The *H. polygyrus* TGF-β Mimic TGM6: A Competitive Inhibitor of TGF-β Signaling

by

Stephen Evans White

B.S., North Carolina State University, 2012

M.S., North Carolina State University, 2014

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

2022

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Stephen Evans White

It was defended on

June 13, 2022

and approved by

Dr. Rieko Ishima, Associate Professor, Department of Structural Biology

Dr. Gordon Rule, Professor, Department of Biological Sciences

Dr. Dario Vignali, Vice Chair and Professor of Immunology, Department of Immunology

Dissertation Director: Dr. Andrew Hinck, Professor, Department of Structural Biology

The *H. polygyrus* TGF-β Mimic TGM6: A Competitive Inhibitor of TGF-β Signaling

Stephen Evans White, PhD

University of Pittsburgh, 2022

The murine helminth parasite *Heligmosomoides polygyrus* expresses a family of structurally novel mimics of TGF-β, first identified through TGF-β Mimic 1 (TGM1), a 5-domain protein which binds and signals through the TGF-β receptors, TβRI and TβRII. Through recent studies, domains 1 and 2 and domain 3 have been shown to bind TβRI and TβRII, respectively. The function of domains 4 and 5 remain undefined. TGM6 is a homologue of TGM1 in which domains 1 and 2 are absent. In this work, biophysical methods and cellular assays were used to characterize its specificity, structure, and function. In accord with TGM1, TGM6 is shown to bind TβRII through domain 3. However, TGM6 does not bind the TGF-β family type I receptors Alk1, Alk2, Alk3, Alk4, or Alk5 or type II receptors ActRII, ActRIIb, or BMPRII. In cellular TGF-β reporter assays, TGM6 is shown to act as a highly potent TGF-β and TGM-1 signaling antagonist, consistent with its ability to bind TβRII, but not other receptors of the TGF-β family previously stated. TGM6, however, inhibits TGF-β signaling in mouse fibroblast cell lines and only does so when domains 4 and 5 are present, suggesting that domains 4, 5, or 4 and 5 are bound by a co-receptor that potentiates the inhibitory activity of TGM6. To gain insights as to how TGM6 specifically binds TβRII, the X-ray crystal structure of the TGM6-D3 bound to TβRII was determined to 1.45 Å. This shows that TGM6-D3 binds TβRII through an interface that is remarkably similar to that of TGF-β:TβRII. These results, together with the finding that TGM6 binds TβRII with high affinity, suggest that TGM6 has specifically adapted its domain structure

and sequence to function as a potent TGF-β and TGM antagonist, possibly to blunt fibrotic damage caused by the parasite as it progresses through its life cycle.

Keywords: TGF-β, *H. polygyrus*, parasite, Hp-TGM, TGM6, NMR, X-ray crystallography, biophysics, structural biology, protein structure, isothermal titration calorimetry, ITC, surface plasmon resonance, SPR, reporter assay

Table of Contents

List of Tables

List of Figures

Acknowledgements

The writing within this dissertation is the culmination of several years of trials, tribulations, failures, and successes. These trials include numerous successful replications of the 1974 paper by Dr. Dennis Upper titled “The unsuccessful self-treatment of a case of ‘writer’s block’” [257]. However, after repeatedly banging my head against a wall, I have arrived at this final accumulation of knowledge that I am calling my doctoral dissertation.

I did not get to this point on my own, however. There were many people who helped me get here and to whom I owe a large debt of gratitude. First, I wish to thank my advisor, Dr. Andrew Hinck, who took me into his lab as a fourth-year student following a few unreconcilable differences between myself and my initial PI. Without Andy’s leadership, guidance, constant pushing, and generosity, I would not have made it through this degree.

I would like to thank my dissertation committee: Dr. Rieko Ishima, Dr. Gordon Rule, and Dr. Dario Vignali. Their experience and guidance helped me avoid several major pitfalls that could have derailed my degree entirely.

Throughout my time in graduate school, I have had a chance to collaborate with many people on projects. I specifically wish to thank Dr. Damodaran Krishnan Achary for his training, support, and eagerness to try new NMR techniques. Secondly, I wish to thank Drs. Rick Maizels and Gareth Inman, without whom we would not have this project or the insights into the biology of these parasite proteins.

When I initially started my degree, I joined the Asher group. Although I did not finish my degree in that group, I am still thankful to the members of that group for their guidance, assistance, and friendship during my time in that group. After joining the Hinck lab, the members of the lab

helped me get up to speed quickly so that I could complete my degree within a timely manner. Specifically, I wish to thank Dr. Cindy Hinck, Dr. Łukasz Wieteska, Chang Byeon, Ananya Mukundan, and Tristin Schwartze. Without their support, I would not have been able to complete my degree.

I wish to also thank my fellow MBSB students and friends who helped keep me sane through this entire process. Specifically, I would like to thank Sara Whitlock, Kirill Lavrenyuk, and Manish Aryal for helping me get through this process and encouraging me to drink in times of celebration. I also wish to also thank Chino Cabalteja and Brenda Ho for being wonderful friends and confidants in both good times and bad.

Finally, I wish to thank my family for their support through this process. I particularly wish to express an immeasurable amount of gratitude to my wife, Justine. Justine was the person who encouraged me the most to follow my dreams and pursue this degree. She witnessed me go through every step of this process, from starting out and wanting to quit to the stressful times when my projects did not want to work to the celebrations after getting papers published. Without her constant encouragement and occasional verbal smack in the face, I would not be here today.

Dedication

I would like to dedicate this work to two people. First, to Ian Peterson, an undergrad friend who died from leukemia at the young age of 21. Second, and most importantly, to Audrey Burchell, my maternal grandmother, whom I watched slowly lose her mental faculties, memory, and ultimately her life to Alzheimer's disease. These two people are my inspiration for pursuing a career in biomedical sciences with a hope that I can work to find treatments, cures, and ultimately preventions for diseases so that others do not have to suffer as these two individuals and their families have. Disease is a war, and it is for those who have lost their battles that I continue to fight.

1.0 Introduction to TGF-β Family Signaling

1.1 The TGF-β Family of Signaling Proteins

The transforming growth factor beta, or TGF-β, family is an ancient family of secreted signaling proteins that has greatly diversified as multicellular eukaryotes have diversified [90]. In worms, there are three family ligands; in flies, there are seven; and in humans and other mammals, there are more than 30 [90]. The ancestors of the family are the bone morphogenetic proteins, or BMPs. The BMPs are found in both vertebrates and invertebrates [90]. In invertebrates, such as worms, they serve as developmental factors responsible for embryonic patterning and organ development. In vertebrates, BMPs also serve as essential developmental factors, but have expanded to take on additional specialized roles, such as regulating bone and muscle mass and vasculogenesis [90]. The TGF-βs themselves are evolutionary latecomers found only in vertebrates. They regulate functions unique to vertebrates, such as repairing wounds and regulating the adaptive immune system [90].

The TGF-β family of signaling proteins has 33 members in humans. This family includes the three isoforms of TGF-β, BMPs, growth and differentiation factors (GDFs), activins, inhibins, nodal, and anti-Müllerian hormone (AMH) [79, 90]. Each gene encodes for a polypeptide comprising of a signal peptide that signals for secretion of the protein, an approximately 250-residue prodomain, and an approximately 110-residue growth factor domain.

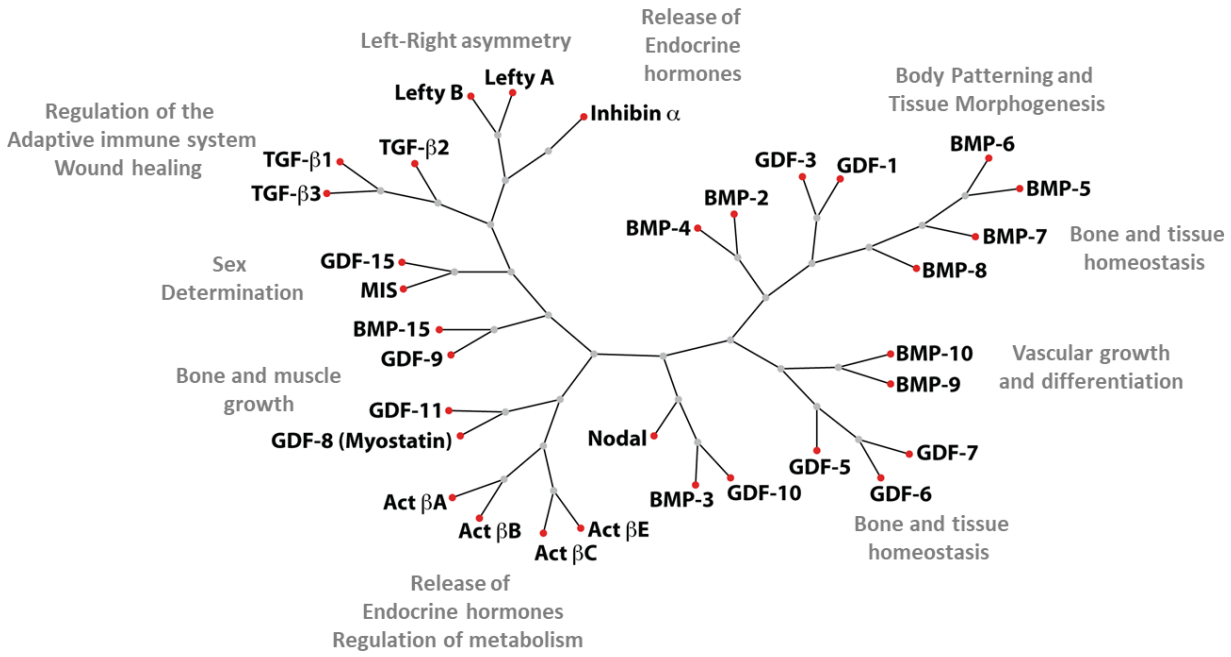

Figure 1. TGF-β family tree. The members of the TGF-β family are shown according to their sequence divergence as a consensus of functional and structural diversification. Additional notes are included regarding specific functions of various branches of the family tree. Figure adapted with permission from Hinck AP. FEBS Lett. 2012;586(14):1860-70. [88]. Copyright 2012 Federation of European Biochemical Societies.

In this chapter, the TGF-β family of signaling proteins, related diseases, the formation and activation of signaling complexes, and the context dependence of the signaling pathways will be briefly reviewed. As the proteins specifically studied in this dissertation interact with the TGF-β pathway, not those of other family members, specific attention will be paid to the related details for the TGF-β isoforms.

1.2 The TGF-β Family and Disease

1.2.1 Inheritable diseases

There is intense interest in deciphering the underlying basis for the unique functions of proteins of the TGF-β family owing to the many essential processes they regulate and the many human diseases that are a consequence of defects in the ligands, their receptors, or accessory proteins that regulate their activity. These range from relatively rare but fatal disorders such as Marfan's Disease and fibrodysplasia ossificans progressiva (FOP) [58, 142], which are caused by excessive TGF-β or BMP signaling that leads to either vascular fragility or bone formation at sites of soft tissue damage, respectively, to relatively more common and non-fatal diseases such as Ehlers-Danlos Syndrome [EDS], which results from a deficiency in the synthesis and processing of collagen [196].

1.2.2 Dysregulation of the signaling pathways

There are also several common diseases that are indirectly promoted by increased family signaling. Due to the pleiotropic nature of these proteins, their functions are tightly regulated. Once these proteins become dysregulated, disease typically follows. The most well-known are a variety of sporadic cancers [238] including those of the breast, brain, lung, and prostate, that are promoted by overexpression of the TGF-βs. TGF-βs are well-known to function as potent growth inhibitors of epithelial and endothelial cells [286], yet most advanced cancer cells become refractory to TGF-

β mediated growth inhibition [229], and the TGF-βs promote tumor progression and metastasis by suppressing host immune surveillance [58, 108] and stimulating epithelial-mesenchymal transitions (EMT) and angiogenesis [58, 108]. TGF-βs also coordinately upregulate the expression of many matrix proteins, such as collagen and fibronectin, and have been shown to play a causative role in the progression of many fibrotic disorders, including kidney, lung, and pulmonary fibrosis [118].

1.2.3 Treating TGF-β-related diseases

TGF-β's disease-promoting activities have made it an important molecular target for treatment of cancer and fibrosis, yet no inhibitors have been FDA-approved [2, 250]. Generally, TGF-β signaling promotes the progression of soft tissue cancers and fibrotic disorders [12, 118, 238].

Although promotion of cancer progression may seem at odds with its tumor suppressor activity, cancer cells dysregulate their cell cycle. This, in turn, antagonizes TGF-β's growth inhibitory activity [12, 30]. Many of the tumor-promoting activities of the TGF-βs, such as their ability to potently suppress immune responses [229] and induce EMTs [108], remain intact.

Using TGF-β inhibitors as therapeutics is one method by which these effects could ultimately be stopped, slowing the growth of cancers within a human body. There are a few recent studies that show that patients who have tumors that are refractory to checkpoint therapies, such as anti-PD-1, anti-PD-L1, and anti-CTLA4, have tumors that are devoid of T-cells [156, 250]. Animal models of these types of cancers have shown that utilizing TGF-β inhibitors is beneficial: TGF-β inhibitors in combination with checkpoint therapy have proved superior to checkpoint therapy alone [53, 54].

However, the pleiotropic nature means that TGF-β can be very difficult to safely target specifically for the treatment of disease. Generally inhibiting TGF-βs throughout the body is thought to cause numerous off-target effects. While providing TGF-β inhibitors to cancer patients, for example, could be beneficial, there are major concerns about the safety of these compounds, specifically around the antagonism of TGF-β2 in the heart [58]. TGF-β2 in the heart is required for the maintenance of the cardiac tissues [58], and inhibiting TGF-β2 in the heart evidently results in cardiotoxicity. Despite the concerns about cardiotoxicity, there are studies that show that TGF-β antagonists are generally safe [281]. Specifically, mice who were exposed to a soluble TGF-β signaling antagonist throughout their lives had no detectable adverse effects due to the treatment [281].

Conversely, using TGF-β signaling agonists as treatments of diseases, such as autoimmune disease, comes with its own concerns. By broadly agonizing TGF-β signaling to induce a down-regulation of the immune system [229], there is an increased risk of fibrotic response throughout the body. The fibrotic disorders driven by TGF-βs include idiopathic pulmonary fibrosis (IPF), renal fibrosis, and cardiac fibrosis; all these forms of fibrosis can be lethal [118]. While the mechanisms triggering fibrosis are diverse, they share a common induction of TGF-β signaling. This induction of TGF-β signaling stimulates the activation and differentiation of fibroblasts into myofibroblasts, which cause aberrant deposition of extracellular matrix (ECM), which in turn leads to scarring and reduced organ function [240].

As targeting TGF-β signaling with either agonists or antagonists could be of great clinical significance, finding ways of specifically targeting compounds to the desired sites that either induce or inhibit TGF-β signaling with minimal off-target effects is strongly desired.

1.3 The TGF-β Signaling Cascade

1.3.1 Structure of the TGF-β signaling protein

The signaling proteins in the TGF-β family have a distinctive fold referred to as the cystine knot growth factor (CKGF) fold. While other proteins contain cystine knots, the presence of the knot does not define the CKGF fold. The fold is also defined by the topological relationship between the β-strand order in the β-sheets and the general amino acid sequence. The overall structure of the CKGF is shown in Figure 2.

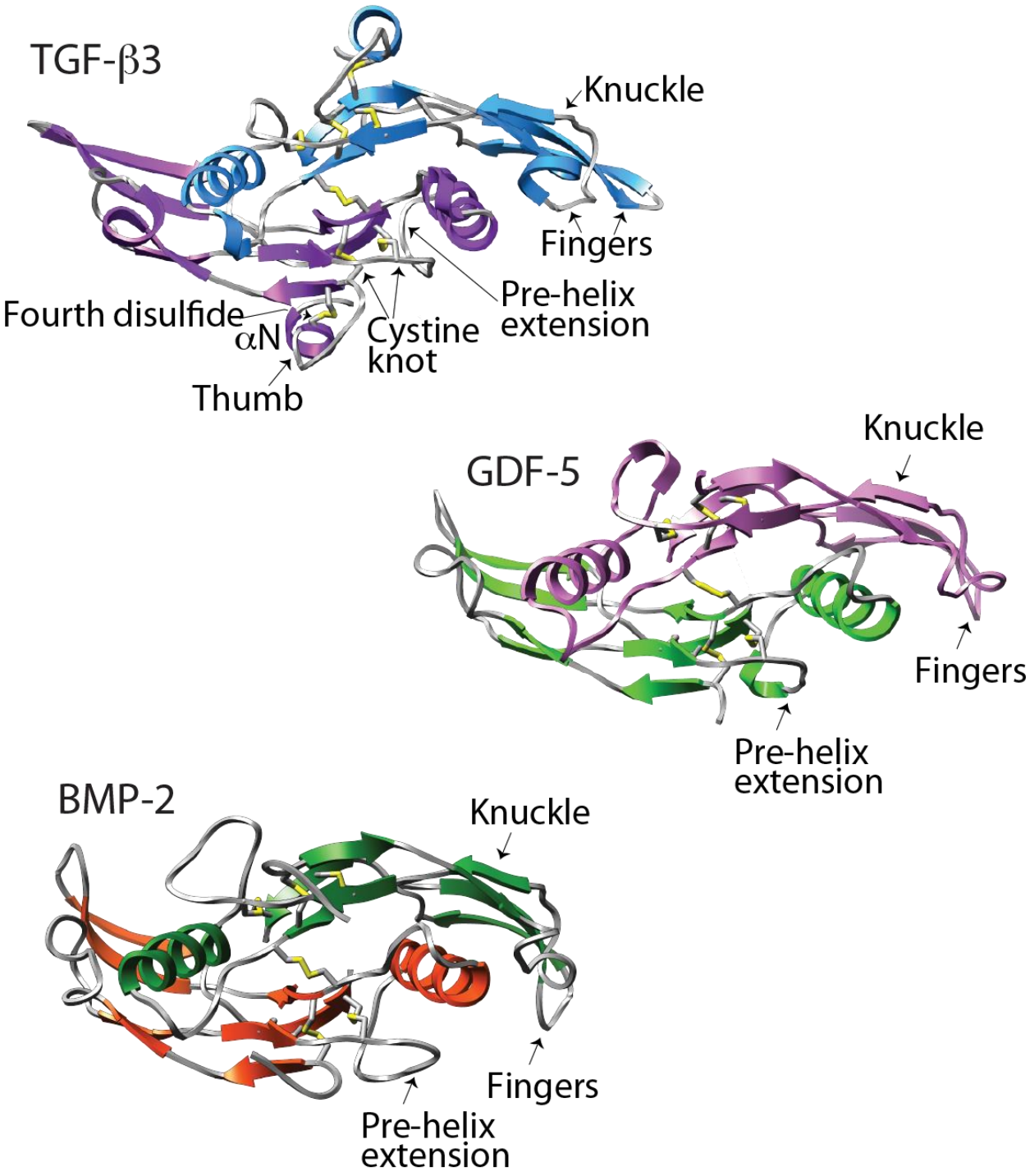

Figure 2. Representative structures of the TGF-β family members.
The structures shown are those of TGF-β3 (PDB 1TGK) [172], GDF-5 (PDB 1WAQ) [187], and BMP-2 (PDB 3BMP) [233]. These structures are representative of the varying structures within the TGF-β family and generally of the CKGF fold. Figure adapted with permission from Hinck AP, Mueller TD, Springer TA. Structural Biology and Evolution of the TGF-beta Family. Cold Spring Harb Perspect Biol. 2016;8(12) [90].

The CKGF contains two long β-ribbons (two anti-parallel β-strands forming a single β-sheet) and a cystine knot. Two closely spaced pairs of cysteines in adjacent and parallel β-strands are disulfide-bonded to form a ring composed of two peptide backbone segments and two disulfides. A third disulfide bond passes through the ring, linking two additional polypeptide segments. Thus, the knot ties together four polypeptide segments that are distal in sequence. Furthermore, this forms a highly stable cystine knot core from which three long loops emanate [86, 90].

All members of the TGF-β family of signaling proteins dimerize, either with themselves forming homodimers (e.g., TGF-β1) or with other family members forming heterodimers (e.g., BMP-2/5 and Inhibin A) [90, 97]. Most of the monomers utilize an interchain disulfide bond to stabilize the dimer, but this is not always the case [90]. Each monomer adopts the shape of a curled left hand, specifically, with regions of the protein referred to as the heel, palm, fingers, and thumb (see Figure 2). The two monomers are joined together with the heel of one packing into the palm of the other. The cystine knot forms the base of the palm. The heel is formed by a three- to four-turn α-helix on one side of the cystine knot. The fingers are formed by two loops that extend from the cystine knot and adopt a β-ribbon conformation in most family members [86, 90]. For the TGF-β isoforms, each β-strand in the β-ribbons is counted as a finger; therefore, the TGF-βs are described as having four fingers. In other family members, each β-ribbon is counted as a finger [86, 90].

1.3.2 The TGF-β signaling complex

As stated previously, the TGF-β family of signaling proteins are extracellular signaling proteins. They must, therefore, transmit signals into the cell to induce function. To do so, each

member of the TGF-β family of signaling ligands ligates a type I and a type II receptor for each monomer in the dimeric signaling protein [94]. Therefore, the full signaling complex is made up of the disulfide-linked dimeric signaling protein, two type I receptors, and two type II receptors, as illustrated in Figure 3.

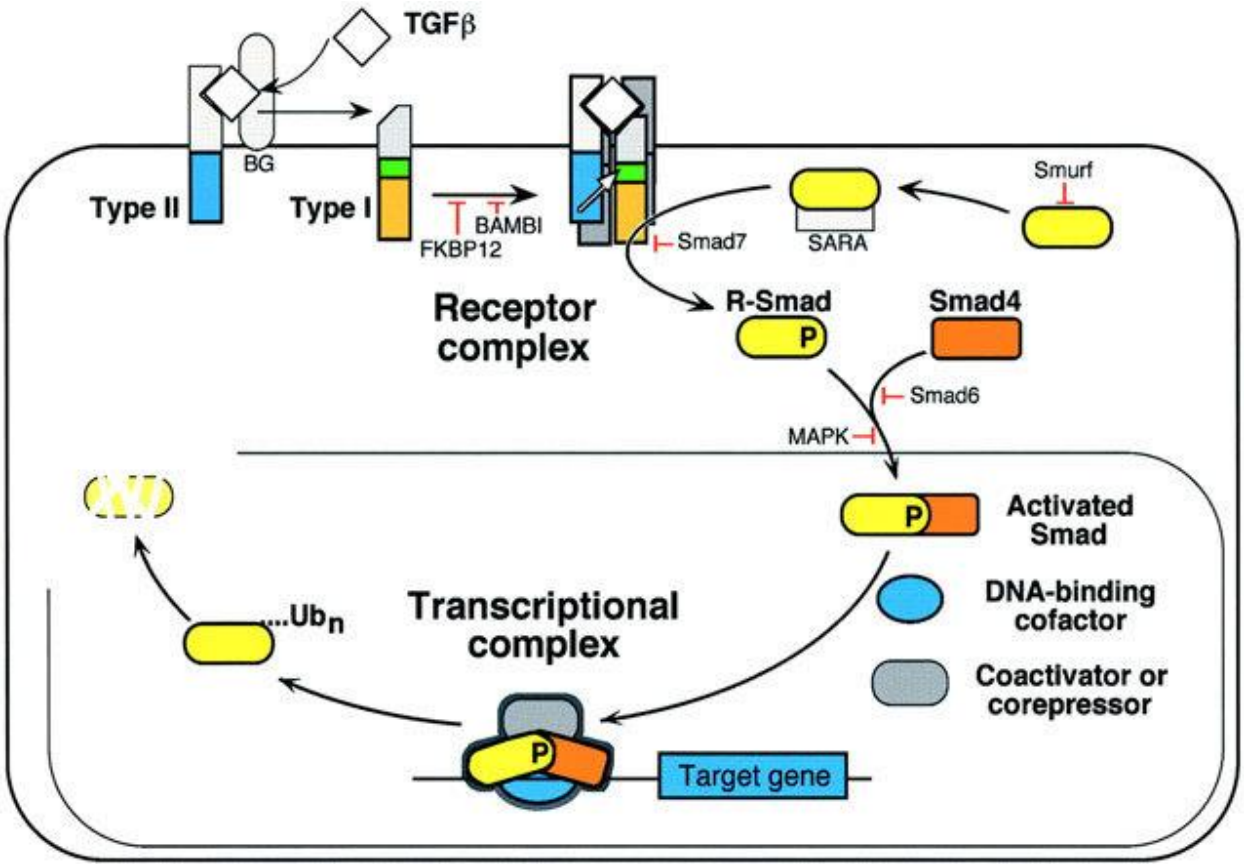

Figure 3. Schematic of the structure and immediate signaling pathway of the 1:2:2 Growth Factor (GF):Type I Receptor:Type II Receptor signaling complex. This system involves a family of membrane receptor protein kinases, coreceptors (i.e. Betaglycan, BG), and a family of receptor substrates (the Smad proteins) that march into the nucleus where they act as transcription factors. The ligand TGF-β assembles a receptor complex that activates Smads, and the Smads assemble multisubunit complexes that regulate transcription. Two general steps suffice to carry the hormonal stimulus to target genes. The central components of this signaling system are indicated along with the sites of action of various positive and negative regulators. Figure used with permission from Massague J, Wotton D. Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. 2000;19(8):1745-54. [160]. Copyright 2000 European Molecular Biology Organization.

There are three major sets of TGF-β family of receptors: the type I receptors, the type II receptors, and the type III receptors. These were initially named as such due to their molecular

weights, with the type I receptors being the smallest and the type III receptors being the largest [28, 162]. Generally, the type I and type II receptors each contain an N-terminal extracellular domain that is responsible for binding the signaling protein, a single-pass transmembrane helix, and a C-terminal intracellular serine/threonine kinase domain [90]. While the intracellular domains of the type I and type II receptors have structural characteristics similar to tyrosine kinases, their main substrates are serine and threonine residues [127].

The type III receptors are coreceptors that have multiple functions. Most commonly, they aid in the presentation of the signaling protein to the type I and type II receptors. For example, TβRIII, has a large a large extracellular domain (ca. 760 amino acids), a single-spanning transmembrane helix, and a short (ca. 40 amino acids) non-catalytic cytoplasmic tail. TβRIII binds all three TGF-β isoforms with high affinity (K_D = 5 – 20 nM) [169], but its primary function is to present TGF-β2 to TβRII. TGF-β2 binds TβRII about 200 times weaker than TGF-β1 and –β3 due to two conserved Arg to Lys substitutions [8, 29, 39, 210] and only signals at supraphysiological concentrations (>50 pM) in the absence of TβRIII [29, 39, 140]. In the presence of TβRIII, cellular responsiveness to TGF-β2 is comparable to that of TGF-β1 and –β3 [29, 47]. TβRIII has been proposed to potentiate assembly of the TGF-β2$(T\beta RII)_2(T\beta RI)_2$ signaling complex by a handoff mechanism [139, 258], with a key component being direct interactions between the ectodomains of TβRI and TβRII that serve not only to stabilize the weakly bound TβRII, but also to drive displacement of TβRIII.

1.3.3 TGF-β is unique in the family

Generally, the order in which the growth factor ligates its receptors is not specific. As long as both a type I and a type II receptor are bound to the signaling protein, a signal is propagated

into the cell. According to cell-based assays [24, 124, 130, 132, 189, 199, 278] and binding studies performed with protein extracellular domains [62, 78, 121, 186, 237], a large assortment of BMP receptors have mixed affinities for their ligands. As examples, ActRII (a type II receptor) has a moderate affinity for BMP-7 and interacts weakly with BMP-2, whereas Alk3 (a type I receptor) binds BMP-2 with high affinity and BMP-7 weakly [237]. Superposition of the BMP-2:Alk3 [122] and BMP-7:ActRII complex structures revealed that the extracellular domains of the two receptors neither interact nor induce significant conformational changes in the growth factors [62]; this was confirmed later by determination of the crystal structures of two BMP ternary complexes [3, 265].

The case for the TGF-β isoforms is different than that of the other TGF-β family members. TGF-β1 and –β3 bind TβRII as their type II receptor with high affinity (K_D = ca. 50 nM); as stated previously, TGF-β2 requires the presence of TβRIII to facilitate the high-affinity binding interaction between the signaling protein and TβRII. However, TGF-β's binding of its type I receptor, Alk5, also known as TβRI, is very weak (K_D = ca. 70 μM) and is hardly detectable by surface plasmon resonance (SPR). However, the affinity of Alk5 to the TGF-β:TβRII binary complex is over 1000-fold higher (K_D = ca. 30 nM) than that of Alk5 for TGF-β in the absence of TβRII [40, 94, 211, 290]. The crystal structure of the TGF-β1:TβRII:Alk5 and TGF-β3:TβRII:Alk5 complexes showed that TβRII bound to the fingertips of TGF-β [64, 211], as compared to the other members of the family which bound their type II receptors to the knuckles [3, 62, 265]. This shift in the position of the type II receptor relative to the signaling ligand allows for the direct contact with TβRI bound to the underside of the fingers and heel helix [64]. Therefore, it was concluded that TGF-β was unique among the members of the family in its recruitment and ligation of its receptors. Unlike interactions between other TGF-β family members

binding their type I receptors, Alk5 binds to TGF-β only in the presence of a combined interface formed between TGF-β and TβRII.

.

1.3.4 Downstream signal transduction

As stated previously, the type I and type II receptors have an extracellular domain to which the growth factor binds, a single-pass transmembrane helix, and an intracellular serine/threonine kinase. The close spatial proximity of the type II and type I kinases in the context of the signaling complex triggers a phosphorylation cascade. In this cascade, the constitutively active type II kinase phosphorylates serine residues within a negative regulatory domain of the type I receptor kinase, leading to its activation [64, 77, 274]. The phospho-type I receptor propagates the signal through the Smad pathway by phosphorylating R-Smads [77]. These R-Smads then translate into the nucleus and weakly bind to Smad Binding Elements (SBEs) in the promoter regions of hundreds of genes, imparting their function to varying cells and tissues throughout the organism [77].

As discussed previously, the number of ligands in the TGF-β family greatly exceeds receptors (humans have 33 signaling proteins, 5 type II receptors, and 7 type I receptors) (Figure 4); however, defined subsets of ligands pair with individual receptors [88]. The receptors furthermore pair with distinct classes of R-Smads that bind to distinct promoter elements [27] and activate distinct sets of genes. As an example, TGF-β principally binds TβRI, but not other type I receptors, such as Alk1, Alk2, Alk3, and Alk6. These receptors (Alk1, Alk2, Alk3, and Alk6) are preferentially bound by BMPs and GDFs. Owing to this and to the strict coupling of TβRI and Alk1/Alk2/Alk3/Alk6 through their kinase domain L3 loops with Smads2/3 and Smads1/5/8,

respectively [134], and the binding of Smads2/3 and Smads1/5/8 to distinct promoter elements, this effectively segregates the signaling activity of the TGF-βs away from the BMPs and GDFs.

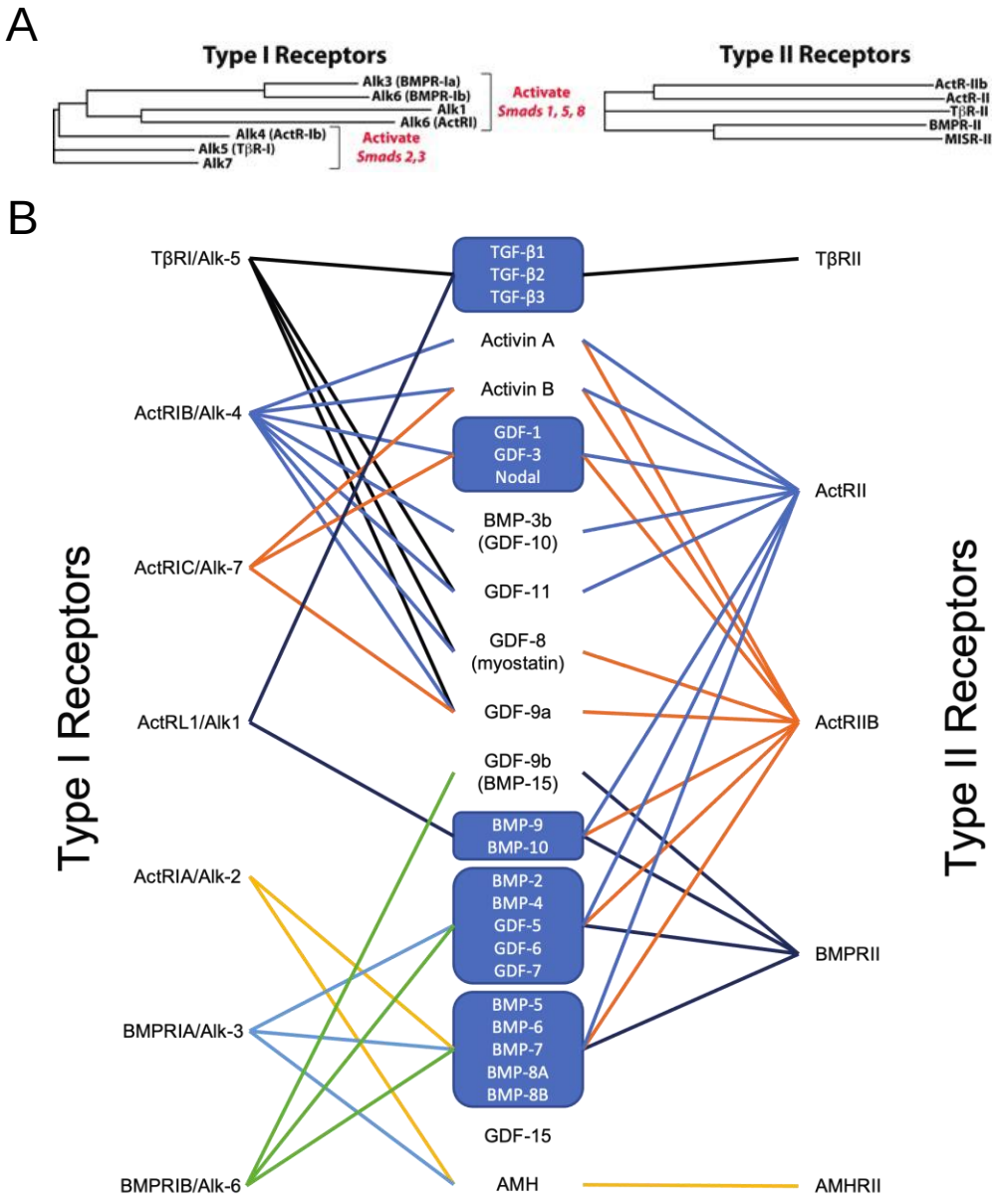

Figure 4. The TGF-β family of type I and type II receptors. (A) The family trees of the TGF-β family of type I and type II receptors. The receptor name is given in black while the Smads activated by the groupings of type I receptors are given in red. Figure adapted with permission from Hinck AP, Mueller TD, Springer TA. Structural Biology and Evolution of the TGF-beta Family. Cold Spring Harb Perspect Biol. 2016;8(12) [90]. Copyright 2016 Cold Spring Harbor Laboratory Press. (B) Schematic of TGF-β receptor binding. The illustration shows the selective binding of members of the TGF-β family to type I and type II receptors. Figure adapted with permission from Heldin CH, Moustakas A. Signaling Receptors for TGF-beta Family Members. Cold Spring Harb Perspect Biol. 2016;8(8). [79]. Copyright 2016 Cold Spring Harbor Laboratory Press.

The identity of the type I receptor segregates the actions of BMPs and GDFs, the ancestors of the family found in vertebrates and invertebrates, from the TGF-βs, which are evolutionary latecomers only found in vertebrates. With the differences in binding of TGF-β to its type I receptor, previously discussed, it can be concluded that the ligands and receptors of the family have coevolved alternative binding modes that segregate the actions of the TGF-βs away from the BMPs and GDFs [64, 162].

Generally, the identity of the type I receptor in the signaling complex determines if the complex signals through the Smad1/5/8 pathway, or the Smad2/3 pathway. The type I receptors Alk1, Alk2, Alk3, and Alk6 activate the Smad1/5/8 pathway, while Alk4, Alk5, and Alk7 activate the Smad2/3 pathway [77]. Notably, this segregation is not strict. Although not nearly as prominent as signaling through Alk5, TGF-βs will signal through Alk1 and activate the Smad1/5/8 branch in some cell lines [57, 77, 212].

Once phosphorylated, the R-Smads activated by the signaling complex form a heterotrimeric complex with Smad4 and are translocated into the nucleus where they interact with other co-regulators to activate or suppress gene expression [77]. This gives rise to the context-dependence of the signals transduced by the signaling complex. This context-dependence gives the growth factors and their receptors the ability to play many roles in cell and tissue physiology.

1.4 Context Dependence and Tissue Specific Functions

As stated previously, the phospho-type I receptor propagates the signal through the Smad pathway by phosphorylating R-Smads 2 and 3, or R-Smads 1, 5 and 8, which then translate into the nucleus and weakly bind to distinct SBEs in the promoter regions of hundreds of genes [77].

Owing to the moderate affinity of phosphorylated R-Smads for binding DNA, they are also dependent on other transcriptional co-activators and co-repressors to effect transcriptional responses [160]. Pronounced pleiotropy and context-dependent signaling that is characteristic of the TGF-βs and other proteins of the family arise from the variation of the co-activators and co-repressors from cell-to-cell [160, 163].

The functional diversity that can be attained through intrinsic differences in signaling is restricted in the TGF-β family since the ligands activate just two classes of transcriptional effectors, Smads 2 and 3 or Smads 1, 5, and 8. Most of the functional diversity instead arises from targeting of the ligands to different cells and tissues, leading to the pronounced context-dependent signaling for ligands of the family [272]. The targeting of ligands is mediated by more than 20 binding proteins which regulate access of the ligands to the signaling receptors [89]. The binding proteins are structurally diverse, ranging from membrane-bound co-receptors, such as betaglycan [120], to soluble matrix proteins, such as follistatin, chordin, noggin, and others [79].

TGF-β family members are highly pleiotropic, and their effects are highly context-dependent. They play essential roles in body patterning and organogenesis in developing embryos [90] and maintenance of the heart, vasculature, and musculoskeletal system in adults [176].

TGF-β itself has three isoforms and is one of the most recently evolved members of the family [90]. The TGF-β isoforms coordinate wound healing [218, 219], maintain the extracellular matrix, and regulate epithelial and endothelial cell growth and differentiation [60, 108]. They also suppress tumor growth by inhibiting the growth of epithelial cells [159, 177] and promote immune tolerance by upregulating suppressive regulatory T-cells (Tregs) [154, 229]. TGF-β isoform-specific null mice are inviable and have distinct phenotypes, indicating that each isoform fulfills a distinct role *in vivo* [106, 125, 205, 228, 242].

TGF-β acts as a potent immunosuppressive cytokine through effects on both differentiation and proliferation of T-cells. Specifically, TGF-β1 promotes the differentiation of naïve T-cells into $CD4^+$ $CD25^+$ regulatory T-cells, or Tregs, by inducing Foxp3 expression in $CD4^+$ T-cells [176]. These Tregs then suppress activation, proliferation, and cytokine production in $CD4^+$ and $CD8^+$ T-cells, inducing a downregulation of a host's immune response and increased immune tolerance.

As much research has been done on the various functions of the TGF-β family members, the following reviews may be of relevance: Activins and Inhibins: Roles in Development, Physiology, and Disease [185]; Bone Morphogenetic Proteins in Vascular Homeostasis and Disease [59]; Regulation of Hematopoiesis and Hematological Disease by TGF-β Family Signaling Molecules [184]; Regulation of the Immune Response by TGF-β: From Conception to Autoimmunity and Infection [229]; TGF-β and the TGF-β Family: Context-Dependent Roles in Cell and Tissue Physiology [176]; The TGF-β Family in the Reproductive Tract [173]; TGF-β Family Signaling in Connective Tissue and Skeletal Diseases [142]; TGF-β Family Signaling in Ductal Differentiation and Branching Morphogenesis [109]; TGF-β Family Signaling in Early Vertebrate Development [288]; TGF-β Family Signaling in Embryonic and Somatic Stem-Cell Renewal and Differentiation [180]; TGF-β Family Signaling in Epithelial Differentiation and Epithelial-Mesenchymal Transition [108]; TGF-β Family Signaling in Mesenchymal Differentiation [60]; TGF-β Family Signaling in Neural and Neuronal Differentiation, Development, and Function [170]; TGF-β Family Signaling in the Control of Cell Proliferation and Survival [286]; TGF-β Family Signaling in Tumor Suppression and Cancer Progression [238]; TGF-β Signaling in Control of Cardiovascular Function [58]; TGF-β, Bone Morphogenetic Protein, and Activin Signaling and the Tumor Microenvironment [201]; TGF-β1 Signaling and Tissue Fibrosis [118].

2.0 Parasites, *H. polygyrus*, and the TGF-β Mimic Proteins

2.1 The Hygiene Hypothesis

Helminth parasites represent one of the greatest challenges to human and animal health worldwide, especially in the absence of effective vaccines [91]. Helminth infections are most prevalent in tropical countries, where resources are limited and sanitation is lacking [55, 206]. With helminths infecting up to a quarter of the world's human population [55, 206], a greater knowledge of the molecular mechanisms that aid in sustaining parasitosis would provide a large advancement toward finding treatments for these infections and improving overall global health.

As developing nations implemented sanitation and public health initiatives in the last century, healthcare workers began observing a decline in the rates of helminth infections. Alongside this decline, they also observed a sharp increase in a suite of inflammatory autoimmune and allergic diseases, such as allergic asthma, rheumatoid arthritis, and inflammatory bowel disease. One possibility, which is suggested by the "hygiene hypothesis" and more recently the "old friends" hypothesis, is that helminths are one of the key environmental factors that dampen immune activity to innocuous allergens, such as dust mites and seasonal pollens released by trees and grasses [19, 138, 150, 232].

The idea of immunogenicity versus immune tolerance is one of the central ideas in the "hygiene hypothesis." The immune system has two arms: an immunogenic arm designed to detect and eliminate entities that are not identified by the tolerogenic arm, the second arm, as self [155]. Maintaining the delicate balance between these two arms is critical for the immune system to not

attack self-identified entities while being able to detect and eliminate entities that are foreign to self, such as bacteria, viruses, and parasites.

There is a plethora of evidence showing that helminths downregulate the host immune system [166]. This suggests that they may play a major role in also regulating human immune disorders, and therefore, experts have postulated that the human immune system has evolved such that it is tuned in such a way that the tolerogenic effects induced by parasites are combined with the self-derived tolerogenic effects [153]. Removing the parasite-derived tolerogenic effects has made the human immune system more immunogenic, and this is hypothesized to be the root of many immune-related diseases.

2.2 Helminth Parasites and Immunity

There is a broadening field of evidence suggesting that a host with a chronic helminth infection develops a form of immunological "tolerance" to the parasites and their antigens [102, 146, 148, 151, 166]. This "tolerance" is important for maintaining the population of helminths within the host [146] while protecting the host from the more pathological outcomes from an infection [276].

Helminths act to modulate host immune functions at many levels. This includes interfering with mechanisms involved with innate antigen sensitization, induction of adaptive immunity, and mobilization of effector mechanisms [166]. One of the most prominent pathways that helminths use to modulate the host immune system is through regulatory T-cells (Tregs) [267]. Tregs are a class of immune cells that are characterized by their expression of the master transcription factor Foxp3 (FOXP3 in humans). These cells can also express surface markers that are key for their

function, including CD25, cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), inducible T-cell co-stimulator (ICOS), and T-cell immunoreceptor with Ig and ITIM domains (TIGIT) [226]. Tregs are often defined as $CD3^+$ $CD4^+$ $CD25^+$ $Foxp3^+$ T-cells in both mouse and human studies.

Tregs are an indispensable component of the immune system. These cells help maintain immunological self-tolerance and homeostasis [105, 225, 241]. Tregs act by secreting regulatory cytokines, such as TGF-β [214], Interleukin 35 (IL-35) [37], and Interleukin 10 (IL-10) [4], and the protease Granzyme B [56] that induce cell cycle arrest and ultimately cell death in effector T-cells [155]. This acts as a suppressive mechanism for effector T-cells and therefore acts to down-regulate the immunogenic arm of the immune system.

As stated previously, there is a plethora of evidence that shows that helminths produce a state of immune downregulation in their mammalian hosts [193]. This downregulation is characterized by the inability of effector T-cells to proliferate or produce inflammatory cytokines when stimulated with both parasite antigens and bystander antigens [145, 193]. Therefore, it was reasonable to hypothesize that helminths acted, in some way, to up-regulate the pathways involved with the differentiation of naïve T-cells into Tregs. In addition to immune downregulation, helminths must also maintain a low burden on their host so that they do not risk killing the host. This includes, for example, stimulating the repair of tissues damaged by the parasites as they transit through the organism in different stages of their lifecycle. Hence, it is likely that there must also be additional non-immune-related mechanisms upon which the parasites rely to complete their lifecycles.

TGF-β is one of the cytokines involved with immune regulation by stimulating the differentiation of regulatory T-cells, which promotes immune tolerance by down-regulating the

immunogenic responses [155, 229]. Therefore, this pathway could potentially be coopted by parasites for inducing their tolerogenic effects.

2.3 *Heligmosomoides polygyrus*: A Model Parasite

Heligmosomoides polygyrus is an intestinal roundworm that is found in wild mouse populations throughout the world. One study of wood mouse (*Apodemus sylvaticus*) populations in Oxfordshire, England reported that 70% of all mice sampled had an *H. polygyrus* infection with an average worm burden of 12 worms per mouse [63]. While wood mice are a natural host for *H. polygyrus*, the laboratory mouse (*Mus musculus*) is also susceptible to infection and is used as the model host for human helminth infections [216]. In this way, *H. polygyrus* can act as a model of other gastrointestinal helminth infections, such as those of *Teladorsagia circumcincta* in sheep and *Ascaris lumbricoides* in humans.

Many helminths secrete proteins that can interact with the different cellular and molecular mechanisms present in the host [73, 82, 151, 152, 224, 239, 269]. In the case of *H. polygyrus*, over 350 proteins have been identified in the excretory-secretory products (HES) [83]. These proteins belong to protein families, such as the venom-allergen-like (VAL) superfamily, apyrases, lysozymes, acetylcholinesterases, proteases, and protease inhibitors [83]. Secretion of proteins belonging to these families has also been noted in other helminths of veterinary (*Haemonchus contortus* [283]; *T. circumcincta* [188]) and medical (human hookworms [181]) importance. This gives further credence that *H. polygyrus* can act specifically as a model parasite for studying the immunology associated with parasitic infections.

H. polygyrus has proven to be an excellent model parasite for studying helminth infections, specifically regarding their mechanisms of immunity and immune evasion in infections [148, 149, 216]. Mice infected with *H. polygyrus* adopt an immunoregulated phenotype, with abated allergic and autoimmune reactions [42, 133, 171, 231]. At the cellular level, infection is accompanied by expanded regulatory T-cell populations, skewed dendritic cell and macrophage phenotypes, B-cell hyperstimulation and multiple localized changes within the intestinal environment [71, 249]. These types of effects are also seen in other mammals with parasitic helminth infections [11, 76, 123, 270], making it an excellent laboratory model of these types of infections.

2.4 The *H. polygyrus* TGF-β Mimic

Several studies have shown that an active *H. polygyrus* infection has protective effects against immunopathology in animal models of autoimmunity [42, 61, 133, 171, 204, 231], colitis [61, 71, 165, 167, 249], and allergy [11, 76, 123, 270]. As a result, there has been a great interest in the potential of using HES proteins to down-modulate autoimmune disease pathology *in vivo* [61, 204], with some protective effects seen after treatment in mouse models [61, 165, 167].

Work from the Maizels Lab at the University of Edinburgh showed that immunological "tolerance" of helminths is in part due to the upregulation of Tregs (Figure 5A) that, in turn, downregulate the host's immune response [251]. They showed that treating *Litosomosoides sigmodontis*-infected mice with an anti-Treg antibody cocktail, specifically anti-CD25 and anti-GITR antibodies, increased host immune responsiveness and resulted in clearance of the parasites (Figure 5B) [251]. This same pattern is observed in mice during *H. polygyrus* infection: mice infected with the parasite have expanded populations of $Foxp3^+$ Tregs during chronic infection

[50], and the parasites are cleared by treatment with an anti-Treg antibody cocktail [245]. Therefore, Tregs must play a critical role in downregulating the host's immune system and sustaining *H. polygyrus* parasitosis within mice.

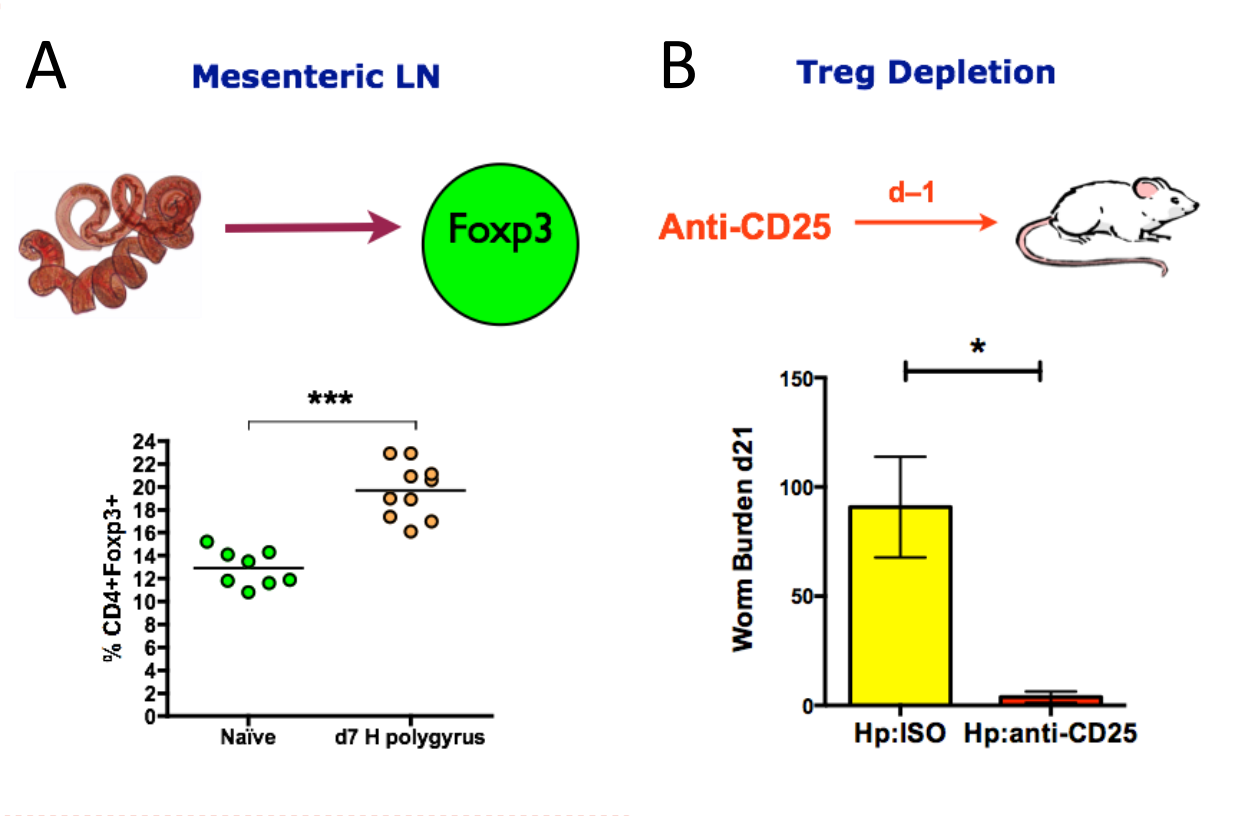

Figure 5. Impact of *H. polygyrus* infection on mouse Tregs. (A) Chronic *H. polygyrus* infection causes an increase in the percentage of $CD4^+$ $Foxp3^+$ Tregs present in the population of $CD4^+$ cells. (B) Depletion of Tregs using an anti-CD25 antibody dramatically decreases worm burden within the host. Figures provided courtesy of the Maizels Lab.

As stated in section 1.4, TGF-β plays a role in the activation and proliferation of Tregs [229]. In addition to the expanded Treg populations, the Maizels Lab observed there was an elevated frequency of $CD4^+$ T-cells that expressed surface TGF-β during chronic *H. polygyrus* infection [50]. Therefore, they hypothesized that TGF-β-like activity was one of the factors responsible for the expansion and proliferation of Tregs during infection.

2.4.1 Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-β pathway

(Note: this sub-section is a summary of the work "Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway" by Grainger, et al., published in 2010 in The Journal of Experimental Medicine [61].)

The Maizels Lab began studying the TGF-β-like activity induced during *H. polygyrus* infection by culturing the adult parasites in growth medium and harvesting the proteins in the medium, as done with mammalian cell cultures [101] (Figure 6). In this fashion, the secreted helminth proteins can be isolated from host proteins, avoiding contaminating proteins and peptides from the mouse host.

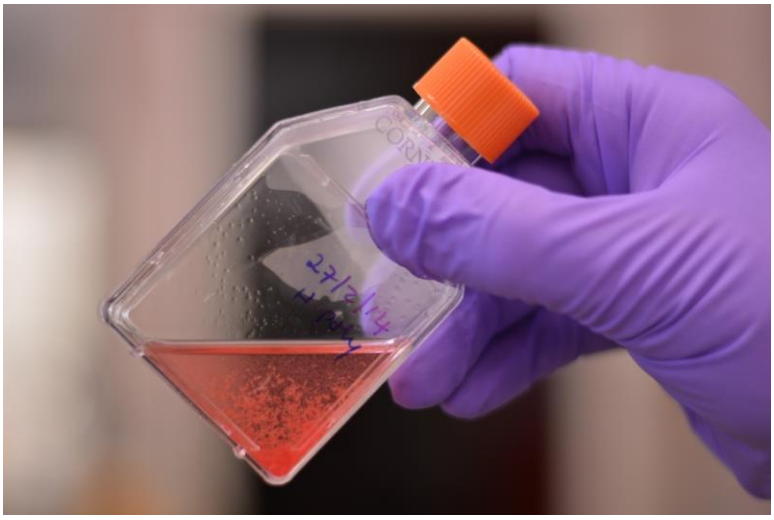

Figure 6. Culture of *H. polygyrus* in serum-free growth medium. Image provided courtesy of the Maizels Lab.

This work resulted in several important findings regarding the interactions between the HES and the mouse immune system. Treatment of naïve mouse splenic T-cells with HES resulted in a larger fraction of $Foxp3^+$ $CD25^+$ cells among the $CD4^+$ cell population, and that the HES acted

on these populations in a dose-dependent manner. They went on to show that HES induces Treg differentiation *in vivo*, even in the absence of a helminth infection.

Compared to a TGF-β1-treated control, there was little difference in the behavior between the HES and TGF-β1 in the induction of Foxp3. Due to the similarities in the effects of HES and TGF-β, they tested the response of HES in a TGF-β reporter cell line. This cell line was derived from murine fibroblasts lacking endogenous TGF-β1 that were transfected with a TGF-β-responsive alkaline phosphatase reporter construct [253]. The assay developed around this reporter cell line is referred to as an MFB-F11 assay, such named for the cell line used (murine fibroblasts; MFB) and the clone that was chosen during screening (F11) [253]. Through these experiments, it was determined that 10 μg mL^{-1} HES activated the reporter cells to a similar extent as 1-5 ng mL^{-1} TGF-β1. Additionally, they tested the ability of HES to covert $CD4^+$ naïve T-cells into $CD4^+$ $Foxp3^+$ Tregs and showed that HES is capable of inducing Foxp3 expression in these cells (Figure 7). This suggested that sequestered mammalian TGF-β, a TGF-β homologue, or a TGF-β mimic may be present within HES.

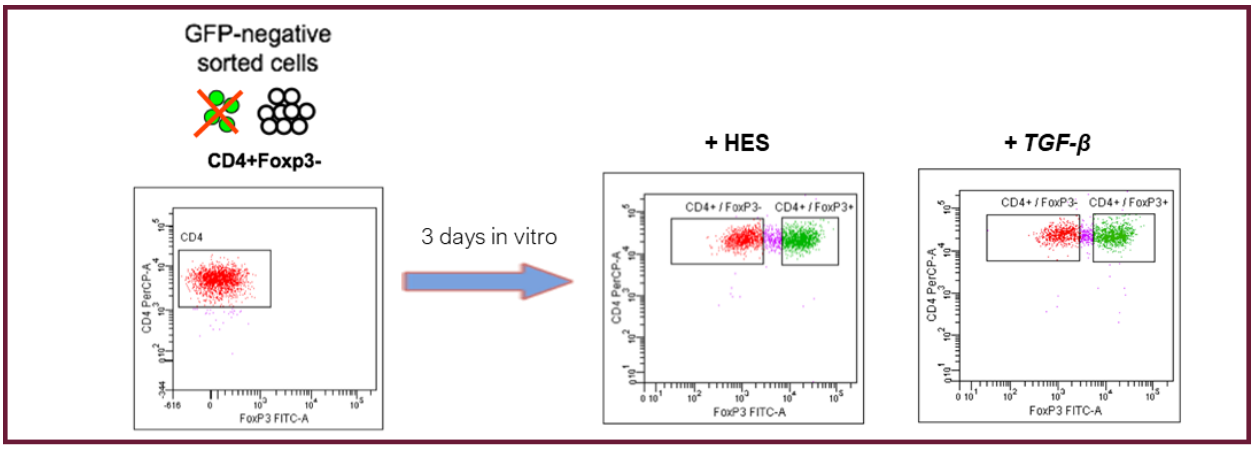

Figure 7. FACS analysis of $CD4^+$ $Foxp3^-$ cells treated with the HES or TGF-β. Treatment with the HES shows similar levels of expression of Foxp3 as cells treated with TGF-β. Figure provided courtesy of the Maizels Lab.

Additionally, they were interested in determining if HES activated the canonical TGF-β signaling pathway downstream of the TGF-β type I and type II receptors. They probed HES- and TGF-β-treated T-cells for phosphorylated Smad2/3 by Western blot, which returned positive results for phosphorylated Smad2/3 for both the HES- and TGF-β-treated T-cells (Figure 8A). Therefore, they concluded that the HES does activate the canonical TGF-β signaling pathway. Furthermore, they showed that the HES did not induce the production of Foxp3 in a cell line lacking the TGF-β type II receptor gene. This response was similar to that of cells expressing basal levels of the TGF-β type II receptor but have their TGF-β signaling blocked by a TGF-β type I receptor inhibitor. Therefore, they concluded that the component in the HES responsible for TGF-β-like signaling was a TGF-β mimic that acted by signaling through the canonical TGF-β pathway.

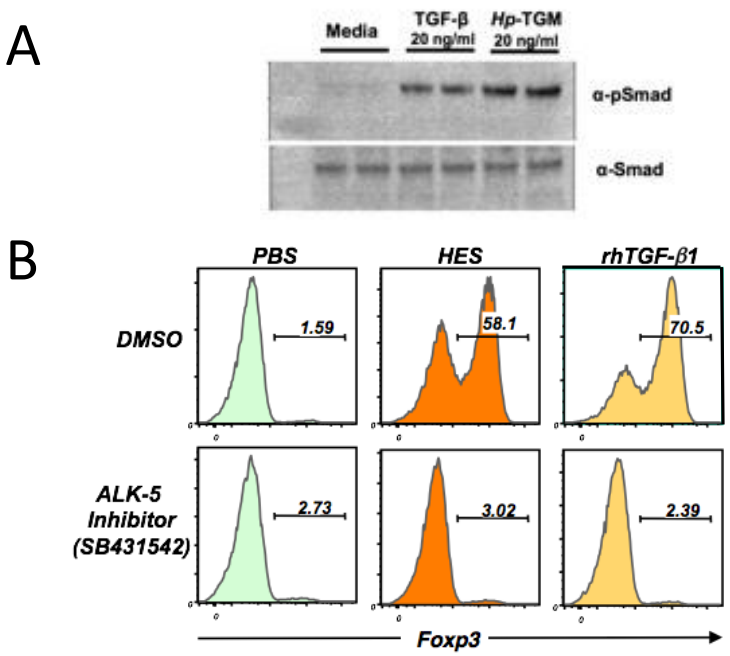

Figure 8. HES signals through the canonical TGF-β signaling pathway. (A) Phosphorylation of Smad2/3 signaling molecules, activated by TGF-β ligation, was measured in C57BL/6 CD4$^+$ cells stimulated with 5 ng/ml TGF-β1, or 10 μg/ml HES. Cell lysates were probed by Western blotting with phospho-Smad2/3-specific antibody (top) and then stripped and reprobed with antibody to SMAD-2/3 peptide backbone (bottom). (B) CD4$^+$ cells were purified from wild-type C57BL/6 mice and were cultured in the presence of PBS, 10 μg/mL HES, or 5 ng/mL TGF-β1. The cells were treated with DMSO (negative control) or the Alk5 inhibitor SB431542. After 72 hours, Foxp3 expression was analyzed by flow cytometry. Figures adapted from Grainger JR, Smith KA, Hewitson JP, McSorley HJ, Harcus Y, Filbey KJ, Finney CA, Greenwood EJ, Knox DP, Wilson MS, Belkaid Y, Rudensky AY, Maizels RM. Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway. J Exp Med. 2010;207(11):2331-41 [61] under the Creative Commons BY-NC-SA 3.0 license.

To further narrow down these possibilities, they performed two sets of experiments. They first tested whether the HES contained a substance that would be reactive in a standard ELISA for mammalian TGF-β. Second, they tested the activity of the HES in the presence of the pan-vertebrate anti-TGF-β blocking antibody 1D11 in the MFB-F11 assay. In both experiments, the HES did not show evidence that it contained TGF-β or a likely TGF-β homologue: the ELISA returned negative results when incubated with up to 50 μg/mL HES (compared to a positive result when incubated with 15 pg mL^{-1} TGF-β1), and the MFB-F11 assay showed TGF-β-like activity

despite the presence of up to 100 μg mL^{-1} of anti-TGF-β blocking antibody. Therefore, they concluded that the component in the HES that was responsible for the TGF-β-like activity was immunologically distinct from TGF-β (neither a TGF-β isoform nor a homologue). Further, they tested serum from chronically infected mice in the same assay. They found that this serum was able to neutralize TGF-β signaling induced by HES, but naïve mouse serum did not interfere with HES signaling. The same antiserum did not impact the action of mammalian TGF-β, confirming that host and parasite ligands are immunologically distinct entities.

2.4.2 A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T cells

(Note: this sub-section is a summary of the work "A structurally distinct TGF-β mimic from an intestinal helminth parasite potently induces regulatory T cells" by Johnston and Smyth, et al., published in 2017 in Nature Communications [103].)

To identify the component in the HES that induced TGF-beta activity, the HES was fractionated by gel filtration and anion exchange chromatography. The fractions were then assayed using the MFB-F11 TGF-β reporter assay (Figure 9) and subjected them to mass spectrometry analysis. Eighteen proteins were identified for which the abundance was highest in the active fractions. Four of these proteins were then cloned and recombinantly produced in HEK-293 cells. The recombinant proteins were again assayed in the TGF-β reporter assay, and one was identified as having a high level of TGF-β-like activity, exceeding that of direct TGF-β stimulation. They named this protein the *H. polygyrus* TGF-β mimic (HpTGM).

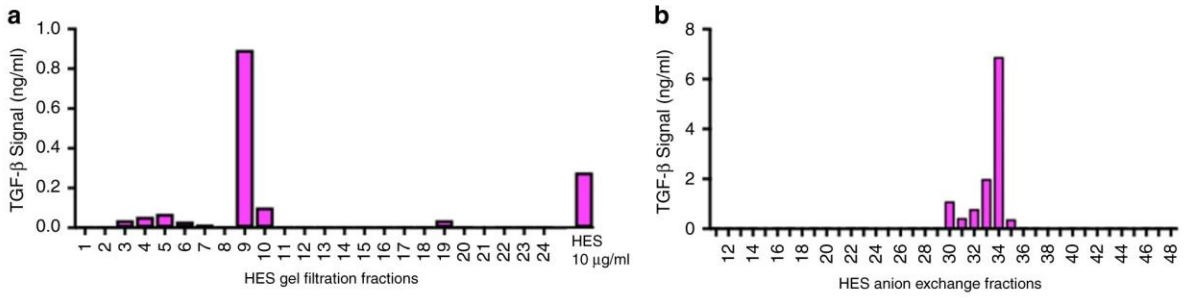

Figure 9. Identification of TGM1 by TGF-β activity in HES fractions. (A) Fractionation of HES by SEC. 1 mg of HES was separated on a Superdex 200 10/300 GL column and 1 ml fractions collected for assay with MFB-F11 reporter cells; responses were calibrated with recombinant human TGF-β1. (B) as (A), fractionation by ion exchange FPLC on a Mono QTM 5/50 G column. Figure adapted from Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741 [103] under the Creative Commons Attribution 4.0 International License.

The amino acid sequence of active HpTGM is 422 residues in length (Figure 10A). The first 18 residues are predicted to form a classical signal peptide with the remainder forming a mature 404-residue protein. The mature protein contains 22 cysteine residues, forming 11 disulfide bonds, and five potential N-linked glycosylation sites. Despite having high levels of TGF-β activity, the mature protein has no sequence similarity to any members of the TGF-β family. However, it does contain five homologous but not identical domains, approximately 80 residues long each, that have distant similarity to the Complement Control Protein (CCP, or Sushi) family (Figure 10A). This similarity arises from the positions of four cysteine residues and conserved tyrosine, glycine, and tryptophan residues in each domain. The mature protein is encoded as an 11-exon gene within the parasite's genome (Figure 10B). This corresponds with the signal peptide (Exon 1) and five pairs of exons whose boundaries exactly match those of the CCP domains.

