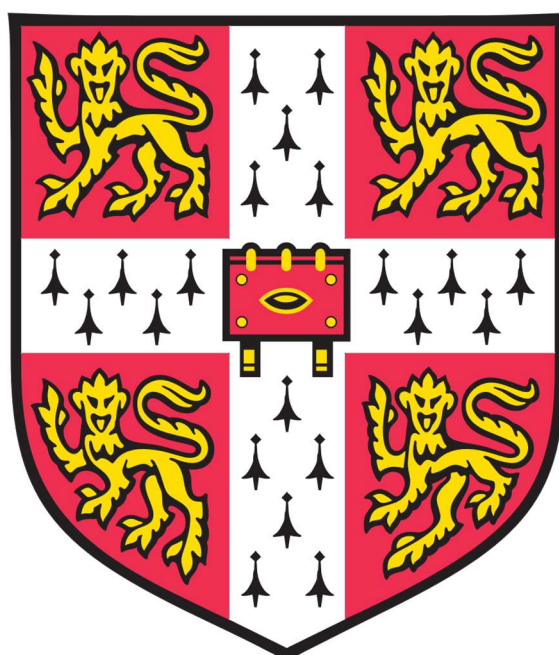


# Developing Chemical Tools for the Inhibition of Activin A Signalling

Joseph Lewis McLoughlin

Supervisor: Dr. Marko Hyvönen



Homerton College

University of Cambridge

This Dissertation is Submitted for the Degree of Doctor of Philosophy

August 2019



## Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.



# Summary

Name: Joseph Lewis McLoughlin

Title: Developing Chemical Tools for the Inhibition of Activin A Signalling

Activin A is a member of the TGF- $\beta$  superfamily, a family of structurally related growth factors that have been implicated in embryogenesis, homeostasis, and cancer. All members consist of pro- and mature domains which are post translationally cleaved. Upon signalling the pro-domain dissociates from the mature growth factor. Signalling in this family requires a type I and a type II receptor. Though there are over thirty growth factors, there are only seven type I and five type II receptors resulting in promiscuity between ligands and receptors. In order to fully elucidate the roles of these proteins and develop therapeutics, specific inhibitors must be generated.

In this thesis, I targeted activin A to achieve specificity. I first targeted mature activin A through the use of XChem, conducting a fragment screen against mature activin A crystals. I describe the identification of twelve fragment hits that bind in the putative type I receptor binding site and the validation and optimisation of one of these fragments.

I then used the structure of the pro-mature complex of activin A to design and generate a series of GB1- fused peptides based on the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of the pro-domain. I determined the  $\alpha$ 1-helix-loop epitope to be key for interactions with the mature domain, before investigating the effect of truncation and mutation on this epitope. I then determined the potency of the fusion GB1-H<sub>1</sub>LH<sub>2</sub> as an inhibitor of activin A signalling ( $IC_{50} = 4.4 \mu M$ ). I further optimised GB1-H<sub>1</sub>LH<sub>2</sub> through dimerization resulting in a large increase in inhibitor potency ( $IC_{50} = 71.5 nM$ ). Lastly, I screened this dimeric inhibitor against several TGF- $\beta$  family members to determine it specifically inhibits both activin A and activin B.

To conclude, I discuss the findings of this thesis before considering how this methodology could be applied to other TGF- $\beta$  family members.



## Acknowledgements

Well here we go...

I would like to start by thanking Marko. You took me on when I was labless and three and a half years later, here we are. I would like to thank you for all the support and scientific guidance you have provided and still continue to provide. You have been a fantastic supervisor and mentor and I could not have done any of this without you. I will be forever grateful.

To Hyvönen Group, past and present. Tom, Xuelu, Matt, Theo, Natalija, Beata (and Michal), Katharina, Luana, Monique, Paul, Aleksei, Marghe, Andrew, Katy, Marcelo, Theresa, Clare, and Gerhard. Your helpful discussions, technical support, and friendship have all contributed to this thesis and made my time in the lab enjoyable. I would also like to thank the Biophysics and Crystallography facilities in the Department of Biochemistry, as well as Graham for keeping my computers functional, and Clova and Tim for the sustenance.

To our collaborators in Chemistry. Sam, Till, Sarah, and Hannah. Thank you for synthesising many of the compounds and peptides used throughout this thesis.

To the BBSRC, thank you for funding my scientific adventure and GSK for allowing me to undertake an internship there. I would also like to thank the XChem facility at the Diamond Light Source for making a large part of this work possible.

I would like to thank all my friends in Cambridge over the years, at Homerton and beyond. Aaron, Thomas, Ari, Michael, Hugo, Fryer, Lucie, Ruth, DTP people, Camilla, Alice, Jack, Em, Amy, Holly, Harry, Vicki, Guy, Ezra, Jack, Maddie, Josh, Dom, Leyla, Flo, Victoria, Nathan, Jordan and many others who I'm sure I've forgotten. Thank you for making my Cambridge experience and for all the fun and laughter. And Lewis, thank you for keeping me company at lunch for four years and indulging my moaning. I hope you one day end up chairing a "Tea Room Improvement Committee".

To Geoff, Melanie, and all the Homerton Staff. You have made Homerton a truly wonderful environment to live and work in. I truly believe we are the best college and it's in no small part down to you. Homerton till I die!

Hannah, this degree allowed me to meet the most incredible person I've ever met – you. Thank you for always being you and loving me every day. I love you more than I could ever express.

Finally, Mum, Dad and Dan. Thank you for getting me through this. It certainly wouldn't have been possible without you.





# Contents

Declaration.....	ii
Summary.....	iv
Acknowledgements.....	vi
Contents.....	1
List of Figures.....	4
List of Tables.....	6
Abbreviations.....	7
Introduction.....	9
1.1 Cell Signalling.....	10
1.2 The TGF- $\beta$ Superfamily and Activin A.....	10
1.3 The Cellular Signalling of the TGF- $\beta$ Superfamily.....	11
1.4 The Structural Features of the TGF- $\beta$ Superfamily and Activin A.....	22
1.5 The Pathology and Physiology of Activin A.....	26
1.6 Inhibitors of TGF- $\beta$ Family Signalling.....	32
1.7 Aims.....	37
Small Molecule Approach.....	41
2.1 Introduction.....	42
2.2 Optimization of Crystallisation conditions for Fragment Screening of Mature Activin A.....	44
2.3 XChem Screening of Mature Activin A.....	48
2.4 Validation of Binding of NM466, HLS327, SG5B8, and SG6B10 to Mature Activin A via X-ray Crystallography.....	59
2.5 Screening of Derivatives of NM466 against Activin A.....	61
2.6 Conclusions.....	63
Peptide Approach.....	65
3.1 Introduction.....	66
3.2 Design of Peptide Fusions.....	67
3.3 Cloning of Peptide Fusions.....	70

3.4	Expression and Purification of Peptide Fusions .....	71
3.5	Identification of the Key Binding Motifs in the Activin A Pro-domain.....	74
3.6	Effect of Chemical Stapling on Binding of Synthetic Peptides to Mature Activin A (In Conjunction with Sam Rowe).....	81
3.7	Determination of the Minimal Binding Peptide of the Activin A Pro-domain.....	84
3.8	Mutagenesis of the Minimal Binding Peptide of the Activin A Pro-domain.....	88
3.9	GB1-H <sub>1</sub> LH <sub>2</sub> and GB1-H <sub>1</sub> L Inhibit Activin A Signalling with Micromolar Efficacy .....	90
3.10	Conclusions .....	92
	Dimeric Approach .....	95
4.1	Introduction .....	96
4.2	Design of PKADD Domain Fused Peptides .....	97
4.3	Cloning of PKADD Domain Fused Peptides.....	98
4.4	Expression and Purification of PKADD Domain Fused Peptides .....	99
4.5	PKADD Domain Fused Peptides are Dimers in Solution .....	100
4.6	PKADD Domain Fused Peptides are More Effective Inhibitors than the Monomeric Peptide Fusions .....	101
4.7	Continuous Fused Helix Peptides Behave in a Similar Manner to GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD .....	105
4.8	Design of Covalently Linked Dimeric Peptide Fusions .....	108
4.9	Introduction of a Disulfide Bridge Results in an Increase in Inhibitor Efficacy .....	112
4.10	Screening of GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD <sub>SS</sub> against Other TGF- $\beta$ Superfamily Members .....	113
4.11	Conclusions .....	117
	Discussion.....	119
5.1	Discussion.....	120
5.2	The Putative Type I Receptor Binding Site as a Target for Small Molecule Inhibition.....	121
5.3	Pro-domain Derived Peptides as Inhibitors of Activin A and the Wider TGF- $\beta$ Family.....	123
5.4	Future Directions of Peptide Inhibitors of Activin A.....	126
5.5	Potential Applications of TGF- $\beta$ Family Pro-domain Derived Peptides .....	128
	Material and Methods .....	131

References .....	139
Appendices.....	165
8.1 Appendix 1: Crystallographic Statistics from Data Collection and Refinement.....	166
8.2 Appendix 2: Expression Construct Protein Sequences .....	168
8.3 Appendix 3: PCR Oligonucleotide Sequences .....	172
8.4 Appendix 4: ITC Isotherms and Thermograms for the Interaction between the Synthetic Peptides / Mutant Minimal Peptide Fusions with Activin A.....	176

## List of Figures

Figure 1: The cellular signalling of activin A.....	13
Figure 2: Topology of pro-TGF- $\beta$ family ligands.....	24
Figure 3: Pro- structures of the TGF- $\beta$ family. ....	25
Figure 4: Common inhibitors TGF- $\beta$ family signalling. ....	34
Figure 5: Binding hotspots on the surface of mature activin A. ....	44
Figure 6: The crystal structure determined for mature activin A grown in the presence of DMSO. ...	46
Figure 7: Crystallisation of mature activin A in the presence of DMSO.....	47
Figure 8: XChem screen of mature activin A.....	57
Figure 9: Alignment of mature activin A and receptor bound structures. ....	58
Figure 10: The “PEGless” structure of mature activin A. ....	59
Figure 11: Binding of NM466 to mature activin A. ....	60
Figure 12: The generic structure of a NM466 derivative.....	61
Figure 13: Binding of TR17 to mature activin A. ....	63
Figure 14: Pro-domain derived peptides of activin A. ....	70
Figure 15: The pOP4BP vector. ....	71
Figure 16: Expression tests of pro-domain derived peptide fusions. ....	72
Figure 17: Purification of the peptide fusions via Ni-affinity and StrepTactin affinity purification.....	74
Figure 18: Identification of the key binding motifs of the activin A pro-domain via BLI. ....	75
Figure 19: $K_d$ determination of GB1-H <sub>1</sub> LH <sub>2</sub> and GB1-H <sub>1</sub> L for mature activin A via BLI and ITC.....	78
Figure 20: Stapling of synthetic activin A derived peptides.....	82
Figure 21: Determination of the minimal binding peptide of the activin A pro-domain via BLI. ....	85
Figure 22: Determination of the $K_d$ for the minimal binding peptide.....	86
Figure 23: End-of- $\alpha$ 1-helix arginine in TGF- $\beta$ family members.....	88
Figure 24: Selected residues for mutation in the activin A minimal peptide. ....	89
Figure 25: Luciferase Assays of GB1, GB1-H <sub>1</sub> LH <sub>2</sub> , and GB1-H <sub>1</sub> L. ....	92
Figure 26: PKADD domain fused peptides. ....	98
Figure 27: The pOP6BPA vector.....	99
Figure 28: Purification of PKADD domain fused peptides. ....	100
Figure 29: SEC-MALS analysis of the PKADD domain fused peptides.....	101
Figure 30: The binding and inhibitory profiles of the PKADD fused peptides. ....	103
Figure 31: Interhelical distance of H <sub>1</sub> LH <sub>2</sub> . ....	105
Figure 32: Binding and inhibitory profiles of the continuous fused helix peptides.....	107

Figure 33: Selected site for disulfide formation in the PKADD domain.....	109
Figure 34: The purification of GB1-H <sub>1</sub> LH <sub>2</sub> -SS.....	110
Figure 35: Cu (ii) catalysed linkage of GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD <sub>SS</sub> . ....	111
Figure 36: Inhibitory profiles of covalently dimerized peptide fusions. ....	113
Figure 37: Alignment of the $\alpha$ 1-helix-loop- $\alpha$ 2-helix region of activin A, human activin B, xenopus activin B, myostatin, and BMP2. ....	115
Figure 38: Specificity screen of GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD <sub>SS</sub> . ....	116
Figure 39: The putative type I receptor binding site as a target for inhibitors.....	122
Figure 40: Comparison of the activin A and myostatin minimal binding peptides.....	125
Figure 41: Potential optimisations of H <sub>1</sub> LH <sub>2</sub> .....	128

## List of Tables

Table 1: ALK2 mutations associated with fibrodysplasia ossificans progressiva .....	31
Table 2: Selected inhibitors of the TGF- $\beta$ family.....	33
Table 3: Fragment hits identified in XChem screen.....	50
Table 4: Screened derivatives of NM466.....	62
Table 5: Pro-domain derived peptides of activin A.....	70
Table 6: Molecular weights of peptide fusions as determined via mass spectrometry .....	74
Table 7: Dissociation constants determined for the interaction between mature activin A and peptide fusions. ....	80
Table 8: Synthetic peptides screened against mature activin A (synthesised by Sam Rowe).....	83
Table 9: N- and C-terminal truncations of H <sub>1</sub> L.....	84
Table 10: Dissociation constant determined for mature activin A and the minimal binding peptide. 86	
Table 11: Dissociation Constants determined for mature activin A and mutant minimal binding peptides. ....	89
Table 12: IC <sub>50</sub> values determined for GB1-H <sub>1</sub> LH <sub>2</sub> and GB1-H <sub>1</sub> L inhibiting activin A signalling. ....	92
Table 13: SEC-MALS analysis of PKADD fused peptides.....	101
Table 14: Dissociation constants and IC <sub>50</sub> values for PKADD domain fused GB1-H <sub>1</sub> LH <sub>2</sub> and GB1-H <sub>1</sub> L with mature activin A.....	102
Table 15: Dissociation constants and IC <sub>50</sub> values obtained for the Fused Helix peptides with mature activin A.....	106
Table 16: Dissociation constants and IC <sub>50</sub> values obtained for the covalently dimerized peptide fusions with mature activin A. ....	112
Table 17: Specificity screen of the TGF- $\beta$ family.....	116
Table 18: Crystallographic statistics.....	166
Table 19: Primer combinations.....	175

## Abbreviations

ActA.....	Activin A
ActB.....	Activin B
ActC.....	Activin C
ActE.....	Activin E
ActRIIA.....	Activin receptor type II A
ActRIIB.....	Activin receptor type II B
ALK2/4.....	Activin-like kinase 2 / 4
AMH.....	Anti-Müllerian hormone
AMHRII.....	Anti-Müllerian hormone receptor type II
BAMBI.....	BMP and activin membrane-bound inhibitor homolog
BLI.....	Biolayer interferometry
BMP.....	Bone morphogenic protein
BMPRII.....	Bone morphogenic protein receptor type II
EC <sub>50</sub> .....	Half maximal effective concentration
ECD.....	Extra-cellular domain
ED <sub>50</sub> .....	Half maximal effective dose
EDTA.....	Ethylenediaminetetraacetic acid
EGF.....	Epidermal growth factor
Erk.....	Extracellular signal regulated kinase
FGF.....	Fibroblast growth factor
FOP.....	Fibrodysplasia Ossificans Progressiva
Fs.....	Follistatin domain
FS.....	Follistatin
FSTL3.....	Follistatin-related protein 3
FSH.....	Follicle stimulating hormone
GB1.....	B1 domain of Streptococcal protein G
GDF.....	Growth and differentiation factor
GST.....	Glutathione-S-transferase
H <sub>1</sub> LH <sub>2</sub> β.....	α1-helix-loop-α2-helix-β-hairpin fragment of the activin A pro-domain
HEK293T.....	Human embryonic kidney cells expressing the SV40 large T antigen

IC<sub>50</sub> ..... Half maximal inhibitory concentration  
IPTG ..... Isopropyl β-D-1-thiogalactopyranoside  
ITC ..... Isothermal calorimetry  
K<sub>d</sub> ..... Equilibrium dissociation constant  
K<sub>i</sub>..... Inhibition constant  
Mstn ..... Myostatin  
PanDDA ..... Pan-Dataset Density Analysis  
PKA ..... Protein kinase A  
PKADD ..... Protein kinase A dimerization and docking  
PKA RIIα ..... Protein kinase A regulatory subunit II α  
SAR ..... Structure activity relationship  
SEC-MALS ..... Size exclusion chromatography – multiangle light scattering  
SPR ..... Surface plasmon resonance  
TβRII ..... TGF-β receptor type II  
TGF-β ..... Transforming growth factor-β  
VEGF..... Vascular endothelial growth factor



---

# Introduction

---

## 1.1 Cell Signalling

The term “cell signalling” can be defined as the transmission of information either within or between cells. It is often initiated by a stimulus (the information), either internal or external to the cell, and results in a biological response. The process in which the information is transmitted, from detection of the stimulus to the response, is through a series of mechanical or biomolecular events that include interactions between proteins, enzymatic processing of substrate, chemical modification of a biomolecule, or interactions between proteins and DNA. Cell signalling can be categorised into five broad categories: intracrine, juxtacrine, autocrine, paracrine, and endocrine.

Intracrine refers to signalling that is initiated in and remains within a single cell. Initiation is commonly caused by one of the other categories of signalling through modulation of membrane receptors or concentration changes of molecules of interest in the cell. Conversely, juxtacrine signalling occurs between two immediately proximal cells and often requires direct contact between cell surface proteins/ glycosides or the use of cell junctions. Autocrine, paracrine and endocrine signalling all require secretion of a signalling molecule that then induces signalling either in the cell from which it was secreted (autocrine); a cell in the local environment of the secreting cell (paracrine); or a cell distant to the site of secretion, using the circulatory system to travel there (endocrine). As cellular signalling directs and determines all aspects of a living system, the interplay between these signalling methods is thus responsible for underpinning the complexity of life.

Cell growth, proliferation, differentiation and apoptosis are four important biological responses that are controlled by cell signalling. Misregulation of these responses in metazoans is the classical phenotype of cancer, thus making tight regulation vital. Growth factors are a group of signalling molecules that regulate these four processes. They can be cytokines or hormones and are usually secreted from the cell, acting in autocrine, paracrine, or endocrine manner. Important families include the fibroblast growth factor (FGF) family, the epidermal growth factor (EGF) family, the vascular endothelial growth factor (VEGF) family, and the transforming growth factor beta (TGF- $\beta$ ) family. In this thesis, I will be looking at the TGF- $\beta$  family, specifically focussing on one of its members – activin A.

## 1.2 The TGF- $\beta$ Superfamily and Activin A

The TGF- $\beta$  superfamily is a family of extracellular growth factors that play vital roles in an organism’s development and homeostasis, especially in embryogenesis, growth, and repair mechanisms.<sup>1</sup> Highly conserved across metazoans,<sup>2</sup> mature TGF- $\beta$  growth factors can signal in both autocrine and paracrine

manners, through the formation of heteromeric complexes with TGF- $\beta$  family receptors on the cell surface. Formation of the growth factor-receptor complex then triggers a phosphorylation cascade that can produce a diverse range of cellular responses. In humans over 30 members of the family have been identified, including the eponymous proteins, bone morphogenic proteins (BMPs), growth differentiation factors (GDFs), anti-müllerian hormone, nodal, and the activins.<sup>3</sup> These proteins all have high structural and sequence similarity with the majority being disulfide linked dimers and consisting of an “active” signalling mature domain and an “inactive” pro-domain (mature domain pairwise identity >30% across the whole family).<sup>4</sup> These domains are post translationally cleaved but with the majority of growth factors, remain in association until signalling where the pro-domain dissociates allowing the mature domain to interact with cell surface receptors. In contrast to the number of growth factors there are only 12 identified TGF- $\beta$  receptors in humans – seven type I and five type II.<sup>5</sup> Signalling requires formation of a heteromeric ligand-type I-type II complex and thus this results in high ligand promiscuity with multiple growth factors signalling through the same receptors (sometimes in different combinations) but inducing different cellular responses.

One well studied member of the superfamily is activin A. Originally isolated by Vale *et al.* in 1986 from porcine ovarian follicular fluid, it was initially identified as having the direct opposite effect to related protein inhibin A, stimulating the secretion of follicle stimulating hormone (FSH) from culture anterior pituitary cells.<sup>6</sup> Homologues were subsequently identified in several species, including humans. Since then activin A signalling has since been found to have a variety of biological functions, including in muscle homeostasis, inflammation and wound healing as well as being implicated in several forms of cancer and in the disease Fibrodysplasia Ossificans Progressiva (FOP).<sup>7-10</sup> Activin A is a homodimer of inhibin- $\beta_A$  subunits that has many typical of features of the TGF- $\beta$  family. Due to its role in disease and its effects on cell differentiation, specific inhibition of activin is desirable. However, this has often been difficult to achieve due to the high structural and sequence similarity with other TGF- $\beta$  superfamily members and the ligand promiscuity of the receptors. In this thesis I will be looking at ways in which to generate specific inhibitors of activin A signalling.

### 1.3 The Cellular Signalling of the TGF- $\beta$ Superfamily

Activin A has been implicated in a diverse range of physiological and pathological processes. Before looking at these processes however, the signalling of activin A must be understood. Being a member of the TGF- $\beta$  superfamily, activin A signalling shares many features with that of the signalling of other family members. Therefore, I will first look at the wider signalling pathway of the TGF- $\beta$  family as a whole, focussing on each component in detail.

### 1.3.1 The Canonical TGF- $\beta$ Signalling Pathway

Unlike the FGF, EGF, and VEGF families, the TGF- $\beta$ s do not primarily signal through a receptor tyrosine kinase pathway, but rather have their own unique serine / threonine receptor kinase pathway.<sup>1,11,12</sup> The canonical mechanism for TGF- $\beta$  family signalling is through a phosphorylation cascade that activates Smad proteins and causes them to induce transcription. In total, four receptors are required for signalling to occur, two type I receptors and two type IIs. In the basal state it is not known if these receptors exist alone or as pre-formed complexes, however FKBP12 binds to the type I receptor, inactivating its kinase domain preventing signalling from occurring.<sup>13</sup> An external stimulus causes dissociation of the pro-domains from a TGF- $\beta$  pro-mature complex. The mature TGF- $\beta$  ligand can then form heteromeric complex with the extracellular domains of two type I and two type II receptors. Formation of this complex causes FKBP12 to dissociate from the receptor, which in turn leads to the constitutively active kinase domain of the type II receptors to phosphorylate the glycine / serine rich (GS) region of the type I receptors. This activates the kinase domain of the type I receptors which then recruits regulatory-Smad proteins (R-Smads) and phosphorylates them. Two phosphorylated R-Smads then typically hetero-oligomerize with Smad4 before translocating to the nucleus, where the complex interacts with other transcription factors and DNA. This triggers transcription and causes a cellular response (Figure 1).<sup>1</sup>

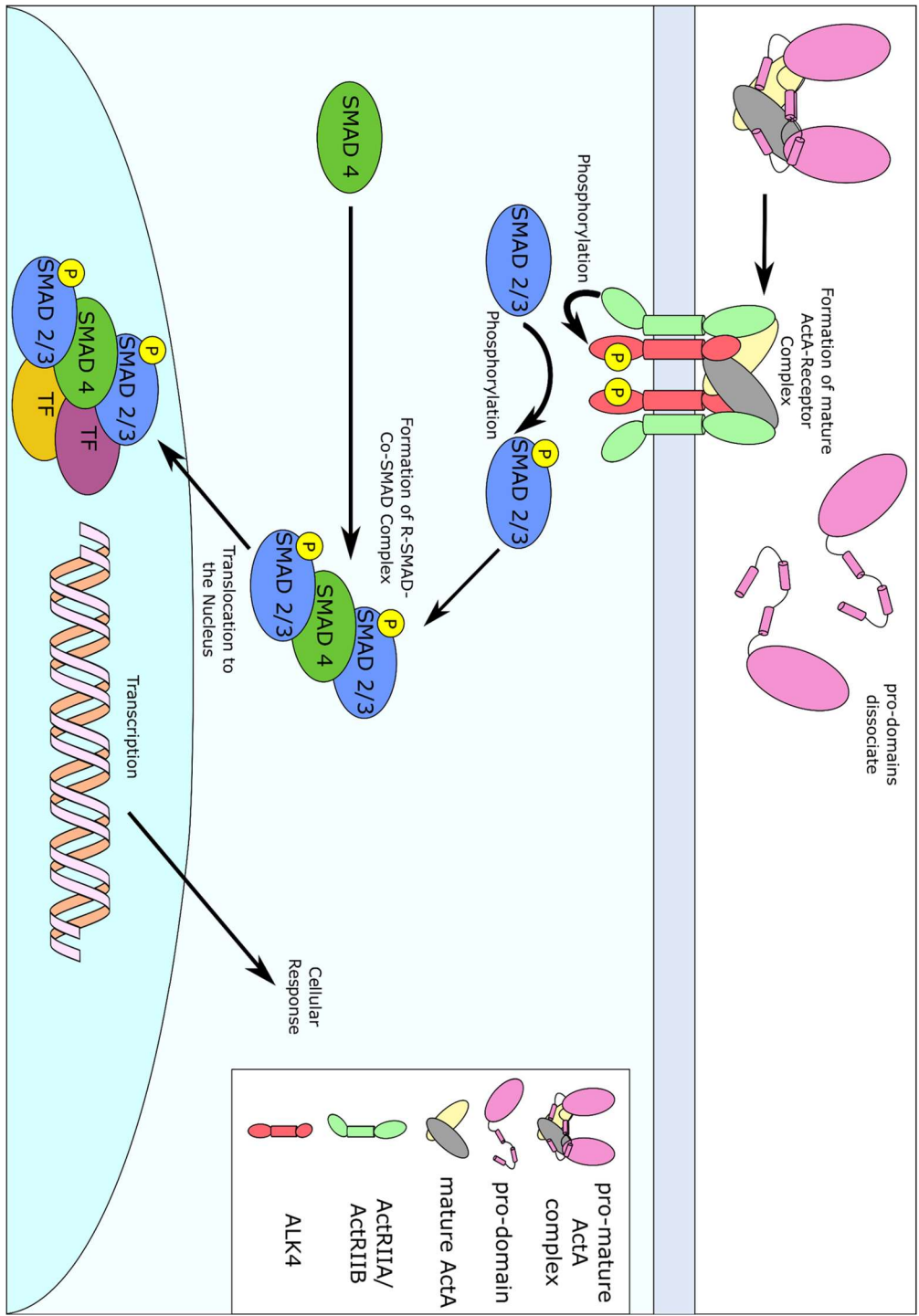


Figure 1: The cellular signalling of activin A. The canonical mechanism of TGF- $\beta$  family signalling shown here using activin A as the example. Thus, the type I, type II receptors and R-Smads shown are ALK4, ActRIIA/B, and Smad2/3 respectively.

### 1.3.2 The TGF- $\beta$ Receptors

All TGF- $\beta$  ligands signal through the formation of a heteromeric complex consisting of two autonomous pairs with two type I and two type II receptors.<sup>14–16</sup> In total there are only seven type I receptors (ALK1–7) and five type II (ActRIIA, ActRIIB, TGF $\beta$ RII, BMPR2, and AMHR2) meaning multiple

ligands share the same receptors and receptor combinations. Within the TGF- $\beta$  family there are two broad subgroups into which the receptors and ligands can be divided based on receptor ligand interactions and their subsequent downstream SMAD phosphorylation cascades – one containing ActRIIA/B and ALK4/5/7 activated by the activins, GDFs 8 and 11, and nodal and the other containing BMPRII and ALK1/2/3/6 activated by the BMPs and the other GDFs.<sup>3,5</sup> Though this rule holds true in the majority of cases there are numerous exceptions. Both BMP2 and BMP4 signal through BMPRII and ActRIIA/B, and activin A has been observed to signal through ALK2 in FOP.<sup>10,17</sup> Unlike most other members of the family, TGF- $\beta$ 1/2/3 only signal through the type II receptor TGF $\beta$ RII. However this can result in the formation of signalling complexes with either ALK1 and ALK5 resulting in two different Smad phosphorylation pathways and mixed Smad complexes.<sup>18</sup> Additionally, AMH signals through a unique type II receptor – AMHRII. It is thought in this case that the type I receptor can be ALK2/3/6 as interactions are observed between AMH and all type Is.<sup>19,20</sup> AMH only signals through the same Smad pathway as the BMPs.

The type I and type II receptors have the same general features, consisting of an extracellular domain, transmembrane helix, and intracellular kinase domain. Type I receptors also have an additional GS region which is phosphorylated by the type II receptors during signalling. One of the key regulators of this phosphorylation is FKBP12. FKBP12 inhibits type I receptor in the basal state. It binds proximal to the GS region, causing a section of the GS region to disrupt the kinase domain resulting in an inactive conformation being adopted.<sup>21</sup> FKBP12 does not prevent association of type II receptors to the type II receptors however it does prevent phosphorylation of the type I receptors by the type II, suggesting it has a role in preventing signalling in the absence of ligand.<sup>13</sup> It is unclear exactly how FKBP12 dissociates from the type I receptor however it is likely to come from the formation of the ligand – type I – type II receptor complex. Dissociation of FKBP12 allows for phosphorylation of the GS region of the type I receptor by the type II receptor, activating the type I receptor kinase domain and thus leading to signalling.

Regulation of the kinase activity of the type II receptor is governed by the phosphorylation state of several key residues in the cytoplasmic region of the protein. Phosphorylation has been shown to be required on serine residue S213 and S409 for signalling to occur, with phosphorylation of S213 occurring through intramolecular autophosphorylation whereas as phosphorylation of S409 can occur through both inter and intra molecular autophosphorylation and is enhanced by receptor dimerization. In addition to the two serine residues, phosphorylation of three tyrosine residues (Y259, Y336, and Y424) has been shown to be vital, with their loss resulting in a reduction in signalling.<sup>22</sup> By contrast the autophosphorylation of residue S416 results in inhibition of signalling, with Luo *et al.* showing mutation to alanine results in a hyperactive receptor. A given receptor may be

phosphorylated on either S409 or S416 but not both allowing for the dynamic regulation of receptor activity and thus signalling.<sup>23</sup>

Two modes of ligand binding exist, with the mechanism being determined by the growth factors affinity for a specific receptor. The first is a cooperative mechanism where the growth factor binds the type II receptors creating a binding interface to which the type I can bind. In this mechanism, direct contacts are formed between the extracellular domain of the type I receptor and the extracellular domain of the type II receptor plus the growth factor, thus binding of the growth factor to the type I receptor is significantly stronger in the presence of the type II. As such, ligands that use this mechanism have a high affinity for the type II receptor and a low affinity for the type I. This is characteristic of the eponymous TGF- $\beta$ s.<sup>24,25</sup> In the second mechanism, the growth factor binds to the type I and type II receptors in independent events, and the extracellular domains do not form direct contacts with each other.<sup>26,27</sup> This is typical of the BMPs which unlike the TGF- $\beta$ s usually have the type I receptors as their high affinity receptors.<sup>28,29</sup> It has been suggested that this second method is due to the existence of pre-formed homodimeric receptor complexes of BMP receptors on the cell surface resulting in an increased affinity due to an avidity effect.<sup>88,91</sup> The complexation of the receptors at the cell surface has important implications for signalling. In the case of the TGF- $\beta$  family there are two possible modes of receptor assembly – either they exist as pre-formed complexes at the cell surface or they form complexes in the presence of ligand. There is currently conflicting evidence as to which of these mechanisms is the case with a number of studies observing different results for the same receptors. For the TGF- $\beta$  receptors ALK5 and T $\beta$ RII, Gilboa *et al.* observed both form homodimeric complexes in the endoplasmic reticulum and that these homodimers were preserved at the cell surface.<sup>30</sup> However, this was not the case with ALK5–T $\beta$ RII complexes which were only detected at low levels.<sup>31</sup> In the case of the BMP receptors, it has been observed that preformed heterodimeric and homodimeric receptor complexes exist on the surface of the cell and both these complexes are significantly enhanced in the presence of ligand.<sup>32</sup> One thing important to note is that these studies used overexpression of the receptors in question. In contrast Zhang *et al.* used fluorescence microscopy to observe that at low expression levels, T $\beta$ RII exists as monomers in the cell membrane, only dimerising upon the addition of ligand.<sup>33</sup> The same group also observed the similar results when looking at ALK5.<sup>34</sup> A fluorescence based method was also used to observe that T $\beta$ RII was trafficked to the cell surface as a monomer.<sup>35</sup> This conflicting evidence for receptor assembly may be due to the differing techniques but also the cell lines used. Different cell lines are known to have differing levels of receptors.<sup>36</sup> In these studies, with Gilboa *et al.* using L6 and COS7 cells and Zhang *et al.* using HeLa and MCF7 cells. Therefore, it may be the case that the existence of pre-formed complexes depends on cellular expression and thus the cell line.

Receptor promiscuity is well described in the TGF- $\beta$  family.<sup>37</sup> As there are only seven type I and five type II receptors, it follows that multiple different ligands must be able to signal through the same receptors and even receptor combinations. There are a number of mechanisms that exist by which cells can direct signalling. These include differences in ligand affinities, such that certain ligand receptor complexes are only formed at high concentrations of the ligand; differences in ligand processing, such that mature ligands are only available for signalling under certain conditions; and the involvement of co-receptors either in an agonistic or antagonistic manner.<sup>38-41</sup> In addition, cells can direct signalling through differences in receptor expression levels between tissues and certain receptor complexes being favoured such that though a TGF- $\beta$  ligand may be able to interact with one of the receptors in a complex, it cannot interact with the complex due to its inability to interact with the other receptor.<sup>17,36</sup> Thus signalling can be directed to produce a diverse range of responses in different tissues.

### 1.3.3 The Smad Proteins

In the canonical TGF- $\beta$  signalling pathway, the Smads are the main substrates for the type I receptors. There are eight Smad proteins in total, five regulatory (Smad1, Smad2, Smad3, Smad5, and Smad8), two inhibitory (Smad6 and Smad7), and one common Smad (Smad4). Similar to that observed with the growth factors and receptors, different R-Smads are only phosphorylated by specific type I receptors – with ALK1/2/3/6 phosphorylating Smad1/5/8 and ALK4/5/7 phosphorylating Smad2/3.<sup>42</sup> They all share a common architecture consisting of an N-terminal MH1 domain, a linker region, and a C-terminal MH2 domain. Both domains can interact with various transcription factors and MH1 can directly interact with DNA,<sup>43</sup> however it is the MH2 domain that interacts with type I receptors. Association to type I receptors is governed by the “basic patch”, a region in the MH2 domain containing a high number of positively charged residues that binds to the phosphorylated GS region of the activated receptor. Phosphorylation then occurs at the C-terminus of the MH2 domain, on an SSXS motif where X is a valine or methionine residue.<sup>44</sup> The phosphorylated SSXS then interacts intramolecularly with the basic patch, causing the Smad protein to dissociate from the receptor.<sup>45</sup> In the case of Smad2 this can be aided by SARA which binds to both the type I receptor and Smad2.<sup>46</sup> Receptor specificity is determined by a combination of the MH2 domain and the receptor kinase domain with the sequence of two loops, the L3 loop of the Smad and the L45 loop of the receptor kinase domain, being vital to determining receptor-Smad interactions. The identity of the X in the SSXS motif has also been identified as playing a crucial role in the generating specificity.<sup>47,48</sup>

The I-Smads, Smad6 and Smad7, act as intracellular inhibitors of TGF- $\beta$  signalling. Smad6 preferentially inhibits BMP signalling whereas Smad7 is more ubiquitous inhibiting both ligands from the BMP and



TGF- $\beta$  / activin pathways.<sup>49,50</sup> The architecture of the I-Smad proteins is similar to that of the R-Smads however there are some key differences. Like all Smads, I-Smads consist of a MH1 domain, linker region and MH2 domain however the MH1 domains are poorly conserved with that of the R-Smads. By contrast their MH2 domains are highly conserved with the R-Smads though they lack an SSXS motif at the C-terminus of their MH2 domains. This means though their basic patch can interact with the type I receptors they cannot be phosphorylated, thus cannot dissociate resulting in the formation of a non-signalling complex. Additional mechanisms of inhibition also exist through interactions with Smad4, interactions with the R-Smads and targeting the receptors for degradation.<sup>51-53</sup>

Smad4 is distinct from the other Smad proteins. Like the I-Smads, it lacks a SSXS motif for phosphorylation however it does not inhibit signalling. Rather its role is to oligomerise with the R-Smads to facilitate translocation to the nucleus.<sup>54</sup> In an unstimulated cell there is constant shuttling of the Smads between the nucleus and cytoplasm however both the R-Smads and Smad4 are primarily cytoplasmic.<sup>55</sup> Upon signalling, Smad4 forms a heterotrimeric complex with two phosphorylated R-Smads through its MH2 domain. This complex then translocates to the nucleus where it is retained and accumulates.<sup>56</sup> The accumulation is due to the import rate of these complexes being four times higher than their export rate.<sup>57</sup> Once in the nucleus, these complexes can either bind to DNA or interact with other transcription factors to induce a cellular response. It is only upon dephosphorylation does the complex dissociate and both the R-Smads and Smad4 are exported.<sup>58</sup>

SARA (Smad Anchor for receptor activation) is a protein that can facilitate Smad signalling through aiding the localisation of R-Smads to the type I receptors.<sup>46</sup> It consists of a FYVE zinc finger domain that allows it to localise to the cell membrane, a Smad binding domain, a PP1c binding domain, and a C-terminal type I receptor binding region. The Smad binding domain has been shown to interact with Smad2 and Smad3 through their MH2 domains, however due to the high conservation between the Smad MH2 domains, interactions with the other Smads may be possible. SARA preferentially binds unphosphorylated Smads, presenting them to the type I receptor for phosphorylation. Upon phosphorylation, the Smads then dissociate from SARA allowing them to translocate to the nucleus. Interestingly, it appears that SARA is not required for TGF- $\beta$  signalling in most cases. In COS7 cells, Goto *et al.* demonstrated that although SARA was required for Smad2 dependent signalling it was not for Smad3.<sup>59</sup> Additionally in HeLa cells and B lymphocytes, Bakkebo *et al.* observed that SARA was not required for phosphorylation of any of the Smads.<sup>60</sup> Therefore it is likely SARA plays more of an aiding role, facilitating TGF- $\beta$  signalling however its full role is yet to be fully elucidated.

### 1.3.4 Non-Smad Mediated TGF- $\beta$ Signalling

Signalling through the Smad pathway is the canonical signalling mechanism for the TGF- $\beta$  superfamily, however it is not the only pathway through which signalling can occur. TGF- $\beta$  ligands have been reported to signal through a number of additional pathways and I will now look at how TGF- $\beta$  signalling occurs through these pathways.<sup>61</sup>

The non-Smad signalling of the TGF- $\beta$  can occur through three primary mechanisms: phosphorylation of non-Smad proteins by the TGF- $\beta$  receptors; interactions between Smads and non-Smad pathway proteins such that the Smad pathway is perturbed; and interactions between Smads and non-Smad pathway proteins such that the non-Smad pathway is perturbed. In the case of receptor phosphorylation, both the type I and type II receptors can phosphorylate non-Smad pathway proteins. Ozdamar *et al.* observed that TGF- $\beta$  can interact with the Rho GTPase pathway through phosphorylation of the protein Par6 by the type II receptor leading to degradation of Rho A and disassembly of the cytoskeleton.<sup>62</sup> Furthermore, Src has been shown to phosphorylate the TFBRII on Y284 allowing for recruitment of Grb2 and Shc allowing for initiation of the Erk pathway.<sup>63</sup> The Erk pathway can also be initiated by the type I receptors through activated type I receptors directly phosphorylating Shc allowing it to recruit Grb2 and Sos initiating the signalling cascade.<sup>64</sup> Another interaction of the type I receptors with a non-Smad pathway is that of the Akt pathway. However, in this case it is the Akt pathway that interacts with the receptor through localising USP4, a deubiquitylating enzyme, to the cell membrane. This results in a reduction in type I receptor degradation thus increasing the effective concentration of type I receptor on the cell surface.<sup>65</sup>

The amount of cross talk between the Smad and non-Smad pathways is often significant and this can be highly important in determining biological outcome. Interactions between Smads and non-Smads usually results in a modulation in their transcriptional activity, either upregulating or downregulating expression. Kretzschmar *et al.* demonstrated that for Smad 1, modulation by the Erk pathway results in a downregulation in Smad transcriptional activity due to a reduction in nuclear accumulation.<sup>66</sup> This is due to the presence of a number of serine residues in the linker region of the R-Smads that can be phosphorylated by Erk pathway kinases. A similar phosphorylation pattern was observed on Smad2 by Erk which resulted in an increase in the duration of Smad2 induced transcription.<sup>67</sup> Interestingly as TGF- $\beta$  can itself activate the Erk pathway, it may therefore be the case that TGF- $\beta$  can modulate its own Smad signalling through non-Smad pathways. In addition to the modification of Smads, modulation of other pathways by the Smads has also been reported. The Smad3-Smad4 complex has been observed to interact with the regulatory subunit of protein kinase A resulting in the release of its catalytic domain independent of cAMP concentrations.<sup>68</sup> Furthermore, an indirect modulation has

been observed for the Jnk and p38 pathways where one mechanism of TGF- $\beta$  activation involves Smad signalling inducing expression of proteins that can trigger these pathways.<sup>69</sup>

Synergism between the Smad and non-Smad pathways is often important in determining a signalling outcome. In the embryogenesis of chickens, BMP5 signals through the Smad and p38 pathway to mediate limb development.<sup>70</sup> Furthermore, the Smad, Erk, and Notch pathways have been shown to work together to promote the epithelial-mesenchymal transition of cells.<sup>71,72</sup> In the case of activin A, its synergism with FGF pathway in mesoderm induction and maintaining the pluripotency of stem cells has been reported.<sup>73,74</sup> All these factors combined mean the true TGF- $\beta$  signalling network in cells is highly complex with many factors and proteins interplaying to produce a specific response.

### 1.3.5 The Cellular Signalling of Activin A

The canonical and non-canonical signalling of activin A follows the generic mechanisms described above with the pro-domains dissociating through an unknown mechanism followed by formation of a heteromeric signalling complex and a phosphorylation cascade through the Smad proteins (Figure 1). However, there are some features that differentiate activin A signalling from that of other TGF- $\beta$  family members. Canonically activin A signals through three of the twelve TGF- $\beta$  receptors in most circumstances – the type II receptors ActRIIA and ActRIIB and the type I receptor ALK4. Of these receptors, activin has a high affinity for the type II receptors and a low affinity for ALK4.<sup>14,75</sup> Mechanistically this means it is likely activin binds to type II receptors first to form a complex which then binds to the type I receptor, however this signalling complex appears to be unstable. Attisano *et al.* demonstrated that when metabolically labelled ALK4 is added to activin A-ActRIIB complex only a small amount of the ternary complex forms, indicating a low affinity of activin A-ActRIIB for ALK4.<sup>14</sup> Therefore, this suggests formation of the ternary complex only occurs transiently however the process may be aided *in vivo* by close spatial proximity of the type I and II receptors. It is currently unknown whether the mechanism through which activin A binds to its receptors occurs through the cooperative mechanism typical of TGF- $\beta$ s or through the independent events mechanism typical of the BMPs, though the latter may be more likely. Goebel *et al.* determined the structure of the ternary GDF11-ActRIIB-ALK5 complex and found that there appear to be no direct contacts between the type I and type II receptors.<sup>76</sup> As activin A is highly similar to GDF11 and signals through the same Smad signalling pathway, it is therefore likely activin A binds its receptors through this mechanism.

Due to the interaction with ALK4, it is conventionally thought of that activin only signals through the Smad 2/3 pathway. However additional activin signalling has been observed through ALK2 and the Smad 1/5/8 pathway under certain circumstances. The disease Fibrodysplasia Ossificans Progressiva is caused by mutations in the intracellular kinase domain of the ALK2 receptor which makes it

responsive to activin signalling.<sup>10</sup> Furthermore Olsen *et al.* observed apoptosis triggered by activin through the Smad 1/5/8 in a number of myeloma cell lines.<sup>36</sup> This indicates that activin A can interact with ALK2 under certain circumstances however does not usually signal through it. This is likely due to the interactions between BMPRII and ALK2. In the case of BMP receptors it is thought that pre-existing heteromeric complexes are formed on the cell surface. As ALK2 is usually a receptor for the BMP pathways it may be the case that the majority of ALK2 is already assembled in a complex with BMPRII in most cell lines, to which activin A cannot bind. In the absence of BMPRII both activin A and B signalling through ALK2 is massively upregulated implying that this is indeed the case. A reverse mechanism is also observed where activin A can inhibit BMP signalling through ActRIIA / ActRIIB due to its higher affinity for those receptors.<sup>17</sup>

### 1.3.6 Regulation of Activin A Signalling

In order for activin A and other members of the TGF- $\beta$  family to produce the diverse range of cellular responses observed, the cell must have mechanisms through which it can dynamically control signalling. In addition to regulation through morphogenic behaviour, variable expression patterns between cell types and the I-Smads, a number of additional mechanisms exist which can regulate signalling. These mechanisms can be broadly broken down into those that interact with activin A and those that interact with the type I and type II receptors. I will now focus on the proteins involved in these mechanisms and how they sequester activin A signalling.

Follistatin is an extracellular glycoprotein that inhibits a number of TGF- $\beta$  family members, including activin A. It exists in three isoforms – FS315, FS303, and FS288, all of which contain four domains Fs0, Fs1, Fs2, and Fs3. Of these domains Fs1 and 2 are crucial for activin inhibition.<sup>77</sup> In addition, the Fs1 domain contains a highly basic region allowing for binding of heparan sulfate.<sup>78</sup> The main difference between the isoforms is at the C-terminus. FS315 contains a highly acidic tail which is lacking in the splice variant FS288. FS303 is a proteolytically processed form of FS315 which partially lacks the C-terminal tail but still contains the acidic sequence. The *in vivo* distribution of these three isoforms varies dramatically, with FS288 and FS300 being primarily found in tissue and follicular fluid whereas FS315 is only detected in serum.<sup>79</sup> One reason for this could be that FS288 has a greater affinity for cell surface proteoglycans than FS303, which in turn has a higher affinity than FS315.<sup>80</sup> It has been suggested that this could be due to the acidic tail of FS315 and FS303 interacting intramolecularly with the basic heparan sulfate binding region, reducing the ability of the isoforms to bind surface proteoglycans and thus aiding transport around the body however there is currently not sufficient evidence for this.

Of the TGF- $\beta$  ligands inhibited by follistatin, it appears to have the strongest affinity for activin A.<sup>81</sup> Activin inhibition is achieved through forming a high affinity complex with the mature growth factor (displacing the pro-domains from the pro-mature complex to achieve this). Follistatin shares the same binding site on activin A as the type II receptors and thus the receptors cannot bind and signalling is inhibited.<sup>77</sup> A related protein to follistatin, FSTL3 also inhibits activin signalling in the same manner. FSTL3 shares very high structural similarity with follistatin and both molecules bind activin A with similar affinities ( $K_d$  approximately 39 pM).<sup>82</sup> Significant differences are found however between the sequences of the N-terminal region and this is thought to have an effect on specificity, with follistatin antagonising a greater number of TGF- $\beta$  family ligands than FSTL3.<sup>80,82</sup> Interestingly, FSTL3 has also been found to localise to the nucleus, though any effect nuclear FSTL3 has on activin or wider TGF- $\beta$  family signalling is unclear.<sup>83</sup>

The second mechanism by which activin signalling is inhibited is through inhibitors binding to its receptors. Inhibin A is the main example of this. Inhibin A was isolated one year before activin though it was first discovered in 1923 as a secreted factor from the testes that could regulate function of the pituitary gland.<sup>84-86</sup> It consists of one inhibin  $\beta$ A subunit similar to activin A, however its second subunit is an inhibin  $\alpha$  chain making inhibin A heterodimer. This allows inhibin to interact with ActRIIA and ActRIIB on the cell surface through its  $\beta$ A subunit however rather than recruit two type II receptors, its  $\alpha$  chain interacts with betaglycan, a type III TGF- $\beta$  receptor. Formation of this complex does not trigger a phosphorylation cascade as the complex is unable to recruit a type I receptor. This competition for type II receptors results in a reduction of available receptors on the cell surface thus resulting in a reduction in activin A signalling.<sup>87</sup>

Similar to inhibin, it is thought that activin C may also compete with activin A for receptor binding. Of the five activin dimers observed in humans (activin A, activin B, activin AB, activin C, and activin E), activin C and E are the least well described with their biological function being poorly known. However, work by Gold *et al.* suggests that the activin C (a homodimer of inhibin  $\beta$ c chains) has an inhibitory role on activin A signalling, specifically in the liver and sex organs, likely through competition for receptors. However the mechanism for this is still poorly defined.<sup>88</sup>

Betaglycan is one example of a TGF- $\beta$  co-receptor. These proteins have a variety of roles including binding growth factors to the cell surface creating a cellular "store", as well as inhibiting signalling through disrupting the formation of an active ligand-receptors complex.<sup>89,90</sup> For activin A betaglycan is an indirect inhibitor, interacting with inhibin to antagonise activin A signalling. Other co-receptors have also been identified that modulate activin signalling, one of which is the pseudo-receptor BAMBI. BAMBI is a transmembrane protein with an extracellular domain that has a high similarity to TGF- $\beta$

type I receptors, however lacks the intracellular kinase domain of type I receptors and thus cannot trigger a phosphorylation cascade.<sup>91</sup> This means BAMBI acts as a decoy for type I receptors on the cell surface interacting with the activin A-ActRIIA/B complex to curtail signalling. Similar to BAMBI, the cell surface protein cripto also has the ability to antagonise activin A signalling through disrupting the ability of the activin A-type II receptor complex to recruit ALK4. Interestingly cripto has the opposite effect on the TGF- $\beta$  ligand nodal which cannot signal without the presence of cripto. Both activin and nodal signal through the Smad 2/3 pathway, therefore cripto may provide a mechanism to prevent cross over between the signalling of these two ligands.<sup>41,92</sup>

## 1.4 The Structural Features of the TGF- $\beta$ Superfamily and Activin A

The TGF- $\beta$  family ligands are typified by the large degree of structural and sequence similarity between members. In this section, I will focus on what the key structural and biochemical features of the ligands are and how they define the interactions between the mature growth factors, their pro-domains and type I and type II receptors. I will specifically focus on the key features of activin A and how this relates to its observed behaviour *in vivo*.

TGF- $\beta$  family proteins are synthesised as large precursors containing an N-terminal signal peptide, a poorly conserved pro-domain that is not actively involved in signalling, and an active highly conserved mature domain (pairwise identity across the whole family, pro-domains = approximately 15 %, mature domains = >30 %).<sup>4</sup> Upon expression, these precursors translocate to the endoplasmic reticulum where they assemble and dimerize. During assembly, domain swapping occurs for some members of the family with the pro-domain of one subunit interacting with the mature domain of the other subunit to adopt the correct dimeric conformation. However, this is not expected to be the case always.<sup>93</sup> Dimerization is typically stabilised through the formation of a disulfide bond between the mature domains. Most products of this dimerization are homodimers however heterodimers do occur, including TGF- $\beta$ 1.2 and activin AB. After dimerization, the pro-TGF- $\beta$  ligand is cleaved by golgi-associated furin-like proteases, cleaving between the pro- and mature domains.<sup>94</sup> However, this does not cause the pro-domains to dissociate and they remain in complex with the mature growth factor. The pro-mature complex is then secreted from the cell where it either binds to the extracellular matrix or surface proteins and is stored ready for use or enters circulation.<sup>95-97</sup>

Extracellular pro-mature TGF- $\beta$  complexes are either secreted in an “active” or “latent” form depending on whether they require additional processing to induce signalling. This additional processing usually takes the form of a second protease cleavage step that cleaves somewhere in the pro-domain allowing the mature domain to be released from the complex.<sup>98</sup> In the case of ligands that

form latent complexes, it is often the case that a ligands pro-domain can inhibit the mature domains signalling with high efficacy. This is the case for the TGF- $\beta$ s, myostatin and GDF11.<sup>39,99,100</sup> By contrast ligands that fall into the active category require no additional processing after secretion and include BMP9 and activin A.<sup>40,101</sup> Inhibition of signalling by the pro-peptide for these proteins may also occur, albeit at significantly higher concentrations of pro-domain than those required for latent growth factors.

Numerous structures exist of TGF- $\beta$  ligands, either as pro- and mature ligands, in complex with inhibitors, or as part of a complex with receptors. They have a “butterfly like” dimeric structure with each monomer having a convex and concave face.<sup>37</sup> Monomers can be broken down into regions named after the human arm. The pro-domains consist of a forearm region containing an  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif that forms the majority of contacts with the mature domain, and a globular shoulder region that sits atop the mature dimer. The mature domains consisting of a wrist helix and a short and long finger region made up of two pairs of twisting anti-parallel  $\beta$ -sheets (Figure 2).<sup>102</sup> A high percentage of residues in mature domains of family members are cysteines (5.0-6.3 %) and these intricately link each chain through a series of inter- and intramolecular disulfide bonds forming a cysteine knot. Due to the similar architecture and sequence identity, receptor binding sites appear to be preserved across the family with type II receptors binding at the fingertip region on the convex face and type I receptors binding to a pocket located at the dimer interface on the concave face.<sup>27</sup> This pocket is highly hydrophobic in nature and exists between the wrist helix of one monomer and the finger region of the other. Activin A follows all these typical features. It is a 426 residue protein consisting of a 20 amino acid signal peptide, a 290 amino acid pro-domain and a 115 amino acid mature domain. The structures of both the pro- and mature forms of activin A are known, as well as the pro-mature complex, the activin-follistatin complex, and the activin-ActRIIB complex.<sup>77,103,104</sup> One peculiarity of these structures is the large degree of flexibility about the dimerization cysteine. In each of the structures this angle is different and can vary anywhere between 108° in the pro-mature complex (PDB ID: 5HLZ) to 57° in the structure of activin A with ActRIIB (PDB ID: 1NYS). This may have some biological relevance, with Greenwald *et al.* suggesting this change in conformation is crucial for allowing ALK4 binding.<sup>103</sup> However, further work is required to fully elucidate this.

