipeline. The first example occurred in 2005 with infection of *Trichuris suis (T. suis)* for treatment of inflammatory bowel disease (IBD)[191]. Initial trials of patients treated with *T. suis* ova every 3 weeks for 28 weeks was promising as all seven patients all showed improvement in disease without adverse side-effects, and six of them went into disease remission[192]. A further randomized control trial of 30 patients showed improvement in 13/30 *T. suis* infected patients as opposed to 4/24 control patients, with no adverse side effects noted[193]. However, despite promising initial results, the large multicenter double-blinded studies for this treatment (Trial identifiers NCT01576471, NCT01279577,

NCT01433471) have concluded no real difference in efficacy between placebo and *T. suis* infection in IBD[194], or have otherwise not posted results. Amongst other parasites, scientists have also attempted to utilize infection with *Necator americanus (N. americanus)*, a human-infecting hookworm for treatment of asthma and inflammatory intestinal disease[191]. While there were no safety issues in either case, double-blinded randomized control trials with larger population sizes showed no significant effect of parasite treatment on primary outcomes of airway responsiveness[195,196] and response to wheat challenge respectively[197].

The failure of these live parasite clinical trials, along with others not mentioned, highlighted two key problems in the use of live helminth parasites towards treatment of asthma and autoimmune disease. First is that treatment of patients with live parasites may need to begin at an earlier age. Patients in populations with high levels of parasitic infection are likely first infected during childhood, prior to the development of autoimmune disease and most certainly prior to enrollment in a clinical trial[170]. It may be that infection with parasites is necessary from a very young age to allow for development of the immune system in a way that mitigates allergy and autoimmunity. Clearly though, it would be ethically dubious to infect very young children with parasites, particularly with individuals that have not developed allergy or autoimmune conditions and following them longitudinally to assess for development of autoimmune conditions would be logistically challenging. The other issue is that due to safety concerns, using an efficacious dose of live parasites may be challenging. These doses are likely much lower than that of natural infection and as such may not be capable of impacting the course of autoimmune disease[170].

As such, the field of helminth therapies has shifted from using live helminth infection to using parasite-derived products. Identifying a specific product allows for more thorough

characterization and mechanistic understanding of the potential immunomodulatory effect, while also potentially avoiding toxicity and safety issues at higher doses. Proteins and parasite-derived small molecules are being adapted for clinical trials, most for anti-inflammatory purposes. However, even with these products, more research needs to be done; many species have different secretory profiles depending on the parasitic stage, and depending on the mode of infection, some parasites may have more systemic rather than localized effects[165]. The main issue with utilizing these parasite-derived products is immunogenicity, though in theory this immunogenicity could be bypassed by use of low, unrepeated dosing. Helminth-derived products provide an exciting avenue for therapeutic discovery, with more and more potential therapeutic products being discovered year by year, particularly as the common parasitic secretomes become more thoroughly annotated.

2.2 Overview of Heligmosomoides polygyrus

2.2.1 Heligmosomoides polygyrus life cycle

Heligmosomoides polygyrus (H. polygyrus) is an intestinal nematode that infects mice and other rodents and is present throughout Europe and North America[198]. These parasites are often utilized in the laboratory setting for studying chronic impacts of parasitic infection on host immune activity. *H. polygyrus* worms form a series of coils and are colored a dark red or brown (Fig. 6). *H. polygyrus* does not have any intermediate hosts, with a direct cycle spanning over approximately two weeks[198]. Eggs are passed from the murine digestive system into feces, with the L1 stage hatching by the first day, and molting to the L2 stage within 2-3 days. Before molting,

the parasite synthesizes the cuticle, and L2 larvae which have retained their cuticle will then molt into the L3 larval stage, which acts to infect the definitive host[198,199]. Mice will then consume the L3 larvae, or will be given L3 larvae via lavage in a laboratory setting[200]. These L3 parasites will invade the intestinal mucosa one day after infection and will molt into the L4 larvae worms four days after infection after this mucosal penetration. The L4 larvae will migrate to the intestinal muscle six days post infection and will molt once more, finally maturing into the adult worm after eight days of infection[200]. By day 10, the adult worms will return to the intestinal lumen to mate, and *H. polygyrus* excretory-secretory products (HES) can be isolated by two weeks post infection[198-200].

The rate of worm expulsion and the burden of disease is highly dependent upon the mouse model, with the F1 cross between the C57/BL6 and CBA strains often used for laboratory study due to their inability to rapidly expel worms and their lack of morbidity[200,201]. To study *H. polygyrus* infection, the worms are administered to 8-week old mice in the L3 larval stage via oral gavage[200]. To collect the adult worms, mice are killed two-weeks post infection and the peritoneal cavity is removed. The small intestine is then excised and the gut lining is scraped to isolate the worms. The resulting mixture is gently agitated to separate the worms from any other detritus; the worms are then soaked in Roswell Park Memorial Institute (RPMI) Medium, along with other supplements, and HES is harvested from the worms at a frequency of no greater than twice per week for a total of three weeks, excluding any HES collected 24 hours post-culture[200].

2.2.2 Host immunity to Heligmosomoides polygyrus

Mice deficient in B-cell and T-cell immunity show an impaired ability to eject *H. polygyrus* and the worms are able to produce more offspring[201]. Infection has been associated with early

increases in IL-3, IL-5, and IL-9 mRNA expression in murine Peyer's patches that was mitigated by monoclonal antibodies against CD4 and CD8[202], along with increased secretion of Th2-associated cytokines IL-4, IL-5, IL-10, and IL-13, and the decreased secretion of Th1-associated cytokines IL-12 and interferon gamma (IFN-γ) in lamina propria mononuclear cells in the small intestine of infected mice[203]. IL-4 seems to be the most important in maintaining parasitic infection, with IL-4 monoclonal antibody treatment leading to a decrease in egg production and an accelerated timeline of worm expulsion in BALB/c mice, and in mouse models of severe combined immunodeficiency (SCID)[204]. This effect is likely associated with follicular T-helper cells, which act to produce IL-4 in reactive lymph nodes, thereby activating T-cell responses to the parasitic infection[205], but which also act to induce B-cell isotype switching to IgG and IgE antibodies[206] and produce IL-21[207], which protects against secondary *H. polygyrus* infection[201,206].

While host mice are largely susceptible to primary *H. polygyrus* infection and find it difficult to expel the worms, secondary challenges are easily overcome[201]. Host immunity to *H. polygyrus* is modulated by the B-cell compartment upon secondary infection; mouse models without B-cells showed a reduction in worm expulsion, though this is ameliorated by administration of immune serum[208]. B-cells drive anti-parasitic immunity through antibody secretion along with differentiation and expansion of Th2 immune cell subsets[208,209]. Polyclonal IgG antibodies against *H. polygyrus* were shown to inhibit parasitic egg production, while multiple infections were required for the affinity maturation of IgG and IgA antibodies that act specifically against the adult worms[210]. Serum from animals that were chronically infected with a single *H. polygyrus* infection did nothing to immunize naïve mice to the parasite, while serum from animals that had been challenged multiple times did afford a protective role to naïve mice[201,211]. Additionally, B-cell immunity to *H. polygyrus* infection is dependent upon the mouse strain

infected, with more resistant strains of mice shown to be correlated with the intensity of B-cell driven immune responses[201]. In a study testing eight different mouse strains, mice with acute or subacute infections, as opposed to chronic parasitic infection, showed higher levels of total serum IgG1 and IgE[212].

The innate immune system also plays a role in immunity to *H. polygyrus*, including alternatively activated macrophages, mast cells, and small intestine epithelial goblet cells, along with others[201]. Alternatively activated, or M2 macrophages, have been shown to oppose pro-inflammatory Th1 immune responses in models of parasitic infection[213]. When mice are depleted of these macrophages, they are unable to even expel *H. polygyrus* parasites during secondary infections[213,214]. Mastocytosis, or overexpression of mast cells, has been shown to be correlated with parasite expulsion[201,215], and IL-9 and IL-10 depletion of mast cells correlates with chronic infection over worm expulsion[215]. Mast cell depleted mice are more prone to higher worm burden than wild type mice, and mast cell depleted mice are unable to challenge secondary infection as readily as wild-type mice[216]. Intestinal immunity has also shown to play a key role in combatting *H. polygyrus* infection, through changes in gut effector cells, such as an increase in epithelial cell permeability, are seen with secondary and not primary infection[217]. Goblet cells act to generate a physical barrier for parasitic and other infection through production of mucus and other defensive proteins[201,218], and during parasitic infection, goblet cell hyperplasia is observed[218,219]. Mice negative for a molecule secreted by goblet cells, RELM-β that acts to regulate epithelial cell permeability have been shown to have a decreased ability to expel *H. polygyrus* adult worms spontaneously and a decreased ability to expel *H. polygyrus* through IL-4 mediated mechanisms[220].

2.2.3 Host immunomodulation by Heligmosomoides polygyrus infection

H. polygyrus infection impacts a variety of cell subsets, including regulatory B-cells, regulatory dendritic cells, and regulatory T-cells. Studies show that CD4- CD19+ B_{regs} from *H.polygyrus*-infected mice are able to impact symptoms in both mouse models of asthma and experimental autoimmune encephalitis (EAE)[221]. Administration of these B_{regs} in an asthma model results in a reduction of airway eosinophilia and IL-5 secretion and overall allergic pathology, while administration in an EAE model shows anti-inflammatory properties. Dendritic cells have been isolated from *H. polygyrus* infected mice, and there is a clear difference in the subsets existing in infected and naïve mice[160,201,222]. Upon infection, there was an expansion in the population of $CD11c^{low}$ CD103- dendritic cells, which are dendritic cells that act to induce Foxp3+ T_{regs}[223].

T_{regs} play a unique role in parasitic infection, as both their spatial and temporal location over the course of infection can alter whether they act to maintain or mitigate parasitic infection. In the initial stage of infection, there is an early increase (days 3-7) in natural T_{regs}, particularly in the Peyer's patches and draining lymph nodes[224]. Helminth infection also induces T_{reg} production through ES products differentiating T_{regs} from naïve CD4+ T-cells and inducing tolerogenic dendritic cells which coax a T_{reg}-promoting environment. Additionally, physical damage to the epithelium caused by parasite entry into the gut promotes M2 macrophage polarization and the release of alarmins and other DAMPs and other cytokines that induce T_{regs}[224]. *H. polygyrus*-induced CD4+ CD25+ T_{regs} are more suppressive against effector T-cells than from naïve mice, and even CD25- CD4+ T-cells are more resistant to suppression from CD4+ CD25+ T_{regs} from naïve mice[225]. These CD4+ CD25+ T_{regs} show higher expression of TGF-β and CD103, which is an integrin involved in cell to cell and cell to matrix interactions[225]. While this subset of T_{regs} is necessary for *H. polygyrus* pathology, depletion of them early in the infection leads to a worsened

outcome for the mice, as these cells are required for the maintenance of gut homeostasis[201,226]. *H. polygyrus* utilizes regulatory T-cell populations to maintain a permissive level of worm burden in the host, a remarkable evolutionary adaptation that takes advantage of existing immune machinery that cannot be easily removed by the host.

2.3 Heligmosomoides polygyrus mimic proteins

H. polygyrus excretory-secretory products (HES), which are products actively secreted by the parasite and diffused from the body of the parasite, are released by the parasite in the lumen of the small intestine[227]. The secretions are immunologically active, and so were analyzed to discover the *H. polygyrus* compound that acts to upregulate the differentiation of T_{regs} from naïve CD4+ T-cells. Adult worms were cultured and the ES harvested and concentrated. Initial characterization of the HES in 2010 showed a four-fold increase in the percentage of CD4+ Foxp3+ T_{regs} from a naïve CD4+ T-cell population over a duration of two days[228]. This ability was mitigated by heat, indicating the ES component that upregulates T_{regs} is heat-sensitive[228]. This HES was shown to be acting in a manner similar to TGF-β, activating a TGF-β specific alkaline phosphatase reporter in MFB-F11 mouse fibroblast cells[228,229].

To isolate the molecule from HES that upregulates T_{regs}, 1mg of HES was fractionated using size-exclusion chromatography (SEC). Fractions were assayed for TGF-β signaling activity via a Smad reporter assay in MFB-F11 fibroblast cells[229], and activity normalized to a TGF-β1 control[230]. HES was also fractionated via ion exchange chromatography and assessed for TGF-β signaling activity. Fractions from both chromatographic methods were then analyzed via mass spectrometry and compared to an existing data base of *H. polygyrus* sequences. From this analysis,

four candidate proteins with high TGF-β signaling activity in both the SEC and ion exchange steps were identified[230]. The genes coding for these four candidates were cloned into mammalian vectors transfected into expi293 cells, and then expressed and purified as recombinant proteins with a C-terminal histidine tag. Out of the candidate proteins, only one of them showed TGF-β signaling activity[230]. The protein was then purified from the cell supernatant using a nickel column and was designated *H. polygyrus* TGF-β mimic protein, or TGM. TGM is a 422-residue protein, the N-terminal 18 residues corresponding to the signal peptide, and the other 404 residues corresponding to five homologous domains with four cysteines each, excepting the fourth domain which includes six cysteines (Fig. 7). The five domains are divided into 11 exons, with Exon 1 corresponding to the signal peptide, and the remaining exons pairing as two exons per domain (Fig. 7).

The TGM domains (ca. 90 amino acids each) belong to the complement control protein (CCP) family, which are β-strand rich domains with two disulfide bonds each. This is in contrast to TGF-β isoforms, which are disulfide-linked homodimeric proteins with 112 amino acids per monomer. Despite a lack of structural similarity to TGF-β, TGM is able to activate the TGF-β signaling pathway. To assess whether TGM was upregulating the expression of TGF-β or whether TGM was directly interacting with the TGF-β receptors TβRI and TβRII, TGM-induced TGF-β signaling was assessed in three separate contexts: 1) with an antibody against TGF-β, 2) with a TβRI inhibitor, SB 431542, and 3) with a TβRII inhibitor, ITD-1. Upon administration of all three inhibitory agents, signaling due to TGF-β was reduced, if not eliminated (Fig. 8A-C). However, when TGM was added, addition of the TβRI or TβRII inhibitors dramatically reduced signaling, but addition of the anti TGF-β antibody did not. Thus TGM is interacting with the TGF-β receptors to signal and does not directly impact the level of TGF-β itself[230]. TGM is also able to achieve a higher maximum signal in the MFB-F11 assay than the maximal concentration of TGF-β. HES is

able to achieve this higher maximum as well, though the concentrations required to do so are several-fold higher than that of recombinant TGM (Fig. 8D).

TGM can induce Foxp3+ T_{regs} *in vitro* in both mouse and human cell lines[230]. Naïve T-cells were isolated from mice and stimulated with IL-2, with the addition of either TGM or human TGF-β1. TGM and TGF-β were able to convert naïve murine CD4+ CD25- T-cells into CD4+ CD25+ Foxp3+ T_{regs} similarly, with 0.78nM of TGM converting 90.65% of the naïve cells into T_{regs} and TGF-β converting 79.65% at similar molar concentrations[230] (Fig. 9A). The mean fluorescence intensity (MFI) of induced Foxp3 was measured as well and found to be slightly higher for TGM than for TGF-β (Fig. 9B). This pattern was similar when TGM and TGF-β were tested against human CD4+ CD25- T-cells; while TGM and TGF-β were able to induce similar amounts of T_{regs}, the MFI of Foxp3 for the TGM-induced cells was much higher than the TGF-β-induced cells, indicating higher levels of Foxp3 expression[230] (Fig. 9C-D).

TGF-β has the ability to differentiate different effector T-cell lines, particularly those that promote a Th1 or pro-inflammatory response, such as Th17 and Th9 cell types[12]. Parasites have been shown to counter such pro-inflammatory mechanisms[165], and thus TGM was also tested for its ability to induce Th9 and Th17 differentiation. TGM was able to induce similar levels of both Th17 and Th9 T-cells from the naïve population as TGF-β. A difference between the function of TGF-β and TGM lies in their ability to induce fibrosis. Despite activating the TGF-β signaling pathway, TGM does not induce fibrosis to the same extent as TGF-β. Administration of TGM induces fibrosis in human lung fibroblasts but TGM is less fibrogenic than TGF-β at equivalent molar concentrations[230]. TGM does not promote the deposition of smooth muscle actin to the same extent as TGF-β, requiring much higher concentrations to achieve similar amounts of actin deposition[230] (Fig. 10A-B). This indicates a potential use of TGM as a TGF-β mimic

therapeutically, as it acts in a similar immunogenic manner but is not as fibrogenic. However, TGM needs to be appropriately characterized: how it binds to the TGF-β receptors, and how this binding may be leveraged for therapeutic benefit needs to be further studied, which is addressed in the next several chapters.

2.4 Tables and Figures

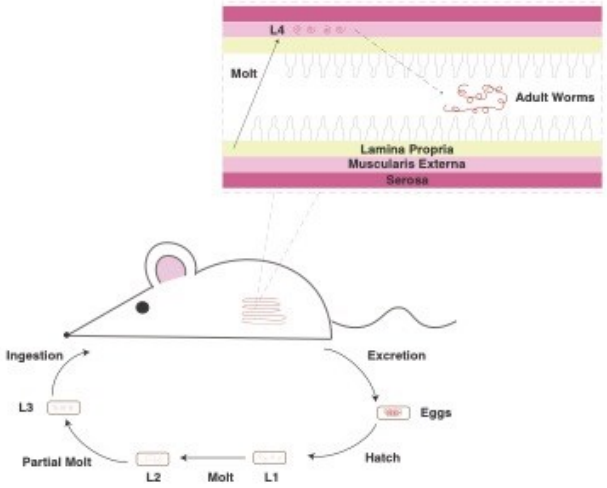

Figure 6: Life cycle of H. polygyrus in its murine host. *H. polygyrus* eggs are shed from mouse fecal matter, hatch into L1 larvae, molt into L2 larvae, and partially molt into the L3 larval stage. The worms are then ingested by the mouse following which they travel into the small intestine and invade the mucosa 24 hours post infection. Following three more days of infection the worms molt into the L4 larval worm stage and invade the intestinal muscle, and will mature into the adult worms over eight more days of infection. At ten days post-infection the worms return to the intestinal lumen and mate with other adult worms. Adapted from "Immunity to the model intestinal helminth parasite *Heligmosomoides polygyrus*"[201], used under CC By 4.0.

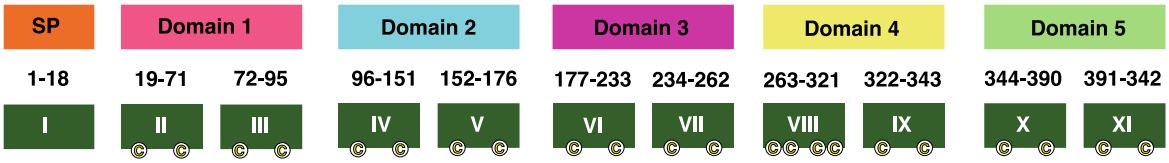

Figure 7: H. polygyrus TGF-β mimic exon partitioning. Exon and intron structure of Hp-TGM as correlated to five domains plus signal peptide, top: domains colored, spanning two exons each exons, middle: residue boundaries for domains noted, bottom: exons indicated with marked with roman numerals, cysteine residue position denoted with circles. Figure from "A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T-cells" by Johnston et. al.[230], used under CC By 4.0.

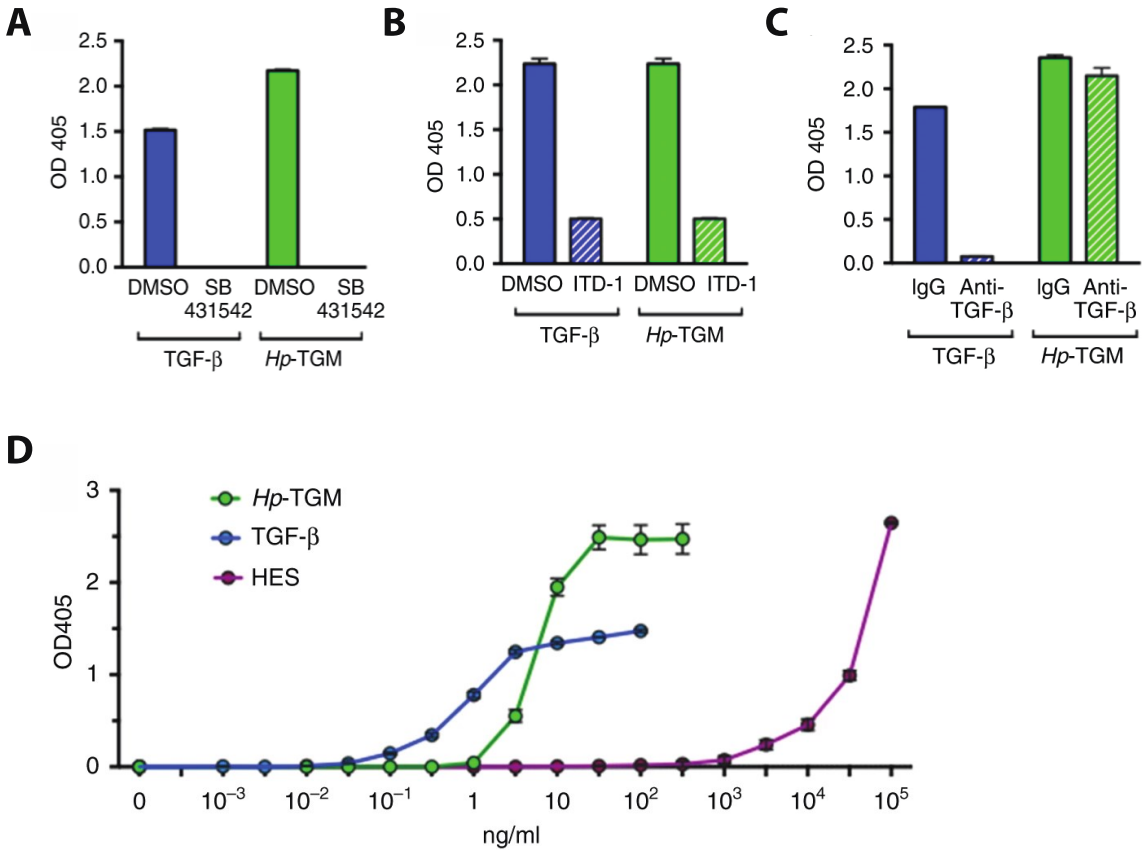

Figure 8: H. polygyrus TGF-β mimic signals through the TGF-β signaling pathway.

A-D. A-C. MFB-F11 bioassay shows TGF-β signaling activity following 24 hours of human TGF-β1 (blue) and TGM (green) incubation after administration of a TβRI inhibitor (A, right panels), a TβRII inhibitor (B, right panels), or an anti-TGF-β monoclonal antibody (C, right panels), in comparison to control (A-B, DMSO left panels, C, IgG left panels). N ≥ 2, independent wells and replicates. D. MFB-F11 bioassay shows TGF-β signaling activity following 24 hours of incubation with either TGF-β (blue), HES (purple), or TGM (green). N ≥ 3, independent wells and replicates. Figure from "A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T-cells" by Johnston et. al.[230], used under CC By 4.0.

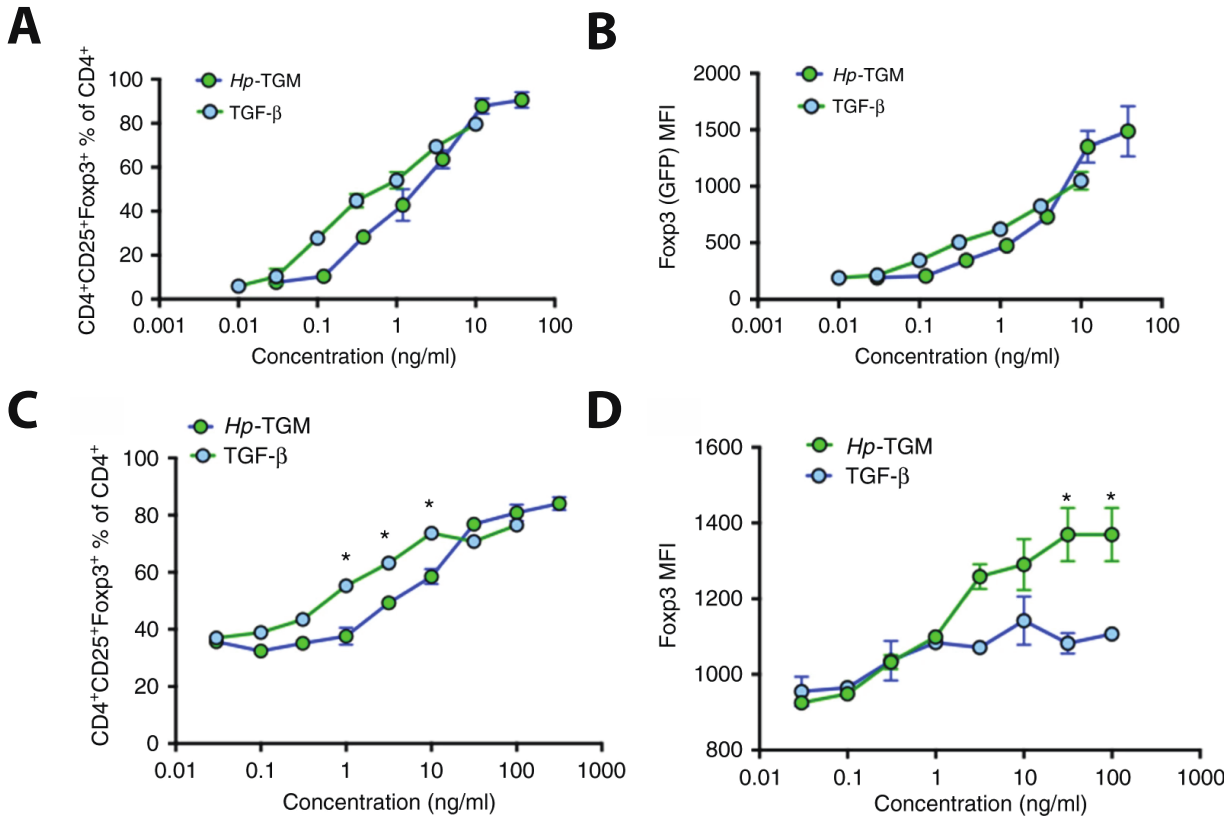

Figure 9: H. polygyrus TGF-β mimic induces the conversion of naïve CD4+ T-cells into Foxp3+ T_{regs}.

A-D. CD4+ T-cells, either murine (A) or human (C), were stimulated with either TGM (green circles) or TGF-β (light blue circles) and percent conversion to Foxp3+ T-regs displayed. B, D. The mean fluorescence intensity of the Foxp3 in the converted Foxp3+ T-cells originated from murine (B) or human (D) T-cells is displayed as a function of TGM (green circles) or TGF-β (light blue circles) concentration. N = 4 independent experiments, 2 technical replicates per concentration (A,B), 2 independent experiments (C, D). Figure from "A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T-cells" by Johnston et. al.[230], used under CC By 4.0.

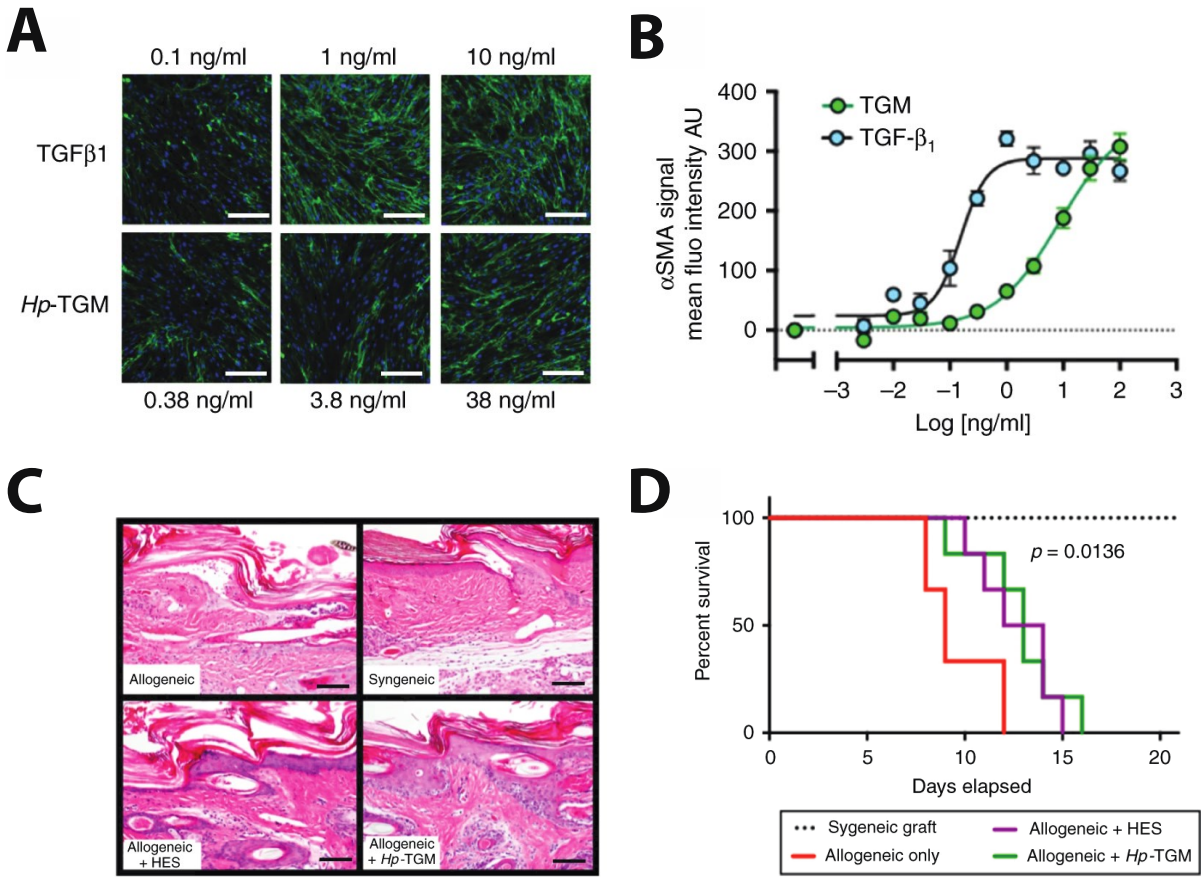

Figure 10: H. polygyrus TGF-β mimic prolongs transplant survival time and is not as fibrotic as TGF-β

A-D. A-B. TGM (green circles) induces less fibrogenesis than TGF-β (light blue circles) in human lung fibroblasts as shown via collagen deposition (A) and smooth muscle actin deposition (B), collagen stained with AlexaFluor 488 (green) and DAPI (blue). N = 3, four technical replicates per experiment. C-D. HES and TGM act to increase lifetime survival of mouse allografts. C. Histology of mouse allografts one week post transplantation, scale bar 100 micrometers. D. Percent survival of mouse grafts upon exposure to HES (purple) and TGM (green) in comparison to no treatment (red) or control syngeneic graft (black), n = 6 for all but syngeneic graft (n = 3). Figure from "A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T-cells" by Johnston et. al.[230], used under CC By 4.0.

3.0 Characterization of TGM Domains and Binding Properties

3.1 Introduction

Dysregulation of TGF-β signaling is related to several pathologic conditions, including multiple hereditary skeletal and connective tissue disorders[231-233], fibrotic disorders[234], familial cancer syndromes[235] and soft tissue cancers[236-238]. However, TGF-β targeting therapeutics have not been successful in clinical trials due to the ligand's pleiotropic nature and off-target side-effects. Neutralizing antibodies have not succeeded in TGF-β inhibition due to the stores of latent extracellular matrix-located TGF-β that is not readily inhibited due to physical location, while the receptor kinase inhibitors have not succeeded due to lack of specificity with inhibition of a wider variety of TGF-β ligands[239-241]. TGM is a potential alternative ligand to act as a TGF-β targeting therapeutic. In previous studies, TGM was not only functional as an immunosuppressive therapeutic, increasing graft survival in a mouse skin allograft model, but also acted in a less fibrotic manner than native TGF-β (Fig. 10). Understanding the structural and functional properties of TGM will aid in the development of medications targeting the TGF-β signaling pathway and in the understanding of parasitic molecular mimicry mechanisms.

The protein that has been previously denoted as TGM will henceforth be designated TGM-1, as TGM-1 is part of a larger family of transforming growth factor beta mimic proteins present within the *H. polygyrus* genome. Six of these proteins are present in the adult worm stage (TGM-1 through TGM-6), while the other three, TGM-7 through TGM-10 were found in the secretome at the L4 larval stage[242]. Of the six adult worm TGM family proteins, TGM-5 is missing domain

4 and TGM-6 is missing domains 1 and 2 (Table 2)[242]. Of the three L4 TGM family proteins,TGM-7 and TGM-8 have two additional domains following domain 3, TGM-9 lacks domains 1 and 2, and TGM-10 lacks domains 3 and 4 (Table 2)[242]. Previous studies have shown that only domains 1-3 are required for signaling through the TGF-β pathway; domain deletion constructs of TGM-1 were tested using the MFB-F11 signaling assay, and deletion constructs with domains 4 and 5 deleted were able to signal, whereas any deletion constructs with domains 1-3 or any combination deleted were unable to activate the TGF-β signaling pathway[242]. This indicates that domains 1-3 are likely responsible for binding to the type I (TβRI) and type II (TβRII) TGF-β receptors, with the role of domains 4 and 5 unknown.

Each of the TGM protein domains shares sequence similarity to the complement control protein (CCP) family, which are rich in β-strands and contain two disulfide bonds. This is in stark structural contrast to the TGF-β isoforms, which are disulfide-linked homodimers. Despite this structural difference, both the TGM proteins and the TGF-β isoforms are capable of binding TβRI and TβRII, indicating that either the TGM proteins have evolved to take advantage of alternate sites on TβRI and TβRII to signal, or that they are taking advantage of the existing binding sites that TGF-β utilizes in order to bind the two proteins.

Thus, the aim of this chapter is to investigate the following questions: 1) which domains, alone or in pairs, of TGM-1 bind to the type I and type II TGF-β receptors, and with what affinity, 2) what is the promiscuity of binding of TGM-1 to the other TGF-β family receptors, and 3) does TGM compete with TGF-β for binding to TβRI and TβRII.

3.2 Methods

Cloning, expression, and purification of TGM domain constructs. DNA inserts coding for the domains of TGM-1 D1, TGM-1 D2, TGM-1 D3, and TGM-1 D12 were inserted into a modified pET32a vector between the KpnI and HindII (AgeI for TGM-1 D3) sites (Table 2). Constructs were expressed in BL21(DE3) cells which were cultured at 37 °C. When light scattering at 600 nm was 0.8, protein expression was induced with 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Unlabeled TGM domain samples were produced in LB, and ^{15}N labeled samples were produced in minimal medium (M9 salts) with the addition of 1g L^{-1} of $^{15}NH_4Cl$ and 15 mL of 20% w/v D-glucose. Ampicillin was added at a concentration of 150 mg L^{-1} for plasmid selection.

Cell pellets were resuspended in lysis buffer (25 mM Na_2HPO_4, 50 mM NaCl, 5 mM imidazole, 10 μM leupeptin, 10 μM pepstatin, 1mM benzamidine, pH 8.0) and sonicated. Following sonication, the solution was centrifuged (20 min, 10000g), and the pellet was resolubilized in the lysis buffer with 0.5 M NaCl. The resultant solution was centrifuged (20 min, 10000g) and the pellet was resolubilized in the original lysis buffer with 1% Triton X-100 and centrifuged. The remaining pellet was resuspended in lysis buffer containing 8 M urea and solubilized overnight. Remaining insoluble material was removed via centrifugation and the resulting solution purified using a 50 mL nickel affinity column. The bound protein was eluted using a gradient of 0 – 0.5 M imidazole.

The protein was then treated with reduced glutathione (GSH), at a final concentration of 2 mM as diluted into the total folding buffer; the total volume of the folding buffer was determined such that the concentration of protein in the buffer was 0.1 mg mL^{-1}. The protein was incubated with the GSH at 25 °C for 30 minutes and then added dropwise to the folding buffer (0.1 M Tris,

1mM EDTA, 0.5 mM oxidized glutathione) and stirred at 4 °C for 16 hours. The solution was concentrated using an Amicon cell with a 5 kDa molecular weight cutoff filter and dialyzed into 25 mM Na_2HPO_4 50 mM NaCl pH 8.0. Thrombin was then added at a concentration of 4U per mg of TGM and incubated at room temperature overnight. The protein was bound to a nickel affinity column (NiNTA, Qiagen) and the column flow-through, along with a column wash with the dialysis buffer was collected. For purification of TGM-1 D1 and TGM-1 D12, the protein was dialyzed to 25 mM CHES pH 9.0 and purified via Source Q ion exchange chromatography in a buffer of 25 mM CHES. The proteins were eluted using a gradient of 0 – 0.5 M NaCl. For purification of TGM-1 D2 and TGM-1 D3, the protein solution was dialyzed to 25 mM sodium acetate pH 5.0 and purified via Source S ion exchange chromatography in a buffer of 25 mM sodium acetate, 2M Urea, pH 5.0. The proteins were eluted using a gradient of 0 – 0.5 M NaCl.

DNA inserts coding for the domains of TGM-1 FL were inserted into a pSecTag vector between the AscI and ApaI sites (Table 2). Constructs were transfected in expi293 cells, and the cells cultured at 37 °C. The conditioned medium was collected after 4-5 days of culture. Insoluble material was removed via centrifugation and the resultant solution was dialyzed into 25 mM Na_2HPO_4 150 mM NaCl 1 mM $NiSO_4$ pH 8.0 and purified via nickel affinity chromatography in a similar manner as performed for the TGM-1 individual domains. PNGASE-F was added at a 1:100 ratio and incubated with the protein solution at 37 °C overnight to remove any glycosylation. The resultant protein was then concentrated and dialyzed into 25 mM Tris 150 mM NaCl pH 8.0 and purified via size exclusion chromatography using a Superdex 16-60 200 column.

Cloning, expression, and purification of TGF-β receptor constructs. DNA inserts coding for the extracellular domains of TβRI and TβRII were inserted into a modified pET15b and pET32a vector respectively between the NdeI and BamHI sites for TβRI and between the NdeI and HindIII

sites for TβRII (Table 3). Protein expression and purification was performed as mentioned in previous protocols[36,243,244]. DNA inserts coding for the extracellular domains of ActRII, ActRIIB, and BMPRII were inserted into a modified pcDNA 3.1+ vector between the NotI and XhoI sites. The pcDNA 3.1 vector was modified with a signal peptide followed by a NotI site. The construct followed a structure of: Signal Peptide-His_6-thrombin cleavage site-Receptor (Table 3). Constructs were expressed in expi293 cells, transfected, and cultured at 37 °C. Purification was performed in a similar manner as mentioned for TGM-1 FL.

Cloning, expression, and purification of biotinylated avi-tagged protein constructs. TβRI and TβRII constructs were altered to include an avi-tag via addition of the sequence "GLNDIFEAQKIEWHE" at the C-terminus. Purification proceeded as reported previously[36,243,244]. Biotinylation was performed using the BirA biotin ligase[245], and the biotinylation reagents were removed during ion-exchange chromatography.

SPR binding analysis. All SPR experiments were performed with a BIAcore X100 system. Neutravidin was coupled to the surface of a CM5 chip and TβRI and TβRII were captured at a maximum density of 150 RU. Neutravidin coupling was performed by EDC-NHS activation of the chip, followed by neutravidin (buffer: sodium acetate, pH 4.5) injection over the surface until the RU increased by 6000-15000 RU. All experiments were run in 25 mM HEPES, 150 mM NaCl, 0.005% P20 surfactant, pH 7.4, at an injection rate of 100 μL min^{-1}. The surface was regenerated in between each injection with a 30 second injection of 1 M guanidine hydrochloride. The data was processed by double-referencing, with both a control cell coupled with neutravidin but without any ligand conjugated to the surface and also with 8 blank buffer injections at the beginning of the run before injection of the samples. The data was analyzed by fitting the results to a kinetic 1:1 model using the SPR analysis software Scrubber.

ITC binding analysis. ITC experiments were performed with a Microcal PEAQ-ITC system. Experiments with TβRI were performed in 25 mM HEPES, 50 mM NaCl, 0.5% NaN_3, pH 7.5 at 25 °C, with 19 2.0 μL injections with a duration of 4 s, a spacing of 150 s, and a reference power of 10. Experiments with TβRII were performed in 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 at 35 °C, with 15 2.5 μL injections with a duration of 5 s, a spacing of 150 s, and a reference power of 10. All samples were dialyzed against the same ITC buffer before loading into the system. Data was globally fit using the programs NITPIC[246], SEDPHAT[247,248], and GUSSI[249].

SEC-MALS analysis. Experiments were performed with a miniDAWN TREOS multi-angle light scattering detector with a Wyatt QELS dynamic light scattering detector and an Optilab T-rEX refractometer. The size-exclusion column attached was a Superdex 10-300 75 column, and the samples loaded were concentrated to 2 mg mL^{-1}. The TGM-1 D2:TβRI and the TGM-1 D12:TβRI samples were run in 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. The TGM-1 D12:TβRI sample was prepared at a 1.6:1 molar ratio, and the TGM-1 D2:TβRI sample was prepared and passed over the column at a 2:1 molar ratio. The column flow rate was 0.5 mL min^{-1}, and all data was analyzed using the Astra software. Samples of TGM-1 D2 alone and TGM-1 D12 alone were also passed over the column, and the elution volumes and scattering values detected and analyzed to calibrate the complex samples.

NMR 2D experiments. All NMR experiments were run in 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0, 310K, with the exception of some of the ^{15}N TGM-1 D1 experiments, which included CHAPS. Each sample was run in a 5 mm susceptibility-matched microtube for data collection. NMR data was collected with Bruker 600, 700, and 800 MHz spectrometers containing a 5 mm ^{1}H (^{13}C, ^{15}N) z-gradient "TCI" cryogenically cooled probe. ^{1}H-^{15}N HSQC spectra were acquired as described, with water flipback[250] and WATERGATE suppression pulses[251]. ZZ-exchange

experiments were recorded as described[252]. NMR spectra were processed, analyzed, and visualized using NMRPipe[253] and NMRFAM-SPARKY[254]. NMR samples were produced by adding each individual component in a 1:1.2 molar ratio, with the labeled component at a concentration of ca. 20-100 μM, in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0.

3.3 Results

3.3.1 TGM domains can be individually isolated and expressed natively in *E. coli*

TGM-1 is a secreted signaling protein composed of five domains of approximately 90 amino acids each, with each domain containing four cysteine residues paired in two disulfide bonds, with the exception of domain 4 ,which contains three disulfide bonds. As these domains shared sequence similarity to the CCP protein family, which are composed of separable units of CCP domains arranged in a semi-linear fashion, it was hypothesized that the TGM-1 domains may be acting as distinct units to bind the TGF-β receptors, rather than a shared interface between multiple domains. The domain boundaries are characterized with respect to the cysteine residues as is characteristic for CCP domains, with the N-terminal boundary one residue prior to the first cysteine and the C-terminal boundary three residues following the fourth cysteine.

Previous studies had used the full five domains of TGM-1 along with the signal peptide, produced recombinantly in eukaryotic cells[230,242], with a C-terminal myc tag and His tag. For ease of isotopic labeling for nuclear magnetic resonance (NMR) assessment of TGF-β receptor binding, individual domains of TGM-1: TGM-1 D1, TGM-1 D2, and TGM-1 D3, were expressed recombinantly in *E. coli*. To increase protein expression, the individual domains were expressed

in a pET32a vector as thioredoxin fusion constructs with an internal His-tag and thrombin cleavage site immediately preceding the region coding for the TGM-1 domains. Additionally, a short 4-linker GSGT tag was added to the N-terminus, to create a short linker region as there is only one residue prior to the first cysteine residue for each of the domains (Table 3).

The proteins were initially produced under conditions to promote partitioning into the soluble fraction via reduced temperature growth at 18 °C to avoid the need for refolding and increase the yield of natively-folded protein. However, this was unsuccessful, thus the proteins were produced as unfolded monomers in inclusion bodies at 37 °C. The inclusion bodies were then solubilized in 8M urea, further purified over a His-tag selecting nickel column and refolded in a Tris buffer at a concentration of 0.1mg/mL using reduced and oxidized glutathione. The thioredoxin tag was cleaved with thrombin, and the final purification to a single species was performed using ion-exchange chromatography. The one exception to this was TGM-1 D3, which when expressed at 37 °C was present in both the soluble and insoluble fraction. The soluble fraction was purified similar to the insoluble fractions, with the exception of the folding step which was omitted. The purifications were validated by mass on an SDS gel, and also using mass spectrometry. As the proteins contained disulfide bonds, the calculated mass was reduced by 4 Da, one for each lost hydrogen during disulfide bond formation. Mass spectrometry for the individual domains showed a mass equivalent to the calculated molecular weight with a variance of 0.5 Da.

3.3.2 Initial NMR characterization of individual domains

Individual domains of TGM-1 were uniformly labeled with ^{15}N using minimal salts medium. The proteins were grown at 37 °C, induced, and purified as mentioned previously.

Isotopic labeling acts as a method to check for native protein folding outside of functional assays. As the TGM domains are sequentially similar to the CCP family, whose domains are structured, with a multitude of β-strands and two disulfide bonds, the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectra should be indicative of folded proteins. For native proteins, the spectra should be well-dispersed upfield and downfield of the random coil region (7.8 – 8.6 ppm in the ^{1}H dimension), with the expected number of amide resonances. TGM-1 D3 contains ninety residues, with nine prolines. Thus, the expected number of backbone amide residues is 80 residues. Upon recording of a ^{1}H-^{15}N HSQC spectra in 25mM phosphate buffer pH 6.0, 37 °C, 80 backbone resonances were observed, which validates the identity of the protein. These signals were well-dispersed, indicating native folding (Fig. 11A).

TGM-1 D2 contains eighty-four residues, seven of which are prolines. The expected number of backbone amide resonances is 76; however, uniquely 106 resonances were observed, which could be due to either 1) disulfide-pairing heterogeneity, or 2) slow conformational dynamics (Fig. 11B, C). The hypothesis of slow conformational dynamics was tested using a ^{1}H-^{15}N ZZ-exchange experiment[255]. This experiment is similar to the ^{1}H-^{15}N HSQC experiment but includes a mixing time with variable delay periods between the indirect and direct detection periods. During this delay the magnetization is aligned to the Z-axis. If there is a singular residue with two resonances at (ω_{1H},ω_{1N}) and (ω_{2H},ω_{2N}) with conformational exchange between the two peaks, as the mixing time increases, cross-peaks at (ω_{1H},ω_{2N}) and (ω_{2H},ω_{1N}) will be detected. This experiment was performed at mixing times of 50, 100, and 250 msec, and twelve pairs demonstrated exchange cross-peaks within the mixing delays. This indicates that TGM-1 D2 is present in two different conformations while in solution. Due to the timescale and the large number

of proline residues, this was hypothesized to be due to proline cis:trans isomerization, though no further testing was performed. The signals are well-dispersed, indicating natively-folded protein.

TGM-1 D1, contains 90 residues, four of which are prolines, with 85 expected amide backbone residues on the HSQC. However, far fewer than the expected number of resonances were detected, and the resonances that were visible were clustered in the random coil region (7.8 -8.6 ppm in the ^{1}H dimension) (Fig. 12A). This could be due to incorrect pairing of cysteines during the folding process or due to sample aggregation. To alleviate any aggregation, the sample concentration was reduced and 10mM CHAPS was added (Fig. 12B-D). When the concentration was reduced from 100 μM to 20 μM, and 10mM CHAPS added, the resonances were more dispersed, with several signals appearing outside of the random coil region, and a total of 85 peaks were detected. This effect was seen to a lesser extent in spectra of ^{15}N TGM-1 D1 at 20 μM without the CHAPS and also with a sample at 200 μM with 10mM CHAPS (Fig. 12B-D). This indicates that the purified TGM-1 D1 was natively folded but aggregated under the initial NMR sample conditions.

3.3.3 TGM-1 domain 3 is responsible for binding TβRII

To check for binding of the TGM-1 domains to TβRII, unlabeled TβRII was added in a slight excess to NMR samples of ^{15}N TGM-1 D1, ^{15}N TGM-1 D2, and ^{15}N TGM-D3 (Fig. 13A-D), with the sample of ^{15}N TGM-1 D1 containing 10mM CHAPS to keep the resonances well-dispersed. The ^{1}H-^{15}N HSQC spectra of ^{15}N TGM-1 D1 and ^{15}N TGM-1 D2 did not show any perturbation in the backbone amide resonances, indicating no binding. However, addition of unlabeled TβRII to ^{15}N TGM-1 D3 caused a visible perturbation in over half of the backbone amide resonances, indicative of binding (Fig. 13C). A titration series was performed, with the ^{1}H-

^{15}N HSQC spectrum of free ^{15}N TGM-1 D3 compared to that of the ^{15}N TGM-1 D3 titrated with 0.35, 0.7, and 1.3 molar equivalents of unlabeled TβRII. For several of the amide resonances in the spectra of the intermediate titration points, peaks originating from both the free and the bound forms of ^{15}N TGM-1 D3 were present, which is characteristic of slow-exchange binding (Fig. 13D). Slow-exchange binding occurs when the rate of exchange between two states is less than the frequency difference between the resonances associated with the two states, thus showing both the unbound and bound versions simultaneously; this is usually indicative of tight binding.

The NMR data demonstrates that TGM-1 D3 is the main binding partner of TβRII, whereas TGM-1 D1 and TGM-1 D2 likely do not bind TβRII. This is confirmed by the reverse titration experiments, in which unlabeled TGM-1 D1, TGM-1 D2, and TGM-1 D3 are separately titrated into ^{15}N TβRII NMR samples. Only upon titration of unlabeled TGM-1 D3 do the backbone resonances of ^{15}N TβRII shift (Fig. 14A-D). Similarly, intermediate titration points (0.4, 0.8, and 1.2 molar equivalents TGM-1 D3:^{15}N TβRII), show that a majority of the residues are in the slow-exchange regime, indicating tight binding. While addition of TGM-1 D2 and TGM-1 D1 do not directly perturb any NMR signals of TβRII, it may be that the 10 mM CHAPS present in the ^{15}N TGM-1 D1:TβRII sample may interfere with potential binding interactions, and that in the ^{15}N TβRII:TGM-1 D1 sample that the concentration of TGM-1 D1 needed to fully titrate the ^{15}N TβRII may have led to formation of an aggregate no longer capable of binding ^{15}N TβRII.

To confirm the binding of TGM-1 D3 to TβRII, and to elucidate the potential role of TGM-1 D1, surface plasmon resonance (SPR) experiments were performed. Avi-tagged biotinylated TβRII was captured onto a streptavidin-coated CM5 chip, and varying concentrations of TGM-1 D1, TGM-1 D2, TGM-1 D3, and full-length TGM-1 (TGM-1 FL) were injected over the surface. A clear concentration-dependent response was present when both TGM-1 D3 and TGM-1 FL were

injected over the immobilized TβRII, but not when either TGM-1 D1 or TGM-1 D2 were injected (Fig. 15), consistent with the conclusion attained by the NMR experiments. K_D measured by a 1:1 kinetic model of the SPR data was determined to be 610 ± 10 nM for TGM-1 FL binding to TβRII and 910 ± 20 nM for TGM-1 D3 binding to TβRII (Table 5). The maximum concentrations of TGM-1 D1 injected over the TβRII surface was 1 μM, which by the ^{1}H-^{15}N NMR spectra of TGM-1 D1 (Fig. 12C) would not be a high enough concentration to form aggregates. Thus, it is evident that TGM-1 D3 is the domain that is mostly responsible for the binding of TGM-1 FL to TβRII.

The one caveat to SPR experiments is that immobilization of one of the proteins on the chip may alter the binding properties. Though a 1:1 model for binding was assumed for the SPR analysis, there are other biophysical methods that are able to determine binding stoichiometry more accurately. Orthogonally, isothermal titration calorimetry (ITC) experiments were performed, in which individual domains of TGM-1, TGM-1 D1, TGM-1 D2, TGM-1 D3, along with TGM-1 FL were titrated into a cell containing TβRII (Fig. 16). The titration of TGM-1 D1 and TGM-1 D2 into TβRII did not produce any measurable change in enthalpy (Fig. 16 A-B), while the titration of both TGM-1 D3 and TGM-1 FL into TβRII produced an exothermic response with large negative changes in enthalpy (Fig. 16). Fitting the binding isotherms globally gave affinities similar to those obtained by the SPR experiments, with the K_D of binding and enthalpies for the TGM-1 FL:TβRII titration and the TGM-1 D3:TβRII titration being 550 nM and -7.1 kcal/mol and 1200 nM and -11 kcal/mol respectively (Table 6). Binding isotherms could not be fit for the experiment titrating TGM-1 D1 into TβRII or the experiment titrating TGM-1 D2 into TβRII, confirming that TGM-1 D3 is the sole domain responsible for the binding of TGM-1 FL to TβRII.

3.3.4 TGM-1 domains 1 and 2 are responsible for binding TβRI

Similar NMR, SPR, and ITC studies were performed to assess the binding of the individual TGM-1 domains to TβRI. Unlabeled TβRI was added in a slight excess of equimolar ratios (1:1.2, TGM-1 domain to TβRI) to NMR samples of ^{15}N TGM-1 D1, ^{15}N TGM-1 D2, and ^{15}N TGM-D3 individually (Fig. 17A-D). The ^{1}H-^{15}N HSQC spectra of ^{15}N TGM-1 D1 and ^{15}N TGM-1 D3 did not show any perturbation in the backbone amide resonances, indicating no binding. The sample of ^{15}N TGM-1 D1 contained 10mM CHAPS to keep the resonances well-dispersed, which may have potentially interfered with the binding (Fig. 17). However, the addition of unlabeled TβRI to ^{15}N TGM-1 D2 caused significant perturbation in over half of the backbone amide resonances, indicative of binding (Fig. 17B). Free ^{15}N TGM-1 D2 was then titrated with increasing amounts of unlabeled TβRI (0.36, 0.8, 1.4 molar equivalents, Fig. 17D), with several of the amide resonances on the slow-exchange timescale. The addition of TβRI to the ^{15}N TGM-1 D2 sample also eliminated the doubling seen with free ^{15}N TGM-1 D2, showing that TβRI is able to stabilize the TGM-1 D2 into one of its conformations (Fig .18). The NMR data demonstrates that TGM-1 D2 is the main binding partner of TβRI, which was further tested by the reverse titration experiments, in which unlabeled TGM-1 D2 and TGM-1 D3 was titrated into ^{15}N TβRI NMR samples, and the resultant spectra recorded (Fig. 19). When unlabeled TGM-1 D2 was titrated into ^{15}N TβRI, there were significant chemical shift perturbations with a pattern characteristic of slow-exchange binding (Fig. 19C). When unlabeled TGM-1 D3 was titrated into ^{15}N TβRI samples, there was no significant chemical shift perturbation (Fig. 19A).

However, when ^{15}N TβRI was titrated with TGM-1 D1 without CHAPS, this was not the case. Though there were no significant chemical shift perturbations (Fig. 20), there was a

significant weakening in intensity for a majority of the amide backbone peaks, and other peaks were mildly perturbed (Fig. 20). This likely indicates that TGM-1 D1 is forming an aggregated species, but that there are enough interactions between TGM-1 D1 and ^{15}N TβRI that the latter is being drawn into the aggregate. As such, it is likely that TGM-1- D1 is binding to TβRI, albeit weakly.

To confirm the roles of TGM-1 D1 and TGM-1 D2 in binding TβRI, SPR experiments were performed with immobilized biotinylated avi-tagged TβRI. A clear concentration-dependent response was present when both TGM-1 D2 and TGM-1 FL were injected onto the chip surface, but not when either TGM-1 D1 or TGM-1 D3 was injected (Fig. 21), consistent with the conclusion that TGM-1 D2 is the main domain responsible for TGM-1 FL binding to TβRI. The binding affinity to TβRI was 13 ± 1 nM for TGM-1 FL and 310 ± 10 nM for TGM-1 D2 (Table 5). From the 24-fold affinity difference, TGM-1 D2 is likely not the sole domain responsible for binding, which is supported by the NMR evidence of binding between TGM-1 D1 and TβRI. To investigate this, a construct containing both TGM-1 D1 and TGM-1 D2, TGM-1 D12, was expressed in *E. coli* (Table 3). This construct was assessed for binding to TβRI and TβRII via SPR. There was no binding of TGM-1 D12 to TβRII (Table 5, Fig. 15E), but there was clear, robust binding of TGM-1 D12 to TβRI by SPR. The K_D was determined as 24 ± 1 nM, which recapitulates the full binding capacity of TGM-1 FL to TβRI.

To validate this, ITC experiments were performed titrating TGM-1 domain constructs into a cell containing TβRI (Fig. 22). The titration of TGM-1 D1 and TGM-1 D3 into TβRI did not produce any measurable change in enthalpy (Fig. 22 A, C), while the titrations of TGM-1 D2, TGM-1 D12, and TGM-1 FL into TβRI all produced an exothermic response with large negative changes in enthalpy (Fig. 22, B, D, E). Fitting these binding isotherms globally gave affinities that

were similar to those obtained by SPR, with the K_D of binding and enthalpies for the TGM-1 FL:TβRI titration, the TGM-1 D2:TβRI, and the TGM-1 D12:TβRI titration being 52 nM and -17 kcal/mol, 1500 nM and -18 kcal/mol, and 25 nM and -19 kcal/mol respectively (Table 6). This data shows that while TGM-1 D2 is responsible for most of the binding affinity to TβRI, TGM-1 D1 is required to fully recapitulate binding. Binding isotherms could not be fit for the experiment titrating TGM-1 D1 into TβRI or the experiment titrating TGM-1 D3 into TβRI, highlighting that TGM-1 D12 are both required for binding to TβRI, but that TGM-1 D2 likely is making the majority of the contacts with the receptor.

3.3.5 TGM-1 binds to TβRI and TβRII in a 1:1:1 stoichiometry

The ITC data shows that TGM-1 D3 binds to TβRII with a 1:1 stoichiometry. Though the NMR titration data does imply a 1:1 stoichiometry between TGM-1 D2 and TβRI (Fig. 17, 19), due to difficulties with fitting the ITC data of the TGM-1 D2:TβRI binding, the binding stoichiometry by ITC was determined to be 0.6, which is closer to a 0.5 stoichiometry. This is likely due to difficulties in fitting the ITC data and is in contrast to the TGM-1 FL:TβRI interaction by ITC, which indicates a 1:1 stoichiometry (Table 6, Fig. 22).

However, to reconcile this discrepant data, SEC-MALS experiments were performed with both the TGM-1 D2:TβRI complex and the TGM-1 D12:TβRI complex (Fig. 23). TGM-D2 and TGM-1 D12 were first run alone on a Superdex 10-300 column, and the masses calculated using multiangle light scattering (MALS). The mass of TGM-1 D2 is expected to be 9.3 kDa, and the mass of TGM-1 D12 is expected to be 18.8 kDa. These masses were then used to calibrate the SEC-MALS experiments for following runs in which TGM-1 D2 and TGM-1 D12 were titrated with TβRI (expected mass 9.6 kDa). Upon titration of TGM-1 D2 with TβRI, two peaks were

visible on the 280 nM absorbance axis (Fig. 23A), with the complex mass calculated to be 17.1 ± 0.8 kDa, which is close to the 1:1 stoichiometry at 18.9 kDa.. The free TGM-1 D2 peak in this chromatogram was calculated as 10.7 ± 0.2 kDa, which is less than 1.5 kDa from the expected mass. There was an additional peak shown on the light-scattering axis earlier in the chromatogram, approximately at 7.2 mL; however, as there is no corresponding peak on the 280 nM absorbance axis, this is likely not due to any protein in the sample.

Upon titration of TGM-1 D12 with TβRI (Fig. 23B), the combined mass was calculated to be 28.8 ± 0.3 kDa, which is close to the 1:1 stoichiometry at 28.4 kDa. The free TGM-1 D12 peak was calculated as 18.5 ± 0.2 kDa, which is very close to the expected mass. There was an additional peak shown on the light-scattering axis earlier in the chromatogram, approximately at 10.1 mL; this corresponds to a small peak on the 280 nM absorbance axis, likely indicating a relatively small amount of aggregated TGM-1 D12 scattering heavily. The SEC-MALS data confirms a 1:1 stoichiometry.

3.3.6 TGM-1 does not bind other TGF-β family receptors

The TGF-β ligand family can be broadly divided into the TGF-β, activin, and BMP arms, comprising a total of five type I and seven type II receptors. As such, there is the potential that TGM-1, along with some of the other members of the TGM family, may activate other TGF-β receptors, including the other type I receptors Alk 1, Alk 2, Alk3, Alk4, Alk6, and Alk7, and the other type II receptors ActRII, ActRIIA, BMPRII, MISRII. Unpublished ELISA assays ran by collaborators in the Maizels lab (data not shown) show that TGM-1 may bind to ActRIIB, one of the type II receptors in the activin arm.

Binding of the extracellular domains of ActRII, ActRIIB and BMPRII to TGM-1 and other TGM isoforms was assessed via native gel (Fig. 24). The native gels show that TGM-1, TGM-2, TGM-4 and TGM-6 do not bind to the above-mentioned receptors. However, while native gels can test for high-affinity binding, they do not have the capability to test for low-binding affinity, and thus weak interactions between the TGM isoforms and the TGF-β family type II receptors cannot be ruled out.

ELISA data showed relatively few potential interactions between TGM-1 and type I TGF- receptors, with only Alk4 potentially binding to TGM-1 FL. To test whether there was binding between TGM-1 FL and Alk4, NMR experiments were performed in which a ^{1}H-^{15}N HSQC experiment was performed for ^{15}N Alk4 free and bound to TGM-1 D2 (Fig. 25). As TGM-1 D2 is the main domain for binding to TβRI, it was hypothesized that similarly TGM-1 D2 would be the main domain that binds to Alk4. NMR experiments show that there is no binding between TGM-1 D2 and Alk4, and while this does not preclude binding of other domains of TGM-1 to Alk4, it does make this interaction much less likely. The data presented thus suggests that TGM-1 is specific for TβRI and TβRII and does not promiscuously bind to other receptors in the TGF-β signaling family.

3.4 Discussion

The results presented in this chapter indicate that the individual domains of TGM-1 can be recombinantly expressed, isolated, and purified. TGM-1 D2 is responsible for the majority of TGM-1's binding affinity to TβRI, and TGM-1 D3 is responsible for the full binding capacity of TGM-1 to TβRII. As TGM-1 D2 does not interact with TβRII and TGM-1 D3 does not interact

with TβRI, it appears that TGM-1 binds to the TGF-β receptors using separate domains, and thus binding of TGM-1 to the TGF-β receptors is modular and separable. Additionally, the binding of TGM-1 D2 to TβRI is enhanced by the weak binding of TGM-1 D1 to TβRI, with the two domains forming a composite interface with which to interact with TβRI.

This is distinct from the binding of TGF-β to TβRI and TβRII; for the native TGF-β ligand, recruitment of TβRII is necessary for binding of TGF-β to TβRI, and the complex forms as a 1:2:2 heterotetramer[256]. The TGF-β homodimer is able to recruit TβRII with a high affinity (ca. 50 nM) and this complex recruits TβRI with a similar affinity (ca. 30 nM). This is distinct from the mode of binding observed with TGM-1, which shows distinct domains responsible for separately binding TβRI and TβRII. TGM-1 binds to TβRI with a higher affinity (K_D ca. 50 nM) than TβRII (K_D ca. 500 nM), which is in direct contrast to how TGF-β recruits the receptors. One TGM-1 monomer binds to one type I receptor and one type II receptor, as shown by biochemical studies. This 1:1:1 stoichiometry, along with the difference in cooperative recruitment of the complex may be what leads to the functional differences between TGM-1 and TGF-β, though further studies into the assembly process would be required to delineate distinctive details.

Finally, the data in this chapter demonstrate that TGM-1 is specific for TβRI and TβRII over the other type I and type II receptors in the TGF-β signaling family. The lack of receptor promiscuity, despite the overall structural similarities between receptors in the TGF-β family suggests that TGM-1 has evolved to bind TβRI and TβRII specifically, rather than to all TGF-β receptors, due to their physiological role in immunosuppressive signaling.

As the TGM-1 domains are separable and capable of binding to TβRI and TβRII individually, domain deletion constructs of TGM-1 may play a potential role as TGF-β antagonists. If a TGM-1 D12 construct can bind TβRI in a way that displaces TGF-β, then TβRI could be

sequestered away from TGF-β and signaling could be prevented. Similarly, if a TGM-1 D3 construct can bind and sequester TβRII, then TGF-β signaling could be blocked via TβRII sequestration. Data in the following chapters will assess whether TGM-1 competes with TGF-β for binding to TβRI and TβRII as well as whether this can then be leveraged towards developing TGF-β antagonistic therapeutics, which would have potential roles in a variety of diseases, including soft-tissue cancers[236-238].

This chapter did not mention domains 4 and 5, as they are not required for binding to TβRI and TβRII or for TGF-β signaling[242]. They do potentiate TGM-1 signaling through the TGF-β pathway, which will be discussed in later chapters. Additionally, though it is known that TGM-1 D12 binds TβRI and TGM-1 D3 binds TβRII, the mechanism has yet to be discussed. These topics will be further discussed in the next several chapters.

3.5 Respective Contributions

All of the NMR experiments were performed by the author, Ananya Mukundan. The SPR experiments were performed by Chang Hyeock-Byeon, a technician in the Hinck laboratory. The ITC experiments for TβRII binding to the TGM-1 domains, both individually as well as the competition experiments were performed by Chang Hyeock-Byeon but processed and analyzed by the author Ananya Mukundan. The ITC experiments for TβRI binding to the TGM-1 domains, both individually as well as the competition binding experiments were performed, processed, and analyzed by the author Ananya Mukundan. Experiments assessing binding of TGM-1 to other TGF-β family proteins were performed, processed, and analyzed in their entirety by the author Ananya Mukundan. The data in Figures 11-14, Figures 17-20, and Figures 22-25 were gathered

and analyzed in their entirety by the author Ananya Mukundan, the data in Figures 15 and 21 were gathered and analyzed by Chang Hyeock-Byeon, and the data in Figure 16 were gathered by Chang Hyeock-Byeon and analyzed by Ananya Mukundan.

3.6 Tables and Figures

Table 2: H. polygyrus extended family of TGM proteins. Amino acid identity (%) with domains denoted with respect to TGM-1. Blank boxes indicate absence of domain. Adapted from "TGF-β mimic proteins form an extended family in the murine parasite *Heligmosomoides polygyrus*" by Smyth et. al.[242], used under CC By 4.0 / Edited from original.

TGM Member	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
TGM-1	100	100	100	100	100
TGM-2	100	100	93	99	81
TGM-3	100	100	100	99	70
TGM-4	71	90	85	73	84
TGM-5	91	90	95		87
TGM-6			66	42	32
TGM-7	64	65	59	20 30 25	33
TGM-8	62	63	60	19 30 22	32
TGM-9			60	36	32
TGM-10	47	50			28

Table 3: Sequences of TGM-1 domain constructs. Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

Construct	Coding region and description (* indicates stop codon)
TGM-1 D1	Residues 16-95 of *H. polygyrus* TGF-β Mimic, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-D1 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTGSGSGSDDSGCMPFSDEAATYKYVAKGPKNIEIPAQIDNSGMYPDYTHVKRFCKGLHGED TTGWFVGICLASQWYYYEGVQECDDR*
TGM-1 D2	Residues 96-176 of *H. polygyrus* TGF-β Mimic, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-D2 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTRCSPLPTNDTVSFEYLKATVNPGIIFNITVHPDASGKYELTYIKRICKNFPTDSNVQGHIIGMC YNAEWQFSSTPTCPAS*
TGM-1 D12	Residues 16-176 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-D12 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTGSGSGSDDSGCMPFSEAATYKYVAKGPKNIEIPAQIDNSGMYPDYTHVKRFCKGLHGEDTT GWFVGICLASQWYYYEGVQECDDRRCSPLPTNDTVSFEYLKATVNPGIIFNITVHPDASGKYPEL TYIKRICKNFP TDSNVQGHIIGMCYNAEWQFSSTPTCPAS*
TGM-1 D3	Residues 177-262 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-D3 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTGCPPLPDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRRCRALSQEADPGEFV AICYKSGTTGESHWEYYKNIGKCPDP*
TGM-1 D13	Residues 16-262 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Signal Peptide-TGM-D1-D3-Linker-Myc-Linker-His_6 METDTLLLWVLLLWVPGSTGDAAQPARRADDSGCMPFSDEAATYKYVAKGPKNIEIPAQIDNSG MYPDYTHVKRFCKGLHGEDTTGWFVGICLASQWYYYEGVQECDDRRCSPLPTNDTVSFEYLKA TVNPGIIFNITVHPDASGKYPELTYIKRICKNFPTDSNVQGHIIGMCYNAEWQFSSTPTCPASGCPPL PDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRRCRALSQEADPGEFVAICYKSGTTG ESHWEYYKNIGKCPDPGPEQKLISEEDLNSAVDHHHHHH*
TGM-1 FL	Residues 16-422 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Signal Peptide-TGM-FL-Linker-Myc-Linker-His_6 METDTLLLWVLLLWVPGSTGDAAQPARRADDSGCMPFSDEAATYKYVAKGPKNIEIPAQIDNSG MYPDYTHVKRFCKGLHGEDTTGWFVGICLASQWYYYEGVQECDDRRCSPLPTNDTVSFEYLKA TVNPGIIFNITVHPDASGKYPELTYIKRICKNFPTDSNVQGHIIGMCYNAEWQFSSTPTCPASGCPPL PDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRRCRALSQEADPGEFVAICYKSGTTG ESHWEYYKNIGKCPDPRCKPLEANESVHYEYFTMTNETDKKKGPPAKVGKSGKYPEHTCVKKV CSKWPYTCSTGGPIFGECIGATWNFTALMECINARGCSSDDLFDKLGFEKVIVRKGEGSDSYKDD FARFYATGSKVIAECGGKTVRLECSNGEWHEPGTKTVHRCTKDGIRTLGPEQKLISEEDLNSAVD HHHHHH*

Table 4: Sequences of TGF-β receptor constructs.