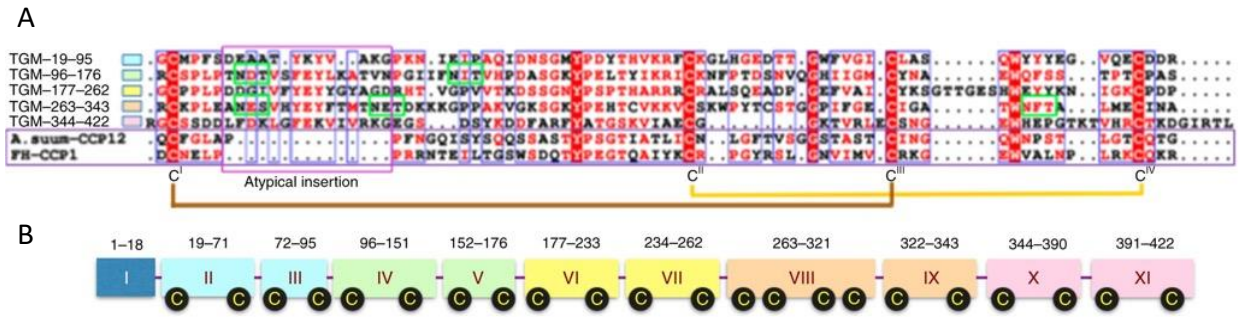

Figure 10. Sequence and domain structure schematic for TGM1. (A) Alignment of five similar domains within Hp-TGM encompassing the entire amino acid sequence apart from the predicted signal peptide (aa 1–18), with conserved cysteine (white on red) and other residues indicated, together with a CCP module from the nematode *Ascaris suum* (domain 12 of ASU_08405, aa 954–1018), and an archetypal CCP domain, human Factor H module 1 (X07523, aa 20–83). Other conserved residues are shown in red, and potential N-glycosylation sites outlined in green. Amino acid positions for each domain of TGM1 are indicated on the left. Note the presence of a 15-aa insertion near the N-terminal of each domain of TGM1 which is not typical of the CCP family. Positions of disulfide bonds in Factor H are shown below the alignment by linked cysteine residues C^{I} – C^{IV}. (B) Exon-intron structure of TGM1 in the *H. polygyrus* genome; domains are colored corresponding to symbols in (A); positions of cysteine residues indicated in black circles Figure adapted from Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741 [103] under the Creative Commons Attribution 4.0 International License.

As previously stated, HpTGM signals through the canonical TGF-β pathway, likely binding to both the TGF-β type I and type II receptors, Alk5 and TβRII, respectively. The performance of HpTGM was then tested in the MFB-F11 assay in comparison to TGF-β and the HES. All three treatments induced activation of the reporter in a dose-dependent manner (Figure 11). The response of the MFB-F11 assay to increasing concentrations of HpTGM reached a maximum signal significantly greater than attained by even the highest concentrations of TGF-β ($A_{405\,nm}$ at 100 ng/ml protein: TGM1 = 2.46 ± 0.16 and TGF-β = 1.48 ± 0.02). The MFB-F11 assay response to HES also exceeded the highest level of the TGF-β-induced signal, but it required more than 1000-fold higher concentrations to achieve the same signal as HpTGM.

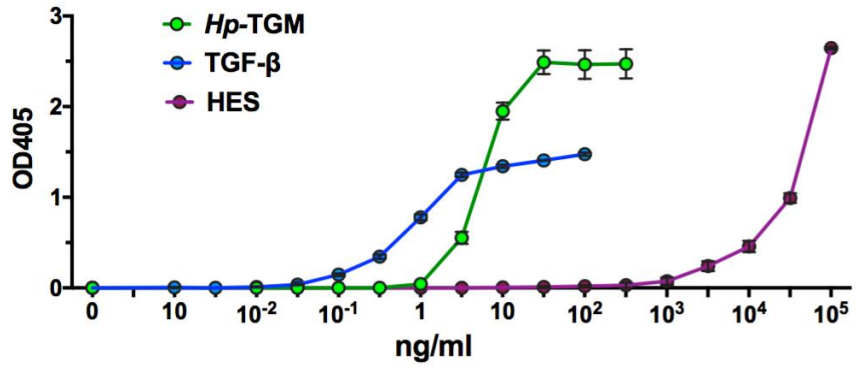

Figure 11. Results of the MFB-F11 assay for TGM1. The results of the MFB-F11 TGF-β-responsive bioassay show that TGM1 is equally as active as if not more active than TGF-β. The complex HES mixture requires a 1000-fold higher concentration to reach the same levels of activity as TGM1. Figure adapted from Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741 [103] under the Creative Commons Attribution 4.0 International License.

In collaboration with the Hinck lab, they used surface plasmon resonance (SPR) to determine that HpTGM bound both the TGF-β type I receptor, Alk5, and the TGF-β type II receptor, TβRII, with low micromolar affinity. The affinities are on the order of those for TGF-β binding to its receptors (Table 1).

Table 1. Binding constants for HpTGM and TGF-β binding to the TGF-β type I and type II receptors. Table adapted from Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741 [103] under the Creative Commons Attribution 4.0 International License.

Ligand	Receptor	K_D (μM)
HpTGM	Alk5	0.351
TGF-β	Alk5	n.d.
HpTGM	TβRII	2.96
TGF-β	TβRII	0.294
HpTGM	Alk5 + TβRII	0.348
TGF-β	Alk5 + TβRII	0.367

n.d. – Not determined due to weak binding. TGF-β does not bind to Alk5 in the absence of TβRII [64, 90].

As done for the work with the HES, it was showed that HpTGM signaling was not inhibited by an anti-TGF-β antibody in the MFB-F11 assay (Figure 12A) [103]. It was also showed that HpTGM signaling, like that for the HES and TGF-β, was inhibited using TGF-β receptor inhibitors in the MFB-F11 assay (Figure 12B,C) [103]. Additionally, overnight incubation of splenocytes from C57BL/6 mice with HpTGM induced phosphorylation of Smad2/3 at least as efficiently as TGF-β (Figure 12D) [103]. Furthermore, incubation of human peripheral blood mononuclear cells with HpTGM produced similar percentages of $CD25^+$ $Foxp3^+$ cells among the total $CD4^+$ cell population as incubation with TGF-β (Figure 12E) [103]. From these sets of experiments, it was concluded that HpTGM not only binds directly to the TGF-β receptors, but it also signals through the canonical TGF-β pathway and activates Tregs.

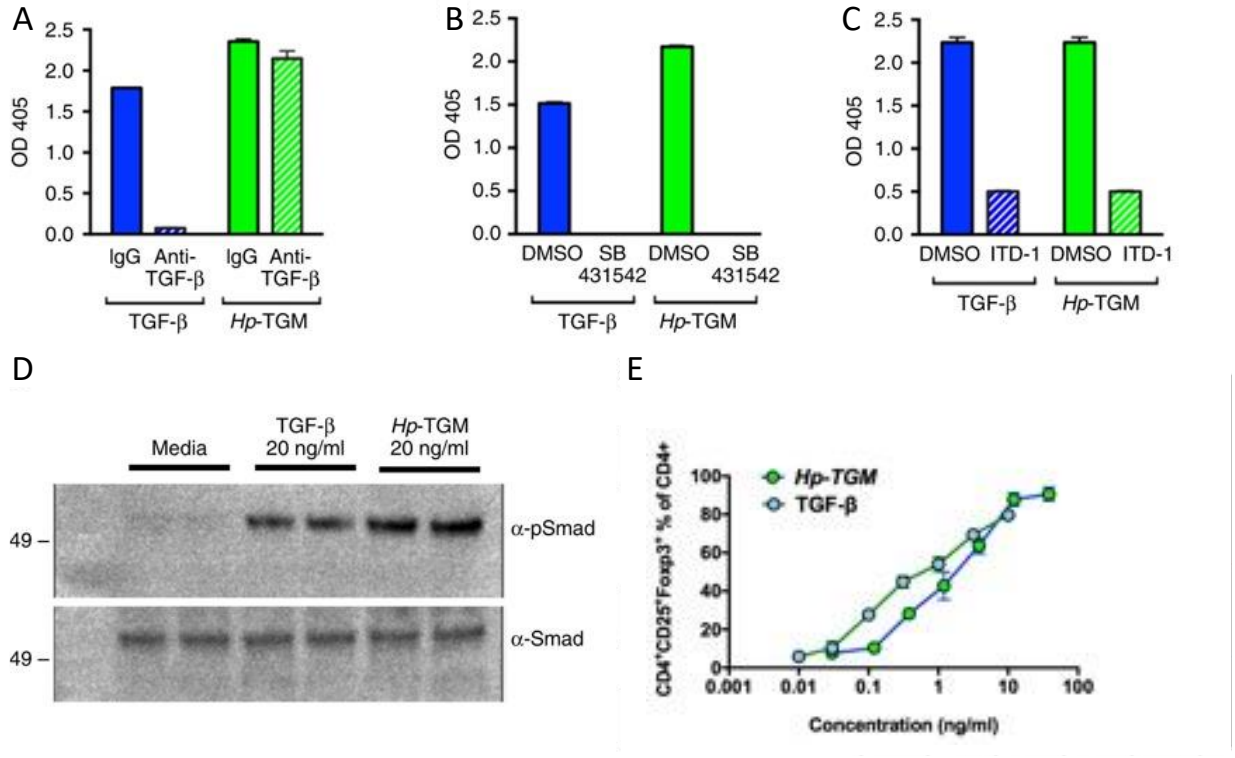

Figure 12. Hp-TGM signals through the TGF-β pathway and activates Tregs. Activity shown from MFB-F11 bioassay after 24 h of culture with TGF-β or HpTGM incubated with anti-TGF-β monoclonal antibody or MOPC31C IgG control. (B,C) Abolition of signaling by inhibitors of the TGF-β receptor kinases. Activity shown from MFB-F11 bioassay after 24 h of culture of TGF-β and HpTGM with: (B) the Alk5 inhibitor, SB431542 or DMSO control; and (C) the TβRII inhibitor, ITD-1. (D) Western blots (Smad2 and phospho-Smad2): cell lysates from C57BL/6 splenocytes following culture at 37 °C for 18 h incubated with TGF-β or HpTGM. (E) Percentage of $CD25^+$ $Foxp3^+$ cells among total $CD4^+$ human peripheral blood mononuclear cells incubated for 96 h with HpTGM or TGF-β. Figure adapted from Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741 [103] under the Creative Commons Attribution 4.0 International License.

2.4.3 TGF-β mimic proteins form an extended gene family in the murine parasite *Heligmosomoides polygyrus*

(Note: this sub-section is a summary of the work "TGF-β mimic proteins form an extended gene family in the murine parasite *Heligmosomoides polygyrus*" by Smyth, et al., published in 2018 in the International Journal of Parasitology [246].)

Proteomics and transcriptomics analysis of *H. polygyrus* [83, 84] showed that HpTGM was not the only member of this protein family. There are at least nine homologues of HpTGM; as such, HpTGM, as the founding member of the family, was renamed TGM1. These homologues are summarized in Figure 13. See Appendix A for sequence alignments between the domains of the individual family members. (Note: the family of homologues will be referred to as TGMs, and HpTGM will be henceforth referred to as TGM1.)

From the adult parasites, three closely related homologues were found and named TGM2, TGM3, and TGM4, having 95.8%, 93.8%, and 80.4% amino acid identity across the mature protein, respectively. Two additional homologues lacking full domains were also found: TGM5, missing domain 4 (D4), shares 90.1% amino acid identity across the other four domains; and TGM6, missing domains 1 and 2 (D1 and D2, respectively), shares 47.2% amino acid identity across the other three domains. Finally, they showed that TGM7, TGM8, TGM9, and TGM10 were secreted exclusively during the larval stage of the parasite since these TGMs had similar sequences and abundances to previously discovered Larval Secreted Proteins (LSPs) [84].

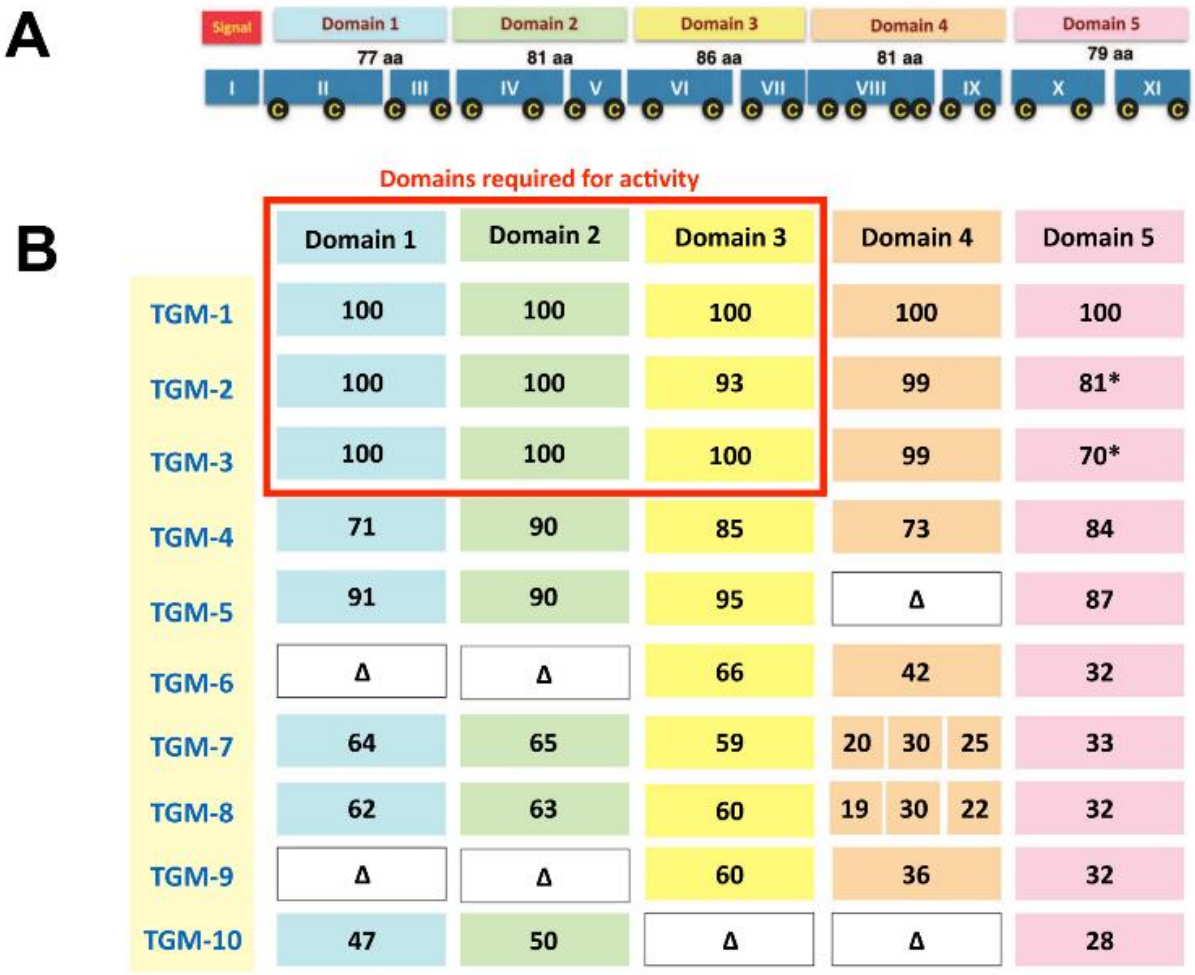

Figure 13. TGM domain structure schematic for the TGM family of proteins. The paired exons that encode the five CCP-like domains (D1-D5) are shown (A). Shown below is the sequence identity comparison of the TGM homologues (B). The percent identity for each domain is given relative to that domain in TGM1. Domains marked with Δ are missing. Figure adapted from Smyth DJ, Harcus Y, White MPJ, Gregory WF, Nahler J, Stephens I, Toke-Bjolgerud E, Hewitson JP, Ivens A, McSorley HJ, Maizels RM. TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus. Int J Parasitol. 2018;48(5):379-85. [246] under the Creative Commons Attribution 4.0 International License.

To further characterize the new members of the TGM family, the Maizels Lab then assayed the different members of the family for TGF-β activity in the MFB-F11 reporter assay (Figure 14A). TGM2 and TGM3 showed strong activity, like TGM1, in the MFB-F11 assay. This activity was blocked by the Alk5 inhibitor SB431542. However, TGM4, TGM6, and TGM7 were not active in the same assay. They also tested the ability of the TGM family members to induce Foxp3

in CD4^{+} cells (Figure 14B). TGM2 showed nearly identical activity compared to TGM1 and TGF-β in inducing Foxp3 expression. While TGM3 had similar activity in TGF-β signaling assays as TGM1 and TGM2, it did not have activity in the Treg conversion assays. Additionally, TGM4 did not have activity in TGF-β signaling assays but did convert naïve T-cells into Tregs with a slightly lower efficiency than TGM1 and TGM2. Finally, TGM6 and TGM7 had no activity in either the TGF-β reporter assays or the Treg conversion assays.

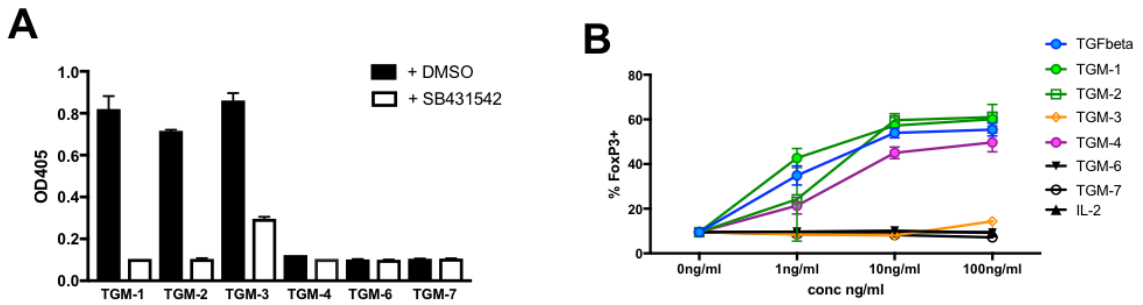

Figure 14. TGF-β reporter assay testing of the TGM family members. (A)Activity detected in the MFB-F11 bioassay after 24 h of culture with TGM family members in the absence or presence of the Alk5 inhibitor, SB431542. TGM-5 is omitted from the assays as it was not possible to express recombinant protein in the same system as the other family member proteins. (B) Percentage of Foxp3 induction induced by TGM-1 and TGM-2 compared with IL2 only (no induction control) and TGF-β (positive control). Figure adapted from Smyth DJ, Harcus Y, White MPJ, Gregory WF, Nahler J, Stephens I, Toke-Bjolgerud E, Hewitson JP, Ivens A, McSorley HJ, Maizels RM. TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus. Int J Parasitol. 2018;48(5):379-85. [246] under the Creative Commons Attribution 4.0 International License.

To further the characterization of TGM1 binding to the TGF-β receptors, the performance of different TGM1 domain truncation variants was measured in the MFB-F11 assay (Figure 15). It was determined that the minimal structure required for TGM1's TGF-β activity was a construct that combined domains 1, 2, and 3. Domains 4 and 5 were not necessary for TGF-β activity, and the absence of domains 1, 2, or 3 abolished all TGF-β activity.

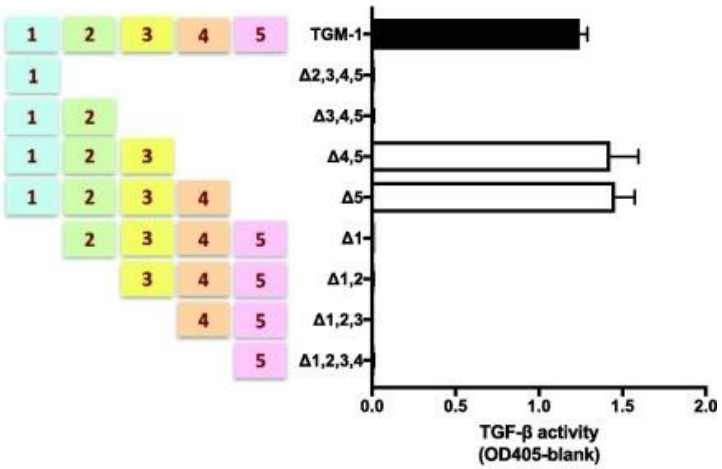

Figure 15. Activity profile of TGM truncations in the MFB-F11 reporter cell assay. Figure adapted from Smyth DJ, Harcus Y, White MPJ, Gregory WF, Nahler J, Stephens I, Toke-Bjolgerud E, Hewitson JP, Ivens A, McSorley HJ, Maizels RM. TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus. Int J Parasitol. 2018;48(5):379-85 [246] under the Creative Commons Attribution 4.0 International License.

2.4.4 Convergent evolution of a parasite-encoded CCP-scaffold to mimic binding of mammalian TGF-β to its receptors, TβRI and TβRII

(Note: this sub-section is a summary of the work "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., published in 2022 in the Journal of Biological Chemistry [179].)

The binding of TβRI and TβRII to TGM1 domains 1, 2, and 3 was then further characterized. Using SPR and ITC, it was shown that TGM1-D1 and -D2, together, bound TβRI, with TGM1-D2 being the primary binding site, and that TGM1-D3 bound TβRII. A combination of competition ITC binding experiments and NMR spectroscopy were then used to further show that TGM1-D3 was a competitor for the TGF-β binding site on TβRII and that TGM1-D12 and the TGF-β:TβRII complex use a similar set of residues when binding TβRI.

To further characterize TGM1, the solution structure of TGM1-D3 was determined by NMR spectroscopy (Figure 16). The domain structure is comprised of four anti-parallel β-strands that form a twisted β-sheet core with two disulfide bonds in a Cys^{I}-Cys^{III} and Cys^{II}-Cys^{IV} pattern. The β1-β2 loop forms an extended and ordered loop structure, referred to as the hypervariable loop, that packs against the face that does not contain the N- and C-termini (the non-NC face). The β2-β3 loop additionally contains a short 3_{10}-helical segment. When the structure of TGM1-D3 was aligned with and compared to previously published CCP family protein structures, the structures showed a close correspondence of the four β-strands and the disulfide bonds that form the core of the protein.

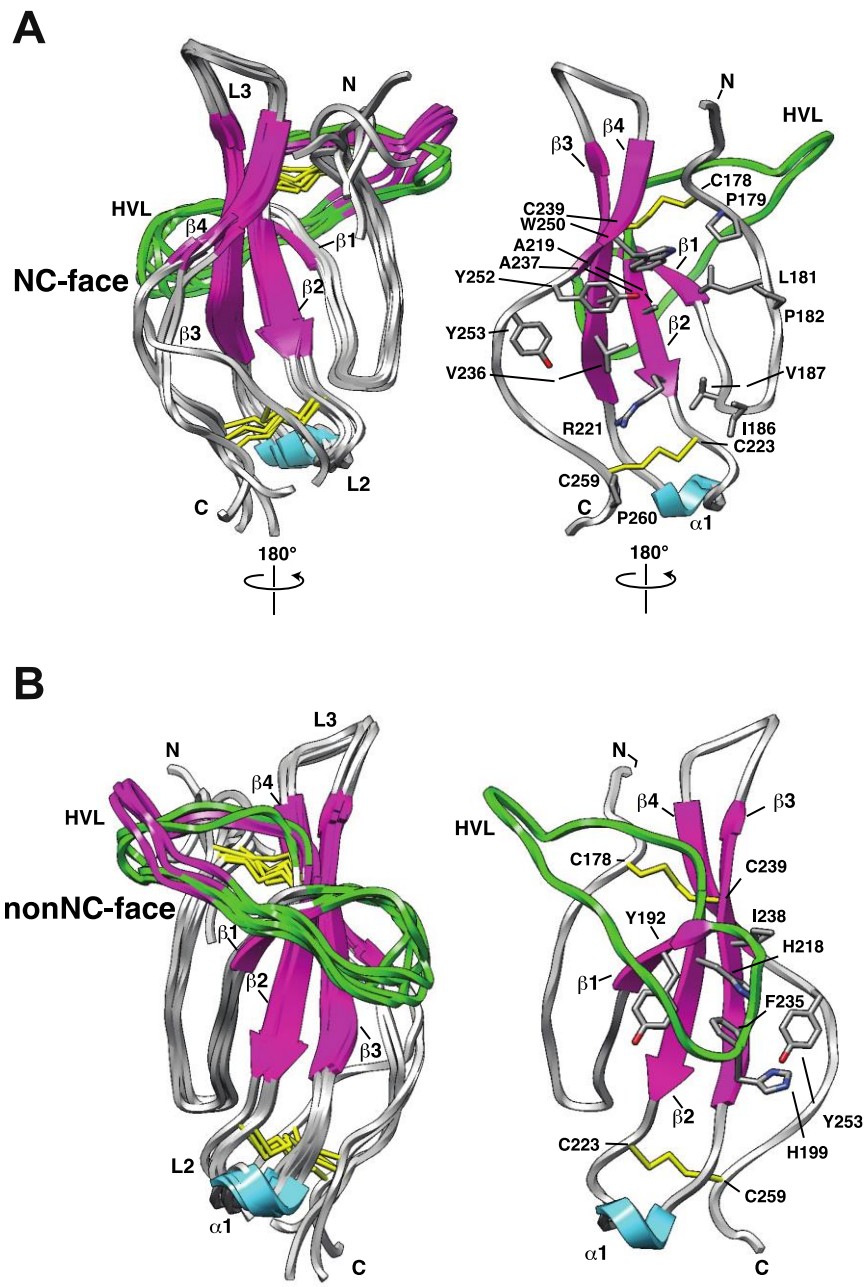

Figure 16. Structure of TGM1-D3. A-B. Shown on the left are an ensemble of the five-lowest energy NMR structures of the unbound form of TGM1-D3: β-strands, magenta; loops, gray; 3_{10} helix, cyan; disulfide bonds, yellow, two conformations of HVL highlighted in green and pink. Key structural features are indicated. Orientations shown differ by a 180-degree rotation around the y-axis, with orientation shown in A highlighting the face of the protein that includes the N- and C-termini (NC face) and the orientation shown in B highlighting the opposite face (non-NC face). Shown to the right are single representative structures, with the four cysteines that form the two disulfide bonds and the side-chains of key residues highlighted. Figure adapted from Mukundan A, Byeon CH, Hinck CS, Cunningham K, Campion T, Smyth DJ, Maizels RM, Hinck AP. 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. 2022:101994 [179] under the Creative Commons CC-BY license.

A combination of NMR chemical shift perturbation measurements and SPR affinity attenuation measurements, the latter of which were performed on single point mutation variants of both TGM1-D3 and TβRII, were then used to determine where in TGM1-D3 and to which specific residues TβRII bound. By mutating key residues to Ala, it was shown that TGM1-D3 binds TβRII along strands β3 and β4 on the non-NC face. The binding interface on TβRII was also assessed by SPR affinity attenuation measurements using single point variants of the TβRII extracellular domain. This showed that TGM1-D3 used many of the same residues as TGF-β when binding TβRII.

2.5 Project Overview and Objectives

To date, most of the work on the TGM family of proteins has been focused on TGM1. Very little is known about TGM6. Therefore, the focus of this dissertation is the study of the structure and function of the *H. polygyrus* TGF-β mimic TGM6.

The objective of this series of studies was to: a) determine whether TGM6 binds a type II receptor of the family, such as ActRII, ActRIIb, BMPRII, or TβRII and if it does so through -D3; b) determine if TGM6 binds a type I receptor of the family, such as Alk1, Alk2, Alk3, Alk4, or Alk5, and if so whether it does so through -D4, -D5, or -D45; c) determine structures of the relevant TGM6 binding domains alone and as bound to their cognate receptors; and d) use cellular assays to determine the function of TGM6. The completion of the proposed studies allows us to better understand the function of this unique member of the TGM family, specifically regarding how it functions and interacts with the TGF-β family type I and type II receptors. To study TGM6, a combination of structural, biological, and biophysical methods was used including, but not limited

to: NMR spectroscopy, X-ray crystallography, isothermal titration calorimetry, surface plasmon resonance, and TGF-β-responsive reporter assays.

Overall, the proposed studies provide molecular-level details into how *H. polygyrus* TGM6 binds to the TGF-β family type I and type II receptors. This information will provide important information into how TGM6 works in concert with the other TGMs to sustain parasitosis and/or to prevent parasite-induced damage to the host.

3.0 The Solution NMR Assignment of TGM6 Domain 3

3.1 Introduction

In a survey of modern structural biology and biophysical methods, structure determination and protein characterization by nuclear magnetic resonance (NMR) spectroscopy is a major technique used in the field. Compared to X-ray crystallography, NMR spectroscopy does not require large amounts of protein, nor does it require extensive sample preparation and crystallization of proteins. Additionally, the structures determined by NMR spectroscopy are more closely related to the native environment in which the protein exists *in vivo*, as these structures are determined in a solution state. However, NMR spectroscopy does require isotopic labelling of samples, which can be expensive, and requires large capital investments for the spectrometers, themselves. However, the benefits of using NMR spectroscopy for protein structure determination, especially for proteins that are difficult or impossible to crystallize, outweigh the drawbacks.

To determine the structure of a protein, there are three general steps that must be completed after the proteins are purified and characterized. First, each resonance in the NMR spectra must be assigned to specific atoms in the protein. This process is key for the second step, which is determining restraints used in structure determination. Once the restraints are determined, structure calculations are performed to determine an ensemble of representative structures determined from the given restraints. While this chapter is not intended as an exhaustive discussion of the process of NMR-based structure determination, the process and NMR spectra used for the assignment and those collected for use in structure determination for TGM6-D3 will be discussed.

3.2 Methods Used for Resonance Assignment

3.2.1 NMR experiments

To assign the resonances present in the NMR spectra, we used a series of multi-dimensional NMR experiments, which are summarized in Table 2. Descriptions of the experiments follow. For more details on the experiments, see the references for each experiment.

Table 2. Summary of NMR experiments used in TGM6-D3 work. Data compiled from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

Experiment	Dimensions	Minimum Labelling	Reference(s)
^{1}H-^{15}N HSQC	2	^{15}N	[20, 26, 175, 202, 244]
^{1}H-^{13}C CT-HSQC	2	^{13}C	[223, 260]
HNCO	3	^{15}N, ^{13}C	[65, 112, 114, 178, 234]
HN(CA)CO	3	^{15}N, ^{13}C	[36, 114]
CBCA(CO)NH	3	^{15}N, ^{13}C	[67, 70, 178]
HNCACB	3	^{15}N, ^{13}C	[66, 178, 271]
CC(CO)NH	3	^{15}N, ^{13}C	[25, 35, 68, 136, 141, 174]
HBHA(CO)NH	3	^{15}N, ^{13}C	[70, 178]
H(CC)(CO)NH	3	^{15}N, ^{13}C	[25, 35, 68, 136, 141, 174]
HCCH-TOCSY	3	^{15}N, ^{13}C	[13, 113, 191]
CB(CGCD)HD	2	^{15}N, ^{13}C	[279]
CB(CGCDCE)HE	2	^{15}N, ^{13}C	[279]
HNHA	3	^{15}N, ^{13}C	[261, 262]
HNHB	3	^{15}N, ^{13}C	[5, 44, 137]
IPAP-HSQC	2	^{15}N	[194, 282]
^{15}N-filtered 1H NOESY	3	^{15}N	[111, 157, 158, 195, 197, 235, 285, 289]
^{13}C-filtered 1H NOESY	3	^{13}C	[111, 157, 158, 195, 197, 235, 285, 289]

For each experiment description, a diagram depicting the flow of the magnetization and an example spectrum are shown. In the magnetization flow diagrams, the nuclei on which evolution occurs are marked in orange while nuclei through which the magnetization is transferred but on

which no evolution occurs are marked in blue for the atoms on the residue containing the amide NH group, i, and its preceding residue, i-1.

3.2.1.1 The ^{1}H-^{15}N HSQC experiment

The first experiment typically performed on protein samples is the ^{1}H-^{15}N Heteronuclear Single Quantum Correlation (HSQC) experiment [20, 26, 175, 202, 244]. This experiment gives a "fingerprint" of the protein, with each peak in the spectrum representing a proton-nitrogen pair. These atom pairs are typically in the protein backbone, but Asn sidechain Nδ-Hδ$_2$, Gln sidechain Nε-Hε$_2$, and Trp sidechain Hε-Nε peaks are also observed. A diagram of the flow of the magnetization for this experiment is shown in Figure 17, and an example spectrum is shown in Figure 18.

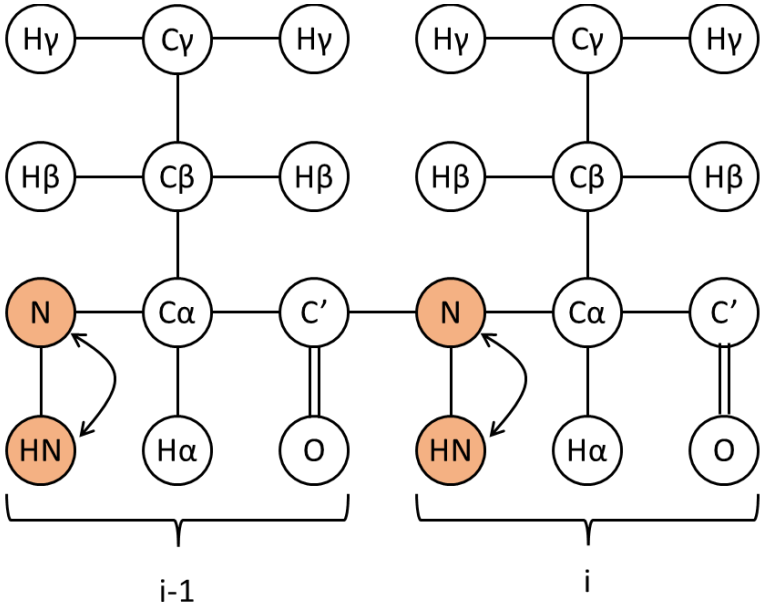

Figure 17. Magnetization flow diagram for the ^{1}H-^{15}N HSQC experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

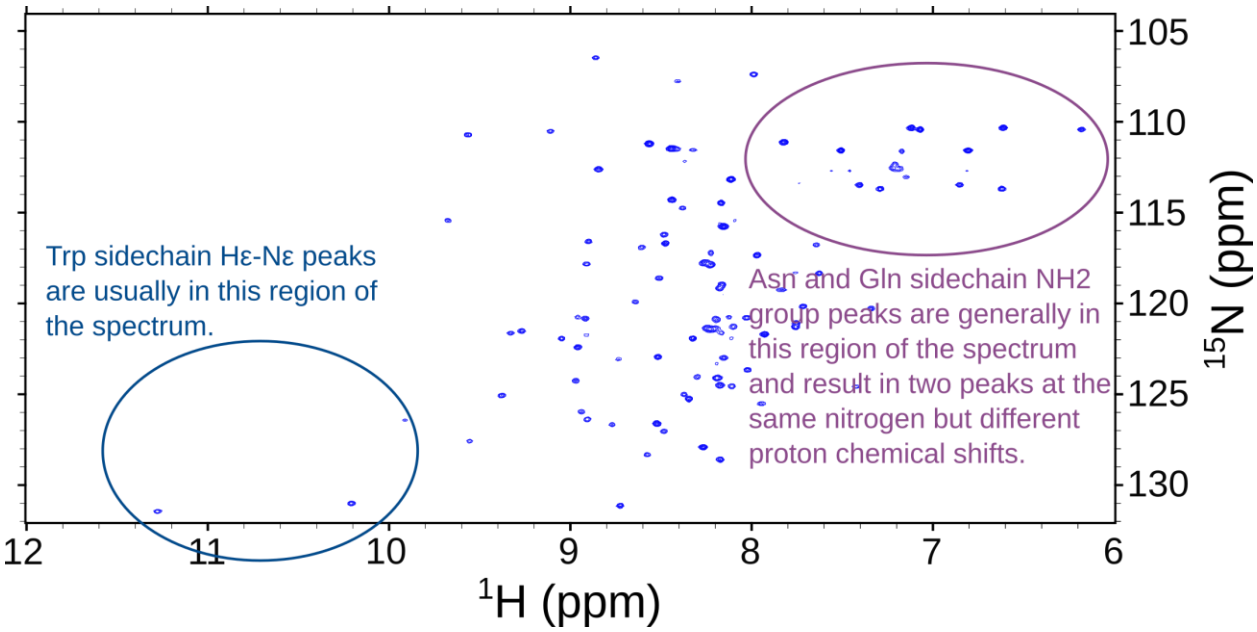

Figure 18. Example spectrum for the ^{1}H-^{15}N HSQC experiment. Data were collected for TGM6-D3.

In this experiment, magnetization begins on the amide proton and is transferred to the attached ^{15}N via the J-coupling. The magnetization evolves on the amide nitrogen prior to being transferred back to the amide proton via the J-coupling. The magnetization is then detected from the amide proton.

From this experiment, a lot of early information can be gleaned regarding the protein's folding state and purity. The number of peaks present in the ^{1}H-^{15}N HSQC spectrum should equal the number of peaks expected based on the protein sequence: each residue aside from the N-terminus, which is an amine instead of an amide, and prolines, which do not have an amide proton, will have one peak each for the backbone; tryptophan residues will have an additional peak for the sidechain Hε-Nε correlation; and asparagine and glutamine residues will each have an additional two peaks for the sidechain Nδ-Hδ$_2$ and Nε-Hε$_2$ correlations, respectively. The positioning of these peaks is shown in Figure 18. Additionally, proteins that are in a "random-coil" conformation will have backbone amide resonances with a proton chemical shift between 7.8 ppm and 8.6 ppm. This

is also a signature of non-natively-folded protein, as natively-folded proteins have backbone amide resonances with proton chemical shifts that are well-dispersed and typically spread between 6.0 ppm and 11.0 ppm [223].

3.2.1.2 The ^{1}H-^{13}C CT-HSQC experiment

The ^{1}H-^{13}C constant-time HSQC (CT-HSQC) experiment [223, 260] is similar to the ^{1}H-^{15}N HSQC experiment in that each peak in the spectrum represents a proton-carbon pair present in the protein. A diagram of the flow of the magnetization for this experiment is shown in Figure 19, and an example spectrum is shown in Figure 20.

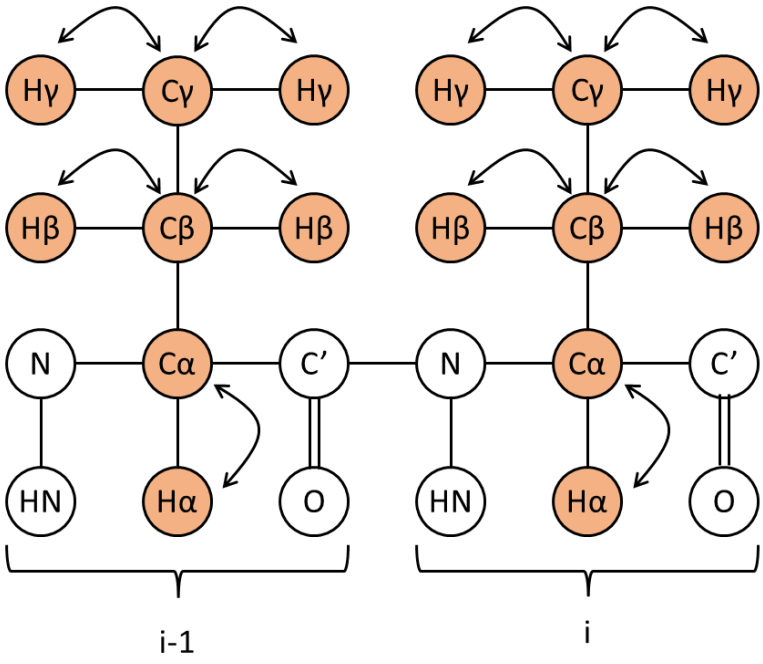

Figure 19. Magnetization flow diagram for the ^{1}H-^{13}C CT-HSQC experiment.

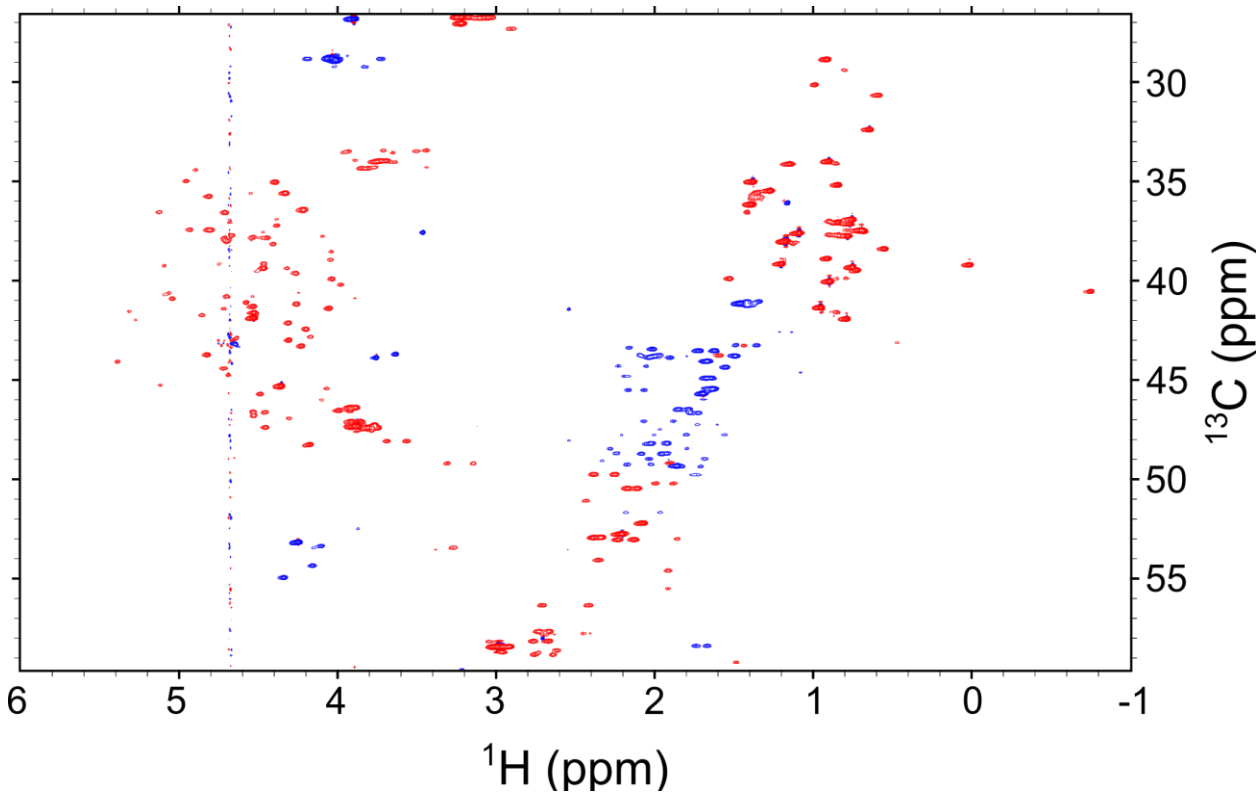

Figure 20. Example spectrum for the ^{1}H-^{13}C CT-HSQC experiment. Data were collected for TGM6-D3 using a 2T value of 26.6 msec. Peaks with positive intensity are shown in blue; peaks with negative intensity are shown in red. The vertical line of peaks at 4.701 ppm is an artifact from the water in the sample.

In this experiment, magnetization begins on the proton and is transferred to the attached ^{13}C via the J-coupling. The magnetization evolves on the carbon nucleus prior to being transferred back to the proton via the J-coupling. The magnetization is then detected from the proton.

Unlike the ^{1}H-^{15}N HSQC, the pulse sequence of the ^{1}H-^{13}C CT-HSQC experiment contains a constant time element that suppresses the effect of the ^{13}C-^{13}C one-bond J-couplings and, therefore, the ^{1}H peak splitting from the spectrum, giving a single peak per ^{1}H-^{13}C pair [260]. The length of the constant time, 2T, is set based on the region of the spectrum being recorded: for alpha protons, 2T = 13.3 ms; for aliphatic protons, 2T = 26.6 ms; for aromatic protons, 2T = 17.6 ms. In the alpha proton CT-HSQC spectrum, all peaks have an intensity of the same sign: positive by convention; however, in the aliphatic proton spectrum, carbons with an odd number of bonded

non-decoupled J-coupled carbons, such as the terminal methyl group in a sidechain, have a positive intensity while carbons with an even number of bonded non-decoupled J-coupled carbons have a negative intensity, such as the beta carbon in the isoleucine sidechain.

Once the signals in this spectrum are assigned, this spectrum is useful for assigning signals in NOESY spectra that are themselves edited using a ^{1}H-^{13}C CT-HSQC, such as in the ^{13}C-filtered NOESY (NOESY-HSQC) experiment.

3.2.1.3 The HNCO and HN(CA)CO experiments

The HNCO [65, 112, 114, 178, 234] and HN(CA)CO [36, 114] experiments are some of the first and simplest three-dimensional NMR experiments performed on protein samples. These experiments allow for the assignment of the backbone carbonyl carbon atoms and sidechain amide carbonyl groups from asparagine and glutamine sidechains. Diagrams of the flow of the magnetization for these experiments are shown in Figure 21 and Figure 22, respectfully, and example spectra are shown in Figure 23.

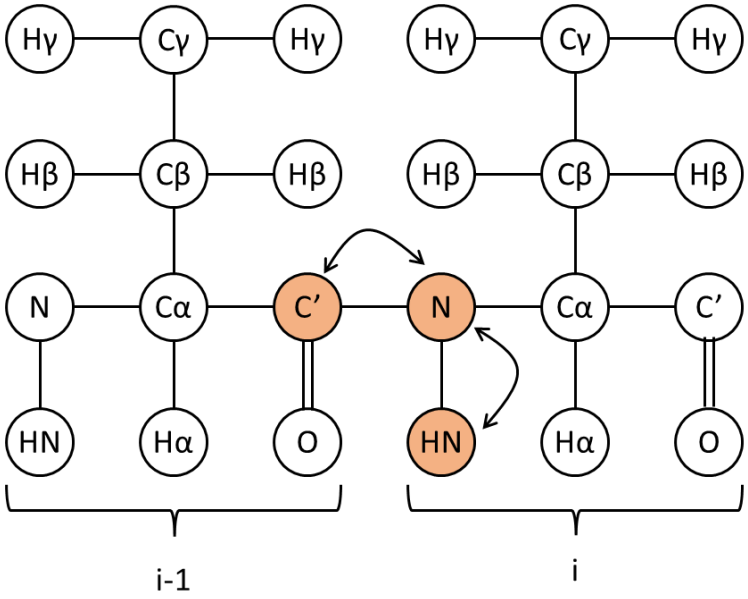

Figure 21. Magnetization flow diagram for the HNCO experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

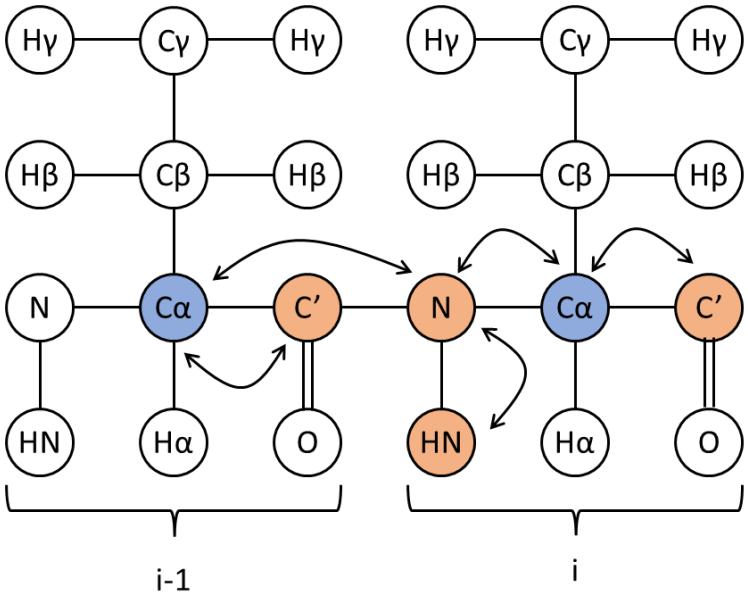

Figure 22. Magnetization flow diagram for the HN(CA)CO experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

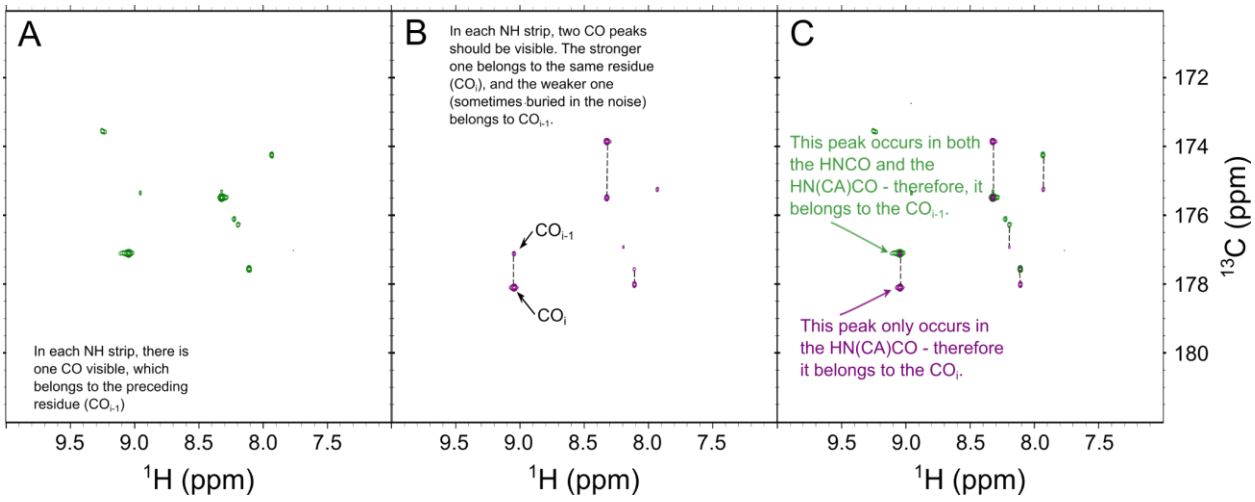

Figure 23. Example spectra for assigning the carbonyl resonances. (A) Example spectrum for the HNCO experiment. (B) Example spectrum for the HN(CA)CO experiment. (C) Overlay of the example HNCO (green) and HN(CA)CO (purple) spectra. Data were collected for TGM6-D3 with a ^{15}N chemical shift of 122 ppm.

For the HNCO experiment, magnetization begins on the amide proton on residue i and is transferred to the attached amide nitrogen ($^{15}N_i$) via the ^{1}H-^{15}N J-coupling. The magnetization evolves on the amide nitrogen prior to being transferred to the attached carbonyl carbon of the proceeding residue (CO_{i-1}) via the ^{15}N-^{13}CO J-coupling. The magnetization then evolves on the carbonyl carbon, resulting in a three-dimensional spectrum. The magnetization is then transferred back to the amide proton through the amide nitrogen via the ^{13}C-^{15}N and ^{1}H-^{15}N one-bond J-couplings, respectively, without further evolution and is detected via the amide proton.

For the HN(CA)CO experiment, magnetization begins on the amide proton of residue I and is transferred to the attached amide nitrogen ($^{15}N_i$) via the ^{1}H-^{15}N J-coupling. The magnetization evolves on the amide nitrogen prior to being transferred to the attached alpha carbons of both the excitation residue ($C\alpha_i$) and the proceeding residue ($C\alpha_{i-1}$) via the ^{15}N-$^{13}C\alpha$ J-couplings. Magnetization is not allowed to evolve on the Cα atoms, and the magnetization is immediately transferred to the carbonyl carbons (CO_i and CO_{i-1}) through the $^{13}C\alpha$-^{13}CO J-couplings. The

magnetization then evolves on the carbonyl carbons, resulting in a three-dimensional spectrum. The magnetization is then transferred back the same way without further evolution for detection: from CO_i and CO_{i-1} to $C\alpha_i$ and $C\alpha_{i-1}$, respectively, via the $^{13}C\alpha$-^{13}CO J-couplings; from $C\alpha_i$ and $C\alpha_{i-1}$ to $^{15}N_i$ via the ^{15}N-$^{13}C\alpha$ J-couplings; and finally, from $^{15}N_i$ to the amide proton through the ^{1}H-^{15}N J-coupling. The magnetization is then detected from the amide proton.

The HNCO and HN(CA)CO experiments allow for the assignment of the backbone carbonyl carbon atoms and sidechain amide carbonyl groups from asparagine and glutamine sidechains. The HNCO and HN(CA)CO spectra, resultant from their respective experiments, are traditionally displayed as ^{1}H-^{13}C planes, with the ^{1}H chemical shift along the x-axis and the ^{13}C chemical shift along the y-axis. The ^{15}N chemical shift axis (the z-axis) is variable such that the correlating ^{1}H-^{13}C planes (the z-axis) can be viewed for each collected ^{15}N frequency.

In the HNCO spectrum (Figure 23A), each vertical strip, representing a ^{1}H-^{15}N pair, has a single visible peak, which belongs to the carbonyl of the proceeding residue ($^{13}CO_{n-1}$). In the HN(CA)CO spectrum (Figure 23B), each vertical NH strip has two visible peaks. The more intense of these two peaks belongs to the carbonyl of the residue correlated to the amide NH pair ($^{13}CO_i$), and the less intense of these two peaks belongs to $^{13}CO_{i-1}$. When the HNCO and HN(CA)CO spectra are overlaid (Figure 23C), the peak in the HNCO spectrum overlays with the weaker of the two peaks in the HN(CA)CO spectrum, as they both represent the same carbonyl carbon. If the two peaks in the HN(CA)CO spectrum are of similar intensity, the HNCO spectrum makes it very easy to distinguish between CO_i and CO_{i-1} for each NH group.

3.2.1.4 The CBCA(CO)NH and HNCACB experiments

While more complex than the HNCO and HN(CA)CO experiments, the CBCA(CO)NH [67, 70, 178] and HNCACB [66, 178, 271] experiments also allow for assignment of resonances

in the protein. Specifically, these experiments allow for the assignment of the alpha and beta carbon atoms. Diagrams of the flow of the magnetization for these experiments are shown in Figure 24 and Figure 25, respectfully, and example spectra are shown in Figure 26.

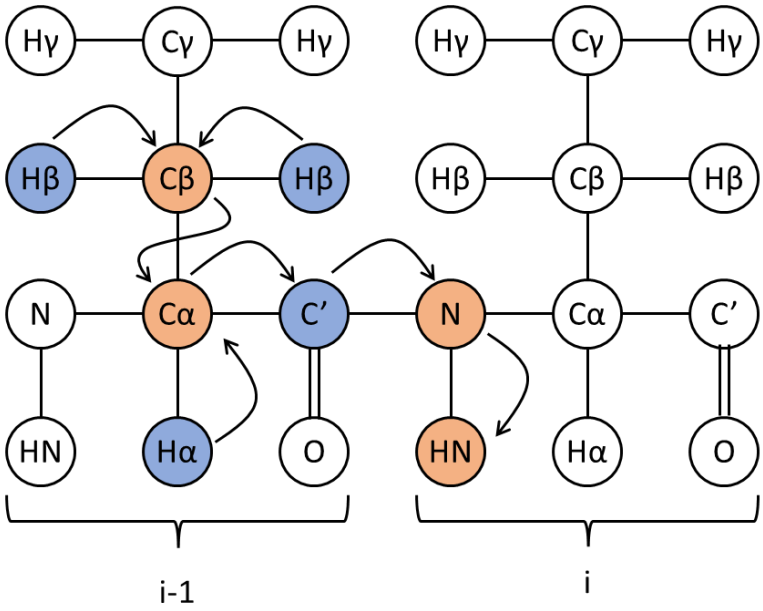

Figure 24. Magnetization flow diagram for the CBCA(CO)NH experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

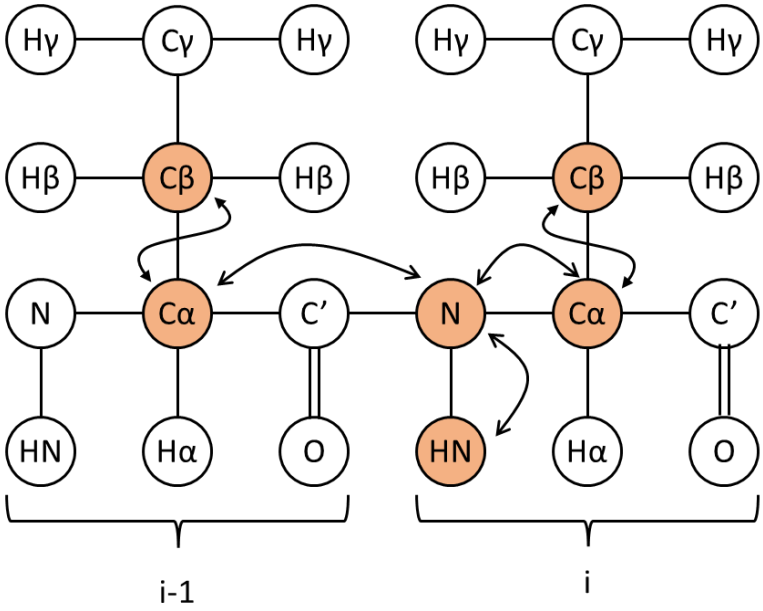

Figure 25. Magnetization flow diagram for the HNCACB experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

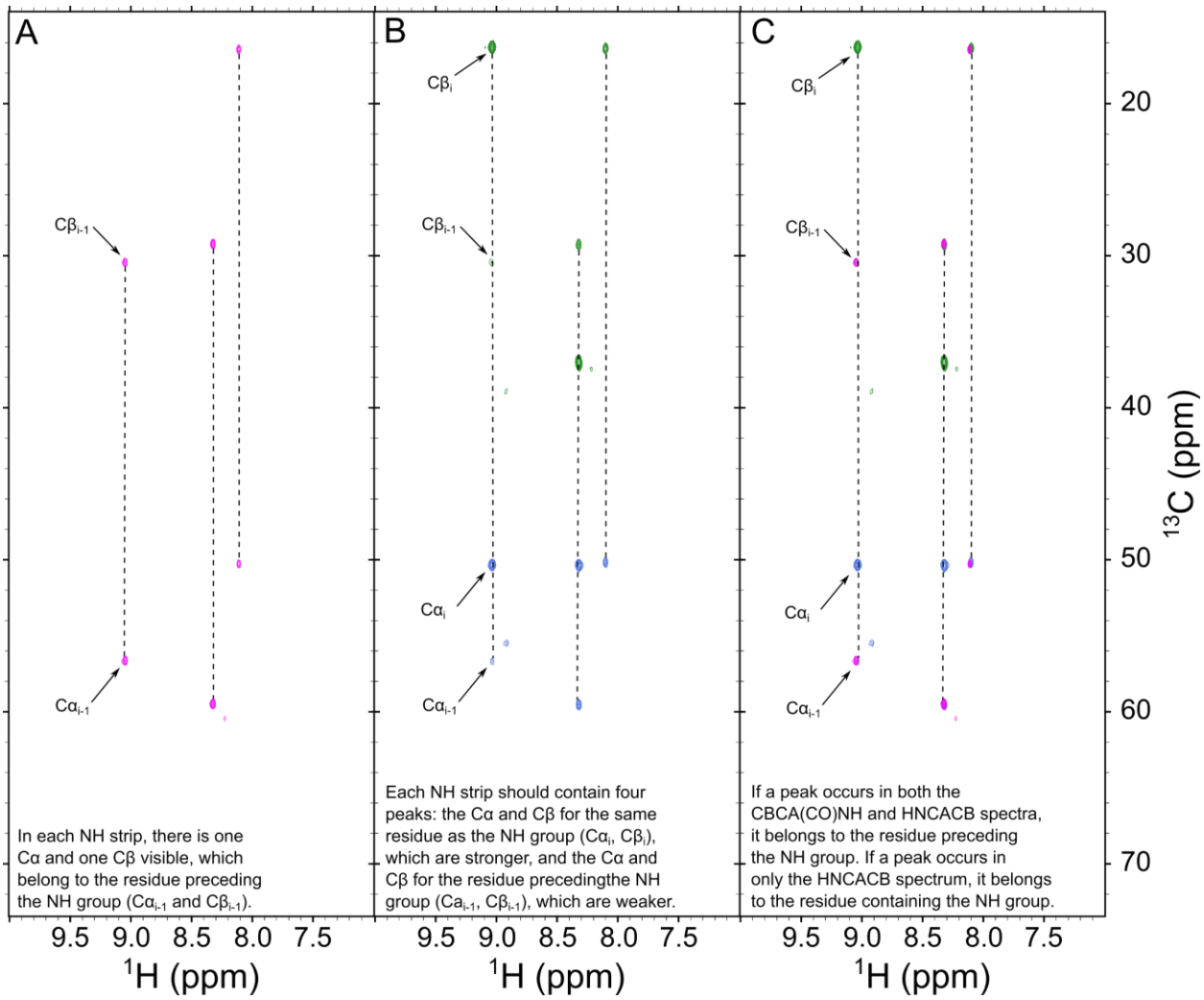

Figure 26. Example spectra for assigning the Cα and Cβ resonances. (A) Example spectrum for the CBCA(CO)NH experiment. (B) Example spectrum for the HNCACB. (C) Overlay of the example CBCA(CO)NH and HNCACB spectra. Data were collected for TGM6-D3 with a ^{15}N chemical shift of 122 ppm.

The CBCA(CO)NH experiment is different than the previously described experiments: the previously described experiments are referred to as "out-and-back" experiments, where the magnetization begins on and is detected from the same nucleus. In this experiment, magnetization begins on the alpha and beta protons of residue i-1 ($^{1}H\alpha_{i-1}$ and $^{1}H\beta_{i-1}$, respectively) and is immediately transferred to the attached alpha and beta carbons of residue i-1 ($^{13}C\alpha_{i-1}$ and $^{13}C\beta_{i-1}$, respectively) through the ^{1}H-^{13}C J-couplings such that there is no evolution on these protons. The

magnetization then evolves simultaneously on the $^{13}C\alpha_{i-1}$ and $^{13}C\beta_{i-1}$ nuclei, resulting in a three-dimensional spectrum with peaks for both the Cα and Cβ resonances in a single NH strip. Following this evolution, the magnetization from the $^{13}C\beta_{i-1}$ is transferred to the $^{13}C\alpha_{i-1}$ nucleus via the $^{13}C\alpha$-$^{13}C\beta$ J-couplings. The magnetization is then transferred from $^{13}C\alpha_{i-1}$ to $^{13}CO_{i-1}$ via the $^{13}C\alpha$-^{13}CO J-coupling and from $^{13}CO_{i-1}$ to $^{15}N_i$ via the ^{13}CO-^{15}N J-coupling without evolution on the ^{13}CO or further evolution on the $^{13}C\alpha$. The magnetization then evolves on the amide nitrogen prior to being transferred to the amide proton via the ^{1}H-^{15}N J-coupling for detection.

For the HNCACB experiment, magnetization begins on the amide proton of residue i and is transferred to the attached amide nitrogen ($^{15}N_i$) via the ^{1}H-^{15}N J-coupling. The magnetization evolves on the amide nitrogen prior to being transferred to the attached alpha carbons of both the excitation residue ($C\alpha_i$) and the proceeding residue ($C\alpha_{i-1}$) via the ^{15}N-$^{13}C\alpha$ J-couplings. The magnetization is partially transferred to the beta carbons of both the excitation residue ($C\beta_i$) and the proceeding residue ($C\beta_{i-1}$) via the $^{13}C\alpha$-$^{13}C\beta$ J-couplings. The magnetization evolves simultaneously on the Cα and Cβ atoms for each residue, resulting in a three-dimensional spectrum with peaks for both the Cα and Cβ resonances in a NH single strip. The magnetization is then transferred back via the reverse path without further evolution for detection: from $C\beta_i$ and $C\beta_{i-1}$ to $C\alpha_i$ and $C\alpha_{i-1}$, respectively, via the $^{13}C\alpha$-$^{13}C\beta$ J-couplings; from $C\alpha_i$ and $C\alpha_{i-1}$ to $^{15}N_i$ via the ^{15}N-$^{13}C\alpha$ J-couplings; and finally, from $^{15}N_i$ to the amide proton through the ^{1}H-^{15}N J-coupling. The magnetization is then detected from the amide proton.

The CBCA(CO)NH and HNCACB experiments allow for the assignment of the alpha and beta carbon atoms for each residue. The CBCA(CO)NH and HNCACB spectra, resultant from their respective experiments, are traditionally displayed as ^{1}H-^{13}C planes, with the ^{1}H chemical shift along the x-axis and the ^{13}C chemical shift along the y-axis. The ^{15}N chemical shift axis (the

z-axis) is variable such that the correlating ^{1}H-^{13}C planes can be viewed for each collected ^{15}N frequency.

In the CBCA(CO)NH spectrum (Figure 26A), each vertical strip, representing a ^{1}H-^{15}N pair, has up to two visible peaks, which belong to the alpha and beta carbons of the proceeding residue ($^{13}C\alpha_{n-1}$ and $^{13}C\beta_{n-1}$, respectively). In the HNCACB spectrum (Figure 26B), each vertical NH strip has up to four visible peaks. The more intense of these peaks belong to the alpha and beta carbons of the residue correlated to the amide NH pair ($^{13}C\alpha_{i}$ and $^{13}C\beta_{i}$, respectively), and the less intense of these two peaks belong to $^{13}C\alpha_{i-1}$ and $^{13}C\beta_{i-1}$, respectively. In the HNCACB spectrum, the peaks for the alpha and beta carbons have opposite signs: alpha carbons will be positive, and beta carbons will be negative, or vice-versa depending on the phasing. When the CBCA(CO)NH and HNCACB spectra are overlaid (Figure 26C), the peaks in the CBCA(CO)NH spectrum overlay with the weaker of the peaks in the HNCACB spectrum, as they both represent the same alpha and beta carbons. If the two peaks in the HNCACB spectrum are of similar intensity, the CBCA(CO)NH spectrum makes it very easy to distinguish between the peaks for the preceding residue and for the detected residue for each NH group.

3.2.1.5 The CC(CO)NH experiment

The CC(CO)NH experiment [25, 35, 68, 136, 141, 174] is similar to CBCA(CO)NH experiment in that it correlates the Cα and Cβ resonances of residue i-1 to the amide proton and nitrogen of residue i, but it also correlates all sidechain carbons with attached protons for residue i-1 to the same amide proton and nitrogen of residue i, giving additional information and assignment. A diagram of the flow of the magnetization for this experiment is shown in Figure 27, and an example spectrum is shown in Figure 28.