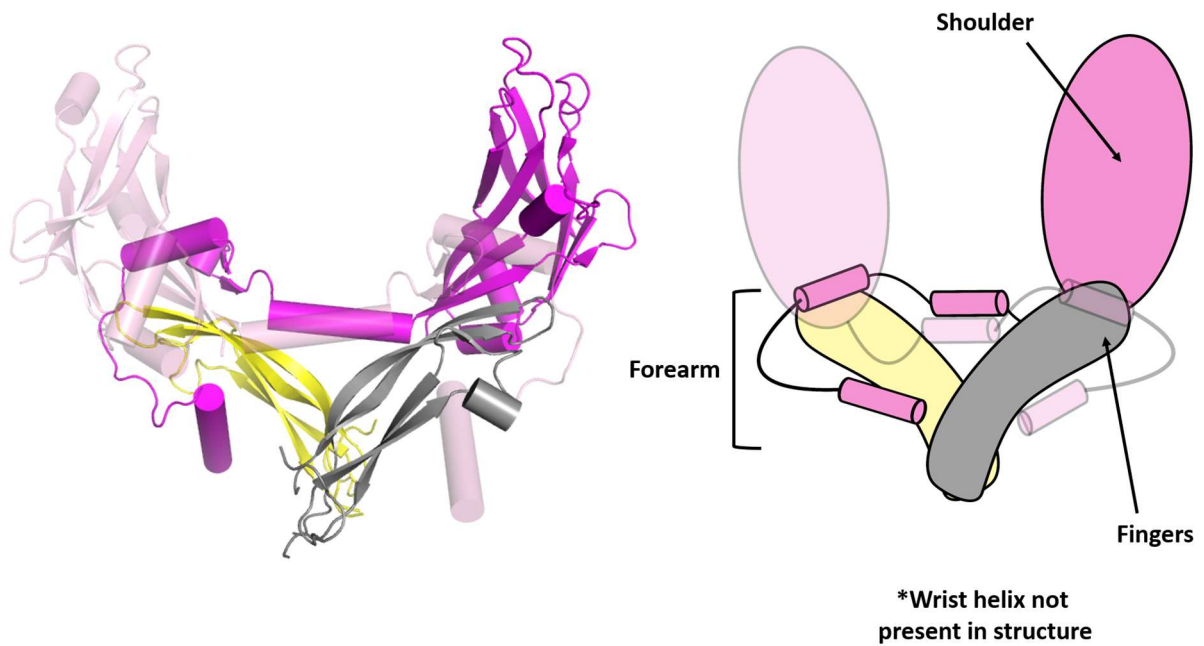


Figure 2: Topology of pro-TGF- $\beta$  family ligands. Here the structure of the pro-activin A is shown as an example (left) with a schematic (right) where the different regions of the ligand are labelled. The pro-domains are shown in pink and the mature domains in yellow and grey. For clarity, one pro-domain has been made transparent.

The role of pro-domains in the TGF- $\beta$  family is highly important. Though not thought to be actively involved in signalling, it aids in dimerization, assembly and interactions with the extracellular matrix.<sup>95,96,105</sup> Furthermore it helps to keep the mature growth factor soluble in solution through the binding of the forearm region to the mature dimer which buries the hydrophobic residues of the mature domains within the structure of the protein.<sup>104,106</sup> In addition to pro-activin A, there are currently three other members of the TGF- $\beta$  family for which the structure of the pro-form is known – TGF- $\beta$ 1, BMP 9 and myostatin (Figure 3).<sup>102,107,108</sup> Comparison of the four structures reveals a number of key similarities. Three of the four pro-domains start with a highly hydrophobic helix ( $\alpha$ 1-helix) followed by a latency lasso (loop) that wraps around the mature domain and then a second helix ( $\alpha$ 2-helix) (the N-terminal region of BMP9 is poorly defined however an additional  $\alpha$ -helix occupies the site occupied by the  $\alpha$ 1-helix in the other three structures). The lasso region contains a high number of prolines (approximately 33 % in most family members), adding rigidity to the region and allowing it to encircle the mature domain in a so-called “straitjacket”. After the second  $\alpha$ -helix, a variable linker connects this “straitjacket” region to the shoulder region. This region has a similar fold in all TGF- $\beta$  members but there is no overall preserved structure. Three of the four structures, pro-TGF- $\beta$ 1, pro-myostatin, and pro-activin A, all show a domain swapped topology with the pro-domain of one subunit interacting primarily with the mature domain, however it is unclear if this is the case for pro-BMP9.<sup>93</sup> Furin cleavage between the pro- and mature domains appears not to significantly affect the overall



structure with no significant changes being observed between the X-ray structures of pro-activin A and the activin A pro-mature complex.<sup>104</sup>

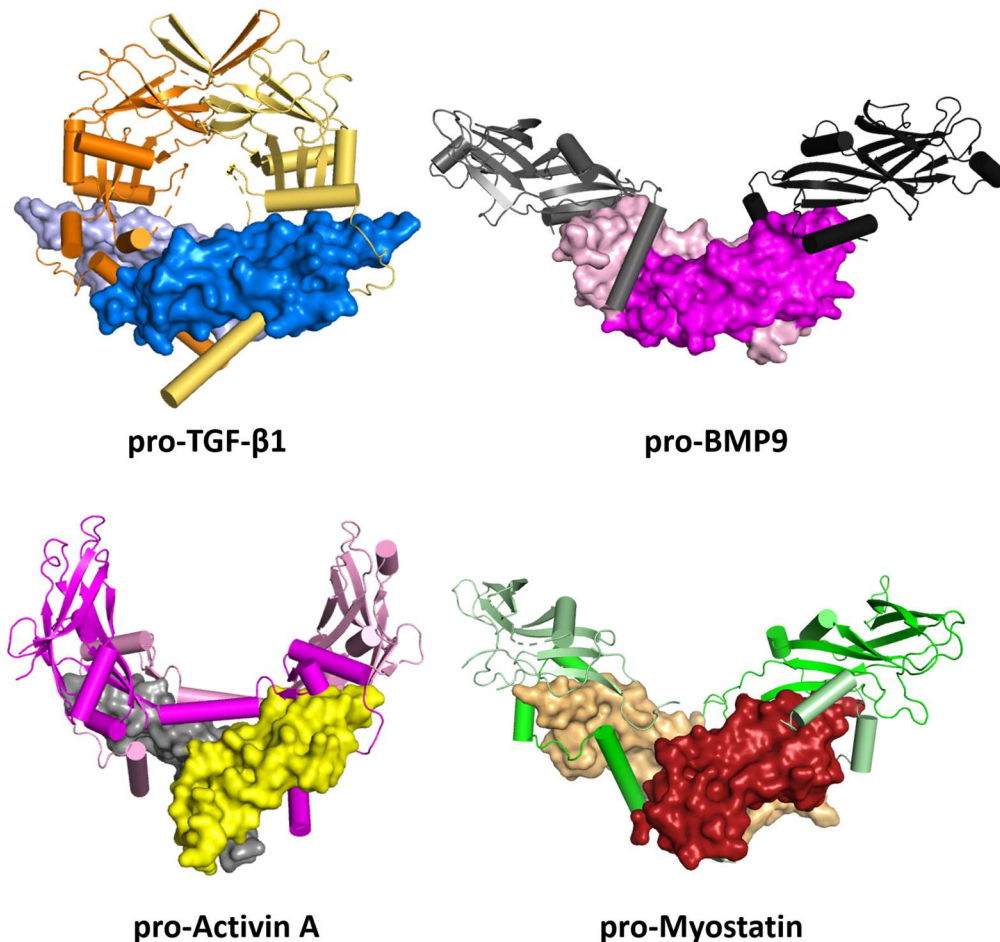


Figure 3: Pro- structures of the TGF- $\beta$  family. A figure showing the available structures of pro-TGF- $\beta$  family members. Here, the mature domains are shown in a surface representation whereas the pro-domains are shown with a cartoon representation.

There are however major differences. There appears to be no overall trend as to whether a particular ligand will be a latent or active complex. Pro-Activin A, myostatin and BMP 9 all appear to adopt an “open arm” “V-like” conformation whereas TGF- $\beta$ 1 adopts a “closed arm” disulfide linked conformation (though Mi *et al.* suggest interconversion between “open arm” and “closed arm” conformations may be possible).<sup>107</sup> Interactions between pro-domains varies between structures as well. In TGF- $\beta$ 1, extensive hydrophobic interactions exist in addition to the inter pro-domain disulfide, whereas in pro-activin A interactions appear to come mainly through two additional helices ( $\alpha$ 3-helix) that also appear to interact with the base of the finger regions. In myostatin and BMP 9 there appear to be no major interactions between the pro-domains. Therefore, the overall poor conservation between pro-domains when compared to the mature domains appears to have a major impact on the tertiary structure of the pro-ligand.

## 1.5 The Pathology and Physiology of Activin A

Now I have covered the main structural features of activin A and how activin A signalling occurs, I will now look at the role of activin A in physiology and disease. The biology of activin A is diverse and has been implicated in many physiological and pathological processes. In this section chapter I will focus on a number of these processes, especially those that are therapeutically or commercially important.

### 1.5.1 Activin A in Development and Stem Cell Biology

The role of activin A in development is complex. Homozygous knockout mice lacking either inhibin  $\beta_A$  or  $\beta_B$  genes both yielded viable pups, though those lacking the inhibin  $\beta_A$  showed craniofacial defects and died within 24 hours of birth.<sup>109</sup> However, maternal activin A is present from an early stage in mammalian embryogenesis in both the oocyte, fallopian tube and womb which may account for survival till birth.<sup>110,111</sup> Furthermore, the expression of activin A or homologues has been detected in several tissues of developing mammals, including the sex organs, the brain and the kidneys.<sup>112-115</sup> The role of activin A in the initial stages of embryogenesis is deeply entwined with that of another TGF- $\beta$  family ligand, nodal.<sup>116</sup> Like activin A, nodal is detected early on in development (pSmad 2/3 detected from 4-egg stage in mice, expression of both detected from blastocyst stage) and both proteins signal through the same receptors (ActRIIA/B and ALK4) and Smad 2/3 pathway, making it difficult to discern which protein is performing which role. However, due to the role of cripto, it is likely activin A signalling is suppressed as nodal signalling increases.<sup>41</sup> Incidentally, disruption of nodal gene expression in mouse embryos has been shown to result in significant abnormalities including lack of mesoderm formation and patterning abnormalities.<sup>117,118</sup> However, it has been observed that both proteins can produce similar phenotypic results in stem cells and early animal models, including maintaining pluripotency and driving mesoderm formation.<sup>119-121</sup>

In stem cell technologies, activin A is often used as a surrogate for nodal due to the more widespread availability of activin A.<sup>122</sup> It has been demonstrated that activin A can maintain human embryonic stem cells in a pluripotent state without the need for feeder layers and can be used to maintain pluripotency in induced pluripotent stem cells as well.<sup>123,124</sup> Mechanistically, activin maintains pluripotency through triggering expression of a number of transcription factors that suppress differentiation including Oct4, Sox2 and Nanog.<sup>125</sup> In addition, a second mechanism for maintaining pluripotency may exist where activin acts in concert with other growth factors to suppress BMP signalling. BMP4 has been shown to initiate differentiation of human embryonic stem cells to the trophoblast, thus by suppressing their signalling pluripotency is maintained.<sup>126</sup>

The key to controlling pluripotency and proliferative potential is activin concentration. At low concentrations of activin A (0.12 – 1.2 nM), pluripotency is maintained. However once activin A levels increase pluripotency and proliferative potential are lost.<sup>123</sup> This is due to the morphogenic nature of activin A, with cells undergoing differential gene expression depending on the concentration of activin A to which they are exposed.<sup>127</sup> Thus, responses and cell fate at different concentrations of activin A will be different. This, along with the method of culturing, additives used, and the time point at which activin A is either added or removed determines the product of activin A induced differentiation.<sup>128</sup> Activin has been reported to drive the formation of the definitive endoderm in stem cells and differentiation into a number of cell types under the influence of activin has been reported including both endoderm and mesoderm cells.<sup>129,130</sup> Therefore by tightly controlling these factors activin can be used to determine the fate of stem cells in a desired manner.

### 1.5.2 Activin A in Inflammation and Wound Repair

Activin is one of the first cytokines released in acute inflammation.<sup>131</sup> Its behaviour is observed to be biphasic, with initial increase in serum levels followed by a drop and another increase. The source of activin A involved in the first phase is likely to be release from the extracellular matrix to which activin A is known to bind, however currently there is insufficient evidence for this.<sup>95</sup> The second increase in activin A levels is due to an increase in expression that is mainly driven by the Toll-like receptor pathway and Toll-like receptor 4.<sup>8</sup> After 4-5 hours, activin levels reduce and the amount of follistatin increases suppressing activin signalling.

Once present, the role of activin A in inflammation and injury is contradictory and depends on the tissue in question. The levels of activin A at a site of inflammation appear to correlate with the magnitude of an immune response, and activin A has been shown to trigger and suppress the levels of inflammatory cytokines depending on the cell type.<sup>132-134</sup> In addition, activin A can both aid and impede regeneration. In the liver, activin A appears to have an inhibitory effect on regeneration after a partial hepatectomy.<sup>135</sup> Furthermore, it has been observed that activin A can increase liver fibrosis and induce hepatocyte apoptosis.<sup>136,137</sup> In inflammatory bowel disease, activin A is expressed at significantly higher levels in the intestine and inhibits epithelial cell proliferation impeding the healing process.<sup>138,139</sup> However in the case of external wounding of the epidermis, increased activin A levels improve the speed of recovery. Activin expression is triggered upon injury to the epidermis and increases the level of granulation tissue and rate of wound repair but also scarring. By contrast, inhibition of activin through overexpression of follistatin results in a slower repair rate, with lower scarring and increased resolution between granulation and surrounding tissue.<sup>140-142</sup> Therefore

modulation of activin activity could be an important therapeutic target for aiding repair and regeneration and minimising scarring.

The anti-inflammatory role of activin A has been described in a number of instances. In the brain activin A has been shown to have a neuroprotective role, increasing survival of and protecting neurons from neurotoxic damage.<sup>143</sup> Furthermore, activin A has been observed to inhibit inflammatory cytokine release in angina patients.<sup>144</sup> Interestingly the opposite effect is observed in healthy patients where activin triggers an inflammatory response. One source of the anti-inflammatory effects of activin likely comes from its ability to antagonise interleukin-6 signalling.<sup>145,146</sup> As interleukins are key regulators of the immune system ultimately this results in suppression of the immune response. However, it is important to note this is just one of many of the complex interactions between activin A and other inflammatory cytokines and a full description of the interplay between these molecules is yet to be described.

### 1.5.3 Activin A in Muscle Homeostasis

Activin A is an important factor in the development and homeostasis of muscle tissue. It acts in concert with myostatin, another member of the TGF- $\beta$  family, to inhibit muscle growth, suppressing muscle fibre hypertrophy.<sup>147</sup> This allows an organism to control its muscle mass and fat metabolism, preventing an uncontrolled increase in the amount of skeletal muscle. Like activin A, myostatin is also inhibited by the extracellular protein follistatin.<sup>148</sup> As such, increased follistatin levels have been associated with an increase in skeletal muscle mass.<sup>149</sup> In healthy individuals, inhibition of activin and myostatin with follistatin helps to maintain muscle homeostasis however if this balance is altered a large increase in muscle mass can ensue. Schuelke *et al.* reports the case of a myostatin deficient boy who exhibited a highly muscular phenotype due to a loss of function mutation in the myostatin gene.<sup>150</sup> Furthermore, several breeds of cattle possess mutations that lower or eliminate myostatin activity.<sup>151</sup> As such they exhibit a significantly more muscular phenotype than other animals.

Activin A has also been implicated in the muscle wastage associated with cancer cachexia. In mouse models, mice overexpression activin A resulted in a significant decrease in muscle mass relative to their starting mass however this change was reversible upon suppression of activin signalling.<sup>152</sup> In humans, high levels of circulating activin A have been associated with cachexia in cancer patients, with patients displaying over a 40% increase in activin A levels when compared to non-cachectic patients. Interestingly, this study found that myostatin levels were decreased by over 35% in cachectic patients.<sup>153</sup> Activin A and myostatin are often implicated as acting in concert to negatively regulate muscle growth however in the case of cachexia it appears that activin A plays a more prominent role.

The general conclusion from a number of mice studies has been that although activin A contributes to muscle development and homeostasis, it is a secondary player when compared to myostatin (GDF8).<sup>147,154,155</sup> However, this may not be the case across mammalia, as shown by Latres *et al.* who studied the effect of specific inhibition of different TGF- $\beta$  family members.<sup>7</sup> It was observed that in mice, specific inhibition of myostatin and activin A both individually caused an increase in muscle mass with myostatin causing a greater increase than activin A. However, when the same experiment was performed in cynomolgus monkeys, specific inhibition of myostatin produced little effect on muscle mass when compared with activin A. They further investigated serum levels of mice, monkeys, and humans to see how they differ and found that activin A levels in humans and monkeys were 3-4-fold higher than those found in mice whereas myostatin levels are 5-18-fold lower. Therefore, targeting activin A specifically may result in a greater increase in muscle mass in humans rather than targeting myostatin.

Targeting activin A as a treatment for muscular dystrophy is one possible therapeutic application of activin A inhibition. Though it is not the causative agent of muscle atrophy inhibition of activin may lead to an increase in muscle mass and muscle hypertrophy, thus combating the disease indirectly. Clinical trials that inhibit activin A for this purpose have already shown some success with Campbell reporting treatment of Duchenne muscular dystrophy patients with the multi-TGF- $\beta$  ligand inhibitor ACE-031 (ActRIIB-ECD-IgG1 fusion) resulted in an increase in the distance of patients completing a six-minute walk test even 2-4 weeks after treatment was stopped. However, it is important to note treatment was discontinued after the second dosage due to concerns about the potential side effects of inhibiting multiple TGF- $\beta$  ligands. Therefore a similar approach with a more specific inhibitor is desirable.<sup>156</sup>

#### 1.5.4 Activin A and Cancer

The role of activin A in cancer biology is complex. In healthy individuals, activin A acts as a tumour suppressor,<sup>157,158</sup> however in cancer it has been implicated as having both oncogenic and anti-oncogenic properties depending on the cellular context. Numerous cancers have been associated with either elevated or suppressed signalling, including breast, prostate, ovarian, lung, and colorectal, and oncogenic signalling can occur through both activin's canonical and non-canonical signalling pathways.<sup>159-164</sup> As such, it is important to take into account many factors when looking into the role of activin A in a specific cancer.

The complex nature of activin A's cancer biology can be demonstrated in breast cancer. Bashir *et al.* used qPCR and immunohistological analysis to demonstrate that activin A expression is highly upregulated in the tumour cells of advanced breast cancer patients compared to normal tissue

samples (11-fold increase) and an increase in pSmad2/3 is also observed (3-fold increase in pSmad2 detected, 2-fold in pSmad3).<sup>159</sup> When applied to the cell line MDA-MB-231, this in turn promoted anchorage independent cell growth, induced epithelial-mesenchymal transition of cancer cells, and promoted cell migration and invasion of cancer cells – all key steps in cancer metastasis. By contrast, in a different breast cancer cell line (T47D) Burdette *et al.* report that activin acts as a tumour suppressor, inhibiting cell proliferation promoting cell cycle arrest.<sup>165</sup> The story is equally complex in prostate cancer where activin is observed to inhibit cell growth prostate cancer yet appears to promote prostate cancer metastasis.<sup>160,166,167</sup> This suggests that the role of activin A in cancer depends on the specific cancer in question, the stage of cancer, and even the cell type involved. Therefore, inhibition of activin may result in tumour suppression in one patient whereas in another it may increase tumorigenesis and metastasis.

The presence of hormonal factors appears to have a large effect on activin's role in cancer cells. In prostate cancer it has been observed that the presence of androgens can eliminate the anti-proliferative effects of activin A.<sup>168</sup> In addition activin contributed towards androgen production which is one of the key drivers of castration resistant prostate cancer (androgens levels also increased levels of activin A expression). In breast cancer a similar interplay is present, however with activin A suppressing proliferation though inhibiting oestrogen production and oestrogen inhibiting expression of activin.<sup>169</sup> Therefore it is important to consider the tumour microenvironment when analysing activin's role.

Resistance to or suppression of activin A signalling is a hallmark of some cancers. Resistance may come at any part of the signalling pathway and may be due to a number of factors. In certain breast cancer cell lines activin resistance was conferred by a reduction in expression of its receptors, thus lowering the total amount of activin signalling the cell undergoes.<sup>170</sup> In addition, Gold *et al.* observed that the anti-proliferative activity of activin A in LNCaP prostate cancer cells was reduced in the presence of the related growth factor activin C suggesting up regulation of this growth factor could induce tumorigenesis in the prostate.<sup>88</sup> This has been hypothesised to be the case with follistatin too, with increased follistatin levels being associated with cell proliferation in the same prostate cancer cell lines.<sup>171</sup> It is also the case that mutations in activin A may alter its signalling. Tournier *et al.* observed that in an ovarian cancer patient, a mutation in the inhibin  $\beta_A$  subunit in the tumour cells decreased the amount of secreted activin A relative to inhibin A.<sup>172</sup> This likely resulted in a reduction in activin A signalling and increased proliferative potential.

### 1.5.5 Activin A and Fibrodysplasia Ossificans Progressiva

Fibrodysplasia Ossificans Progressiva (FOP) is an extremely rare autosomal dominant disease. It is caused by a gain-of-function mutation on chromosome 2q23-24 which often occurs spontaneously in the gametes prior to fertilisation.<sup>173</sup> This region contains the ACVR1 gene which codes for the type I receptor ALK2, which is generally associated with BMP signalling, phosphorylating SMADS 1, 5 and 8. Surprisingly, in FOP patients, mutations in the intracellular GS region or kinase domain of ALK2 cause it to become responsive to activin A signalling. Several mutations have been reported however the most common by far is R206H that occurs in >95 % of FOP patients (Table 1).<sup>174,175</sup> Incidentally, mutations on the same residues have also been reported in approximately 25 % of childhood cases of diffuse intrinsic pontine glioma, suggesting activin A may also play a role in this type of cancer too.<sup>176</sup>

**Table 1: ALK2 mutations associated with fibrodysplasia ossificans progressiva.**

<b>Mutation</b>	<b>Location</b>
P197_F198del insL	GS Region
R206H	GS Region
Q207E	GS Region
G328E	Kinase Domain
G328R	Kinase Domain
G328W	Kinase Domain
G356D	Kinase Domain
R375P	Kinase Domain

Any injury or inflammation to fibrous tissue in a FOP patient results in BMP-like signalling in the tissue due to the physiological release of activin A (which now activates ALK2). This results in the gradual ossification of the muscular system, resulting in the progressive immobilisation and eventual death of the diseased individual.<sup>177</sup> There is currently no cure and treatment options are limited. Surgery can be used to remove ectopic bone, but the resulting inflammation from surgery increases bone formation and thus only aggravates the condition in the long run.

The mechanism through which activin A becomes able to signal through ALK2 is at first glance unclear. Mutations causing FOP occur in the intracellular kinase domain of ALK2, whereas activin A is an extracellular growth factor interacting with the extracellular domain of the receptor only. Furthermore, the extracellular domains of ALK2 and ALK4 (the most common type I receptor for activin A) have low sequence similarity (sequence similarity = 33 %, identity = 20 %), and though activin A is observed to bind to ALK2, it usually does not signal through it.

However, Botello-Smith *et al.* suggests that activation occurs not due to a change in the binding properties of activin A or ALK2, but due to the mutations inducing leaky signalling. The key interaction in ALK2 that prevents ATP binding to the kinase domain in the basal state is a salt bridge blocking the ATP binding site. Upon activation by a type II receptor, this salt bridge is destabilised allowing ATP to co-ordinate and the kinase domain to phosphorylate downstream proteins. Through molecular dynamics simulations, this group observed that mutations associated with FOP disrupt this salt bridge to break due to the breakdown of an allosteric hydrogen bonding network that aids salt bridge formation thus activating the kinase domain in the absence of ligand and the type II receptor. Additionally, several mutations can also bypass inhibition by FKBP12, an interacting partner of the TGF- $\beta$  family type I receptors that binds to the GS region and acts as a secondary mechanism to prevent leaky signalling.<sup>178</sup> However, despite this ALK2 does not signal in the absence of the type II receptor. Bagarova *et al.* showed that BMP2 mediated ALK2 signalling\* and ALK2<sub>R206H</sub> / ALK2<sub>Q207D</sub> signalling was significantly reduced in the absence of BMPRII and ActRIIA.<sup>180</sup> Incidentally, restoration of these receptors resulted in the restoration of ALK2 signalling, though in the case of ALK2<sub>R206H</sub> / ALK2<sub>Q207D</sub> this was independent of ligand binding. It has been shown that the affinity of the type I receptors for the SMADs is greatly enhanced if the receptor is activated by the type II receptor.<sup>181</sup> Thus, this observation may explain the role of activin A in FOP, with formation of the heteromeric ActA-ActRIIA/B-ALK2 complex ligand-receptor complex localising ActRIIA/B to ALK2 resulting in the induction of signalling.

Though this offers a good explanation as to why downstream phosphorylation occurs, it still does not explain why ALK2 requires binding of activin A to induce signalling if the receptor is constitutively active. This can be explained by the requirement of SMAD proteins. It has been shown that the affinity of the type I receptors for the SMADs is greatly enhanced if the receptor is activated by the type II receptor.<sup>180,181</sup> Therefore, binding of activin is required to form the heteromeric ligand-receptor complex such that the type I receptor activation can occur and signalling be induced.

## 1.6 Inhibitors of TGF- $\beta$ Family Signalling

The specific inhibition of activin A, and more generally the ligands of the TGF- $\beta$  family, is highly desirable due to potential research and clinical applications. Numerous methods have been attempted to inhibit members of the TGF- $\beta$  family targeting formation of ligand-receptor complexes and phosphorylation of the Smad proteins. In this section several of these methods will be discussed. For

---

\* In addition to activin A, the leaky signalling of ALK2<sub>R206H</sub> makes it hypersensitive to BMP signalling. As such, several BMP ligands, including BMP2, have been shown to produce similar results to activin A in biochemical assays.<sup>179</sup>



ease, all inhibitors mentioned by name in this section are listed in Table 2. A figure indicating how the signalling pathway is targeted is shown in Figure 4.

**Table 2: Selected inhibitors of the TGF- $\beta$  family. Where inhibitors have been reported for numerous targets, preference will be given to the activin A system.**

Inhibitor	Class	Target	$K_d$ / nM	IC <sub>50</sub> / nM (Reporter Assay)	Reference
Anti-activin A antibodies	Antibody	Activin A	<0.005	<0.08	Gromada <i>et al.</i> <sup>182</sup>
Bimagrumab	Antibody	ActRIIA	0.973	0.58	Forvan <i>et al.</i> <sup>183</sup>
		ActRIIB	0.016	0.58	
ActRIIA–mFc	Ligand Trap (Receptor)	Activin A	0.003	0.275	Pearsall <i>et al.</i> <sup>184</sup>
ActRIIB.hFc	Ligand Trap (Receptor)	Activin A	0.036	1.63	Sako <i>et al.</i> <sup>185</sup>
Follistatin-288-Fc	Ligand Trap (Follistatin)	Activin A	0.196	0.070	Castonguay <i>et al.</i> <sup>186</sup>
Modified ActA	Ligand Trap (Pro-domain)	Activin A	–	5	Chen <i>et al.</i> <sup>187</sup>
AT propeptide	Pro-domain Mutant	Activin A	–	2.6	Makanji <i>et al.</i> <sup>188</sup>
Peptide 7	Peptide (Pro-domain)	Myostatin	29.7	3500	Takayama <i>et al.</i> <sup>189</sup>
Peptide 11	Stapled Peptide (Pro-domain)	Myostatin		260	Rentier <i>et al.</i> <sup>190</sup>
Pro- $\alpha$ peptide	Peptide (Pro-domain)	Inhibin A*	47.5	255	Walton <i>et al.</i> <sup>191</sup>
		Inhibin B*	28.3	150	
nsc119889, 910, 911, 913, 915	Small Molecule	TGF- $\beta$ 1 or T $\beta$ RII	–	~10000	Burmester <i>et al.</i> <sup>192</sup>
NUCC-555	Small Molecule	Activin A	–	5300	Zhu <i>et al.</i> <sup>193</sup>
Galunisertib	Small Molecule	ALK5	–	221	Yingling <i>et al.</i> <sup>194</sup>
LDN-214117	Small Molecule	ALK2		100	Mohedas <i>et al.</i> <sup>195</sup>
ar2mini	Peptide (Phage Display)	ActRIIB	1800	5000-12000	Sakamoto <i>et al.</i> <sup>196</sup>

\* Denotes glycosylated form

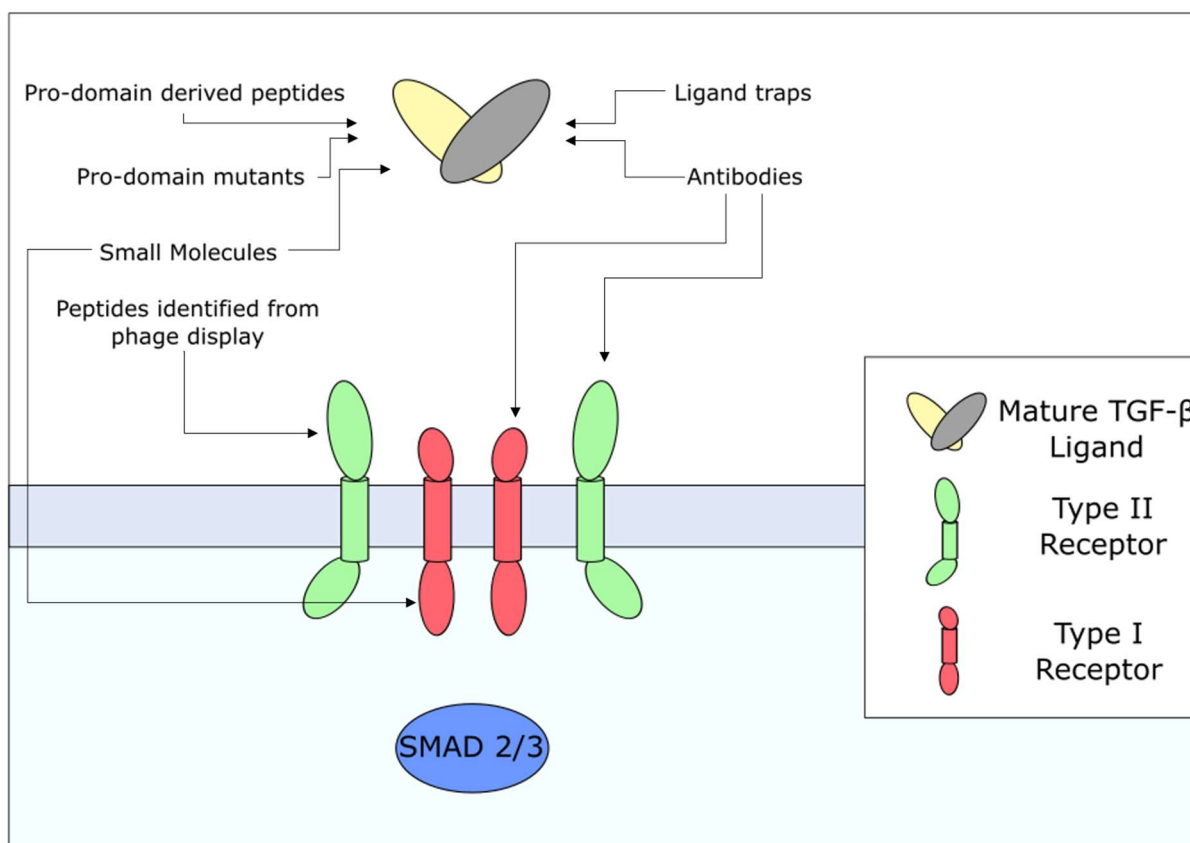


Figure 4: Common inhibitors TGF- $\beta$  family signalling. Here, a selection of common methods of inhibiting TGF- $\beta$  family signalling are shown. The specific protein targets for each method are also listed.

The most common target for TGF- $\beta$  family inhibitors is the complexation of the ligand with its type I and type II receptors. By disrupting formation of this complex, the type II receptor is unable to phosphorylate the type I receptor thus a phosphorylation cascade cannot be triggered. The extracellular nature of this complex is advantageous for inhibitor development as any potential inhibitors would not have to cross the cell membrane. There are two main protein targets when using this strategy – the extracellular domains of the receptors and the growth factors themselves. Antibodies have been described that use both these approaches with activin A, targeting either mature activin A, ActRIIA/B and ALK4.<sup>197</sup> When compared to other methods of inhibition antibodies are highly specific, often binding their target with picomolar affinity, and have low off target toxicity, however they are commercially expensive to produce, requiring mammalian expression systems. Anti-activin A antibodies were first described in 1996 by Funaba *et al.* who used fragments of activin A to generate polyclonal antibodies.<sup>198</sup> The same group later used chickens to develop an antibody against mature activin A that could inhibit activin's ability to stimulate FSH secretion from pituitary cells and suppress the development of mouse embryos.<sup>199</sup> Since then a number of patents have been filed for neutralising antibodies that target activin A (Gromada *et al.* patent filed for antibodies with  $K_d$  values  $< 5$  pM and  $IC_{50}$  values  $< 80$  pM).<sup>182,200</sup> Antibodies that target activin's receptors often target the

extracellular domains. Bimagrumab, an anti-ActRIIA/B antibody binds to both their extracellular domains with a preference for ActRIIB (ActRIIA  $K_d$  = 973 pM, ActRIIB  $K_d$  = 16 pM) through the same binding site as activin A and other TGF- $\beta$  ligands.<sup>183</sup> Currently, antibodies make up the vast majority of commercially available inhibitors of activin A.

A similar approach to antibody inhibitors has been to develop ligand traps. These proteins are designed such that a dimeric Fc fragment of antibody is fused to a “trap” that can form an inactive complex with the ligand in question – either the extracellular domain of a receptor, an extracellular inhibitor such as follistatin, or a ligand’s pro-domain. While affinities of the individual binders can be relatively weak, dimerization through Fc part provides avidity and thus increases the apparent affinity for the target(s). This approach has been used for inhibition of activin A using ActRIIA, ActRIIB and follistatin as the traps as well as activin’s pro-domain.<sup>184–187</sup> Both ActRIIA/B and follistatin traps have been shown to inhibit activin A signalling with high efficacy (ActRIIA–mFc  $IC_{50}$  = 275 pM, ActRIIB.hFc  $IC_{50}$  = 1.63 nM, Follistatin-288-Fc  $IC_{50}$  = 70.3 pM) however both have a broad ranging specificity and also inhibit activin B, myostatin and GDF11. However, this broad specificity has been advantageous when targeting muscle growth due to the complimentary roles of activin A and myostatin, with both ActRIIA/B (Sotatercept and Luspatercept) and follistatin (ACE-083) traps currently in clinical trials.<sup>201–204</sup> Pro-domain based ligand traps have an advantage over the other two classes in that due to the poor conservation between pro-domains, they are more likely to be specific against a given ligand. This was the case in Chen *et al.* who developed ligand traps based on the pro-domain of activin A ( $IC_{50}$  = 5 nM) that had >20-fold specificity for activin A over activin B and did not inhibit myostatin or GDF11 signalling at the concentrations tested.<sup>187</sup>

The pro-domains of TGF- $\beta$  family members represent a good starting point for developing inhibitors. Pro-domains bind their mature growth factors with relatively high affinity (<1  $\mu$ M),<sup>104,107,189</sup> have good solubility to prevent the mature growth factors from precipitating out and in some cases can actually inhibit signalling of the mature ligand.<sup>104,205</sup> This is the case for activin A, whose pro-domain binds its mature domain with an affinity of 5 nM (as determined via biolayer interferometry). However, as pro-activin A is processed to form an active complex, the pro-domain only appears to inhibit activin A signalling at high concentrations. In addition to ligand traps, a pro-domain approach for activin A has already been attempted. Makanji *et al.* created a fusion of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix of activin A to the shoulder region of TGF- $\beta$ 1 pro-domain (AT Propeptide).<sup>188</sup> As the shoulder region of TGF- $\beta$ 1 is a disulfide linked dimer this resulted in a dimeric inhibitor, not dissimilar from a ligand trap, that was a potent inhibitor of activin A and B signalling ( $IC_{50}$  10.3 nM).

Another pro-domain based method that has been utilised to generate inhibitors is to use pro-domain based peptides. This approach has seen success for both myostatin and inhibin. In the case of myostatin, the pro-domain is a highly potent inhibitor of myostatin signalling with an  $IC_{50}$  value of 0.9 nM.<sup>108</sup> Based on this, Jiang *et al.* identified the region responsible for this inhibition to be in the N-terminus between residues 42 and 115.<sup>206</sup> Takayama *et al.* subsequently developed a synthetic 23 residue high affinity minimum peptide ( $K_d = 29.7$  nM via SPR) based on the pro-domain of mouse myostatin that could inhibit myostatin signalling with an  $IC_{50}$  value of 3.5  $\mu$ M (Peptide 7).<sup>189</sup> This sequence has since been identified as corresponding to the pro-domain's  $\alpha$ 1-helix and the start of the loop region.<sup>108</sup> Since then the same group has sequentially optimised this inhibitor through acetylating the peptide ( $IC_{50} = 1.2$   $\mu$ M), used SAR to mutate key residues ( $IC_{50} = 320$  nM), and introduced stapling to the peptide ( $IC_{50} = 260$  nM) (Peptide 11).<sup>190,207–209</sup> Interestingly, Fc-peptide fusions developed by Ohsawa *et al.*, similar to those reported in Takayama *et al.*, bound not only to mature myostatin and GDF11 but also to the type I and type II receptors.<sup>205</sup> This suggests interactions between the pro-domain and receptors may play an important role in the initiation of myostatin signalling.

A similar approach was attempted by Walton *et al.* who generated an inhibitory peptide for inhibin A and B based on the pro-domain of the inhibin  $\alpha$  subunit (inhibin A  $IC_{50} = 255$  nM, inhibin B  $IC_{50} = 150$  nM).<sup>191</sup> This peptide was highly specific, interacting only with the inhibins, and not with either activin A or activin B, indicating the peptide only interacted with the  $\alpha$  subunit of the mature inhibin dimer. Pepinsky *et al.* also attempted a peptide approach with GDF 11. Unlike myostatin and inhibin, the GDF11 pro-domain derived peptide did not inhibit GDF11 signalling at the concentrations used (though these concentrations were significantly lower than those used with the myostatin peptides by Takayama *et al.*<sup>189</sup>). Though no inhibition was observed, the minimal peptide dramatically increased the solubility of mature GDF 11 suggesting a potential growth factor delivery application of pro-domain derived peptides.<sup>106</sup>

Small molecules are the other major category of TGF- $\beta$  family inhibitors. Small molecules have been reported that disrupt ligand-receptor complex formation. Burmester *et al.* conducted a high throughput screen against binding of TGF- $\beta$ 1 to solubilised T $\beta$ RII which led to the identification of five small molecules that block this interaction (Most potent antagonist  $ED_{50} =$  approximately 10  $\mu$ M).<sup>192</sup> In addition Zhu *et al.* used virtual high throughput *in silico* screening against the crystal structure of mature activin A to identify two inhibitors that bind in the hydrophobic pocket of putative type I receptor binding site, the most potent of which, NUCC-555 had an  $IC_{50}$  value of 5.3  $\mu$ M.<sup>193</sup> However the most common target for small molecule inhibitors is the kinase domains of the type I receptors. Targeting these kinase domains has several advantages when compared with other proteins and domains involved in the signalling pathway. They have an obvious binding pocket, a substrate that can

be used as a scaffold for inhibitor development, and many assays exist for testing kinase specificity thus off target toxicity can be more easily determined. Numerous inhibitors have been developed for the kinase domains of the type I receptors, including Galunisertib for ALK4 and ALK5, and LDN-214117 for ALK2.<sup>195,210–214</sup> These molecules are typically based on either a dihydropyrrolopyrazole or an imidazole scaffold and have a high specificity for the receptor they are targeting – even being able to distinguish between the different type I receptors. However, on-target toxicity is often a concern due to multiple ligands signalling through the same receptor.<sup>215</sup> Despite this, the potent ALK4 / ALK5 inhibitor Galunisertib ( $K_i = 86$  nM,  $IC_{50} = 221$  nM in cell based reporter assay) has shown success in clinical trials with no on-target toxicity or dose limiting event observed.<sup>194,213</sup>

Alongside the methods already discussed, a number of novel methods of inhibition have been described. Harrison *et al.* generated mutants of mature activin A that could form a complex with activin A's type II receptors but not ALK4, thus blocking the type II receptors and preventing signalling from being initiated.<sup>216</sup> In addition, Sakamoto *et al.* generated peptide inhibitors for the extracellular domain of ActRIIB through phage display.<sup>196</sup> Interestingly several of the peptides displayed different ligand specificities, with one inhibiting a broad range of ligands (activin A, myostatin, GDF11, BMP9) that signal through ActRIIB ( $IC_{50}$  values = 5-12  $\mu$ M) whereas another only inhibited myostatin signalling ( $IC_{50}$  value = 40  $\mu$ M). The mode of binding of the peptides or the reason for this specificity however was not speculated on. In the larger TGF- $\beta$  family, other approaches have included the use of antisense oligonucleotides, small molecule inhibitors that bind sequence specifically to a gene promoter, and upregulation of I-Smads through gene transfer.<sup>217–219</sup>

Specificity is an important concept when designing inhibitors for the TGF- $\beta$  family. Due to the sequence and structural similarity of the ligands and the promiscuity of the receptors, it can be difficult to achieve specific inhibition that doesn't have off target effects. This is especially true when using inhibitors that are based on or affect either the type I or type II receptors. In two clinical trials that used receptor-based ligand traps to target activin A signalling, numerous side effects were observed including bone pain and gingival bleeding that is likely caused by inhibition of other TGF- $\beta$  family members such as BMP9.<sup>220,221</sup> Therefore, to minimise such effects it may be the case that targeting the growth factors themselves is the best approach overall.

## 1.7 Aims

The key question this thesis aimed to address is “is it possible to generate specific inhibitors for activin A signalling through targeting the mature activin A dimer alone?”. To answer this, the objectives of this thesis were as follows:

1) Identify and characterise fragments that bind to mature activin A (Small Molecule Approach)

The identification of fragments that bind to mature activin A could be an important potential starting point to developing small molecule inhibitors of activin A signalling. To achieve this, I crystallised mature activin A as previously described by Harrington *et al.*<sup>77</sup> and optimised the conditions for fragment soaking. I then performed a fragment screen via X-ray crystallography utilising the XChem platform at the Diamond Light Source to identify binding molecules. I further validated several of the binding molecules through X-ray crystallography and derivatised them to generate binders with more optimised interactions.

2) Identify and develop peptide inhibitors based on the key sites of interactions between activin A's pro-and mature domains (Peptide Approach)

In order to generate inhibitors based on the pro-domain of activin A, it is important to first identify the key sites of interaction between the pro- and mature domains. To this end, I developed a series of peptide fusions containing different epitopes of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif and used biolayer interferometry (BLI) to determine which epitopes are crucial for binding. I then determined the affinity of these epitopes for the mature domain using BLI and isothermal calorimetry (ITC) and attempted to identify the key interacting residues, before investigating the inhibitory properties of the peptide fusions via a Smad reporter luciferase assay.

3) Optimise pro-domain derived peptide inhibitors through dimerization (Dimeric Approach)

As the mature activin A dimer binds two pro-domains, it is possible that dimerization of any pro-domain derived peptide inhibitor will provide a route to optimization and increased potency. As such, I utilised the dimerization and docking domain of the RII $\alpha$  subunit of protein kinase A, and disulfide bonds to induce dimerization of the inhibitory peptides. Similarly to their monomeric counterparts, I characterised their binding properties to mature activin A via ITC, and effect on activin A signalling through luciferase assays in an iterative process to generate a highly potent inhibitor.

#### 4) Screen pro-domain derived peptide inhibitors for specificity against other members of the TGF- $\beta$ superfamily (Dimeric Approach)

Due to the structural and sequence similarity of TGF- $\beta$  ligands and receptor promiscuity, it is important to determine the specificity of any inhibitor generated in this thesis for activin A. Accordingly, I determined an appropriate selection of TGF- $\beta$  ligands to generate a wide and robust screen. I then used the selected ligands in a luciferase assay to determine the effect of the most potent peptide inhibitor developed for activin A on their signalling.





---

# Small Molecule Approach

---

## 2.1 Introduction

Small molecule drugs refer to molecules with low molecular weights (usually less than 1 kDa) that have therapeutic properties, usually through agonism or antagonism of signalling or activation or inhibition of a specific biological process. Overall, they account for over 90% of all commercially available drugs and have several advantages over protein-based therapeutics, namely they are cheaper, easier to manufacture, and usually have a longer shelf life.<sup>222,223</sup> The identification of a hit compound is the first step in generating a small molecule therapeutic agent for a target protein. In some cases, a rational starting point is present (such as the substrate analogue of an enzyme), however for many targets this is not the case. As such, large numbers of compounds are often screened against a protein to identify binding molecules in a high-throughput screen. The nature of these molecules is often diverse, with a range of molecular weights (typically between 300-1000 Da) and physicochemical properties, allowing for a wide coverage of chemical space. However, major limitations can arise from unoptimized fits of the molecule into the binding site, meaning interacting motifs can be overlooked due to steric or electrostatic repulsion of the large molecule. An alternate method that aims to overcome this limitation is fragment screening. In contrast to high-throughput screening, fragment screening is a low-throughput screen that aims to still cover a large chemical but with significantly fewer molecules. This is achieved through the use of smaller molecules (typically around 200 Da) that often consist of a one or two motifs.<sup>224</sup> The hit rates varies from protein to protein but is often high, and multiple hits in a binding site are common. Motifs from these hits can then be either grown, merged, or linked together to create an optimised molecule for a specific binding site. As hits can be combined in numerous different ways, this makes the exploration of chemical space efficient despite the library size. Major limitations of fragment screening can arise from a high hit rate making it difficult to select which compound to further optimise, or weak interactions. Due to their low size and low number of interacting motifs, fragments usually only interact with a protein weakly making detection problematic.

X-ray crystallography is one method through which fragment screening can be conducted. This is advantageous compared to other methods as it allows for structural information to be obtained allowing for identification of the binding site of a fragment the interacting residues. Additionally, it is more sensitive than many other screening techniques allowing for fragment binding to be more easily detected.<sup>225</sup> However, crystallographic screening is time and resource intensive and requires that the protein of interest to be crystallisable. The crystals need to have unliganded, accessible target site, diffract to high resolution, and they need to be robust so as not to be destroyed upon high

concentration fragment soaking. Furthermore, it is still difficult to detect weakly binding fragments. This is because the electron density observed in maps is averaged across the entire crystal lattice. If a molecule binds strongly, it occupies a large percentage of sites in the crystal lattice, generating a large signal in electron density maps. However, if a molecule binds weakly, the ligand bound state only corresponds to a small percentage of states in the crystal, thus the signal is weak and difficult to discern from noise. Furthermore, if a ligand binds in a number of conformations, averaging will lead to a weak overall signal. As such these signals are often difficult to interpret or go undetected when analysing electron density maps. The XChem platform developed at the Diamond Light Source by Frank von Delft aims to overcome the problems of detecting weak fragment hits. This methodology, implemented in software package PanDDA, uses more sophisticated analysis of crystallographic data to detect weakly bound fragments. The PanDDA program aligns all “event” electron density maps (those obtained from fragment soaks) to the map generated from a ground state map and structure (in this case a map and structure of the protein without a bound ligand). It then creates an average electron density map from the entire data set and assigns a Background Correction Density factor to it. This corrected average map has a much lower signal to noise ratio than any individual map obtained and thus, when it is subtracted from each event map, can reveal the presence of weak density (for further explanation of the methodology, see Pearce *et al.*).<sup>226</sup> As such it can identify weakly binding ligands or uncommon states that are not normally detectable. This method has already been shown to have success in identifying small molecule binders for several proteins including the deubiquitinase OTUB2, the pyrophosphatase NUDT7 and the histone deacetylase HDAC6.<sup>227,228</sup> Therefore, I decided to apply this methodology to develop small molecule inhibitors of mature activin A.

The choice of activin A as a target for small molecule screening at first glance is a challenging one. As activin A is an extracellular signalling protein, its primary interactions are protein-protein interactions. Disrupting these interactions is often difficult due to large area of interaction and lack of obvious binding pockets that can be targeted.<sup>229</sup> Plus, sequence and structural similarity between the TGF- $\beta$  family growth factors means that achieving specific binding of any small molecule will be challenging. However, this system does have some advantages. Firstly, the structure of mature activin A at 2.0 Å resolution is known, having been determined via x-ray crystallography by Harrington *et al.*<sup>77</sup> Analysis of this structure shows the presence of a hydrophobic binding groove in the putative type I receptor binding site. This site was used by Zhu *et al.* to identify a number of small molecule binders via *in silico* screening that were subsequently developed into inhibitors (most potent  $IC_{50} = 5.3 \mu M$ ).<sup>193</sup> As this site is also the putative type I receptor binding site, it represents a good target site for developing inhibitors. In addition to this site, analysis performed by Chris Radoux (University of Cambridge, unpublished work) using methodology as described in Radoux *et al.*<sup>230</sup> on the mature activin A

structure identified the presence of a potential second binding site at the dimer interface (Figure 5). This was through the identification of two hydrophobic hotspots close to the dimerization cysteine. Greenwald *et al.* have suggested that flexibility about the dimerization cysteine may play an important role in activin A signalling.<sup>103</sup> As such, a molecule that binds here may be able to conformationally lock the activin A dimer, thus antagonising its signalling. These factors considered, I decided to conduct a fragment screen on mature activin A.

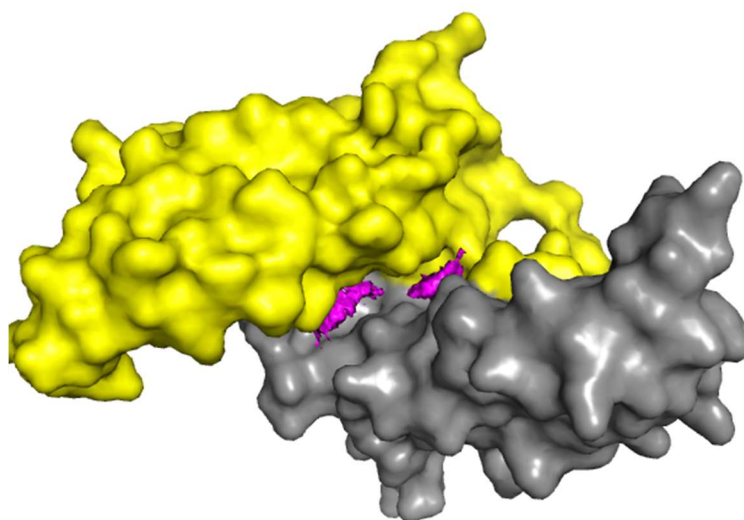


Figure 5: Binding hotspots on the surface of mature activin A. Here mature activin A is shown in yellow and grey and the hotspots in pink. Figure adapted from work performed by Chris Radoux.

In this section I will describe the fragment screening of mature activin A crystals using the XChem platform. I will describe how crystallisation conditions were optimised for fragment screening before identifying small molecule binders through the screening process. I will then discuss how several of the binding molecules were further validated via x-ray crystallography and how their in-solution binding was analysed via ligand-based NMR spectroscopy.

## 2.2 Optimization of Crystallisation conditions for Fragment Screening of Mature Activin A

In order to conduct a fragment screen on mature activin A, crystallisation conditions first needed to be optimised from those first described in Harrington *et al.* for soaking.<sup>77</sup> Crystallisation under the original conditions (1.2-1.5 M  $(\text{NH}_4)_2\text{SO}_4$ , 2-3 % PEG300, 100 mM HEPES, pH 7.4-7.8, 6 mg/mL activin A in 20 % acetonitrile) did not contain DMSO (the solvent most commonly used in fragment libraries). The addition of DMSO to crystals has been reported to result in osmotic shock in a number of systems which leads to degradation of the crystal. In order to minimise such a shock, the effect of DMSO on crystallisation was investigated. 24-well screens were set up, where the concentrations of  $(\text{NH}_4)_2\text{SO}_4$ ,

PEG 300, and mature activin A were varied against an increasing concentration of DMSO (100 mM Hepes pH 7.4 was present in all conditions). Crystals grew in most wells, with the protein first precipitating before crystallising from the precipitate. Nucleation was observed to increase with increasing concentration of  $(\text{NH}_4)_2\text{SO}_4$ , PEG 300, DMSO and activin A. Of these, DMSO had the most potent effect with unsuitably small crystals being observed at concentrations higher than 2 %. The only condition that produced a low number of crystals of reasonable size was 1.65 M  $(\text{NH}_4)_2\text{SO}_4$ , 4 % PEG 300, 100 mM Hepes pH 7.4, 2 % DMSO, 3.6 mg/mL activin A. A crystal from this condition was tested for diffraction properties.

Due to the concentration of  $(\text{NH}_4)_2\text{SO}_4$  and presence of DMSO in the crystallisation conditions, no additional cryoprotectant was used when harvesting the crystal. The crystal was shot at i03 beamline at Diamond Light Source using X-ray wavelength of 0.9763 Å, diffracting to 2.1 Å (for crystallographic statistics, see Appendix 1). This is similar to the resolution observed by Harrington *et al.* for crystals that do not contain DMSO (2.0 Å)<sup>77</sup> suggesting crystallising activin A in the presence of 2 % DMSO does not significantly affect diffraction quality. Analysis of the structures of DMSO-containing condition and original structure (PDB:2ARV) showed no significant structural differences between the two with a PEG molecule binding in the hydrophobic pocket thought to be the type I receptor binding site (Figure 6). Density corresponding to DMSO was only observed at one site on the protein, in the hydrophobic pocket present on the other subunit where it interacted with W338 and Y403. Based on this it was decided that mature activin A crystals grown in the presence of DMSO were appropriate for fragment screening.