Construct	Coding region and description (* indicates stop codon)
TβRI 25-125	Residues 25-125 of the extracellular domain of human TGF-β type I receptor, NCBI NP_001394345 Linker-His_6-Linker-Thrombin Cleavage Site-Linker-Initiating Methionine-TβRI MGSSHHHHHHSSGLVPRGSHMAALLPGATALQCFCHLCTK DNFTCVTDGLCFVSVTETTD KVIHNSMCIAEIDLIPRDRPFVCAPSSKTGSVTTTYCCNQDHCNKIELPTTVKSSPGLGPVE *
TβRII 38-159	Residues 38-159 of the extracellular domain of human TGF-β type II receptor, NCBI NP_001394056 Initiating Methionine-TβRII MVTDNNGAVKFPQLCKFCDVRFSTCDNQKSCMSNCSITSICEKPQEVCVAVWRKNDENITLET VCHDPKLPYHDFILEDAASPKCIMKEKKKPGETFFMCSCSSDECNDNIIFSEEYNTSNPD
ActRII 23-123	Residues 23-123 of the extracellular domain of human Activin receptor type-2A, NCBI NP_001265508.1 Signal Peptide-Linker-His_6-Linker-Thrombin Cleavage Site-Linker-Initiating Methionine-ActRII MKWVTFLLLLFISGSAFSAAAGSSHHHHHHSSGLVPRGSHMGRSETQECLFFNANWEKDR TNQTGVEPCYGDKDKRRHCFATWKNISGSIEIVKQGCWLDDINCYDRTDCVEKKDSPEVY FCCCEGNMCNEKFSYFPEMEVT
ActRIIB 25-117	Residues 25-117 of the extracellular domain of human Activin receptor type-2B, NCBI EAW64536 Signal Peptide-Linker-His_6-Linker-Thrombin Cleavage Site-Linker-ActRIIB MKWVTFLLLLFISGSAFSAAAGSSHHHHHHSSGLVPRGSHMLEDPVPETRECIYYNANWELER TNQSGLERCEGEQDKRLHCYASWRNSSGTIELVKKGCWLDDFNCYDRQECVATEENPQVYFC CCEGNFCNERFTHLP
BMPRII 29-132	Residues 29-132 of the extracellular domain of human Activin receptor type-2B, NCBI XP_011509989 Signal Peptide-Linker-His_6-Linker-Thrombin Cleavage Site-Linker-Initiating Methionine-BMPRII MKWVTFLLLLFISGSAFSAAAGSSHHHHHHSSGLVPRGSHMNQERLCAFKDPYQQDLGIG ESRISHENGTILCSKGSTCYGLWEKSKGDINLVKQGCWSHIGDPQECHYEECVVTTTPPS IQNGTYRFCCCSTDLCNVNFTENFP

Table 5: TGM-1 binding to TβRI and TβRII as assessed by surface plasmon resonance. Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0

Surface	Analyte	Fitted Parameters[a]			
		k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d (nM)	R_{max} (RU)
TβRI	TGM-1 D1	ND[a]	ND[a]	ND[a]	ND[a]
TβRI	TGM-1 D2	$(3.0 \pm 0.1) \times 10^{5}$[b]	$(9.1 \pm 0.1) \times 10^{-2}$[b]	310 ± 10[b]	89.6 ± 0.7[b]
TβRI	TGM-1 D3	ND[b]	ND[b]	ND[b]	ND[a]
TβRI	TGM-1 D12	$(6.7 \pm 0.1) \times 10^{4}$[b]	$(1.6 \pm 0.1) \times 10^{-3}$[b]	24 ± 1[b]	429 ± 1[b]
TβRI[c]	TGM-1 FL	$(5.9 \pm 0.1) \times 10^{4}$[b]	$(7.8 \pm 0.2) \times 10^{-4}$[b]	13 ± 1[b]	193 ± 1[b]
TβRII	TGM-1 D1	ND[b]	ND[b]	ND[b]	ND[b]
TβRII	TGM-1 D2	ND[b]	ND[b]	ND[b]	ND[b]
TβRII	TGM-1 D12	ND[b]	ND[b]	ND[b]	ND[b]
TβRII	TGM-1 D3	$(6 \pm 1) \times 10^{5}$[b]	0.6 ± 0.1[b]	910 ± 20[b]	33.0 ± 0.4[b]
TβRII	TGM-1 FL	$(2 \pm 6) \times 10^{7}$[b]	$(1 \pm 4) \times 10^{-1}$[b]	610 ± 10[b]	215 ± 2[b]

[a]Not determined due to weak signal
[b]Fitted parameters were derived from kinetic analysis of a single injection series
[c]Measured on a lower density chip compared to that used for TβRI:TGM-D2 and TβRI:TGM-D12

Table 6: TGM-1 binding to TβRI and TβRII as assessed by isothermal titration calorimetry

Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0

Cell	TβRI	TβRI	TβRI	TβRI	TβRI	TβRII	TβRII	TβRII	TβRII
Syringe	TGM-1 D1	TGM-1 D2	TGM-1 D3	TGM-1 D12	TGM-1 FL	TGM-1 D1	TGM-1 D2	TGM-1 D3	TGM-1 FL
Cell concentration (μM)	7.5	7.5	7.5	7.5	7.5	15	15	15	15
Syringe concentration (μM)	150	150	135	100	58	300	300	300	320
Temperature (°C)	25	25	25	25	25	35	35	35	35
K_D (nM)	ND[a]	1500 (500 – 4600)[bc]	ND[a]	25 (11, 48)[bc]	52 (29 – 90)[bc]	ND[a]	ND[a]	1200 (900, 1500)[bd]	550 (260, 1080)[be]
ΔH (kcal mol^{-1})	ND[a]	-18 (-27– -13)[b]	ND[a]	-19 (-20 – -18)[b]	-17 (-18 – -15)[b]	ND[a]	ND[a]	-11 (-11, -10)[b]	-7.1 (-7.9, -6.6)[b]
ΔG (kcal mol^{-1})	ND[a]	-8.0	ND[a]	-11	-9.9	ND[a]	ND[a]	-8.4	-8.8
-TΔS (kcal mol^{-1})	ND[a]	9.7	ND[a]	8.3	6.8	ND[a]	ND[a]	2.4	-1.7
Stoichiometry (n)	ND[a]	0.54[f]	ND[a]	1.2[f]	0.96[f]	ND[a]	ND[a]	1.1[f]	0.84[f]

[a]Not determined due to weak signal
[b]Uncertainty reported as 68.3% confidence interval
[c]Fit for one replicate
[d]Global fit of three replicates
[e]Global fit of two replicates
[f]Number of sites determined by incompetent fraction value on sedphat; set to '1' for K_D analysis

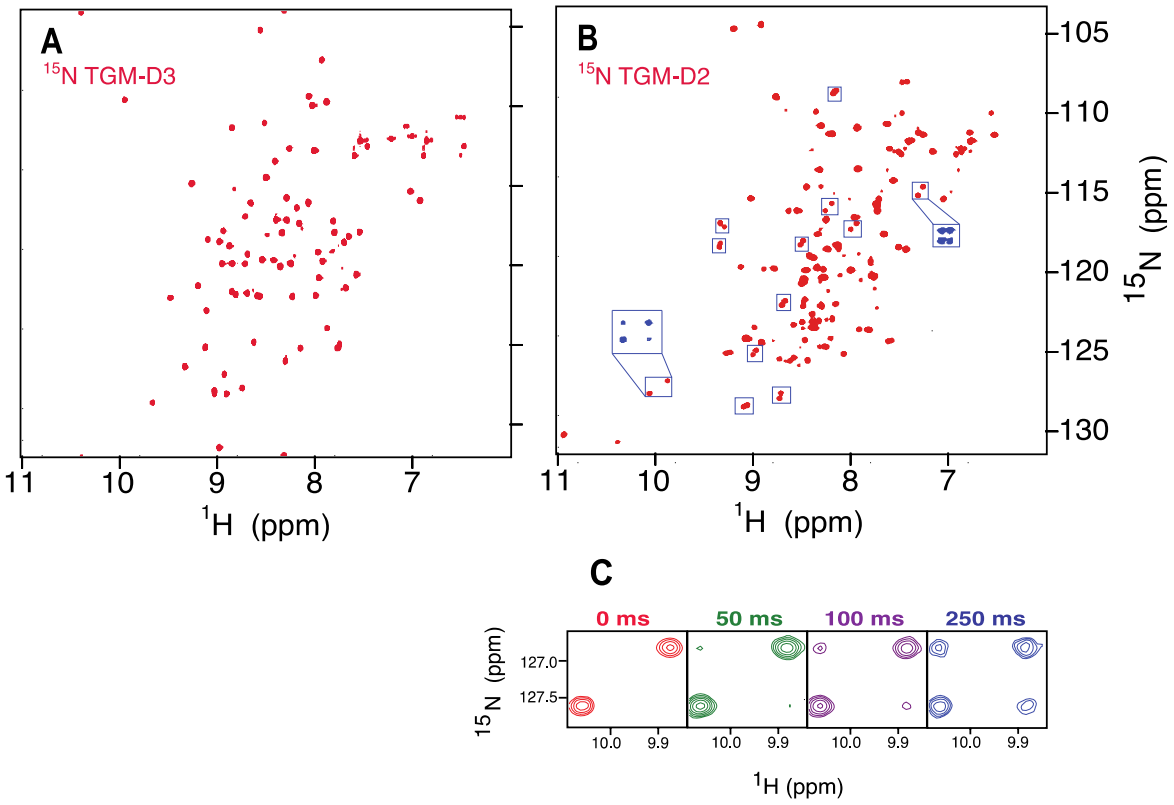

Figure 11: Recombinantly produced TGM-1 D2 and TGM-1 D3 fold natively under NMR sample conditions. A-C. ^{1}H-^{15}N HSQC spectrum of ^{15}N TGM-1 D3 (A) and ^{15}N TGM-1 D2 (B). Blue boxes mark peaks in conformational exchange with each other as designated by ZZ-exchange. An expansion of this ZZ-exchange experiment as a function of mixing time is shown (C) with one representative pair of peaks. Spectra recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

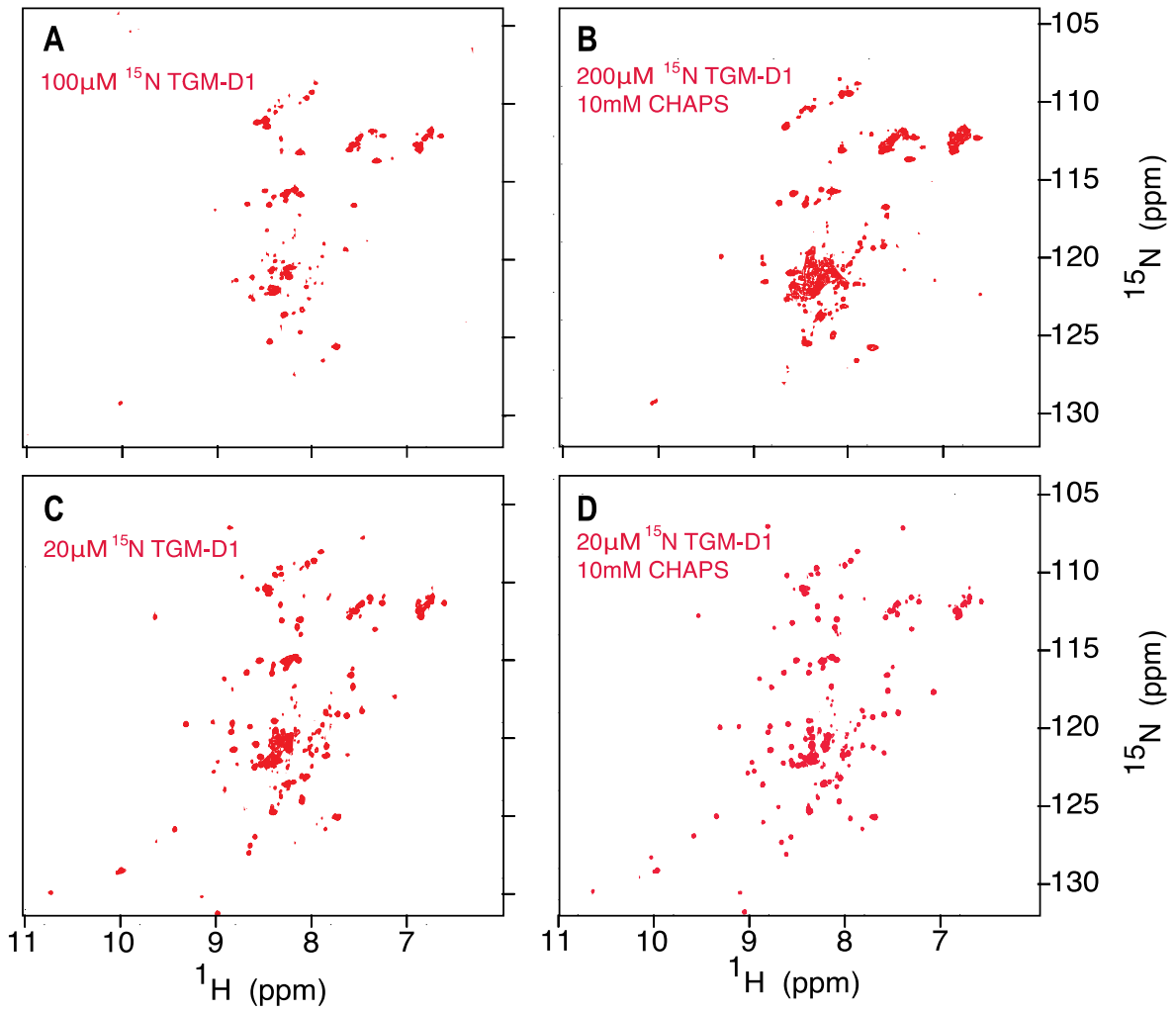

Figure 12: Recombinantly produced TGM-1 D1 aggregates under NMR sample conditions. A-D. ^{1}H-^{15}N spectra of ^{15}N TGM-1 D1 (A) in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K, with (B, D) and without (A, C) the addition of 10mM CHAPS. ^{15}N TGM-1 D1 ^{1}H-^{15}N HSQC spectra gathered over concentrations of 20 µM (C, D), 100 µM (A), and 200 µM (B). Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

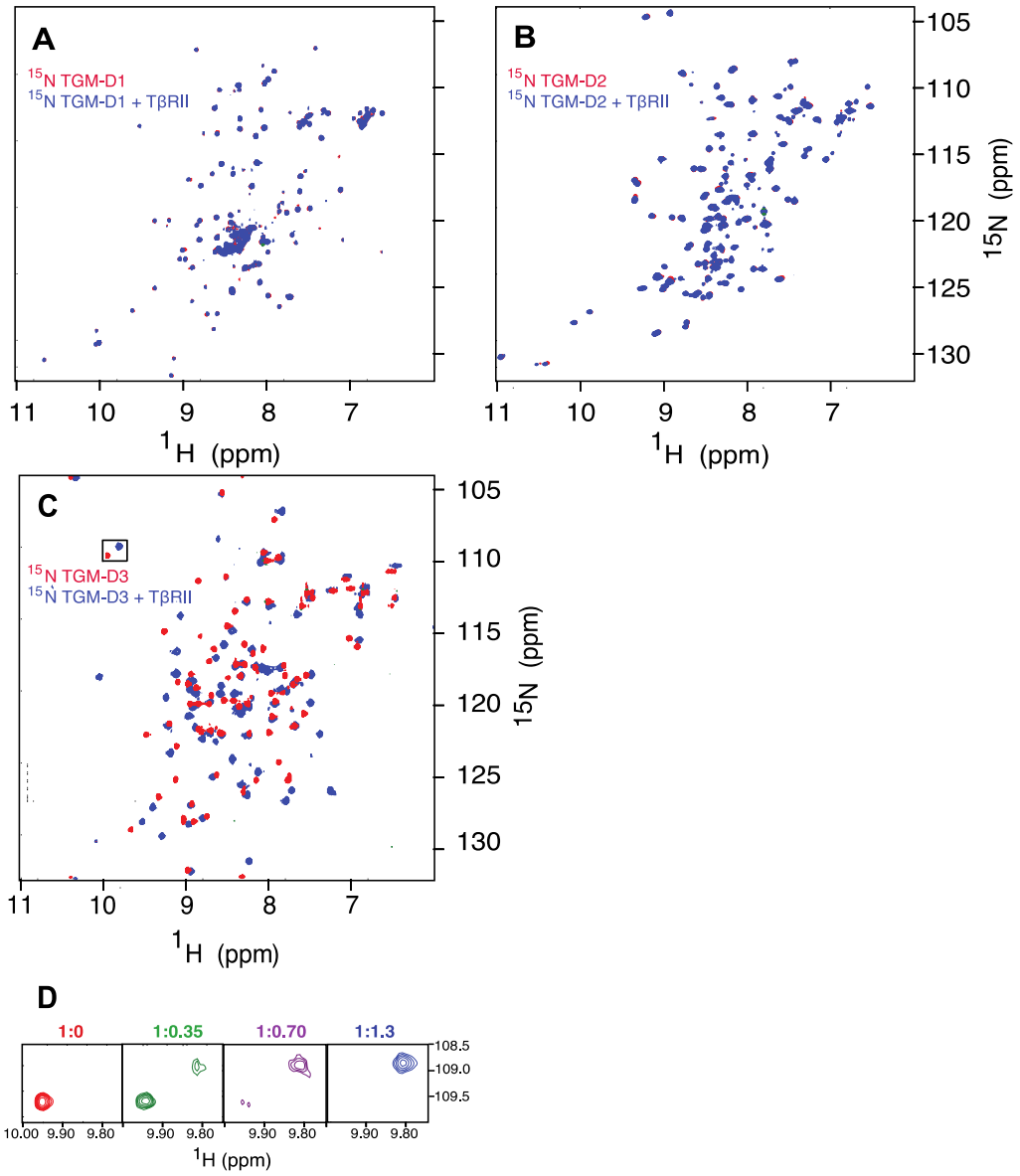

Figure 13: TβRII binds to TGM-1 D3 under NMR sample conditions. A-D. ^{1}H-^{15}N spectra of ^{15}N TGM-1 D1 (A), ^{15}N TGM-1 D2 (B), and ^{15}N TGM-1 D3 (C,D) free (red) and overlaid with the ^{1}H-^{15}N spectra of the same proteins bound to 1.2 molar equivalents of unlabeled TβRII (blue). Expansion of the boxed region in panel C at 0.35, 0.7, and 1.3 molar equivalents is shown below in part D. All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

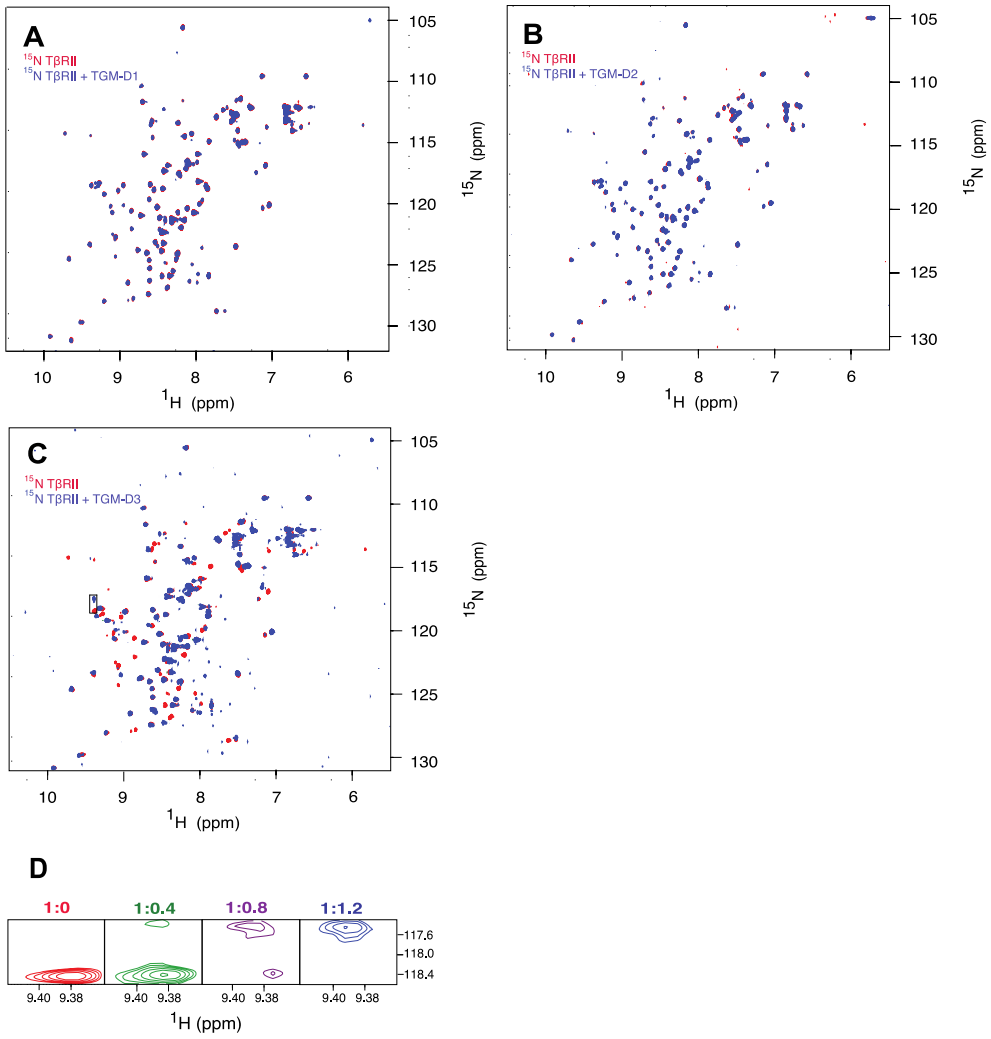

Figure 14: TGM-1 D3 binds to TβRII under NMR sample conditions. A-D. ^{1}H-^{15}N spectra of ^{15}N TβRII (red) overlaid with the ^{1}H-^{15}N spectra of the same proteins bound to 1.2 molar equivalents of unlabeled TGM-1 D1 (A), TGM-1 D2 (B), and TGM-1 D3 (C,D) (blue). Expansion of the boxed region in panel C at 0.4, 0.8, and 1.2 molar equivalents is shown below in part D. All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

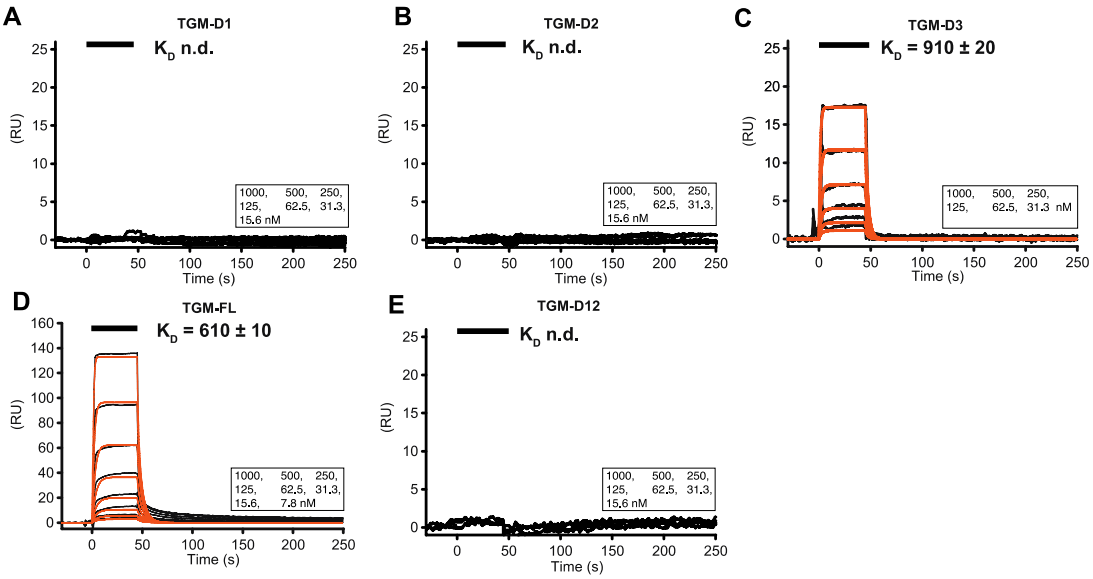

Figure 15: TGM-1 D3 is the sole domain responsible for binding TβRII as shown by SPR. A-E. SPR sensorgrams upon injection of TGM-1 D1 (A), TGM-1 D2 (B), TGM-1 D3 (C), TGM-1 D12 (D), and TGM-1 FL (E) over biotinylated avi-tagged TβRII immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom rihgt of each sensorgram. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

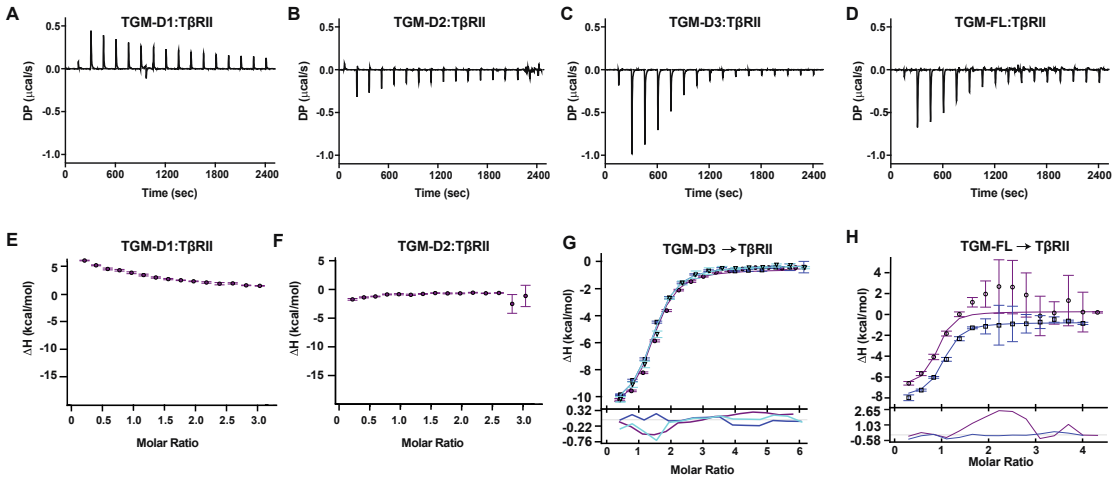

Figure 16: TGM-1 D3 is the sole domain responsible for binding TβRII as shown by ITC. A-H. Raw thermograms for the injection of TGM-1 D1 (A) TGM-1 D2 (B) TGM-1 D3 (C) or TGM-1 FL (D) into TβRII. Corresponding integrated heats for the injection of TGM-1 D1 (E), TGM-1 D2 (F), TGM-1 D3 (G), TGM-1 FL (H), with the TGM-1 D3 and TGM-1 FL data fit to a 1:1 binding model. Fits correspond to the global fit over the data sets with residuals below. Error bars correspond to undertainty in the estimation of integrated heats by NITPIC. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

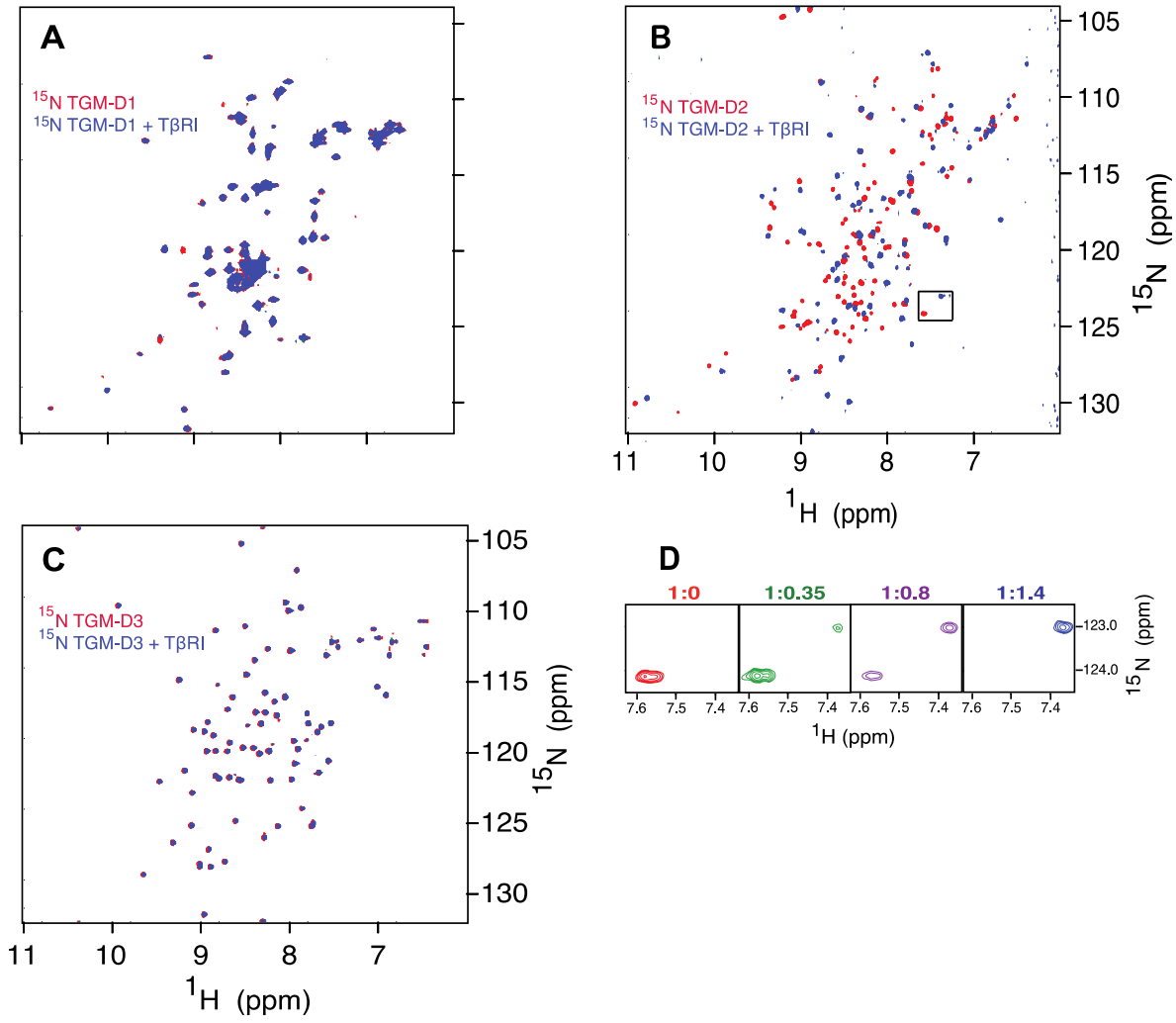

Figure 17: TβRI binds to TGM-1 D2 under NMR sample conditions. A-D. ^{1}H-^{15}N spectra of ^{15}N TGM-1 D1 (A), ^{15}N TGM-1 D2 (B), and ^{15}N TGM-1 D3(C,D) free (red) and overlaid with the ^{1}H-^{15}N spectra of the same proteins bound to an excess of unlabeled TβRI (blue). Expansion of the boxed region in panel B at 0.35, 0.7, and 1.3 molar equivalents is shown below in part D. All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

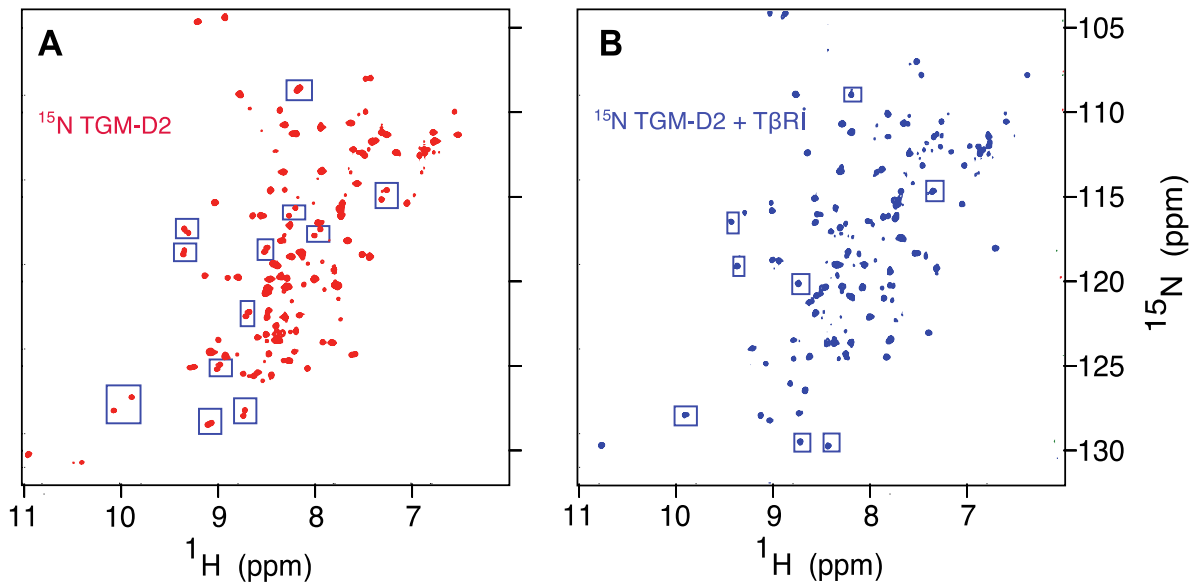

Figure 18: Conformational doubling of TGM-1 D2 is resolved upon binding to TβRI. A-B. ^{1}H-^{15}N HSQC spectra of TGM-D2 alone (A, red) or with 1.2 molar equivalents of unlabeled TβRI added (B, blue). Peaks in conformational exchange marked on the spectra of ^{15}N TGM-D2 alone (A, blue boxes) and the resolved conformational exchange marked on the spectra of ^{15}N TGM-D2 bound to TβRI (B, blue boxes). All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

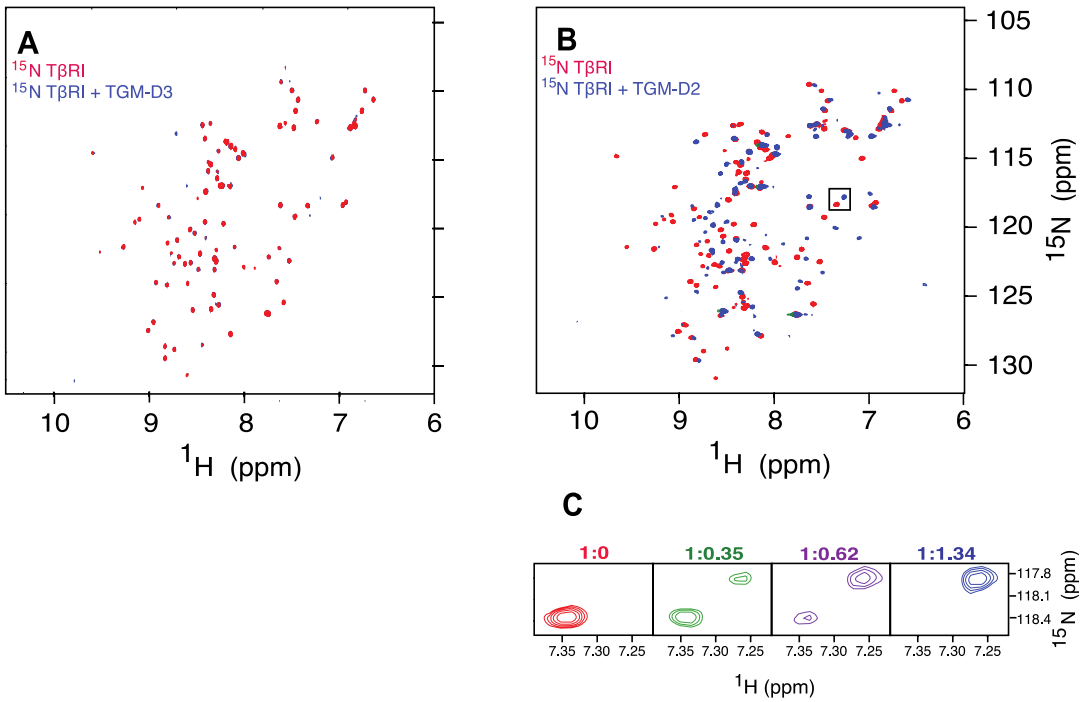

Figure 19: TGM-1 D2 binds to TβRI under NMR sample conditions. A-C ^{1}H-^{15}N spectra of ^{15}N TβRI (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.2 molar equivalents of unlabeled TGM-1 D3 (A), and 1.34 equivalents of TGM-1 D2 (B). Expansion of the boxed region in panel B at 0.35, 0.62, and 1.34 molar equivalents is shown below in part C. All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

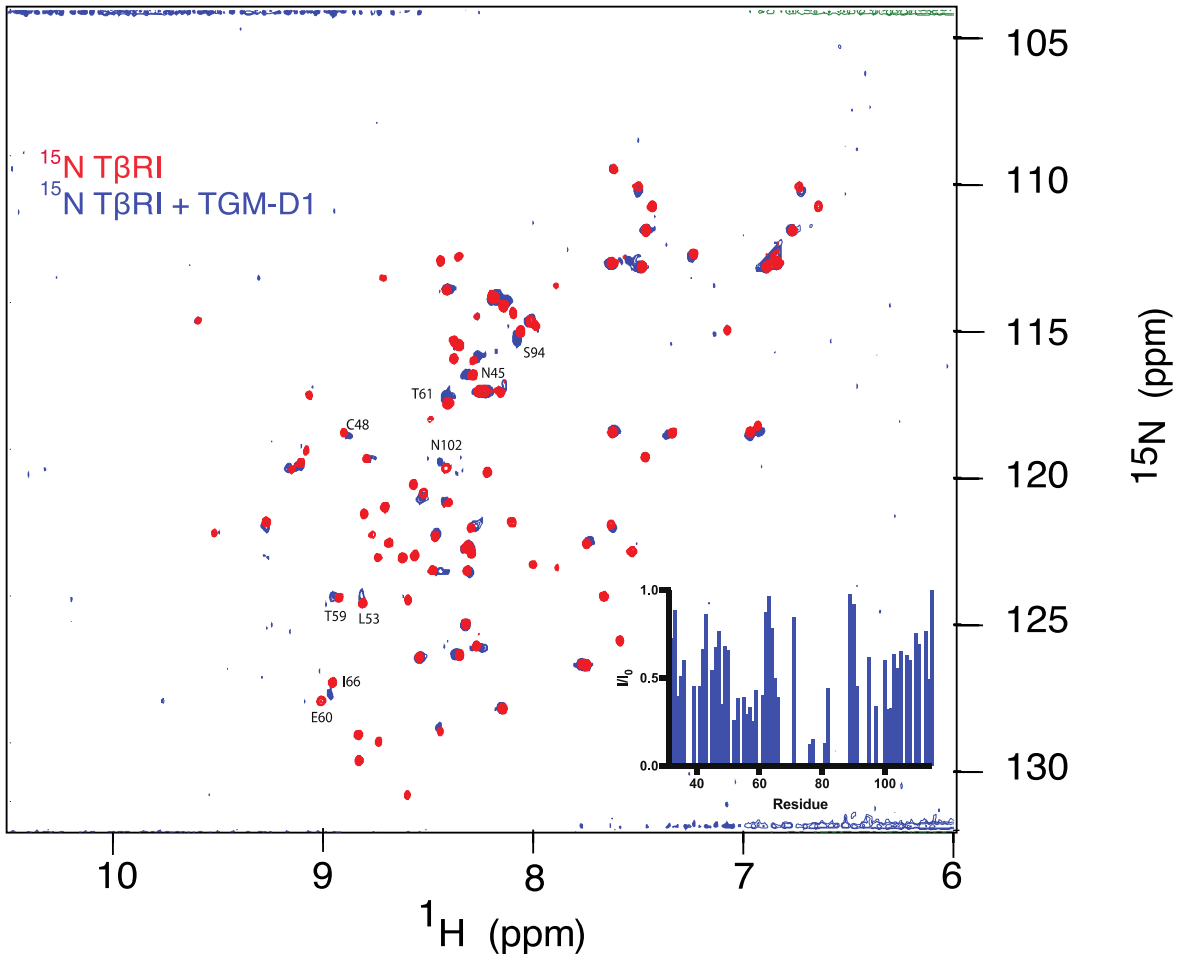

Figure 20: TGM-1 D1 binds to TβRI under NMR sample conditions. ^{1}H-^{15}N spectra of ^{15}N TβRI (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.5 molar equivalents of unlabeled TGM-1 D1. The box in the lower right section of the spectrum corresponds to the per residue ratio of the free TβRI peak intensity compared to the TβRI peak intensity upon binding of TGM-1 D1 The spectrum was recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

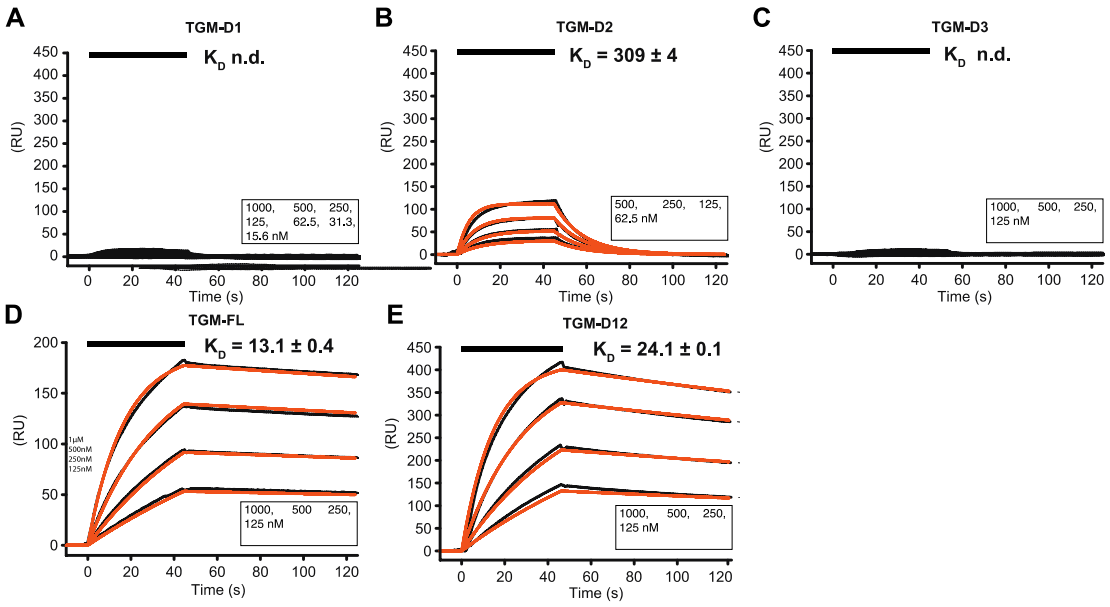

Figure 21: TGM-1 D1 and TGM-1 D2 are jointly responsible for binding TβRI as shown by SPR. A-E. SPR sensorgrams upon injection of TGM-1 D1 (A), TGM-1 D2 (B), TGM-1 D3 (C), TGM-1 D12 (D), and TGM-1 FL (E) over biotinylated avi-tagged TβRI immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom right of each sensorgram. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

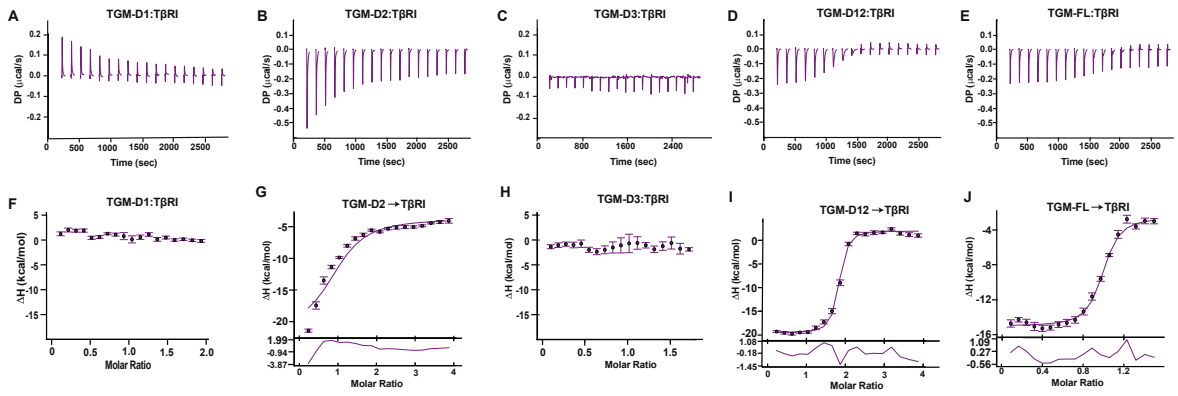

Figure 22: TGM-1 D1 and TGM-1 D2 are jointly responsible for binding TβRI as shown by ITC. A-J. Raw thermograms for the injection of TGM-1 D1 (A) TGM-1 D2 (B) TGM-1 D3 (C) or TGM-1 FL (D) into TβRII. Corresponding integrated heats for the injection of TGM-1 D1 (E), TGM-1 D2 (F), TGM-1 D3 (G), TGM-1 FL (H), with the TGM-1 D2, TGM-1 D12, and TGM-1 FL data fit to a 1:1 binding model. Fits correspond to the global fit over the data sets with residuals below. Error bars correspond to undertainty in the estimation of integrated heats by NITPIC. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

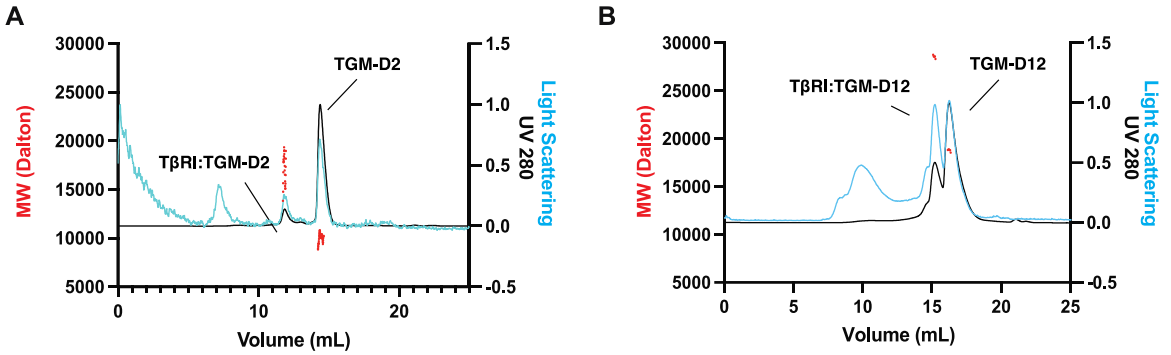

Figure 23: TGM-1 D2 and TGM-1 D12 bind to TβRI with a 1:1 stoichiometry. A-B. SEC chromatogram of TGM-1 D2 (A) and TGM-1 D12 (B) partially titrated with TβRI. Left-most axis indicates molecular weight which in both figures is represented by the red dots at the position of the free TGM-1 D2 (A) / TGM-1 D12 (B) and the bound TGM-1 D2:TβRI (A) / TGM-1 D12:TβRI (B), and right-most axis indicates UV 280 (black) and light scattering (blue).

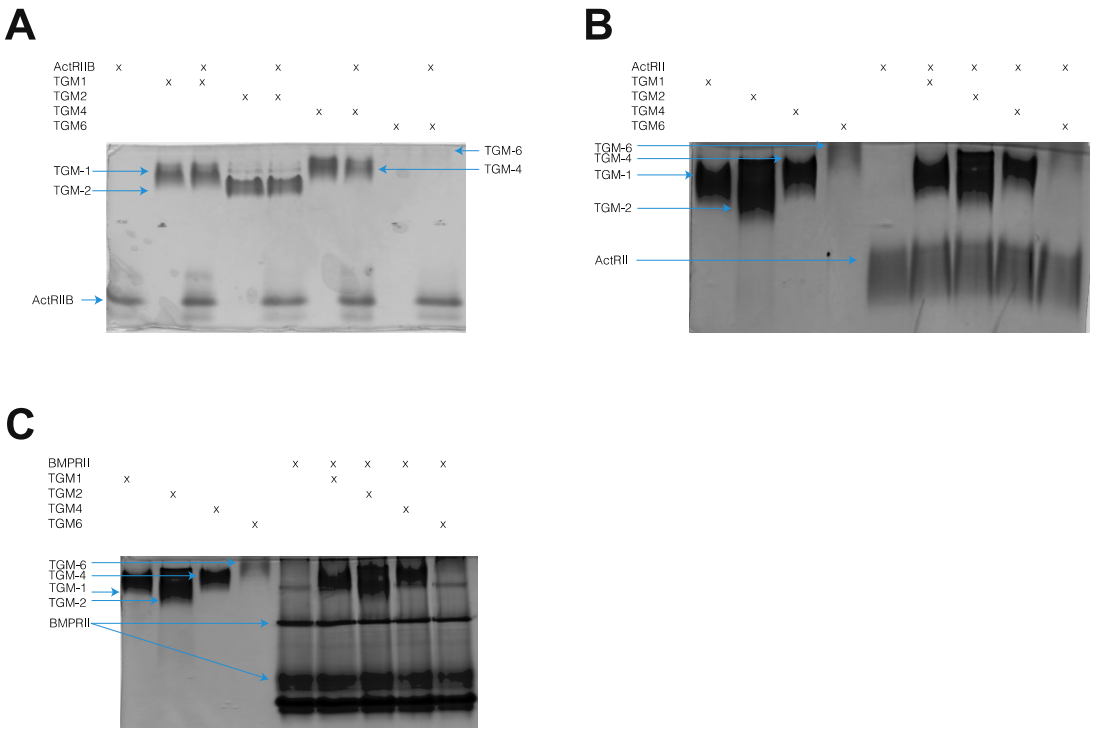

Figure 24: TGM-1 FL, TGM-2 FL, TGM-4 FL, and TGM-6 FL do not bind to the other type II receptors of the TGF-β family as assessed by native gel. A-C. Native gels assessing binding of the extracellular domains ActRIIB (A), ActRII (B), and BMPRII (C) added in large molar excess to members of the TGM family of proteins: TGM-1, TGM-2, TGM-4, and TGM-6 FL. Arrows indicate positions of protein bands on the native gel.

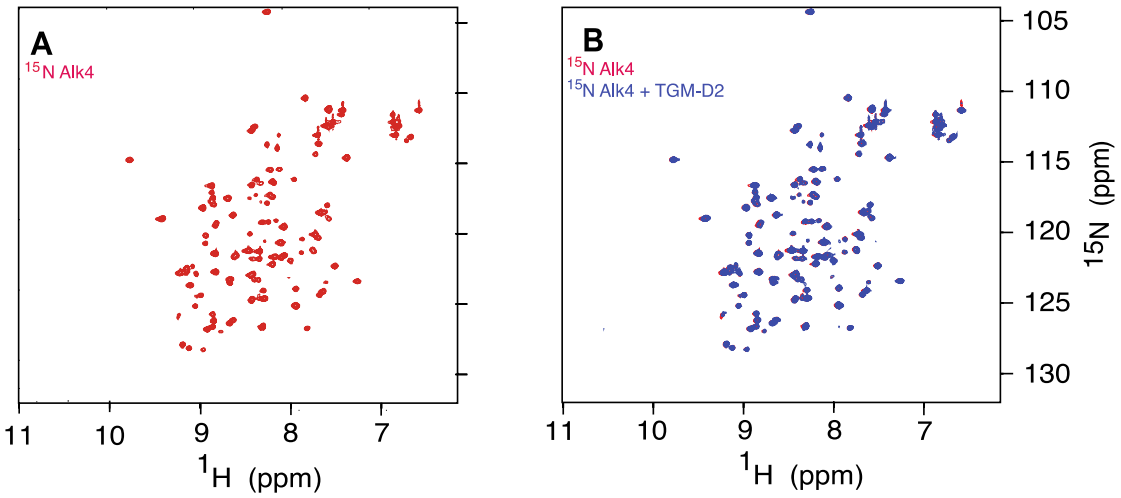

Figure 25: TGM-1 D2 does not bind to Alk4 under NMR sample conditions. ^{1}H-^{15}N spectra of ^{15}N Alk4 (A, red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.2 molar equivalents of unlabeled TGM-1 D2 (B, blue). The spectrum was recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 7.0, 298K.

4.0 TGM-1 D3 Binding to TβRII: Structural and Functional Properties

4.1 Introduction

The previous chapter answered several questions about the interactions between TGM-1, TβRI, and TβRII. It was made apparent that the binding of TGM-1 to the TGF-β receptors is modular and separable, with domains 1 and 2 responsible for binding the type I receptor through a shared binding interface, not allostery, while domain 3 is responsible for binding the type II receptor. TGM-1 acts as a functional mimic of TGF-β, and though its sequence suggests that it is structurally distinct from TGF-β, it is unknown as to whether TGM-1 engages the same receptor binding sites as TGF-β.

There are two possibilities, one is that TGM-1 D3 uses the same interface as TGF-β to bind to TβRII, thus alone would be predicted to be a competitive inhibitor of TGF-β. The other possibility is that TGM-1 D3 uses a different interface from TGF-β to bind to TβRII, which may or may not interfere with TGF-β binding to TβRII. The structure of TGM-1 D3 is also unknown and deciphering this would show how the parasite may have adapted the canonical CCP domain to generate a domain that binds TβRII, as well as give insight into the structures of the other TGM-1 domains, due to their high sequence similarity.

Thus, this chapter aims to investigate whether TGM-1 D3 and TGF-β compete for binding to TβRII, and how the structure of TGM-1 D3 has adapted from the canonical CCP domain to enable binding of TGM-1 D3 to TβRII.

4.2 Methods

Cloning, expression, and purification of mmTGFβ7M. DNA inserts coding for the domains of mmTGFβ27M were inserted into a modified pET32a vector between the NdeI and HindII) sites (Table 7). Constructs were expressed in BL21 (DE3) cells, and cultured at 37 °C. When the light scattering at 600 nm was detected to be 0.8 protein expression was induced by adding 0.8 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG),and were produced in LB. The purification protocol proceeded as in previous publications[258].

Cloning, expression, and purification of ^{15}N ^{13}C TβRII. Constructs, expression protocols, and protein purification protocols were the same as stated in Chapter 3 with one exception. In the M9 medium used to grow ^{15}N ^{13}C TβRII, rather than utilize 1g L^{-1} of $^{15}NH_4Cl$ along with 15 mL L^{-1} 20% w/v D-glucose, both 1g L^{-1} of $^{15}NH_4Cl$ and 15 mL 20% U-^{13}C-D-glucose was used. Otherwise, there was no difference in protein purification.

Cloning, expression, and purification of TGM-1 D3 and TβRII mutants. Mutants of TGM-1 D3 and TβRII were produced using the same constructs mentioned in the methods section of Chapter 3, with modifications made to produce the desired residue substitutions. Site-directed mutagenesis was performed using Phusion polymerase, and the clones sequenced to confirm presence of the mutation.

ITC competition experiments. All ITC experiments were performed with a Microcal PEAQ-ITC system (Malvern Instruments). Cell and syringe conditions are indicated in Table 7. All experiments were performed in 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 at 35 °C, with 13 3.0 μL injections with a duration of 5 s, a spacing of 150 s, and a reference power of 10. All samples

were dialyzed against the same ITC buffer before loading into the system. Data was globally fit using the programs NITPIC[246], SEDPHAT[247,248], and GUSSI[249].

NMR backbone assignment. For assignments of the TβRII backbone residues as bound to TGM-1 D3, samples of ^{13}C ^{15}N TβRII:TGM-1 D3 were produced by adding each individual component in a 1:1.2 molar ratio, with the labeled component at a concentration of ca. 200 μM, in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. Assignments were made by recording, processing, and analyzing the following triple-resonance spectra: HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA, HNCO, and HN(CA)CO, along with the 2D ^{1}H-^{15}N HSQC spectrum, all of which were recorded at 310K. Assignment was performed semi-automatically using PINE[259,260] and then manually corrected using NMRFAM-SPARKY[254] and PECAN[261].

NMR chemical shift perturbation calculations. The backbone residues of TβRII were assigned in the unbound and TGM-1 D3 bound forms. The magnitude of the difference in chemical shift value was calculated for the five nuclei: $^{15}N^{H}$, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C^{O}$, and normalized to the largest chemical shift perturbation value. These normalized values were summed up for all of the nuclei per residue and then normalized once more by dividing the number of nuclei contributing to the total chemical shift perturbation. The same methodology was performed for the reverse set of proteins, with the backbone residues of TGM-1 D3 assigned in the unbound and in the TβRII-bound forms and the chemical shift perturbations normalized and calculated.

NMR solvent accessible surface area calculations. Solvent accessible surface area was calculated for the structure of the extracellular domain of TβRII using PDB 1KTZ using the program Chimera[262]. The solvent accessible surface area was calculated for TβRII in the presence of TGF-β and in the absence, and the difference plotted per-residue of TβRII.

NMR structure determination of TGM-1 D3. The structure of TGM-1 D3 was determined using the program XPLOR-NIH[263], which allows for input of NMR restraints – NOE distance, RDC restraints, TALOS[264] phi and psi and psi angle restraints, and J-couplings. to determine an ensemble of energetically-favored structures (Appendix A.1.1); the final determined structure was submitted to the PDB under the accession code 7SXB, complete with validation statistics. For the full sidechain assignment of TGM-1 D3, further NMR spectra were gathered including: 2D ^{1}H-^{13}C CT-HSQC, 3D HNHB, HCCH-TOCSY, HBHA(CO)NH, 2D CB(CGCDCD)HD, ^{1}H-^{13}C HSQC-TOCSY, CB(CGCDCD)HE, 3D H(CCCO)NH, and CC(CONH). NOE distance restraints were gathered using data collected from 3D ^{1}H-^{15}N NOESY spectra and ^{1}H-^{13}C aliphatic and aromatic NOESY spectra, and the distances derived using the program CCP-NMR[265]. Three sets of residual dipolar coupling (RDC) restraints were gathered: ^{1}H-^{15}N, $^{13}C^{\alpha}$-$^{1}H^{\alpha}$, and $^{13}C^{O}$-^{15}N RDCs. Experiments were run on ^{15}N ^{13}C-labeled TGM-1 D3 in the usual buffer: 25 mM Na_2HPO_4 50 mM NaCl pH 6.0 to determine the unaligned dipolar couplings, and also in 10 mg mL^{-1} Pf1 phage[266] to determine the aligned dipolar couplings, with the difference taken as the RDCs. These RDCs were measured using a 2D IPAP-HSQC experiment[267], a 3D $^{1}H^{\alpha}$-coupled HN(CO)CA experiment, and a 3D $^{13}C^{\alpha}$-coupled HNCO experiment, for the ^{1}H-^{15}N, $^{13}C^{\alpha}$-$^{1}H^{\alpha}$, and $^{13}C^{O}$-^{15}N RDCs respectively. The J-coupling measurements were made by assessing the ratio of the cross-peak to the diagonal in a 3D HNHA experiment, as was described previously[268]. The program PROCHECK[269,270] was used to perform the Ramachandran analysis and RMSD analysis were performed using Chimera.

NMR ^{15}N T_2 relaxation time calculations. The ^{15}N T_2 backbone relaxation times were measured in 25 MES, 50 mM NaCl, pH 6.0, 310K using an interleaved HSQC pulse sequence[271].

Data was collected using 8 delay times from 16 – 240 msec, and the per-residue relative peak intensity was fit using a two-parameter decaying exponential with respect to the delay time.

SPR of mutated TGM-1 D3 and TβRII constructs. Experiments with variant TGM-1 D3 or TβRII constructs were analyzed by SPR in a similar manner as discussed in Chapter 3, with the ligands being either biotinylated avi-tagged TβRII or biotinylated avi-tagged TGM-1 D3 respectively, captured at a density of 50-150 RU onto a neutravidin-coated sensor chip. All SPR experiments were performed in a buffer of 25 mM HEPES, 150 mM NaCl, 0.05% P20 surfactant, pH 7.4, with an injection speed of 100 μL min^{-1}. Regeneration was performed using a 30 second injection 100 mM – 200 mM guanidine hydrochloride. Baseline correction was performed using buffer injections and using a reference cell without the biotinylated protein captured. All mutated constructs were run in duplicate or triplicate and analyzed using Scrubber.

ITC of mutated TGM-1 D3 and TβRII constructs. Experiments with variant TGM-1 D3 constructs were analyzed by ITC in a similar manner as discussed in Chapter 4. Wild-type TβRII was injected via syringe into cells containing variant TGM-1 D3 (concentrations listed in Table 11). All experiments were performed in 25 mM Na_2HPO_4 50 mM NaCl pH 6.0 at 35 °C with 19 2.0 μL injections with a duration of 4 s, a spacing of 150 s, and a reference power of 10. All samples were dialyzed against the same ITC buffer before loading into the system. Data was globally fit using the programs NITPIC[246], SEDPHAT[247,248], and GUSSI[249].

4.3 Results

4.3.1 TGM-1 D3 and TGF-β compete for binding with TβRII

Competition between TGM-1 D3 and TGF-β for binding to the ectodomain of TβRII was assessed using isothermal titration calorimetry (ITC). To perform the competitive inhibition experiment, both a weak and strong-affinity binder for TβRII, with differing enthalpies, were needed. TGM-1 D3 binds to TβRII with an affinity of 1.2 μM (Table 6) as assessed by ITC, whereas TGF-β3 binds to TβRII with an affinity of close to 50 nM. As such, TGM-1 D3 and TGF-β can take the role of strong and weak-affinity binders to TβRII respectively. An initial experiment was performed without any TGM-1 D3, with TβRII in the cell and an engineered TGF-β monomer in the syringe, with subsequent competition experiments including TβRII and varying concentrations of TGM-1 D3 in the cell, and the engineered TGF-β monomer in the syringe.

There is one caveat to this design, which is that TGF-β isoforms are known to be highly insoluble, unless at either extremes of the pH scale, at very acidic (4.0) or basic (11.0) solution conditions[258]. This makes ITC experiments difficult, as they rely on all experimental components being soluble under the same solution condition. Thus rather than run experiments in which TGF-β is titrated into TβRII, and TGM-1 D3 added as a competitive inhibitor, an engineered TGM-1 monomer, mmTGFβ27M was used instead[258]. This engineered monomer lacks the heel helix of TGF-β and is much more soluble at neutral pH. SPR analysis showed that mmTGFβ27M was able to bind TβRII with an affinity identical to that of TGF-β3[258].

mmTGFβ27M was injected from the syringe at a concentration of 150 μM into a cell containing TβRII (15 μM) and either 0, 6, or 12 μM TGM-1 D3. Addition of mmTGFβ27M into

the cell containing TβRII resulted in a large negative enthalpy and a strong binding affinity. Addition of 6 μM and 12 μM TGM-1 D3 progressively decreased the slope of the transition between the free and bound forms of TβRII (Fig. 26), as well as decreased the magnitude of ΔH, all of which is indicative of competitive inhibition. The data from the three isotherms, 0 μM, 6 μM, and 12 μM TGM-1 D3 were globally fit to a competitive binding model, with the high-affinity mmTGFβ27M:TβRII interaction K_D determined to be 35 nM, with a ΔH of -7.4 kcal mol^{-1}, which is in line with previous SPR measurements of this affinity[258] (Table 8). Overall, the data clearly shows that TGM-1 D3 and TGF-β compete with binding for TβRII.

4.3.2 TGM-1 D3 and TGF-β bind to similar residues on TβRII

TGM-1 D3 and TGF-β compete for binding with TβRII, but whether their binding sites partially or completely overlap is in question. To assess the residues of TβRII that TGM-1 D3 binds to, NMR chemical shift perturbation experiments (CSP) were used. CSP experiments are based off of the premise that the chemical shift perturbations between the unbound and bound forms of a labeled protein can be used to identify residue-specific binding sites[272]. The chemical shifts for the backbone of the protein are assigned in the unbound form, specifically the amide nitrogen and hydrogen shifts, carbonyl shifts, and alpha and beta carbon shifts. Subsequently, the ^{15}N ^{13}C labeled protein is titrated with the unlabeled binder, and the chemical shifts for the aforementioned atoms re-assigned. The difference in the chemical shift value per atom is calculated and normalized per residue. The higher the value, the more the residue of the labeled protein is perturbed, whether by direct binding or indirect binding effects, by the binding of the unlabeled binder. TβRII was ^{15}N and ^{13}C-labeled and bound to 1.2 equivalents of unlabeled TGM-

1 D3. Backbone chemical shift values for the unbound ^{15}N ^{13}C TβRII had already been assigned, and is deposited under BMRB code 4779[273] (Fig. 27A). The chemical shift values for backbone residues were assigned to the TGM-1 D3 bound TβRII (Fig. 27B), and the difference in chemical shifts between the unbound and bound forms per atom were calculated (Fig. 28).

The plot of composite shift per residue for ^{15}N ^{13}C TβRII is striking in its lack of significant composite shift perturbation detected (Fig. 28A). On a scale normalized to 1.0, only five residues have a normalized composite shift perturbation higher than 0.2, and of those five, three of them are localized to a set of three residues, T74, S75, and I76. This signifies that there is likely one major region of TβRII involved in binding TGM-1 D3, the region between residues 73-77, which corresponds to β4 of TβRII. As there are few other regions perturbed by binding to TGM-1 D3, it is likely that this spike is indicative of direct binding and not indirect binding effects. The other regions of TβRII perturbed, to a normalized value of in between 0.1 and 0.2 likely represent indirect effects from binding and not a direct binding site, though there may be interactions with the sidechains of TβRII that this method does not detect.

To compare the binding interface between TGM-1 D3:TβRII and TGF-β:TβRII, the per residue composite shift perturbations from the binding of TGM-1 D3 to TβRII were compared to the difference in per residue solvent accessible surface area (ΔSAS) of residues of TβRII in the unbound and TGF-β bound forms, which was used to demonstrate the residues key in the TGF-β:TβRII binding interface (Fig. 28B). The plot of per residue ΔSAS shows that there are three key regions of TβRII whose surface is buried upon binding to TGF-β: 1) the first β-strand and the first turn following said β-strand (residues 50-55), 2) β4 (residues 73-77), and 3), turn 8 (residues 141-142). The first region includes residue D55 of TβRII, which forms a hydrogen-bonded ion pair with R394 of TGF-β, and the third region includes residue E141 of TβRII, which forms a similar

hydrogen-bonded ion pair with R325 of TGF-β. The composite shift perturbation plot of TGM-1 D3 bound TβRII lacks these specific residue contacts, though it includes the main contact of the edge β-strand. This indicates that while there are some shared residues between the TGM-1 D3:TβRII and the TGF-β:TβRII binding interfaces, TGF-β likely contacts more residues of TβRII throughout the entire protein, as opposed to the edge β-strand (β4) alone. The hypothesis that the TGM-1 D3:TβRII interface is not as extensive as the TGF-β:TβRII interface would explain the reduced affinity of TGM-1 D3 to TβRII (ca. 1 μM) as opposed to the higher affinity of TGF-β (ca. 50 nM). However, it could also be that there are interactions between TGM-1 D3 and the sidechain atoms of TβRII that are not readily detected by CSP analysis.