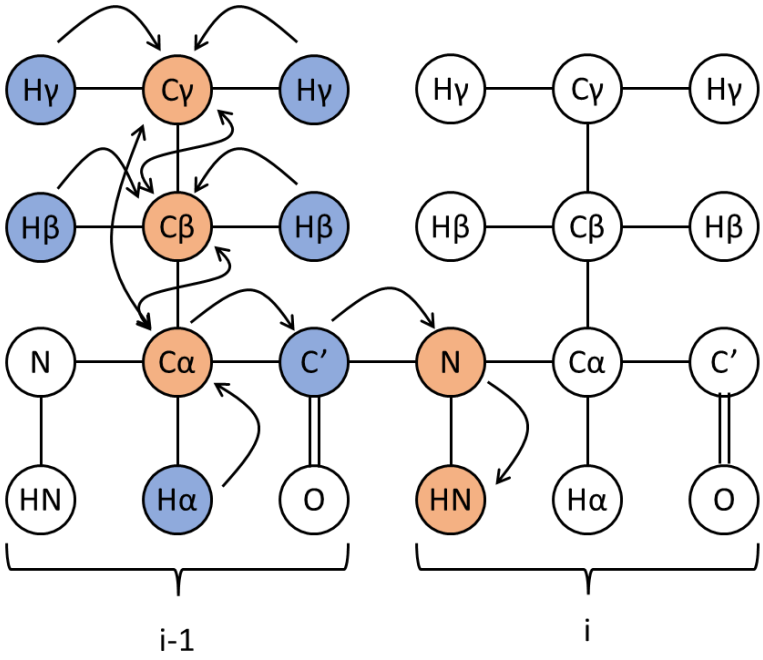

Figure 27. Magnetization flow diagram for the CC(CO)NH experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

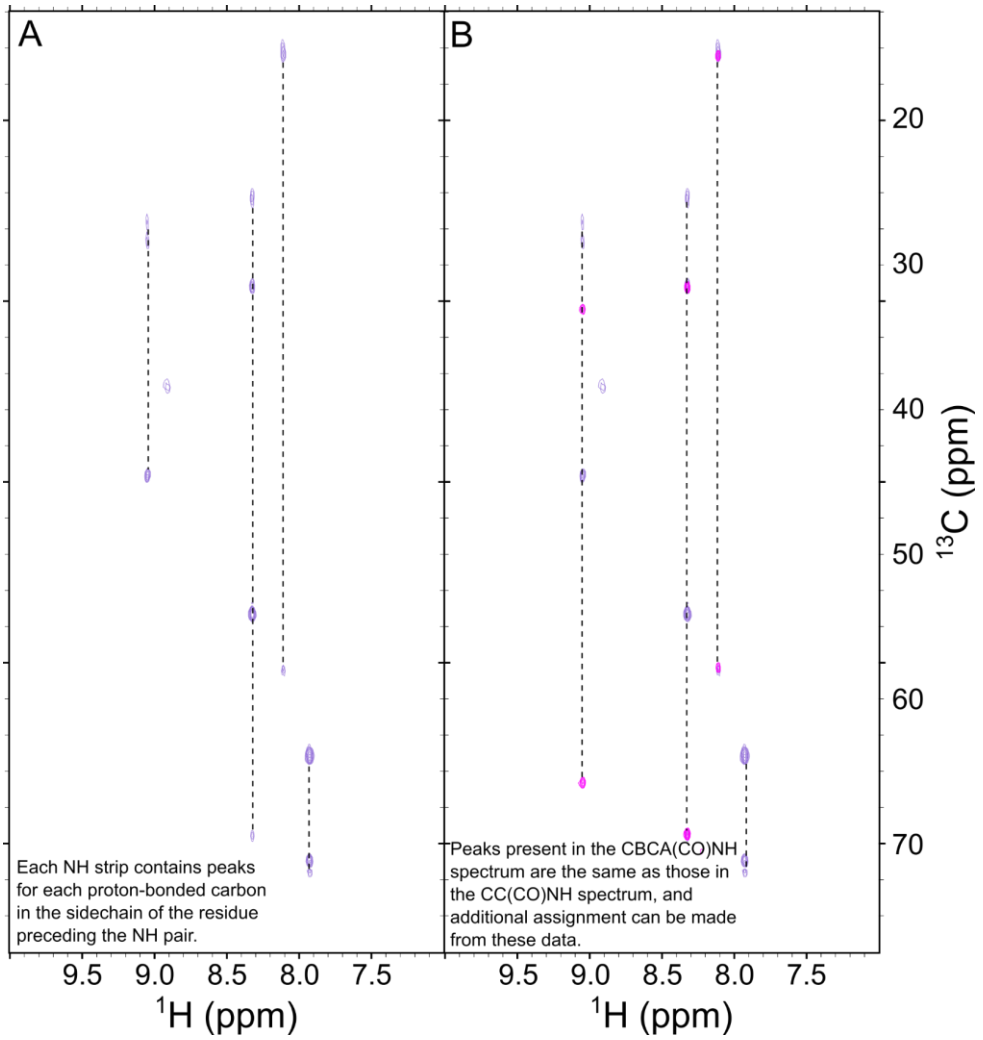

Figure 28. Example spectrum for the CC(CO)NH experiment. (A) Example spectrum for the CC(CO)NH experiment. (B) Overlay of the example CC(CO)NH and CBCA(CO)NH spectra. Data were collected for TGM6-D3

Similar to the CBCA(CO)NH experiment, in this experiment, magnetization begins on the aliphatic protons of residue i-1 ($^1H\alpha_{i-1}$, $^1H\beta_{i-1}$, $^1H\gamma_{i-1}$, etc.) and is immediately transferred to the attached carbons of residue i-1 ($^{13}C\alpha_{i-1}$, $^{13}C\beta_{i-1}$, $^{13}C\gamma_{i-1}$, etc., respectively) through the ^1H-^{13}C J-couplings such that there is no evolution on these protons. The magnetization then evolves on the carbon nuclei. Following this evolution, the magnetization from the carbon nuclei is transferred to the $^{13}C\alpha_{i-1}$ nucleus via the ^{13}C-^{13}C J-couplings. The magnetization is then transferred from $^{13}C\alpha_{i-1}$ to $^{13}CO_{i-1}$ via the $^{13}C\alpha$-^{13}CO J-coupling and from $^{13}CO_{i-1}$ to $^{15}N_i$ via the ^{13}CO-^{15}N J-coupling

without evolution on the ^{13}CO or further evolution on the $^{13}C\alpha$. The magnetization then evolves on the amide nitrogen prior to being transferred to the amide proton via the ^{1}H-^{15}N J-coupling for detection.

The CC(CO)NH experiment allows for the assignment of the aliphatic carbon atoms for each residue. The CC(CO)NH spectrum, resultant from the experiment, is traditionally displayed as ^{1}H-^{13}C planes, with the ^{1}H chemical shift along the x-axis and the ^{13}C chemical shift along the y-axis. The ^{15}N chemical shift axis (the z-axis) is variable such that the correlating ^{1}H-^{13}C planes can be viewed for each collected ^{15}N frequency.

In the CC(CO)NH spectrum (Figure 28A), each vertical strip, representing a ^{1}H-^{15}N pair, has multiple visible peaks, which belong to the aliphatic carbons bonded to a proton for the residue preceding the amide NH pair ($^{13}Cx_{n-1}$). After assignment of the Cα and Cβ resonances using the CBCA(CO)NH and HNCACB experiments, these assignments can be transferred to the CC(CO)NH spectrum (Figure 28B), and the remaining peaks can be assigned based on the characteristic frequency ranges in which the carbons appear.

3.2.1.6 The HBHA(CO)NH and H(CC)(CO)NH experiments

Similar to the CBCA(CO)NH and CC(CO)NH experiments, the HBHA(CO)NH [70, 178] and H(CC)(CO)NH [25, 35, 68, 136, 141, 174] experiments allow for assignment of proton resonances in the protein. Specifically, these experiments allow for the assignment of the aliphatic protons in the protein sidechains. Diagrams of the flow of the magnetization for these experiments are shown in Figure 29 and Figure 30, respectfully, and example spectra are shown in Figure 31.

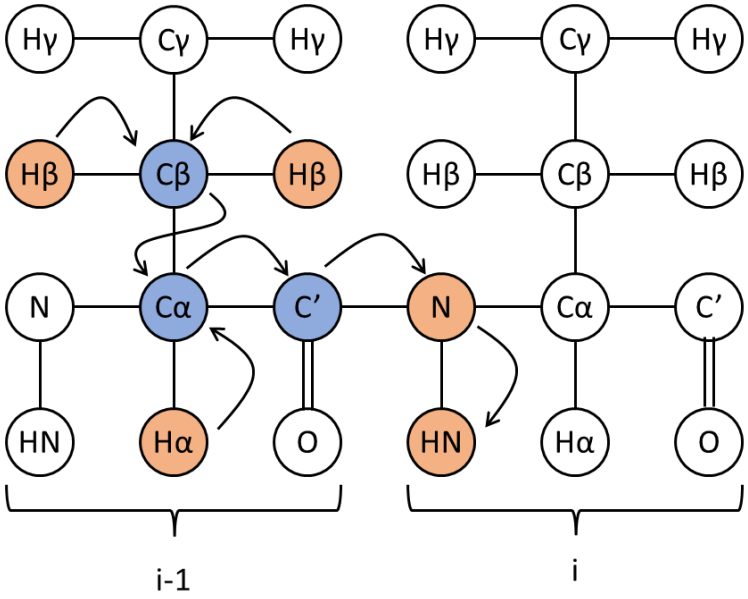

Figure 29. Magnetization flow diagram for the HBHA(CO)NH experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

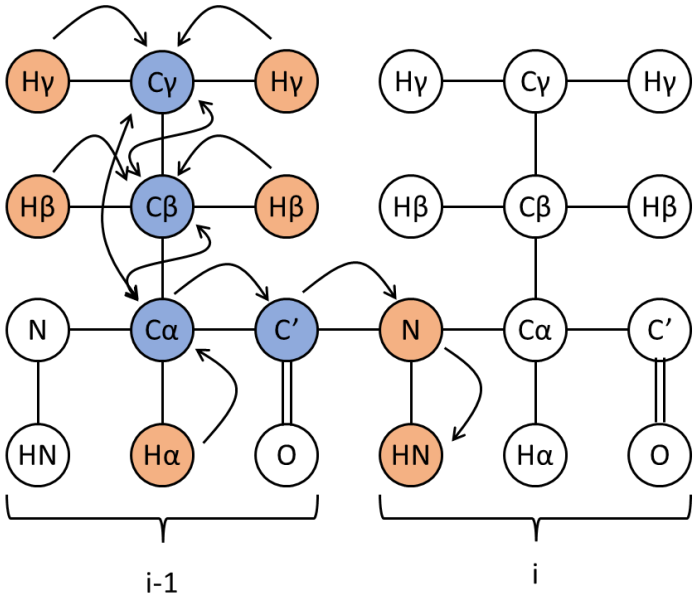

Figure 30. Magnetization flow diagram for the H(CC)(CO)NH experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

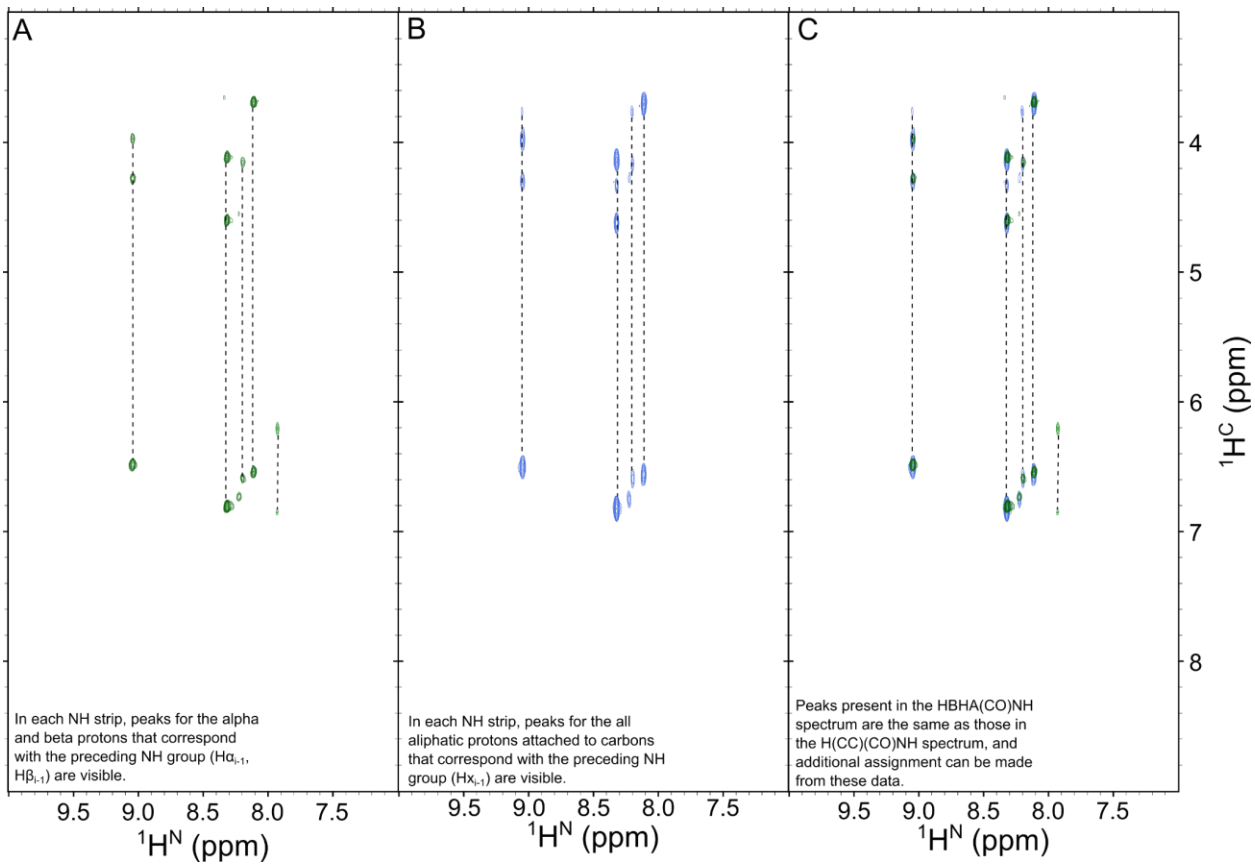

Figure 31. Example spectra for assigning the aliphatic proton resonances. (A) Example spectrum for the HBHA(CO)NH experiment. (B) Example spectrum for the H(CC)(CO)NH experiment. (C) Overlay of the example HBHA(CO)NH and H(CC)(CO)NH spectra. Data were collected for TGM6-D3.

The HBHA(CO)NH experiment is similar to the CBCA(CO)NH experiment: in the CBCA(CO)NH experiment, the spectra correlate the sidechain alpha and beta carbons to the amide nitrogen and proton in the following residue; in the HBHA(CO)NH experiment, the spectra correlate the sidechain alpha and beta protons to the amide nitrogen and proton in the following residue. In this experiment, magnetization begins on the alpha and beta protons of residue i-1 ($^1H\alpha_{i-1}$ and $^1H\beta_{i-1}$, respectively) and is allowed to evolve. The magnetization is then transferred to the attached alpha and beta carbons of residue i-1 ($^{13}C\alpha_{i-1}$ and $^{13}C\beta_{i-1}$, respectively) through the ^1H-^{13}C J-couplings and then is immediately transferred to the $^{13}C\alpha_{i-1}$ nucleus such that there is no evolution on these carbons. The magnetization is then transferred from $^{13}C\alpha_{i-1}$ to $^{13}CO_{i-1}$ via the

$^{13}C\alpha$-^{13}CO J-coupling and from $^{13}CO_{i-1}$ to $^{15}N_i$ via the ^{13}CO-^{15}N J-coupling without evolution on the $^{13}CO_{i-1}$. The magnetization then evolves on the amide nitrogen prior to being transferred to the amide proton via the ^{1}H-^{15}N J-coupling for detection.

Similar to the HBHA(CO)NH experiment, in H(CC)(CO)HN experiment, magnetization begins on the aliphatic protons of residue i-1 ($^{1}H\alpha_{i-1,}$ $^{1}H\beta_{i-1}$, $^{1}H\gamma_{i-1}$, etc.) and is allowed to evolve on these nuclei. The magnetization is then transferred to the attached carbons of residue i-1 ($^{13}C\alpha_{i-1}$, $^{13}C\beta_{i-1}$, $^{13}C\gamma_{i-1}$, etc., respectively) through the ^{1}H-^{13}C J-couplings and is then immediately transferred to the $^{13}C\alpha_{i-1}$ nucleus such that there is no evolution on these carbons. The magnetization is then transferred from $^{13}C\alpha_{i-1}$ to $^{13}CO_{i-1}$ via the $^{13}C\alpha$-^{13}CO J-coupling and from $^{13}CO_{i-1}$ to $^{15}N_i$ via the ^{13}CO-^{15}N J-coupling without evolution on the $^{13}CO_{i-1}$. The magnetization then evolves on the amide nitrogen prior to being transferred to the amide proton via the ^{1}H-^{15}N J-coupling for detection.

The HBHA(CO)NH and H(CC)(CO)NH experiments allow for the assignment of the sidechain protons for each residue. The HBHA(CO)NH and H(CC)(CO)NH spectra, resultant from their respective experiments, are traditionally displayed as $^{1}H^{N}$-$^{1}H^{C}$ planes, with the amide proton ($^{1}H^{N}$) chemical shift along the x-axis and the aliphatic proton ($^{1}H^{C}$) chemical shift along the y-axis. The ^{15}N chemical shift axis (the z-axis) is variable such that the correlating $^{1}H^{N}$-$^{1}H^{C}$ planes can be viewed for each collected ^{15}N frequency.

In the HBHA(CO)NH spectrum (Figure 31A), each vertical strip, representing a $^{1}H^{N}$-^{15}N pair has up to three visible peaks, which belong to the alpha and beta carbons of the residue that precedes the NH pair ($^{1}H\alpha_{n-1}$ and $^{1}H\beta_{n-1}$, respectively). In the H(CC)(CO)NH spectrum (Figure 31B), each vertical strip, representing a $^{1}H^{N}$-^{15}N pair, has multiple visible peaks, which belong to the aliphatic protons bonded to a carbon for the residue preceding the amide NH pair ($^{1}Hx_{n-1}$).

After assignment of the $H\alpha_{i-1}$ and $H\beta_{i-1}$ resonances using the HBHA(CO)NH experiment, these assignments can be transferred to the H(CC)(CO)NH spectrum (Figure 31C), and the remaining peaks can be assigned based on the characteristic frequency ranges in which the protons appear.

3.2.1.7 The HCCH-TOCSY experiment

The HCCH-TOCSY experiment [13, 113, 191] is a three-dimensional experiment that correlates the proton-carbon pairs within each residue with the other proton-carbon pairs in the same residue. A diagram of the flow of the magnetization for this experiment is shown in Figure 32, and an example spectrum is shown in Figure 33.

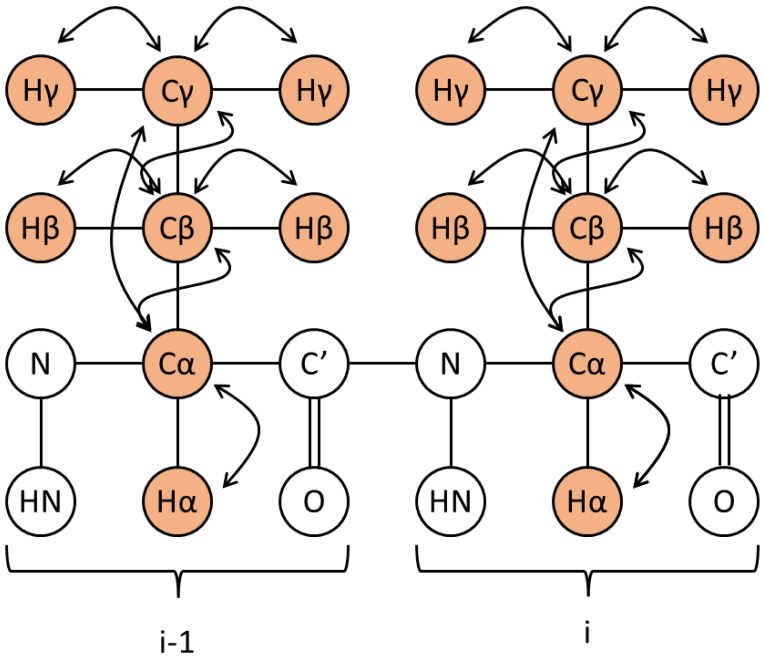

Figure 32. Magnetization flow diagram for the HCCH-TOCSY experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

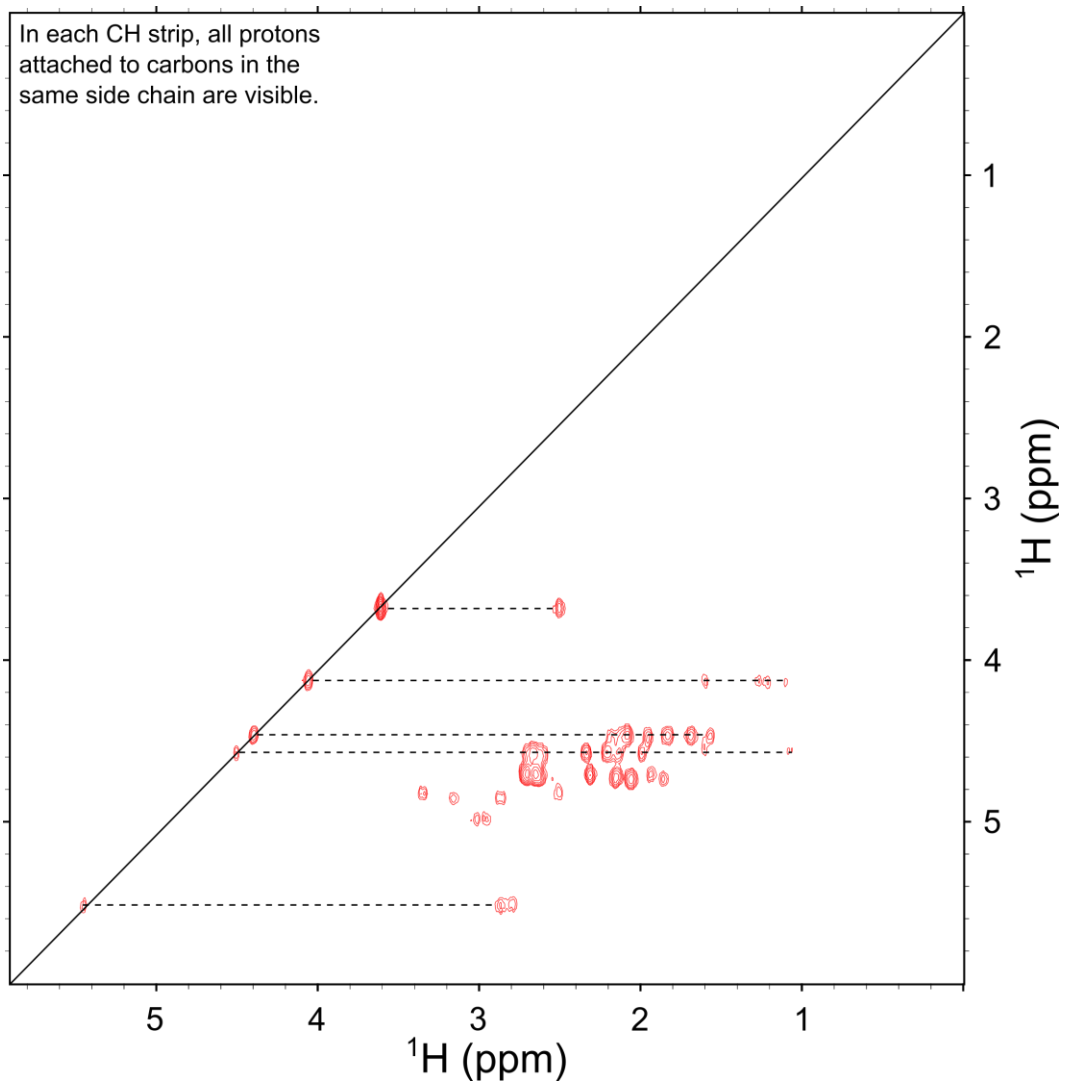

Figure 33. Example spectrum for the HCCH-TOCSY experiment. Data were collected for TGM6-D3. Note: the data shown in this figure are processed such that peaks belonging to one sidechain appear in a horizontal strip.

Magnetization begins on the sidechain protons, where it evolves. The magnetization is then transferred to the attached carbon nuclei and immediately undergoes isotropic mixing between the ^{13}C nuclei. Following mixing, the magnetization evolves on the ^{13}C nuclei and is then transferred back to the attached protons for detection.

This spectrum is used for side-chain assignment. It yields strips at each carbon frequency in the sidechain in which all sidechain proton resonances are visible. Note that the spectrum is

usually displayed in such a way that the detected ^{1}H dimension is shown along the y-axis rather than the x-axis. This means that the peaks belonging to one sidechain appear in a vertical strip rather than a horizontal strip.

3.2.1.8 The CB(CGCD)HD and CB(CGCDCE)HE experiments

The CB(CGCD)HD and CB(CGCDCE)HE experiments [279] generate two-dimensional NMR spectra that correlate the Cβ with the Hδ and the Cβ with the Hε, respectively, within the same residue. These are primarily used for assigning aromatic sidechains, whose ring resonances are not observed in the H(CC)(CO)NH or HCCH-TOCSY spectra. Diagrams of the flow of the magnetization for these experiments are shown in Figure 34 and Figure 35, respectfully.

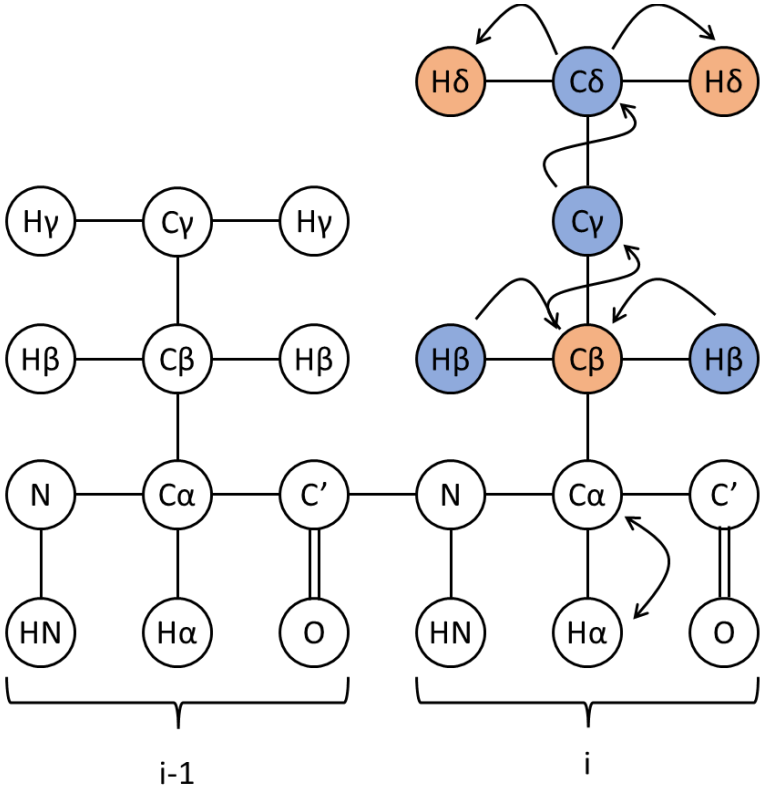

Figure 34. Magnetization flow diagram for the CB(CGCD)HD experiment. Note the missing protons attached to Cγ.

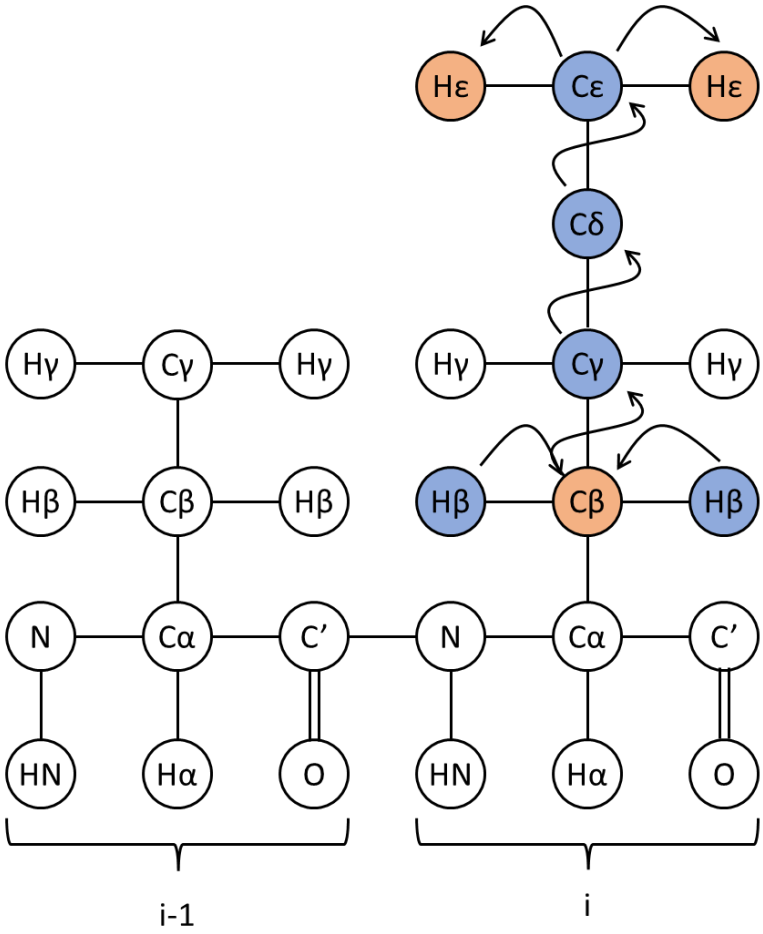

Figure 35. Magnetization flow diagram for the CB(CGCDCE)HE experiment. Note the missing protons attached to Cδ.

The pulse sequences for these experiments are very similar. Magnetization begins on the beta protons (Hβ) and is immediately transferred to the beta carbon. The magnetization does not evolve on the Hβ, but it is allowed to evolve on the Cβ. Following evolution, the magnetization is transferred to the gamma carbon (Cγ) followed by the delta carbon (Cδ); for the CB(CGCDCE)HE experiment, magnetization is further transferred to the epsilon carbon (Cε). Evolution of the magnetization does not occur on any of these carbon atoms. Finally, the magnetization is transferred to the delta protons (Hδ) for the CB(CGCD)HD experiment or the epsilon protons (Hε) for the CB(CGCDCE)HE experiment where the magnetization is detected.

As stated previously, the CB(CGCD)HD and CB(CGCDCE)HE experiments generate two-dimensional NMR spectra that correlate the Cβ with the Hδ and the Cβ with the Hε, respectively, within the same residue. The structures of the sidechains of phenylalanine, tryptophan, and tyrosine residues are such that there is a discontinuous sequence of protons for the transfer of magnetization that is used in the H(CC)(CO)NH and HCCH-TOCSY experiments. Therefore, the aromatic ring resonances cannot be correlated with the backbone resonances for these residues. The CB(CGCD)HD and CB(CGCDCE)HE experiments transfer magnetization through the carbon chain, not relying on the protons for correlation of the resonances. Therefore, these experiments can provide a link between the backbone (more specifically, the Cβ just off the backbone) and the aromatic components of Phe, Trp, and Tyr residues, allowing for identification and assignment of these aromatic resonances.

3.2.2 Resonance assignment

Once the appropriate data are collected, such as those from the experiments described in the previous section, the next step is to assign the peaks to the appropriate atoms in the protein. The following sections describe the processes for manual and automatic assignment of these peaks.

3.2.2.1 Manual resonance assignment

The general workflow of assigning the peaks in NMR spectra is shown in Figure 36.

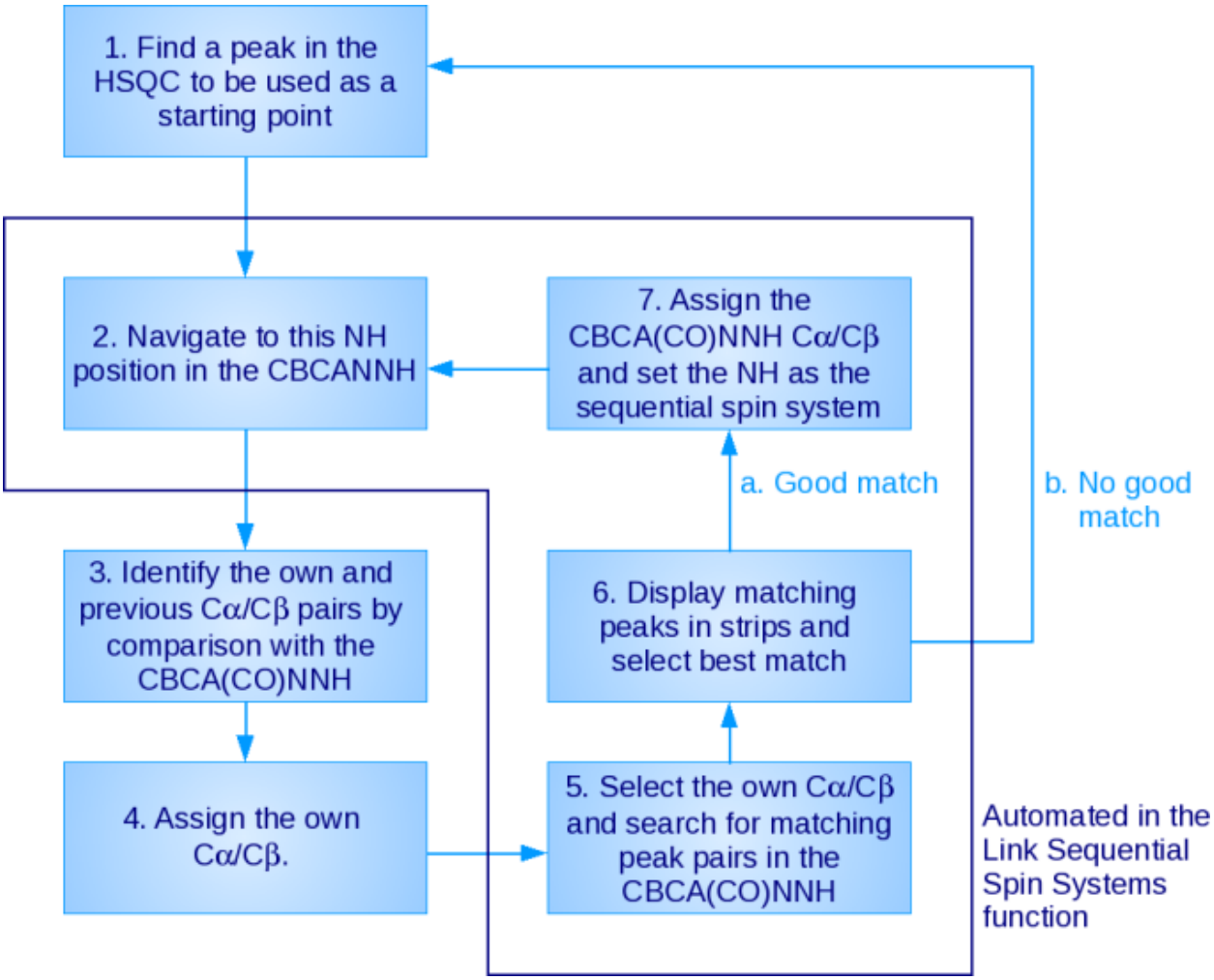

Figure 36. General workflow for manual assignment of protein resonances. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

Standard triple resonance backbone assignment of proteins is based on the HNCACB and CBCA(CO)NH spectra. The idea is that the HNCACB correlates each NH group with the Cα and Cβ chemical shifts of its own residue (strongly) and of the residue preceding (weakly). The CBCA(CO)NH only correlates the NH group to the preceding Cα and Cβ chemical shifts. The Figure below (Figure 37) shows how this can be used to link one NH group to the next into a long chain.

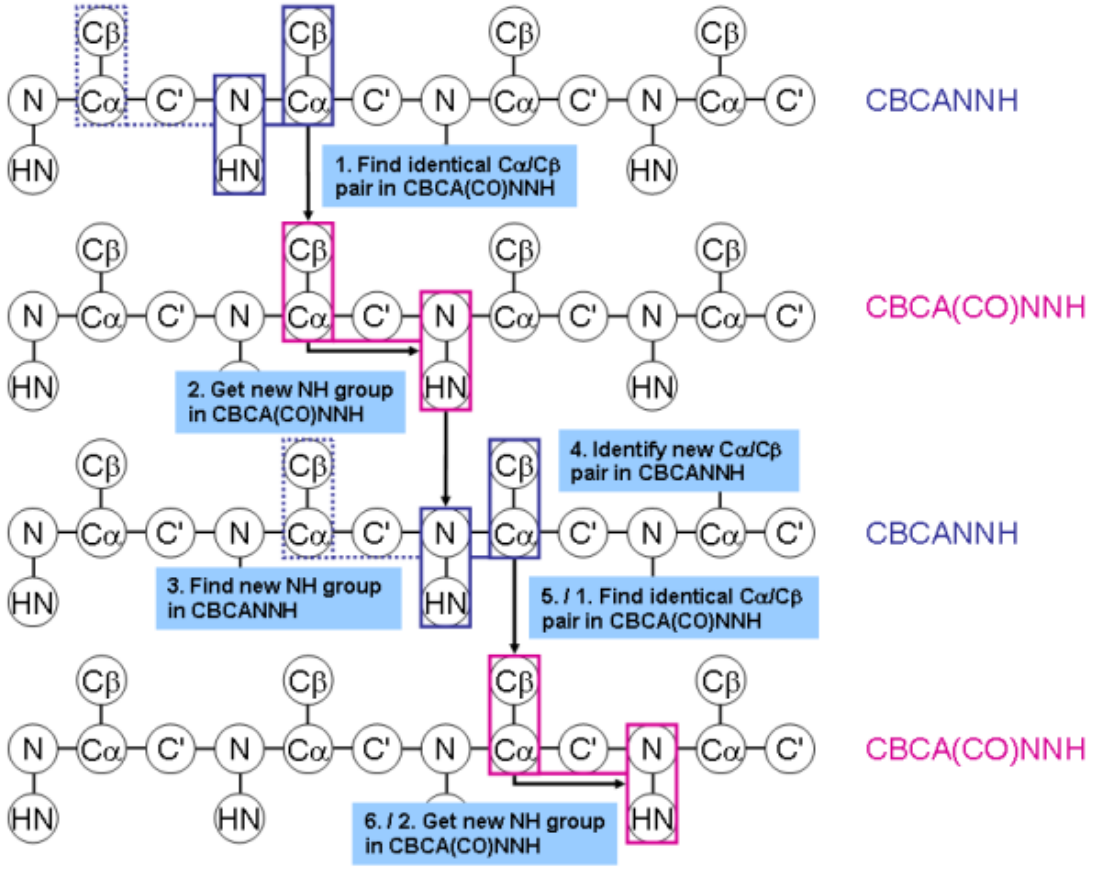

Figure 37. Workflow for identification and correlation of chemical shifts between the HNCACO and CBCA(CO)NH experiments. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

In practice, using the HNCACB and CBCA(CO)NH spectra looks like this (Cαs are shown in dark blue, Cβs in light blue) (Figure 38):

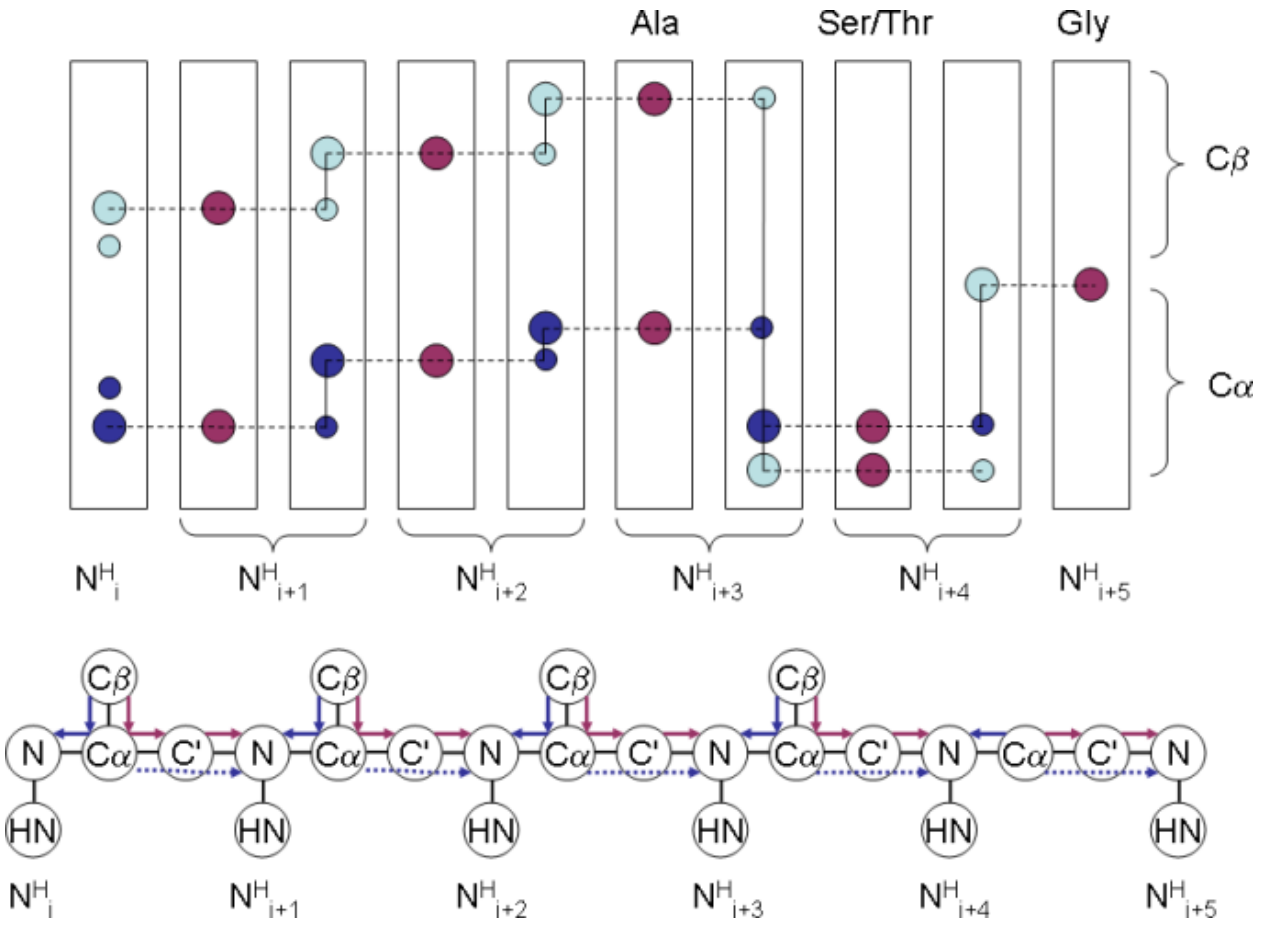

Figure 38. Spectral strips for identification and correlation of chemical shifts between the HNCACO and CBCA(CO)NH experiments. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

Alternatively, some software packages (such as CCPNmr Analysis) allow for two spectra to be superimposed, where strips will look like this (Figure 39):

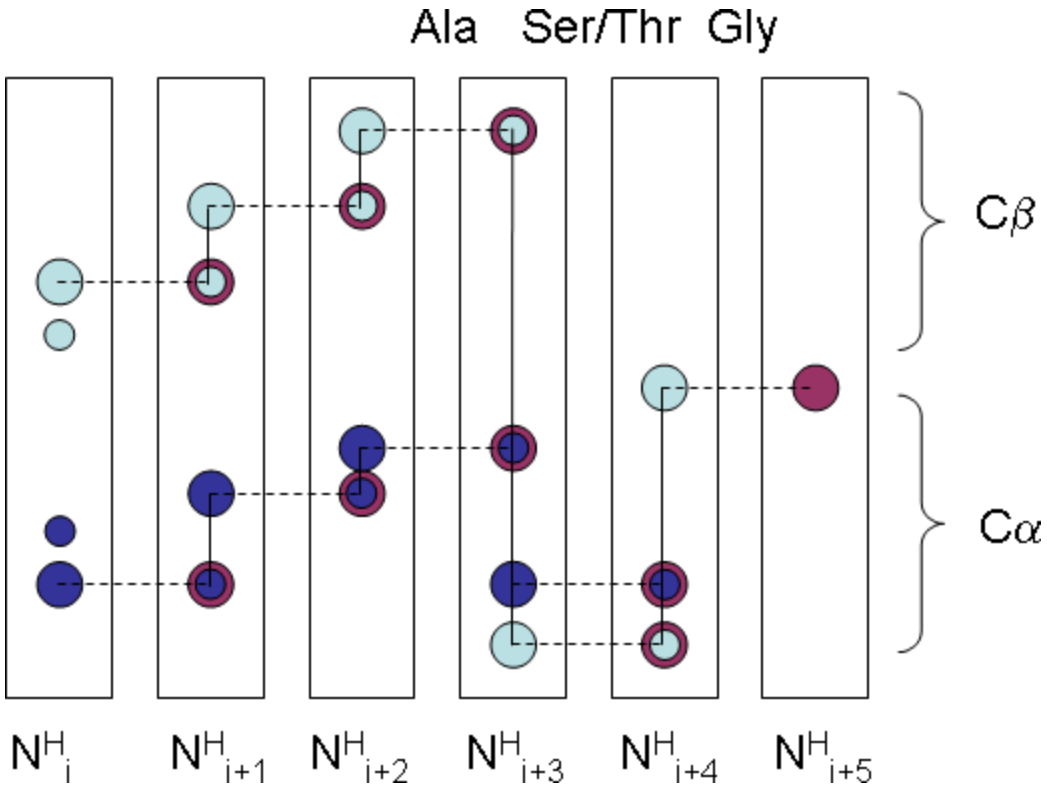

Figure 39. Superimposed spectral strips for identification and correlation of chemical shifts between the HNCACO and CBCA(CO)NH experiments. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

The Cα and Cβ chemical shifts adopt values characteristic of the amino acid type. Some of these, such as Alanine, Serine, Threonine and Glycine are very easy to spot as their Cβ chemical shifts are very different to those of the other amino acids (and in the case of Glycine there is no Cβ). Valine, Isoleucine and Proline are also likely to stand out by the fact that they have lower than normal Cα chemical shifts. Once a chain of NH groups with their corresponding Cα and Cβ chemical shifts has been built, then the identification of some of the amino acid types makes it possible to match this string to the sequence. For example: a string of shifts may have been found that corresponds to xxxSxxAx – if this sequence only appears once in the sequence of the protein in question, then sequence-specific assignment can be made. This process is referred to as

"walking the backbone," and should be iterated until all resonances and/or amino acids in the protein sequence are assigned.

In some cases, in particular if the protein is fairly large (>200 residues, for example), the quality of the HNCACB and CBCA(CO)NH spectra may not be very good. The Cβ resonances, for example, may not be visible above the noise level. In this case it is possible to use the Cα and CO chemical shifts rather than the Cα and Cβ chemical shifts, as those which you use to walk from one residue to the next. The HNCA and HN(CO)CA experiments give you the same information as the HNCACB and CBCA(CO)NH spectra, except without the Cβ resonances. To complement this, the HNCO and HN(CA)CO experiments can be used. These link each NH_i group with the CO_{i-1} (HNCO) or with CO_i and CO_{i-1} (HN(CA)CO). However, the small ^{15}N-$^{13}C\alpha$ one-bond J-coupling leads to long J-coupling transfer times in the experiment; this means that the HN(CA)CO experiment is not very sensitive, especially compared to the CBCA(CO)NH and HN(CO)CA experiments, and typically requires long overall acquisition times.

Using the HNCA, HN(CO)CA, HNCO, and HN(CA)CO experiments, the residues are now linked up in the following manner (Figure 40):

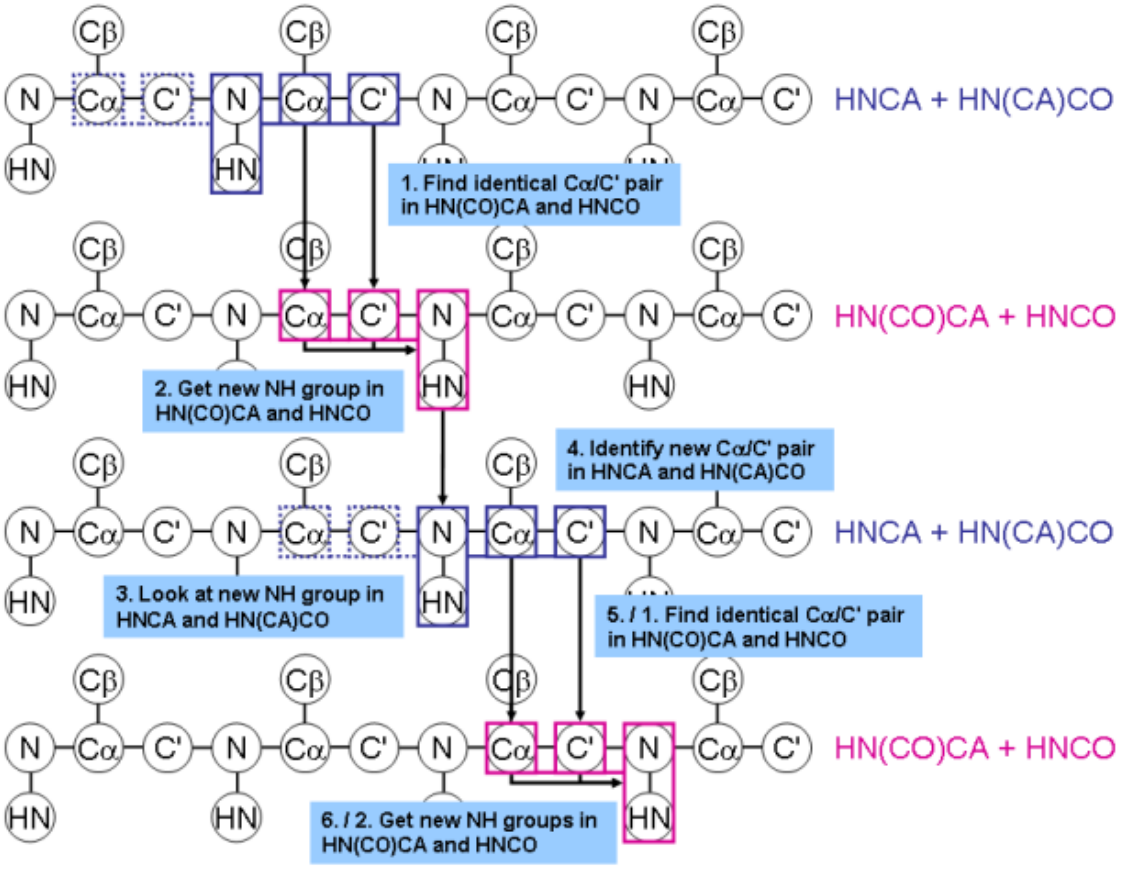

Figure 40. Workflow for identification and correlation of chemical shifts between the HNCA, HN(CA)CO, HN(CO)CA, and HNCO experiments. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

The advantage of using the HNCO and HNCA-based spectra is that they are more sensitive than the HNCACB-type and thus the spectral quality should improve. The disadvantage is that the Cα and CO chemical shifts provide less information about the amino acid type than the Cβ chemical shift and are less disperse.

Once the backbone atoms are assigned, the sidechains can be assigned. Various methods and spectra are available for this task; the chosen experiment(s) depend on the size of the protein and the amount of spectrometer time available as to what spectra are used. A straight forward

method is to begin with a set of HBHA(CO)NH, H(CC)(CO)NH, and CC(CO)NH spectra. These will provide the hydrogen and carbon side-chain chemical shifts for the residue preceding each NH group. For longer side chains not all peaks may necessarily be visible, so that this may not be sufficient. In some cases, it may also be difficult to distinguish between Hβ and Hγ shifts. Furthermore, the connectivity of which hydrogen is attached to which carbon is also not provided. This is, for instance, relevant in the case of Valines (Figure 41) where there are two methyl groups: it may be possible to identify both methyl carbon and both methyl hydrogen chemical shifts, but it will not be known which is bound to which.

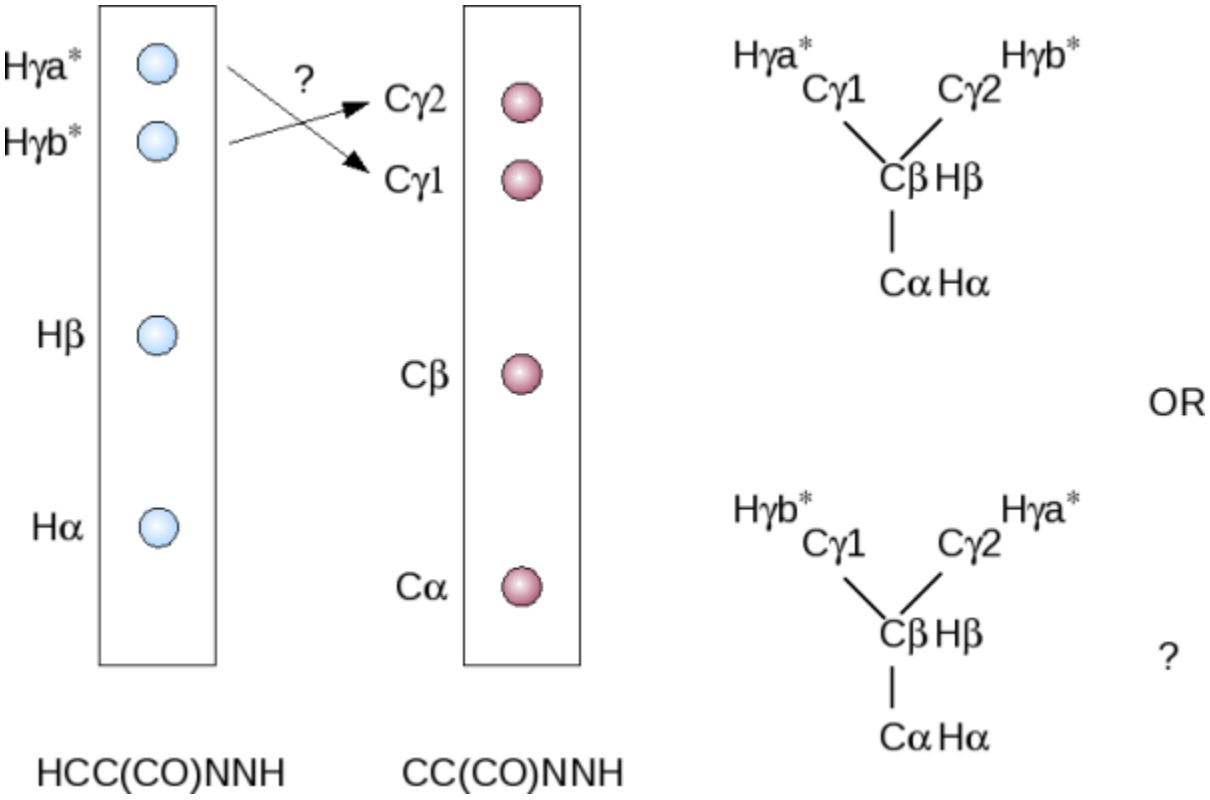

Figure 41. Example of proton and carbon resonance assignment in valine sidechains using C(CC)(CO)NH and CC(CO)NH experiments. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

The most useful spectrum for side-chain assignment is an HCCH-TOCSY spectrum (Figure 42). The HCCH-TOCSY will, at any one carbon position, show in one dimension the chemical shift of the hydrogen which is attached to the carbon and in another the other hydrogens belonging to that side chain. There is thus a huge amount of information in this spectrum and for large proteins it may become rather crowded. The following figure (Figure 42) shows the strips you should be able to see for a Valine residue in the HCCH-TOCSY spectrum.

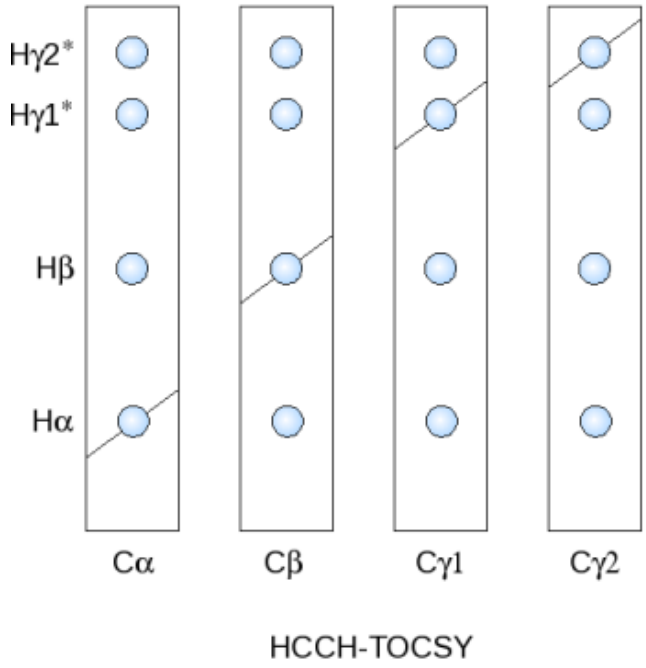

Figure 42. Example of HCCH-TOCSY strips for valine sidechains. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

The general principle behind using the HCCH-TOCSY spectrum is as follows (Figure 43): Using the known Cα and Cβ chemical shifts from the backbone assignment, go to these points in the spectrum. From this, the Hα and Hβ chemical shifts can be immediately identified by finding strips at each carbon shift which have peaks at the same hydrogen ppm values. Additional peaks

for the Hγ and Hδ atoms (if present in that particular amino acid type) should also be visible. By navigating to these new hydrogen shifts, the shifts of the carbons to which they are attached should also be identifiable.

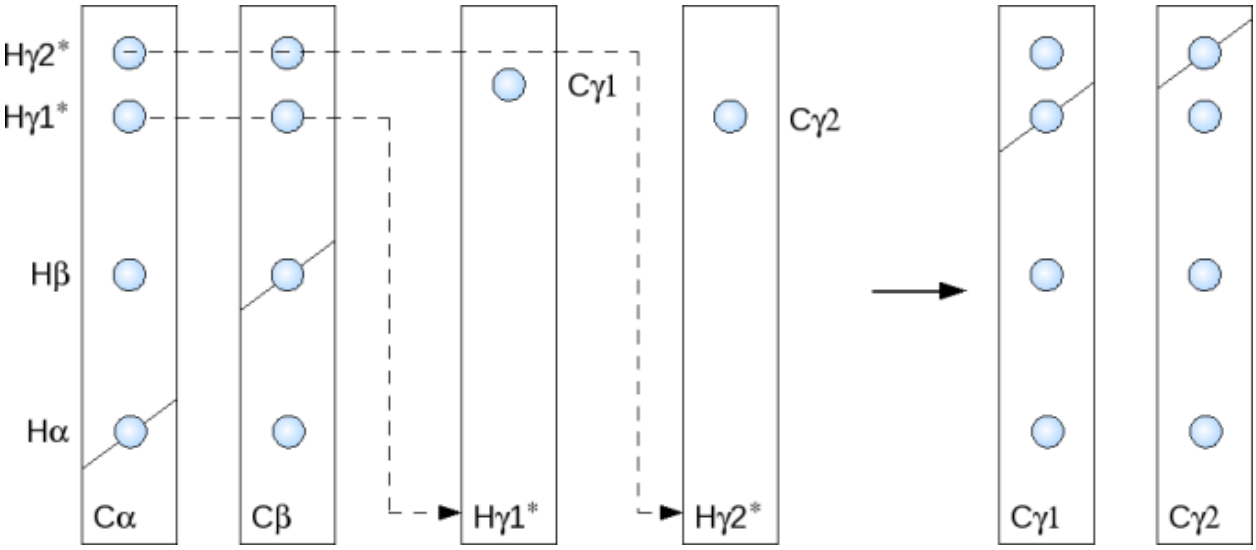

Figure 43. General process of assigning HCCH-TOCSY strips for valine sidechains. Figure used with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

While the previous example was specific to valine sidechains, the same process can and should be applied to all sidechains in the protein.

3.2.2.2 Automated resonance assignment

While manual assignment is arguably more reliable than automated assignment, automated assignment of the resonances is much faster than manual assignment. Various methods exist for automated assignment. The probabilistic interaction network of evidence (PINE) algorithm for automated assignment of backbone and side chain chemical shifts is one of the more popular and reliable methods for automated assignment. PINE utilizes a Bayesian probabilistic network model

that analyzes sequence data and peak lists from multiple NMR experiments to assign individual peaks in NMR spectra to the atoms of a specific protein [9].

As of 2019, the PINE software suite has been updated and upgraded, which the authors call the Integrative PINE, or I-PINE, and includes a freely accessible webserver interface. I-PINE supports more types of NMR experiments than PINE (including three-dimensional nuclear Overhauser enhancement and four-dimensional J-coupling experiments) along with more comprehensive visualization of chemical shift-based analysis of protein structure and dynamics. The I-PINE server is freely accessible at http://i-pine.nmrfam.wisc.edu [129].

3.3 Assignment of TGM6-D3

The previously described NMR experiments and methods were used to assign the backbone and sidechain carbon and proton resonances of TGM6-D3. These assignments are shown in the ^{1}H-^{15}N HSQC spectrum in Figure 44.

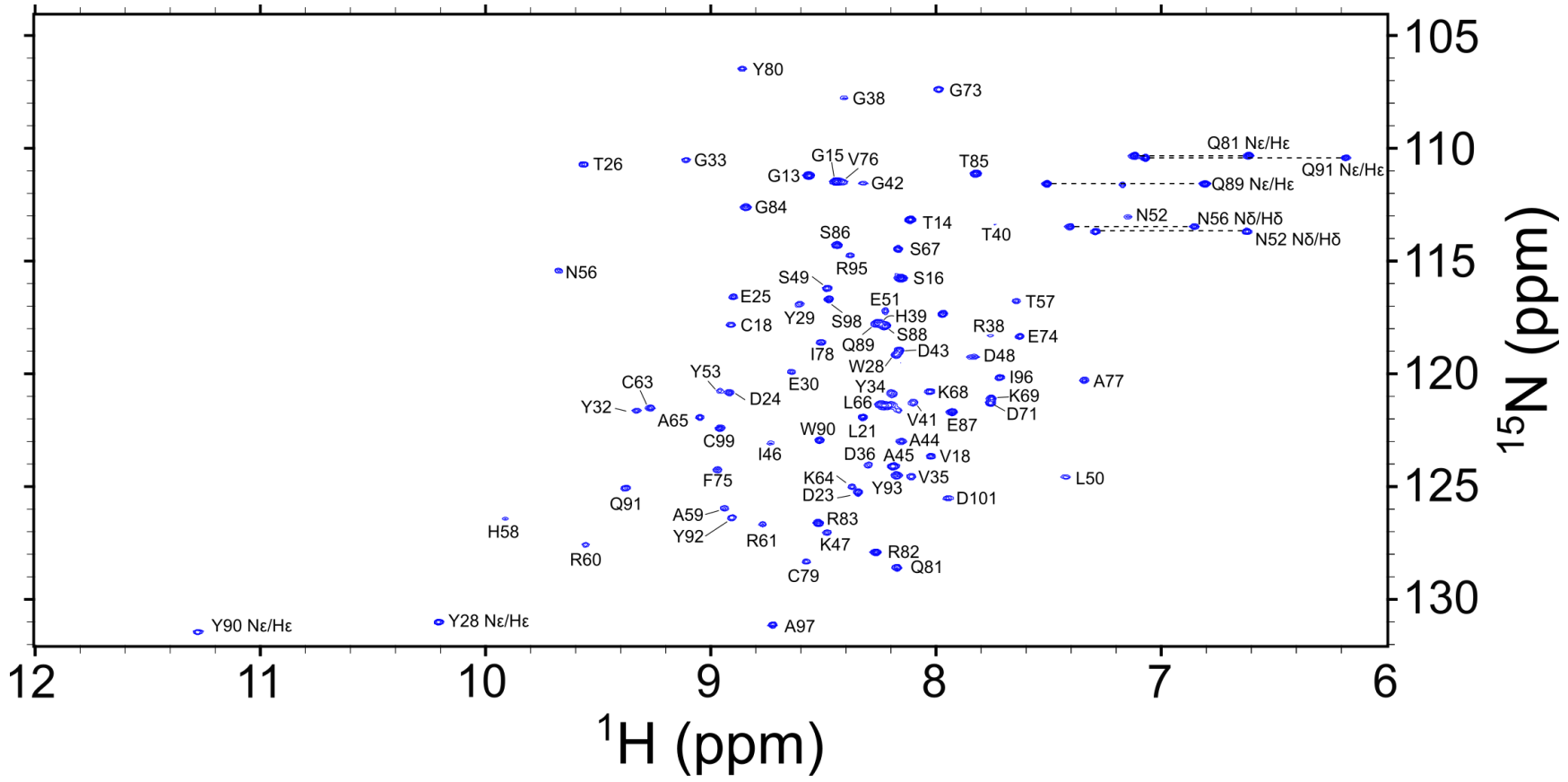

Figure 44. Resonance assignments for the TGM6-D3 ^{1}H-^{15}N HSQC spectrum.

3.4 Toward Structure Determination

In this section, the process of determining restraints for NMR structure determination is discussed. While this process was started for determining the solution structure of TGM6-D3, this work was not completed.

3.4.1 Three-bond J-couplings

J-couplings, or scalar couplings, are indirect dipolar couplings between two nuclei that are mediated through chemical bonds. A commonly used J-coupling is the heteronuclear one-bond amide ^{1}H-^{15}N J-coupling. The specific values of different one-bond J-couplings are shown in Figure 45.

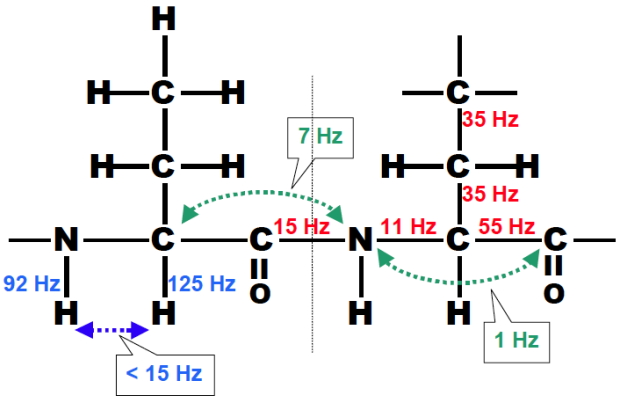

Figure 45. J-coupling constants for protein samples. Figure provided courtesy of A. Hinck.