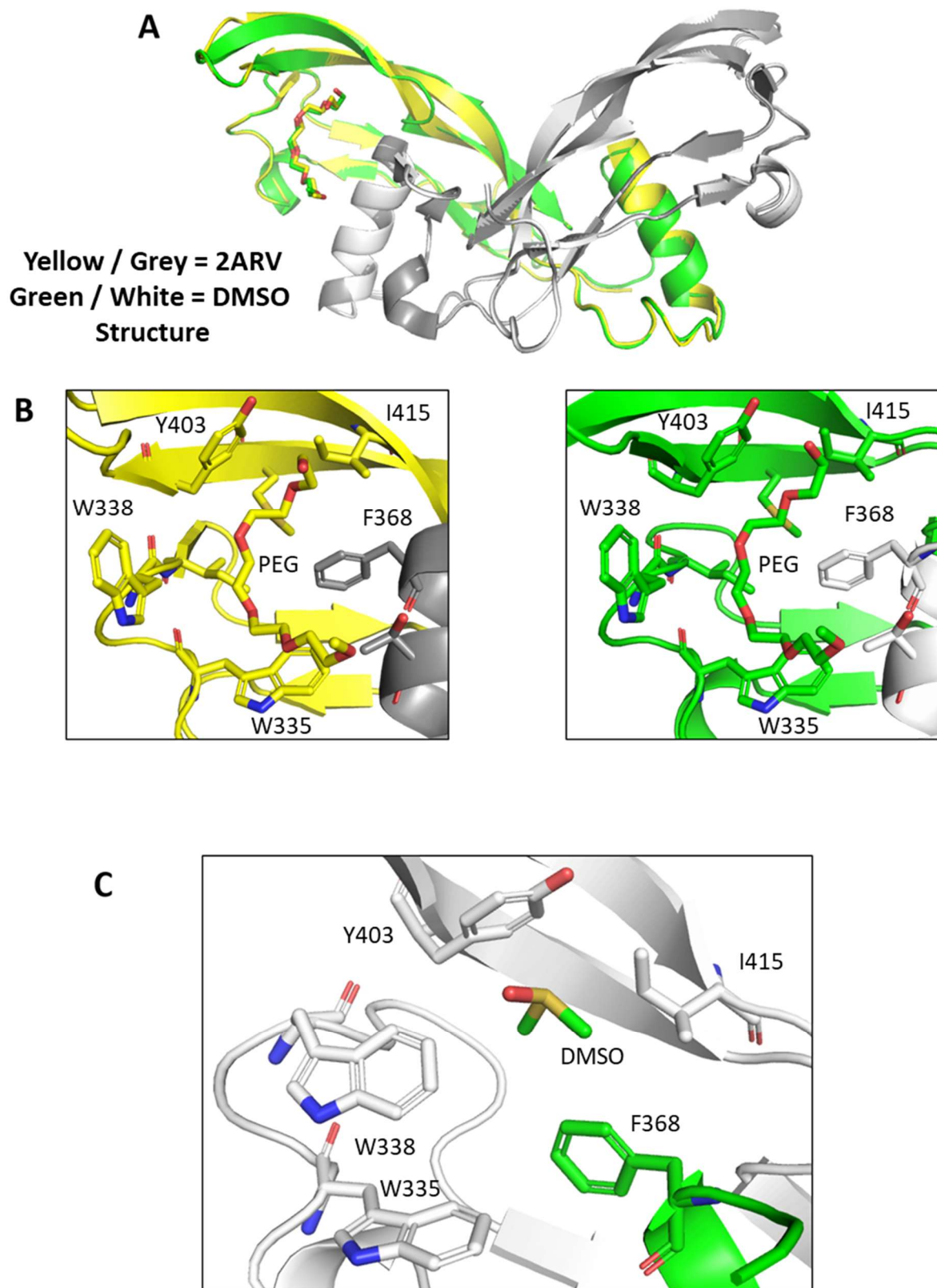
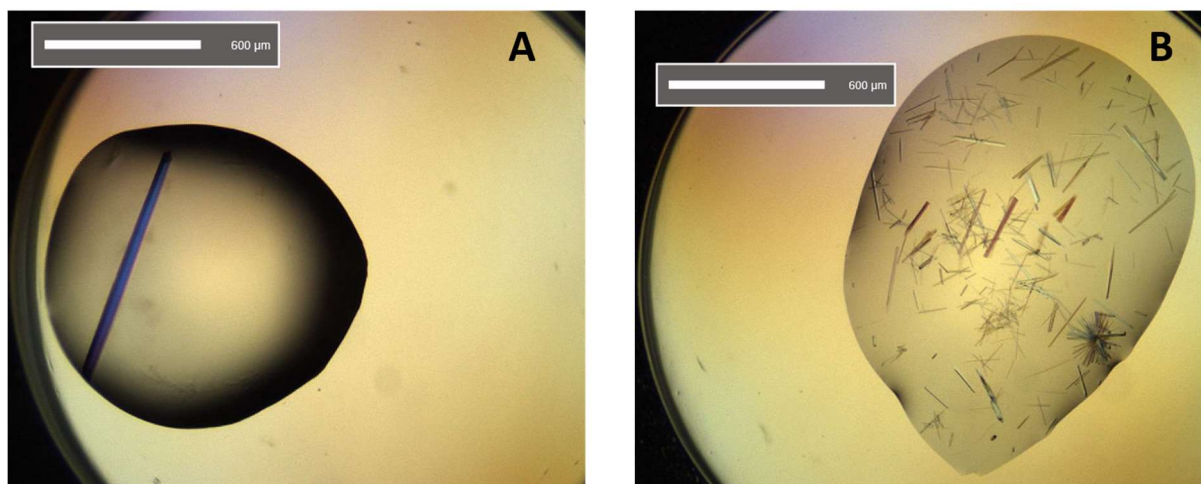


Figure 6: The crystal structure determined for mature activin A grown in the presence of DMSO. Here the structure of the DMSO mature activin A is shown in green (subunit A) and white (subunit B), and the structure of mature activin A (PDB ID: 2ARV) is shown in yellow (subunit A) and white (subunit B) A) Structure of the DMSO mature activin A structure aligned to the structure of mature activin A. B) The PEG binding sites of 2ARV (left) and the DMSO structure (right). Here we see no significant difference between the two. C) The putative type I receptor

*binding site unoccupied by PEG in the DMSO structure. Here we see a DMSO molecule interacting with residues W338 and Y403 on chain A.*

Fragment screening using the XChem platform is designed to use 96-well 3-drop Swiss CI plates. As such, conditions needed to be optimised from those reported above for a 96 well format. A screen varying  $(\text{NH}_4)_2\text{SO}_4$  concentration against DMSO concentration was set up in 96-well 3-drop Swiss CI plates. Conditions were again selected based on crystals being large and uniform with low levels of nucleation occurring. Generic trends were seen similar to those observed in the 24 well format, with increasing  $(\text{NH}_4)_2\text{SO}_4$  and DMSO resulting in increased nucleation. Based on visual inspection of the plate, a crystallisation condition of 1.55 M  $(\text{NH}_4)_2\text{SO}_4$ , 8 % DMSO produced a single needle-like crystal, measuring approximately 700  $\mu\text{m}$  in the longest dimension and uniform thickness (Figure 7). Thus, mature activin A crystals crystallised under the following conditions were taken forward for XChem screening:

- 1.55 M  $(\text{NH}_4)_2\text{SO}_4$
- 100 mM Hepes pH 7.4
- 4 % PEG 300
- 8 % DMSO
- 3.6 mg/mL activin A in 20 % acetonitrile



*Figure 7: Crystallisation of mature activin A in the presence of DMSO. Crystals of mature activin A grown in the 96 well crystal screen as described above. A) A crystal generated in the condition selected for XChem screening. Here we see only one large, unclustered, needle-like crystal is produced. B) crystals produced from a condition containing a higher concentration of  $(\text{NH}_4)_2\text{SO}_4$  and DMSO. Here we see many crystals that are smaller in nature with more clustering.*

## 2.3 XChem Screening of Mature Activin A

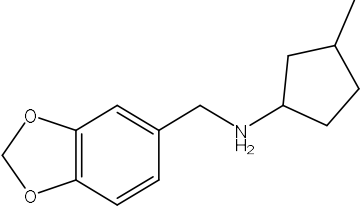
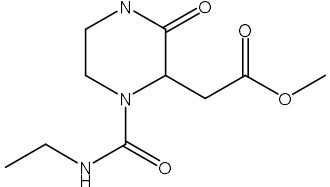
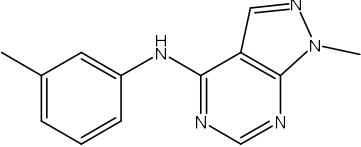
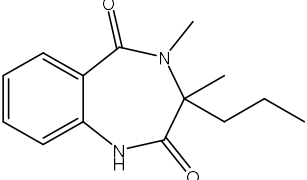
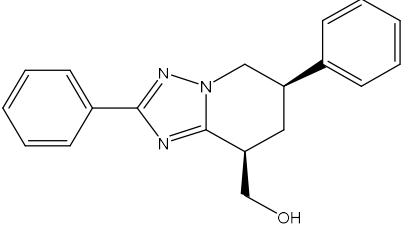
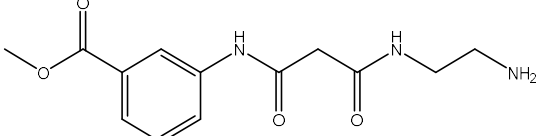
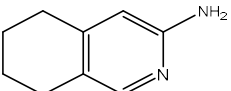
The first step of conducting the fragment screen was to establish the DMSO tolerance of the crystals. This would allow for the determination of the highest possible concentration of fragment a crystal could be soaked with as the library stock solutions are at concentrations of 250 mM or 500 mM compound in DMSO. Crystals drops were soaked with DMSO in increasing increments up to 20 %. Crystals were allowed to equilibrate at room temperature for 1 hour before harvesting and data collection. No significant increase or decrease was seen in resolution trend across the crystals suggesting that concentrations of up to 20 % DMSO did not affect resolution. To maximise concentration of fragment, initial test was done using 5-fold diluted fragments with final concentration of 50 or 100 mM fragment and 20% DMSO. Despite good tolerance for DMSO, in the presence of high concentrations of fragment a large number of crystals dissolved or lost diffraction. To avoid this, the soaking concentration was lower by half which resulted in no loss of crystal number or reduction of diffraction; most crystals diffracted to approximately 2.0-2.4 Å resolution. Thus, a fragment concentration of 25 mM / 50 mM was selected for the full fragment screen.

Crystals were then screened in a similar manner with fragments from different libraries. Libraries were selected based on their coverage of chemical space and the ease with which derivatives and similar molecules could be obtained. In total 345 compounds from three libraries, were successfully screened: Edleris Keychemical Fragments, Kidd 3D and DSPL.<sup>231,232</sup> PanDDA requires a series of refined structures and associated diffraction data to perform its analysis. As such, structures of each successfully diffracting crystal that had higher resolution than 2.5 Å were solved and refined using the Dimple pipeline (part of CCP4)<sup>233</sup> and datasets were analysed via PanDDA. Analysis of the datasets via PanDDA revealed that of these 345 compounds, 13 were detected in two main sites – the putative type I receptor binding site and the interfacial site identified by Chris Radoux (



Table 3, Figure 8). Of these sites, the putative type I receptor binding site was significantly more favoured with 12 of the 13 compounds interacting here. Interestingly, despite being a symmetrical protein with two identical binding sites (be it two putative type I receptor binding sites or one interfacial site with two identical epitopes), binding was only observed in each case in one of the binding sites, likely due to accessibility in the crystal. In addition, different fragments were observed to bind to the same site on different chains of the dimeric protein, further validating these as legitimate hits. This was the case for molecules that bind to the type I receptor binding site, with three hits binding to one chain and nine binding to the other. As such, it may depend on the specific molecule as to which site is more accessible.

**Table 3: Fragment hits identified in XChem screen for mature activin along with the interacting residues for each fragment. TIRB refers to the putative type I receptor binding site whereas IF refers to the interfacial site.**

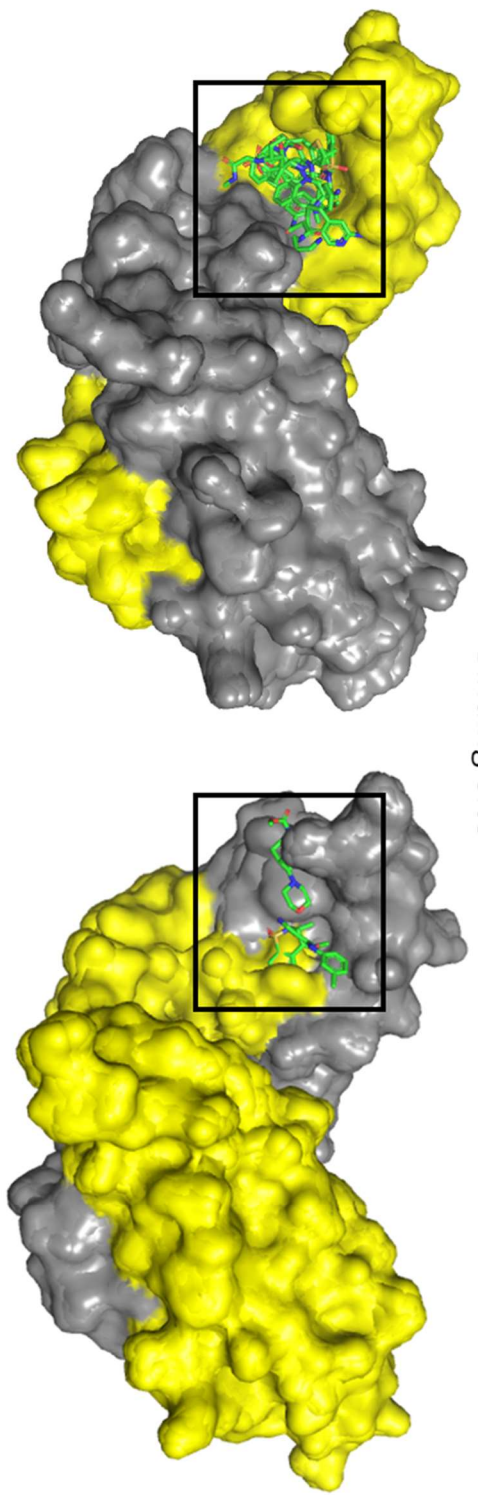
Name	Structure	Binding Site	Interacting Residues Identified in the XChem Screen
FMOPL000043a		TIRB	W335, I339, F368, M401, Y403
FMOPL000578a		IF	S362, P393, T394, L396
PKTTA024495b		TIRB*	W335, W338, F368, M401, Y403, I415
NM466		TIRB	F368, T371, M401, Y403, I415, M418
HLS327		TIRB	W335, W338, I339, F368, T371, M401, Y403, I415
SG5B8		TIRB	S367, F368, T371, Y403, I415
SG6B10		TIRB	F368, Y403, K413, I415, Q416

FMOOA000513a		TIRB	W335, W338, I339, F368, M401, Y403
FMOOA000520a		TIRB*	W335, W338, I339, M401, Y403, I415
FMOOA000553a		TIRB	W338, I339, F368, M401, Y403
FMOOA000565a		TIRB	W335, L366, F368, T371
FMOOA000580a		TIRB	W335, W338, F368, T371, Y403
FMOOA000598a		TIRB*	W338, Y403, D406

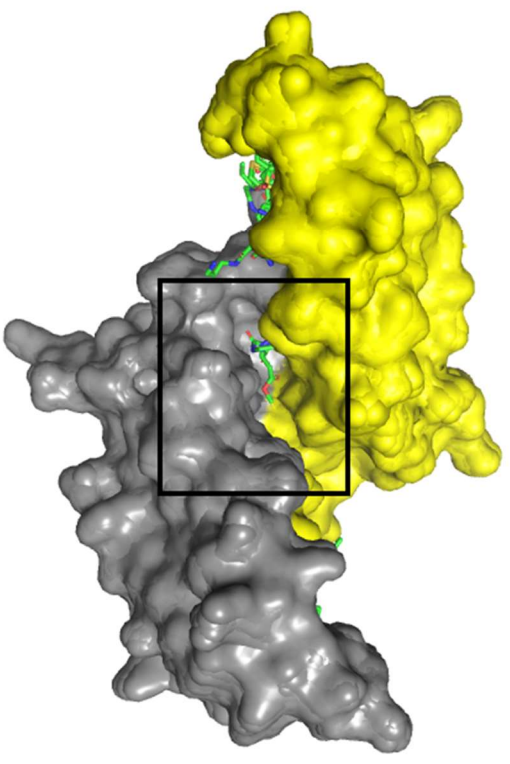
\*denotes binds to the putative type I receptor binding site on the other mature activin A subunit.

A

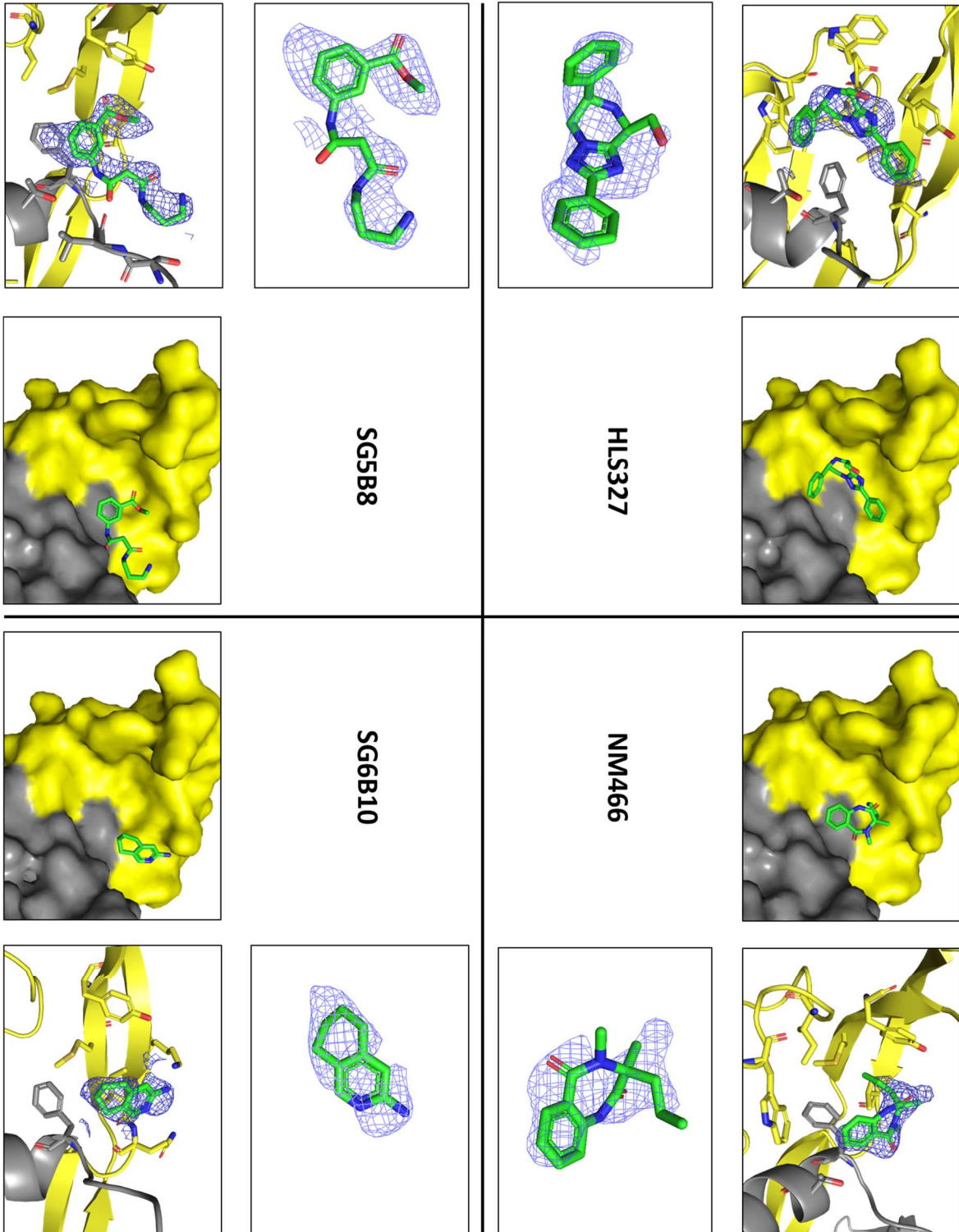
Putative Type I Receptor  
Binding Site

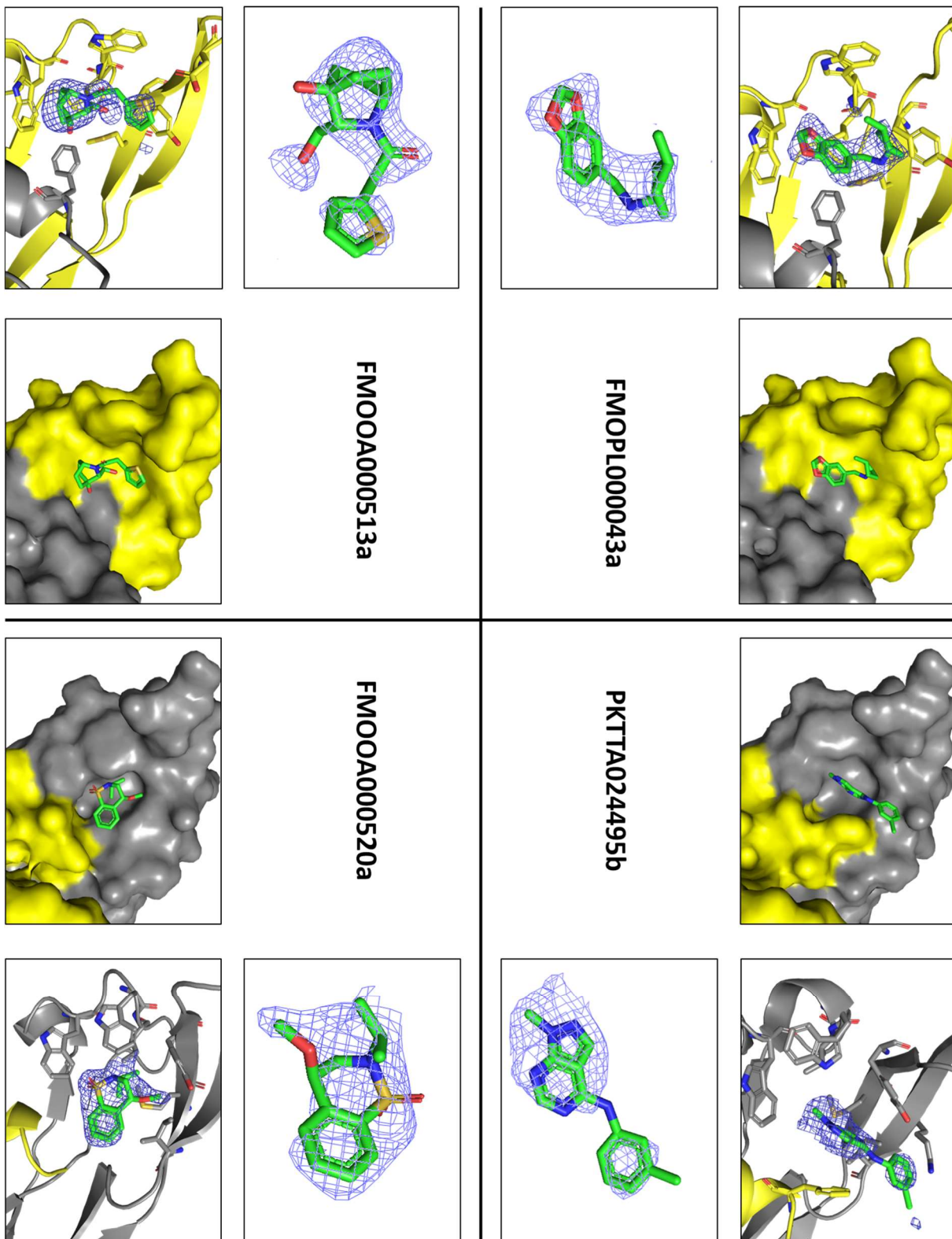


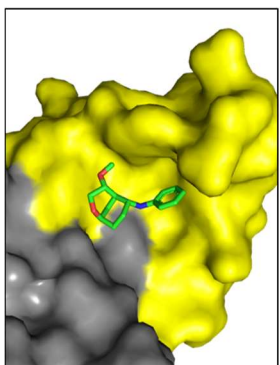
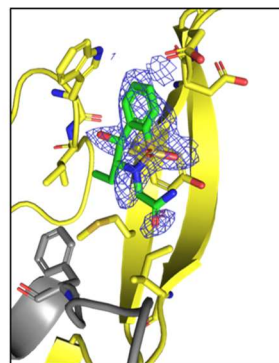
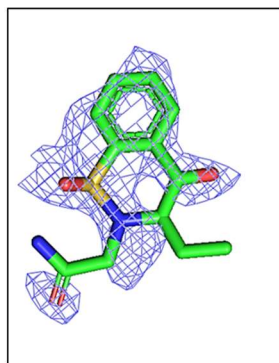
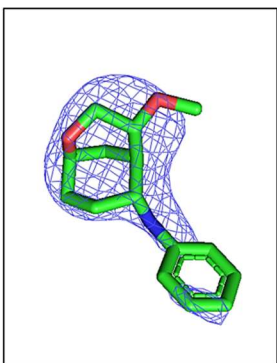
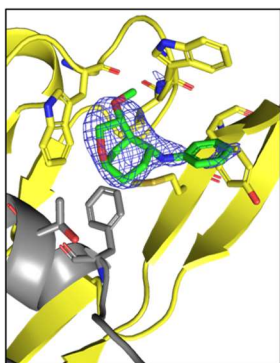
Interfacial Site



**B**

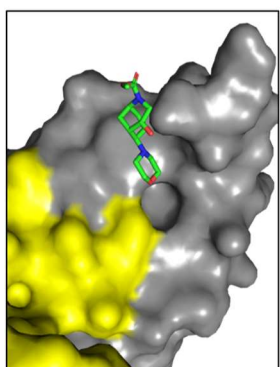
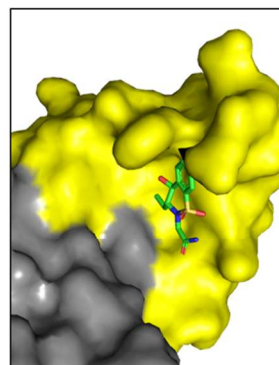






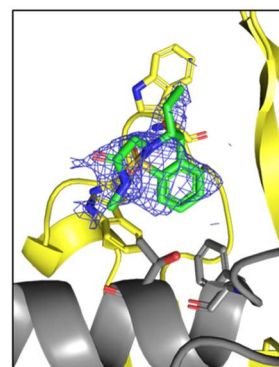
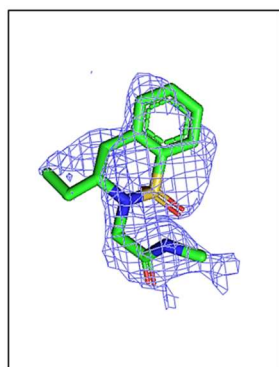
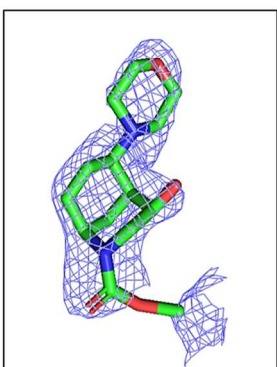
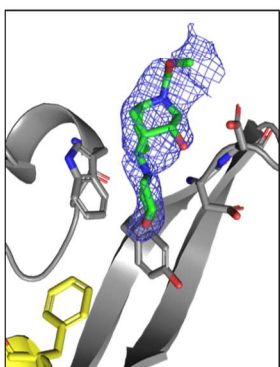
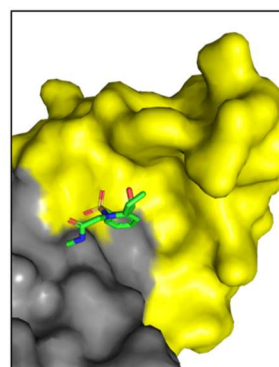
**FM00A000580a**

**FM00A000553a**



**FM00A000598a**

**FM00A000565a**



C

FMOP L000578a

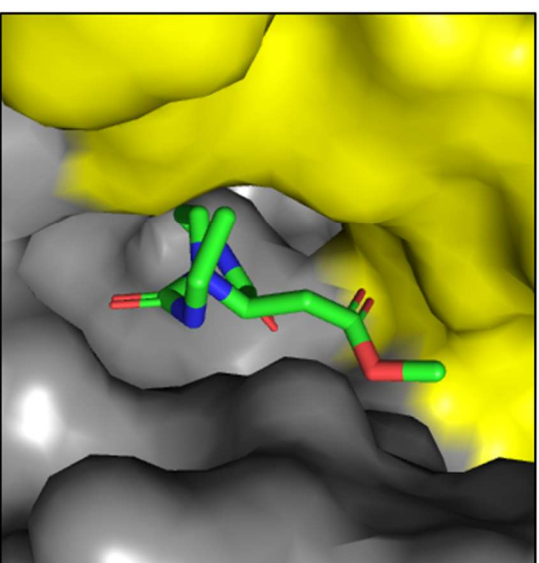
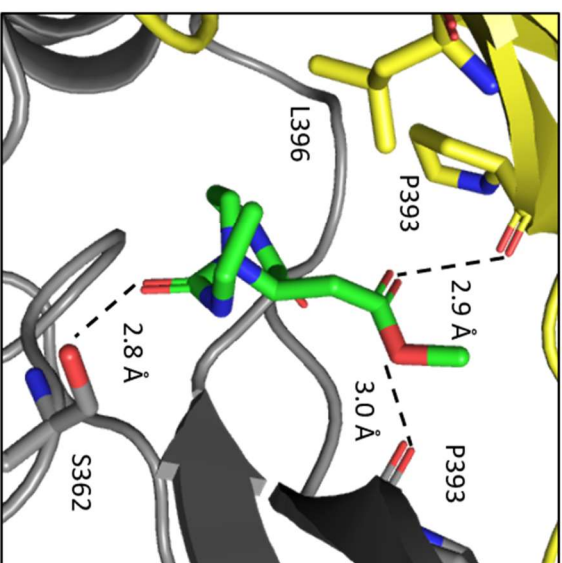
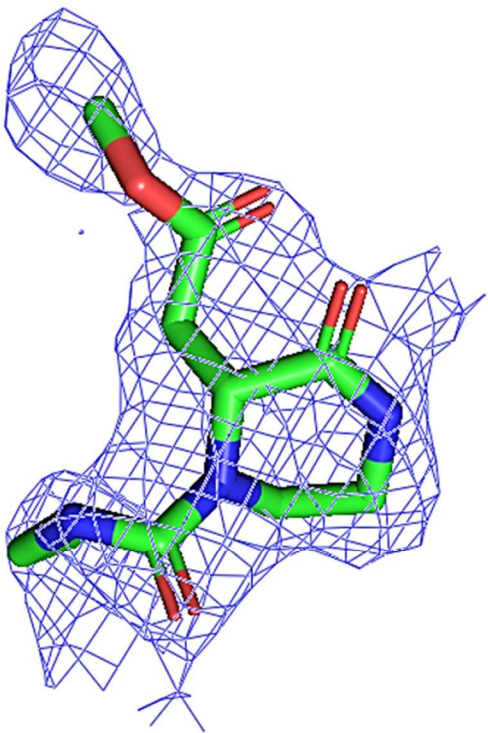
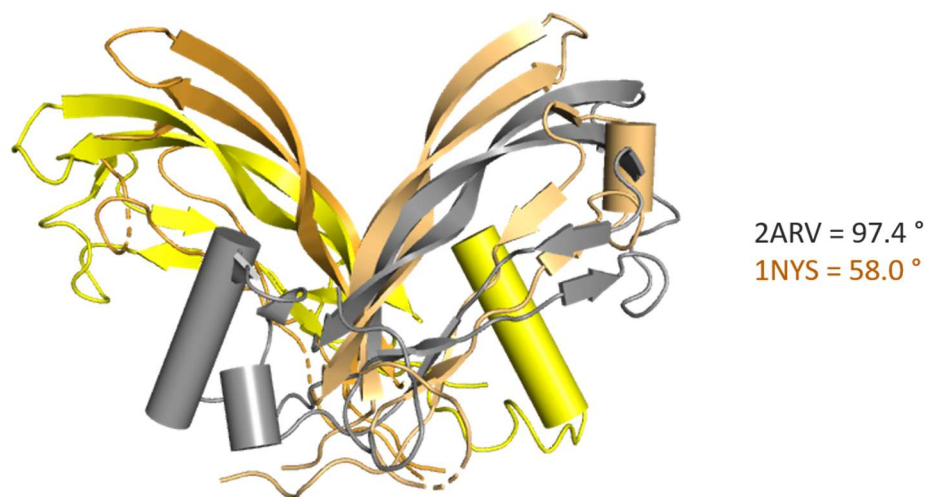




Figure 8: XChem screen of mature activin A. A) The overlaid alignments of all structures obtained for fragment hits from the XChem screen to the DMSO structure of activin A. For clarity, the surface of the original DMSO structure is shown. B) The structure of all 12 fragment hits bound to mature activin A in the putative type I receptor binding site as identified from the XChem screen. Here the structure of each fragment fitted to the density is shown along the fragment fitted to the density in the binding site and fitted with a surface representation. C) The structure of FMOPL000578a bound to mature activin A in the interfacial site as identified from the XChem screen. Here the structure of each fragment fitted to the density is shown along the fragment fitted to the density in the binding site and fitted with a surface representation.

The most common binding site observed for fragments was the groove in the putative type I receptor binding site (Figure 8B). Molecules that bound at this site were generally hydrophobic in nature, often containing large aromatic or aliphatic rings with few hydrogen bonding groups. As such, the nature of the interactions with residues in this site were usually hydrophobic, with interactions with the side chains of W335, W338, M401, Y403, I415, and M418 of chain A being the most common, as well as F368 of chain B. Incidentally, this is advantageous for inhibitor development as Harrison *et al.* identified M401, I415, and M418 as being crucial residues for mediating ALK4 binding thus by interacting with these residues, a molecule may inhibit signalling.<sup>216</sup> Though not directly implicated as being crucial to binding by Harrison *et al.*, residues F368 and Y403 sit in proximity to M401, I415, and M418 at the entrance to a binding surface containing these residues. As such, molecules that interact with these residues appear to block access to the binding surface and so may be able to prevent ALK4 from binding. Incidentally, though interactions with M401, I415, and M418 are not seen for every hit, interactions with F368 and/ or Y403 are observed in every structure.

Only one molecule from the entire screen was identified to bind at the dimer interface, FMOPL000578a (Figure 8C). This site is of interest due to the role of Activin A's flexibility about its dimerization cysteine in binding to type II receptors. The structure obtained by Greenwald *et al.* (PDB ID: 1NYS) of activin A bound to two ActRIIB extracellular domains both show a reduction in the angle about the dimerization cysteine compared to the structure of mature activin A (58.0° vs 97.4°) (Figure 9).<sup>103</sup> As such, it is postulated that this change in conformation is crucial in creating an interface that allows the type I receptor to bind. Therefore, it is possible that a molecule that binds in the interfacial site may reduce the mobility of the activin A subunits, thus preventing the creation of the type I binding interface and allosterically inhibiting signalling.



*Figure 9: Alignment of mature activin A and receptor bound structures. Alignment of the mature activin A structure to the structure of activin A (yellow / grey) bound to the extracellular domains of ActRIIB (orange) (ECDs of ActRIIB removed). Angles are calculated between C $\alpha$  of Y403, C $\alpha$  of C390, and C $\alpha$  of Y403 of the other subunit.*

In contrast to the molecules that bind in the vicinity of the type I receptor binding site, FMOPL000578a appears more hydrophilic, containing several hydrogen bonding groups. Indeed, the majority of interactions here appear to be hydrogen bonds, formed with the sidechain of S362 of chain B, as well as the backbone of P393 and T394 of chain A and P393 of chain B. A hydrophobic interaction does appear to be present between L396 of chain A and the hydrophobic face of the 6-membered ring though this is the only clear major hydrophobic interaction observed. This indicates that the nature of this pocket is significantly less hydrophobic than the type I receptor binding site. This is also reflected in the detection of water molecules in this site in the apo structure. Similar to the type I receptor binding site hits, the majority of density corresponding to the molecule is only observed in one half of the symmetrical binding pocket. Due to the symmetrical nature of this pocket, one possible optimisation route would be to take the key binding motif of the hit and develop a molecule containing two of these motifs joined by a linker. However due to the time constraints of the project, further optimisation or targeting of this binding site was not attempted.

Of the binders identified, the four hits from the Kidd 3D library that bound in the putative type I receptor binding site were selected for further analysis – NM466, HLS327, SG5B8, SG6B10. This was due to the availability of local chemistry support for these four fragments.

## 2.4 Validation of Binding of NM466, HLS327, SG5B8, and SG6B10 to Mature Activin A via X-ray Crystallography

In order to validate the binding of NM466, HLS327, SG5B8, and SG6B10 to mature activin A each molecule was screened via X-ray crystallography again.

Before soaking, crystallisation conditions were further optimised to improve data quality. The ability of PEG 300 to bind to the same site as the four fragments (or the binding site on the opposite side of the molecule) could be problematic. Fragments may have to compete with PEG for binding which could lead to difficult-to-interpret density. Therefore, it was decided to attempt crystallisation in the absence of PEG 300. Initial attempts to crystallise activin A in the absence of PEG 300 failed, however microseeding with PEG 300 containing crystals facilitated crystal formation with needle like crystals, similar to those described previously, forming (for crystallographic statistics, see Appendix 1). Analysis of the structure obtained shows a key difference in the fragment binding site compared to the PEG 300 containing structure (Figure 10). In the “PEGless” structure, W338 adopts a difference conformation to that seen in the PEG containing structure, with the side chain being flipped such that it sits adjacent to the side chain of W335 appearing to undergo  $\pi$ -stacking interaction. Similar to the non-PEG containing binding site in the DMSO structure, W338 also appears to hydrophobically interact with a DMSO molecule. This is in opposition to the PEG containing structure where the sidechain adopts a different conformation to accommodate PEG binding. As such, this residue can act as a useful tool for screening, as monitoring the conformation of this residue may indicate binding of a compound.

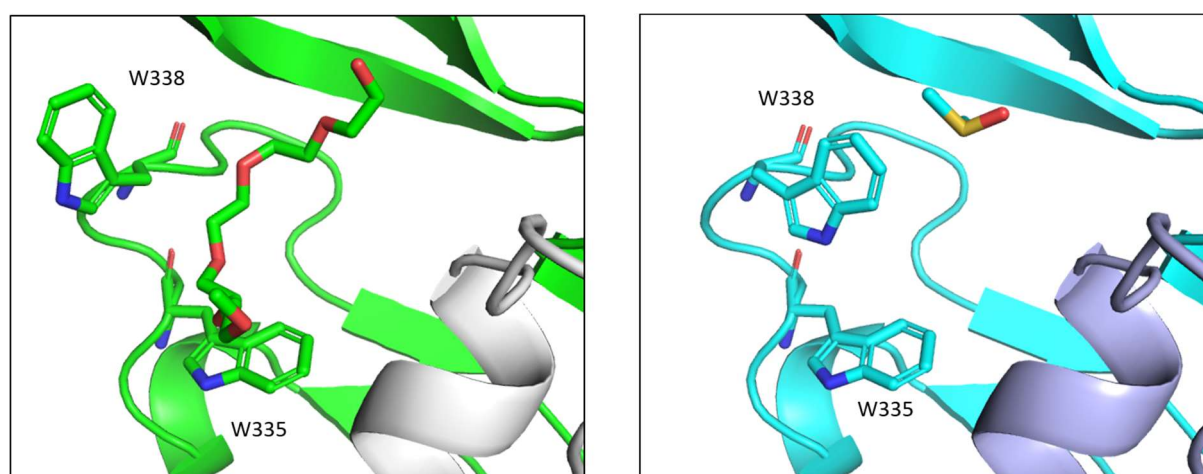


Figure 10: The “PEGless” structure of mature activin A. The structure of the PEG binding site in the “PEG containing” (left) and “PEGless” (right) structures. Here, the “PEG containing structure is shown in green and white and the “PEGless” structure is shown in cyan and pale blue. W338 and interacting residues are highlighted.

Crystals of mature activin A were soaked with 40 mM NM466, HLS327, SG5B8, and SG6B10 and diffraction data was collected. All four crystals diffracted with resolutions of 2.0 Å respectively (for crystallographic statistics, see Appendix 1). After refinement only NM466 showed the presence of a large unoccupied density at the molecule binding site. As chiral molecules in the Kidd 3D library exist as racemic mixtures, it was important to determine which of the two enantiomers were binding. As such, both enantiomers were fitted into the density in separate refinements. From fitting these molecules, it was determined that the (R)-enantiomer corresponded much better to the observed density than the (S)-enantiomer and thus the (R)-enantiomer is interacting with mature activin A (Figure 11A).

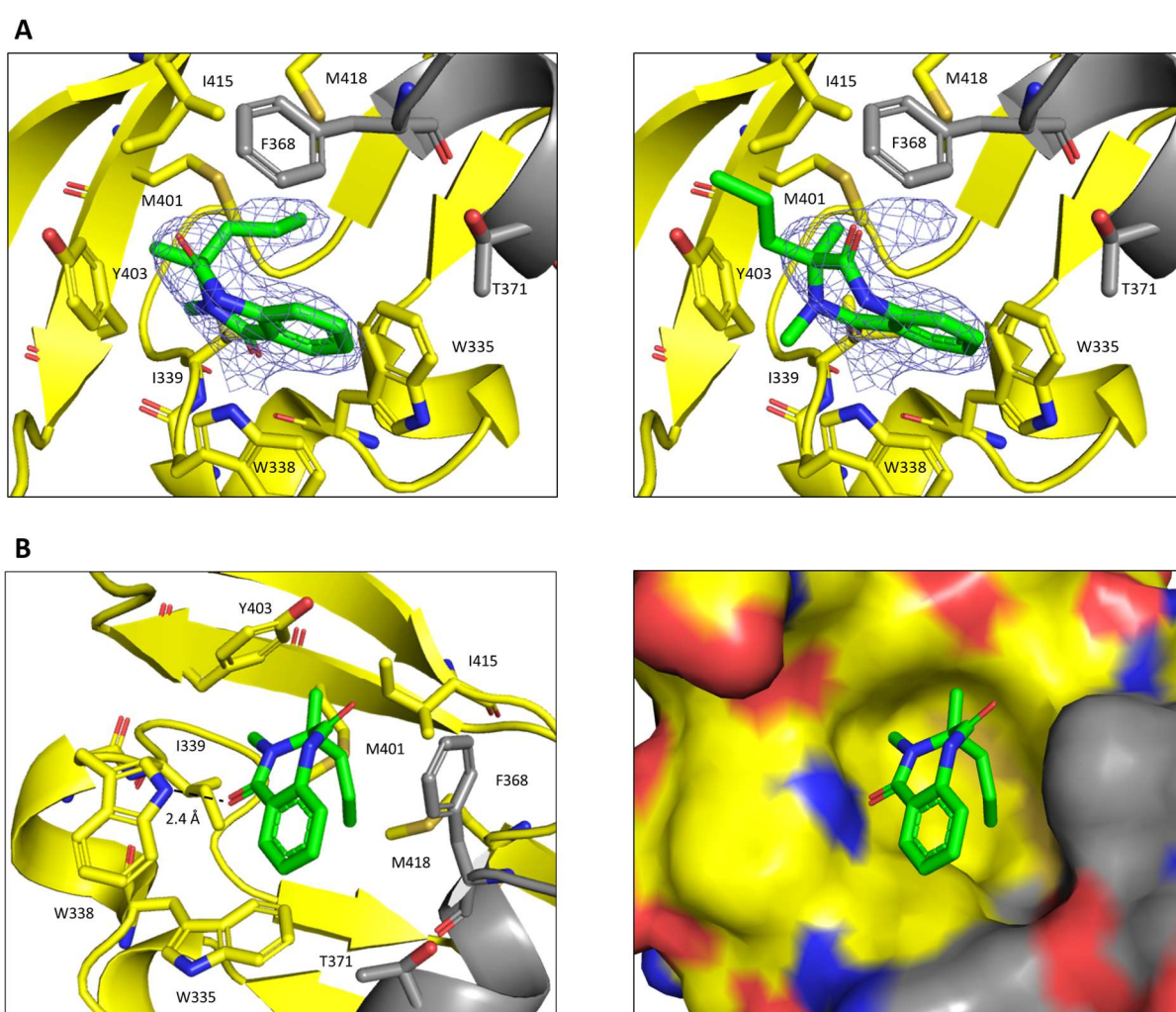


Figure 11: Binding of NM466 to mature activin A. A) The structure of NM466 soaked mature activin A solved with the (S)- (left) and (R)- (right) enantiomers B) NM466 bound to the putative type I receptor binding site. A ribbon representation with the interacting residues is shown on the left whereas a surface representation is shown on the right. Here carbon atoms are shown in yellow and grey, oxygen atoms are coloured in red, nitrogen in blue, and sulfur in gold.

NM466 utilises a compact, folded conformation when binding to mature activin A, with the 7-membered bi-lactam ring adopting a boat-like conformation (Figure 11B). As such the hydrophobic

propyl chain is positioned axially to the ring allowing for maximisation of the hydrophobic interactions between the molecule and activin A. The majority of the interactions are hydrophobic with NM466 interacting hydrophobically with W335, I339, M401, Y403, I415 and M418 of chain A plus F368 and T371 of chain B. W338 adopts a similar conformation to that observed in the structure without PEG, creating the pocket to which NM466 binds. A  $\pi$ -stacking interaction between the phenyl ring of NM466 and W335 also appears to be present. Hydrophilic interactions between NM466 and mature activin A are significantly less common than hydrophobic ones, with only one hydrogen bonding interaction appearing occur. This occurs between the nitrogen of W338's indole and the carbonyl of one of the amide groups of NM466 measuring 2.4 Å. Based on these observed interactions, a series of derivatives of NM466 were designed and synthesised in collaboration with Prof David Spring's group (Department of Chemistry, University of Cambridge) to validate binding to mature activin A and explore SAR. (Figure 12, Table 4).

## 2.5 Screening of Derivatives of NM466 against Activin A

A total of 10 derivatives of NM466 were soaked into activin A crystals for screening via X-ray diffraction. Of these, only one, TR17, showed density corresponding to the molecule binding with no obvious density being observed for the others.

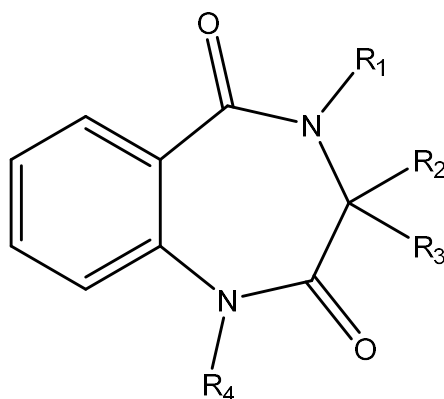
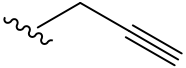
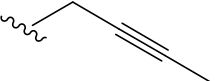


Figure 12: The generic structure of a NM466 derivative.

**Table 4: Screened derivatives of NM466.**

Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	Observed Binding?
NM466	Me	Me	Pr	H	Yes
T17	Et	Me	Pr	H	Yes
TR19	Me	Ph	Pr	H	No
TR22	H	Me	Pr	H	No
TR28F1	Me	Me		H	No
TR28F2	Me	Me		H	No
TR61	Me	Et	Pr	H	No
TR75	Me	H	H	H	No
TR78	Pr	Me	Pr	H	No
TR80	Me	Me	Pr	Me	No
TR81	Bu	Me	Pr	H	No

TR17 differs in only one position from NM466, with an ethyl group replacing the methyl group at position R<sub>1</sub>. Both molecules bind in a similar manner with adopting the same conformation and interacting with the same residues on activin A. The binding of TR17 is slightly rotated compared to that of NM466 to accommodate the increased interactions between the ethyl group at R<sub>1</sub> and M401 and Y403 (Figure 13). The identity of R<sub>1</sub> appears to have a large influence on binding. For molecules with either a methyl or ethyl group placed at position R<sub>1</sub>, binding to mature activin A was observed. However, for molecules with propyl or butyl groups in the same position it was not – likely due to the steric clashes between mature activin A and these larger groups. Interestingly, TR22 which has a hydrogen atom positioned at R<sub>1</sub> did not bind to mature activin A. This may be due to an alternate conformation of TR22 being more energetically favourable than the observed binding conformation of NM466 and TR17, thus there are a lower number of molecules in the binding conformation so binding is not detected. Of the other derivatives, it is likely the majority of molecules did not bind due to steric clashes of their larger or more structurally rigid groups with the surface of mature activin A.

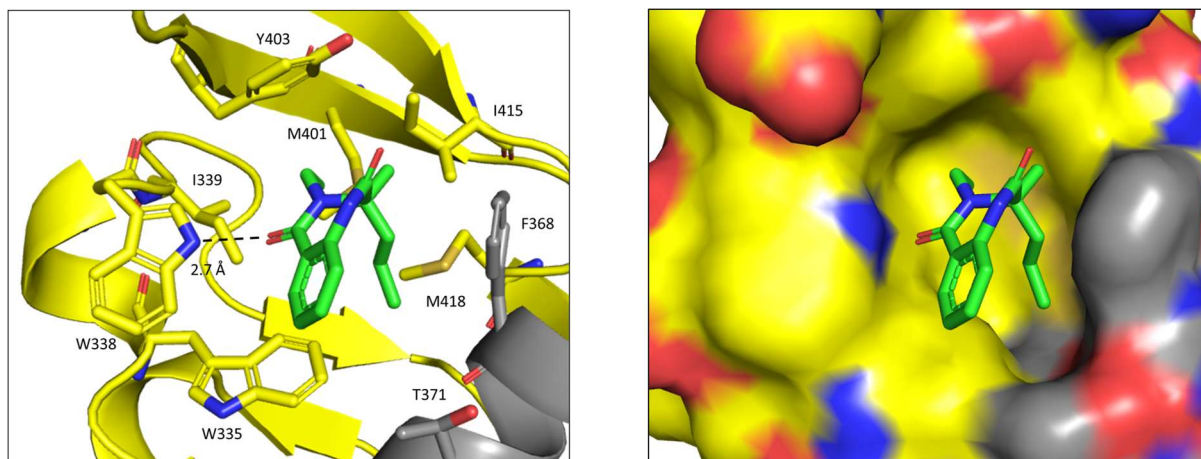


Figure 13: Binding of TR17 to mature activin A. A ribbon representation with the interacting residues is shown on the left whereas a surface representation is shown on the right. Here carbon atoms are shown in yellow and grey, oxygen atoms are coloured in red, nitrogen in blue, and sulfur in gold.

Ligand-based NMR was attempted to determine if NM466 and TR17 bound to activin A however the data was inconclusive (data not shown). Due to the time constraints of the project, no further derivatives were screened.

## 2.6 Conclusions

In this section, I have described how the crystallisation conditions of mature activin A were optimised for fragment soaking. I then described how a 345-fragment screen was conducted against activin A using the XChem platform with PanDDA analysis to identify a total of 13 binding fragments. These fragments bound in two distinct sites – the putative type I receptor binding site, and an interfacial site adjacent to the dimerization cysteine. I then took four of these compounds, NM466, HLS327, SG5B8, and SG6B10, forward for further validation via X-ray crystallography. From this, I only observed density for NM466 and determined that the (R)-enantiomer was interacting with activin A. Derivatives of this molecule were then screened via X-ray crystallography resulting in the identification of one further binding compound, TR17, which bound in a similar manner. Ligand-based NMR spectroscopy was then used to analyse the in-solution binding of NM466 and TR17 however the data was inconclusive.

Based on these results, I would say the viability of this pocket is questionable. The pocket is relatively large (surface area =  $1700 \text{ \AA}^2$ , volume =  $950 \text{ \AA}^3$ ) and the highly hydrophobic nature of the pocket means that specific interactions would be difficult to achieve. As such, binding molecules must depend on a number of non-specific hydrophobic interactions spread across a large surface area. Therefore, it is unlikely that binding in this pocket could lead to a highly specific or potent inhibitor of mature activin A.

In the next section, I will discuss how I used an alternate approach for developing inhibitors of activin A – the use of peptide fusions derived from the pro-domain of activin A.



---

# Peptide Approach

---

### 3.1 Introduction

In addition to a small molecule approach, a peptide-based approach was also attempted to achieve inhibition of activin A signalling. The generation of peptide inhibitors has certain advantages when compared to small molecule, especially with regards to protein-protein interactions. As larger molecules, peptides can take advantage a number of weak interactions across a wide surface area to create high affinity binding. Additionally, due to the large number of interactions, interactions can be tuned at specific sites to create highly specific binding. Thus, peptides are less likely to suffer from off-target interactions and events. This is highly desirable in the case of the TGF- $\beta$  family where the structural and sequence similarity of the ligands and ligand promiscuity make high specificity a requirement of any inhibitor. However, issues can arise from peptides' lack of membrane permeability and susceptibility to proteolysis.<sup>234,235</sup>

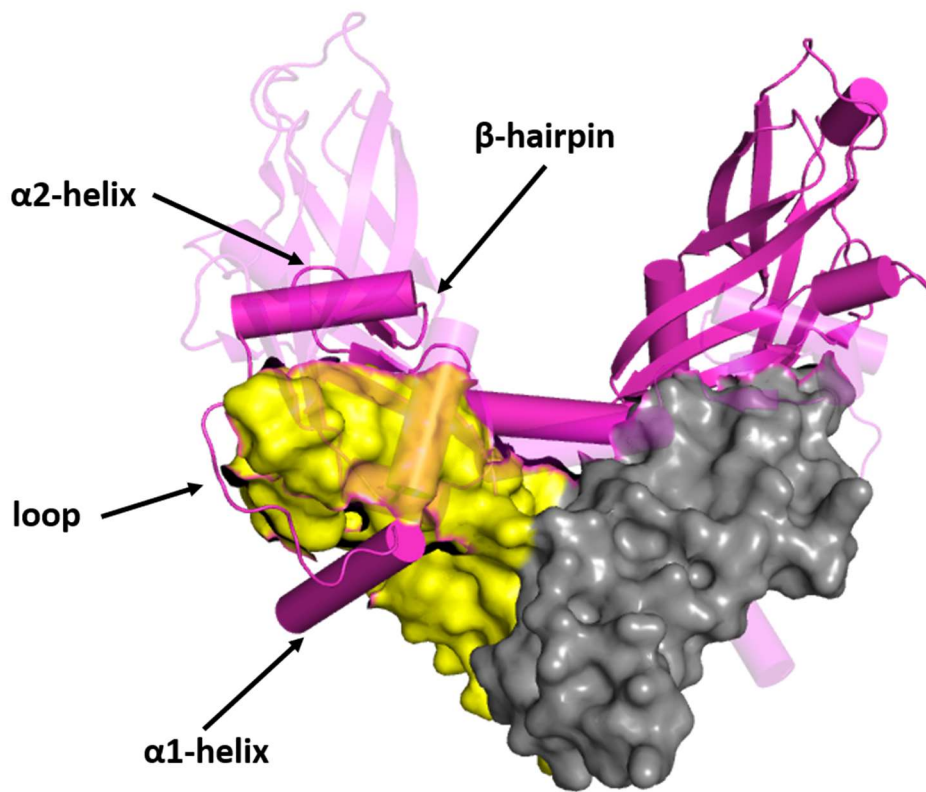
The TGF- $\beta$  family ligands are good candidates for peptide-based inhibition. As they are extracellular proteins, there are no requirements for an inhibitor to cross the cell membrane. Furthermore, as only the mature growth factor is required for signalling, the TGF- $\beta$  family has a natural starting point for designing peptide inhibitors – the pro-domains. Each dimeric TGF- $\beta$  ligand is made up of two pro-domains and two mature domains. After expression, the pro-domains are cleaved from the mature domains but remain in association.<sup>94</sup> Though important for solubility, stability, and storage, pro-domains are not believed to be actively involved in signalling, dissociating to allow receptor binding to occur.<sup>95,104,106</sup> Pro-domains usually bind their mature growth factor with high affinity and it has been shown that for several members of the family the pro-domain can also inhibit signalling, though the potency of this inhibition is highly dependent on the growth factor in question.<sup>99,108</sup> Pro-domains have already been used as inhibitors with pro-domain-based ligand traps being used to inhibit activin A signalling (See “1.6 Inhibitors of TGF- $\beta$  Family Signalling”) (Figure 4).<sup>236</sup> Mankanji et al. also developed a chimeric inhibitor through fusing the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of activin A to the shoulder region of the TGF- $\beta$ 1 pro-domain. This inhibitor inhibited activin A signalling with an IC<sub>50</sub> of 2.6 nM.<sup>188</sup> Peptide inhibitors based on the N-terminal region of myostatin pro-domain have been shown to successfully inhibit myostatin signalling. Takayama *et al.* developed inhibitors based on the mouse pro-domain, identifying a minimal binding peptide with an IC<sub>50</sub> value of 3.5  $\mu$ M in a cell-based assay and a K<sub>d</sub> value of 30 nM for the mature domain obtained via surface plasmon resonance (SPR). Since then the peptide has been optimised through increased interactions with the mature domain and introduction of a staple. This has led to an approximately 13-fold increase in inhibitor potency with an IC<sub>50</sub> value of 260 nM for the most potent inhibitor.<sup>189,190</sup>

The region that all these inhibitors are based on is N-terminal segment of the pro-domains. This region is of particular interest as it contains the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif that was discussed previously (see “1.4 The Structural Features of the TGF- $\beta$  Family and Activin A”). Analysis of the individual residues in this region show the region is highly hydrophobic, with residues being conserved across the family. The currently available pro-domain structures also show how residues in this region form the majority of contacts with the mature domains (Figure 14A). The pro-domain of activin A has a high affinity for the mature growth factor (5 nM measured via BLI) and it is likely that the affinity is due to the interactions in this motif.<sup>104</sup> Therefore, I decided to use the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of the activin A pro-domain to develop inhibitors for activin A signalling.

