

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

The Human Myostatin Precursor Protein: Structure, Function and Amyloid Formation.

Implications for the muscle wastage disease sporadic inclusion body myositis

A dissertation presented in partial fulfillment of the requirements for the degree
of

Doctor of Philosophy
in
Biochemistry

at Massey University, Palmerston North, New Zealand

Carlene Sheree Starck
2010

Abstract

Myostatin is a major player in the regulation of mammalian muscle growth and development, maintaining the balance between proliferation and differentiation prenatally and the quiescence of satellite cells in adults. An absence or overexpression of myostatin results in double-muscling and cachexia respectively, placing myostatin as a promising target in the treatment of muscle wastage diseases.

As a transforming growth factor- β superfamily member, myostatin is produced as a precursor protein, consisting of a propeptide region N-terminal to the growth factor domain. Cleavage of the precursor between the domains forms the myostatin latent complex, an inhibitory structure which is exported from the cell where a second cleavage event releases the active myostatin growth factor. The precursor protein, propeptide, and latent complex play important roles in the regulation of myostatin. However, their structure and function are poorly understood, and a possible role for the myostatin precursor protein in the muscle wastage disease sporadic inclusion body myositis, suggests that pre-growth factor forms of myostatin may be additional important therapeutic targets.

This thesis presents an investigation into the structure and function of the myostatin precursor protein, the latent complex, and the propeptide region within these, with comparisons to a mutant form of myostatin responsible for the naturally-occurring double-muscled phenotype of the Piedmontese cattle breed. Results suggest that the diverse functions of the propeptide region are facilitated by regions of intrinsic disorder within a primarily structured domain, and that conformational alterations accompany the precursor to latent complex transition, resulting in a tighter inhibitory structure. Comparative analyses between the wild-type and mutant proteins suggest that the Piedmontese phenotype is due to a reduced capacity for covalent dimerisation and significant structural alterations within the type I receptor-binding domain. Investigation into misfolded myostatin precursor protein found that the precursor is able to form cytotoxic amyloid aggregates and mature fibrils under partially denaturing conditions, suggesting a possible mechanism for the role of the myostatin precursor in sporadic inclusion body myositis.

Together, these novel results contribute important information towards an understanding of myostatin structure, function and regulation in both normal and disease scenarios.

~ The answer to a question begins with the question itself ~

Acknowledgements

A PhD is more than a degree. It is a journey of self-discovery, a rite of passage. While there is much to be gained, some things are discarded. At the end you look back and discover you are in a much better place then when you began.

Luckily my journey was not made alone. So, there are many people I need to acknowledge and thank.

First and foremost, to my supervisor Dr. Andrew Sutherland-Smith, who somehow managed to deal with all the highs and lows and yet keep me grounded. Most of all, he gave me the space to enter a completely new area, which turned out to be so very worthwhile. Thankyou.

To everyone at the Centre for Structural Biology, especially Greg Sawyer, who played the role of counsellor as well as scientific expert and Trevor, who explained things to me far too many times without showing any frustration. To all my ‘science’ friends, especially Kristy, Komala, Meekyung, Jan, Ava, Corey and Nick, who made even the bad days bearable. To Doug Hopcroft at the MMIC, for his multitude of help with TEM, as well as all the laughs. To Ann, Cynthia, Pat, Katrina, and Tara for all the behind the scenes help and organization. To Jeremy Hyams and Barry Scott, it means a lot to have support from the top.

And a gigantic thankyou to Kathryn Stowell. You have been a constant tower of support and advice from my very first year at Massey University and have become a huge role-model for me. I can honestly say that I would not have had my successes without you.

To my mum. You are the only person in the world who I know will always be there and think I am amazing even when I am breaking down.

To Chris. You only went part of the journey with me but the lessons I learned from you will stay with me a lifetime.

To my ‘non-science’ friends, especially Marti and Lauren, for keeping the balance. And finally to Bryant. I have no idea how you do it but your strength and support have never faltered. You are my rock.