4.3.3 The structure of TGM-1 D3 shows distinct adaptations from the canonical CCP domain

4.3.3.1 Structure of the canonical CCP domain

CCP domains are present in eukaryotes, viruses, and bacteria alike, but are most well-known for their presence in the regulators of complement activation (RCA) family of proteins. Examples of solved CCP structures in human and non-human animals include CCP modules in the rat GAG(B)R1a receptor (PDB 1SRZ), the CCP modules of the vaccinia virus protein (PCB 1VVC), and baboon and chimp isoforms of complement receptor 1. RCA proteins are known for their repeating CCP domains, such as C4-binding protein, decay-accelerating factor, and factor H[274]. The CCP sequence is characterized by a domain length of ca. 60 amino acids, with four conserved cysteines forming two disulfide bonds in a Cys^{I}-Cys^{III}, Cys^{II}-Cys^{IV} topology, a highly conserved tryptophan at the C-terminus of the domain, and conserved prolines, glycines, and hydrophobic regions (Fig. 29). The 3D structure is characterized by a highly β-strand rich structure,

with a laterally-extended loop between β2 and β3 ranging from 2-10 residues called the hypervariable loop[274,275]. This hypervariable loop is present amongst all CCP domains, and as the name would suggest, varies in sequence content and length amongst even CCP domains within the same protein. In addition, CCP domains are known for their malleability, with wide sequence and structural variations that leave the domain the perfect scaffold for adaptation[275].

Two examples of CCP domain-containing proteins are CD46 (decay activating factor, DAF) and CD55 (membrane co-factor protein, MCP); the former is a glycoprotein while the latter is important for the C3/C4 complement cascade[275]. CCP domains are elongated, with the β-strands oriented in the direction of the longitudinal axis. The N- and C-termini are oriented at opposite ends of the longitudinal face, and the core of the protein is compact and hydrophobic, with an invariably conserved tryptophan residue in the hydrophobic core (Fig. 29, 30). The CCP domains usually contain six β-strands: β1- β4, β' and β'' (Fig. 30). The two disulfide bonds encircle the hydrophobic core: Cys^{I} which is present at the N-terminus pairs with Cys^{III} which usually occurs right after or at the C-terminus of β3, Cys^{II} which usually occurs right after or on the C-terminus of β2 pairs with Cys^{IV} which occurs either on β'' or at the C-terminal tail. Some CCP domains include a couple of other β-strands, one which occurs right at the N-terminus, N-terminal of β1, and another than occurs right after β2, for a maximum of 8 β-strands (Fig. 30).

4.3.3.2 TGM-1 D3 NMR structure and dynamics

TGM-1 D3 has a molecular weight of 9800 Da, and is highly soluble, even upwards of concentrations of 5 mM or 50 mg mL^{-1}. It is also purified in *E. coli* after refolding and can be readily isotopically labeled, as has been shown in previous sections. As such, NMR was the appropriate method for structure determination, over crystallography, on account of its high solubility, or Cryo-EM, on account of its small size. Near-complete assignments were made of the

atoms in the backbone, along with the sidechain hydrogen and carbon atoms (deposition to BioMagResBank under accession code 51083). Further restraints included ^{1}H-^{1}H NOE distance restraints, (1139 total) RDC restraints – from ^{1}H-^{15}N, $^{13}C^{\alpha}$-$^{1}H^{\alpha}$, and $^{13}C^{O}$-^{15}N experiments (209 total), TALOS dihedral restraints (120 total), along with $^{3}J^{HN\text{-}H\alpha}$ J-coupling restraints (39 total) (Table 9).

TGM-1 D3 contains 90 residues with the addition of the noncanonical GSGT-N terminal linker. This corresponds to residues 172-262 of TGM-1 (NCBI accession code MG099712). Structure calculations from XPLOR-NIH[263] yielded an ensemble of 10 lowest energy structures (Fig. 31). TGM-1 D3 forms a twisted half β-barrel structure, with a concave face containing both N- and C-termini and a convex face, and the N and C-termini on opposite ends of the long-axis of the structure. The four β-strands include residues Tyr[189]-Arg[221], Thr[217]-Arg[221], Glu[234]-Lys[241], Ser[248]-Tyr[252], forming a twisted β-sheet of antiparallel β-strands. The first β-strand is present in a majority of the 10 lowest-energy structures, but not in all of them. Similarly, there is an α-helix between β2 and β3 from Gln[228]-Ala[230] that is a present in most, but not all of the lowest energy structures. The secondary structure elements account for twenty seven of the 90 residues, while the core of the protein resides from residues Cys[178]-Tyr[253], or 84% of the structure. The core is formed by the two disulfide bonds, one connecting Cys[178]-Cys[239] and the other connecting Cys[223]-Cys[259]. The first disulfide bond connects the N-terminus to the C-terminal half of β3 while the second connects the C-terminus to the loop connecting β2 and β3 (Fig. 31C,D), and splitting the protein into two faces, the concave face containing the N and C-termini and the convex face. The core includes several hydrophobic residues on both faces of the protein, including Leu[181] and Pro[182] from the region of the protein N-terminal to the first β-strand, Ile[186], Val[187], Ile[189] from β1, Ala[219] and Arg[221] from β2, Val[236] and Ala[237] from β3, and Trp[250] and Tyr[252] from β4 (Fig. 31).

The program PECAN was used to corroborate this secondary structure analysis, with chemical shift analysis predicting secondary structures in the following regions: β-strands from residues 177-179, 184-191, 201-206, 216-222, 234-241, 248-252, and 226-228, though the regions from residues 177-179 and 201-206 are predicted to be structured with lower probability than the others, and an α-helix from residues 226-228 (Fig. 32). This secondary structure prediction correlates with the XPLOR-NIH calculated structures, with the exceptions of the two lower-probability β-strand regions of residues 177-179 and 201-206; the former corresponds to the N-terminal residue of TGM-1 D3, excluding the GSGT tag, the first cysteine, and the residue immediately following, while the latter to a set of residues within the hypervariable loop (HVL), between β1 and β2 (Fig. 31). The HVL begins at the C-terminus of β1, where it projects perpendicularly against the top half of the protein, contacting hydrophobic residues on N-terminal half β3 including Phe^{235} and Ile^{238}, and then wrapping around the face of the protein, contacting Tyr^{191}, Tyr^{192}, and Pro^{180}, to attach at its C-terminus to β2, and contacting hydrophobic residues on the C-terminal half of β3, including Tyr^{240} and Lys^{241} (Fig. 31C, D). The HVL converges into two distinct conformations in the final ensemble, one in which the C-terminal segment of the HVL ascends as it contacts β2, and one in which it descends.

The backbone root-mean-square deviation (RMSD) for the ensemble of 10 lowest energy structures was calculated as 0.4 angstroms when aligned to secondary structure regions, or 2.3 angstroms when aligned over the entire core region (Table 9). The HVL has an average backbone pairwise RMSD of 1.74 angstroms, which is lower than the ordered regions, and is likely attributable to this conformational flexibility, though it could also be due to a low restraint density in this region which could make an ordered region appear disordered. There are a couple of

Ramachandran outliers in the structure as well, but these occur in the unstructured N- and C-terminus and the residue in which the divergence of the C-terminal portion of the HVL begins.

To characterize the domain dynamics, the backbone ^{15}N T_2 relaxation times were assessed using a modified interleaved sequence[271]. These relaxation times are sensitive to amino acid backbone fluctuations (ns-ps timescale) and can be used to assess flexibility. The analysis showed increased T_2 relaxation times (ms) in the N-terminal tail and in the regions of the protein connecting secondary structure, specifically the loops connecting the β2 and β3 strands, the β3 and β4 strands, along with the C-terminal residues (Fig. 33). This increase in T_2 is indicative of residue flexibility. The N-terminal tail has drastically larger T_2 relaxation time in comparison to the rest of the structure, even the more flexible loop regions, which makes sense as these regions are well-converged in the final ensemble. The regions of secondary structure show lower T_2 relaxation times in comparison to these loop regions, which is congruent with their lesser flexibility from being structured. The residues in the HVL have relatively low T_2 values as well, with the exception of the C-terminal portion of the HVL, which is the divergent segment in the HVL.

4.3.3.3 The structure of TGM-1 D3 is distinct from the canonical CCP domain

The structure of TGM-1 D3 was compared to canonical CCP domain structures using DALI, which is a server that queries the coordinates of an input PDB file and compares them to the existing protein structures in the PDB[276-278]. All of the structures found comparable to the ensemble of TGM-1 D3 solution structures were determined to be CCP-domain containing proteins, which is expected based off of initial bioinformatic studies during TGM-1's discovery[230]. The top hit on DALI for TGM-1 D3 was human CD46 (PDB 1CKL), and other top hits include human IL-15Rα (PDB 2PSM), human CD55 (PDB 1H2P), human CR1 (PDB 5FO9) and human DAF (PDB 5FOA), proteins that belong to the regulators of complement activation family.

Comparison of TGM-1 D3 to these CCP domains highlighted key similarities and differences from TGM-1 D3 to the canonical CCP structure. TGM-1 D3 and the aforementioned CCP domains all contain β1-β4, along with the expected Cys^{I}-Cys^{III}, Cys^{II}-Cys^{IV} disulfide topology (Fig. 34). However, despite the core being similar, TGM-1 D3 lacks two of the β-strands present in the CCP domains surveyed, the β' and β'' strands. The β' strand is located in the loop connecting β2 and β3, and the β'' strand is located at the C-terminus, just past the β4 strand (Fig. 34 C-F). This is surprising as TGM-1 D3 is 90 residues long, approximately 15-25 residues longer than a canonical CCP domain (65-75 residues).

The β' and β'' strands in the canonical CCP domain form a second antiparallel β-sheet, and along with the Cys^{I} – Cys^{IV} disulfide, pull the C-terminal half of the protein inwards towards the β2-β3 loop. The core of the protein is closed off, with the hydrophobic residues in the core inaccessible. This lies in direct contrast to TGM-1 D3 which due to the lack of β' and β'' strands is laterally expanded. Since TGM-1 D3 lacks β'' at the C-terminus, the C-terminal region is expanded away from the core of the protein. This partially exposes hydrophobic residues on both the NC face (Fig. 34 A,C) and the nonNC face (Fig. 34 B,D) of TGM-1 D3. Partially exposed residues include NC face Ile^{186} and Val^{187}, both of which occur in the ordered segment just prior to β1, and Tyr^{252}, which is at the C-terminal end of β4, and nonNC face residues Tyr^{192} on β1, Phe^{235} on the N-terminal half of β3, and Tyr^{253}, which directly follows β4.

The second key difference between TGM-1 D3 and the canonical CCP domain is in the HVL. In the canonical CCP domain, the HVL between β1 and β2 is approximately 4-10 residues and runs parallel to the face of the protein (Fig. 34 C, F). This is in contrast to TGM-1 D3, which shows an expanded HVL of 23 residues that lies perpendicular to the faces of the protein, packing against β1 and the N-terminal half of β2 (Fig. 3, Fig34 A,B). This potentially blocks hydrophobic

residues on the N-terminus, β1, and β2 from involvement in binding interfaces, such as Tyr^{191} on β1, and Pro^{180} and Pro^{182} on the N-terminal segment.

4.3.4 Mutagenesis of key residues of TGM-1 D3 suggests a role for the adaptations to the CCP domain

Backbone chemical shifts of ^{15}N ^{13}C TGM-1 D3 were assigned with and without the presence of equimolar equivalents of TβRII (Fig. 35). Similar to what was done with the assignments of free and TGM-1 D3 bound TβRII, composite shift perturbations were calculated for bound TGM-1 D3. Several regions of TGM-1 D3 are perturbed upon binding of TβRII; the regions that are maximally perturbed are in the C-terminal half of the protein, specifically correlating to β3, β4, and the C-terminal tail (Fig. 36A), residues 243-243 and residues 249-257. Regions that are less perturbed map to the N-terminal half of β2 and the N-terminal region of the HVL, residues 214-219 and residues 193-200. The direct binding interface likely involves the residues maximally perturbed, residues 252-256.

The residues that are maximally perturbed do not seem to be limited to one of the faces of the protein, including Tyr^{252} and Val^{236} on the NC face and Ile^{238}, Tyr^{253}, and Ile^{256} on the nonNC face (Fig. 36B). From this data, it is unclear which of the faces of the protein, if not both faces, are involved in binding TβRII, and which residues specifically. From the structure and the backbone chemical shift perturbations of TGM-1 D3 upon binding TβRII, a set of TGM-1 D3 residues were mutated to alanine. These mutants were purified similarly to the wild-type protein and were checked for native folding using 1D ^{1}H NMR spectra (Fig. 37) and the mutagenesis confirmed by mass spectrometry. The residues mutated were spread across the NC and nonNC faces, including Val^{236} and Tyr^{252} on the NC face, and Arg^{198}, His^{199}, Phe^{235}, Ile^{238}, and Tyr^{253} on the nonNC face.

Binding of mutated constructs were assessed using SPR. Biotinylated avi-tagged TβRII was captured on a streptavidin-conjugated CM5 SPR chip, and mutated TGM-1 D3 constructs were injected over the surface. The K_D values were derived by fitting the sensorgrams to a 1:1 kinetic model (Table 10). All of the mutants were fit to a 1:1 binding model, except for the Tyr^{253}Ala mutant, which did not have a significant enough response, despite injected concentrations of the ligand running up to 100 μM (Table 10, Fig. 38). Other mutants, including Arg^{198}Ala, Phe^{235}Ala, Val^{236}Ala, Ile^{238}Ala, and Tyr^{252}Ala had attenuated responses with respect to wild-type TGM-1 D3, but were still fit, indicating a weaker binding affinity. The Arg^{198}Ala, Ile^{238}Ala, Tyr^{252}Ala, Tyr^{253}Ala, and Lys^{254}Ala mutants showed a weaker-affinity K_D by 20-fold or more, particularly the Arg^{198}Ala mutant with a 70 μM binding affinity and the Tyr^{253}Ala mutant, whose binding could not be measured. The Phe^{235}Ala and Val^{236}Ala mutants also showed weakening in binding affinity, with a 4-fold difference from wild-type.

The sidechains of Arg^{198}, Ile^{238}, Tyr^{253} and Lys^{254} are all present on the nonNC face of the protein (Fig. 31D), and the sidechain of Tyr^{252} is located adjacent to Tyr^{253} on the NC face of the protein (Fig. 31B). Other mutated residues include His^{199}, Asn^{255}, Ile^{256}, and Lys^{258}, the latter three of which are close to Tyr^{253} and show significant chemical shift perturbations (Fig. 36) but do not impact binding to TβRII (Table 10). This seems to suggest that the binding site of TβRII on TGM-1 D3 is located on the nonNC face of the protein, and the chemical shift perturbations of residues on the NC face, particularly Tyr^{252}, are due to contacts between the sidechains of residues on the NC face and backbone residues directly adjacent, or due to indirect effects of binding through perturbations of other secondary structural elements that are then transmitted to other residues.

To orthogonally method to confirm these results, ITC was used. ^{15}N constructs of the TGM-1 D3 $I^{238}A$, $Y^{252}A$, $Y^{253}A$, and $I^{256}A$ mutants were produced, shown to be native via ^{1}H-^{15}N

HSQC NMR, and ITC was used to confirm binding of these constructs to TβRII (Fig. 39, Table 11). NMR spectra of these mutants show a well-dispersed set of signals, with signals spanning from ^{1}H proton shifts of 6-11 ppm, and few residues clumped in the random coil region (7.8-8.6 ppm) (Fig. 39 A-H). The titration of the $Y^{253}A$ mutant into TβRII did not produce any measurable change in enthalpy (Fig. 39 K), while the titration of the other three mutants into TβRII produced an exothermic response quantifiable via a binding isotherm showing large negative changes in enthalpy. The largest change in enthalpy was present in the $I^{256}A$ titration, followed by the $I^{238}A$ titration, with the $Y^{252}A$ titration showing a barely quantifiable change of enthalpy. Fitting these binding isotherms globally gave binding affinities similar to those obtained by the SPR experiments, with the K_D of binding and enthalpies for the $I^{236}A$ titration being 15 μM and -10.3 kcal/mol, the titration of the $Y^{252}A$ titration being 28 μM and -4.3 kcal/mol and the titration of the $I^{256}A$ titration being 1.4 μM and -8.2 kcal/mol respectively (Table 11). Binding isotherms could not be fit for the experiment titrating the $Y^{253}A$ mutation into TβRII.

Reverse experiments were performed as well, testing to see which residues of TβRII are significant in the TGM-1 D3:TβRII binding interfaces. The residues in the edge β-strand of TβRII were mutated: $Ile^{73}Ala$, $Ser^{76}Leu$, $Ile^{76}Ala$ in β4 and binding assessed via SPR (Fig. 40, Table 12). The mutated residues in β4 all showed drastic changes in binding affinity, with the $S^{75}L$ mutation weakening the binding affinity by over 100-fold. The $I^{76}A$ mutation reduced the binding affinity by 20-fold, and the $I^{73}A$ mutation reduced the binding affinity by 6-fold (Table 12). Similarly, the residues of TβRII involved in hydrogen bonding with the fingertip residues of TGF-β, Asp^{55} and Glu^{142} were mutated to Asn and Gln respectively. Mutating Asp^{55} to Asn weakened the binding affinity by 60-fold and mutating Glu^{142} to Gln weakened the binding affinity by 15-fold (Table 12). Thus, it is likely that TGM-1 D3 forms hydrogen bonding or electrostatic interactions with

these residues in TβRII as well as interacting with β4 of TβRII. These residues may be potentially interacting with TGM-1 D3 Arg^{198} and Lys^{254}, mutations were shown to greatly impact binding to TβRII, though these potential interactions have yet to be confirmed (Table 11).

4.4 Discussion

The ITC experiments in this chapter show clear evidence of competition between TGF-β and TGM-1 D3 for binding to TβRII, which is supported by the NMR CSP experiments, which indicate that the main segment of TβRII perturbed by binding to TGM-1 D3 is the edge β-strand (β4). This region is canonically used to bind TGF-β, with the strand inserted into a hydrophobic pocket between the fingers of TGF-β. These experiments suggest that the parasite has evolved to utilize similar binding residues of TβRII as the human ligand, though interestingly, the contacts seem to be less extensive in the TGM-1 D3:TβRII interface as compared to the TGF-β:TβRII interface, which may account for the weaker binding affinity of TGM-1 for TβRII.

TGM-1 D3 belongs to the CCP family of proteins, whose domains are elongated with the N- and C-termini on opposite sides of the protein core, characterized by two disulfide bonds and a preponderance of twisted antiparallel β-strands. Most CCP domains have anywhere from 65-75 residues, whereas TGM-1 D3 has a total of 85 residues (90 residues when the noncanonical linker is added). These extra residues result in an extension in the hypervariable loop of the protein; this HVL lies perpendicular to the main face of the protein, as opposed to the lateral HVLs in the canonical CCP domains, and is structurally ordered, as shown via the ^{15}N T_2 measurements. The perpendicular orientation of the HVL allows for several residues, particularly in the N-terminal

segment of the HVL, such as Tyr[194], Arg[198], and His[199], to pack against hydrophobic residues on the convex, or nonNC face of the protein, such as His[218], Phe[235], and Ile[238].

The other distinct difference between TGM-1 D3 and other CCP proteins is in the lateral expansion of the C-terminal half of TGM-1 D3. Canonical CCP domains contain a minimum of 6 β-strands in two sheets, one including β1-β4 and comprising the main face of the protein, and the other including β' and β'', which cluster in a separate sheet at the C-terminal end of the protein. β' exists between β2 and β3, while β'' exists at the C-terminal tail of the protein. These two β-strands are lacking in TGM-1 D3, which allows for a lateral expansion of the C-terminal half of the protein away from the N-terminal half. This leads to hydrophobic residues on both faces of TGM-1 D3 being partially exposed to solvent.

These specific adaptations are likely the reason TGM-1 D3 is capable of binding TβRII, with the lateral expansion allowing for the edge β-strand of TβRII to contact the hydrophobic cleft of the nonNC face, and specifically Tyr[253]. Asp[55] and Glu[142] of TβRII likely form stabilizing interactions with residues of TGM-1 D3 in a manner similar to TGF-β, though this has not yet been confirmed. Overall, the structure of TGM-1 D3 suggests that *H. polygyrus* has adapted the canonical CCP domain structure in order to mimic the TβRII binding ability of TGF-β in a similar manner as the native ligand, though this has yet to be proven.

Comparisons of the sequences of the domains in TGM-1 show a high sequence similarity, with all the domains predicted to belong to the CCP family (Fig. 41). The main difference between the domains of TGM-1 is within a six-residue loop extension between β3 and β4. In the TGM-1 D3 structure, this set of residues comprises an extended turn connecting β3 and β4, and in other TGM-1 domains this loop is six residues shorter. This change may alter the lateral expansion of the C-terminal half of the protein, altering residue eposure to allow for binding to other proteins.

The residues that most impact binding to TβRII : Phe^{235}, Val^{236}, Ile^{238}, Tyr^{252}, Tyr^{253} are different in all of the other TGM-1 domains, with the exception of domain 1. It is likely that this lack of conservation enables TGM-1 D3 to bind to TβRII, but not the other domains.

4.5 Respective Contributions

The ITC competitions were performed by Chang Hyeock-Byeon and analyzed by Ananya Mukundan. The TGM-1 D3 unbound and bound backbone assignments were all completed by Ananya Mukundan. The TGM-1 D3 sidechain assignments, RDC analysis, and intramolecular NOE assignment was performed by Chang Hyeock-Byeon and confirmed by Chang Hyeock-Byeon and Ananya Mukundan. Initial refinement of the TGM-1 D3 structure was performed by Chang Heycok-Byeon and the final refinement was performed by Ananya Mukundan. The TβRII backbone assignments were completed by Ananya Mukundan. Additionally, all SPR experiments were completed by Ananya Mukundan. The data in Figures 27-30, 33, and 35-40 were gathered and analyzed by the Ananya Mukundan, the data in Figure 26 was gathered by Chang Hyeock-Byeon and analyzed by Ananya Mukundan, and the data in Figures 31-32, and 34 were gathered by Chang Hyeock-Byeon and analyzed by Chang Hyeock-Byeon and Ananya Mukundan.

4.6 Tables and Figures

Table 7: Sequences of TGF-β and TGF-β receptor constructs.

Construct	Coding region and description (* indicates stop codon)
mmTGF-β27M	Residues 331-442 of human TGF-β2 proprotein, with deletion of residues 381-400 and mutation of the following residues K357R, R358K, L382R, A406K, C409S, L421V, I424V, N426R, T427K, I430V, NCBI NP_001129071 Initiating Methionine-mmTGFβ27M MALDAAYCFRNVQDNCCLRPLYIDFRKDLGWKWIHEPKGYNANFCAGACPYRASKSPSCVSQDLEPLTIVYYVGRKPVEQLSNMIVKSCKCS*

Table 8: Competitive binding between TGM-1 and TGF-β to TβRI and TβRII as assessed by ITC

Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0

Cell	TβRI	TβRI	TβRII
Syringe	TGF-β($T\beta RII)_2$	TGM-1 D12	mmTGF-β27M
Competitor[a]	None	6 μM TGF-β($T\beta RII)_2$	0, 6, or 12 μM TGM-1 D3
Cell concentration (μM)	5	10	15
Syringe concentration (μM)	100	110	150
Temperature (°C)	30	25	35
K_D (nM)	61 (36 – 97)[b]	ND[c]	35 (17 – 64)[b,d]
ΔH (kcal mol^{-1})	-4.2 (-4.5 - - 4.0)[b]	ND[c]	-7.4 (-7.7 – -7.0)[b,d]
ΔG (kcal mol^{-1})	-10	ND[c]	-11[e]
-TΔS (kcal mol^{-1})	-5.8	ND[c]	-3.2[e]

[a]Competitor was added to the sample cell
[b]Fit for one replicate
[c]Unable to be fitted
[d]K_D and ΔH correspond to the parameters, derived from the global fit, for TβRII:mmTGF-β27M binding in the absence of the competitor, uncertainty determined by 68.3% confidence interval
[e]ΔG and -TΔS correspond to the parameters, derived from the global fit, for TβRII:mmTGF-β27M binding in the absence of the competitor, calculated from ΔG = ΔH - TΔS, along with the globally fitted values for K_D and ΔH.

Table 9: TGM-1 D3 NMR structural statistics. Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

NOE	
Intramolecular NOE: i-j = 0	465
Sequential NOE: i-j = 1	323
Short-Range NOE:1 < i-j <5	104
Long-Range: i-j ≥ 5	247
Angle	
TALOS (ϕ,φ) dihedral constraints	120
$^3J_{HNHA}\alpha$	39
RDC	12
RDC: N-H	69
RDC: Hα-Cα	74
RDC: Cα-CO	66
RMSD (Deviations)	
Bonds (Å)	0.008 ± 0.000
Impropers (°)	1.067 ± 0.159
Angles (°)	1.032 ± 0.035
Dihedral (°)	4.171 ±0.441
HBDA (Å)	0.025 ± 0.009
$^3J_{HNHA}\alpha$ (Hz)	1.349 ± 0.092
Ramachandran[a]	
Most Favored	81.2%
Additionally Allowed	11.6%
Generously Allowed	1.4%
Disallowed	2.9%
RMSD[b]	
Secondary Structure[c]	
Backbone	0.68Å
Heavy	1.14Å
Core[d]	
Backbone	1.00Å
Heavy	1.48Å

[a]Ramachandran values from the ten lowest-energy structures
[b]RMSD values are computed from a mean structure
[c]Residues 189-93, 217-221, 228-230, 234-241, 248-252
[d]Residues 178-253

Table 10: TGM-1 D3 variant constructs bind to TβRII as assessed by surface plasmon resonance. Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

Surface	Analyte	Fitted Parameters[a]			
		k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d (μM)	R_{max} (RU)
TβRII	WT TGM-1 D3	(1.6 ± 0.1) x 10^5	0.26 ± 0.01	1.6 ± 0.1	120 ± 10
TβRII	R198A	(1.1 ± 0.1) x 10^5	0.78 ± 0.01	70 ± 1	260 ± 10
TβRII	H199A	(3.3 ± 0.1) x 10^5	0.98 ± 0.01	3.0 ± 0.1	310 ± 10
TβRII	F235A	(4.5 ± 0.1) x 10^5	1.8 ± 0.2	4.1 ± 0.1	63 ± 1
TβRII	V236A	(3.8 ± 0.1) x 10^5	1.8 ± 0.1	4.6 ± 0.1	84 ± 1
TβRII	I238A	(9.4 ± 0.1) x 10^4	2.3 ± 0.1	25 ± 1	140 ± 10
TβRII	Y252A	(7.8 ± 0.2) x 10^4	1.7 ± 0.1	21 ± 1	150 ± 10
TβRII	Y253A	ND[b]	ND[b]	ND[b]	ND[a]
TβRII	K254A	(3 ± 2) x 10^5	12 ± 6	35 ± 1	310 ± 10
TβRII	N255A	(4.8 ± 0.1) x 10^5	1.3 ± 0.1	2.7 ± 0.1	150 ± 10
TβRII	I256A	(8.8 ± 0.1) x 10^5	1.4 ± 0.1	1.6 ± 0.1	130 ± 10
TβRII	K258A	(4.3 ± 0.1) x 10^5	1.2 ± 0.1	2.7 ± 0.1	250 ± 10

[a]Fitted parameters were derived from kinetic analysis of a single injection series
[b]Not determined due to weak signal

Table 11: TGM-1 D3 variant constructs bind to TβRII as assessed by isothermal titration calorimetry

Cell	TGM-1 D3 $I^{236}A$	$Y^{252}A$	$Y^{253}A$	$I^{256}A$
Syringe	TβRII	TβRII	TβRII	TβRII
Cell concentration (μM)	15	8.2	15	15
Syringe concentration (μM)	300	165	300	320
Temperature (°C)	35	35	35	35
K_D (μM)	15 (7 – 121)[ab]	28[c]	ND[d]	1.4 (1.0 – 2.1)[ab]
ΔH (kcal mol^{-1})	-10.3 (-17.2 – 9.9)[ab]	-4.3[c]	ND[d]	-8.2 (-8.8 - -7.6)[ab]
ΔG (kcal mol^{-1})	-6.9	-6.4	ND[d]	-8.2
-TΔS (kcal mol^{-1})	3.5	-2.1	ND[d]	-0.1
Stoichiometry (n)	1.2[e]	ND[d]	ND[d]	1.4[e]

[a]Uncertainty reported as 68.3% confidence interval
[b]Global fit of three replicates
[c]Uncertainty could not be fit due to low signal
[d]Not determined due to weak signal
[e]Number of sites determined by incompetent fraction value on sedphat; set to '1' for K_D analysis

Table 12: TβRII variant constructs bind to TGM-1 D3 as assessed by surface plasmon resonance. Table from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

Surface	Analyte	Fitted Parameters[a]			
		k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d (μM)	R_{max} (RU)
TGM-1 D3	WT TβRII	(4.1 ± 0.1) x 10^5	0.7 ± 0.1	1.6 ± 0.1	240 ± 10
TGM-1 D3	D55N	(5.0 ± 0.2) x 10^4	3.1 ± 0.1	63 ± 1	200 ± 10
TGM-1 D3	I73A	(1.6 ± 0.1) x 10^5	1.1 ± 0.1	6.9 ± 0.1	220 ± 10
TGM-1 D3	S75L	(1.3 ± 0.1) x 10^4	3.9 ± 0.9	4.1 ± 30	250 ± 20
TGM-1 D3	I76A	(4.7 ± 0.1) x 10^4	1.2 ± 0.1	26 ± 1	430 ± 10
TGM-1 D3	E142Q	(4.1 ± 0.1) x 10^5	10 ± 10	17 ± 1	130 ± 10
[a]Fitted parameters were derived from kinetic analysis of a single injection series					

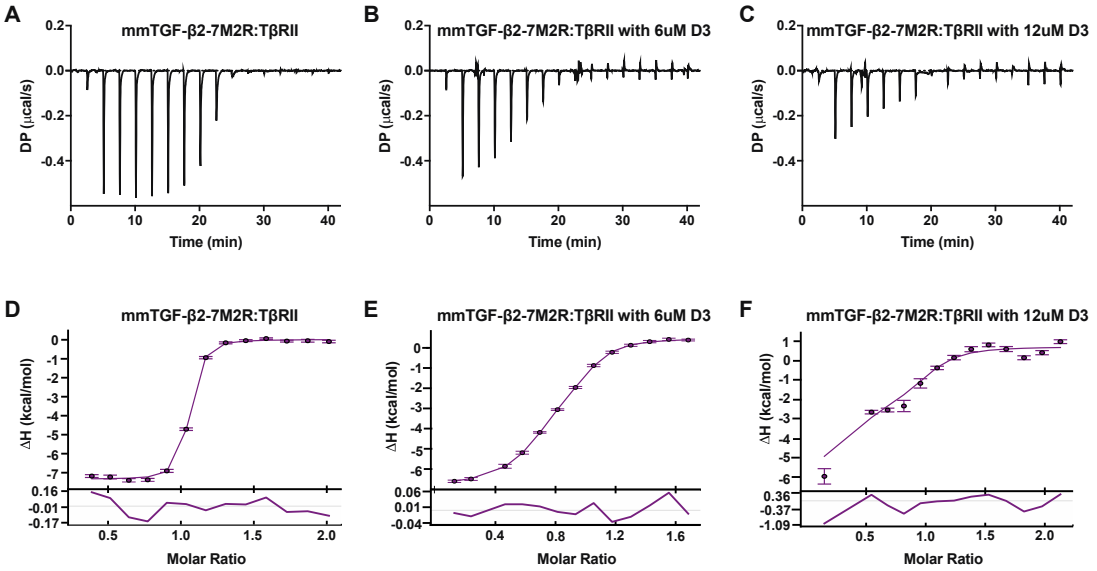

Figure 26: TGM-1 D3 competes with mmTGFβ7M for binding to TβRII as shown by ITC. A-H. Raw thermograms for the injection of mmTGFβ7M into a cell containing TβRII and injection of 0 µM TGM-1 D3 (A), 6 µM TGM-1 D3 (B), and 12 µM TGM-1 D3 (C). Corresponding integrated heats for the injection of 0 µM TGM-1 D3 (D), 6 µM TGM-1 D3 (E), and 12 µM TGM-1 D3 (F) data fit to a 1:1 binding model. Fits correspond to the global fit over the data sets with residuals below. Error bars correspond to undertainty in the estimation of integrated heats by NITPIC. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

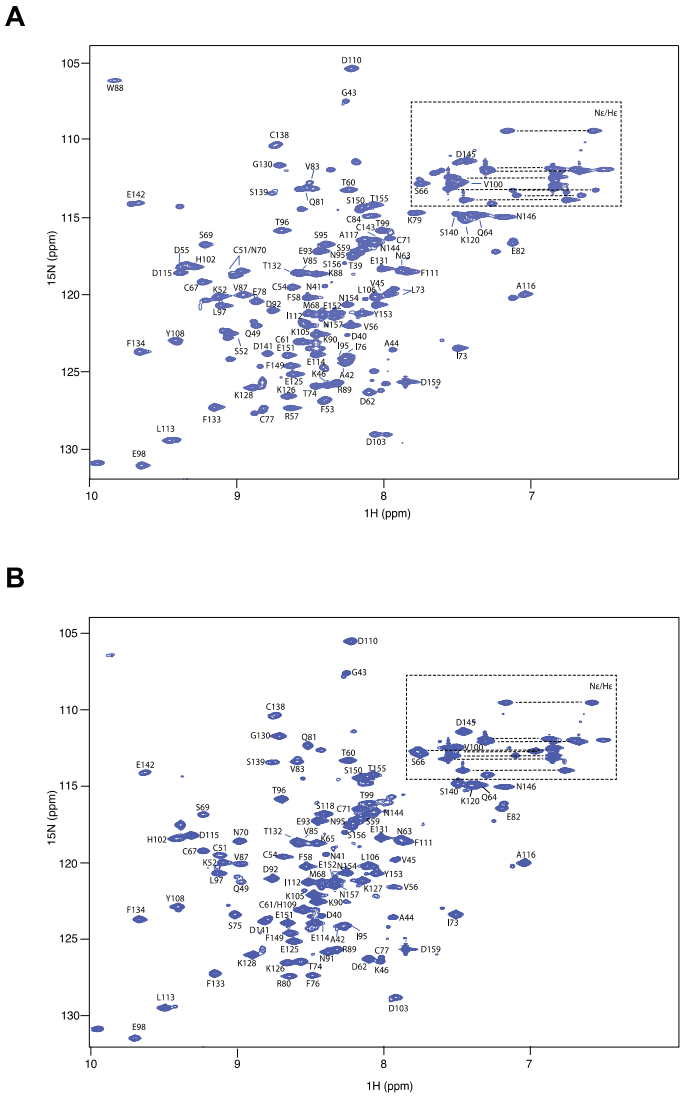

Figure 27: ^{1}H-^{15}N Assignment of TβRII. ^{1}H-^{15}N HSQC of free TβRII (A) and TβRII bound to TGM-1 D3 (B), assigned peaks displayed on the spectra. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K. Dashed line indicating region of HSQC corresponding to sidechain Asn/Gln residue amino groups. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

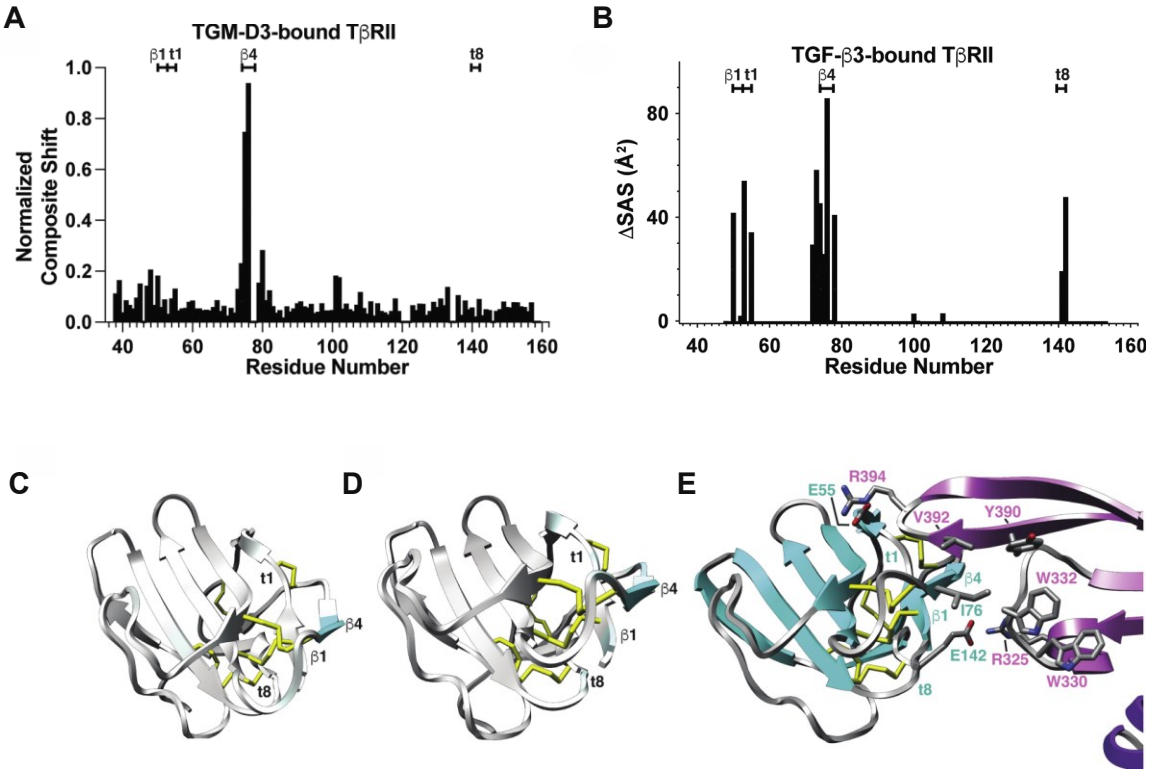

Figure 28: TGM-1 D3 binds to TβRII using similar residues to TGF-β. Plot of the normalized composite shift perturbations of TGM-1 D3 bound TβRII (A), with corresponding depiction onto the structure of TβRII (C) with white indicating little perturbation and blue indicating close to 1.0 normalized composite shift perturbation. Plot of the difference in solvent exposed surface area between free TβRII and TGF-β bound TβRII (B), with corresponding depiction onto the structure of TβRII (D) with white indicating little composite shit perturbation and blue indicating close to 1.0 normalized composite shift perturation. A depiction of the TβRII:TGF-β3 interface is shown in panel E with relevant structural features annotated, TβRII colored in blue and TGF-β monomer colored in purple. and TβRII:TGF-β structure from PDB 1KTZ. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

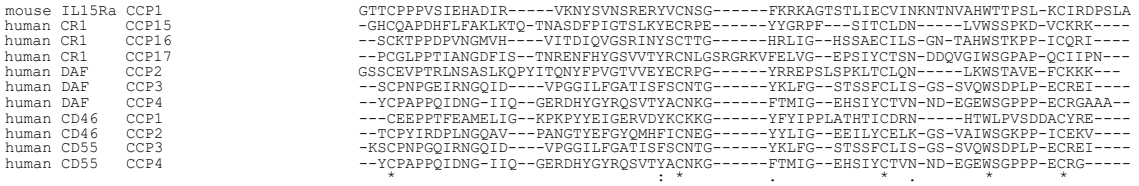

Figure 29: Sequence alignment of CCP domains. Alignment of CCP domains for mouse IL-15 receptor, human CR1 CCP domains 15 – 17, human DAF CCP domains 2 – 4, human CD46 CCP domains 1 – 2, and human CD55 CCP domains 3 – 4. '*****' indicates invariable residues, '**:**' indicates conservation between residues with similar properties, and '**.**' indicates conservation between residues with less similar properties. All analysis completed using Clustal Omega[279] and subsequent manual adjustment.

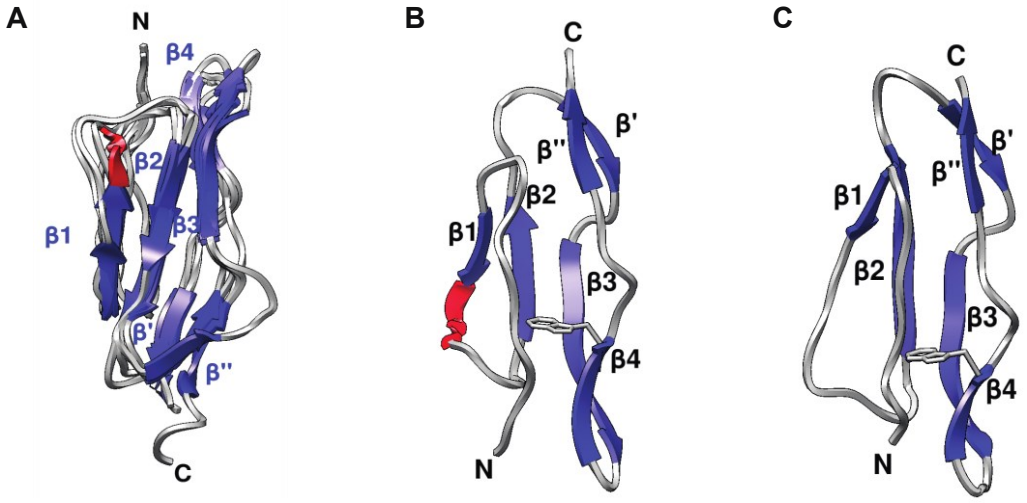

Figure 30: Structures of Representative CCP domains. A. Alignment of multiple CCP domains (PDB: 2psm, 1ckl, 1h2p, 5fo9, 5foa) aligned using UCSF Chimera: β-strands, blue; loops, gray; 310 helix, red. B-C. Representative structure of one of the CCP domains of CD46 (B, PDB: 1ckl), and CD55 (C, PDB: 1h2p), with the same secondary structure coloring. Key structural features are demonstrated, and conserved tryptophans shown.

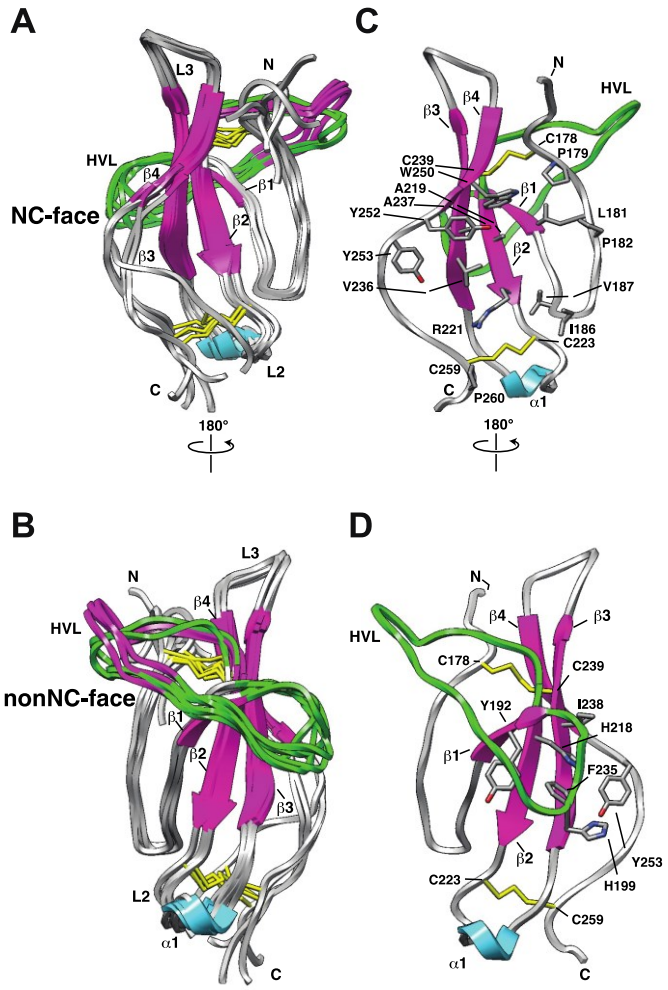

Figure 31: Solution NMR structure of TGM-1 D3. A-D. Ensemble of the five lowest-energy conformers of the unbound TGM-1 D3, with appropriate secondary structures indicated: β-strands, pink; loops, gray; 310 helix, light blue, disulfide bonds; yellow, two different conformations of the HVL in green and in pink. Key residues are indicated. Orientations are shown with respect to a 180 rotation about the y-axis. The orientation containing both the N- and C-termini, the NC face is shown in (A), with a 180 rotation to the nonNC face shown in (C). Singular representative structures from the ensemble shown to the right in B and D respectively. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

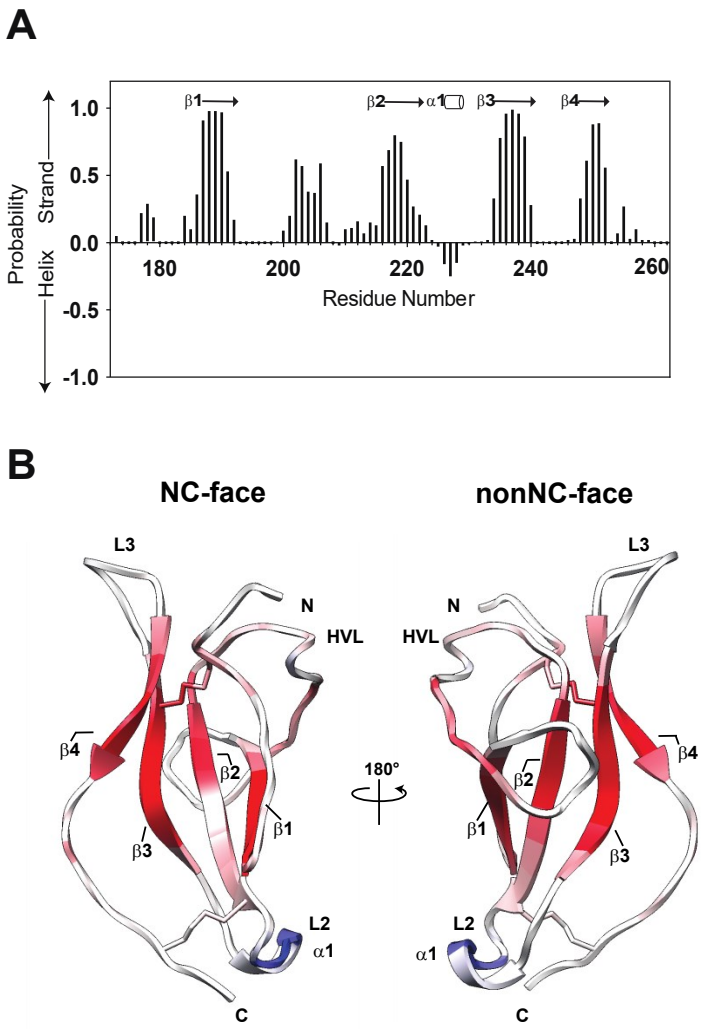

Figure 32: Solution NMR structure of TGM-1 D3 in comparison to PECAN predictions. A-B. PECAN prediction of TGM-1 D3 secondary structure, with positive values indicating strand probaiblity and negative values indicating helical probability. Representative of the unbound TGM-1 D3 structure, with key structural features indicated (B) colored according to the PECAN probability, with positive or β-strand probability indicated by a red gradient and a negative or helical probability indicated by a blue gradient . Orientations are shown with respect to a 180 rotation about the y-axis with the NC face on the left and the nonNC face on the right.. Adapted from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

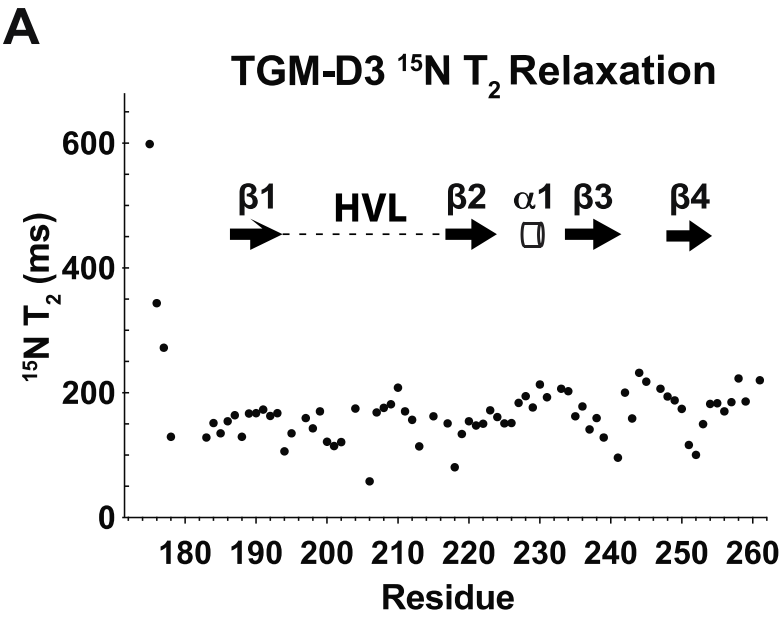

Figure 33: ^{15}N T_2 relaxation times for TGM-1 D3.. Per residue relaxation times plotted, structural features are indicated above the graph. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

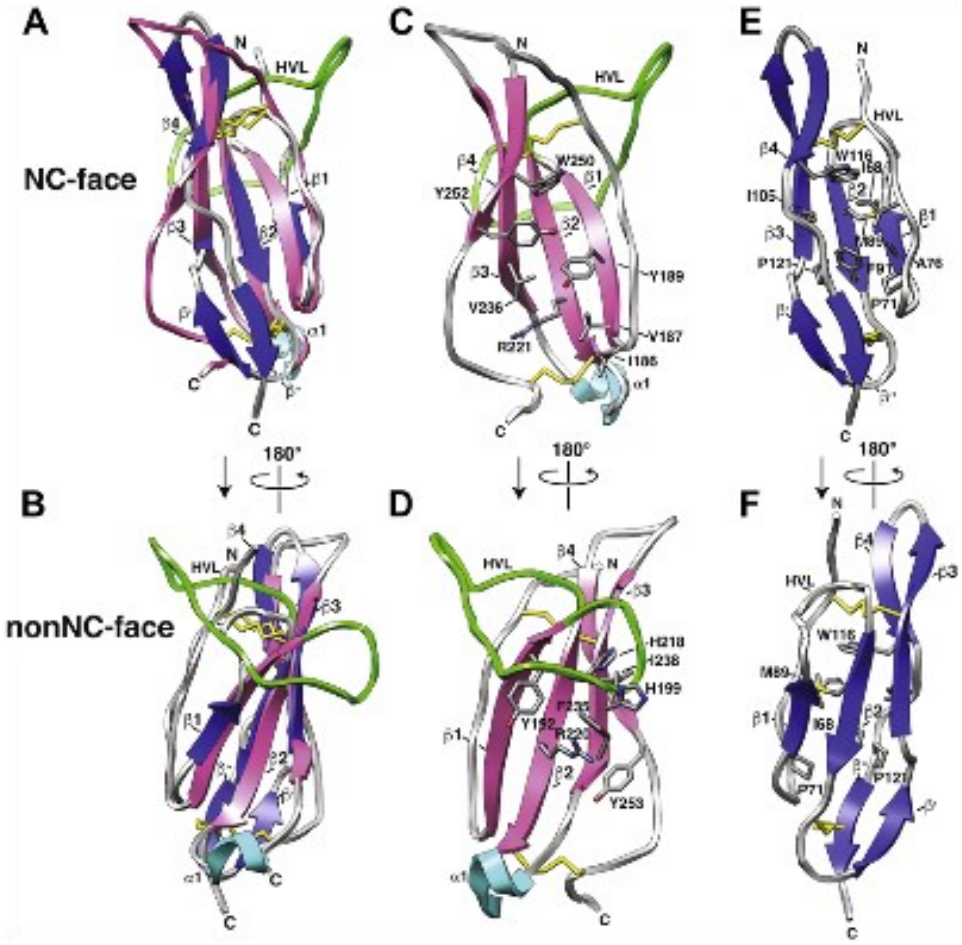

Figure 34: Solution NMR structure of TGM-1 D3 in comparison to canonical CCP domains. A-F. Alignment of TGM-1 D3 secondary structure, in comparison to the top DALI hit human CD46 (PDB 1CKL) from the orientation of the NC face (A) and nonNC face (B), or in those same orientations without the CCP overlay (C, D), with key structural features indicated: β-strands, pink; loops, gray; 310 helix, light blue, disulfide bonds; yellow, HVL; green. Human CD46 (PDB 1CKL) in the same orietnations as the overlay (E, F) with key structural features indicated: β-strands, dark blue; loops, gray; disulfide bonds; yellow. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

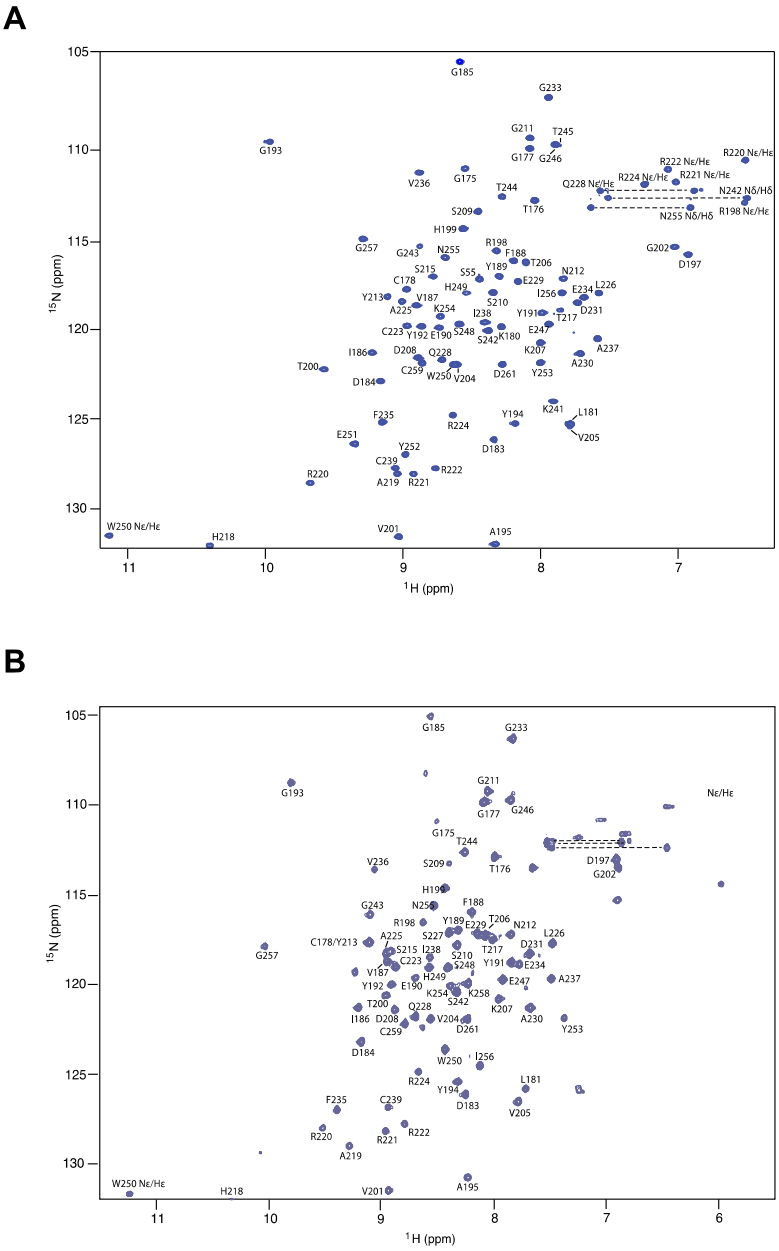

Figure 35: ^{1}H-^{15}N Assignment of TGM-1 D3. ^{1}H-^{15}N HSQC of free TGM-1 D3 (A) and TGM-1 D3 bound to TβRII (B), assigned peaks displayed on the spectra. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K. Dashed lines indicating region of HSQC corresponding to sidechain Asn/Gln residue amino groups. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

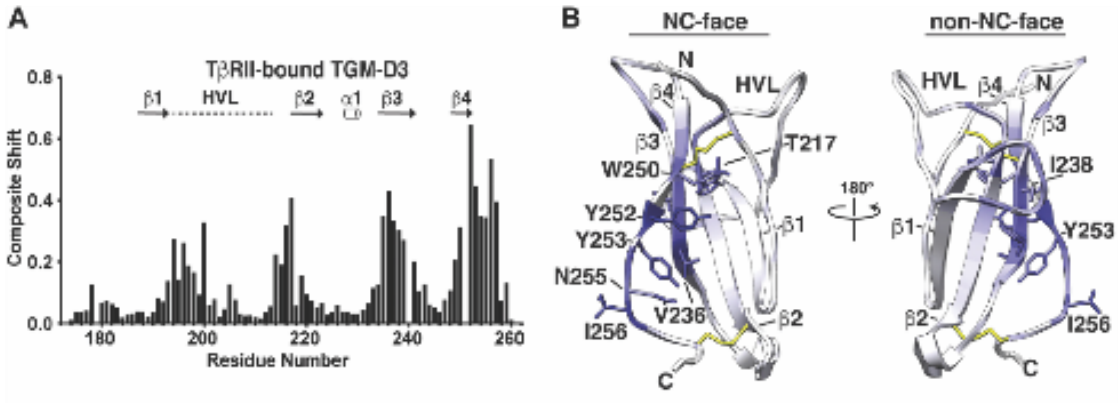

Figure 36: Composite shift perturbations of TGM-1 D3 upon binding TβRII. Plot of the normalized composite shift perturbations of TβRII bound TGM-1 D3 (A), with corresponding depiction onto the structure of TGM-1 D3 (B) on the NC face (left) and nonNC face (right) with white indicating little-no composite shift perturbation and blue indicating close to 1.0 normalized composite shift perturbation. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

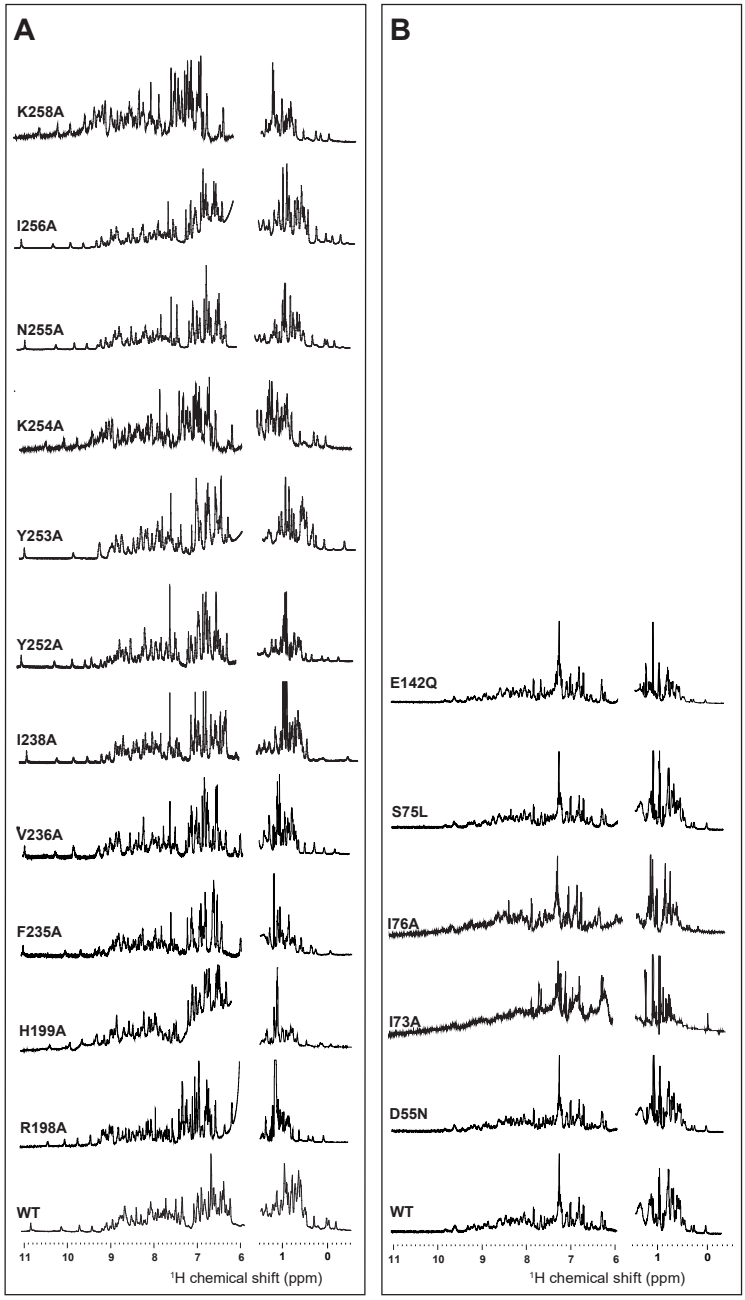

Figure 37: ^{1}H spectra of TGM-1 D3 and TβRII mutants. ^{1}H NMR spectra of the amide (left) and methyl (right) regions of TGM-1 D3 variants (A) and TβRII variants (B) compared to wild-type. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

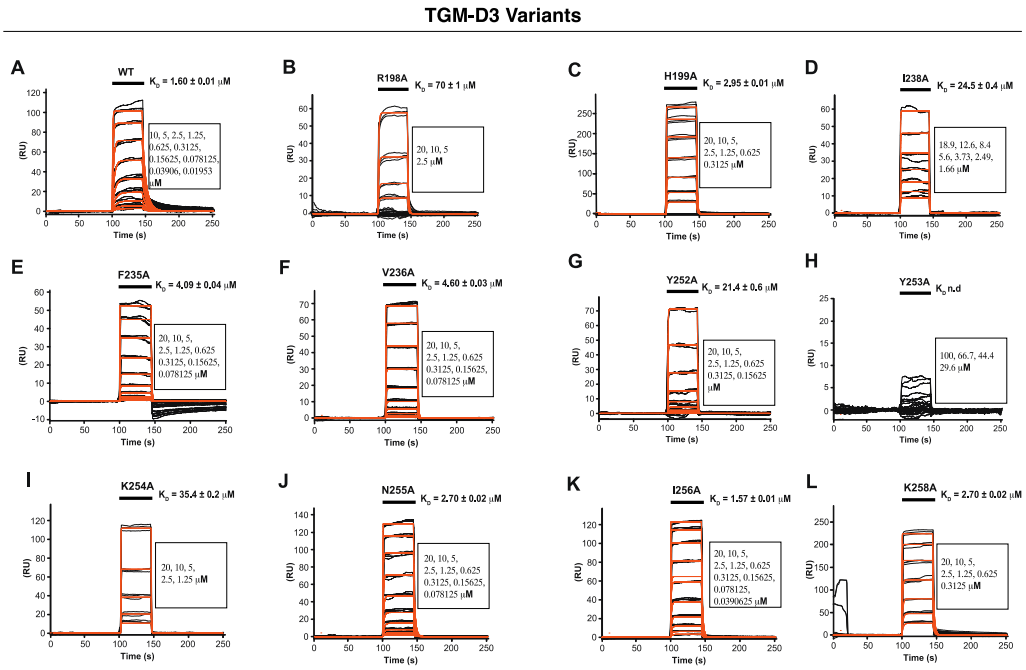

Figure 38: Binding of mutated TGM-1 D3 constructs to TβRII as shown by SPR. A-. SPR sensorgrams upon injection of TGM-1 D3 WT (A), $R^{198}A$ (B), H^{199A} (C), $I^{238}A$ (D), $F^{235}A$ (E), $V^{236}A$ (F), $Y^{252}A$ (G), $Y^{253}A$ (H), $K^{254}A$ (I), $N^{255}A$ (J), $I^{256}A$ (K), and $K^{258}A$ (L) over biotinylated avi-tagged TβRII immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom rihgt of each sensorgram. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

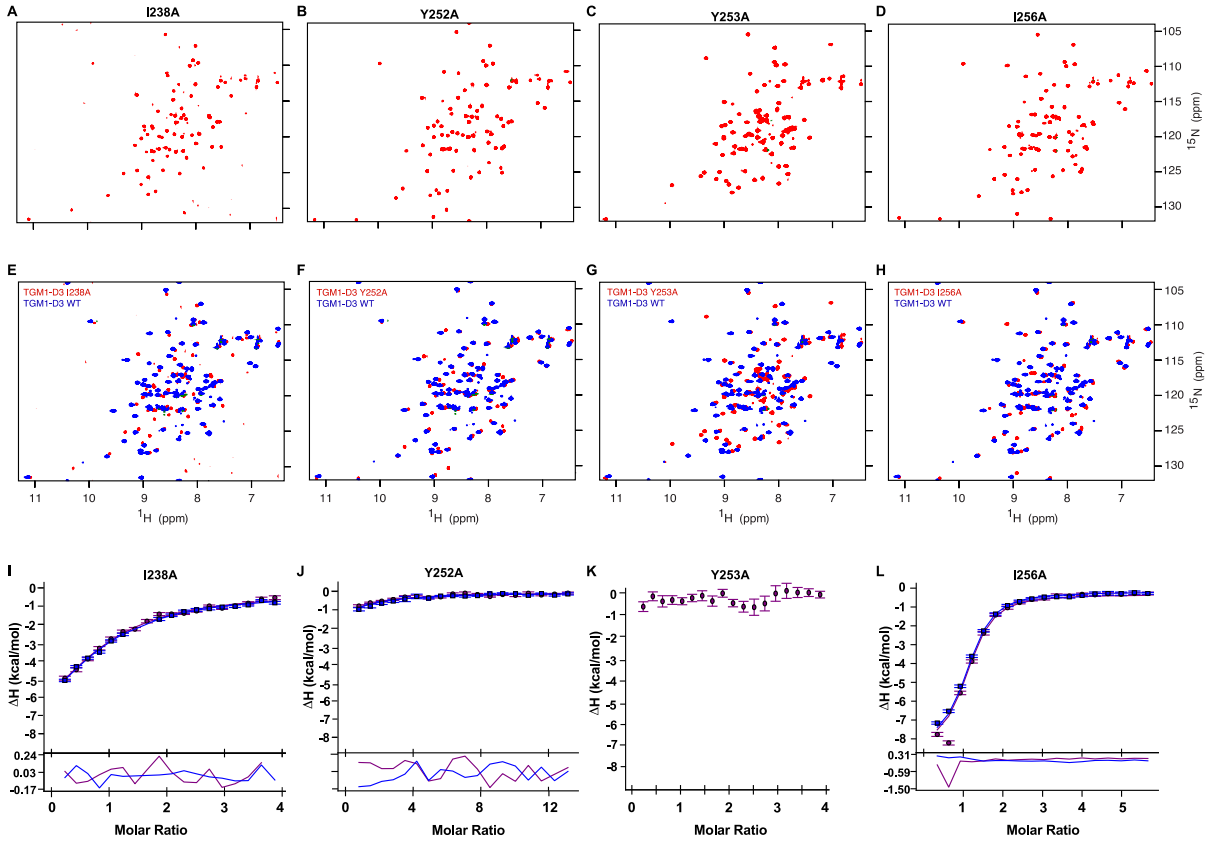

Figure 39: Binding of mutated TGM-1 D3 constructs to TβRII as shown by ITC. A-L. ^{1}H-^{15}N HSQC of free TGM-1 D3 I238A (A) Y252A (B), Y253A, (C), and I256A (D), alone and overlaid with TGM-1 D3 WT (E-H) respectively. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K. ITC isotherms upon injection of TβRII into cells contaning TGM-1 D3 I238A (I), Y252A (J), Y253A (K), and I256A (L). Fits correspond to the global fit over the data sets with residuals below. Error bars correspond to undertainty in the estimation of integrated heats by NITPIC.