## 3.2 Design of Peptide Fusions

In order to generate inhibitors for activin A based on its pro-domain, it was first important to identify the key motifs that were vital for mature domain binding. Analysis of the structure of the activin A pro-mature complex shows the majority of the interacting pro-domain residues occurs between Q51-E104 (Figure 14). This sequence contains the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif typical of TGF- $\beta$  family pro-domains along with the highly conserved hydrophobic residues present across the pro-domains of all TGF- $\beta$  family members. In this epitope, the  $\alpha$ 1-helix binds to the putative type I receptor binding site on the concave face of the mature domain and the  $\alpha$ 2-helix binds to the convex face in the vicinity of the type II receptor binding site. A  $\beta$ -hairpin that follows the  $\alpha$ 2-helix also appears to make a number of contacts on the convex face. Based on this, five peptides were designed incorporating motifs that may be key to binding (Table 5).

**A**



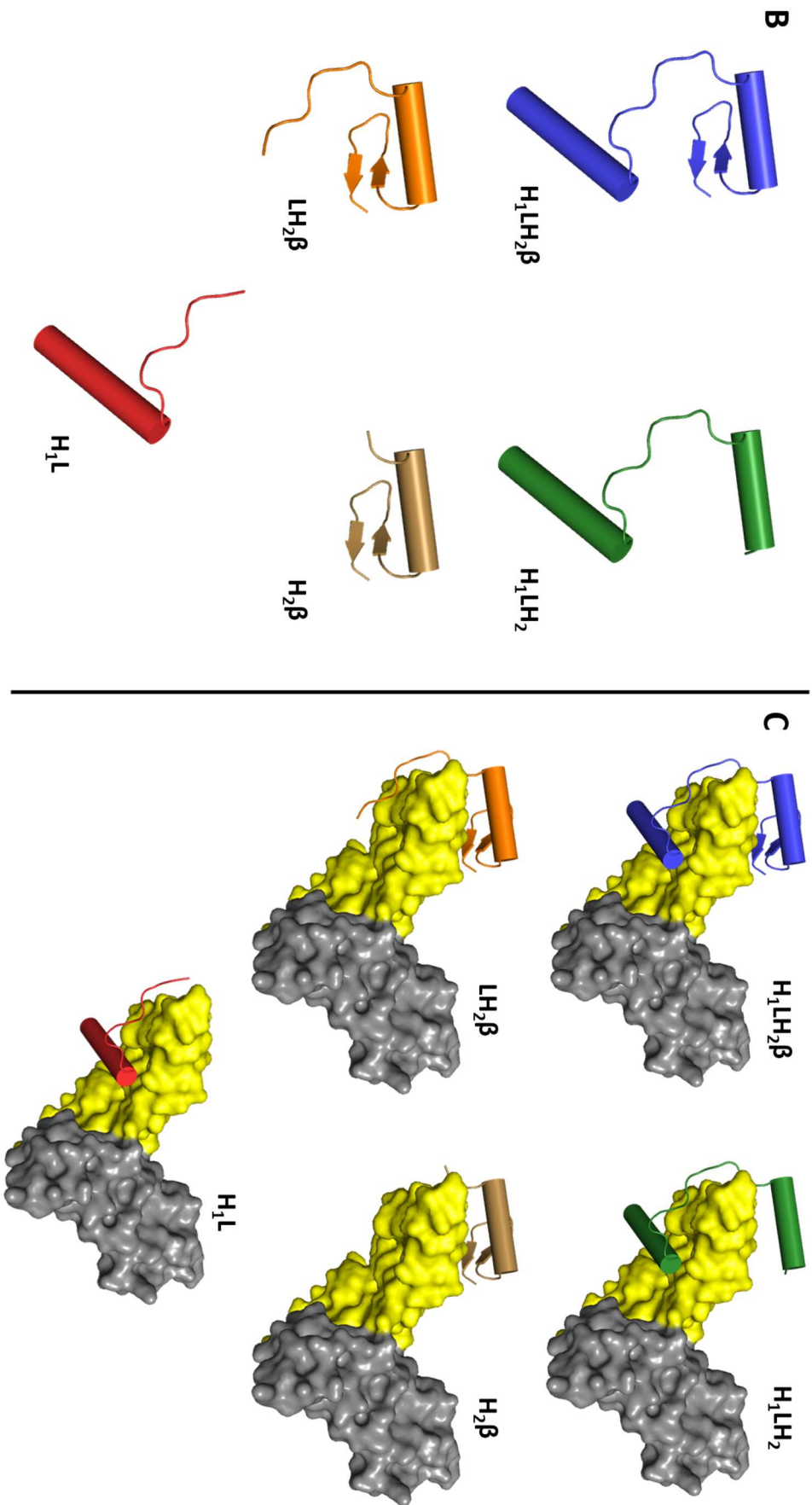


Figure 14: Pro-domain derived peptides of activin A. A) The structure of the pro-mature complex of activin A. Here one of the pro-domains (light pink) has been made transparent and the  $\alpha$ 1-helix, loop,  $\alpha$ 2-helix and the  $\beta$ -Hairpin of the other (magenta) highlighted. The mature domains are shown in yellow and grey B) Representations of the peptides used in this project based on the pro-mature structure. Here  $H_1LH_2\beta$  is shown in blue,  $H_1LH_2$  in green,  $LH_2\beta$  in orange,  $H_2\beta$  in gold, and  $H_1L$  in red. C) Representations of the peptides used in this project based on the pro-mature structure bound to mature activin A.

**Table 5: Pro-domain derived peptides of activin A.**

Peptide	Features	Residue Range
$H_1LH_2\beta$	$\alpha$ 1-helix-loop- $\alpha$ 2-helix- $\beta$ -hairpin	Q51-E104
$H_1LH_2$	$\alpha$ 1-helix-loop- $\alpha$ 2-helix	Q51-H91
$LH_2\beta$	loop- $\alpha$ 2-helix- $\beta$ -hairpin	K69-E104
$H_2\beta$	$\alpha$ 2-helix- $\beta$ -hairpin	Q76-E104
$H_1L$	$\alpha$ 1-helix-loop	Q51-Q76

Due to the length of the designed peptides (26-74 amino acids) it would be uneconomical and inefficient to purchase them commercially or synthesise them via solid phase peptide synthesis. I therefore decided that the peptides would be generated as fusion proteins with an N-terminal B1 domain from streptococcal protein G (GB1). GB1 is well characterised as a fusion partner and its C-terminus is positioned such that the peptide inhibitors are left exposed for interactions in solution. In addition, an N-terminal His-tag and C-terminal StrepTag II were incorporated to allow a two-step purification where the protein is first purified by Ni affinity chromatography followed by StrepTactin affinity chromatography.<sup>237</sup> The exposed nature of the pro-domain epitopes means that proteolytic degradation by cellular proteases during the expression phase is likely. Therefore the dual purification system allows for any product that has undergone degradation at the N-terminus to be removed in the first step and any that has undergone degradation at the C-terminus to be removed in the second step, allowing for the isolation of the desired full-length peptide fusion only (for full construct sequences, see Appendix 2).

### 3.3 Cloning of Peptide Fusions

To generate the peptide fusions desired, it was first necessary to obtain a vector that incorporated the desired tags and fusions partners. The vector pOP5BP, used already in the Hyvönen Group, contained DNA encoding an N-terminal His-tag, GB1 domain multiple cloning site, and a C-terminal Avi tag, however lacked the desired C-terminal Strep-tag II. Thus, pOP5BP was restriction digested to remove the Avi-tag. Primers encoding Strep-tag II with 5' and 3' restriction sites for incorporation into pOP5BP were the directly annealed and phosphorylated before being ligated into digested pOP5BP to generate

pOP4BP (Figure 15). The success of the ligation for this step and all subsequent ligation steps was determined via Sanger sequencing.

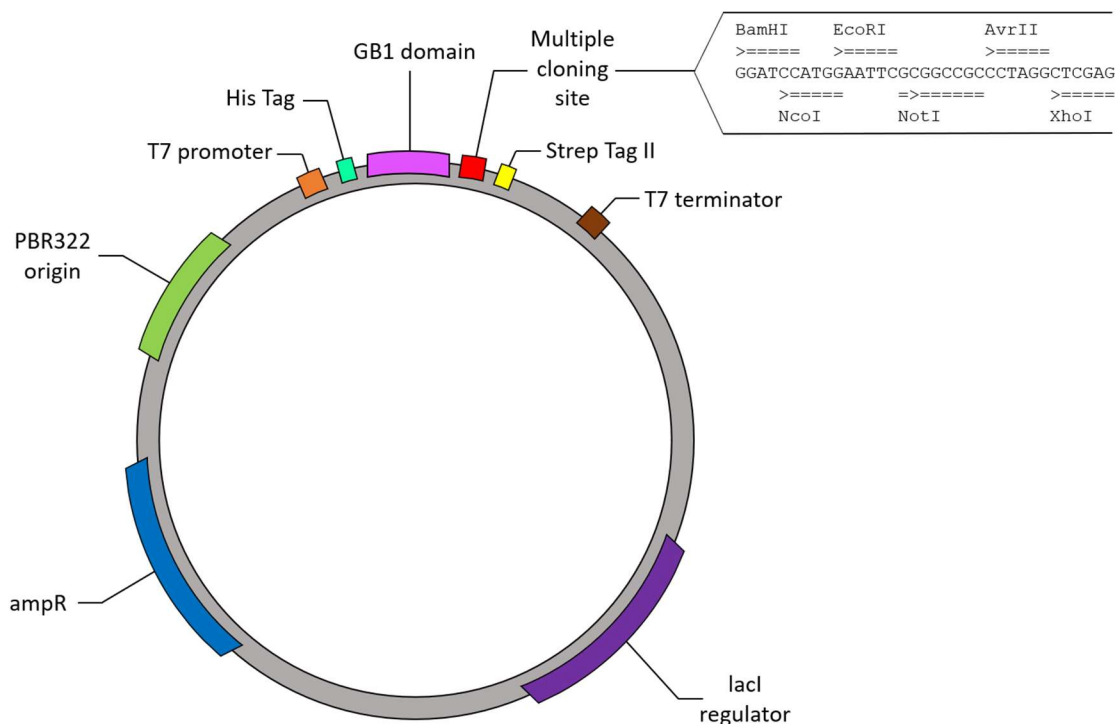


Figure 15: The pOP4BP vector. Here, the design of the pOP4BP vector employed in this study is shown. The multiple cloning site has been expanded so the restriction enzyme cutting sites are displayed.

Once the vector was generated. It was necessary to generate the peptide fusions themselves. DNA encoding the sequences for the peptides was amplified via polymerase chain reaction using appropriate primers with 5' and 3' restriction sites for incorporation into pOP4BP, and a template of pHAT2-actinin A (residues 21-426, Uniprot P08476). Amplified DNA was then ligated into pOP4BP. All subsequent peptide fusions described in this section were generated in the same manner.

### 3.4 Expression and Purification of Peptide Fusions

Successfully cloned peptide fusions and doubly tagged GB1 were transformed into chemically competent *E. coli* BL21 (DE3) cells for expression. In order to determine the optimal temperature for expression, test expressions were conducted at 37 °C, 25 °C and 16 °C (Figure 16). Analysis of the subsequent polyacrylamide gel revealed prominent bands in the soluble fraction of each sample at 25 °C corresponding to the approximate molecular weight of the desired products. Though higher amounts of expression are observed for GB1-LH<sub>2</sub>β at 37 °C, expression appears lower in the majority of cases with the amounts of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>2</sub>β being significantly lower than that observed at 25 °C. Samples grown at 16 °C show similar expression to those grown at 25 °C, however due to the

quicker expression time, 25 °C was selected as the expression temperature. Based on this, expression in 1 l cultures was then conducted at 25 °C.

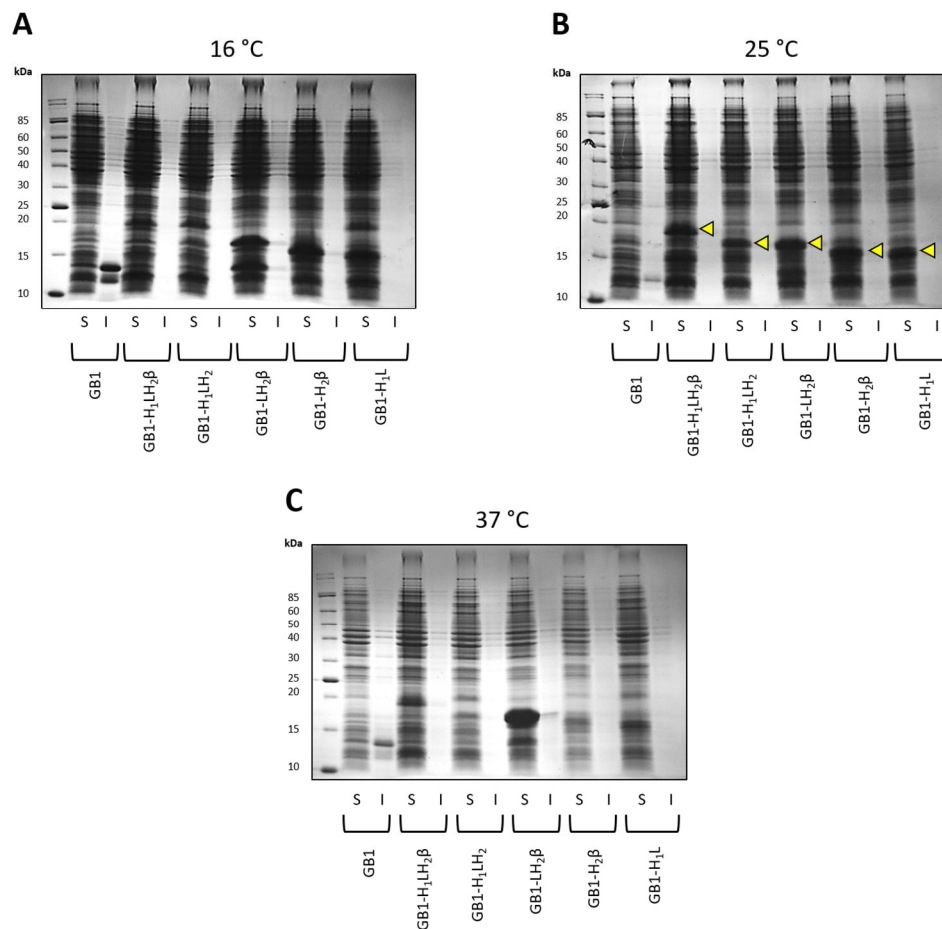


Figure 16: Expression tests of pro-domain derived peptide fusions. Expression tests conducted in BL21 (DE3) cells of the peptide fusions at 16 °C (A), 25 °C (B), and 37 °C (C). Here we see prominent bands in the soluble fraction for all peptides at 25 °C.

After expression, the peptide fusions were successfully purified via a two-step Ni-affinity-StrepTactin affinity purification. Analysis of the SDS-PAGE gels of the Ni-affinity purification shows that all the peptide fusions were successfully expressed however significant degradation products are present in each sample at lower molecular weights (Figure 17). To isolate the desired product from this mixture, each Ni-affinity purified sample was further purified via StrepTactin affinity purification. Here there is a single prominent band in each sample at the expected molecular weight indicating each protein had been successfully purified. To confirm the identity of the peptide fusions, each sample was analysed via intact MALDI-Tof mass spectrometry (Table 6). The molecular weight of each sample determined via mass spectrometry compares well with that of the calculated mass, differing only in each case by 2 Da (likely due to protonation and deamidation). Therefore, it was determined that each peptide fusion was purified successfully.



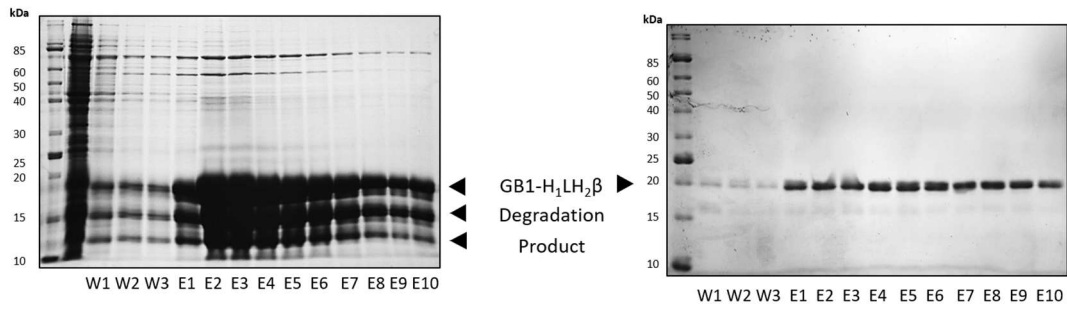
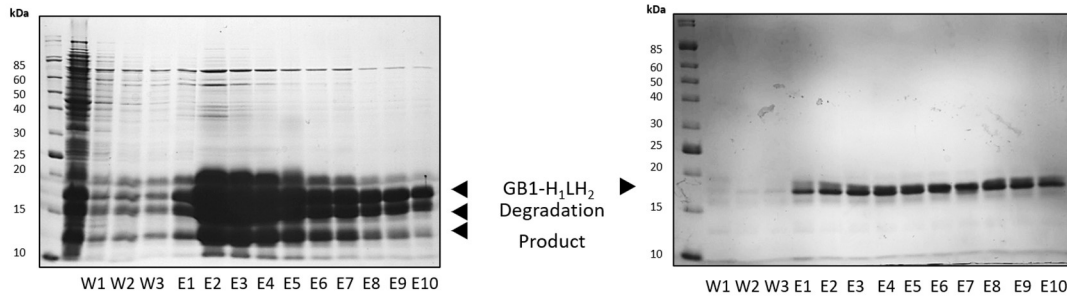
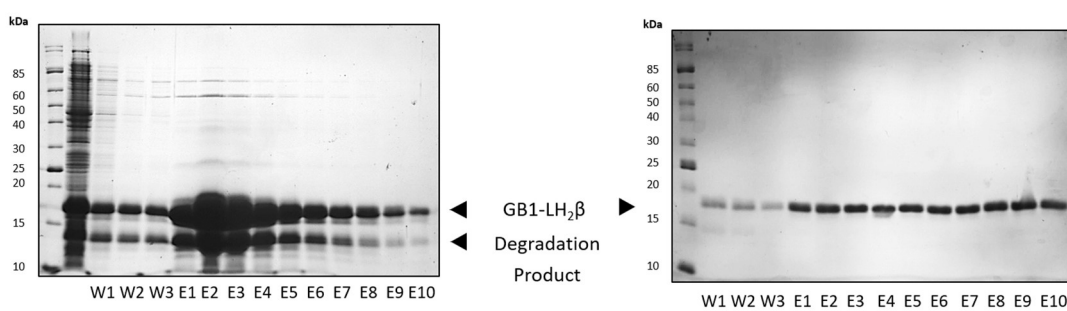
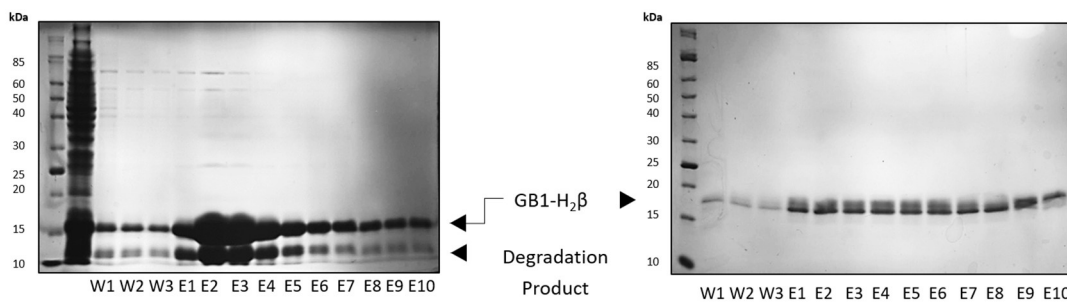
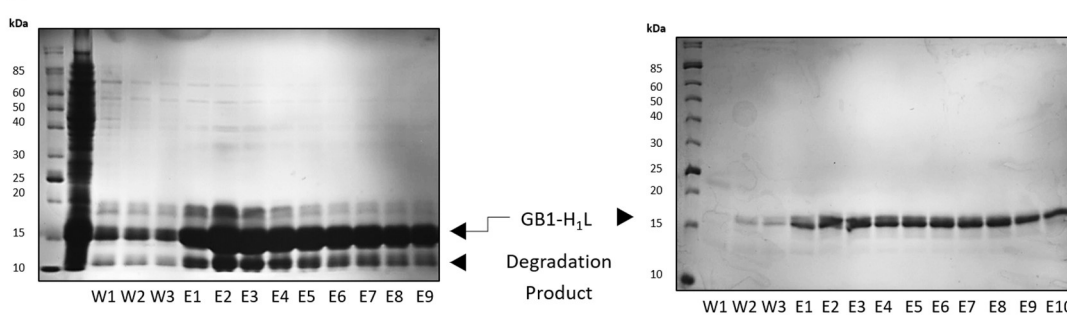
**A****B****C****D****E**

Figure 17: Purification of the peptide fusions via Ni-affinity and StrepTactin affinity purification. Here, SDS-PAGE gels are shown of the Ni-affinity and StrepTactin affinity purifications of each peptide. In each case we see the degradation product present after Ni-affinity purification is removed after StrepTactin affinity purification.

**Table 6: Molecular weights of peptide fusions as determined via mass spectrometry.**

Peptide Fusion	Calculated M <sub>w</sub> / Da	Observed M <sub>w</sub> (Mass Spec) / Da
GB1	11508.47	11510.20
GB1-H <sub>1</sub> LH <sub>2</sub> β	16837.82	16840.22
GB1-H <sub>1</sub> LH <sub>2</sub>	15463.30	15465.14
GB1-LH <sub>2</sub> β	14725.24	14727.64
GB1-H <sub>2</sub> β	13900.26	13902.21
GB1-H <sub>1</sub> L	13840.28	13842.89

### 3.5 Identification of the Key Binding Motifs in the Activin A Pro-domain

To determine which pro-domain motifs were vital for mature domain binding, the purified peptide fusions were analysed via biolayer interferometry (BLI). Peptide fusions were immobilised on anti-His tips and immersed in solutions of mature activin A at 100 nM and 1 μM. Doubly tagged GB1, was deemed to have a small but non-zero interaction with mature activin A, however this was at a significantly lower value than interacting motifs. As such, it was used as a reference to confirm binding was specific and all fusions were normalised to this (Figure 18 – Raw data: A, Normalised data: B).

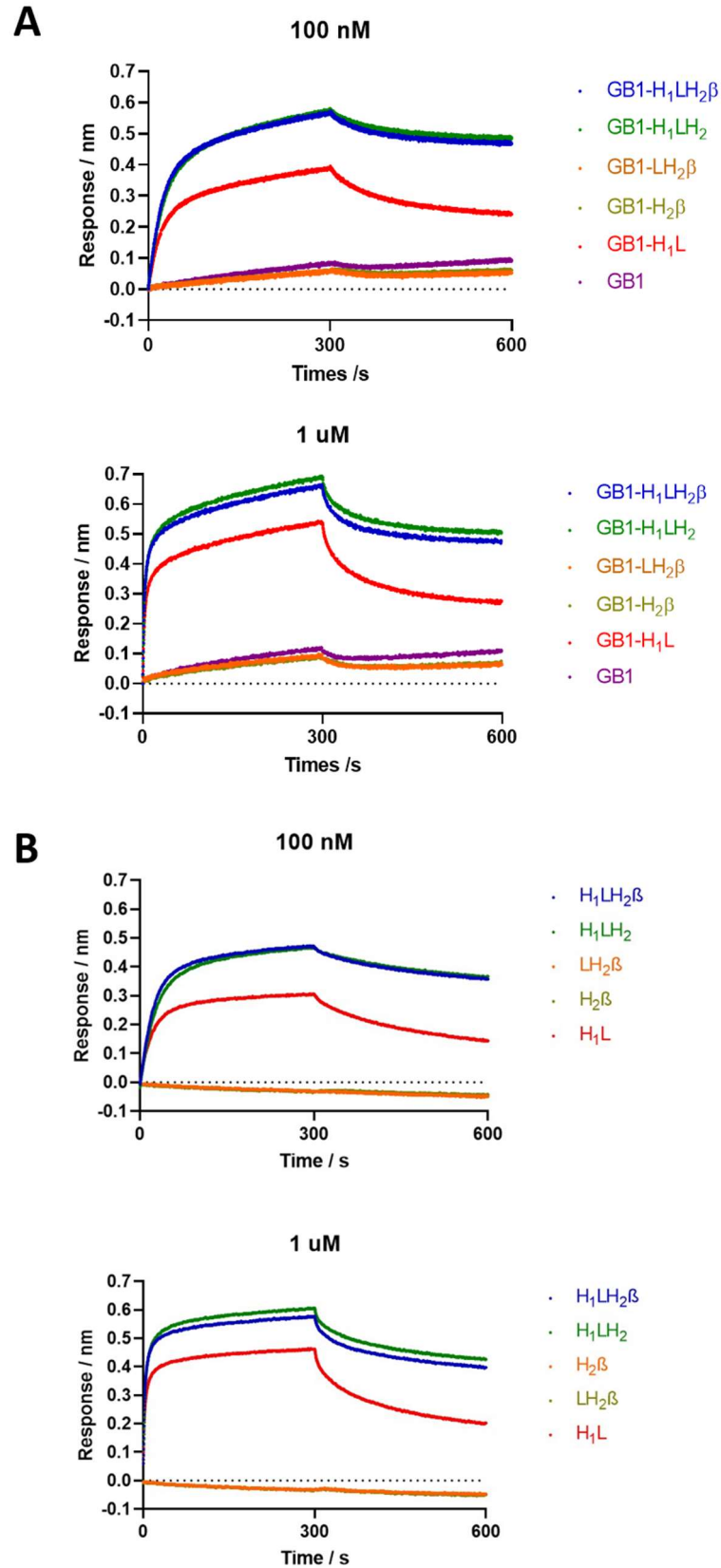
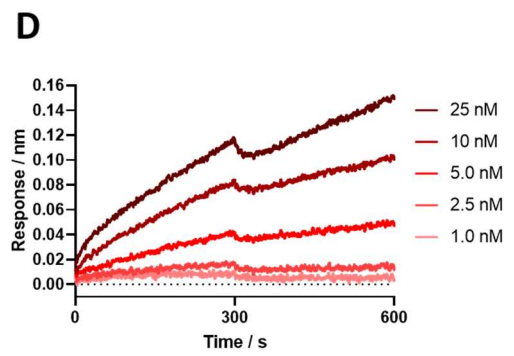
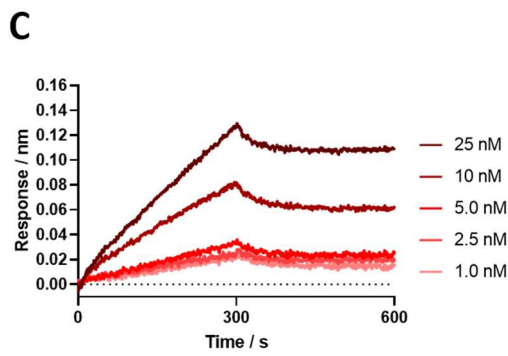
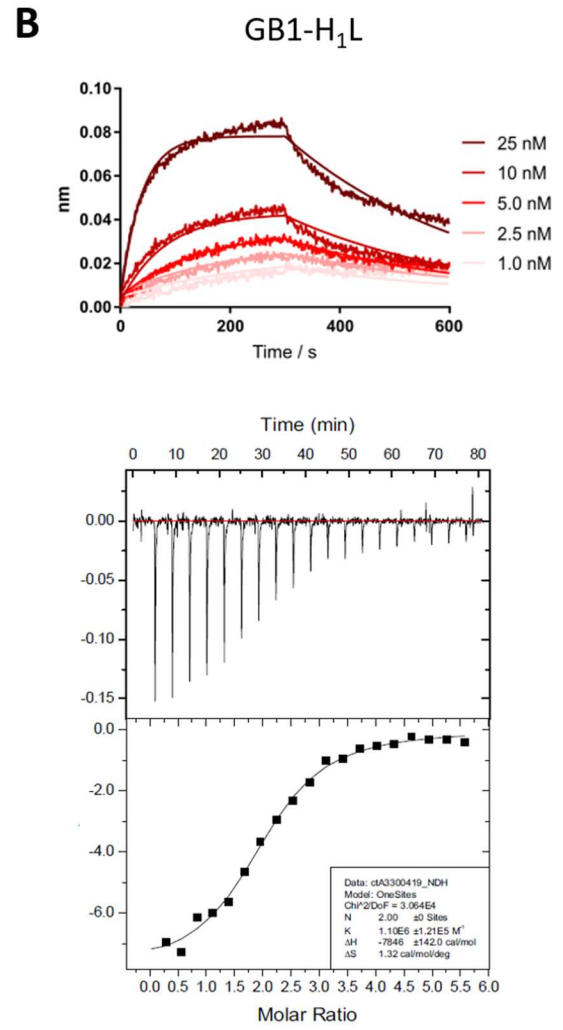
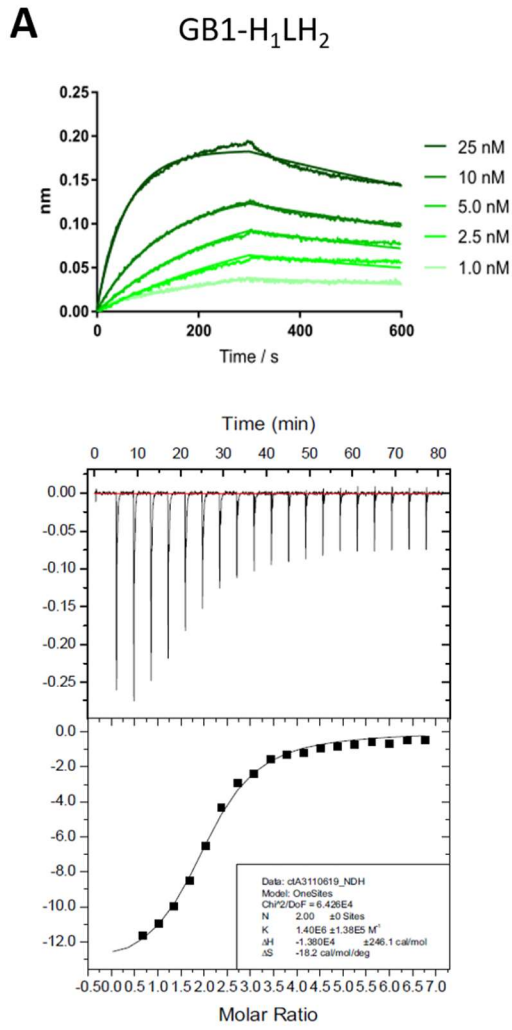


Figure 18: Identification of the key binding motifs of the activin A pro-domain via BLI. A) Unnormalized BLI traces of the peptide fusions at binding to 100 nM mature activin A. B) Normalized BLI traces of the peptide fusions

*binding to 100 nM and 1  $\mu$ M of mature activin A. Here the association phase occurs between 0-300 s and the dissociation phase between 300-600 s.*

Analysis of the binding traces shows that GB1-H<sub>1</sub>LH<sub>2</sub> $\beta$ , GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L showed significant binding relative to the GB1 control whereas GB1-LH<sub>2</sub> $\beta$  and GB1-H<sub>2</sub> $\beta$  did not, demonstrating the  $\alpha$ 1-helix and possibly the loop are important for mature domain binding. This compares well with observations made for myostatin and that the  $\alpha$ 1-helix and the initial residues of the loop are vital for mature domain binding.<sup>189</sup> The data reveals also that the  $\beta$ -hairpin does not appear to make a significant contribution to binding, with the binding profiles of GB1-H<sub>1</sub>LH<sub>2</sub> $\beta$  and GB1-H<sub>1</sub>LH<sub>2</sub> showing minimal, if any, difference. Additionally, though the  $\alpha$ 2-helix does not bind to mature activin A alone, it does appear to contribute towards binding, with the response magnitude and dissociation rates of GB1-H<sub>1</sub>LH<sub>2</sub> $\beta$  and GB1-H<sub>1</sub>LH<sub>2</sub> being larger and slower than that of GB1-H<sub>1</sub>L. Overall, this indicates that the  $\alpha$ 1-helix and the loop are the key motifs of the pro-domain for binding to the mature activin A.

Interestingly, closer inspection of the BLI data showed that plateau was not reached in the association phase for any of the peptides. Instead the response increased at an apparent constant rate. In addition, full dissociation of activin from the peptides did not occur in any instance, with signals not returning to the baseline in the dissociation phase. When the raw data is analysed, both observations are still present, only more severely. Initial attempts at determining the  $K_d$  of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L proved successful however, the kinetic model poorly fit the data (Figure 19A and B). There are two major possibilities for this. The first is that the anti-His antibodies used to immobilise the peptide fusions are holding two peptides in close spatial proximity, thus creating an avidity effect. As such, two association and dissociation events are being observed, the binding of one peptide fusion to activin A followed by a second. These interactions will have different kinetic rates and thus the measured change in response will be a combination of these two kinetic events resulting in the drift in the association phase and lack of return to baseline in the dissociation phase. To see if this was indeed the case, an additional BLI experiment was performed where biotinylated activin A was immobilised to streptavidin tips and immersed in a solution of GB1-H<sub>1</sub>L (Figure 19C). As GB1-H<sub>1</sub>L is monomeric with only one site of interaction for activin A, the drift in the association phase and lack of return to baseline in the dissociation phase should not be observed. However, this was not the case and continuous upward drift and lack of return to baseline were observed thus making this explanation unlikely.



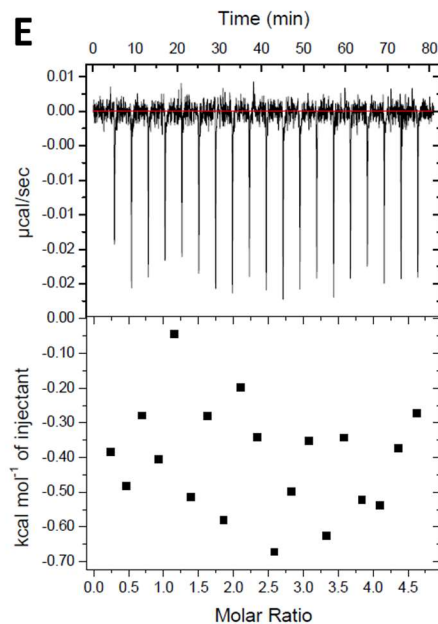


Figure 19:  $K_d$  determination of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L for mature activin A via BLI and ITC. A) The binding profiles generated for the interaction between GB1-H<sub>1</sub>LH<sub>2</sub> and mature activin A from BLI (top) and ITC (bottom). B) The binding profiles generated for the interaction between GB1-H<sub>1</sub>L and mature activin A from BLI (top) and ITC (bottom). In both cases the processed data along with the fitting is shown. C) The binding profiles generated for the interaction between GB1-H<sub>1</sub>L and mature activin A from BLI. This time mature activin A is immobilised on streptavidin tips and immersed in a solution of GB1-H<sub>1</sub>L D) The binding profiles generated for the interaction between GB1-H<sub>1</sub>L and mature activin A from BLI. This time mature activin A is directly immobilised onto AR2G tips and immersed in a solution of GB1-H<sub>1</sub>L E) Binding profile of double tagged GB1 to mature activin A. Here we see there is no significant binding observed.

The second possibility is that the peptide fusions or activin A undergo a secondary non-specific interaction with either the antibodies to which they are immobilised or the dextran matrix at the end of the fibre optic cable, thus resulting in a second kinetic event. To test if this was the case, activin A was directly immobilised onto AR2G tips and immersed into a solution of GB1-H<sub>1</sub>L (Figure 19D). As there are no antibodies present, if a secondary kinetic event is still observed it is likely due to interactions between the peptide fusions or activin A and the matrix. Again, this time there was a continuous upward drift throughout the experiment making it likely that either the peptide fusions or activin A interact non-specifically with the dextran matrix. As such, I decided though BLI may be appropriate for qualitative experiments, alternative methods should be used for quantitative experiments.

The alternate method chosen was isothermal titration calorimetry (ITC). Unlike BLI which uses kinetic methods to determine  $K_d$ , ITC determines the thermodynamic parameters of an interaction, from which  $K_d$  can be calculated. Furthermore, it is an in-solution measurement that requires no additional antibodies or dextran matrices. Peptide fusions at a concentration of 100  $\mu$ M were injected into an experimental cell containing solution of mature activin A at 4  $\mu$ M (here all mature activin A

concentrations refer to the mature activin A dimer and not a single protomer chains). From this, the heats of interaction were measured and a binding profile generated (Figure 19A and B). If the stoichiometry determined from this profile was within 10 % of a whole number, the ligand concentration was adjusted such that the stoichiometry was a whole number and  $K_d$  was calculated accordingly. Doubly tagged GB1 was again used as a control and it was determined that there was no significant interaction between it and mature activin A at the concentrations used (Figure 19E).

Table 7: Dissociation constants determined for the interaction between mature activin A and peptide fusions.

Peptide Fusion	$K_d$ / nM (BLI)	$K_d$ / nM (ITC)	Stoichiometry (ITC)	$\Delta H$ / kcal mol <sup>-1</sup>	$\Delta S$ / cal mol <sup>-1</sup> K <sup>-1</sup>	Unadjusted $K_d$ / nM (ITC)	Unadjusted Stoichiometry (ITC)
Pro-domain <sup>†</sup>	5.0 ± 0.2 <sup>104</sup>	—	—	—	—	—	—
GB1-H <sub>1</sub> LH <sub>2</sub>	1.25 ± 0.01	642 ± 76.1*	2	-13.8 ± 0.2*	-18.2*	624 ± 142*	1.86 ± 0.05*
GB1-H <sub>1</sub> L	3.11 ± 0.04	909 ± 112	2	-7.85 ± 0.14	1.32	909 ± 131	2.02 ± 0.04

<sup>†</sup> For GB1- H<sub>1</sub>LH<sub>2</sub>, the  $K_d$  and  $N$  value obtained is an average of three identical ITC experiments.

Analysis of the  $K_d$  values determined from ITC shows a stark difference to those determined via BLI (Table 7). The values obtained for GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L from ITC are over 100-fold lower affinity than those obtained via BLI (640 nM and 900 nM vs 1.25 nM and 3.11 nM). Affinity measured for GB1-



H<sub>1</sub>LH<sub>2</sub> (640 nM) is slightly stronger than that measured for GB1-H<sub>1</sub>L (900 nM) with the additional affinity possibly arising from increased interactions with the mature domain due to the presence of the  $\alpha$ 2-helix. However, the similarity in values confirms what was observed via BLI, that the  $\alpha$ 2-helix only contributes weakly towards binding. The stoichiometries determined for both fusions also show that in both cases, two moles of peptide fusion interact with every mole of activin A. This is expected due to the dimeric nature of activin A meaning two peptide binding sites are present on every molecule of protein.

Comparison of the  $K_d$  values obtained from ITC with the reported  $K_d$  value of the pro-domain is poor (5 nM – obtained via BLI).<sup>104</sup> This disparity may be due to additional interactions resulting in a higher affinity between the pro-domain and mature domain or due to the difference in methodology. Incidentally, the  $K_d$  values obtained via BLI compare very well with that reported for the pro-domain (1.3 nM and 3.1 nM vs 5.0 nM). Therefore, unless a  $K_d$  value for the pro-domain is obtained via ITC it is difficult to make a full comparison of the system. Interestingly, the  $K_d$  values obtained via ITC do compare well with the predicted  $K_d$  as reported in Walton *et al.* (100 nM) however this value is obtained from projections made based on inhibition data and no kinetic or thermodynamic experiments were performed.<sup>238</sup>

### 3.6 Effect of Chemical Stapling on Binding of Synthetic Peptides to Mature Activin A (In Conjunction with Sam Rowe)

In addition to the peptide fusions described so far, the effect of chemical stapling on the binding of synthetic peptides to activin A was also investigated. These synthetic peptides were based on H<sub>1</sub>L, lacking only the two C-terminal residues. The staple that was chosen for this work is N-(4,6-divinyl-1,3,5-triazin-2-yl)-N-methylglycine. This molecule contains two alkenes which can undergo a Michael addition with cysteine residues to form thioethers, therefore by including two cysteine residues in the peptide sequence a staple can be introduced (Figure 20A). To select which residues should be mutated to cysteines, a structural analysis was performed (Figure 20B). Analysis of the structure of the pro-mature activin A complex reveals that the residues that could likely be stapled with minimum disruption to interactions between the pro- and mature domains were residues 54 and 61 in an i+7 stapling. This was supported by modelling performed by Tan Yaw Sing (A\*STAR, Singapore, unpublished data),<sup>239</sup> that also suggested the most optimal residues for introducing a staple to be M54 and H61. Additionally, this analysis also suggested that the mutations V58I and H61F (assuming residues 61 is not the site of stapling) may improve the affinity of the peptides for activin through an optimised spatial fit and increased hydrophobic interactions.

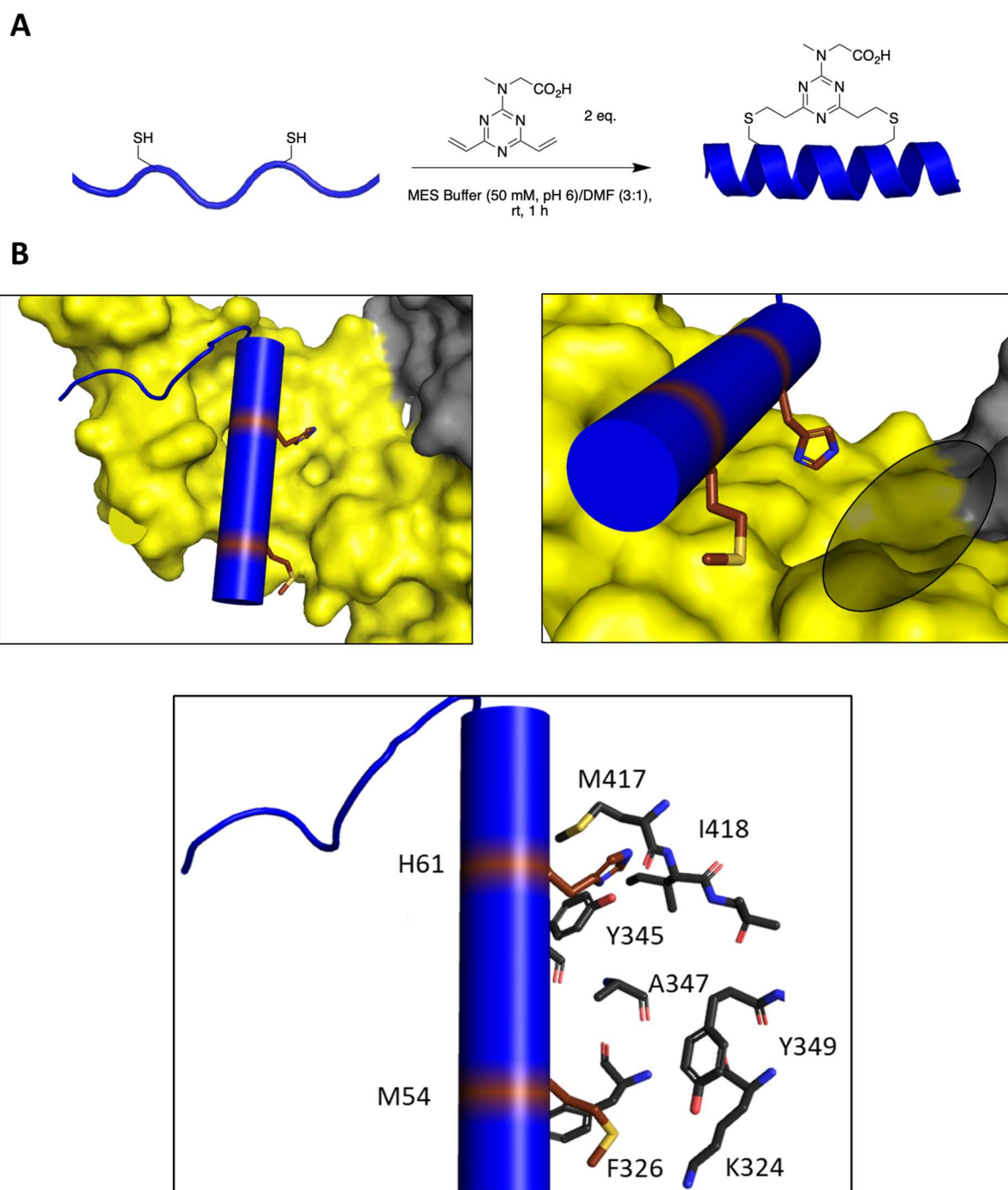


Figure 20: Stapling of synthetic activin A derived peptides. A) Stapling of a peptide through the Michael addition of *N*-(4,6-divinyl-1,3,5-triazin-2-yl)-*N*-methylglycine to two cysteine residues (Figure provided by Sam Rowe). B) Structural analysis of the pro-mature activin A complex using a representation of the full-length peptide sequence (blue) with residues M54 and H61 highlighted. In the right picture, a possible surface for the staple to interact with activin A is highlighted with shading. In the below picture, potential interacting residues on mature activin A are shown (here carbon atoms are displayed in black).

**Table 8: Synthetic peptides screened against mature activin A (synthesised by Sam Rowe). Here X denotes the site of stapling (highlighted in red), and all mutations are highlighted in cyan.  $K_d$  values were determined via ITC.**

Peptide	Sequence	$K_d$ / nM	Stoichiometry	$\Delta H$ / kcal mol <sup>-1</sup>	$\Delta S$ / cal mol <sup>-1</sup> K <sup>-1</sup>
H <sub>1</sub> L	QPEMVEAVKKHILNMLHLKKRPDVTQ	–	–	–	–
Ac-H <sub>1</sub> ΔL	Ac-QPEMVEAVKKHILNMLHLKKRPDV-NH <sub>2</sub>	395 ± 77.1	2	-10.6 ± 0.2	-6.23
Ac-H <sub>1</sub> ΔL <sub>mut</sub>	Ac-QPEMVEAIVKKHILNMLHLKKRPDV-NH <sub>2</sub>	194 ± 23.4	2	-11.1 ± 0.1	-6.63
Ac-H <sub>1</sub> ΔL-53St	Ac-QPEMVEAVKKHILNMLHLKKRPDV-NH <sub>2</sub>	–	–	–	–
Ac-H <sub>1</sub> ΔL-54St	Ac-QPEMVEAVKKHILNMLHLKKRPDV-NH <sub>2</sub>	–	–	–	–
Ac-H <sub>1</sub> ΔL-56St	Ac-QPEMVEAVKKHILNMLHLKKRPDV-NH <sub>2</sub>	2050 ± 291	2	-7.97 ± 0.25	-0.711


In total, five peptides were provided by Sam Rowe (Department of Chemistry, University of Cambridge) at >95 % purity, of which three contained a chemical staple (Table 8). The binding of these peptides to mature activin A was then analysed via ITC (Table 8, Appendix 4). The  $K_d$  values obtained for Ac-H<sub>1</sub>ΔL<sub>mut</sub> (190 nM) and Ac-H<sub>1</sub>ΔL (400 nM) are slightly lower than that of GB1-H<sub>1</sub>L (900 nM), binding to mature activin A with approximately 2-fold/ 4-fold higher affinity. This may be due to a reduced conformational change afforded by the smaller peptide upon binding. The  $K_d$  observed Ac-H<sub>1</sub>ΔL<sub>mut</sub> compared to Ac-H<sub>1</sub>ΔL supports the modelling data suggesting that the mutations V58I and H61F increase interactions between the peptide and mature activin A through a more spatially optimised fit. However, caution must be taken when drawing conclusions from all these values due to their similarity. Interestingly, the addition of staples to Ac-H<sub>1</sub>ΔL<sub>mut</sub> appear to reduce its ability to interact with the mature domain. Of the three stapled peptides tested, no significant binding was observed for two of them, with the third Ac-H<sub>1</sub>ΔL-56St, showing a 10-fold decrease in affinity (2.1 μM). Surprisingly Ac-H<sub>1</sub>ΔL-54St shows no binding to the mature domain despite the structural analysis and computational modelling identifying residues M54 and H61 as the optimal sites for stapling. This suggests that additional factors that were not considered in these analyses must be present, such as the presence of steric clashes between the staple and mature domain that are not adequately modelled using the pro-mature structure. In addition, no binding was seen for the second stapled peptide Ac-H<sub>1</sub>ΔL-53St. The pro-mature structure suggests in this peptide, the staple would be highly solvent exposed, which given the hydrophobic nature of the staple, may be energetically unfavourable and hinder binding. The only observed binder of the three, Ac-H<sub>1</sub>ΔL-56St, shows a reduced affinity which again could be due to steric clashes between the staple and C-terminal residues of the peptide, reducing its ability to bind. Another possibility is that the staple introduces a conformation constraint that significantly hinders binding. Indeed, this is a possibility for all three peptides though further experimentation would be required to determine if this is the case. Overall however, stapling with N-(4,6-divinyl-1,3,5-triazin-2-yl)-N-methylglycine results in a destabilisation of the interaction between the peptide and mature activin A compared to the unstapled peptide.

In addition to the ITC experiments, co crystallisation was attempted with all the peptides discussed in this section (both stapled and unstapled). However, these experiments failed to produce any crystals of a peptide-mature activin A complex.

### 3.7 Determination of the Minimal Binding Peptide of the Activin A Pro-domain

Previously, I have shown that the  $\alpha$ 1-helix and the loop are the key motifs for high affinity binding. However, this was the shortest peptide used that contained both these epitopes. In order to determine what the minimal binding peptide for high affinity binding to mature activin A was, a series of N-terminal and C-terminal truncations were generated (Table 9).

**Table 9: N- and C-terminal truncations of H<sub>1</sub>L.**



Peptide	Sequence
GB1-H <sub>1</sub> L	QPEMVEAVKKKHI LNMLHLKKRPDVTQ
GB1-P52-Q76	PEMVEAVKKKHI LNMLHLKKRPDVTQ
GB1-E53-Q76	EMVEAVKKKHI LNMLHLKKRPDVTQ
GB1-M54-Q76	MVEAVKKKHI LNMLHLKKRPDVTQ
GB1-V55-Q76	VEAVKKKHI LNMLHLKKRPDVTQ
GB1-E56-Q76	EAVKKKHI LNMLHLKKRPDVTQ
GB1-Q51-T75	QPEMVEAVKKKHI LNMLHLKKRPDVT
GB1-Q51-V74	QPEMVEAVKKKHI LNMLHLKKRPDV
GB1-Q51-D73	QPEMVEAVKKKHI LNMLHLKKRPD
GB1-Q51-P72	QPEMVEAVKKKHI LNMLHLKKRP
GB1-Q51-R71	QPEMVEAVKKKHI LNMLHLKKR
GB1-H <sub>1</sub>	QPEMVEAVKKKHI LNMLHL

In addition, to confirm whether the  $\alpha$ 1-helix alone bound to mature activin A, an additional construct corresponding to the residues Q51-H67 (GB1-H<sub>1</sub>) was also used. These peptide fusions, along with GB1-H<sub>1</sub>L, were then analysed via biolayer interferometry using solutions of 100 nM and 1  $\mu$ M of mature activin A (Figure 21). Analysis of the binding profiles of the C-terminal deletions reveals a significant loss in affinity at R71, with P72 being observed to be the last residue that results in high affinity binding. As P72 is present in the loop rather than the  $\alpha$ 1-helix, this indicates that the initial loop residues are required for binding to mature activin A. This is further supported by the binding

profile of GB1-H<sub>1</sub> which shows no significant binding to the mature domain. As such, it was decided that the C-terminal residue of the minimum peptide should be P72.

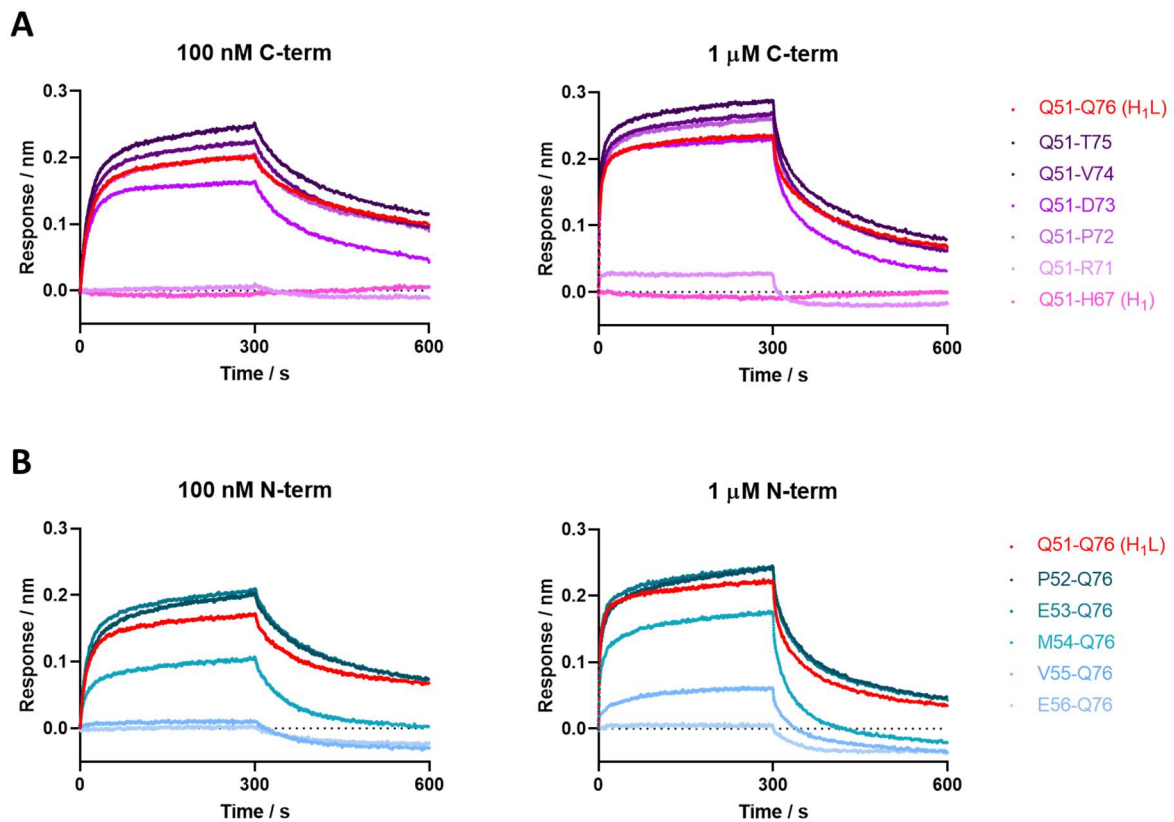


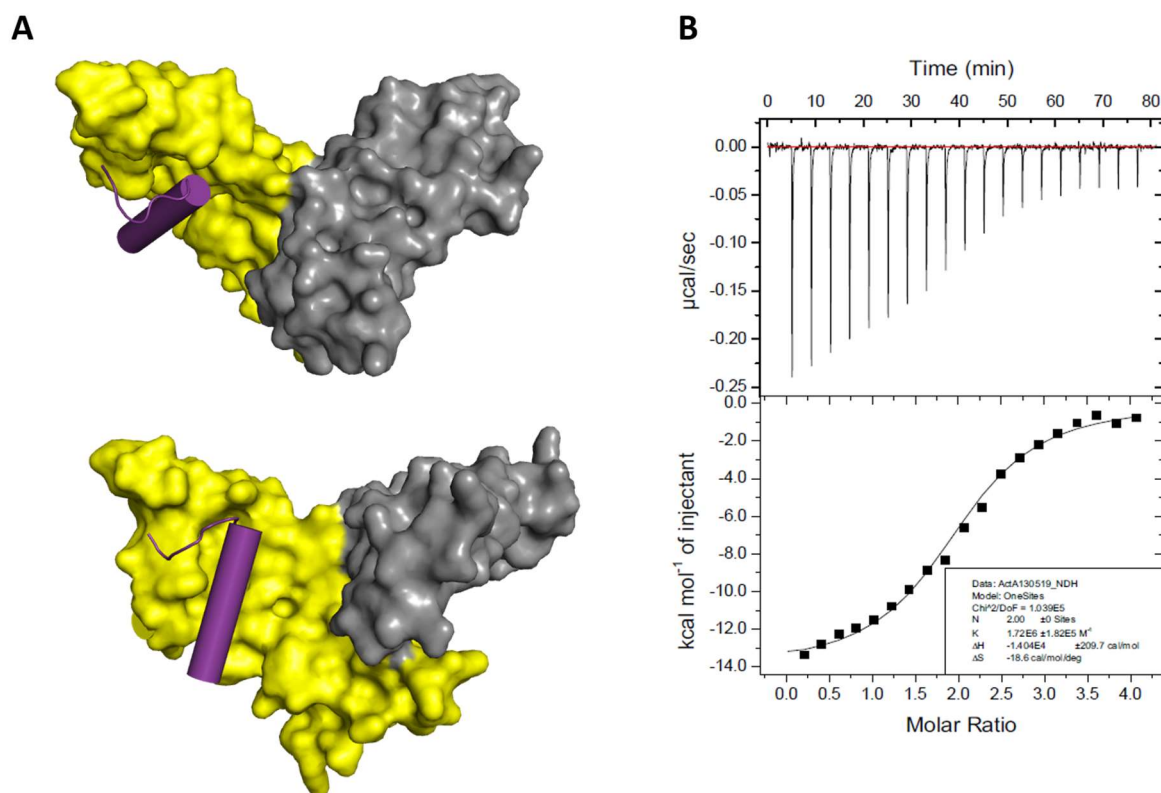
Figure 21: Determination of the minimal binding peptide of the activin A pro-domain via BLI. A) Binding profiles for the C-terminal deletion peptides to 100 nM and 1 μM mature activin A as determined via BLI. B) Binding profiles for the N-terminal deletion peptides to 100 nM and 1 μM mature activin A as determined via BLI.