Thankyou.

Table of Contents

<i>List of Figures</i>	<i>xiii</i>
<i>List of Tables</i>	<i>xvii</i>
<i>Abbreviations.....</i>	<i>xviii</i>
1 <i>Introduction</i>.....	1
1.1 Myostatin	1
1.1.1 Myostatin as a TGF- β family member	1
1.1.2 Production of the myostatin growth factor	4
1.1.3 Myostatin function	6
Naturally occurring myostatin-null phenotypes.....	8
Myostatin in non-muscle tissues.....	9
Myostatin signalling.....	10
1.1.4 Structural characteristics.....	11
The myostatin growth factor.....	11
The myostatin precursor, propeptide and latent complex.....	13
1.1.5 Regulation of myostatin.....	15
Transcriptional regulation	16
Post-translational processing and modification	16
Intracellular binding partners.....	17
Extracellular binding partners.....	17
Autoregulation by myostatin	18
1.1.6 Myostatin as a therapeutic target.....	18
Myostatin inhibitors.....	19
1.2 Amyloid.....	21
1.2.1 History of amyloid	21
1.2.2 Characteristics of amyloid fibrils	23
Tinctorial properties.....	23
Electron microscopy	24
X-ray fibre diffraction	24
Small angle X-ray scattering	25
Solid-state NMR	25
X-ray microcrystallography.....	27
Sequence determinants of amyloid formation	27
1.2.3 The pathway to amyloid formation	28
Structural characterisation of intermediates.....	32

1.2.4	The role of amyloid in disease	34
Mechanisms of cytotoxicity.....		37
1.2.5	Functional amyloid.....	39
1.2.6	Sporadic inclusion body myositis and myostatin	40
Sporadic inclusion body myositis		40
A role for myostatin?		40
Summary		42
2	Materials and Methods	43
2.1	Water, chemicals and media	43
2.1.1	Water and chemicals	43
2.1.2	Media for <i>E.coli</i> growth.....	43
2.2	Electrophoresis methods	43
2.2.1	Agarose gel electrophoresis	43
2.2.2	SDS-PAGE	43
2.2.3	Western blotting.....	44
2.3	Cloning.....	45
2.3.1	PCR	45
2.3.2	Restriction digest.....	45
2.3.3	Ligation.....	46
2.3.4	Transformation.....	46
2.3.5	Colony PCR	46
2.3.6	Sequencing.....	47
2.3.7	Site-directed mutagenesis	47
2.4	Protein Expression and Purification.....	48
2.4.1	Expression of recombinant MstnPP	48
2.4.2	Inclusion body isolation, solubilisation and refolding of MstnPP	49
2.4.3	Periplasmic extraction	50
2.4.4	Purification of refolded MstnPP dimer and soluble aggregates	50
Heparin affinity chromatography.....		50
Size exclusion chromatography		51
2.4.5	Ni-NTA affinity chromatography	51
2.4.6	Furin cleavage.....	51
2.5	Amyloid formation	52
2.5.1	Generation of insulin amyloid fibrils	52
2.5.2	Generation of MstnPP amyloid aggregates and fibrils	52

2.6 Cell culture	52
2.6.1 C2C12 growth and maintenance	52
2.6.2 C2C12 proliferation assays.....	53
2.6.3 C2C12 cytotoxicity assays.....	53
2.6.4 WST-1 colorimetric assay for cell proliferation and protein cytotoxicity	53
2.7 Protein analysis.....	54
2.7.1 Circular dichroism spectroscopy.....	54
2.7.2 CD thermal denaturation	54
2.7.3 Thermal shift assays.....	54
2.7.4 X-ray crystallography.....	55
2.7.5 Protease resistance analysis	55
2.7.6 Negative-stained transmission electron microscopy	56
2.7.7 Thioflavin T binding assays.....	56
2.7.8 One dimensional NMR.....	56
2.8 In silico analyses	57
2.8.1 Sequence alignment and tree building.....	57
2.8.2 Primary sequence	57
2.8.3 Secondary structure.....	57
2.8.4 Intrinsic disorder	57
2.8.5 Aggregation and amyloid formation propensity	58
2.8.6 Structural modelling.....	58
Overview of Results	59
3 Production of recombinant human myostatin	60
3.1 Cloning of human myostatin	60
3.2 Expression.....	61
Expression trials	62
pTUM4 expression	62
Refolding.....	64
3.3 Purification of refolded MstnPP	68
3.4 Evidence of refolding success: production of active human MstnGF	71
3.4.1 Furin digest	71
3.4.2 Activity assays	73
4 Biochemical and biophysical analysis of the myostatin precursor protein and latent complex.....	76