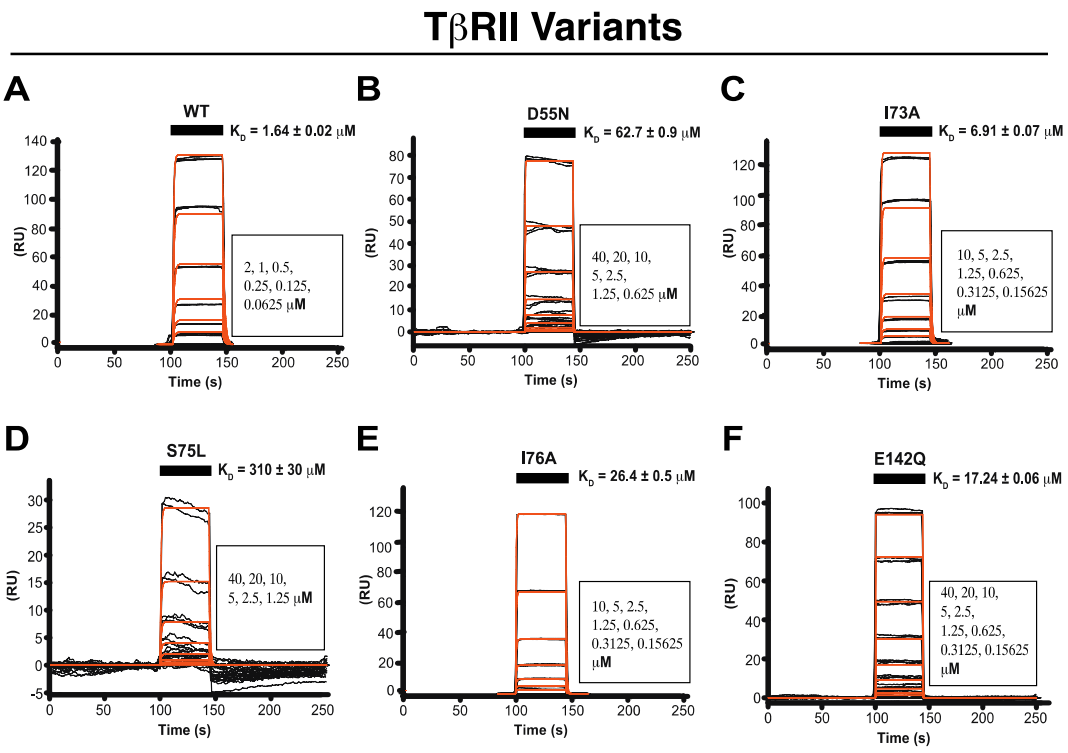

Figure 40: Binding of mutated TβRII constructs to TGM-1 D3 as shown by SPR. A-F. SPR sensorgrams upon injection of TβRII WT (A), $D^{55}N$ (B), $I^{73}A$ (C), $S^{75}L$ (D), $I^{76}A$ (E), and $E^{142}Q$ (F) over biotinylated avi-tagged TGM-1 D3 immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom right of each sensorgram. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

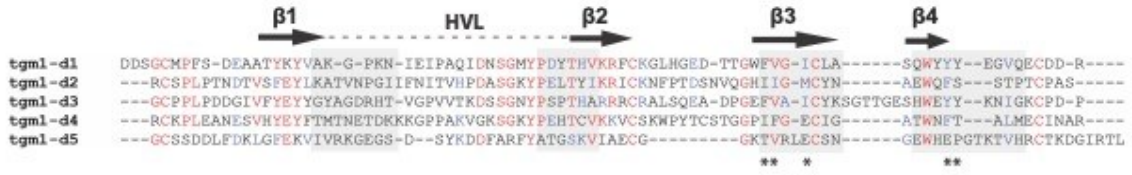

Figure 41: Sequence alignment of TGM-1 domains. Alignment of TGM-1 domains 1-5. Red indicates conserved resiues and blue indicates residues with similar properties. On top are the secondary structure features of TGM-1 D3 and gray highlights correspond to regions with composite chemical shift perturbation values greater than 0.1 (from TGM-1 D3 binding to TβRII). Asterix indicate residues of TGM-1 D3 mutation to alanine of which greatly impacts binding to TβRII. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

5.0 Structure of the TGM-1 D3:TβRII Complex

5.1 Introduction

The lateral expansion of TGM-1 D3 caused by the lack of β' and β'' strands and the elongated, structured HVL likely act in concert to contribute to the binding of TGM-1 D3 to TβRII. Mutagenesis of residues of TGM-1 D3 in both the HVL (Arg^{198}) and the C-terminal β-strands (Ile^{238}, Tyr^{253}) led to a greater than 20-fold reduction in the binding affinity to TβRII as assessed by SPR (Table 11). Mutagenesis of residues of TβRII known to be necessary for binding to TGF-β showed that TGM-1 D3 was binding to the same residues of TβRII as the native ligand, making it likely that selective pressure has allowed for these specific adaptations of TGM-1 D3 in order to bind TβRII.

A comparison of the binding interfaces of canonical CCP domains shows no singular common surface or set of residues is used by these proteins to bind with their ligands (Fig. 42). The top PDB hits for proteins that were structurally similar to TGM-1 D3 were assessed for their binding interfaces. The difference in solvent accessible surface area was calculated for the free CCP domains in comparison to the bound CCP domains and modeled onto the PDB structure of the domains (Fig. 42). From these structures, it is apparent that there is no one shared binding interface used for binding; the proteins utilize the extended N-terminal region, β1, β2, and the β2-β3 loop, along with the interdomain connections, in order to bind their ligands (Fig. 42).

From this, the following questions need to be answered: 1) what interfaces form the TGM-1 D3:TβRII complex, 2) how are these interfaces similar and dissimilar from the TGF-β:TβRII

interface, and 3) how the two major adaptations of the TGM-1 D3 CCP domain might play a role in this. Understanding the similarities and differences between the binding interfaces may also help in developing TβRII-specific binders that can act to inhibit TGF-β signaling by sequestering TβRII. The field of TGF-β inhibitors is vast, but subdomains of TGM-1 have potential to act as inhibitors through its modular binding. Thus TGM-1 D3 or modified versions of TGM-1 D3 may be able to be used as a TβRII binder and a TGF-β inhibitor.

Previous experimentation has shown that only TGM-1, TGM-2 and TGM-3 are able to activate TGF-β signaling via the MFB-F11 assay[242]. However, despite a high sequence identity, TGM-4 does not signal via the MFB-F11 assay, though it is able to convert naïve CD4+ T-cells into T_{regs}[242]. This implies a differential ability in the TGM family members to bind to the TGF-β receptors, which is somewhat surprising as there is a high sequence identity between domains within the family (Table 2). It is clear that a complex structure between TGM-1 D3 and TβRII would not only solidify the inferences made by the structural and biophysical experiments made in the previous chapter but would also generate an understanding of the TGM family of parasitic proteins, potentially explaining why the parasite needs such a large family of highly similar proteins, to which receptors these proteins bind, and how this family aids in parasitic pathology.

5.2 Methods

Cloning, expression, and purification of TβRII 46-153, unlabeled and ^{15}N-labeled lanthanide-binding tag TGM-1 D3, and lanthanide-binding tag TβRII constructs. Constructs, expression protocols, and protein purification protocols were the same as stated in Chapter 3[36,243,244].

Cloning, expression, and purification of ^{2}H ^{15}N TβRII and ^{2}H ^{15}N TGM-1 D3. Constructs, expression protocols, and protein purification protocols were the same as stated above with one exception; expression was performed in M9 media using $^{2}H_2O$ rather than $^{1}H_2O$. Following purification, the final solution was solubilized in 25 mM Na_2HPO_4 50 mM NaCl pH 6.0 5% $^{2}H_2O$.

Analytical size exclusion chromatography assessment of the formation of the TGM-1 D3:TβRII complex. The size-exclusion columns utilized were Superdex 10-300 75 and Superdex 10-300 200 columns. Samples of TGM-1 D3, TβRII, and TGM-1 D3:TβRII were run with an excess of TGM-1 D3 in buffer conditions of 25 mM Tris 150 mM NaCl pH 7.5. The column was run at a rate of 0.5 mL min^{-1} and relevant fractions were collected, concentrated, and analyzed by SDS-PAGE gel to assess for protein composition.

2D NMR experiments. All NMR experiments were run in the buffers mentioned in the figures, usually 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0, 310K, excepting any experiments involving constructs with lanthanide-binding tags which were run in 25 mM MES pH 6.0. Each sample was run in a 5 mm susceptibility-matched microtube for data collection. NMR data was collected with Bruker 600, 700, and 800 MHz spectrometers containing a 5 mm ^{1}H (^{13}C, ^{15}N) z-gradient "TCI" cryogenically cooled probe. ^{1}H-^{15}N HSQC spectra were acquired as described,

with water flipback[250] and WATERGATE suppression pulses[251]. NMR spectra were processed, analyzed, and visualized using NMRPipe[253] and NMRFAM-SPARKY[254].

Determination of the pseudocontact shift tensors and generation of intermolecular PCS restraints. ^{15}N TGM-1 D3 loop protein was purified and then dialyzed into 25 mM MES pH 6.0. The sample was prepared at a concentration of ca. 60 μM in a buffer of 25 mM MES 5% $^{2}H_2O$, with the addition of either La^{2+} or Tb^{3+} at a molar ratio of 1:1.1 (^{15}N TGM-1 D3 loop protein: lanthanide metal). Subsequently the ^{1}H-^{15}N HSQC spectra were recorded at 310K in the 25 mM MES 5% $^{2}H_2O$ buffer. This same process was repeated for the ^{15}N TβRII DE construct as well, using the same buffer and the same ratio of La^{2+} and Tb^{3+}, though at a concentration of 100 μM. Differences in the ^{1}H chemical shift (Tb^{3+} - La^{2+}) were calculated on a per residue basis for the ^{15}N TGM-1 D3 loop LBT and the ^{15}N TβRII DE LBT samples.

These values, along with the PDB structure of either TGM-1 D3 (7SXB) or TβRII (1PLO), were input into the program Numbat[280], which utilizes these $\Delta\delta_{ppm}$ to calculate the tensor and gives a list of calculated PCS values. The metal position was initially constrained to a 6Å sphere around the residues replaced by the LBT tag. If the calculated PCS value deviated significantly from the observed PCS value, the residue assignment and $\Delta\delta_{ppm}$ was scrutinized and then discarded if the residue assignment was determined to be unclear. This process was iterated until a final tensor was determined, which included a definitive metal position. The PCS values used then formed the intramolecular TGM-1 D3 Loop LBT PCS restraint list and the intramolecular TβRII DE LBT PCS restraint list.

For the generation of intermolecular PCS values, two sets of samples were produced. ^{15}N TGM-1 D3, ^{15}N TβRII, TGM-1 D3 loop LBT, and TβRII DE LBT samples were produced. The complex ^{15}N TGM-1 D3:TβRII DE LBT sample was produced by adding each individual

component in a 1:1.2 molar ratio, with the labeled component at a concentration of ca. 80 μM, in a buffer of 25 mM MES pH 6.0. Either La^{2+} or Tb^{3+} at a molar ratio of 1:1.1, and the ^{1}H-^{15}N HSQC spectrum were recorded at 310K in 25 mM MES 5% $^{2}H_2O$. This same process was repeated for the ^{15}N TβRII:TGM-1 D3 Loop LBT sample as well, using the same molar ratio of 1:1.2, at a concentration of 40 μM, the same buffer, and the same ratio of La^{2+} and Tb^{3+}, as well as the paramagnetic lanthanides europium and thulium. Differences in the ^{1}H chemical shift were calculated on a per residue basis for both samples and used as the intermolecular PCS restraints.

Intermolecular NOE restraints for the TGM-1 D3:TβRII complex. To calculate NOE distance restraint data, two separate samples were made: ^{2}H ^{15}N TGM-1 D3:TβRII, and ^{2}H ^{15}N TβRII:TGM-1 D3, both at concentrations of 250 μM in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. NOE distance restraints were gathered using data collected from 3D ^{1}H-^{15}N NOESY spectra and the distances derived using the program CCP-NMR[265]. To assign the aforementioned distance restraints, further NMR spectra were gathered for a sample of ^{15}N ^{13}C TβRII:TGM-1 D3, at a concentration of 400 μM, in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. The spectra gathered include 2D ^{1}H-^{13}C CT-HSQC, 3D HCCH-TOCSY, 3D HBHA(CO)NH, 3D H(CCCO)NH, and 3D CC(CONH). All experiments were gathered at 310K.

Once the NOE restraints were assigned, distance restraints were calculated using the isolated spin-pair approximation, which linearizes the relationship between NOE peak intensity and distance[281]. In order to do so, the peak volumes for both the diagonal peaks (^{1}H-^{15}N amide hydrogen and nitrogen chemical shifts) and the NOE cross peaks were calculated. To calculate the NOE distance, the ratio of the cross to diagonal peak intensities was used. For the purposes of this set of experiments, a distance of 3.2 angstroms was set to a normalized peak intensity ratio of 0.06.

The experimental NOE distance was thus calculated as such: $r_{exp} = r_0 \left(\frac{v_0}{v_{exp}}\right)^c$, where r_0 is the reference distance of 3.2 angstroms, v_0 is the normalized peak intensity ratio of 0.06, v_{exp} is the ratio or cross peak to diagonal peak intensities. C is a constant theoretically set to 1/6, but practically set to 1/5.2 due to spin diffusion. The accuracy of the cross-peak volume measurement was additionally used to generate upper and lower bounds for the distances.

RDC analysis of the TGM-1 D3:TβRII complex. Two sets of residual dipolar coupling (RDC) restraints were gathered. These RDCs were measured using a 2D IPAP-HSQC experiment[267], as described previously. Separate samples of ^{2}H ^{15}N TGM-1 D3:TβRII and ^{2}H ^{15}N TβRII:TGM-1 D3 were buffered with 10 mg mL^{-1} Pf1 phage[266] to determine the anisotropic J_{app}. Similarly, separate samples of ^{1}H ^{15}N TGM-1 D3:TβRII and ^{1}H ^{15}N TβRII:TGM-1 D3 I84A were buffered into 25 mM Na_2HPO_4 50 mM NaCl pH 6.0 to determine the isotropic J_{app}. The difference between the two per residue was taken as the RDCs.

NMR Structure determination. PCS, RDC, and NOE restraints were input into the program Xplor-NIH[263], and structures calculated using a rigid-docking scheme, where the backbone residues are kept rigid and the sidechain residues allowed to vary in position. The rigid docking scheme was edited to include PCS, RDC, and NOE restraints (Appendix A.1.2). Structure calculations were performed using the PDB structures for free TGM-1 D3 (PDB: 7SXB) and TβRII (PDB: 1PLO). The program PROCHECK[269,270] was used to perform the Ramachandran analysis and RMSD analysis were performed using Chimera.

5.3 Results

5.3.1 Structure determination of the TGM-1 D3:TβRII complex using X-ray crystallography

The previous chapter showed that the size of the TGM-1 D3:TβRII complex is just under 25 kDa. As such, the complex is well-within the size limits to be determined by NMR. However, the previous two chapters also suggested few intermolecular interactions between the two proteins, as shown via the per-residue composite shift perturbation of the structure of TβRII upon binding to TGM-1 D3 (Fig. 28A, C). Structure determination of the complex by NMR may potentially be difficult, owing to limited intermolecular NOE distance restraints. As such, initial structure determination was attempted by X-ray crystallography.

The construct for TGM-1 D3 used was the same as in previous chapters, beginning with four residues comprising a GSGT tag, followed by the 85 residues of TGM-1 D3 (Table 3). The construct was not truncated. Even though the N and C-terminal tails of the protein were showed to have longer T_2 relaxation times and thus higher levels of flexibility, cleaving N-terminal residues may impact the ability of TGM-1 D3 to fold. N-terminal residues of TGM-1 D3 are adjacent to the invariant cysteine residues involved in disulfide bond formation, and cleaving C-terminal residues may impact the binding to TβRII. The construct for TβRII used was truncated, including residues 46-153 (Table 13) of the extracellular domain as opposed to residues 38-159 as in previous chapters. These residues comprise highly flexible and disordered regions as shown in previous structures PDB: 1PLO, 1M9Z[273,282], with a construct containing residues 41-159 having previously been used to generate a crystal structure of TβRII (PDB: 1M9Z) [282], and a construct

containing residues 42-149 used for the TGF-β:TβRI$_2$:TβRII$_2$ heterotetrameric complex crystal structure (PDB: 2PJY)[49].

The binding affinity for the TβRII:TGM-1 D3 interaction is somewhat weak at 0.8 μM, with a rapid on and off-rate (Fig. 15, Table 5). The ability of the two proteins to form a complex via size-exclusion chromatography (SEC) was first tested using analytical SEC. TGM-1 D3 and TβRII were run in a 1.1:1 molar ratio on an analytical Superdex 10-300 200 column. The chromatogram shows a single peak with a lagging overlapping peak at 16.5 mL and 17.5 mL respectively, which via SDS gel correspond to TβRII (16.5 mL) and TGM-1 D3 (17.5 mL) (Fig. 43A). This aligns with how the proteins run individually, with the TGM-1 D3 eluting at 17.5 mL and the TβRII eluting at 16.5 mL (Fig. 43A-B), congruent with the MW of TGM-1 at ca. 10 kDa and the MW of TβRII at ca. 13 kDa. Hence TβRII and TGM-1 D3 cannot be isolated from a peak do not form a stable complex that can be isolated separately from peaks containing TβRII or TGM-1 D3.

To generate the crystallographic stock, TGM-1 D3 and TβRII were dialyzed into the same buffer: 25 mM Tris, 50 mM NaCl, pH 8.0. Following this the two proteins were added in equimolar amounts and concentrated, to a total concentration of 50 mg mL^{-1}. The two proteins are highly soluble under those solution conditions and no precipitate was visible. The concentrated stock was assessed for protein identity via SDS gel; both proteins were clearly present (Fig. 43C). The concentrated stock was then sent for crystallographic screening (1200 conditions); however, the proteins did not crystallize under any of the conditions tested.

However it should be noted that after the stock was sent for crystallization screening, attempts to isolate the TGM-1 D3:TβRII complex via analytical SEC were re-performed, specifically using the 10-300 75 Superdex column, which did end up showing separation of the

complex from its individual components, with the TGM-1 D3:TβRII complex eluting at 11.7 mL, TβRII eluting at 13.0 mL, and TGM-1 D3 eluting at 13.8 mL which corroborates with the previously stated molecular weights of each of the individual proteins (Fig. 43D-E).

5.3.2 NMR was utilized for the structure determination of the TGM-1 D3:TβRII complex

Due to the inability to obtain crystals of the TGM-1 D3:TβRII complex under the conditions screened, NMR was utilized to determine the structure of the complex by docking the two individual proteins. This could be feasibly done as both the structure of TGM-1 D3 (PDB 7SXB)[257] and the structure of TβRII (PDB: 1PLO, 1M9Z)[273,282] have been already determined. Both of the partner proteins can be easily produced with isotopic labeling in sufficient quantities in *E. coli*, as has been mentioned in the previous chapters. The stoichiometry is 1:1 and so the total molecular weight is less than 30 kDa, making determining the complex of the structure achievable. The construct of TGM-1 D3 used for all measurements was the wild-type construct as has been previously described (Table 3), while the construct of TβRII used was the wild-type construct from residues 38-159 (Table 4). For the structure calculation, the lowest energy conformer from the previously published TGM-1 D3 structure was used (PDB: 7SXB), as was a modified version of a previously published structure of TβRII (PDB: 1PLO). The TβRII structure was modified to start at residue Lys^{46} rather than at Val^{39} due to the lack of convergence between the NMR conformers in this region. The NMR structure (1PLO) was used as it contained residues 153-159, which were not present in the crystallographically-derived TβRII structure (1M9Z).

Three different sets of restraints were used to determine the complex structure via NMR: pseudocontact shift (PCS) restraints, which gives information about per-residue distance from a

paramagnetic center[283,284], residual dipolar coupling (RDC) restraints[285], and a few intermolecular nuclear Overhauser enhancement (NOE) restraints[286,287]. The existing TGM-1 D3 and TβRII structures, along with the aforementioned restraints, were docked using rigid-body conditions for structural determination, which has been shown to be fruitful even when few intermolecular NOE restraints and RDC restraints are available[286,288] using the program Xplor-NIH[263].

5.3.2.1 Pseudocontact shifts were obtained for both partner proteins

For the purposes of NMR, paramagnetism is relevant as the presence of unpaired electrons result in an electronic spin ($\vec{S}$) that is then associated with a magnetic moment ($\vec{\mu}$) via the equation $\vec{\mu} = cg\vec{S}$, where c is a constant and g is the electron tensor[284]. This tensor can be isotropic or anisotropic, the latter occurring when the spin magnetic moment couples to the electron orbital magnetic moment[284]. Pseudocontact shifts occur due to the dipolar interactions between an anisotropic paramagnetic metal and NMR-active nuclear spins, interacting through-space over a distance of up to 40 $\dot{A}$.

The change in chemical shift can be expressed as a function of the axial and rhombic components of the anisotropy tensor[284]. Eight constraints are needed to define the PCS tensor: coordinates of the paramagnetic center (x, y, z position), the axial and rhombic components of the PCS tensor (Δx_{ax} and Δx_{rh} respectively), and the orientation of the PCS tensor with respect to the protein of interest (α, β, γ)[283]. Additionally, paramagnetic metals can also result in paramagnetic relaxation enhancement (PRE), which leads to signal broadening. Thus, to utilize PCS restraints for structure determination, metals with high anisotropy but minimal PRE should be chosen, which is the case for anisotropic lanthanides, which have a free electron in the f orbital.

For practical utilization of PCS restraints, a protein is labeled with a lanthanide-binding tag (LBT, sequence: YIDTNNDGWIEGDELY, PDB 1TJB)[289] and the residue chemical shifts are determined for the protein in the presence of a diamagnetic lanthanide (lanthanum) and a paramagnetic lanthanide (terbium, thulium, etc.). The per-residue chemical shift perturbation is determined, and the PCS is calculated from the difference in the chemical shift in the presence of a diamagnetic and paramagnetic lanthanide. The eight tensor parameters are then determined by fitting the measured PCS and the known structure to the PCS-fitting equation:

$$\Delta\delta_{PCS} = \frac{1}{12\pi {r_i}^3}\left(\frac{\Delta x_{ax}\left(2z_i^2 - x_i^2 - y_i^2\right)}{{r_i}^2} + \frac{3}{2}\frac{\Delta x_{rh}\left(x_i^2 - y_i^2\right)}{r_i^2}\right)$$

where r_i is the distance between the spin and the lanthanide metal, the coordinates of the lanthanide metal are (x_i, y_i, z_i), and the axial and rhombic components of the PCS tensor are Δx_{ax} and Δx_{rh} respectively. The Euler angles α, β, and γ are further used as parameters to orient the tensor with respect to the protein[290].

The partner protein is labeled with NMR-active isotopes (usually ^{15}N is sufficient for the acquisition of ^{1}H-^{15}N HSQC spectra), titrated with equivalent molar amounts of the unlabeled partner but modified with the lanthanide-binding tag. The chemical shifts of the partner protein are then calculated with and without the paramagnetic lanthanide to derive PCS and the tensor is utilized to generate distance restraints. This process can then be repeated with the partner protein, allowing for two sets of intramolecular and intermolecular PCS restraints.

5.3.2.1.1 TGM-1 D3 was modified with a lanthanide-binding tag to generate PCS restraints

To generate a set of intra- and intermolecular PCS restraints, a high affinity lanthanide-binding peptide[289] was inserted into TGM-1 D3. The previously determined interactions between

TGM-1 D3 and TβRII, along with the structure of TGM-1 D3 (Fig. 31, 38), indicated that TGM-1 D3 is likely utilizing the C-terminal half of the non-NC face of the protein, specifically β4 and the residues Arg^{198}, Ile^{238}, Tyr^{253}, and Lys^{254}, to bind TβRII. Thus, we hypothesized that replacing the β3-β4 loop (6 residues) with an LBT tag would not disrupt the overall protein structure nor interfere with the residues involved in binding TβRII, while also bringing the LBT tag somewhat close to the expected binding residues. The construct, called the TGM-1 D3 Loop LBT construct, was cloned into *E. coli* and the protein was expressed and purified in a similar manner as has been previously described for the purification of wild-type TGM-1 D3 (Chapter 3.2).

To calculate the tensor parameters, ^{15}N TGM-1 D3 Loop LBT was expressed and purified in *E. coli*. ^{1}H-^{15}N spectra were recorded for the construct in the presence of equimolar amounts of either diamagnetic lanthanum (La^{3+}) or paramagnetic terbium (Tb^{2+}). Using the amide hydrogen and nitrogen assignments of free TGM-1 D3, the amide hydrogen chemical shifts of select residues of the lanthanum and terbium-bound TGM-1 D3 loop LBT were assigned, and pseudocontact shifts between the two metal-bound forms of TGM-1 D3 loop LBT were calculated per residue as the difference in chemical shift between the terbium and lanthanide-bound forms.

From the ^{1}H-^{15}N HSQC spectra (Fig. 44A), it was apparent that with the addition of terbium there were several peaks that shifted, mostly upfield in both the ^{1}H and ^{15}N axes. As the ^{1}H-^{15}N HSQC spectra of the TGM-1 D3 loop LBT construct was similar to that of the wild-type TGM-1 D3 (Fig. 44B), several of the backbone amide and hydrogen chemical shifts could be assigned, enabling over 40 ^{1}H intramolecular restraints to be used to determine the tensor parameters. These values were input into the program Numbat[280] which fits the anisotropic magnetic susceptibility tensors to the pseudocontact shift data using the protein structure (TGM-1 D3 PDB: 7SXB). The LBT was inserted in the loop between the β3 and β4 strands, deleting residues T^{244}, T^{245}, and G^{246}

in the original sequence followed by a seventeen-residue insertion: GYIDTNNDGWIEGDELY. For tensor calculations, the metal position was restricted to a 6Å sphere about the residues replaced by the LBT tag. The final calculated metal position from the PCS values is distal to the faces of the protein but still distinctly close to the β3-β4 loop (Fig. 45, Table 14).

^{15}N TβRII was titrated with 1.2 molar equivalents of unlabeled TGM-1 D3 loop LBT, and the ^{1}H-^{15}N HSQC spectra of the resultant protein complex in the presence of 1.1 molar equivalents of lanthanum and terbium were collected (Fig. 44C). From the spectra, it was apparent that with the addition of terbium, there were several peaks that shifted, with slight upfield or downfield shifts in both the ^{1}H and ^{15}N axes. As the ^{1}H-^{15}N HSQC spectra of the TGM-1 D3 loop LBT bound TβRII was similar to that of TβRII bound to wild-type TGM-1 D3 (Fig. 44C), several of the backbone amide and hydrogen chemical shifts could be assigned, and so over thirty ^{1}H intramolecular restraints were used as intermolecular restraints in structure calculation using the PCS modules in Xplor-NIH derived from PyPara Tools[291].

5.3.2.1.2 TβRII was modified with a lanthanide-binding tag to generate intermolecular PCS restraints

To generate a second set of intra- and intermolecular PCS restraints, the same high affinity lanthanide-binding peptide[289] was inserted into TβRII. TβRII likely binds to TGM-1 D3 using the edged β-strand (β4) along with Glu142 and Asp55. Thus, without disruption of the overall protein structure nor interference with residues involved in binding TGM-1 D3, while also being somewhat close to the expected binding site, residues D^{102} and E^{103} in the β5-β6 loop were replaced with the LBT: GYIDTNNDGWIEGDELY. The construct, called the TβRII DE LBT construct,

was cloned into *E. coli* and the protein was expressed and purified in a manner similar to wild-type TβRII (Chapter 3.2).

^{15}N TβRII DE was expressed and purified in *E. coli*, and from the ^{1}H-^{15}N HSQC spectrum (Fig. 46A) it was apparent that with the addition of terbium most of the peaks experienced significant chemical shift perturbations compared to the lanthanum-bound form. However, only a few of the backbone amide and hydrogen chemical shifts could be easily assigned due to assignment ambiguities. As such, nine ^{1}H intramolecular restraints were used to determine the tensor parameters for the TβRII DE construct. For tensor calculations, the metal position was restricted to a 6Å sphere about the residues replaced by the LBT tag. The calculated metal position from the PCS values is distal to the faces of the protein, but still distinctly close to the β5-β6 loop where the LBT tag was inserted (Fig. 47, Table 14). For PCS tensor determination, the NMR-derived structure of TβRII was utilized as this structure contains hydrogen atoms as opposed to an X-ray structure of TβRII.

After determining the TβRII DE tensor upon titration with terbium (Tb^{3+}), ^{15}N TGM-1 D3 was titrated with 1.2 molar equivalents of unlabeled TβRII DE, and the ^{1}H-^{15}N HSQC spectra of the resultant protein complex in the presence of 1.1 molar equivalents of lanthanum and terbium were collected (Fig. 46D). From the spectra, it was apparent that with the addition of terbium, there were several peaks that shifted, however the magnitude of the shifts were quite small, making it easy to assign a majority of the backbone amide and hydrogen chemical shifts. Thus over sixty ^{1}H intramolecular restraints were used as intermolecular restraints in structure calculation using the PCS modules in Xplor-NIH derived from PyPara Tools[291].

Along with determining the tensor using the terbium-bound form of TβRII DE, ^{15}N TβRII DE was titrated with thulium and europium, both of which are paramagnetic lanthanides, and ^{15}N

TGM-1 D3 bound to TβRII DE was also titrated with both thulium and europium (Fig. 46 B-C, E-F). However, even though there were measurable shifts in the ^{15}N TβRII DE ^{1}H-^{15}N HSQC spectra upon titration with thulium and europium, the intermolecular PCS with wild-type TGM-1 D3 were hard to detect and thus these PCS were not utilized as restraints.

5.3.2.2 Intermolecular NOEs and residual dipolar couplings restraints were obtained

In addition to the PCS restraints, a few intermolecular NOE distance restraints were used for the complex structure determination. Two separate samples were made: ^{2}H ^{15}N TGM-1 D3:TβRII, and ^{2}H ^{15}N TβRII:TGM-1 D3, both at concentrations of 250 μM in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. NOE distance restraints were gathered using data collected from 3D ^{1}H-^{15}N NOESY spectra and the distances derived using the program CCP-NMR[265]. To assign the aforementioned distance restraints, further NMR spectra were gathered for a sample of ^{15}N ^{13}C TβRII:TGM-1 D3, at a concentration of 400 μM, in a buffer of 25 mM Na_2HPO_4 50 mM NaCl pH 6.0. The spectra gathered include 2D ^{1}H-^{13}C CT-HSQC, 3D HCCH-TOCSY, 3D HBHA(CO)NH, 3D H(CCCO)NH, and 3D CC(CONH). All experiments were gathered at 310K.

Seven intermolecular NOEs were assigned from backbone amide hydrogen atoms of TGM-1 D3 to sidechain hydrogen atoms of TβRII (Table 15). Once the NOEs were assigned, distance restraints were calculated using the isolated spin-pair approximation, which linearizes the relationship between NOE peak intensity and distance[281]. The backbone amide atoms of TGM-1 D3 that were close enough to sidechain hydrogen atoms of TβRII were clearly limited in the complex structure, with two restraints from Ile^{238} of TGM-1 D3 and three restraints from Tyr^{253} of TGM-1 D3 to sidechain hydrogen atoms of Ile^{76} of TβRII, and two restraints from Lys^{254} of TGM-1 D3 to sidechain hydrogen atoms of Asp^{55} and Phe^{53} of TβRII. Distances measured ranged from

3.0 – 5.0 Å. There were other potential intermolecular NOEs identified, but only these eight could be accurately assigned. The residues involved are consistent with the mutagenesis SPR data in Chapter 4, as Ile^{238} and Tyr^{253} were key residues for maintaining the binding affinity between TGM-1 D3 and TβRII. That the intermolecular partner of these residues was Ser^{75} and Ile^{76} of TβRII was equally as expected, as these residues belong to β4 of TβRII, the edged β-strand that TβRII uses to bind TGF-β.

Along with the aforementioned distance restraints orientation restraints were also used for structure calculation, specifically ^{15}N-^{1}H RDCs. Two sets of residual dipolar coupling (RDC) restraints were gathered. These RDCs were measured using a 2D IPAP-HSQC experiment[267], as described previously. Separate samples of ^{2}H ^{15}N TGM-1 D3:TβRII and ^{2}H ^{15}N TβRII:TGM-1 D3 were buffered with 10 mg mL^{-1} Pf1 phage[266] to determine the anisotropic J_{app}. Similarly, separate samples of ^{1}H ^{15}N TGM-1 D3:TβRII and ^{1}H ^{15}N TβRII:TGM-1 D3 I84A were buffered into 25 mM Na_2HPO_4 50 mM NaCl pH 6.0 to determine the isotropic J_{app}. The difference between the two per residue was taken as the RDCs, with over 50 RDCs used for the final structure calculation.

5.3.3 The structure of the TGM-1 D3:TβRII complex demonstrates how TGM-1 D3 acts as a structural mimic of TGF-β

5.3.3.1 Structure of the TGM-1 D3:TβRII Complex

Structure calculations of the TGM-1 D3:TβRII complex from XPLOR-NIH[263] yielded an ensemble of 10 lowest energy structures, five of which are plotted in Fig. 48. Of the ten lowest ensemble structures, nine converged to a common fold while one deviated slightly, though with the same general orientation of the two proteins at the interface, likely due to the fewer NOE

distance restraints utilized for structure determination. The backbone root-mean-square deviation (RMSD) for the ensemble of nine converging lowest energy structures was calculated as 0.46 Å when aligned to secondary structure regions, or 0.84 Å when aligned over the entire core region, while the heavy atoms RMSD was calculated as 2.83 Å when aligned to secondary structure regions and 2.49 Å when aligned over the entire core (Table 16). Experimental versus calculated RDC values and PCS values for the lowest-energy structure were plotted as well (Fig. 49), with major outliers noted. While there were outliers, the general trend for this singular structure yields a correlation between the experimentally derived RDC and PCS values and the XPLOR-NIH calculated values.

Per the structure calculation method, the backbone structures of TGM-1 D3 and TβRII in the complex did not differ from the free solution structures of the individual proteins while the side chains were allowed to vary. From the structures there are three distinct interactions that form the composite interface of the TGM-1 D3:TβRII structure: 1) Arg^{198} of TGM-1 D3 and Lys^{79} and Pro^{80} of TβRII, 2) Ile^{238} and Tyr^{253} of TGM-1 D3 and Ile^{76} of TβRII, and 3) Lys^{254} of TGM-1 D3 and Asp^{55} and Phe^{53} of TβRII (Fig. 50A-D). One thing to note is that there were limited restraints constraining the sidechain orientations. The only restraints that constrained sidechain residues were the H^N NOEs from TGM-1 D3 that correlated to sidechain hydrogen residues of TβRII. Thus, it was up to the NOE and repulsion terms to govern sidechain orientation. From this, there are a few residues within the interface that varied slightly in position amongst low energy conformers, specifically Tyr^{253} and Arg^{198} of TGM-1 D3 and Glu^{78}, Lys^{79}, and Gln^{81} of TβRII.

Mutagenesis of Arg^{198} of TGM-1 D3 had been shown to reduce the binding affinity by over 50-fold (Table 12) and thus it had been hypothesized to form a hydrogen bond with Glu^{142} of TβRII. Surprisingly this was not the case, with TGM-1 D3 Arg^{198} interfacing with Lys^{79} and Pro^{80}

of TβRII (Fig. 50B). The stretch of residues of TβRII is Glu^{78}-Lys^{79}-Pro^{80}-Gln^{81}. Ideally, Arg^{198} would interact with Glu^{78}, electrostatically and with a hydrogen bond. However, the distance between these two residues is too great, making Arg^{198} much closer to Lys^{79}-Pro^{80}-Gln^{81}. As a result, the sidechain of Arg^{198} is placed further from the sidechain of Lys^{79} to avoid electrostatic repulsion, and thus forming hydrophobic interactions with Pro^{80}. In the ten conformers, the locations of the Arg^{198} and Glu^{78}-Lys^{79}-Pro^{80}-Gln^{81} sidechains vary slightly, but still follow the pattern of minimizing electrostatic repulsion by similarly charged sidechains.

The second set of interactions is between Ile^{238} and Tyr^{253} of TGM-1 D3 and Ile^{76} of TβRII. From the lowest energy structure, it is clear that Ile^{76} of TβRII inserts itself into the hydrophobic core of TGM-1 D3, bracketed by Ile^{238} of TGM-1 D3 and Tyr^{253} of TGM-1 D3. However, while the sidechain position of Ile^{238} of TGM-1 D3 remains constant amongst conformers, the position of Tyr^{253} of TGM-1 D3 does not. In other conformers, Tyr^{253} points away from the two Ile residues but towards other residues, interacting with Glu^{78} of TβRII and Phe^{53} of TβRII, belonging to the first and third interface respectively. This positional variation suggests that either there are simply not sufficient constraints in the calculations to determine the sidechain orientation of Tyr^{253} or that Tyr^{253} acts transiently to support multiple binding interactions.

The final set of interactions is between Lys^{254} of TGM-1 D3 and Asp^{55} and Phe^{53} of TβRII. The proximity of Lys^{254} and Asp^{55}, along with the dramatic impact of mutagenesis of either residue on the binding of TGM-1 D3 to TβRII suggests that there is a hydrogen bond or electrostatic interaction between the two sidechains. In two of the ten lowest-energy conformers, the distance between the two sidechains is appropriate for both a hydrogen bond and a salt bridge, with another conformer being the appropriate length for a salt bridge but not a hydrogen bond. Thus though it is not definitive that there is a hydrogen bond or electrostatic interaction between these two

residues, due to the drastic impact of mutagenesis on binding, along with this initial structural data, this interaction is likely to occur as well. In addition, the sidechain of Phe^{53} is close to that of Lys^{254} and likely stabilizes the residue via hydrophobic interactions with the carbon sidechain of Lys^{254} while the charged nitrogen interacts with Asp^{55}.

5.3.3.2 Comparison of the TGM-1 D3:TβRII and TGF-β:TβRII binding interfaces

As shown in Figure 28, there are three key interactions forming the composite interface utilized by TGF-β in order to bind TβRII: 1) Asp^{55} of TβRII and ^{Arg394} of TGF-β, 2) Ile^{76} of TβRII and Tyr^{390}, Trp^{332}, and Trp^{330} of TGF-β, and 3) Glu^{142} of TβRII and Arg^{324} of TGF-β. TGF-β forms two hydrogen-bonded ion pairs with TβRII, through the Asp^{55}-Arg^{394} pair and the Glu^{142}-Arg^{325} pair (Fig. 51). This is both similar and distinct from the binding of TGM-1 D3 to TβRII as in the case of the latter, TGM-1 D3 only forms a potential bond with Asp^{55} of TβRII and does not seem to interact with Glu^{142} of TβRII. This may be in part why the affinity of TGM-1 D3 for TβRII is weaker than the affinity of TGF-β for TβRII. What the two interfaces have in common is the insertion of Ile^{76} into a hydrophobic region on the ligand, being Val^{392},Tyr^{390},Trp^{3330}, and Trp^{332} of TGF-β and Ile^{238} and Tyr^{253} of TGM-1 D3 (Fig. 51). Additionally, there is a minor interface with the TGM-1 D3:TβRII complex that is not recapitulated in the TGF-β:TβRII complex, that of Arg^{198} of TGM-1 D3 and Glu^{78}-Lys^{79}-Pro^{80}-Gln^{81} of TβRII. Due to the charged nature of this segment, the sidechain of Arg^{198} is guided due to steric hindrance and charge repulsion towards Pro^{80}, likely forming hydrophobic interactions with Pro^{80} and with the carbon sidechains of Lys^{79} and Gln^{81}, generating an additional binding interface to further stabilize the complex.

5.3.3.3 Mutagenesis of key residues of TβRII is in agreement with structural information

The plot of the composite shift per residue for ^{15}N ^{13}C TβRII was shown in the previous chapter (Fig. 28A). In particular, on a scale normalized to 1.0, only one region showed significant chemical shift perturbation, including residues Thr^{74} – Ile^{76}. This corroborates with the structural data of the TGM-1 D3:TβRII complex. Along with this TGF-β also interacts with TβRII Asp^{55} and Glu^{142}, forming two hydrogen-bonded ion pairs with these residues. To assess the roles of these residues on the TβRII:TGM-1 D3 interface, binding of mutated constructs were assessed using SPR. Biotinylated avi-tagged TGM-1 D3 was captured on a streptavidin-conjugated CM5 SPR chip, and mutated TβRII constructs were injected over the surface. Five mutants were tested: $D^{55}N$, $I^{73}A$, $S^{75}L$, $I^{76}A$, and $E^{142}Q$. The K_D values were derived by fitting the sensorgrams to a 1:1 kinetic model (Table 12, Fig. 40). All of the mutants showed attenuated responses with respect to wild-type TβRII, but were still fit, indicating a weaker binding affinity. The $Asp^{55}Asn$, $Ser^{75}Leu$, and the $Ile^{76}Ala$ mutants showed a weaker-affinity K_D by 20-fold or more, at a 60, 310, and 25 μM binding affinity respectively, in comparison to 1.5 μM for wild type. This, along with the structural data, indicates that disruption to the interactions between TβRII's edged β-strand and the hydrophobic cleft of TGM-1 D3 formed by the hydrophobic residues Ile^{238} and Tyr^{253} will disrupt the binding interface. Additionally, the significant impact of the $Asp^{55}Asn$ indicates the importance of the Asp^{55}-Lys^{254} interaction, which is likely so significant for the TGM-1 D3:TβRII interaction to make up for the lack of interaction with the Glu^{142} residue of TβRII.

Mutagenesis of $Ile^{73}Ala$ and $Glu^{142}Gln$ were less significant, with a four-fold difference and a fifteen-fold difference in affinity respectively. Surprising was the impact of the $Glu^{142}Gln$ mutant, as the NMR structure did not indicate an interaction between TGM-1 D3 and TβRII Glu^{142}. It may be that there is a transient interaction that was overlooked due to the lack of intermolecular

NOE restraints used in the structure calculation, or that the mutation of $Glu^{142}Gln$ may cause a charged interaction that would disrupt nearby interactions with TGM-1 D3. Overall, the data corroborates the structural information, highlighting the importance of the edge β-strand and of the Asp^{55}-Lys^{254} interaction. To gain a more thorough understanding of the functional impacts of TβRII residues, further SPR experiments may be performed in the future, particularly assessing the role of the N-terminal segments of TβRII and the region immediately C-terminal to the edge β-strand in interacting with the N-terminal half of the hypervariable loop of TGM-1 D3.

5.4 Discussion

The previous chapter noted distinct differences between the structure of TGM-1 D3 and that of the canonical CCP domain, mainly the extension of the hypervariable loop of the protein by over ten residues, and the lateral expansion of the C-terminal half of TGM-1 D3. Canonical CCP domains contain a minimum of 6 β-strands in two sheets, two of which cluster in a separate sheet at the C-terminal end of the protein. These two β-strands are not present in TGM-1 D3, which allows for a lateral expansion of the C-terminal half of the protein away from the N-terminal half. This leads to hydrophobic residues on both faces of TGM-1 D3 being partially exposed to solvent. It had thus been hypothesized, along with the inclusion of SPR data, that these specific adaptations are the reason why TGM-1 D3 can bind TβRII. The lateral expansion would allow for the edge β-strand of TβRII to contact the hydrophobic cleft of the nonNC face.

The goal of this chapter was thus to determine the interacting interfaces between TGM-1 D3 and TβRII and compare that to those between TGF-β and TβRII. The binding of TGM-1 D3

to TβRII is both similar and distinct from the binding of TGF-β to TβRII. TGM-1 D3 interfaces with TβRII using three interactions, one involving Lys^{79} and Pro^{80} of TβRII, another involving Ile^{76} of TβRII, and a final interaction including Asp^{55} and Phe^{53} of TβRII. TGF-β also shares with the latter two interaction but not the former, while having an additional interaction with Glu^{142} of TβRII that TGM-1 D3 lacks. While there are some slight differences, the sidechain of Ile^{76} of TβRII inserts into a hydrophobic pocket formed by Ile^{238} and Tyr^{253} of TGM-1 D3 in a similar manner as this same sidechain would insert between the hydrophobic fingers of TGF-β. Similarly, as TGF-β would form a hydrogen-bonded ion pair with Asp^{55} of TβRII, TGM-1 D3 Lys^{254} likely forms a salt bridge with Asp^{55} that is stabilized by the hydrophobic sidechain of Phe^{53}. In addition, Tyr^{253} likely acts to stabilize the Lys^{254}-Asp^{55} interaction, thereby serving a dual role. Thus the structure of the TGM-1 D3:TβRII complex highlights the structural mimicry of the TGM proteins. Though there are some minor differences in how TGM-1 D3 and TGF-β bind to TβRII, the main binding interface, the insertion o the edge β-strand of TβRII into a hydrophobic cleft, remains the same, indicating that TGM-1 D3 has evolutionarily adapted to binding TβRII, which is supported by mutagenesis data.

5.5 Respective Contributions

Most experiments in this chapter, and all analyses of these experiments were performed in their entirety by the author Ananya Mukundan, with the exception of the RDC experiments and some of the 3D experiments. 2H ^{15}N protein samples for the 3D NOESY experiments were prepared and spectra recorded by Chang-Hyeock Byeon, and samples for the HCCH-TOCSY

experiment were prepared by Chang-Hyeock Byeon and the spectra recorded by Andrew Hinck. 2H ^{15}N protein samples for the anisotropic RDC experiments were produced by Chang Hyeock-Byeon and recorded by Andrew Hinck, and subsequent analyses performed by Ananya Mukundan. Samples for the isotropic RDC experiments were produced, recorded, and analyzed by Ananya Mukundan.

5.6 Tables and Figures

Table 13: Sequences of lanthanide-binding tag domain constructs, and truncated TβRII constructs.

Construct	Coding region and description (* indicates stop codon)
TGM-1 D3 Loop Lanthanide Binding Tag	Residues 177-262 of *H. polygyrus* TGF-β Mimic, with residues 243-246 substituted with a lanthanide-binding tag, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-D3-Lanthanide Binding Tag-TGM-D3 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTGCPPLPDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRRCRALSQEADPGEFV AICYKSGYIDTNNDGWIEGDELYESHWEYYKNIGKCPDP*
TβRII DE Lanthanide Binding Tag	Residues 38-154 of the extracellular domain of human TGF-β type II receptor, with residues 102-103 substituted with a lanthanide-binding tag, NCBI NP_001394056 Initiating Methionine-TβRII-Lanthanide Binding Tag-TβRII MVTDNNGAVKFPQLCKFCDVRFSTCDNQKSCMSNCSITSICEKPQEVCVAVWRKNGYIDTNNDG WIEGDELYNITLETVCHDPKLPYHDFILEDAASPKCIMKEKKKPGETFFMCSCSSDECNDNIIFSEE YN*
TβRII 46-153	Residues 46-153 of the extracellular domain of human TGF-β type II receptor, NCBI NP_001394056 Initiating Methionine-TβRII MKFPQLCKFCDVRFSTCDNQKSCMSNCSITSICEKPQEVCVAVWRKNDENITLETVCHDPKLPYH DFILEDAASPKCIMKEKKKPGETFFMCSCSSDECNDNIIFSEEY

Table 14: TGM-1 D3:TβRII Complex LBT tensor parameters.

TGM-1 D3 Loop LBT	
Δx_{ax} (10^{-32} m^3)	6.63
Δx_{rh} (10^{-32} m^3)	3.98
Xi	31.84
Y_i	23.02
Z_i	-11.15
α (°)	26.89
β (°)	153.96
γ (°)	105.08
TβRII DE LBT	
Δx_{ax} (10^{-32} m^3)	21.25
Δx_{rh} (10^{-32} m^3)	16.71
Xi	-15.98
Y_i	-9.36
Z_i	20.00
α (°)	323.61
β (°)	12.71
γ (°)	8.42

Table 15: TGM-1 D3:TβRII Complex NOE restraints.

TGM-1 D3	TβRII	Distance	Upper Bound	Lower Bound
Ile 238 HN	Ile 76 HG2	5.4	3.6	1.3
Ile 238 HN	Ile 76 HD1	4.1	2.3	1.3
Tyr 253 HN	Ile 76 HG1	4.2	2.4	1.3
Tyr 253 HN	Ile 76 HG2	4.6	2.8	1.3
Tyr 253 HN	Ile 76 HD1	4.9	3.1	1.3
Lys 254 HN	Asp 55 HB	3.6	1.8	0.3
Lys 254 HN	Phe 53 HB	3.4	1.6	0.3

Table 16: TGM-1 D3:TβRII Complex NMR structural statistics.

NOE	
Intermolecular NOE restraints	7
PCS	
D3 Loop (Intramolecular, Intermolecular)	56
TβRII DE (Intramolecular, Intermolecular)	70
RDC	
RDC: N-H TGM-1 D3	65
RDC: N-H TβRII	24
RMSD (Deviations)	
Bonds (Å)	0.012 ± 0.003
Impropers (°)	5.188 ± 1.059
Angles (°)	2.149 ± 0.282
Ramachandran[a]	
Most Favored	78.0%
Additionally Allowed	16.2%
Generously Allowed	2.9%
Disallowed	2.9%
RMSD[b]	
Secondary Structure[c]	
Backbone (Cα, N, Cβ)	0.46Å
Heavy	2.83Å
Core[d]	
Backbone	0.84Å
Heavy	2.49Å

[a]Ramachandran values from the lowest-energy structure
[b]RMSD values are computed from the lowest-energy structure
[c]Residues TGM-1 D3: 189-93, 217-221, 228-230, 234-241, 248-252, TβRII: 50-52, 55-58, 74-76, 83-89, 95-103
[d]Residues TGM-1 D3: 178-253, TβRII: 46-148

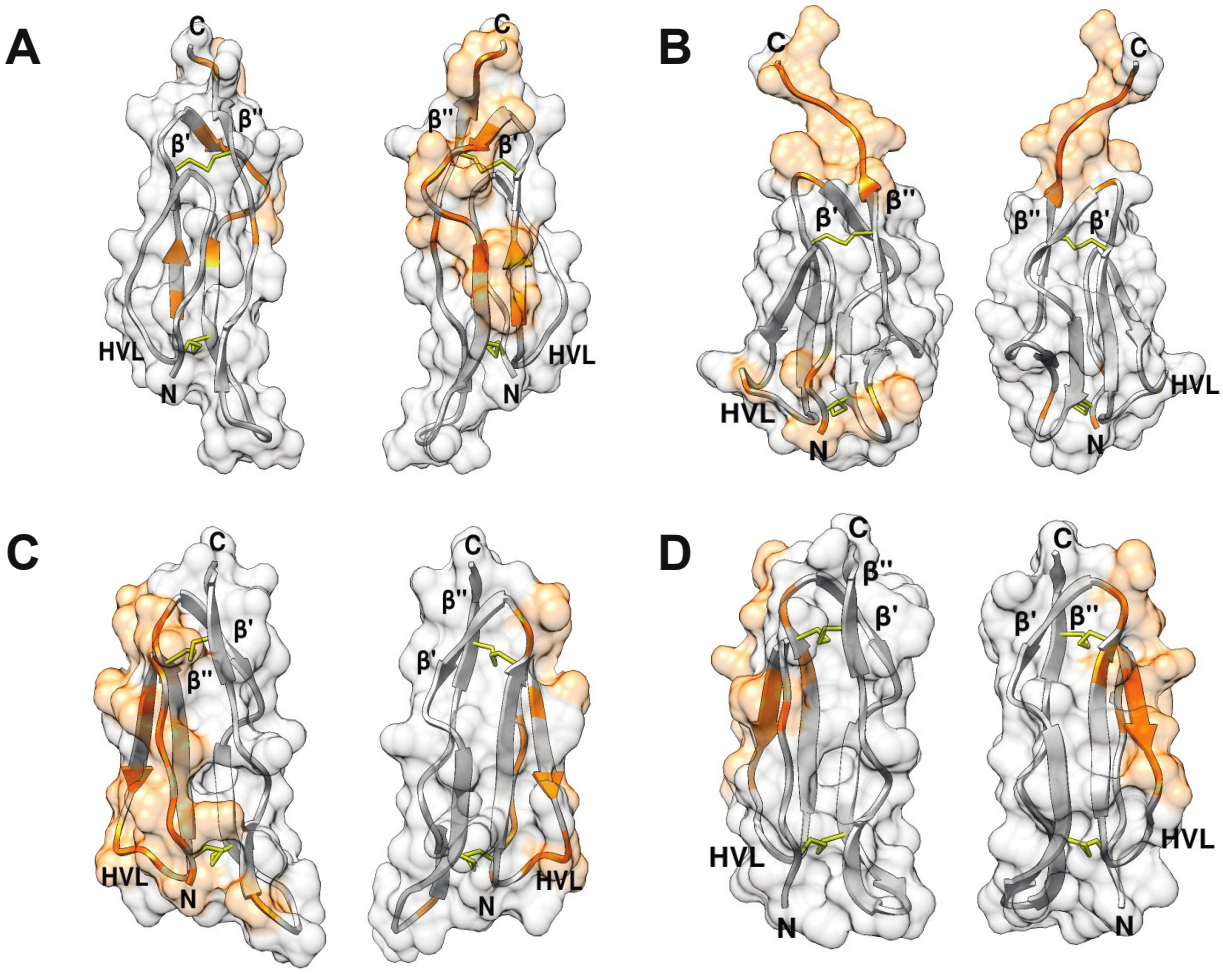

Figure 42: CCP domain binding interfaces. A-D. Structures of CCP domains (IL-15, PDB: 2psm (A). MASP, PDB:4fxg (B). CR1, PDB: 5fo9 (C). CD55, PDB: 6la5 (D)) with the NC (left) or nonNC (right) faces forward with amino acids involved in binding surfaces colored by threshold of dSAS > 20Å^2.

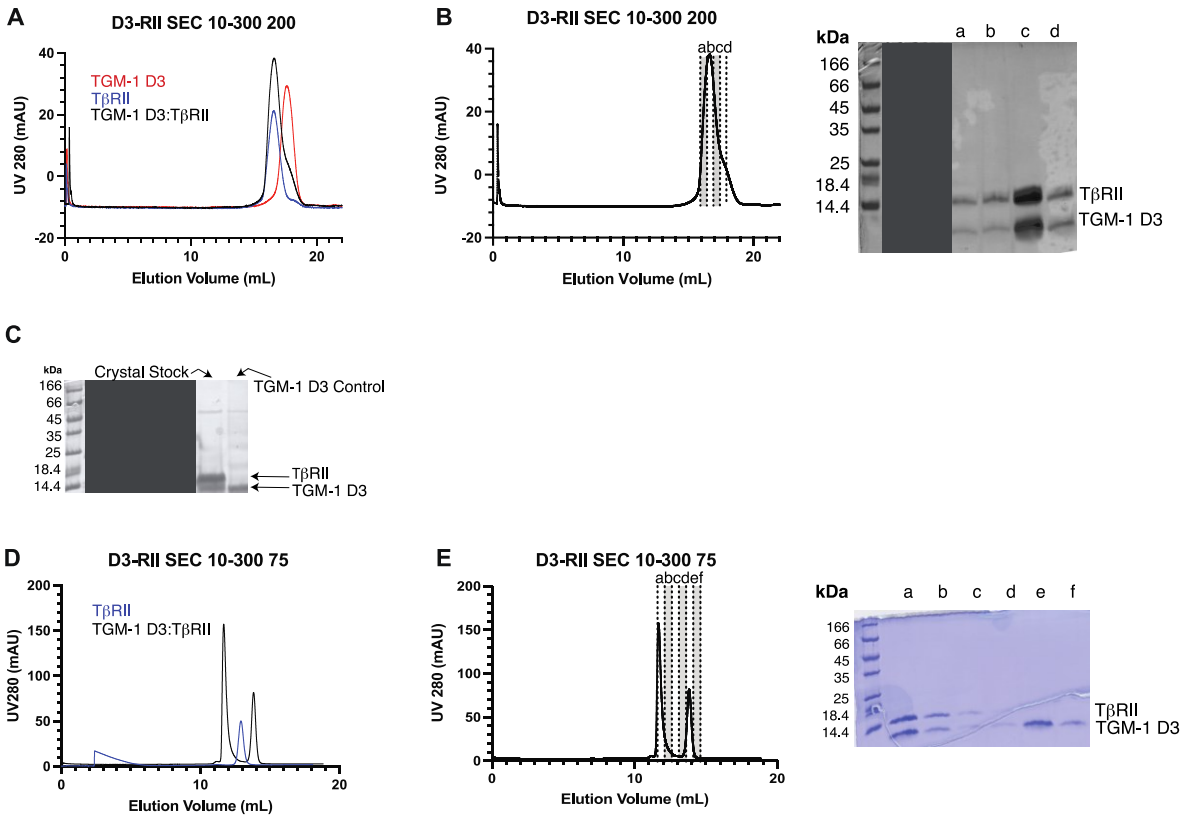

Figure 43: Isolation of the TGM-1 D3:TβRII complex by analytical SEC. A-E. A. SEC chromatogram of TGM-1 D3 (red), TβRII (blue), and the TGM-1 D3:TβRII complex (black) in a 1.1:1 titration on a 10-300 200 column. Left axis indicates absorbance at 280 nM, which is plotted with respect to elution volume. B. SEC chromatogram (left) of the TGM-1 D3:TβRII complex on a 10-300 200 column with dashed lines indicating fractions of the chromatogram assessed via SDS gel (right), labeled as fractions a-d. C. SDS gel confirming composition of crystallization stock, with TGM-1 D3 control to the right. D. SEC chromatogram of TβRII (blue), and the TGM-1 D3:TβRII complex (black) on a 10-300 75 column. Left axis indicates absorbance at 280 nM, which is plotted with respect to elution volume. E. SEC chromatogram (left) of the TGM-1 D3:TβRII complex (black) with an excess of TGM-1 D3 on a 10-300 75 column with dashed lines indicating fractions of the chromatogram assessed via SDS gel (right) labeled as fractions a-f.

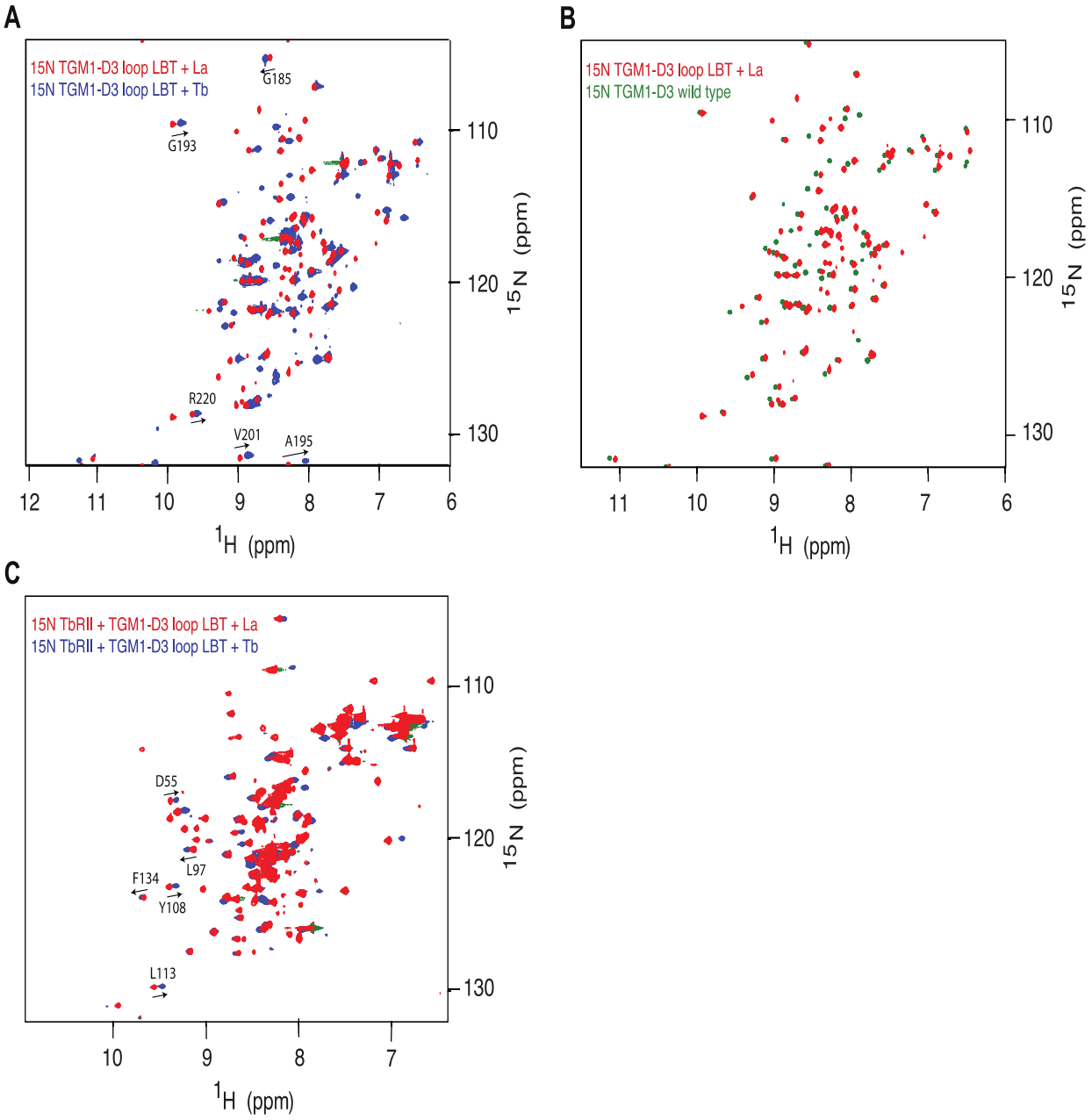

Figure 44: Addition of metals to TGM-1 D3 loop LBT construct demonstrates measurable pseudocontact shifts in ^{1}H-^{15}N HSQC spectra. A-C. A-B. ^{1}H-^{15}N spectra of ^{15}N TGM-1 D3 loop LBT construct with 1.1 molar equivalents of lanthanum (La^{3+}) (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.1 molar equivalents of terbium (Tb^{2+}) (blue, A) or overlaid with the spectra of wild-type TGM-1 D3 (green, B). C. ^{1}H-^{15}N spectra of ^{15}N TβRII bound to 1.2 molar equivalents of TGM-1 D3 loop LBT with 1.1 molar equivalents of lanthanum (La^{3+}) (red) overlaid with the ^{1}H-^{15}N spectra of the same protein complex bound to 1.1 molar equivalents of terbium (Tb^{2+}) (blue). Arrows in A and C indicate pseudocontact shifts of marked residues. All spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K.

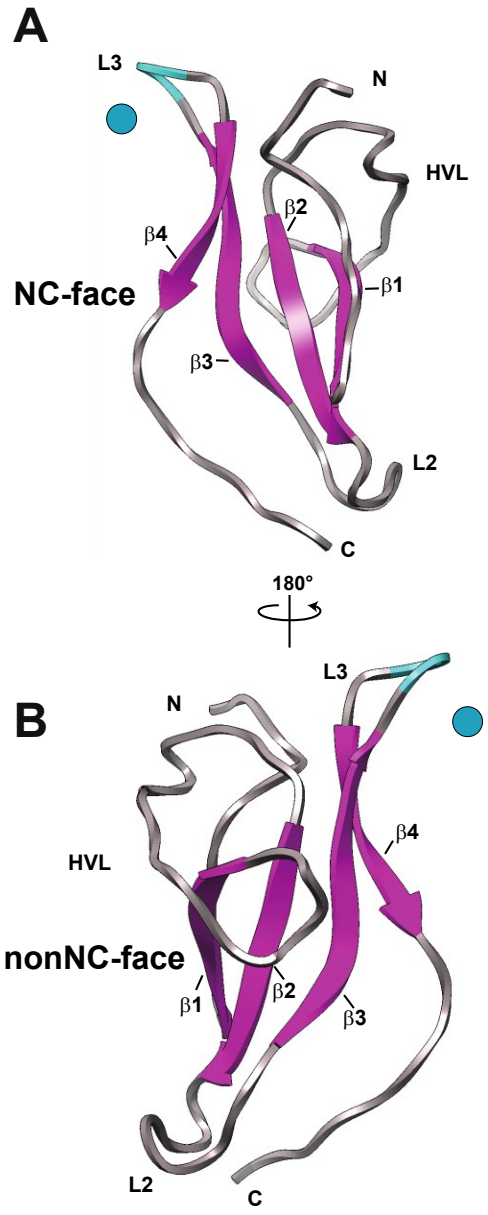

Figure 45: NMR structure of TGM-1 D3 with position of metal indicated. A-B. Structure of unbound TGM-1 D3, with appropriate secondary structures indicated: β-strands, pink; loops, gray, lanthanide metal, light blue, residues of TGM-1 D3 replaced by lanthanide binding tag, light blue. Orientations are shown with respect to a 180 rotation about the y-axis. The orientation containing both the N- and C-termini, the NC face is shown in (A), with a 180 rotation to the nonNC face shown in (B).

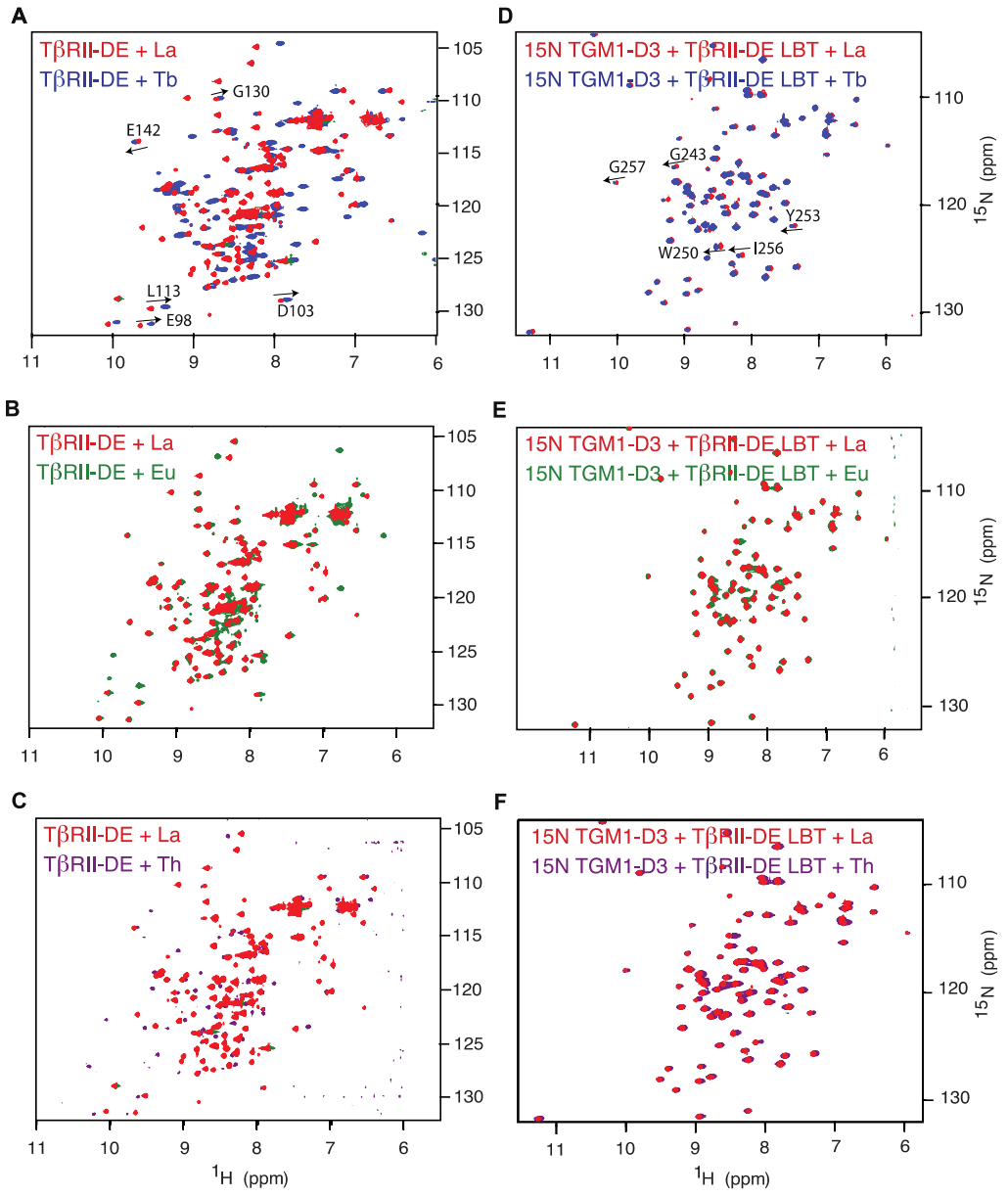

Figure 46: Addition of metals to TβRII DE LBT construct demonstrates measurable pseudocontact shifts in ^{1}H-^{15}N HSQC spectra. A-F. A-C. ^{1}H-^{15}N spectra of ^{15}N TβRII DE loop LBT construct with 1.1 molar equivalents of lanthanum (La^{3+}) (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.1 molar equivalents of terbium (Tb^{2+}) (blue, A) europium (green, B), or thulium (purple, C). D-F. ^{1}H-^{15}N spectra of ^{15}N TβRII bound to 1.2 molar equivalents of TβRII DE LBT with 1.1 molar equivalents of lanthanum (La^{3+}) (red) overlaid with the ^{1}H-^{15}N spectra of the same protein complex bound to 1.1 molar equivalents of terbium (Tb^{2+}) (blue, D), europium (green, E), or thulium (purple, F). Arrows in A and C indicate pseudocontact shifts of marked residues. All spectra were recorded in 25mM MES, 5% $^{2}H_2O$, pH 6.0, 310K.