The binding profiles of the N-terminal deletions are more difficult to interpret. Of the five peptides, GB1-P52-Q76 and GB1-E53-Q76 show a similar binding profile to GB1-H<sub>1</sub>L (albeit with a slightly higher signal) in both 100 nM and 1 μM of mature activin A. By contrast, the shortest peptide GB1-E56-Q76 shows very low binding at both concentrations. It is the two next shortest peptides however, GB1-M54-Q76 and GB1-V55-Q76, that make determining the minimal peptide challenging. GB1-M54-Q76 shows binding to mature activin A at both concentrations, however the association response is lower and the dissociation quicker than observed for the longer peptides. GB1-V55-Q76 on the other hand shows very low binding at 100 nM however this binding shows an increase at 1 μM. This indicates that the loss of E53 and to a lesser extent M54 does not prevent the peptide from being able to bind to mature activin A, however it does likely reduce the affinity. Thus, as GB1-E53-Q76 was the shortest peptide that closely mimicked the binding of GB1-H<sub>1</sub>L, it was decided that the minimal peptide should be designated as E53-P72.

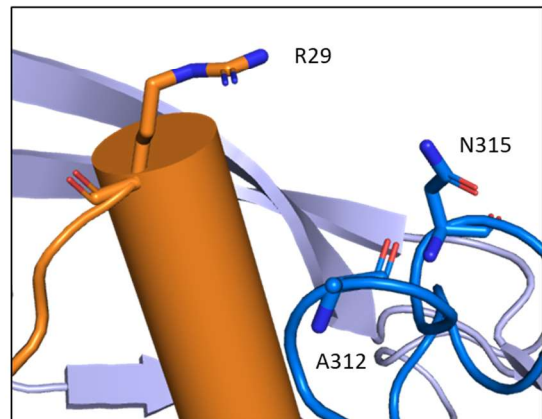
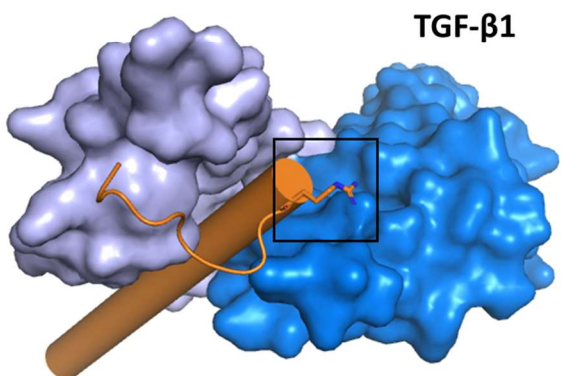
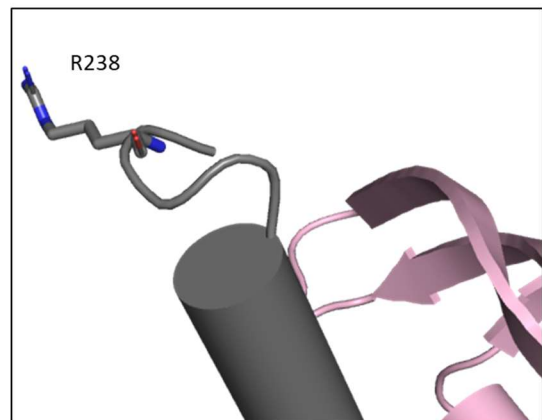
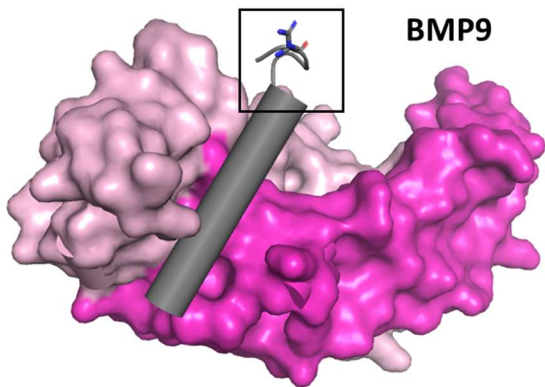
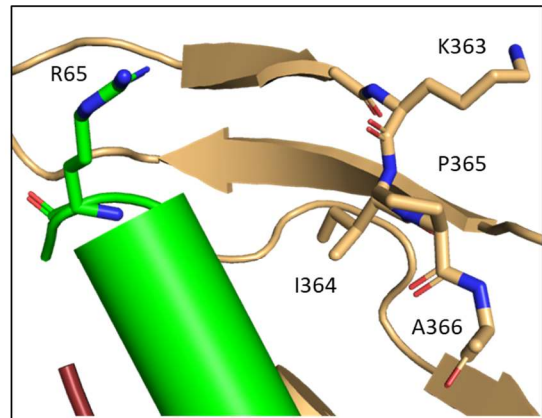
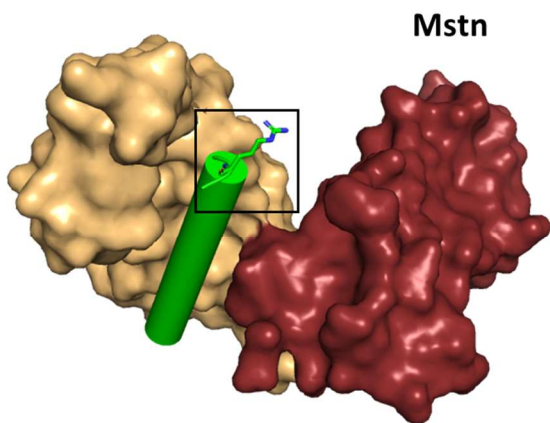
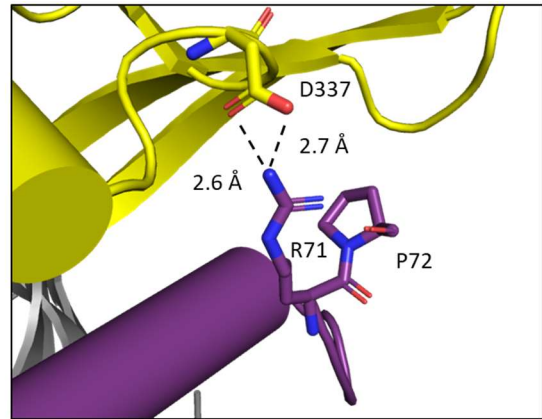
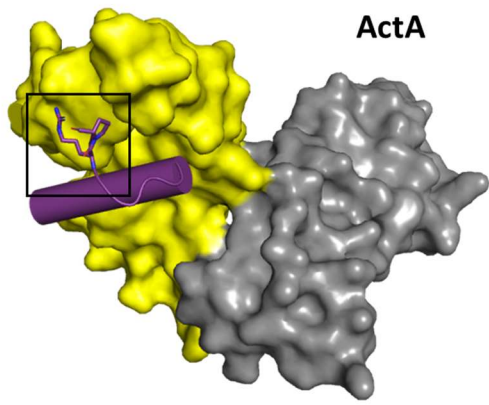
To confirm if GB1-E53-P72 was indeed the minimal high affinity binding peptide, this peptide fusion was then generated and binding to mature activin A was confirmed via ITC as described previously. From this a  $K_d$  was determined of 580 nM (Table 10,). This value compares well with that of GB1-H<sub>1</sub>L (900 nM) and the unmutated synthetic peptide Ac-Q51-V74-NH<sub>2</sub> (395 nM). This therefore suggests that E53-P72 is the minimal binding peptide for binding to mature activin A.

**Table 10: Dissociation constant determined for mature activin A and the minimal binding peptide as determined via ITC.**

Peptide Fusion	Sequence	$K_d$ / nM	Stoichio-metry	$\Delta H$ / kcal mol <sup>-1</sup>	$\Delta S$ / cal mol <sup>-1</sup> K <sup>-1</sup>
GB1-E53-P72	EMVEAVKKHILNMLHLKGRP	581 ± 68.8	2	-14.0 ± 0.2	-18.6



**Figure 22: Determination of the  $K_d$  for the minimal binding peptide..** A) Representation of E53-P72 (purple) interacting with mature activin A (yellow and grey) based on the structure of the pro-mature activin A complex. B) The binding isotherm and thermogram generated for GB1-E53-P72 interacting with mature activin A.



*Figure 23: End-of-  $\alpha$ 1-helix arginine in TGF- $\beta$  family members. Arginine residues at the terminus of the  $\alpha$ 1-helix of activin A, myostatin, BMP9, and TGF- $\beta$ 1. Apart from activin A, the other TGF-  $\beta$  family members have no aspartate or glutamate residues in their immediate vicinity to form a salt bridge with. In each case, the  $\alpha$ 1-helix and the initial loop residues of the pro-domain have been extracted. Nearby residues are also represented as sticks.*

In addition to determining the minimal binding peptide, the C-terminal truncations also revealed that the initial loop residues are vital for binding, with truncations not containing residue P72 showing a large drop in affinity for mature activin A. Analysis by Tan Yaw Sing (A\*STAR, unpublished data) of the energetic contributions to binding of each residue in the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix region indicated that contribution comes from the formation of a salt bridge between R71 and D337. Therefore, it is possible that the loss of P72 may affect the orientation of these two residues thus resulting in destabilisation of the salt bridge and loss of binding. The presence of an arginine residue at the end of the  $\alpha$ 1-helix or start of the loop is not uncommon for TGF- $\beta$  family members, however an arginine-aspartate salt bridge interaction appears to be relatively unique to activin A (Figure 23). The structures of pro-TGF- $\beta$ 1 (PDB ID: 3RJR<sup>102</sup>), pro-BMP9 (PDB ID: 4YCG<sup>107</sup>) and pro-myostatin (PDB ID: 5NTU<sup>108</sup>) show the presence of an arginine residue at the end of the  $\alpha$ 1-helix that is either not proximal to an aspartate / glutamate residue to form a salt bridge or oriented away from the mature domain. Alignment of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix region of activin A with other TGF- $\beta$  family members shows that an arginine residue at the end of the  $\alpha$ 1-helix is conserved across the activins and many BMPs. Due to the high sequence similarity of the activins and ability of H<sub>1</sub>LH<sub>2</sub> to interact with activin B in a similar manner to activin A, it is likely this interaction is preserved here. However, with lack of structures or alanine scanning for more of the BMPs, it is difficult to say whether this interaction occurs. As such this salt bridge may be a unique interaction to the activins that facilitates interactions between their pro- and mature domains.

### 3.8 Mutagenesis of the Minimal Binding Peptide of the Activin A

#### Pro-domain

To test if the minimal binding peptide could be optimised further, a structural analysis was performed. This was to determine whether mutations could be introduced that optimise the spatial fit of the peptide into its binding groove and maximise interactions. Analysis of residues E53-P7Z of the pro-mature complex structure identified three sites for optimisation – The presence of a relatively flat hydrophobic surface adjacent to M54, a large channel adjacent to V55, and hydrophobic pocket that may not be spatially optimized adjacent to V58 (Figure 24). Based on this, five mutant constructs were designed (Table 11).



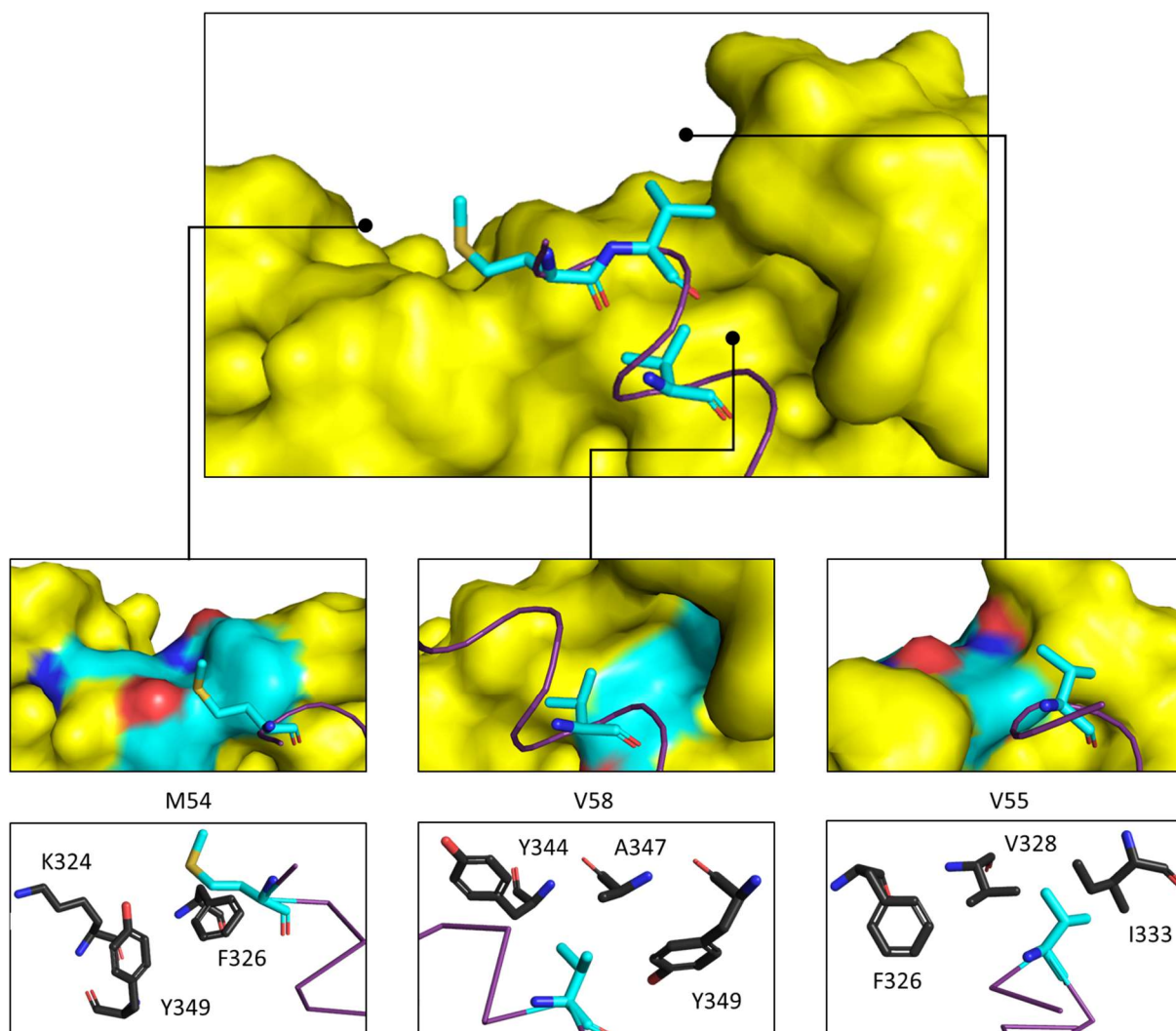


Figure 24: Selected residues for mutation in the activin A minimal peptide. A structural Analysis of the pro-mature activin A complex using only pro-domain residues E53-P72. Here we see the three sites proposed for optimisation highlighted. Here, residues selected are shown in cyan. Sites for optimisation on the mature domain are also shown in cyan (carbon), red (oxygen) and blue (nitrogen) respectively. In the below figures, carbon atoms in the mature domain are shown as black to distinguish them from those in the mutated peptide.

Table 11: Dissociation Constants determined for mature activin A and mutant minimal binding peptides via ITC.

Mutation	Sequence	$K_d$ / nM	Stoichiometry	$\Delta H$ / kcal mol <sup>-1</sup>	$\Delta S$ / cal mol <sup>-1</sup> K <sup>-1</sup>
GB1-E53-P72	EMVEAVKKHILNMLHLKGRP	581 ± 68.8	2	-14.0 ± 0.2	-18.6
M54Y	EYVEAVKKHILNMLHLKGRP	1590 ± 338	2	-14.4 ± 0.5	-21.7
M54W	EWVEAVKKHILNMLHLKGRP	1360 ± 131	2	-11.8 ± 0.2	-12.8
V58I	EMVEAVKKHILNMLHLKGRP	654 ± 85.0	2	-14.1 ± 0.2	-19.0
V58L	EMVEAVKKHILNMLHLKGRP	585 ± 43.3	2	-12.2 ± 0.1	-12.3
V55Y	EMVEAVKKHILNMLHLKGRP	690 ± 87.3	2	-12.3 ± 0.2	-12.9

These mutants were then screened via ITC as described previously (Table 11, Appendix 4). Analysis of the resulting binding profiles shows that the mutation of methionine 54 to tyrosine or tryptophan appeared to destabilise peptide binding, with an approximate three-fold reduction in affinity. This suggests that this surface is not as available for the large aromatic groups as our analysis suggested. By contrast, the  $K_d$  values observed for the mutations V55Y, V58I, and V58L were highly similar to that of the original peptide (690 nM, 650 nM, and 590 nM vs 580 nM). The similarity means the mutations introduced do not significantly optimise interactions. However, it may be the case that combinations of mutations would have a greater impact on affinity.

Incidentally, the mutation V58I is one of the two mutations present in the synthetic peptide Ac-Q51-V74-NH<sub>2</sub>\*<sup>‡</sup> which had a  $K_d$  value of 190 nM. This peptide had a two-fold higher affinity for mature activin A than the equivalent native sequence. As there is no significant difference between the mutant V58I and the native sequence in this experiment, this suggests that it may be the mutation H61F that contributes the most to the observed higher affinity of the synthetic peptide. However, it is difficult to say due to the common observation that the interactions of multiple mutations accumulate to produce a more pronounced effect on affinity, which may also be the case here. Due to the time constraints of the project it was not possible to test the mutation H61F in the GB1 fusion format.

### 3.9 GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L Inhibit Activin A Signalling with Micromolar Efficacy

Based on the binding data, it was decided to investigate whether GB1-H<sub>1</sub>LH<sub>2</sub>, GB1-H<sub>1</sub>L and the synthetic peptide Ac-H1ΔL<sub>mut</sub> could inhibit mature activin A signalling. As these peptides correspond to the key binding motifs of the pro-domain and bind in the vicinity of the receptor binding sites, they should compete with the type I and II receptors for mature domain binding. Therefore, these peptides may be able to inhibit activin A signalling.

To evaluate the ability of these peptides to inhibit activin A signalling, I used an activin A responsive luciferase assay in HEK293T cells. In this assay the cells were transfected with plasmids encoding a Smad 2/3 responsive firefly luciferase and a constitutively expressed *Renilla* luciferase gene. As activin A signals through the Smad 2/3 pathway the expression levels of firefly luciferase are dependent on the level of activin A signalling, allowing for the effect of the pro-domain fragments to be investigated quantitatively. Constitutive expression of *Renilla* luciferase allows normalisation of the data and takes into account variations in cell numbers and transfection efficiency. The assay was performed at the EC<sub>80</sub> concentration of 100 pM as determined by Wang *et al.* for the same system,<sup>104</sup> and a range of

concentrations of GB1-H<sub>1</sub>LH<sub>2</sub>, GB1-H<sub>1</sub>L and doubly tagged GB1 were used from 20 μM – 0.9 nM (for Ac-H1ΔL<sub>mut</sub>, a concentration range of 100 μM – 0.9 nM was used).

Analysis of the inhibition profile of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L shows a general inhibitory trend for both (Figure 25), however IC<sub>50</sub> values could only be approximated as the full inhibition curve could not be obtained (cell death occurred at higher concentrations of peptide). These profiles show GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L inhibit with low micromolar efficacy with IC<sub>50</sub> values of approximately 4.4 μM and 3.8 μM respectively (Table 12). By contrast, doubly tagged GB1 lacking activin A peptide showed no change in activin A response, as expected. This indicates that the His-tag, Strep-tag II, and GB1 domain do not contribute towards inhibition and that inhibition was solely due to H<sub>1</sub>LH<sub>2</sub> and H<sub>1</sub>L sequences, thus confirming that the pro-domain derived peptides inhibit activin A signalling. In the case of Ac-H1ΔL<sub>mut</sub>, a general inhibitory trend was observed, however it was not possible to obtain an IC<sub>50</sub> value with any accuracy. Furthermore, analysis of the normalised responses for each concentration of Ac-H1ΔL<sub>mut</sub> suggests it appears to be a much weaker inhibitor of activin A signalling than GB1-H<sub>1</sub>LH<sub>2</sub> or GB1-H<sub>1</sub>L. As such, subsequent further analysis of Ac-H1ΔL<sub>mut</sub> was not performed.

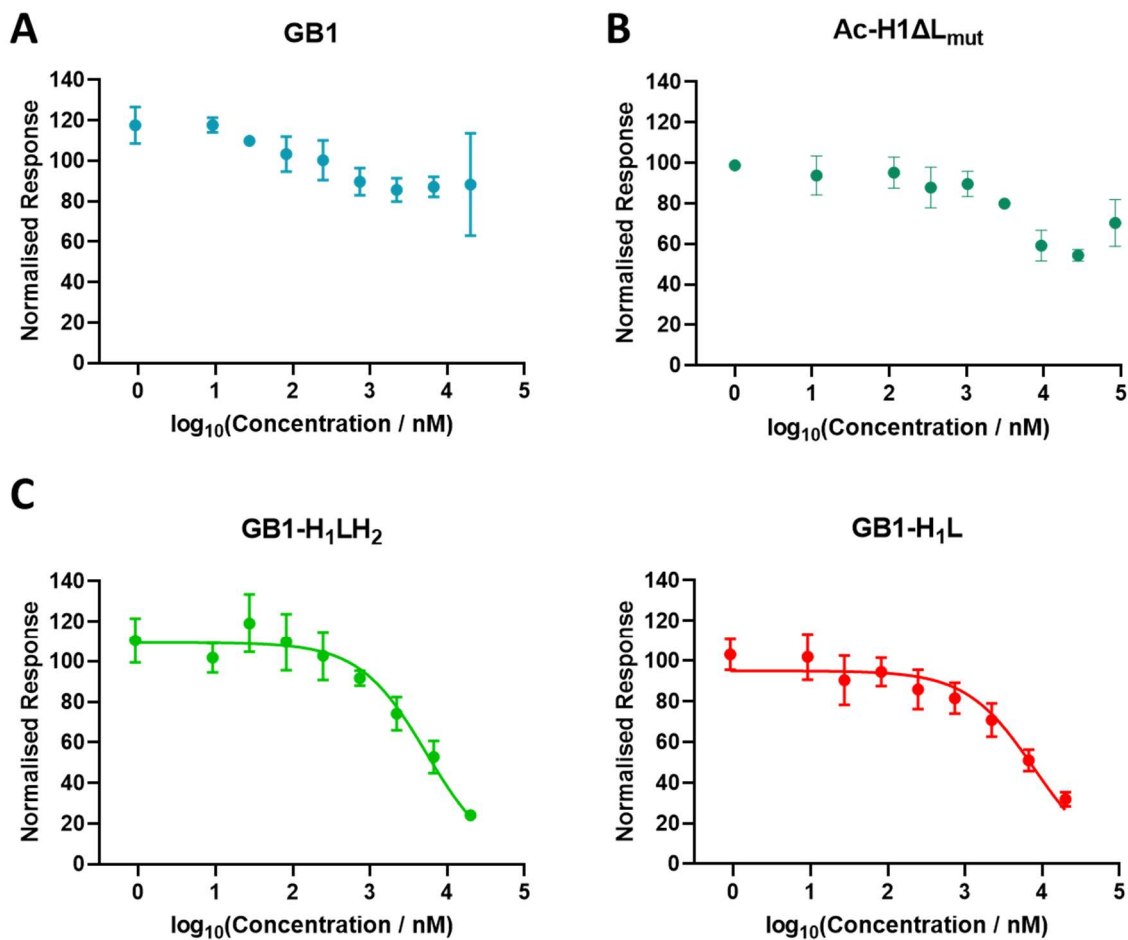


Figure 25: Luciferase Assays of GB1, GB1-H<sub>1</sub>LH<sub>2</sub>, and GB1-H<sub>1</sub>L. A) The inhibitory profile of doubly tagged GB1 inhibiting mature activin A signalling in HEK293T cells. B) The inhibitory profile of Ac-H1ΔL<sub>mut</sub> inhibiting mature activin A signalling in HEK293T cells. C) The inhibitory profiles of GB1-H<sub>1</sub>LH<sub>2</sub>, and GB1-H<sub>1</sub>L inhibiting mature activin A signalling in HEK293T cells. In both cases the log<sub>10</sub> of the concentration in nM is plotted against the normalised ratios of firefly luminescence: renilla luminescence. Inhibitory profiles are generated from three technical repeats. Error bars correspond to one standard deviation of uncertainty.

**Table 12: IC<sub>50</sub> values determined for GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L inhibiting activin A signalling.**

Inhibitor	Log <sub>10</sub> (IC <sub>50</sub> / nM)	IC <sub>50</sub> / nM
GB1-H <sub>1</sub> LH <sub>2</sub>	3.64 ± 0.14	4,400
GB1-H <sub>1</sub> L	3.58 ± 0.16	3,800

The similarity between the IC<sub>50</sub> values of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L suggests that the α2-helix contributes little towards inhibition. Taking this into account, it is likely that inhibition occurs through the α1-helix of the peptides preventing the type I receptors from binding to the type II-ActA complex. This is due to the binding site of the α1-helix and the putative binding site of ALK4 overlapping on the mature domain. The observation that the main inhibitory motifs are the α1-helix and the loop compares well with the previous observation that the motifs required for high affinity binding are the same epitopes. This indicates that these motifs must be present to achieve inhibition when designing inhibitors based on the pro-domain. Overall, this data shows that peptides derived from the pro-domain can inhibit activin A signalling, and the key motif responsible for this inhibition is the α1-helix and the loop.

Incidentally, the IC<sub>50</sub> values of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L may be a good approximation of that of activin A's pro-domain. Previous attempts to measure IC<sub>50</sub> values for the pro-domain have failed with no inhibitory trend often being reported. This is despite evidence suggesting the pro-domain can inhibit the downstream effects of activin A. The limitation in both attempts has been the highest concentration of pro-domain (10 nM and 100 nM respectively).<sup>236,240</sup> Wang *et al.* has shown that the pro-domain alone is likely monomeric.<sup>104</sup> As GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L are both monomeric and contain the key interacting pro-domain residues, it is likely that the IC<sub>50</sub> value of the pro-domain is similar to that obtained for these peptides (4.4 μM and 3.8 μM).

### 3.10 Conclusions

In this section I have described the generation of a series of peptide fusions based on the α1-helix-loop-α2-helix-β-hairpin motif of activin A's pro-domain and characterised their binding affinity for mature activin A. From this I have determined that the α1-helix and the loop play a crucial role allowing the pro-domain to bind the mature domain with high affinity whereas the α2-helix and the β-hairpin appear to contribute little. I then tested a series of stapled peptides based on the α1-helix and the

loop, finding that stapling appeared to have a negative effect on binding, before determining the minimal binding peptide to be residues E53-P72. Finally, I tested the effect of two of the original peptide fusions, GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L, on activin A signalling. From this I found that they both show an inhibitory trend, appearing to inhibit activin signalling with micromolar affinity (approximately 4.4  $\mu$ M and 3.8  $\mu$ M respectively). This therefore shows that pro-domain derived peptides are a viable strategy for developing inhibitors of activin A signalling.

There are multiple ways in which the optimisation of these peptides could be attempted. However, in the next section I will discuss how dimerization impacts the inhibitory properties of these peptide fusions. I will test both covalent and non-covalent optimisations based on GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L before going on to investigate the specificity of the one of the peptide fusions against other members of the TGF- $\beta$  family.



---

# Dimeric Approach

---

## 4.1 Introduction

In the previous section I looked how peptide fusions based on the key interacting epitopes of the activin A pro-domain bound to and inhibited mature activin A signalling. However, this inhibition was observed only at low micromolar level in cellular assays. One reason for this low efficacy may be due to the monomeric nature of the peptides. In “3.5 Identification of the Key Binding Motifs in the Activin A Pro-domain”, it was shown that both GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L interact monomerically with mature activin A, with a stoichiometry of 2. As these peptides are thought to bind mature activin A in the vicinity of its type I and type II receptors binding sites, it may be the case that they have difficulty competing with the multiple receptors on the cell surface for activin A binding due to an avidity effect created by receptor clustering. Therefore, dimerization of the peptides may moderate this difficulty. It has long been demonstrated that for molecules with multiple binding sites, increasing the valency of a binding partner can result in an increase in affinity due to avidity, and in the case of the pro-domain derived peptides, this may also lead to an increase in inhibitor efficacy. Therefore, to improve the efficacy of the peptides, I decided to explore further optimisation through dimerization.

The use of dimeric inhibitors to inhibit activin A signalling has been reported in a number of instances. Pro-domain ligand traps such as those reported in Chen *et al.* are dimeric due to the dimeric Fc fragment to which the pro-domain is fused (IC<sub>50</sub> value of most potent ligand trap = 5 nM).<sup>236</sup> Furthermore, Makanji *et al.* achieved highly potent inhibition of activin A signalling through the fusion of the N-terminal helices of the activin A pro-domain to the covalently dimeric TGF-β1 shoulder region (IC<sub>50</sub> = 10.3 nM).<sup>188</sup> Outside of activin A inhibition, dimeric inhibitors have also been demonstrated to be efficacious inhibitors for the related TGF-β family ligand, myostatin. Jiang *et al.* observed that GST-fused peptides based on myostatin pro-domain showed similar inhibitory activity to that of myostatin pro-domain (IC<sub>50</sub> value of the pro-domain = 0.9 nM, no IC<sub>50</sub> value was reported for the GST-fused peptide).<sup>108,206</sup> This contrasts with Takayama *et al.* whose monomeric peptide inhibitors of myostatin were significantly weaker than the pro-domain (IC<sub>50</sub> value of the original mouse derived peptide = 3.5 μM).<sup>189</sup> Therefore, optimisation through dimerization may present a viable strategy for increasing inhibitor efficacy for activin A.

There are several possible ways to induce dimerization, both covalently and non-covalently. One such way is to create a fusion with a protein or motif that dimerizes. The dimerization and docking domain of the cAMP-dependent protein kinase type II-α regulatory subunit (PKADD domain) is one such potential fusion. This small domain dimerizes with a second regulatory subunit and binds to the PKA binding motif of A-kinase anchor proteins. This localises the regulatory subunits and the inactive protein kinase A complex in the cell for signalling.<sup>241,242</sup> PKADD domain consists of a disordered region



followed by two helices linked by a short linker, each with a highly hydrophobic and highly hydrophilic face. In solution, these helices dimerize to form an X-type four helical bundle. This dimerization is facilitated by the formation of a hydrophobic core by the hydrophobic faces of the helices.<sup>243</sup> This X-type helical arrangement means the two N-termini of the PKADD domain subunits are oriented towards the opposite direction, thus making the system ideal for connecting two peptides that are also oriented in opposite directions. Furthermore, with this region only corresponding to 45 residues, the PKADD domain represents a relatively small fusion that can be added to proteins to induce dimerization.

Another strategy to induce dimerization is to introduce a disulfide bridge between the two molecules. This can be generated through the addition of a cysteine residues to a peptide and placing it in oxidising conditions. This strategy has been shown to have success in the case of HIV-1, where facilitating dimerization through the introduction of a disulfide bond resulted in an increase in anti-HIV-1 activity.<sup>244</sup> One common problem when introducing cysteines is the potential for multiple products due to disulfide formation with other cysteines in the protein. However, in the case of the activin peptide fusions, neither GB1-H<sub>1</sub>LH<sub>2</sub>, GB1-H<sub>1</sub>L or the PKADD contain any cysteine residues and therefore disulfide mediated dimerization should be very feasible.

In this section I will look at the optimisation of the peptide fusions as inhibitors of activin A signalling through dimerization before testing specificity of the most potent inhibitors against other TGF- $\beta$  family members.

## 4.2 Design of PKADD Domain Fused Peptides

In order to investigate the effect of dimerization on inhibitor potency, the two peptides GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L first needed to be dimerized. The method selected for this was to introduce the PKADD domain to the C-terminus of the peptide fusions (for mock model, see Figure 26). Fusions were designed such that residues 2-46 of the PKADD domain (Uniprot P13861) consisting of an eight-residue linker region and the two helices that form an X-helical bundle were inserted between the pro-domain derived peptide and the C-terminal purification tag. Due to the vector employed, the termini of the purification tags were swapped such that the N-terminal tag was Strep-tag II and the C-terminal tag a His-tag. The resulting peptides were subsequently named GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and GB1-H<sub>1</sub>L-PKADD (for full construct sequences, see Appendix 2).

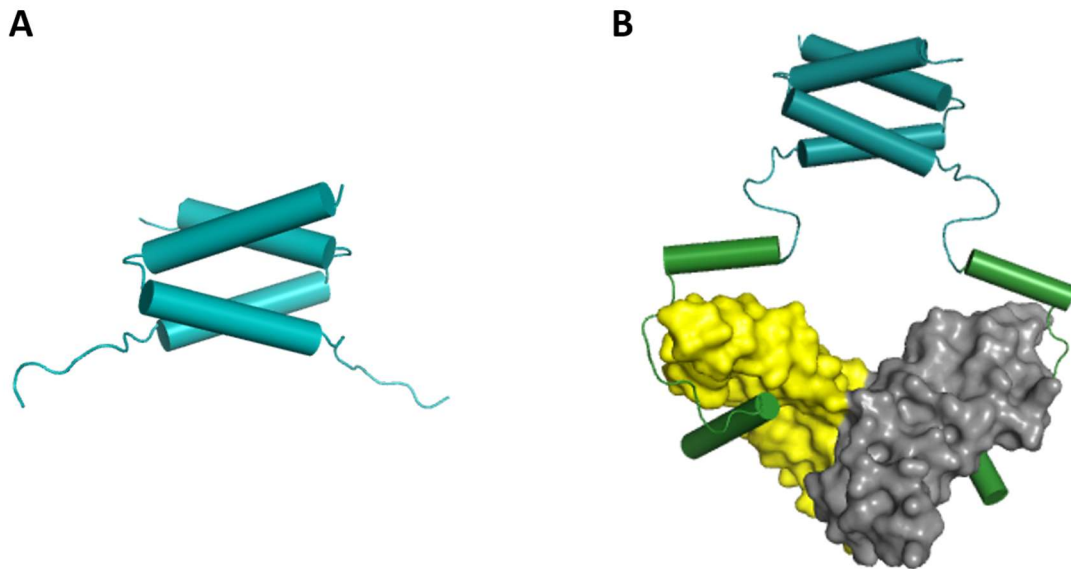


Figure 26: PKADD domain fused peptides. A) The PKADD domain dimer (PDB ID: 2DRN). B) A schematic mock model of H<sub>1</sub>LH<sub>2</sub>-PKADD generated from the pro-mature activin A structure (PDB ID: 5HLZ) and PKADD domain dimer. Here mature activin A is shown in yellow and grey, H<sub>1</sub>LH<sub>2</sub> in green, and the PKADD domain in teal.

### 4.3 Cloning of PKADD Domain Fused Peptides

To generate the PKA fused peptides, an appropriate vector was first generated. This vector was designed as discussed in “4.2 Design of PKADD Domain Fused Peptides” with an N-terminal Strep-tag II, followed by a GB1 domain, a multiple cloning site, the PKADD domain and a C-terminal His-tag (Figure 27). Generation of the vector was achieved through modification to the vector pOP2H (used in the Hyvönen Group). Primers were designed with appropriate 5’ and 3’ restriction site to generate DNA encoding the Strep-tag II, GB1 domain and the PKADD domain. DNA was then amplified via PCR using pOP4BP and pGEX-PKADD as a template. The resulting PCR products were subsequently ligated into pOP2H to generate the vector pOP6BPA. DNA encoding H<sub>1</sub>LH<sub>2</sub> and H<sub>1</sub>L was generated in a similar manner using GB1- H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L as templates and ligated into pOP6BPA.

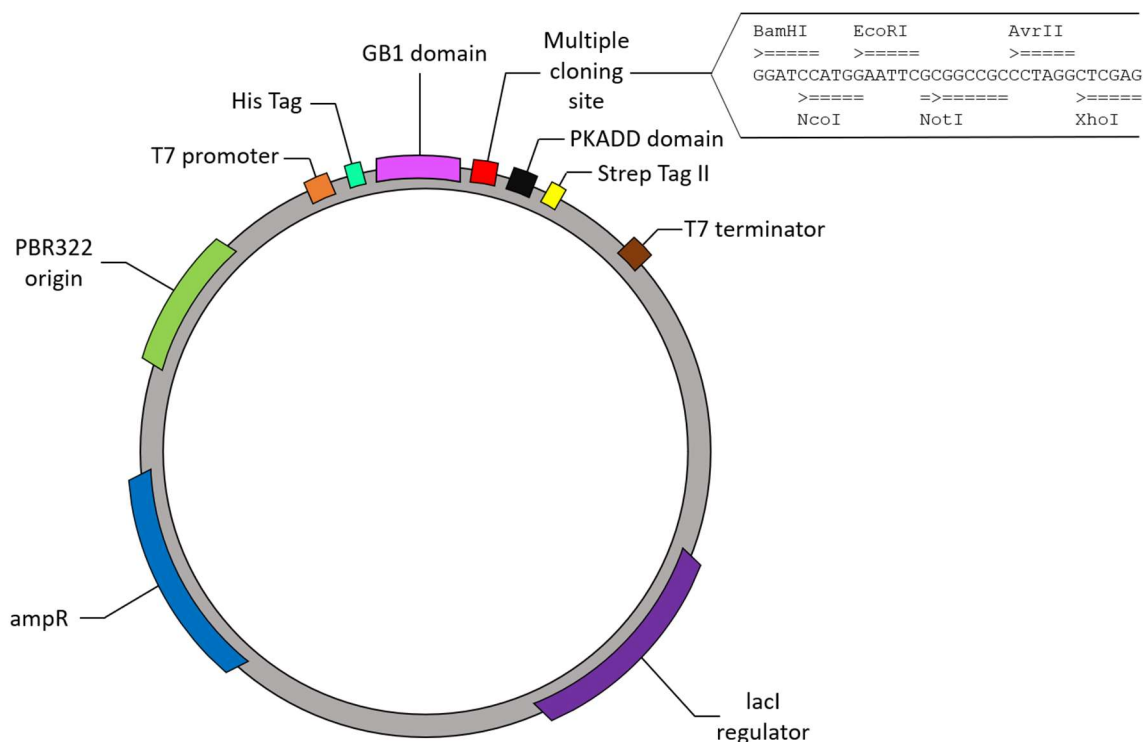


Figure 27: The pOP6BPA vector. Design of the pOP4BP vector employed in this study. Here the multiple cloning site has been expanded so the restriction enzyme cutting sites are displayed.

#### 4.4 Expression and Purification of PKADD Domain Fused Peptides

Like with the monomeric peptide fusions, expression tests were first conducted to determine the optimal temperature for expression. Expression tests on chemically competent BL21(DE3) cells containing DNA encoding for GB1-PKADD, GB1-H<sub>1</sub>L-PKADD, and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD revealed significant soluble expression of the peptide fusions occurred at 37 °C. Large scale expression was thus conducted at 37 °C and the peptide fusions were purified using the same dual purification system as described before (see 3.4 Expression and Purification of Peptide Fusions) with similar results being observed (Figure 28). Subsequent analysis via mass spectrometry shows molecular weights were observed for all three fusions at 130 Da lower than expected. This is most likely due to the loss of the initial methionine thus indicating that all three were successfully generated (Table 13).

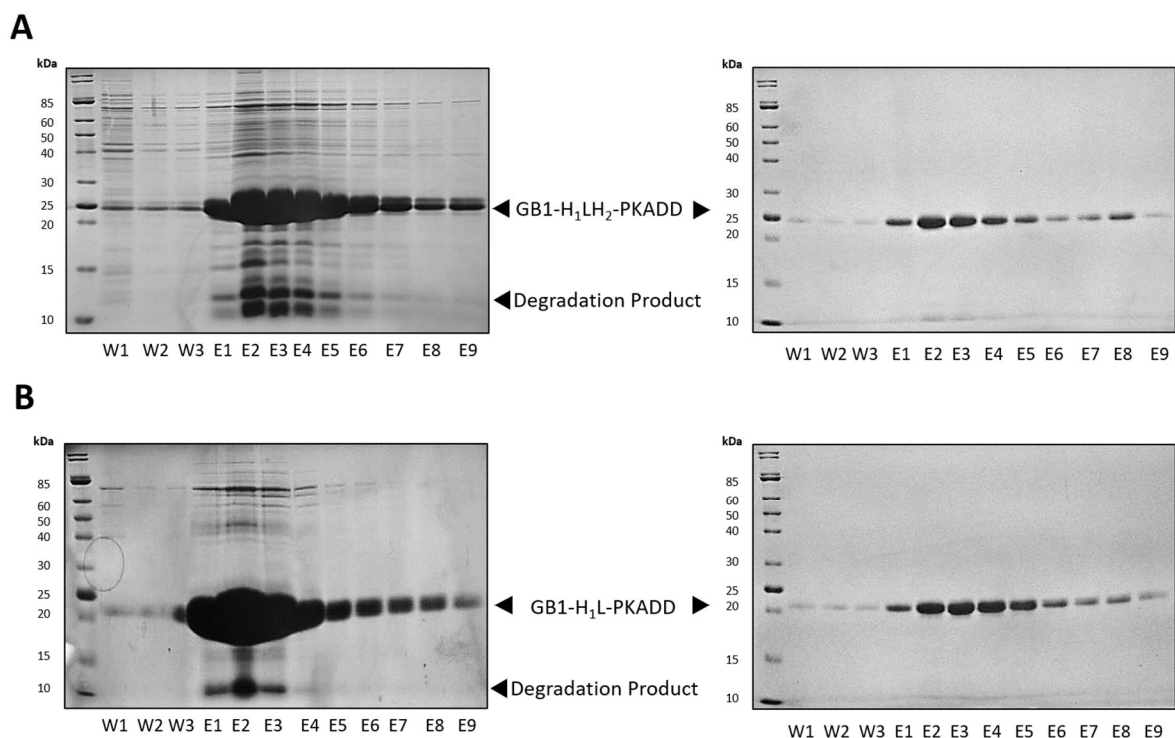


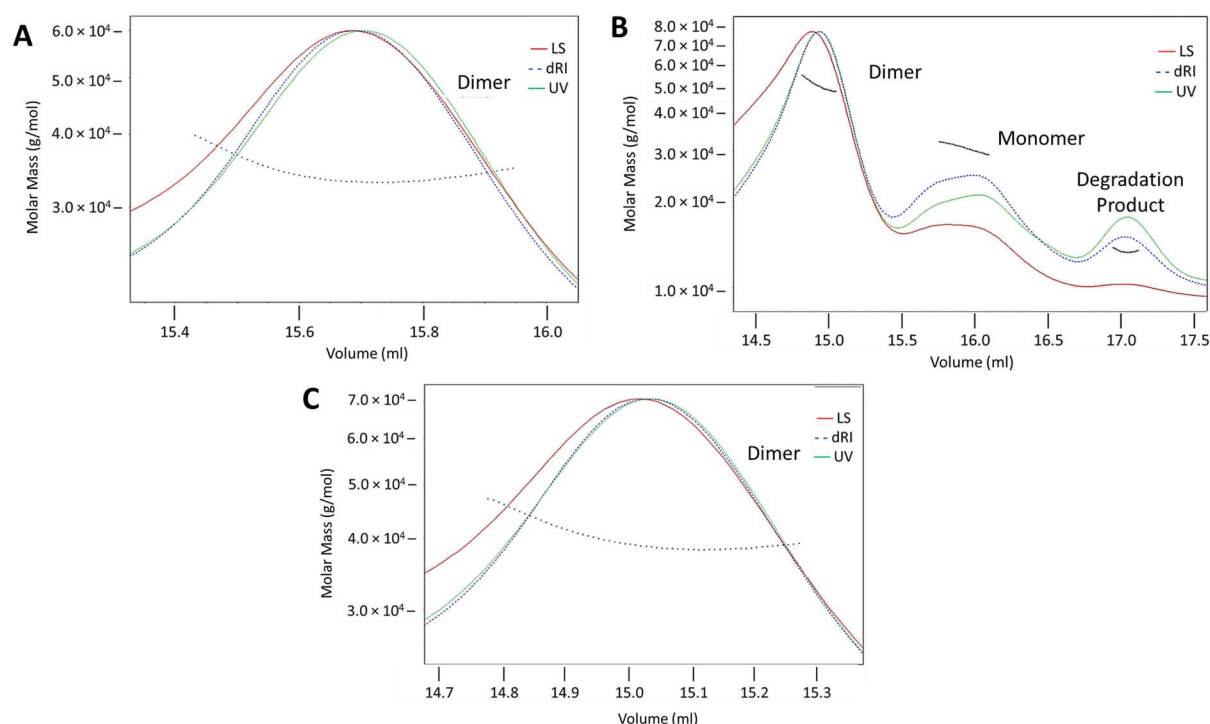
Figure 28: Purification of PKADD domain fused peptides. A) SDS-PAGE of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD after metal affinity purification with Ni-NTA beads (left) and StrepTactin Beads (right). B) SDS-PAGE of GB1-H<sub>1</sub>L-PKADD after metal affinity purification with Ni-NTA beads (left) and StrepTactin Beads (right). After, StrepTactin affinity purification we now see only one prominent band corresponding to the PKADD fused peptides indicating the purification has been successful.

## 4.5 PKADD Domain Fused Peptides are Dimers in Solution

In order to confirm whether the PKA fused peptides were dimeric in solution, SEC-MALS analysis was performed. Each sample was loaded onto a Superdex 75 increase 10/300 column (GE Healthcare) and the light scattering and refractive index change was measured. From the light scattering, the molecular mass of each species was calculated (Table 13, Figure 29). Analysis of the resulting chromatograms for GB1-PKADD and GB1-H<sub>1</sub>L-PKADD both show single peaks with masses of 34 kDa and 40 kDa, respectively, indicating both exist as dimers in solution. The chromatogram of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD shows the elution of three peaks with assigned molecular weight values of 51 kDa, 31 kDa, and 13 kDa. In this instance it is likely that these three peaks correspond to the dimer, monomer and a lower weight degradation product that does not dimerize. This suggests GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD may be less stable as a dimer than either GB1-PKADD and GB1-H<sub>1</sub>L-PKADD, though it is not clear as to why this may be the case. Overall this data shows that GB1-PKADD, and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and GB1-H<sub>1</sub>L-PKADD exist as dimers in solution.

**Table 13: Mass spectrometry and SEC-MALS analysis of PKADD fused peptides.**

Peptide Fusion	Calculated $M_w$ (Da)	$M_w$ (MALDI-MS)	$M_w$ (SEC MALS)	$M_n$ (SEC MALS)	Polydispersity
GB1-PKADD	16853.9	16724.3	$34500 \pm 25.5$	$34410 \pm 25.5$	$1.003 \pm 0.001$
GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD	20978.6	20848.6	$51030 \pm 354.1$	$50950 \pm 356.1$	$1.002 \pm 0.001$
			$30810 \pm 515.8$	$30780 \pm 515.3$	$1.001 \pm 0.024$
			$13390 \pm 203.5$	$13390 \pm 203.7$	$1.000 \pm 0.022$
GB1-H <sub>1</sub> L-PKADD	19355.6	19225.9	$40070 \pm 42.3$	$39940 \pm 42.3$	$1.003 \pm 0.002$



*Figure 29: SEC-MALS analysis of the PKADD domain fused peptides. The chromatograms of doubly tagged GB1 (A), GB1-H<sub>1</sub>L (B), and GB1-H<sub>1</sub>LH<sub>2</sub> (C) generated via SEC MALS. Here the refractive index is shown in blue, the light scattering in red, and the absorbance at 280 nm in green. The molecular weight analysis of the eluted peaks is shown as dotted black lines.*

#### 4.6 PKADD Domain Fused Peptides are More Effective Inhibitors than the Monomeric Peptide Fusions

In order to determine if the PKA fused peptides were better than their monomeric counterparts at binding to and inhibiting mature activin A signalling, ITC experiments and luciferase assays were performed as described previously on GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and GB1-H<sub>1</sub>L-PKADD (Table 14, Figure 30). Similar to the previous experiments, an ITC experiment and luciferase assay was performed on doubly tagged GB1-PKADD as a control. No significant binding to mature activin A or inhibitory properties

were observed suggesting the PKADD had no significant effect on the system apart from dimerizing the pro-domain derived peptides.

**Table 14: Dissociation constants and IC<sub>50</sub> values for PKADD domain fused GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L with mature activin A.**

Peptide Fusion	K <sub>d</sub> / nM	Stoichiometry	ΔH / kcal mol <sup>-1</sup>	ΔS / cal mol <sup>-1</sup> K <sup>-1</sup>	Log <sub>10</sub> (IC <sub>50</sub> / nM)	IC <sub>50</sub> / nM
GB1-H <sub>1</sub> LH <sub>2</sub>	642 ± 76.1*	2	-13.8 ± 0.2*	-18.2*	3.64 ± 0.14	4400
GB1-H <sub>1</sub> L	909 ± 112	2	-7.85 ± 0.14	1.32	3.58 ± 0.16	3800
GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD	58.6 ± 10.6*	1	-24.0 ± 0.5*	-49.2*	2.58 ± 0.06	380
GB1-H <sub>1</sub> L-PKADD	156 ± 33.9	1	-29.4 ± 0.4	-65.4	3.56 ± 0.09	3600

\* denotes that the K<sub>d</sub> and N values obtained are an average of three identical ITC experiments.

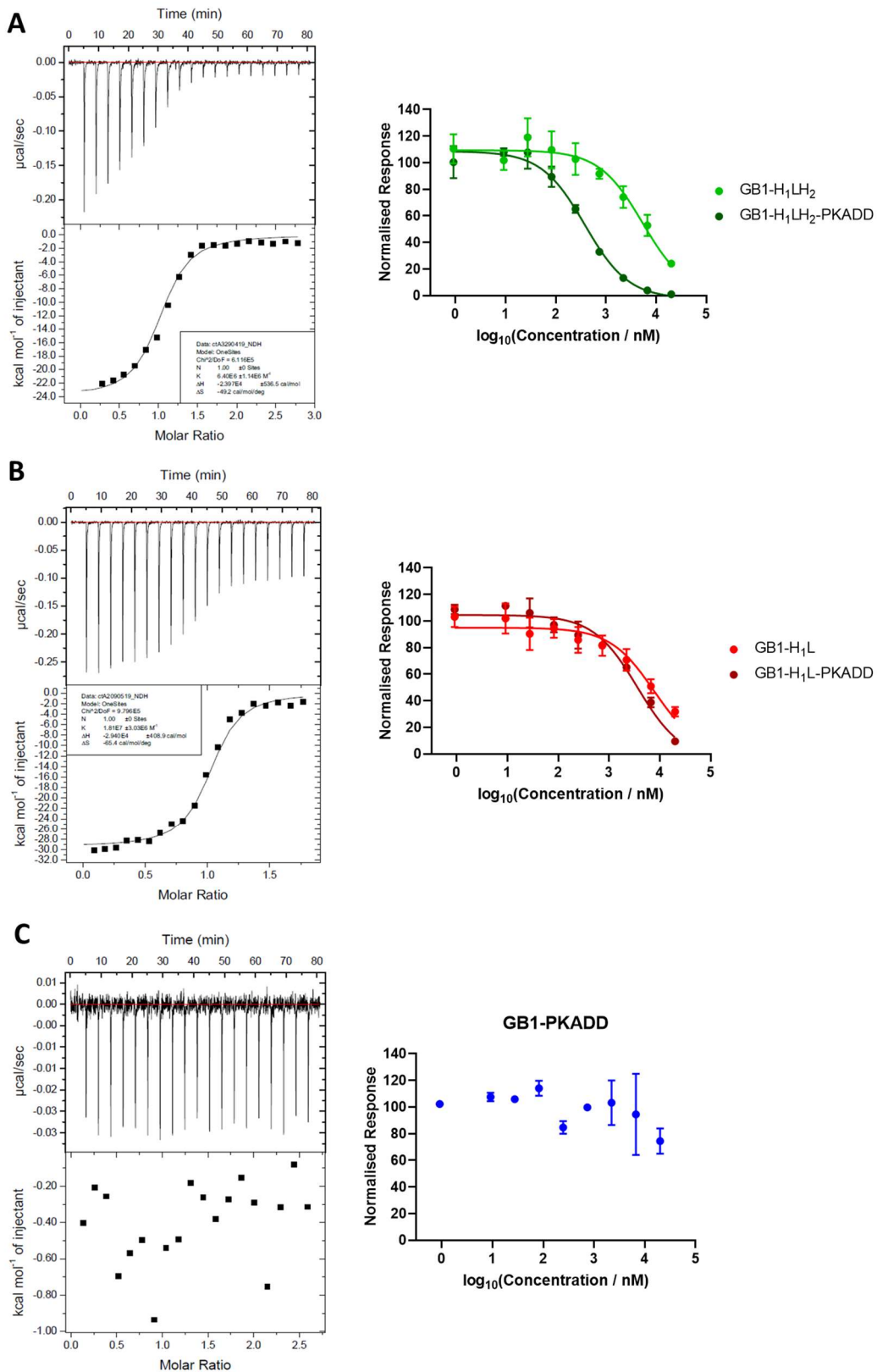
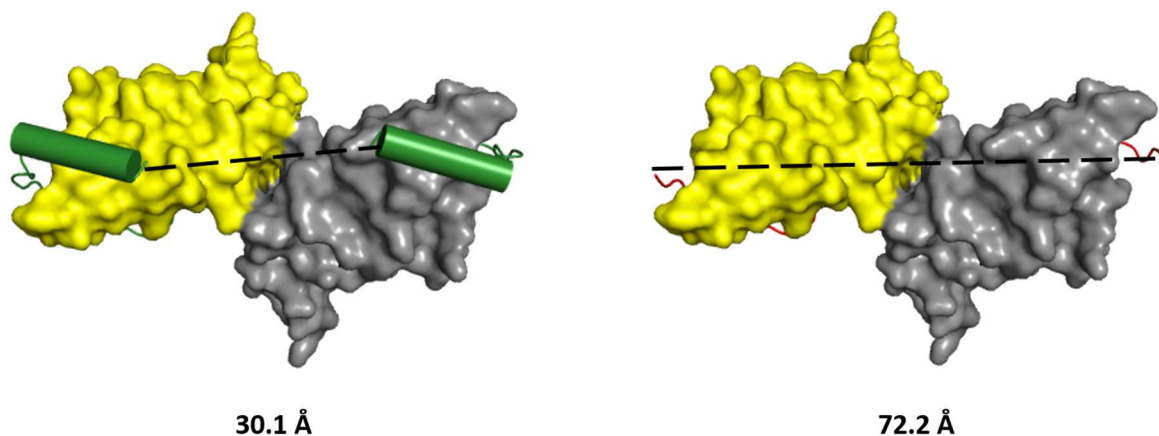


Figure 30: The binding and inhibitory profiles of the PKADD fused peptides. A) The binding profile generated via ITC for the interaction between GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and mature activin A (left) and the inhibition profile of GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD (right). B) The binding profile generated via ITC for the interaction between GB1-H<sub>1</sub>L-PKADD and mature activin A (left) and the inhibition profile of GB1-H<sub>1</sub>L and GB1-H<sub>1</sub>L-PKADD (right). C) The binding profile generated via ITC for the interaction between GB1-PKADD and mature activin A (left) and the

*inhibition profile of GB1 -PKADD (right). All inhibitory profiles are generated from three technical repeats. Error bars correspond to one standard deviation of uncertainty.*

Analysis of the ITC data for GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD reveals a stoichiometry of 1, meaning the fusion binds to mature activin A in a 1:1 ratio. This suggests GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD occupies both binding sites on mature activin A, thus interacting as a dimer. The approximate 10-fold increase in affinity observed for GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD compared to GB1-H<sub>1</sub>LH<sub>2</sub> (58.6 nM vs 642 nM) is likely due to this dimeric binding, with the increase in affinity resulting from the increase in avidity. Inhibition data shows a similar story with activin A signalling. A similar approximate 10-fold increase in inhibitor potency is observed in cell-based reporter assay upon dimerization, with IC<sub>50</sub> values of 380 nM and 4.4 μM for the dimeric and monomeric constructs, respectively. This shows that non-covalent dimerization, as predicted, is a valid method of optimising the inhibition of activin A signalling with pro-domain derived peptide fusions.