4.1	<i>In silico</i> analysis of MstnPP.....	76
4.1.1	Primary sequence analysis	76
4.1.2	Secondary structure prediction.....	78
4.2	Biophysical analysis of MstnPP.....	80
4.2.1	Circular dichroism spectroscopy	80
4.2.2	X-ray crystallography.....	81
4.3	Stability of MstnPP.....	82
4.3.1	Fluorescence-based thermal shift assay (Sypro Orange)	82
4.3.2	CD thermal denaturation	84
4.3.3	Limited proteolysis.....	86
4.4	The latent complex	89
4.4.1	Circular dichroism spectroscopy	89
4.4.2	CD thermal denaturation	90
4.4.3	Gel filtration chromatography.....	92
5	<i>Discussion: The structure and function of the human myostatin precursor and the latent complex.....</i>	96
5.1	The structure and properties of human MstnPP	96
5.1.1	Summary.....	96
5.1.2	MstnPP as an intrinsically disordered protein (IDP)	98
Hallmarks of intrinsic disorder	98	
Structural characteristics	99	
Pinpointing intrinsically disordered regions of MstnPP	100	
Intrinsic disorder and amyloid	102	
Formation of the myostatin latent complex	103	
5.2	Implications for MstnPP and latent complex function <i>in vivo</i>	104
5.2.1	β-Aggregation by MstnPP	104
5.2.2	The propeptide as a chaperone.....	105
5.2.3	Export of myostatin	106
5.2.4	Inhibition by the myostatin propeptide.....	107
5.3	The structure of the latent complex	108
5.3.1	Structural modelling.....	108
5.3.2	The role of cysteine residues in the propeptide domain	112
5.3.3	One or two propeptides per growth factor?	113
5.3.4	Model: Formation of the myostatin latent complex from a partially structured precursor.....	114

Summary.....	116
6 Analysis of C313Y, the Piedmontese myostatin mutation	117
6.1 Production of the C313Y precursor protein (C313Y MstnPP)	117
6.1.1 Site-directed mutagenesis	117
6.1.2 Expression	118
6.1.3 Refolding	119
6.1.4 Purification	120
6.2 Verification of refolding success.....	121
6.2.1 Production of the C313Y growth factor	121
6.2.2 Activity assays	123
6.3 Biochemical analysis of C313Y MstnPP and comparison to WT MstnPP....	124
6.3.1 Secondary structure	124
6.3.2 Thermal stability	125
Thermal shift assay.....	125
CD thermal denaturation	127
6.3.3 Protease resistance.....	129
6.4 The C313Y latent complex	130
6.4.1 Circular dichroism spectroscopy.....	130
6.4.2 CD thermal denaturation	132
6.4.3 Gel filtration analysis	134
6.5 Structural modelling of the C313Y mutation	135
7 Discussion: Analysis of C313Y, the Piedmontese myostatin mutation.....	138
7.1 The human myostatin precursor as a model system.....	138
7.2 Biochemical consequences of C313Y.....	139
7.2.1 Dimerisation.....	139
7.2.2 Stability and flexibility	142
Structural differences	142
Thermal stability.....	144
Proteolysis	145
7.2.3 Processing, secretion and biological activity	145
7.3 Implications for myostatin function <i>in vivo</i>	147
Summary.....	149
8 Amyloid formation by MstnPP.....	150