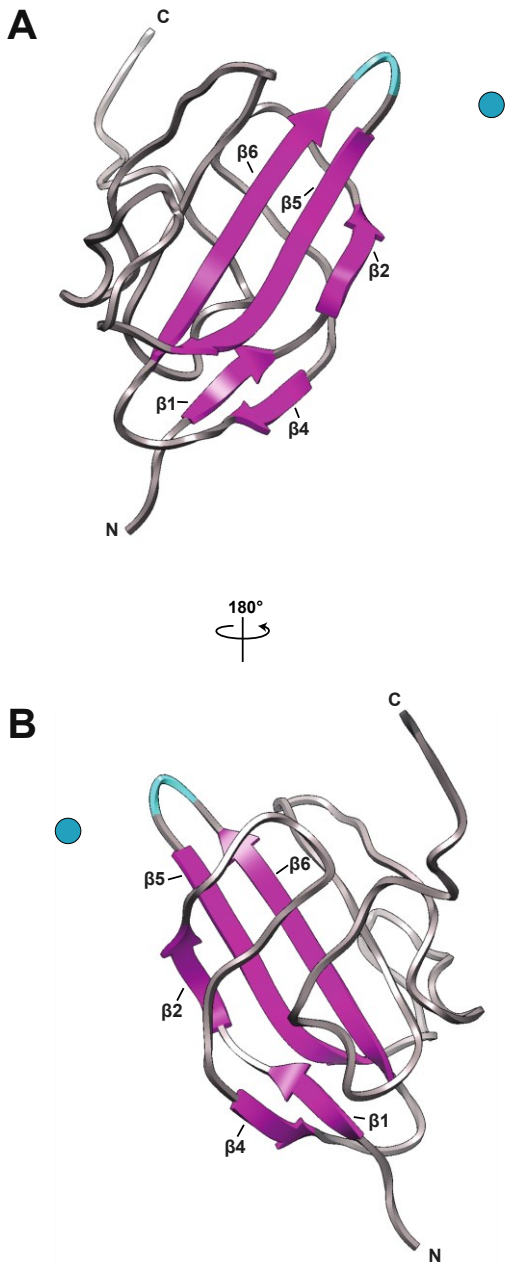

Figure 47: NMR structure of TβRII with position of metal indicated. A-B. Structure of unbound TβRII (PDB 1PLO), with appropriate secondary structures indicated: β-strands, pink; loops, gray, lanthanide metal, light blue, residues of TβRII replaced by lanthanide binding tag, light blue. Orientations are shown with respect to a 180 rotation about the y-axis.

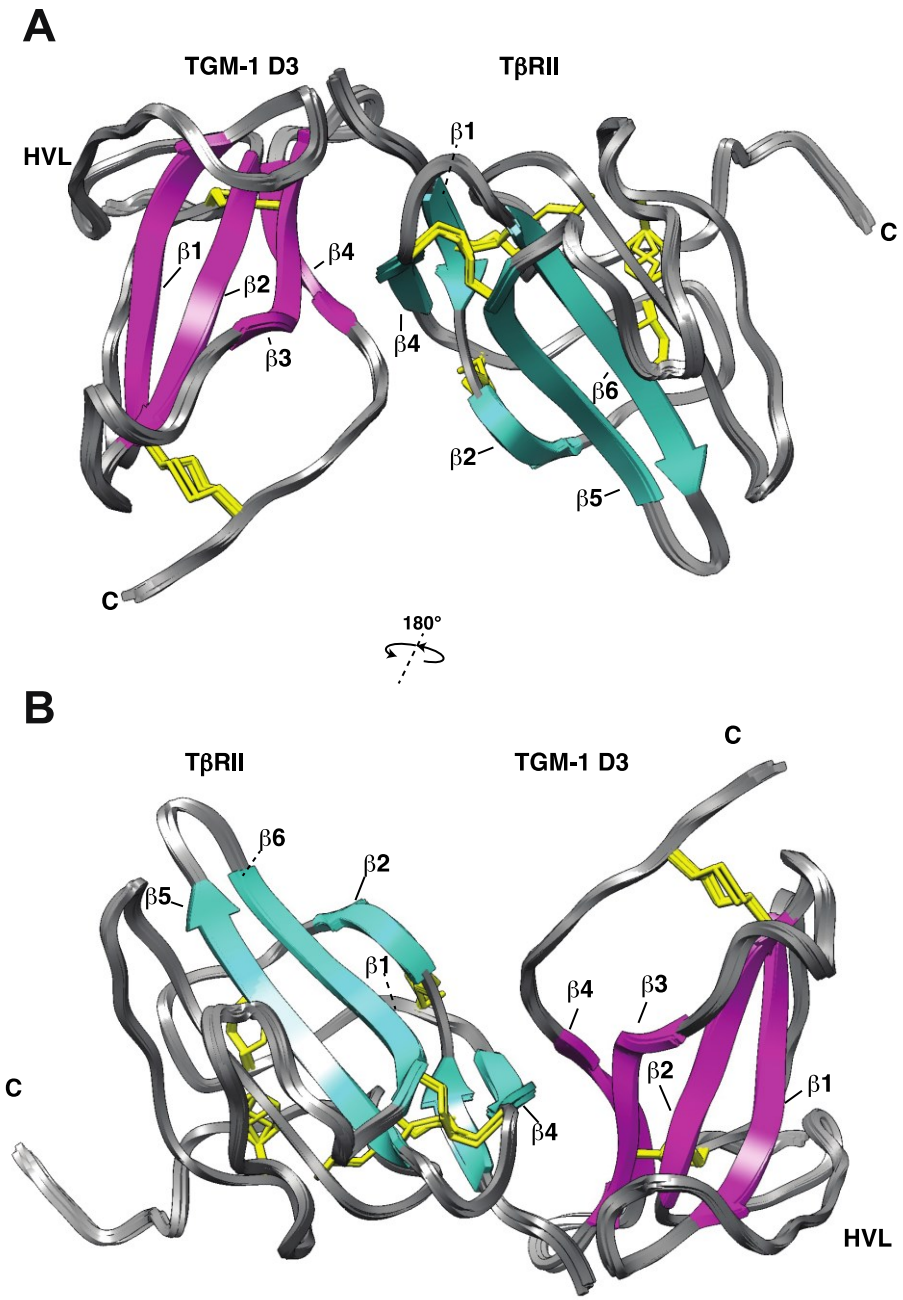

Figure 48: Ensemble of solution NMR structures of the TGM-1 D3:TβRII Complex. A-B. Ensemble of the five lowest-energy conformers of the TGM-1 D3:TβRII (labeled) complex, with appropriate secondary structures indicated. Orientations are shown with respect to a 180 rotation about the z-axis.

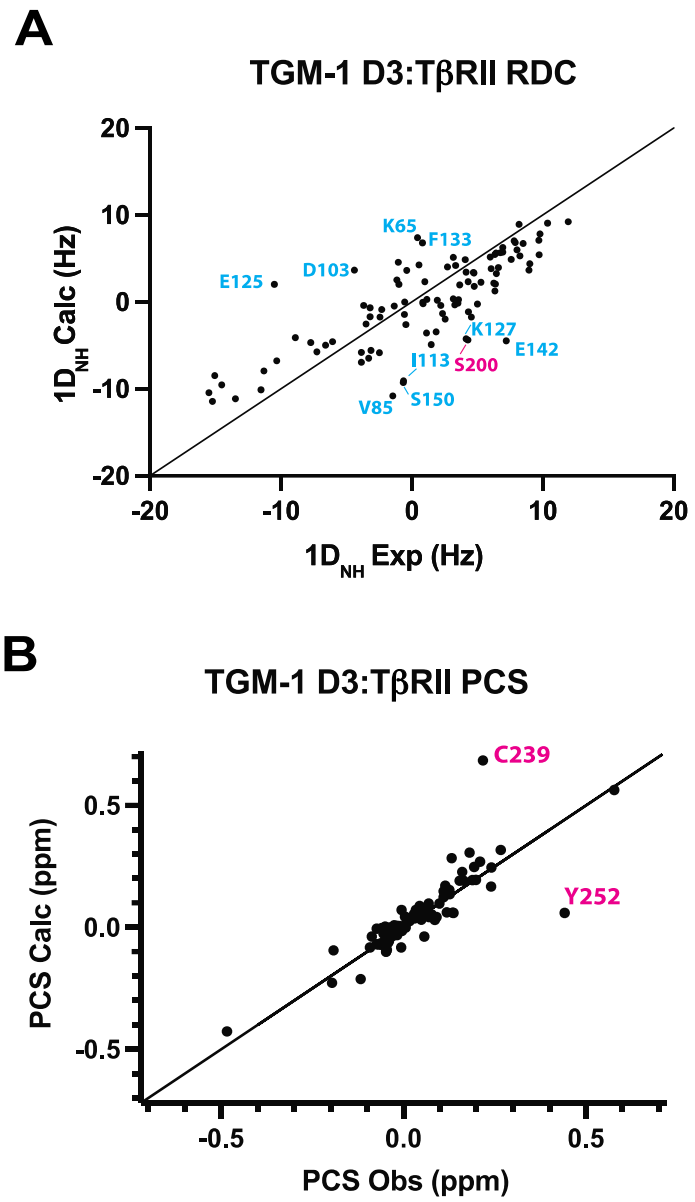

Figure 49: Plots of experimental vs. calculated RDCs and PCS values. A. Experimental v. Calculated NH RDC values for the lowest energy structure generated from the top 10 energetically favored structures were plotted, with outlier residues noted in blue (TβRII) or magenta (TGM-1 D3). B. Experimental v. Calculated PCS values for the lowest energy structure generated from the top 10 energetically favored structures were plotted, with outlier residues noted in magenta (TGM-1 D3 loop LBT construct) or blue (TβRII DE LBT construct).

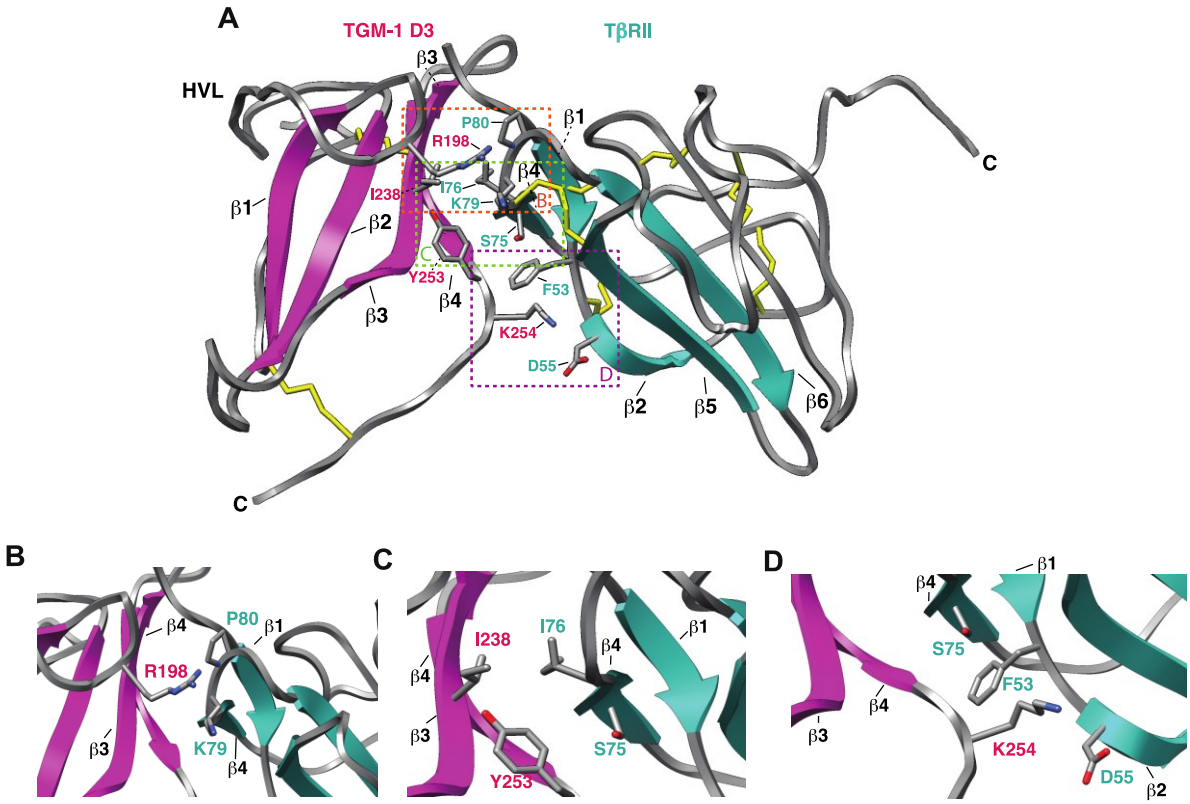

Figure 50: Structure of the TGM-1 D3:TβRII Complex. A-D. A. Lowest-energy conformer of the TGM-1 D3 (magenta) :TβRII (blue) complex, with appropriate secondary structures indicated. Dotted boxes indicate regions that are expanded on in B (orange) C (green) D (purple). Labels for residues of TGM-1 D3 are colored magenta while labels for residues of TβRII are colored blue.

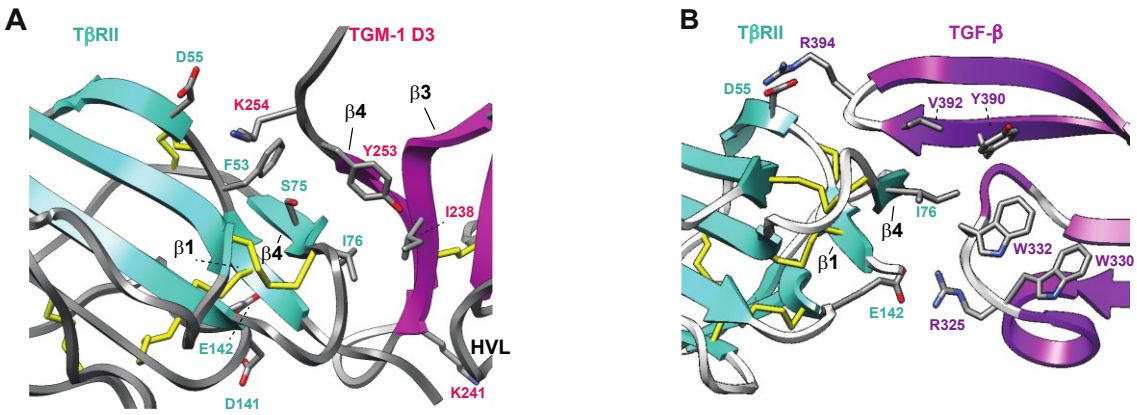

Figure 51: Comparison of the TGM-1 D3:TβRII and TGF-β:TβRII Complexes. A. Lowest-energy conformer of the TGM-1 D3 (magenta) :TβRII (blue) complex, with appropriate structural features indicated. B. Structure of the TGF-β (purple): TβRII (blue) complex, with appropriate structural features indicated.

6.0 TGM-1 D12 binding to TβRI

6.1 Introduction

The binding of TGM-1 to the TGF-β receptors is modular and separable, i.e. domains 1 and 2 are responsible for binding the type I receptor through a shared binding interface. Though it is known that the TGM-1 D3 contacts the same residues of TβRII as TGF-β, it is unknown whether TGM-1 is contacting the same residues of TβRI as TGF-β. If so, TGM-1 D12 could act as a competitive inhibitor of TGF-β, which would make TGM-1 of high interest towards development of a TβRI-specific inhibitor.

The details of the TGM-1 D12:TβRI structure would give insight into the details of TGF-β receptor specificity and show how the parasite has adapted to bind specific TGF-β receptors. This chapter looks to investigate the following questions: 1) do TGM-1 D12 and TGF-β:TβRII compete for binding to TβRI, 2) if so, are TGM-1 D12 and TGF-β:TβRII binding to similar or different sites on TβRI, 3) how has the structure of the TGM-1 D12 adapted from the canonical CCP domain to enable binding of TGM-1 D12 to TβRI over TβRII, and 4) how does the structure of the TGM-1 D12:TβRI complex differ from that of the TGF-β:TβRII complex.

6.2 Methods

Cloning, expression, purification, and assembly of the TGF-β3:TβRII$_2$ binary complex. DNA inserts coding for TGF-β3 were inserted into a modified pET32a vector between the NdeI and HindII sites (Table 17). Constructs were expressed in BL21 (DE3) cells and cultured at 37 °C. When the light scattering at 600 nm was 0.8, protein expression was induced by adding 0.8 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Purification of TGF-β from LB was performed in a similar manner as the purification of mmTGFβ27M (Chapter 4.2). Expression and purification of TβRII was performed as described in previous chapters (Chapter 3.2). Complex formation was performed by adding two molar equivalents of monomeric TβRII for every one molar equivalent of homodimeric TGF-β3.

Cloning, expression, and purification of TGM-1 D123, TGM-1 D12, TGM-1 D2, and TβRI. Expression and purification of TGM-1 D123, TGM-1 D12, TGM-1 D2, and TβRI were performed as described in previous chapters (Chapter 3.2).

ITC Competition Experiments. All ITC experiments were performed with a Microcal PEAQ-ITC system. Cell and syringe conditions are indicated in Table 8. All experiments were performed in 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 at 25 °C (for the TβRI, TGM-D12, TGF-β:TβRII experiment) on 30°C (for the TβRI, TGF-β:TβRII experiment) with 13 3.0 μL injections with a duration of 5 s, a spacing of 150 s, and a reference power of 10. All samples were dialyzed against the same ITC buffer before loading into the system. Data was globally fit using the programs NITPIC[246], SEDPHAT[247,248], and GUSSI[249].

NMR backbone assignment. Samples were comprised of ^{15}N ^{13}C TβRI:TGM-1 D2 in a 1:1.2 molar ratio at a concentration of ca. 200 μM in buffer 25 mM HEPES, 50 mM NaCl, 0.02%

sodium azide, 5% $^{2}H_2O$ pH 6.0, and all spectra recorded at 300K. Each sample was run in a 5 mm susceptibility-matched microtube for data collection. NMR data was collected with Bruker 600, 700, and 800 MHz spectrometers containing a 5 mm ^{1}H (^{13}C, ^{15}N) z-gradient "TCT" cryogenically cooled probe. Assignments of the TβRI backbone residues as bound to TGM-1 D2 were made by recording, processing, and analyzing the following triple-resonance spectra: 3D HNCACB, 3D CBCA(CO)NH, 3D HNCA, 3D HN(CO)CA, 3D HNCO, and 3D HN(CA)CO, along with the 2D ^{1}H-^{15}N HSQC spectrum. ^{1}H-^{15}N HSQC spectra were acquired as described, with water flipback[250] and WATERGATE suppression pulses[251]. Assignment was performed semi-automatically using PINE[259,260] and then manually corrected using NMRFAM-SPARKY[254] and PECAN[261].

NMR chemical shift perturbation calculations. The backbone residues of TβRI were assigned in the TGM-1 D2-bound form. The magnitude of the difference in chemical shift value between the unbound and bound forms of TβRI was calculated for the five backbone nuclei: $^{15}N^{H}$, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C^{O}$, and normalized to the largest chemical shift perturbation value per nuclei. These normalized values were summed up for all of the nuclei per residue and then normalized once more by dividing the number of nuclei contributing to the total chemical shift perturbation.

NMR solvent accessible surface area calculations. Solvent accessible surface area was calculated for the structure of the extracellular domain of TβRI using PDB 2PJY using the program UCSF Chimera[262]. The solvent accessible surface area was calculated for TβRI in the presence of the TGF-β:TβRII complex and in the absence, and the difference plotted per-residue of TβRI. Chimera calculates the solvent accessible surface using an approximation of spheres.

Analytical size exclusion chromatography isolation of the TGM-1 D123:TβRI:TβRII complex. Samples of TGM-1 D123 and TβRII were combined with a 1.1 molar excess of TGM-1 D123 at pH 7.5 on a Superdex 10-300 200 column. The column was run at a rate of 0.5 mL min^{-}

[1] in a buffer consisting of 25 mM Tris 150 mM NaCl and relevant fractions were collected, concentrated, and analyzed by SDS-PAGE gel to assess for protein composition. Similarly, samples of TGM-1 D123 and TβRI were combined on a Superdex 10-300 75 column with a molar excess of TGM-1 D123 in 25 mM Tris 150 mM NaCl pH 7.5, at a rate of 0.5 mL min^{-1} and relevant fractions were collected, concentrated, and analyzed by SDS-PAGE gel to assess for protein composition.

Analytical size exclusion chromatography isolation of the TGM-1 D12:TβRI and the TGM-1 D2:TβRI complexes. Samples of TGM-1 D12 and TβRI were run with a 2-fold molar excess of TGM-1 D12 in 25 mM Tris 150 mM NaCl pH 7.5 on a Superdex 10-300 75 column. The column was run at a rate of 0.5 mL min^{-1}, and relevant fractions were collected, concentrated, and analyzed by SDS-PAGE gel to assess for protein composition.

Crystal screening of the TGM-1 D12:TβRI and the TGM-1 D2:TβRI complexes. Crystal screening was performed using hanging drop vapor diffusion. Drops of the 15.9 mg mL^{-1} TGM-1 D12:TβRI stock and the 8.9 mg mL^{-1} stock of TGM-1 D2:TβRI were set against conditions in the Hampton ProComplex screen. 400 nL of protein stock was added to 400 nL of buffer with a well volume of 300 μL. Samples were left at 25 °C and assessed for crystal growth after 1 day, 2 days, 3 days, 1 week, 2 weeks, and 3 weeks.

6.3 Results

6.3.1 TGM-1 D12 and the TGF-β:TβRII binary complex compete for binding to TβRI

Competition between TGM-1 D12 and the TGF-β:TβRII complex for binding to the extracellular domain of TβRI was assessed using isothermal titration calorimetry (ITC). For this set of experiments, ligands included TGM-1 D12 and the TGF-β:TβRII complex. The TGF-β:TβRII complex was titrated into a cell containing TβRI (Table 8, Fig. 52). The K_D was fitted to be 61 nM with a fitted ΔH of -4.2 kcal mol^{-1}. This high binding affinity is in line with what has previously been measured[36,49,258] (Table 6, Table 8). The binding affinity for this interaction was quite similar, with TGM-1 D12 binding to TβRI with an affinity of 25 nM (11 – 48 nM, 95% CI) (Table 6). Despite the similarities in binding affinities, due to the difference in the change in enthalpy, assessment of competition was performed by titrating TGM-1 D12 into a cell containing TβRI and TGF-β:TβRII. When this experiment was performed, there was no heat detected, indicating competition between TGM-1 D12 and the TGF-β;TβRII binary complex for binding to TβRI (Fig. 52).

6.3.2 TGM-1 D2 and the TGF-β:TβRII binary complex bind to similar residues on TβRI

TGM-1 D12 and the TGF-β:TβRII complex compete for binding with TβRI. To identify the residues of TβRI that TGM-1 D12 binds, NMR chemical shift perturbation experiments (CSP) were performed. TGM-1 D12 has a molecular weight of 18.8 kDa, while TβRI has a molecular mass of 9.6 kDa. At a 1:1 stoichiometry the combined mass is nearly 30 kDa, which while

permissible for NMR experiments, could make residue assignment difficult. TGM-1 D2 is half the molecular weight of TGM-1 D12 at 9.3 kDa, which makes assignment of the TGM-1 D2 bound TβRI easier. In addition, TGM-1 D12, including a GSGTGSGSGS tag, contains 174 residues, eleven of which are prolines, thus having 162 expected amide backbone residues visible on HSQC. However, not only were far fewer than the expected number of resonances detected, but the resonances that were visible were clustered in the random coil region (7.8 -8.6 ppm in the ^{1}H dimension) (Fig. 53A). This is likely due to sample aggregation as was the case for TGM-1 D1 (Fig. 12). To alleviate any potential aggregation, 10 mM CHAPS was added (Fig. 12B-D). Upon the addition of 10 mM CHAPS, the resonances were more dispersed, with several signals appearing outside of the random coil region. Though this was not tested, it is likely that reduction of the sample concentration would perform a similar effect, indicating that the TGM-1 D12 was natively folded but aggregated under NMR sample conditions.

The propensity of TGM-1 D12 to aggregate would make assignment of TGM-1 D12 bound TβRI difficult, as CHAPS would be needed. This is shown with an overlay of the ^{1}H-^{15}N HSQC spectra of TGM-1 D12 bound ^{15}N TβRI and free ^{15}N TβRI (Fig. 54). When ^{15}N TβRI was titrated with TGM-1 D12 without the presence of CHAPS, there was a visible change in the ^{1}H-^{15}N HSQC of TβRI. Though there were significant chemical shift perturbations for several of the residues, (Fig. 54), more importantly there was a weakening in intensity for a majority of the amide backbone peaks. This indicates that TGM-1 D12 is forming an aggregated species, but that there are enough interactions between TGM-1 D12 and ^{15}N TβRI that the latter is being drawn into this aggregate. Without the presence of CHAPS, it is unlikely that a well-dispersed spectra of TGM-1 D12 bound ^{15}N TβRI could be produced.

For these reasons, CSP experiments of ^{15}N TβRI were performed with binding to TGM-1 D2. TβRI was ^{15}N and ^{13}C-labeled and bound to 1.2 equivalents of unlabeled TGM-1 D2. Backbone chemical shift values for the unbound ^{15}N ^{13}C TβRI had already been assigned, deposited under BMRB code 17276[48] (Fig 55A). The chemical shift values for the amide nitrogen and hydrogen, along with the carbonyl, alpha and beta carbons, were assigned to the TGM-1 D2 bound TβRI (Fig. 56B). The plot of composite shift per residue for ^{15}N ^{13}C TβRI is striking in that there are several regions in which there is significant chemical shift perturbation detected (Fig. 56A). On a scale normalized to 1.0, three distinct regions have a normalized composite shift perturbation higher than 0.4, spread across the entire face of the protein, though more so in the C-terminal half. This likely signifies that there are multiple regions of TβRI involved in binding TGM-1 D2: the C-terminal end of β1 and the subsequent turn (t1) comprising amino acids 32-40, turn 5, aka the pre-helix extension mentioned previously (t5) and the subsequent helix comprising amino acids 78-87, and β5, turn 6, and the C-terminal tail (ect), comprising residues 97-110.

To compare the binding interface between TGM-1 D2:TβRI and TGF-β:TβRII:TβRI, the per residue composite shift perturbations from the binding of TGM-1 D2 to TβRI were compared to the difference in per residue solvent accessible surface area (ΔSAS) of residues of TβRI in the unbound and TGF-β:TβRII bound forms, which was used to demonstrate the residues key in the TGF-β:TβRII binding interface (Fig. 56B). The plot of per residue ΔSAS shows that there are two regions of TβRI whose surfaces are buried upon binding to TGF-β:TβRII: 1) the region comprising the pre-helix extension and the following helix (residues 77-84), and 2) the region including β5, turn 5 and the reaming C-terminal region (residues 98-106). The first region, the prehelix extension and the subsequent turn, specifically residues R77 and D78, contacts the TGF-β:TβRII shared interface, including D141 of TβRII and K397 of TGF-β. The second C-terminal region contacts

the ordered N-terminal tail of TβRII, specifically V45 and F47 of TβRII. The composite shift perturbation plot of TGM-1 D2 bound TβRI has these specific residue contacts, but also includes a third contact region β1 and the subsequent turn t1, that the TGF-β:TβRII interface lacks. This indicates that while there are some shared residues between the TGM-1 D2:TβRI binding interface and the TGF-β:TβRII:TβRI binding interfaces, that TGM-1 D2 likely contacts more residues of TβRI throughout the entire protein. The hypothesis that the TGM-1 D2:TβRII interface is more extensive as the TGF-β:TβRII:TβRII interface would explain the increased affinity of TGM-1 to TβRI over the TGF-β:TβRII complex.

6.3.3 Structure determination of TGM-1:TβRI by X-ray crystallography

6.3.3.1 Crystallization of the TGM-1 D123:TβRI:TβRII Complex

Attempts were made to crystallize the TGM-1:TβRI complex for structure determination. The first attempt used the TGM-1 D123:TβRI:TβRII complex. Not only would this confirm the TGM-1 D3:TβRII structure determined via NMR, but this would also confirm whether there is any interaction between TβRI and TβRII in the TGM-1 D123:TβRI:TβRII complex, and how the TGM-1 domains are oriented with respect to each other. Previous attempts had been made to crystallize TGM-1 D123, a construct containing the first three domains of TGM-1, but this had been a construct that included a his-tag and a myc-tag (Table 3), and so for the determination of a TGM-1 D123:TβRI:TβRII complex, a separate construct of TGM-1 D123 without the myc-tag and with a thrombin cleavable his-tag was expressed and produced in mammalian cells.

TGM-1 D123 was assessed for complex formation with TβRI and TβRII via analytical SEC. Analytical 10-300 200 SEC was run for the TGM-1 D123:TβRII complex (Fig. 57A). As

indicated, TGM-1 D123 eluted at 14.2 mL, while TβRII eluted at 16.6 mL. The elution volume of the TGM-1 D123:TβRII complex peak was unchanged from that of the TGM-1 D123 peak eluting at 14.3 mL, indicating a lack of separable complex via SEC. Attempts to isolate the TGM-1 D123:TβRI complex were subsequently performed on the 10-300 75 column. TGM-1 D123 eluted at 11.7 mL in this instance and the complex peak eluted at 12.0 mL (Fig. 57B). Since the complex could not be isolated via SEC, in order to generate the TGM-1 D123:TβRI:TβRII complex, equimolar amounts of TGM-1 D123, TβRI, and TβRII were combined and concentrated, generating a stock at a concentration of 7.6 mg mL^{-1} (Fig. 57C).

The complex was sent for crystallization screening of 1200 conditions over a variety of screens, with a robot pipetting 200 nL of the protein stock and 200 nL of the condition in question. Of the conditions screened, one of them produced a crystal under the conditions: 0.056 M sodium phosphate monobasic monohydrate ($NaH_2PO_4 \bullet H_2O$) 1.344 M potassium phosphate dibasic K_2HPO_4 pH 8.2. The crystal diffracted to 5-6 angstroms, but the crystal structure could not be solved due to lattice translocation defects.

6.3.3.2 The TGM-1 D12:TβRI and the TGM-1 D2:TβRI complexes are being optimized

As mentioned previously, on analytical SEC 10-300 75 the TGM-1 D12:TβRI complex eluted at 11.0 mL, while TGM-1 D12 eluted at 12.2 mL, and the TβRI eluted at 13.8 mL (Fig. 23B). The elution volume of the TGM-1 D12:TβRI complex peak was distinct from that of the TGM-1 D12 indicating a separable complex formation via SEC. A stock of TGM-1 D12:TβRI complex was generated at a concentration of 6.1 mg mL^{-1}. The complex was sent for crystallization screening of 1200 conditions over a variety of screens, with a robot pipetting 200 nL of the protein

stock and 200 nL of the condition in question. One condition produced a crystal that diffracted to 12 angstroms, which was not deemed sufficient for further attempts at optimization.

More TGM-1 D12:TβRI complex was produced (Fig. 58) but at a much higher concentration, as at neutral pH, both components of the complex are soluble. TGM-1 D12:TβRI complex was generated at a concentration of 15.9 mg mL^{-1} and tested over a variety of crystallization conditions. At this stock concentration, over the course of 2-3 days at room temperature, several conditions produced thin needle-like crystals, and some small plate-like crystals. None of the looped crystals diffracted. The conditions that produced the most well-formed crystals are currently being optimized for growth.

Similarly, the TGM-1 D2:TβRI complex was produced at a concentration of 8.9 mg mL^{-1} and tested over a variety of crystallization conditions. At this stock concentration, over the course of 1-2 days at room temperature, while most of the wells were clear or produced light precipitate, two conditions produced plate-like crystals: 1) 0.1 M lithium chloride, 0.1 M HEPES pH 7.5, 20% PEG 400, and 2) 0.1M ammonium sulfate, 0.1 M Tris pH 7.5, 20% PEG 1500. The crystals formed in the latter condition were tested to see if they contained one, or both of the proteins present (Fig. 58C), which they clearly did, while the crystals in the latter condition were looped as the condition included the cryoprotectant PEG 400. None of the looped crystals diffracted. Optimization has been limited, due to difficulties in reproducing the crystal conditions.

6.3.4 Subsequent attempts at TGM-1 D12:TβRI complex structure determination were performed using NMR

The backbone residues of TβRI had been previously assigned (Fig. 56), both in the unbound and TGM-1 D2 bound forms, indicating complex structure determination may be permissive to NMR methods. The difficulties in doing so comes from the conformational heterogeneity of free TGM-1 D2, likely due to cis-trans proline isomerization, that leads to observable conformational doubling (Fig. 18). This doubling makes assignment of TGM-1 D2 difficult, though not impossible. Currently, the backbone assignments of free and TβRI-bound TGM-1 D2 are in progress.

With the CSP determination for both the TGM-1 D2-bound TβRI and the TβRI-bound TGM-1 D2, along with the crystal structure of TβRI and the alphafold generated structure of TGM-1 D2, docking simulations could be performed, using a program such as Haddock[292]. Without any intermolecular restraints, PCS, NOE, etc., this would generate a varying set of potential structures rather than a converged set. However, were the backbone assignment to be successful, PCS restraints and RDC restraints could be generated in the future, keeping in mind that attempts to modify TβRI with lanthanide-binding tags would likely prove difficult, yielding constructs that would not readily fold into a native conformation, even less so than wild-type TβRI. Still, NMR structure determination of this complex would potentially be feasible if crystallography was to be unsuccessful.

6.4 Discussion

Currently, structure determination of the TGM-1:TβRI complex is in progress and will likely include more months of crystal growth optimization and NMR restraint generation. Thus, the structure determination remains incomplete, and so the specific residues involved in the TGM-1 D12:TβRI interface cannot yet be determined with confidence. Nonetheless, this chapter did yield some insights about the formation of the TGM-1:TβRI:TβRII complex.

The ITC experiments show clear evidence of competition, supported by negligible change in enthalpy upon injection of TGM-1 D12 into a cell containing both the TGF-β:TβRII binary complex and TβRI. This is supported by the NMR CSP experiments comparing the chemical shifts of TGM-1 D2 bound TβRI to free TβRI, which indicate that there are three regions of TβRI perturbed by binding to TGM-1 D2: 1) the C-terminal end of β1 and the subsequent turn (t1) comprising amino acids 32-40, 2), turn 5, aka the pre-helix extension mentioned previously (t5) and the subsequent helix comprising amino acids 78-87, and 3) β5, turn 6, and the C-terminal tail (ect), comprising residues 97-110. The latter two regions are canonically used to bind the TGF-β:TβRII complex, with the prehelix extension inserting into the interface made by TGF-β and TβRII, and the C-terminal tail and β5 strand contacting the N-terminal residues of TβRII. These experiments suggest that the parasite has evolved to utilize similar binding residues of TβRII as the human ligand, though according to the CSPs the contacts seem to be more extensive in the TGM-1 D2:TβRI interface as compared to the TGF-β:TβRII:TβRI interface, which may account for the ability of TGM-1 to bind to TβRI without the need of TβRII.

These results highlight that TGM-1 is a true molecular mimic of TGF-β, as it binds to similar residues of TβRI as does TGF-β. Still there are several unanswered questions about the

characterization of the TGM family proteins, not only in the structure of the TGM-1:TβRI complex, but with other matters such as the functional characterization of domains 4 and 5, which are not required for binding to TβRI and TβRII. The role of these domains will be discussed in Chapter 8. Additionally, discussion of other TGM family members will be present in the following chapter.

6.5 Respective Contributions

All experiments in this chapter, and all analyses of these experiments were performed in their entirety by the author Ananya Mukundan, with the exception of the ^{15}N TGM-1 D12 ^{1}H-^{15}N HSQC experiments (Fig. 12), which were performed by Andrew Hinck. In addition, the initial crystal screen of the TGM-1 D123:TβRI:TβRI complex, as well as the crystal looping and diffraction of the crystal produced by this complex, as well as the initial crystal screen of the ca. 5mg mL^{-1} TGM-1 D12:TβRI complex were performed by Alexander Taylor at the University of Texas San Antonio.

6.6 Tables and Figures

Table 17: Sequences of constructs used in this chapter.

Construct	Coding region and description (* indicates stop codon)
TGF-β3	Residues 331-442 of human TGF-β3 protein, NCBI ABQ59024 Initiating Methionine-TGF-β3 MALDAAYCFRNVQDNCCLRPLYIDFRKDLGWKWIHEPKGYNANFCAGACPYRASKSPSCVSQ DLEPLTIVYYVGRKPVEQLSNMIVKSCKCS*
TGM-1 D123, cleavable His tag	Residues 16-262 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Signal Peptide- His_8- Linker-Thrombin Cleavage Site-TGM-D1-D3 MKWVTFLLLLFISGSAFSGSGSHHHHHHHHGSLVPRGSDDSGCMPFSDEAATYKYVAKGPKNI EIPAQIDNSGMYPDYTHVKRFCKGLHGEDTTGWFVGICLASQWYYYEGVQECDDRRCSPLPT NDTVSFEYLKATVNPGIIFNITVHPDASGKYPELTYIKRICKNFPTDSNVQGHIIGMCYNAEWQF SSTPTCPASGCPPLPDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRRCRALSQEAD PGEFVAICYKSGTTGESHWEYYKNIGKCPDP*

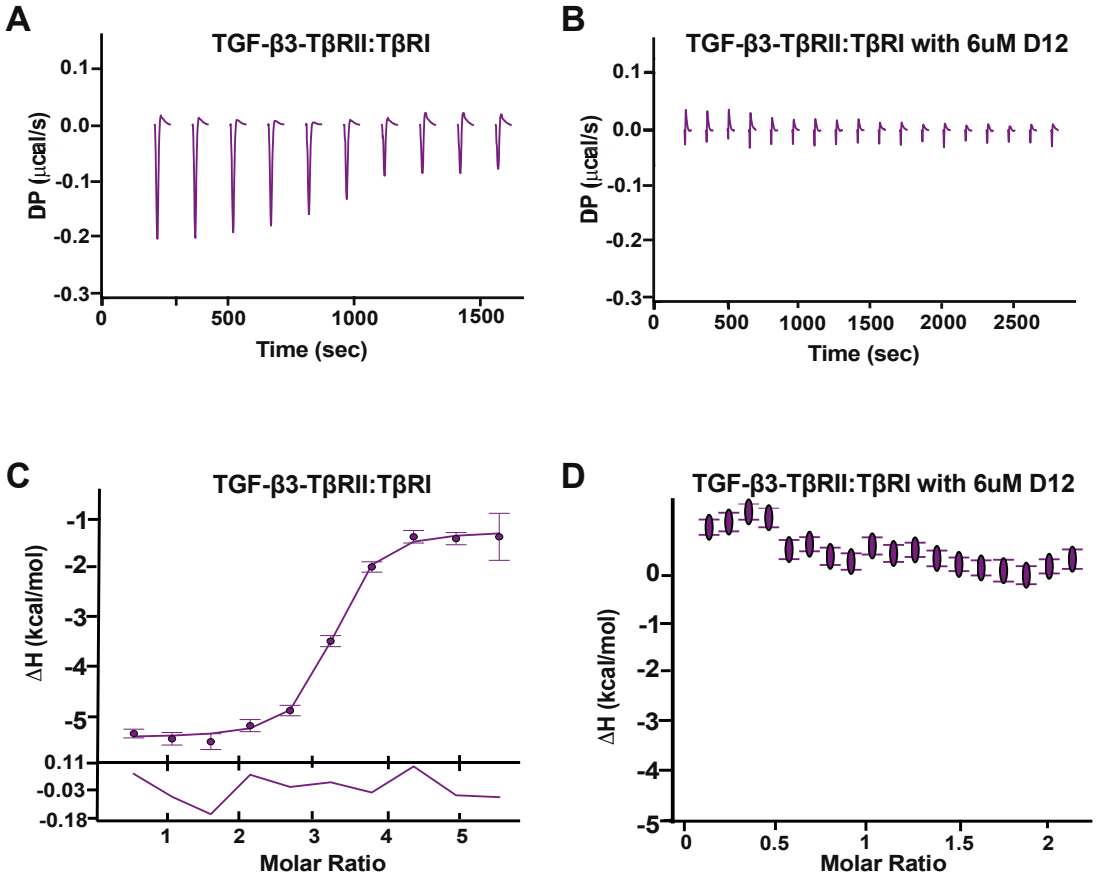

Figure 52: TGM-1 D12 competes with the TGF-β;TβRII complex for binding to TβRII as shown by ITC. A-D. Raw thermograms for the injection of the TGF-β:TβRII complex into a cell containing TβRI (A) and injection of TGM-1 D12 into a cell containing TβRI and the TGF-β:TβRII complex (B). Corresponding integrated heats for the aforementioned data (A-C, B-D) fit to a 1:1 binding model. Fits correspond to the global fit over the data sets with residuals below. Error bars correspond to undertainty in the estimation of integrated heats by NITPIC. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0 / Edited from original.

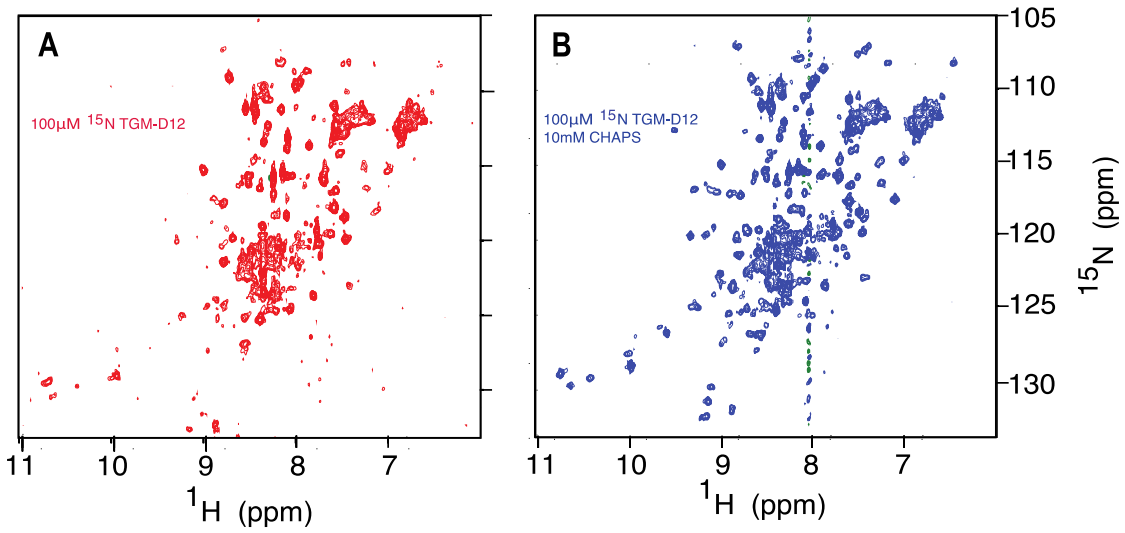

Figure 53: Recombinantly produced TGM-1 D12 aggregates under NMR sample conditions. A-B. ^{1}H-^{15}N spectra of ^{15}N TGM-1 D12 (A) in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K, with (B) and without (A) the addition of 10mM CHAPS.

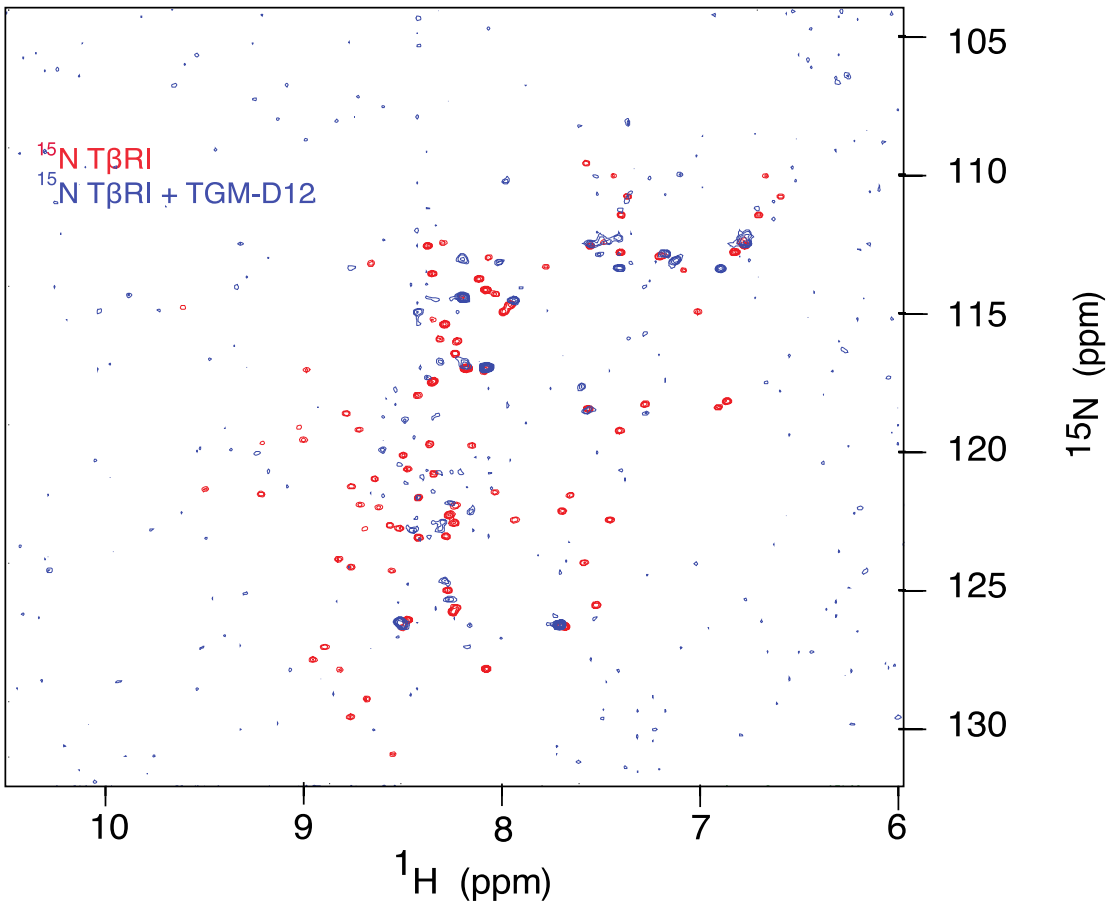

Figure 54: TGM-1 D12 binds to TβRI under NMR sample conditions. ^{1}H-^{15}N spectra of ^{15}N TβRI (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.3 molar equivalents of unlabeled TGM-1 D12. The spectrum was recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K.

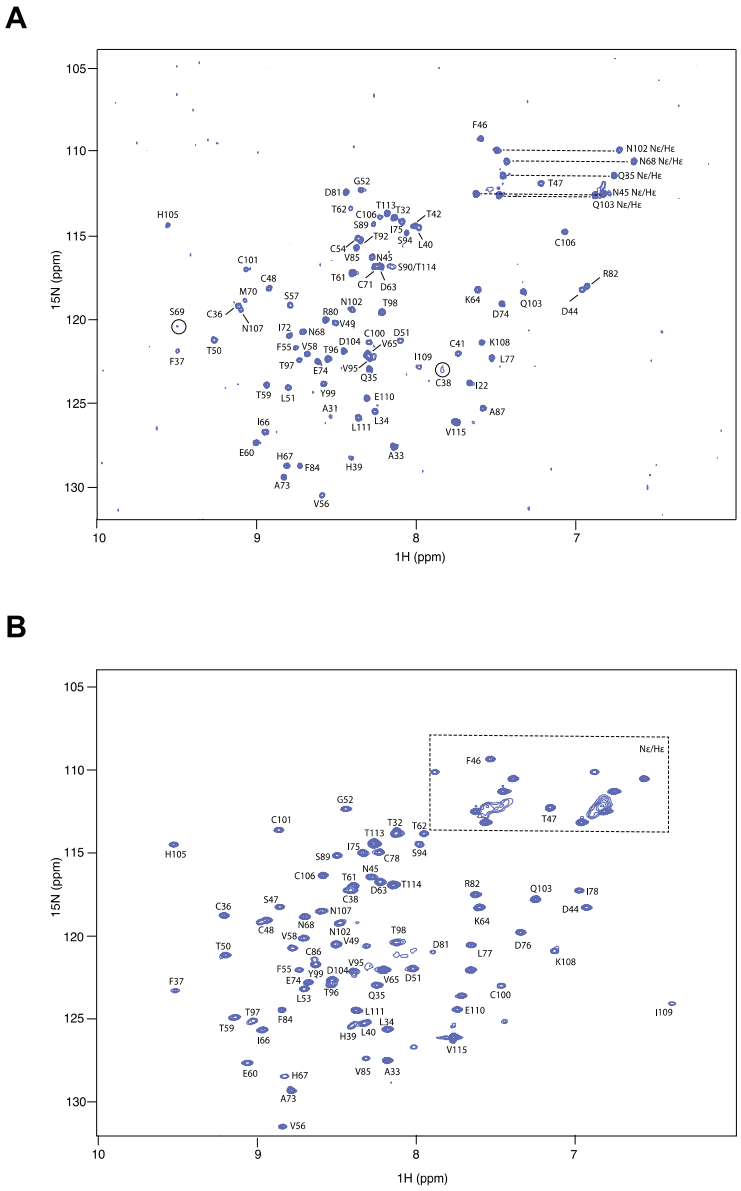

Figure 55: ^{1}H-^{15}N Assignment of TβRI. ^{1}H-^{15}N HSQC of free TβRI (A) and TβRI bound to TGM-1 D2 (B), assigned peaks displayed on the spectra. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K. Dashed line indicating region of HSQC corresponding to sidechain Asn/Gln residue amides. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0.

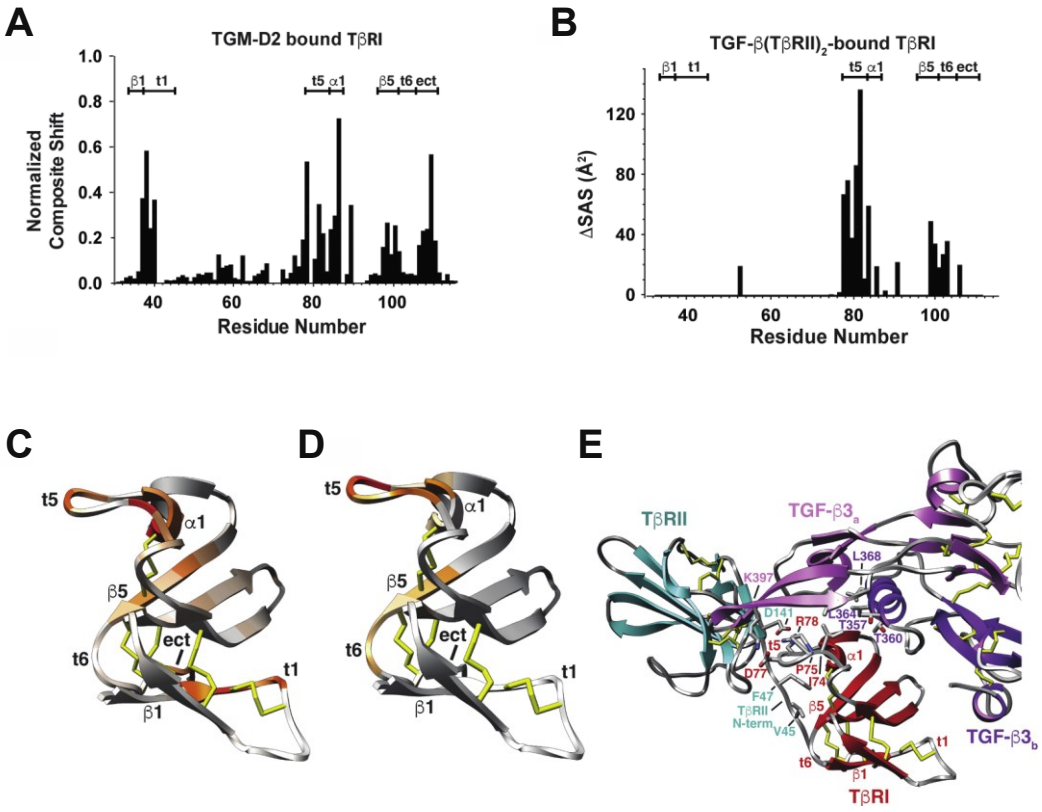

Figure 56: TGM-1 D2 binds to TβRI using similar residues to the TGF-β:TβRII complex. Plot of the normalized composite shift perturbations of TGM-1 D2 bound TβRI (A), with corresponding depiction onto the structure of TβRI (C) with white indicating little-no composite shift perturbation and red indicating close to 1.0 normalized composite shift perturbation. Plot of the difference in solvent exposed surface area between free TβRI and TGF-β:TβRII bound TβRI (B), with corresponding depiction onto the structure of TβRI (D) with white indicating little composite shit perturbation and red indicating close to 1.0 normalized composite shift perturation. A depiction of the TβRI:TβRII:TGF-β3 interface is shown in panel E with relevant structural features annotated, TβRI colored in red, TβRII colored in blue, and TGF-β monomer colored in purple. and TβRI:TβRII:TGF-β structure from PDB 2PJY. Figure from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0. / Edited from original.

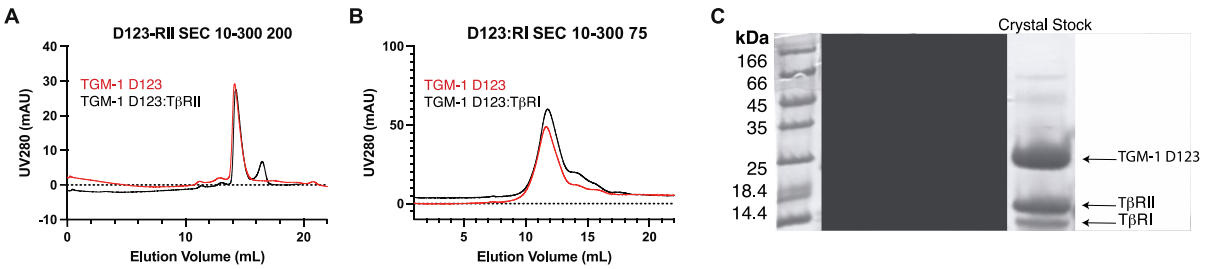

Figure 57: The TGM-1 D123:TβRII and TGM-1 D123:TβRI complexes do not readily separate from their individual components on 10-300 200 analytical SEC. A-C. A. SEC chromatogram of TGM-1 D123 (red) and the TGM-1 D312:TβRII complex (black) in a near 1:1 titration run on a 10-300 200 column. Left axis indicates absorbance at 280nM, which is plotted with respect to elution volume. B. SEC chromatogram of TGM-1 D123 (red) and the TGM-1 D312:TβRI complex (black) in a near 1:1 titration run on a 10-300 75 column. Left axis indicates absorbance at 280nM, which is plotted with respect to elution volume. C SDS gel confirming composition of crystallization stock, with labels as indicated.

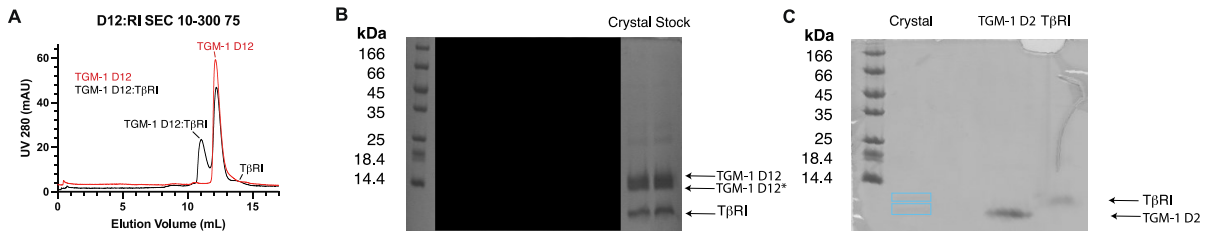

Figure 58: The TGM-1 D12:TβRI complex readily separates from its individual components on 10-300 75 analytical SEC A-C. A. SEC chromatogram of TGM-1 D12 (red) and the TGM-1 D12:TβRII complex (black) in a near 1:1 titration run on a 10-300 75 column. Left axis indicates absorbance at 280nM, which is plotted with respect to elution volume. B. SDS gel confirming composition of TGM-1 D12:TβRI crystallization stock, with labels as indicated. TGM-1 D12* indicates TGM-1 D12 with minor proteolytic cleavage. C. SDS gel confirming composition of TGM-1 D2:TβRI crystal, with labels as indicated, blue boxes surrounding bands in lane containing resolubilized crystal.

7.0 Characterization of the functional properties of TGM-4

7.1 Introduction

TGM-1 is part of a larger family of *H. polygyrus* proteins, the first six of which (TGM-1 through TGM-6) are present in the adult worm, while the other three proteins, TGM-7 through TGM-10 are found in the L4 larval proteome[242]. To identify the members of this protein family, mass spectrometric data of peptide fragments were compared to a currently existing transcriptomic database[242]. The first three adult worm TGM family proteins (TGM-1 through TGM-3) share over 90% identity, and all three arise from the same genetic locus. TGM-4 has 80% identity TGM-1, TGM-2, and TGM-3, and contains all five domains D1-D5. TGM-5 and TGM-6 contain one or two fewer domains, TGM-5 lacking domain 4 and TGM-6 lacking domain 1-2[242]. The larval TGMs differ dramatically in composition; TGM-7 and TGM-8 contain two extra domains following domain 3, and TGM-9 and TGM-10 lack two domains each, domains 1-2 for TGM-9 and domains 3-4 for TGM-10. This indicates the adaptability of the TGM family, with different domain compositions likely contributing to different TGF-β receptor pairings.

Along with these differences in domain and sequence identity, the other proteins in the TGM family differentially activate the TGF-β signaling pathway, as assessed via the MFB-F11 alkaline phosphatase reporter assay mentioned previously[229,242]. Of the ten TGM family proteins, only TGM-1 through TGM-3 are able to activate the TGF-β signaling pathway and convert naïve CD4+ T-cells into T_{regs}[242]. However, despite TGM-4 being unable to activate TGF-β signaling, it is able to induce the conversion of naïve T-cells to Tregs [242]. Additionally, initial studies have suggested that TGM-4 binds to TβRI with high affinity (Fig. 59). Thus, there are several potential

hypotheses to clear these discrepancies: 1) it could be that TGM-4 interacts with activin receptors, as activin signaling has been correlated to T_{reg} induction[293,294], or 2) TGM-4 may not be able to bind with high enough affinity to TβRII in the MFB-F11 assay, but may be able to utilize avidity to stimulate conversion of CD4+ T-cells into T_{regs}. The former hypothesis is supported by studies showing that the ES products of a fox tapeworm *Echinococcus multilocularis* contain an activin homologue that expands host T_{reg} populations[295].

The difference in TGF-β receptor binding between the two TGM isoforms indicates that the different TGM family members have likely evolved towards different roles in the parasite. TGM-4 offers an exciting opportunity for therapeutic benefit. Due to its high affinity for TβRI and domains 1 and 2 of TGM-4 may be used to bind TβRI with higher efficacy in a cell context than TGM-1. Thus, chimeras of TGM-4 and TGM-1 may be produced that act to stimulate the TGF-β signaling pathway more efficiently than either TGF-β or TGM-1. Thus the goals of this chapter are to clarify the binding properties of TGM-4, and then to attempt to engineer TGM-1 and TGM-4, from what is currently known about their structural and functional properties, into developing an efficacious TGF-β signaling therapeutic. Due to the modular nature of the TGM domains and depending on the affinity of TGM-4 for the TGF-β receptors, therapeutics targeting different TGF-β receptors may be developed.

7.2 Methods

Cloning, expression, and purification of TGM-4 D2 and TGM-4 D3. DNA inserts coding for TGM-4 were inserted into a modified pET32a vector between the NdeI and HindII) sites (Table 18). Constructs were expressed in BL21 (DE3) cells and cultured at 37 °C. When the light scattering at 600 nm was detected to be 0.8 protein expression was induced by adding 0.8 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG),and were produced in LB. Purification of TGM-4 D2 and TGM-4 D3 was performed in a similar manner as the purification of TGM-1 D2 and TGM-1 D3 respectively (Chapter 3.2).

Cloning, expression, and purification of TGM-1 FL and TGM-4 FL. DNA inserts coding for TGM-1 FL and TGM-4 FL were inserted into a pSecTag vector between the AscI and ApaI sites (Table 18). Constructs were expressed in expi293 cells, transfected, and cultured at 37 °C. The solutions were collected after one week. Purification was performed in a similar manner as described for in previous sections (Chapter 3.2 Methods).

Cloning, expression, and purification of TGF-β family type II receptors. DNA inserts coding for the extracellular domains of ActRII, ActRIIB, and BMPRII, preceded by a hexahistidine tag and a thrombine cleavage site, were inserted into a pcDNA 3.1+ vector between the NotI and XhoI sites. The pcDNA 3.1 vector was modified with a signal peptide followed by a NotI site. The resultant construct had the following components: Signal Peptide-His_6-thrombin cleavage site-Receptor (Table 3). Constructs were expressed in expi293 cells, transfected, and cultured at 37 °C. Purification was performed in a similar manner as mentioned for TGM-1 FL.

NMR 2D experiments. All NMR experiments were run in the buffers mentioned in the figures/text, usually 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0, 310K. Each sample was run in a 5

mm susceptibility-matched NMR microtube for data collection. NMR data was collected with Bruker 600, 700, and 800 MHz spectrometers containing a 5 mm ^{1}H (^{13}C, ^{15}N) z-gradient "TCI" cryogenically cooled probe. ^{1}H-^{15}N HSQC spectra were acquired as described, with water flipback[250] and WATERGATE suppression pulses[251]. NMR spectra were processed, analyzed, and visualized using NMRPipe[253] and NMRFAM-SPARKY[254].

SPR measuring TGM-4 binding to TβRI and TβRII. All SPR experiments were performed with a BIAcore X100 system. Neutravidin was coupled to the surface of a CM5 chip and biotinylated avi-tagged TβRI and biotinylated avi-tagged TβRII were captured onto the chip surface, at a maximum density of 150 RU. Neutravidin coupling was performed by EDC-NHS activation of the chip, followed by neutravidin (buffer: sodium acetate, pH 4.5) injection over the surface until the RU increased by 6000-15000 RU. All experiments were run in the same buffer: 25 mM HEPES, 150 mM NaCl, 0.005% P20 surfactant, pH 7.4, at an injection rate of 100 μL min^{-1}. The surface was regenerated in between each injection with a 30 second injection of 1 M guanidine hydrochloride. The experimental sensorgrams were obtained with double referencing with a control cell coated similarly with neutravidin but lacking the captured receptor and 8 blank buffer injections at the beginning of the run before injection of the samples. The data was analyzed by fitting the results to a 1:1 kinetic model using the SPR analysis software Scrubber .

ITC measuring TGM-4 FL binding to type II TGF-β receptors. All ITC experiments were performed with a Microcal PEAQ-ITC system. All experiments were performed in 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 at 35 °C, with 15 2.5 μL injections with a duration of 5 s, a spacing of 150 s, and a reference power of 10. All samples were dialyzed against the same ITC buffer before loading into the system. Data was globally fit using the programs NITPIC[246], SEDPHAT[247,248], and GUSSI[249].

MFB-F11 TGF-β stimulatory assays. The MFB-F11 TGF-β bioassay was previously reported by Tesseur et al.[229]. Confluent cells (8 x 10^5 cells/mL) were resuspended in DMEM with 2.5% FCS, 100 U/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine. To each cell of a 96-well plate, $4x10^4$ cells were added. Purified proteins at the stated concentrations were added to the wells (volume 50 µL) in triplicate and incubated overnight at 37 °C. After incubation, 20 µL were added to an ELISA plate with 180 µL of reconstituted Sigma FastTM p-nitrophenyl phosphate substrate and incubated at room temperature in the dark for 12 h – 18 h. Plates were read on at 405 nm on an E_{max} precision microplate reader.

Regulatory T-cell conversion assays. The T_{reg} conversion assay was performed as previously described[230]. CD4+ CD25- GFP- CD62Lhi cells were sorted by FACS, washed and resuspended in RPMI, and added to CD3/CD28-coated 24 well plates. IL-2 and variable amounts of TGM family members were added to the plates and cells analyzed by flow cytometry after 96 hrs of incubation.

7.3 Results

7.3.1 TGM-4 domains 2 and 3 can be expressed and purified as native proteins in E. coli

In previous chapters, TGM-1 D2 was shown to be the main binding partner of TβRI, and TGM-1 D3 was shown to be the primary binding partner of TβRII. Thus, it was hypothesized that TGM-4 D2 would be the primary binding partner of a TGF-β family type I receptor, and that TGM-4 D3 would be the primary binding partner of a TGF-β family type II receptor. To test this hypothesis, TGM-4 D2 and TGM-4 D3 were expressed recombinantly in E. coli. In a similar manner as in chapter 3, the individual domains were expressed as thioredoxin fusions with a thrombin-cleavable internal His-tag. (Table 18).

The proteins were produced in a similar manner as the native TGM-1 domains, and due to their high sequence identity, their soluble-insoluble partitioning as well as their charge properties were similar. The proteins were produced in inclusion bodies via incubation at 37 °C and were expressed as unfolded monomers. The inclusion bodies were solubilized in 8M Urea, 25 mM Na_2HPO_4 and 50 mM NaCl pH 8.0, and the resultant solute purified on an NiNTA column. The fractions corresponding to the thioredoxin fusion domain were folded by diluting into a Tris buffer using reduced and oxidized glutathione, 2 mM and 0.5 mM respectively. Following the folding, the thioredoxin tag was cleaved from the protein and the resulting protein was further purified to homogeneity using ion-exchange chromatography.

The results were validated by SDS and mass spectrometry, the former for homogeneity, and the latter for protein identity. Mass spectrometry for the individual domains showed a mass equivalent to the calculated molecular weight with a variance of 0.5 Da. The proteins were then

validated for native folding via ^{1}H NMR. Both domains showed distinct shifts in the expected methyl and amide regions without any evidence of nonnative structures, as would be evident by shifts condensed in the 0.8-1.0 ppm and in the 7.8-8.6 ppm region (Fig. 60).

7.3.2 TGM-4 binds TβRI

To test whether TGM-4 binds TβRI, SPR was utilized. This was done as previous attempts at testing TGM-4 FL binding to TβRI via ITC were inconclusive due to small changes in enthalpy that prevented accurate quantitation of the binding affinity. The sequence of TGM-4 diverges from TGM-1 mostly in domains 1 with 71% sequence identity of TGM-4 D1 to TGM-1 D1. However, the expected binding domain to TβRI, domain 2, shares 90% identity with TGM-1. To measure the binding affinity of TGM-4 to TβRI, biotinylated avi-tagged TβRI was captured on a neutravidin-coated SPR chip, and TGM-4 FL was injected as the analyte. Due to the high affinity of TβRI to TGM-1 FL and the hypothesized high affinity of TGM-4 FL to TβRI, the analyte concentration used varied between 12.5 nM and 250 nM (Fig. 59A-D).

A clear concentration-dependent response was present when both TGM-1 FL and TGM-4 FL were injected over the TβRI chip surface, (Fig. 59). The binding affinity as derived by fitting the sensorgrams to a 1:1 binding model were 70 ± 6 nM for the former and ranging from 3 – 5 nM for the latter (Table 19). The binding affinity of the TGM-1 Fl to TβRI interaction is consistent with previously attained ITC data in Chapter 3, which shows a binding affinity of 52 nM (Table 6). From this data, it is clear that TGM-4 FL not only binds to TβRI, but does so with a higher affinity than TGM-1 FL by roughly an order of magnitude.

7.3.3 TGM-4 does not bind the TGF-β family type I receptor Alk4

While TGM-1 has been shown to not be promiscuous for other TGF-β receptors (Fig. 24), unpublished ELISA data suggested that there was a potential interaction between TGM-1 and Alk4 (data not shown). This is supported by structural similarities between Alk4 and TβRI, both of which contain a prehelix extension, albeit with a somewhat different amino acid composition: PAGKP for Alk4 and PRDRP for TβRI. To test for binding between TGM-4 and Alk4, ^{1}H-^{15}N HSQC experiments were performed with ^{15}N-labeled Alk4 in the absence or presence of TGM-4 D2 (Fig. 61). As TGM-1 D2 was the primary domain responsible for the TGM-1 FL:TβRI interaction, it was hypothesized that were TGM-4 FL to bind to Alk4, TGM-4 D2 would be the primary domain responsible. However, the NMR results clearly show that there is no binding between TGM-4 D2 and Alk4, which is consistent with the lack of binding between TGM-1 D2 and Alk4. Although this doesn't preclude binding of other domains of TGM-4 to Alk4, this is unlikely based on previous data. Binding to the other type I receptors was not tested due to the difficulties in producing adequate amounts of these receptors over a reasonable timescale, but these receptors are unlikely to bind to TGM-4 as well, as they are less similar in sequence and structure to TβRI than Alk4. This data suggests that TGM-4 FL, in a manner similar to that of TGM-1 FL, is both specific for TβRI over other type I receptors and binds with a very high affinity.