The observations made for GB1-H<sub>1</sub>L-PKADD interactions with mature activin A are more complex. The ITC data shows an increase in affinity compared to the monomeric version (156 nM and 909 nM) with a stoichiometry of 1 however there is no significant difference between the IC<sub>50</sub> values obtained for GB1-H<sub>1</sub>L and GB1-H<sub>1</sub>L-PKADD (3.8 μM vs 3.6 μM). The structure of the pro-mature complex indicates a large distance between the two C-terminal residues of the loop, 72.2 Å, with them being oriented away from each other on different faces of the mature dimer (Figure 31). This contrasts with the distance between the C-terminal residues of H<sub>1</sub>LH<sub>2</sub>, 30.1 Å, and their orientation towards one another. As such, the large distance and suboptimal orientation of GB1-H<sub>1</sub>L-PKADD may weaken the interaction between the two PKADD domains resulting in a partial monomeric and partial dimeric character. This would explain the similarity between the IC<sub>50</sub> values of GB1-H<sub>1</sub>L and GB1-H<sub>1</sub>L-PKADD as well as the K<sub>d</sub> values, however evidence for this is currently weak and further investigation is required.





*Figure 31: Interhelical distance of H<sub>1</sub>LH<sub>2</sub>. The inter-residue distance between C $\alpha$  of the C-terminal residues of H<sub>1</sub>LH<sub>2</sub> and H<sub>1</sub>L based on the structure of pro-mature activin A. Here mature activin A is shown in yellow and grey, H<sub>1</sub>LH<sub>2</sub> in green, and H<sub>1</sub>L in red.*

Overall this section has demonstrated for the H<sub>1</sub>LH<sub>2</sub> containing fusion, dimerization results in both an increase in affinity and inhibitor potency for activin A, confirming the validity of a dimeric approach.

## 4.7 Continuous Fused Helix Peptides Behave in a Similar Manner to GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD

GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD inhibits activin A signalling with the highest affinity observed so far (380 nM) representing approximately a 10-fold increase in potency compared to the other three peptides fusion. However, it is desirable to optimise this peptide further. The target selected for this optimisation was the linker between the  $\alpha$ 2-helix and the PKADD. The length of linker (eight residues) means that the  $\alpha$ 2-helix- $\alpha$ 2-helix distance between the dimeric subunits may be large when the PKADD domain is taken into account. Therefore, this may have a negative effect on affinity and potency as the avidity affect may be reduced due to distance and conformational changes being required for mature activin A binding. Thus, in order to try to spatially optimise GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and increase its structural rigidity, I decided to eliminate the linker and fuse the C-terminus of the  $\alpha$ 2-helix directly to the N-terminal helix of PKADD. In order to achieve this, the C-terminal residue of the  $\alpha$ 2-helix and the N-terminal residue of the PKADD domain first needed to be identified. Analysis of the respective pro-mature activin A complex and PKADD domain structures revealed these residues to be L90 of the  $\alpha$ 2-helix (L90<sub>H2</sub>) and L10 of the PKADD domain (L10<sub>PKADD</sub>). Next, additional alternating serine and alanine spacer residues were added to allow for the testing of whether the relative geometry of the  $\alpha$ 2-helix and PKADD to each other had an impact on either affinity or efficacy. From this, three peptide fusions were generated – GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub>, GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>, and GB1-H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub>. Of these, GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub> contained one serine residue as a spacer between L90<sub>H2</sub> and L10<sub>PKADD</sub>, GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub> an alanine followed by a serine, and GB1-H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub> a serine followed by an alanine and a serine. JPred4 and Spider2 secondary structure prediction software was used to predict whether helicity would be retained throughout the fused helix which it did in all cases.<sup>245,246</sup> Each peptide was then expressed and purified and ITC experiments and luciferase assays were performed on each (Table 15).

**Table 15: Dissociation constants and IC<sub>50</sub> values obtained for the Fused Helix peptides with mature activin A.**

Peptide Fusion	K <sub>d</sub> / nM	Stoichiometry	ΔH / kcal mol <sup>-1</sup>	ΔS / cal mol <sup>-1</sup> K <sup>-1</sup>	Log <sub>10</sub> (IC <sub>50</sub> / nM)	IC <sub>50</sub> / nM
GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD	58.6 ± 10.6	1	-24.0 ± 0.5	-49.2	2.580 ± 0.061	380.4
GB1-H <sub>1</sub> LH <sub>2</sub> -S-PKADD <sub>F</sub>	33.3 ± 11.7	1	-27.0 ± 0.5	-56.3	2.405 ± 0.050	253.9
GB1-H <sub>1</sub> LH <sub>2</sub> -AS-PKADD <sub>F</sub>	69.9 ± 17.3	1	-25.8 ± 0.4	-53.8	2.468 ± 0.079	293.9
GB1-H <sub>1</sub> LH <sub>2</sub> -SAS-PKADD <sub>F</sub>	45.7 ± 7.53	1	-32.8 ± 0.3	-76.2	3.448 ± 0.109	2803

The ITC data obtained for these fusions shows that all three exhibit similar binding properties to that of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD binding both in a dimeric manner (stoichiometry = 1 in all instances) with similar affinities (K<sub>d</sub> = 33.3 nM, 69.9 nM, and 45.7 nM vs 58.6 nM) (Figure 32). Analysis of the luciferase assays of GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub> and GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub> show a similar story, with both inhibiting signalling in a similar manner to GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD (IC<sub>50</sub> = 254 nM and 294 nM vs 380 nM). These similarities suggest that fusing the α2-helix directly to the N-terminal helix of the PKADD domain has a positive effect on mature domain binding and inhibitor potency, though the magnitude of this effect is small.

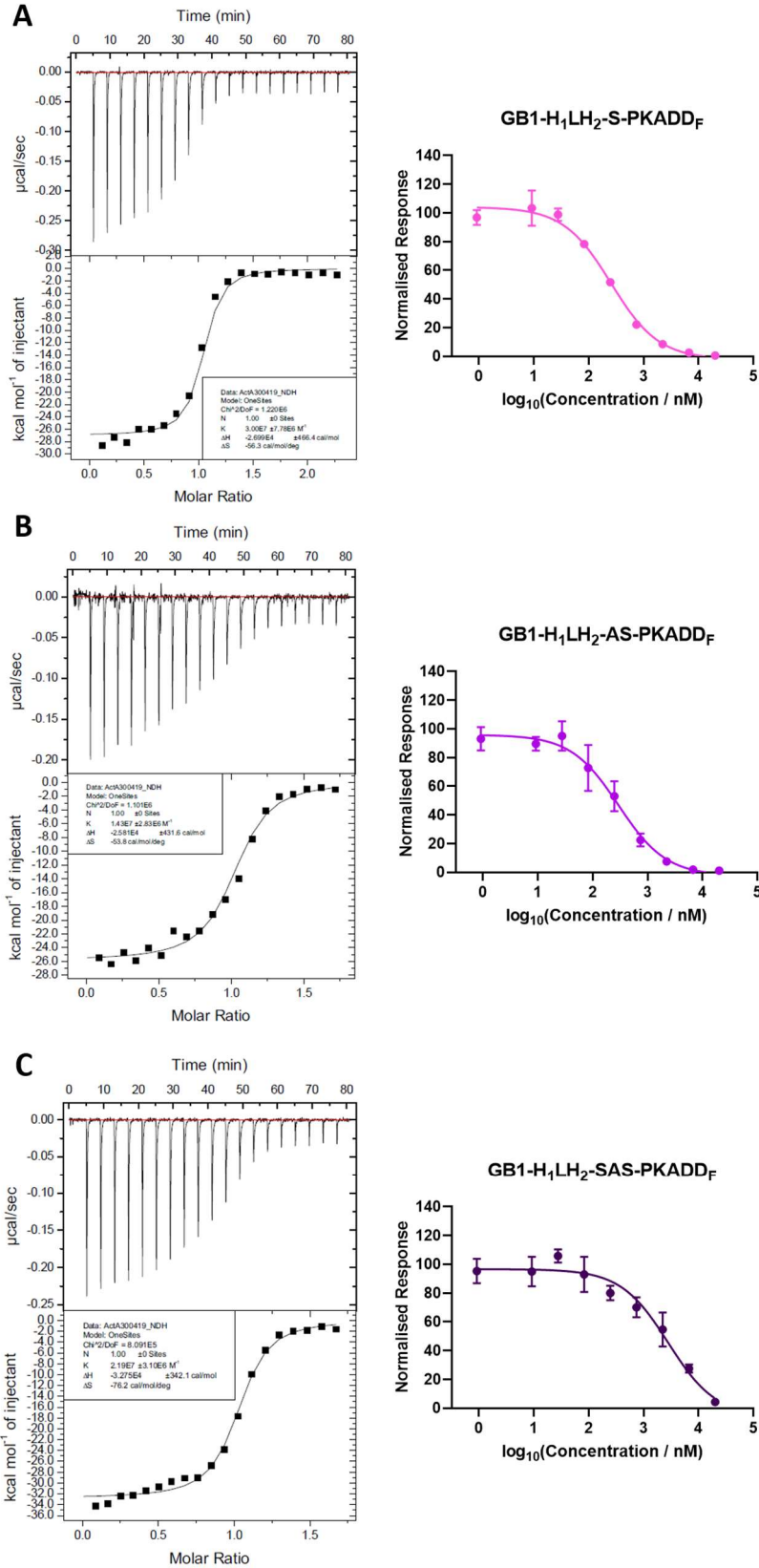


Figure 32: Binding and inhibitory profiles of the continuous fused helix peptides. A) The binding and inhibition profiles generated for GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub> with mature activin A via ITC and luciferase assay. B) The binding and inhibition profiles generated for GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub> with mature activin A via ITC and luciferase assay. C) The binding and inhibition profiles generated for GB1-H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub> with mature activin A via ITC and

*luciferase assay. All inhibitory profiles are generated from three technical repeats. Error bars correspond to one standard deviation of uncertainty.*

Interestingly, the IC<sub>50</sub> value obtained for GB1-H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub> is significantly lower than that of GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub> and GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>. This may be due to the position of the X-helical bundle of PKA relative to the α2-helix, causing it to clash sterically with the mature domain of activin A, leading to a partial monomeric character similar to that suggested for GB1-H<sub>1</sub>L-PKADD however again, evidence is currently weak and further investigation would be required. Overall, this data shows that eliminating the linker between the PKADD domain and the α2-helix and fusing the two helices together results in a small increase in potency for GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub> and GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>.

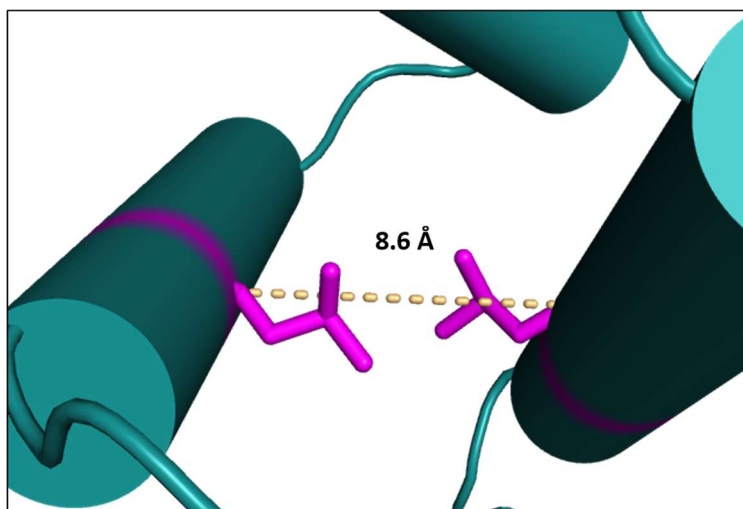
## 4.8 Design of Covalently Linked Dimeric Peptide Fusions

The effect of dimerization on inhibitor potency so far has been positive, with non-covalent dimerization resulting in a 10-20-fold decrease in IC<sub>50</sub> values depending on the peptide in question. However, the non-covalent nature of this dimerization may be a limitation. The K<sub>d</sub> of dimerization for PKADD is unknown, however it is possible at the lower inhibitor concentrations used in the luciferase assays, the PKADD fused peptides monomerise thus reducing the efficacy of the inhibitor. Therefore, to overcome this I decided to adopt a covalent dimerization approach. For this, two constructs were designed – GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>.

To generate GB1-H<sub>1</sub>LH<sub>2</sub>-SS, three residues containing one cysteine were added to GB1-H<sub>1</sub>LH<sub>2</sub> between the α2-helix and the Strep-tag II such that a single disulfide could be formed. Though the combined average chain length of these six additional residues is likely less than that of the interhelical distance of 30.1 Å (assuming each residue has an end to end distance of between 3.6-4.0 Å), however activin A has been observed to be flexible about its dimerization cysteine so this should not prevent dimeric binding. An equivalent GB1-H<sub>1</sub>L-SS was not generated due to the observation that GB1-H<sub>1</sub>L-PKADD did not show a significant increase in inhibitor potency compared to GB1-H<sub>1</sub>L.

GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> was generated through introduction the mutation L14C into the PKADD domain of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD. Residue L14 is an interfacial residue in the N-terminal helix of the PKADD domain that interacts hydrophobically with the corresponding L14 residue on the N-terminal helix of the other monomer. Therefore, only one mutation is required to generate the desired disulfide product, thus eliminating the problem of potential by-products associated with introducing multiple cysteine mutations. Measuring the distance between the Cαs of the two residues in the PKADD structure reveals they are 8.6 Å apart (Figure 33). This distance is longer than most disulfide bonds, which usually

fall in the range 4.3-6.4 Å. As such, some amount of conformational change may be required to facilitate bond formation, however it was decided to proceed non the less.



*Figure 33: Selected site for disulfide formation in the PKADD domain. The interatomic distance between Ca of residues L14 on both PKADD domain chains. Here L14 is highlighted in magenta.*

Unlike the other peptide fusions, GB1 H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> were purified via Ni affinity purification using an increasing concentration of imidazole. As dimerization forms a covalent link between two monomers, the resulting dimer will contain two His-tags rather than one. Due to the increase in avidity, the dimer should therefore interact more strongly with the resin than the monomer. Therefore, the monomeric and dimeric forms should elute at different concentrations of imidazole. Analysis of the chromatogram of GB1 H<sub>1</sub>LH<sub>2</sub>-SS shows the elution of two clear broad peaks at much higher absorbance than that of the increasing the imidazole between approximately 70-100 mL (Fractions A7-C2) (Figure 34). SDS-PAGE of these fractions shows that both the dimer and monomer are successfully expressed in large amounts and elute at different enough concentrations of imidazole that the products can be separated easily via this purification. As such, the fractions containing the dimer were then pooled and purified further with StrepTactin beads. Analysis of this purification shows the presence of a prominent band at the expected molecular weight for GB1-H<sub>1</sub>LH<sub>2</sub>-SS, however there are fainter bands at higher molecular weights. Analysis via mass spectrometry (data not shown) revealed that only monomeric and dimeric GB1-H<sub>1</sub>LH<sub>2</sub>-SS were detected indicating these bands correspond to oligomerisation, though how this occurs is unclear. These higher molecular weight species could not be adequately separated from dimeric GB1-H<sub>1</sub>LH<sub>2</sub>-SS thus the sample was taken forward for inhibition assays regardless.

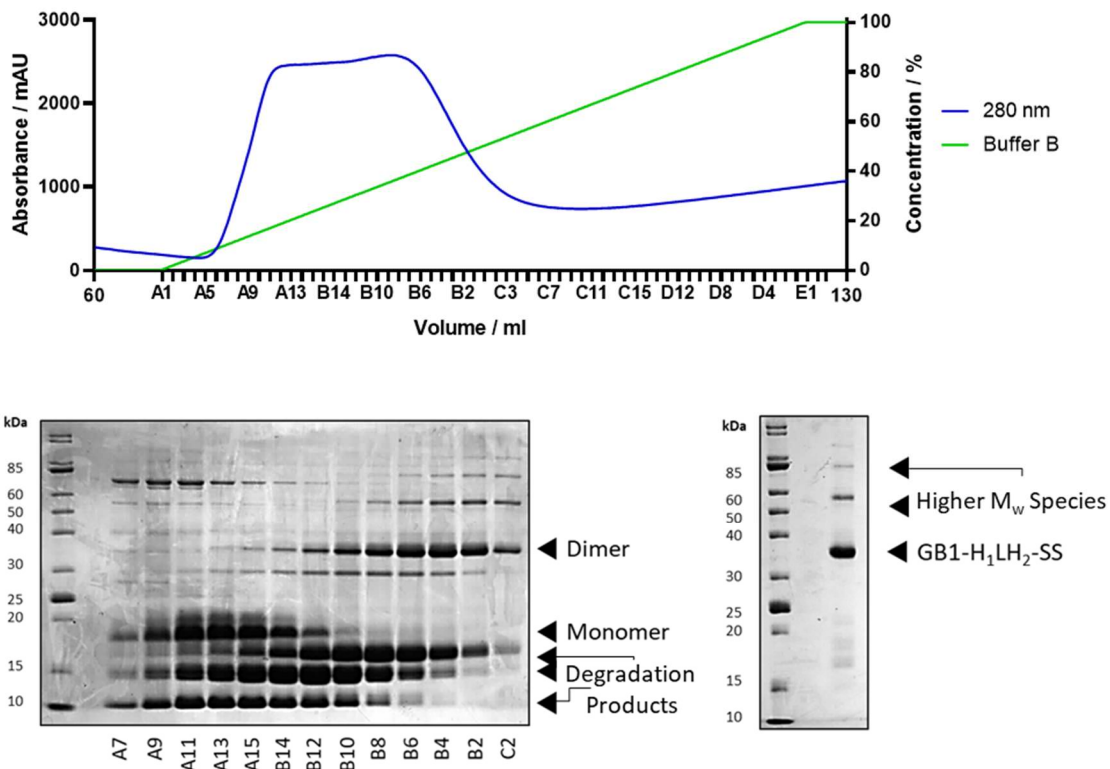


Figure 34: The purification of GB1-H<sub>1</sub>LH<sub>2</sub>-SS. A) Chromatogram of the metal affinity purification of GB1-H<sub>1</sub>LH<sub>2</sub>-SS with increasing concentration of imidazole. B) SDS-PAGE of the metal affinity purification of GB1-H<sub>1</sub>LH<sub>2</sub>-SS. C) StrepTactin affinity purification of the pooled fractions of GB1-H<sub>1</sub>LH<sub>2</sub>-SS containing the covalent dimer.

In the case of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>, analysis of the StrepTactin affinity purification reveals the presence of both the covalent and non-covalent forms after expression with the non-covalent form being more prevalent. Lu *et al.* has shown the ability of Cu (ii) ions to catalyse the formation of disulfide bonds in denatured human granulocyte colony-stimulating factor resulting in protein folding.<sup>247</sup> Therefore, in order to maximise covalent dimerization rapidly, the sample was incubated with 1 mM CuSO<sub>4</sub> at 4 °C overnight. Analysis of the subsequent SDS PAGE shows that after 16 hours most of the sample had covalently dimerised with only traces of the non-covalent dimer being detected (Figure 35A).

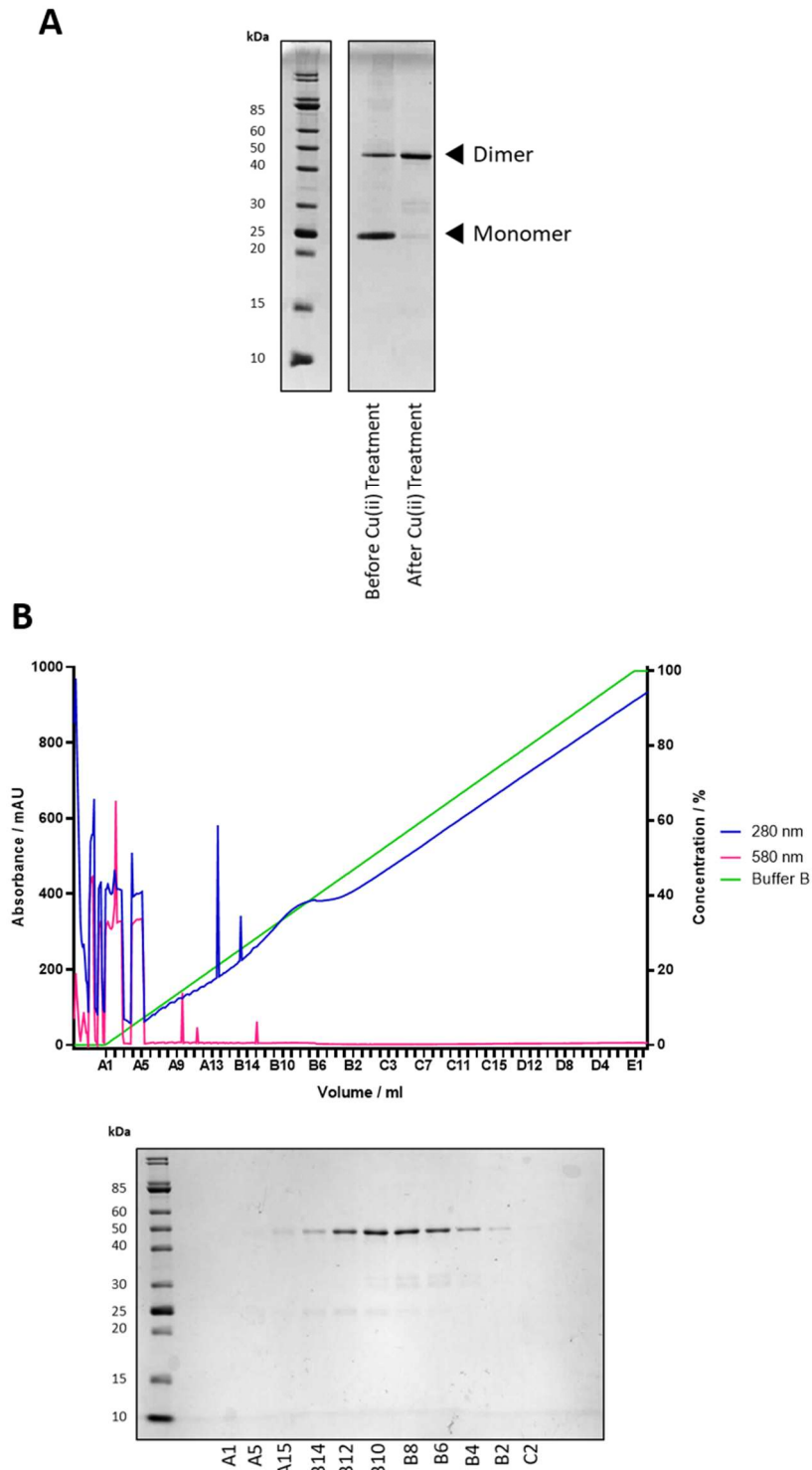


Figure 35: Cu (ii) catalysed linkage of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>. A) SDS-PAGE of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> before and after treatment with CuSO<sub>4</sub>. Here we see that the addition of CuSO<sub>4</sub> results in almost complete formation of the covalent dimer. B) Chromatogram of the metal affinity purification of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> with increasing concentration of imidazole (top). SDS-PAGE of the metal affinity purification of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>. Here we see the presence of a single band corresponding to the covalent dimer (bottom).

The sample was then purified in a similar manner to GB1-H<sub>1</sub>LH<sub>2</sub>-SS. The purification was performed in 6 M urea to monomerise any non-covalently linked GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> which should exhibit a

different elution profile to covalently linked GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> due to the difference in the number of His-tags. In addition, the absorbance at 580 nm was measured due to the absorbance of Cu (ii) ions at this wavelength (Figure 35B). Analysis of the chromatogram shows zero absorbance at 580 nm in the elution phase suggesting all the Cu (ii) ions have been removed (here the sharp spikes likely correspond to air in the column). A perturbation of the increasing absorbance at 280 nm associated with the increasing imidazole concentration occurs at approximately 90 – 110 mL (Fractions B14-B2). SDS-PAGE of these fractions revealed the presence of a single band at the molecular weight of the dimer. Thus, these fractions were pooled, concentrated and taken forward for luciferase assays.

#### 4.9 Introduction of a Disulfide Bridge Results in an Increase in Inhibitor Efficacy

To determine the effect covalent dimerization has on inhibitor potency, luciferase assays were performed on both GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> (Table 16, Figure 36). Analysis of the inhibition curves generated for each peptide fusion shows a marked increase in inhibitor potency compared with GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD. Here the IC<sub>50</sub> values for GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> are 45.2 nM and 71.5 nM. This is around 5-10-fold more potent than the non-covalent dimer GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD and 50-100-fold more potent than the monomeric GB1-H<sub>1</sub>LH<sub>2</sub>. This therefore suggests that covalent dimerization is the best method of improving inhibitor potency through dimerization. An additional ITC titration was also performed with GB1-H<sub>1</sub>LH<sub>2</sub>-SS as described previously. The *K<sub>d</sub>* value obtained from this experiment is around 2-fold stronger than that of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD with a stoichiometry of 1 confirming GB1-H<sub>1</sub>LH<sub>2</sub>-SS interacts as a dimer with mature activin A.

**Table 16: Dissociation constants and IC<sub>50</sub> values obtained for the covalently dimerized peptide fusions with mature activin A.**

Peptide Fusion	<i>K<sub>d</sub></i> / nM	Stoichiometry	$\Delta H$ / kcal mol <sup>-1</sup>	$\Delta S$ / cal mol <sup>-1</sup> K <sup>-1</sup>	Log <sub>10</sub> (IC <sub>50</sub> / nM)	IC <sub>50</sub> / nM
GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD	58.6 ± 10.6	1	-24.0 ± 0.5	-49.2	2.58 ± 0.06	380
GB1-H <sub>1</sub> LH <sub>2</sub> -SS	30.5 ± 12.0	1	-39.0 ± 0.8*	-96.3*	1.66 ± 0.04	45.2
GB1-H <sub>1</sub> LH <sub>2</sub> -PKADD <sub>SS</sub>	–	–	–	–	1.85 ± 0.06	71.5



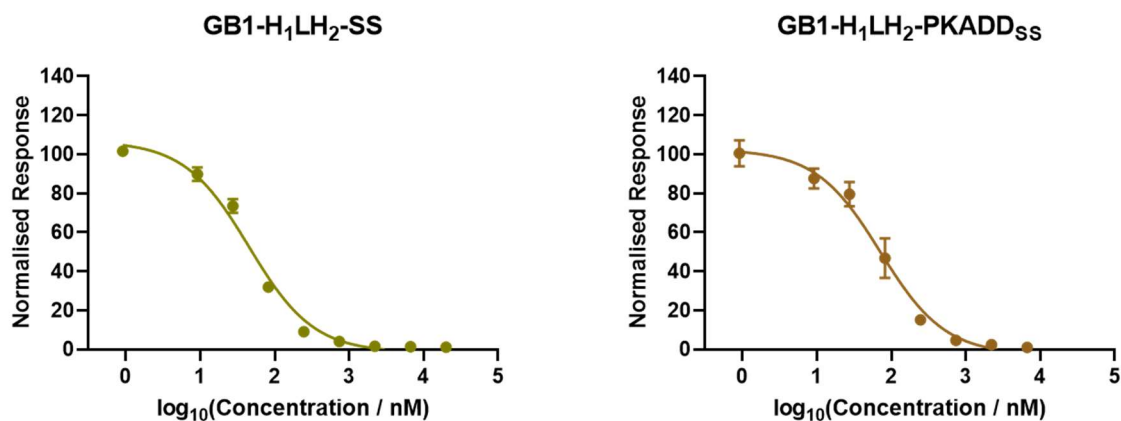
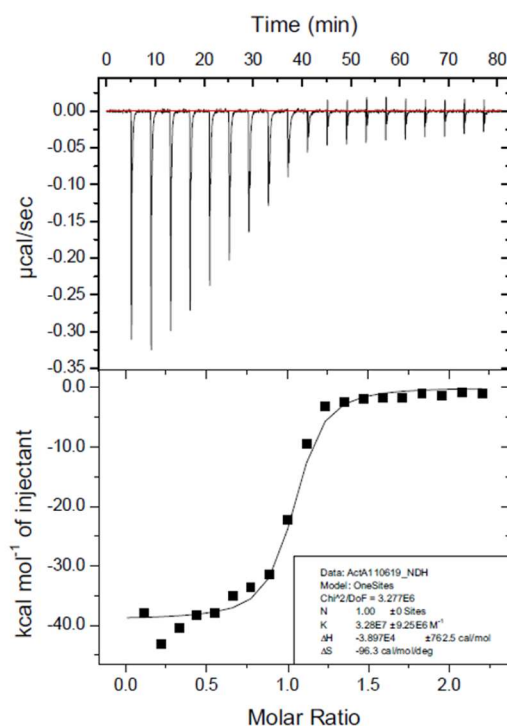
**A****B**

Figure 36: Inhibitory profiles of covalently dimerized peptide fusions. A) The inhibition profiles of GB1-H<sub>1</sub>LH<sub>2</sub>-SS (left) and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> (right) generated via luciferase assays. Inhibitory profiles are generated from three technical repeats. Error bars correspond to one standard deviation of uncertainty. B) The binding profile of GB1-H<sub>1</sub>LH<sub>2</sub>-SS to mature activin A generated via ITC.

#### 4.10 Screening of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> against Other TGF-β Superfamily Members

The covalent dimerization of peptide fusions through a disulfide bond generated two highly potent inhibitors of activin A signalling, GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>. However, due to the high

structural and sequence similarity of TGF- $\beta$  ligands and receptor promiscuity within the family, it was important to test one of these fusions against other ligands to determine specificity.

The first step of the determining specificity was to select which members of the TGF- $\beta$  family would be screened. To provide a robust screen, three ligands were selected of varying sequence similarity to activin A and that signal through two different Smad pathways. The first two were activin B and myostatin. Both these ligands are commonly used for specificity screens on activin A due to their similarities with activin A.<sup>188</sup> Like activin A, activin B and myostatin signal through the ActRIIA / B and ALK4 receptors, triggering the Smad 2/3 pathway (though they can also signal through ALK7 and ALK5 respectively).<sup>248,249</sup> Furthermore, the sequences of the mature domains of activin B and myostatin have a high similarity with that of activin A. Activin B shares 82.8 % similarity (63.8 % identity) with the key receptor interacting residues being conserved and myostatin 49.6 similarity. Functional studies have also implicated activin B and myostatin as acting in concert with activin A in tissues such as the sex organs and skeletal muscle thus making distinguishing their specificity highly important for therapeutics.<sup>7,250</sup>

The final ligand selected for screening was BMP2. BMP2 is a more distant relative of activin A, appearing to have diverged from a common ancestor much earlier than either activin B or myostatin. As such, BMP2 primarily signals through the Smad 1/5/8 pathway and appears not to interact with ALK4, interacting with other type I receptors instead (however BMP2 does interact with ActRIIA and ActRIIB as well as the type II receptor BMPRII).<sup>251</sup> Therefore, this ligand was selected to increase the broadness of the screen.

To confirm the appropriateness of this selection, the similarity of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix region between the three selected ligands and activin A was compared. Alignment was performed using ClustalX (Figure 37).<sup>252</sup> The minimal binding peptide for activin A obtained in this thesis (see 3.7 Determination of the Minimal Binding Peptide of the Activin A Pro-domain) tells us the key interacting residues of H<sub>1</sub>LH<sub>2</sub> are in the region E53-P72. This region includes a crucial salt bridge formed by residue R71, as well as a number of hydrophobic residues in the  $\alpha$ 1-helix (V55, V58, I62, L63, L66 and L68), that appear to form the majority of contacts between the peptide and the mature domain. Analysis of the alignment reveals that these hydrophobic residues are highly conserved across activin B, myostatin, and BMP2, however subtle differences exist and substitution of similar aliphatic amino acids is common. However, in the case of BMP2 the differences are larger with substitution to phenylalanine occurring at two of these positions. This suggests larger, more hydrophobic binding pockets may be present on the mature BMP2 surface, thus making inhibition by H<sub>1</sub>LH<sub>2</sub> less likely. R71, which forms the crucial salt bridge in activin A, is conserved in activin B however not in myostatin,

meaning it is likely H<sub>1</sub>LH<sub>2</sub> will have a higher specificity for activin B over myostatin.<sup>†</sup> Furthermore, comparison of the aligned sequences reveals activin B to have a much higher similarity to activin A than myostatin or BMP2 (sequence similarity activin B = 82.1 %, myostatin = 47.7 %, BMP2 = 46.3 %). As such it is likely H<sub>1</sub>LH<sub>2</sub> will have discernible differences in specificity for these three ligands.

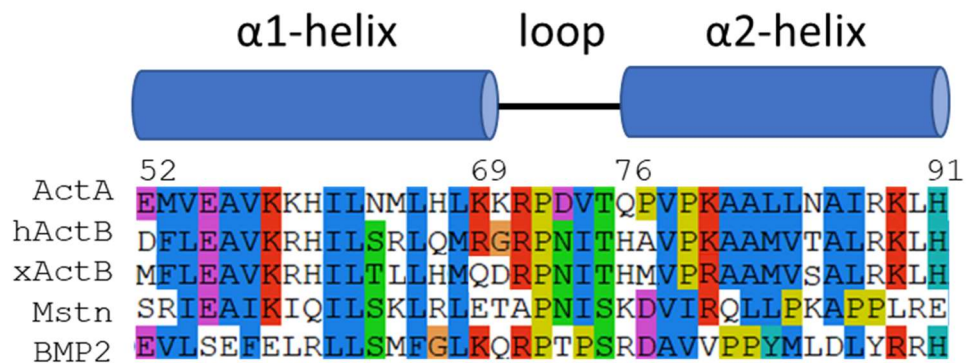


Figure 37: Alignment of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix region of activin A, human activin B, xenopus activin B, myostatin, and BMP2. Here highly conserved hydrophobic residues are shown in blue, aromatic residues in turquoise, acidic residues in purple, basic residues in red and polar residues in green. Glycine and proline residues are also shown in orange and yellow respectively. Above the residue numbers are shown for activin A along with the helix boundaries.

Analysis of the  $\alpha$ 2-helix of activin A and B reveals sequence conservation is high. However, compared to the  $\alpha$ 1-helix, there is significantly lower conservation between family members and a greater degree of sequence variation is seen. Though this region of H<sub>1</sub>LH<sub>2</sub> was not crucial for mature domain binding, it still forms contacts with the mature domain in the pro-mature ActA structure and as such may also contribute towards specificity of the peptide.

To test specificity, inhibitory luciferase assays were performed on activin B, myostatin and BMP2 using GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub> as the inhibitor (Table 17, Figure 38). As both activin B and myostatin signal through the Smad 2/3 pathway, the assays could be performed as described previously. However, this is not the case for BMP2, which usually signals through the Smad 1/5/8 pathway. As such, HEK293T cells containing a BMP responsive element (BRE), which expresses firefly luciferase upon signalling through the Smad 1/5/8 pathway, were used.

The three assays were all performed at the EC<sub>80</sub> of each respective ligand (activin B = 80 pM, myostatin = 250 pM, BMP2 = 2.5 nM). The activin B used here was not the human sequence, rather Xenopus. However, the human and Xenopus sequence is highly conserved with 96.5% identity and 98.3% similarity and thus likely to behave in a similar manner. Both the myostatin and BMP2 used were

<sup>†</sup> R71 is also conserved in BMP2, however due to the phylogenetic distance between BMP2 and activin A, it is difficult to predict whether it forms a salt bridge.

human. An additional positive control was used in the BMP2 assay – the extracellular BMP inhibitor gremlin.

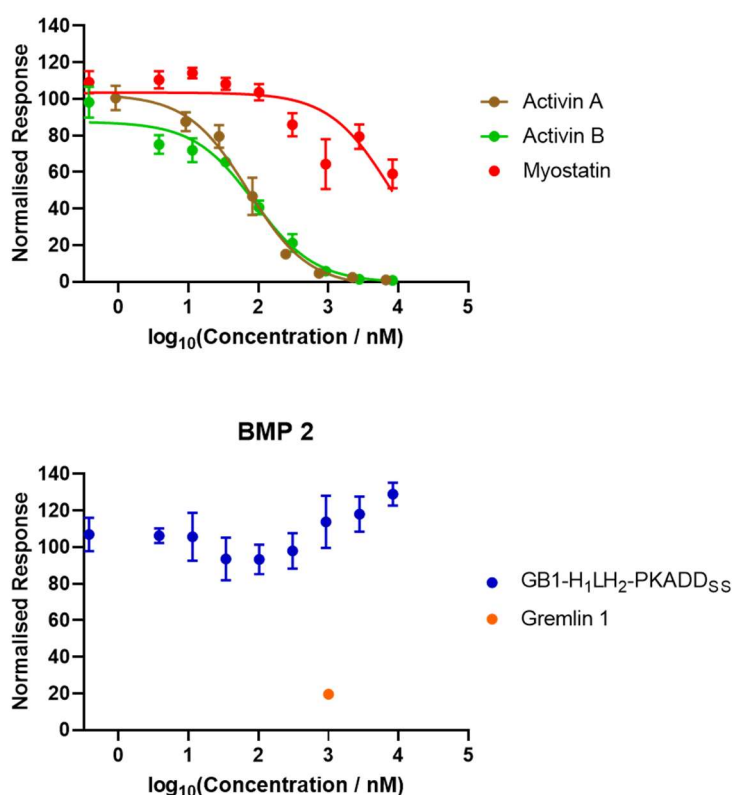


Figure 38: Specificity screen of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub>. The inhibitory profile of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub> against activin A, activin B, myostatin and BMP2 as determined via luciferase assay. Inhibitory profiles are generated from three technical repeats. Error bars correspond to one standard deviation of uncertainty.

Table 17: Specificity screen of the TGF- $\beta$  family.

Ligand	Log <sub>10</sub> (IC <sub>50</sub> / nM)	IC <sub>50</sub> / nM
Activin A	1.85 ± 0.06	71.5
Activin B	1.94 ± 0.08	88.0
Myostatin	3.88 ± 0.11	7600
BMP2	–	–

Analysis of the inhibition curves shows that GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub> inhibits activin B with an IC<sub>50</sub> value of 88.0 nM. This value is highly similar to that obtained for activin A with the same inhibitor (71.5 nM), suggesting GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub> inhibits both proteins in a similar manner. This is to be expected due to the sequence similarity of the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of activins A and B (82.1 %) and the similarity of the mature domains (82.8 %). The inhibitory trend for myostatin is weaker, with a full inhibition not being obtained within the concentration range of the experiment. Extrapolation of an

IC<sub>50</sub> value from this data yields a value of approximately 7.6 μM, meaning GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> is around 100-fold more selective for activin A and B than myostatin. The curve obtained for BMP2 shows maximum signal throughout with the signal being significantly higher at all concentration points than that obtained for gremlin. As such GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> does not inhibit BMP2 signalling. Therefore, GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>, and thus the H<sub>1</sub>LH<sub>2</sub> motif of activin A's pro-domain, is a specific inhibitor of activin A and B signalling.

## 4.11 Conclusions

In this section I described the optimisation of the inhibitory peptides described in the section “Peptide Approach” through dimerization. I described the generation and characterisation of non-covalent PKADD domain fusions based on GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L and show that both are dimeric in solution. I then showed that their affinity for mature activin A is higher than that of their monomeric counterparts (58.6 nM and 156 nM vs 642 nM and 909 nM), likely due to the increase in avidity. I demonstrated that a similar effect is observed for GB1-H<sub>1</sub>LH<sub>2</sub> on the inhibition of activin A signalling (380 nM vs 4.4 μM) but not for GB1-H<sub>1</sub>L (3.6 μM vs 3.8 μM). I then attempted to optimise GB1-H<sub>1</sub>LH<sub>2</sub> through the design of a fused helix system, however these fusions showed no great increase in either affinity for mature activin A or inhibition of mature activin signalling. I designed two covalent constructs – one directly fusing two GB1-H<sub>1</sub>LH<sub>2</sub> at the C-terminus through a disulfide bond, and one introducing a disulfide into the PKADD domain of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD. These molecules had significantly higher inhibitory activity compared to both the monomeric and non-covalent dimeric peptides (GB1-H<sub>1</sub>LH<sub>2</sub>-SS = 45.2 nM, GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> = 71.5 nM). Finally, I established the specificity of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> against 3 other members of the TGF-β family – activin B, myostatin, and BMP2. From this I determined GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> inhibited activin B in a similar manner to activin A, whereas it had approximately 100-fold selectivity over myostatin, and even greater over BMP2.



---

# Discussion

---

## 5.1 Discussion

Activin A signalling has been implicated as a key mechanism in the pathology in several diseases. As well as driving the proliferation and epithelial-mesenchymal transitions in a number of breast and prostate cancers, activin A signalling has been identified as the main cause of fibrodysplasia ossificans progressiva.<sup>159,167,253</sup> Furthermore, it plays a crucial role in muscular homeostasis as a major negative regulator of muscle growth.<sup>7</sup> As such, activin A is an important therapeutic target. However, specific inhibition of activin A signalling has often been difficult to achieve, due to structural and sequence similarity between activin A and other TGF- $\beta$  family members and ligand promiscuity in TGF- $\beta$  family signalling. Indeed, commonly used methods either target activin's receptors or utilise non-specific extracellular inhibitors resulting in convoluted phenotypes and side effects associated with the inhibition of multiple TGF- $\beta$  ligands.<sup>186,210</sup> In this project, I aimed to develop inhibitors that specifically target activin A and its signalling.

In order to achieve this, two approaches were attempted in parallel – a small molecule approach and a peptide approach. The small molecule approach used XChem fragment screening to identify 13 molecules that bound to mature activin A, of which four were analysed further. Validation through a second crystallographic screen led me to focus on the molecule NM466, and a number of NM466 derivatives were designed. Of these derivatives, only one, TR17, was validated as an activin A binding molecule. The in-solution binding of NM466 and TR17 was then analysed through ligand-based NMR, however the data were inconclusive.

In parallel to the small molecule approach, the peptide approach succeeded in the generation of potent activin A inhibitors. First the structure of the pro-mature activin A complex was analysed to identify key motifs for the interaction between activin A's pro- and mature domains. Peptide fusions based on these motifs were then generated and their interactions with mature activin A investigated. From this, it was determined that the  $\alpha$ 1-helix and the loop were the key interacting motifs. It was then found that two peptide fusions containing these motifs, GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L, inhibited activin A signalling with IC<sub>50</sub> values of approximately 4.4  $\mu$ M and 3.8  $\mu$ M respectively. Optimisation of these fusions through non-covalent C-terminal dimerization utilising the dimerization and docking domain of the PKA RII $\alpha$  subunit resulted in an approximate 12-fold increase in inhibitor potency for H<sub>1</sub>LH<sub>2</sub> to 380 nM. Finally, stabilisation of the H<sub>1</sub>LH<sub>2</sub> dimers through the introduction of either a C-terminal disulfide bond or a disulfide linkage in the PKADD domain resulted in further increases in inhibitor potency to 45.2 nM for GB1-H<sub>1</sub>LH<sub>2</sub>-SS, and 71.5 nM for GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>. The specificity of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> was assessed through screening against three other TGF- $\beta$  family members, activin B, myostatin, and BMP2. I showed that GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> inhibited both activin A and activin B in a



similar manner (71.5 nM vs 88.0 nM) with a 100-fold specificity over myostatin (approximately 7.6  $\mu$ M). No inhibition was observed for BMP2 signalling.

## 5.2 The Putative Type I Receptor Binding Site as a Target for Small Molecule Inhibition

Due to their importance in activin A signalling, the type I and type II receptor binding sites make obvious targets for developing inhibitors. The type II binding site has been structurally identified by Greenwald *et al.* and Thompson *et al.* to exist on the convex face of the monomeric subunit of activin A, in the fingertip region.<sup>103,254</sup> Conversely the site of the type I receptor is believed to be within a hydrophobic groove on the concave face of the monomeric subunit. Though there is currently no structural data available for the interaction between ALK4 and activin A, structural data for the related family member GDF11 and ALK5 shows this hydrophobic groove is the site of interaction, thus increasing confidence in the case of activin A.<sup>76</sup> It was in this site that Zhu *et al.* designed several small molecule binders via *in silico* screening, of which the most potent, NUCC-555, inhibited activin A signalling with an IC<sub>50</sub> value of 5.3  $\mu$ M.<sup>193</sup>

Incidentally, the XChem screen reported in this thesis identified 12 fragment hits that bound in this same groove, by far the most commonly occupied site. One of these hits NM466 and its derivative TR17 were found to interact with mainly through hydrophobic interactions with residues W335, I339, F368, T371, M401, Y403, I415, and M418 forming one hydrogen bond with W338 (Figure 39A). Modelling of these residues found that several of these residues also form the majority of hydrophobic interactions with NUCC-555 (W335, W338, F368, Y403, and I415) with only additional hydrogen bonding interactions with K413 and N417 being observed. NUCC-555 was observed to compete with ALK4 for activin A binding suggesting that NM466 and TR17 may also do the same.

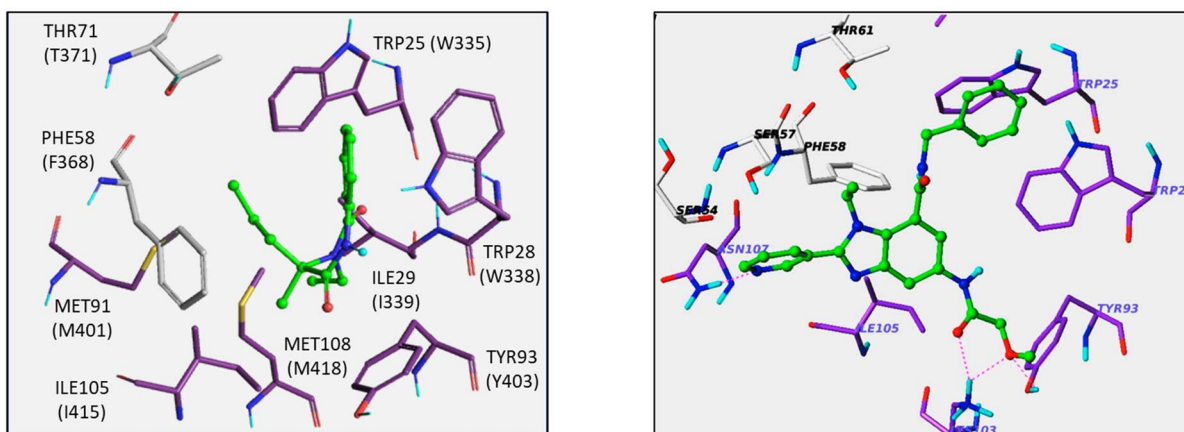
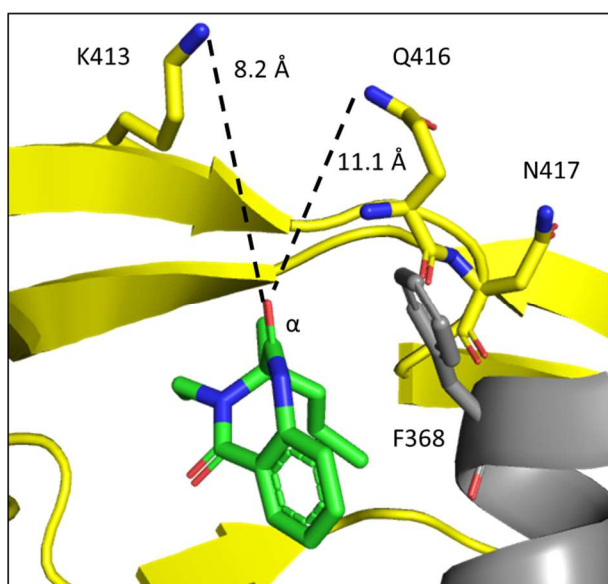
**A****B**

Figure 39: The putative type I receptor binding site as a target for inhibitors. A) Comparison of the binding site of TR17 (left, here shown as an example) with that modelled for NUCC-555 with activin A (right) by Zhu et al. (Figure of modelled NUCC-555 structure adapted from Zhu et al.<sup>193</sup>). Note mature domain numbering is shown in the part of the figure adopted from Zhu et al. B) The structure of TR17 bound to mature activin A with residues K413, Q416 and N417 highlighted. Here, the amide designated “ $\alpha$ ” of TR17 highlighted.

Many of the derivatives of NM466 screened were not observed to bind to activin A. Most of these molecules aimed to spatially optimise the weak hydrophobic interactions in NM466’s binding. However, due to the small number of derivatives that bound, it may be prudent look beyond the immediate binding site of NM466 and try to form new electrostatic and hydrogen bonding interactions. As such interactions beyond the immediate vicinity of NM466’s binding site may need to be utilised for optimisation. The hydrogen bonding interactions of NUCC-555 could be used as a guide for this. The two key hydrogen bonding interactions between NUCC-555 and activin A occur at residues K413 and N417.<sup>193</sup> Analysis of the crystal structure of NM466 and TR17 reveals the carbonyl of amide  $\alpha$  is oriented towards K413, 8.2 Å away, appearing to undergo no obvious interactions with any residue

of activin A. As such replacing this group with an alkyl chain containing hydrogen bonding groups (such as alcohols and carboxylic acids) would allow for hydrogen bonding with K413 to occur. Due to the positioning of F58 upon NM466/ TR17 binding, N417 is significantly less accessible than in the NCC-555-activin A model and thus it is not a reasonable target for interactions. However, Q416 is proximal to NM466/ TR17 (11.1 Å away) and more accessible than N417 meaning further derivatisation of the chain may also allow for hydrogen bonding to occur with this residue (Figure 39B). These increased interactions of a NM466/TR17 derivative will likely increase affinity for activin A and as such become a useful next step in inhibitor development.

### 5.3 Pro-domain Derived Peptides as Inhibitors of Activin A and the Wider TGF- $\beta$ Family

The inhibitors generated in this project represent a new chapter in the inhibition of activin A. Current methods that either utilise or target activin receptors or follistatin usually have specificity issues due to ligand promiscuity.<sup>186,210</sup> Antibodies overcome this problem, as do full pro-domain-based inhibitors, and both are highly specific, however they are synthesised using mammalian expression systems which are significantly more expensive than *E. coli*-based systems.<sup>188,198,236</sup> Thus, the most potent peptides reported in this thesis represent a happy medium, with relatively high specificity (100-fold more specific for activin A / B than myostatin) compared to receptor and follistatin approaches, yet utilising a rapid, low cost production method that results in a high inhibitor yield.

Compared to the AT pro-peptide (the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of activin A fused to the shoulder region of TGF- $\beta$ 1) and pro-domain ligand traps reported in Makanji *et al.* and Chen *et al.*, the two covalently dimerised peptide inhibitors have a lower potency – up to 14-fold lower than that of the most potent ligand trap (IC<sub>50</sub> GB1-H<sub>1</sub>LH<sub>2</sub>-SS = 45.2 nM, GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> = 71.5 nM, AT peptide = 10.3 nM, most potent ligand trap = 5 nM).<sup>188,236</sup> Interestingly, the Fc fusion reported in Chen *et al.* is specific for activin A signalling, showing a ~100-fold specificity over activin B, whereas the AT pro-peptide and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> are not (4-fold and approximately the same specificity over activin B respectively). This indicates that additional epitopes in the C-terminus of the pro-domain of activin A may be important in determining its specificity for activin A and not activin B. However, the identity of these epitopes is yet to be elucidated.

To my knowledge, this thesis represents the first attempt to use peptides derived from activin A's pro-domain to inhibit activin A signalling. However, it is not the first reported peptide-based approach to inhibit TGF- $\beta$  family ligands, with pro-domain derived peptides also being reported for myostatin, inhibin A and B, and GDF11.<sup>106,189,191</sup>

The trends observed for both binding and inhibition by the peptide fusions generated in this project mirror those observed by Takayama *et al.* for peptides based on myostatin's pro-domain.<sup>189</sup> Two peptides developed by Takayama *et al.* based on the murine pro-domain inhibited human myostatin signalling with IC<sub>50</sub> values of 4.1 μM and 3.5 μM respectively, highly similar to the IC<sub>50</sub> values obtained for GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L of ~4.4 μM and ~3.8 μM. This is not where the similarities end. The minimal binding peptide that I determined for activin A is the 19mer GB1-E53-P72 whereas the minimal binding peptide identified by Takayama for myostatin is a 23mer. Sequence alignment of the two sequences shows that the majority of the highly conserved hydrophobic residues across TGF-β family members are present in both (Figure 40A). Analysis of the structures of human pro-mature activin A and human pro-mature myostatin with the minimal peptides highlighted reveals that both correspond to the majority of the α1-helix and initial loop residues, however the surface area of interaction on the mature domain is slightly higher for myostatin (activin A = 1845 Å<sup>2</sup>, myostatin = 2046 Å<sup>2</sup>) and the myostatin α1-helix has a higher number of interacting hydrophobic residues (activin A = 6 hydrophobic residues, myostatin = 9) (Figure 40B). These differences may be the explanation behind the differences in measured affinities, with the affinities for GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L being ~10-fold weaker than that of the myostatin peptides (642 nM and 909 nM vs 35.9 nM and 29.7 nM). Dimerization has not been employed with the myostatin peptides and the most potent inhibitor to date is a monomeric stapled peptide that inhibits myostatin signalling with an IC<sub>50</sub> value of 260 nM.<sup>190</sup> This is still significantly less potent than the dimeric inhibitors developed in this thesis, GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>, which inhibited signalling with IC<sub>50</sub> values of 45.2 nM and 71.5 nM respectively.