J-couplings are not limited to one bond. Three-bond J-couplings are present in protein samples between the amide proton and the alpha proton (H^{N}-H^{α}) and between the amide nitrogen and the beta proton (N^{H}-H^{β}). Additionally, these three-bond J-couplings can be used as restraints in protein structure determination. Work from the Bax Lab [261, 262] showed that the value of the

three-bond H^N-H^α J-coupling is quantitative for the backbone dihedral ϕ angle according to the Karplus curve:

$$J = A\cos^2(\phi - 60) + B\cos(\phi - 60) + C \quad [1]$$

where A = 6.51, B = -1.76, C = 1.60. On the other hand, the three-bond N^H-H^β J-coupling can be related to the dihedral angle χ_1 in the regular manner [5]:

$$J = A\cos^2(\chi_1) + B\cos(\chi_1) + C \quad [2]$$

where A = 9.5 ± 0.3, B = -1.6 ± 0.2, C = 1.8 ± 0.6.

3.4.2 J-coupling NMR experiments

3.4.2.1 The HNHA and HNHB experiments

The HNHA [261, 262] and HNHB [5, 44, 137] experiments allow for quantitation of the homonuclear three-bond H^N-H^α and $^{15}N^H$-H^β J-coupling constants, respectively. Diagrams of the flow of the magnetization for these experiments are shown in Figure 46 and Figure 47, respectfully, and example spectra are shown in Figure 48.

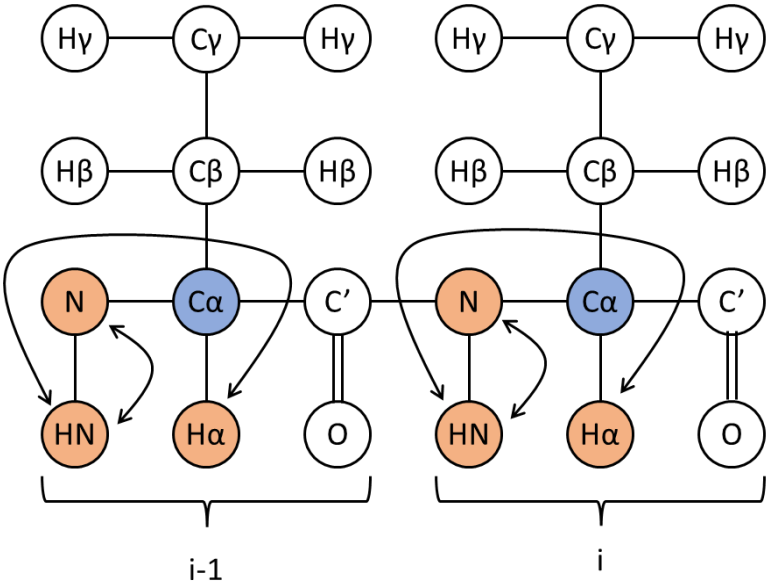

Figure 46. Magnetization flow diagram for the HNHA experiment.

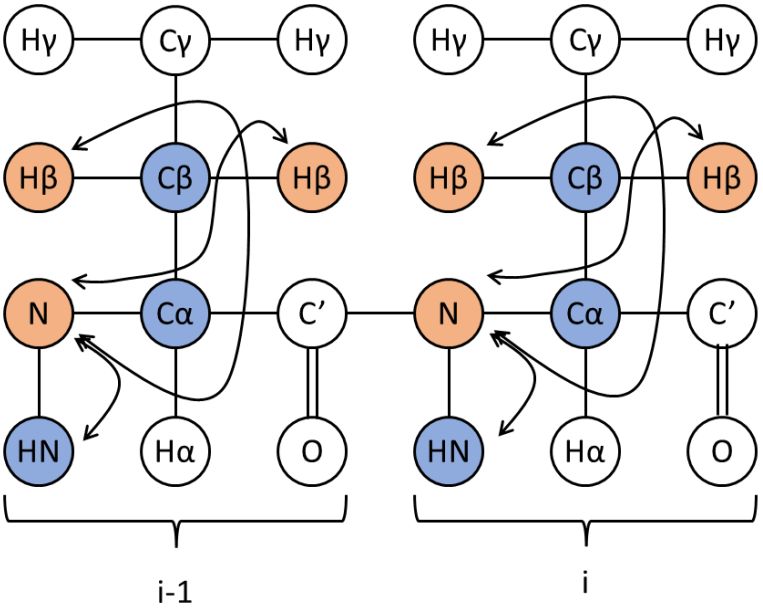

Figure 47. Magnetization flow diagram for the HNHB experiment.

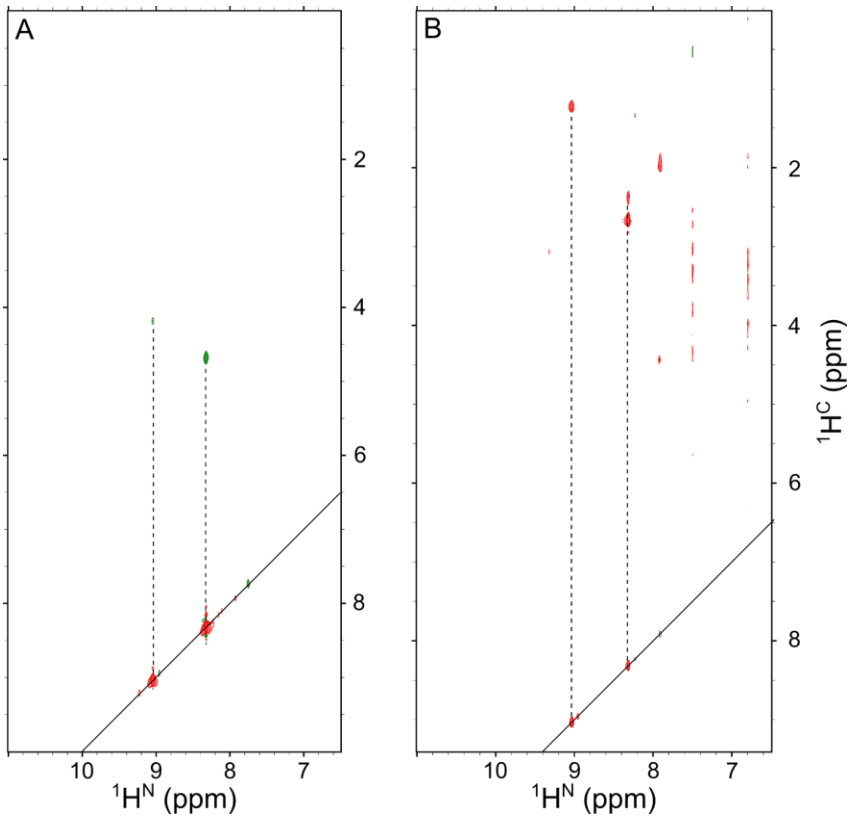

Figure 48. Example spectra for the J-coupling NMR experiments. (A) Example spectrum for the HNHA experiment. (B) Example spectrum for the HNHB experiment. Data were collected for TGM6-D3 with a ^{15}N chemical shift of 121.9 ppm.

For the HNHA experiment (Figure 46), magnetization begins on the amide proton and is modulated by the alpha proton via the three-bond H^N-H^α J-coupling. The magnetization is then transferred to the attached ^{15}N via the one-bond N-H J-coupling using a heteronuclear multiple-quantum correlation (HMQC) sequence, allowed to evolve on the amide nitrogen, and then transferred back to the attached ^{15}N via the one-bond N-H J-coupling using a HMQC sequence. The magnetization is then detected on the amide proton. For a more complete description of the pulse sequence, see [261].

For the HNHB experiment (Figure 47), magnetization begins on the amide proton and is transferred to the attached ^{15}N via the one-bond N-H J-coupling. The magnetization evolves on the amide nitrogen prior to being modulated by the beta proton via the three-bond N^H-H^β J-coupling

using a HMQC sequence. Finally, the magnetization is transferred back to the amide proton via the one-bond N-H J-coupling. The magnetization is then detected from the amide proton. For a more complete description of the pulse sequence, see [5].

As stated previously, the HNHA and HNHB experiments correlate the amide proton with the alpha and beta protons, respectively. Therefore, they can be useful for identifying the alpha and beta protons in a sample, as the HBHA(CO)NH experiment does. However, these experiments provide more information than just identities of protons. The ratio of the intensities between the amide and alpha or beta resonance peaks in the spectra is proportional to the dihedral ϕ and χ_1 angles, respectively, according to the Karplus relationships shown in Equation 1 and Equation 2. These dihedral angles are then used as restraints during NMR structure calculations.

3.4.3 Residual dipolar couplings

One of the general descriptions of NMR spectroscopy involves the consideration of the spinning nuclei as analogous to bar magnets, each with a north and south pole. These magnets are considered dipolar. With multiple nuclei in a sample, these spinning magnets interact with one another, inducing an effect on one another. One of these effects is a modulation of the one-bond J-couplings between two bonded nuclei, such as the one-bond $^{15}N^{H}$-$^{1}H^{N}$ J-coupling, the one-bond $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ J-coupling, or the one-bond $^{13}C^{\alpha}$-$^{13}C^{O}$ J-coupling. The magnitude of these dipolar interactions is dependent on the orientation of the protein but independent of the magnetic field strength, as shown below:

$$D_{PQ} = -\frac{\gamma_P \gamma_Q \hbar \mu_0}{4\pi r^3} \left\langle \frac{1}{2}(3\cos^2\theta - 1) \right\rangle \quad [3]$$

where: D_{PQ} is the magnitude of the dipolar interaction between nuclei P and Q; $\gamma_P$and γ_Q are the gyromagnetic ratios for nuclei P and Q, respectively; $\hbar$ is Planck's Constant divided by 2π; μ_0 is the permittivity of vacuum; r is the distance between the nuclei; and θ is the angle of the dipolar interaction relative to the static magnetic field [15, 255].

For the entire sample, this angle of the dipolar interaction is averaged over all nuclei in the sample. For a sample that freely tumbles in solution, this value averages to zero, and there are no net effects due to the orientation of the sample. However, for samples whose alignment is constrained relative to the static magnetic field, also known as an anisotropic sample, the angle of the dipolar interaction relative to the static magnetic field does have an effect on the overall magnitude of the dipolar interaction. Therefore, the difference between an unaligned, or isotropic, dipolar coupling and an aligned, or anisotropic, dipolar coupling, is indicative of the angle of the chemical bond between the two coupled nuclei relative to the static magnetic field. These differences are referred to as the Residual Dipolar Coupling, RDC. This technique, therefore, can be used to generate restraints that are useful for NMR structure calculations [14, 15, 18, 255, 284].

There are multiple options for the medium used to align the protein samples. These include, but are not limited to: Pf1 filamentous phage [72, 292]; acrylamide gels [96, 168, 230, 256]; and liquid crystalline materials [263, 291]. Each alignment medium provides unique constraints for the molecular tumbling of a protein, and therefore provides unique RDC values that can be used as restraints in structure determination. Therefore, multiple sets of RDC measurements provide non-degenerate restraints for the structure determination process.

As a general rule, each set of RDC measurements, regardless of the experiment, must be performed at least twice: once under isotropic (unaligned) conditions, and once under each

anisotropic (aligned) condition with the chosen alignment medium. Therefore, the calculation of RDC values for each chosen alignment medium requires two independent sets of spectra.

3.4.4 RDC NMR experiments

3.4.4.1 The IPAP-HSQC experiment

The in-phase/anti-phase heteronuclear single-quantum coherence (IPAP-HSQC) experiment [194, 282] is designed to measure the amide $^{1}H^{N}$-$^{15}N^{H}$ RDC [194]. The experiment is similar to the standard ^{1}H-^{15}N HSQC experiment; however, the pulse sequence for the IPAP-HSQC experiment is set up such that the amide nitrogen resonance splitting induced by the proton is not suppressed. An example spectrum for the IPAP-HSQC experiment, following processing of both the in-phase and anti-phase components, is shown in Figure 49.

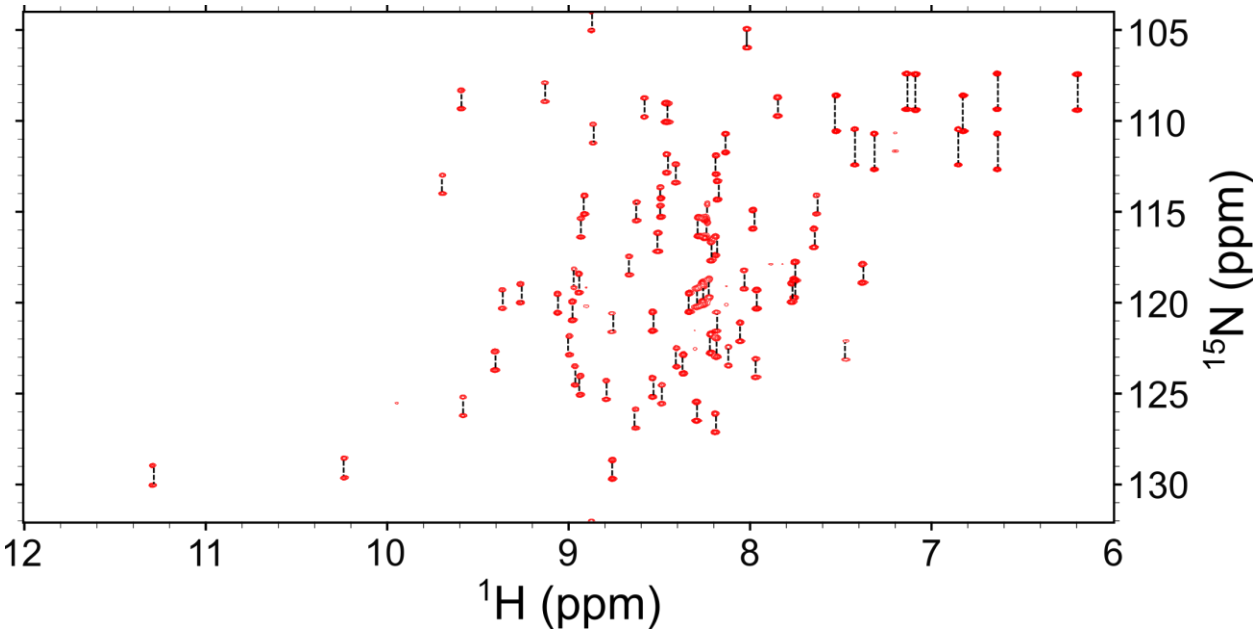

Figure 49. Example spectrum for the IPAP-HSQC experiment. Data were collected for TGM6-D3.

As with the standard ^{1}H-^{15}N HSQC experiment, the IPAP-HSQC experiment generates two-dimensional NMR spectra that correlate the amide proton with the amide nitrogen within the same residue. Two peaks are present at each proton chemical shift in the IPAP-HSQC spectrum for each peak present in the ^{1}H-^{15}N HSQC spectrum; this is due to the ^{1}H-induced splitting of the ^{15}N resonance. Assignment of the IPAP-HSQC is straightforward once the ^{1}H-^{15}N HSQC spectrum is assigned: the assignments can be easily transferred to the IPAP-HSQC spectrum.

The $^{1}H^{N}$-$^{15}N^{H}$ RDC is measured from the IPAP-HSQC spectrum by means of a change in the one-bond J-coupling between the amide proton and amide nitrogen. For each peak in the IPAP-HSQC spectrum, the J-coupling constant, or distance between the two NH peaks, is measured for both the isotropic and isotropic samples. The difference in the J-coupling constants between the isotropic and anisotropic samples is then calculated. This difference is the RDC that is induced by alignment of the sample. This $^{1}H^{N}$-$^{15}N^{H}$ RDC value is then used as a restraint in NMR structure calculations.

3.4.4.2 The HA-coupled HN(CO)CA experiment

The HA-coupled HN(CO)CA experiment [36, 114] is designed to measure the $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ RDC. The experiment is similar to the HN(CO)CA experiment; however, the pulse sequence for the HA-coupled HN(CO)CA experiment is set up such that the alpha carbon resonance splitting induced by the alpha proton is not suppressed. An example spectrum for the HA-coupled HN(CO)CA experiment is shown in Figure 50.

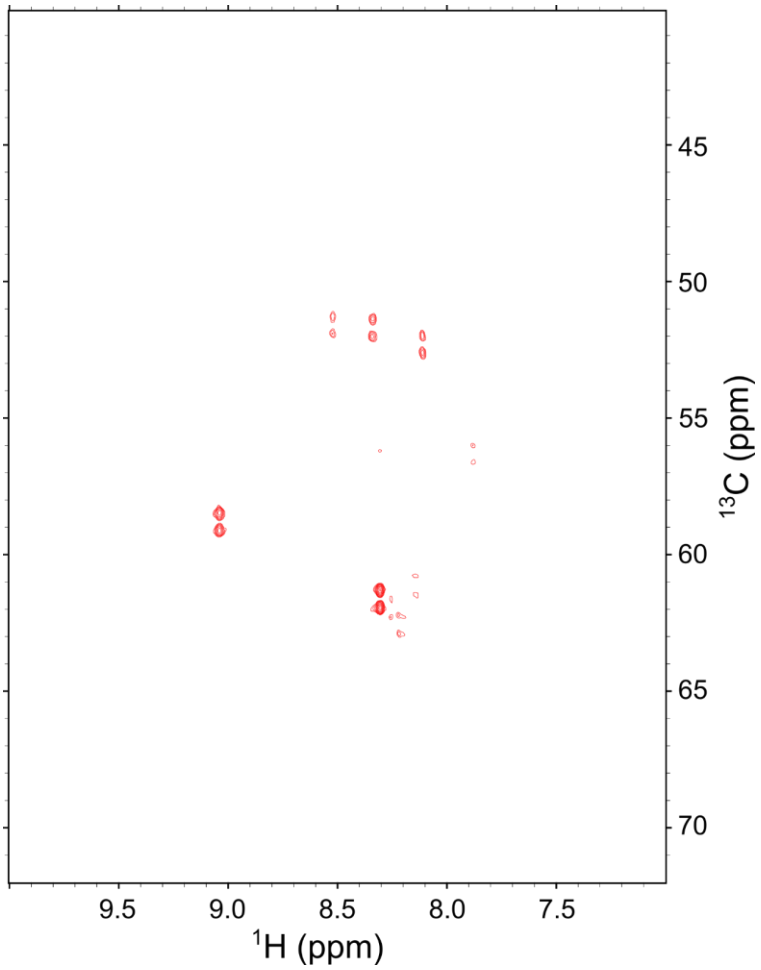

Figure 50. Example spectrum for the HA-coupled HN(CO)CA experiment. Data were collected for TGM6-D3 with a ^{15}N chemical shift of 121.9 ppm.

As with the standard HN(CO)CA experiment, the HA-coupled HN(CO)CA experiment generates three-dimensional NMR spectra that correlate the alpha carbon to the amide-nitrogen pair of the following residue. Two peaks are present at each NH chemical shift in the HA-coupled HN(CO)CA spectrum for each peak present in the standard HN(CO)CA; this is due to the ^{1}H-induced splitting of the $^{13}C^{\alpha}$ resonance. Assignment of the HA-coupled HN(CO)CA spectrum is straightforward once the HN(CO)CA or CBCA(CO)NH spectrum is assigned: the assignment can easily be transferred to the HA-coupled HN(CO)CA spectrum.

The $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ RDC is measured from the HA-coupled HN(CO)CA spectrum by means of a change in the one-bond J-coupling between the alpha proton and alpha carbon. For each peak in the HA-coupled HN(CO)CA spectrum, the J-coupling constant, or distance between the two carbon peaks, is measured for both the isotropic and anisotropic samples. The difference in the J-coupling constants between the isotropic and anisotropic samples is then calculated. This difference is the RDC that is induced by alignment of the sample. This $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ RDC value is then used as a restraint in NMR structure calculations.

3.4.4.3 The CA-coupled HNCO experiment

The CA-coupled HNCO experiment [65, 112, 114, 178, 234] is designed to measure the $^{13}C^{\alpha}$-$^{13}C^{O}$ RDC. The experiment is similar to the HNCO experiment; however, the pulse sequence for the CA-coupled HNCO experiment is set up such that the carbonyl carbon resonance splitting induced by the alpha proton is not suppressed. An example spectrum for the CA-coupled HNCO experiment is shown in Figure 51.

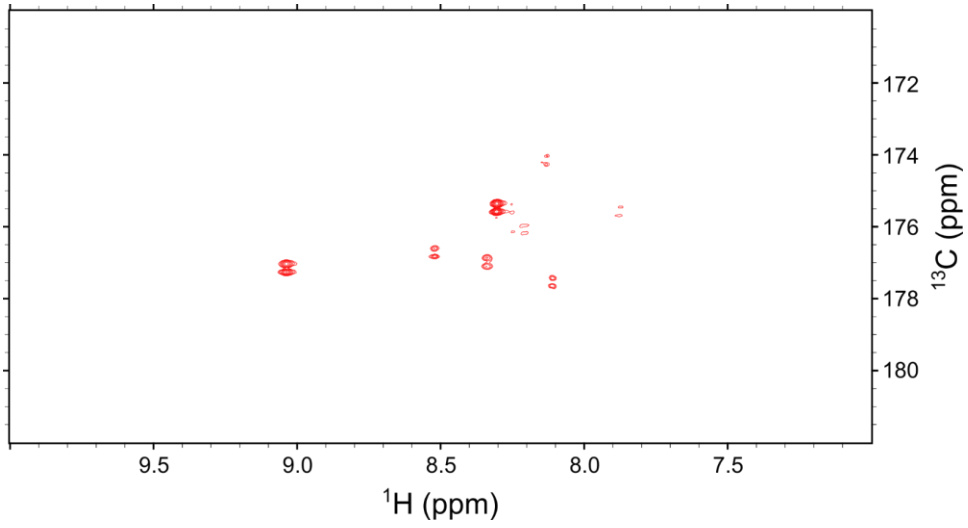

Figure 51. Example spectrum for the CA-coupled HNCO experiment. Data were collected for TGM6-D3 with a ^{15}N chemical shift of 121.9 ppm.

As with the standard HNCO experiment, the CA-coupled HNCO experiment generates three-dimensional NMR spectra that correlate the carbonyl carbon to the amide-nitrogen pair of the following residue. Two peaks are present at each NH chemical shift in the CA-coupled HNCO spectrum for each peak present in the standard HNCO; this is due to the $^{13}C^{\alpha}$-induced splitting of the $^{13}C^{O}$ resonance. Assignment of the CA-coupled HNCO spectrum is straightforward once the HNCO spectrum is assigned: the assignment can easily be transferred to the CA-coupled HNCO spectrum.

The $^{13}C^{\alpha}$-$^{13}C^{O}$ RDC is measured from the CA-coupled HNCO spectrum by means of a change in the one-bond J-coupling between the alpha carbon and carbonyl carbon. For each peak in the CA-coupled HNCO spectrum, the J-coupling constant, or distance between the two carbon peaks, is measured for both the isotropic and anisotropic samples. The difference in the J-coupling constants between the isotropic and anisotropic samples is then calculated. This difference is the RDC that is induced by alignment of the sample. This $^{13}C^{\alpha}$-$^{13}C^{O}$ RDC value is then used as a restraint in NMR structure calculations.

3.4.5 The Nuclear Overhauser Effect

Simply, the Nuclear Overhauser Effect, or Nuclear Overhauser Enhancement, (NOE) is a transfer of spin polarization in a through-space manner due to cross-relaxation and dipolar interactions between two nuclei [115, 223]. Practically, this effect is measured as a change in the integrated intensity, either positive or negative, of one resonance that occurs when another is saturated by an RF pulse. The change in resonance intensity of a nucleus is a consequence of the nucleus being close in space to those directly affected by the RF perturbation. The NOE is particularly important in the assignment of NMR resonances for small proteins and for the

elucidation and confirmation of the structures or configurations of proteins. The ^{1}H two-dimensional NOE Spectroscopy (NOESY) experiment and its extensions, such as the three- and four-dimensional filtered ^{1}H NOESY experiments, are important for the determination of through-space relationships for various protons within a protein sample, as the NOE acts in a through-space manner.

3.4.6 NOE NMR experiments

3.4.6.1 The ^{15}N-edited ^{1}H NOESY experiment

The ^{15}N-edited ^{1}H NOESY experiment [111, 157, 158, 195, 197, 235, 285, 289] is a three-dimensional NMR experiment that allows for the determination of through-space ^{1}H-^{1}H contacts, from which distances can be determined. A diagram of the flow of the magnetization for this experiment is shown in Figure 52, and an example spectrum is shown in Figure 53.

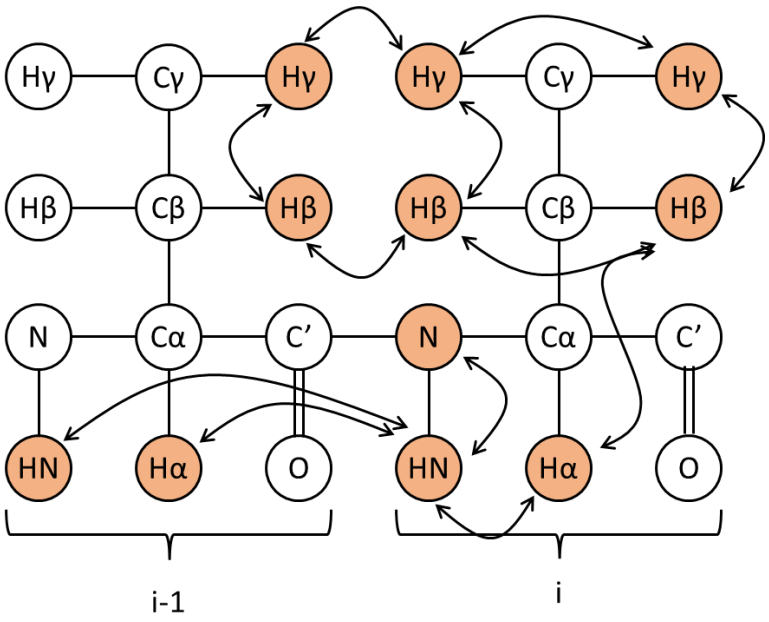

Figure 52. Magnetization flow diagram for the ^{15}N-edited ^{1}H-NOESY HSQC experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

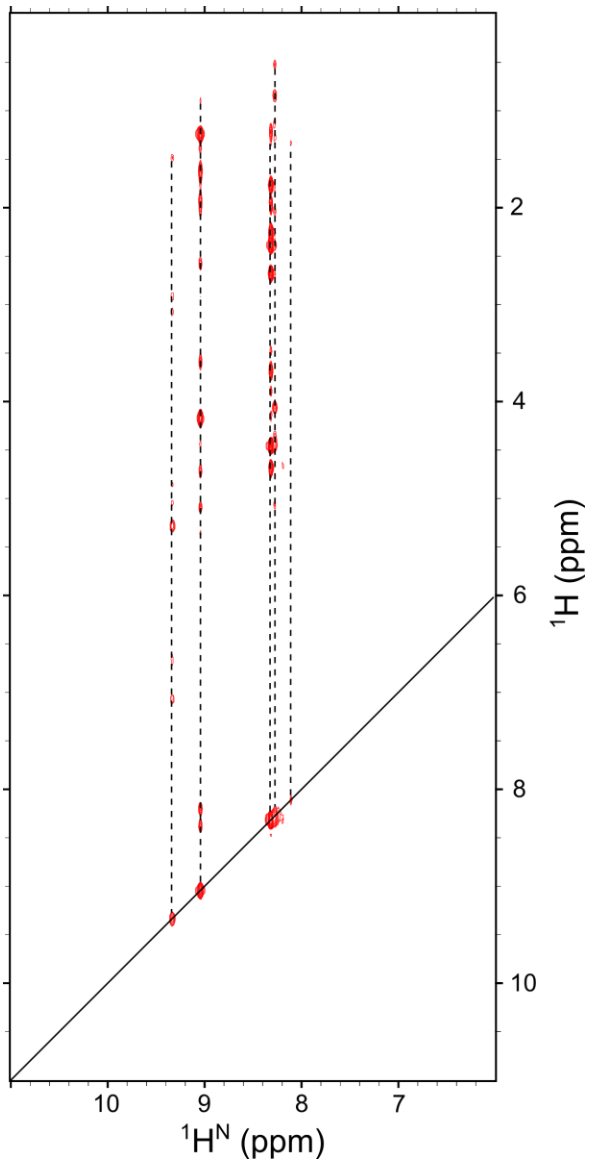

Figure 53. Example spectrum for the ^{15}N-edited ^{1}H-NOESY HSQC experiment. Data were collected for the aliphatic region of TGM6-D3 with a ^{15}N chemical shift of 121.9 ppm.

In this experiment, magnetization begins on all protons in the protein. The magnetization is then exchanged between all protons using the NOE. At this point, the magnetization undergoes a ^{1}H-^{15}N HSQC for detection.

In this experiment, the magnetization is NOESY transferred between the protons in the protein sample. In this way, all protons can "see" their neighbors in a through-space manner. This

includes the amide protons being able to "see" the aliphatic protons in the protein sidechains. By following the NOESY transfer with a ^{1}H-^{15}N HSQC sequence, the protons visible to the amide proton are selectively filtered from the other resonances.

As stated previously, the ^{15}N-edited ^{1}H NOESY experiment provides data used in determining distances between the detected protons. The ^{15}N-edited ^{1}H NOESY spectrum, resultant from the experiment, is traditionally displayed as $^{1}H^{N}$-^{1}H planes, with the amide proton ($^{1}H^{N}$) chemical shift along the x-axis and the "seen" proton (^{1}H) chemical shift along the y-axis. The ^{15}N chemical shift axis (the z-axis) is variable such that the correlating $^{1}H^{N}$-^{1}H planes can be viewed for each collected ^{15}N frequency.

In the ^{15}N-edited ^{1}H NOESY spectrum (Figure 53), each vertical strip, representing a $^{1}H^{N}$-^{15}N pair, has visible peaks representing the NOEs from protons in physical proximity to the amide ^{1}H-^{15}N pair. Once the resonances for the atoms in the protein are assigned by the methods previously discussed, the NOE peaks can be assigned, and the relative intensities of the peaks can be used as restraints for structure calculations.

3.4.6.2 The ^{13}C-edited ^{1}H NOESY experiment

The ^{13}C-edited ^{1}H NOESY experiment [111, 157, 158, 195, 197, 235, 285, 289] is another three-dimensional NMR experiment that allows for the determination of through-space ^{1}H-^{1}H contacts, from which distances can be determined. A diagram of the flow of the magnetization for this experiment is shown in Figure 54, and an example spectrum is shown in Figure 55.

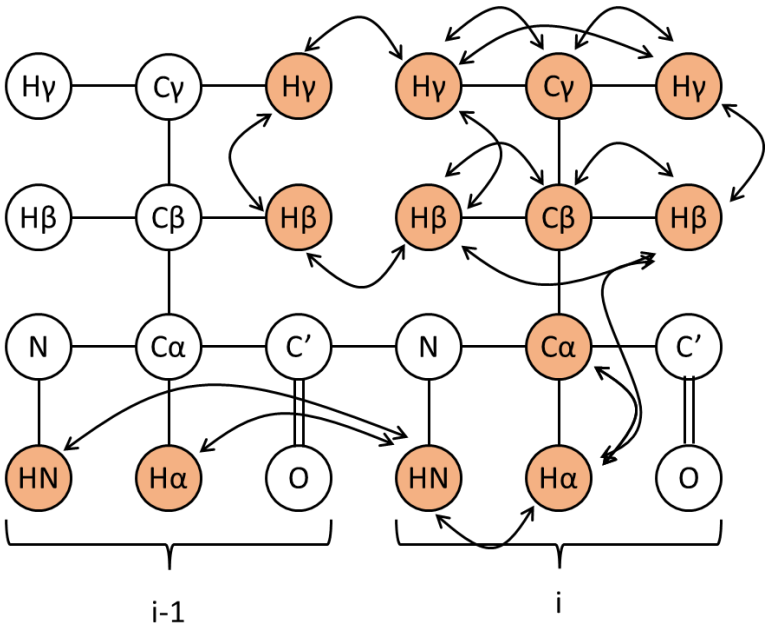

Figure 54. Magnetization flow diagram for the ^{13}C-edited ^{1}H NOESY HSQC experiment. Figure adapted with permission from Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/. [85].

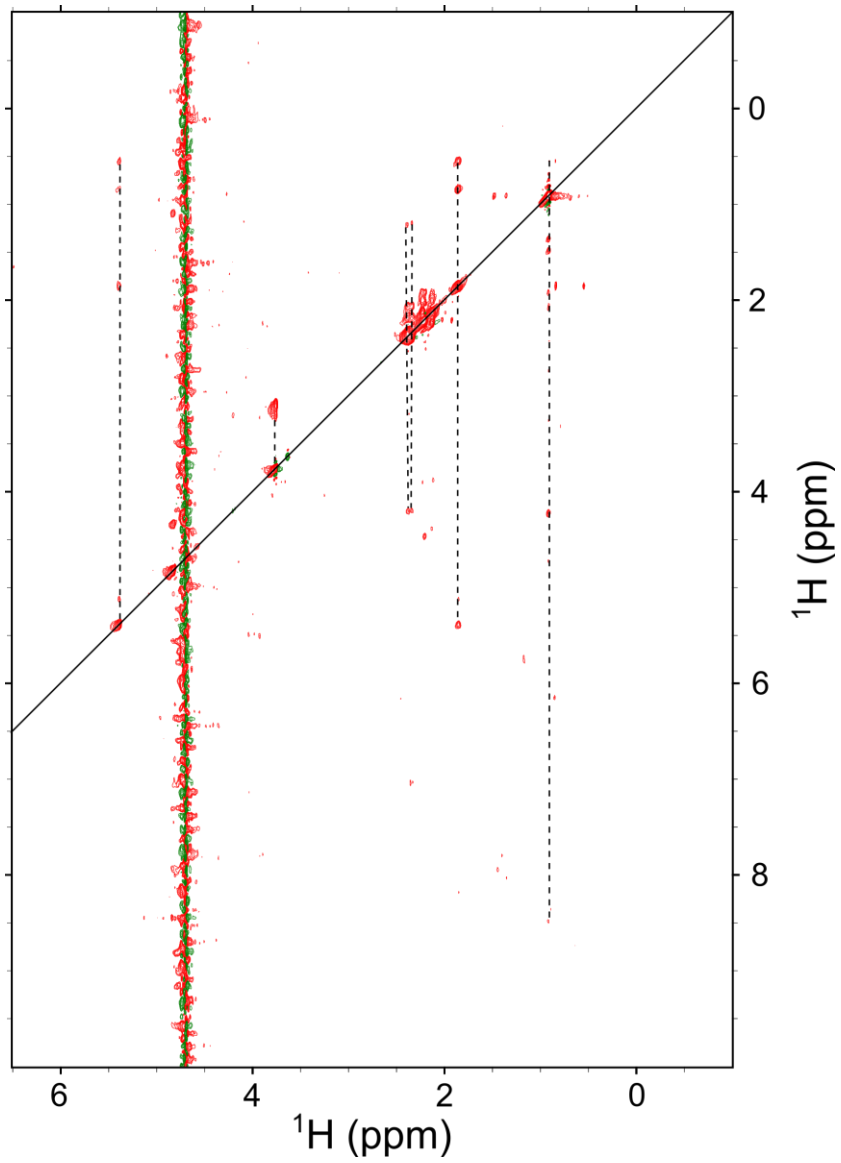

Figure 55. Example spectrum for the ^{13}C-edited ^{1}H-NOESY HSQC experiment. Data were collected for the aliphatic region of TGM6-D3 with a ^{13}C chemical shift of 31.4 ppm.

Similar to the ^{15}N-edited ^{1}H NOESY experiment, in this experiment, magnetization begins on all protons in the protein. The magnetization is then exchanged between all protons using the NOE. At this point, the magnetization undergoes a ^{1}H-^{13}C CT-HSQC for detection.

In this experiment, the magnetization is NOESY transferred between the protons in the protein sample. In this way, all protons can "see" their neighbors in a through-space manner. By

following the NOESY transfer with a ^{1}H-^{13}C CT-HSQC sequence, the protons visible to those bonded to specific carbon nuclei are selectively filtered from the other resonances.

As stated previously, the ^{13}C-edited ^{1}H NOESY experiment provides data used in determining distances between the detected protons. The ^{13}C-edited ^{1}H NOESY spectrum, resultant from the experiment, is traditionally displayed as $^{1}H^{C}$-^{1}H planes, with the carbon-attached proton ($^{1}H^{C}$) chemical shift along the x-axis and the "seen" proton (^{1}H) chemical shift along the y-axis. The ^{13}C chemical shift axis (the z-axis) is variable such that the correlating $^{1}H^{C}$-^{1}H planes can be viewed for each collected ^{13}C frequency.

In the ^{13}C-edited ^{1}H NOESY spectrum (Figure 55), each vertical strip, representing a $^{1}H^{C}$-^{13}C pair, has visible peaks representing the NOEs from protons in physical proximity to the $^{1}H^{C}$-^{13}C pair. In addition to the single set of peaks in a strip, as seen with the assignment spectra and the ^{15}N-edited ^{1}H NOESY, this spectrum also contains strips with doubled peaks. These doubled peaks arise from CH_2 groups present in the protein sidechains, where there two protons attached to the same carbon have different chemical shifts. Once the resonances for the atoms in the protein are assigned by the methods previously discussed, the NOE peaks can be assigned, and the relative intensities of the peaks can be used as restraints for structure calculations.

3.5 Conclusions

NMR spectroscopy is a powerful method for the determination of protein structures and characterization of protein interactions. While the process for determining an NMR structure is not trivial, it can provide a lot of insight into the solution state structure of a protein, relative to that of a crystallized protein. This chapter shows that the backbone and sidechain carbon and

proton resonances were assigned for TGM6-D3, and that spectra were collected and assignments began for the determination of restraints for structure determination. At this point, the determination of restraints and calculation of the NMR structure for TGM6-D3 is incomplete. However, this task can be completed with the data on hand.

3.6 Materials and Methods

3.6.1 Expression and purification of TGM6-D3

Details of the TGM6-D3 construct used is provided in Table 19. The DNA fragment corresponding to TGM6-D3 was inserted between the *Kpn*I and *Hind*III sites of a modified form of the pET32a vector (EMD-Millipore, Danvers, MA). The pET32a vector was modified to include a His10 tag instead of the standard His6 and a *Kpn*I site immediately following the sequence for the thrombin cleavage site. The resulting constructs were designed as follows: thioredoxin-decahistidine tag-thrombin cleavage site-TGM6-D3 coding cassette. Note: the expressed, un-cleaved TGM6 protein described in this section is referred to as the TGM6 fusion protein, and the cleaved protein is referred to as TGM6-D3.

The construct was overexpressed in BL21(DE3) cells (EMD-Millipore, Danvers, MA) cultured at 37°C. Unlabeled samples for binding studies were produced on 2x tris-buffered Lauria Broth (LB); ^{15}N and ^{13}C, ^{15}N isotopically labeled samples for NMR studies were produced using M9 minimal medium containing 0.1% $^{15}NH_4Cl$ (Sigma-Aldrich, St. Louis, MO) or 0.1% $^{15}NH_4Cl$ and 0.5% U-^{13}C-D-glucose (Cambridge Isotope Laboratories, Tewksbury, MA). To select for cells bearing the expression plasmid, carbenicillin or ampicillin was added to all media at 100 μg mL^{-1}

or 150 μg mL $^{-1}$, respectively. Protein expression was induced by adding 200 mg L^{-1} IPTG when the optical density of the culture at 600 nm reached ~1.0. Expression occurred for 3-4 hours after induction at 37°C, and the cultures were harvested by centrifugation (7,300 g, 15 minutes, 25°C).

Cell pellets from 3 L of culture were resuspended in 100 mL lysis buffer (50 mM Na_2HPO_4, 100 mM NaCl, 5 mM imidazole, pH 8.0) and sonicated for 5 minutes at 50% duty cycle. Following centrifugation (30,000 g, 20 minutes, 25°C), the supernatant was discarded, the pellet was resuspended in lysis buffer containing 500 mM NaCl, and the solution was re-centrifuged (30,000 g, 20 minutes, 25°C). Following the second centrifugation, the supernatant was again discarded, the pellet was resuspended in lysis buffer containing 1% (v/v) Triton X-100, and the solution was centrifuged for a third time (30,000 g, 20 minutes, 25°C). Following the third centrifugation, the supernatant was discarded, the pellet was resuspended in resuspension buffer (50 mM Na_2HPO_4, 8 M urea, 100 mM NaCl, 5 mM imidazole, pH 8.0), and the solution was stirred vigorously overnight at 25°C. The remaining insoluble material was removed by centrifugation (30,000 g, 20 minutes, 25°C), and the supernatant was loaded onto a 30 mL metal affinity column (Ni^{++} loaded chelating Sepharose resin, GE Healthcare, Piscataway, NJ) pre-equilibrated with 100 mL of resuspension buffer. The column was washed with 100 mL of resuspension buffer, and the bound protein was eluted using a linear gradient of resuspension buffer containing 0 – 0.5 M imidazole.

Protein from the eluted peak was treated with reduced glutathione (GSH) such that the final GSH concentration once the protein is diluted into the folding buffer was 2 mM. Following a 30-minute incubation at 25°C, the protein was slowly diluted into folding buffer (100 mM tris, 1 mM EDTA, 0.5 mM oxidized glutathione (GSSG), pH 8.0) to a final concentration of less than 0.1 mg mL^{-1} with a urea concentration of less than 300 mM. The folding solution was then stirred for 12 – 18 hours at 4°C and subsequently concentrated by ultrafiltration. To the concentrated protein,

solid thrombin was added to a final concentration of 1 U mg^{-1} TGM6 fusion protein. The protein mix was dialyzed against 50 mM Na_2HPO_4, 100 mM NaCl, 5 mM imidazole, pH 8.0 for 24 – 36 hours at 4°C. Cleavage was stopped by adding 10 μM leupeptin hemisulfate, 10 μM pepstatin, and 100 μM PMSF, and the protein solution was dialyzed against 25 mM $NaCH_3COO$, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 4.8 at 4°C. The dialyzed protein solution was then sterile-filtered, bound to a Source S column (GE Healthcare, Piscataway, NJ) equilibrated in 25 mM $NaCH_3COO$, 2 M urea, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 4.8, and eluted with a 0 – 0.35 M NaCl gradient.

Masses of the TGM6-D3 protein were confirmed by liquid chromatography electrospray ionization time-of-flight mass spectroscopy (LC-ESI-TOF-MS; Bruker Micro TOF, Billerica, MA) (Figure 71). Folding of the TGM6 domain proteins was assessed through NMR spectroscopy using 2D ^{1}H-^{15}N HSQC spectra, as described in the main text.

3.6.2 NMR sample preparation and 2D experiments

Samples of TGM6-D3 were prepared at a concentration of 0.2 to 1.2 mM in 25 mM Na_2HPO_4, 50 mM NaCl, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 5.5, and transferred to 5-mm susceptibility-matched microtubes (Shigemi) for data collection. NMR data were collected at 30°C using a Bruker 600, 700, 800, or 900 MHz spectrometer equipped with a 5-mm ^{1}H (^{13}C,^{15} N) z-gradient "TCI" cryogenically cooled probe (Bruker Biospin, Billerica, MA). 2D ^{1}H-^{15}N HSQC spectra were recorded with sensitivity enhancement [111], water flip-back pulses [69], and WATERGATE water suppression pulses [203]. NMR data were processed using NMRPipe [41] and analyzed using a combination of NMRFAM-SPARKY [128] and CcpNmr Analysis Version 2 [259].

3.6.3 NMR backbone assignment

Backbone resonances were assigned by recording and analyzing 2D ^{1}H-^{15}N HSQC and 3D HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA, HNCO, and HN(CA)CO triple resonance datasets. Proton and side chain resonances were assigned by recording and analyzing 2D ^{1}H-^{13}C CT-HSQC and 3D CC(CO)NH, HBHA(CO)NH, HCCH-TOCSY, H(CC)(CO)NH, HNHA, and HNHB datasets. NMR data were processed using nmrPipe [41] and analyzed using a combination of NMRFAM-SPARKY [128], CcpNmr Analysis Version 2 [259], PINE [9, 129], and PECAN [45].

3.6.4 NMR structure measurements of TGM6-D3

^{1}H-^{1}H NOEs, ^{1}H-^{15}N, $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ and $^{13}C^{\alpha}$–$^{13}C^{O}$ residual dipolar couplings (RDCs), TALOS-derived phi and psi restraints (79), hydrogen bond restraints, and $^{3}J^{HN\text{-}H\alpha}$ J-coupling restraints were measured as input for calculation of the solution structure of TGM6-D3. The ^{1}H-^{1}H distance restraints were derived from manually peak-picked 3D ^{15}N-edited and 3D ^{13}C-edited NOESY datasets using the program CcpNmr Analysis Version 2 [259], with distance restraints derived using routines provided by CcpNmr. The RDCs were recorded using a sample containing 12.5 mg mL^{-1} Pf1 phage [72] or soaked into a 5% polyacrylamide gel [230] for alignment and were measured using a 2D IPAP-HSQC [194] for ^{1}H-^{15}N RDCs, a 3D $^{13}C^{\alpha}$-coupled HNCO for $^{13}C^{\alpha}$–$^{13}C^{O}$ RDCs, and a 3D H^{α}-coupled HN(CO)CA for $^{1}H\alpha$-$^{13}C\alpha$ RDCs. The $^{3}J^{HN\text{-}H\alpha}$ was measured from the ratio of the crosspeak to diagonal in a 3D H^{N}-H^{α} experiment as described [261]. Note: structure calculations for TGM6-D3 were not performed for this project as of the completion of this dissertation.

3.7 Acknowledgements

We would like to thank Mr. Mike Delk for his assistance with the NMR instrumentation. This research was supported by the NIH (GM58670 and AI53915). 2D NMR spectra were plotted using NMRFAM-SPARKY [128].

4.0 The *H. polygyrus* TGF-β Mimic TGM6 Functions as a Competitive TGF-β Signaling Inhibitor in Murine Fibroblasts

At the time of the completion of this dissertation, this chapter was being prepared as a manuscript for peer review and publication. White SE, Schoenherr C, Hinck CS, Cunningham K, Campion T, Schwartze T, White MPJ, Inman G, Maizels RM, Hinck AP. The H. polygyrus TGF-β Mimic TGM6 Functions as a Competitive TGF-β Signaling Inhibitor in Murine Fibroblasts. In preparation. 2022 [268].

Author contributions: SEW conceptualized experiments, produced proteins, performed and analyzed SPR, ITC, and NMR experiments, prepared protein crystals, analyzed X-ray diffraction data, determined the crystal structure, and wrote the manuscript; CS, CSH, KC, TC, and MPJW performed TGM6 inhibition assays; TS prepared protein crystals; and GI, RMM, and APH oversaw the project. All authors edited the manuscript.

4.1 Introduction

Helminth parasites are a continuing human health burden in the developing world, especially in Africa, Asia and South America [91, 92, 206]. The longevity of the parasites reflects a refined ability to evade the host immune system through molecular strategies that are only now being elucidated [152, 224, 269]. One prominent mechanism that these parasites have evolved is the secretion of proteins that act by modulating the host's immune response [73, 82, 151, 152, 239]. For example, helminth infections are often associated with the activation of regulatory T-

cells (Tregs), which dampen the inflammatory response and reduce anti-parasitic immune responses [49, 135, 151]. This is achieved through two mechanisms: expansion of the host's pre-existing Treg populations and/or inducing *de novo* differentiation of peripheral T-cells into Tregs [103, 147, 252, 266, 267].

The murine intestinal parasite *Heligmosomoides polygyrus* is widely used as a model parasite to study host-parasite interactions [16, 101, 148]. Upon infection of a mouse with this parasite, there is an observable increase in the number of Tregs [51, 213]; conversely, antibody-mediated depletion of Tregs from infected mice results in significantly reduced parasite burdens [245]. *H. polygyrus* secretes a protein that mimics the function of the mammalian cytokine Transforming Growth Factor Beta (TGF-β) [103], the cytokine known to drive Treg differentiation in the immune system [131, 222]. This parasite protein was isolated from *H. polygyrus* excretory secretory products, also known as HES, and was subsequently referred to as the *H. polygyrus* TGF-β mimic, TGM.

The three mammalian TGF-β isoforms control and influence many pathways in cellular differentiation [106, 227, 242] and immune homeostasis [117, 229, 242]. Specifically, TGF-β induces the differentiation of naïve $CD4^+$ T-cells into $CD4^+$ $CD25^+$ $Foxp3^+$ Tregs, which are essential for peripheral immune tolerance [34, 229]. A lack of endogenous TGF-β1 in mice is characterized by the development of autoimmune disorders, such as multi-organ inflammatory disease, and death after maternal TGF-β is depleted [242]. Furthermore, dysregulation of TGF-β signaling has been shown to be involved with the pathogenesis of several human diseases, including inflammatory bowel disease [95], renal and cardiac fibrosis [93, 118], and cancer [118, 161, 238].

TGF-β growth factors are comprised of two elongated cystine-knotted monomers held together by a single interchain disulfide bond [90]. The growth factor signals by assembling a heterotetrameric complex with two independently signaling pairs of serine/threonine kinase receptors, known as the TGF-β type I and type II receptors, TβRI (or Alk5, as it is referred to in this paper) and TβRII, respectively [64, 94, 273, 275]. For signaling to occur, the constitutively active TβRII phosphorylates Alk5, which then in turn phosphorylates the downstream effector molecules, SMAD2 and SMAD3 [77].

TGM is a disulfide-rich 422-amino acid protein with an N-terminal signal peptide and five homologous domains [103]. Each domain has approximately 85 to 90 amino acids with either two or three disulfide bonds. These domains bear no homology to TGF-β or other TGF-β family members; instead, the individual domains are distantly related to the complement control protein (CCP) or Sushi domain family [103]. Further work showed that there were at least nine homologs of the TGM protein, identified in the secreted proteome and transcriptome of *H. polygyrus* [246], which were numbered TGM2 through TGM10, with the founding member, TGM, being numbered TGM1. Among this family of proteins, six (TGM1 through TGM6) are expressed primarily in the adult stages of the parasite while the remaining four (TGM7 through TGM10) are expressed exclusively in the larval stages [246]. Previous work also showed that domains 1, 2, and 3 (D1, D2, and D3, respectively) of TGM1 are necessary and sufficient for signaling in reporter cells [246]; specifically, TGM1-D1 and -D2 bind Alk5 while -D3 binds TβRII [179]. This combination of receptor pairings mimics the signaling complex created by TGF-β. However, this signaling complex differs in that monomeric TGM1 bivalently binds and assembles a TβRII:Alk5 heterodimer, while TGF-β is a dimer with two identical binding sites for TbRII:Alk5 heterodimers that results in a hexameric signaling complex.

Among the TGMs expressed during the adult stages of the parasite, TGM6 is unique in that it lacks D1 and D2 [103]. Previous work has shown that TGM6 does not signal in TGF-β functional assays, unlike TGM1 [246]. In this work, the function of TGM6 is described. NMR spectroscopy and isothermal titration calorimetry (ITC) experiments show that TGM6 does not bind the TGF-β family type I receptors Alk1, Alk2, Alk3, Alk4, and Alk5 and only binds TβRII as a type II receptor. The thermodynamics of binding between TGM6 and TβRII show that TGM6-D3 is necessary and sufficient for binding TβRII and that TGM6-D3 competes with TGF-β and TGM1-D3 for the TGF-β binding site on TβRII. This is confirmed by determining the structure of the TGM6-D3:TβRII binary complex and testing the key interactions between the two binding partners by surface plasmon resonance (SPR) and ITC. To further explore the function of TGM6, cellular signaling assays show that full-length TGM6 is an inhibitor of TGF-β signaling in mouse fibroblasts. Finally, the mechanism by which TGM6 can act as an inhibitor in mouse fibroblasts but not human kidney cells is proposed.

4.2 Results

4.2.1 Isolation and characterization of TGM6-D3 and -D45

TGM6 is a 254-amino acid protein consisting of an 18-amino acid classical signal peptide followed by three homologous domains, each of about 85 amino acids [246]. TGM6 is homologous to TGM1, the founding member of the TGM family; however, it was shown that TGM6 lacks domains 1 and 2 relative to TGM1 [246]. Therefore, the three domains were numbered as 3, 4, and 5 (TGM6-D3, -D4, and -D5, respectively) due to their sequence homology to their TGM1

counterparts. As these domains have a distant homology to the CCP family or Sushi domain family of proteins [103], they are expected to have internal disulfide bonds (two in TGM6-D3 and -D5, which have four cysteines each, and three in TGM6-D4, which has six cysteines). While the structures of the TGM6 domains have not been previously characterized, the disulfide bonding patterns and overall fold are hypothesized to similar to that of CCP/Sushi domains in general (i.e. Cys I-III and Cys II-IV), and TGM1-D3 in particular.

Constructs for the TGM6-D3 and TGM6-D45 proteins (Table 19) were over-expressed in *E. coli* with N-terminal thioredoxin and decahistidine tags separated from the TGM6 proteins by a thrombin cleavage site. The proteins were isolated from both the lysed *E. coli* supernatant and insoluble inclusion bodies, reconstituted in 8 M urea, and refolded in the presence of a glutathione redox couple prior to thrombin cleavage to remove the N-terminal tags and subsequent high-resolution anion exchange chromatography. Protein folding was characterized by ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled TGM6 domains in 25 mM phosphate buffer at pH 5.5 (Figure 68). The CCP family of proteins, to which TGM6 is homologous, is characterized by four or more β-strands [52, 215]; additionally, TGM1, the founding member of the TGM family, is sequentially and hypothesized to be structurally homologous to TGM6 [179]. Therefore, in addition to having the expected number of peaks indicative of homogeneous pairing of cysteines, it is expected that the proteins would have well-dispersed peaks with minimal clustering in the random-coil region (7.8 to 8.6. ppm in the ^{1}H dimension). For TGM6-D3, 83 backbone amide resonances were observed with excellent signal dispersion, in close accord with the 82 signals expected (Figure 68A). This indicates that the protein is conformationally homogenous and natively folded. For TGM6-D45, 128 of the expected 148 backbone amide resonances were observed with excellent signal dispersion (Figure 68B); from this, it is concluded that the protein is also is conformationally

homogenous and natively folded. The absence is of some of the backbone amide signals is likely due to spectral overlap in the ^{1}H-^{15}N 2D spectrum.

4.2.2 TGM6-D3 holds the full TβRII binding capacity

Previously published results showed that TGM1 binds TβRII with a K_D of approximately 1 μM [103, 179]. To determine if TGM6 bound TβRII, an ITC experiment was performed in which TβRII (Table 20) was titrated into full-length TGM6 (flTGM6) (Figure 56A, top). This titration gave a strong exothermic response, indicating binding between TGM6 and TβRII. Fitting the integrated heats to a standard binding isotherm yielded a K_D of 220 ± 100 nM (average ± 1σ) (Figure 56A, middle and bottom; Table 3).

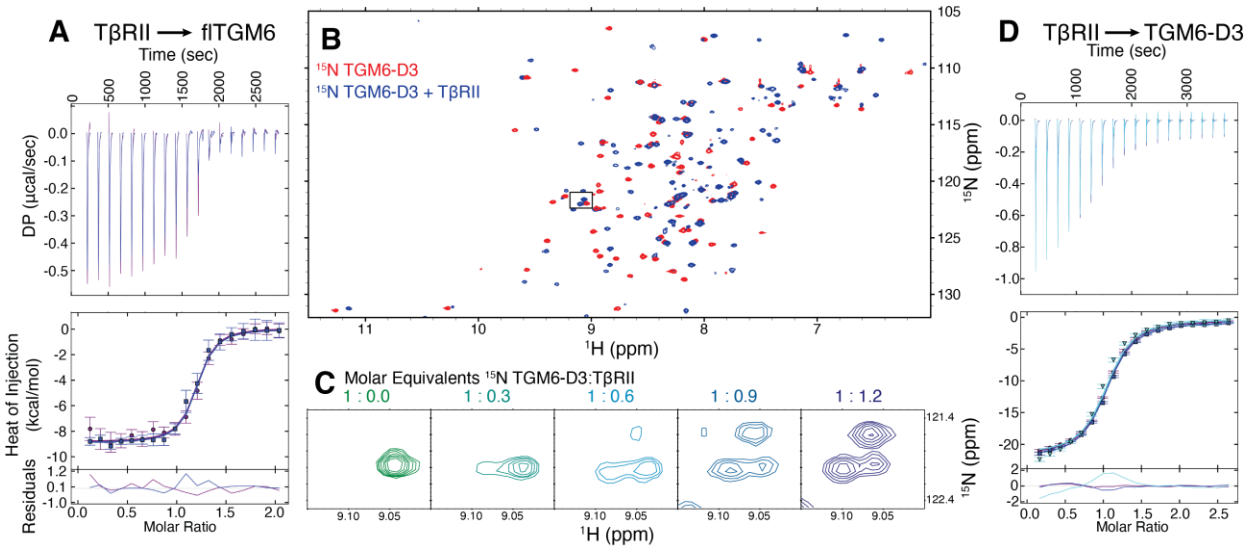

Figure 56. TGM6 binds TβRII through -D3. (A,D) ITC thermograms (top) obtained upon injection of TβRII into flTGM6 (A) or TGM6-D3 (D). Thermograms are overlaid as two (A) or three (D) replicates. Integrated heats for the panels are shown immediately below the thermograms with the residuals as a function of the molar ratio of flTGM6:TβRII (A) or TGM6-D3:TβRII (D). The data points correspond to the integrated heats, and the colored lines correspond to a global fit of the data to a 1:1 binding model. (B) 1H-15N HSQC spectra of ^{15}N TGM6-D3 alone (red) overlaid with the spectrum of the same sample containing a 1.2-fold molar excess of unlabeled TβRII. Spectra were recorded at 303.15 K in 25 mM sodium phosphate, 0.05% NaN_3, 5% ^2H_2O, pH 5.5. (C) Expansion of the boxed region of the spectra in panel A with all titration points, as labeled in molar equivalents of ^{15}N TGM6-D3:TβRII.

Table 3. TGM6:TβRII binding as assessed by ITC.

Cell	32.3 µM flTGM6	14.7 µM TGM6-D3
Syringe	339.7 µM TβRII	200.0 µM TβRII
Temperature (°C)	25	35
N (sites)	1.16	1.01
K_D (nM) [a]	220 ± 100 [b,c]	440 ± 80 [b,d]
ΔH (kcal mol^{-1}) [a]	-8.9 ± 0.2 [b]	-21 ± 0.65 [b]
ΔG (kcal mol^{-1}) [a]	-9.1	-9.0
-TΔS (kcal mol^{-1}) [a]	0.2	12

a Number of sites set to "1" for analysis.
b Uncertainty reported as ± 1σ.
c Global fit of two replicates.
d Global fit of three replicates.

Domain 3 is the TGM1 domain responsible for binding TβRII in TGM1 [179] and has 66% amino acid identity compared to TGM6-D3 [246]. With this similarity between the two TGM

domains, the binding of TGM6 to TβRII was assessed by NMR spectroscopy. To do this, ^{1}H-^{15}N HSQC spectra of ^{15}N TGM6-D3 were recorded with increasing amounts of unlabeled TβRII. This addition caused significant perturbations in over half of the amide proton signals of TGM6-D3 (Figure 56B). Through the course of the titration, peaks corresponding to both the unbound and bound forms of TGM6-D3 were observed at all intermediate titration points (1:0.3, 1:0.6, and 1:0.9) (Figure 56C). This is indicative of slow-exchange binding and suggests that TGM6-D3 binds to TβRII with relatively high affinity.

Using ITC, the binding of TGM6-D3 to TβRII was quantified: as with the flTGM6 experiment, TβRII was titrated into TGM6-D3. Again, the titration gave a strong exothermic response (Figure 56D top), which confirms binding between TGM6-D3 and TβRII. Fitting the integrated heats of interactions to a standard binding isotherm yielded a K_D of 440 ± 80 nM (average ± 1σ) (Figure 56D middle and bottom; Table 3). The difference between the K_D for flTGM6 and TGM6-D3 is not statistically significant (Unpaired t-test p-value = 0.1826; assuming n = 2 for flTGM6 and n = 3 for TGM6-D3 due to replicate experiment count). Therefore, the calculated affinity of TGM6 for TβRII (ca. 320 nM) is roughly three-fold higher than that of TGM1 for TβRII (ca. 1 μM) [179].

Additionally, the binding of TGM6-D3 to ActRII, ActRIIb, and BMPRII was assessed by ITC. For this experiment, each type II receptor was titrated into TGM6-D3 or buffer (Table 21). All three of these type II receptors gave weak responses when titrated into TGM6-D3 (Figure 69A,C,E) that were similar to titrating them into buffer (Figure 69B,D,F). Therefore, TGM6 domain 3 holds the full binding capacity for TβRII and either does not bind ActRII, ActRIIb, or BMPRII or does but so weakly that it is not likely biologically relevant. Thus, TGM6 domain 3 is

necessary and sufficient for binding TβRII as it appears to be the only relevant type II receptor binding partner.

4.2.3 TGM6 does not bind a type I receptor

As stated previously, TGM6 lacks domains 1 and 2. In TGM1, these domains were shown to be required for TGF-β signaling activity [246] and binding TβRI [179]. Additionally, it was shown that TGM6 does not activate signaling in a TGF-β reporter cell line [246]. Therefore, one possibility is that TGM6 does not signal through the TGF-β pathway, presumably as a consequence of the lack of domains 1 and 2 and its inability to compensate by binding TβRI through domains 4 and 5. However, the homology of domains 4 and 5 to the remaining domains does not rule out the possibility that TGM6 might bind another TGF-β family type I receptor through domains 4 and 5.

The hypothesis that TGM6 domains 4 and 5 might bind another TGF-β family type I receptor was tested by NMR spectroscopy. ^{15}N-labeled type I receptor ectodomains (Table 22) were titrated with unlabeled TGM6-D45, and the ^{1}H-^{15}N HSQC spectra of the receptor alone were compared with that of the receptor with TGM6-D45 added in excess. For the five type I receptors tested (Alk1, Alk2, Alk3, Alk4, and Alk5; Figure 57A-E, respectively), there were small (less than 0.02 ppm in the ^{1}H dimension) or no chemical shift perturbations present upon the addition of TGM6-D45, whereas much larger perturbations, for at least a portion of the residues, would be expected if a complex stably formed. Since these titrations were performed with 100 μM receptor and 112.5 μM TGM6-D45, yet the shift perturbations were minimal, binding, if any is present, is in the millimolar range or weaker. Therefore, TGM6-D45 either does not bind any of the type I receptors tested or does but so weakly it is not likely biologically relevant.

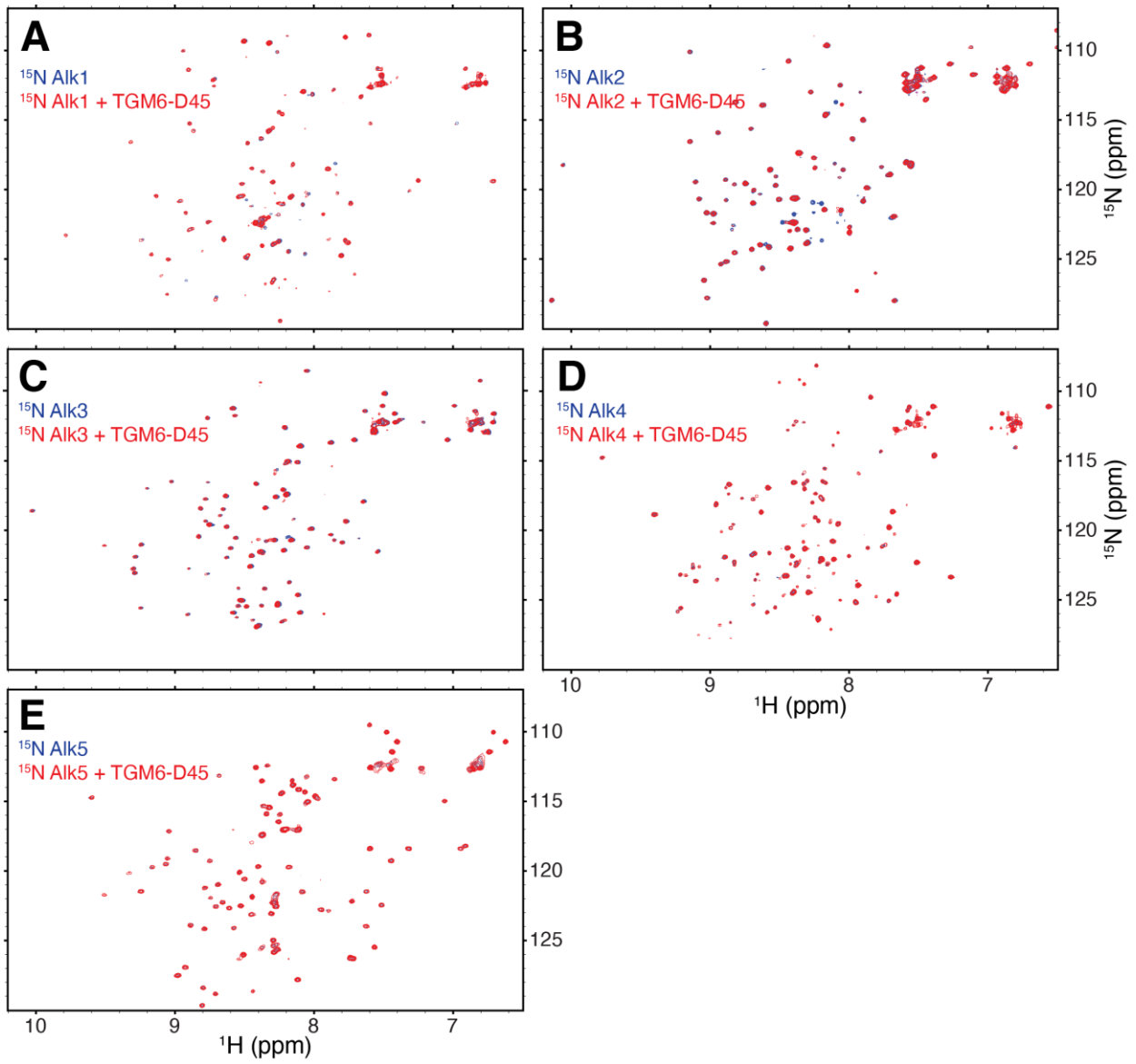

Figure 57. TGM6-D45 does not bind to any type I receptors.^{1}H-^{15}N HSQC spectra of ^{15}N-labeled type I receptors alone (blue) and as bound to an excess of unlabeled TGM6-D45 (red). The receptors tested were: (A) ^{15}N Alk1; (B) ^{15}N Alk2; (C) ^{15}N Alk3; (D) ^{15}N Alk4; and (E) ^{15}N Alk5. None of the NMR signals exhibited changes to their chemical shifts larger than 0.002 ppm in the ^{1}H dimension, whereas shifts of greater than 0.005 ppm in the ^{1}H dimension would be expected in the event of binding.