7.3.4 TGM-4 weakly binds TβRII

TGM-4 FL is unable to activate TGF-β signaling in the MFB-F11 alkaline phosphatase reporter assay yet is still able to convert naïve CD4+ T-cells into T_{regs} (Fig. 62). As TGM-4 binds

to TβRI with even higher than TGM-1, this cannot be because of a lack of type I receptor binding. The prevailing hypothesis then was that TGM-4 does not bind to TβRII or binds to TβRII so weakly that it is unable to activate TGF-β signaling in MFB-F11 cells. However, the regulatory T-cell conversion assay requires incubation with the TGM protein over a period of 96 hours, rather than the 12-18 hours of the MFB-F11 assay. Additionally, the T_{reg} conversion assay utilizes other co-factors to prime the CD4+ T-cells such as IL-2, which may aid in maximizing T_{reg} conversion despite weaker TGF-β signaling.

There is unpublished data to support this hypothesis, which shows a difference in signaling ability dependent upon the construct of TGM-4 used. When monomeric, or wild-type TGM-4 is used, there is little to no TGF-β signaling detected via the MFB-F11 assay. However, when a dimeric construct of TGM-4 is made using an Fc-fusion, there is detectable TGF-β signaling via the assay. This signaling is not as high as that of monomeric TGM-1 FL, but the detection of any signaling at all would indicate that TGM-4 does bind to TβRII, albeit with a very weak affinity, and that this affinity is likely strengthened by avidity.

To evaluate this hypothesis, TGM-4 was tested for binding to TβRII by NMR and by SPR. NMR was initially used to determine whether there was binding to TβRII. To do this, TGM-4 D3 was isotopically labeled with ^{15}N and a ^{1}H-^{15}N HSQC was recorded on this sample with and without the presence of equimolar equivalents of TβRII (Fig. 63A-B). This showed that some backbone amide signals of the ^{1}H-^{15}N spectrum of ^{15}N TGM-4 D3 shift upon the addition of wild-type TβRII, albeit slightly. All of the peaks are in fast exchange, as shown by the change in average peak position rather than the presence of two peaks, indicative of weak binding (Fig. 63C). Following this, to quantitate the binding of TGM-4 to TβRII, varying concentrations of full-length TGM-4 were injected over captured biotinylated avi-tagged TβRII on a neutravidin-coated chip.

Though high concentrations of TGM-4 analyte were required to attain SPR sensorgrams, based on the concentrations evaluated, the binding affinity of the TGM-4 FL:TβRII interaction is over 100 μM (Fig. 63B). Notably this K_D was determined using the SPR sensorgram and the program Scrubber but it is not present in the SPR data as the K_D is an approximation based on the maximum response over the sensor surface, which was determined based off of the maximum response of a similarly sized protein over the same chip, TGM-1, which the sample series never reached. However, from the NMR and SPR data, it is apparent that while TGM-4 does bind to TβRII, using TGM-4 D3, it does so weakly and TGM-4 is likely only able to signal due to avidity or other in-cell effects that effectively maximize the local concentration.

7.3.5 TGM-4 does not bind any other type II TGF-β receptors

TβRII clearly binds TGM-4, albeit weakly, and thus is it possible that TGM-4 may bind another type II TGF-β family receptor, such as ActRII, ActRIIB, BMPRII, or MISRII. Due to the role of MISRII in the reproductive system, it is unlikely to be a target of parasitic molecular mimicry, and thus binding of TGM-4 to MISRII was not assessed. In addition, due to the high sequence identity between the extracellular domains of ActRII and ActRIIB, only binding of TGM-4 to ActRII was assessed. Binding of the extracellular domains of ActRIIB and BMPRII to TGM-1 and other TGM isoforms was assessed via ITC (Fig. 64). TGM-4 FL was placed in the cell at a concentration of 10 – 15 μM. ActRII and BMPRII were placed in the syringe at 150 – 200 μM and titrated into the cell. It is clear from the ITC thermograms that there is no change in heat detected upon titration of the receptors into TGM-4 FL, indicating a lack of binding. This likely indicates that TGM-4 FL does not bind to any other type II TGF-β receptors outside of TβRII,

though there is the possibility that the binding is weak enough that under the concentrations tested, the change in enthalpy over the course of the injection series was too small to be detected.

7.3.6 TGM-1:TGM-4 chimeras and TGF-β signaling

TGM-4 FL binds to both TβRI and TβRII, the former with an order of magnitude stronger binding affinity than TGM-1 FL, and the latter two order of magnitudes weaker than TGM-4 FL. Additionally, assessment of binding to other TGF-β family receptors indicates that TGM-4 FL is specific for binding to TβRI and TβRII, similar to TGF-β and TGM-1. TGM-4 FL has a very high affinity for TβRI, and while TGM-1 FL has a weaker affinity for TβRII than TGF-β, the affinity is still submillimolar. Thus, a chimeric construct of TGM-1 and TGM-4 could be developed to generate a high affinity binder of the TGF-β receptors, and thus a high affinity TGF-β mimic.

To test this, collaborators in the Maizels lab generated a series of TGM-1 and TGM-4 mutants and chimeras, with both positive (TGM-1 FL WT, TGF-β) and negative controls (TGM-4 FL WT, non-transfected cells). The fusion constructs tested were: TGM-4 D12:TGM-1 D3:TGM-4 D45 and TGM-1 D12:TGM-4 D3:TGM-1 D45. These proteins were expressed in HEK293 cells, and the resultant supernatant transfected into MFB-F11 cells. The hypothesis was that the TGM-4 D12:TGM-1 D3:TGM-4 D45 construct would signal with greater amplitude than TGM-1, due to inclusion of the domains with the high affinity to both TβRI and TβRII, and the TGM-1 D12:TGM-4 D3:TGM-1 D45 construct would likely signal with lesser amplitude than TGM-1, due to the minimal affinity of TGM-4 D3 for TβRII. However, the actual data showed the opposite (Fig. 65). The TGM-4 D12:TGM-1 D3:TGM-4 D45 construct signaled only minimally through the assay. There was some detectable signal at the highest concentrations of this construct,

which was more than seen for the TGM-4 FL alone, but the signal was minimal and not comparable to native TGF-β signaling nor to TGM-1 signaling. The TGM-1 D12:TGM-4 D3:TGM-1 D45 construct signaled in a similar manner to wild type TGM-1 (Fig. 65).

7.4 Discussion

TGM-4 is unique as in previous studies it was shown to not be able to activate the Smad-sensitive alkaline phosphatase reporter in the MFB-F11 murine fibroblast cell line yet was still able to convert naïve CD 4+ T-cells into T_{regs} (Fig. 63). The initial hypothesis was that TGM-4 was binding a type I or a type II receptor in the activin arm of the TGM family, thus activating T_{regs} via a different mechanism than TGM-1. However, upon testing of the promiscuity of binding of TGM-4 to the various receptors of the TGF-β family, it was apparent that TGM-4 did bind TβRI and TβRII, and not the activin type I receptor Alk4 or the activin type II receptor ActRIIB, though more in-depth studies into promiscuity should be completed. TGM-4 binds to TβRI with an order of magnitude stronger binding affinity than TGM-1, and approximately two orders of magnitude weaker affinity to TβRII. This answers the initially posed question, as TGM-4 still binds both receptors, TβRI and TβRII, and is thus able to trigger the downstream effects of TGF-β signaling, but due to the weak affinity to TβRII, signaling via the MFB-F11 reporter assays showed no effect.

TGM-4 binds to TβRI with high affinity, but only barely binds TβRII, while TGM-1 binds to both TβRI and TβRII with submicromolar or micromolar affinity. A chimeric construct of TGM-1 and TGM-4, including domain 3 of TGM-1 but all other domains of TGM-4, would thus be hypothesized to be even better at stimulating TGF-β signaling than TGM-1, and likely be better

at converting naïve CD 4+ T-cells into T_{regs}. To test this hypothesis, a TGM-4 D12:TGM-1 D3:TGM-4 D45 construct was tested for signaling via the MFB-F11 assay. This hypothesis was proven to be mostly false as addition of supernatant containing the TGM-4 D12:TGM-1 D3:TGM-4 D45 protein did not signal sufficiently through the TGF-β pathway; there was some Smad2/3 detected, but only barely stimulating above negative controls and TGM-4 FL. The reverse chimera, TGM-1 D12:TGM-4 D3:TGM-1 D45, unexpectedly, signaled as well or comparably as TGM-1 FL.

This would indicate that there is some element of the interaction between TGM proteins and their receptors that is being overlooked. For example, there may be a promiscuous binder for TGM-4 D12 such that it does not bind solely TβRI and is thus taken up by other partners, minimizing TGF-β signaling. That the substitution of TGM-4 D3 into TGM-1 did not interrupt signaling drastically also indicates that the strength of the binding interaction between the TGM and TβRII is not relevant, only that there is a binding interaction, indicating a potential role for avidity. This indicates a potential role for domains 4 and 5 which have thus far not been studied. It is likely that domains 4 and 5 of the TGM proteins act to bind a coreceptor, and likely one that targets the TGM to the cell surface, allowing for avidity, thus overcoming weaker binding affinities such as the TGM-4 D3:TβRII interaction.

7.5 Respective Contributions

A majority of the experiments in this chapter were performed in their entirety by the author Ananya Mukundan, with the exception of the in vitro TGF-β signaling assays: the stimulation of

the TGF-β signaling pathway by TGM-1 TGM-4 chimeric constructs were performed by Tiffany Campion, and the T_{reg} conversion and MFB-F11 stimulation assays of full-length TGM family members were performed by Danielle Smyth. Thus the data in figures 60-62, and 64-65 was gathered by the author Ananya Mukundan, the data in figure 63 was gathered by Danielle Smythe formerly of the Maizels lab, and the data in figure 66 was gathered by Tiffany Campion in the Maizels lab.

7.6 Tables and Figures

Table 18: Sequences of TGM-4 domain constructs

Construct	Coding region and description (* indicates stop codon)
TGM-4 D2	Residues 97-176 of *H. polygyrus* TGF-β Mimic 4, NCBI MG429739 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-4 D2 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTCSPLPTNDTVTYEYLKATVNAGINFNITVHPDASGKYPELTYIKRICKNFPADSKVQGHIIG MCYNAEWRFSSTPTCPPS*
TGM-4 D3	Residues 177-262 of *H. Polygyrus* TGF-β Mimic 4, NCBI MG429739 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-TGM-4 D3 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNP GTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR GSGTGCPPLPDDGIVFYEYYGYAGNRHTVGRAVSKDSSGNYPPQTHARRRCRALSQKADPGEFV GICYKSGTTGESHWDYYSHIRKCPDP*
TGM-4 FL	Residues 16-422 of *H. Polygyrus* TGF-β Mimic 4, NCBI MG429739 Signal Peptide-TGM-FL-Linker-Myc-Linker-His_6 MLLIVLIGLLEVAATDASGCMPFSDETASYKYLTERSRNDETPAQNDSSGAYPDHTHVKRFCKGL HGEEKTGRYVGICLGSEWVYYQGVQECQDRRCSPLPTNDTVTYEYLKATVNAGINFNITVHPDA SGKYPELTYIKRICKNFPADSKVQGHIIGMCYNAEWRFSSTPTCPPSGCPPLPDDGIVFYEYYGYA GNRHTVGRAVSKDSSGNYPPQTHARRRCRALSQKADPGEFVGICYKSGTTGESHWDYYSHIRKC PDPRCKPLETNVSVHYEYFTMTNETGRKEGTPAEVDKGGKYPQHTCVRKFCDKSPYTCSVKGPIF GECLDGQWNFTALDECLNARGCNSDDLFDKLGFEGVMVREEEGSDSYKDDFVRFYATGSKVNA ECKGKTVQLECSDGEWHDPGTKTVHRCTKEGIRALAAARGGPEQKLISEEDLNSAVDHHHHHH*

Table 19: TGM-1 and TGM-4 binding to TβRI as assessed by surface plasmon resonance.

Surface	Analyte (Concentrations)	Fitted Parameters[a]			
		k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d (nM)	R_{max} (RU)
TβRI[a]	TGM-1 FL (31.25 nM – 1000 nM)	(6.4 ± 0.6) x 10^4	(4.4 ± 0.1) x 10^{-3}	70 ± 6	232 ± 2
TβRI[a]	TGM-4 FL (62.5 nM – 250 nM)	(2.2 ± 0.1) x 10^5	(7.3 ± 0.1) x 10^{-4}	3.2 ± 0.1	222 ± 8
TβRI[a]	TGM-4 FL (31.25 nM – 250 nM)	(1.9 ± 0.1) x 10^5	(9.6 ± 0.7) x 10^{-4}	5.0 ± 0.1	184 ± 1
TβRI[ab]	TGM-4 FL (12.5 nM – 100 nM)	(4.6 ± 0.6) x 10^5	(1.8 ± 0.1) x 10^{-3}	3.8 ± 1	67 ± 1

[a]Fitted parameters were derived from kinetic analysis of a single injection series
[b]Measured with lower analyte concentrations compared to that used for the other experiments

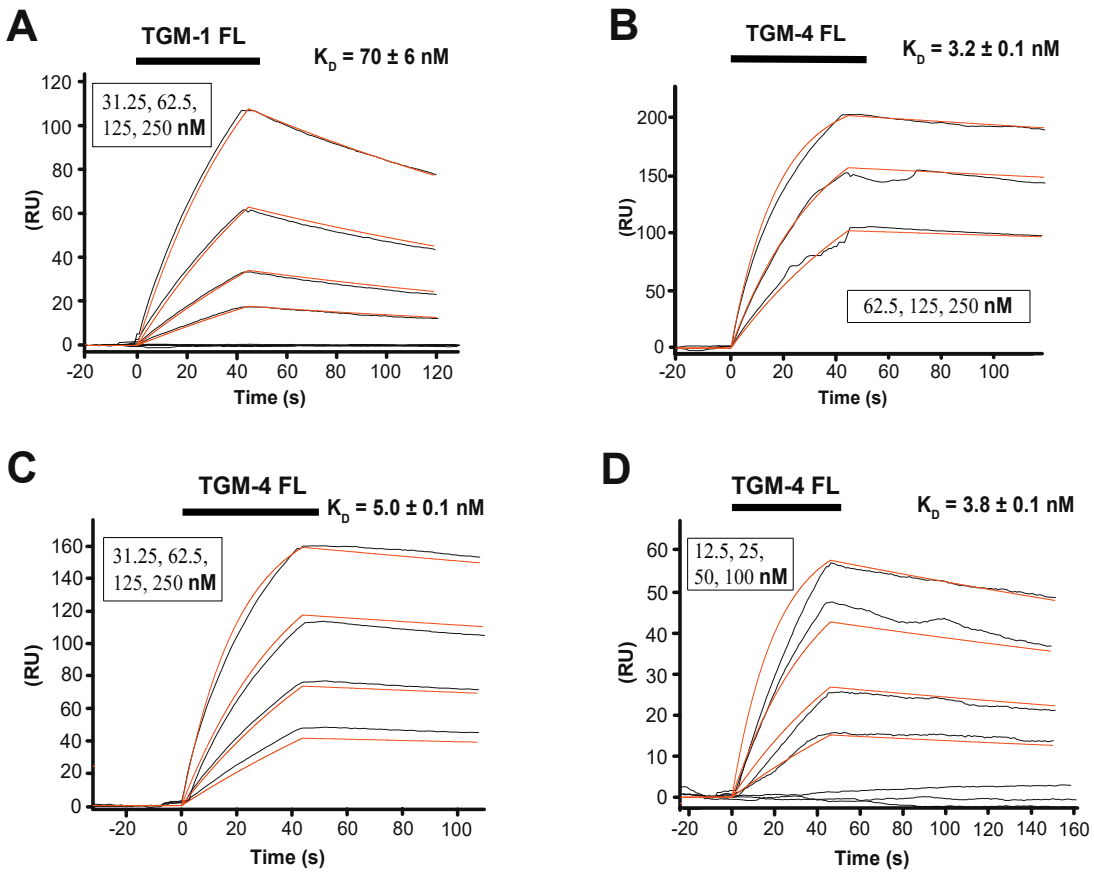

Figure 59: Binding of TGM-1 and TGM-4 to TβRI as shown by SPR. A-D. SPR sensorgrams upon injection of TGM-1 FL (A), and three repeats of TGM-4 FL (B-D) over biotinylated avi-tagged TβRI immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom right of each sensorgram.

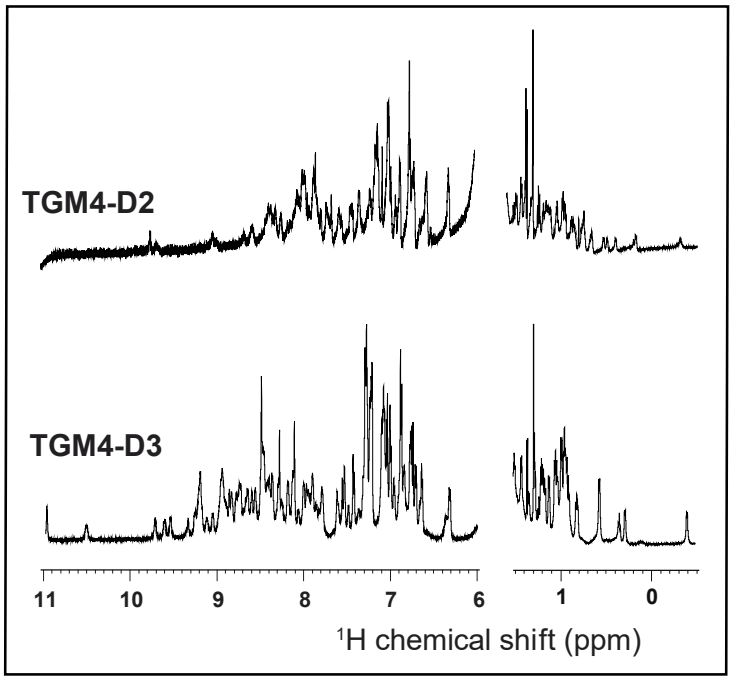

Figure 60: ^{1}H spectra of TGM-4 domains. ^{1}H NMR spectra of the amide (left) and methyl (right) regions of TGM-4 D2 and TGM-4 D3. All spectra recorded in 25 mM Na_2HPO_4, 50 mM NaCl, 5% $^{2}H_2O$, 310K.

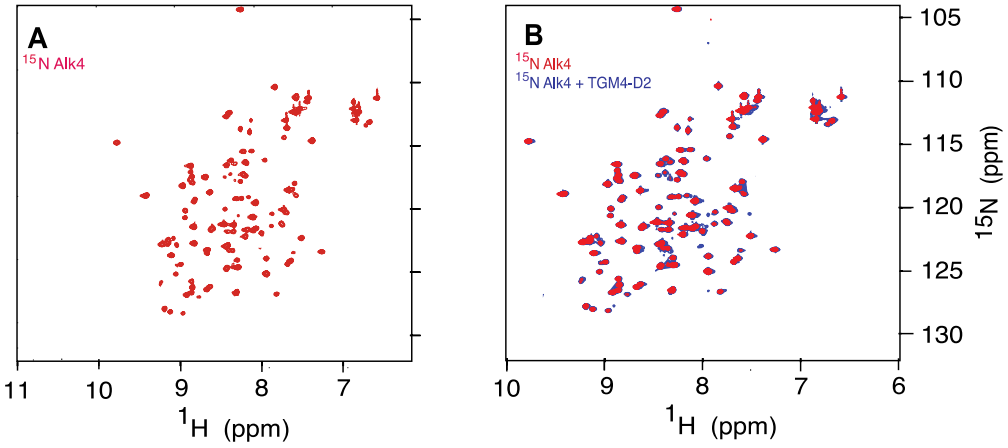

Figure 61: TGM-4 D2 does not bind to Alk4 under NMR sample conditions. ^{1}H-^{15}N spectra of ^{15}N Alk4 (A, red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.2 molar equivalents of unlabeled TGM-4 D2 (B, blue). The spectrum was recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K.

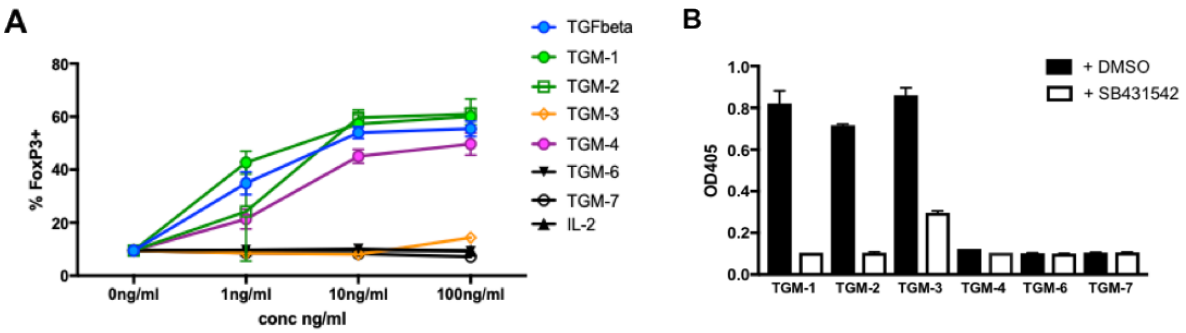

Figure 62: TGM family member conversion of T_{regs} and signaling in the MFB-F11 assay . A. Conversion of naïve CD4+ T-cells into regulatory T-cells upon administration of TGM family members in a concentration dependent manner. B. TGF-β stimulation in MFB-F11 cells as measured via a secreted phosphatase alkaline reporter upon administration of different TGM family members. All data collected by Danielle Smyth of the Maizels lab, and reported here with permission of the Maizels lab.

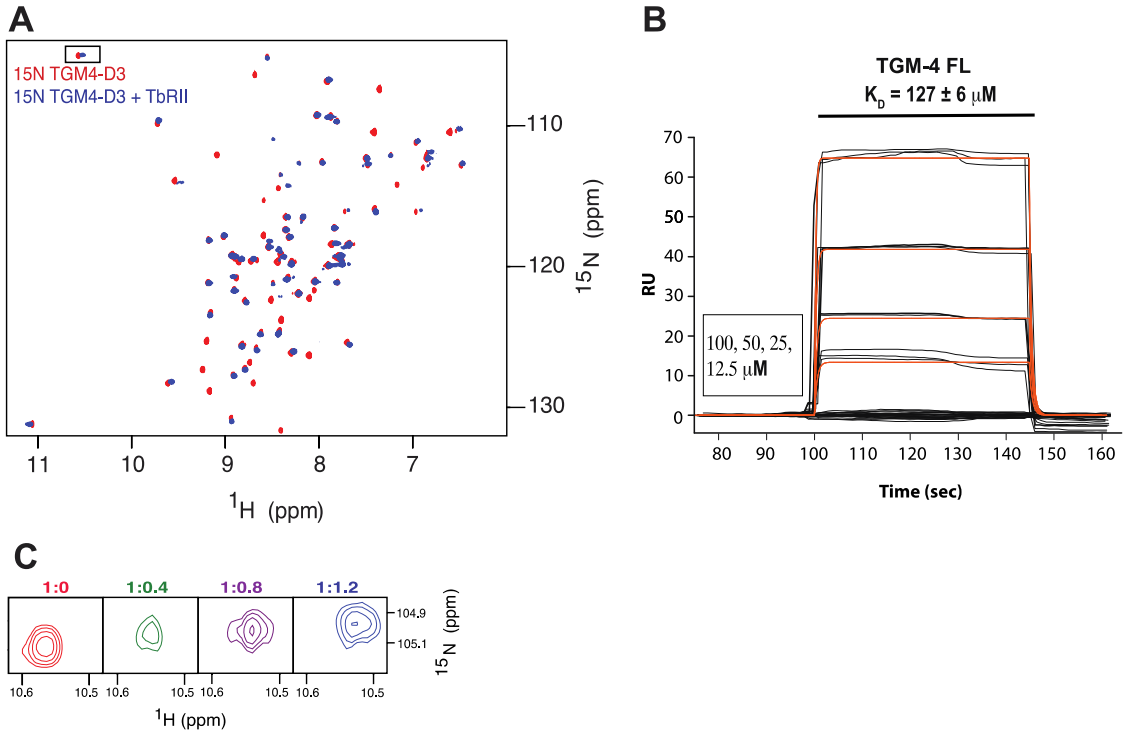

Figure 63: TGM-4 D3 binds to TβRII. A-C A. ^{1}H-^{15}N spectra of ^{15}N TGM-4 D3 (red) overlaid with the ^{1}H-^{15}N spectra of the same protein bound to 1.2 molar equivalents of unlabeled TβRII (B, blue). Expansion of intermediate titration points (1:0, 1:0.4, 1:0.8, 1:1.2 ^{15}N TGM-4 D3:TβRII) of boxed residue in panel C. The spectra were recorded in 25mM sodium phosphate, 50mM sodium chloride, 5% $^{2}H_2O$, pH 6.0, 310K. B. SPR sensorgram upon injection of TGM-4 FL over biotinylated avi-tagged TβRII immobilized on a streptavidin chip. Injections were performed as a two-fold dilution series and are shown in black, with the orange traces over the raw data showing curves fitted to a 1:1 model, when possible. The black bars over the top of the sensorgrams correlates to the injection period, and the injection concentrations are on the bottom right of each sensorgram.

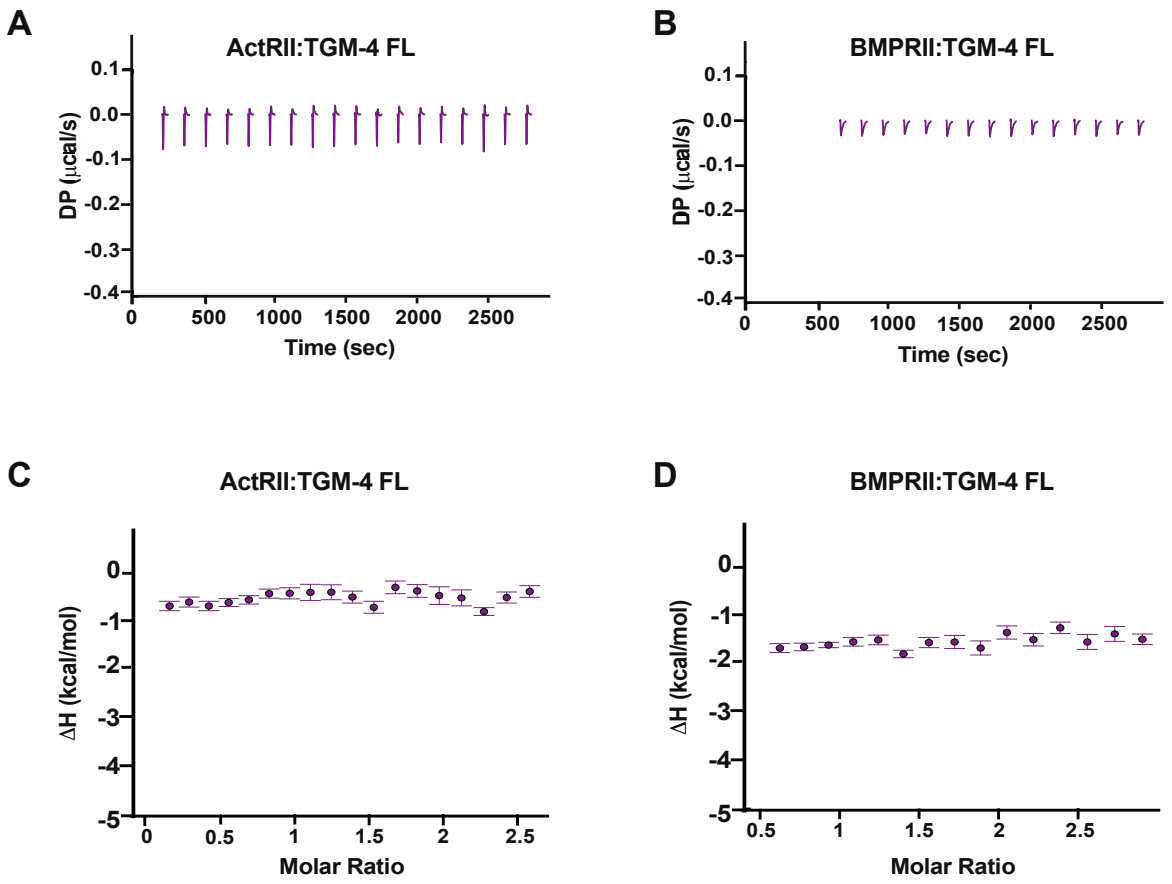

Figure 64: TGM-4 does not bind to ActRII nor BMPRII as shown by ITC. A-B. Raw thermograms for the injection of ActRII (A) or BMPRII (B) into a cell containing TGM-4 FL. Corresponding integrated heats for the aforementioned data (A-C, B-D) fit to a 1:1 binding model**.**

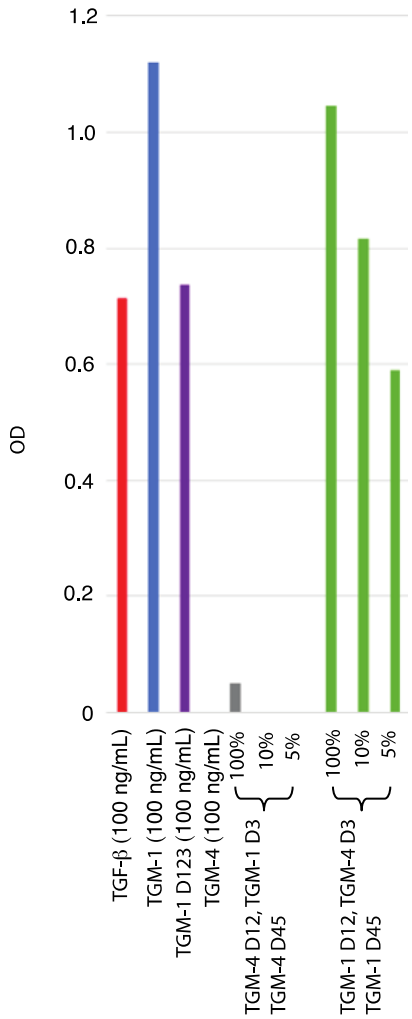

Figure 65: TGM-1 TGM-4 chimeras and their signaling in the MFB-F11 Assay. MFB-F11 assay measurement of TGF-β signaling for TGF-β (100 ng/mL, red), TGM-1 (100ng/mL blue), TGM-1 D123 (100ng/mL, purple), 100%, 10%, and 5%of cells transfected with TGM-4 D12:TGM-1 D3:TGM-4 D45 (gray), and 100%, 10%, and 5% of cells transfected with TGM-1 D12:TGM-4 D3:TGM-1D45 (green). Data collected by Tiffany Campion of the Maizels lab, and reported here with permission from the Maizels lab. Data the result of one biological replicate.

8.0 Co-receptor studies

8.1 Introduction

The previous chapters have characterized the function of proteins in the TGM family, along with determining the structure of free TGM-1 D3 and the TGM-1 D3:TβRII complex. The structure of free TGM-1 D3 confirmed that TGM-1 D3 has high structural similarities to proteins in the CCP domain family, containing several β-strands, along with conserved tryptophan residues and disulfide bonds. The differences between the structure of TGM-1 D3 and the canonical CCP domain show how TGM-1 D3 had two main adaptations: the extended hypervariable loop, and the lack of the β' and β'' strands, both of which allow for the exposure of key residues involved in binding to TβRII. The structure of the TGM-1 D3:TβRII complex shows that TGM-1 D3 closely mimics TGF-β and uses this exposed surface to engage the same structural motifs of TβRII as TGF-β.

This data, along with the high affinity binding of TGM-4 to TβRI, led to the hypothesis that TGM-1 – TGM-4 chimeras could be generated to stimulate TGF-β signaling with a potentcy that exceeded that of TGM-1. However, TGM-1 – TGM-4 chimeras which replaced the domain 3 of one TGM with the other did not significantly impact TGF-β signaling as measured via the MFB-F11 assay. Thus, it became clear that there was some portion of the mechanism that had not yet been understood that was contributing to these differences.

To account for these discrepancies, it was hypothesized that there is an alternate receptor that binds domains 4 and 5 of TGM-1, and that this alternate receptor targets TGM-1 to the cell

membrane. This would allow for a higher local concentration of TGM-1 on the cell membrane, thus increasing the protein-receptor affinity through avidity. This would explain why a chimera of TGM-1 D12:TGM-4 D3:TGM-1 D45 would still act as a TGF-β stimulator despite the low affinity of TGM-4 D3 to TβRII. Likely the avidity imbued by TGM-1 D45 would make up for the 100-fold lesser affinity of TGM-4 D3 to TβRII. Similarly, the TGM-4 D12:TGM-1 D3:TGM-4 D45 would likely have an alternate co-receptor, with domain 4 of TGM-4 being 73% identical to TGM-1 D4 and domain 5 of TGM-4 is 84% to TGM-1 D5. This alternate co-receptor would likely not be present on MFB-F11 cells and thus not increase the potency of the TGM-4 D12:TGM-1 D3:TGM-4 D45 chimera in comparison to wild-type TGM-4 FL.

Thus this chapter aims to do the following: identify the binding partner for TGM-1 domains 4 and 5, demonstrate whether this binding partner increases the inhibitory efficacy of individual TGM domain constructs, and attempt to utilize these potential co-receptor binding domains for the development of TGF-β inhibitory therapeutics.

8.2 Methods

Cloning, expression, and purification of TGM-1 D45, TGM-1 D123, TGM-4 D45, and TGM-6 D45. DNA inserts coding for TGM-1 D45, TGM-1 D123, TGM-4 D45, and TGM-6 D45 were inserted into a pSecTag vector between the AscI and ApaI sites (Table 20). Constructs were transfected into expi293 cells and cultured at 37 °C. The conditioned media was collected after 4-5 days. Purification was performed in a similar manner as described in previous sections (Chapter 3.2 Methods).

Cloning, expression, and purification of mCD44 and hCD44. DNA inserts coding for the ordered hyaluronan-binding domains of mCD44 and hCD44 (residues 21-159 NCBI XP_047283851 and residues 21-169 NCBI XP_047283851 respectively) were inserted into a pcDNA vector between the BamHI and XhoI sites (Table 20) and cloned downstream of a hexahistidine tag and a thrombin cleavage site. Constructs were transfected into expi293 cells, transfected, and cultured at 37 °C. The conditioned media as collected after 4-5 days. Purification was performed in a similar manner as described in previous sections (Chapter 3.2 Methods).

Cloning, expression, and purification of R2MPB:TGM-1 D45. DNA inserts coding for R2MPB:TGM-1 D45 fusion protein were inserted into a modified pET32a vector between the KpnI and HindIII sites (Table 20). Constructs were expressed in BL21(DE3) cells and cultured at 37 °C. When the light scattering at 600 nm was 0.8, protein expression was induced by adding 0.8 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Purification was performed in a similar manner as described for TGM-1 D3, as described in previous sections (Chapter 3.2 Methods).

Cloning, expression, and purification of 7M2R:TGM-1 D45. DNA inserts coding for 7M2R:TGM-1 D45 fusion protein were inserted into a modified pET32a vector between the NdeI and HindIII sites (Table 20). Constructs were expressed in BL21 (DE3) cells and cultured at 37 °C. When the light scattering at 600 nm was 0.8, protein expression was induced by adding 0.8 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Purification was performed in a similar manner as described for 7M2R, as described in in previous sections (Chapter 4.2 Methods).

MFB-F11 TGF-β stimulatory assays. The MFB-F11 TGF-β bioassay was previously reported by Tesseur et al.[229]. Confluent cells (8 x 10^5 cells/mL) were resuspended in DMEM with 2.5% FCS, 100 U/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine. To each cell of a 96-well plate, 4x10^4 cells were added. Purified proteins at the stated concentrations were added

to the wells (volume 50 µL) in triplicate and incubated overnight at 37 °C. After incubation, 20 µL were added to an ELISA plate with 180 µL of reconstituted Sigma FastTM p-nitrophenyl phosphate substrate and incubated at room temperature in the dark for 12 h – 18 h. Plates were read on at 405 nm on an E_{max} precision microplate reader.

HEK293T TGF-β inhibitory assays. TGF-β inhibitory assays were performed in a similar manner to the MFB-F11 TGF-β stimulatory assays, with the exception of a 30-minute incubation period with varying concentrations of 7M2R, 7M2R-45, R2MPB, or R2MPB-45 prior to overnight incubation with 10 pM TGF-β3. This HEK293T cell line utilizes a CAGA-LUC reporter line stably transfected with a TGF-β sensitive CAGA-LUC expression cassette[258].

8.3 Results

8.3.1 Domains 4 and 5 of TGM-1 play a key role in *in vitro* signaling

Due to the modular nature of the TGM domains, they can be potentially used to develop TGF-β inhibitors, by using some, but not all of the domains to sequester either TβRI or TβRI. Thus, TGM-1 D12 and TGM-1 D3 FC-linked dimers were generated and tested for inhibition, with the hypothesis being that TGM-1 D12 would be able to inhibit TGF-β and TGM-1 signaling more potently than TGM-1 D3. Inhibition of TGF-β signaling by either TGM-1 D12 or TGM-1 D3 was tested via the MFB-F11 assay. Addition of TGM-1 D12 or TGM-1 D3 in increasing amounts to MFB-F11 cells stimulated TGF-β did not inhibit TGF-β signaling at all in the case of TGM-1 D12, and only at the highest of concentrations, in the case of TGM-1 D3 (Fig. 66 A-B). This was surprising due to the two-fold order of magnitude stronger binding affinity in the case of the TGM-1 D12:TβRI interaction as compared to the TGM-1 D3:TβRII interaction. TGM-1 D3 was only able to inhibit at a concentration greater than 10 μM incongruous with its micromolar K_D. Thus the data shows that the TβRI and TβRII-binding domains of TGM-1 cannot by themselves be used as TGF-β inhibitors.

That neither domain deletion constructs were able to inhibit TGF-β signaling points to a role for domains 4 and 5 in binding a co-receptor that would target the TGM proteins to the cell surface and increase potency by avidity. This was further suggested by the TGF-β signaling functionality of TGM-1 D3 mutants. The I238A and Y252A mutants were shown by SPR to decrease binding affinity to TβRII by ca. 20-fold while the Y253A mutant nearly completely abrogated binding (Table 10, Fig. 38). The I256A mutant was shown to have no effect on binding

and as such was used as a positive control. These mutations were studied in two different contexts, one in which the mutations were made in the full-length protein, and another in which the mutations were made in a TGM construct lacking domains 4 and 5, TGM-1 D123.

In the truncated TGM-1 background (TGM-1 D123), mutagenesis of either I238A or Y252A moderately decreased TGF-β signaling potency with respect to the wild-type control (EC_{50} of as 4.6 ng mL^{-1} for the wild-type, 11.0 ng mL^{-1} for the I238A mutation, and 8.2 ng mL^{-1} for the Y252A mutation, while Y253A nearly completely abrogated signaling, except at the highest concentrations tested. The positive control, I256A, had an EC_{50} nearly the same as wild-type of 5.1 ng mL^{-1} (Fig. 67, Table 21). However, in the background of full-length TGM-1, the I238A, Y252A, and I256A mutants showed little apparent change in signaling potency in comparison to wild-type TGM-1 FL (Fig. 67, Table 22). The EC_{50} of the wild-type TGM-1 FL, I238A, Y252A, and I256 were 0.56 ng mL^{-1}, 1.73 ng mL^{-1}, 2.15 ng mL^{-1}, and 0.43 ng mL^{-1} respectively. The Y253A mutant showed a moderate decrease in potency with an EC_{50} of 4.69 ng mL^{-1}, approximately 10-fold weaker than wild-type TGM-1 FL. The clear attenuation of the effects of the TGM-1 D3 mutations in the full-length background relative to the truncated background strongly suggests that TGM-1 D45 binds to a coreceptor that amplifies TGM signaling.

8.3.2 CD44 acts as a co-receptor for TGM-1

The next step in studying was to test for the identity of the potential TGM co-receptor. This was spear-headed by collaborators in the ten Dijke lab who approached this using orthogonal methods. Initially mass spectrometry was utilized to generate a set of multiple potential co-receptors. Biotinylated TGM-1 D45 was added to NM18 cells (murine epithelial cell line) and the resultant complex was pulled from the lysate using neutravidin. Mass spectrometry-based

proteomics found the TGM-1 D45-binding receptor to be CD44, a transmembrane glycoprotein expressed near ubiquitously in human tissues, but with particularly high expression levels on immune cell types. Confirmation of binding to CD44 was performed using immunoprecipitation with an anti-CD44 antibody, which showed staining for TGM-1 FL pulldowns but not for TGM-1 D123 pulldowns.

These binding interactions however could be an artifact, with CD44 potentially interacting with TβRI, as has been previously suggested[296]. Thus the potential binding between TGM-1 and CD44 was investigated via biophysical and biochemical methods. TGM-1 was tested for binding to both mouse and human CD44 using isothermal titration calorimetry (ITC). Both mouse and human CD44 ectodomains were expressed in expi293 cells (Table 20) and purified in a manner similar to what has been described in previous methods sections. These proteins, along with TGM-1 FL, TGM-1 D45, and TGM-1 D123, were dialyzed into 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 and ITC experiments were performed with the concentrated CD44 in the syringe and with the TGM-1 constructs in the cell (Table 23, Fig. 68-69). These experiments confirmed binding of TGM-1 to CD44, with TGM-1 FL binding with a K_D of 64 nM and 92 nM for mouse and human CD44 respectively, and with TGM-1 D45 binding with a K_D of 62 nM and 140 nM respectively. Though this data does show a weaker binding affinity for human over mouse CD44, this difference is within a 68.3 % confidence interval (Table 23), and thus not largely significant. Additionally, as expected, there was no binding detected between TGM-1 D123 and either CD44 ectodomain, highlighting that TGM-1 D45 is the main binding partner for CD44. The stoichiometry was approximately 1.0 for all of these interactions, indicating a 1:1:1:1 complex of TGM-1:TβRI:TβRII:CD44. Overall this data shows that CD44 is the coreceptor for TGM-1, and binds to TGM-1 D45 with high affinity, comparable to the affinity of TGM-1 for TβRI.

8.3.3 CD44 binds to TGM-4, but not to TGM-6

A subsequent question then arose as to whether other proteins in the TGM family bound CD44. TGM-4 domains 4 and 5 are 73% and 84% identical to TGM-1 domains 4 and 5 respectively, suggesting that TGM-4 could possibly bind CD44. TGM-6 domains 4 and 5, share only 42% and 32% identity with TGM-1 domains 4-5, and thus are not as likely to bind CD44. TGM-4 and TGM-6 were therefore tested for binding to both mouse and human CD44 using isothermal titration calorimetry (ITC). TGM-4 D45, and TGM-6 D45, along with mouse and human CD44 were dialyzed into 25 mM Na_2HPO_4, 50 mM NaCl, pH 6.0 and ITC experiments were performed with the concentrated CD44 in the syringe and with the TGM constructs in the cell (Table 24, Fig. 70-71). These experiments confirmed binding of TGM-4 to CD44, with TGM-4 D45 binding with a K_D of 20 nM and 56 nM for mouse and human CD44 respectively. However, there was no measurable binding between TGM-6 and CD44 (Fig. 70-71). Though this data does show a weaker binding affinity for human over mouse CD44, this difference is within a 68.3 % confidence interval (Table 24) and thus is not largely significant. These values, additionally, are not statistically significant from that of the binding of TGM-1 D45 to CD44 (Table 20). The stoichiometry was approximately 1.0 for all of these interactions, indicating a 1:1:1:1 complex of TGM-4:TβRI:TβRII:CD44. Overall, this data shows that CD44 can bind TGM-4, though whether this acts as a coreceptor to potentiate TGM-4 function has yet to be determined. This data, however, also shows that CD44 does not bind TGM-6 D45 and thus does not likely function as a co-receptor for TGM-6.

8.3.4 Murine, but not human CD44 acts to promote TGM-induced TGF-β signaling

The next question then became as to whether the presence of CD44 would potentiate TGM-1 signaling. This is hypothesized from the data showing that the TGM-1 D3 $Y^{253}A$ mutant nearly abolished TGM-1 signaling in the background of truncated TGM-1 but not full-length TGM-1 (Fig. 68, Table 21-22). This data suggests that the binding interaction between TGM-1 D45 and CD44 is sufficient to overcome weak binding between the TGM-1 D3 $Y^{253}A$ mutant and TβRII.

The next step was to test whether human and mouse cell lines were equally responsive to TGM-1, and whether addition of the co-receptor potentiated TGM-1 signaling in multiple cell lines. TGM-1 signaling was assessed in a variety of both mouse and human cell lines via gel blot intensity of phosphorylated SMAD2, with TGF-β utilized as a control sample (Fig. 72). Of the six human cell lines and eight mouse cell lines tested for phosphorylated SMAD2, the majority of the mouse cell lines showed intense phosphorylated SMAD2 bands for both TGF-β and TGM-1 ligands, whereas the human cell lines did not or did so weakly (Fig. 72). There are multiple explanations for why this may have occurred, including a potential lack of extracellular CD44 expression in the human cell lines. However, this would not explain the weak intensity of the TGF-β-stimulated controls. For this reason, for future experiments a control cell line for which high expression of TβRI and TβRII is known, such as the MFB-F11 cell line mentioned previously, should be used, if only for there to be another positive control to determine the significance of the phosphorylated SMAD2 band intensity in other human cell lines.

Knowing that these human cell lines were less responsive to TGM-1 stimulation, the next logical step was to determine whether the addition of exogenous CD44 would increase the maximum amplitude of TGM-1 stimulation from baseline in these cell lines, or whether the reverse

would occur – that the deletion of CD44 from cell lines would decrease the amplitude of TGM-1 stimulation. NM18 cells, a murine epithelial cell line known to express CD44, were modified to be CD44 knockouts, and then tested for signaling via incubation of both TGM-1 and TGF-β. Both the wild-type and CD44 KO cells were able to signal via TGF-β, as measured using a fluorescent reporter cell line; however, with TGM-1, while it was apparent that TGM-1 was able to signal in the two CD44 KO cell lines, the signaling was closed to three times more intense in the wild-type compared to the CD44 KO cells (Fig. 73A). To test this dependence on CD44 further, HEK293 cells, a human embryonic kidney cell line with low CD44 expression, were transfected with a doxycycline-inducible CD44 construct, and tested for signaling with TGF-β or TGM-1. TGF-β signaling showed no dependence on doxycycline concentration, but TGM-1 showed an increase in fluorescence intensity as the doxycycline concentration increased (Fig. 73B). Despite the lack of a significant difference in binding affinity between TGM-1 D45:mCD44 and TGM-1 D45:hCD44, when the HEK293 cells were transfected with a doxycycline-dependent human CD44 construct, very little TGM-1 signaling was present, except at the highest concentrations of both TGM-1 and doxycycline tested. Thus, in this context, hCD44 has a lessened ability to potentiate signaling, the reason for which is yet to be determined (Fig. 73C).

8.3.5 TGM1D45 fusion constructs inhibit TGF-β signaling in HEK293T cell lines

The previous sections showed that mouse and human CD44 bind to TGM-1 D45 and TGM-4 D45, but not to TGM-6 D45. Overexpression of mouse CD44 was also shown to increase TGM-1 signaling output. Though the mechanism by which CD44 potentiates TGM-1 signaling has not been investigated, hypotheses include potentiation via avidity or via improving binding affinity to

TβRI and/or TβRII. Additionally, from the limited data it can be hypothesized that different proteins in the TGM family utilize different co-receptors, as TGM-6 D45 did not show a measurable binding affinity to CD44 by ITC (Fig. 70-71). Thus, the utility of domains 4 and 5 is potentially two-fold: the ability to potentiate TGF-β signaling through avidity, and more interestingly the ability to target the TGMs to different cell types that express the cognate co-receptor for that TGM.

TGF-β is a highly pleiotropic cytokine. Untargeted inhibition of TGF-β is undesirable due to the potential off-target side-effects. Thus, one aspect currently missing from TGF-β therapeutics is the ability to target TGF-β inhibition to a specific cell type or set of cells. This is where the TGM domains 4 and 5 may play a role. The addition of TGM-1 D45 to an existing TGF-β inhibitory therapeutic could potentiate the IC_{50} of these therapeutics in cell types with high CD44 expression. Similarly, were the coreceptor targets for other TGM isoforms identified, multiple TGF-β inhibitors with different cell targets could be generated, creating a set of cell-type specific TGF-β inhibitors.

To test this hypothesis, two separate TGF-β inhibitory molecules were utilized: mmTGF-β27M2R, which had been previously mentioned in Chapter 4, and R2MPB. mmTGF-β27M2R, otherwise called 7M2R, is an engineered monomer of TGF-β2, with residue substitutions to improve binding to TβRII to that of TGF-β1 or TGF-β3. 7M2R only binds TβRII with high-affinity, thus acting as an inhibitor of TGF-β signaling. Structurally, 7M2R acts in a similar manner to TGF-β, interacting with β4 of TβRII, along with forming electrostatic interactions with Asp55 and Glu142 of TβRII (PDB 5TX4). R2MPB, a ca. 14 kDa protein, is a computationally designed TGF-β inhibitor that acts to sequester TβRII through high-affinity binding with a similar

mechanism to that of TGF-β/7M2R. This protein was developed by collaborators in the Baker lab. Structurally, R2MPB engages the edge β-strand, β4 of TβRII, and Ile76 specifically (PDB unreleased). Due to the known binding mechanism and the ability to readily produce these proteins, these two inhibitors were selected for the generation of TGM-1 D45 fusion constructs.

7M2R and the TGM-1 D45 fused 7M2R (7M2R-45, Table 20) were purified in a similar manner. *E. coli* cells were transformed with a pET32a plasmid and cells were grown at 37 °C. At OD 0.8 the cells were induced with 200 mg L^{-1} IPTG and spun down after six hours. The cell pellets were solubilized, and the inclusion bodies solubilized in 8M urea. This solution was reduced with 50 mM DTT, and after acidification, an initial purification was performed using an SP sepharose column (pH 4.2) with a salt gradient (0 – 0.5M NaCl). Appropriate fractions were identified by SDS-PAGE gel, and the combined fractions were folded at a final concentration of 0.1 mg mL^{-1} in a solution of 100 mM Tris, 30 mM CHAPS, 5mM reduced glutathione, 1M NaCl, 20% w/v DMSO pH 9.5. The folding mixture was concentrated after 48 hours at 4 °C and final purification was performed via ion-exchange chromatography and reverse phase chromatography.

R2MPB and the TGM-1 D45 fused R2MPB (R2MPB-45, Table 20) were also purified similarly. *E. coli* cells were transformed with a pET32a plasmid and cells were grown at 37 °C. At OD 0.8 the cells were induced with 200 mg L^{-1} IPTG and spun down after six hours. The cell pellets were solubilized and the inclusion bodies were solubilized in 8M urea. The protein was initially purified using a NiNTA column (pH 8.0) with an imidazole elution gradient (0 – 0.5M imidazole). Appropriate fractions were identified by SDS-PAGE gel and the combined fractions folded at a concentration of 0.1 mg mL^{-1} in a solution of 100 mM Tris, 1 mM EDTA, 0.5 mM oxidized glutathione, pH 8.0 after initial reduction with 2 mM reduced glutathione. The folding

mixture was concentrated after 12 hours at 4 °C and final purification was performed via ion-exchange chromatography.

All of these proteins were produced in minimal media with ^{15}N isotopic labeling and thus confirmation of native folding and identity was performed using ^{15}N-^{1}H HSQC. To attain an NMR spectrum with well-dispersed signals, the 7M2R-D45 was prepared in 10 mM Na_2HPO_4, 20 mM CHAPS pH 4.7 and the spectrum recorded at 310K. Two samples were taken from the final chromatogram and assessed for native folding. 7M2R-D45 is a 29 kDa protein with 259 residues. While not all of these residues were visible on the HSQC spectra, both samples assessed showed a wide dispersion of signals ranging from 6 – 11 ppm on the ^{1}H axis (Fig. 74). Similarly, R2MPB-D45 was buffered at 25mM CHES 50mM NaCl pH 8.5 and the spectrum recorded at 310K. Two samples were taken from the final chromatogram and assessed for native folding. R2MPB-D45 is a 30 kDa protein with 274 residues. While not all of these residues were visible on the HSQC spectra, both samples assessed showed a wide dispersion of signals ranging from 6 – 11 ppm on the ^{1}H axis (Fig. 75).

Whether the addition of domains 4 and 5 of TGM-1 would increase the inhibitory capacity of 7M2R and R2MPB was assessed using a TGF-β CAGA-LUC reporter assay in which HEK293T cells were stimulated with 10 pM TGF-β3 and increasing amounts of inhibitor added. All of the samples tested showed inhibition of TGF-β signaling to varying degrees. The control 7M2R protein inhibited at an IC_{50} of 31.6 nM (95% CI 18.4 – 50.9 nM), and the two 7M2R-D45 samples inhibited at an IC_{50} of 11.6 nM (95% CI 9.2 – 15.5 nM) and 44.1 nM (95% CI 21.8 – 124.1 nM) respectively (Fig. 76A, Table 25). The confidence intervals suggest that there is little significant difference in the IC_{50} of the 7M2R and the 7M2R-D45 samples, which corresponds with the low cell surface expression of CD44 on HEK293T cells. Similarly, the control R2MPB protein

inhibited at an IC_{50} of 270.7 nM (95% CI 107.8 – 728.5 nM), and the two R2MPB-D45 samples inhibited at an IC_{50} of 9.7 nM (5.1 – 19.8 nM) and 26.6 nM (95% CI 14.1 – 48.6 nM) (respectively (Fig. 76B, Table 25). While it appears as though the R2MPB-D45 samples are much more potent than R2MPB alone, based on the reported IC_{50} values, it is clear that the nonlinear fit to the data was suboptimal, and thus this experiment should be repeated for a more accurate quantification of the R2MPB IC_{50}.

Thus this data reveals that 7M2R and R2MPB are both capable of inhibiting TGF-β signaling, and that addition of the fused TGM-1 D45 domains does not negatively impact the ability of either protein to inhibit TGF-β signaling in a low CD44 expression cell line. The CD44 expression on the HEK293T cells is likely too low to significantly impact the IC_{50}, which would explain why the D45 domains did not potentiate inhibition for either fused construct. In addition, the 7M2R may be a strong enough inhibitor that addition of domains 4-5 would not increase inhibition, particularly if the cell expression of CD44 is low. Further studies into the inhibition of TGF-β signaling by these constructs are needed, including measurement of binding affinity of R2MPB and 7M2R to TβRII, as well as testing of these fused constructs in a variety of cell types, including those with higher levels of CD44 expression, before any definitive statements can be made. It could also be that since HEK293T cells are human cells, and that previous data has shown that potentiation of signaling only occurs with mouse CD44 that switching to a murine cell line may show more signal potentiation.

8.4 Discussion

It was hypothesized based on the results presented in earlier chapters that TGM-1 could be isolated into its modular domains and be targeted to bind one of the two TGF-β receptors: TβRI for domains 1-2 and TβRII for domain 3. Binding only one of the two TGF-β receptors would thus lead to signaling inhibition, with domain deletion constructs of TGM-1 acting as TGF-β inhibitory molecules. However, when TGM-1 D12 and TGM-1 D3 Fc-linked dimers were generated and tested for inhibition in a TGF-β signaling assay, none of the constructs were shown to inhibit signaling, with the TGM-1 D3 dimer construct only inhibiting at the maximal concentration which was orders of magnitude greater than its binding affinity to TβRII (Fig. 64).

Following this, TGM-1 D3 mutations were made in the full-length TGM-1 background and a truncated (TGM-1 D123) background, which showed that only in the truncated TGM-1 D123 background did the mutations significantly impact the EC_{50}. TGM-1 D3 $Y^{253}A$ mutation in the full-length background weakened the EC_{50} by four-fold whereas in the truncated background this same mutation nearly abrogated binding (Fig. 65, Table 21-22). This data, along with the lack of inhibition of TGF-β signaling by the TGM-1 D12 and TGM-1 D3 dimers, suggested a role for domains 4 and 5 of TGM-1. Thus, it was hypothesized that TGM domains 4 and 5 played an alternate role that would allow for increased signaling potency in an *in vitro* context. Specifically, it was hypothesized that TGM-1 domains 4 and 5, and likely domains 4 and 5 of other TGM family proteins, was involved in binding a co-receptor that would target TGM-1 to the cell surface, increasing the local concentration of TGM-1 through avidity and improving potency.

The putative co-receptor was identified by mass spectrometry and immunoprecipitation to be CD44, a transmembrane glycoprotein expressed heavily on immune cell types. Using ITC it

was confirmed that CD44 binds to TGM-1 D45 with a high affinity, and does not bind to TGM-1 D123 (Table 23). Both mouse and human CD44 were shown to bind to TGM-1 D45, with ca. 60 nM and 90 nM affinity respectively (Table 23, Fig. 69-70). The difference in binding affinity between the mouse and human CD44 proteins for TGM-1 D45 was shown to be minimal, with TGM-1 D45 binding with a slightly weaker affinity to human CD44. Mouse CD44 was additionally shown through cell studies to potentiate TGM signaling in a concentration-dependent manner (Fig. 73). Interestingly, while TGM-4 D45 was shown to be able to bind to human and mouse CD44, TGM-6 D45 was not, with studies in-progress to identify coreceptors for TGM-4 D45 and TGM-6 D45 under way (Table 23, Fig. 71-72). This data led to the hypothesis that different members of the TGM family bind to different co-receptors, and that these co-receptors target the parasitic protein to different cell types depending on the expression of the co-receptor.

As the TGM proteins are modular with separable domains, and as the D45 domains are hypothesized to target the TGM family members to different cell types, it was hypothesized that conjugating TGF-β inhibitors to TGM-1 D45 would improve the inhibitory capacity of these inhibitors. Specifically, two inhibitors, R2MPB and 7M2R, both of which bind to TβRII, were studied. These two proteins were purified natively as a fusion protein containing an additional TGM-1 D45 separated by a short linker (Fig. 75-76). These proteins were then tested for TGF-β inhibitory capacity in comparison to the parent proteins without fusion to the TGM-1 D45 fusion. Neither fusing TGM-1 D45 to 7M2R nor fusing TGM-1 D45 to R2MPB showed any gross impact on IC_{50} in HEK293T cells, neither weakening nor strengthening the IC_{50} (Fig. 77), likely due to the low cell surface CD44 expression. Future experiments would thus involve testing in cells with higher expression of CD44, such as the MFB-F11 cell line. This chapter opens up questions as to the potential use of TGM domains in acting as targeting mechanisms for TGF-β inhibitors. TGF-

β inhibition has been shown to be broadly ineffective due to pleiotropic effects depending on the cell-context. Development of fusion inhibitors utilizing the domains of TGM-1 could potentially eliminate this problem, as it would direct the inhibitor to a cell type with high levels of expression of the co-receptor of interest, thus eliminating these pleiotropic effects.

8.5 Respective Contributions

The experiments in this chapter were performed in combination by the author Ananya Mukundan and by collaborators. The inhibition of the TGF-β inhibitory pathway by TGM-1 D12 and by TGM-1 D3, was performed by Cynthia Hinck of the Hinck laboratory (Fig. 67). TGM-1 binding assays in Fig. 68 were performed by Kyle Cunningham of the Maizels laboratory. The ITC studies (Fig. 69-72) were performed by Ananya Mukundan, as was the purification of the R2MPB-D45 and the 7M2R-D45 constructs (Fig. 75-76). The cell studies assessing the ability of mouse and human CD44 to potentiate TGM signaling were performed by Maarten van Dinther of the ten Dijke laboratory (Fig. 73-74), as was the identification of CD44 as the TGM-1 co-receptor. Finally, the 7M2R-D45 and the R2MPB-D45 inhibition assays in HEK293T cells were performed by Cynthia Hinck of the Hinck lab (Fig. 77).

8.6 Tables and Figures

Table 20: Sequences of constructs used in this chapter.

Construct	Coding region and description (* indicates stop codon)
TGM-1 D45	Residues 263-422 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Signal Peptide-TGM-D45-Linker-Myc-Linker-His_6 METDTLLLWVLLLWVPGSTGDAAQPARRACKPLEANESVHYEYFTMTNETDKKKGPPAKVGKSGKYPEHTCVKKVCSKW PYTCSTGGPIFGECIGATWNFTALMECINARGCSSDDLFDKLGFEKVIVRKGEGSDSYKDDFARFYATGSKVIAECGGKTVR LECSNGEWHEPGTKTVHRCTKDGIRTLGPEQKLISEEDLNSAVDHHHHHH*
TGM-1 D123, cleavable His tag	Residues 16-262 of *H. Polygyrus* TGF-β Mimic, NCBI MG099712 Signal Peptide- His_8- Linker-Thrombin Cleavage Site-TGM-D1-D3 MKWVTFLLLLFISGSAFSGSGSHHHHHHHHHGSLVPRGSDDSGCMPFSDEAATYKYVAKGPKNIEIPAQIDNSGMYPDYTHV KRFCKGLHGEDTTGWFVGICLASQWYYYEGVQECDDRRCSPLPTNDTVSFEYLKATVNPGIIFNITVHPDASGKYPELTYIK RICKNFPTDSNVQGHIIGMCYNAEWQFSSTPTCPASGCPPLPDDGIVFYEYYGYAGDRHTVGPVVTKDSSGNYPSPTHARRR CRALSQEADPGEFVAICYKSGTTGESHWEYYKNIGKCPDP*
mCD44	Residues 23-174 of mouse CD44 isoform X4, NCBI XP_006498712 Signal Peptide- His_8- Linker-Thrombin Cleavage Site-mouse CD44 MKWVTFLLLLFISGSAFSGSGSHHHHHHHHHGSLVPRGSHQQIDLNVTCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPT MDQMKLALSKGFETCRYGFIEGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNASAPPEEDCTSVTDLPNSFDGPVT ITIVNRDGTRYSKKGEYRTHQEDID*
hCD44	Residues 21-169 of human CD44 isoform X27, NCBI XP_054226560 Signal Peptide- His_8- Linker-Thrombin Cleavage Site-human CD44 MKWVTFLLLLFISGSAFSGSGSHHHHHHHHHGSLVPRGSQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMA AQMEKALSIGFETCRYGFIEGHVVIPRIHPNSICAANNTGVYILTSNTSQYDTYCFNASAPPEEDCTSVTDLPNAFDGPITITIV NRDGTRYVQKGEYRTNPEDIY*
TGM-4 D45	Residues 263-422 of *H. Polygyrus* TGF-β Mimic 4, NCBI MG429739 Signal Peptide-TGM-4 D45-Linker-Myc-Linker-His_6 METDTLLLWVLLLWVPGSTGDAAQPARRARCKPLETNVSVHYEYFTMTNETGRKEGTPAEVDKGGKYPQHTCVRKFCDK SPYTCSVKGPIFGECLDGQWNFTALDECLNARGCNSDDLFDKLGFEGVMVREEEGSDSYKDDFVRFYATGSKVNAECKGK TVQLECSDGEWHDPGTKTVHRCTKEGIRALGPEQKLISEEDLNSAVDHHHHHH*
TGM-6 D45	Residues 103-254 of *H. Polygyrus* TGF-β Mimic 6, NCBI MG429741 Signal Peptide-TGM-6 D45-Linker-Myc-Linker-His_6 METDTLLLWVLLLWVPGSTGDAAQPARRA RCKPLEKNDSVSYEYFTKPTKGLKMGSITKPDKSGKYPEETFVRRYCNDLPRNSLAQGKTYAECLDSEWKLKNLPDCRFAA GCDEEYLLEKLMFVDISYWGKDAAKFSDDKTYRYYRPGSKVTAKCKGKSVKLTCVDGGYWVTVDGRKALCTGPEQKLIS EEDLNSAVDHHHHHH*
R2MPB- TGM-1 D45	Residues 1-101 of synthetic TβRII binding protein fused to Residues 263-423 of *H. polygyrus* TGF-β Mimic, with residues 243-246 substituted with a lanthanide-binding tag, NCBI MG099712 Thioredoxin-His_6-Linker-Thrombin Cleavage Site-Linker-R2MPB-Linker-TGM-1 D45 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFK NGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPRGSGTGLKELLKELNKAIASGDTETVRRILEELLEL LKEAFEKGDYDLAISIASMAVKAASYIGDTETLKELLEILKKIKEKLKKEGDEAALKAVERNIKVVEKVAGGSGGSGGRCKP LEANESVHYEYFTMTNETDKKKGPPAKVGKSGKYPEHTCVKKVCSKWPYTCSTGGPIFGECIGATWNFTALMECINARGCS SDDLFDKLGFEKVIVRKGEGSDSYKDDFARFYATGSKVIAECGGKTVRLECSNGEWHEPGTKTVHRCTKDGIRTL
7M2R-TGM-1 D45	Residues 331-442 of human TGF-β2 proprotein, with deletion of residues 381-400 and mutation of the following residues K357R, R358K, L382R, A406K, C409S, L421V, I424V, N426R, T427K, I430V, NCBI NP_001129071 fused to Residues 263-423 of *H. polygyrus* TGF-β Mimic, with residues 243-246 substituted with a lanthanide-binding tag, NCBI MG099712 Initiating Methionine-7M2R- Linker-TGM-1 D45 MALDAAYCFRNVQDNCCLRPLYIDFRKDLGWKWIHEPKGYNANFCAGACPYRASKSPSCVSQDLEPLTIVYYVGRKPVEQ LSNMIVKSCKCSGSGTGNRCKPLEANESVHYEYFTMTNETDKKKGPPAKVGKSGKYPEHTCVKKVCSKWPYTCSTGGPIFG ECIGATWNFTALMECINARGCSSDDLFDKLGFEKVIVRKGEGSDSYKDDFARFYATGSKVIAECGGKTVRLECSNGEWHEP GTKTVHRCTKDGIRTL

Table 21: TGF-β Signaling of mutated TGM-1 D123 constructs

Construct	TGM-1 WT	TGM-1 I238A	TGM-1 Y252A	TGM-1 Y253A	TGM-1 I256A
Concentrations (ng mL^{-1})	0.0625, 0.625, 1.25, 2.5, 25, 50, 100	0.0625, 0.625, 1.25, 2.5, 25, 50, 100	0.0625, 0.625, 1.25, 2.5, 25, 50, 100	0.0625, 0.625, 1.25, 2.5, 25, 50, 100	0.0625, 0.625, 1.25, 2.5, 25, 50, 100
EC50 (ng mL^{-1})	4.62	11.03	8.19	ND[b]	5.07
EC50 95% CI (ng mL^{-1})	(3.45, 6.38)	(8.32, 14.70)	(5.91, 11.58)	ND[b]	(3.91, 6.75)
[a]Global fit of three replicates [b]Not determined due to weak signal					

Table 22: TGF-β Signaling of mutated TGM-1 FL constructs

Construct	TGM-1 WT	TGM-1 I238A	TGM-1 Y252A	TGM-1 Y253A	TGM-1 I256A
Concentrations (ng mL^{-1})	1, 2.5, 5, 10, 25, 50, 100	1, 2.5, 5, 10, 25, 50, 100	1, 2.5, 5, 10, 25, 50, 100	1, 2.5, 5, 10, 25, 50, 100	1, 2.5, 5, 10, 25, 50, 100
EC50 (ng mL^{-1})	0.56	1.73	2.15	4.69	0.43
EC50 95% CI (ng mL^{-1})	(0.17, 1.06)	(0.95, 2.86)	(1.43, 3.10)	(3.51, 6.24)	(0.16, 0.75)
[a]Global fit of three replicates					

Table 23: TGM-1 binding to mouse and human CD44 as assessed by isothermal titration calorimetry

Cell	mCD44	hCD44	mCD44	hCD44	mCD44	hCD44
Syringe	TGM-1 FL	TGM-1 FL	TGM-1 D45	TGM-1 D45	TGM-1 D123	TGM-1 D123
Cell concentration (μM)	15	10	12	8	10	10
Syringe concentration (μM)	100	100	90	100	90	100
Temperature (°C)	35	35	35	35	35	35
K_D (nM)	64 (28, 123)[ab]	92 (32, 219)[ab]	62 (28, 117)[ab]	140 (86, 221)[ab]	ND[c]	ND[c]
ΔH (kcal mol^{-1})	-17.5 (-18.8, -16.3)[ab]	-10.3 (-11.8, -9.1)[ab]	-19.4 (-20.5, -18.4)[ab]	-9.6 (-10.2, -9.0)[ab]	ND[c]	ND[c]
ΔG (kcal mol^{-1})	-10.1[ab]	-9.9[ab]	-10.2[ab]	-9.7[ab]	ND[c]	ND[c]
-TΔS (kcal mol^{-1})	7.3[ab]	0.4[ab]	9.2[ab]	-0.1[ab]	ND[c]	ND[c]
Stoichiometry (n)	0.8[d]	1.2[d]	1.4[d]	1.2[d]	ND[c]	ND[c]

[a]Uncertainty reported as 68.3% confidence interval
[b]Global fit of two replicates
[c]Not determined due to weak signal
[d]Number of sites determined by incompetent fraction value on sedphat; set to '1' for K_D analysis

Table 24: TGM-4 and TGM-6 binding to mouse and human CD44 as assessed by isothermal titration calorimetry

Cell	mCD44	hCD44	mCD44	hCD44
Syringe	TGM-4 D45	TGM-4 D45	TGM-6 D45	TGM-6 D45
Cell concentration (μM)	5.5	5	8	10
Syringe concentration (μM)	60	75	90	100
Temperature (°C)	35	35	35	35
K_D (nM)	20 (7, 41)[ab]	56 (28, 99)[ab]	ND[c]	ND[c]
ΔH (kcal mol^{-1})	-21.7 (-23.0, -20.5)[ab]	-16.6 (-17.9, -15.5)[ab]	ND[c]	ND[c]
ΔG (kcal mol^{-1})	-10.9[ab]	-10.2[ab]	ND[c]	ND[c]
-TΔS (kcal mol^{-1})	10.8[ab]	6.4[ab]	ND[c]	ND[c]
Stoichiometry (n)	1.2[d]	1.5[d]	ND[c]	ND[c]

[a]Uncertainty reported as 68.3% confidence interval
[b]Global fit of two replicates
[c]Not determined due to weak signal
[d]Number of sites determined by incompetent fraction value on sedphat; set to '1' for K_D analysis

Table 25: Inhibition of TGF-β signaling by 7M2R and R2MPB Fusion Constructs

Inhibitor	7M2R	7M2R-45 Sample 1	7M2R-45 Sample 2	R2MPB	R2MPB-45 Sample 1	R2MPB-45 Sample 2
Inhibitor Concentrations (nM)	0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000	0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000	0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000	0.1, 0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000	0.1, 0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000	0.1, 0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200, 10000
IC50 (nM)	31.56	11.63	44.15	270.70	9.83	26.34
IC50 95% CI (nM)	(18.40, 50.88)	(9.2, 15.50)	(21.80, 124.1)	(107.8,728.5)	(5.09, 19.76)	(14.10, 48.60)
[a]TGF-β concentration at 10 pM [b]Fit of three technical replicates						

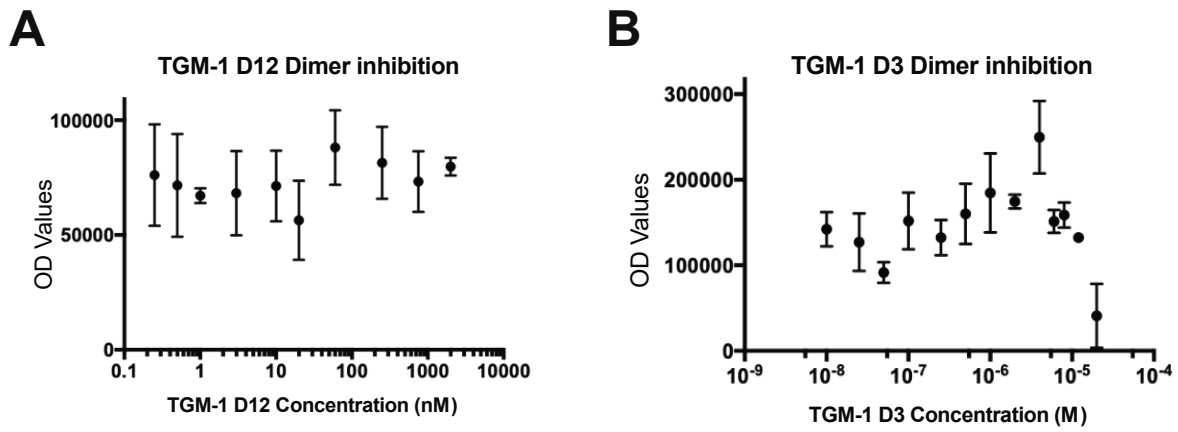

Figure 66: TGM-1 domain deletion dimers do not inhibit TGF-β signaling . A-B. Inhibition of TGF-β signaling as measured via the MFB-F11 assay, with constant amounts of TGF-β (10pM) and increasing amounts of dimeric TGM-1 D12 (A) or dimeric TGM-1 D3 (B). Data collected and analyzed by Dr. Cynthia Hinck, Hinck Lab.