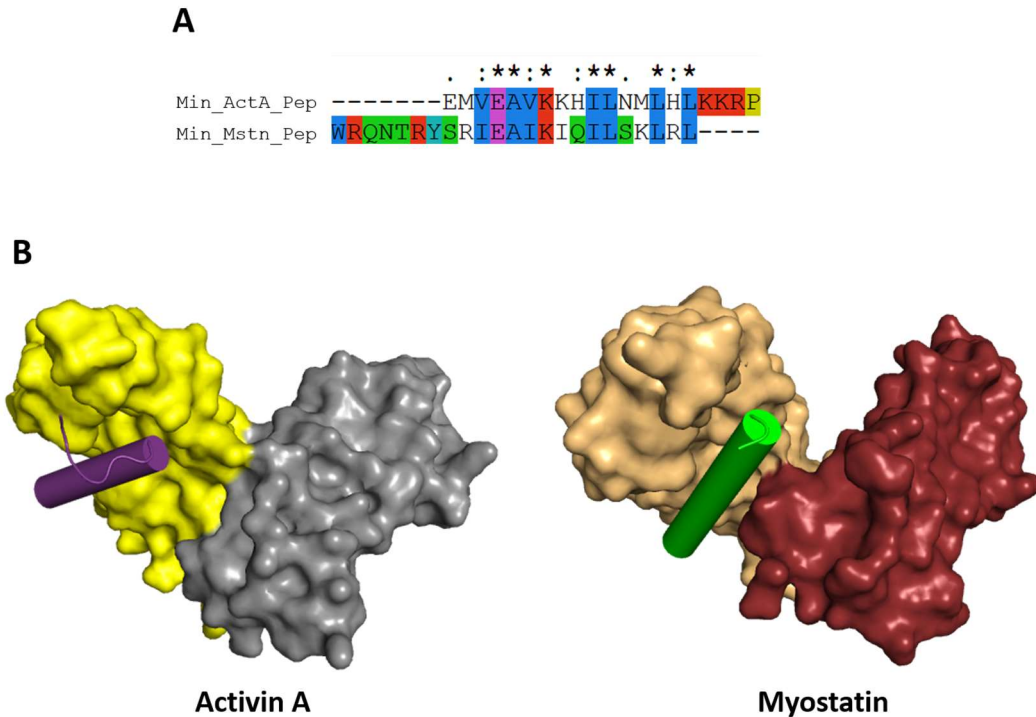


Figure 40: Comparison of the activin A and myostatin minimal binding peptides. A) Alignment of the activin A and myostatin minimal binding peptides performed using ClustalX2. Here hydrophobic residues are shown in blue. Polar residues are shown in green, basic residues in red and proline residues in yellow. Here “\*” denotes identical residues whereas “:” denotes highly similar residues. B) Representations of the minimal binding peptide of the activin A pro-domain (left, purple) and myostatin pro-domain (right, green) bound to their respective mature domains.

Similar trends are seen in terms of specificity for the activin and myostatin systems as well. In Ohsawa *et al.*, Fc fusions of a 29mer containing the humanised version of the peptides reported in Takayama *et al.* preferentially inhibited myostatin and GDF11 over TGF- $\beta$ 1 and activin A signalling (though no IC<sub>50</sub> values were reported so this specificity is difficult to quantify).<sup>205</sup> Similarly, GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>55</sub> had a 100-fold specificity for activin A and B over myostatin. This therefore suggests, pro-domain derived peptides are a good approach for generating TGF- $\beta$  family inhibitors with a high specificity compared to other methods.

In addition to myostatin, pro-domain derived peptides have also been reported for inhibin.<sup>191</sup> Walton *et al.* describe the generation of a 43mer peptide based on the inhibin  $\alpha$  subunit that specifically inhibited inhibin A and B signalling with IC<sub>50</sub> values of 255 and 150 nM respectively. These IC<sub>50</sub> values are significantly higher than those reported for unmodified monomeric activin A and myostatin peptides, however it is likely due to the inhibin derived peptide preventing association of inhibin A/B to the co-receptor betaglycan rather than a type I receptor, thus the values are not directly comparable. Incidentally, it is possible that the inhibitors reported in this thesis may also have an inhibitory effect on inhibin signalling due to their interactions with the  $\beta_A$  and  $\beta_B$  subunits, however

the efficacy of this inhibition is likely to be similar to that observed for the monomeric peptides GB1-H<sub>1</sub>LH<sub>2</sub> and GB1-H<sub>1</sub>L due to the inhibin being a heterodimer of  $\alpha$  and  $\beta$  subunits.

In contrast to activin A, inhibin and myostatin pro-domain derived peptides, Pepinsky *et al.* observed that a sequence corresponding to the  $\alpha$ 1-helix-loop- $\alpha$ 2-helix motif of GDF11 had no inhibitory effect on GDF11 signalling at the concentrations tested.<sup>106</sup> This is unexpected as, like myostatin, the pro-mature GDF11 is a latent and the pro-domain of GDF11 is a known antagonist of GDF11 signalling (IC<sub>50</sub> value unreported). However, the peptide concentrations used in this study were significantly lower than those used in this thesis or Takayama *et al.* Interestingly, it was observed that this GDF11 derived peptide, PDP<sub>60-114</sub>, was able to improve the solubility of mature GDF11, especially at neutral pH. This was not tested with the peptide fusions described in this thesis for activin A but given the similarity between the two systems, it is possible that the activin A derived peptides may also be able to aid the solubility mature activin A. Like mature GDF11, mature activin A has poor solubility at physiological pH. As such, this could have applications in experimental biology, such as in biochemical assays and animal studies, where the presence of soluble activin A for an extended period of time is required.

## 5.4 Future Directions of Peptide Inhibitors of Activin A

In this thesis, I have demonstrated the optimisation of pro-domain derived peptide inhibitors of activin A through covalent dimerization, resulting in an approximately 50-100-fold increase in inhibitor potency compared to the monomeric form. However, to improve their viability as chemical tools it is still desirable to optimise these inhibitors further.

So far, I have shown the key pro-domain epitopes for the inhibition of activin A signalling to be the  $\alpha$ 1-helix and the loop. Furthermore, it is within this region where the key residues for interaction with the mature domain are contained. Consequently, the  $\alpha$ 2-helix does not contribute significantly to inhibition or binding, however it is present in the two most potent peptide inhibitors, GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>. Therefore, increasing interactions between this region and the mature domain may lead to a more potent inhibitor of activin signalling. One way of achieving this would be to mimic the interactions of follistatin with mature activin A. Like the  $\alpha$ 2-helix, the primary interactions between follistatin and mature activin A occur in the vicinity of the type II receptor binding site, with interactions in the type I site not being necessary to achieve inhibition.<sup>77</sup> Follistatin is a much more potent inhibitor of activin A signalling with an IC<sub>50</sub> value of 90 pM (FS288 isoform). As such, mimicking the crucial interactions between follistatin and the type II site could increase the potency of H<sub>1</sub>LH<sub>2</sub> inhibitors. Harrington *et al.* identifies one of the crucial interactions in the Fs2 domain to be the interaction between R192 of follistatin with D337 and Y404. Alignment of the structure of the Fs12-

ActA complex (PDB ID: 2ARP) to the pro-mature complex reveals residue of K80 of the  $\alpha 2$  helix to be occupy a similar spatial position to R192 relative to Y404 of the mature domain (Figure 41). Thus, the most promising mutation would be K80R. It is however important to note that D337 forms a key salt bridge with R71 and thus these mutations may also destabilise the peptide-activin A complex.

A second potential mechanism for optimising inhibition would be to reduce the conformational flexibility of the peptide. Inspiration for this comes from Chen *et al.* who incorporated fastener residues into the pro-domain of activin A. Highly conserved across the latent TGF- $\beta$ s, fastener residues occur in the  $\alpha 1$ -helix, shoulder region and mature domains that, through a network of interactions with one another, reduce the conformational flexibility of the  $\alpha 1$ -helix. This stabilises binding of the pro-domain to the mature domain. Though H<sub>1</sub>LH<sub>2</sub> does not contain the region where the majority of these residues lie, “locking” H<sub>1</sub>LH<sub>2</sub> through another mechanism may result in an increase in inhibitor potency. The most promising site to do this would be between the  $\alpha 1$ -helix and the loop. This is the region in which most of the crucial interactions between H<sub>1</sub>LH<sub>2</sub> and mature activin A occur, so conformationally restricting these residues may minimise the entropic penalty to binding. The most obvious method of “locking” the conformation would be the use of a chemical staple. Stapling has already been shown to improve potency in the myostatin system with the most potent peptide inhibitor of myostatin to date containing a diester staple.<sup>190</sup> Though stapling resulted in a reduction in affinity when attempted in this thesis, this is likely due to the position or nature of the staple and more work is needed to explore the feasibility and benefit of stapling for activin-derived peptides. Analysis of the crystal structure suggests that K60 and K70 are the most promising sites to introduce a staple due to spatial proximity and lack of obvious interactions with the mature domain (Figure 41B). Thus, introducing a staple here may represents the next avenue for optimisation of the dimeric peptide fusions reported in this thesis.

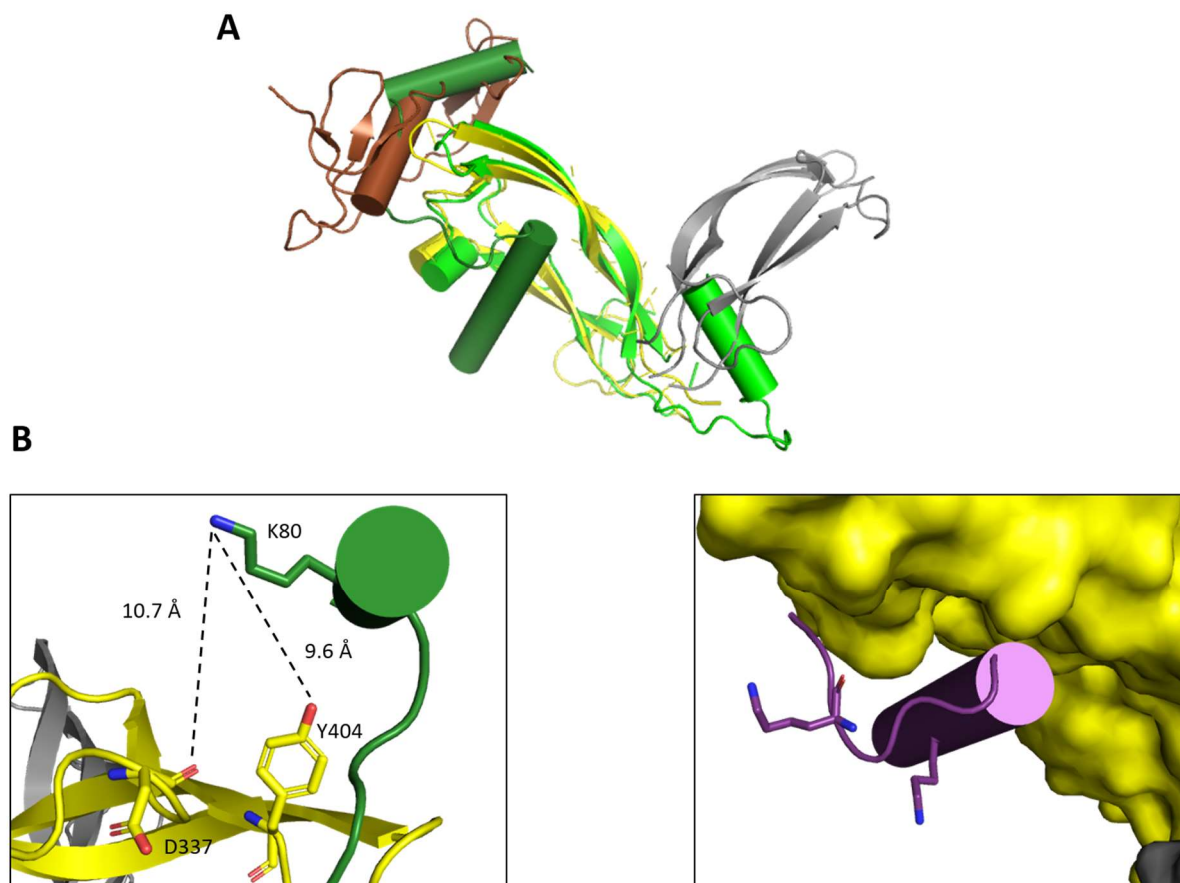


Figure 41: Potential optimisations of  $H_1LH_2$ . A) Alignment of the  $Fs_{12}$ -ActA structure to the representation of  $H_1LH_2$  bound to mature activin A based on the pro-mature ActA complex. Here the  $Fs$ -ActA complex is shown in green/ white (mature activin A), and light brown ( $Fs_2$ ). B) The  $\alpha_2$ -helix of pro-activin A bound to mature activin A with residues K80, D337, and Y404 highlighted (left).  $H_1LH_2$  bound to mature activin A with potential sites for stapling, K60 and K70, highlighted (right).

## 5.5 Potential Applications of TGF- $\beta$ Family Pro-domain Derived Peptides

Overall, this project has demonstrated a rational method for designing and generating specific inhibitors of activin A and B signalling via a peptide approach, using the interacting motifs of activin A's pro-domain as a basis for inhibitor design. To date, this represents only the third member of the TGF- $\beta$  family for which a pro-domain derived peptide approach has been thoroughly investigated, along with myostatin and inhibin (the concentrations tested for the GDF11 derived peptide were significantly lower than those reported in this thesis or Takayama *et al.* for pro-domain derived peptides).<sup>106,189,191</sup> As such, there is potential for expansion of this methodology to other members of the family. In the case of the myostatin, inhibin and activin A, specificity was high with only very closely related ligands being inhibited in each case. It is still highly difficult to generate specific inhibitors for TGF- $\beta$  family signalling, so this approach represents a viable means of achieving this. Expansion to the



wider TGF- $\beta$  family could result in therapeutic applications. TGF- $\beta$  signalling has been reported to have oncogenic effects in a number of cancers, facilitating epithelial-mesenchymal transitions, suppressing immune response, and aiding distal metastasis, and is thus an important clinical target.<sup>255</sup> Furthermore, though activin signalling is the main causative agent of FOP, the ossifying TGF- $\beta$  proteins BMP2 and BMP4 have been observed to have elevated levels in FOP.<sup>256</sup> Thus, using a pro-domain based peptide method to develop inhibitors for these proteins, combined with the H<sub>1</sub>LH<sub>2</sub> activin peptides reported here could be a viable approach for developing FOP therapeutics.

Furthermore, to the best of the author's knowledge, this thesis represents the first reported application of the PKADD domain as a fusion partner to induce dimerization in a system outside of the PKA RII $\alpha$  subunit itself. This could be especially important in developing inhibitors of other protein targets that contain two sites of interest that are appropriately spaced and oriented away from each other, including other TGF- $\beta$  family members. Therefore, the work described here may aid in a greater understanding of TGF- $\beta$  family signalling and represent a new approach for targeting multivalent systems.

Comparing the two approaches taken in this thesis, it is clear that the peptide approach was much more successful than the small molecule approach. Whereas the small molecule approach failed to identify any molecules that bound to activin A with high affinity, the peptide approach not only identified peptides that bind activin A with high affinity but can also inhibit its signalling with nanomolar potency. Furthermore, due to the structural and sequence similarity between the mature domains TGF- $\beta$  family members, a peptide approach is much more likely to yield a specific inhibitor due to the larger number of interactions that can be fine-tuned for specificity. As such, based on the evidence presented here, I believe it would be much more prudent for a peptide-based approach to be taken rather than a small molecule approach when designing inhibitors that target other ligands in the TGF- $\beta$  family.



---

## Material and Methods

---

## 6.0 Materials

All materials described in this thesis were obtained from the communal stocks of the Hyvönen Lab, University of Cambridge unless stated otherwise. All solutions and buffers were made from these stocks.

### 6.1 Crystallisation of Mature Activin A

Purified mature activin A for all experiments was kindly provided by Mrs Katharina Ravn. Screens were performed around the original conditions as described in Harrington *et al.* in a 24-well format.<sup>71</sup> Lyophilised mature activin A was resuspended in 20 % acetonitrile and pipetted onto a glass cover slip in a 1  $\mu$ L : 1  $\mu$ L with well solution. The slips were then placed over the wells of an XRL pater (Molecular Dimensions) 24 well plate containing 1 mL well solution and sealed with high vacuum grease (Dow Corning). Crystals of activin A of grew best in 1.65 M  $(\text{NH}_4)_2\text{SO}_4$ , 4 % PEG 300, 100 mM Hepes pH 7.4, 2 % DMSO, 3.6 mg/mL activin A.

For crystals grown in a 96-well format, lyophilised mature activin A was resuspended in 20 % acetonitrile to a concentration of 3.6 mg/mL. The protein was then crystallised in 1.55 M  $(\text{NH}_4)_2\text{SO}_4$ , 100 mM Hepes pH 7.4, 4 % PEG 300, 8 % DMSO at a ratio of 200 nl well solution : 200 nl protein using a Mosquito Crystal (TTP Labtech) crystallization robot. Crystal drops were either set up in MRC 2 Well Crystallization Plates in UVP (Jena Bioscience) or 96-well 3-drop Swissci plates (Molecular Dimensions) for soaking at XChem facility.

For crystals containing no PEG 300, a microseed stock was generated from mature activin A crystals grown as previously described. A crystal drop containing 5 or fewer crystals was added to 50  $\mu$ L well solution in a tube containing a single Microseed Bead (Molecular Dimensions). The mixture was then vortexed in 3  $\times$  30 s intervals with incubation on ice for 30 s in between. 96 well plates were then set up using a Mosquito Crystal (TTP Labtech) crystallization robot with a ratio of 200 nl protein (3.3 mg/mL): 150 nl well solution (1.55 M  $(\text{NH}_4)_2\text{SO}_4$ , 100 mM Hepes pH 7.4, 8% DMSO) : 50 nl microseed stock.

In all cases crystals were allowed to grow for a minimum of 3 days at 18 °C before being flash cooled in liquid nitrogen at 98 K for storage before data collection.

### 6.2 Structural determination of mature activin A crystals

All crystals were shot using either the i03, i04-1, or i24 beamlines at the Diamond Light Source, Harwell, UK. Diffraction data was processed using the autoProc<sup>257</sup> (Global Phasing) automatic data processing pipeline within IsPyB.<sup>258</sup> All data was managed using CCP4i. Structures were solved using

molecular replacement with either 2ARV (DMSO structure) or the DMSO-containing structure (all other datasets) as search model in using Phaser MR.<sup>259</sup> Models were refined using refmac5<sup>260</sup> and Coot.<sup>261</sup> Ligands were modelled using Grade (Global Phasing). All crystals belonged to the space group I222.

### 6.3 XChem Screening

Crystals of mature activin A were soaked at with fragments of the DSPL, Kidd3D and Edelris Keychemical Fragment libraries at concentrations of 25 mM/ 50 mM (library dependent) in DMSO using an ECHO acoustic liquid handler (Labcyte). Crystals were then allowed to equilibrate at 18 °C for 1 hour before being vitrified in liquid nitrogen at 98 K. Diffraction data was then collected on the i04-1 beamline at the Diamond Light Source, Harwell, UK. Data was managed using XChemExplorer and the apo structure of mature activin A grown in 2 % DMSO was used as the model reference. Data was then processed using PanDDA Analyse as described in Pearce *et al.*<sup>142</sup> and models refined using PanDDA Inspect.

### 6.4 Fragment soaking of mature activin A

Compounds for fragment soaking were dissolved to a final concentration of 40 mM in 1.55 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM Hepes pH 7.4, 8% DMSO. 5 µL of this solution was added to wells containing mature activin A crystals and the crystals allowed to equilibrate for 1 hour at 18 °C. Crystals were then harvested and vitrified in liquid nitrogen at 98 K.

### 6.5 Cloning of Peptide Fusions and Vectors

To generate the vectors required in this project primers were designed such as to generate the DNA encoding the relevant tags and dimerization domains with appropriate 5' and 3' restriction sites (Appendix 3).

pOP4BP was generated from the pOP5BP vector (<http://hyvonen.bioc.cam.ac.uk/pOP-vectors/>) with the C-terminal Avi Tag being replaced by a Strep-tag II. To achieve this, pOP5BP was restriction digested with *Xho*I and *Hind*III (New England Biolabs) to remove the Avi Tag from the vector, before being dephosphorylated with shrimp alkaline phosphatase (GE Healthcare) and purified using Gel Extraction (ThermoScientific). Primers encoding the Strep Tag II with appropriate 5' and 3' restriction sites were then mixed and heated to 98 °C for 60 s before allowing to cool to room temperature. The annealed primers were then phosphorylated using polynucleotide kinase and ligated into the digested vector using Quick ligase (New England Biolabs). The resulting ligation mixture was transformed into 50 µL of DH5α E.coli competent cells. Cells were grown overnight at 37 °C on agar plates containing

100 µg/mL ampicillin before a single colony from each plate was removed and added to 5 mL LB with 100 µg/mL ampicillin. This mixture was then grown overnight at 37 °C with shaking. Plasmid DNA was extracted from the pellet using a Miniprep Kit (ThermoScientific). DNA concentration was determined by measuring the absorbance at 260 nm using a NanoDrop 2000 (Thermoscientific).

In the case of pOP6BPA and all peptide fusions, DNA encoding the various tags, fusion partners and pro-domain derived sequences were generated via PCR. Primers were designed containing appropriate 5' and 3' restriction sites (Restrictions sites used are listed alongside the primers in Appendix 3). PCR was performed using a template of pOP4BP (Strep-tag II + GB1 domain) and pGEX-PKADD (PKADD domain) (provided by Dr. Katherine Stott) for pOP6BPA. To confirm the identity of PCR products, 5 µL product was analysed on a 1% agarose gel at 80 V for 50 minutes. PCR products were then purified via a PCR Purification Kit (ThermoScientific) before being ligated into appropriately digested pOP2H and transformed as described above.

Similarly, PCR for the amplification of DNA encoding the peptides was performed using a template of pHAT2-pro-activin A (provided by Dr. Xuelu Wang, Department of Biochemistry, University of Cambridge). Continuous fused helix peptides PCR was performed according to in vitro assembly cloning protocol described in García-Nafria *et al.* using pHAT2-pro-activin A and pOP6BPA as a template.<sup>262</sup> Identity was confirmed, and the purified products were ligated into appropriately digested pOP4BP or pOP6BPA depending on whether they were PKADD domain fusions or not. Again, transformation was performed as described above.

All constructs were verified by DNA sequencing at Source Bioscience.

## 6.6 Expression Tests

Plasmids encoding the peptide fusions were transformed into BL21 (DE3) cells and added to 4 mL LB with 100 µg/mL ampicillin before growing for 3 hours. 900 µL of each sample was then added to 100 µL 80% glycerol and mixed thoroughly to generate glycerol stocks before storing at -80 °C. The remaining culture was then split into 3 × 1 mL samples and IPTG was added to a concentration of 400 µM. The individual samples were then grown shaking at 16 °C overnight, 25 °C for 4 hours, and 37 °C for 3 hours respectively. 500 µL aliquots of each sample were then centrifuged at 17,000 × g for 3 mins and the supernatant removed. 50 µL of a lysis master mix (1 mL BugBuster, 5 µL Lysonase (Novagen)) was then added to each sample and vortexed thoroughly before being incubated at room temperature for 10 minutes. Samples were then centrifuged at 17,000 × g for 3 mins and the supernatant removed and kept for further analysis. The pellets were then resuspended in 20% lysis master mix and centrifuged at 17,000 × g for 3 mins and the supernatant removed. The pellets were then resuspended

in 20% lysis master mix taken forward for analysis via SDS-PAGE along with the initial supernatant after lysis.

## 6.7 SDS-PAGE

All proteins samples were analysed via SDS-PAGE using 15 % polyacrylamide gels (acrylamide : bisacrylamide = 37.5 : 1). Gel solution (5 ml) was polymerised (240) 10 % w/v ammonium persulfate (240  $\mu$ l) in water and Tetramethylethylenediamine (150  $\mu$ l). A 5 % stacking gel was similarly polymerised atop the 15% polyacrylamide gel to allow for sample loading. Samples were diluted 1:1 with either non-reducing (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.02% bromophenol blue) or reducing (+ 8%  $\beta$ -mercaptoethanol) loading dye and heated to 95 °C for 5 minutes before 10  $\mu$ l of each sample was loaded onto the gel and run for 55 minutes in 1 $\times$  Laemmli Running Buffer (25 mM Tris HCl, 192 mM glycine, 0.1% w/v SDS) at 200 V, 400 mA. The gels were then stained for 60 minutes with Coomassie Blue Stain (0.1% w/v Coomassie Brilliant Blue R, 10% ethanol, 10% mL acetic acid) before destaining with 10% acetic acid.

## 6.8 Large Scale Expression

To express fusion proteins on a large scale, glycerol stocks of each fusion was streaked on ampicillin agar plates and incubated overnight at 37 °C. A single colony from each plate was then added to 10 mL LB with 100  $\mu$ g/mL ampicillin before being incubated at 37 °C shaking overnight. 10 mL of culture was then added to 1 L 2 $\times$ YT with 100  $\mu$ g/mL ampicillin and incubated shaking at 37 °C until an optical density of 0.7-1.0 was reached. IPTG was added to a final concentration of 400  $\mu$ M IPTG and each sample was grown at 25 °C/ 37 °C shaking for a further 3/4 hours or overnight depending on the optimal temperature determined in the expression test. Cultures were then centrifuged at 4000  $\times$  g for 20 minutes at 4 °C and the supernatant removed. The pellet of each sample was then resuspended in 25 mL Resuspension Buffer (50 mM NaPi pH 7.0, 300 mM NaCl, 10 mM imidazole, 0.5 cComplete mini EDTA free protease inhibitor cocktail tablet) and passed through a French press (Avestin) six times at approximately 5 MPa. Samples were then centrifuged at 16,600  $\times$  g for 30 minutes at 4 °C and the supernatant collected for purification.

## 6.9 Metal Affinity Purification

To purify the fusion proteins from the supernatant, the supernatant was then incubated with Ni-NTA agarose beads (Cube Biotech) at 4 °C for 1 hour. Beads were then loaded onto gravity flow columns and washed with 2  $\times$  10 mL of Ni Wash Buffer (50 mM NaPi pH 7.0, 300 mM NaCl, 10 mM imidazole) followed by a further 3 mL wash. The protein was then eluted with 10 mL Ni Elution Buffer (50 mM

NaPi pH 7.0, 300 mM NaCl, 250 mM imidazole) and the sample was analysed via SDS-PAGE. Protein concentration was calculated by measuring absorbance at 280 nm using a NanoDrop 2000 (ThermoScientific). Concentration was then determined via the Beer Lambert law using extinction coefficients as calculated by ProtParam (ExpASY).<sup>263</sup>

Metal affinity purification was also performed GB1-H<sub>1</sub>LH<sub>2</sub>-SS and GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> using a gravity flow column assembled as described above with an ÄKTApurifier (GE Healthcare). Samples loaded onto the column was purified using two buffers, Buffer A (50 mM NaPi pH 7.0, 300 mM NaCl, 10 mM imidazole) and Buffer B (50 mM NaPi pH 7.0, 300 mM NaCl, 500 mM imidazole), where an increasing concentration of Buffer B was used from 0% - 100%. 1 mL fractions were collected and the absorbance at 280 nm was measured. In the case of GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub> both Buffer A and Buffer B contained 6 M urea and absorbance at 580 nm was also measured. Fractions were then analysed via SDS-PAGE.

#### 6.10 StrepTactin Affinity Purification

Pooled elution fractions from the metal affinity purification (5 mL) were then added to 1 mL of StrepTactin Sepharose High Performance beads (GE Healthcare) suspended in 10 mL StrepTactin Wash Buffer (100 mM Tris HCl pH 8.0, 300 mM NaCl) and incubated for 1 hour at room temperature. The beads were then loaded onto gravity flow columns and washed with 2 × 10 mL of StrepTactin Wash Buffer before eluting with 5 mL StrepTactin Elution Buffer (100 mM Tris HCl pH 8.0, 300 mM NaCl, 2.5 mM d-desthiobiotin). Samples were then buffer exchanged into 100 mM Tris HCl pH 8.0, 5 mM EDTA using a PD-10 column (GE Healthcare) and concentrated using an Amicon Ultra-15 centrifugal filter MWCO 3k or 10k (Millipore) as appropriate.

#### 6.11 Biolayer Interferometry

Biolayer Interferometry experiments were performed using an Octet RED96 instrument (ForteBio). Anti-penta-HIS biosensors were regenerated 3 times in 10 mM glycine pH 1.7 followed by Kinetic Buffer (PBS, 0.1% BSA, 0.02% Tween 20) and baselined in Kinetic Buffer for 60 s. Peptide fusions at a concentration of 1 µg/mL in Kinetic Buffer were then loaded onto the sensors for 90 s. The loaded biosensors were then immersed in Kinetic Buffer for 60 s to establish a baseline before being immersed in a solution of mature activin A at 100 nM or 1 µM in Kinetic Buffer for 300 s to determine the association rate. The sensors were then moved back to the baseline position for 300 s to determine the dissociation rate. All experiments were performed at 30 °C. The data was then analysed using Prism 8 software (GraphPad) according to the equations as described by Motulsky.<sup>264</sup>



## 6.12 Isothermal Calorimetry

Isothermal Calorimetry Experiments were performed using a Microcal iTC200 instrument (GE Healthcare). Peptide Fusions were exchanged into 100 mM Tris HCl pH 8.0, 5 mM EDTA to a concentration of approximately 100  $\mu$ M for monomeric peptides or 50  $\mu$ M for dimeric peptides. Lyophilised activin (250  $\mu$ g) was resuspended in 100  $\mu$ L of 10 mM HCl before being diluted to a concentration of approximately 4.5  $\mu$ M in 400  $\mu$ L of 100 mM Tris HCl pH 8.0, 5 mM EDTA, 3.2  $\mu$ M BSA. The 200  $\mu$ L of activin solution was then loaded into the cell and bubbles removed before loading the peptide 50  $\mu$ L of cell solution into the syringe. The cell was then set to 25  $^{\circ}$ C and the stirring speed set to 750 rpm. The syringe was placed into the cell and the cell was allowed to equilibrate until the measured power difference between the experimental and references cells was  $> 0.4$   $\mu$ cal/s. The experiment was then run according to the following protocol.

Injection	Volume / $\mu$ L	Spacing / s
1	0.2	240
2-20	2	240

Data analysis was performed using Origin 7.0 software (Origin). After initial analysis, ligand concentration was adjusted to give a stoichiometry of the nearest whole number if the original stoichiometry was within 0.2 of that number.  $K_d$  was then determined from these adjusted values.

In the case of synthetic peptides, the syringe and cell solutions also contained 2% DMSO.

## 6.13 Luciferase Assays

In order to measure  $IC_{50}$  values of the peptides, luciferase assays were performed. HEK293T cells were cultured in DMEM (Life Technologies) + 10% v/v Foetal Calf Serum (Life Technologies) in a 37  $^{\circ}$ C in a humidified incubator with 5%  $CO_2$  until a confluency of 80% was achieved. The cells were then diluted to a density of  $8 \times 10^4$  cells and 100  $\mu$ L was added to every well of a polylysine (Sigma Aldrich) treated Nunc Microwell 96F plate (ThermoScientific) and incubated overnight as previously described. After 24 hours, 33 ng of pGL3-CAGA (carrying the activin A responsive firefly luciferase gene) and 17 ng of pRL-SV40 (Promega) (with constitutively expressed renilla luciferase) were added to every well with 200 nl of FuGENE HD (Promega) to transfect the cells. After a further 24 hours each well was washed with 100  $\mu$ L PBS (Sigma Aldrich) and cultured in DMEM containing 0.5 % FCS and 100 pM activin. Each well contained peptide fusion with concentrations ranging from 90 pM to 20  $\mu$ M and each fusion was analysed in triplicate. Three wells were additionally cultured using 100 pM activin A without inhibitor and three wells were cultured without activin or inhibitor. The cells were then incubated for 8 hours before washing with 100  $\mu$ L PBS and lysed with shaking using 20  $\mu$ L of Passive Lysis Buffer (Promega)

for 15 minutes. 5  $\mu$ L from each well was then added to a 384 well low volume assay plate (Corning). Each well was injected with 25  $\mu$ L of either firefly or renilla luciferase substrate using a Pherastar plate reader (BMG Labtech) and luminescence was recorded. The ratio of firefly to renilla response was then determined and this ratio was normalised to the average ratio in 3 wells that contained 100 pM activin A only and 3 wells that contained no activin A. Data was then analysed using Prism 8 (GraphPad)

For activin B, myostatin, and BMP2, concentrations of 80 pM, 250 pM, and 2.5 nM were used respectively.

#### 6.14 SEC-MALS analysis

SEC-MALS analysis was performed using a Superdex 75 increase 10/300 column (GE Healthcare), connected to the DWAN Helios light scattering detector (Wyatt Technology) and the Optilab T-rEX refractive index detector (Wyatt technology). The system was equilibrated overnight in StrepTactin Wash Buffer (100 mM Tris pH 8.0, 300 nM NaCl) and calibrated using BSA (Sigma Aldrich). After calibration, 50  $\mu$ L of protein sample at a concentration of 2.5 mg/mL were analysed and the experimental data was processed using ASTRA (Wyatt Technology) software.

#### 6.15 $\text{CuSO}_4$ catalysed disulfide formation of GB1-H<sub>1</sub>HL<sub>2</sub>-SS and GB1-H<sub>1</sub>HL<sub>2</sub>-PKADD<sub>SS</sub>

5 mL StrepTactin purified protein was incubated with 1 mM  $\text{CuSO}_4$  at 4 °C with exposure to air overnight. Urea was then added to the samples to a final concentration of 6 M. Samples were then purified using metal affinity purification via an ÄKTApurifier (GE Healthcare) as described above.

#### 6.16 Mass Spectrometry

All mass spectrometry experiments described in this thesis were performed via MALDI-ToF by Dr. Mike Deery or Dr. Anja Andrejeva at the Cambridge Centre for Proteomics, Cambridge Systems Biology Centre, University of Cambridge

---

## References

---

1. Wu, M. Y. & Hill, C. S. TGF- $\beta$  Superfamily Signaling in Embryonic Development and Homeostasis. *Dev. Cell* **16**, 329–343 (2009).
2. Herpin, A., Lelong, C. & Favrel, P. Transforming growth factor- $\beta$  related proteins : an ancestral and widespread superfamily of cytokines in metazoans. *Dev. Comp. Immunol.* **28**, 461–485 (2004).
3. Feng, X.-H. & Derynck, R. Specificity and Versatility in Tgf-B Signaling Through Smads. *Annu. Rev. Cell Dev. Biol.* **21**, 659–693 (2005).
4. Cotton, T. R. The Structure of Human pro-Myostatin and Molecular Basis of Latency. (University of Cambridge, 2018).
5. Heldin, C.-H. & Moustakas, A. Signaling Receptors for TGF- $\beta$  Family Members. *Cold Spring Harb. Perspect. Biol.* **8**, (2016).
6. Vale, W. *et al.* Purification and characterization of an FSH releasing protein from porcine ovarian follicular fluid. *Nature* **321**, 776–779 (1986).
7. Latres, E. *et al.* Activin A more prominently regulates muscle mass in primates than does GDF8. *Nat. Commun.* **8**, 1–13 (2017).
8. Jones, K. L. *et al.* Activin A is a critical component of the inflammatory response, and its binding protein, follistatin, reduces mortality in endotoxemia. *Proc. Natl. Acad. Sci.* **104**, 16239–16244 (2007).
9. Antsiferova, M. & Werner, S. The bright and the dark sides of activin in wound healing and cancer. *J. Cell Sci.* **125**, 3929–3937 (2012).
10. Hatsell, S. J. *et al.* ACVR1R206H receptor mutation causes fibrodysplasia ossificans progressiva by imparting responsiveness to activin A. *Sci. Transl. Med.* **7**, 1–12 (2015).
11. Lemmon, M. A. & Schlessinger, J. Cell Signaling by Receptor Tyrosine Kinases. *Cell* **141**, 1117–1134 (2010).
12. Ferrara, N., Gerber, H.-P. & LeCouter, J. The biology of VEGF and its receptors. *Nat. Med.* **9**, 669–676 (2003).

13. Chen, Y. G., Liu, F. & Massague, J. Mechanism of TGFbeta receptor inhibition by FKBP12. *EMBO J.* **16**, 3866–3876 (1997).
14. Attisano, L., Wrana, J. L., Montalvo, E. & Massagué, J. Activation of signalling by the activin receptor complex. *Mol. Cell. Biol.* **16**, 1066–1073 (1996).
15. Wrana, J. L. *et al.* TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* **71**, 1003–1014 (1992).
16. Huang, T. *et al.* TGF- $\beta$  signalling is mediated by two autonomously functioning T $\beta$ RI:T $\beta$ RII pairs. *EMBO J.* **30**, 1263–1276 (2011).
17. Olsen, O. E. *et al.* Activin A inhibits BMP-signaling by binding ACVR2A and ACVR2B. *Cell Commun. Signal.* **13**, 1–7 (2015).
18. Goumans, M. J. *et al.* Balancing the activation state of the endothelium via two distinct TGF- $\beta$  type I receptors. *EMBO J.* **21**, 1743–1753 (2002).
19. di Clemente, N., Josso, N., Gouédard, L. & Belville, C. Components of the anti-Müllerian hormone signaling pathway in gonads. *Mol. Cell. Endocrinol.* **211**, 9–14 (2003).
20. Belville, C., Jamin, S. P., Picard, J.-Y., Josso, N. & di Clemente, N. Role of type I receptors for anti-Müllerian hormone in the SMAT-1 Sertoli cell line. *Oncogene* **24**, 4984–4992 (2005).
21. Huse, M., Chen, Y. G., Massagué, J. & Kuriyan, J. Crystal structure of the cytoplasmic domain of the type I TGF beta receptor in complex with FKBP12. *Cell* **96**, 425–436 (1999).
22. Lawler, S. *et al.* The Type II Transforming Growth Factor- $\beta$  Receptor Autophosphorylates Not Only on Serine and Threonine but Also on Tyrosine Residues. *J. Biol. Chem.* **272**, 14850–14859 (1997).
23. Luo, K. & Lodish, H. F. Positive and negative regulation of type II TGF- $\beta$  receptor signal transduction by autophosphorylation on multiple serine residues. *EMBO J.* **16**, 1970–1981 (1997).

24. Groppe, J. *et al.* Cooperative Assembly of TGF- $\beta$  Superfamily Signaling Complexes Is Mediated by Two Disparate Mechanisms and Distinct Modes of Receptor Binding. *Mol. Cell* **29**, 157–168 (2008).
25. Radaev, S. *et al.* Ternary Complex of Transforming Growth Factor- $\beta$ 1 Reveals Isoform-specific Ligand Recognition and Receptor Recruitment in the Superfamily. *J. Biol. Chem.* **285**, 14806–14814 (2010).
26. Sebald, W., Nickel, J., Zhang, J. & Mueller, T. D. Molecular recognition in bone morphogenetic protein (BMP)/receptor interaction. *Biol. Chem.* **385**, 697–710 (2004).
27. Allendorph, G. P., Vale, W. W. & Choe, S. Structure of the ternary signaling complex of a TGF-beta superfamily member. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7643–8 (2006).
28. Koenig, B. B. *et al.* Characterization and cloning of a receptor for BMP-2 and BMP-4 from NIH 3T3 cells. *Mol. Cell. Biol.* **14**, 5961–5974 (1994).
29. Knaus, P. & Sebald, W. Cooperativity of binding epitopes and receptor chains in the BMP/TGF $\beta$  superfamily. *Biol. Chem.* **382**, 1189–1195 (2001).
30. Gilboa, L., Wells, R. G., Lodish, H. F. & Henis, Y. I. Oligomeric Structure of Type I and Type II Transforming Growth Factor  $\beta$  Receptors: Homodimers Form in the ER and Persist at the Plasma Membrane. *J. Biol. Chem.* **140**, 767–777 (1998).
31. Wells, R. G. *et al.* Transforming Growth Factor- $\beta$  Induces Formation of a Dithiothreitol-resistant Type I / Type II Receptor Complex in Live Cells. *J. Biol. Chem.* **274**, 5716–5722 (1999).
32. Gilboa, L. *et al.* Bone Morphogenetic Protein Receptor Complexes on the Surface of Live Cells : A New Oligomerization Mode for Serine / Threonine Kinase Receptors. *Mol. Biol. Cell* **11**, 1023–1035 (2000).
33. Zhang, W. *et al.* Single-molecule imaging reveals transforming growth factor- $\beta$  -induced type II receptor dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15679–15683 (2009).
34. Zhang, W. *et al.* Monomeric type I and type III transforming growth factor-B receptors and their dimerization revealed by single-molecule imaging. *Cell Res.* **20**, 1216–1223 (2010).

35. Zhang, M. *et al.* Quantitative Characterization of the Membrane Dynamics of Newly Delivered TGF- $\beta$  Receptors by Single-Molecule Imaging. *Anal. Chem.* **90**, 4282–4287 (2018).
36. Olsen, O. E. *et al.* BMPR2 inhibits activin and BMP signaling via wild-type ALK2. *J. Cell Sci.* **131**, (2018).
37. Mueller, T. D. & Nickel, J. Promiscuity and specificity in BMP receptor activation. *FEBS Lett.* **586**, 1846–1859 (2012).
38. Kalter, V. G. & Brody, A. R. Receptors for Transforming Growth Factor- $\beta$  (TGF- $\beta$ ) on Rat Lung Fibroblasts Have Higher Affinity for TGF- $\beta$ 1 than for TGF- $\beta$ 2. *Am. J. Respir. Cell Mol. Biol.* **4**, 397–407 (1991).
39. Wolfman, N. M. *et al.* Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. *Proc. Natl. Acad. Sci.* **100**, 15842–15846 (2003).
40. Mason, A. J., Farnworth, P. G. & Sullivan, J. Characterization and determination of the biological activities of noncleavable high molecular weight forms of inhibin A and activin A. *Mol. Endocrinol.* **10**, 1055–1065 (1996).
41. Kelber, J. A., Shani, G., Booker, E. C., Vale, W. W. & Gray, P. C. Cripto is a noncompetitive activin antagonist that forms analogous signaling complexes with activin and nodal. *J. Biol. Chem.* **283**, 4490–4500 (2008).
42. Massagué, J., Seoane, J. & Wotton, D. Smad transcription factors. *Genes Dev.* **19**, 2783–2810 (2005).
43. Shi, Y. *et al.* Crystal Structure of a Smad MH1 Domain Bound to DNA: Insights on DNA Binding in TGF- $\beta$  Signaling. *Cell* **94**, 585–594 (1998).
44. Abdollah, S. *et al.* T $\beta$ RI Phosphorylation of Smad2 on Ser465 and Ser467 Is Required for Smad2-Smad4 Complex Formation and Signaling. *J. Biol. Chem.* **272**, 27678–27685 (1997).
45. Wu, J.-W. *et al.* Crystal Structure of a Phosphorylated Smad2: Recognition of Phosphoserine by the MH2 Domain and Insights on Smad Function in TGF- $\beta$  Signaling. *Mol. Cell* **8**, 1277–1289 (2001).

46. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L. & Wrana, J. L. SARA, a FYVE Domain Protein that Recruits Smad2 to the TGF $\beta$  Receptor. *Cell* **95**, 779–791 (1998).
47. Lo, R. S., Chen, Y.-G., Shi, Y., Pavletich, N. P. & Massagué, J. The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF- $\beta$  receptors. *EMBO J.* **17**, 996–1005 (1998).
48. Persson, U. *et al.* The L45 loop in type I receptors for TGF- $\beta$  family members is a critical determinant in specifying Smad isoform activation. *FEBS Lett.* **434**, 83–87 (1998).
49. Hanyu, A. *et al.* The N domain of Smad7 is essential for specific inhibition of transforming growth factor- $\beta$  signaling. *J. Cell Biol.* **155**, 1017–1028 (2001).
50. Ishisaki, A. *et al.* Differential Inhibition of Smad6 and Smad7 on Bone Morphogenetic Protein- and Activin-mediated Growth Arrest and Apoptosis in B Cells. *J. Biol. Chem.* **274**, 13637–13642 (1999).
51. Hata, a., Lagna, G., Massague, J. & Hemmati-Brivanlou, a. Smad6 inhibits BMP/Samd1 signaling by specifically competing with the Samd4 tumor suppressor. *Genes Dev* **12**, 186–197 (1998).
52. Kamiya, Y., Miyazono, K. & Miyazawa, K. Smad7 Inhibits Transforming Growth Factor- $\beta$  Family Type I Receptors through Two Distinct Modes of Interaction. *J. Biol. Chem.* **285**, 30804–30813 (2010).
53. Yan, X. *et al.* Human BAMBI Cooperates with Smad7 to Inhibit Transforming Growth Factor- $\beta$  Signaling. *J. Biol. Chem.* **284**, 30097–30104 (2009).
54. Nakao, A. *et al.* TGF- $\beta$  receptor-mediated signalling through Smad2, Smad3 and Smad4 includes activins and bone morphogenetic proteins Masahiro Kawabata 2 , Akira Ishisaki, differentiation, migration and apoptosis of many different Kohei Miyazono 2 and Peter ten Dijke. *EMBO J.* **16**, 5353–5362 (1997).
55. Nicolás, F. J., Bosscher, K. D., Schmierer, B. & Hill, C. S. Analysis of Smad nucleocytoplasmic shuttling in living cells. *J. Cell Sci.* **117**, 4113–4125 (2004).



56. Pierreux, C. E., Nicolás, F. J. & Hill, C. S. Transforming Growth Factor  $\beta$ -Independent Shuttling of Smad4 between the Cytoplasm and Nucleus. *Mol. Cell. Biol.* **20**, 9041–9054 (2000).
57. Li, Y., Luo, W. & Yang, W. Nuclear Transport and Accumulation of Smad Proteins Studied by Single-Molecule Microscopy. *Biophys. J.* **114**, 2243–2251 (2018).
58. Schmierer, B. & Hill, C. S. Kinetic Analysis of Smad Nucleocytoplasmic Shuttling Reveals a Mechanism for Transforming Growth Factor  $\beta$ -Dependent Nuclear Accumulation of Smads. *Mol. Cell. Biol.* **25**, 9845–9858 (2005).
59. Goto, D. *et al.* Interaction between Smad Anchor for Receptor Activation and Smad3 Is Not Essential for TGF- $\beta$ /Smad3-Mediated Signaling. *Biochem. Biophys. Res. Commun.* **281**, 1100–1105 (2001).
60. Bakkebo, M. *et al.* SARA is dispensable for functional TGF- $\beta$  signaling. *FEBS Lett.* **586**, 3367–3372 (2012).
61. Zhang, Y. E. Non-Smad signaling pathways of the TGF- $\beta$  family. *Cold Spring Harb. Perspect. Biol.* **9**, (2017).
62. Ozdamar, B. *et al.* Regulation of the polarity protein Par6 by TGF $\beta$  receptors controls epithelial cell plasticity. *Science* **307**, 1603–1609 (2005).
63. Galliher, A. J. & Schiemann, W. P. Src phosphorylates Tyr284 in TGF- $\beta$  type II receptor and regulates TGF- $\beta$  stimulation of p38 MAPK during breast cancer cell proliferation and invasion. *Cancer Res.* **67**, 3752–3758 (2007).
64. Lee, M. K. *et al.* TGF- $\beta$  activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J.* **26**, 3957–3967 (2007).
65. Zhang, L. *et al.* USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF- $\beta$  type I receptor. *Nat. Cell Biol.* **14**, 717–726 (2012).
66. Kretzschmar, M., Doody, J. & Massagué, J. Opposing BMP and EGF signalling pathways converge on the TGF- $\beta$  family mediator Smad1. *Nature* **389**, 618–622 (1997).

67. Hough, C., Radu, M. & Doré, J. J. E. TGF-beta induced Erk phosphorylation of smad linker region regulates smad signaling. *PLoS ONE* **7**, 26–28 (2012).
68. Zhang, L. *et al.* A transforming growth factor beta-induced Smad3/Smad4 complex directly activates protein kinase A. *Mol. Cell. Biol.* **24**, 2169–80 (2004).
69. Takekawa, M. *et al.* Smad-dependent GADD45 $\beta$  expression mediates delayed activation of p38 MAP kinase by TGF- $\beta$ . *EMBO J.* **21**, 6473–6482 (2002).
70. Zuzarte-Luís, V. *et al.* A new role for BMP5 during limb development acting through the synergic activation of Smad and MAPK pathways. *Dev. Biol.* **272**, 39–52 (2004).
71. Oft, M., Akhurst, R. J. & Balmain, A. Metastasis is driven by sequential elevation of H-ras and Smad2 levels. *Nat. Cell Biol.* **4**, 487–494 (2002).
72. Zavadil, J., Cermak, L., Soto-Nieves, N. & Böttinger, E. P. Integration of TGF- $\beta$ /Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition. *EMBO J.* **23**, 1155–1165 (2004).
73. Cornell, R. A. & Kimelman, D. Activin-mediated mesoderm induction requires FGF. *Trends Genet.* **10**, 150 (2003).
74. Vallier, L., Alexander, M. & Pederson, R. A. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J. Cell Sci.* **118**, 4495–4509 (2005).
75. Lebrun, J. J. & Vale, W. W. Activin and inhibin have antagonistic effects on ligand-dependent heteromerization of the type I and type II activin receptors and human erythroid differentiation. *Mol. Cell. Biol.* **17**, 1682–91 (1997).
76. Goebel, E. J. *et al.* Structural characterization of an activin class ternary receptor complex reveals a third paradigm for receptor specificity. *Proc. Natl. Acad. Sci.* **116**, 15505–15513 (2019).
77. Harrington, A. E. *et al.* Structural basis for the inhibition of activin signalling by follistatin. *EMBO J.* **25**, 1035–1045 (2006).

78. Inouye, S., Ling, N. & Shimasaki, S. Localization of the heparin binding site of follistatin. *Mol. Cell. Endocrinol.* **90**, 1–6 (1992).
79. Schneyer, A. L., Wang, Q., Sidis, Y. & Sluss, P. M. Differential distribution of follistatin isoforms: Application of a new FS315-specific immunoassay. *J. Clin. Endocrinol. Metab.* **89**, 5067–5075 (2004).
80. Sidis, Y. *et al.* Biological activity of follistatin isoforms and follistatin-like-3 is dependent on differential cell surface binding and specificity for activin, myostatin, and bone morphogenetic proteins. *Endocrinology* **147**, 3586–3597 (2006).
81. Sidis, Y. *et al.* Follistatin-Related Protein and Follistatin Differentially Neutralize Endogenous vs. Exogenous Activin. *Endocrinology* **143**, 1613–1624 (2002).
82. Stampler, R. *et al.* The Structure of FSTL3·Activin A Complex. *J. Biol. Chem.* **283**, 32831–32838 (2008).
83. Saito, S., Sidis, Y., Mukherjee, A., Xia, Y. & Schneyer, A. Differential biosynthesis and intracellular transport of follistatin isoforms and follistatin-like-3. *Endocrinology* **146**, 5052–5062 (2005).
84. Robertson, D. M. *et al.* Isolation of Inhibin from Bovine Follicular Fluid. *Biochem. Biophys. Res. Commun.* **126**, 220–226 (1985).
85. Rivier, J., Spiess, J., McClintock, R., Vaughan, J. & Vale, W. Purification and Partial Characterisation of Inhibin from Porcine Follicular Fluid. *Biochem. Biophys. Res. Commun.* **133**, 120–127 (1985).
86. Mottram, J. C. & Cramer, W. On the General Effects of Exposure to Radium on Metabolism and Tumour Growth in the Rat and the Special Effects on Testis and Pituitary. *Exp. Physiol.* **13**, 209–226 (1923).
87. Lewis, K. A. *et al.* Betaglycan binds inhibin and can mediate functional antagonism of activin signalling. *Nature* **404**, 411–414 (2000).

88. Gold, E. *et al.* Activin C antagonizes activin A in vitro and overexpression leads to pathologies in vivo. *Am. J. Pathol.* **174**, 184–195 (2009).
89. López-Casillas, F., Wrana, J. L. & Massagué, J. Betaglycan presents ligand to the TGF $\beta$  signaling receptor. *Cell* **73**, 1435–1444 (1993).
90. López-Casillas, F., Payne, H. M., Andres, J. L. & Massagué, J. Betaglycan can act as a dual modulator of TGF- $\beta$  access to signaling receptors: Mapping of ligand binding and GAG attachment sites. *J. Cell Biol.* **124**, 557–568 (1994).
91. Onichtchouk, D. *et al.* Silencing of TGF- $\beta$  signalling by the pseudoreceptor BAMBI. *Nature* **401**, 480–485 (1999).
92. Gray, P. C., Harrison, C. A. & Vale, W. Cripto forms a complex with activin and type II activin receptors and can block activin signaling. *Proc. Natl. Acad. Sci.* **100**, 5193–5198 (2003).
93. Zhao, B., Xu, S., Dong, X., Lu, C. & Springer, T. A. Prodomain–growth factor swapping in the structure of pro-TGF- $\beta$ 1. *J. Biol. Chem.* **293**, 1579–1589 (2018).
94. Dubois, C. M., Laprise, M.-H., Blanchette, F., Gentry, L. E. & Leduc, R. Processing of Transforming Growth Factor 1 Precursor by Human Furin Convertase. *J. Biol. Chem.* **270**, 10618–10624 (1995).
95. Li, S. *et al.* Activin A Binds to Perlecan through Its Pro-region That Has Heparin/Heparan Sulfate Binding Activity. *J. Biol. Chem.* **285**, 36645–36655 (2010).
96. Sengle, G., Ono, R. N., Sasaki, T. & Sakai, L. Y. Prodomains of Transforming Growth Factor  $\beta$  (TGF $\beta$ ) Superfamily Members Specify Different Functions EXTRACELLULAR MATRIX INTERACTIONS AND GROWTH FACTOR BIOAVAILABILITY. *J. Biol. Chem.* **286**, 5087–5099 (2011).
97. Wakefield, L. M. *et al.* Recombinant latent transforming growth factor beta 1 has a longer plasma half-life in rats than active transforming growth factor beta 1, and a different tissue distribution. *J. Clin. Invest.* **86**, 1976–1984 (1990).
98. Khalil, N. TGF- $\beta$ : from latent to active. *Microbes Infect.* **1**, 1255–1263 (1999).