A lack of binding to TGM6-D45, however, does not rule out the possibility that TGM6 binds a type I receptor through cooperative binding. This type of binding is seen in the TGF-β:TβRII:Alk5 complex, where TGF-β and TβRII come together to form the high-affinity binding site for Alk5 [64, 211]. Alk5 does not bind to TGF-β or TβRII alone with high affinity, but requires

a composite binding interface created between TGF-β and TβRII to bind and form the ternary complex.

The possibility of a composite binding interface was also tested by titrating ^{15}N-labeled type I receptor ectodomains with unlabeled flTGM6:TβRII binary complexes and comparing the ^{1}H-^{15}N HSQC spectra of the unbound receptor with that of the receptor with the flTGM6:TβRII binary complex added in excess. For the five type I receptors tested (Alk1, Alk2, Alk3, Alk4, and Alk5; Figure 70A-E, respectively), there were little to no chemical shift perturbations present upon the addition of the flTGM6:TβRII binary complex. As with the TGM6-D45 titrations, the small perturbations are most likely due to weak or non-specific binding and are not of biological relevance. Therefore, TGM6 does not bind a type I receptor, either through domains 4 and 5 or through a composite interface created upon the binding of TGM6 to TβRII.

4.2.4 TGM6-D3 competes for the TGF-β binding site on TβRII

Competition between TGM1-D3 and TGF-β for the TGF-β binding site on TβRII was previously reported [179]. Since TGM6-D3 has close homology to TGM1-D3 and binds TβRII with high affinity, it was hypothesized that it also competes for the TGF-β binding site on TβRII. Therefore, ITC experiments were performed to test these hypotheses. An engineered TGF-β monomer, known as mmTGF-β2-7M2R (Table 23), was used instead of TGF-β1 or -β3. Like these TGF-β isoforms, mmTGF-β2-7M2R has an intact finger region that binds TβRII with the same affinity as TGF-β1 and -β3 (ca. 50 nM); but unlike TGF-β1 or -β3, mmTGF-β2-7M2R is highly soluble at neutral pH [118]. Additionally, TGM6-D3 was used instead of flTGM6 since it was previously shown that TGM6-D3 holds the full binding capacity for TβRII.

TβRII was titrated into the sample cell loaded with mmTGF-β2-7M2R alone or mmTGF-β2-7M2R with TGM6-D3 (Figure 58A). The addition of TGM6-D3 both increased the extent of curvature in the binding isotherms and reduced the overall enthalpy of the reaction. This is consistent with the behavior expected for competitive binding [243]. To quantify this interaction, the integrated heats from the two experiments were globally fit to a simple competitive binding model. The previously measured K_D and interaction enthalpy for the lower-affinity TGM6-D3:TβRII interaction were held constant, and the binding parameters for the high-affinity mmTGF-β2-7M2R:TβRII interaction in the absence of the competitor were derived (Table 4). The K_D for the high-affinity mmTGF-β2-7M2R:TβRII was calculated to be 19 nM ($\pm 1\sigma$ confidence interval: 7.8 nM – 40 nM), which is in rough agreement with the previously determined K_D between mmTGF-β2-7M and TβRII of ca. 50 nM [119]. Therefore, TGM6-D3 and TGF-β compete for the same binding site on TβRII.

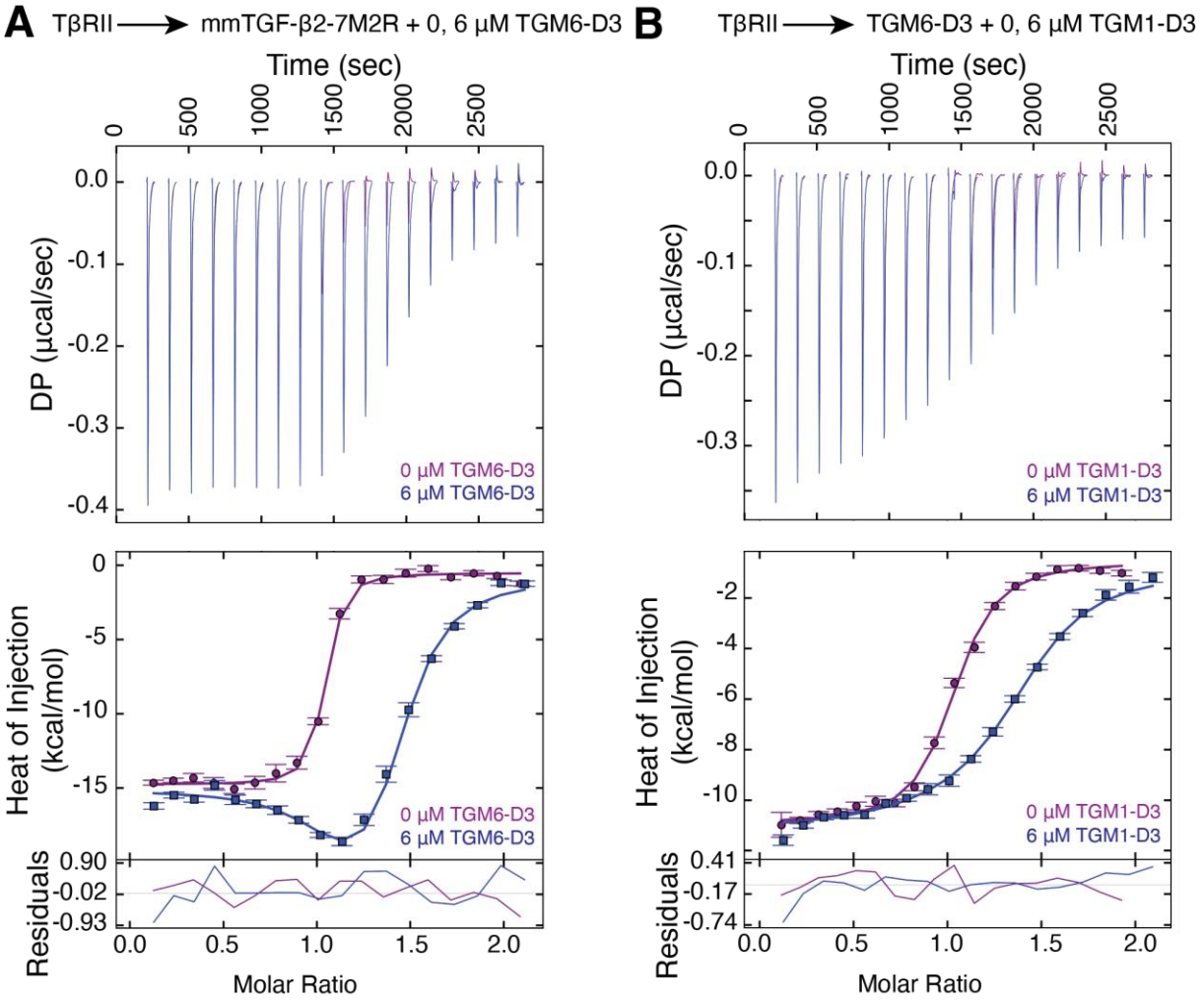

Figure 58. TGM6-D3 competes with TGF-β and TGM1-D3 for the same binding site on TβRII. ITC thermograms show changes to binding affinity and interaction enthalpy when increasing the amount of low-affinity binder. (A) TβRII is titrated into mmTGF-β2-7M2R in the absence (purple) or presence (blue) of 6 µM TGM6-D3 as the lower-affinity binder. (B) TβRII is titrated into TGM6-D3 in the absence (purple) or presence (blue) of 6 µM TGM1-D3 as the lower-affinity binder. Each panel includes the thermograms (top), fitted isotherms (middle) and fitting residuals (bottom) for the associated titrations. Note: the error for some data points is small, and therefore the error bars are smaller than the marker for the data.

Table 4. TβRII competition binding as assessed by ITC

Cell	15 μM mmTGF-β2-7M2R	15 μM TGM6-D3
Syringe	105 μM TβRII	150 μM TβRII
Competitor [a]	0.0 or 6.0 μM TGM6-D3	0.0 or 6.0 μM TGM1-D3
Temperature (°C)	35	25
K_D (nM)	19 (7.8, 40)[b]	190 (110, 310) [c]
ΔH (kcal mol^{-1})	-14 (-15, -13)[b]	-10 (-11, -9.9) [c]
ΔG (kcal mol^{-1})	-11[d]	-9.2[e]
-TΔS (kcal mol^{-1})	3.4 [d]	1.2[e]

[a] Competitor was added to the sample cell.
[b] K_D and ΔH correspond to the parameters, derived from the global fit, for TβRII:mmTGF-β2-7M2R binding in the absence of competitor; uncertainty is reported as the limits of the $\pm 1\sigma$ confidence interval.
[c] K_D and ΔH correspond to the parameters, derived from the global fit, for TGM6-D3:TβRII binding in the absence of competitor; uncertainty is reported as the limits of the $\pm 1\sigma$ confidence interval.
[d] ΔG and -TΔS correspond to those for TβRII:mmTGF-β2-7M binding in the absence of competitor calculated from $\Delta G = \Delta H - T\Delta S$ and globally fitted values for K_D and ΔH.
[e] ΔG and -TΔS correspond to those for TβRII:TGM6-D3 binding in the absence of competitor calculated from $\Delta G = \Delta H - T\Delta S$ and globally fitted values for K_D and ΔH.

As stated previously, TGM1-D3 also competes for the TGF-β binding site on TβRII; therefore, a competition experiment between TGM6-D3 and TGM1-D3 was also performed by titrating TβRII into the sample cell loaded with TGM6-D3 alone or TGM6-D3 with TGM1-D3 (Figure 58B). The changes to the isotherm upon addition of TGM1-D3 again indicated competitive binding [243]. The interaction was quantified by globally fitting to a simple competitive binding model and determining the parameters for the higher-affinity TGM6-D3:TβRII interaction while holding the previously reported lower-affinity TGM1-D3:TβRII interaction parameters constant [179]. From this binding model, the binding constant for the high-affinity TGM6-D3:TβRII interaction in the absence of the competitor (Table 4) was determined to be 190 nM ($\pm 1\sigma$ confidence interval: 110 nM – 310 nM). This agrees with the previously determined K_D between TGM6-D3 and TβRII (Table 3). Therefore, TGM6-D3 and TGM1-D3 compete for the same binding site on TβRII. Additionally, this confirms that TGM6-D3 is a competitor for the TGF-β binding site on TβRII since TGM1-D3 and mmTGF-β2-7M also compete for the TGF-β binding site on TβRII [179].

4.2.5 Structural characterization of TGM6-D3 as bound to TβRII

To compare the binding interactions between TβRII and TGM6-D3 with those between TβRII and TGF-β, the TβRII:TGM6-D3 binary complex was crystallized, and its structure was determined. The binary complex was prepared at 50 mg mL^{-1} due to the high solubility of the protein complex, and crystals were grown under hanging drop vapor diffusion conditions in 0.1 M sodium cacodylate, 25% (w/v) PEG 4000, pH 6.5. Crystals appeared in less than two days and grew to a maximum size in about five days. The crystals that grew under these conditions were clusters of flat, planar crystals emanating from a central point in a star-burst-like pattern. Molecular replacement was used to phase the X-ray density map using the 1.1 Å TβRII X-ray structure (PDB 1M9Z) [21] and the TGM1-D3 NMR structural ensemble (PDB 7SXB) [179] as search models. The final binary complex structure was determined to 1.45 Å (Table 5, Figure 59).

Table 5. Crystallographic data, phasing, and refinement of the TGM6-D3:TβRII complex

λ (Å)	1.00000
Resolution Range (Å)[a]	34.35 – 1.45 (1.502 – 1.45)
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	55.877, 130.663, 29.755
α, β, γ (°)	90.00, 90.00, 90.00
Number of observations	509436 (40410)
Number of unique reflections	39024 (3566)
Multiplicity	13.1 (11.3)
Completeness (%)[a]	98.01 (90.61)
Mean I/sigma(I)	13.38 (0.86)
Wilson B-factor	24.66
R-merge	0.08526 (1.761)
R-meas	0.08889 (1.844)
R-pim	0.02469 (0.5345)
CC1/2	0.998 (0.714)
CC*	0.999 (0.913)
Refinement	
Reflections used in refinement	38 912 (3 549)
Reflections used for R-free	1 942 (177)
R_{work} (%)[b]	0.2445 (0.4005)
R_{free} (%)[c]	0.2664 (0.3817)
CC(work)	0.940 (0.813)
CC(free)	0.942 (0.818)
Number of:	
Non-hydrogen atoms	1 747
Macromolecules	1 602
Ligands	0
Solvent	145
Protein Residues	19
R.m.s. deviation	
Bonds (Å)	0.008
Angles (°)	1.15
Ramachandran Statistics:	
favored, allowed, outliers (%)	95.83, 4.17, 0.00
Rotamer outliers (%)	3.12
Clashscore	8.66
B-factors:	
Average	40.83
Macromolecules	40.66
Solvent	42.69

[a] The numbers in parentheses correspond to the last shell.

[b] $R_{work} = \Sigma||Fo,p| - |Fc,p|| / \Sigma|Fo,p|$.

[c] $R_{free} = \Sigma||Fo,p| - |Fc,p|| / \Sigma|Fo,p|$, where |Fo,p| is from a test set not used in the structural refinement (36 973 reflections).

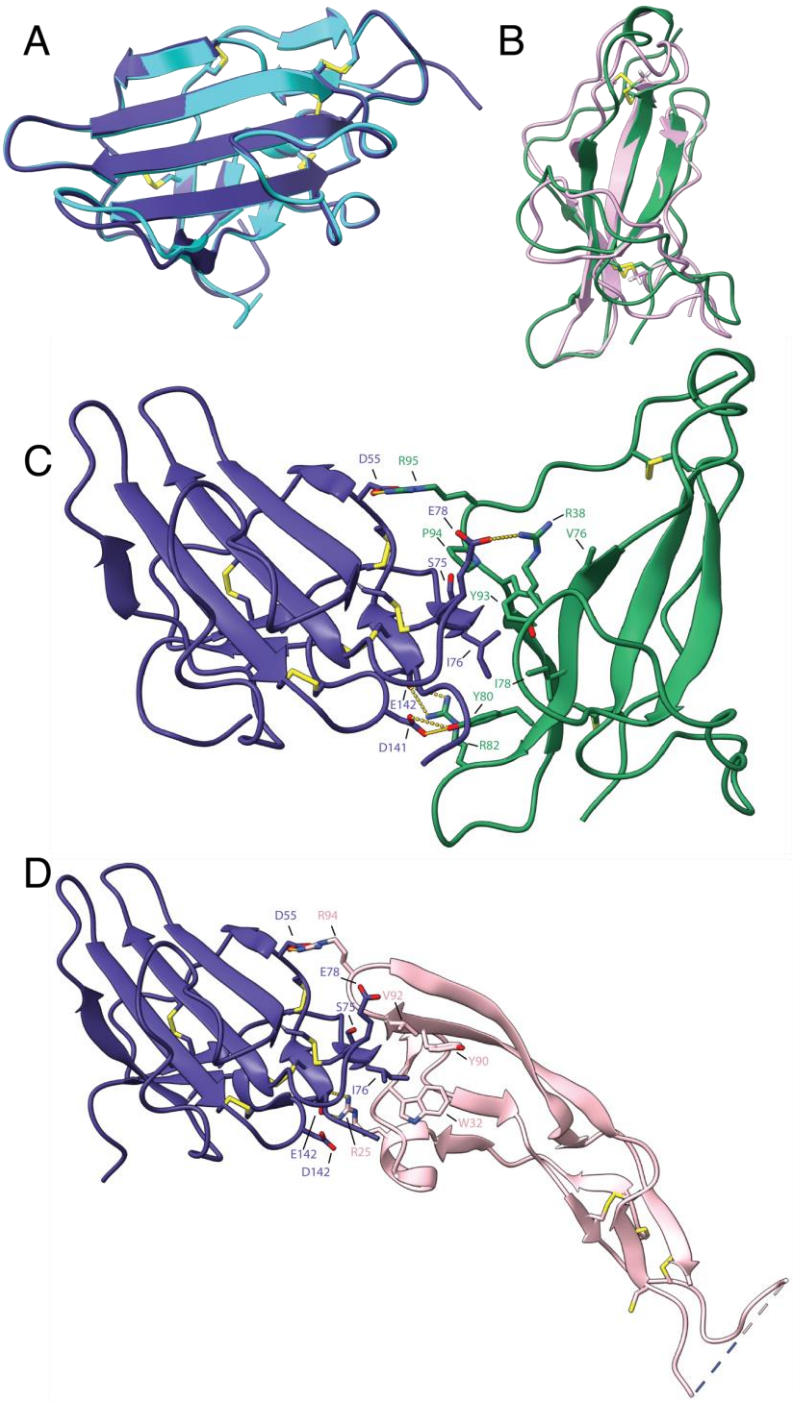

Figure 59. Crystal structure of the TGM6-D3:TβRII binary complex solved to 1.45Å. (A) The TβRII component of the crystal structure, purple, overlaid with unbound TβRII, cyan (PDB ID 1M9Z [21]). (B) The TGM6-D3 component of the crystal structure, green, overlaid with a representative TGM1-D3 structure, pink (PDB ID 7SXB [179]). (C) TβRII, purple, bound to TGM6-D3, green. (D) TβRII, purple, bound to TGF-β3, pink (PDB ID 1KTZ [75]). TGM6-D3 mimics the binding of TGF-β at the level of individual residues. Note: C and D are in the same orientation relative to the TβRII structures.

Residues 46-153 of the TβRII construct are visible in the X-ray density. There is excellent structural agreement between the TβRII component of the TGM6-D3:TβRII crystal structure and the unbound TβRII structure (PDB ID 1M9Z) [21] (Figure 59A).

All residues (16-102) of the TGM6-D3 construct are visible in the X-ray density, including the loops between the various β-strands. The TGM6-D3 component of the binary complex adopts a four-strand anti-parallel β-sheet fold similar to that of TGM1-D3 (PDB 7SXB) [179] (Figure 59B), complete with the ordered hypervariable loop between β-strands 1 and 2 and the short α-helical segment between loop between β-strands 2 and 3.

The TβRII:TGM6-D3 binary complex structure (Figure 59C) is shown in comparison to the TβRII:TGF-β3 structure (Figure 59D, PDB 1KTZ [75]). The specific interactions between TβRII and TGM6-D3 are an excellent mimic of those between TβRII and TGF-β. On the TβRII side, the binding interface is made up of various loops in the structure, spanning the length of the entire sequence. On the TGM6-D3 side, the binding interface is made up primarily from β-strand 4 with its following structured loop toward the C-terminus of the protein. To assess the relative contribution of individual residues of TGM6-D3 or TβRII to binding, single residues were substituted and their binding to the partner protein as characterized using either surface plasmon resonance (SPR) and/or ITC. The SPR and ITC binding data are summarized in Table 6 and Table 7, respectively, with the SPR sensorgrams and the ITC thermograms and integrated heats shown in Figure 73 and Figure 74, respectively.

Table 6. TGM6-D3 variant binding to a Biotin-Avi-TβRII surface as assessed by SPR

Surface	Analyte	Kinetic Fitted Parameters			
		k_{on} (M^{-1} sec^{-1})	k_{off} (s^{-1})	K_D (nM)	R_{max} (RU)
TβRII	WT TGM6-D3	1.34×10^6	0.428	0.32	241.4
TβRII	R38A	1.93×10^5	1.74	9.2	141.25
TβRII	V76A	1.46×10^6	2.925	2.2	125.3
TβRII	I78A	2.01×10^5	1.373	7.0	123.4
TβRII	Y80A	ND[a]	ND[a]	ND[a]	ND[a]
TβRII	Y80F	9.96×10^5	1.381	1.4	109.3
TβRII	Q91A	1.62×10^6	0.479	0.31	185.1
TβRII	Y93A	1.41×10^4	0.636	41	99.5
TβRII	R95A	5.56×10^5	2.09	3.5	98.6
TβRII	P94K,R95N	1.71×10^6	0.374	0.21	207.2

[a] Not determined due to artifacts; see Figure 73.

Table 7. WT TGM6-D3:TβRII variant and WT TβRII:TGM6-D3 variant binding as assessed by ITC at 25°C

Syringe	Cell	N (sites)	K_D (μM)	ΔH (kcal mol^{-1})	ΔG (kcal mol^{-1})	-TΔS (kcal mol^{-1})
426 μM TGM6-D3 WT	25 μM TβRII WT	0.68[a]	0.35 (0.28 – 0.43)[b]	-16.3 (-16.8 – -15.9)[b]	-8.8	7.5
290 μM[a] TGM6-D3 WT	16 μM TβRII D55N	0.62[a]	1.57 (1.37 – 1.79)[b]	-11.1 (-11.5 – -10.7)[b]	-7.9	3.2
290 μM[a] TGM6-D3 WT	24.5 μM TβRII S75L	1.55[a]	1.96 (1.52 – 2.53)[b]	-10.3 (-11.0 – -9.8)[b]	-7.7	2.5
290 μM[a] TGM6-D3 WT	18.4 μM TβRII I76A	0.69[a]	7.35 (6.40 – 8.51)[b]	-10.6 (-11.4 – -10.0)[b]	-7.0	3.6
290 μM[a] TGM6-D3 WT	18.9 μM TβRII D141A	0.75[a]	0.94 (0.87 – 1.00)[b]	-11.5 (-11.7 – -11.4)[b]	-8.2	3.3
400 μM[a] TGM6-D3 WT	30.5 μM TβRII E142Q	1.38[a]	3.94 (3.58 – 4.34)[b]	-8.7 (-8.9 – -8.4)[b]	-7.4	1.3
300 μM TβRII WT	10 μM TGM6-D3 WT	0.87[a]	0.36 (0.31 – 0.41)[b]	-11.3 (-11.6 – -11.1)[b]	-8.8	2.5
667 μM TβRII WT	25 μM TGM6-D3 R38A	0.85[a]	8.03 (7.61 – 8.48)[b]	-17.0 (-17.4 – -16.6)[b]	-7.0	10.0
150 μM TβRII WT	10 μM TGM6-D3 V76A	0.78[a]	1.94 (1.50 – 2.54)[b]	-18.4 (-20.3 – -16.9)[b]	-7.8	10.6
667 μM TβRII WT	25 μM TGM6-D3 I78A	0.71[a]	5.56 (5.24 – 5.71)[b]	-17.5 (-17.6 – -17.3)[b]	-7.2	10.3
667 μM TβRII WT	25 μM TGM6-D3 Y80A	0.33[a]	48.4 (38.1 – 61.8)[b]	-21.5 (-26.6 – -17.7)[b]	-5.9	15.6
150 μM TβRII WT	10 μM TGM6-D3 Y80F	0.67[a]	1.25 (1.04 – 1.49)[b]	-9.6 (-10.2 – -9.1)[b]	-8.1	1.6
150 μM TβRII WT	10 μM TGM6-D3 Q91A	0.81[a]	0.46 (0.36 – 0.60)[b]	-9.4 (-9.9 – -9.0)[b]	-8.6	0.8
667 μM TβRII WT	25 μM TGM6-D3 Y93A	0.17[a]	39.0 (30.4 – 50.8)[b]	-19.2 (-23.2 – -16.3)[b]	-6.0	13.2
150 μM TβRII WT	10 μM TGM6-D3 R95A	1.0[a]	3.82 (1.50 – 15.18)[b]	-13.3 (-30.3 – -9.4)[b]	-7.4	5.9
150 μM TβRII WT	10 TGM6-D3 P94K R95N	0.90[a]	0.27 (0.23 – 0.30)[b]	-13.4 (-13.6 – -13.1)[b]	-9.0	4.4

[a] Number of sites determined by incompetent fraction value on Sedphat; set to '1' for thermodynamic analysis.

[b] Uncertainty reported as 68.3% confidence interval.

The overall architectures of the interface between TGM6-D3 and TβRII mirrors that between TGF-β1/-β3 and TβRII, with a central hydrophobic portion flanked by polar/charge interactions at the periphery. On one side of the interface, TβRII Asp-55 forms a salt bridge with TGM6-D3 Arg-95, which closely mimics the interaction between TβRII Asp-55 and TGF-β Arg-94. Mutation of TGM6-D3 Arg-95 to Ala (R95A) results in about a 12-fold decrease (WT K_D = 320 nM, R95A K_D = 3.8 μM) in affinity between the two proteins. Mutation of TβRII Asp-55 to Asn (D55N) is also disruptive, but is only about 6-fold, not 12-fold (WT K_D = 320 nM, D55N K_D = 1.4 μM).

The residues of TGM6-D3 that interact with TβRII are conserved in TGM1-D3, apart from TGM6-D3 Arg-95, which substituted with an asparagine in TGM1-D3. While the functional group of the asparagine sidechain in TGM1-D3 can form a hydrogen bond with the TβRII aspartate, the side chain is much shorter in length than the arginine present in TGM6-D3. This would prevent the formation of the hydrogen bonds due to the distance between the two groups. The amide sidechain is furthermore uncharged, and thus would be unable to electrostatically interact with the TβRII asparate. In TGM6, the preceding proline residue in TGM6-D3 is compact and nicely accommodated in a hydrophobic pocket created by Ser-75, Phe-53, and the Cys-77-Cys-110 disulfide bond. In TGM1-D3, the equivalent residue is a lysine, which is large and has a charged functional group at the end of its sidechain. Though it is conceivable that there might be some type of rearrangement that allows the Lys-Asn dipeptide in TGM1 to interact with TβRII Asp-55 in a way that provides a similar contribution to binding as Pro-Arg dipeptide in TGM6, it seems unlikely. Thus, we hypothesized that substitution of Pro-Arg dipeptide in TGM6-D3 with the Lys-Asn dipeptide in TGM1-D3 was responsible for its approximate 4-fold weaker affinity for TβRII. To test this hypothesis, the TGM6-D3 P94K R95N double mutant was generated and its affinity

for TβRII was measured by SPR (Table 6, Figure 73J). These mutations did not impair binding as expected (WT K_D = 320 nM; P94K R95N K_D = 220 nM), thus this hypothesis is incorrect. This suggests that the differences between the TGM proteins that bind TβRII are non-obvious regarding the binding interface between TβRII and the TGM domain, regardless of the high degree of homology between the TGM domains.

In the center of the interface, TβRII Ile-76 inserts into a hydrophobic pocket created by TGM6-D3 residues Ile-78, Tyr-80, and Tyr-93. This interaction mimics the interaction between TβRII Ile-76 and the hydrophobic patch present between the fingers of TGF-β, characterized by Trp-32 emanating from the loop connecting fingers 1-2 and Val-92 and Tyr-90 emanating from finger 3. Mutation of TGM6-D3 Tyr-80 or Tyr-93 to Ala (Y80A or Y93A, respectively) resulted in a greater than 30-fold decrease in affinity (WT K_D = 320 nM, Y80A K_D = 25 μM, Y93A K_D = 13 μM) and a large attenuation of the interaction enthalpy (WT ΔH = -11 kcal mol^{-1}, Y80A ΔH = -6.0 kcal mol^{-1}, Y93A ΔH = -2.6 kcal mol^{-1}). Mutation of TβRII Ile-76 to Ala (I76A) decreased the affinity by 22-fold (WT K_D = 320 nM, I76A K_D = 7.2 μM) but did not have a large attenuation of the interaction enthalpy (WT ΔH = -11 kcal mol^{-1}, I76A ΔH = -11 kcal mol^{-1}).

In TGM6-D3, residues Arg-38, Val-76, and Gln-91, which are nearby Ile-78, Tyr-80 and Tyr-93, were also each mutated to Ala (R38A, V76A, and Q91A, respectively). The R38A mutation led to a 28-fold decrease in affinity (WT K_D = 320 nM, R38A K_D = 9.0 μM), which suggests that may be responsible for orienting Tyr-93 to optimally interact with TβRII as the Arg-38 sidechain sterically hinders movement of Tyr-93 in the structure of the TGM6-D3:TβRII complex. TGM6-D3 Arg-38 forms a salt bridge with TβRII Glu-78, holding it in place as well. The V76A mutation led to a 6-fold decrease in affinity (WT K_D = 320 nM, V76A K_D = 2.0 μM), suggesting that it was also responsible for orienting Ile-78 and Tyr-80 to optimally interact with

TβRII; Val-76 is a core residue, but being to the N-terminal side of both Ile-78 and Tyr-80, it is important for the arrangement of these two residues in the TGM6-D3:TβRII structure. Although nearby, the substitution of Gln-91 with alanine (Q91A) had a negligible impact on the binding affinity (WT K_D = 320 nM, Q91A K_D = 200 nM).

In TβRII, the residue Ser-75, which is proximal to Ile-76, was substituted with leucine (S75L). This mutation led to a 6-fold decrease in affinity (WT K_D = 320 nM, S75L K_D = 1.97 μM).

On the side of the interface opposite the TGM6-D3 Arg-95:TβRII Asp-55 interaction, TβRII Asp-141 forms a hydrogen-bonded ion pair with the hydroxyl group of the TGM6-D3 Tyr-80 sidechain, likely holding Tyr-80 in place as part of the hydrophobic pocket. This interaction mimics the hydrogen-bonded salt bridge between TβRII Glu-142 and TGF-β Arg-25. Substitution of Tyr-80 with Phe (Y80F), which preserves the hydrophobic interaction with TβRII I75 but removes the hydrogen bond interaction, decreased the affinity 4-fold (WT K_D = 320 nM, Y80F K_D = 1.4 μM). Mutation of TβRII Asp-141 to Ala (D141A) yielded similar result (WT K_D = 320 nM, D141A K_D = 935 nM): a 3-fold decrease in affinity. This confirms that the phenolic hydroxyl group plays a role in organizing the hydrophobic pocket and aids in the specificity of the interaction.

TβRII Glu-142 also plays a role in the interaction between TβRII and TGM6-D3. TβRII Glu-142 forms a salt bridge with TGM6-D3 Arg-82. Mutation of TβRII Glu-142 to Gln (E142Q) led to a 9-fold reduction in the affinity between TβRII and TGM6-D3 (WT K_D = 320 nM, E142Q K_D = 2.78 μM).

4.2.6 Full-length TGM6 acts as a TGF-β signaling antagonist in murine fibroblasts

TGM6 binds TβRII, but it does not bind the TGF-β type I receptor, Alk5, or any of the other type I and type II receptors tested. This is consistent with previously published results that TGM6 does not signal through the TGF-β pathway in the MFB-F11 TGF-β reporter cell line or convert näive T-cells into Tregs [246]. This prompted the hypothesis that TGM6 might function as a TGF-β or TGM1 signaling antagonist. Both TβRI and TβRII are required for TGF-β and TGM1 signaling, and therefore, occupancy of cell surface of TβRII by TGM6 would serve to compete TβRII away from TGF-β and TGM1.

To test this hypothesis, signaling inhibition assays were performed. Mouse NIH-3T3 fibroblasts stably transfected with the $CAGA_{12}$ TGF-β reporter [254] and mouse embryonic fibroblasts (MFB-F11) containing a TGF-β-sensitive secreted alkaline phosphatase (SEAP) reporter were incubated with differing amounts of flTGM6 and stimulated with either TGF-β or TGM1. Both cell lines showed a dose-dependent decrease in TGF-β signaling as the flTGM6 concentration was increased (Figure 60A,B). In both sets of assays, the IC_{50} value was calculated. TGM6 has an IC_{50} of 0.25 nM for inhibition of both TGF-β1- and TGM1-induced TGF-β signaling in both cell lines. This inhibitory concentration is 1300-fold lower than its affinity for TβRII, ca. 320 nM, suggesting that other cell-endogenous factors influence its inhibitory activity against TGF-β1 and TGM1.

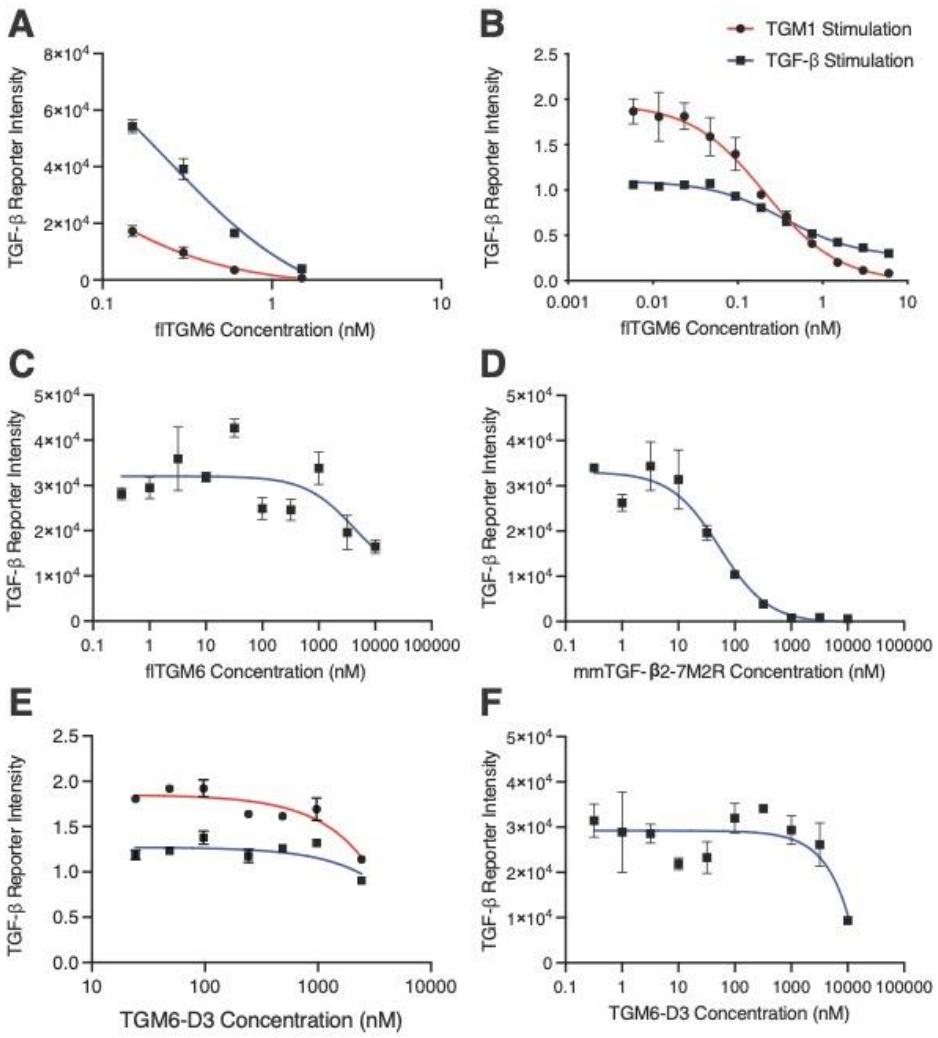

Figure 60. flTGM6, but not TGM6-D3, is a potent inhibitor of TGF-β signaling in murine fibroblasts. (A,B) flTGM6 inhibits TGM1-induced (red, circles) (100 nM TGM1 stimulation) and TGF-β-induced (blue, squares) (200 nM TGF-β1 stimulation) TGF-β signaling in NIH-3T3 (A) and MFB-F11 (B) cells containing TGF-β-sensitive reporters with an IC_{50} value of 0.25 nM. (C) flTGM6 does not inhibit TGF-β-induced (10 pM TGF-β3 stimulation) TGF-β signaling in HEK-293 cells except at the highest concentrations tested. (D) The TGF-β signaling inhibitor mmTGF-β2-7M2R inhibits TGF-β-induced (10 pM TGF-β3 stimulation) TGF-β signaling with an IC_{50} approximately equal to that of its K_D for TβRII, 50 nM. (E,F) TGM6-D3 does not inhibit TGF-β- or TGM1-induced TGF-β signaling in MFB-F11 (E) (200 nM TGF-β1 stimulation; 100 nM TGM1 stimulation) or HEK-392 (F) (10 pM TGF-β3 stimulation) cells containing a TGF-β-sensitive reporter at concentrations below 1000 nM. Note: the error for some data points is small, and therefore the error bars are smaller than the marker for the data.

The assay was repeated in human embryonic kidney (HEK-293) cells containing the stably transfected $CAGA_{12}$ TGF-β reporter [254], incubating the cells with differing amounts of flTGM6 and stimulating them with TGF-β3. However, the reporter did not show inhibition of TGF-β

signaling at flTGM6 concentrations lower than 1000 nM (IC_{50} > 1000 nM) (Figure 60C). This was surprising since it was hypothesized that flTGM6 acted as an inhibitor regardless of the cell line. These results were compared to the previously reported mmTGF-β2-7M2R inhibitor [119], which showed inhibition in HEK-293 cells with an IC_{50} of 50 nM (Figure 60D). This is equivalent to its K_D for TβRII. Therefore, flTGM6 functions as a potent TGF-β signaling antagonist in mouse fibroblast cells, but not human kidney cells.

4.2.7 TGM6-D3 does not antagonize TGF-β signaling

To further investigate the hypothesis that TGM6 binds TβRII and sequesters the type II receptor away from the signaling proteins, the activity assays were repeated with TGM6-D3 alone. TGM6-D3 was expected, at best, to inhibit TGF-β- and TGM1-induced TGF-β signaling with an IC_{50} approximately equal to its K_D for TβRII, approximately 320 nM. The inhibition assays, performed in the MFB-F11 and HEK293 reporter cell lines, however, showed no inhibition below 1000 nM TGM6-D3 (IC_{50} > 1000 nM) (Figure 60E,F). This is similar to the behavior exhibited by flTGM6 when it was tested as an inhibitor in the HEK-293 cell line.

4.3 Discussion

Among the mammalian TGF-β family members, the signaling molecules require the binding of both a type I receptor and a type II receptor for signaling [126]. This paradigm is true for the three isoforms of TGF-β, TGF-β1, -β2, and -β3, which bring together Alk5 as the type I receptor and TβRII as the type II receptor [64, 90]. For TGM1, it was previously shown that

TGM1-D1 and -D2 are responsible for binding Alk5 while TGM1-D3 binds TβRII, thus forming a complex that mimics the function of the TGF-β:Alk5:TβRII complex [179]. These data are consistent with additional reports that TGM1-D1, -D2, and -D3 are necessary and sufficient for TGF-β signaling activity [246].

TGM6 is unique among the TGM family members that are expressed during the adult stages of the parasite's lifecycle in that it lacks domains 1 and 2 [246]. Although TGM6 is similar in sequence to TGM1 (45.6% identity over its full-length and 69.5% identity for D3 alone), it does not signal through the canonical TGF-β pathway, as determined in the TGF-β-responsive MFB-F11 reporter assay, nor does it convert naïve T-cells into Tregs [246]. As shown here, TGM6 only binds TβRII, but not the other TGF-β family type II receptors ActRII, ActRIIB, and BMPRII. Additionally, TGM6 neither binds Alk5, nor other type I receptors of the TGF-β family, Alk1, Alk2, Alk3, Alk4, and Alk5, either through -D45 or with a composite interface formed by the flTGM6:TβRII binary complex. Although TGM6 has been tested against the majority of the TGF-β family type II and type I receptors, it has not been tested against the type II receptor AMHRII or the type I receptors Alk6 and Alk7, thus binding to these receptors cannot be ruled out.

TGM6-D3 holds the full binding capacity for TβRII with an affinity of ca. 320 nM. Additionally, TGM6-D3 competes with TGF-β and TGM1-D3 for the TGF-β binding site on TβRII, as shown by competition ITC experiments. This was further verified by determining the structure of the TGM6-D3:TβRII binary complex. TGM6-D3 engages nearly the same set of residues of TβRII as TGF-β and does so using an interface with a similar overall architecture. TβRII Asp-55 forms a doubly-bonded ion pair with TGF-β residue Arg-94, which is mimicked by TGM6-D3 residue Arg-95. This sidechain of TGM6-D3 Arg-95 is held in an extended conformation to ideally pair with the TβRII Asp-55 by the phenyl ring of the TβRII Phe-52

sidechain, as the sidechain of Arg-94 in the TβRII:TGF-β complex is. The two hydrogen bonds between TβRII residue Glu-142 and TGF-β residue Arg-25 are mimicked by TβRII residue Asp-141 and the phenolic hydroxyl group of TGM6-D3 Tyr-80. Finally, the hydrophobic pocket in TGF-β into which TβRII residue Ile-75 inserts is mimicked by TGM6-D3 residues Ile-78, Tyr-80, and Tyr-93. This is a further example of the TGM family's remarkable mimicry of TGF-β while having no evolutionary homology to the mammalian cytokine.

It is interesting to note that the substitution of TβRII Ser-75 with leucine (S75L) did not result in a large attenuation of the binding affinity, only a 6-fold decrease. In the TGM1-D3:TβRII interaction, this substitution caused a much greater attenuation (ca. 300-fold decrease) of the affinity between these two proteins [179]. This suggests that TβRII Ser-75 could play a role in the arrangement of the residues in the binding interface but does not directly play a significant role in binding between TGM6-D3 and TβRII. This is in direct contrast to TβRII Ser-75 playing a significant role in binding between TGM1-D3 and TβRII: due to the differences in the sequence between TGM6-D3 and TGM1-D3, it is likely that the TβRII S75L substitution leads to steric clashes in the interface with TGM1-D3 that are not observed with TGM6-D3 due to the differences in the protein sequences.

In contrast to TGM1, which is a signaling protein, full-length TGM6 acts as a potent inhibitor of TGF-β- and TGM1-induced TGF-β signaling in mouse fibroblast cells. Full-length TGM6 has an IC_{50} of about 0.25 nM against both ligands, which is about 1300-fold lower than the measured K_D between TβRII and both full-length TGM6 and TGM6-D3, about 320 nM. This is unexpected as the previously reported TGF-β-based signaling inhibitor mmTGF-β2-7M2R, which is thought to function in the same manner, has an IC_{50} comparable to its affinity for TβRII, about 50 nM [119]. TGM6 also has other unexpected properties that do not adhere to a simple model of

competitive inhibition, including its inability to inhibit TGF-β signaling in human kidney cells and the inability of TGM6-D3 to inhibit TGF-β signaling in any of the cell lines tested, except at micromolar concentrations or higher.

One possible explanation for the observed behavior is that the inhibitory activity of TGM6 is that domains 4 and/or 5 bind to a co-receptor that increases the overall affinity to TβRII and the cell surface by avidity. However, in light of the cell-based inhibitory data that showed that TGM6-D3 alone was a significantly poorer inhibitor than expected, these data suggest the co-receptor is not only increasing the potency by avidity through binding to domains 4 and 5, but also serving in some way to unmask the inhibitory potential of TGM6-D3. The molecular mechanism responsible for this is not known but might be some type of repulsive interaction of -D3 with the cell surface that is overcome by the presence of -D4 and -D5, together with co-receptor binding. Physiologically, the co-receptor may function as an "address" for the inhibitor, guiding it to specific cells in which TGF-β- or TGM1-induced TGF-β signaling would be detrimental to the parasite or where a lack of TGF-β signaling would be beneficial to the parasite.

The notion of the requirement of a co-receptor for effective inhibition by TGM6 is also suggested by the finding that flTGM6 potently inhibited TGF-β- and TGM1-induced TGF-β signaling in mouse fibroblasts, but not human kidney cells. This could be explained by cell lineage-specific or species-specific expression of the co-receptor and might reflect the parasite taking advantage of these differences to reduce fibrotic activity as it transitions through its life cycle in which it initially inhabits the intestinal epithelium after animals consume dung on the forest floor containing parasite larvae. Upon maturation in the intestinal epithelium, the parasites burrow through the intestinal wall where they mate, produce offspring, and re-enter the intestinal epithelium before being released back to the environment in the feces as larvae. In light of the

considerable tissue damage that would occur in this process, and the well-established role of TGF-β and TGM1 in contributing to wound repair by stimulating deposition of type I collagen [103, 118, 218], but also in tissue fibrosis if signaling is dysregulated [103] it makes sense that TGM6 would competitively reduce the amount of TGF-β signaling taking place in fibroblasts, which in turn would reduce the amount of collagen deposition and fibrosis.

In conclusion, our understanding of the *H. polygyrus* TGF-β mimic family of proteins is expanded by showing that TGM6 is a potent inhibitor of TGF-β- and TGM1-induced TGF-β signaling in mouse fibroblast cells. It was shown that, like TGM1, TGM6 domain 3 is wholly responsible for binding TβRII and competes with TGF-β and TGM1-D3 for the TGF-β binding site on TβRII. Additionally, the structure of the TGM6-D3:TβRII binary complex shows remarkable mimicry of the TGF-β:TβRII interactions. Finally, it was proposed that TGM6 requires a mouse- or tissue-specific co-receptor to direct its inhibitory activity, lending specificity and potency to its inhibitory activity and potentially targeting TGM6 to fibroblasts to minimize tissue damage caused by the parasite and it transitions through its life cycle.

4.4 Materials and Methods

4.4.1 Expression and purification of TGM proteins

Details of the TGM protein constructs used are provided in Table 19. DNA fragments corresponding to TGM6-D3, the TGM6-D3 variants, and TGM6-D45 were inserted between the *Kpn*I and *Hind*III sites for TGM6-D3 and its variants or the *Kpn*I and *Xho*I sites for TGM6-D45 in a modified form of the pET32a vector (EMD-Millipore, Danvers, MA). The pET32a vector was

modified to include a His10 tag instead of the standard His6 and a *Kpn*I site immediately following the sequence of the thrombin cleavage site. The resulting constructs were designed as follows: thioredoxin-decahistidine tag-thrombin cleavage site-TGM6 domain coding cassette. Note: the expressed, un-cleaved TGM6 proteins described in this section are referred to as the TGM6 fusion proteins, and the cleaved proteins are referred to as the TGM6 domain proteins.

The constructs were overexpressed in chemically-competent *E. coli* BL21(DE3) cells (EMD-Millipore, Danvers, MA) cultured at 37°C. Unlabeled samples for binding studies were produced on tris-buffered Lauria Broth (LB); ^{15}N isotopically labeled samples for NMR studies were produced using M9 minimal medium containing 0.1% $^{15}NH_4Cl$ (Sigma-Aldrich, St. Louis, MO). To select for cells bearing the expression plasmid, carbenicillin or ampicillin was added to all media at 100 μg mL^{-1} or 150 μg mL^{-1}, respectively. Protein expression was induced by adding 200 mg L^{-1} IPTG when the optical density of the culture at 600 nm reached 1.0. Expression occurred for 3-4 hours after induction at 37°C, and the cultures were harvested by centrifugation (7,300 g, 15 minutes, 25°C).

Cell pellets from 3 L of culture were resuspended in 100 mL lysis buffer (50 mM Na_2HPO_4, 100 mM NaCl, 5 mM imidazole, pH 8.0) and sonicated for 5 minutes at 50% duty cycle. Following centrifugation (30,000 g, 20 minutes, 25°C), the supernatant was transferred to a clean container, the pellet was resuspended in lysis buffer containing 500 mM NaCl, and the solution was re-centrifuged (30,000 g, 20 minutes, 25°C). Following the second centrifugation, the supernatant was again discarded, and solid urea was dissolved into the solution to a final concentration of 8 M. The pellet was resuspended in lysis buffer containing 1% (v/v) Triton X-100, and the solution was centrifuged for a third time (30,000 g, 20 minutes, 25°C). Following the third centrifugation, the supernatant was discarded, the pellet was resuspended in the 8 M urea-containing lysate, and the

solution was stirred vigorously overnight at 25°C. The remaining insoluble material was removed by centrifugation (30,000 g, 20 minutes, 25°C), and the supernatant was loaded onto a 30 mL metal affinity column (Ni^{++} loaded chelating Sepharose resin; GE Healthcare, Piscataway, NJ) pre-equilibrated with 100 mL of resuspension buffer (25 mM NaH_2PO_4, 8 M urea, 100 mM NaCl, 5 mM imidazole, pH 8.0). The column was washed with 100 mL of resuspension buffer, and the bound protein was eluted using a linear gradient of resuspension buffer containing 0.5 M imidazole.

Protein from the eluted peak was treated with reduced glutathione (GSH) such that the final GSH concentration once the protein is diluted into the folding buffer was 2 mM. Following a 30-minute incubation at 25°C, the protein was slowly diluted into folding buffer (100 mM tris, 1 mM EDTA, 0.5 mM oxidized glutathione (GSSG), pH 8.0) to a final concentration of less than 0.1 mg mL^{-1}. The folding solution was then stirred for 12 – 18 hours at 4°C and subsequently concentrated. To the concentrated protein, solid thrombin was added to a final concentration of 1-2 units mg^{-1} TGM6 fusion protein. The protein mix was dialyzed against 50 mM Na_2HPO_4, 100 mM NaCl, 5 mM imidazole, pH 8.0 for 24 – 36 hours at 4°C. Cleavage was stopped by adding 10 μM leupeptin hemisulfate, 10 μM pepstatin, and 100 μM PMSF. The protein solution was then passed over a Ni^{++} chelating Sepharose column equilibrated with the dialysis buffer (50 mM Na_2HPO_4, 100 mM NaCl, 5 mM imidazole, pH 8.0), and the column was then washed with additional dialysis buffer. The column flow-through and subsequent wash were collected as they contained the cleaved TGM6 domain protein.

For TGM6-D3, the flow-through and wash pool was dialyzed against 25 mM $NaCH_3COO$, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 4.8 at 4°C. The protein solution was then sterile-filtered, bound to a Source S column (GE Healthcare, Piscataway,

NJ) equilibrated in 25 mM $NaCH_3COO$, 2 M urea, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 4.8, and eluted with a 0 – 0.35 M NaCl gradient.

For TGM6-D45, the flow-through and wash pool was dialyzed against 25 mM tris, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 7.5 at 4°C. The protein was then sterile-filtered, bound to a Source S column (GE Healthcare, Piscataway, NJ) equilibrated in 25 mM Tris, 2 M urea, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, pH 7.5, and eluted with a 0 – 0.2 M NaCl gradient.

Masses of the TGM6 domain proteins were confirmed by liquid chromatography electrospray ionization time-of-flight mass spectroscopy (LC-ESI-TOF-MS; Bruker Micro TOF, Billerica, MA) (Figure 71). Folding of the TGM6 domain proteins was assessed through NMR spectroscopy using 2D ^{1}H-^{15}N HSQC spectra, as described in the main text.

TGM1-D3 was overexpressed in *E. coli* at 37°C in the form of insoluble inclusion bodies, refolded, and purified as described previously [43].

flTGM6 was expressed in expi293 cells (Promega, USA) and initially purified by metal affinity chromatography (Ni^{++} loaded chelating Sepharose resin, GE Healthcare, Piscataway, NJ). The bound protein was eluted using a linear gradient of buffer containing 0.5 M imidazole. The eluted protein peak was then concentrated, deglycosylated by treatment with PNGAse F, and further purified by size exclusion chromatography (HiLoad 26/60 Superdex 75 column, GE Healthcare, Piscataway, NJ).

4.4.2 Expression and purification of type I receptors

Details of the Type I receptor constructs used are provided in Table 22.

Alk1, Alk3, and Alk5 were expressed in the form of insoluble inclusion bodies, refolded, and purified as described previously [144, 290].

^{15}N Alk2 was overexpressed in chemically-competent *E. coli* BL21(DE3) cells. Cells were grown at 37°C to an OD_{600} of 0.4 and then transferred to an incubator set to 14°C, where they were grown to an OD_{600} of 0.6, induced with 200 mg L^{-1} IPTG, and allowed to express overnight. Cells were harvested by centrifugation and lysed. The expressed Alk2 protein was harvested from the lysis supernatant in its soluble form using a Ni^{++} loaded chelating Sepharose resin (GE Healthcare, Piscataway, NJ) and eluted with a 0.5 M imidazole gradient, collecting fractions of the elution. The fractions were pooled and dialyzed against 25 mM CHES, pH 9.0 and further purified by ion exchange chromatography (Source 15Q; GE Healthcare, Piscataway, NJ) followed by reverse-phase purification on a C18 semi-preparative column (Jupiter 5μ C18 300A; Phenomenex, Torrance, CA).

^{15}N Alk4 was expressed in the form of insoluble inclusion bodies in chemically-competent *E. coli* BL21(DE3) cells. Cells were grown at 37°C to an OD_{600} of 0.8 and induced with 200 mg L^{-1} IPTG. Expression occurred for 4 hours, after which the cells were harvested by centrifugation and lysed. The inclusion bodies were harvested from the lysate, washed with buffer containing 0.5 M NaCl followed by buffer containing 1% Triton X-100, and solubilized in buffer containing 8 M urea overnight. The expressed Alk4 protein was isolated from the solubilization mixture using a Ni^{++} loaded chelating Sepharose resin (GE Healthcare, Piscataway, NJ) and eluted with a 0.5 M imidazole gradient, collecting fractions of the elution. The fractions were then pooled and refolded

in the presence of a glutathione redox couple for 24 hours at 4°C. The folding solution was then concentrated, cleaved with thrombin, and dialyzed against 25 mM Tris, pH 8.0 containing 10 μM leupeptin hemisulfate, 10 μM pepstatin, and 100 mg L^{-1} benzamidine. The Alk4 ectodomain monomers were then purified from the mixture by subsequent fractionation on high-resolution ion exchange chromatography (Source 15Q; GE Healthcare, Piscataway, NJ). The protein was eluted from the column using buffer containing 2 M urea over a 0.35 M NaCl gradient. The monomers were further purified using reverse-phase purification on a C18 semi-preparative column (Jupiter 5μ C18 300A; Phenomenex, Torrance, CA).

4.4.3 Expression and purification of type II receptors

Details of the type II receptor constructs used are provided in Table 20.

ActRII and ActRIIb were expressed in the form of insoluble inclusion bodies in chemically-competent *E. coli* BL21(DE3) cells. ActRII was grown on Lauria Broth (LB); ActRIIb was grown on M9 minimal media supplemented with 1 g L^{-1} $^{15}NH_4Cl$ (Sigma-Aldrich, St. Louis, MO). Both proteins were grown at 37°C to an OD_{600} of 0.8 and induced with 200 mg L^{-1} IPTG. Expression occurred for 4 hours, after which the cells were harvested by centrifugation and lysed. The inclusion bodies were harvested from the lysate, washed with buffer containing 1.0 M NaCl followed by buffer containing 1% Triton X-100, and solubilized in buffer containing 8 M urea overnight. The expressed proteins were isolated from the solubilization mixture using a Ni^{++} loaded chelating Sepharose resin (GE Healthcare, Piscataway, NJ) and eluted with a 0.5 M imidazole gradient, collecting fractions of the elution. The fractions were then pooled and refolded in the presence of a glutathione redox couple for 24 hours at 4°C. The folding solution was then concentrated, cleaved with thrombin, and dialyzed against 25 mM sodium phosphate, pH 6.6

containing 10 μM leupeptin hemisulfate, 10 μM pepstatin, and 100 mg L^{-1} benzamidine. The protein monomers were then purified from the mixture by subsequent fractionation on high-resolution ion exchange chromatography (Source 15Q; GE Healthcare, Piscataway, NJ). The protein was eluted from the column over a 0.15 M NaCl gradient.

BMPRII was expressed in expi293 cells (Promega, USA) and initially purified by metal affinity chromatography (Ni^{++} loaded chelating Sepharose resin, GE Healthcare, Piscataway, NJ). The bound protein was eluted using a linear gradient of buffer containing 0.5 M imidazole. The eluted protein peak was then concentrated, deglycosylated by treatment with PNGAse F, and further purified by size exclusion chromatography (HiLoad 26/60 Superdex 75 column, GE Healthcare, Piscataway, NJ).

The TGF-β type II receptor, TβRII, and its variants were overexpressed in chemically-competent *E. coli* BL21(DE3) at 37°C in the form of insoluble inclusion bodies, refolded, and purified as described previously [87]. Unlabeled samples for binding studies were produced on tris-buffered Lauria Broth (LB); ^{15}N isotopically labeled samples for NMR studies were produced using M9 minimal medium containing 0.1% $^{15}NH_4Cl$ (Sigma-Aldrich, St. Louis, MO).

All masses were verified by LC-ESI-TOF-MS, and native-like folding was assessed through NMR spectroscopy using 1D ^{1}H spectra for unlabeled samples and 2D ^{1}H-^{15}N HSQC spectra for ^{15}N-labeled samples (Figure 72).

4.4.4 Expression and purification of growth factors

Details of the growth factor constructs used are provided in Table 23. TGF-β3 and the TGF-β2 mini monomer (mm-TGF-β2-7M) were overexpressed in chemically-competent *E. coli* BL21(DE3) cells at 37°C in the form of inclusion bodies, refolded, and purified as previously

described [119]. All masses were verified by LC-ESI-TOF-MS, and native-like folding was assessed through NMR spectroscopy using 1D ^{1}H spectra.

4.4.5 Expression and purification of biotinylated avi-tagged TβRII

Avi-tagged TβRII was produced using constructs modified to include the amino acid sequence "GLNDIFEAQKIEWHE" at the N-terminus. Protein expression and purification were carried out using the same procedures as described previously for the non-tagged proteins. Biotinylation was performed using BirA biotin ligase as previously described [38]. Constructs were validated by LC-ESI-TOF-MS, where the addition of a single biotin increases the protein mass by 226.3 Da. Following biotinylation, the proteins were re-purified using size-exclusion chromatography to remove the biotinylation reagents.

4.4.6 NMR data collection

Samples of ^{15}N TGM6-D3 and its corresponding complex with TβRII were prepared at a concentration of 150 μM in 25 mM Na_2HPO_4, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, 0.05% (w/v) NaN_3, 5% $^{2}H_2O$, pH 5.5. Samples of ^{15}N Alk1, ^{15}N Alk2, ^{15}N Alk3, ^{15}N Alk4, and ^{15}N Alk5 and their corresponding samples containing 1.125 molar equivalents of TGM6-D45 or the flTGM6:TβRII binary complex were prepared at a concentration of 100 μM ^{15}N-labelled receptor in 25 mM Na_2HPO_4, 10 μM leupeptin hemisulfate, 10 μM pepstatin, 100 mg L^{-1} benzamidine, 0.05% (w/v) NaN_3, 5% $^{2}H_2O$, pH 6.0.

All NMR samples were transferred to 5 mm susceptibility-matched microtubes (Shigemi, Allison Park, PA) for data collection. NMR data were collected at 303.15 K using Bruker 600,

700, or 800 MHz spectrometers equipped with 5 mm ^{1}H (^{13}C,^{15}N) z-gradient "TCI" cryogenically cooled probes (Bruker Biospin, Billerica, MA). 2D ^{1}H-^{15}N HSQC spectra were recorded with sensitivity enhancement [111], water flip-back pulses [69], and WATERGATE water suppression pulses [203]. NMR data were processed using NMRPipe [41] and analyzed using NMRFAM-SPARKY [128].

4.4.7 ITC measurements

ITC data were generated using a Microcal PEAQ-ITC instrument (Malvern Instruments, Westborough, MA). All experiments were performed in ITC buffer (25 mM HEPES, 150 mM NaCl, 0.05% NaN_3, pH 7.4) at 25°C with the exception of the TGM6-D3:TβRII replicates (Table 3) and the TβRII competition binding experiments between TGM6-D3 and mmTGF-β2-7M2R, which were performed at 35°C. The proteins in the syringe and sample cell for each experiment are given in Table 3 and Table 7. Prior to each experiment, all proteins were dialyzed three times against ITC buffer and were concentrated or diluted as necessary before being loaded into the sample cell or syringe. For each experiment, nineteen 2.0 μL injections were performed with an injection duration of 4 sec, a spacing of 150 sec, and a reference power of 10. Integration and data fitting were performed using Nitpic [116] and Sedphat [23, 287]. No more than two outlier data points were removed from any one ITC data set for analysis. The flTGM6:TβRII binding experiment was globally fit to a simple binding model from two replicates; the TGM6-D3:TβRII binding experiment was globally fit to a simple binding model from three replicates. The TGM6-D3 variant and TβRII variant binding experiments were fit to a simple binding model from one replicate per variant.

Competition experiments performed by ITC were set up such that TβRII was in the syringe and both competitors were in the syringe with the concentration of the higher-affinity competitor held constant. The proteins in the syringe and sample cell for each experiment as well as experimental conditions are given in Table 4. All proteins were concurrently dialyzed into 25 mM HEPES, 150 mM NaCl, 0.05% NaN_3, pH 7.4 prior to the experiments. For each experiment, nineteen 2.0 μL injections were performed with an injection duration of 4 sec, a spacing of 150 sec, and a reference power of 10. The data were globally fit using a simple competitive binding model with one replicate per condition.

4.4.8 X-ray structure determination

TGM6-D3 (residues 16-102 of the full-length construct) and TβRII 46-155 were mixed in a 1.1-to-1.0 ratio, with TGM6-D3 being in slight excess. The binary complex was fractionated by SEC using a HiLoad Superdex 75 26/60 column (GE Healthcare, Piscataway, NJ) in 25 mM Tris, 100 mM NaCl, 0.05% NaN_3, pH 8.0. The fractions containing the binary complex were pooled and concentrated to 50 mg mL^{-1} for crystallization. The binary complex was crystallized in 0.1 M sodium cacodylate, 25% (w/v) PEG 4000, pH 6.5. Large star-burst-like crystal clusters with plate-like arms grew at ambient temperature in space group $P2_12_12$ with cell dimensions a = 55.88 Å, b = 130.67 Å, c = 29.76 Å in about three days.

Harvested crystals were briefly soaked in mother liquor containing 14% glycerol for cryoprotection and mounted in undersized nylon loops with excess mother liquor wicked off. The looped crystals were then flash-cooled in liquid nitrogen prior to data collection. Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory and integrated and scaled using XDS

[107]. The structure was determined by the molecular replacement method implemented in PHASER [164] using the 1.1 Å TβRII X-ray structure (PDB 1M9Z) [21] and the TGM1-D3 NMR structural ensemble (PDB 7SXB) [179] as search models. Coordinates were refined using PHENIX [1], including simulated annealing with torsion angle dynamics and alternated with manual rebuilding using COOT [46]. Final refinement was performed using the PDB-Redo webserver [104]. Data collection and refinement statistics are shown in Table 5.

4.4.9 SPR measurements

SPR datasets with TGM6-D3 variants binding to TβRII were generated using a BIAcore X100 instrument (GE Lifesciences, Piscataway, NJ). Biotinylated Avi-tagged TβRII was captured onto a neutravidin-coated CM-5 sensor chip (GE Lifesciences, Piscataway, NJ) at a density of 100 – 250 RU. Neutravidin-coated sensor chips for capture of biotinylated avi-tag receptors were made by activating the surface of a CM-5 chip with EDC and NHS, followed by injection of neutravidin (Pierce, Rockford, IL) diluted into sodium acetate at pH 4.5 until the surface density reached 6000 – 15000 RU. Kinetic binding assays were performed by injections of the analytes in 25 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% sodium azide, 0.05% surfactant P20 (Pierce, Rockford, IL) at 100 μL min^{-1}. Regeneration of the surface was achieved by an injection of 100 mM guanidine hydrochloride at 100 μL min^{-1} at the conclusion of each run. Baseline correction was performed by subtracting the response from both the reference surface with no immobilized ligand and 5 – 10 blank buffer injections. Kinetic analyses were performed by fitting the results to a simple 1:1 model using the program Scrubber (Biologic Software, Canberra, Australia).

4.4.10 TGM6 inhibition assays

The TGF-β inhibition assays utilizing NIH-3T3 cells were performed using the C2C12 BRE-Luc BMP reporter cell line stably transfected with pGL3(BRE)-luciferase reporter construct as previously reported [81]. Briefly, C2C12-BRE cells were plated at a concentration of 2×10^4 cells per well in 24 well plates containing Dulbecco's modified Eagle's medium (DMEM) plus 10% fetal calf serum (FCS) and allowed to attach for 18 hours. Cells were washed with PBS and refed with 0.5 ml of DMEM plus 0.1% FCS for 7 hours. After this initial incubation, increasing concentrations of either flTGM6 or TGM6-D3 were added to the wells. Recombinant growth factors were added to cells at the appropriate concentrations for 15 hours and then cells were washed with PBS and lysed using 100 µl of reporter lysis buffer (Promega, Madison, USA). To measure luciferase activity, 40 µL of lysate was added to 40 µL Luciferase Assay Reagent (Promega). The protein concentration of each lysate was analyzed using Bio-Rad protein assay reagent according to the manufacturer's instructions (Biorad, USA). Luciferase units obtained were normalized to the protein content of each well. All experiments were performed at least three times with four independent wells per condition. Cell numbers were counted prior to lysis. Chemiluminescence was detected with a Tecan Infinite M200 Plate Reader and normalized to cell counts.

The TGF-β inhibition assays utilizing MFB-F11 cells containing a TGF-β-responsive alkaline phosphatase reporter [254] was performed as previously reported [103]. Briefly, confluent cells were detached with trypsin, and resuspended in DMEM containing 2.5% FCS, 100 U mL^{-1} penicillin, 100 mg mL^{-1} streptomycin, and 2 mM L-glutamine at a concentration of $8x10^5$ cells ml^{-1}. Cells were plated at $4x10^4$ cells per well of a 96-well flat-bottomed plate and left to incubate at 37°C for 2 h. After this initial incubation, increasing concentrations of either full-

length TGM6 or TGM6-D3 were added to the wells as 25 μL. After 30 minutes, cells were stimulated with 5 μg/mL TGF-β or TGM1 in a volume of up to 50 μL and incubated for another twenty-four hours at 37°C. The final volume in each well was 100 μL. After the second incubation, 20 μL of supernatant was aspirated from each well, added to an ELISA plate (Nalge Nunc International, USA) with 180 μL of reconstituted Sigma FastTM p-nitrophenyl phosphate substrate and incubated at room temperature in the dark for up to four hours. Plates were read at 405 nm on an Emax precision microplate reader (Molecular Devices, USA). All conditions were set up in triplicate and repeated at least twice.