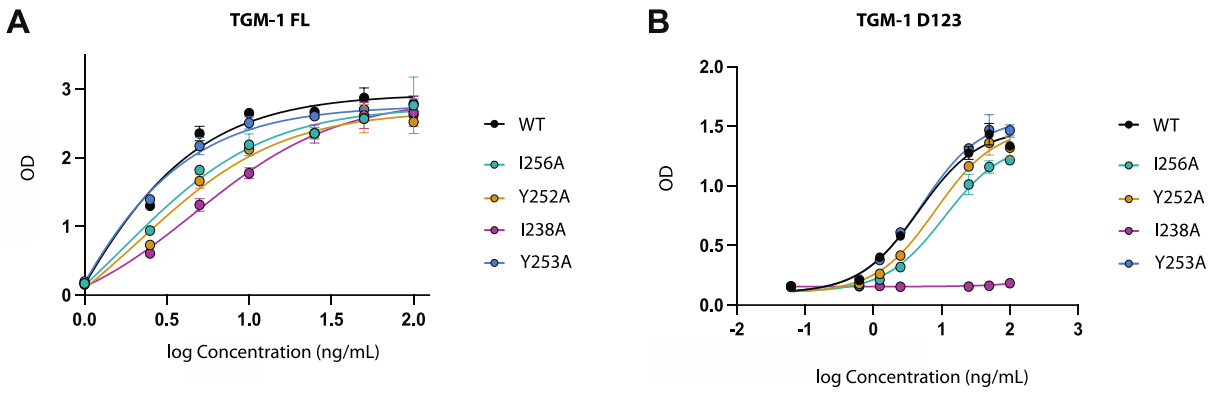

Figure 67: TGM-1 D3 mutations functionally impact TGM-1 signaling. A-B. Effect of TGM-1 D3 mutations on TGM-1 FL (A) or TGM-1 D123 (B) signaling, measured via the MFB-F11 TGF-β responsivve bioassay. Relevant mutations are noted in the key. Adapted from "Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII" by Mukundan et. al.,[257] used under CC By 4.0 / Edited from original

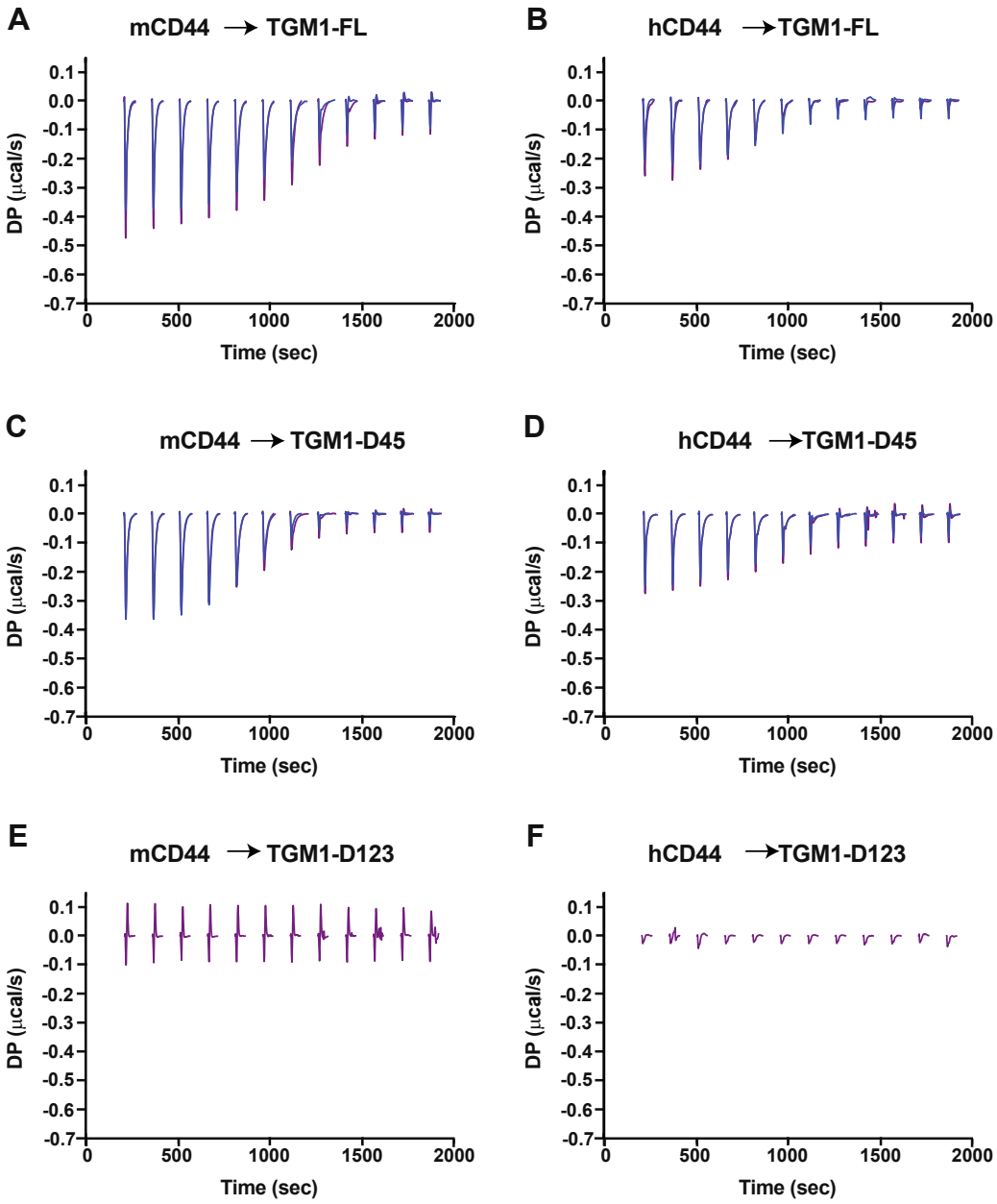

Figure 68: Raw thermograms of TGM-1 binding to mouse and human CD44 by ITC. A-F. Thermogram for the injection of mCD44 into TGM-1 FL (A), TGM-1 D45 (C), TGM-1 D123 €, or the injection of hCD44 into TGM-1 FL (B), TGM-1 D45 (D), or TGM-1 D123 (F).

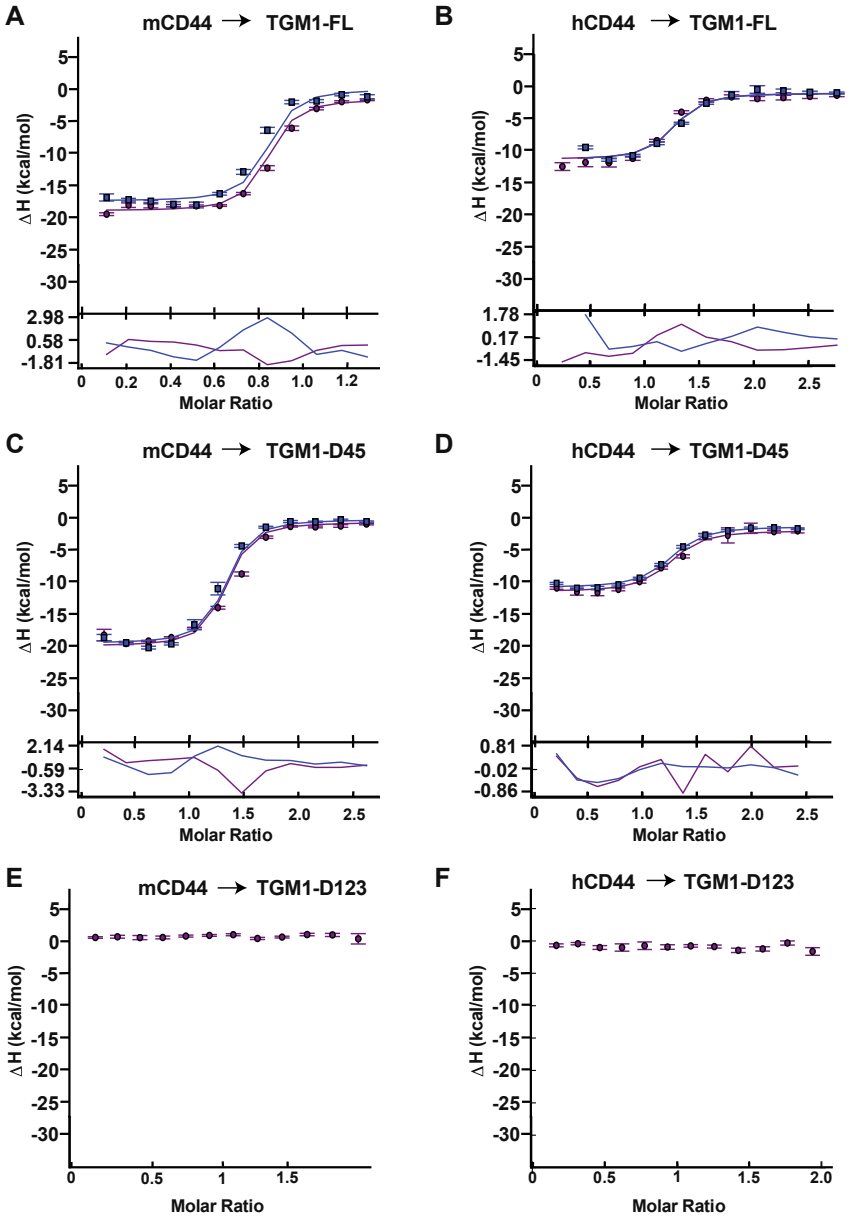

Figure 69: Isotherms of TGM-1 binding to mouse and human CD44 by ITC. A-F. Integrated heats for the injection of mCD44 into TGM-1 FL (A), TGM-1 D45 (C), TGM-1 D123 €, or the injection of hCD44 into TGM-1 FL (B), TGM-1 D45 (D), or TGM-1 D123 (F), together with the fit (smooth line) and residuals (below) to a 1:1 binding model. Error bars indicate bias in the NITPIC estimation of the integrated heats.

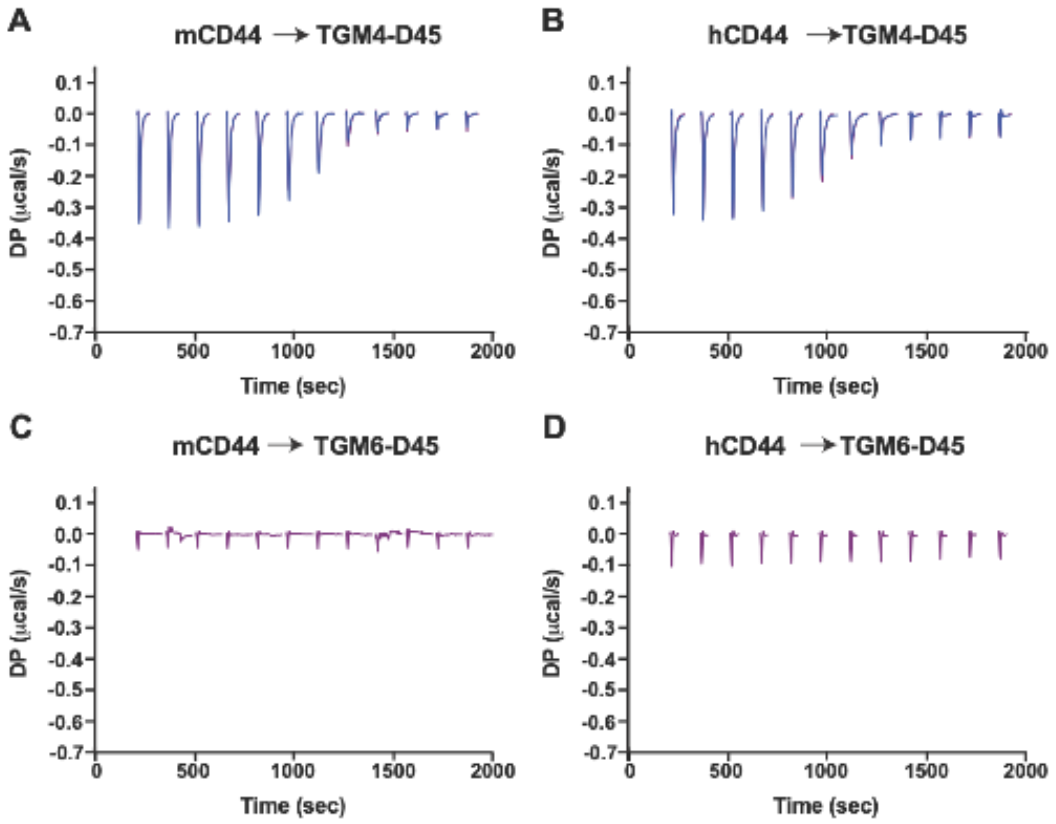

Figure 70: Raw thermograms of TGM-4 and TGM-6 binding to mouse and human CD44 by ITC. A-D. Thermogram for the injection of mCD44 into TGM-4 D45 (A), TGM-6 D34 (B) or the injection of hCD44 into TGM-4 D45 (C), TGM-6 D45 (D).

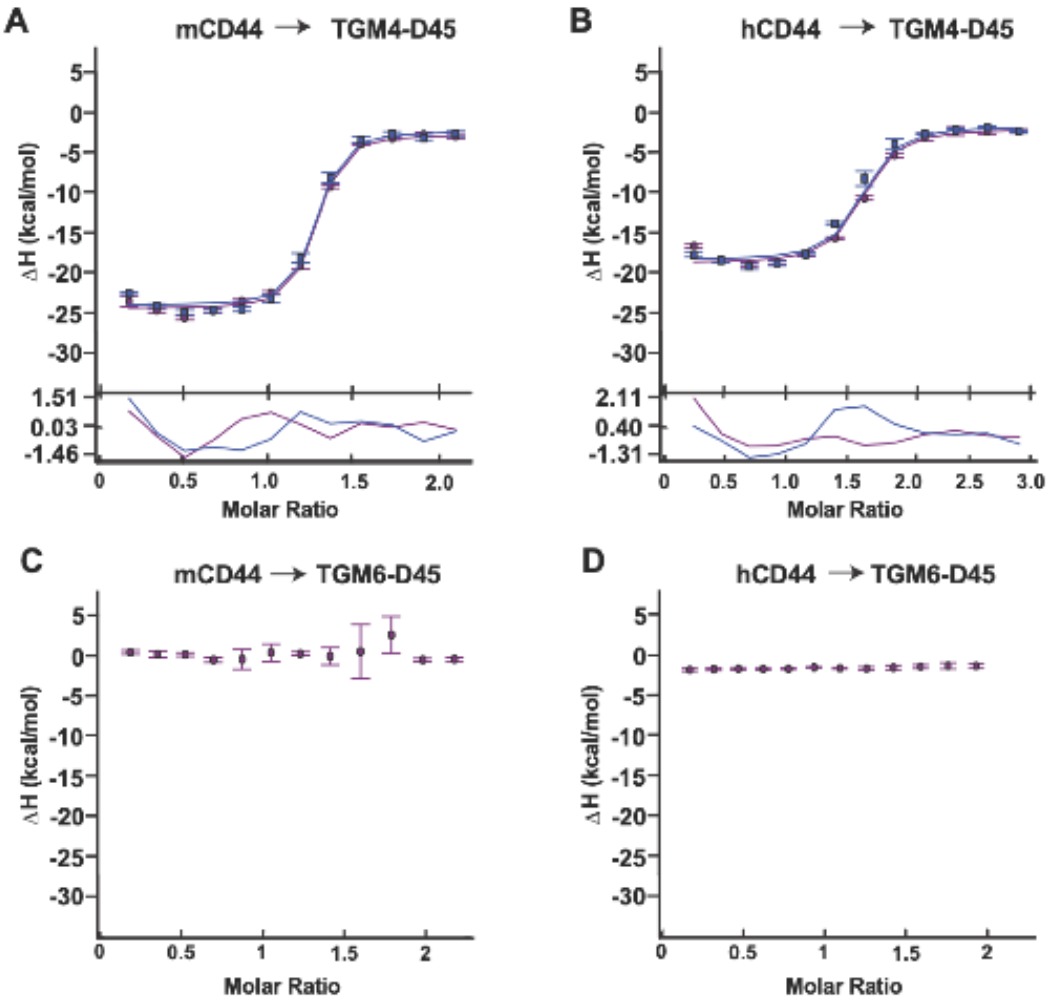

Figure 71: Isotherms of TGM-4 and TGM-6 binding to mouse and human CD44 by ITC. A-D. Integrated heats for the injection of mCD44 into TGM-4 D45 (A), TGM-6 D45 (C), or the injection of hCD44 into TGM-4 D45 (B), TGM-6 D45 (D), together with the fit (smooth line) and residuals (below) to a 1:1 binding model. Error bars indicate bias in the NITPIC estimation of the integrated heats.

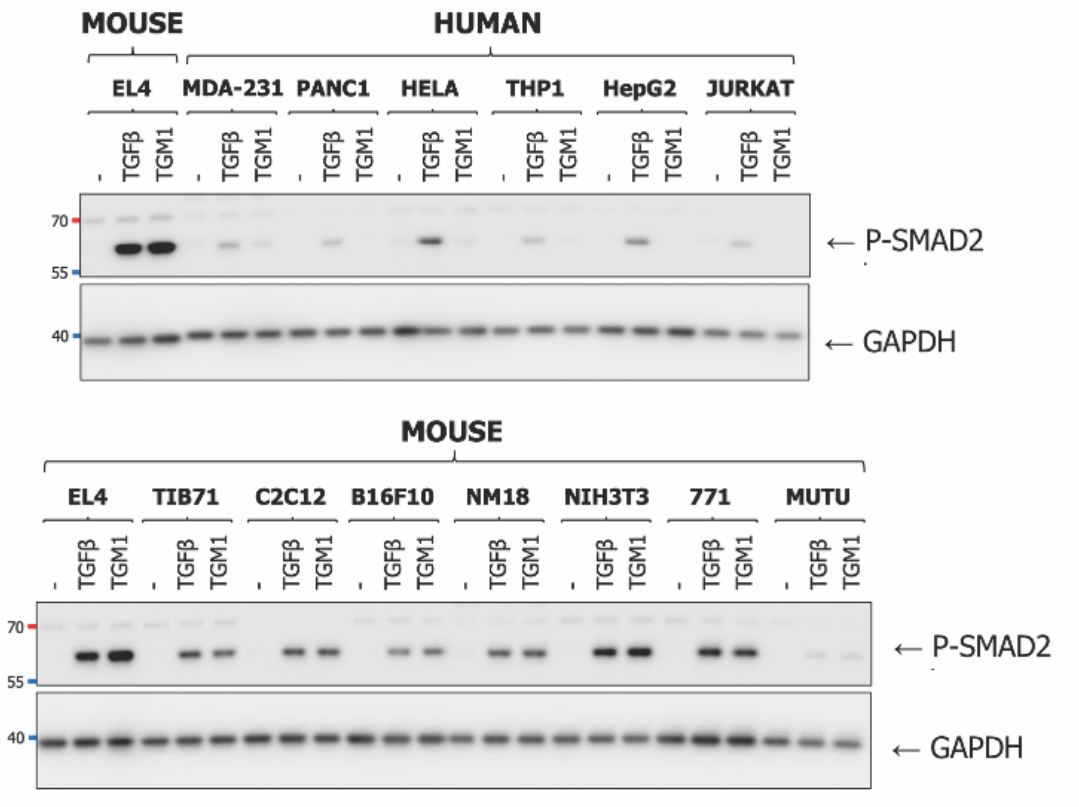

Figure 72: TGM-1 and TGF-β signaling in various mouse and human cell lines. Cell lines incubated with either media, TGF-β (1ng mL^{-1}), or TGM-1 (10 ng mL^{-1}) for 1 hour, and the resultant lysates identified on Western Blot using an anti-phospho SMAD antibody or a control GADPH antibody. Data collected and analyzed by the Maizels lab and the ten Dijke lab submitted for revision in PNAS under the title "CD44 acts as a co-receptor for cell-specific enhancement of signaling and regulatory T cell induction by TGM1, a parasite TGF-β mimic"

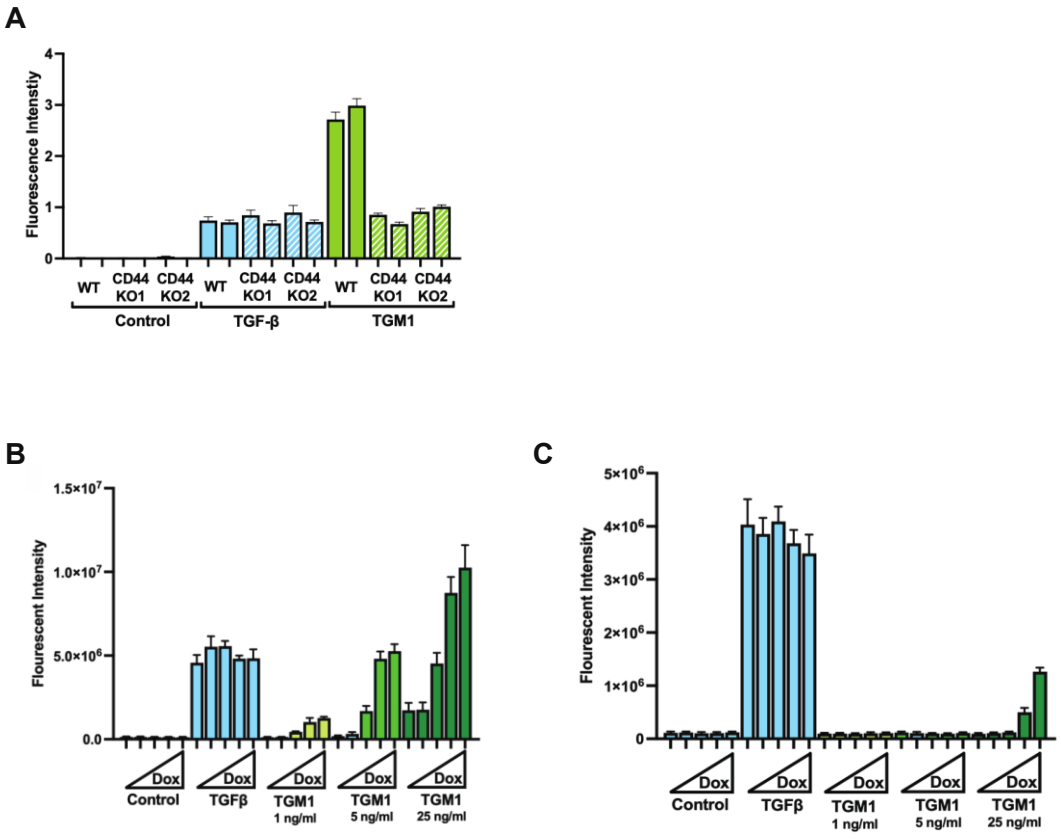

Figure 73: Role of CD44 in TGF-β and TGM-1 signaling in vitro. A-C. A. mCHERRY reporter NM18 cells, CD44 KO or WT, incubated with either control media, TGF-β, or TGM-1. B-C. mCHERRY reporter with doxycycline inducible mouse CD44 (B) or human CD44 (C) HEK293 cells incubated with either control media, TGF-β, or TGM-1 (1 ng ml^{-1}, 5 ng mL^{-1}, or 25 ng mL^{-1}) Data collected and analyzed by the Maizels lab and the ten Dijke lab, submitted for revision in PNAS under the title "CD44 acts as a co-receptor for cell-specific enhancement of signaling and regulatory T cell induction by TGM1, a parasite TGF-β mimic".

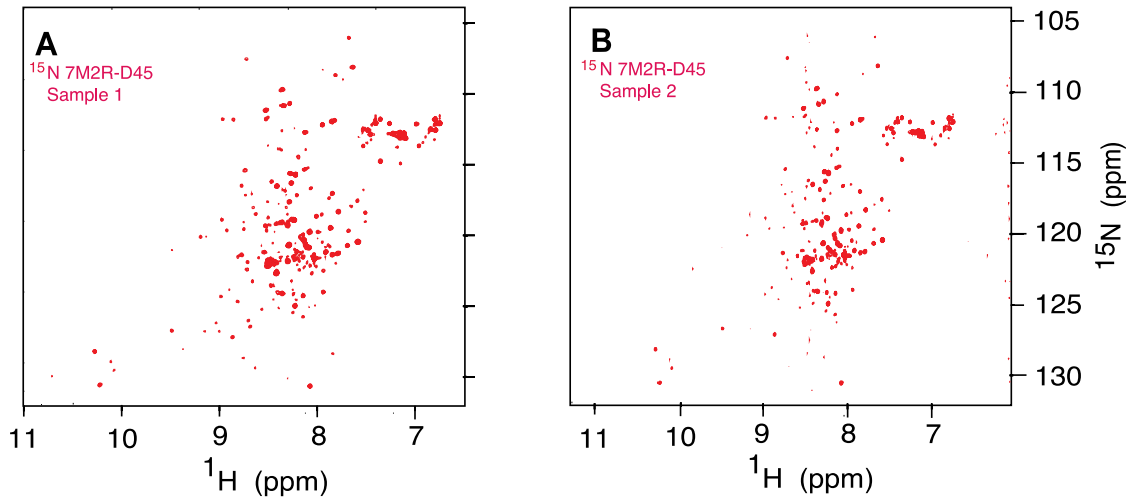

Figure 74: Native 7M2R-D45 can be purified to homogeneity. ^{1}H-^{15}N spectra of ^{15}N 7M2R-D45 sample 1 (A, red) and sample 2 (B, red). The spectrum was recorded in 10mM sodium phosphate, 20mM CHAPS, 5% $^{2}H_2O$, pH 4.7, 310K.

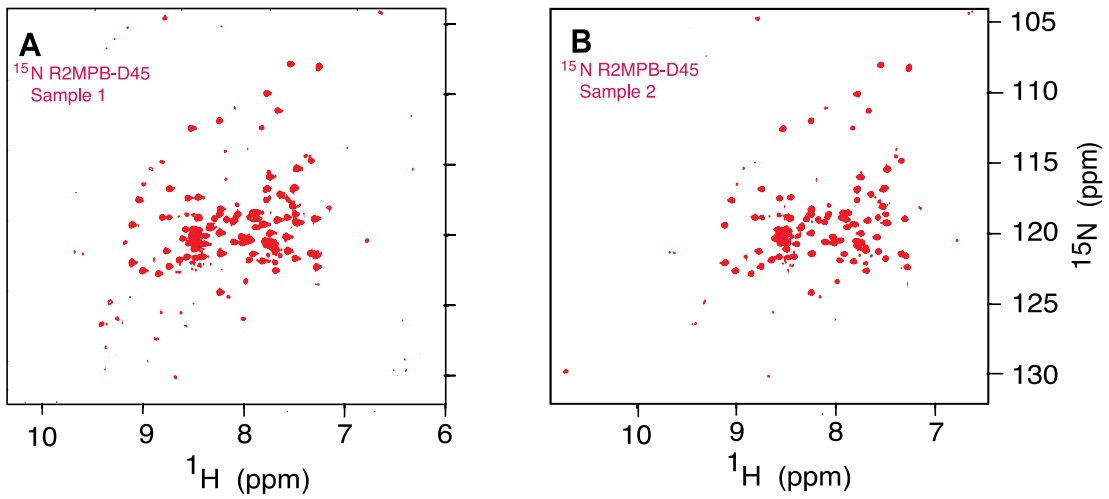

Figure 75: Native R2MPB-D45 can be purified to homogeneity. ^{1}H-^{15}N spectra of ^{15}N R2MPB-D45 sample 1 (A, red) and sample 2 (B, red). The spectrum was recorded in 25mM CHES, 60mM NaCl 5% $^{2}H_2O$, pH 8.5, 310K.

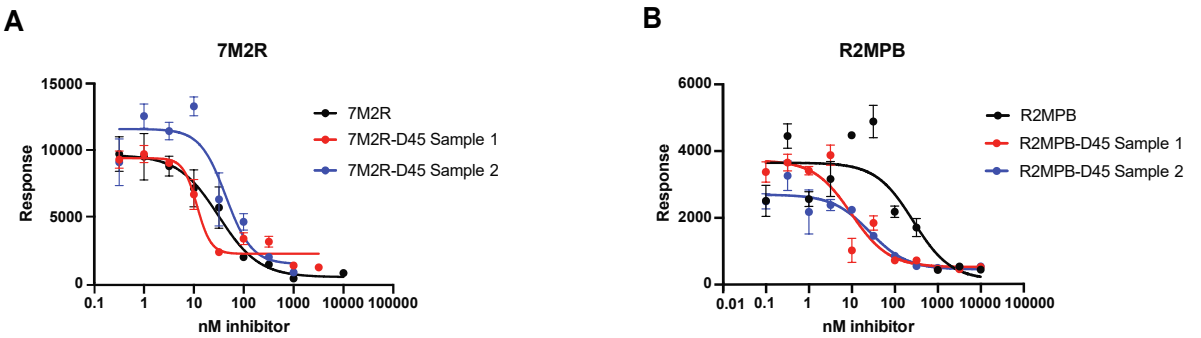

Figure 76: 7M2R and R2MPB fusion constructs inhibit TGF-β signaling. A-B. Inhibition of TGF-β signaling as measured via the CAGA-LUC assay, with constant amounts of TGF-β and increasing amounts of 7M2R and 7M2R-D45 (A), along with R2MPB and R2MPB-D45 (B). N=1 with three technical replicates. Data collected by Dr. Cynthia Hinck, Hinck Lab.

9.0 Conclusions and Future Directions

H. polygyrus TGM or TGF-β mimic is aptly named as it functionally mimics the activity of mammalian TGF-β. Despite broad sequence dissimilarity between TGM-1 and TGF-β, analysis of the TGM-1 D3 structure, both alone and in a complex with TGF-β type 2 receptor TβRII showed that TGM has adapted from the canonical CCP domain structure in order to bind TβRII. TGM-1 D3 lacks two of the canonical CCP domain β-strands, which allows for an expansion of the lateral half of the protein and partial exposure of several core hydrophobic residues. In particular, exposure of both I^{238} and Y^{253} is key as these residues interact with I^{76} of TβRII, creating the main binding interactions. This interaction mimics the insertion of I^{76} of TβRII into a hydrophobic cleft formed between the fingers of TGF-β, including residues such as Trp^{330} and Tyr^{390}. Thus, TGM-1 D3 mimics the contact residues of TGF-β, utilizing similar hydrophobic residues to bracket the inserted I^{76} from TβRII. Additional mimicked interactions include the TGF-β R^{394} - TβRII D^{55} hydrogen bonded ion-pair with the TGM-1 D3 K^{254} - TβRII D^{55} ion pair, while the similar hydrogen bonded ion pair with TGF-β R^{325} - TβRII E^{142} is not mimicked by TGM-1 D3, indicating that there are some slight differences between the two interfaces, though with enough similarities that the two ligands compete for binding with TβRII. TGM-1 was also shown to compete with the TGF-β:$TβRII_2$ complex for binding to TβRI; though the interface has yet to be elucidated, it is likely that TGM-1 is similarly mimicking the binding residues of TGF-β in order to bind TβRI.

Unlike TGF-β, the binding ability of TGM-1 can be separated by its distinct domains. TGM-1 D3 is the sole binding domain for TβRII while TGM-1 D12 are the two binding domains responsible for binding TβRI in a multivalent manner, with TGM-1 D2 being the main binding

partner. This means that the ability of TGM-1 to bind the TGF-β receptors is modular and domain-separable, particularly as the individual domains of TGM-1 can be purified separately. This has implications for the potential of developing TGM-1 as a TGF-β targeting therapeutic. This can occur in two manners: as an inhibitory or a stimulatory therapeutic. The latter is evident, with the use of constructs containing TGM-1 D123 to bind to both TβRI and TβRII and thus act to stimulate the TGF-β signaling pathway, which may be of use in pathogenic contexts with downregulated TGF-β signaling, such as wound healing contexts. For the inhibitory context, if the receptor binding capabilities of TGM can be distinctly isolated, then soluble TGM-1 constructs can be produced and utilized to sequester away either the type I or the type II receptor, thus prohibiting TGF-β signaling. This has clinical applications in a variety of pathologic contexts including overexpression of TGF-β in the context of autoimmune disease and soft-tissue cancers. Inhibition of TGF-β has been shown to be a highly effective adjuvant to PD-1 and PD-L1 therapy in comparison to monotherapy *in vivo*[141,142], and engineered domain deletion constructs of TGM-1 may be able to perform a similar role as a chemotherapeutic or immunosuppressive adjuvant.

With this in mind, different constructs of TGM-1 were studied for efficacy as either TGF-β stimulatory or inhibitory therapeutics. This was done in part as one of the members of the family, TGM-4 was shown to bind to TβRI with a close to 10-fold higher affinity than TGM-1, despite binding with very weak affinity to TβRII. First, the potential utility of TGM domain deletion constructs as inhibitors was tested. Constructs of TGM-1 D12, TGM- D12, and TGM-1 D3 were tested for their ability to inhibit TGF-β signaling. However, despite these hypotheses, none of these constructs inhibited TGF-β signaling, indicating a separate component to the signaling mechanism that has yet to be determined. Additionally, in theory, a combination of domains from TGM-1 and TGM-4 would create a construct that would be maximized for binding to the TGF-β receptors and

thus act as a more potent signaling molecule. Using the MFB-F11 assay, chimeric constructs of TGM-4 D12: TGM-1 D3: TGM-4 D45 and TGM-1 D12: TGM-4 D3: TGM-1 D45 constructs were created. It was hypothesized that the former construct would maximize TGF-β signaling due to the high affinity of TGM-4 D12 for TβRI and of TGM-1 D3 for TβRII. Thus conversely the latter construct, due to the particularly weak affinity of TGM-4 D3 for TβRII would minimally signal through the TGF-β pathway. However, the TGM-4 D12: TGM-1 D3: TGM-4 D45 construct did not show significant activity in the MFB-F11 TGF-β signaling assay, while the TGM-1 D12: TGM-4 D3: TGM-1 D45 construct did show significant activity in the assay, despite the weak binding affinity of TGM-4 D3 to TβRII. From this data, it was clear that domains 4 and 5 were key to the activity of TGM-1, and that this role needed to be studied.

It was hypothesized that TGM-1 domains 4-5, and in consequence, domains 4 and 5 of all of the TGM family members, bound to a coreceptor separate from TβRI and TβRII but integral to the *in vitro* signaling, likely due to effects of avidity. Mass spectrometry, immunoblotting, and ITC experiments demonstrated that CD44, a glycoprotein heavily expressed on T-cells, was this co-receptor, binding to TGM-1 D45 and TGM-4 D45, but not TGM-6 D45. A cell line was generated with dox-inducible CD44, both the human and the mouse isoforms. As the amount of doxycycline increased, corresponding to increased expression of CD44, TGM-1 signal was potentiated, in a doxycycline-dependent manner, indicating the potentiation of TGM-1 signaling by CD44, though this effect was most pronounced for the mouse CD44 isoform over the human CD44 isoform. With this in mind, it was clear that TGM-1 was binding a co-receptor whose purpose appeared to be the further potentiation of TGM-1 signaling in a cell context. The TGM family members, additionally, were hypothesized to bind to different coreceptors as well, as TGM-6 D45 showed no binding to CD44, and current in-progress studies have identified an alternate co-

receptor. While CD44 is ubiquitously expressed, albeit moreso in T-cells, the potential TGM-6 coreceptor is more highly expressed in fibroblasts. This then led to the hypothesis that the purpose of domains 4 and 5 in the TGM family was to target the TGM protein of interest to certain cell types, thus acting as a targeting mechanism for different parasitic proteins.

Chimeric proteins were generated by fusing known TGF-β inhibitors with TGM-1 D45. Two chimeric proteins were tested: R2MPB-D45, and 7M2R-D45; both TGF-β inhibitors act by sequestering TβRII. The chimeric proteins were tested for TGF-β signaling inhibition in in HEK293T cells, a human embryonic cell line with low levels of CD44 expression. The HEK293T cells were stimulated with TGF-β3 and then increasing amounts of either R2MPB, R2MPB-D45, 7M2R, or 7M2R-45 were added. Both 7M2R and R2MPB are both capable of inhibiting TGF-β signaling, but addition of the fused TGM-1 D45 domains did not seem to potentiate inhibition in a low CD44 expression cell line. However, these are initial studies, and further studies in cells with higher surface expression of CD44 needs to be performed in order to assess inhibitory capacity of these fusion proteins.

Future experiment include further efforts to characterize the TGM domains and their interactions with the TGF-β receptors. Though NMR CSP experiments were performed showing that the residues of TβRI perturbed by binding to TGM-1 D2 were similar to the residues perturbed by binding to the TGF-β:$T\beta RII_2$ complex, the specific residues involved in the TGM-1 D2:TβRI interface have yet to be elucidated. Future experimentation involves the reverse set of chemical shift perturbation experiments looking at the chemical shift differences between apo and TβRI-bound TGM-1 D2, and then utilizing biophysical methods to assess for key interface residues, along with either crystallographic or NMR methods for further complex determination. Ideally,

this would lead to the development of a TβRI-targeting therapeutic, though much more protein engineering work would have to be involved for TGM-1 D12 to be of any use therapeutically.

Additionally, the interaction between CD44 and TGM-1 D45 must be further studied. There has been little biophysical work done to characterize the mCD44:TGM-1 D45 complex, outside of some SEC and native gel experimentation. The residues involved in the binding interface should be elucidated, either via NMR or crystallographic methods supported by SPR or ITC biophysical binding studies. This would help in the generation of CD44 binders, whether these are based on TGM-1 D45 or otherwise. There also needs to be more *in vitro* work involved in the characterization of this interaction. Though initial studies support a model of avidity, suggesting that the potentiation of TGM-1 signaling from CD44 comes from its ability to target TGM-1 to the cell surface and thus increase the local concentration and facilitate receptor binding, this has not yet been mechanistically proven, nor has the question of why mouse CD44 is so much more potent than human CD44 in cell-based experiments. Future directions involve studying this in more detail.

Future experimentation will also involve further studying of the TGM-1 D45 fusion domain proteins and their ability to inhibit TGF-β signaling. Cells with higher levels of CD44 should be studied, as should cells with inducible CD44, to further understand the ability of the fusion proteins to inhibit TGF-β signaling. If this proves successful, then *in vivo* studies may be possible further down the line in immunosuppressed pathophysiologies, such as in several cancer varieties. The co-receptors of other TGM proteins should be studied as well. If co-receptors can be identified for TGM-2 through TGM-10, it is likely that more potential cell targets can be found. This would expand the repertoire of targets for TGF-β therapeutics, with different TGM domain 4-5 constructs leading to different cell types, thus allowing for a hopefully wider array of cell targets, and generating a novel paradigm for the development of TGF-β targeting therapeutics.

Appendix A : NMR Supplement

Appendix A.1 Xplor-NIH scripts

Appendix A.1.1 TGM-1 D3 structure determination script

The script for the determination of the TGM-1 D3 structure is based off of the script provided with Xplor-NIH version 2.52 by Charles D. Schweiters, the director of the Computational Biomolecular Magnetic Resonance Core at NIDDK for slow docking of simulated annealing from an extended structure.

****

```
xplor.requireVersion("2.52")


# slow cooling protocol in torsion angle space for protein G. Uses
# NOE, RDC, J-coupling restraints.
#
# this script performs annealing from an extended structure.
# It is faster than the original anneal.py
#
# CDS 2009/07/24

# this checks for typos on the command-line. User-customized arguments can
# also be specified.
#
(opts,args) = xplor.parseArguments(["quick"])

quick=False
for opt in opts:
   if opt[0]=="quick":  #specify -quick to just test that the script runs
      quick=True
      pass
   pass
```

```
# filename for output structures. This string must contain the STRUCTURE
# literal so that each calculated structure has a unique name. The SCRIPT
# literal is replaced by this filename (or stdin if redirected using <),
# but it is optional.
#
outFilename = "SCRIPT_STRUCTURE.sa"
numberOfStructures=500   #usually you want to create at least 20

if quick:
   numberOfStructures=3
   pass

# protocol module has many high-level helper functions.
#
import protocol
protocol.initStruct('generate4/tgmd3.psf') #read in PSF file (such that disulfide bonds are included)
protocol.initParams('protein') #read covalent and nonbonded parameters from parameter file(s)
protocol.initRandomSeed()   #set random seed - by time

command = xplor.command

# generate random extended initial structure with correct covalent geometry
protocol.genExtendedStructure()

#
# a PotList contains a list of potential terms. This is used to specify which
# terms are active during refinement.
#
from potList import PotList
potList = PotList()

# parameters to ramp up during the simulated annealing protocol
#
from simulationTools import MultRamp, StaticRamp, InitialParams, IVMAction

rampedParams=[]
highTempParams=[]

# IVM setup
#   the IVM is used for performing dynamics and minimization in torsion-angle
#   space, and in Cartesian space.
#
from ivm import IVM
dyn  = IVM()
minc = IVM() # minc used for final cartesian minimization
```

```
# initialize ivm topology for torsion-angle dynamics

# compare atomic Cartesian rmsd with a reference structure
#  backbone and heavy atom RMSDs will be printed in the output
#  structure files
#
#from posDiffPotTools import create_PosDiffPot
#refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
#                        pdbFile='g_xray.pdb',
#                        cmpSel="not name H*")

# orientation Tensor - used with the dipolar coupling term
#  one for each medium
#   For each medium, specify a name, and initial values of Da, Rh.
#
from varTensorTools import create_VarTensor
media=290
#                    medium  Da   rhombicity
for (medium,Da,Rh) in [ ('t',     8.54, 0) ]:
#                    ('b',   -9.9, 0.23) ]:
   oTensor = create_VarTensor(medium)
   oTensor.setDa(Da)
   oTensor.setRh(Rh)
   media[medium] = oTensor
   pass

from varTensorTools import calcTensorOrientation
highTempParams.append(StaticRamp("""
for medium in media.values():
  calcTensorOrientation(medium)
  """) )

dyn.addConfigAction( IVMAction("""
for m in media.values():
   m.setFreedom("fixDa, fixRh")       #fix tensor Rh, Da, vary orientation
   """) )

minc.addConfigAction( IVMAction("""
for m in media.values():
   m.setFreedom("varyDa, varyRh")     #vary tensor Rh, Da, vary orientation
   """) )

# dipolar coupling restraints for protein amide NH.
#
# collect all RDCs in the rdcs PotList
#
```

```
# RDC scaling. Three possible contributions.
#   1) gamma_A * gamma_B / r_AB^3 prefactor. So that the same Da can be used
#      for different expts. in the same medium. Sometimes the data is
#      prescaled so that this is not needed. scale_toNH() is used for this.
#      Note that if the expt. data has been prescaled, the values for rdc rmsd
#      reported in the output will relative to the scaled values- not the expt.
#      values.
#   2) expt. error scaling. Used here. A scale factor equal to 1/err^2
#      (relative to that for NH) is used.
#   3) sometimes the reciprocal of the Da^2 is used if there is a large
#      spread in Da values. Not used here.
#
from rdcPotTools import create_RDCPot,scale_toNH,correctGyromagneticSigns
#correctGyromagneticSigns()
rdcs = PotList('rdc')
for (medium,expt,file,                scale) in \
    [('t','NH' ,'tgmd3_rdc_nh_CHB.tbl'       ,1),
     ('t','CAHA','tgmd3_rdc_caha_ave_v4.tbl'    ,0.1),
     ('t','CACO','tgmd3_rdc_caco_ave_v3.tbl'     ,20)
     #('t','CAHA','gb1-dipo-caha4.tbl'    ,1),
     #('t','CACO','gb1-dipo-cac4.tbl'     ,1)
     #('b','NH' ,'bicelles_new_nh.tbl' ,1),
     #('b','NCO','bicelles_new_nc.tbl' ,.05),
     #('b','HNC','bicelles_new_hnc.tbl',.108)
     ]:
    rdc = create_RDCPot("%s_%s"%(medium,expt),file,media[medium])

    #1) scale prefactor relative to NH
    #   see python/rdcPotTools.py for exact calculation

    scale_toNH(rdc) # not needed for these datasets -

    #                     but non-NH reported rmsd values will be wrong.
    #3) Da rescaling factor (separate multiplicative factor)
    # scale *= ( 9.9 / rdc.oTensor.Da(0) )**2
    rdc.setScale(scale)
    rdcs.append(rdc)
    pass
potList.append(rdcs)
rampedParams.append( MultRamp(0.01,1.0, "rdcs.setScale( VALUE )") )


# set up NOE potential
noe=PotList('noe')
potList.append(noe)
from noePotTools import create_NOEPot
```

```
for (name,scale,file) in [('all',1,"noe-4-16june-nodup.tbl"),
                          #add entries for additional tables
                          ]:
   pot = create_NOEPot(name,file)
   # pot.setPotType("soft") # if you think there may be bad NOEs
   pot.setScale(scale)
   noe.append(pot)
rampedParams.append( MultRamp(2,30, "noe.setScale( VALUE )") )

# set up J coupling - with Karplus coefficients
from jCoupPotTools import create_JCoupPot
jCoup = create_JCoupPot("jcoup","jhnhacoup4.tbl",
                        A=7.09,B=-1.42,C=1.55,phase=-60.0)
potList.append(jCoup)

# Set up dihedral angles
from xplorPot import XplorPot
dihedralRestraintFilename="tgmd3-dihed-nov22-2021.tbl"
protocol.initDihedrals(dihedralRestraintFilename,
                       #useDefaults=False  # by default, symmetric sidechain
                                           # restraints are included
                       )
potList.append( XplorPot('CDIH') )
highTempParams.append( StaticRamp("potList['CDIH'].setScale(10)") )
rampedParams.append( StaticRamp("potList['CDIH'].setScale(200)") )
# set custom values of threshold values for violation calculation
#
potList['CDIH'].setThreshold( 5 )

# gyration volume term
#
from gyrPotTools import create_GyrPot
gyr = create_GyrPot("Vgyr",
                    "resid 4:85") # selection should exclude disordered tails
potList.append(gyr)
rampedParams.append( MultRamp(.002,1,"gyr.setScale(VALUE)") )

# hbda - distance/angle bb hbond term
#
protocol.initHBDA('HBDA-5.tbl')
potList.append( XplorPot('HBDA') )

# hbdb - hbond database-based term
#  CDS - 2017/04/24 - found to make precision worse - maybe better to use
#  during refinement only
#protocol.initHBDB()
```

```
#potList.append( XplorPot('HBDB') )

#New torsion angle database potential
#
from torsionDBPotTools import create_TorsionDBPot
torsionDB = create_TorsionDBPot('torsionDB', system='protein')
potList.append( torsionDB )
rampedParams.append( MultRamp(.002,2,"torsionDB.setScale(VALUE)") )

#
# setup parameters for atom-atom repulsive term. (van der Waals-like term)
#
from repelPotTools import create_RepelPot,initRepel
repel = create_RepelPot('repel')
potList.append(repel)
rampedParams.append( StaticRamp("initRepel(repel,use14=False)") )
rampedParams.append( MultRamp(.004,4,  "repel.setScale( VALUE)") )
# nonbonded interaction only between CA atoms
highTempParams.append( StaticRamp("""initRepel(repel,
                                    use14=True,
                                    scale=0.004,
                                    repel=1.2,
                                    moveTol=45,
                                    interactingAtoms='name CA'
                                    )""") )

# Selected 1-4 interactions.
import torsionDBPotTools
repel14 = torsionDBPotTools.create_Terminal14Pot('repel14')
potList.append(repel14)
highTempParams.append(StaticRamp("repel14.setScale(0)"))
rampedParams.append(MultRamp(0.004, 4, "repel14.setScale(VALUE)"))


potList.append( XplorPot("BOND") )
potList.append( XplorPot("ANGL") )
potList['ANGL'].setThreshold( 5 )
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)") )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold( 5 )
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)") )



# Give atoms uniform weights, except for the anisotropy axis
#
```

```
protocol.massSetup()


protocol.torsionTopology(dyn)

protocol.initMinimize(minc)

protocol.cartesianTopology(minc)



# object which performs simulated annealing
#
from simulationTools import AnnealIVM
init_t  = 3500.     # Need high temp and slow annealing to converge
cool = AnnealIVM(initTemp =init_t,
                 finalTemp=25,
                 tempStep =12.5,
                 ivm=dyn,
                 rampedParams = rampedParams)

#cart_cool is for optional cartesian-space cooling
cart_cool = AnnealIVM(initTemp =init_t,
              finalTemp=25,
      tempStep =12.5,
                    ivm=minc,
                    rampedParams = rampedParams)

def calcOneStructure(loopInfo):
    """ this function calculates a single structure, performs analysis on the
    structure, and then writes out a pdb file, with remarks.
    """

    # generate a new structure with randomized torsion angles
    #
    from monteCarlo import randomizeTorsions
    randomizeTorsions(dyn)

    # set torsion angles from restraints
    #
    from torsionTools import setTorsionsFromTable
    setTorsionsFromTable(dihedralRestraintFilename)
    protocol.fixupCovalentGeom(maxIters=100,useVDW=1)
    protocol.writePDB(loopInfo.filename()+".init")

    # initialize parameters for high temp dynamics.
```

```
InitialParams( rampedParams )
# high-temp dynamics setup - only need to specify parameters which
#   differfrom initial values in rampedParams
InitialParams( highTempParams )

# high temp dynamics
#
protocol.initDynamics(dyn,
                potList=potList, # potential terms to use
                bathTemp=init_t,
                initVelocities=1,
                finalTime=100,   # stops at 100ps or 1000 steps
                numSteps=1000,   # whichever comes first
                printInterval=100)

dyn.setETolerance( init_t/100 )  #used to det. stepsize. default: t/1000
dyn.run()

# initialize integrator for simulated annealing
#
protocol.initDynamics(dyn,
                numSteps=200,       #at each temp: 100 steps or
                finalTime=.2 ,      # .2ps, whichever is less
                printInterval=100)

# perform simulated annealing
#
cool.run()


# final torsion angle minimization
#
protocol.initMinimize(dyn,
                printInterval=50)
dyn.run()

# optional cooling in Cartesian coordinates
#
protocol.initDynamics(minc,
                potList=potList,
                numSteps=200,       #at each temp: 100 steps or
                finalTime=.4 ,      # .2ps, whichever is less
                printInterval=100)
#cart_cool.run()
# final all- atom minimization
#
```

```
    protocol.initMinimize(minc,
                          potList=potList,
                          dEPred=10)
    minc.run()

    #do analysis and write structure when function returns
    pass



from simulationTools import StructureLoop, FinalParams
StructureLoop(numStructures=numberOfStructures,
              pdbTemplate=outFilename,
              structLoopAction=calcOneStructure,
              doWriteStructures=True,
              genViolationStats=True,
              averageTopFraction=0.02, #report stats on best 20% of structs
              averageContext=FinalParams(rampedParams),
#               averageCrossTerms=refRMSD,
              averageSortPots=[potList['BOND'],potList['ANGL'],potList['IMPR'],
                               noe,rdcs,potList['CDIH']],
              averageFilename="SCRIPT_ave.pdb",
              averagePotList=potList).run()
```

Appendix A.1.2 TGM-1 D3:TβRII structure determination script

The script for the determination of the TGM-1 D3:TβRII complex structure is based off of the script provided with Xplor-NIH version 2.52 by Charles D. Schweiters, the director of the Computational Biomolecular Magnetic Resonance Core at NIDDK for slow docking of two rigid structures.

```
xplor.requireVersion("2.52")

# slow docking protocol for TbRII and TGM6D3 complex with
# few NOE, and PCS restrains.


# filename for output structures.
outFilename = "SCRIPT_STRUCTURE.sa"
```

```
numberOfStructures=200
dockNumber = 100

import protocol

protocol.initRandomSeed(1215)   #set random seed - by time
command = xplor.command
# selection chains
mol1Sel="resid 1:91"   #D3
mol2Sel="resid 115:237"  #RII


######### Importing data #############

# Read PSF file(s).
protocol.initStruct(('tgmd3.psf', 'rii-t.psf', 'metal.psf'))

# Load paramaters.
protocol.initParams(['protein', 'metal'])

# Read input structure(s).
protocol.initCoords(('D3-2.pdb','RII-t3.pdb','metal.pdb'))

# Maybe I should delete these unknown ?
xplor.simulation.deleteAtoms("not known")

# Creating energy list
from potList import PotList
potList = PotList()

# parameters to ramp up during the simulated annealing protocol
from simulationTools import MultRamp, StaticRamp, InitialParams, IVMAction

rampedParams=[]
highTempParams=[]

##IVM for performing dynamics in torsion-angle space and in cartesian space

from ivm import IVM
dyn = IVM()
minc = IVM() #minc for final cartesian minimization

#----------------------getting in RDC data-------------------------------#

#------------setting tensor-----------------------------#
```

```
from rdcPotTools import correctGyromagneticSigns
correctGyromagneticSigns()

from varTensorTools import create_VarTensor
media=290

for (medium,Da,Rh) in [ ('RII', -6.8, 0.347),
('D3', -3.5, 0.665)
]:
   VTensor = create_VarTensor(medium)
   VTensor.setDa(Da)
   VTensor.setRh(Rh)
   media[medium] = VTensor
   pass

from varTensorTools import calcTensorOrientation
highTempParams.append(StaticRamp("""
for medium in media.values():
  calcTensorOrientation(medium)
  """) )

rdcData = [("RII", "HN", "RDC/HNRDC_phage_RII-A.tbl", 1),
          ("D3", "HN", "RDC/HNRDC_phage_D3-C.tbl", 1)
  ]

from rdcPotTools import create_RDCPot, scale_toNH
rdcs = PotList("rdc")
for (medium, exp, table, scale ) in rdcData:
name = " % s_ % s" % (exp, medium)
rdc = create_RDCPot(name, table, media[medium])
rdc.setScale(scale)
scale_toNH(rdc)
rdcs.append(rdc)
pass

potList.append(rdcs)
rampedParams.append(MultRamp(0.01, 1.0, "rdcs.setScale(VALUE)"))
rdc.setThreshold(0)

#----------------------getting in PCS data-----------------------------#

#------------setting tensor----------------------------#


pcs = PotList('pcs')
potList.append(pcs)
```

```
from rdcPotTools import create_RDCPot
from varTensorTools import create_VarTensor, calcTensor



# Tensor set up.
tensor = create_VarTensor('tensor')
tensor.setDaMax(1e5)


# List with PCS data.
 #           name              table
pcsData =  [('intra', 'PCS/new2/DE-intramolecular-Tb6.tbl'),
          ('inter', 'PCS/new2/DE-intermolecular-Tb5.tbl')]

# PCS term set up.

pcs_intra = PotList('pcs_intra')  # to calculate tensor

for (name, table) in pcsData:
   pot = create_RDCPot(name, table, oTensor=tensor)
   pot.setUseDistance(True)
   pot.setAveType('sum')
   pot.setShowAllRestraints(True)
   if name == 'intra':
      pot.setScale(10)
      pcs_intra.append(pot)
      pcs.append(pot) # adding intermolecular PCSs as one of the energy terms
   else:
      pot.setScale(100)
      pcs.append(pot)

calcTensor(tensor, expts=pcs_intra)  # calculate tensor
print(tensor.Da(), tensor.Rh())
tensor.setDa(5636.0)
tensor.setRh(0.666)
tensor.setFreedom("fixDa, fixRh")
print(tensor.Da(), tensor.Rh())
rampedParams.append(MultRamp(1, 100, "pcs.setScale(VALUE)"))

pcs.setThreshold(0)

#----------------------getting in PCS data-----------------------------#
```

```
#------------setting tensor-----------------------------#


pcs2 = PotList('pcs2')
potList.append(pcs2)


from rdcPotTools import create_RDCPot
from varTensorTools import create_VarTensor, calcTensor

# Tensor set up.
tensor2 = create_VarTensor('tensor2')
tensor2.setDaMax(1e5)


# List with PCS data.
#            name              table
pcs2Data =  [('intra2', 'PCS/new2/D3-loop-intramolecular-3.tbl'),
          ('inter2', 'PCS/new2/D3-loop-intermolecular-3.tbl')]

# PCS term set up.

pcs2_intra2 = PotList('pcs2_intra2')  # to calculate tensor

for (name, table) in pcs2Data:
   pot = create_RDCPot(name, table, oTensor=tensor2)
   pot.setUseDistance(True)
   pot.setAveType('sum')
   pot.setShowAllRestraints(True)
   if name == 'intra2':
      pot.setScale(10)
      pcs2_intra2.append(pot)
      pcs2.append(pot) # adding intermolecular PCSs as one of the energy terms
   else:
      pot.setScale(100)
      pcs2.append(pot)

calcTensor(tensor2, expts=pcs2_intra2)  # calculate tensor
print(tensor2.Da(), tensor2.Rh())
tensor2.setDa(1758.1)
tensor2.setRh(0.600)
tensor2.setFreedom("fixDa, fixRh")
print(tensor2.Da(), tensor2.Rh())
rampedParams.append(MultRamp(1, 100, "pcs2.setScale(VALUE)"))

pcs2.setThreshold(0)
```

```
# set up NOE potential
import noePotTools
noe = noePotTools.create_NOEPot(name="noe", file="NOE/noe_20feb23.tbl")
potList.append(noe)
#noe.setPotType("soft") # in case of bad NOEs
rampedParams.append( MultRamp(2,30, "noe.setScale( VALUE )") )

from xplorPot import XplorPot

from torsionDBPotTools import create_TorsionDBPot
torsionDB = create_TorsionDBPot('torsionDB', system='protein')
potList.append( torsionDB )
rampedParams.append( MultRamp(.002,2,"torsionDB.setScale(VALUE)") )

from repelPotTools import create_RepelPot,initRepel
repel = create_RepelPot('repel')
potList.append(repel)
rampedParams.append( StaticRamp("initRepel(repel,use14=False)") )
rampedParams.append( MultRamp(.004,4,  "repel.setScale( VALUE)") )
 #nonbonded interaction only between CA atoms
highTempParams.append( StaticRamp("""initRepel(repel,
                                    use14=True,
                                    scale=0.004,
                                    repel=1.2,
                                    moveTol=45,
                                    interactingAtoms='name CA'
                                    )""") )

#Selected 1-4 interactions.
import torsionDBPotTools
repel14 = torsionDBPotTools.create_Terminal14Pot('repel14')
potList.append(repel14)
highTempParams.append(StaticRamp("repel14.setScale(0)"))
rampedParams.append(MultRamp(0.004, 4, "repel14.setScale(VALUE)"))#

potList.append( XplorPot("BOND") )
potList.append( XplorPot("ANGL") )
potList['ANGL'].setThreshold( 5 )
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)") )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold( 5 )
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)") )


from atomSelLang import nameSelection
```

```
nameSelection('backbone'," or ".join(["name %s"% n
                        for n in "C CA N HNO".split()]))

# IVM setup#########################################################################
from ivm import IVM
dyn  = IVM()
dynRigid  = IVM()
minc = IVM() # minc used for final cartesian minimization


dynRigid.fix(mol2Sel)    #completely rigid docking partners
dynRigid.group(mol1Sel)

dyn.fix("resid 237 or (recall backbone and (%s))" % mol2Sel)
dyn.group("resid 91 or (recall backbone and (%s))" % mol1Sel)


#tensor.setFreedom("varyDa, varyRh")
#tensor2.setFreedom("varyDa, varyRh")
VTensor.setFreedom("varyDa, varyRh")

protocol.torsionTopology(dyn)
protocol.torsionTopology(dynRigid)
protocol.initMinimize(minc)
protocol.cartesianTopology(minc)

# Give atoms uniform weights, except for the anisotropy axis
#
protocol.massSetup()


# simulated annealing
from simulationTools import AnnealIVM
init_t  = 3500.
cool = AnnealIVM(initTemp =init_t,
             finalTemp=25,
             tempStep =12.5,
             ivm=dyn,
             rampedParams = rampedParams)

#cart_cool is for optional cartesian-space cooling
cart_cool = AnnealIVM(initTemp =init_t,
          finalTemp=25,
     tempStep =12.5,
                ivm=minc,
                rampedParams = rampedParams)
```

```
def calcOneStructure(loopInfo):
   """ One structure with n dockings"""

   InitialParams( rampedParams ) # somehow still important.
   InitialParams( highTempParams ) # 'Listining 27' not really easy for me to grasp the concept - maybe during next reading iteration !

   atomPos0 = xplor.simulation.atomPosArr()
   atomPosMin = xplor.simulation.atomPosArr()
   energyMin =  1e9# big number
   k=0

   # now is the loop adopted from 'eginput/dock_dipolar_chemshift/dock.py' that actually does the initial rigid docking
   # first with RDC, the both RDC and noe + repel energ term
   while k < dockNumber:

   xplor.simulation.setAtomPosArr(atomPos0)

      from atomAction import randomizeDomainPos
      randomizeDomainPos(mol1Sel)# randomizing second partner position
   protocol.initMinimize(dynRigid,potList=[pcs,pcs2,rdc,noe],
                         printInterval=50)
   dynRigid.run() ; dynRigid.run()# gradient minimization in torsion angle space with RDC energy term

   protocol.initMinimize(dynRigid,potList=[pcs,pcs2,rdc,noe,repel],
                         printInterval=50)
   dynRigid.run() ; dynRigid.run()# gradient minimization in torsion angle space with RDC, NOE and repel energy terms

      print 'iter: %d    energy: %f' % (k,potList.calcEnergy()) # just printing energies

if potList.calcEnergy() < energyMin :# not sure about this criterion, but sound like 'if the energy is not ridiculusly high (e.g structure overlap), then is ok'
         atomPosMin = xplor.simulation.atomPosArr()  # the coordinates then are saved
         energyMin = potList.calcEnergy()# and the minimal energy is overwirtten,
   pass# - if only the new iteration provides better values, the energy and coordinates will be overwiritten again

print "#####"+str(k)+"#######"

k += 1
      pass
```

```
   ######################## Now getting back into more 'standard' protocol
########################################
   # accpting the position from best run from the loop
   xplor.simulation.setAtomPosArr(atomPosMin)

   #randomize sidechain torsions
   from monteCarlo import randomizeTorsions
   randomizeTorsions(dyn)

   InitialParams( rampedParams )
   InitialParams( highTempParams ) # somehow this need to be called again

   # high temp dynamics - not sure if this step is required ?
   protocol.initDynamics(dyn,
                  potList=potList, # potential terms to use
                  bathTemp=init_t,
                  initVelocities=1,
                  finalTime=100,   # stops at 100ps or 1000 steps
                  numSteps=1000,   # whichever comes first
                  printInterval=100)

   dyn.setETolerance( init_t/100 )  #used to det. stepsize. default: t/1000
   dyn.run()

   InitialParams( rampedParams ) # and again

   # simulated annealing - is it even used ?
   protocol.initDynamics(dyn,
                  numSteps=100,       #at each temp: 100 steps or
                  finalTime=.2 ,      # .2ps, whichever is less
                  printInterval=100)

   # perform simulated annealing - or isn't that he one specified earlier ?
   cool.run()


   # final torsion angle minimization
   protocol.initMinimize(dyn,
                  printInterval=50)
   dyn.run()

#   # optional cooling in Cartesian coordinates
#   #
#   protocol.initDynamics(minc,
#                    potList=potList,
#                    numSteps=100,      #at each temp: 100 steps or
```

```
#                       finalTime=.4 ,       # .2ps, whichever is less
#                       printInterval=100)
#   #cart_cool.run()

   # final all-atom minimization
   #
#   protocol.initMinimize(minc,
#                         potList=potList,
#                         dEPred=10)
#   minc.run()

   #do analysis and write structure when function returns
   pass

from simulationTools import StructureLoop, FinalParams
StructureLoop(numStructures=numberOfStructures,
              pdbTemplate=outFilename,
              structLoopAction=calcOneStructure,
              doWriteStructures=True,
              genViolationStats=True,
              averageTopFraction=0.05, #report stats on best 10% of structs
              averageContext=FinalParams(rampedParams),
              averageFilename="SCRIPT_ave.pdb",
              averageFitSel="(%s) and recall backbone" % mol1Sel,
              averagePotList=potList).run()
```

Bibliography

1 Sporn, M. B. & Todaro, G. J. Autocrine secretion and malignant transformation of cells. *N Engl J Med* **303**, 878-880 (1980). https://doi.org:10.1056/NEJM198010093031511

2 Moses, H. L., Roberts, A. B. & Derynck, R. The Discovery and Early Days of TGF-beta: A Historical Perspective. *Cold Spring Harb Perspect Biol* **8** (2016). https://doi.org:10.1101/cshperspect.a021865

3 Roberts, A. B., Anzano, M. A., Lamb, L. C., Smith, J. M. & Sporn, M. B. New class of transforming growth factors potentiated by epidermal growth factor: isolation from non-neoplastic tissues. *Proc Natl Acad Sci U S A* **78**, 5339-5343 (1981). https://doi.org:10.1073/pnas.78.9.5339

4 Derynck, R. *et al.* Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells. *Nature* **316**, 701-705 (1985). https://doi.org:10.1038/316701a0

5 Cheifetz, S. *et al.* The transforming growth factor-beta system, a complex pattern of cross-reactive ligands and receptors. *Cell* **48**, 409-415 (1987). https://doi.org:10.1016/0092-8674(87)90192-9

6 Derynck, R. *et al.* A new type of transforming growth factor-beta, TGF-beta 3. *EMBO J* **7**, 3737-3743 (1988). https://doi.org:10.1002/j.1460-2075.1988.tb03257.x

7 ten Dijke, P., Hansen, P., Iwata, K. K., Pieler, C. & Foulkes, J. G. Identification of another member of the transforming growth factor type beta gene family. *Proc Natl Acad Sci U S A* **85**, 4715-4719 (1988). https://doi.org:10.1073/pnas.85.13.4715

8 Proetzel, G. *et al.* Transforming growth factor-beta 3 is required for secondary palate fusion. *Nat Genet* **11**, 409-414 (1995). https://doi.org:10.1038/ng1295-409

9 Kaartinen, V. *et al.* Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. *Nat Genet* **11**, 415-421 (1995). https://doi.org:10.1038/ng1295-415

10 Shull, M. M. *et al.* Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. *Nature* **359**, 693-699 (1992). https://doi.org:10.1038/359693a0

11 Sanford, L. P. *et al.* TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. *Development* **124**, 2659-2670 (1997).