8.1	<i>In silico</i> predictions.....	150
8.2	The amyloid-like properties of MstnPP soluble aggregates	151
8.2.1	Transmission electron microscopy	151
8.2.2	Thioflavin T assays	152
8.3	Formation of MstnPP amyloid fibrils at acidic pH and elevated temperature	153
8.3.1	ThT and TEM	153
8.3.2	Circular dichroism spectroscopy.....	157
8.3.3	Resistance to proteolytic digest by trypsin	159
8.3.4	SDS resistance and DTT.....	160
8.4	Alternative conditions for amyloid formation	163
8.4.1	Physiological temperature	163
8.4.2	Alternative pH, alternative pathway.....	165
8.5	Formation of amyloid by the refolded MstnPP dimer.....	166
8.6	Biological activity of MstnPP aggregates.....	168
9	<i>Discussion: Amyloid formation by MstnPP</i>	170
9.1	<i>In vivo implications of MstnPP amyloid formation and sporadic inclusion body myositis.....</i>	170
9.1.1	Implications for the folding of MstnPP <i>in vivo</i>	170
9.1.2	Implications for sporadic inclusion body myositis.....	171
How might myostatin/MstnPP be involved in the pathogenesis of sIBM?	172	
Increased activity of the myostatin growth factor	172	
MstnPP aggregation	172	
9.1.3	MstnPP cytotoxicity	173
ER stress-induced upregulation of MstnPP	174	
9.2	Model for amyloid formation	175
9.2.1	Summary of results.....	175
9.2.2	The role of non-native disulphides	176
9.2.3	Overall model.....	178
9.3	Implications of the model for disease	181
Summary	182	
10	<i>Summary and Conclusions</i>	183
The native human myostatin precursor protein can be refolded <i>in vitro</i> after insoluble expression in <i>E.coli</i>	183	

The human myostatin precursor protein may contain functional intrinsic disorder	184
Formation of the human myostatin latent complex may involve structural rearrangement ...	184
The structure of the human myostatin latent complex.....	185
The Piedmontese mutant myostatin has reduced dimerisation capacity.....	185
The Piedmontese mutant myostatin may have structural perturbations and increased flexibility in the Type I receptor-binding domain.....	186
The myostatin precursor protein can form cytotoxic amyloid aggregates and fibrils <i>in vitro</i> with implications for sporadic inclusion body myositis.....	186
Overall summary and conclusions: the big picture	188
11 Future work	189
11.1 Structure and function of MstnPP and latent complex	189
11.1.1 Structural analysis	189
Crystallisation.....	189
Other biophysical analyses	190
Structural changes on complex formation	190
Intrinsic disorder.....	191
11.1.2 Stability	192
11.1.3 Functional analysis.....	192
The role of propeptide cysteine residues	192
Mapping of the propeptide/growth factor binding domains	193
A role for calcium?	194
11.2 Structure of C313Y and implications for function	194
11.2.1 Effects of the C313Y mutation	194
Dimerisation.....	194
Receptor binding.....	195
Thermodynamic stability.....	195
11.2.2 Crystallisation of C313Y	195
11.3 Amyloid.....	196
11.3.1 MstnPP amyloid structure	196
11.3.2 The MstnPP amyloid formation pathway	196
Which regions of MstnPP are involved in amyloid formation?	196
The kinetics of MstnPP amyloid formation.....	196
The effects of disulphide-bonding on MstnPP amyloid formation	197
11.3.3 Implications for sporadic inclusion body myositis (sIBM)	197
Studies in sIBM tissue.....	197
The mechanism of MstnPP amyloid cytotoxicity	198
Summary	199