The TGF-β inhibition assays utilizing HEK-293 cells stably transfected with the $CAGA_{12}$ TGF-β reporter [254] were performed as previously reported [119]. Cells were maintained in DMEM containing 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. Cells were plated at $3x10^4$ cells per well in a treated 96-well plate. After 24 hours, the media was removed and replaced with fresh DMEM containing 0.1% bovine serum albumin (BSA) and a concentration series of inhibitor (mmTGF-β2-7M2R, flTGM6, or TGM6-D3). After 30 minutes, cells were stimulated with 10 pM TGF-β3. Twenty-four hours after stimulation, the cells were lysed, and luciferase activity was measured using Luciferin.

IC_{50} values were calculated in Prism 9 (Graphpad Software, Inc.) by globally fitting the replicates of each inhibition assay to a nonlinear dose-response inhibition model.

4.4.11 Statistical analyses

Statistical analyses were performed using Prism 9 (Graphpad Software, Inc.) or Sedphat [23, 287], as appropriate. For comparisons of two groups, a Student's unpaired two-tailed t-test was used, assuming unequal variance. *P* values of < 0.05 were considered statistically significant.

Sample sizes were chosen empirically based on the laboratory's previous experience in the calculation of experimental variability; sample sizes for each experiment were not pre-determined by individual power calculations.

4.5 Acknowledgments

We would like to thank Mr. Chang Byeon for his assistance in producing and purifying proteins used in the binding studies. We would like to thank Mr. Mike Delk for his assistance with the NMR instrumentation. We would like to thank Dr. Matthew Whitley for his assistance with crystal structure data analysis.

For APH, this research was supported by the NIH (GM58670 and AI53915). For RMM, this work was supported by the Wellcome Trust through an Investigator Award (Ref 1219530) and the Wellcome Trust core-funded Wellcome Centre for Investigative Parasitology (Ref 104111).

X-ray data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. SER-CAT is supported by its member institutions, and equipment grants (S10_RR25528, S10_RR028976 and S10_OD027000) from the National Institutes of Health.

2D NMR spectra were plotted using NMRFAM-SPARKY [128]. ITC thermograms were plotted using GUSSI [22]. Structure figures were created using UCSF ChimeraX [200]. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

5.0 Project Conclusions and Future Work

5.1 What Have Parasites Taught Us About TGF-β Signaling?

The TGF-β family of signaling proteins and their receptors present a plethora of opportunities for continued understanding of how the human body functions at a molecular level. Additionally, studying the signaling mechanisms and pathways of the TGF-β family of proteins presents ample opportunities to combat diseases, such as cancer and fibrosis. Furthermore, understanding how parasites, such as *H. polygyrus*, have coopted these signaling pathways can help us better understand how the signaling pathways work, but also provide potential new opportunities for the development of therapeutics, including anti-parasitics for human or livestock parasites and adapting the proteins parasites have evolved for treating human TGF-β family-related diseases.

5.1.1 Use of TGM proteins as therapeutics and laboratory tools

While the TGM family of proteins is still being highly investigated, there is already evidence that these proteins could be directly used as therapeutics. The Maizels and Mayfield labs investigated the use of TGM1 as a therapeutic in a DSS model of colitis in mice. They showed that feeding mice with algae engineered to express TGM1 could minimize the symptoms associated with the DSS colitis model [247]. This suggests that TGM1 treatment could improve the overall symptoms of gut-associated inflammatory diseases, such as Crohn's Disease.

As stated previously, on-target adverse effects of TGF-β agonists or antagonists are a major concern for the overall safety profile of these types of therapeutics. However, engineering therapeutic proteins based on the TGM domains could provide safe and effective targeting of the TGF-β signaling pathway. One method of doing so would be to use the receptor-binding domains (TGM1-D12, TGM1-D3, or TGM6-D3) fused to a specific targeting protein. For example, fusing the variable region of an antibody against a cancer target with TGM6-D3 would allow for specific delivery of a TGF-β signaling antagonist to a tumor. This combination would provide a sink for TβRII in the tumor environment, which could prevent down-regulation of the immune system and therefore allow the immune system to fight the tumor without interfering with TGF-β signaling in other parts of the body, such as the heart.

Another example of using TGM-based therapeutics would be to use the individual TGM domains as scaffolds for targeting other TGF-β family receptors or mimicking other TGF-β family members. As most of the TGF-β family signaling proteins bind to their receptors in similar ways, TGM6-D3 could be subjected to engineering, mutagenesis, and affinity maturation against the extracellular domains of other family type II receptors. In this way, a treatment for Fibrodysplasia Ossificans Progressiva (FOP) could be created by mutating TGM6-D3 such that it bound and sequestered BMPRII or ActRIIb in lymphoblastoid cells to offset the receptor-associated BMP4 signaling dysregulation [110].

Finally, the TGM proteins themselves and their derivatives could become powerful laboratory tools. As shown in this dissertation, TGM6 is a potent inhibitor of TGF-β signaling [268]. This inhibitor can be used as a competitor to probe and/or block the interactions between the TGF-β isoforms and TβRII in the development of new therapeutics.

5.1.2 Applications toward human and livestock parasites

H. polygyrus, being a model parasite that only infects mice, is relatively safe and easy to study in the lab. While *H. polygyrus* is a mouse parasite, and its secreted proteins will not necessarily impact other species, knowledge of the TGM family of proteins and their functions is invaluable for the field of parasitology. The identification and characterization of these proteins can aid in the characterization of parasites that are of human or veterinary importance, such as *Teladorsagia circumcincta* in sheep and *Ascaris lumbricoides* in humans. In preliminary work, proteins secreted by *T. circumcincta* fourth-stage larvae had activity in TGF-β reporter assays that was not blocked by the anti-TGF-β blocking antibody 1D11 [61]. The identification of immune modulatory proteins secreted from other parasites, specifically proteins that act through the TGF-β pathway, can now be identified and better characterized. Specifically, this could lead to the development of anti-parasitic compounds once the structures of the key proteins are determined.

5.2 Project Conclusions

In this dissertation, the initial work to characterize the function and structure of the *H. polygyrus* TGF-β mimic TGM6 was reported. The work shows that TGM6, like its parent family member TGM1, binds the TGF-β type II receptor (TβRII) through domain 3 with a 320 nM affinity, and that TGM6-D3 is necessary and sufficient for binding TβRII. Further, TGM6 does not bind another TGF-β family type II receptor (with ActRII, ActRIIb, and BMPRII tested) nor does it bind a TGF-β family type I receptor (with Alk1, Alk2, Alk3, Alk4, and Alk5 tested) either

through domains 4 and 5 or through a combined interface created by the TGM6:TβRII binary complex.

ITC experiments showed that TGM6-D3 bound TβRII in a competitive manner, using either the same or similar residues as those that TGF-β uses to bind the receptor. Determining the structure of the TGM6-D3:TβRII binary complex by X-ray crystallography showed that TGM6-D3 not only bound the same site as TGF-β on TβRII, but that TGM6-D3 was a remarkable mimic of TGF-β in its binding of TβRII as it bound the same residues in the same manner as TGF-β does on TβRII. Mutation of key residues on TGM6-D3 and TβRII confirmed the specific interactions on both binding partners.

Finally, initial work of characterizing the functional role of TGM6 was completed. In TGF-β reporter assays, full-length TGM6 acted as potent antagonist of TGF-β- and TGM1-induced TGF-β signaling in murine fibroblast cell lines, inhibiting with an IC_{50} greater than 1000-fold lower than the K_D between TGM6-D3 and TβRII. However, TGM6-D3, which binds and sequesters TβRII, did not inhibit in these cell lines. Furthermore, neither full-length TGM6 nor TGM6-D3 inhibited TGF-β-induced TGF-β signaling in human embryonic kidney cells. From this, it was hypothesized that TGM6 binds a coreceptor through domain 4, domain 5, or both domains 4 and 5 that provided specificity for its antagonistic activity.

5.3 Future Directions

Despite the knowledge gained by this work to characterize the structure and function of TGM6, significant knowledge gaps remain. As TGM6 does not appear to function as an inhibitor without domains 4 and 5 physically attached to domain 3 [268], it is likely that TGM6 binds a cell-

or species-specific coreceptor through domains 4 and/or 5. Therefore, identification of the coreceptor(s) would provide details on the specificity of TGM6 toward potential target cells and tissues. Additionally, assaying various cell types from lineages throughout the animal kingdom and from varying tissue types for TGM6-induced inhibition of TGF-β signaling would aid in the identification of the coreceptor(s) and give further details about the tissues in which TGM6-induced TGF-β signaling inhibition is important. Once the identity of the coreceptor(s) is known, the structure of full-length TGM6 as bound to both TβRII and the coreceptor(s) can be determined. With a deeper understanding of the interacting partners for TGM6, its function within the parasitic lifecycle can be better understood.

Tissue-specific antagonism of TGF-β signaling is highly desired for the treatment of diseases like cancer and fibrosis. As shown in this dissertation, TGM6 is a highly specific and highly potent TGF-β- and TGM1-induced TGF-β signaling inhibitor. The development of this protein into a therapeutic would provide such a treatment, provided domains 4 and/or 5 bind coreceptors present on the appropriate tissues. At present, much development and engineering would be required, as these proteins are highly immunogenic in mice [247].

There are three potential paths forward with this development. First, the affinity and kinetics of the interaction between TGM6-D3 and TβRII could be improved by means of affinity maturation. This would improve the dose requirement and increase the residence time of the protein in the targeted tissues, reducing the required dose and improving the pharmacokinetics of the treatment. Second, TGM6-D3 could be fused with a different protein, such as the variable region of an antibody, to provide preferential targeting to cells and tissues with receptors which domains 4 and 5 do not naturally bind. Finally, TGM6-D3 could be mutated and affinity-matured against other TGF-β family type II receptors, such as BMPRII and ActRIIb, to create signaling

antagonists for these receptors as well. As all TGF-β family members aside from the TGF-βs bind a type II receptor in a similar manner [90], TGM6-D3 would be a logical scaffold for these binding partners. In this way, additional receptors can be targeted, and signaling through these receptors can be antagonized. While there would be concern about off-target effects by generally antagonizing signaling through any of the TGF-β family type II receptors, these new proteins would be of great benefit to the research community at large.

Finally, the work on the TGM family of proteins is still ongoing. As the proteins themselves have been identified [246], all the receptors and coreceptors to which these proteins bind and through which they act still have not been identified. Furthermore, the characterization of these interactions is also incomplete. Therefore, one major step forward in understanding how *H. polygyrus* acts on a host is to complete these identifications and characterizations of these interactions, including the characterization of the thermodynamic properties, the affinities, and determination of the structures of the appropriate complexes with specific attention on determining the key interacting residues on both the parasite protein domains and their binding partners.

In better understanding the function and interactions of the TGM family of proteins, we can begin to better understand the molecular details of how various proteins impact the lifecycle of a parasite. As this work has been done with proteins secreted by *H. polygyrus*, we could apply this knowledge to other parasites of importance to either livestock or human health. Identification of proteins secreted by other parasites that interact with the TGF-β family of signaling proteins and/or their receptors present a major target in fighting parasitic infections in the developing world.

Appendix A Sequence Alignments for the Members of the TGM Family

In Chapter 2.0, the sequence alignments between different TGM domains and various members of the TGM family are discussed. Table 8 provides the NCBI accession numbers for the proteins in these sequence alignments. Sequence alignments were performed using Clustal Omega [143]. Sequence alignment figures were created using Jalview 2 [264].

Table 8. NCBI accession numbers for proteins in sequence alignments.

Gene Name	Length of Protein	NCBI Accession Number	Reference
Human Factor H	449	X07523	[217]
TGM1	422	MG099712	[103, 246]
TGM2	430	MG429737	[246]
TGM3	429	MG429738	[246]
TGM4	422	MG429739	[246]
TGM5	341	MG429740	[246]
TGM6	254	MG429741	[246]
TGM7	599	MG429742	[246]
TGM8	599	MG429743	[246]
TGM9	252	MG429744	[246]
TGM10	251	MG429745	[246]

The color key used in the sequence alignments shown in this appendix are summarized in Table 9.

Table 9. Color codes used in protein sequence alignments.

Color	Property	Amino Acid(s)
Magenta	Nonpolar aromatic	FWY
Red	Negatively charged	DE
Yellow	Cysteine	C
Green	Nonpolar aliphatic	GAVLIM
Cyan	Polar uncharged	STCYNQP
Blue	Positively charged	KRH

The sequence alignment between the individual TGM1 domains and an archetypal CCP domain, human Factor H module 1 is shown in Figure 61. The TGM1 domains have a distant sequence identity between one another, as summarized in Table 10. There is distant homology between the TGM1 domains and the archetypal CCP domain [103], as summarized in Table 11. Alignment of the five similar domains within TGM1 encompasses the entire amino acid sequence apart from the predicted signal peptide (aa 1–18), with conserved cysteine (yellow) and other residues indicated, together with an archetypal CCP domain, human Factor H module 1 (aa 20–83). Note the presence of a 15-aa insertion near the N-terminus of each domain of TGM1, which is not typical of the CCP family.

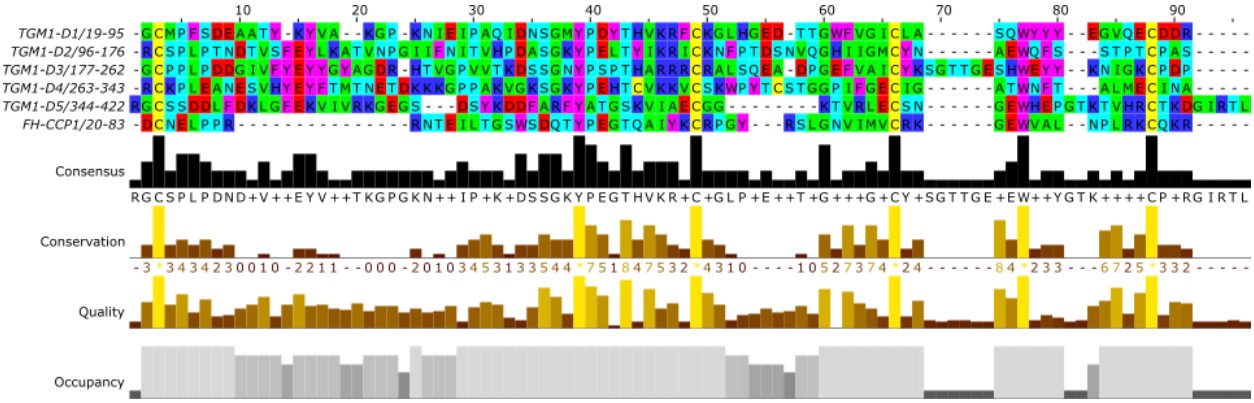

Figure 61. Sequence alignment between the individual TGM1 domains and the archetypal CCP domain from human human Factor H.

Table 10. Percent identity matrix for TGM1 domains.

	TGM1-D1	TGM1-D2	TGM1-D3	TGM1-D4	TGM1-D5
TGM1-D1	100.00	26.32	35.06	27.63	20.90
TGM1-D2	26.32	100.00	27.85	33.33	20.00
TGM1-D3	35.06	27.85	100.00	24.05	14.29
TGM1-D4	27.63	33.33	24.05	100.00	14.29
TGM1-D5	20.90	20.00	14.29	14.29	100.00

Table 11. Percent identity matrix for TGM1 domains compared to FH-CCP1.

	FH-CCP1
TGM1-D1	20.63
TGM1-D2	23.81
TGM1-D3	25.40
TGM1-D4	19.05
TGM1-D5	21.43
FH-CCP1	100.00

The sequence alignment of the TGM family of proteins is shown in Figure 62 through Figure 67, broken up by individual domains for clarity. The percent identity matrix for the full-length proteins is given in Table 12, and the percent identity matrices for each domain across the TGM family are given in Table 14 through Table 18.

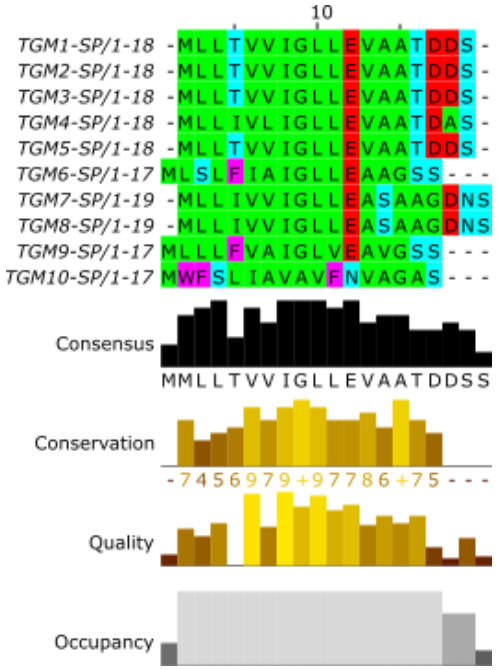

Figure 62. Sequence alignment between the signal peptides of the members of the TGM family.

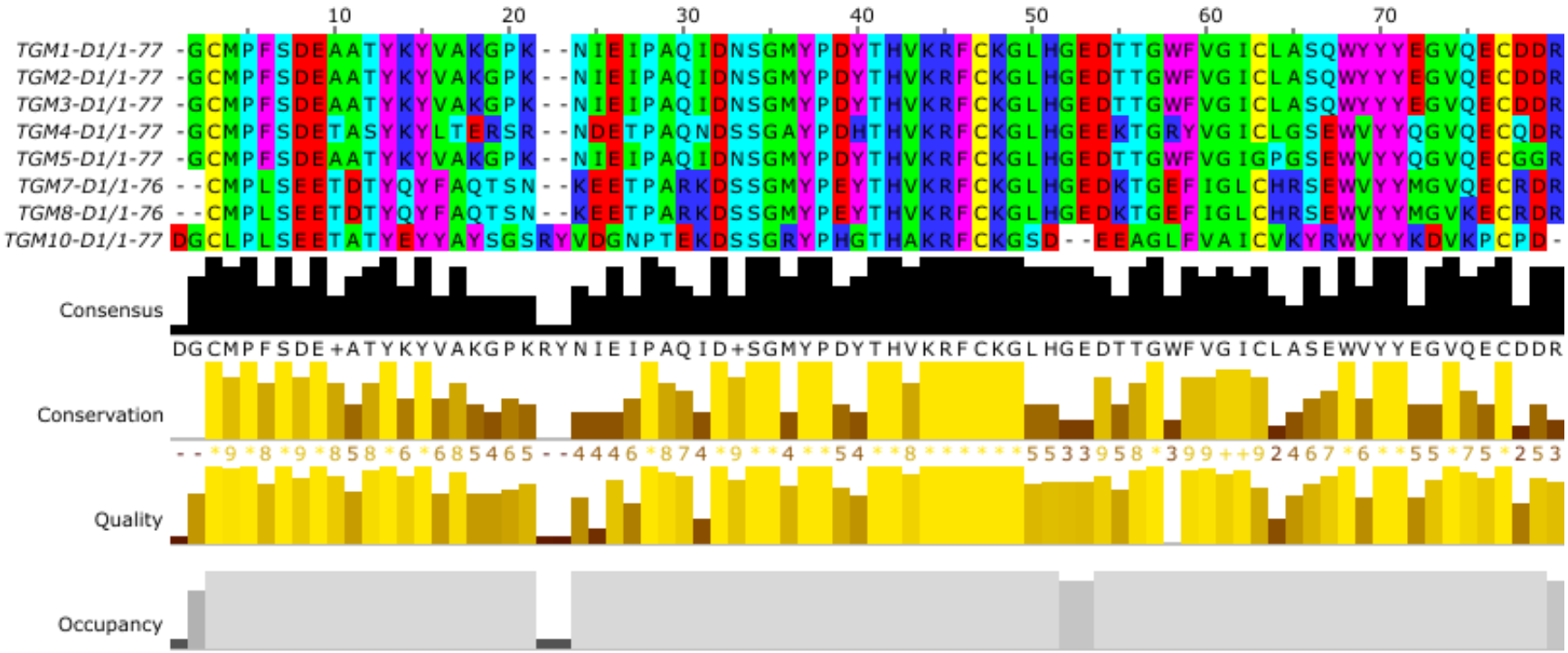

Figure 63. Sequence alignment between domain 1 of the members of the TGM family. TGM6 and TGM9 are lacking domain 1 and therefore do not appear in this alignment.

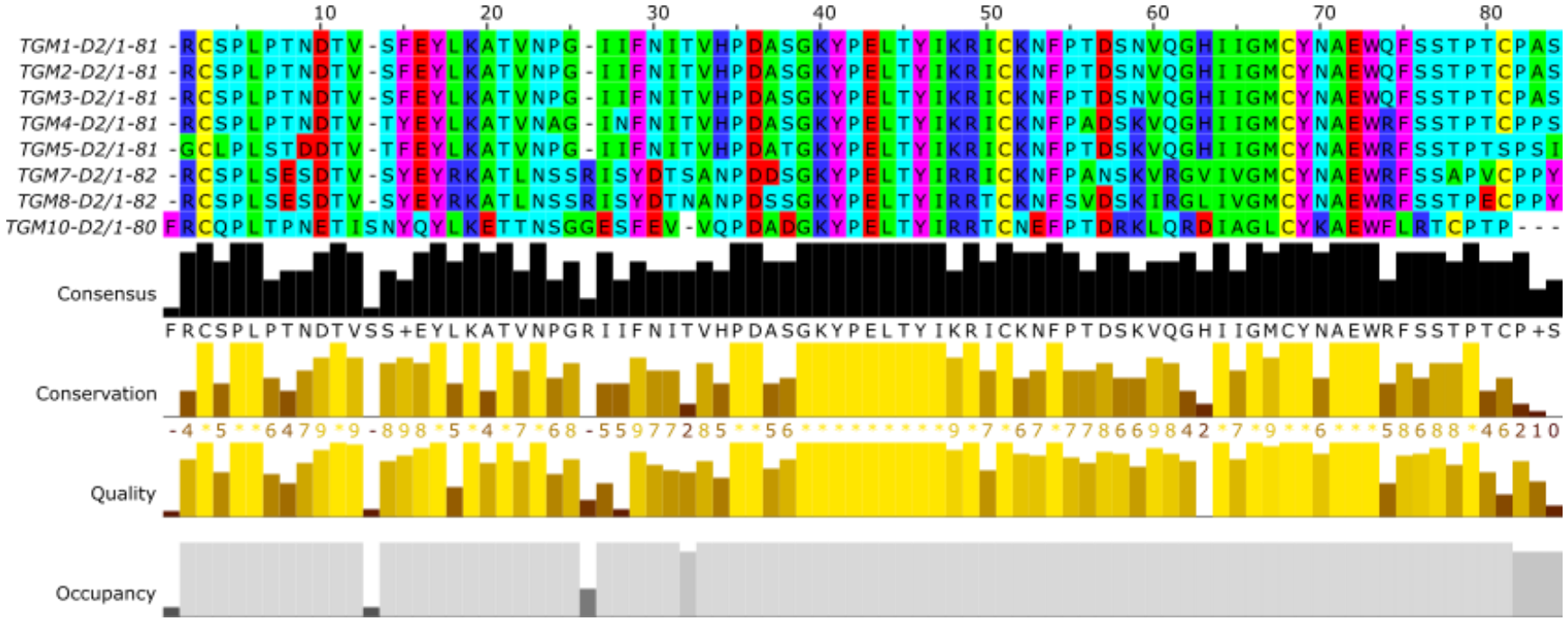

Figure 64. Sequence alignment between domain 2 of the members of the TGM family. TGM6 and TGM9 are lacking domain 2 and therefore do not appear in this alignment.

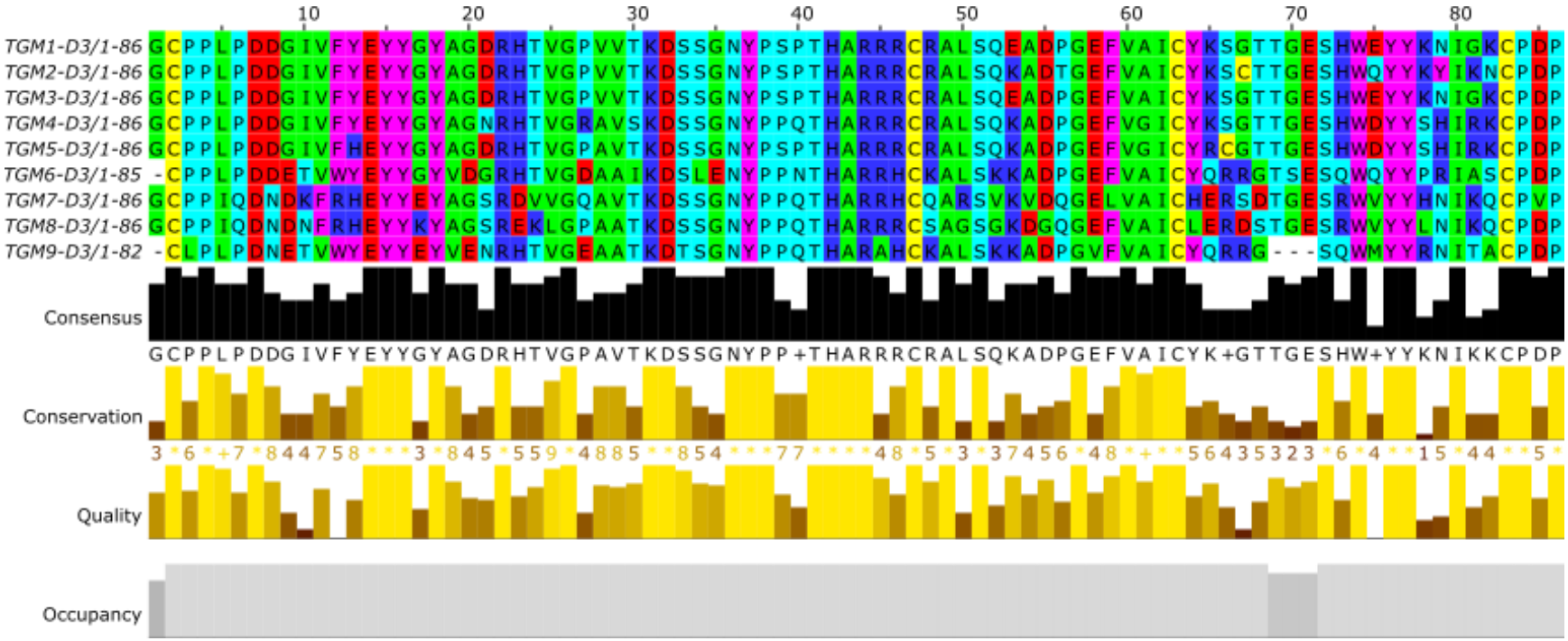

Figure 65. Sequence alignment between domain 3 of the members of the TGM family. TGM10 is lacking domain 3 and therefore does not appear in this alignment.

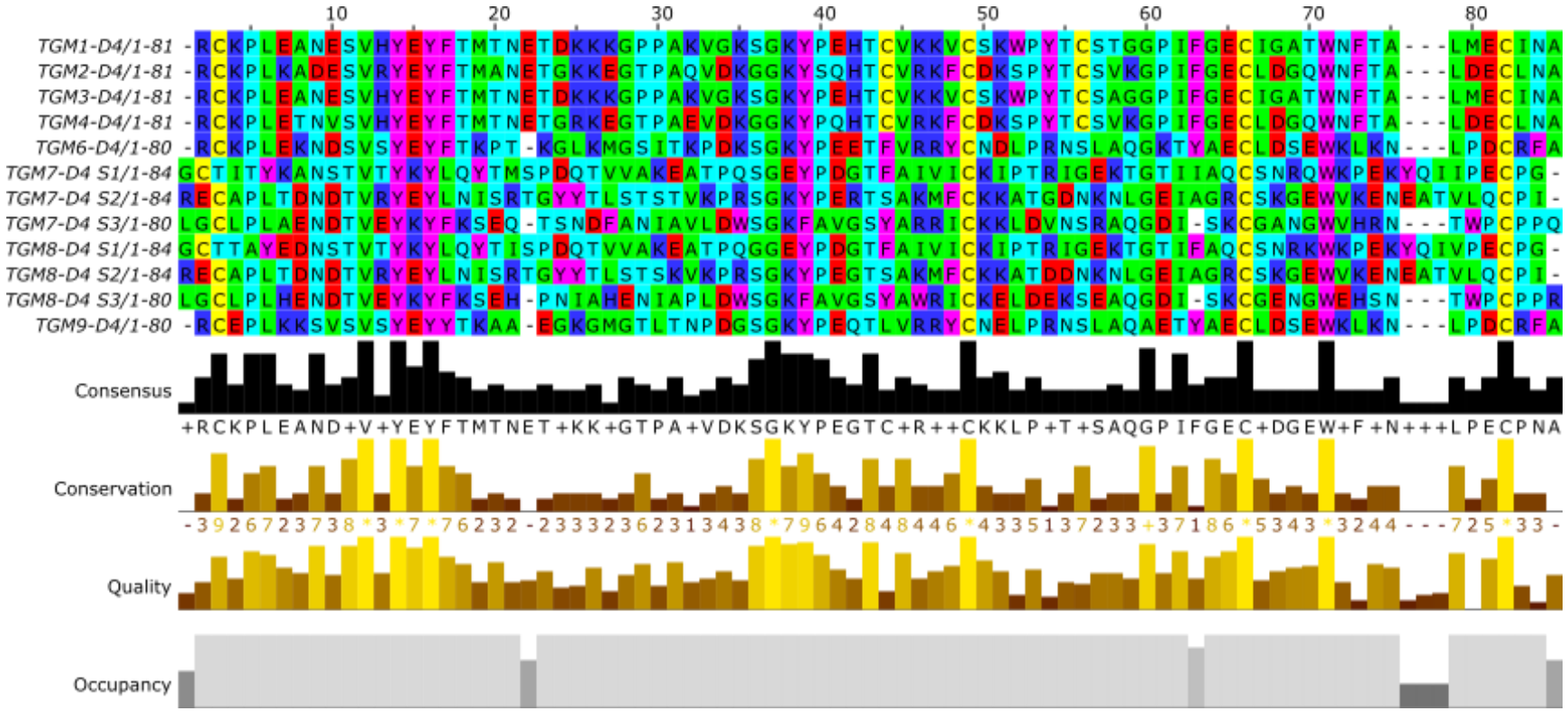

Figure 66. Sequence alignment between domain 4 of the members of the TGM family. TGM5 and TGM10 are lacking domain 4 and therefore do not appear in this alignment. TGM7 and TGM8 each have three sub-domains in place of one domain 4; these sub-domains are noted as S1, S2, and S3.

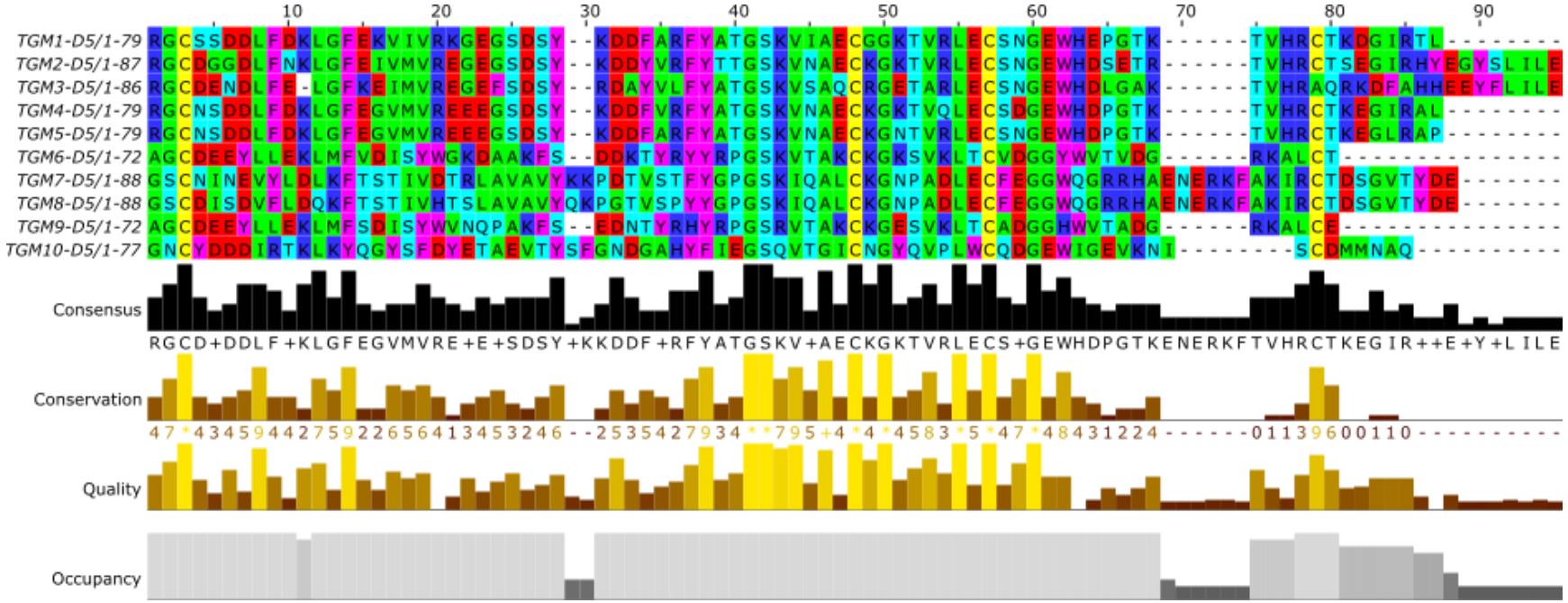

Figure 67. Sequence alignment between domain 5 of the members of the TGM family.

Table 12. Percent identity matrix for full-length TGM proteins. TGM7 and TGM8 each have three sub-domains in place of D4, noted as S1, S2, and S3.

	TGM1	TGM2	TGM3	TGM4	TGM5	TGM6	TGM7 S1	TGM7 S2	TGM7 S3	TMG8 S1	TGM8 S2	TGM8 S3	TGM9	TGM10
TGM1	100.00	87.91	92.16	80.81	87.68	45.63	49.05	50.71	49.76	48.34	50.00	48.81	40.96	40.98
TGM2	87.91	100.00	85.78	82.70	85.92	46.03	48.70	50.83	50.12	47.99	50.35	49.17	42.97	40.57
TGM3	92.16	85.78	100.00	76.48	82.65	45.42	48.58	50.24	49.52	48.10	49.76	48.81	41.94	39.34
TGM4	80.81	82.70	76.48	100.00	85.04	47.62	50.00	52.13	51.43	49.05	50.71	49.52	43.78	42.62
TGM5	87.68	85.92	82.65	85.04	100.00	45.93	56.01	56.01	56.01	54.84	54.84	54.84	42.60	40.57
TGM6	45.63	46.03	45.42	47.62	45.93	100.00	33.46	36.61	37.70	33.86	36.61	37.30	78.09	25.61
TGM7 S1	49.05	48.70	48.58	50.00	56.01	33.46	100.00	86.44	85.85	88.51	77.24	76.57	32.27	37.50
TGM7 S2	50.71	50.83	50.24	52.13	56.01	36.61	86.44	100.00	87.47	77.24	89.89	77.03	35.86	37.50
TGM7 S3	49.76	50.12	49.52	51.43	56.01	37.70	85.85	87.47	100.00	76.10	78.19	86.31	35.34	37.50
TMG8 S1	48.34	47.99	48.10	49.05	54.84	33.86	88.51	77.24	76.10	100.00	86.90	85.85	31.08	38.31
TGM8 S2	50.00	50.35	49.76	50.71	54.84	36.61	77.24	89.89	78.19	86.90	100.00	86.77	35.06	38.31
TGM8 S3	48.81	49.17	48.81	49.52	54.84	37.30	76.57	77.03	86.31	85.85	86.77	100.00	34.54	38.31
TGM9	40.96	42.97	41.94	43.78	42.60	78.09	32.27	35.86	35.34	31.08	35.06	34.54	100.00	19.51
TGM10	40.98	40.57	39.34	42.62	40.57	25.61	37.50	37.50	37.50	38.31	38.31	38.31	19.51	100.00

Table 13. Percent identity matrix for TGM protein signal peptides.

	TGM1-SP	TGM2-SP	TGM3-SP	TGM4-SP	TGM5-SP	TGM6-SP	TGM7-SP	TGM8-SP	TGM9-SP	TGM10-SP
TGM1-SP	100.00	100.00	100.00	83.33	100.00	43.75	66.67	66.67	43.75	12.50
TGM2-SP	100.00	100.00	100.00	83.33	100.00	43.75	66.67	66.67	43.75	12.50
TGM3-SP	100.00	100.00	100.00	83.33	100.00	43.75	66.67	66.67	43.75	12.50
TGM4-SP	83.33	83.33	83.33	100.00	83.33	43.75	61.11	61.11	43.75	12.50
TGM5-SP	100.00	100.00	100.00	83.33	100.00	43.75	66.67	66.67	43.75	12.50
TGM6-SP	43.75	43.75	43.75	43.75	43.75	100.00	43.75	43.75	76.47	35.29
TGM7-SP	66.67	66.67	66.67	61.11	66.67	43.75	100.00	100.00	50.00	6.25
TGM8-SP	66.67	66.67	66.67	61.11	66.67	43.75	100.00	100.00	50.00	6.25
TGM9-SP	43.75	43.75	43.75	43.75	43.75	76.47	50.00	50.00	100.00	23.53
TGM10-SP	12.50	12.50	12.50	12.50	12.50	35.29	6.25	6.25	23.53	100.00

Table 14. Percent identity matrix for domain 1 of the TGM proteins. TGM6 and TGM9 are lacking domain 1 and therefore do not appear in this matrix.

	TGM1-D1	TGM2-D1	TGM3-D1	TGM4-D1	TGM5-D1	TGM7-D1	TGM8-D1	TGM10-D1
TGM1-D1	100.00	100.00	100.00	70.13	89.61	64.47	63.16	47.30
TGM2-D1	100.00	100.00	100.00	70.13	89.61	64.47	63.16	47.30
TGM3-D1	100.00	100.00	100.00	70.13	89.61	64.47	63.16	47.30
TGM4-D1	70.13	70.13	70.13	100.00	71.43	65.79	64.47	50.00
TGM5-D1	89.61	89.61	89.61	71.43	100.00	64.47	63.16	45.95
TGM7-D1	64.47	64.47	64.47	65.79	64.47	100.00	98.68	50.68
TGM8-D1	63.16	63.16	63.16	64.47	63.16	98.68	100.00	52.05
TGM10-D1	47.30	47.30	47.30	50.00	45.95	50.68	52.05	100.00

Table 15. Percent identity matrix for domain 2 of the TGM proteins. TGM6 and TGM9 are lacking domain 2 and therefore do not appear in this matrix.

	TGM1-D2	TGM2-D2	TGM3-D2	TGM4-D2	TGM5-D2	TGM7-D2	TGM8-D2	TGM10-D2
TGM1-D2	100.00	100.00	100.00	90.12	86.42	65.43	64.20	54.55
TGM2-D2	100.00	100.00	100.00	90.12	86.42	65.43	64.20	54.55
TGM3-D2	100.00	100.00	100.00	90.12	86.42	65.43	64.20	54.55
TGM4-D2	90.12	90.12	90.12	100.00	85.19	70.37	67.90	55.84
TGM5-D2	86.42	86.42	86.42	85.19	100.00	62.96	61.73	53.25
TGM7-D2	65.43	65.43	65.43	70.37	62.96	100.00	87.80	47.44
TGM8-D2	64.20	64.20	64.20	67.90	61.73	87.80	100.00	48.72
TGM10-D2	54.55	54.55	54.55	55.84	53.25	47.44	48.72	100.00

Table 16. Percent identity matrix for domain 3 of the TGM proteins. TGM10 is lacking domain 3 and therefore does not appear in this matrix.

	TGM1-D3	TGM2-D3	TGM3-D3	TGM4-D3	TGM5-D3	TGM6-D3	TGM7-D3	TGM8-D3	TGM9-D3
TGM1-D3	100.00	91.86	100.00	86.05	88.37	65.88	59.30	60.47	63.41
TGM2-D3	91.86	100.00	91.86	83.72	86.05	67.06	60.47	61.63	62.20
TGM3-D3	100.00	91.86	100.00	86.05	88.37	65.88	59.30	60.47	63.41
TGM4-D3	86.05	83.72	86.05	100.00	90.70	68.24	60.47	60.47	65.85
TGM5-D3	88.37	86.05	88.37	90.70	100.00	65.88	60.47	61.63	62.20
TGM6-D3	65.88	67.06	65.88	68.24	65.88	100.00	55.29	55.29	78.05
TGM7-D3	59.30	60.47	59.30	60.47	60.47	55.29	100.00	77.91	58.54
TGM8-D3	60.47	61.63	60.47	60.47	61.63	55.29	77.91	100.00	57.32
TGM9-D3	63.41	62.20	63.41	65.85	62.20	78.05	58.54	57.32	100.00

Table 17. Percent identity matrix for domain 4 of the TGM proteins. TGM5 and TGM10 are lacking domain 4 and therefore do not appear in this matrix. TGM7 and TGM8 each have three sub-domains in place of domain 4, noted as S1, S2, and S3.

	TGM1-D4	TGM2-D4	TGM3-D4	TGM4-D4	TGM6-D4	TGM7-D4 S1	TGM7-D4 S2	TGM7-D4 S3	TGM8-D4 S1	TGM8-D4 S2	TGM8-D4 S2	TGM9-D4
TGM1-D4	100.00	70.37	98.77	72.84	43.75	23.75	30.00	25.32	23.75	30.00	22.78	35.00
TGM2-D4	70.37	100.00	70.37	88.89	42.50	20.00	28.75	25.32	18.75	28.75	21.52	41.25
TGM3-D4	98.77	70.37	100.00	72.84	45.00	23.75	30.00	26.58	23.75	30.00	24.05	36.25
TGM4-D4	72.84	88.89	72.84	100.00	46.25	21.25	30.00	26.58	23.75	30.00	22.78	40.00
TGM6-D4	43.75	42.50	45.00	46.25	100.00	24.05	30.38	34.18	24.05	29.11	31.65	76.25
TGM7-D4 S1	23.75	20.00	23.75	21.25	24.05	100.00	29.76	24.05	89.29	30.95	26.58	21.52
TGM7-D4 S2	30.00	28.75	30.00	30.00	30.38	29.76	100.00	31.65	30.95	96.43	27.85	29.11
TGM7-D4 S3	25.32	25.32	26.58	26.58	34.18	24.05	31.65	100.00	22.78	32.91	77.50	27.85
TGM8-D4 S1	23.75	18.75	23.75	23.75	24.05	89.29	30.95	22.78	100.00	32.14	25.32	18.99
TGM8-D4 S2	30.00	28.75	30.00	30.00	29.11	30.95	96.43	32.91	32.14	100.00	29.11	27.85
TGM8-D4 S3	22.78	21.52	24.05	22.78	31.65	26.58	27.85	77.50	25.32	29.11	100.00	26.58
TGM9-D4	35.00	41.25	36.25	40.00	76.25	21.52	29.11	27.85	18.99	27.85	26.58	100.00

Table 18. Percent identity matrix for domain 5 of the TGM proteins.

	TGM1-D5	TGM2-D5	TGM3-D5	TGM4-D5	TGM5-D5	TGM6-D5	TGM7-D5	TGM8-D5	TGM9-D5	TGM10-D5
TGM1-D5	100.00	74.68	58.97	83.54	83.54	31.94	30.38	29.11	27.78	27.03
TGM2-D5	74.68	100.00	65.12	79.75	78.48	34.72	31.25	31.25	30.56	24.32
TGM3-D5	58.97	65.12	100.00	60.26	61.54	29.58	29.11	27.85	29.58	19.18
TGM4-D5	83.54	79.75	60.26	100.00	92.41	34.72	31.65	29.11	30.56	29.73
TGM5-D5	83.54	78.48	61.54	92.41	100.00	31.94	32.91	30.38	29.17	29.73
TGM6-D5	31.94	34.72	29.58	34.72	31.94	100.00	29.17	29.17	80.56	27.54
TGM7-D5	30.38	31.25	29.11	31.65	32.91	29.17	100.00	87.50	25.00	22.08
TGM8-D5	29.11	31.25	27.85	29.11	30.38	29.17	87.50	100.00	23.61	22.08
TGM9-D5	27.78	30.56	29.58	30.56	29.17	80.56	25.00	23.61	100.00	23.19
TGM10-D5	27.03	24.32	19.18	29.73	29.73	27.54	22.08	22.08	23.19	100.00

Appendix B Supplementary Information for Chapter 3.0

This appendix contains the supplementary information prepared for submission with the manuscript presented in Chapter 3.0.

Table 19. *H. polygyrus* TGM constructs used in this study

Construct	Construct features	Sequence
flTGM6	Expressed as Igk Signal Peptide-flTGM6-Linker-Myc Tag-Linker-His6 fusion	METDTLLLWV LLLWVPGSTG DAAQPARRAS CPPLPDDETV WYEYYGYVDG RHTVGDAAIK DSLENYPPNT HARRHCKALS KKADPGEFVA ICYQRRGTSE SQWQYYPRIA SCPDPRCKPL EKNDSVSYEY FTKPTKGLKM GSITKPDKSG KYPEETFVRR YCNDLPRNSL AQGKTYAECL DSEWKLKNLP DCRFAAGCDE EYLLEKLMFV DISYWGKDAA KFSDDKTYRY YRPGSKVTAK CKGKSVKLTC VDGGYWVTVD GRKALCTAAA RGGPEQKLIS EEDLNSAVDH HHHHH
TGM6-D3	Expressed as a Thioredoxin- His10-Linker-Thrombin-Linker-TGM6-D3 fusion	MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMSSGH HHHHHHHHHS SGGSGLVPR|G SGTGSSCPPL PDDETVWYEY YGYVDGRHTV GDAAIKDSLE NYPPNTHARR HCKALSKKAD PGEFVAICYQ RRGTSESQWQ YYPRIASCPD P
TGM6-D45	Expressed as a Thioredoxin- His10-Linker-Thrombin-Linker-TGM-D45 fusion	MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMSSGH HHHHHHHHHS SGGSGLVPR|G SGTRCKPLEK NDSVSYEYFT KPTKGLKMGS ITKPDKSGKY PEETFVRRYC NDLPRNSLAQ GKTYAECLDS EWKLKNLPDC RFAAGCDEEY LLEKLMFVDI SYWGKDAAKF SDDKTYRYYR PGSKVTAKCK GKSVKLTCVD GGYWVTVDGR KALCT
TGM1-D3	Expressed as a Thioredoxin- His6-Linker-Thrombin-Linker-TGM-D3 fusion	MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMHHHH HHSSGLVPR|G SGTGCPPLPD DGIVFYEYYG YAGDRHTVGP VVTKDSSGNY PSPTHARRRC RALSQEADPG EFVAICYKSG TTGESHWEYY KNIGKCPDP

Table 20. Type II receptor constructs used in this study

Construct	Construct features	Sequence
ActRII	Expressed as Linker-His6-Linker-Thrombin-Linker-ActRII fusion	MGSSHHHHHH SSGLVPR|GSH MAILGRSETQ ECLFFNANWE KDRTNQTGVE PCYGDKDKRR HCFATWKNIS GSIEIVKQGC WLDDINCYDR TDCVEKKDSP EVYFCCCEGN MCNEKFSYFP EME
ActRIIb	Expressed as Linker-His6-Linker-Thrombin-Linker-ActRIIb fusion	MGSSHHHHHH SSGLVPR|GSH MLEDPVPETR ECIYYNANWE LERTNQSGLE RCEGEQDKRL HCYASWRNSS GTIELVKKGC WLDDFNCYDR QECVATEENP QVYFCCCEGN FCNERFTHLP
BMPRII	Expressed as Signal-Linker-His6-Linker-Thrombin-Linker-BMPRII fusion	MKWVTFLLLL FISGSAFSAA AGSSHHHHHH SSGLVPR|GSH MNQERLCAFK DPYQQDLGIG ESRISHENGT ILCSKGSTCY GLWEKSKGDI NLVKQGCWSH IGDPQECHYE ECVVTTTPPS IQNGTYRFCC CSTDLCNVNF TENFPP
TβRII	Expressed as TβRII alone, with no tags or otherwise	MVTDNNGAVK FPQLCKFCDV RFSTCDQKSC MSNCSITSIC EKPQEVCVAV WRKNENITLE TVCHDPKLPY HDFILEDAAS PKCIMKEKKK PGETFFMCSC SSDECNDNII FSEEY

Table 21. Type II receptor binding to TGM6-D3 as assessed by ITC

Cell	TGM6-D3	Buffer	TGM6-D3	Buffer	TGM6-D3	Buffer
Syringe	ActRII	ActRII	ActRIIb	ActRIIb	BMPRII	BMPRII
Cell concentration (µM)	7.35	0.00	10.0	0.00	20.0	0.00
Syringe concentration (µM)	60.5	60.5	100	100	215	215
Temperature (°C)	35	35	25	25	35	35
N (sites)	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]
KD (nM)	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]
ΔH (kcal mol-1)	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]
ΔG (kcal mol-1)	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]
-TΔS (kcal mol-1)	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]	ND[a]

[a] Not determined due to weak or no signal

Table 22. Type I receptor constructs used in this study

Construct	Construct features	Sequence
Alk1 (TSRI)	Expressed as a Linker-His6-Linker-Thrombin-Linker-Alk1 fusion	MGSSHHHHHH SSGLVPR|GSH MDPVKPSRGP LVTCTCESPH CKGPTCRGAW CTVVLVREEG RHPQEHRGCG NLHRELCRGR PTEFVNHYCC DSHLCNHNVS LVLEATQPPS EQPGTDGQ
Alk2 (ActRIA)	Expressed as a Thioredoxin- His6-Linker-Thrombin-Linker-Alk2 fusion	MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMHHHH HHSSGLVPR|G SGTMEDEKPK VNPKLYMCVC EGLSCGNEDH CEGQQCFSSL SINDGFHVYQ KGCFQVYEQG KMTCKTPPSP GQAVECCQGD WCNRNITAQL PTKGKSFPGT QNF
Alk3 (BMPRIA)	Expressed as a Thioredoxin- His6-Linker-Thrombin-Linker-Alk3 fusion	MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL KEFLDANLAG SGSGHMHHHH HHSSGLVPR|G SGTQNLDSML HGTGMKSDSD QKKSENGVTL APEDTLPFLK CYCSGHCPDD AINNTCITNG HCFAIIEEDD QGETTLASGC MKYEGSDFQC KDSPKAQLRR TIECCRTNLC NQYLQPTLPP VVIGPFFDGS IR
Alk4 (ActRIB)	Expressed as a Linker-His6-Linker-Thrombin-Linker-Alk4 fusion	MGSSHHHHHH SSGLVPR|GSH MVQALLCACT SCLQANYTCE TDGACMVSIF NLDGMEHHVR TCIPKVELVP AGKPFYCLSS EDLRNTHCCY TDYCNRIDLR
Alk5 (TβRI)	Expressed as Linker-His6-Linker-Thrombin-Linker-TbRI fusion	MGSSHHHHHH SSGLVPR|GSH MAALLPGATA LQCFCHLCTK DNFTCVTDGL CFVSVTETTD KVIHNSSCIA EIDLIPRDRP FVCAPSSKTG SVTTTYCCNQ DHCNKIELPT TVKSSPGLGP VE

Table 23. Growth factor constructs used in this study

Construct	Construct features	Sequence
TGF-β3	Expressed as TGF-β3 alone, with no tags or otherwise	ALDTNYCFRN LEENCCVRPL YIDFRQDLGW KWVHEPKGYY ANFCSGPCPY LRSADTTHST VLGLYNTLNP EASASPCCVP QDLEPLTILY YVGRTPKVEQ LSNMVVKSCK CS
mmTGF-β2-7M2R	Residues 303-352 and 377-414 of mouse TGF-β2 (NCBI NP_0033393) connected by an engineered loop C379R substitution renders the protein monomeric; K327R, R328K, V381R, L391V, I394V, K396R, T397K, and I400V substitutions enable high affinity TβRII binding and high solubility Expressed as mmTGF-β2-7M2R alone, with no tags or otherwise	ALDAAYCFRN VQDNCCLRPL YIDFRKDLGW KWIHEPKGYN ANFCAGACPY RASKSPRCRS QDLEPLTIVY YVGRKPKVEQ LSNMIVKSCK CS

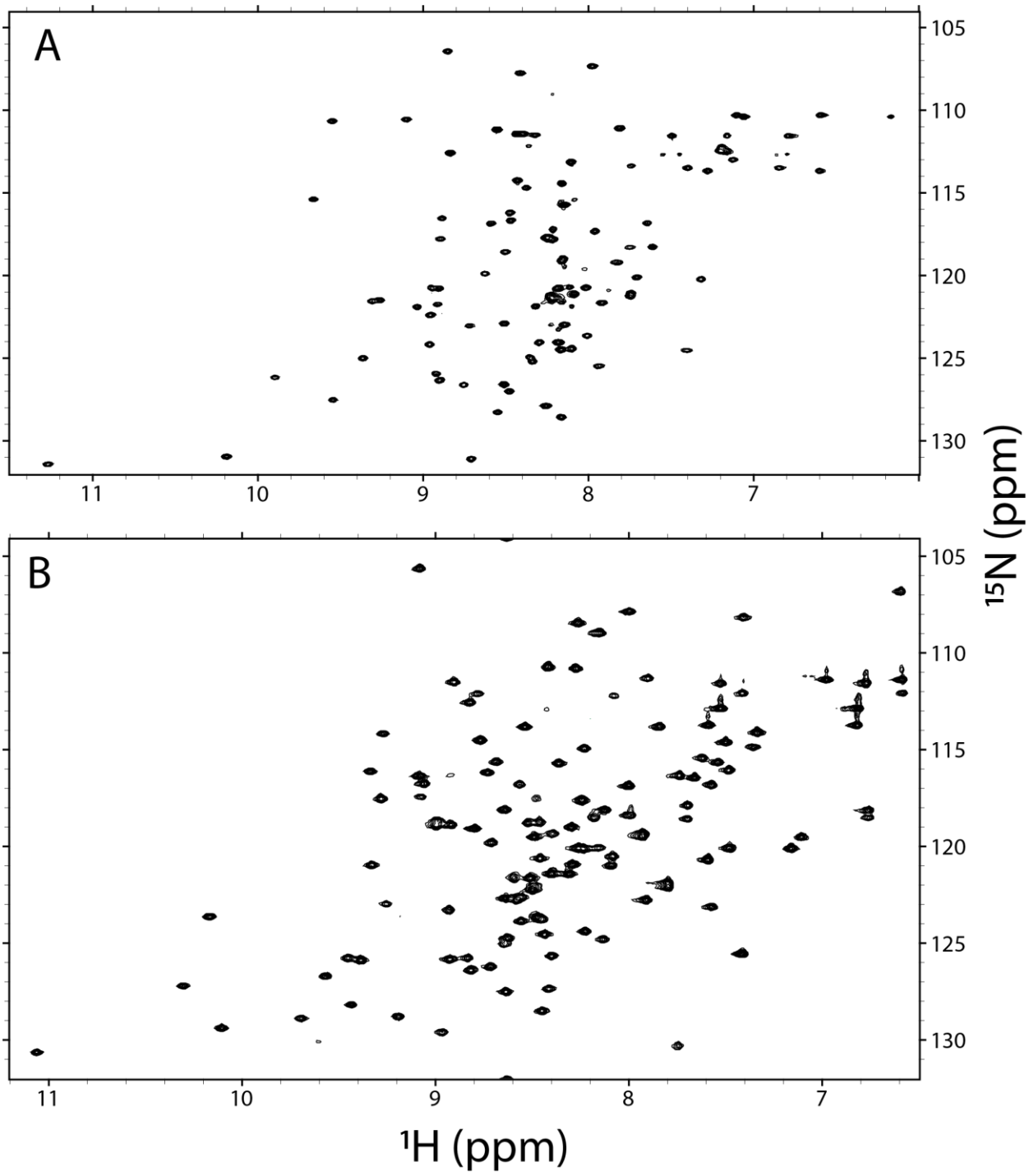

Figure 68. TGM6-D3 and -D45 are expressed as natively-folded proteins. ^{1}H-^{15}N HSQC spectra of (A) TGM6-D3 and (B) TGM6-D45 show well-dispersed peaks outside of the random coil region (7.8 – 8.6 ppm ^{1}H). This is indicative of a natively-folded protein.

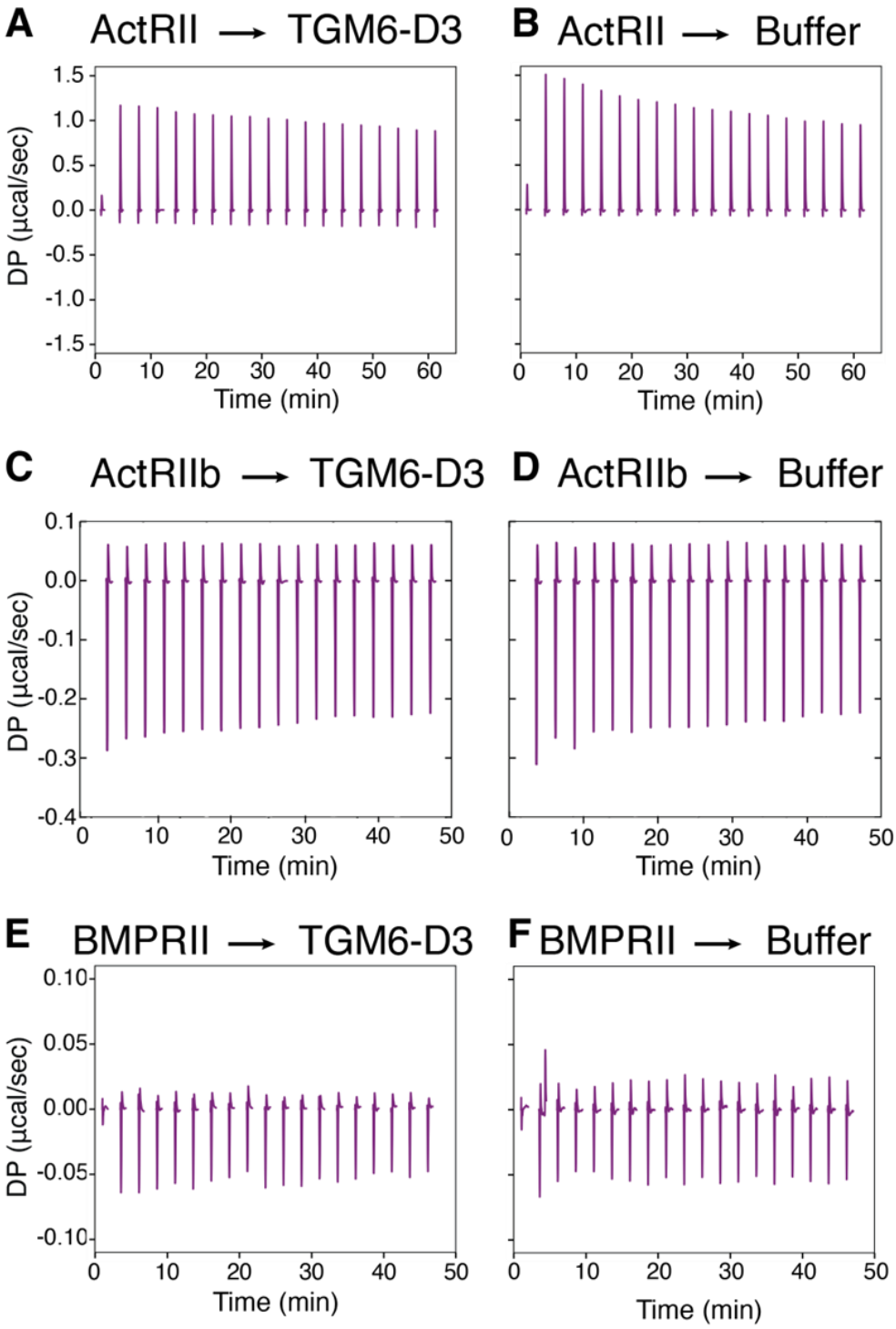

Figure 69. TGM6-D3 does not bind ActRII, ActRIIb, or BMPRII. Thermograms obtained upon the injection of ActRII, ActRIIb, or BMPRII into TGM6-D3 or Buffer. Panels A, C, and E correspond to the injection of ActRII, ActRIIb, and BMPRII into TGM6-D3, respectively; panels B, D, and F correspond to the injection of ActRII, ActRIIb, and BMPRII into buffer, respectively.

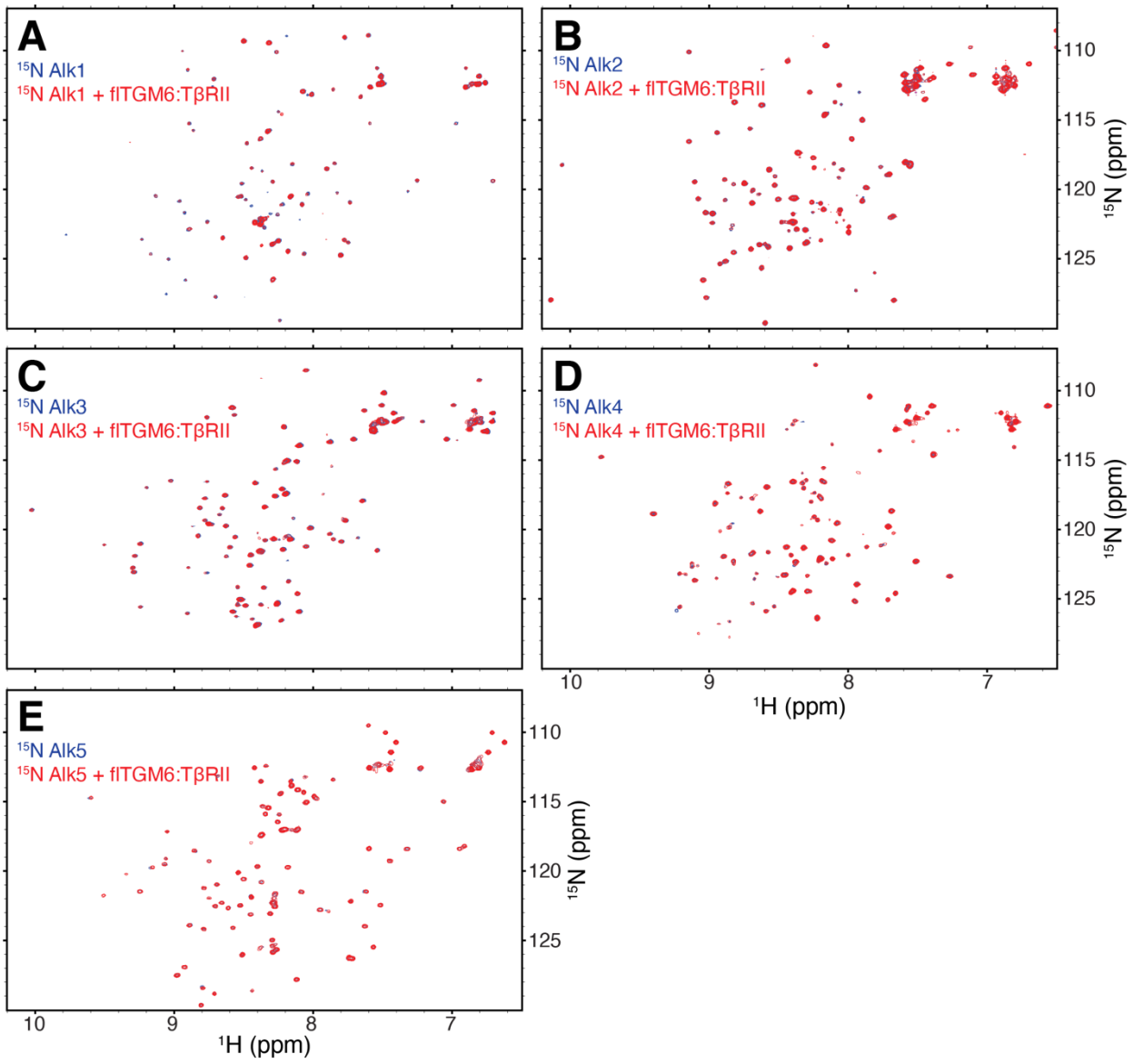

Figure 70. The flTGM6:TβRII binary complex does not bind to any type I receptors. ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled type I receptors alone (blue) and as bound to an excess of unlabeled flTGM6:TβRII binary complex (red). The receptors tested were: (A) ^{15}N Alk1; (B) ^{15}N Alk2; (C) ^{15}N Alk3; (D) ^{15}N Alk4; and (E) ^{15}N Alk5. None of the NMR signals exhibited changes to their chemical shifts larger than 0.002 ppm in the ^{1}H dimension, whereas shifts of greater than 0.005 ppm in the ^{1}H dimension would be expected in the event of binding.