99. Gentry, L. E. & Nash, B. W. The pro domain of pre-pro-transforming growth factor beta 1 when independently expressed is a functional binding protein for the mature growth factor. *Biochemistry* **29**, 6851–6857 (1990).
100. Ge, G., Hopkins, D. R., Ho, W. & Greenspan, D. S. GDF11 Forms a Bone Morphogenetic Protein 1-Activated Latent Complex That Can Modulate Nerve Growth Factor-Induced Differentiation of PC12 Cells. *Mol. Cell. Biol.* **25**, 5846–5858 (2005).
101. Brown, M. A. *et al.* Crystal Structure of BMP-9 and Functional Interactions with Pro-region and Receptors. *J. Biol. Chem.* **280**, 25111–25118 (2005).
102. Shi, M. *et al.* Latent TGF- $\beta$  structure and activation. *Nature* **474**, 343–349 (2011).
103. Greenwald, J. *et al.* A flexible activin explains the membrane-dependent cooperative assembly of TGF- $\beta$  family receptors. *Mol. Cell* **15**, 485–489 (2004).
104. Wang, X., Fischer, G. & Hyvönen, M. Structure and activation of pro-activin A. *Nat. Commun.* 12052 (2016).
105. Gray, A. M. & Mason, A. J. Requirement for activin A and transforming growth factor--beta 1 pro-regions in homodimer assembly. *Science* **247**, 1328–1330 (1990).
106. Pepinsky, B. *et al.* A Prodomain Fragment from the Proteolytic Activation of Growth Differentiation Factor 11 Remains Associated with the Mature Growth Factor and Keeps It Soluble. *Biochemistry* **56**, 4405–4418 (2017).
107. Mi, L.-Z. *et al.* Structure of bone morphogenetic protein 9 procomplex. *Proc. Natl. Acad. Sci.* **112**, 3710–3715 (2015).
108. Cotton, T. R. *et al.* Structure of the human myostatin precursor and determinants of growth factor latency. *EMBO J.* **37**, 367–383 (2018).
109. Matzuk, M. M. *et al.* Functional analysis of activins during mammalian development. *Lett. Nat.* **374**, 354–356 (1995).

110. James, D., Levine, A. J., Besser, D. & Hemmati-brivanlou, A. TGF  $\beta$  / activin / nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **132**, 1273–1282 (2005).
111. Lu, R. Z., Matsuyama, S., Nishihara, M. & Takahashi, M. Developmental Expression of Activin / Inhibin  $\beta$ A ,  $\beta$ B , and  $\alpha$  Subunits , and Activin Receptor-IIB Genes in Preimplantation Mouse Embryos. *Biol. Reprod.* **49**, 1163–1169 (1993).
112. Da Silva, S. J. M. *et al.* Expression of activin subunits and receptors in the developing human ovary: Activin A promotes germ cell survival and proliferation before primordial follicle formation. *Dev. Biol.* **266**, 334–345 (2004).
113. Mendis, S. H. S., Meachem, S. J., Sarraj, M. A. & Loveland, K. L. Activin A Balances Sertoli and Germ Cell Proliferation in the Fetal Mouse Testis. *Biol. Reprod.* **84**, 379–391 (2010).
114. Roberts, V. J. & Barth, S. L. Expression of messenger ribonucleic acids encoding the inhibin/activin system during mid- and late-gestation rat embryogenesis. *Endocrinology* **134**, 914–923 (1994).
115. Tuuri, T., Erämaa, M., Hildén, K. & Ritvos, O. The tissue distribution of activin beta A- and beta B-subunit and follistatin messenger ribonucleic acids suggests multiple sites of action for the activin-follistatin system during human development. *J. Clin. Endocrinol. Metab.* **78**, 1521–1524 (1994).
116. Papanayotou, C. & Collignon, J. Activin/Nodal signalling before implantation: setting the stage for embryo patterning. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, (2014).
117. Iannaccone, P. M., Zhou, X., Khokha, M., Boucher, D. & Kuehn, M. R. Insertional mutation of a gene involved in growth regulation of the early mouse embryo. *Dev. Dyn.* **194**, 198–208 (1992).
118. Lowe, L. A., Yamada, S. & Kuehn, M. R. Genetic dissection of nodal function in patterning the mouse embryo. *Dev. Camb. Engl.* **128**, 1831–1843 (2001).
119. Vallier, L. *et al.* Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* **136**, 1339–1349 (2009).

120. Thomsen, G. *et al.* Activins Are Expressed Early in *Xenopus* Embryogenesis and Can Induce Axial Mesoderm and Anterior Structures. *Cell* **63**, 485–493 (1990).
121. Jones, C. M., Kuehn, M. R., Hogan, B. L., Smith, J. C. & Wright, C. V. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651–3662 (1995).
122. Lanza, R. *et al.* *Essentials of Stem Cell Biology*. (Academic Press, 2009).
123. Tomizawa, M. *et al.* Activin A maintains pluripotency markers and proliferative potential of human induced pluripotent stem cells. *Exp. Ther. Med.* **2**, 405–408 (2011).
124. Beattie, G. M. *et al.* Activin A Maintains Pluripotency of Human Embryonic Stem Cells in the Absence of Feeder Layers. *Stem Cells* **23**, 489–495 (2005).
125. Xu, R. H. *et al.* NANOG Is a Direct Target of TGF $\beta$ /Activin-Mediated SMAD Signaling in Human ESCs. *Cell Stem Cell* **3**, 196–206 (2008).
126. Xu, R.-H. *et al.* BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* **20**, 1261–1264 (2002).
127. Saka, Y. & Smith, J. C. A mechanism for the sharp transition of morphogen gradient interpretation in *Xenopus*. *BMC Dev. Biol.* **7**, 1–9 (2007).
128. Sulzbacher, S., Schroeder, I. S., Truong, T. T. & Wobus, A. M. Activin a-induced differentiation of embryonic stem cells into endoderm and pancreatic progenitors-the influence of differentiation factors and culture conditions. *Stem Cell Rev. Rep.* **5**, 159–173 (2009).
129. Touboul, T. *et al.* Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology* **51**, 1754–1765 (2010).
130. Sa, S. & McCloskey, K. E. Activin A and BMP4 signaling for efficient cardiac differentiation of H7 and H9 human embryonic stem cells. *J. Stem Cells Regen. Med.* **8**, 198–202 (2012).

131. Jones, K. L., de Kretser, D. M., Clarke, I. J., Scheerlinck, J. P. Y. & Phillips, D. J. Characterisation of the rapid release of activin A following acute lipopolysaccharide challenge in the ewe. *J. Endocrinol.* **182**, 69–80 (2004).
132. Yamashita, N. *et al.* Effects of activin A on IgE synthesis and cytokine production by human peripheral mononuclear cells. *Clin. Exp. Immunol.* **94**, 214–9 (1993).
133. Nüsing, R. M. & Barsig, J. Induction of prostanoid, nitric oxide, and cytokine formation in rat bone marrow derived macrophages by activin A. *Br. J. Pharmacol.* **127**, 919–926 (1999).
134. Sugama, S., Takenouchi, T., Kitani, H., Fujita, M. & Hashimoto, M. Activin as an anti-inflammatory cytokine produced by microglia. *J. Neuroimmunol.* **192**, 31–39 (2007).
135. Kogure, K., Zhang, Y. Q., Shibata, H. & Kojima, I. Immediate onset of DNA synthesis in remnant rat liver after 90% hepatectomy by an administration of follistatin. *J. Hepatol.* **29**, 977–984 (1998).
136. Sugiyama, M. *et al.* Expression of Activin A is increased in cirrhotic and fibrotic rat livers. *Gastroenterology* **114**, 550–558 (1998).
137. Schwall, R. H. *et al.* Activin induces cell death in hepatocytes in vivo and in vitro. *Hepatology* **18**, 347–356 (1993).
138. Hubner, G., Brauchle, M., Gregor, M. & Werner, S. Activin A: a novel player and inflammatory marker in inflammatory bowel disease? *Lab. Investig. J. Tech. Methods Pathol.* **77**, 311–318 (1997).
139. Dignass, A. U. Functional Relevance of Activin A in the Intestinal Epithelium. (2002).
140. Werner, S. & Alzheimer, C. Roles of activin in tissue repair , fibrosis , and inflammatory disease. *Cytokine Growth Factor Rev.* **17**, 157–171 (2006).
141. Munz, B. *et al.* Overexpression of activin A in the skin of transgenic mice reveals new activities of activin in epidermal morphogenesis, dermal fibrosis and wound repair. *EMBO J.* **18**, 5205–5215 (1999).



142. Wankell, M. *et al.* Impaired wound healing in transgenic mice overexpressing the activin antagonist follistatin in the epidermis. *EMBO J.* **20**, 5361–5372 (2001).
143. Kriegstein, K., Suter-crazzolaro, C., Fischerl, W. H. & Unsicker, K. TGF- $\beta$  superfamily members promote survival of midbrain dopaminergic neurons and protect them against MPP + toxicity TGF-132. *EMBO J.* **14**, 736–742 (1995).
144. Smith, C. *et al.* Potential anti-inflammatory role of activin A in acute coronary syndromes. *J. Am. Coll. Cardiol.* **44**, 369–375 (2004).
145. Hedger, M. P., Phillips, D. J. & De Kretser, D. M. Divergent cell-specific effects of activin-A on thymocyte proliferation stimulated by phytohemagglutinin, and interleukin 1 $\beta$  or interleukin 6 in vitro. *Cytokine* **12**, 595–602 (2000).
146. Brosh, N. *et al.* The Plasmacytoma Growth Inhibitor Restrictin-P Is an Antagonist of Interleukin 6 and Interleukin 11. *J. Biol. Chem.* **270**, 29594–29600 (2002).
147. Chen, J. L. *et al.* Specific targeting of TGF-  $\beta$  family ligands demonstrates distinct roles in the regulation of muscle mass in health and disease. *Proc. Natl. Acad. Sci.* **114**, (2017).
148. Zheng, H. *et al.* Follistatin N terminus differentially regulates muscle size and fat in vivo. *Exp. Mol. Med.* **49**, (2017).
149. Chang, F. *et al.* The transgenic expression of human follistatin-344 increases skeletal muscle mass in pigs. *Transgenic Res.* **26**, 25–36 (2017).
150. Schuelke, M. *et al.* Myostatin Mutation Associated with Gross Muscle Hypertrophy in a Child. *N. Engl. J. Med.* **350**, 2682–2688 (2004).
151. Fiems, L. O. Double Muscling in Cattle: Genes, Husbandry, Carcasses and Meat. 472–506 (2012) doi:10.3390/ani2030472.
152. Chen, J. L. *et al.* Elevated expression of activins promotes muscle wasting and cachexia. *FASEB J.* **28**, 1711–1723 (2014).
153. Loumaye, A. *et al.* Role of activin A and myostatin in human cancer cachexia. *J. Clin. Endocrinol. Metab.* **100**, 2030–2038 (2015).

154. Lee, S.-J. *et al.* Regulation of Muscle Mass by Follistatin and Activins. *Mol. Endocrinol.* **24**, 1998–2008 (2010).
155. Gilson, H. *et al.* Follistatin induces muscle hypertrophy through satellite cell proliferation and inhibition of both myostatin and activin. *Am. J. Physiol.-Endocrinol. Metab.* **297**, E157–E164 (2009).
156. Campbell, C. *et al.* Myostatin inhibitor ACE-031 treatment of ambulatory boys with Duchenne muscular dystrophy: Results of a randomized, placebo-controlled clinical trial. *Muscle Nerve* **55**, 458–464 (2017).
157. Bloise, E. *et al.* Activin A in Mammalian Physiology. *Physiol. Rev.* **99**, 739–780 (2018).
158. Spencer, S. J., Mesiano, S., Lee, J. Y. & Jaffe, R. B. Proliferation and apoptosis in the human adrenal cortex during the fetal and perinatal periods: Implications for growth and remodeling. *J. Clin. Endocrinol. Metab.* **84**, 1110–1115 (1999).
159. Bashir, M., Damineni, S., Mukherjee, G. & Kondaiah, P. Activin-a signaling promotes epithelial–mesenchymal transition, invasion, and metastatic growth of breast cancer. *Npj Breast Cancer* **1**, 1–13 (2015).
160. Dalkin, A. C., Gilrain, J. T., Bradshaw, D. & Myers, C. E. Activin Inhibition of Prostate Cancer Cell Growth: Selective Actions on Androgen-Responsive LNCaP Cells. *Endocrinology* **137**, 5230–5235 (1996).
161. Do, T. V. *et al.* The role of activin A and Akt/GSK signaling in ovarian tumor biology. *Endocrinology* **149**, 3809–3816 (2008).
162. Marini, K. D. *et al.* Inhibition of activin signaling in lung adenocarcinoma increases the therapeutic index of platinum chemotherapy. *Sci. Transl. Med.* **10**, (2018).
163. Wildi, S. *et al.* Overexpression of activin A in stage IV colorectal cancer. *Gut* **49**, 409–417 (2001).
164. Jeruss, J. S., Sturgis, C. D., Rademaker, A. W. & Woodruff, T. K. Down-regulation of activin, activin receptors, and Smads in high-grade breast cancer. *Cancer Res.* **63**, 3783–3790 (2003).

165. Burdette, J. E., Jeruss, J. S., Kurley, S. J., Lee, E. J. & Woodruff, T. K. Activin A mediates growth inhibition and cell cycle arrest through Smads in human breast cancer cells. *Cancer Res.* **65**, 7968–7975 (2005).
166. Wang, Q., Tabatabaei, S., Planz, B., Lin, C. & Sluss, P. M. Identification of an activin-follistatin growth modulatory system in the human prostate. *J. Urol.* **161**, 1378–1384 (2003).
167. Kang, H. Y. *et al.* Activin A enhances prostate cancer cell migration through activation of androgen receptor and is overexpressed in metastatic prostate cancer. *J. Bone Miner. Res.* **24**, 1180–1193 (2009).
168. Hofland, J. *et al.* Activin a stimulates AKR1C3 expression and growth in human prostate cancer. *Endocrinology* **153**, 5726–5734 (2012).
169. Burdette, J. E. & Woodruff, T. K. Activin and estrogen crosstalk regulates transcription in human breast cancer cells. *Endocr. Relat. Cancer* **14**, 679–689 (2007).
170. Kalkhoven, E. *et al.* Resistance to transforming growth factor and activin due to reduced receptor expression in human breast tumor cell lines. *Cell Growth Differ. Mol. Biol. J. Am. Assoc. Cancer Res.* **6**, 1151 (1995).
171. Vaarala, M. H., Porvari, K., Kyllönen, A. & Vihko, P. Differentially Expressed Genes in Two LNCaP Prostate Cancer Cell Lines Reflecting Changes during Prostate Cancer Progression. *Lab. Invest.* **80**, 1259–1268 (2000).
172. Tournier, I. *et al.* Germline mutations of inhibins in early-onset ovarian epithelial tumors. *Hum. Mutat.* **35**, 294–297 (2014).
173. Shore, E. M. *et al.* A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and sporadic fibrodysplasia ossificans progressiva. *Nat. Genet.* **38**, 525–527 (2006).
174. Kaplan, F. S. *et al.* Classic and atypical fibrodysplasia ossificans progressiva (FOP) phenotypes are caused by mutations in the bone morphogenetic protein (BMP) type I receptor ACVR1. *Hum. Mutat.* **30**, 379–390 (2009).

175. Alessi Wolken, D. M., Idone, V., Hatsell, S. J., Yu, P. B. & Economides, A. N. The obligatory role of Activin A in the formation of heterotopic bone in Fibrodysplasia Ossificans Progressiva. *Bone* **109**, 210–217 (2018).
176. Taylor, K. R., Vinci, M., Bullock, A. N. & Jones, C. ACVR1 mutations in DIPG: lessons learned from FOP. *Cancer Res.* **74**, 4565–4570 (2014).
177. Connor, J. M. & Evans, D. A. Fibrodysplasia ossificans progressiva. The clinical features and natural history of 34 patients. *J. Bone Joint Surg. Br.* **64**, 76–83 (1982).
178. Botello-Smith, W. M. *et al.* Polymodal allosteric regulation of Type 1 Serine/Threonine Kinase Receptors via a conserved electrostatic lock. *PLoS Comput. Biol.* **13**, e1005711 (2017).
179. Fukuda, T. *et al.* Constitutively Activated ALK2 and Increased SMAD1/5 Cooperatively Induce Bone Morphogenetic Protein Signaling in Fibrodysplasia Ossificans Progressiva. *J. Biol. Chem.* **284**, 7149–7156 (2009).
180. Bagarova, J. *et al.* Constitutively active ALK2 receptor mutants require type II receptor cooperation. *Mol. Cell. Biol.* **33**, 2413–2424 (2013).
181. Huse, M. *et al.* The TGF $\beta$  Receptor Activation Process. *Mol. Cell* **8**, 671–682 (2004).
182. Gromada, J., Latres, E., Murphy, A., Yancopoulos, G. & Morton, L. Anti-Activin A Antibodies and Uses Thereof. (2017).
183. Morvan, F. *et al.* Blockade of activin type II receptors with a dual anti-ActRIIA / IIB antibody is critical to promote maximal skeletal muscle hypertrophy. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12448–12453 (2017).
184. Pearsall, R. S. *et al.* A soluble activin type IIA receptor induces bone formation and improves skeletal integrity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7082–7087 (2008).
185. Sako, D. *et al.* Characterization of the Ligand Binding Functionality of the Extracellular Domain of Activin Receptor Type IIB. *J. Biol. Chem.* **285**, 21037–21048 (2010).
186. Castonguay, R. *et al.* Follistatin-288-Fc Fusion Protein Promotes Localized Growth of Skeletal Muscle. *J. Pharmacol. Exp. Ther.* **368**, 435–445 (2019).

187. Chen, J. L. *et al.* Development of Novel Activin-Targeted Therapeutics. *Mol. Ther.* **23**, 434–444 (2015).
188. Makanji, Y. *et al.* Generation of a Specific Activin Antagonist by Modification of the Activin A Propeptide. *Endocrinology* **152**, 3758–3768 (2011).
189. Takayama, K. *et al.* Identification of the minimum peptide from mouse myostatin prodomain for human myostatin inhibition. *J. Med. Chem.* **58**, 1544–1549 (2015).
190. Rentier, C. *et al.* Design and synthesis of potent myostatin inhibitory cyclic peptides. *Bioorg. Med. Chem.* **27**, 1437–1443 (2019).
191. Walton, K. L., Kelly, E. K., Chan, K. L., Harrison, C. A. & Robertson, D. M. Inhibin Biosynthesis and Activity Are Limited by a Prodomain-Derived Peptide. *Endocrinology* **156**, 3047–3057 (2015).
192. Burmester, J. K., Salzman, S. A., Zhang, K. Q. & Dart, R. A. Small molecule antagonists of the TGF- $\beta$ 1/TGF- $\beta$  receptor binding interaction. *Med. Oncol.* **23**, 553–562 (2006).
193. Zhu, J. *et al.* Virtual High-Throughput Screening to Identify Novel Activin Antagonists. *J. Med. Chem.* **58**, 5637–5648 (2015).
194. Yingling, J. M. *et al.* Preclinical assessment of galunisertib (LY2157299 monohydrate), a first-in-class transforming growth factor- $\beta$  receptor type I inhibitor. in *Oncotarget* (2017). doi:10.18632/oncotarget.23795.
195. Mohedas, A. H. *et al.* Structure–Activity Relationship of 3,5-Diaryl-2-aminopyridine ALK2 Inhibitors Reveals Unaltered Binding Affinity for Fibrodysplasia Ossificans Progressiva Causing Mutants. *J. Med. Chem.* **57**, 7900–7915 (2014).
196. Sakamoto, K. *et al.* Identification of ligand-selective peptidic ActRIIB-antagonists using phage display technology. *Biochem. Biophys. Rep.* **11**, 33–39 (2017).
197. Lebrun, J., Takabe, K., Chen, Y. & Vale, W. Roles of Pathway-Specific and Inhibitory Smads in Activin Receptor Signaling. *Mol. Endocrinol.* **13**, 15–23 (1999).

198. Funaba, M. *et al.* Unique Inhibin Recognition Subunits of Activin and Inhibin by Polyclonal Antibodies to Inhibin Subunits. *J. Biochem. (Tokyo)* **119**, 953–960 (1996).
199. Murata, T. *et al.* Anti-activin A Antibody (IgY) Specifically Neutralizes Various Activin A Activities. *Proc. Soc. Exp. Biol. Med.* **211**, 100–107 (1996).
200. Han, H. Q., Chen, Q., Kwak, K. S. N. & Zhou, X. Anti-Activin A Antibodies and Uses Thereof. (2012).
201. Komrokji, R. *et al.* Sotatercept with long-term extension for the treatment of anaemia in patients with lower-risk myelodysplastic syndromes: a phase 2, dose-ranging trial. *Lancet Haematol.* **5**, e63–e72 (2018).
202. Piga, A. *et al.* Luspatercept (ACE-536) Reduces Disease Burden, Including Anemia, Iron Overload, and Leg Ulcers, in Adults with Beta-Thalassemia: Results from a Phase 2 Study. *Blood* **126**, 752–752 (2015).
203. Shy, M. *et al.* CMT AND NEUROGENIC DISEASE: P.339Preliminary phase 2 results for ACE-083, local muscle therapeutic, in patients with CMT1 and CMTX. *Neuromuscul. Disord.* **28**, S132–S133 (2018).
204. Statland, J. *et al.* FSHD / OPMD / EDMD / DMI: P.365Results for a dose-escalation phase 2 study to evaluate ACE-083, a local muscle therapeutic, in patients with facioscapulohumeral muscular dystrophy. *Neuromuscul. Disord.* **28**, S140 (2018).
205. Ohsawa, Y. *et al.* The inhibitory core of the myostatin prodomain: Its interaction with both type I and II membrane receptors, and potential to treat muscle atrophy. *PLoS ONE* **10**, 1–18 (2015).
206. Jiang, M. S. *et al.* Characterization and identification of the inhibitory domain of GDF-8 propeptide. *Biochem. Biophys. Res. Commun.* **315**, 525–531 (2004).
207. Asari, T. *et al.* Structural basis for the effective myostatin inhibition of the mouse myostatin prodomain-derived minimum peptide. *ACS Med. Chem. Lett.* **8**, 113–117 (2017).
208. Takayama, K. *et al.* Development of Potent Myostatin Inhibitory Peptides through Hydrophobic Residue-Directed Structural Modification. *ACS Med. Chem. Lett.* **8**, 751–756 (2017).

209. Takayama, K. *et al.* Effect of N-Terminal Acylation on the Activity of Myostatin Inhibitory Peptides. *ChemMedChem* **11**, 845–849 (2016).
210. Inman, G. J. *et al.* SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **62**, 65–74 (2002).
211. Cho, H. *et al.* Identification of the First Selective Activin Receptor-Like Kinase 1 Inhibitor, a Reversible Version of L-783277. *J. Med. Chem.* **60**, 1495–1508 (2017).
212. Ehata, S. *et al.* Ki26894, a novel transforming growth factor- $\beta$  type I receptor kinase inhibitor, inhibits in vitro invasion and in vivo bone metastasis of a human breast cancer cell line. *Cancer Sci.* **98**, 127–133 (2007).
213. Valcarcel, D. *et al.* Phase 2 Study of Monotherapy Galunisertib (LY2157299 Monohydrate) in Very Low-, Low-, and Intermediate-Risk Patients with Myelodysplastic Syndromes. *Blood* **126**, 1669–1669 (2015).
214. Carvalho, D. *et al.* ALK2 inhibitors display beneficial effects in preclinical models of ACVR1 mutant diffuse intrinsic pontine glioma. *Commun. Biol.* **2**, 1–10 (2019).
215. de Gramont, A., Faivre, S. & Raymond, E. Novel TGF- $\beta$  inhibitors ready for prime time in onco-immunology. *Oncoimmunology* **6**, (2016).
216. Harrison, C. A. *et al.* An Activin Mutant with Disrupted ALK4 Binding Blocks Signaling via Type II Receptors. **279**, 28036–28044 (2004).
217. Liebert, M. A. *et al.* Intracerebral and Intrathecal Infusion of the TGF- $\beta$ 2 Specific Antisense Phosphorothioate Oligonucleotide AP 12009 in Rabbits and Primates: Toxicology and Safety. *Oligonucleotides* **15**, 94–104 (2005).
218. Chen, M. *et al.* Pretranscriptional Regulation of TGF- $\beta$ 1 by PI Polyamide Prevents Scarring and Accelerates Wound Healing of the Cornea After Exposure to Alkali. *Mol. Ther.* **18**, 519–527 (2009).

219. Chen, H. Y. *et al.* The Protective Role of Smad7 in Diabetic Kidney Disease: Mechanism and Therapeutic Potential. *Diabetes* **60**, 590–601 (2011).
220. Cappellini, M. D. *et al.* Sotatercept, a novel transforming growth factor  $\beta$  ligand trap, improves anemia in  $\beta$ -thalassemia: A phase II, open-label, dose-finding study. *Haematologica* **104**, 477–484 (2019).
221. Tao, J. J. *et al.* First-in-Human Phase I Study of the Activin A Inhibitor, STM 434, in Patients with Granulosa Cell Ovarian Cancer and Other Advanced Solid Tumors. *Clin. Cancer Res.* clincanres.1065.2019 (2019) doi:10.1158/1078-0432.CCR-19-1065.
222. Bayer Pharmaceuticals. Pharmaceuticals | Bayer - Small and large molecules. *Pharmaceuticals | Bayer - Small and large molecules* <http://pharma.bayer.com/en/innovation-partnering/technologies-and-trends/small-and-large-molecules> (2019).
223. Ngo, H. X. & Garneau-Tsodikova, S. What are the drugs of the future? *MedChemComm* **9**, 757–758 (2018).
224. Lamoree, B. & Hubbard, R. E. Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem.* **61**, 453–464 (2017).
225. Davies, T. G. & Tickle, I. J. Fragment Screening Using X-Ray Crystallography. in *Fragment-Based Drug Discovery and X-Ray Crystallography* (eds. Davies, T. G. & Hyvönen, M.) 33–59 (Springer Berlin Heidelberg, 2012). doi:10.1007/128\_2011\_179.
226. Pearce, N. M. *et al.* A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
227. Resnick, E. *et al.* Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).
228. Harding, R. J. *et al.* Small Molecule Antagonists of the Interaction between the Histone Deacetylase 6 Zinc-Finger Domain and Ubiquitin. *J. Med. Chem.* **60**, 9090–9096 (2017).
229. Fischer, G., Rossmann, M. & Hyvönen, M. Alternative modulation of protein-protein interactions by small molecules. *Curr. Opin. Biotechnol.* **35**, 78–85 (2015).



230. Radoux, C. J., Olsson, T. S. G., Pitt, W. R., Groom, C. R. & Blundell, T. L. Identifying Interactions that Determine Fragment Binding at Protein Hotspots. *J. Med. Chem.* **59**, 4314–4325 (2016).
231. Mateu, N. *et al.* Synthesis of Structurally Diverse N-Substituted Quaternary-Carbon-Containing Small Molecules from  $\alpha,\alpha$ -Disubstituted Propargyl Amino Esters. *Chem. – Eur. J.* **24**, 13681–13687 (2018).
232. Cox, O. B. *et al.* A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330 (2016).
233. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
234. Yang, N. J. & Hinner, M. J. Getting Across the Cell Membrane: An Overview for Small Molecules, Peptides, and Proteins. *Methods Mol. Biol. Clifton NJ* **1266**, 29–53 (2015).
235. Wójcik, P. & Berlicki, Ł. Peptide-based inhibitors of protein–protein interactions. *Bioorg. Med. Chem. Lett.* **26**, 707–713 (2016).
236. Chen, J. L. *et al.* Development of novel activin-targeted therapeutics. *Mol. Ther. J. Am. Soc. Gene Ther.* **23**, 434–444 (2015).
237. Schmidt, T. G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2**, 1528–1535 (2007).
238. Walton, K. L., Makanji, Y. & Harrison, C. A. New insights into the mechanisms of activin action and inhibition. *Mol. Cell. Endocrinol.* **359**, 2–12 (2012).
239. Sing, Y. Modelling of stapled peptides binding to mature activin A via Molecular Dynamics (Unpublished work). (2017).
240. Walton, K. L. *et al.* A common biosynthetic pathway governs the dimerization and secretion of inhibin and related transforming growth factor beta (TGFbeta) ligands. *J. Biol. Chem.* **284**, 9311–9320 (2009).
241. Newlon, M. G. *et al.* The molecular basis for protein kinase A anchoring revealed by solution NMR. *Nat. Struct. Biol.* **6**, 222 (1999).

242. Newlon, M. G. *et al.* A novel mechanism of PKA anchoring revealed by solution structures of anchoring complexes. *EMBO J.* **20**, 1651–1662 (2001).
243. Morikis, D., Roy, M., Newlon, M. G., Scott, J. D. & Jennings, P. A. Electrostatic properties of the structure of the docking and dimerization domain of protein kinase A II $\alpha$ . *Eur. J. Biochem.* **269**, 2040–2051 (2002).
244. Kobayakawa, T. *et al.* Dimeric C34 Derivatives Linked through Disulfide Bridges as New HIV-1 Fusion Inhibitors. *ChemBioChem*  
<https://onlinelibrary.wiley.com/doi/abs/10.1002/cbic.201900187> (2019)  
doi:10.1002/cbic.201900187.
245. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4 : a protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).
246. Yang, Y. *et al.* SPIDER2: A Package to Predict Secondary Structure, Accessible Surface Area, and Main-Chain Torsional Angles by Deep Neural Networks. *Methods Mol. Biol. Clifton NJ* **1484**, 55–63 (2017).
247. Lu, H. S. *et al.* Folding and oxidation of recombinant human granulocyte colony stimulating factor produced in *Escherichia coli*. Characterization of the disulfide-reduced intermediates and cysteine----serine analogs. *J. Biol. Chem.* **267**, 8770–8777 (1992).
248. Bernard, D. J., Lee, K. B. & Santos, M. M. Activin B can signal through both ALK4 and ALK7 in gonadotrope cells. *Reprod. Biol. Endocrinol. RBE* **4**, 52 (2006).
249. Rebbapragada, A., Benchabane, H., Wrana, J. L., Celeste, A. J. & Attisano, L. Myostatin Signals through a Transforming Growth Factor  $\beta$ -Like Signaling Pathway To Block Adipogenesis. *Mol. Cell. Biol.* **23**, 7230–7242 (2003).
250. Mason, A. J., Berkemeier, L. M., Schmelzer, C. H. & Schwall, R. H. Activin B: Precursor Sequences, Genomic Structure and *in Vitro* Activities. *Mol. Endocrinol.* **3**, 1352–1358 (1989).
251. Bai, L. *et al.* SMAD1/5 mediates bone morphogenetic protein 2-induced up-regulation of BAMBI expression in human granulosa-lutein cells. *Cell. Signal.* **37**, 52–61 (2017).

252. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
253. Hatsell, S. J. *et al.* ACVR1R206H receptor mutation causes fibrodysplasia ossificans progressiva by imparting responsiveness to activin A. *Sci. Transl. Med.* **7**, 303ra137-303ra137 (2015).
254. Thompson, T. B., Woodruff, T. K. & Jardetzky, T. S. Structures of an ActRIIB:activin A complex reveal a novel binding mode for TGF-beta ligand:receptor interactions. *EMBO J.* **22**, 1555–66 (2003).
255. Massagué, J. TGFβ in Cancer. *Cell* **134**, 215–230 (2008).
256. Gannon, F. H. *et al.* Bone morphogenetic protein 2/4 in early fibromatous lesions of fibrodysplasia ossificans progressiva. *Hum. Pathol.* **28**, 339–343 (1997).
257. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).
258. Delagenière, S. *et al.* ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics* **27**, 3186–3192 (2011).
259. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2006).
260. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
261. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
262. García-Nafría, J., Watson, J. F. & Greger, I. H. IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. *Sci. Rep.* **6**, (2016).
263. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExpASY Server. in *The Proteomics Protocols Handbook* (ed. Walker, J. M.) 571–607 (Humana Press, 2005).  
doi:10.1385/1-59259-890-0:571.

264. H.J. Motulsky. 'Equation: Association then dissociation', GraphPad Curve Fitting Guide.

[https://www.graphpad.com/guides/prism/8/curve-fitting/reg\\_equaton\\_association\\_then\\_disso.html](https://www.graphpad.com/guides/prism/8/curve-fitting/reg_equaton_association_then_disso.html).

---

# Appendices

---

## 8.1 Appendix 1: Crystallographic Statistics from Data Collection and Refinement

The statistics generated from the crystallographic data collection and structural refinement of the DMSO, PEGless, and NM466 and TR17 soaked structures reported in this thesis (Table 18). All beamlines used were at the Diamond Light Source, Harwell, UK. Values in parentheses denote the high-resolution shell.

**Table 18: Data collection and refinement statistics for all structures reported in this thesis (excluding those obtained from XChem).**

Data Processing*	DMSO	No PEG	NM466 Soak	TR17 Soak
Beamline	i03	i04-1	i24	i24
Wavelength / Å	0.9763	0.9159	0.9786	0.9786
Space group	I 2 2 2	I 2 2 2	I 2 2 2	I 2 2 2
Unit cell (a, b, c / Å, $\alpha$ , $\beta$ , $\gamma$ / °)	63.43 96.41 118.11 / 90 90 90	63.21 95.35 117.98 / 90 90 90	64.25 96.95 120.27 / 90 90 90	64.05 96.94 117.24 / 90 90 90
Resolution limits / Å	59.05-2.12 (2.14-2.12)	55.72-2.14 (2.18-2.14)	56.67-2.03 (2.07-2.03)	74.71-2.04 (2.07-2.04)
Number of total / unique reflections	135187 (20803)	160168 (18416)	214143 (24727)	199275 (23146)
Multiplicity	6.5 (6.8)	8.7 (8.6)	8.7 (9.0)	8.6 (8.9)
Anomalous Multiplicity	3.4 (3.5)	4.6 (4.5)	4.6 (4.7)	4.6 (4.6)
$R_{\text{meas}}$	0.071 (0.973)	0.166 (4.699)	0.194 (3.698)	0.206 (3.655)
$I/\sigma I$	14.7 (2.03)	10.2 (0.6)	8.3 (0.8)	8.7 (0.8)
$CC_{1/2}$	1.0 (0.9)	1.0 (0.5)	1.0 (0.4)	1.0 (0.4)
Completeness / %	99.8 (99.5)	92.2 (99.9)	99.6 (100.0)	97.2 (100.0)
Anomalous Completeness / %	98.8 (99.1)	91.8 (99.9)	99.3 (99.9)	96.6 (99.9)
Refinement				
$R_{\text{factor}}$	0.2091	0.2268	0.2240	0.2233
$R_{\text{factor}}$	0.2410	0.2797	0.2640	0.2540
R.m.s.d. bonds	0.011	0.007	0.018	0.018
R.m.s.d. angles	1.660	1.571	1.961	2.165

## 8.2 Appendix 2: Expression Construct Protein Sequences

Here are listed the protein sequences of all peptide fusions described and used in this thesis. For clarity, His-tags are highlighted in cyan, GB1 domains in pink, activin derived peptides in green, PKADD domains in teal, and Strep-tag IIs in grey. Cysteine residues are highlighted in yellow. The extinction absorption coefficients ( $\epsilon$ ) at 280 nm are reported below.<sup>263</sup>

### GB1

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSMEFAAALGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub> $\beta$

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSQPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLHVGKVGGENGYVEIEGSSWSHPQFEK

$$\epsilon_{280} = 16960 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSQPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLHGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-LH<sub>2</sub> $\beta$

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSKKRPDVTQVPVKAALLNAIRKLHVGKVGGENGYVEIEGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>2</sub> $\beta$

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSQVPVKAALLNAIRKLHVGKVGGENGYVEIEGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>L

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSQPEMVEAVKKHILNMLHLKKRPDVTQGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSQPEMVEAVKKHILNMLHGSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-P52-Q76

MNGSHHHHHHHHTSTYKLIILNGKTLKGETTTEAVDDATAEKVFKQYANDNGVDGEWYDDATKFTVTETGSGTSGSTLEVL  
FQGPGSPEMVEAVKKHILNMLHLKKRPDVTQGSSWSHPQFEK



$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-E53-Q76

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS EMVEAVKKHILNMLHLKKRPDVTQ GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-M54-Q76

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS MVEAVKKHILNMLHLKKRPDVTQ GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-V55-Q76

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS VEA VKKHILNMLHLKKRPDVTQ GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-E56-Q76

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS EAVKKHILNMLHLKKRPDVTQ GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-Q51-T75

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS QPEMVEAVKKHILNMLHLKKRPDVT GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-Q51-V74

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS QPEMVEAVKKHILNMLHLKKRPDV GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-Q51-D73

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS QPEMVEAVKKHILNMLHLKKRPD GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-Q51-P72

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS QPEMVEAVKKHILNMLHLKKRP GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

#### GB1-Q51-R71

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTETGSGTS GSTLEVL  
FQGP GS QPEMVEAVKKHILNMLHLKKR GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EMVEAVKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72<sub>M54W</sub>

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EWVEAVKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 20970 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72<sub>M54Y</sub>

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EYVEAVKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 16960 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72<sub>V58I</sub>

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EMVEAIKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72<sub>V58L</sub>

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EMVEALKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 15470 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-E53-P72<sub>V55Y</sub>

MNGS HHHHHHHH TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL  
FQGP GS EYVEAVKKHILNMLHLKGRP GSSWSHPQFEK

$$\epsilon_{280} = 16960 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-PKADD

MGSWSHPQFEK TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL F  
QGP GS MEFAAALG SHIQIPPGLTELLQGYTVEVLRQQPPDLVEFAVEYFTRLREARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD

MGSWSHPQFEK TSTYKLI L NGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKFTFTVETGSGTSGSTLEVL F  
QGP GS QPEMVEAVKKHILNMLHLKRPDVTQPVPKAALLNAIRKLG SHIQIPPGLTELLQGYTVEVLRQQPPDLVEFAV  
EYFTRLREARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>L-PKADD

MGSWSHPQFEK TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
QGP GS QPEMVEAVKKHILNMLHLKKRPDVTQGLG SHIQIPPGLTELLQGYTVEVLRQQPPDLVEFAVEYFTRLREARAP GSG  
SSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub>

MGSWSHPQFEK TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
QGP GSMEF QPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLS LTELLQGYTVEVLRQQPPDLVEFAVEYFTRLR  
EARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>

MGSWSHPQFEK TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
QGP GSMEF QPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLS AS LTELLQGYTVEVLRQQPPDLVEFAVEYFTRLR  
EARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub>

MGSWSHPQFEK TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
QGP GSMEF QPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLS SAS LTELLQGYTVEVLRQQPPDLVEFAVEYFTRLR  
REARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-SS

MNGS HHHHHHHH TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
FQGP GS QPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLS GSGSSWSHPQFEK

$$\epsilon_{280} = 30940 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

### GB1-H<sub>1</sub>LH<sub>2</sub>-PKADD<sub>SS</sub>

MGSWSHPQFEK TSTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTETGSGTS GSTLEVL F  
QGP GS QPEMVEAVKKHILNMLHLKKRPDVTQVPVKAALLNAIRKLS HGLG SHIQIPPGLTELLQGYTVEVLRQQPPDLVEFAV  
EYFTRLREARAP GSGSSP HHHHHHHH

$$\epsilon_{280} = 36900 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$$

## 8.3 Appendix 3: PCR Oligonucleotide Sequences

Here are listed the DNA sequences of the primers used to generate all vectors and constructs used in this thesis. In each case, all primers are shown 5'-3'. The restriction enzyme used for restriction digest is also included in parentheses and the restriction site highlighted in yellow. In some cases primers from a previous section will have been used to amplify the desired sequence. For, clarity primer combinations will be listed at the bottom (Table 19).

### Strep-tag II Primers

ST2\_FP (*XhoI*)

TCGAGCTGGAGCCATCCGCAGTTCGAAAAATA

ST2\_RP (*HindIII*)

AGCTTATTTTTTCGAACTGCGGATGGCTCCAGC

### Peptide Primers

4Q51\_FP (*BamHI*)

TATATGGATCCAGCCAGAGATGGTGGAGGCCGTC

4K69\_FP (*BamHI*)

TATATGGATCCAAGAAGAGACCCGATGTCACCCAG

4Q76\_FP (*BamHI*)

TATATGGATCCAGCCGGTACCCAAGGCGGCGCTTC

4E104\_RP (*XhoI*)

ATATACTCGAGCCCTCTATCTCCACATACCCGTTCTC

4H91\_RP (*XhoI*)

ATATACTCGAGCCATGAAGCTTTCTGATCGCGTTCAG

6Q76\_RP (*XhoI*)

ATATACTCGAGCCCTGGGTGACATCGGGTCTCTTCTT

### Minimal Peptide Primers

4P52\_FP (*BamHI*)

TATATGGATCCAGAGATGGTGGAGGCCGTCAAG

4E53\_FP (*BamHI*)

TATATGGATCCGAGATGGTGGAGGCCGTCAAGAAG

4M54\_FP (*BamHI*)

TATATGGATCCATGGTGGAGGCCGTCAAGAAGCAC

4V55\_FP (*BamHI*)

TATATGGATCCGTGGAGGCCGTCAAGAAGCACATT

4E56\_FP (*Bam*HI)

TATATGGATCCGAGGCCGTCAAGAAGCACATTTTA

4H67\_FP (*Xho*I)

ATATACCTCGAGCCGTGCAGCATGTTTAAAATGTGCTT

4R71\_FP (*Xho*I)

ATATACCTCGAGCCTCTCTTCTTCAAGTGCAGCATGTT

4P72\_FP (*Xho*I)

ATATACCTCGAGCCGGGTCTCTTCTTCAAGTGCAGCAT

4D73\_FP (*Xho*I)

ATATACCTCGAGCCATCGGGTCTCTTCTTCAAGTGCAG

4V74\_FP (*Xho*I)

ATATACCTCGAGCCGACATCGGGTCTCTTCTTCAAGTG

4T75\_FP (*Xho*I)

ATATACCTCGAGCCGGTGACATCGGGTCTCTTCTTCAA

#### **Mutant Minimal Peptide Primers**

4E53\_M54Y\_FP (*Bam*HI)

TATATGGATCCGAGTATGTGGAGGCCGTCAAGAAG

4E53\_M54W\_FP (*Bam*HI)

TATATGGATCCGAGTGGGTGGAGGCCGTCAAGAAG

4E53\_V58I\_FP (*Bam*HI)

TATATGGATCCGAGATGGTGGAGGCCATCAAGAAG

4E53\_V58L\_FP (*Bam*HI)

TATATGGATCCGAGATGGTGGAGGCCCTGAAGAAG

4E53\_V55Y\_FP (*Bam*HI)

TATATGGATCCGAGATGTATGAGGCCGTCAAGAAG

#### **pOP6BPA Primers**

ST2GB1\_FP (*Bam*HI)

TATATGGATCCGTGGAGCCATCCGCAGTTCGAAAAAACTAGTACCTACAAA

GB1\_RP (*Eco*RI)

ATATAGAATTCATGGATCCTGGGCCCTGAAACAG

PKADD\_FP (*Avr*II)

TATATCCTAGGCAGCCACATCCAGATCCCGCCGGGG

PKADD\_RP (*Xho*I)

ATATACCTCGAGCCGCTCCCTGGGGCGCGGGCCTCGCGCAG

### **PKADD Fused Peptide Primers**

6Q51\_FP (*Eco*RI)

TATATGAATTCAGCCAGAGATGGTGGAGGCCGTC

6H91\_RP (*Avr*II)

ATATACCTAGGCCATGAAGCTTTCTGATCGCGTTCAG

6Q76\_RP (*Avr*II)

ATATACCTAGGCCCTGGGTGACATCGGGTCTCTTCTT

### **Continuous Fused Helix Primers**

H<sub>1</sub>LH<sub>2</sub>\_Insert\_FP

ATGGAATTCAGCCAGAGATGGTGGAGGCCGTC

H<sub>1</sub>LH<sub>2</sub>\_Vector\_RP

GGCTGGAATTCATGGAGCCTGGGCCCTG

H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub>\_Vector\_FP

AAGCTTAGCCTCACGGAGCTGCTGCAGGGCTA

H<sub>1</sub>LH<sub>2</sub>-S-PKADD<sub>F</sub>\_Insert\_RP

CGTGAGGCTAAGCTTTCTGATCGCGTTCAGAAGC

H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>\_Vector\_FP

CTTGCGAGCCTCACGGAGCTGCTGCAGG

H<sub>1</sub>LH<sub>2</sub>-AS-PKADD<sub>F</sub>\_Insert\_RP

TGAGGCTCGCAAGCTTTCTGATCGCGTTCAG

H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub>\_Vector\_FP

TAGCGGAGCCTCACGGAGCTGCTGCAG

H<sub>1</sub>LH<sub>2</sub>-SAS-PKADD<sub>F</sub>\_Insert\_RP

AGGCTCGCGCTAAGCTTTCTGATCGCGTTCAG

### **Covalent Dimerization Primers**

4H91\_GSC\_RP

ATATACTCGAGCCGACGAGCCATGAAGCTTTCTGATCGCGTTCAG

TELCQGYT\_mut\_FP

ATATAACGGAGCTGTGTCAGGGCTACACG

TELCQGYT\_mut\_RP

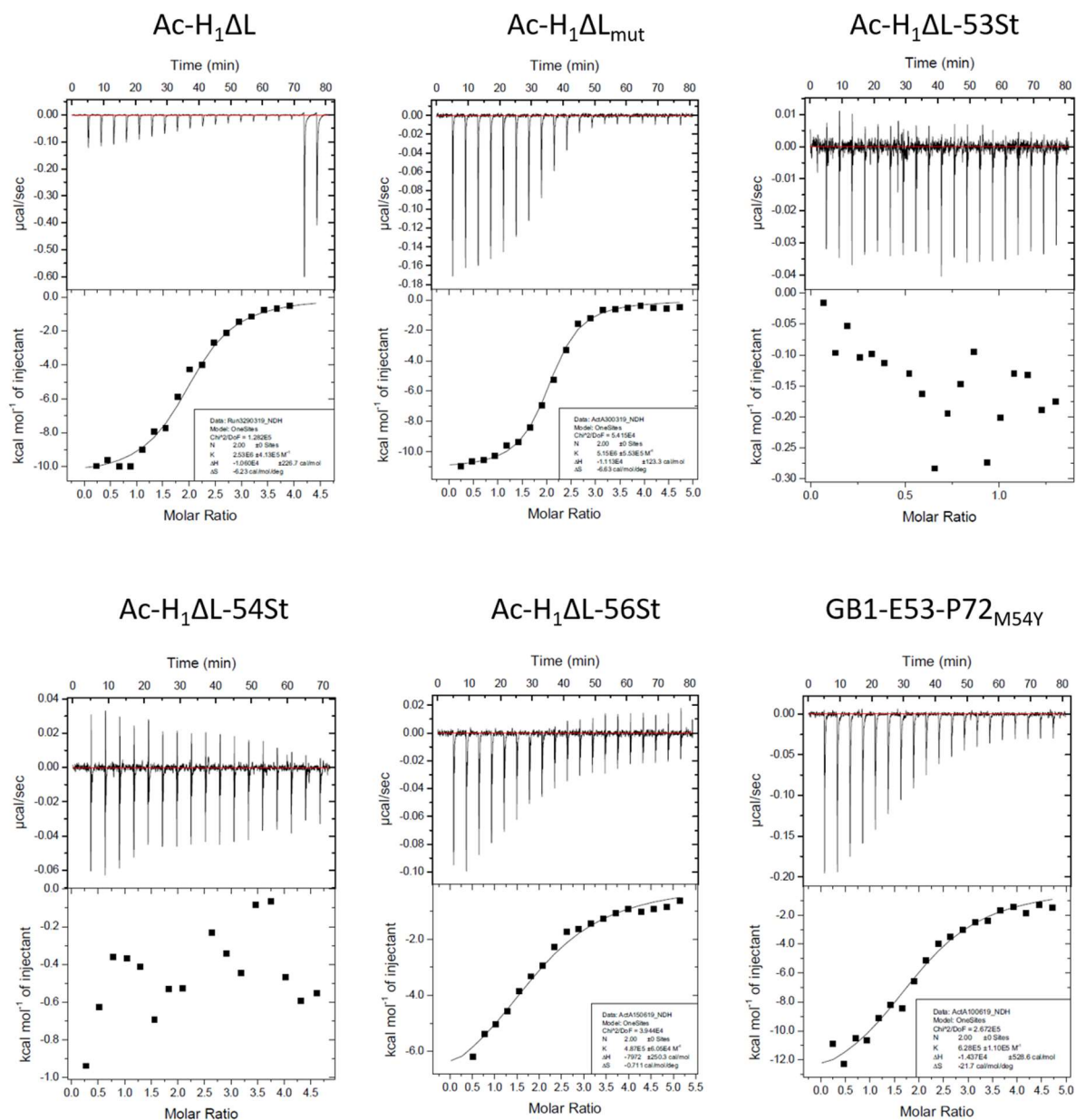
TATATTGCCTCGACACAGTCCCGATGTGC

**Table 19: Primer combinations.**

Construct	Forward Primer	Reverse Primer	Additional Primers	Vector Cloned into
Strep-tag II	ST2_FP	ST2_RP		pOP5BP
H <sub>1</sub> LH <sub>2</sub> β	4Q51_FP	4E104_RP		pOP4BP
H <sub>1</sub> LH <sub>2</sub>	4Q51_FP	4H91_RP		pOP4BP
LH <sub>2</sub> β	4K69_FP	4E104_RP		pOP4BP
H <sub>2</sub> β	4Q76_FP	4E104_RP		pOP4BP
H <sub>1</sub> L	4Q51_FP	4Q76_RP		pOP4BP
H <sub>1</sub>	4Q51_FP	4H67_RP		pOP4BP
P52-Q76	4P52_FP	4Q76_RP		pOP4BP
E53-Q76	4E53_FP	4Q76_RP		pOP4BP
M54-Q76	4M54_FP	4Q76_RP		pOP4BP
V55-Q76	4V55_FP	4Q76_RP		pOP4BP
E56-Q76	4E56_FP	4Q76_RP		pOP4BP
Q51-R71	4Q51_FP	4R71_RP		pOP4BP
Q51-P72	4Q51_FP	4P72_RP		pOP4BP
Q51-D73	4Q51_FP	4D73_RP		pOP4BP
Q51-V74	4Q51_FP	4V74_RP		pOP4BP
Q51-T75	4Q51_FP	4T75_RP		pOP4BP
E53-P72	4E53_FP	4P72_RP		pOP4BP
E53-P72 <sub>M54Y</sub>	4E53_M54Y_FP	4P72_RP		pOP4BP
E53-P72 <sub>M54W</sub>	4E53_M54W_FP	4P72_RP		pOP4BP
E53-P72 <sub>V58I</sub>	4E53_V58I_FP	4P72_RP		pOP4BP
E53-P72 <sub>V58L</sub>	4E53_V58L_FP	4P72_RP		pOP4BP
E53-P72 <sub>V55Y</sub>	4E53_V55Y_FP	4P72_RP		pOP4BP
Strep-tag II – GB1	ST2GB1_FP	GB1_RP		pOP2H
PKADD	PKADD_FP	PKADD_RP		pOP2H
H <sub>1</sub> LH <sub>2</sub>	6Q51_FP	6H91_RP		pOP6BPA
H <sub>1</sub> L	6Q51_FP	6Q76_RP		pOP6BPA
H <sub>1</sub> LH <sub>2</sub> -S-PKADD <sub>F</sub>	H <sub>1</sub> LH <sub>2</sub> _Insert_FP	H <sub>1</sub> LH <sub>2</sub> -S-PKADD <sub>F</sub> _Insert_RP	H <sub>1</sub> LH <sub>2</sub> -S-PKADD <sub>F</sub> _Vector_FP H <sub>1</sub> LH <sub>2</sub> _Vector_RP	pOP6BPA
H <sub>1</sub> LH <sub>2</sub> -AS-PKADD <sub>F</sub>	H <sub>1</sub> LH <sub>2</sub> _Insert_FP	H <sub>1</sub> LH <sub>2</sub> -AS-PKADD <sub>F</sub> _Insert_RP	H <sub>1</sub> LH <sub>2</sub> -AS-PKADD <sub>F</sub> _Vector_FP H <sub>1</sub> LH <sub>2</sub> _Vector_RP	pOP6BPA
H <sub>1</sub> LH <sub>2</sub> -SAS-PKADD <sub>F</sub>	H <sub>1</sub> LH <sub>2</sub> _Insert_FP	H <sub>1</sub> LH <sub>2</sub> -SAS-PKADD <sub>F</sub> _Insert_RP	H <sub>1</sub> LH <sub>2</sub> -SAS-PKADD <sub>F</sub> _Vector_FP H <sub>1</sub> LH <sub>2</sub> _Vector_RP	pOP6BPA
H <sub>1</sub> LH <sub>2</sub> -SS H <sub>1</sub> LH <sub>2</sub> -PKADD <sub>SS</sub>	4Q51_FP 6Q51_FP	4H91_GSCGS_RP PKADD_RP	TELCQGYT_mut_FP TELCQGYT_mut_RP	pOP6BPA

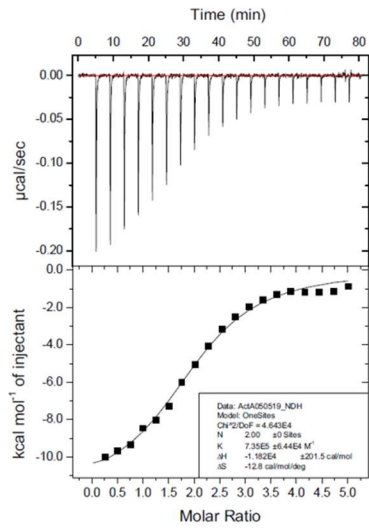
## 8.4 Appendix 4: ITC Isotherms and Thermograms for the Interaction between the Synthetic Peptides / Mutant Minimal Peptide Fusions with Activin A

Here are shown all the isotherms and thermograms for the binding of the interaction between mature activin A and the synthetic peptides and the mutant peptides based on GB1-E53-P72, as determined via ITC. Here *mut* corresponds to the mutations V58I and H61F. *St* corresponds to a N-(4,6-divinyl-1,3,5-triazin-2-yl)-N-methylglycine staple.

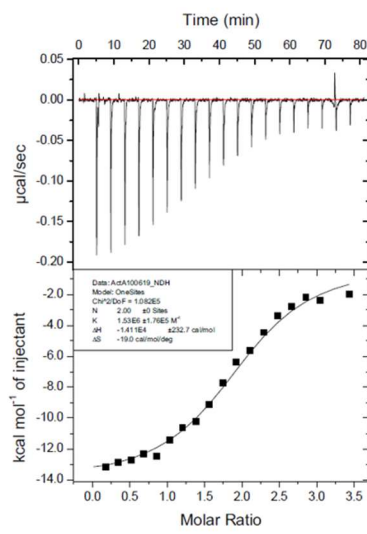




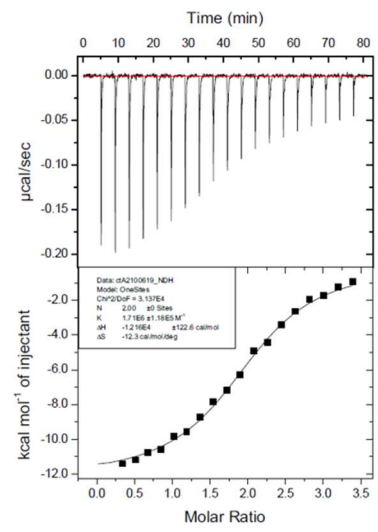
GB1-E53-P72<sub>M54W</sub>



GB1-E53-P72<sub>V58I</sub>



GB1-E53-P72<sub>V58L</sub>



GB1-E53-P72<sub>V55Y</sub>