12 Sanjabi, S., Oh, S. A. & Li, M. O. Regulation of the Immune Response by TGF-beta: From Conception to Autoimmunity and Infection. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022236

13 Geiser, A. G. *et al.* Transforming growth factor beta 1 (TGF-beta 1) controls expression of major histocompatibility genes in the postnatal mouse: aberrant histocompatibility antigen expression in the pathogenesis of the TGF-beta 1 null mouse phenotype. *Proc Natl Acad Sci U S A* **90**, 9944-9948 (1993). https://doi.org:10.1073/pnas.90.21.9944

14 Kulkarni, A. B. *et al.* Transforming growth factor beta 1 null mutation in mice causes excessive inflammatory response and early death. *Proc Natl Acad Sci U S A* **90**, 770-774 (1993). https://doi.org:10.1073/pnas.90.2.770

15 Belghith, M. *et al.* TGF-beta-dependent mechanisms mediate restoration of self-tolerance induced by antibodies to CD3 in overt autoimmune diabetes. *Nat Med* **9**, 1202-1208 (2003). https://doi.org:10.1038/nm924

16 Chen, W. *et al.* Conversion of peripheral CD4+CD25- naive T cells to CD4+CD25+ regulatory T cells by TGF-beta induction of transcription factor Foxp3. *J Exp Med* **198**, 1875-1886 (2003). https://doi.org:10.1084/jem.20030152

17 Aoki, C. A. *et al.* Transforming growth factor beta (TGF-beta) and autoimmunity. *Autoimmun Rev* **4**, 450-459 (2005). https://doi.org:10.1016/j.autrev.2005.03.006

18 Hyytiainen, M., Penttinen, C. & Keski-Oja, J. Latent TGF-beta binding proteins: extracellular matrix association and roles in TGF-beta activation. *Crit Rev Clin Lab Sci* **41**, 233-264 (2004). https://doi.org:10.1080/10408360490460933

19 Ruiz-Ortega, M., Rodriguez-Vita, J., Sanchez-Lopez, E., Carvajal, G. & Egido, J. TGF-beta signaling in vascular fibrosis. *Cardiovasc Res* **74**, 196-206 (2007). https://doi.org:10.1016/j.cardiores.2007.02.008

20 Leask, A. TGFbeta, cardiac fibroblasts, and the fibrotic response. *Cardiovasc Res* **74**, 207-212 (2007). https://doi.org:10.1016/j.cardiores.2006.07.012

21 Kishigami, S. & Mishina, Y. BMP signaling and early embryonic patterning. *Cytokine Growth Factor Rev* **16**, 265-278 (2005). https://doi.org:10.1016/j.cytogfr.2005.04.002

22 Salazar, V. S., Gamer, L. W. & Rosen, V. BMP signalling in skeletal development, disease and repair. *Nat Rev Endocrinol* **12**, 203-221 (2016). https://doi.org:10.1038/nrendo.2016.12

23 Corrigan, A. Z. *et al.* Evidence for an autocrine role of activin B within rat anterior pituitary cultures. *Endocrinology* **128**, 1682-1684 (1991). https://doi.org:10.1210/endo-128-3-1682

24 Wang, Q. F., Farnworth, P. G., Findlay, J. K. & Burger, H. G. Effect of purified 31K bovine inhibin on the specific binding of gonadotropin-releasing hormone to rat anterior pituitary

cells in culture. *Endocrinology* **123**, 2161-2166 (1988). https://doi.org:10.1210/endo-123-5-2161

25 Namwanje, M. & Brown, C. W. Activins and Inhibins: Roles in Development, Physiology, and Disease. *Cold Spring Harb Perspect Biol* **8** (2016). https://doi.org:10.1101/cshperspect.a021881

26 Rebbapragada, A., Benchabane, H., Wrana, J. L., Celeste, A. J. & Attisano, L. Myostatin signals through a transforming growth factor beta-like signaling pathway to block adipogenesis. *Mol Cell Biol* **23**, 7230-7242 (2003). https://doi.org:10.1128/MCB.23.20.7230-7242.2003

27 Schier, A. F. Nodal morphogens. *Cold Spring Harb Perspect Biol* **1**, a003459 (2009). https://doi.org:10.1101/cshperspect.a003459

28 Huminiecki, L. *et al.* Emergence, development and diversification of the TGF-beta signalling pathway within the animal kingdom. *BMC Evol Biol* **9**, 28 (2009). https://doi.org:10.1186/1471-2148-9-28

29 Liu, S. *et al.* Molecular Evolution of Transforming Growth Factor-beta (TGF-beta) Gene Family and the Functional Characterization of Lamprey TGF-beta2. *Front Immunol* **13**, 836226 (2022). https://doi.org:10.3389/fimmu.2022.836226

30 Hinck, A. P. Structural studies of the TGF-betas and their receptors - insights into evolution of the TGF-beta superfamily. *FEBS Lett* **586**, 1860-1870 (2012). https://doi.org:10.1016/j.febslet.2012.05.028

31 Franchini, A., Malagoli, D. & Ottaviani, E. Cytokines and invertebrates: TGF-beta and PDGF. *Curr Pharm Des* **12**, 3025-3031 (2006). https://doi.org:10.2174/138161206777947524

32 Jensen, P. A., Zheng, X., Lee, T. & O'Connor, M. B. The Drosophila Activin-like ligand Dawdle signals preferentially through one isoform of the Type-I receptor Baboon. *Mech Dev* **126**, 950-957 (2009). https://doi.org:10.1016/j.mod.2009.09.003

33 Hinck, A. P., Mueller, T. D. & Springer, T. A. Structural Biology and Evolution of the TGF-beta Family. *Cold Spring Harb Perspect Biol* **8** (2016). https://doi.org:10.1101/cshperspect.a022103

34 Wrana, J. L. *et al.* TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* **71**, 1003-1014 (1992). https://doi.org:10.1016/0092-8674(92)90395-s

35 Yamashita, H., ten Dijke, P., Franzen, P., Miyazono, K. & Heldin, C. H. Formation of hetero-oligomeric complexes of type I and type II receptors for transforming growth factor-beta. *J Biol Chem* **269**, 20172-20178 (1994).

36 Zuniga, J. E. *et al.* Assembly of TbetaRI:TbetaRII:TGFbeta ternary complex in vitro with receptor extracellular domains is cooperative and isoform-dependent. *J Mol Biol* **354**, 1052-1068 (2005). https://doi.org:10.1016/j.jmb.2005.10.014

37 Wrana, J. L., Attisano, L., Wieser, R., Ventura, F. & Massague, J. Mechanism of activation of the TGF-beta receptor. *Nature* **370**, 341-347 (1994). https://doi.org:10.1038/370341a0

38 Massague, J. & Wotton, D. Transcriptional control by the TGF-beta/Smad signaling system. *EMBO J* **19**, 1745-1754 (2000). https://doi.org:10.1093/emboj/19.8.1745

39 Bakin, A. V., Rinehart, C., Tomlinson, A. K. & Arteaga, C. L. p38 mitogen-activated protein kinase is required for TGFbeta-mediated fibroblastic transdifferentiation and cell migration. *J Cell Sci* **115**, 3193-3206 (2002). https://doi.org:10.1242/jcs.115.15.3193

40 Bakin, A. V., Tomlinson, A. K., Bhowmick, N. A., Moses, H. L. & Arteaga, C. L. Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. *J Biol Chem* **275**, 36803-36810 (2000). https://doi.org:10.1074/jbc.M005912200

41 Iyer, S. & Acharya, K. R. Tying the knot: the cystine signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. *FEBS J* **278**, 4304-4322 (2011). https://doi.org:10.1111/j.1742-4658.2011.08350.x

42 Gentry, L. E. & Nash, B. W. The pro domain of pre-pro-transforming growth factor beta 1 when independently expressed is a functional binding protein for the mature growth factor. *Biochemistry* **29**, 6851-6857 (1990). https://doi.org:10.1021/bi00481a014

43 Shi, M. *et al.* Latent TGF-beta structure and activation. *Nature* **474**, 343-349 (2011). https://doi.org:10.1038/nature10152

44 Taipale, J., Miyazono, K., Heldin, C. H. & Keski-Oja, J. Latent transforming growth factor-beta 1 associates to fibroblast extracellular matrix via latent TGF-beta binding protein. *J Cell Biol* **124**, 171-181 (1994). https://doi.org:10.1083/jcb.124.1.171

45 Zhao, B., Xu, S., Dong, X., Lu, C. & Springer, T. A. Prodomain-growth factor swapping in the structure of pro-TGF-beta1. *J Biol Chem* **293**, 1579-1589 (2018). https://doi.org:10.1074/jbc.M117.809657

46 Goebel, E. J., Hart, K. N., McCoy, J. C. & Thompson, T. B. Structural biology of the TGFbeta family. *Exp Biol Med (Maywood)* **244**, 1530-1546 (2019). https://doi.org:10.1177/1535370219880894

47 Huang, T., Schor, S. L. & Hinck, A. P. Biological activity differences between TGF-beta1 and TGF-beta3 correlate with differences in the rigidity and arrangement of their component monomers. *Biochemistry* **53**, 5737-5749 (2014). https://doi.org:10.1021/bi500647d

48 Zuniga, J. E. *et al.* The TbetaR-I pre-helix extension is structurally ordered in the unbound form and its flanking prolines are essential for binding. *J Mol Biol* **412**, 601-618 (2011). https://doi.org:10.1016/j.jmb.2011.07.046

49 Groppe, J. *et al.* Cooperative assembly of TGF-beta superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Mol Cell* **29**, 157-168 (2008). https://doi.org:10.1016/j.molcel.2007.11.039

50 Townson, S. A. *et al.* Specificity and structure of a high affinity activin receptor-like kinase 1 (ALK1) signaling complex. *J Biol Chem* **287**, 27313-27325 (2012). https://doi.org:10.1074/jbc.M112.377960

51 Mahlawat, P., Ilangovan, U., Biswas, T., Sun, L. Z. & Hinck, A. P. Structure of the Alk1 extracellular domain and characterization of its bone morphogenetic protein (BMP) binding properties. *Biochemistry* **51**, 6328-6341 (2012). https://doi.org:10.1021/bi300942x

52 Attisano, L., Wrana, J. L., Montalvo, E. & Massague, J. Activation of signalling by the activin receptor complex. *Mol Cell Biol* **16**, 1066-1073 (1996). https://doi.org:10.1128/MCB.16.3.1066

53 Liu, F., Ventura, F., Doody, J. & Massague, J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Mol Cell Biol* **15**, 3479-3486 (1995). https://doi.org:10.1128/MCB.15.7.3479

54 Nishitoh, H. *et al.* Identification of type I and type II serine/threonine kinase receptors for growth/differentiation factor-5. *J Biol Chem* **271**, 21345-21352 (1996). https://doi.org:10.1074/jbc.271.35.21345

55 De Crescenzo, G. *et al.* Three key residues underlie the differential affinity of the TGFbeta isoforms for the TGFbeta type II receptor. *J Mol Biol* **355**, 47-62 (2006). https://doi.org:10.1016/j.jmb.2005.10.022

56 Goebel, E. J. *et al.* Structural characterization of an activin class ternary receptor complex reveals a third paradigm for receptor specificity. *Proc Natl Acad Sci U S A* **116**, 15505-15513 (2019). https://doi.org:10.1073/pnas.1906253116

57 Hata, A. & Chen, Y. G. TGF-beta Signaling from Receptors to Smads. *Cold Spring Harb Perspect Biol* **8** (2016). https://doi.org:10.1101/cshperspect.a022061

58 Feng, X. H. & Derynck, R. A kinase subdomain of transforming growth factor-beta (TGF-beta) type I receptor determines the TGF-beta intracellular signaling specificity. *EMBO J* **16**, 3912-3923 (1997). https://doi.org:10.1093/emboj/16.13.3912

59 Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massague, J. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature* **383**, 832-836 (1996). https://doi.org:10.1038/383832a0

60 Miyazawa, K. & Miyazono, K. Regulation of TGF-beta Family Signaling by Inhibitory Smads. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022095

61 McGonigle, S., Beall, M. J., Feeney, E. L. & Pearce, E. J. Conserved role for 14-3-3epsilon downstream of type I TGFbeta receptors. *FEBS Lett* **490**, 65-69 (2001). https://doi.org:10.1016/s0014-5793(01)02133-0

62 Engel, M. E., McDonnell, M. A., Law, B. K. & Moses, H. L. Interdependent SMAD and JNK signaling in transforming growth factor-beta-mediated transcription. *J Biol Chem* **274**, 37413-37420 (1999). https://doi.org:10.1074/jbc.274.52.37413

63 Zhang, Y. E. Non-Smad Signaling Pathways of the TGF-beta Family. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022129

64 Morikawa, M., Derynck, R. & Miyazono, K. TGF-beta and the TGF-beta Family: Context-Dependent Roles in Cell and Tissue Physiology. *Cold Spring Harb Perspect Biol* **8** (2016). https://doi.org:10.1101/cshperspect.a021873

65 Tucker, R. F., Shipley, G. D., Moses, H. L. & Holley, R. W. Growth inhibitor from BSC-1 cells closely related to platelet type beta transforming growth factor. *Science* **226**, 705-707 (1984). https://doi.org:10.1126/science.6093254

66 Roberts, A. B. *et al.* Type beta transforming growth factor: a bifunctional regulator of cellular growth. *Proc Natl Acad Sci U S A* **82**, 119-123 (1985). https://doi.org:10.1073/pnas.82.1.119

67 Siegel, P. M. & Massague, J. Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* **3**, 807-821 (2003). https://doi.org:10.1038/nrc1208

68 Jahn, S. C., Law, M. E., Corsino, P. E. & Law, B. K. TGF-beta antiproliferative effects in tumor suppression. *Front Biosci (Schol Ed)* **4**, 749-766 (2012). https://doi.org:10.2741/s297

69 Reynisdottir, I., Polyak, K., Iavarone, A. & Massague, J. Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. *Genes Dev* **9**, 1831-1845 (1995). https://doi.org:10.1101/gad.9.15.1831

70 Moustakas, A. & Kardassis, D. Regulation of the human p21/WAF1/Cip1 promoter in hepatic cells by functional interactions between Sp1 and Smad family members. *Proc Natl Acad Sci U S A* **95**, 6733-6738 (1998). https://doi.org:10.1073/pnas.95.12.6733

71 Iavarone, A. & Massague, J. Repression of the CDK activator Cdc25A and cell-cycle arrest by cytokine TGF-beta in cells lacking the CDK inhibitor p15. *Nature* **387**, 417-422 (1997). https://doi.org:10.1038/387417a0

72 Yamasaki, K. *et al.* Keratinocyte growth inhibition by high-dose epidermal growth factor is mediated by transforming growth factor beta autoinduction: a negative feedback

mechanism for keratinocyte growth. *J Invest Dermatol* **120**, 1030-1037 (2003). https://doi.org:10.1046/j.1523-1747.2003.12239.x

73 Meyers, E. A. & Kessler, J. A. TGF-beta Family Signaling in Neural and Neuronal Differentiation, Development, and Function. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022244

74 Mullen, A. C. & Wrana, J. L. TGF-beta Family Signaling in Embryonic and Somatic Stem-Cell Renewal and Differentiation. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022186

75 Kahata, K., Dadras, M. S. & Moustakas, A. TGF-beta Family Signaling in Epithelial Differentiation and Epithelial-Mesenchymal Transition. *Cold Spring Harb Perspect Biol* **10** (2018). https://doi.org:10.1101/cshperspect.a022194

76 Schock, F. & Perrimon, N. Molecular mechanisms of epithelial morphogenesis. *Annu Rev Cell Dev Biol* **18**, 463-493 (2002). https://doi.org:10.1146/annurev.cellbio.18.022602.131838

77 Mercado-Pimentel, M. E. & Runyan, R. B. Multiple transforming growth factor-beta isoforms and receptors function during epithelial-mesenchymal cell transformation in the embryonic heart. *Cells Tissues Organs* **185**, 146-156 (2007). https://doi.org:10.1159/000101315

78 Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* **127**, 2021-2036 (1994). https://doi.org:10.1083/jcb.127.6.2021

79 Nawshad, A., LaGamba, D. & Hay, E. D. Transforming growth factor beta (TGFbeta) signalling in palatal growth, apoptosis and epithelial mesenchymal transformation (EMT). *Arch Oral Biol* **49**, 675-689 (2004). https://doi.org:10.1016/j.archoralbio.2004.05.007

80 Moustakas, A. & Heldin, C. H. Mechanisms of TGFbeta-Induced Epithelial-Mesenchymal Transition. *J Clin Med* **5** (2016). https://doi.org:10.3390/jcm5070063

81 Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* **15**, 178-196 (2014). https://doi.org:10.1038/nrm3758

82 Nawshad, A., Medici, D., Liu, C. C. & Hay, E. D. TGFbeta3 inhibits E-cadherin gene expression in palate medial-edge epithelial cells through a Smad2-Smad4-LEF1 transcription complex. *J Cell Sci* **120**, 1646-1653 (2007). https://doi.org:10.1242/jcs.003129

83 Tsapara, A. *et al.* The RhoA activator GEF-H1/Lfc is a transforming growth factor-beta target gene and effector that regulates alpha-smooth muscle actin expression and cell migration. *Mol Biol Cell* **21**, 860-870 (2010). https://doi.org:10.1091/mbc.E09-07-0567

84 Azhar, M. *et al.* Transforming growth factor beta in cardiovascular development and function. *Cytokine Growth Factor Rev* **14**, 391-407 (2003). https://doi.org:10.1016/s1359-6101(03)00044-3

85 Iwata, J., Parada, C. & Chai, Y. The mechanism of TGF-beta signaling during palate development. *Oral Dis* **17**, 733-744 (2011). https://doi.org:10.1111/j.1601-0825.2011.01806.x

86 Sporn, M. B. *et al.* Polypeptide transforming growth factors isolated from bovine sources and used for wound healing in vivo. *Science* **219**, 1329-1331 (1983). https://doi.org:10.1126/science.6572416

87 Ignotz, R. A. & Massague, J. Transforming growth factor-beta stimulates the expression of fibronectin and collagen and their incorporation into the extracellular matrix. *J Biol Chem* **261**, 4337-4345 (1986).

88 Edwards, D. R. *et al.* Transforming growth factor beta modulates the expression of collagenase and metalloproteinase inhibitor. *EMBO J* **6**, 1899-1904 (1987). https://doi.org:10.1002/j.1460-2075.1987.tb02449.x

89 Shah, M., Foreman, D. M. & Ferguson, M. W. Neutralising antibody to TGF-beta 1,2 reduces cutaneous scarring in adult rodents. *J Cell Sci* **107 (Pt 5)**, 1137-1157 (1994). https://doi.org:10.1242/jcs.107.5.1137

90 Shah, M., Foreman, D. M. & Ferguson, M. W. Neutralisation of TGF-beta 1 and TGF-beta 2 or exogenous addition of TGF-beta 3 to cutaneous rat wounds reduces scarring. *J Cell Sci* **108 (Pt 3)**, 985-1002 (1995). https://doi.org:10.1242/jcs.108.3.985

91 Ferguson, M. W. *et al.* Prophylactic administration of avotermin for improvement of skin scarring: three double-blind, placebo-controlled, phase I/II studies. *Lancet* **373**, 1264-1274 (2009). https://doi.org:10.1016/S0140-6736(09)60322-6

92 Ouyang, W. *et al.* TGF-beta cytokine signaling promotes CD8+ T cell development and low-affinity CD4+ T cell homeostasis by regulation of interleukin-7 receptor alpha expression. *Immunity* **39**, 335-346 (2013). https://doi.org:10.1016/j.immuni.2013.07.016

93 Ouyang, W., Beckett, O., Ma, Q. & Li, M. O. Transforming growth factor-beta signaling curbs thymic negative selection promoting regulatory T cell development. *Immunity* **32**, 642-653 (2010). https://doi.org:10.1016/j.immuni.2010.04.012

94 Liu, Y. *et al.* A critical function for TGF-beta signaling in the development of natural CD4+CD25+Foxp3+ regulatory T cells. *Nat Immunol* **9**, 632-640 (2008). https://doi.org:10.1038/ni.1607

95 Tzachanis, D. *et al.* Tob is a negative regulator of activation that is expressed in anergic and quiescent T cells. *Nat Immunol* **2**, 1174-1182 (2001). https://doi.org:10.1038/ni730

96 Brabletz, T. *et al.* Transforming growth factor beta and cyclosporin A inhibit the inducible activity of the interleukin-2 gene in T cells through a noncanonical octamer-binding site. *Mol Cell Biol* **13**, 1155-1162 (1993). https://doi.org:10.1128/mcb.13.2.1155-1162.1993

97 Sanjabi, S., Zenewicz, L. A., Kamanaka, M. & Flavell, R. A. Anti-inflammatory and pro-inflammatory roles of TGF-beta, IL-10, and IL-22 in immunity and autoimmunity. *Curr Opin Pharmacol* **9**, 447-453 (2009). https://doi.org:10.1016/j.coph.2009.04.008

98 Bouchard, C., Fridman, W. H. & Sautes, C. Effect of TGF-beta1 on cell cycle regulatory proteins in LPS-stimulated normal mouse B lymphocytes. *J Immunol* **159**, 4155-4164 (1997).

99 Warner, G. L., Ludlow, J. W., Nelson, D. A., Gaur, A. & Scott, D. W. Anti-immunoglobulin treatment of murine B-cell lymphomas induces active transforming growth factor beta but pRB hypophosphorylation is transforming growth factor beta independent. *Cell Growth Differ* **3**, 175-181 (1992).

100 Kee, B. L., Rivera, R. R. & Murre, C. Id3 inhibits B lymphocyte progenitor growth and survival in response to TGF-beta. *Nat Immunol* **2**, 242-247 (2001). https://doi.org:10.1038/85303

101 Cazac, B. B. & Roes, J. TGF-beta receptor controls B cell responsiveness and induction of IgA in vivo. *Immunity* **13**, 443-451 (2000). https://doi.org:10.1016/s1074-7613(00)00044-3

102 Coffman, R. L., Lebman, D. A. & Shrader, B. Transforming growth factor beta specifically enhances IgA production by lipopolysaccharide-stimulated murine B lymphocytes. *J Exp Med* **170**, 1039-1044 (1989). https://doi.org:10.1084/jem.170.3.1039

103 Gomez, G. *et al.* TGF-beta 1 inhibits mast cell Fc epsilon RI expression. *J Immunol* **174**, 5987-5993 (2005). https://doi.org:10.4049/jimmunol.174.10.5987

104 Ganeshan, K. & Bryce, P. J. Regulatory T cells enhance mast cell production of IL-6 via surface-bound TGF-beta. *J Immunol* **188**, 594-603 (2012). https://doi.org:10.4049/jimmunol.1102389

105 Wahl, S. M. *et al.* Transforming growth factor type beta induces monocyte chemotaxis and growth factor production. *Proc Natl Acad Sci U S A* **84**, 5788-5792 (1987). https://doi.org:10.1073/pnas.84.16.5788

106 Li, M. O., Wan, Y. Y., Sanjabi, S., Robertson, A. K. & Flavell, R. A. Transforming growth factor-beta regulation of immune responses. *Annu Rev Immunol* **24**, 99-146 (2006). https://doi.org:10.1146/annurev.immunol.24.021605.090737

107 Marie, J. C., Liggitt, D. & Rudensky, A. Y. Cellular mechanisms of fatal early-onset autoimmunity in mice with the T cell-specific targeting of transforming growth factor-beta receptor. *Immunity* **25**, 441-454 (2006). https://doi.org:10.1016/j.immuni.2006.07.012

108 Gorelik, L. & Flavell, R. A. Abrogation of TGFbeta signaling in T cells leads to spontaneous T cell differentiation and autoimmune disease. *Immunity* **12**, 171-181 (2000). https://doi.org:10.1016/s1074-7613(00)80170-3

109 Biancheri, P. *et al.* The role of transforming growth factor (TGF)-beta in modulating the immune response and fibrogenesis in the gut. *Cytokine Growth Factor Rev* **25**, 45-55 (2014). https://doi.org:10.1016/j.cytogfr.2013.11.001

110 Feagins, L. A. Role of transforming growth factor-beta in inflammatory bowel disease and colitis-associated colon cancer. *Inflamm Bowel Dis* **16**, 1963-1968 (2010). https://doi.org:10.1002/ibd.21281

111 Fuss, I. J., Boirivant, M., Lacy, B. & Strober, W. The interrelated roles of TGF-beta and IL-10 in the regulation of experimental colitis. *J Immunol* **168**, 900-908 (2002). https://doi.org:10.4049/jimmunol.168.2.900

112 Di Giacinto, C., Marinaro, M., Sanchez, M., Strober, W. & Boirivant, M. Probiotics ameliorate recurrent Th1-mediated murine colitis by inducing IL-10 and IL-10-dependent TGF-beta-bearing regulatory cells. *J Immunol* **174**, 3237-3246 (2005). https://doi.org:10.4049/jimmunol.174.6.3237

113 Glocker, E. O. *et al.* Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med* **361**, 2033-2045 (2009). https://doi.org:10.1056/NEJMoa0907206

114 Naviglio, S. *et al.* Severe inflammatory bowel disease associated with congenital alteration of transforming growth factor beta signaling. *J Crohns Colitis* **8**, 770-774 (2014). https://doi.org:10.1016/j.crohns.2014.01.013

115 Kang, S. S. *et al.* An antibiotic-responsive mouse model of fulminant ulcerative colitis. *PLoS Med* **5**, e41 (2008). https://doi.org:10.1371/journal.pmed.0050041

116 Bira, Y. *et al.* Transforming growth factor beta stimulates rheumatoid synovial fibroblasts via the type II receptor. *Mod Rheumatol* **15**, 108-113 (2005). https://doi.org:10.1007/s10165-004-0378-2

117 Gonzalo-Gil, E. & Galindo-Izquierdo, M. Role of transforming growth factor-beta (TGF) beta in the physiopathology of rheumatoid arthritis. *Reumatol Clin* **10**, 174-179 (2014). https://doi.org:10.1016/j.reuma.2014.01.009

118 Johns, L. D., Flanders, K. C., Ranges, G. E. & Sriram, S. Successful treatment of experimental allergic encephalomyelitis with transforming growth factor-beta 1. *J Immunol* **147**, 1792-1796 (1991).

119 Racke, M. K. *et al.* Prevention and treatment of chronic relapsing experimental allergic encephalomyelitis by transforming growth factor-beta 1. *J Immunol* **146**, 3012-3017 (1991).

120 Mokhtarian, F. *et al.* Defective production of anti-inflammatory cytokine, TGF-beta by T cell lines of patients with active multiple sclerosis. *J Immunol* **152**, 6003-6010 (1994).

121 Gutcher, I. *et al.* Autocrine transforming growth factor-beta1 promotes in vivo Th17 cell differentiation. *Immunity* **34**, 396-408 (2011). https://doi.org:10.1016/j.immuni.2011.03.005

122 Ye, Z. & Hu, Y. TGF-beta1: Gentlemanly orchestrator in idiopathic pulmonary fibrosis (Review). *Int J Mol Med* **48** (2021). https://doi.org:10.3892/ijmm.2021.4965

123 Zheng, X., Qi, C., Zhang, S., Fang, Y. & Ning, W. TGF-beta1 induces Fstl1 via the Smad3-c-Jun pathway in lung fibroblasts. *Am J Physiol Lung Cell Mol Physiol* **313**, L240-L251 (2017). https://doi.org:10.1152/ajplung.00523.2016

124 Tanjore, H. *et al.* Alveolar epithelial cells undergo epithelial-to-mesenchymal transition in response to endoplasmic reticulum stress. *J Biol Chem* **286**, 30972-30980 (2011). https://doi.org:10.1074/jbc.M110.181164

125 Khan, S., Joyce, J., Margulies, K. B. & Tsuda, T. Enhanced bioactive myocardial transforming growth factor-beta in advanced human heart failure. *Circ J* **78**, 2711-2718 (2014). https://doi.org:10.1253/circj.cj-14-0511

126 Isaka, Y. Targeting TGF-beta Signaling in Kidney Fibrosis. *Int J Mol Sci* **19** (2018). https://doi.org:10.3390/ijms19092532

127 Wu, C. F. *et al.* Transforming growth factor beta-1 stimulates profibrotic epithelial signaling to activate pericyte-myofibroblast transition in obstructive kidney fibrosis. *Am J Pathol* **182**, 118-131 (2013). https://doi.org:10.1016/j.ajpath.2012.09.009

128 Yamamoto, T. *et al.* Expression of types I, II, and III TGF-beta receptors in human glomerulonephritis. *J Am Soc Nephrol* **9**, 2253-2261 (1998). https://doi.org:10.1681/ASN.V9122253

129 Kubiczkova, L., Sedlarikova, L., Hajek, R. & Sevcikova, S. TGF-beta - an excellent servant but a bad master. *J Transl Med* **10**, 183 (2012). https://doi.org:10.1186/1479-5876-10-183

130 Biswas, S. *et al.* Transforming growth factor beta receptor type II inactivation promotes the establishment and progression of colon cancer. *Cancer Res* **64**, 4687-4692 (2004). https://doi.org:10.1158/0008-5472.CAN-03-3255

131 Xu, Q., Norman, J. T., Shrivastav, S., Lucio-Cazana, J. & Kopp, J. B. In vitro models of TGF-beta-induced fibrosis suitable for high-throughput screening of antifibrotic agents. *Am J Physiol Renal Physiol* **293**, F631-640 (2007). https://doi.org:10.1152/ajprenal.00379.2006

132 Brown, J. A. *et al.* TGF-beta-Induced Quiescence Mediates Chemoresistance of Tumor-Propagating Cells in Squamous Cell Carcinoma. *Cell Stem Cell* **21**, 650-664 e658 (2017). https://doi.org:10.1016/j.stem.2017.10.001

133 Mumm, J. B. & Oft, M. Cytokine-based transformation of immune surveillance into tumor-promoting inflammation. *Oncogene* **27**, 5913-5919 (2008). https://doi.org:10.1038/onc.2008.275

134 Sharma, A. & Rudra, D. Emerging Functions of Regulatory T Cells in Tissue Homeostasis. *Front Immunol* **9**, 883 (2018). https://doi.org:10.3389/fimmu.2018.00883

135 Ohue, Y. & Nishikawa, H. Regulatory T (Treg) cells in cancer: Can Treg cells be a new therapeutic target? *Cancer Sci* **110**, 2080-2089 (2019). https://doi.org:10.1111/cas.14069

136 Reiss, M. & Barcellos-Hoff, M. H. Transforming growth factor-beta in breast cancer: a working hypothesis. *Breast Cancer Res Treat* **45**, 81-95 (1997). https://doi.org:10.1023/a:1005865812918

137 Junker, U. *et al.* Transforming growth factor beta 1 is significantly elevated in plasma of patients suffering from renal cell carcinoma. *Cytokine* **8**, 794-798 (1996). https://doi.org:10.1006/cyto.1996.0105

138 Krasagakis, K. *et al.* Elevated plasma levels of transforming growth factor (TGF)-beta1 and TGF-beta2 in patients with disseminated malignant melanoma. *Br J Cancer* **77**, 1492-1494 (1998). https://doi.org:10.1038/bjc.1998.245

139 Wikstrom, P., Stattin, P., Franck-Lissbrant, I., Damber, J. E. & Bergh, A. Transforming growth factor beta1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer. *Prostate* **37**, 19-29 (1998). https://doi.org:10.1002/(sici)1097-0045(19980915)37:1<19::aid-pros4>3.0.co;2-3

140 Teixeira, A. F., Ten Dijke, P. & Zhu, H. J. On-Target Anti-TGF-beta Therapies Are Not Succeeding in Clinical Cancer Treatments: What Are Remaining Challenges? *Front Cell Dev Biol* **8**, 605 (2020). https://doi.org:10.3389/fcell.2020.00605

141 Principe, D. R. *et al.* TGFbeta Blockade Augments PD-1 Inhibition to Promote T-Cell-Mediated Regression of Pancreatic Cancer. *Mol Cancer Ther* **18**, 613-620 (2019). https://doi.org:10.1158/1535-7163.MCT-18-0850

142 Sow, H. S., Ren, J., Camps, M., Ossendorp, F. & Ten Dijke, P. Combined Inhibition of TGF-beta Signaling and the PD-L1 Immune Checkpoint Is Differentially Effective in Tumor Models. *Cells* **8** (2019). https://doi.org:10.3390/cells8040320

143 Kim, B. G., Malek, E., Choi, S. H., Ignatz-Hoover, J. J. & Driscoll, J. J. Novel therapies emerging in oncology to target the TGF-beta pathway. *J Hematol Oncol* **14**, 55 (2021). https://doi.org:10.1186/s13045-021-01053-x

144 Fujiwara, Y. *et al.* Phase 1 study of galunisertib, a TGF-beta receptor I kinase inhibitor, in Japanese patients with advanced solid tumors. *Cancer Chemother Pharmacol* **76**, 1143-1152 (2015). https://doi.org:10.1007/s00280-015-2895-4

145 Wick, A. *et al.* Phase 1b/2a study of galunisertib, a small molecule inhibitor of transforming growth factor-beta receptor I, in combination with standard temozolomide-based radiochemotherapy in patients with newly diagnosed malignant glioma. *Invest New Drugs* **38**, 1570-1579 (2020). https://doi.org:10.1007/s10637-020-00910-9

146 Yap, T. A. *et al.* First-In-Human Phase I Study of a Next-Generation, Oral, TGFbeta Receptor 1 Inhibitor, LY3200882, in Patients with Advanced Cancer. *Clin Cancer Res* **27**, 6666-6676 (2021). https://doi.org:10.1158/1078-0432.CCR-21-1504

147 Welsh, B. T. *et al.* Nonclinical Development of SRK-181: An Anti-Latent TGFbeta1 Monoclonal Antibody for the Treatment of Locally Advanced or Metastatic Solid Tumors. *Int J Toxicol* **40**, 226-241 (2021). https://doi.org:10.1177/1091581821998945

148 Mitra, M. S. *et al.* A Potent Pan-TGFbeta Neutralizing Monoclonal Antibody Elicits Cardiovascular Toxicity in Mice and Cynomolgus Monkeys. *Toxicol Sci* **175**, 24-34 (2020). https://doi.org:10.1093/toxsci/kfaa024

149 Morris, J. C. *et al.* Phase I study of GC1008 (fresolimumab): a human anti-transforming growth factor-beta (TGFbeta) monoclonal antibody in patients with advanced malignant melanoma or renal cell carcinoma. *PLoS One* **9**, e90353 (2014). https://doi.org:10.1371/journal.pone.0090353

150 Tolcher, A. W. *et al.* A phase 1 study of anti-TGFbeta receptor type-II monoclonal antibody LY3022859 in patients with advanced solid tumors. *Cancer Chemother Pharmacol* **79**, 673-680 (2017). https://doi.org:10.1007/s00280-017-3245-5

151 Varricchio, L. *et al.* TGF-beta1 protein trap AVID200 beneficially affects hematopoiesis and bone marrow fibrosis in myelofibrosis. *JCI Insight* **6** (2021). https://doi.org:10.1172/jci.insight.145651

152 Strauss, J. *et al.* Phase I Trial of M7824 (MSB0011359C), a Bifunctional Fusion Protein Targeting PD-L1 and TGFbeta, in Advanced Solid Tumors. *Clin Cancer Res* **24**, 1287-1295 (2018). https://doi.org:10.1158/1078-0432.CCR-17-2653

153 Paz-Ares, L. *et al.* Bintrafusp Alfa, a Bifunctional Fusion Protein Targeting TGF-beta and PD-L1, in Second-Line Treatment of Patients With NSCLC: Results From an Expansion Cohort of a Phase 1 Trial. *J Thorac Oncol* **15**, 1210-1222 (2020). https://doi.org:10.1016/j.jtho.2020.03.003

154 Massague, J. TGFbeta signalling in context. *Nat Rev Mol Cell Biol* **13**, 616-630 (2012). https://doi.org:10.1038/nrm3434

155 Greenwood, B. M. Autoimmune disease and parasitic infections in Nigerians. *Lancet* **2**, 380-382 (1968). https://doi.org:10.1016/s0140-6736(68)90595-3

156 Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. *Lancet* **351**, 1225-1232 (1998).

157 Greenwood, B. M., Herrick, E. M. & Voller, A. Suppression of autoimmune disease in NZB and (NZB x NZW) F1 hybrid mice by infection with malaria. *Nature* **226**, 266-267 (1970). https://doi.org:10.1038/226266a0

158 Strachan, D. P. Hay fever, hygiene, and household size. *BMJ* **299**, 1259-1260 (1989). https://doi.org:10.1136/bmj.299.6710.1259

159 Maizels, R. M., McSorley, H. J. & Smyth, D. J. Helminths in the hygiene hypothesis: sooner or later? *Clin Exp Immunol* **177**, 38-46 (2014). https://doi.org:10.1111/cei.12353

160 Maizels, R. M. & McSorley, H. J. Regulation of the host immune system by helminth parasites. *J Allergy Clin Immunol* **138**, 666-675 (2016). https://doi.org:10.1016/j.jaci.2016.07.007

161 Zakeri, A., Hansen, E. P., Andersen, S. D., Williams, A. R. & Nejsum, P. Immunomodulation by Helminths: Intracellular Pathways and Extracellular Vesicles. *Front Immunol* **9**, 2349 (2018). https://doi.org:10.3389/fimmu.2018.02349

162 van Liempt, E. *et al.* Schistosoma mansoni soluble egg antigens are internalized by human dendritic cells through multiple C-type lectins and suppress TLR-induced dendritic cell activation. *Mol Immunol* **44**, 2605-2615 (2007). https://doi.org:10.1016/j.molimm.2006.12.012

163 Zakeri, A., Borji, H. & Haghparast, A. Interaction Between Helminths and Toll-Like Receptors: Possibilities and Potentials for Asthma Therapy. *Int Rev Immunol* **35**, 219-248 (2016). https://doi.org:10.3109/08830185.2015.1096936

164 Thawer, S. *et al.* Surfactant Protein-D Is Essential for Immunity to Helminth Infection. *PLoS Pathog* **12**, e1005461 (2016). https://doi.org:10.1371/journal.ppat.1005461

165 Maizels, R. M., Smits, H. H. & McSorley, H. J. Modulation of Host Immunity by Helminths: The Expanding Repertoire of Parasite Effector Molecules. *Immunity* **49**, 801-818 (2018). https://doi.org:10.1016/j.immuni.2018.10.016

166 Ramos-Benitez, M. J., Ruiz-Jimenez, C., Aguayo, V. & Espino, A. M. Recombinant Fasciola hepatica fatty acid binding protein suppresses toll-like receptor stimulation in response to multiple bacterial ligands. *Sci Rep* **7**, 5455 (2017). https://doi.org:10.1038/s41598-017-05735-w

167 Figueroa-Santiago, O. & Espino, A. M. Fasciola hepatica fatty acid binding protein induces the alternative activation of human macrophages. *Infect Immun* **82**, 5005-5012 (2014). https://doi.org:10.1128/IAI.02541-14

168 Abdelaziz, M. H. *et al.* Alternatively activated macrophages; a double-edged sword in allergic asthma. *J Transl Med* **18**, 58 (2020). https://doi.org:10.1186/s12967-020-02251-w

169 Robinson, M. W. *et al.* A family of helminth molecules that modulate innate cell responses via molecular mimicry of host antimicrobial peptides. *PLoS Pathog* **7**, e1002042 (2011). https://doi.org:10.1371/journal.ppat.1002042

170 McSorley, H. J. & Maizels, R. M. Helminth infections and host immune regulation. *Clin Microbiol Rev* **25**, 585-608 (2012). https://doi.org:10.1128/CMR.05040-11

171 Taylor, M. D. *et al.* CTLA-4 and CD4+ CD25+ regulatory T cells inhibit protective immunity to filarial parasites in vivo. *J Immunol* **179**, 4626-4634 (2007). https://doi.org:10.4049/jimmunol.179.7.4626

172 Catalan, D. *et al.* Immunosuppressive Mechanisms of Regulatory B Cells. *Front Immunol* **12**, 611795 (2021). https://doi.org:10.3389/fimmu.2021.611795

173 Smits, H. H. *et al.* Protective effect of Schistosoma mansoni infection on allergic airway inflammation depends on the intensity and chronicity of infection. *J Allergy Clin Immunol* **120**, 932-940 (2007). https://doi.org:10.1016/j.jaci.2007.06.009

174 Amu, S. *et al.* Regulatory B cells prevent and reverse allergic airway inflammation via FoxP3-positive T regulatory cells in a murine model. *J Allergy Clin Immunol* **125**, 1114-1124 e1118 (2010). https://doi.org:10.1016/j.jaci.2010.01.018

175 Lv, K., Zhang, Y., Zhang, M., Zhong, M. & Suo, Q. Galectin-9 promotes TGF-beta1-dependent induction of regulatory T cells via the TGF-beta/Smad signaling pathway. *Mol Med Rep* **7**, 205-210 (2013). https://doi.org:10.3892/mmr.2012.1125

176 Kim, J. Y. *et al.* Inhibition of dextran sulfate sodium (DSS)-induced intestinal inflammation via enhanced IL-10 and TGF-beta production by galectin-9 homologues isolated from intestinal parasites. *Mol Biochem Parasitol* **174**, 53-61 (2010). https://doi.org:10.1016/j.molbiopara.2010.06.014

177 Manoury, B., Gregory, W. F., Maizels, R. M. & Watts, C. Bm-CPI-2, a cystatin homolog secreted by the filarial parasite Brugia malayi, inhibits class II MHC-restricted antigen processing. *Curr Biol* **11**, 447-451 (2001). https://doi.org:10.1016/s0960-9822(01)00118-x

178 Lustigman, S., Brotman, B., Huima, T., Prince, A. M. & McKerrow, J. H. Molecular cloning and characterization of onchocystatin, a cysteine proteinase inhibitor of Onchocerca volvulus. *J Biol Chem* **267**, 17339-17346 (1992).

179 Driss, V. *et al.* The schistosome glutathione S-transferase P28GST, a unique helminth protein, prevents intestinal inflammation in experimental colitis through a Th2-type response with mucosal eosinophils. *Mucosal Immunol* **9**, 322-335 (2016). https://doi.org:10.1038/mi.2015.62

180 Norton, R. S., Pennington, M. W. & Wulff, H. Potassium channel blockade by the sea anemone toxin ShK for the treatment of multiple sclerosis and other autoimmune diseases. *Curr Med Chem* **11**, 3041-3052 (2004). https://doi.org:10.2174/0929867043363947

181 Chhabra, S. *et al.* Kv1.3 channel-blocking immunomodulatory peptides from parasitic worms: implications for autoimmune diseases. *FASEB J* **28**, 3952-3964 (2014). https://doi.org:10.1096/fj.14-251967

182 Dittrich, A. M. *et al.* Helminth infection with Litomosoides sigmodontis induces regulatory T cells and inhibits allergic sensitization, airway inflammation, and hyperreactivity in a murine asthma model. *J Immunol* **180**, 1792-1799 (2008). https://doi.org:10.4049/jimmunol.180.3.1792

183 Pacifico, L. G. *et al.* Schistosoma mansoni antigens modulate experimental allergic asthma in a murine model: a major role for CD4+ CD25+ Foxp3+ T cells independent of interleukin-10. *Infect Immun* **77**, 98-107 (2009). https://doi.org:10.1128/IAI.00783-07

184 Arrais, M. *et al.* Helminth infections and allergic diseases: Systematic review and meta-analysis of the global literature. *J Allergy Clin Immunol* **149**, 2139-2152 (2022). https://doi.org:10.1016/j.jaci.2021.12.777

185 Sur, L. M. *et al.* Antinuclear Antibodies: Marker of Diagnosis and Evolution in Autoimmune Diseases. *Lab Med* **49**, e62-e73 (2018). https://doi.org:10.1093/labmed/lmy024

186 Mutapi, F. *et al.* Schistosome infection intensity is inversely related to auto-reactive antibody levels. *PLoS One* **6**, e19149 (2011). https://doi.org:10.1371/journal.pone.0019149

187 Osada, Y., Shimizu, S., Kumagai, T., Yamada, S. & Kanazawa, T. Schistosoma mansoni infection reduces severity of collagen-induced arthritis via down-regulation of pro-inflammatory mediators. *Int J Parasitol* **39**, 457-464 (2009). https://doi.org:10.1016/j.ijpara.2008.08.007

188 He, Y. *et al.* The inhibitory effect against collagen-induced arthritis by Schistosoma japonicum infection is infection stage-dependent. *BMC Immunol* **11**, 28 (2010). https://doi.org:10.1186/1471-2172-11-28

189 Hubner, M. P., Stocker, J. T. & Mitre, E. Inhibition of type 1 diabetes in filaria-infected non-obese diabetic mice is associated with a T helper type 2 shift and induction of FoxP3+ regulatory T cells. *Immunology* **127**, 512-522 (2009). https://doi.org:10.1111/j.1365-2567.2008.02958.x

190 Hubner, M. P. *et al.* Helminth protection against autoimmune diabetes in nonobese diabetic mice is independent of a type 2 immune shift and requires TGF-beta. *J Immunol* **188**, 559-568 (2012). https://doi.org:10.4049/jimmunol.1100335

191 Maizels, R. M. Parasitic helminth infections and the control of human allergic and autoimmune disorders. *Clin Microbiol Infect* **22**, 481-486 (2016). https://doi.org:10.1016/j.cmi.2016.04.024

192 Summers, R. W. *et al.* Trichuris suis seems to be safe and possibly effective in the treatment of inflammatory bowel disease. *Am J Gastroenterol* **98**, 2034-2041 (2003). https://doi.org:10.1111/j.1572-0241.2003.07660.x

193 Summers, R. W., Elliott, D. E., Urban, J. F., Jr., Thompson, R. A. & Weinstock, J. V. Trichuris suis therapy for active ulcerative colitis: a randomized controlled trial. *Gastroenterology* **128**, 825-832 (2005). https://doi.org:10.1053/j.gastro.2005.01.005

194 Weinstock, J. V. & Elliott, D. E. Translatability of helminth therapy in inflammatory bowel diseases. *Int J Parasitol* **43**, 245-251 (2013). https://doi.org:10.1016/j.ijpara.2012.10.016

195 Feary, J. *et al.* Safety of hookworm infection in individuals with measurable airway responsiveness: a randomized placebo-controlled feasibility study. *Clin Exp Allergy* **39**, 1060-1068 (2009). https://doi.org:10.1111/j.1365-2222.2009.03187.x

196 Feary, J. R. *et al.* Experimental hookworm infection: a randomized placebo-controlled trial in asthma. *Clin Exp Allergy* **40**, 299-306 (2010). https://doi.org:10.1111/j.1365-2222.2009.03433.x

197 Daveson, A. J. *et al.* Effect of hookworm infection on wheat challenge in celiac disease--a randomised double-blinded placebo controlled trial. *PLoS One* **6**, e17366 (2011). https://doi.org:10.1371/journal.pone.0017366

198 Marchiondo, A. A., Cruthers, L. R. & Reinemeyer, C. R. *Nematoda*. Vol. 2 135-335 (2019).

199 Page, A. P. & Johnstone, I. L. The cuticle. *WormBook*, 1-15 (2007). https://doi.org:10.1895/wormbook.1.138.1

200 Johnston, C. J. *et al.* Cultivation of Heligmosomoides polygyrus: an immunomodulatory nematode parasite and its secreted products. *J Vis Exp*, e52412 (2015). https://doi.org:10.3791/52412

201 Reynolds, L. A., Filbey, K. J. & Maizels, R. M. Immunity to the model intestinal helminth parasite Heligmosomoides polygyrus. *Semin Immunopathol* **34**, 829-846 (2012). https://doi.org:10.1007/s00281-012-0347-3

202 Svetic, A. *et al.* A primary intestinal helminthic infection rapidly induces a gut-associated elevation of Th2-associated cytokines and IL-3. *J Immunol* **150**, 3434-3441 (1993).

203 Setiawan, T. *et al.* Heligmosomoides polygyrus promotes regulatory T-cell cytokine production in the murine normal distal intestine. *Infect Immun* **75**, 4655-4663 (2007). https://doi.org:10.1128/IAI.00358-07

204 Urban, J. F., Jr., Maliszewski, C. R., Madden, K. B., Katona, I. M. & Finkelman, F. D. IL-4 treatment can cure established gastrointestinal nematode infections in immunocompetent and immunodeficient mice. *J Immunol* **154**, 4675-4684 (1995).

205 King, I. L. & Mohrs, M. IL-4-producing CD4+ T cells in reactive lymph nodes during helminth infection are T follicular helper cells. *J Exp Med* **206**, 1001-1007 (2009). https://doi.org:10.1084/jem.20090313

206 King, I. L., Mohrs, K. & Mohrs, M. A nonredundant role for IL-21 receptor signaling in plasma cell differentiation and protective type 2 immunity against gastrointestinal helminth infection. *J Immunol* **185**, 6138-6145 (2010). https://doi.org:10.4049/jimmunol.1001703

207 Frohlich, A. *et al.* IL-21 receptor signaling is integral to the development of Th2 effector responses in vivo. *Blood* **109**, 2023-2031 (2007). https://doi.org:10.1182/blood-2006-05-021600

208 Liu, Q. *et al.* B cells have distinct roles in host protection against different nematode parasites. *J Immunol* **184**, 5213-5223 (2010). https://doi.org:10.4049/jimmunol.0902879

209 Wojciechowski, W. *et al.* Cytokine-producing effector B cells regulate type 2 immunity to H. polygyrus. *Immunity* **30**, 421-433 (2009). https://doi.org:10.1016/j.immuni.2009.01.006

210 McCoy, K. D. *et al.* Polyclonal and specific antibodies mediate protective immunity against enteric helminth infection. *Cell Host Microbe* **4**, 362-373 (2008). https://doi.org:10.1016/j.chom.2008.08.014

211 Williams, D. J. & Behnke, J. M. Host protective antibodies and serum immunoglobulin isotypes in mice chronically infected or repeatedly immunized with the nematode parasite Nematospiroides dubius. *Immunology* **48**, 37-47 (1983).

212 Wahid, F. N. & Behnke, J. M. Immunological relationships during primary infection with Heligmosomoides polygyrus (Nematospiroides dubius): parasite specific IgG1 antibody responses and primary response phenotype. *Parasite Immunol* **15**, 401-413 (1993). https://doi.org:10.1111/j.1365-3024.1993.tb00625.x

213 Anthony, R. M. *et al.* Memory T(H)2 cells induce alternatively activated macrophages to mediate protection against nematode parasites. *Nat Med* **12**, 955-960 (2006). https://doi.org:10.1038/nm1451

214 Chang, N. C. *et al.* A macrophage protein, Ym1, transiently expressed during inflammation is a novel mammalian lectin. *J Biol Chem* **276**, 17497-17506 (2001). https://doi.org:10.1074/jbc.M010417200

215 Behnke, J. M. *et al.* Immunological relationships during primary infection with Heligmosomoides polygyrus (Nematospiroides dubius): downregulation of specific cytokine secretion (IL-9 and IL-10) correlates with poor mastocytosis and chronic survival of adult worms. *Parasite Immunol* **15**, 415-421 (1993). https://doi.org:10.1111/j.1365-3024.1993.tb00626.x

216 Hepworth, M. R. *et al.* Mast cells orchestrate type 2 immunity to helminths through regulation of tissue-derived cytokines. *Proc Natl Acad Sci U S A* **109**, 6644-6649 (2012). https://doi.org:10.1073/pnas.1112268109

217 Shea-Donohue, T. *et al.* The role of IL-4 in Heligmosomoides polygyrus-induced alterations in murine intestinal epithelial cell function. *J Immunol* **167**, 2234-2239 (2001). https://doi.org:10.4049/jimmunol.167.4.2234

218 Kim, J. J. & Khan, W. I. Goblet cells and mucins: role in innate defense in enteric infections. *Pathogens* **2**, 55-70 (2013). https://doi.org:10.3390/pathogens2010055

219 Hashimoto, K. *et al.* Depleted intestinal goblet cells and severe pathological changes in SCID mice infected with Heligmosomoides polygyrus. *Parasite Immunol* **31**, 457-465 (2009). https://doi.org:10.1111/j.1365-3024.2009.01123.x

220 Herbert, D. R. *et al.* Intestinal epithelial cell secretion of RELM-beta protects against gastrointestinal worm infection. *J Exp Med* **206**, 2947-2957 (2009). https://doi.org:10.1084/jem.20091268

221 Wilson, M. S. *et al.* Helminth-induced CD19+CD23hi B cells modulate experimental allergic and autoimmune inflammation. *Eur J Immunol* **40**, 1682-1696 (2010). https://doi.org:10.1002/eji.200939721

222 Maizels, R. M. *et al.* Immune modulation and modulators in Heligmosomoides polygyrus infection. *Exp Parasitol* **132**, 76-89 (2012). https://doi.org:10.1016/j.exppara.2011.08.011

223 Blum, A. M. *et al.* Heligmosomoides polygyrus bakeri induces tolerogenic dendritic cells that block colitis and prevent antigen-specific gut T cell responses. *J Immunol* **189**, 2512-2520 (2012). https://doi.org:10.4049/jimmunol.1102892

224 White, M. P. J., McManus, C. M. & Maizels, R. M. Regulatory T-cells in helminth infection: induction, function and therapeutic potential. *Immunology* **160**, 248-260 (2020). https://doi.org:10.1111/imm.13190

225 Finney, C. A., Taylor, M. D., Wilson, M. S. & Maizels, R. M. Expansion and activation of CD4(+)CD25(+) regulatory T cells in Heligmosomoides polygyrus infection. *Eur J Immunol* **37**, 1874-1886 (2007). https://doi.org:10.1002/eji.200636751

226 Rausch, S. *et al.* Establishment of nematode infection despite increased Th2 responses and immunopathology after selective depletion of Foxp3+ cells. *Eur J Immunol* **39**, 3066-3077 (2009). https://doi.org:10.1002/eji.200939644

227 Hewitson, J. P., Grainger, J. R. & Maizels, R. M. Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Mol Biochem Parasitol* **167**, 1-11 (2009). https://doi.org:10.1016/j.molbiopara.2009.04.008

228 Grainger, J. R. *et al.* Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway. *J Exp Med* **207**, 2331-2341 (2010). https://doi.org:10.1084/jem.20101074

229 Tesseur, I., Zou, K., Berber, E., Zhang, H. & Wyss-Coray, T. Highly sensitive and specific bioassay for measuring bioactive TGF-beta. *BMC Cell Biol* **7**, 15 (2006). https://doi.org:10.1186/1471-2121-7-15

230 Johnston, C. J. C. *et al.* A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. *Nat Commun* **8**, 1741 (2017). https://doi.org:10.1038/s41467-017-01886-6

231 Janssens, K., ten Dijke, P., Ralston, S. H., Bergmann, C. & Van Hul, W. Transforming growth factor-beta 1 mutations in Camurati-Engelmann disease lead to increased signaling by altering either activation or secretion of the mutant protein. *J Biol Chem* **278**, 7718-7724 (2003). https://doi.org:10.1074/jbc.M208857200

232 Hara, H. *et al.* Activation of TGF-beta signaling in an aortic aneurysm in a patient with Loeys-Dietz syndrome caused by a novel loss-of-function variant of TGFBR1. *Hum Genome Var* **6**, 6 (2019). https://doi.org:10.1038/s41439-019-0038-x

233 Harradine, K. A. & Akhurst, R. J. Mutations of TGFbeta signaling molecules in human disease. *Ann Med* **38**, 403-414 (2006). https://doi.org:10.1080/07853890600919911

234 Biernacka, A., Dobaczewski, M. & Frangogiannis, N. G. TGF-beta signaling in fibrosis. *Growth Factors* **29**, 196-202 (2011). https://doi.org:10.3109/08977194.2011.595714

235 Brosens, L. A., Langeveld, D., van Hattem, W. A., Giardiello, F. M. & Offerhaus, G. J. Juvenile polyposis syndrome. *World J Gastroenterol* **17**, 4839-4844 (2011). https://doi.org:10.3748/wjg.v17.i44.4839

236 Kim, K. K., Sheppard, D. & Chapman, H. A. TGF-beta1 Signaling and Tissue Fibrosis. *Cold Spring Harb Perspect Biol* **10** (2018). https://doi.org:10.1101/cshperspect.a022293

237 Massague, J. TGFbeta in Cancer. *Cell* **134**, 215-230 (2008). https://doi.org:10.1016/j.cell.2008.07.001

238 Seoane, J. & Gomis, R. R. TGF-beta Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022277

239 Ge, R. *et al.* Selective inhibitors of type I receptor kinase block cellular transforming growth factor-beta signaling. *Biochem Pharmacol* **68**, 41-50 (2004). https://doi.org:10.1016/j.bcp.2004.03.011

240 Singh, J. *et al.* Transforming the TGFbeta pathway: convergence of distinct lead generation strategies on a novel kinase pharmacophore for TbetaRI (ALK5). *Curr Opin Drug Discov Devel* **7**, 437-445 (2004).

241 Akhurst, R. J. Targeting TGF-beta Signaling for Therapeutic Gain. *Cold Spring Harb Perspect Biol* **9** (2017). https://doi.org:10.1101/cshperspect.a022301

242 Smyth, D. J. *et al.* TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus. *Int J Parasitol* **48**, 379-385 (2018). https://doi.org:10.1016/j.ijpara.2017.12.004

243 Walker, R. G. *et al.* Structural basis for potency differences between GDF8 and GDF11. *BMC Biol* **15**, 19 (2017). https://doi.org:10.1186/s12915-017-0350-1

244 Huang, T. & Hinck, A. P. Production, Isolation, and Structural Analysis of Ligands and Receptors of the TGF-beta Superfamily. *Methods Mol Biol* **1344**, 63-92 (2016). https://doi.org:10.1007/978-1-4939-2966-5_4

245 Cull, M. G. & Schatz, P. J. Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol* **326**, 430-440 (2000). https://doi.org:10.1016/s0076-6879(00)26068-0

246 Keller, S. *et al.* High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem* **84**, 5066-5073 (2012). https://doi.org:10.1021/ac3007522

247 Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT--a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137-148 (2015). https://doi.org:10.1016/j.ymeth.2014.11.012

248 Brautigam, C. A., Zhao, H., Vargas, C., Keller, S. & Schuck, P. Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc* **11**, 882-894 (2016). https://doi.org:10.1038/nprot.2016.044

249 Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol* **562**, 109-133 (2015). https://doi.org:10.1016/bs.mie.2015.05.001

250 Grzesiek, S. & Bax, A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *Journal of the American Chemical Society* **115**, 12593-12594 (1993).

251 Pitto, M., Saudek, V., Sklenar, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* **2**, 661-665 (1992).

252 Farrow, N. A., Zhang, O. W., Formankay, J. D. & Kay, L. E. A Heteronuclear Correlation Experiment for Simultaneous Determination of N-15 Longitudinal Decay and Chemical-Exchange Rates of Systems in Slow Equilibrium. *Journal of Biomolecular NMR* **4**, 727-734 (1994).

253 Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

254 Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015). https://doi.org:10.1093/bioinformatics/btu830

255 Wider, G., Neri, D. & Wüthrich, K. Studies of slow conformational equilibria in macromolecules by exchange of heteronuclear longitudinal 2-spin-order in a 2D difference correlation experiment. *J Biomol NMR* **1**, 93-98 (1991). https://doi.org:https://doi.org/10.1007/BF01874572

256 Huang, T. *et al.* TGF-beta signalling is mediated by two autonomously functioning TbetaRI:TbetaRII pairs. *EMBO J* **30**, 1263-1276 (2011). https://doi.org:10.1038/emboj.2011.54

257 Mukundan, A. *et al.* Convergent evolution of a parasite-encoded complement control protein-scaffold to mimic binding of mammalian TGF-beta to its receptors, TbetaRI and TbetaRII. *J Biol Chem* **298**, 101994 (2022). https://doi.org:10.1016/j.jbc.2022.101994

258 Kim, S. K. *et al.* An engineered transforming growth factor beta (TGF-beta) monomer that functions as a dominant negative to block TGF-beta signaling. *J Biol Chem* **292**, 7173-7188 (2017). https://doi.org:10.1074/jbc.M116.768754

259 Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005). https://doi.org:10.1002/prot.20449

260 Bahrami, A., Assadi, A. H., Markley, J. L. & Eghbalnia, H. R. Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy. *PLoS Comput Biol* **5**, e1000307 (2009). https://doi.org:10.1371/journal.pcbi.1000307

261 Eghbalnia, H. R., Wang, L., Bahrami, A., Assadi, A. & Markley, J. L. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J Biomol NMR* **32**, 71-81 (2005). https://doi.org:10.1007/s10858-005-5705-1

262 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004). https://doi.org:10.1002/jcc.20084

263 Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65-73 (2003). https://doi.org:10.1016/s1090-7807(02)00014-9

264 Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009). https://doi.org:10.1007/s10858-009-9333-z

265 Skinner, S. P. *et al.* Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallogr D Biol Crystallogr* **71**, 154-161 (2015). https://doi.org:10.1107/S1399004714026662

266 Hansen, M. R., Hanson, P. & Pardi, A. Pf1 filamentous phage as an alignment tool for generating local and global structural information in nucleic acids. *J Biomol Struct Dyn* **17 Suppl 1**, 365-369 (2000). https://doi.org:10.1080/07391102.2000.10506642

267 Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* **131**, 373-378 (1998). https://doi.org:10.1006/jmre.1998.1361

268 Bax, A. *et al.* Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol* **239**, 79-105 (1994). https://doi.org:10.1016/s0076-6879(94)39004-5

269 Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486 (1996). https://doi.org:10.1007/BF00228148

270 Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364 (1992). https://doi.org:10.1002/prot.340120407

271 Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989). https://doi.org:10.1021/bi00449a003

272 Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* **73**, 1-16 (2013). https://doi.org:10.1016/j.pnmrs.2013.02.001

273 Deep, S., Walker, K. P., 3rd, Shu, Z. & Hinck, A. P. Solution structure and backbone dynamics of the TGFbeta type II receptor extracellular domain. *Biochemistry* **42**, 10126-10139 (2003). https://doi.org:10.1021/bi034366a

274 Soares, D. C. & Barlow, P. N. in *Structural Biology of the Complement System* Vol. 1 (eds D. Morikis & J. D. Lambris) Ch. 2, 21-62 (CRC Press, Taylor & Francis Group, 2005).

275 Kirkitadze, M. D. & Barlow, P. N. Structure and flexibility of the multiple domain proteins that regulate complement activation. *Immunol Rev* **180**, 146-161 (2001). https://doi.org:10.1034/j.1600-065x.2001.1800113.x

276 Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res* **50**, W210-215 (2022). https://doi.org:10.1093/nar/gkac387

277 Holm, L. Using Dali for Protein Structure Comparison. *Methods Mol Biol* **2112**, 29-42 (2020). https://doi.org:10.1007/978-1-0716-0270-6_3

278 Holm, L. DALI and the persistence of protein shape. *Protein Sci* **29**, 128-140 (2020). https://doi.org:10.1002/pro.3749

279 Sievers, F. & Higgins, D. G. The Clustal Omega Multiple Alignment Package. *Methods Mol Biol* **2231**, 3-16 (2021). https://doi.org:10.1007/978-1-0716-1036-7_1

280 Schmitz, C., Stanton-Cook, M. J., Su, X. C., Otting, G. & Huber, T. Numbat: an interactive software tool for fitting Deltachi-tensors to molecular coordinates using pseudocontact shifts. *J Biomol NMR* **41**, 179-189 (2008). https://doi.org:10.1007/s10858-008-9249-z

281 Gronenborn, A. M. & Clore, G. M. Investigations into the solution structures of short nucleic acid fragments by means of nuclear Overhauser enhancement measurements. *Prog Nucl Magn Reson Spectrosc* **17** (1985).

282 Boesen, C. C., Radaev, S., Motyka, S. A., Patamawenu, A. & Sun, P. D. The 1.1 A crystal structure of human TGF-beta type II receptor ligand binding domain. *Structure* **10**, 913-919 (2002). https://doi.org:10.1016/s0969-2126(02)00780-3

283 Pintacuda, G., John, M., Su, X. C. & Otting, G. NMR structure determination of protein-ligand complexes by lanthanide labeling. *Acc Chem Res* **40**, 206-212 (2007). https://doi.org:10.1021/ar050087z

284 Muntener, T., Joss, D., Haussinger, D. & Hiller, S. Pseudocontact Shifts in Biomolecular NMR Spectroscopy. *Chem Rev* **122**, 9422-9467 (2022). https://doi.org:10.1021/acs.chemrev.1c00796

285 Purslow, J. A., Khatiwada, B., Bayro, M. J. & Venditti, V. NMR Methods for Structural Characterization of Protein-Protein Complexes. *Front Mol Biosci* **7**, 9 (2020). https://doi.org:10.3389/fmolb.2020.00009

286 Clore, G. M. Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear overhauser enhancement data and dipolar couplings by rigid body minimization. *Proc Natl Acad Sci U S A* **97**, 9021-9025 (2000). https://doi.org:10.1073/pnas.97.16.9021

287 Gradwell, M. J. & Feeney, J. Validation of the use of intermolecular NOE constraints for obtaining docked structures of protein-ligand complexes. *J Biomol NMR* **7**, 48-58 (1996). https://doi.org:10.1007/BF00190456

288 Tang, C. & Clore, G. M. A simple and reliable approach to docking protein-protein complexes from very sparse NOE-derived intermolecular distance restraints. *J Biomol NMR* **36**, 37-44 (2006). https://doi.org:10.1007/s10858-006-9065-2

289 Nitz, M. *et al.* Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide. *Angew Chem Int Ed Engl* **43**, 3682-3685 (2004). https://doi.org:10.1002/anie.200460028

290 Bertini, I., Luchinat, C. & Parigi, G. Magnetic susceptibility in paramagnetic NMR. *Prog NMR Spectrosc* **40**, 249-273 (2002). https://doi.org:https://doi.org/10.1016/S0079-6565(02)00002-X

291 Stanton-Cook, M. J., Su, X. C. & Otting, G.

292 Dominguez, C., Boelens, R. & Bonvin, A. M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**, 1731-1737 (2003). https://doi.org:10.1021/ja026939x

293 De Martino, M. *et al.* Activin A Promotes Regulatory T-cell-Mediated Immunosuppression in Irradiated Breast Cancer. *Cancer Immunol Res* **9**, 89-102 (2021). https://doi.org:10.1158/2326-6066.CIR-19-0305

294 De Martino, M., Daviaud, C. & Vanpouille-Box, C. Activin A backs-up TGF-ss to promote regulatory T cells. *Oncoimmunology* **10**, 1883288 (2021). https://doi.org:10.1080/2162402X.2021.1883288

295 Nono, J. K., Lutz, M. B. & Brehm, K. Expansion of Host Regulatory T Cells by Secreted Products of the Tapeworm Echinococcus multilocularis. *Front Immunol* **11**, 798 (2020). https://doi.org:10.3389/fimmu.2020.00798

296 Anido, J. *et al.* TGF-beta Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating Cell Population in Human Glioblastoma. *Cancer Cell* **18**, 655-668 (2010). https://doi.org:10.1016/j.ccr.2010.10.023