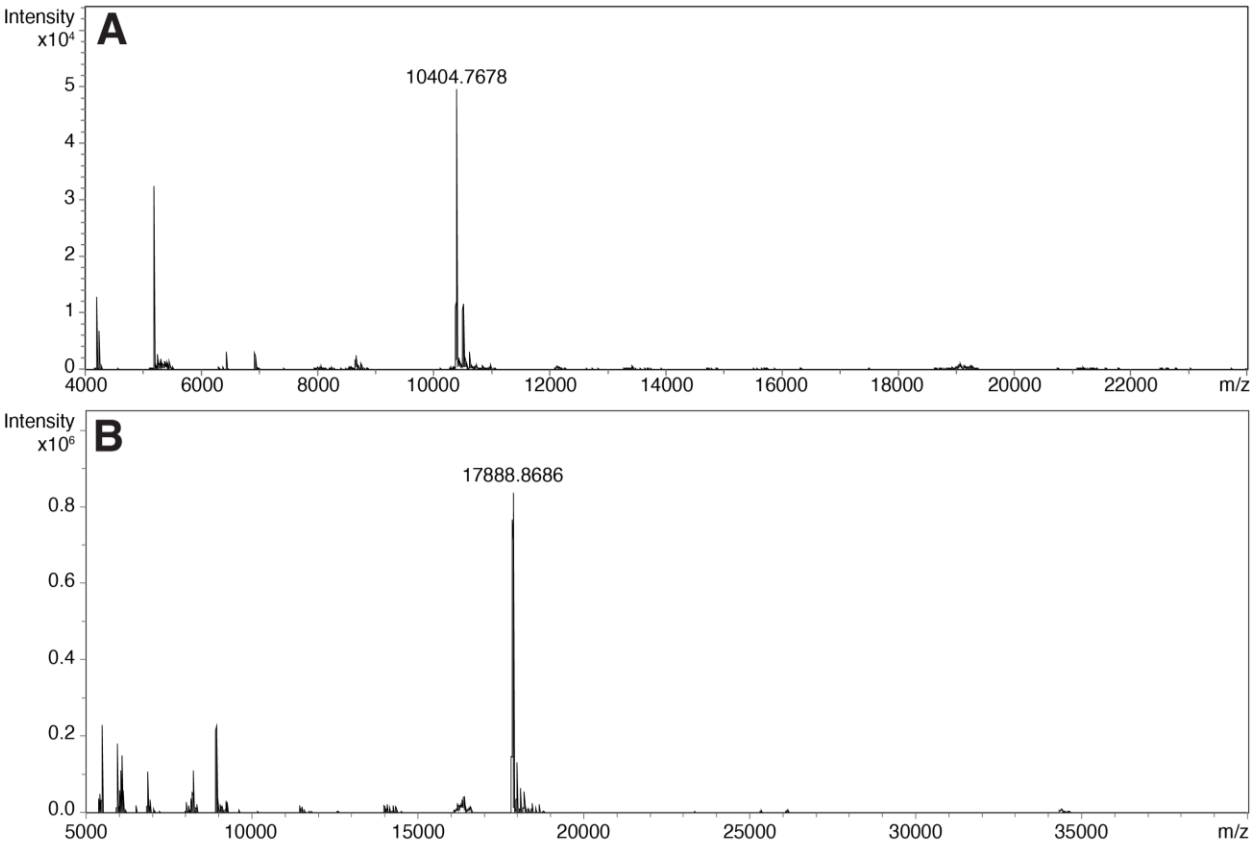

Figure 71. Mass spectra of (A) ^{15}N TGM6-D3 and (B) ^{15}N TGM6-D45. The calculated mass of ^{15}N TGM6-D3 is 10,408 Da. The measured mass of 10404.8 Da equates to an isotopic labelling efficiency of 97.6%. The calculated mass of ^{15}N TGM6-D45 is 17,901 Da. The measured mass of 17888.9 Da equates to an isotopic labeling efficiency of 94.3%.

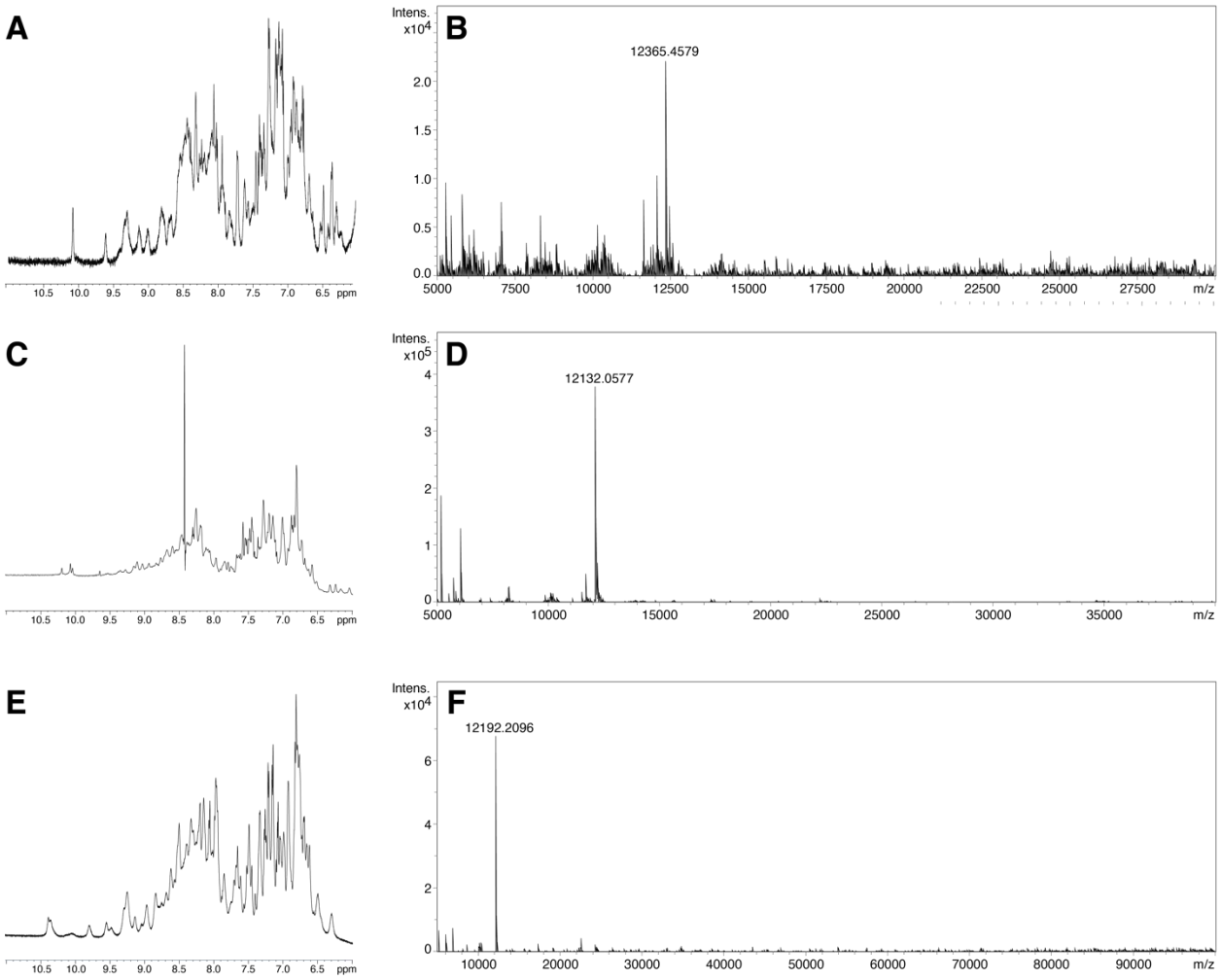

Figure 72. 1D 1H NMR spectra (A, C, E) and mass spectra (B, D, F) confirming the identity and native folding of (A, B) ActRII, (C, D) ActRIIb, and (E, F) BMPRII used in the ITC experiments.

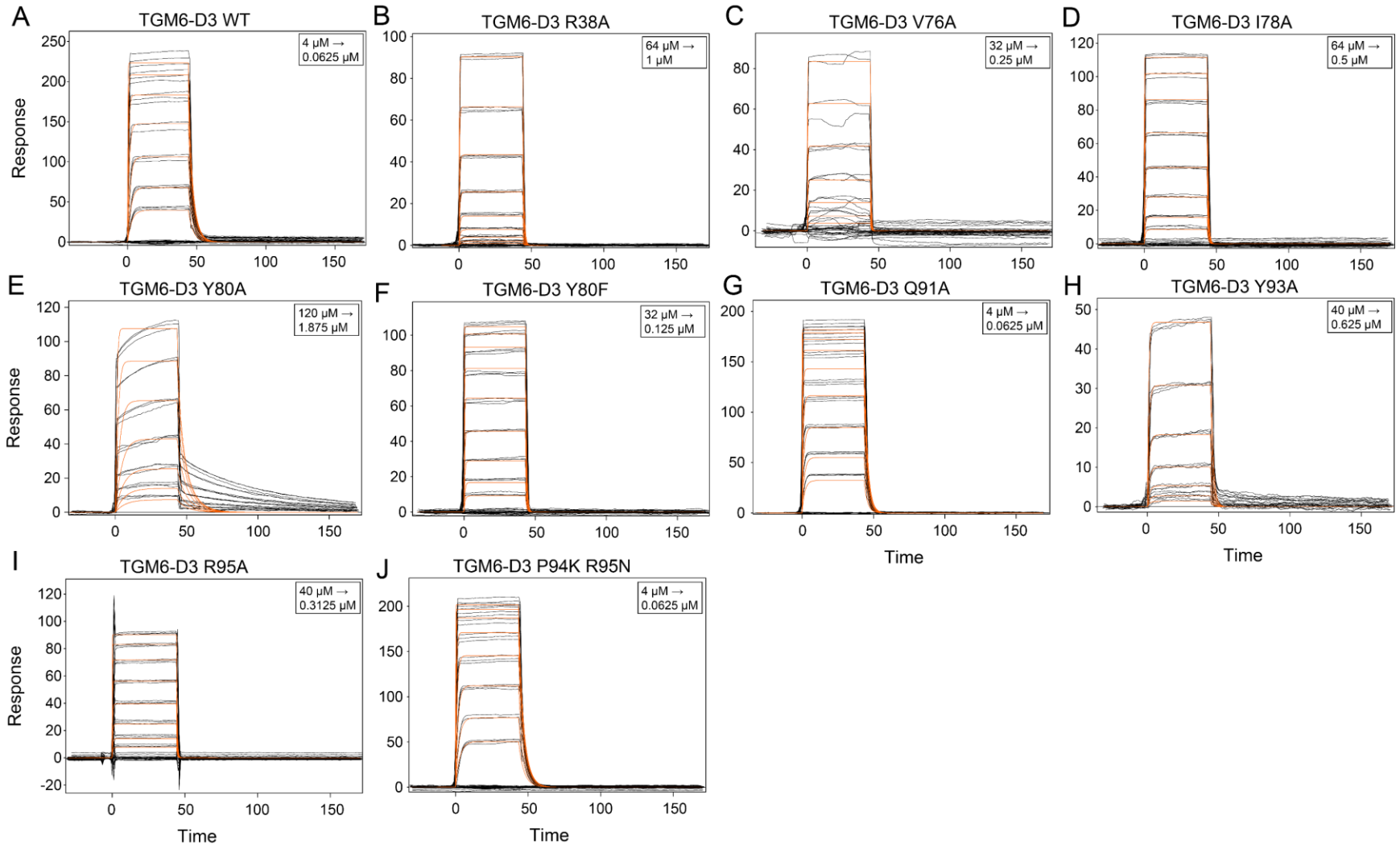

Figure 73. SPR traces for TGM6-D3 variants over a Bio-Avi-TβRII surface. (A) WT TGM6-D3, (B) TGM6-D3 R28A, (C) TGM6-D3 V66A, (D) TGM6-D3 I68A, (E) TGM6-D3 Y70A, (F) TGM6-D3 Y70F, (G) TGM6-D3 Q81A, (H) TGM6-D3 Y83A, (I) TGM6-D3 R85A, and (J) TGM6-D3 P84K R85N.

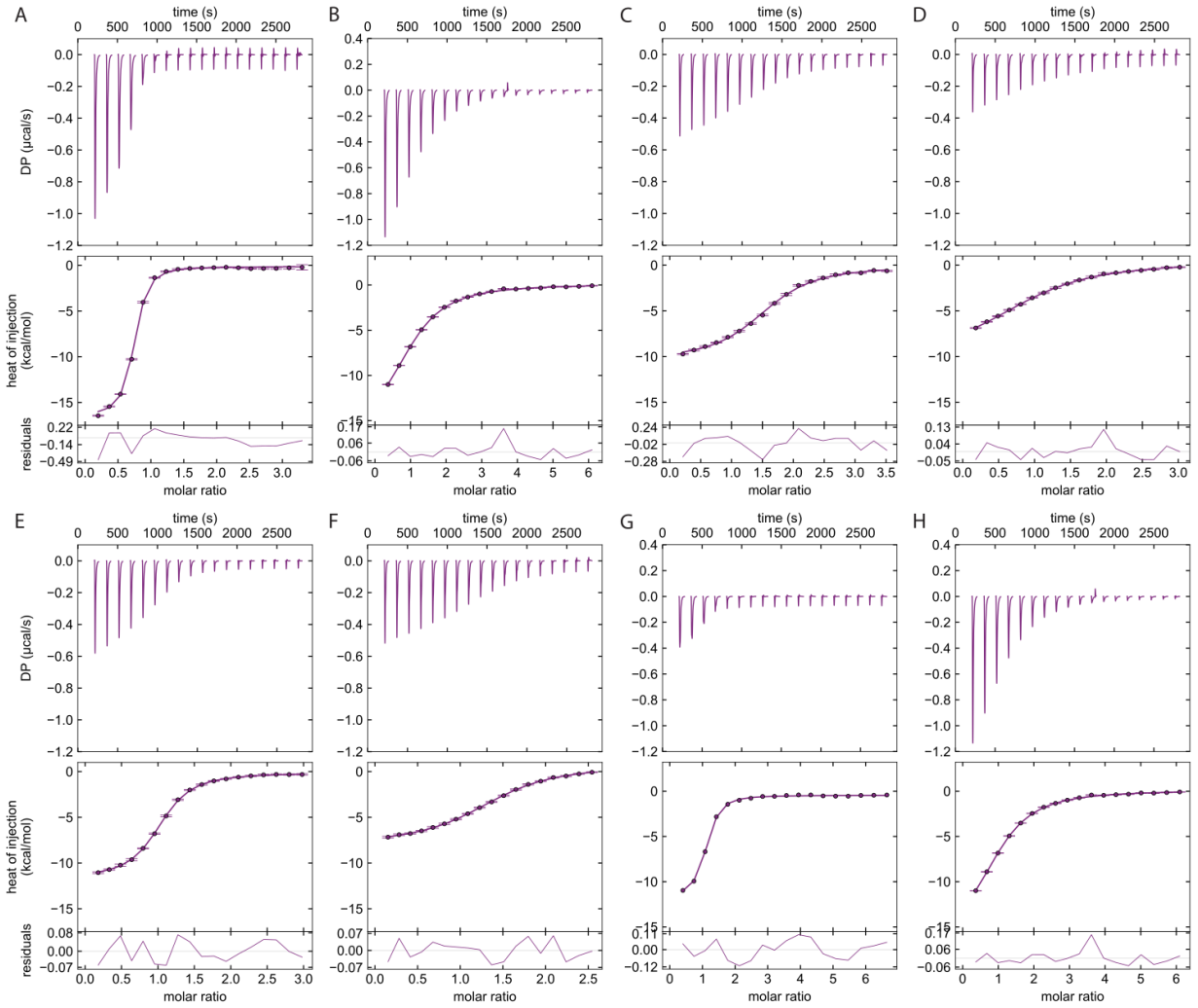

Figure 74. ITC thermograms obtained for WT TGM6-D3 injected into TβRII variants (A-F) and WT TβRII injected into TGM6-D3 variants (G-P). (A) TβRII WT (B) TβRII D55N, (C) TβRII S75L, (D) TβRII I76A, (E) TβRII D141A, (F) TβRII E142Q, (G) TGM6-D3 WT, (H) TGM6-D3 R28A, (I) TGM6-D3 V66A, (J) TGM6-D3 I68A, (K) TGM6-D3 Y70A, (L) TGM6-D3 Y70F, (M) TGM6-D3 Q81A, (N) TGM6-D3 Y83A, (O) TGM6-D3 R85A, and (P) TGM6-D3 P84K R85N.

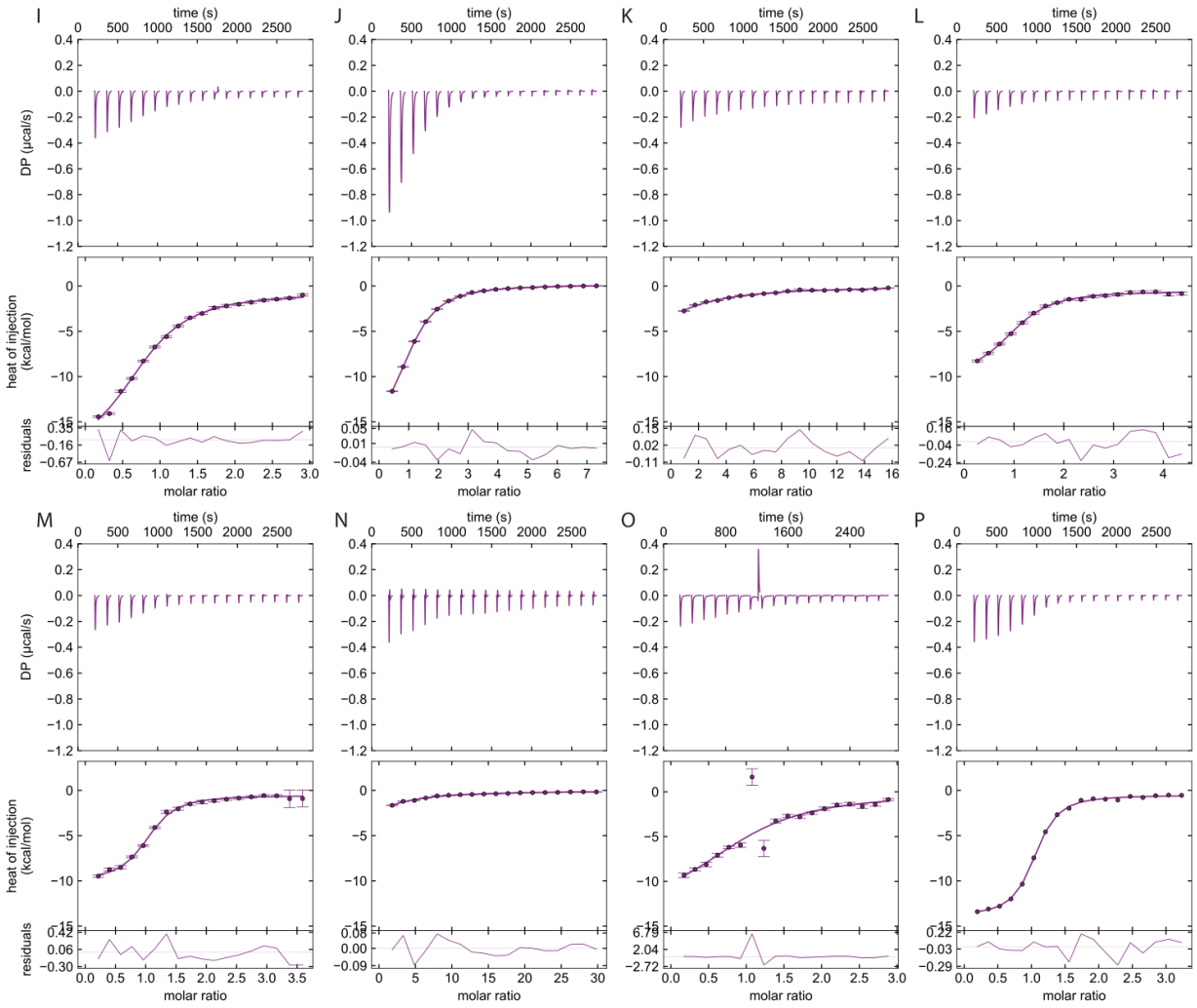

Figure 74 Continued.

Appendix C Summary of Polyglutamine Structure Project

The work summarized in this appendix was performed between May 2017 and November 2019 under the guidance of Dr. Sanford Asher of the Department of Chemistry at the University of Pittsburgh. In accordance with the agreement made with Dr. Asher, this work is presented in an appendix as an extended summary of the peer-reviewed, published work performed during the author's time in the Asher Lab.

Appendix C.1 UV Resonance Raman Spectroscopy

One of the basic facts that students are taught in high school chemistry is that all matter has motion at the atomic level, even if it doesn't seem like it's moving. This molecular motion can be broken down into different scales depending on the physical size of what component is moving and over what distance this component is moving. For proteins, these motions are summarized in Table 24. For example, the rotation of amino acid sidechains involves multiple atoms. These rotations occur over the picosecond time scale and over the distance of about an angstrom (0.1 nm). In contrast, the movement of a protein domain after the binding of a small molecule ligand occurs in the microsecond time scale and over much larger distances. Therefore, various scientific methods have been developed to characterize different types of motion.

Table 24. Timescale and length for atomic and molecular motions in proteins. Table adapted from reference [80].

Motions	Time Scale	Distance Scale
Local e.g. Atomic vibrations, sidechain motions, short loop motions	psec (10^{-15} – 10^{-1} sec)	0.01 – 5 Å
Rigid body motions e.g. Helix motion, domain motion, subunit motion	nsec (10^{-9} – 1 sec)	1 – 10 Å
Medium scale e.g. Helix-coil transitions, dissociation/association	μsec	>5 Å
Large scale e.g. Protein folding/unfolding, protein interactions	msec	>5 Å

There are two methods commonly used to measure atomic and molecular vibrations: infrared (IR) spectroscopy and Raman spectroscopy [74]. The differences in these modes of spectroscopy are summarized in Figure 75. IR spectroscopy is an absorption spectroscopy method. Infrared light is passed through the sample of interest, and the light is absorbed at frequencies corresponding to vibrational transitions of the sample. The transmission spectrum then shows the vibrational frequencies of the sample of interest [74]. There is no electronic transition associated with IR spectroscopy.

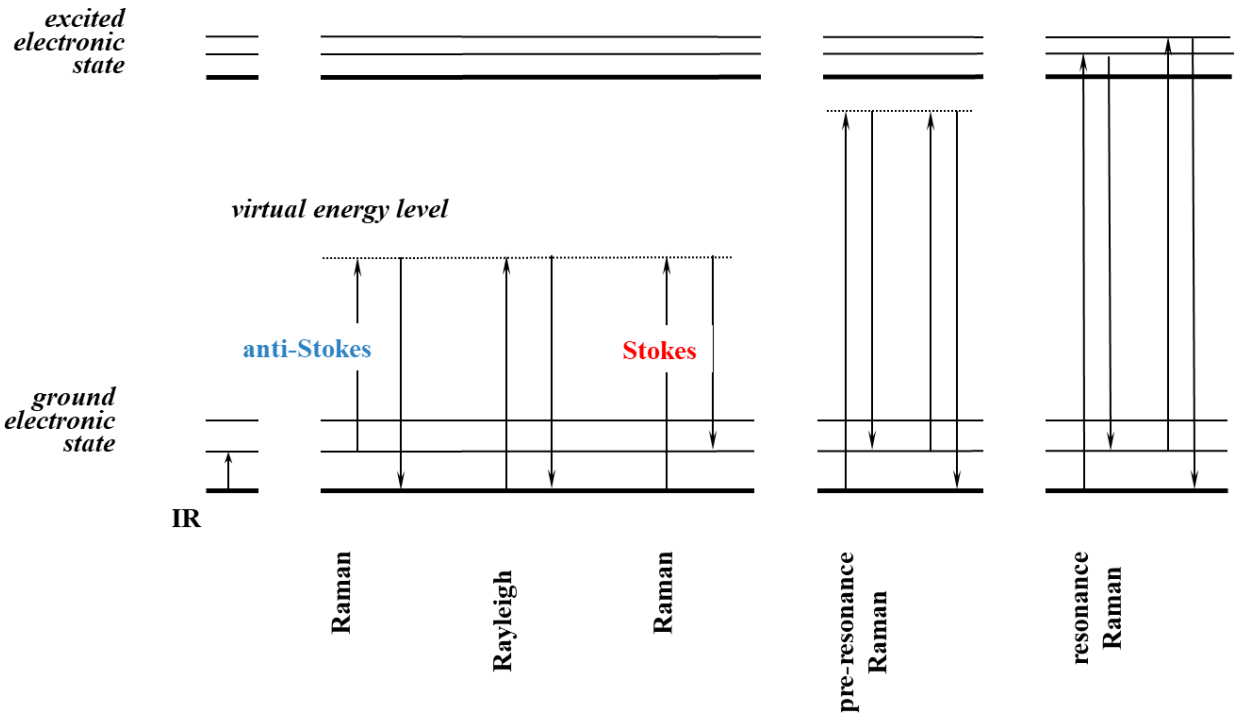

Figure 75. Energy level diagram of IR and Raman spectroscopy modes.

In contrast, Raman spectroscopy is a light scattering method [74]. A diagram of how a molecule scatters light is shown in Figure 76. A monochromatic beam of light with wavelength ν_0 is incident upon a sample of interest. This light is then scattered. The majority of this light (99.99%) is scattered with the same wavelength as the incident beam, ν_0. However, approximately one out of every 1000 photons is scattered with a frequency difference associated with a vibrational motion of the sample of interest, ν_v. If the frequency of the scattered light is less than that of the incident frequency $(\nu_0 - \nu_v)$, it is known as Stokes scattering; if the frequency of the scattered light is greater than that of the incident frequency $(\nu_0 + \nu_v)$, known as Anti-Stokes scattering. Together, the Stokes and Anti-Stokes scattering are collectively referred to as Raman scattering.

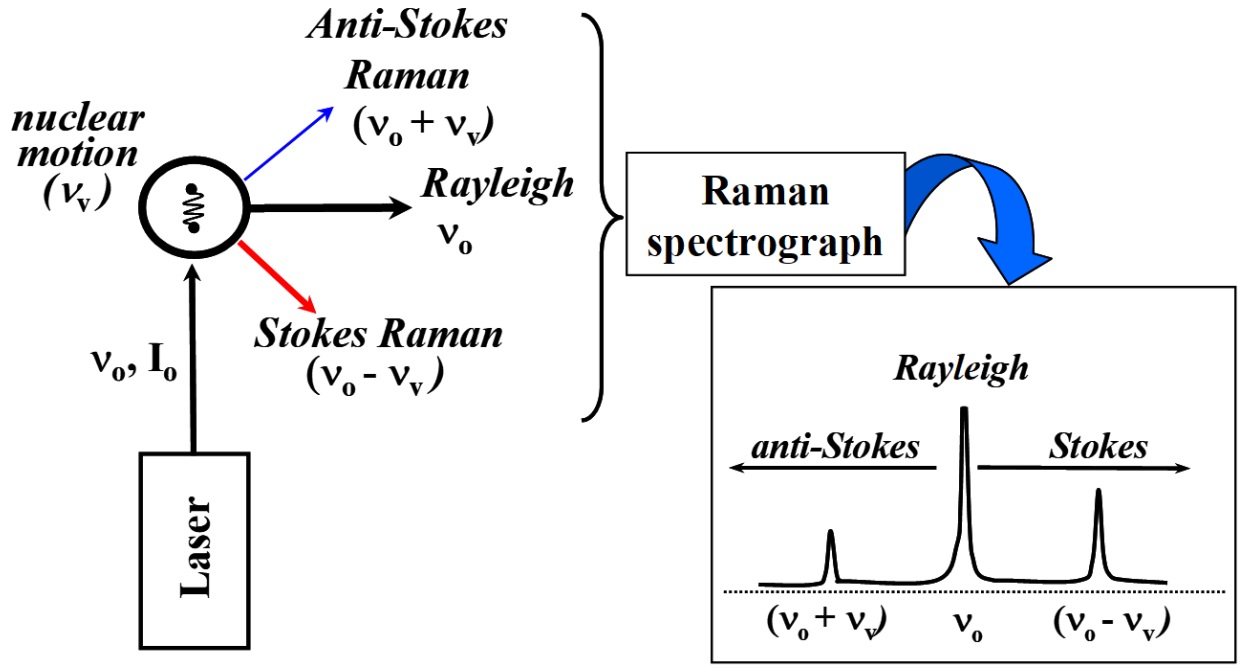

Figure 76. Modes of light scattering associated with Raman spectroscopy with a representative spectrum.

In "traditional" Raman spectroscopy, a monochromatic light beam of any wavelength can be used. This referred to as non-resonance, or simply, Raman spectroscopy. These conditions show vibrational modes without enhancing those associated with an electronic transition. However, using a light beam with a wavelength that is within an electronic absorption band enhances the vibrations that are coupled to that electronic absorption. This method is referred to as resonance Raman spectroscopy. Taking advantage of this coupling produces enhancements as high as 10^8-fold over non-resonance conditions for the vibrational bands coupled to the electronic transition within the chosen absorption band (Figure 77). By using multiple wavelengths of light, characterizations of different chromophores and their vibrational modes can be performed.

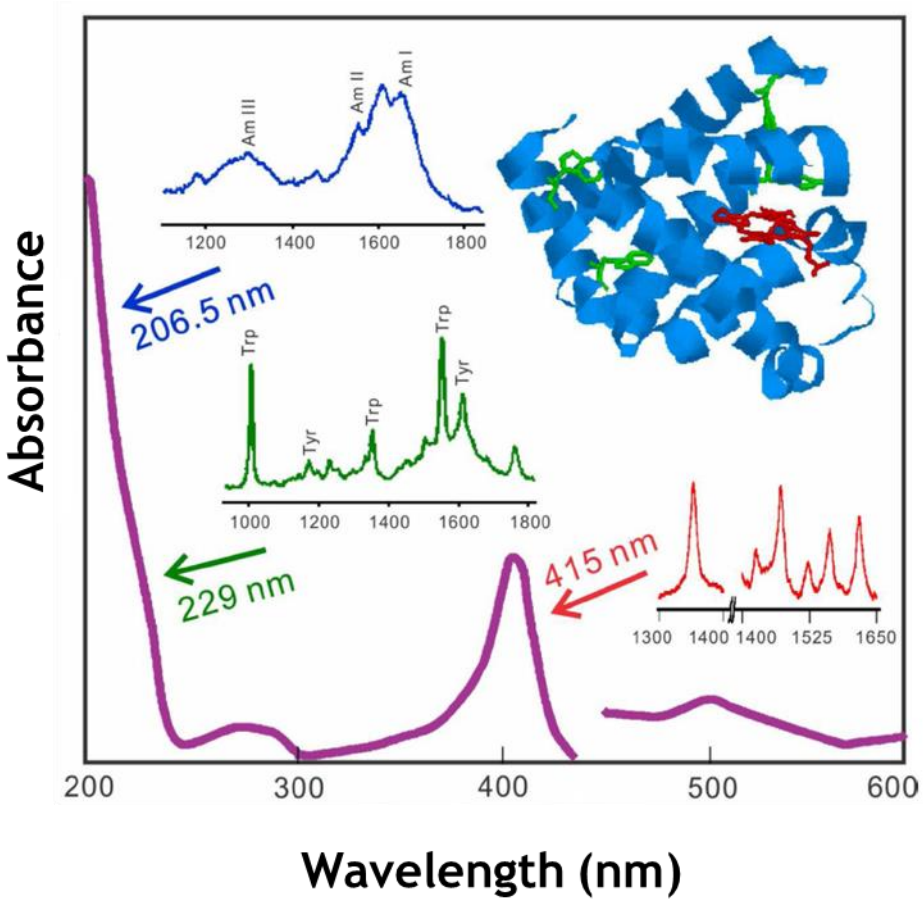

Figure 77. An example of three resonance Raman spectra at different incident wavelengths. 206.5 nm light enhances amide backbone vibrations, 229 nm light enhances aromatic ring vibrations, and 415 nm light enhances heme ring vibrations. This figure was reproduced with permission from Oladepo SA, Xiong K, Hong Z, Asher SA, Handen J, Lednev IK. UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. Chemical Reviews. 2012;112(5):2604-28. Copyright 2012 American Chemical Society.

As shown in Figure 77, proteins have multiple electronic absorption bands that can be utilized for resonance Raman spectroscopy. The Asher Lab has developed methods to characterize proteins based on these varying absorption bands. For reviews on this topic, see references [7, 98, 190]. Briefly, there are two sets of electronic absorptions used to characterize protein structure: 204, 210, and 214 nm excitations are in resonance with the primary, secondary, and tertiary amide vibrations, respectively, present in the protein backbone and glutamine and asparagine sidechains [7, 43, 190, 207, 277]; and 229 nm excitation is in resonance with the aromatic ring vibrations of

the tryptophan and tyrosine sidechains [6, 190]. These examples are chromophores in the ultraviolet (UV) range of the electromagnetic spectrum. Raman spectroscopy using UV excitation is referred to as UV resonance Raman (UVRR) spectroscopy. Additional protein chromophores, such as the heme ring in myoglobin, can also be explored by resonance Raman spectroscopy using 415 nm excitation [10]; however, this excitation wavelength is not in the UV.

Protein structure and folding are of great interest in the biomedical research community, as the structure of a protein provides clues to its function and vice-versa [236]. Therefore, tools that can probe the conformation and dynamics of proteins are valuable to the field. UV absorption spectroscopy [220] and circular dichroism (CD) [48] are able to detect changes in protein conformation, but only provide low-resolution structural information. X-ray crystallography yields atomic-level details about a protein's structure but is not able to easily give information on protein dynamics and requires that the protein be crystallizable [248]. NMR spectroscopy is a powerful tool for probing protein structure and dynamics but requires expensive isotopic labelling, high protein concentrations, and long measurement times [17].

Employing UVRR spectroscopy in the structural and functional exploration of proteins has advantages over the more traditional methods: the method is fast, requires no special sample preparation, and can be done with low analyte concentrations. However, there are three main drawbacks to the technique: 1) the structural data provided by UVRR are not atomic resolution but are instead at the ensemble level; 2) the UVRR signals can be noisy and difficult to interpret if not collected and processed carefully; and 3) the learning curve for employing UVRR spectroscopy can be steep for those not familiar with vibrational spectroscopy due to the nature of the technical setup and data processing.

Appendix C.2 Polyglutamine Peptide Repeat Disorders

Polyglutamine peptides are of clinical interest due to their association with a class of neurodegenerative diseases like Huntington's Disease. In this protein, the polyglutamine, or polyQ, tract is expanded to greater than 36 glutamine residues in patients who present clinical symptoms of the disease. These expanded polyQ tracts have a propensity to aggregate and form amyloid-like fibrils [192]. While the mechanism of toxicity and the identity of the toxic species are still debated [183, 221, 280], the common factor for polyQ-associated neurodegenerative diseases is the presence of an expanded polyQ tract in the proteins identified in the disease [192]. Clinical presentation of Huntington's Disease occurs for patients with polyQ tracts greater than 36 residues long, with longer polyQ tracts correlated with an earlier disease symptom age of onset [198]. Therefore, studying proteins and peptides of disease-length is of great interest.

Appendix C.3 Ultraviolet Resonance Raman Spectroscopic Markers for Protein Structure and Dynamics

Summarized from: Jakubek RS, Handen J, White SE, Asher SA, Lednev IK. Ultraviolet resonance Raman spectroscopic markers for protein structure and dynamics. TrAC Trends in Analytical Chemistry. 2018;103:223-9. doi: 10.1016/j.trac.2017.12.002.

In the review *Ultraviolet resonance Raman spectroscopic markers for protein structure and dynamics* [98], we discuss the current state of the use of UVRR spectroscopy as applied to protein structure and dynamics. Specifically, we show how the frequency and inhomogeneous

broadening of the amide III_3 band is used to determine the distribution of the protein backbone dihedral ψ angle. Additionally, the review discusses the use of Raman bands that describe the solvation state of protein side chains, such as the tyrosine and tryptophan aromatic rings and the arginine guanidinium groups. Furthermore, we show how the Amide III_3 band can also be used to determine the glutamine and asparagine side chain χ_2 and χ_3 dihedral distributions. Finally, the review discusses the application of these markers to determining the solution state and fibril structures of a polyglutamine peptide with a glutamine repeat length of 10, the method of two-dimensional correlation deep UV resonance Raman spectroscopy, and how hydrogen-deuterium exchange can be combined with UVRR spectroscopy for probing the structure of the core of polyglutamine fibrils.

Appendix C.4 UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide

Summarized from: Jakubek RS, White SE, Asher SA. UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide. J Phys Chem B. 2019;123(8):1749-63. Epub 2019/02/06. doi: 10.1021/acs.jpcb.8b10783. PubMed PMID: 30717595.

The fibrillization of polyQ is thought to result from the peptide's insolubility in aqueous solutions; longer polyQ tracts show decreased aqueous solution solubility, which is thought to lead to faster fibrillization kinetics. However, few studies have characterized the structure(s) of polyQ peptides with low solubility. In the work titled *UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide* [99], we used UVRR spectroscopy to investigate the

secondary structures, backbone hydrogen bonding, and sidechain hydrogen bonding for a variety of solution-state, solid, and fibril forms of the polyQ peptide $D_2K_{20}K_2$ (Q20). These forms are summarized in Figure 78.

In this work, we showed that Q20 is insoluble in water and has a β-strand-like conformation with extensive inter- and intrapeptide hydrogen bonding in both dry and aqueous environments. We found that Q20 has weaker backbone−backbone and backbone−side chain hydrogen bonding and is less ordered compared to that of polyQ fibrils. Interestingly, we also found that the insoluble Q20 will form fibrils when incubated in water at room temperature for ∼5 hours. Furthermore, Q20 was prepared using a well-known disaggregation procedure [32] to produce a water-soluble polyproline II- (PPII-)like conformation. This water-soluble form had negligible inter- and intrapeptide hydrogen bonding and a resistance to aggregation.

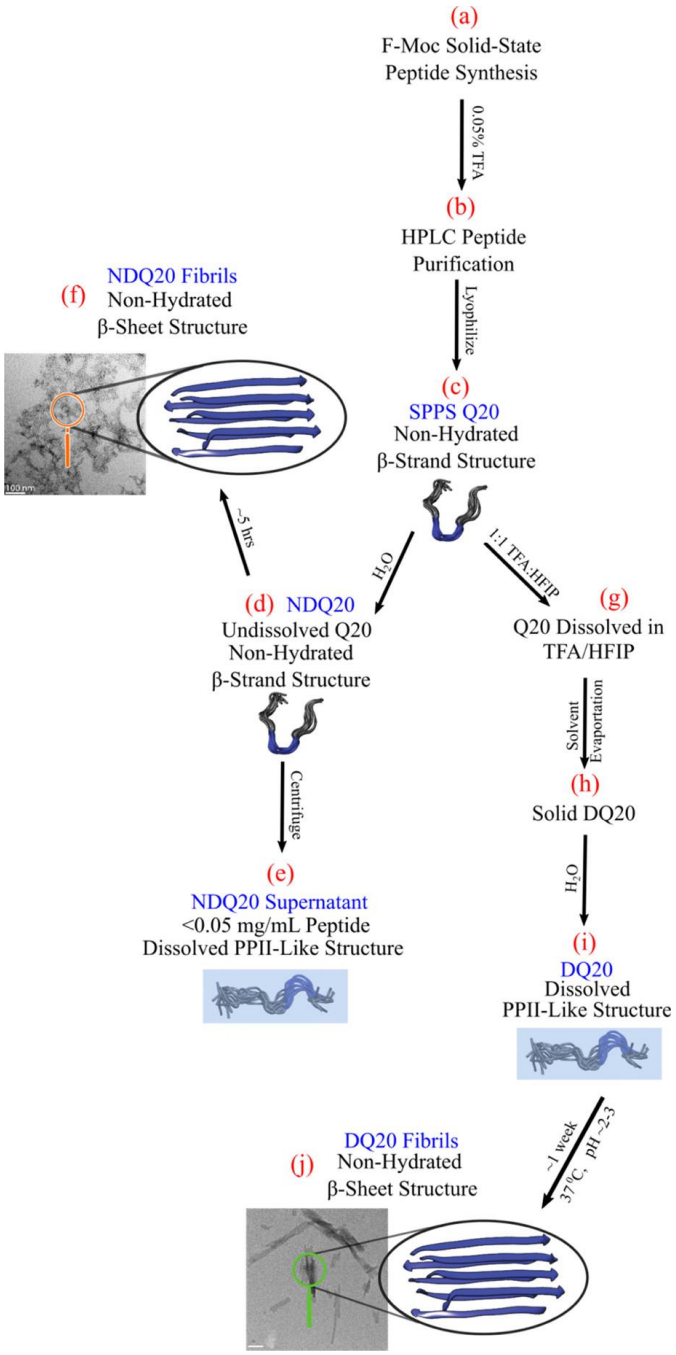

Figure 78. Summary of the forms of Q20 examined in reference [99]. Each letter (a−j) indicates a form of the Q20 peptide. The blue text indicates the nomenclature for this particular form of Q20. For descriptions of the forms, structures, and procedures required to make the samples, please see the original text. This figure was reproduced with permission from Jakubek RS, White SE, Asher SA. UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide. J Phys Chem B. 2019;123(8):1749-63. Copyright 2019 American Chemical Society.

Appendix C.5 Polyglutamine Solution-State Structural Propensity is Repeat Length Dependent

Summarized from: Jakubek RS, Workman RJ, White SE, Asher SA. Polyglutamine Solution-State Structural Propensity Is Repeat Length Dependent. J Phys Chem B. 2019;123(19):4193-203. Epub 2019/04/23. doi: 10.1021/acs.jpcb.9b01433. PubMed PMID: 31008597.

As stated previously, longer polyQ tracts correlate with faster protein aggregation kinetics and a decreased age of onset for polyQ disease symptoms [31, 33, 182, 198]. In the work titled *Polyglutamine Solution-State Structural Propensity is Repeat Length Dependent* [100], we used UVRR spectroscopy, CD spectroscopy, and metadynamics simulations to investigate the solution-state structures of the $D_2Q_{15}K_2$ (Q15) and $D_2Q_{20}K_2$ (Q20) polyQ peptides. Using metadynamics, we explored the conformational energy landscapes of Q15 and Q20 and investigated the relative energies and activation barriers between these low-energy structures, shown in Figure 79. We compared the solution-state structures of $D_2Q_{10}K_2$ (Q10) (previously studied by Punihaole, et al. [208, 209]), Q15, and Q20 to determine the dependence of polyQ structure on the glutamine tract length. We showed that these peptides can adopt two distinct monomeric conformations: an aggregation-resistant PPII-like conformation and an aggregation-prone β-strand-like conformation. We found that longer polyQ peptides have an increased preference for the aggregation-prone β-strand-like conformation. This preference may play an important role in the increased aggregation rate of longer polyQ peptides that is thought to lead to decreased neurodegenerative disease age of onset for polyQ disease patients.

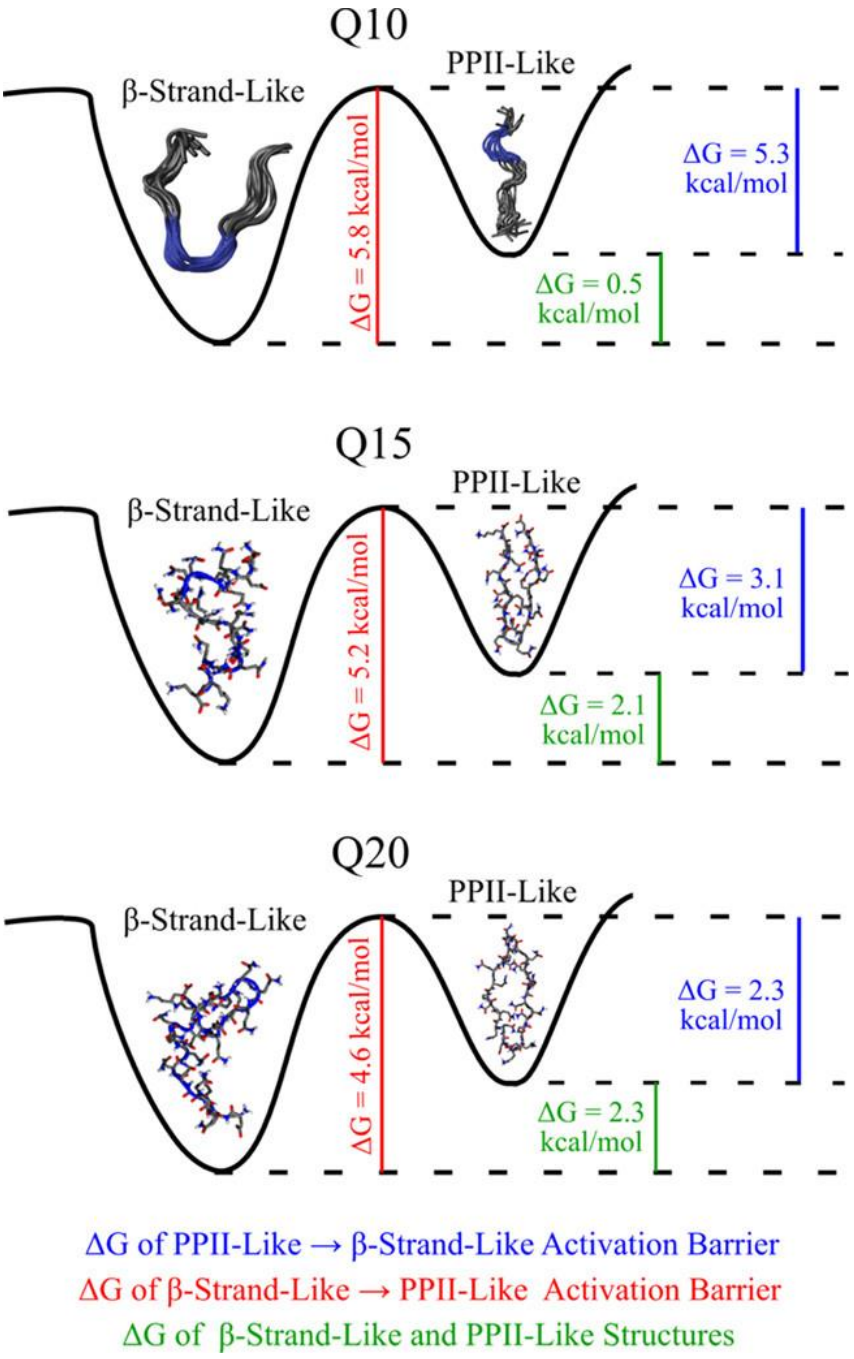

Figure 79. Depiction of relative energies and activation barriers for the Q10–20 PPII-like and β-strand-like conformations. The activation barrier ΔG for the β-strand → PPII transitions is shown in red, the activation barrier ΔG for the PPII → β-strand transitions is shown in blue, and the relative ΔG of the β-strand-like and PPII-like minima is shown in green. This figure was reproduced with permission from Jakubek RS, Workman RJ, White SE, Asher SA. Polyglutamine Solution-State Structural Propensity Is Repeat Length Dependent. J Phys Chem B. 2019;123(19):4193-203. Copyright 2019 American Chemical Society.

Appendix C.6 Conclusions

The works summarized here show the current state of the development of UVRR spectroscopy for use in characterization of protein structures with examples of how these methods can be applied to polyQ peptides in various solution-state, solid, and fibril forms. We used UVRR spectroscopy and metadynamics simulations to characterize the structures of polyQ peptides in these various forms. Additionally, we showed that polyQ peptides with a polyQ tract of less than 20 residues have a propensity for forming fibrils while in a β-strand-like conformation but are resistant to aggregation and fibril formation when in a PPII-like conformation. This work has helped improve upon the understanding of the aggregation mechanism behind the proteins implicated in polyglutamine repeat disorders.

Appendix C.7 Author Contributions

For the work titled *Ultraviolet resonance Raman markers for protein structure and dynamics* [98], SEW contributed the general outline and structure of the introduction and discussion sections, assisted with figure creation, and reviewed and edited the manuscript prior to submission for review.

For the work titled *UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide* [99], SEW worked with RSJ (first author) to prepare samples and to collect and analyze UVRR spectra of DQ20 and NDQ20 samples. Additionally, SEW contributed the TEM images of DQ20 and NDQ20 fibrils, assisted with the preparation of the manuscript, and reviewed and edited the manuscript prior to submission for review.

For the work titled *Polyglutamine Solution-State Structural Propensity Is Repeat Length Dependent* [100], SEW worked with RSJ (first author) to prepare samples and to collect and analyze UVRR and CD spectra of DQ15, NDQ15, DQ20, and NDQ20 samples. Additionally, SEW assisted with the preparation of the manuscript and reviewed and edited the manuscript prior to submission for review.

Appendix D List of Abbreviations and Acronyms

^{1}H – Proton

^{13}C – Carbon-13

^{15}N – Nitrogen-15

aa – amino acid

ActRI – Activin receptor like kinase type I; Alk1

ActRIA – Activin receptor like kinase type IA; Alk2

ActRIB – Activin receptor like kinase type IB; Alk4

ActRII – Activin receptor type II

ActRIIb – Activin receptor type II-b

Ala – Alanine

Alk1 – activin A receptor type II-like kinase 1; ActRI

Alk2 – activin A receptor type II-like kinase 2; ActRIA

Alk3 – activin A receptor type II-like kinase 3; BMPRIA

Alk4 – activin A receptor type II-like kinase 4; ActRIB

Alk5 – activin A receptor type II-like kinase 5; TβRI

Alk6 – activin A receptor type II-like kinase 6; BMPRIB

Alk7 – activin A receptor type II-like kinase 7; ActRIC

AMH – Anti-Müllerian Hormone

Arg – Arginine

Asn – Asparagine

Asp – Aspartate/Aspartic acid

BMP – Bone Morphogenic Protein

BMPRIA – BMP receptor type IA

BMPRII – BMP receptor type II

CCP – Complement control protein

CD3 – Cluster of differentiation 3

CD4 – Cluster of differentiation 4

CD8 – Cluster of differentiation 8

CD25 – Cluster of differentiation 25; Interleukin-2 receptor alpha chain

CI – Confidence interval

CKGF – Cystine knot growth factor

CT-HSQC – Constant-Time Heteronuclear Single Quantum Correlation

CTLA-4 – Cytotoxic T-lymphocyte-associated protein 4

Cys – Cysteine

D1 – Domain 1

D12 – Domains 1 and 2

D2 – Domain 2

D3 – Domain 3

D4 – Domain 4

D45 – Domains 4 and 5

D5 – Domain 5

DMEM – Dulbecco's modified Eagle's medium

DQ15 – Disaggregated $D_2Q_{15}K_2$ peptide

DQ20 – Disaggregated $D_2Q_{20}K_2$ peptide

ECM – Extracellular matrix

EMT – epithelial-to-mesenchymal transition

FBS – Fetal bovine serum

FCS – Fetal calf serum

flTGM6 – full-length TGF-β mimic 6

FOP – Fibrodysplasia Ossificans Progressiva

Foxp3 – Forkhead box protein 3

GDF – Growth and Differentiation Factor

Gln – Glutamine

Glu – Glutamate/Glutamic acid

Gly – Glycine

GSH – reduced glutathione

GSSG – oxidized glutathione

HES – *H. polygyrus* excretory-secretory products

HHT – Hereditary Hemorrhagic Telangiectasia

His – Histidine

HMQC – Heteronuclear Multiple-Quantum Correlation

HpTGM – *H. polygyrus* TGF-β mimic

HSQC – Heteronuclear Single Quantum Correlation

I-PINE – Integrative PINE

ICOS – Inducible T-cell co-stimulator

IL-10 – Interleukin 10

IL-35 – Interleukin 35

Ile – Isoleucine

IPAP-HSQC – In-Phase/Anti-Phase Heteronuclear Single Quantum Correlation

IPF – Idiopathic pulmonary fibrosis

IR - Infrared

ITC – Isothermal titration calorimetry

LB – Lauria Broth

LC-ESI-TOF-MS – liquid chromatography electrospray ionization time-of-flight mass spectroscopy

Leu – Leucine

LSP – Larval secreted protein

Lys – Lysine

Met – Methionine

NDQ15 – Non-disaggregated $D_2Q_{15}K_2$ peptide

NDQ20 – Non-disaggregated $D_2Q_{20}K_2$ peptide

NMR – Nuclear magnetic resonance

NOE – Nuclear Overhauser Effect

NOESY – Nuclear Overhauser Effect Spectroscopy

PBS – phosphate-buffered saline

PD-1 – Programmed cell death protein 1

PD-L1 – Programmed death-ligand 1

PDB – Protein databank

Phe – Phenylalanine

PINE - probabilistic interaction network of evidence

PPII – polyproline II

polyQ – polyglutamine

Pro – Proline

Q10 – $D_2Q_{10}K_2$ peptide

Q15 – $D_2Q_{15}K_2$ peptide

Q20 – $D_2Q_{20}K_2$ peptide

R-Smad – Regulatory Smad

RDC – Residual Dipolar Coupling

SBE – Smad Binding Element

Ser – Serine

SPR – Surface plasmon resonance

TβRI – TGF-β receptor type I; Alk5

TβRII – TGF-β receptor type II

TβRIII – TGF-β receptor type III

TGF-β – Transforming Growth Factor beta

TGM – TGF-β mimic

TGM1 – TGF-β mimic 1

TGM2 – TGF-β mimic 2

TGM3 – TGF-β mimic 3

TGM4 – TGF-β mimic 4

TGM5 – TGF-β mimic 5

TGM6 – TGF-β mimic 6

TGM7 – TGF-β mimic 7

TGM8 – TGF-β mimic 8

TGM9 – TGF-β mimic 9

TGM10 – TGF-β mimic 10

Thr – Threonine

TIGIT – T-cell Immunoreceptor with Ig and ITIM domains

TOCSY – Total Correlation Spectroscopy

Tregs – Regulatory T-cells

Trp – Tryptophan

Tyr – Tyrosine

UV – ultraviolet

UVRR – ultraviolet resonance Raman

VAL – venom-allergen-like

Val – Valine

Works Cited

1. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D. 2010;66(2):213-21. doi: 10.1107/S0907444909052925.

2. Akhurst RJ. Targeting TGF-beta Signaling for Therapeutic Gain. Cold Spring Harb Perspect Biol. 2017;9(10). Epub 2017/03/02. doi: 10.1101/cshperspect.a022301. PubMed PMID: 28246179; PMCID: PMC5630004.

3. Allendorph GP, Vale WW, Choe S. Structure of the ternary signaling complex of a TGF-β superfamily member. Proceedings of the National Academy of Sciences. 2006;103(20):7643-8. doi: 10.1073/pnas.0602558103.

4. Annacker O, Asseman C, Read S, Powrie F. Interleukin-10 in the regulation of T cell-induced colitis. Journal of Autoimmunity. 2003;20(4):277-9. doi: 10.1016/s0896-8411(03)00045-3.

5. Archer SJ, Ikura M, Torchia DA, Bax A. An Alternative 3D NMR Technique for Correlating Backbone 15N with Side Chain Hbeta Resonances in Larger Proteins. J Magn Reson. 1991;95:636-41.

6. Asher SA, Ludwig M, Johnson CR. UV resonance Raman excitation profiles of the aromatic amino acids. Journal of the American Chemical Society. 1986;108(12):3186-97.

7. Asher SA. UV resonance Raman spectroscopy for analytical, physical, and biophysical chemistry. Part 2. Analytical chemistry. 1993;65(4):201A-10A.

8. Baardsnes J, Hinck CS, Hinck AP, O'Connor-McCourt MD. TbetaR-II discriminates the high- and low-affinity TGF-beta isoforms via two hydrogen-bonded ion pairs. Biochemistry. 2009;48(10):2146-55. Epub 2009/01/24. doi: 10.1021/bi8019004. PubMed PMID: 19161338; PMCID: PMC2801812.

9. Bahrami A, Assadi AH, Markley JL, Eghbalnia HR. Probabilistic Interaction Network of Evidence Algorithm and its Application to Complete Labeling of Peak Lists from Protein NMR Spectroscopy. PLoS Computational Biology. 2009;5(3):e1000307. doi: 10.1371/journal.pcbi.1000307.

10. Balakrishnan G, Hu Y, Oyerinde OF, Su J, Groves JT, Spiro TG. A conformational switch to β-sheet structure in cytochrome c leads to heme exposure. Implications for cardiolipin peroxidation and apoptosis. Journal of the American Chemical Society. 2007;129(3):504-5.

11. Bashir ME, Andersen P, Fuss IJ, Shi HN, Nagler-Anderson C. An enteric helminth infection protects against an allergic response to dietary antigen. J Immunol. 2002;169(6):3284-92. Epub 2002/09/10. doi: 10.4049/jimmunol.169.6.3284. PubMed PMID: 12218148.

12. Batlle E, Massagué J. Transforming Growth Factor-β Signaling in Immunity and Cancer. Immunity. 2019;50(4):924-40. doi: 10.1016/j.immuni.2019.03.024.

13. Bax A, Clore GM, Gronenborn AM. 1H-1H correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins. Journal of Magnetic Resonance (1969). 1990;88(2):425-31. doi: 10.1016/0022-2364(90)90202-k.

14. Bax A, Kontaxis G, Tjandra N. [8] - Dipolar Couplings in Macromolecular Structure Determination. In: James TL, Dötsch V, Schmitz U, editors. Methods in Enzymology: Academic Press; 2001. p. 127-74.

15. Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Science. 2003;12(1):1-16. doi: 10.1110/ps.0233303.

16. Behnke JM, Menge DM, Noyes H. Heligmosomoides bakeri: a model for exploring the biology and genetics of resistance to chronic gastrointestinal nematode infections. Parasitology. 2009;136(12):1565-80. doi: 10.1017/s0031182009006003.

17. Billeter M. Comparison of protein structures determined by NMR in solution and by X-ray diffraction in single crystals. Quarterly reviews of biophysics. 1992;25(3):325-77.

18. Blackledge M. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. Progress in Nuclear Magnetic Resonance Spectroscopy. 2005;46(1):23-61. doi: 10.1016/j.pnmrs.2004.11.002.

19. Bloomfield SF, Rook GA, Scott EA, Shanahan F, Stanwell-Smith R, Turner P. Time to abandon the hygiene hypothesis: new perspectives on allergic disease, the human microbiome, infectious disease prevention and the role of targeted hygiene. Perspect Public Health. 2016;136(4):213-24. Epub 2016/06/30. doi: 10.1177/1757913916650225. PubMed PMID: 27354505; PMCID: PMC4966430.

20. Bodenhausen G, Ruben DJ. Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. Chem Phys Lett. 1980;69(1):185-9.

21. Boesen CC, Radaev S, Motyka SA, Patamawenu A, Sun PD. The 1.1 Å Crystal Structure of Human TGF-β Type II Receptor Ligand Binding Domain. Structure. 2002;10(7):913-9. doi: 10.1016/S0969-2126(02)00780-3.

22. Brautigam CA. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. In: Cole JL, editor. Methods in Enzymology: Academic Press; 2015. p. 109-33.

23. Brautigam CA, Zhao H, Vargas C, Keller S, Schuck P. Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. Nature Protocols. 2016;11(5):882-94. doi: 10.1038/nprot.2016.044.

24. Brummel TJ, Twombly V, Marqués G, Wrana JL, Newfeld SJ, Attisano L, Massagué J, O'Connor MB, Gelbart WM. Characterization and relationship of dpp receptors encoded by the saxophone and thick veins genes in Drosophila. Cell. 1994;78(2):251-61. doi: 10.1016/0092-8674(94)90295-X.

25. Carlomagno T, Maurer M, Sattler M, Schwendinger MG, Glaser SJ, Griesinger C. PLUSH TACSY: Homonuclear planar TACSY with two-band selective shaped pulses applied to C(alpha),C' transfer and C (beta),C (aromatic) correlations. J Biomol NMR. 1996;8(2):161-70. Epub 1996/09/01. doi: 10.1007/BF00211162. PubMed PMID: 22911140.

26. Cavanagh J, Fairbrother WJ, Palmer III AG, Skelton NJ. Protein NMR spectroscopy: principles and practice: Elsevier; 1995.

27. Chaikuad A, Bullock AN. Structural Basis of Intracellular TGF-beta Signaling: Receptors and Smads. Cold Spring Harb Perspect Biol. 2016;8(11). Epub 2016/11/03. doi: 10.1101/cshperspect.a022111. PubMed PMID: 27549117; PMCID: PMC5088531.

28. Cheifetz S, Weatherbee JA, Tsang MLS, Anderson JK, Mole JE, Lucas R, Massagué J. The transforming growth factor-β system, a complex pattern of cross-reactive ligands and receptors. Cell. 1987;48(3):409-15. doi: 10.1016/0092-8674(87)90192-9.

29. Cheifetz S, Hernandez H, Laiho M, ten Dijke P, Iwata KK, Massagué J. Distinct transforming growth factor-beta (TGF-beta) receptor subsets as determinants of cellular responsiveness to three TGF-beta isoforms. Journal of Biological Chemistry. 1990;265(33):20533-8. doi: 10.1016/S0021-9258(17)30535-5.

30. Chen C-R, Kang Y, Massagué J. Defective repression of c-myc in breast cancer cells: A loss at the core of the transforming growth factor β growth arrest program. Proceedings of the National Academy of Sciences. 2001;98(3):992-9. doi: 10.1073/pnas.98.3.992.

31. Chen S, Berthelier V, Yang W, Wetzel R. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. J Mol Biol. 2001;311(1):173-82. doi: 10.1006/jmbi.2001.4850. PubMed PMID: 11469866.

32. Chen S, Wetzel R. Solubilization and disaggregation of polyglutamine peptides. Protein Sci. 2001;10(4):887-91. doi: 10.1110/ps.42301. PubMed PMID: 11274480; PMCID: PMC2373971.

33. Chen S, Ferrone FA, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci U S A. 2002;99(18):11884-9. Epub 20020819. doi: 10.1073/pnas.182276099. PubMed PMID: 12186976; PMCID: PMC129363.

34. Chen W, Jin W, Hardegen N, Lei KJ, Li L, Marinos N, McGrady G, Wahl SM. 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. 2003;198(12):1875-86. doi: 10.1084/jem.20030152. PubMed PMID: 14676299; PMCID: PMC2194145.

35. Clowes RT, Boucher W, Hardman CH, Domaille PJ, Laue ED. A 4D HCC(CO)NNH experiment for the correlation of aliphatic side-chain and backbone resonances in 13C/15N-labelled proteins. Journal of Biomolecular NMR. 1993;3(3):349-54. doi: 10.1007/bf00212520.

36. Clubb RT, Thanabal V, Wagner G. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C′ chemical shifts in 15N-13C-labelled proteins. Journal of Magnetic Resonance (1969). 1992;97(1):213-7. doi: 10.1016/0022-2364(92)90252-3.

37. Collison LW, Workman CJ, Kuo TT, Boyd K, Wang Y, Vignali KM, Cross R, Sehy D, Blumberg RS, Vignali DA. The inhibitory cytokine IL-35 contributes to regulatory T-cell function. Nature. 2007;450(7169):566-9. Epub 2007/11/23. doi: 10.1038/nature06306. PubMed PMID: 18033300.

38. Cull MG, Schatz PJ. [26] Biotinylation of proteins in vivo and in vitro using small peptide tags. Methods in Enzymology: Academic Press; 2000. p. 430-40.

39. De Crescenzo G, Hinck CS, Shu Z, Zuniga J, Yang J, Tang Y, Baardsnes J, Mendoza V, Sun L, Lopez-Casillas F, O'Connor-McCourt M, Hinck AP. Three key residues underlie the differential affinity of the TGFbeta isoforms for the TGFbeta type II receptor. J Mol Biol. 2006;355(1):47-62. Epub 2005/11/23. doi: 10.1016/j.jmb.2005.10.022. PubMed PMID: 16300789.

40. Deep S, Walker KP, 3rd, Shu Z, Hinck AP. Solution structure and backbone dynamics of the TGFbeta type II receptor extracellular domain. Biochemistry. 2003;42(34):10126-39. Epub 2003/08/27. doi: 10.1021/bi034366a. PubMed PMID: 12939140.

41. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995;6(3):277-93. Epub 1995/11/01. doi: 10.1007/BF00197809. PubMed PMID: 8520220.

42. Donskow-Lysoniewska K, Krawczak K, Doligalska M. Heligmosomoides polygyrus: EAE remission is correlated with different systemic cytokine profiles provoked by L4 and adult nematodes. Exp Parasitol. 2012;132(2):243-8. Epub 2012/08/18. doi: 10.1016/j.exppara.2012.07.009. PubMed PMID: 22898371.

43. Dudik JM, Johnson CR, Asher SA. UV resonance Raman studies of acetone, acetamide, and N-methylacetamide: models for the peptide bond. The Journal of Physical Chemistry. 1985;89(18):3805-14.

44. Dux P, Whitehead B, Boelens R, Kaptein R, Vuister GW. Measurement of (15)N- (1)H coupling constants in uniformly (15)N-labeled proteins: Application to the photoactive yellow protein. J Biomol NMR. 1997;10(3):301-6. Epub 1997/10/01. doi: 10.1023/A:1018393225804. PubMed PMID: 20700833.

45. Eghbalnia HR, Wang L, Bahrami A, Assadi A, Markley JL. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. Journal of Biomolecular NMR. 2005;32(1):71-81. doi: 10.1007/s10858-005-5705-1.

46. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallographica Section D. 2010;66(4):486-501. doi: 10.1107/S0907444910007493.

47. Esparza-López J, Montiel JL, Vilchis-Landeros MM, Okadome T, Miyazono K, López-Casillas F. Ligand Binding and Functional Properties of Betaglycan, a Co-receptor of the Transforming Growth Factor-β Superfamily. Journal of Biological Chemistry. 2001;276(18):14588-96. doi: 10.1074/jbc.m008866200.

48. Fasman GD. Circular dichroism and the conformational analysis of biomolecules: Springer Science & Business Media; 2013.

49. Finlay CM, Walsh KP, Mills KHG. Induction of regulatory cells by helminth parasites: exploitation for the treatment of inflammatory diseases. Immunological Reviews. 2014;259(1):206-30. doi: 10.1111/imr.12164.

50. Finney CA, Taylor MD, Wilson MS, Maizels RM. Expansion and activation of CD4(+)CD25(+) regulatory T cells in Heligmosomoides polygyrus infection. Eur J Immunol. 2007;37(7):1874-86. Epub 2007/06/15. doi: 10.1002/eji.200636751. PubMed PMID: 17563918; PMCID: PMC2699425.

51. Finney CAM, Taylor MD, Wilson MS, Maizels RM. Expansion and activation of CD4+CD25+ regulatory T cells in Heligmosomoides polygyrus infection. European Journal of Immunology. 2007;37(7):1874-86. doi: 10.1002/eji.200636751.

52. Gaboriaud C, Rossi V, Bally I, Arlaud GJ, Fontecilla-Camps JC. Crystal structure of the catalytic domain of human complement C1s: a serine protease with a handle. The EMBO Journal. 2000;19(8):1755-65. doi: 10.1093/emboj/19.8.1755.

53. Gajewski TF. The Next Hurdle in Cancer Immunotherapy: Overcoming the Non–T-Cell–Inflamed Tumor Microenvironment. Seminars in Oncology. 2015;42(4):663-71. doi: 10.1053/j.seminoncol.2015.05.011.

54. Ganesh K, Massagué J. TGF-β Inhibition and Immunotherapy: Checkmate. Immunity. 2018;48(4):626-8. doi: 10.1016/j.immuni.2018.03.037.

55. GBD 2017 SDG Collaborators. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries

and territories: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):2091-138. doi: 10.1016/s0140-6736(18)32281-5.

56. Gondek DC, Lu LF, Quezada SA, Sakaguchi S, Noelle RJ. Cutting edge: contact-mediated suppression by CD4+CD25+ regulatory cells involves a granzyme B-dependent, perforin-independent mechanism. J Immunol. 2005;174(4):1783-6. Epub 2005/02/09. doi: 10.4049/jimmunol.174.4.1783. PubMed PMID: 15699103.