12 Appendices.....	200
12.1 Appendix 1: The human myostatin growth factor	200
12.1.1 Expression	200
12.1.2 Refolding	201
12.1.3 The production of myostatin growth factor from MstnPP.....	203
12.1.4 Structural and biochemical analysis.....	205
X-ray crystallography	205
Circular Dichroism spectroscopy.....	205
12.1.6 Discussion.....	208
Summary	209
12.2 Appendix 2: pProEX-Htb Vector Map	210
12.3 Appendix 3: Gel filtration calibration curves	211
12.4 Appendix 4: CDDN Deconvolution tables.....	213
12.5 Appendix 5: <i>In silico</i> raw output.....	214
12.6 Appendix 6: MstnGF refolding screen.....	216
13 References.....	217

Publication: Cytotoxic aggregation and amyloid formation by the myostatin precursor protein. (2010). PLoS One. 5: e9170.

List of Figures

Figure 1.1 Phylogenetic tree of the human TGF- β superfamily members.....	2
Figure 1.2 Conservation of myostatin across species.....	4
Figure 1.3 The processing of the myostatin precursor protein.....	5
Figure 1.4 Naturally occurring myostatin-null phenotypes produce a double-muscled phenotype.....	9
Figure 1.5 Canonical signaling pathway of the TGF- β superfamily.....	10
Figure 1.6 Structural characteristics of the myostatin growth factor.....	12
Figure 1.7 A conformational change accompanies TGF- β family latent complex formation.....	14
Figure 1.8 The characteristics of amyloid fibrils.....	26
Figure 1.9 The general pathway to amyloid formation.....	29
Figure 1.10 Structural characterisation of prefibrillar oligomeric intermediates.....	32
Figure 1.11 Variation in amyloid deposition and localisation between diseases.....	35
Figure 1.12 Models for the mechanism of amyloid cytotoxicity.....	38
Figure 2.1 Colony PCR of cells carrying pProEX-MstnF19/F21/F25 constructs.....	47
Figure 2.2 Splice-overlap extension (SOE) mutagenesis protocol.....	48
Figure 2.3 Purification of inclusion bodies containing MstnPP.....	50
Figure 3.1 Cloning of F19, F21, F25 and F267 myostatin cDNA.....	60
Figure 3.2 Expression analysis of MstnF21 in <i>E.coli</i>	61
Figure 3.3 The pTUM4 system for production of soluble MstnF21.....	63
Figure 3.4 Refolding of the myostatin precursor protein dimer.....	66
Figure 3.5 Sequence alignment of human, pig and zebrafish myostatin.....	67
Figure 3.6 Purification of the refolded His-tagged myostatin precursor protein by Ni-NTA chromatography.....	68
Figure 3.7 Purification of refolded MstnPP by heparin affinity and gel filtration chromatography.....	70
Figure 3.8 Furin digest of MstnPP.....	72
Figure 3.9 Activity of the myostatin growth factor by C2C12 proliferation assays ..	74
Figure 4.1 Secondary structure predictions for MstnPP.....	79
Figure 4.2 Circular dichroism spectrum of MstnPP.....	80
Figure 4.3 Preliminary crystals of MstnPP.....	81