57. Goumans MJ, Valdimarsdottir G, Itoh S, Lebrin F, Larsson J, Mummery C, Karlsson S, ten Dijke P. Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. Mol Cell. 2003;12(4):817-28. doi: 10.1016/s1097-2765(03)00386-1. PubMed PMID: 14580334.

58. Goumans MJ, Ten Dijke P. TGF-beta Signaling in Control of Cardiovascular Function. Cold Spring Harb Perspect Biol. 2018;10(2). Epub 2017/03/30. doi: 10.1101/cshperspect.a022210. PubMed PMID: 28348036; PMCID: PMC5793760.

59. Goumans MJ, Zwijsen A, Ten Dijke P, Bailly S. Bone Morphogenetic Proteins in Vascular Homeostasis and Disease. Cold Spring Harb Perspect Biol. 2018;10(2). Epub 2017/03/30. doi: 10.1101/cshperspect.a031989. PubMed PMID: 28348038; PMCID: PMC5793761.

60. Grafe I, Alexander S, Peterson JR, Snider TN, Levi B, Lee B, Mishina Y. TGF-beta Family Signaling in Mesenchymal Differentiation. Cold Spring Harb Perspect Biol. 2018;10(5). Epub 2017/05/17. doi: 10.1101/cshperspect.a022202. PubMed PMID: 28507020; PMCID: PMC5932590.

61. Grainger JR, Smith KA, Hewitson JP, McSorley HJ, Harcus Y, Filbey KJ, Finney CA, Greenwood EJ, Knox DP, Wilson MS, Belkaid Y, Rudensky AY, Maizels RM. Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway. J Exp Med. 2010;207(11):2331-41. Epub 2010/09/30. doi: 10.1084/jem.20101074. PubMed PMID: 20876311; PMCID: PMC2964568.

62. Greenwald J, Groppe J, Gray P, Wiater E, Kwiatkowski W, Vale W, Choe S. The BMP7/ActRII Extracellular Domain Complex Provides New Insights into the Cooperative Nature of Receptor Assembly. Molecular Cell. 2003;11(3):605-17. doi: 10.1016/S1097-2765(03)00094-7.

63. Gregory RD. On the interpretation of host-parasite ecology:Heligmosomoides polygyrus(Nematoda) in wild wood mouse (Apodemus sylvaticus) populations. Journal of Zoology. 1992;226(1):109-21. doi: 10.1111/j.1469-7998.1992.tb06130.x.

64. Groppe J, Hinck CS, Samavarchi-Tehrani P, Zubieta C, Schuermann JP, Taylor AB, Schwarz PM, Wrana JL, Hinck AP. Cooperative assembly of TGF-beta superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. Mol Cell. 2008;29(2):157-68. Epub 2008/02/05. doi: 10.1016/j.molcel.2007.11.039. PubMed PMID: 18243111.

65. Grzesiek S, Bax A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. Journal of Magnetic Resonance (1969). 1992;96(2):432-40. doi: 10.1016/0022-2364(92)90099-s.

66. Grzesiek S, Bax A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. Journal of Magnetic Resonance (1969). 1992;99(1):201-7. doi: 10.1016/0022-2364(92)90169-8.

67. Grzesiek S, Bax A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. Journal of the American Chemical Society. 1992;114(16):6291-3. doi: 10.1021/ja00042a003.

68. Grzesiek S, Anglister J, Bax A. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. Journal of Magnetic Resonance, Series B. 1993;101(1):114-9. doi: 10.1006/jmrb.1993.1019.

69. Grzesiek S, Bax A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. J Am Chem Soc. 1993;115(26):12593-4.

70. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. J Biomol NMR. 1993;3(2):185-204. Epub 1993/03/01. doi: 10.1007/BF00178261. PubMed PMID: 8477186.

71. Hang L, Blum AM, Setiawan T, Urban JP, Jr., Stoyanoff KM, Weinstock JV. Heligmosomoides polygyrus bakeri infection activates colonic Foxp3+ T cells enhancing their capacity to prevent colitis. J Immunol. 2013;191(4):1927-34. Epub 2013/07/16. doi: 10.4049/jimmunol.1201457. PubMed PMID: 23851695; PMCID: PMC3790665.

72. Hansen MR, 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. 2000;17 Suppl 1:365-9. Epub 2000/01/01. doi: 10.1080/07391102.2000.10506642. PubMed PMID: 22607445.

73. Harnett W. Secretory products of helminth parasites as immunomodulators. Molecular and Biochemical Parasitology. 2014;195(2):130-6. doi: 10.1016/j.molbiopara.2014.03.007.

74. Harris DC, Bertolucci MD. Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy. 1st ed: Dover Publications; 1989.

75. Hart PJ, Deep S, Taylor AB, Shu Z, Hinck CS, Hinck AP. Crystal structure of the human TbetaR2 ectodomain--TGF-beta3 complex. Nat Struct Biol. 2002;9(3):203-8. Epub 2002/02/19. doi: 10.1038/nsb766. PubMed PMID: 11850637.

76. Hartmann S, Schnoeller C, Dahten A, Avagyan A, Rausch S, Lendner M, Bocian C, Pillai S, Loddenkemper C, Lucius R, Worm M, Hamelmann E. Gastrointestinal nematode infection interferes with experimental allergic airway inflammation but not atopic

dermatitis. Clin Exp Allergy. 2009;39(10):1585-96. Epub 2009/06/11. doi: 10.1111/j.1365-2222.2009.03290.x. PubMed PMID: 19508324.

77. Hata A, Chen YG. TGF-beta Signaling from Receptors to Smads. Cold Spring Harb Perspect Biol. 2016;8(9). Epub 2016/07/28. doi: 10.1101/cshperspect.a022061. PubMed PMID: 27449815; PMCID: PMC5008074.

78. Hatta T, Konishi H, Katoh E, Natsume T, Ueno N, Kobayashi Y, Yamazaki T. Identification of the ligand-binding site of the BMP type IA receptor for BMP-4. Peptide Science. 2000;55(5):399-406. doi: 10.1002/1097-0282(2000)55:5<399::AID-BIP1014>3.0.CO;2-9.

79. Heldin CH, Moustakas A. Signaling Receptors for TGF-beta Family Members. Cold Spring Harb Perspect Biol. 2016;8(8). Epub 2016/08/03. doi: 10.1101/cshperspect.a022053. PubMed PMID: 27481709; PMCID: PMC4968163.

80. Hernández-Rodríguez M, Rosales-Hernández CM, Mendieta-Wejebe EJ, Martínez-Archundia M, Basurto CJ. Current Tools and Methods in Molecular Dynamics (MD) Simulations for Drug Design. Current Medicinal Chemistry. 2016;23(34):3909-24. doi: 10.2174/0929867323666160530144742.

81. Herrera B, Inman GJ. A rapid and sensitive bioassay for the simultaneous measurement of multiple bone morphogenetic proteins. Identification and quantification of BMP4, BMP6 and BMP9 in bovine and human serum. BMC Cell Biology. 2009;10(1):20. doi: 10.1186/1471-2121-10-20.

82. Hewitson JP, Grainger JR, Maizels RM. Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. Mol Biochem Parasitol. 2009;167(1):1-11. Epub 2009/05/02. doi: 10.1016/j.molbiopara.2009.04.008. PubMed PMID: 19406170; PMCID: PMC2706953.

83. Hewitson JP, Harcus Y, Murray J, van Agtmaal M, Filbey KJ, Grainger JR, Bridgett S, Blaxter ML, Ashton PD, Ashford DA, Curwen RS, Wilson RA, Dowle AA, Maizels RM. Proteomic analysis of secretory products from the model gastrointestinal nematode Heligmosomoides polygyrus reveals dominance of venom allergen-like (VAL) proteins. J Proteomics. 2011;74(9):1573-94. Epub 2011/07/05. doi: 10.1016/j.jprot.2011.06.002. PubMed PMID: 21722761; PMCID: PMC4794625.

84. Hewitson JP, Ivens AC, Harcus Y, Filbey KJ, McSorley HJ, Murray J, Bridgett S, Ashford D, Dowle AA, Maizels RM. Secretion of protective antigens by tissue-stage nematode larvae revealed by proteomic analysis and vaccination-induced sterile immunity. PLoS Pathog. 2013;9(8):e1003492. Epub 2013/08/24. doi: 10.1371/journal.ppat.1003492. PubMed PMID: 23966853; PMCID: PMC3744408.

85. Higman VA. Protein NMR: A Practical Guide 2019 [updated 14 Aug 2019; cited 16 Jun 2022]. Available from: https://www.protein-nmr.org.uk/.

86. Hinck AP, Archer SJ, Qian SW, Roberts AB, Sporn MB, Weatherbee JA, Tsang ML, Lucas R, Zhang BL, Wenker J, Torchia DA. Transforming growth factor beta 1: three-dimensional structure in solution and comparison with the X-ray structure of transforming growth factor beta 2. Biochemistry. 1996;35(26):8517-34. Epub 1996/07/02. doi: 10.1021/bi9604946. PubMed PMID: 8679613.

87. Hinck AP, Walker KP, 3rd, Martin NR, Deep S, Hinck CS, Freedberg DI. Sequential resonance assignments of the extracellular ligand binding domain of the human TGF-beta type II receptor. J Biomol NMR. 2000;18(4):369-70. Epub 2001/02/24. doi: 10.1023/a:1026775321886. PubMed PMID: 11200535.

88. Hinck AP. Structural studies of the TGF-betas and their receptors - insights into evolution of the TGF-beta superfamily. FEBS Lett. 2012;586(14):1860-70. Epub 2012/06/02. doi: 10.1016/j.febslet.2012.05.028. PubMed PMID: 22651914.

89. Hinck AP, Huang T. TGF-beta antagonists: same knot, but different hold. Structure. 2013;21(8):1269-70. Epub 2013/08/13. doi: 10.1016/j.str.2013.07.015. PubMed PMID: 23931139; PMCID: PMC3809007.

90. Hinck AP, Mueller TD, Springer TA. Structural Biology and Evolution of the TGF-beta Family. Cold Spring Harb Perspect Biol. 2016;8(12). Epub 2016/09/18. doi: 10.1101/cshperspect.a022103. PubMed PMID: 27638177; PMCID: PMC5131774.

91. Hotez P. Hookworm and poverty. Ann N Y Acad Sci. 2008;1136:38-44. Epub 2007/10/24. doi: 10.1196/annals.1425.000. PubMed PMID: 17954674.

92. Hotez PJ, Molyneux DH. Tropical anemia: one of Africa's great killers and a rationale for linking malaria and neglected tropical disease control to achieve a common goal. PLoS Negl Trop Dis. 2008;2(7):e270. Epub 20080730. doi: 10.1371/journal.pntd.0000270. PubMed PMID: 18665256; PMCID: PMC2474697.

93. Hu HH, Chen DQ, Wang YN, Feng YL, Cao G, Vaziri ND, Zhao YY. New insights into TGF-beta/Smad signaling in tissue fibrosis. Chem Biol Interact. 2018;292:76-83. Epub 20180711. doi: 10.1016/j.cbi.2018.07.008. PubMed PMID: 30017632.

94. Huang T, David L, Mendoza V, Yang Y, Villarreal M, De K, Sun L, Fang X, Lopez-Casillas F, Wrana JL, Hinck AP. TGF-beta signalling is mediated by two autonomously functioning TbetaRI:TbetaRII pairs. Embo j. 2011;30(7):1263-76. Epub 2011/03/23. doi: 10.1038/emboj.2011.54. PubMed PMID: 21423151; PMCID: PMC3094126.

95. Ihara S, Hirata Y, Koike K. TGF-beta in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota. J Gastroenterol. 2017;52(7):777-87. Epub 20170522. doi: 10.1007/s00535-017-1350-1. PubMed PMID: 28534191.

96. Ishii Y, Markus MA, Tycko R. Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. J Biomol NMR. 2001;21(2):141-51. Epub 2001/12/01. doi: 10.1023/a:1012417721455. PubMed PMID: 11727977.

97. Israel DI, Nove J, Kerns KM, Kaufman RJ, Rosen V, Cox KA, Wozney JM. Heterodimeric Bone Morphogenetic Proteins Show Enhanced Activity In Vitro and In Vivo. Growth Factors. 1996;13(3-4):291-300. doi: 10.3109/08977199609003229.

98. Jakubek RS, Handen J, White SE, Asher SA, Lednev IK. Ultraviolet resonance Raman spectroscopic markers for protein structure and dynamics. TrAC Trends in Analytical Chemistry. 2018;103:223-9. doi: 10.1016/j.trac.2017.12.002.

99. Jakubek RS, White SE, Asher SA. UV Resonance Raman Structural Characterization of an (In)soluble Polyglutamine Peptide. J Phys Chem B. 2019;123(8):1749-63. Epub 2019/02/06. doi: 10.1021/acs.jpcb.8b10783. PubMed PMID: 30717595.

100. Jakubek RS, Workman RJ, White SE, Asher SA. Polyglutamine Solution-State Structural Propensity Is Repeat Length Dependent. J Phys Chem B. 2019;123(19):4193-203. Epub 2019/04/23. doi: 10.1021/acs.jpcb.9b01433. PubMed PMID: 31008597.

101. Johnston CJ, Robertson E, Harcus Y, Grainger JR, Coakley G, Smyth DJ, McSorley HJ, Maizels R. Cultivation of Heligmosomoides polygyrus: an immunomodulatory nematode parasite and its secreted products. J Vis Exp. 2015(98):e52412. Epub 2015/04/14. doi: 10.3791/52412. PubMed PMID: 25867600; PMCID: PMC4401400.

102. Johnston CJ, Smyth DJ, Dresser DW, Maizels RM. TGF-beta in tolerance, development and regulation of immunity. Cell Immunol. 2016;299:14-22. Epub 2015/12/01. doi: 10.1016/j.cellimm.2015.10.006. PubMed PMID: 26617281; PMCID: PMC4711336.

103. Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, Kildemoes AO, Le Bihan T, Soares DC, Anderton SM, Brenn T, Wigmore SJ, Woodcock HV, Chambers RC, Hinck AP, McSorley HJ, Maizels RM. A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells. Nat Commun. 2017;8(1):1741. Epub 2017/11/25. doi: 10.1038/s41467-017-01886-6. PubMed PMID: 29170498; PMCID: PMC5701006.

104. Joosten RP, Long F, Murshudov GN, Perrakis A. ThePDB_REDOserver for macromolecular structure model optimization. IUCrJ. 2014;1(4):213-20. doi: 10.1107/s2052252514009324.

105. Josefowicz SZ, Lu LF, Rudensky AY. Regulatory T cells: mechanisms of differentiation and function. Annu Rev Immunol. 2012;30:531-64. Epub 2012/01/10. doi: 10.1146/annurev.immunol.25.022106.141623. PubMed PMID: 22224781; PMCID: PMC6066374.

106. Kaartinen V, Voncken JW, Shuler C, Warburton D, Bu D, Heisterkamp N, Groffen J. Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. Nat Genet. 1995;11(4):415-21. doi: 10.1038/ng1295-415. PubMed PMID: 7493022.

107. Kabsch W. XDS. Acta Crystallographica Section D. 2010;66(2):125-32. doi: 10.1107/S0907444909047337.

108. Kahata K, Dadras MS, Moustakas A. TGF-beta Family Signaling in Epithelial Differentiation and Epithelial-Mesenchymal Transition. Cold Spring Harb Perspect Biol. 2018;10(1). Epub 2017/03/02. doi: 10.1101/cshperspect.a022194. PubMed PMID: 28246184; PMCID: PMC5749157.

109. Kahata K, Maturi V, Moustakas A. TGF-beta Family Signaling in Ductal Differentiation and Branching Morphogenesis. Cold Spring Harb Perspect Biol. 2018;10(3). Epub 2017/03/16. doi: 10.1101/cshperspect.a031997. PubMed PMID: 28289061; PMCID: PMC5830900.

110. Kaplan FS, Fiori J, De La Peña LS, Ahn J, Billings PC, Shore EM. Dysregulation of the BMP-4 Signaling Pathway in Fibrodysplasia Ossificans Progressiva. Annals of the New York Academy of Sciences. 2006;1068(1):54-65. doi: 10.1196/annals.1346.008.

111. Kay L, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. Journal of the American Chemical Society. 1992;114(26):10663-5. doi: 10.1021/ja00052a088.

112. Kay LE, Ikura M, Tschudin R, Bax A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. Journal of Magnetic Resonance (1969). 1990;89(3):496-514. doi: 10.1016/0022-2364(90)90333-5.

113. Kay LE, Xu GY, Singer AU, Muhandiram DR, Formankay JD. A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain 1H and 13C Correlations in H2O Samples of Proteins. Journal of Magnetic Resonance, Series B. 1993;101(3):333-7. doi: 10.1006/jmrb.1993.1053.

114. Kay LE, Xu GY, Yamazaki T. Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H2O Saturation. Journal of Magnetic Resonance, Series A. 1994;109(1):129-33. doi: 10.1006/jmra.1994.1145.

115. Keeler J. Understanding NMR Spectroscopy. 2nd ed: Wiley; 2011.

116. Keller S, Vargas C, Zhao H, Piszczek G, Brautigam CA, Schuck P. High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. Analytical Chemistry. 2012;84(11):5066-73. doi: 10.1021/ac3007522.

117. Kelly A, Houston SA, Sherwood E, Casulli J, Travis MA. Regulation of Innate and Adaptive Immunity by TGFbeta. Adv Immunol. 2017;134:137-233. Epub 20170210. doi: 10.1016/bs.ai.2017.01.001. PubMed PMID: 28413021.

118. Kim KK, Sheppard D, Chapman HA. TGF-beta1 Signaling and Tissue Fibrosis. Cold Spring Harb Perspect Biol. 2018;10(4). Epub 2017/04/23. doi: 10.1101/cshperspect.a022293. PubMed PMID: 28432134.

119. Kim SK, Barron L, Hinck CS, Petrunak EM, Cano KE, Thangirala A, Iskra B, Brothers M, Vonberg M, Leal B, Richter B, Kodali R, Taylor AB, Du S, Barnes CO, Sulea T, Calero G, Hart PJ, Hart MJ, Demeler B, Hinck AP. An engineered transforming growth factor

beta (TGF-beta) monomer that functions as a dominant negative to block TGF-beta signaling. J Biol Chem. 2017;292(17):7173-88. Epub 2017/02/24. doi: 10.1074/jbc.M116.768754. PubMed PMID: 28228478; PMCID: PMC5409485.

120. Kim SK, Henen MA, Hinck AP. Structural biology of betaglycan and endoglin, membrane-bound co-receptors of the TGF-beta family. Exp Biol Med (Maywood). 2019;244(17):1547-58. Epub 2019/10/12. doi: 10.1177/1535370219881160. PubMed PMID: 31601110.

121. Kirsch T, Nickel J, Sebald W. BMP-2 antagonists emerge from alterations in the low-affinity binding epitope for receptor BMPR-II. The EMBO Journal. 2000;19(13):3314-24. doi: 10.1093/emboj/19.13.3314.

122. Kirsch T, Sebald W, Dreyer MK. Crystal structure of the BMP-2–BRIA ectodomain complex. Nature Structural Biology. 2000;7(6):492-6. doi: 10.1038/75903.

123. Kitagaki K, Businga TR, Racila D, Elliott DE, Weinstock JV, Kline JN. Intestinal helminths protect in a murine model of asthma. J Immunol. 2006;177(3):1628-35. Epub 2006/07/20. doi: 10.4049/jimmunol.177.3.1628. PubMed PMID: 16849471.

124. Koenig BB, Cook JS, Wolsing DH, Ting J, Tiesman JP, Correa PE, Olson CA, Pecquet AL, Ventura F, Grant RA, Chen G-X, Wrana JL, Massagué J, Rosenbaum JS. Characterization and cloning of a receptor for BMP-2 and BMP-4 from NIH 3T3 cells. Molecular and Cellular Biology. 1994;14(9):5961-74. doi: 10.1128/MCB.14.9.5961.

125. Kulkarni AB, Huh CG, Becker D, Geiser A, Lyght M, Flanders KC, Roberts AB, Sporn MB, Ward JM, Karlsson S. Transforming growth factor beta 1 null mutation in mice causes excessive inflammatory response and early death. Proceedings of the National Academy of Sciences. 1993;90(2):770-4. doi: 10.1073/pnas.90.2.770.

126. Laiho M, Weis FM, Boyd FT, Ignotz RA, Massagué J. Responsiveness to transforming growth factor-beta (TGF-beta) restored by genetic complementation between cells defective in TGF-beta receptors I and II. Journal of Biological Chemistry. 1991;266(14):9108-12. doi: 10.1016/s0021-9258(18)31558-8.

127. Lawler S, Feng XH, Chen RH, Maruoka EM, Turck CW, Griswold-Prenner I, Derynck R. The type II transforming growth factor-beta receptor autophosphorylates not only on serine and threonine but also on tyrosine residues. J Biol Chem. 1997;272(23):14850-9. Epub 1997/06/06. doi: 10.1074/jbc.272.23.14850. PubMed PMID: 9169454.

128. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2014;31(8):1325-7. doi: 10.1093/bioinformatics/btu830.

129. Lee W, Bahrami A, Dashti HT, Eghbalnia HR, Tonelli M, Westler WM, Markley JL. I-PINE web server: an integrative probabilistic NMR assignment system for proteins. J Biomol NMR. 2019;73(5):213-22. Epub 2019/06/06. doi: 10.1007/s10858-019-00255-3. PubMed PMID: 31165321; PMCID: PMC6579641.

130. Letsou A, Arora K, Wrana JL, Simin K, Twombly V, Jamal J, Staehling-Hampton K, Hoffmann FM, Gelbart WM, Massagué J, O'Connor MB. Drosophila Dpp signaling is mediated by the punt gene product: A dual ligand-binding type II receptor of the TGFβ receptor family. Cell. 1995;80(6):899-908. doi: 10.1016/0092-8674(95)90293-7.

131. Li MO, Sanjabi S, Flavell R. Transforming Growth Factor-β Controls Development, Homeostasis, and Tolerance of T Cells by Regulatory T Cell-Dependent and -Independent Mechanisms. Immunity. 2006;25(3):455-71. doi: 10.1016/j.immuni.2006.07.011.

132. Liu F, Ventura F, Doody J, Massagué J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. Molecular and Cellular Biology. 1995;15(7):3479-86. doi: 10.1128/MCB.15.7.3479.

133. Liu Q, Sundar K, Mishra PK, Mousavi G, Liu Z, Gaydo A, Alem F, Lagunoff D, Bleich D, Gause WC. Helminth infection can reduce insulitis and type 1 diabetes through CD25- and IL-10-independent mechanisms. Infect Immun. 2009;77(12):5347-58. Epub 2009/09/16. doi: 10.1128/IAI.01170-08. PubMed PMID: 19752032; PMCID: PMC2786463.

134. Lo RS, Chen Y-G, Shi Y, Pavletich NP, Massagué J. The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF-β receptors. The EMBO Journal. 1998;17(4):996-1005. doi: https://doi.org/10.1093/emboj/17.4.996.

135. Logan J, Navarro S, Loukas A, Giacomin P. Helminth-induced regulatory T cells and suppression of allergic responses. Current Opinion in Immunology. 2018;54:1-6. doi: 10.1016/j.coi.2018.05.007.

136. Logan TM, Olejniczak ET, Xu RX, Fesik SW. A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. J Biomol NMR. 1993;3(2):225-31. Epub 1993/03/01. doi: 10.1007/BF00178264. PubMed PMID: 8477187.

137. Lohr F, Ruterjans H. Alternative E.COSY techniques for the measurement of 3J(C (i) (') (-1),C (i) (beta)) and (3) J(H (i) (N) ,C (i) (beta)) coupling constants in proteins. J Biomol NMR. 1999;13(3):263-74. Epub 1999/03/01. doi: 10.1023/A:1008378719908. PubMed PMID: 20700819.

138. Loke P, Lim YA. Helminths and the microbiota: parts of the hygiene hypothesis. Parasite Immunol. 2015;37(6):314-23. Epub 2015/04/15. doi: 10.1111/pim.12193. PubMed PMID: 25869420; PMCID: PMC4428757.

139. López-Casillas F, Wrana JL, Massagué J. Betaglycan presents ligand to the TGF beta signaling receptor. Cell. 1993;73(7):1435-44. doi: 10.1016/0092-8674(93)90368-z. PubMed PMID: 8391934.

140. López-Casillas F, Wrana JL, Massagué J. Betaglycan presents ligand to the TGFβ signaling receptor. Cell. 1993;73(7):1435-44. doi: 10.1016/0092-8674(93)90368-Z.

141. Lyons BA, Montelione GT. An HCCNH Triple-Resonance Experiment Using Carbon-13 Isotropic Mixing for Correlating Backbone Amide and Side-Chain Aliphatic Resonances in Isotopically Enriched Proteins. Journal of Magnetic Resonance, Series B. 1993;101(2):206-9. doi: 10.1006/jmrb.1993.1034.

142. MacFarlane EG, Haupt J, Dietz HC, Shore EM. TGF-beta Family Signaling in Connective Tissue and Skeletal Diseases. Cold Spring Harb Perspect Biol. 2017;9(11). Epub 2017/03/02. doi: 10.1101/cshperspect.a022269. PubMed PMID: 28246187; PMCID: PMC5666637.

143. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;47(W1):W636-W41. Epub 2019/04/13. doi: 10.1093/nar/gkz268. PubMed PMID: 30976793; PMCID: PMC6602479.

144. Mahlawat P, Ilangovan U, Biswas T, Sun LZ, Hinck AP. Structure of the Alk1 extracellular domain and characterization of its bone morphogenetic protein (BMP) binding properties. Biochemistry. 2012;51(32):6328-41. Epub 2012/07/18. doi: 10.1021/bi300942x. PubMed PMID: 22799562; PMCID: PMC3448977.

145. Maizels RM, Lawrence RA. Immunological tolerance: The key feature in human filariasis? Parasitol Today. 1991;7(10):271-6. Epub 1991/01/01. doi: 10.1016/0169-4758(91)90093-4. PubMed PMID: 15463386.

146. Maizels RM, Balic A, Gomez-Escobar N, Nair M, Taylor MD, Allen JE. Helminth parasites--masters of regulation. Immunol Rev. 2004;201:89-116. Epub 2004/09/14. doi: 10.1111/j.0105-2896.2004.00191.x. PubMed PMID: 15361235.

147. Maizels RM, Smith KA. Regulatory T cells in infection. Adv Immunol. 2011;112:73-136. Epub 2011/11/29. doi: 10.1016/b978-0-12-387827-4.00003-6. PubMed PMID: 22118407.

148. Maizels RM, Hewitson JP, Murray J, Harcus YM, Dayer B, Filbey KJ, Grainger JR, McSorley HJ, Reynolds LA, Smith KA. Immune modulation and modulators in Heligmosomoides polygyrus infection. Exp Parasitol. 2012;132(1):76-89. Epub 2011/08/31. doi: 10.1016/j.exppara.2011.08.011. PubMed PMID: 21875581; PMCID: PMC6485391.

149. Maizels RM, Hewitson JP, Smith KA. Susceptibility and immunity to helminth parasites. Curr Opin Immunol. 2012;24(4):459-66. Epub 2012/07/17. doi: 10.1016/j.coi.2012.06.003. PubMed PMID: 22795966; PMCID: PMC3437973.

150. Maizels RM, McSorley HJ, Smyth DJ. Helminths in the hygiene hypothesis: sooner or later? Clin Exp Immunol. 2014;177(1):38-46. Epub 2014/04/23. doi: 10.1111/cei.12353. PubMed PMID: 24749722; PMCID: PMC4089153.

151. Maizels RM, McSorley HJ. Regulation of the host immune system by helminth parasites. J Allergy Clin Immunol. 2016;138(3):666-75. Epub 2016/08/02. doi: 10.1016/j.jaci.2016.07.007. PubMed PMID: 27476889; PMCID: PMC5010150.

152. Maizels RM, Smits HH, McSorley HJ. Modulation of Host Immunity by Helminths: The Expanding Repertoire of Parasite Effector Molecules. Immunity. 2018;49(5):801-18. Epub 2018/11/22. doi: 10.1016/j.immuni.2018.10.016. PubMed PMID: 30462997; PMCID: PMC6269126.

153. Maizels RM. Regulation of immunity and allergy by helminth parasites. Allergy. 2020;75(3):524-34. doi: 10.1111/all.13944.

154. Maizels RM. Chapter One - The multi-faceted roles of TGF-β in regulation of immunity to infection. In: Alt FW, Murphy KM, editors. Advances in Immunology: Academic Press; 2021. p. 1-42.

155. Male D, Peebles RS, Male V. Immunology 9th edition: Elsevier; 2020.

156. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel Iii EE, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, Guan Y, Chalouni C, Ziai J, Şenbabaoğlu Y, Santoro S, Sheinson D, Hung J, Giltnane JM, Pierce AA, Mesh K, Lianoglou S, Riegler J, Carano RAD, Eriksson P, Höglund M, Somarriba L, Halligan DL, Van Der Heijden MS, Loriot Y, Rosenberg JE, Fong L, Mellman I, Chen DS, Green M, Derleth C, Fine GD, Hegde PS, Bourgon R, Powles T. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018;554(7693):544-8. doi: 10.1038/nature25501.

157. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. Biochemistry. 1989;28(15):6150-6. Epub 1989/07/25. doi: 10.1021/bi00441a004. PubMed PMID: 2675964.

158. Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. Journal of the American Chemical Society. 2002;111(4):1515-7. doi: 10.1021/ja00186a066.

159. Markowitz SD, Roberts AB. Tumor suppressor activity of the TGF-β pathway in human cancers. Cytokine & Growth Factor Reviews. 1996;7(1):93-102. doi: 10.1016/1359-6101(96)00001-9.

160. Massague J, Wotton D. Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. 2000;19(8):1745-54. Epub 2000/04/25. doi: 10.1093/emboj/19.8.1745. PubMed PMID: 10775259; PMCID: PMC302010.

161. Massague J. TGFbeta in Cancer. Cell. 2008;134(2):215-30. doi: 10.1016/j.cell.2008.07.001. PubMed PMID: 18662538; PMCID: PMC3512574.

162. Massague J. A very private TGF-beta receptor embrace. Mol Cell. 2008;29(2):149-50. Epub 2008/02/05. doi: 10.1016/j.molcel.2008.01.006. PubMed PMID: 18243107.

163. Massagué J, Chen Y-G. Controlling TGF-β signaling. Genes & Development. 2000;14(6):627-44. doi: 10.1101/gad.14.6.627.

164. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. Journal of Applied Crystallography. 2007;40(4):658-74. doi: 10.1107/S0021889807021206.

165. McSorley HJ, O'Gorman MT, Blair N, Sutherland TE, Filbey KJ, Maizels RM. Suppression of type 2 immunity and allergic airway inflammation by secreted products of the helminth Heligmosomoides polygyrus. Eur J Immunol. 2012;42(10):2667-82. Epub 2012/06/19. doi: 10.1002/eji.201142161. PubMed PMID: 22706967; PMCID: PMC4916998.

166. McSorley HJ, Hewitson JP, Maizels RM. Immunomodulation by helminth parasites: defining mechanisms and mediators. Int J Parasitol. 2013;43(3-4):301-10. Epub 2013/01/08. doi: 10.1016/j.ijpara.2012.11.011. PubMed PMID: 23291463.

167. McSorley HJ, Blair NF, Smith KA, McKenzie AN, Maizels RM. Blockade of IL-33 release and suppression of type 2 innate lymphoid cell responses by helminth secreted products in airway allergy. Mucosal Immunol. 2014;7(5):1068-78. Epub 2014/02/06. doi: 10.1038/mi.2013.123. PubMed PMID: 24496315; PMCID: PMC4016792.

168. Meier S, Haussinger D, Grzesiek S. Charged acrylamide copolymer gels as media for weak alignment. J Biomol NMR. 2002;24(4):351-6. Epub 2003/01/11. doi: 10.1023/a:1021609207024. PubMed PMID: 12522299.

169. Mendoza V, Vilchis-Landeros MM, Mendoza-Hernandez G, Huang T, Villarreal MM, Hinck AP, Lopez-Casillas F, Montiel JL. Betaglycan has two independent domains required for high affinity TGF-beta binding: proteolytic cleavage separates the domains and inactivates the neutralizing activity of the soluble receptor. Biochemistry. 2009;48(49):11755-65. Epub 2009/10/22. doi: 10.1021/bi901528w. PubMed PMID: 19842711; PMCID: PMC2796082.

170. Meyers EA, Kessler JA. TGF-beta Family Signaling in Neural and Neuronal Differentiation, Development, and Function. Cold Spring Harb Perspect Biol. 2017;9(8). Epub 2017/01/29. doi: 10.1101/cshperspect.a022244. PubMed PMID: 28130363; PMCID: PMC5538418.

171. Mishra PK, Patel N, Wu W, Bleich D, Gause WC. Prevention of type 1 diabetes through infection with an intestinal nematode parasite requires IL-10 in the absence of a Th2-type response. Mucosal Immunol. 2013;6(2):297-308. Epub 2012/07/19. doi: 10.1038/mi.2012.71. PubMed PMID: 22806101.

172. Mittl PR, Priestle JP, Cox DA, McMaster G, Cerletti N, Grutter MG. The crystal structure of TGF-beta 3 and comparison to TGF-beta 2: implications for receptor binding. Protein Sci. 1996;5(7):1261-71. Epub 1996/07/01. doi: 10.1002/pro.5560050705. PubMed PMID: 8819159; PMCID: PMC2143453.

173. Monsivais D, Matzuk MM, Pangas SA. The TGF-beta Family in the Reproductive Tract. Cold Spring Harb Perspect Biol. 2017;9(10). Epub 2017/02/15. doi: 10.1101/cshperspect.a022251. PubMed PMID: 28193725; PMCID: PMC5554737.

174. Montelione GT, Lyons BA, Emerson SD, Tashiro M. An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. Journal of the American Chemical Society. 1992;114(27):10974-5. doi: 10.1021/ja00053a051.

175. Mori S, Abeygunawardana C, Johnson MO, van Zijl PC. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. J Magn Reson B. 1995;108(1):94-8. Epub 1995/07/01. doi: 10.1006/jmrb.1995.1109. PubMed PMID: 7627436.

176. 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. 2016;8(5). Epub 2016/05/04. doi: 10.1101/cshperspect.a021873. PubMed PMID: 27141051; PMCID: PMC4852809.

177. Moses HL. TGF-? regulation of epithelial cell proliferation. Molecular Reproduction and Development. 1992;32(2):179-84. doi: 10.1002/mrd.1080320215.

178. Muhandiram DR, Kay LE. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. Journal of Magnetic Resonance, Series B. 1994;103(3):203-16. doi: 10.1006/jmrb.1994.1032.

179. Mukundan A, Byeon CH, Hinck CS, Cunningham K, Campion T, Smyth DJ, Maizels RM, Hinck AP. 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. 2022:101994. Epub 20220429. doi: 10.1016/j.jbc.2022.101994. PubMed PMID: 35500648.

180. Mullen AC, Wrana JL. TGF-beta Family Signaling in Embryonic and Somatic Stem-Cell Renewal and Differentiation. Cold Spring Harb Perspect Biol. 2017;9(7). Epub 2017/01/22. doi: 10.1101/cshperspect.a022186. PubMed PMID: 28108485; PMCID: PMC5495062.

181. Mulvenna J, Hamilton B, Nagaraj SH, Smyth D, Loukas A, Gorman JJ. Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, Ancylostoma caninum. Mol Cell Proteomics. 2009;8(1):109-21. Epub 2008/08/30. doi: 10.1074/mcp.M800206-MCP200. PubMed PMID: 18753127.

182. Myers RH. Huntington's disease genetics. NeuroRx. 2004;1(2):255-62. doi: 10.1602/neurorx.1.2.255. PubMed PMID: 15717026; PMCID: PMC534940.

183. Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, Urade Y, Goto Y, Naiki H, Toda T. A toxic monomeric conformer of the polyglutamine protein. Nat Struct Mol Biol. 2007;14(4):332-40. Epub 20070318. doi: 10.1038/nsmb1215. PubMed PMID: 17369839.

184. Naka K, Hirao A. Regulation of Hematopoiesis and Hematological Disease by TGF-beta Family Signaling Molecules. Cold Spring Harb Perspect Biol. 2017;9(9). Epub 2017/02/15. doi: 10.1101/cshperspect.a027987. PubMed PMID: 28193723; PMCID: PMC5585852.

185. Namwanje M, Brown CW. Activins and Inhibins: Roles in Development, Physiology, and Disease. Cold Spring Harb Perspect Biol. 2016;8(7). Epub 2016/06/23. doi: 10.1101/cshperspect.a021881. PubMed PMID: 27328872; PMCID: PMC4930927.

186. Natsume T, Tomita S, Iemura S, Kinto N, Yamaguchi A, Ueno N. Interaction between Soluble Type I Receptor for Bone Morphogenetic Protein and Bone Morphogenetic Protein-4*. Journal of Biological Chemistry. 1997;272(17):11535-40. doi: 10.1074/jbc.272.17.11535.

187. Nickel J, Kotzsch A, Sebald W, Mueller TD. A single residue of GDF-5 defines binding specificity to BMP receptor IB. J Mol Biol. 2005;349(5):933-47. Epub 2005/05/14. doi: 10.1016/j.jmb.2005.04.015. PubMed PMID: 15890363.

188. Nisbet AJ, Smith SK, Armstrong S, Meikle LI, Wildblood LA, Beynon RJ, Matthews JB. Teladorsagia circumcincta: activation-associated secreted proteins in excretory/secretory products of fourth stage larvae are targets of early IgA responses in infected sheep. Exp Parasitol. 2010;125(4):329-37. Epub 2010/03/09. doi: 10.1016/j.exppara.2010.02.014. PubMed PMID: 20206168.

189. Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, Rosenbaum JS. Identification of a Human Type II Receptor for Bone Morphogenetic Protein-4 That Forms Differential Heteromeric Complexes with Bone Morphogenetic Protein Type I Receptors (∗). Journal of Biological Chemistry. 1995;270(38):22522-6. doi: 10.1074/jbc.270.38.22522.

190. Oladepo SA, Xiong K, Hong Z, Asher SA, Handen J, Lednev IK. UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. Chemical Reviews. 2012;112(5):2604-28. doi: 10.1021/cr200198a.

191. Olejniczak ET, Xu RX, Fesik SW. A 4D HCCH-TOCSY experiment for assigning the side chain 1H and 13C resonances of proteins. J Biomol NMR. 1992;2(6):655-9. Epub 1992/11/01. doi: 10.1007/BF02192854. PubMed PMID: 1283353.

192. Orr HT, Zoghbi HY. Trinucleotide repeat disorders. Annu Rev Neurosci. 2007;30:575-621. doi: 10.1146/annurev.neuro.29.051605.113042. PubMed PMID: 17417937.

193. Ottesen EA, Weller PF, Heck L. Specific cellular immune unresponsiveness in human filariasis. Immunology. 1977;33(3):413-21. Epub 1977/09/01. PubMed PMID: 332621; PMCID: PMC1445652.

194. Ottiger M, Delaglio F, Bax A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson. 1998;131(2):373-8. Epub 1998/05/08. doi: 10.1006/jmre.1998.1361. PubMed PMID: 9571116.

195. Palmer AG, Cavanagh J, Wright PE, Rance M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. Journal of Magnetic Resonance (1969). 1991;93(1):151-70. doi: https://doi.org/10.1016/0022-2364(91)90036-S.

196. Parapia LA, Jackson C. Ehlers-Danlos syndrome – a historical review. British Journal of Haematology. 2008;141(1):32-5. doi: https://doi.org/10.1111/j.1365-2141.2008.06994.x.

197. Pascal SM, Muhandiram DR, Yamazaki T, Formankay JD, Kay LE. Simultaneous Acquisition of 15N- and 13C-Edited NOE Spectra of Proteins Dissolved in H2O. Journal of Magnetic Resonance, Series B. 1994;103(2):197-201. doi: 10.1006/jmrb.1994.1031.

198. Penney JB, Jr., Vonsattel JP, MacDonald ME, Gusella JF, Myers RH. CAG repeat number governs the development rate of pathology in Huntington's disease. Ann Neurol. 1997;41(5):689-92. doi: 10.1002/ana.410410521. PubMed PMID: 9153534.

199. Penton A, Chen Y, Staehling-Hampton K, Wrana JL, Attisano L, Szidonya J, Cassill JA, Massagué J, Hoffmann FM. Identification of two bone morphogenetic protein type I receptors in Drosophila and evidence that Brk25D is a decapentaplegic receptor. Cell. 1994;78(2):239-50. doi: 10.1016/0092-8674(94)90294-1.

200. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. UCSF ChimeraX : Structure visualization for researchers, educators, and developers. Protein Science. 2021;30(1):70-82. doi: 10.1002/pro.3943.

201. Pickup MW, Owens P, Moses HL. TGF-beta, Bone Morphogenetic Protein, and Activin Signaling and the Tumor Microenvironment. Cold Spring Harb Perspect Biol. 2017;9(5). Epub 2017/01/08. doi: 10.1101/cshperspect.a022285. PubMed PMID: 28062564; PMCID: PMC5411701.

202. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR. 1992;2(6):661-5. Epub 1992/11/01. doi: 10.1007/BF02192855. PubMed PMID: 1490109.

203. Pitto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR. 1992;2:661-5.

204. Pritchard DI, Lawrence CE, Appleby P, Gibb IA, Glover K. Immunosuppressive proteins secreted by the gastrointestinal nematode parasite Heligmosomoides polygyrus. International Journal for Parasitology. 1994;24(4):495-500. doi: 10.1016/0020-7519(94)90140-6.

205. Proetzel G, Pawlowski SA, Wiles MV, Yin M, Boivin GP, Howles PN, Ding J, Ferguson MWJ, Doetschman T. Transforming growth factor–β3 is required for secondary palate fusion. Nature Genetics. 1995;11(4):409-14. doi: 10.1038/ng1295-409.

206. Pullan RL, Smith JL, Jasrasaria R, Brooker SJ. Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. Parasit Vectors. 2014;7:37. Epub

2014/01/23. doi: 10.1186/1756-3305-7-37. PubMed PMID: 24447578; PMCID: PMC3905661.

207. Punihaole D, Hong Z, Jakubek RS, Dahlburg EM, Geib S, Asher SA. Glutamine and Asparagine Side Chain Hyperconjugation-Induced Structurally Sensitive Vibrations. J Phys Chem B. 2015;119(41):13039-51. Epub 20150930. doi: 10.1021/acs.jpcb.5b07651. PubMed PMID: 26392216; PMCID: PMC5065012.

208. Punihaole D, Workman RJ, Hong Z, Madura JD, Asher SA. Polyglutamine Fibrils: New Insights into Antiparallel beta-Sheet Conformational Preference and Side Chain Structure. J Phys Chem B. 2016;120(12):3012-26. Epub 20160318. doi: 10.1021/acs.jpcb.5b11380. PubMed PMID: 26947327.

209. Punihaole D, Jakubek RS, Workman RJ, Marbella LE, Campbell P, Madura JD, Asher SA. Monomeric Polyglutamine Structures That Evolve into Fibrils. J Phys Chem B. 2017;121(24):5953-67. Epub 20170608. doi: 10.1021/acs.jpcb.7b04060. PubMed PMID: 28531354.

210. Qian SW, Burmester JK, Tsang ML, Weatherbee JA, Hinck AP, Ohlsen DJ, Sporn MB, Roberts AB. Binding affinity of transforming growth factor-beta for its type II receptor is determined by the C-terminal region of the molecule. J Biol Chem. 1996;271(48):30656-62. Epub 1996/11/29. doi: 10.1074/jbc.271.48.30656. PubMed PMID: 8940041.

211. Radaev S, Zou Z, Huang T, Lafer EM, Hinck AP, Sun PD. Ternary complex of transforming growth factor-beta1 reveals isoform-specific ligand recognition and receptor recruitment in the superfamily. J Biol Chem. 2010;285(19):14806-14. Epub 2010/03/09. doi: 10.1074/jbc.M109.079921. PubMed PMID: 20207738; PMCID: PMC2863181.

212. Ramachandran A, Vizán P, Das D, Chakravarty P, Vogt J, Rogers KW, Müller P, Hinck AP, Sapkota GP, Hill CS. TGF-β uses a novel mode of receptor activation to phosphorylate SMAD1/5 and induce epithelial-to-mesenchymal transition. Elife. 2018;7. Epub 20180129. doi: 10.7554/eLife.31756. PubMed PMID: 29376829; PMCID: PMC5832415.

213. Rausch S, Huehn J, Kirchhoff D, Rzepecka J, Schnoeller C, Pillai S, Loddenkemper C, Scheffold A, Hamann A, Lucius R, Hartmann S. Functional Analysis of Effector and Regulatory T Cells in a Parasitic Nematode Infection. Infection and Immunity. 2008;76(5):1908-19. doi: 10.1128/IAI.01233-07.

214. Read S, Malmstrom V, Powrie F. Cytotoxic T lymphocyte-associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. J Exp Med. 2000;192(2):295-302. Epub 2000/07/19. doi: 10.1084/jem.192.2.295. PubMed PMID: 10899916; PMCID: PMC2193261.

215. Reid KBM, Day AJ. Structure-function relationships of the complement components. Immunology Today. 1989;10(6):177-80. doi: 10.1016/0167-5699(89)90317-4.

216. Reynolds LA, Filbey KJ, Maizels RM. Immunity to the model intestinal helminth parasite Heligmosomoides polygyrus. Semin Immunopathol. 2012;34(6):829-46. Epub

2012/10/12. doi: 10.1007/s00281-012-0347-3. PubMed PMID: 23053394; PMCID: PMC3496515.

217. Ripoche J, Day AJ, Harris TJ, Sim RB. The complete amino acid sequence of human complement factor H. Biochem J. 1988;249(2):593-602. Epub 1988/01/15. doi: 10.1042/bj2490593. PubMed PMID: 2963625; PMCID: PMC1148743.

218. Roberts AB, Sporn MB, Assoian RK, Smith JM, Roche NS, Wakefield LM, Heine UI, Liotta LA, Falanga V, Kehrl JH. Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. Proceedings of the National Academy of Sciences. 1986;83(12):4167-71. doi: 10.1073/pnas.83.12.4167.

219. Roberts AB. Transforming growth factor-β: activity and efficacy in animal models of wound healing. Wound Repair and Regeneration. 1995;3(4):408-18. doi: 10.1046/j.1524-475X.1995.30405.x.

220. Rosenheck K, Doty P. The far ultraviolet absorption spectra of polypeptide and protein solutions and their dependence on conformation. Proceedings of the National Academy of Sciences of the United States of America. 1961;47(11):1775.

221. Ross CA, Poirier MA. What is the role of protein aggregation in neurodegeneration? Nature reviews Molecular cell biology. 2005;6(11):891-8.

222. Rubtsov YP, Rudensky AY. TGFβ signalling in control of T-cell-mediated self-reactivity. Nature Reviews Immunology. 2007;7(6):443-53. doi: 10.1038/nri2095.

223. Rule GS, Hitchens TK. Fundamentals of Protein NMR Spectroscopy. 1st ed. Pittsburgh, PA: Springer; 2006.

224. Ryan SM, Eichenberger RM, Ruscher R, Giacomin PR, Loukas A. Harnessing helminth-driven immunoregulation in the search for novel therapeutic modalities. PLoS Pathog. 2020;16(5):e1008508. Epub 2020/05/15. doi: 10.1371/journal.ppat.1008508. PubMed PMID: 32407385.

225. Sakaguchi S. Naturally arising CD4+ regulatory t cells for immunologic self-tolerance and negative control of immune responses. Annu Rev Immunol. 2004;22:531-62. Epub 2004/03/23. doi: 10.1146/annurev.immunol.21.120601.141122. PubMed PMID: 15032588.

226. Sakaguchi S, Mikami N, Wing JB, Tanaka A, Ichiyama K, Ohkura N. Regulatory T Cells and Human Disease. Annu Rev Immunol. 2020;38:541-66. Epub 2020/02/06. doi: 10.1146/annurev-immunol-042718-041717. PubMed PMID: 32017635.

227. Sanford LP, Ormsby I, Gittenberger-de Groot AC, Sariola H, Friedman R, Boivin GP, Cardell EL, Doetschman T. TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. Development.

1997;124(13):2659-70. doi: 10.1242/dev.124.13.2659. PubMed PMID: 9217007; PMCID: PMC3850286.

228. Sanford LP, Ormsby I, Gittenberger-De Groot AC, Sariola H, Friedman R, Boivin GP, Cardell EL, Doetschman T. TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. Development. 1997;124(13):2659-70. doi: 10.1242/dev.124.13.2659.

229. Sanjabi S, Oh SA, Li MO. Regulation of the Immune Response by TGF-beta: From Conception to Autoimmunity and Infection. Cold Spring Harb Perspect Biol. 2017;9(6). Epub 2017/01/22. doi: 10.1101/cshperspect.a022236. PubMed PMID: 28108486; PMCID: PMC5453394.

230. Sass H-J, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. Journal of Biomolecular NMR. 2000;18(4):303-9. doi: 10.1023/A:1026703605147.

231. Saunders KA, Raine T, Cooke A, Lawrence CE. Inhibition of autoimmune type 1 diabetes by gastrointestinal helminth infection. Infect Immun. 2007;75(1):397-407. Epub 2006/10/18. doi: 10.1128/IAI.00664-06. PubMed PMID: 17043101; PMCID: PMC1828378.

232. Schaub B, Lauener R, von Mutius E. The many faces of the hygiene hypothesis. J Allergy Clin Immunol. 2006;117(5):969-77; quiz 78. Epub 2006/05/06. doi: 10.1016/j.jaci.2006.03.003. PubMed PMID: 16675321.

233. Scheufler C, Sebald W, Hulsmeyer M. Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution. J Mol Biol. 1999;287(1):103-15. Epub 1999/03/13. doi: 10.1006/jmbi.1999.2590. PubMed PMID: 10074410.

234. Schleucher J, Sattler M, Griesinger C. Coherence Selection by Gradients without Signal Attenuation: Application to the Three-Dimensional HNCO Experiment. Angewandte Chemie International Edition in English. 1993;32(10):1489-91. doi: 10.1002/anie.199314891.

235. Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletzky O, Glaser SJ, Sørensen OW, Griesinger C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. Journal of Biomolecular NMR. 1994;4(2):301-6. doi: 10.1007/BF00175254.

236. Schulz GE, Schirmer RH. Principles of protein structure: Springer Science & Business Media; 2013.

237. Sebald W, Nickel J, Zhang JL, Mueller TD. Molecular recognition in bone morphogenetic protein (BMP)/receptor interaction. Biol Chem. 2004;385(8):697-710. doi: 10.1515/BC.2004.086. PubMed PMID: 15449706.

238. Seoane J, Gomis RR. TGF-beta Family Signaling in Tumor Suppression and Cancer Progression. Cold Spring Harb Perspect Biol. 2017;9(12). Epub 2017/03/02. doi: 10.1101/cshperspect.a022277. PubMed PMID: 28246180; PMCID: PMC5710110.

239. Shepherd C, Navarro S, Wangchuk P, Wilson D, Daly NL, Loukas A. Identifying the immunomodulatory components of helminths. Parasite Immunology. 2015;37(6):293-303. doi: 10.1111/pim.12192.

240. Sheppard D. Transforming Growth Factor : A Central Modulator of Pulmonary and Airway Inflammation and Fibrosis. Proceedings of the American Thoracic Society. 2006;3(5):413-7. doi: 10.1513/pats.200601-008aw.

241. Shevach EM. Regulatory T cells in autoimmmunity. Annu Rev Immunol. 2000;18:423-49. Epub 2000/06/03. doi: 10.1146/annurev.immunol.18.1.423. PubMed PMID: 10837065.

242. Shull MM, Ormsby I, Kier AB, Pawlowski S, Diebold RJ, Yin M, Allen R, Sidman C, Proetzel G, Calvin D, et al. Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. Nature. 1992;359(6397):693-9. doi: 10.1038/359693a0. PubMed PMID: 1436033; PMCID: PMC3889166.

243. Sigurskjold BW. Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry. Analytical Biochemistry. 2000;277(2):260-6. doi: 10.1006/abio.1999.4402.

244. Sklenar V, Piotto M, Leppik R, Saudek V. Gradient-Tailored Water Suppression for 1H-15N HSQC Experiments Optimized to Retain Full Sensitivity. Journal of Magnetic Resonance, Series A. 1993;102(2):241-5. doi: 10.1006/jmra.1993.1098.

245. Smith KA, Filbey KJ, Reynolds LA, Hewitson JP, Harcus Y, Boon L, Sparwasser T, Hammerling G, Maizels RM. Low-level regulatory T-cell activity is essential for functional type-2 effector immunity to expel gastrointestinal helminths. Mucosal Immunol. 2016;9(2):428-43. Epub 2015/08/20. doi: 10.1038/mi.2015.73. PubMed PMID: 26286232; PMCID: PMC4677460.

246. Smyth DJ, Harcus Y, White MPJ, Gregory WF, Nahler J, Stephens I, Toke-Bjolgerud E, Hewitson JP, Ivens A, McSorley HJ, Maizels RM. TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus. Int J Parasitol. 2018;48(5):379-85. Epub 2018/03/07. doi: 10.1016/j.ijpara.2017.12.004. PubMed PMID: 29510118; PMCID: PMC5904571.

247. Smyth DJ, Ren B, White MPJ, McManus C, Webster H, Shek V, Evans C, Pandhal J, Fields F, Maizels RM, Mayfield S. Oral delivery of a functional algal-expressed TGF-β mimic halts colitis in a murine DSS model. Journal of Biotechnology. 2021;340:1-12. doi: 10.1016/j.jbiotec.2021.08.006.

248. Strandberg B, Dickerson RE, Rossmann MG. 50 years of protein structure analysis. Journal of Molecular Biology. 2009;392(1):2-32.

249. Sutton TL, Zhao A, Madden KB, Elfrey JE, Tuft BA, Sullivan CA, Urban JF, Jr., Shea-Donohue T. Anti-Inflammatory mechanisms of enteric Heligmosomoides polygyrus infection against trinitrobenzene sulfonic acid-induced colitis in a murine model. Infect Immun. 2008;76(10):4772-82. Epub 2008/07/23. doi: 10.1128/IAI.00744-07. PubMed PMID: 18644879; PMCID: PMC2546858.

250. Tauriello DVF, Sancho E, Batlle E. Overcoming TGFβ-mediated immune evasion in cancer. Nature Reviews Cancer. 2022;22(1):25-44. doi: 10.1038/s41568-021-00413-6.

251. Taylor MD, LeGoff L, Harris A, Malone E, Allen JE, Maizels RM. Removal of regulatory T cell activity reverses hyporesponsiveness and leads to filarial parasite clearance in vivo. J Immunol. 2005;174(8):4924-33. Epub 2005/04/09. doi: 10.4049/jimmunol.174.8.4924. PubMed PMID: 15814720.

252. Taylor MD, van der Werf N, Maizels RM. T cells in helminth infection: the regulators and the regulated. Trends Immunol. 2012;33(4):181-9. Epub 2012/03/09. doi: 10.1016/j.it.2012.01.001. PubMed PMID: 22398370.

253. Tesseur I, Zou K, Berber E, Zhang H, Wyss-Coray T. Highly sensitive and specific bioassay for measuring bioactive TGF-beta. BMC Cell Biol. 2006;7:15. Epub 2006/03/22. doi: 10.1186/1471-2121-7-15. PubMed PMID: 16549026; PMCID: PMC1479809.

254. Thies RS, Chen T, Davies MV, Tomkinson KN, Pearson AA, Shakey QA, Wolfman NM. GDF-8 propeptide binds to GDF-8 and antagonizes biological activity by inhibiting GDF-8 receptor binding. Growth Factors. 2001;18(4):251-9. Epub 2001/08/25. doi: 10.3109/08977190109029114. PubMed PMID: 11519824.

255. Tolman JR, Ruan K. NMR Residual Dipolar Couplings as Probes of Biomolecular Dynamics. Chemical Reviews. 2006;106(5):1720-36. doi: 10.1021/cr040429z.

256. Tycko R, Blanco FJ, Ishii Y. Alignment of Biopolymers in Strained Gels: A New Way To Create Detectable Dipole−Dipole Couplings in High-Resolution Biomolecular NMR. Journal of the American Chemical Society. 2000;122(38):9340-1. doi: 10.1021/ja002133q.

257. Upper D. The unsuccessful self-treatment of a case of "writer's block". J Appl Behav Anal. 1974;7(3):497. Epub 1974/10/01. doi: 10.1901/jaba.1974.7-497a. PubMed PMID: 16795475; PMCID: PMC1311997.

258. Villarreal MM, Kim SK, Barron L, Kodali R, Baardsnes J, Hinck CS, Krzysiak TC, Henen MA, Pakhomova O, Mendoza V, O'Connor-McCourt MD, Lafer EM, Lopez-Casillas F, Hinck AP. Binding Properties of the Transforming Growth Factor-beta Coreceptor Betaglycan: Proposed Mechanism for Potentiation of Receptor Complex Assembly and Signaling. Biochemistry. 2016;55(49):6880-96. Epub 2016/12/14. doi: 10.1021/acs.biochem.6b00566. PubMed PMID: 27951653; PMCID: PMC5551644.

259. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: Development of a

software pipeline. Proteins: Structure, Function, and Bioinformatics. 2005;59(4):687-96. doi: https://doi.org/10.1002/prot.20449.

260. Vuister GW, Bax A. Resolution Enhancement and Spectral Editing of Uniformly 13C-Enriched Proteins by Homonuclear Broadband 13C Decoupling. J Magn Reson. 1992;98:428-35.

261. Vuister GW, Bax A. Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins. Journal of the American Chemical Society. 1993;115(17):7772-7. doi: 10.1021/ja00070a024.

262. Vuister GW, Bax A. Measurement of four-bond HN-H alpha J-couplings in staphylococcal nuclease. J Biomol NMR. 1994;4(2):193-200. Epub 1994/03/01. doi: 10.1007/BF00175247. PubMed PMID: 8019134.

263. Wang Y-X, Marquardt JL, Wingfield P, Stahl SJ, Lee-Huang S, Torchia D, Bax A. Simultaneous Measurement of 1H− 15N, 1H− 13C ‘, and 15N− 13C ‘Dipolar Couplings in a Perdeuterated 30 kDa Protein Dissolved in a Dilute Liquid Crystalline Phase. Journal of the American Chemical Society. 1998;120(29):7385-6.

264. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189-91. Epub 2009/01/20. doi: 10.1093/bioinformatics/btp033. PubMed PMID: 19151095; PMCID: PMC2672624.

265. Weber D, Kotzsch A, Nickel J, Harth S, Seher A, Mueller U, Sebald W, Mueller TD. A silent H-bond can be mutationally activated for high-affinity interaction of BMP-2 and activin type IIB receptor. BMC Structural Biology. 2007;7(1):6. doi: 10.1186/1472-6807-7-6.

266. White MP, Smyth DJ, Cook L, Ziegler SF, Levings MK, Maizels RM. The parasite cytokine mimic Hp -TGM potently replicates the regulatory effects of TGF-β on murine CD4 + T cells. Immunology & Cell Biology. 2021. doi: 10.1111/imcb.12479.

267. White MPJ, McManus CM, Maizels RM. Regulatory T-cells in helminth infection: induction, function and therapeutic potential. Immunology. 2020. Epub 2020/03/11. doi: 10.1111/imm.13190. PubMed PMID: 32153025.

268. White SE, Schoenherr C, Hinck CS, Cunningham K, Campion T, Schwartze T, White MPJ, Inman G, Maizels RM, Hinck AP. The H. polygyrus TGF-β Mimic TGM6 Functions as a Competitive TGF-β Signaling Inhibitor in Murine Fibroblasts. In preparation. 2022.

269. Wiedemann M, Voehringer D. Immunomodulation and Immune Escape Strategies of Gastrointestinal Helminths and Schistosomes. Front Immunol. 2020;11:572865. Epub 20200917. doi: 10.3389/fimmu.2020.572865. PubMed PMID: 33042153; PMCID: PMC7527441.

270. Wilson MS, Taylor MD, Balic A, Finney CA, Lamb JR, Maizels RM. Suppression of allergic airway inflammation by helminth-induced regulatory T cells. J Exp Med. 2005;202(9):1199-212. Epub 2005/11/09. doi: 10.1084/jem.20042572. PubMed PMID: 16275759; PMCID: PMC2213237.

271. Wittekind M, Mueller L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. Journal of Magnetic Resonance, Series B. 1993;101(2):201-5. doi: 10.1006/jmrb.1993.1033.

272. Wotton D, Massagué J. Smad Transcriptional Corepressors in TGFβ Family Signaling. In: Privalsky ML, editor. Transcriptional Corepressors: Mediators of Eukaryotic Gene Repression. Berlin, Heidelberg: Springer Berlin Heidelberg; 2001. p. 145-64.

273. Wrana JL, Attisano L, Cárcamo J, Zentella A, Doody J, Laiho M, Wang X-F, Massague J. TGFβ signals through a heteromeric protein kinase receptor complex. Cell. 1992;71(6):1003-14. doi: 10.1016/0092-8674(92)90395-S.

274. Wrana JL, Attisano L, Wieser R, Ventura F, Massague J. Mechanism of activation of the TGF-beta receptor. Nature. 1994;370(6488):341-7. Epub 1994/08/04. doi: 10.1038/370341a0. PubMed PMID: 8047140.

275. Wrana JL, Attisano L, Wieser R, Ventura F, Massagué J. Mechanism of activation of the TGF-β receptor. Nature. 1994;370(6488):341-7. doi: 10.1038/370341a0.

276. Wynn TA, Thompson RW, Cheever AW, Mentink-Kane MM. Immunopathogenesis of schistosomiasis. Immunol Rev. 2004;201:156-67. Epub 2004/09/14. doi: 10.1111/j.0105-2896.2004.00176.x. PubMed PMID: 15361239.

277. Xiong K, Punihaole D, Asher SA. UV resonance Raman spectroscopy monitors polyglutamine backbone and side chain hydrogen bonding and fibrillization. Biochemistry. 2012;51(29):5822-30. Epub 20120712. doi: 10.1021/bi300551b. PubMed PMID: 22746095; PMCID: PMC3415266.

278. Yamashita H, Ten Dijke P, Huylebroeck D, Sampath TK, Andries M, Smith JC, Heldin CH, Miyazono K. Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects. Journal of Cell Biology. 1995;130(1):217-26. doi: 10.1083/jcb.130.1.217.

279. Yamazaki T, Forman-Kay JD, Kay LE. Two-dimensional NMR experiments for correlating carbon-13.beta. and proton.delta./.epsilon. chemical shifts of aromatic residues in 13C-labeled proteins via scalar couplings. Journal of the American Chemical Society. 1993;115(23):11054-5. doi: 10.1021/ja00076a099.

280. Yang W, Dunlap JR, Andrews RB, Wetzel R. Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. Hum Mol Genet. 2002;11(23):2905-17. doi: 10.1093/hmg/11.23.2905. PubMed PMID: 12393802.

281. Yang Y-A, Dukhanina O, Tang B, Mamura M, Letterio JJ, Macgregor J, Patel SC, Khozin S, Liu Z-Y, Green J, Anver MR, Merlino G, Wakefield LM. Lifetime exposure to a soluble TGF-β antagonist protects mice against metastasis without adverse side effects2002;109(12):1607-15. doi: 10.1172/jci15333.

282. Yao L, Ying J, Bax A. Improved accuracy of 15N-1H scalar and residual dipolar couplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins. J Biomol NMR. 2009;43(3):161-70. Epub 2009/02/12. doi: 10.1007/s10858-009-9299-x. PubMed PMID: 19205898; PMCID: PMC2753394.

283. Yatsuda AP, Krijgsveld J, Cornelissen AW, Heck AJ, de Vries E. Comprehensive analysis of the secreted proteins of the parasite Haemonchus contortus reveals extensive sequence variation and differential immune recognition. J Biol Chem. 2003;278(19):16941-51. Epub 2003/02/11. doi: 10.1074/jbc.M212453200. PubMed PMID: 12576473.

284. Zeng J, Boyles J, Tripathy C, Wang L, Yan A, Zhou P, Donald BR. High-resolution protein structure determination starting with a global fold calculated from exact solutions to the RDC equations. Journal of Biomolecular NMR. 2009;45(3):265-81. doi: 10.1007/s10858-009-9366-3.

285. Zhang O, Kay LE, Olivier JP, Forman-Kay JD. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. Journal of Biomolecular NMR. 1994;4(6):845-58. doi: 10.1007/BF00398413.

286. Zhang Y, Alexander PB, Wang XF. TGF-beta Family Signaling in the Control of Cell Proliferation and Survival. Cold Spring Harb Perspect Biol. 2017;9(4). Epub 2016/12/07. doi: 10.1101/cshperspect.a022145. PubMed PMID: 27920038; PMCID: PMC5378054.

287. Zhao H, Piszczek G, Schuck P. SEDPHAT – A platform for global ITC analysis and global multi-method analysis of molecular interactions. Methods. 2015;76:137-48. doi: 10.1016/j.ymeth.2014.11.012.

288. Zinski J, Tajer B, Mullins MC. TGF-beta Family Signaling in Early Vertebrate Development. Cold Spring Harb Perspect Biol. 2018;10(6). Epub 2017/06/11. doi: 10.1101/cshperspect.a033274. PubMed PMID: 28600394.

289. Zuiderweg ER, Fesik SW. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. Biochemistry. 1989;28(6):2387-91. Epub 1989/03/21. doi: 10.1021/bi00432a008. PubMed PMID: 2730871.

290. Zuniga JE, Ilangovan U, Mahlawat P, Hinck CS, Huang T, Groppe JC, McEwen DG, Hinck AP. The TbetaR-I pre-helix extension is structurally ordered in the unbound form and its flanking prolines are essential for binding. J Mol Biol. 2011;412(4):601-18. Epub 2011/08/09. doi: 10.1016/j.jmb.2011.07.046. PubMed PMID: 21821041; PMCID: PMC3576881.

291. Zweckstetter M, Bax A. Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR. Journal of the American Chemical Society. 2000;122(15):3791-2. doi: 10.1021/ja0000908.

292. Zweckstetter M, Bax A. Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage. J Biomol NMR. 2001;20(4):365-77. Epub 2001/09/21. doi: 10.1023/a:1011263920003. PubMed PMID: 11563559.