Figure 4.4 Thermal shift analysis of MstnPP by Sypro Orange binding and fluorescence.	83
Figure 4.5 CD thermal denaturation of MstnPP.	85
Figure 4.6 Limited proteolysis of MstnPP.	87
Figure 4.7 CD comparison of myostatin latent complex and MstnPP.	89
Figure 4.8 CD thermal denaturation of the furin-digested myostatin latent complex.	91
Figure 4.9 Gel filtration analysis of the furin-digested latent complex.	92
Figure 4.10 Gel filtration analysis of the acid-treated latent complex.	93
Figure 5.1 Intrinsic disorder versus secondary structure in MstnPP.	101
Figure 5.2 Hydrophobic cluster analysis for MstnPP.	102
Figure 5.3 The propeptide as a chaperone.	105
Figure 5.4 Interactions of myostatin with follistatin, TGF- β family receptors and the myostatin propeptide.	109
Figure 5.5 Myostatin interaction domains of follistatin versus the myostatin propeptide.	111
Figure 5.6 Model for the structural characteristics of myostatin latent complex formation.	115
Figure 6.1 Nucleotide sequencing of C313Y cDNA.	118
Figure 6.2 Expression of the myostatin C313Y precursor protein.	118
Figure 6.3 Refolding of the C313Y and WT precursor proteins.	119
Figure 6.4 Purification of the refolded C313Y precursor protein.	120
Figure 6.5 Furin digest of the C313Y precursor protein.	122
Figure 6.6 Activity of the mutant myostatin growth factor by C2C12 proliferation assays.	123
Figure 6.7 Circular dichroism of the C313Y precursor protein.	124
Figure 6.8 Thermal shift analysis of C313Y and WT precursor proteins by Sypro Orange binding and fluorescence.	126
Figure 6.9 CD thermal denaturation of the C313Y precursor protein.	128
Figure 6.10 Limited proteolysis of C313Y and WT precursor proteins.	129
Figure 6.11 Circular dichroism of the C313Y latent complex.	131
Figure 6.12 CD thermal denaturation of the C313Y latent complex.	133
Figure 6.13 Gel filtration analysis of acid-treated C313Y latent complex.	134
Figure 6.14 Modelling of the C313Y mutation in the myostatin growth factor crystal structure.	135

Figure 6.15 Rotamers available to C313Y.....	136
Figure 7.1 Sequence alignment of human and bovine myostatin precursor proteins.....	139
Figure 7.2 Molecular structure of VEGF mutant equivalent to C313Y.....	141
Figure 7.3 Location of the VEGF ΔIII mutant and associated structural changes.....	143
Figure 7.4 Trypsin cleavage sites of the C313Y growth factor.....	145
Figure 7.5 Putative type I and type II receptor interactions for the myostatin growth factor.....	148
Figure 8.1 <i>In silico</i> predictions of propensity for β-sheet aggregation and amyloid formation.....	151
Figure 8.2 Characterisation of MstnPP soluble aggregates by negative-stain transmission electron microscopy.....	152
Figure 8.3 Thioflavin T fluorescence of MstnPP soluble aggregates.....	153
Figure 8.4 Characterisation of MstnPP prefibrillar oligomers and amyloid fibrils by ThT binding and TEM.....	154
Figure 8.5 ThT fluorescence at pH 7.5.....	155
Figure 8.6 Circular dichroism spectra showing the structural transitions that occur in amyloid fibril formation.....	158
Figure 8.7 MstnPP resistance to trypsin digestion by SDS-PAGE analysis and TEM.....	159
Figure 8.8 The effects of DTT on amyloid formation shown by TEM, ThT binding and CD spectroscopy.....	162
Figure 8.9 Characterisation of MstnPP amyloid formation at 37°C by CD, ThT binding and TEM.....	164
Figure 8.10 Characterisation of MstnPP amyloid formation at alternative pH values by TEM.....	165
Figure 8.11 Characterisation of amyloid formation by the refolded MstnPP dimer by TEM.....	167
Figure 8.12 MstnPP amyloid C2C12 cytotoxicity assay.....	169
Figure 9.1 The putative role of disulphide bonding in amyloid formation.....	177
Figure 9.2 Model of the amyloid formation pathway for MstnPP.....	179
Figure 12.1 Expression analysis of the myostatin growth factor by reducing SDS-PAGE.....	201
Figure 12.2 The MstnGF refolding screen.....	202

Figure 12.3 Purification of the myostatin growth factor post-furin digest by RP-HPLC.....	204
Figure 12.4 Preliminary crystals of the myostatin growth factor.	205
Figure 12.5 CD spectra of the myostatin growth factor in different purification conditions.....	206
Figure 12.6 The pProEX-Htb vector map (Invitrogen).	210
Figure 12.7 Superdex S200 calibration curves.	212
Figure 12.8 MstnPP disorder predictions: raw data.	214
Figure 12.9 MstnPP β -aggregation and amyloid formation propensity predictions: raw data.	215
Figure 12.10 MstnGF refolding screen.	216

List of Tables

Table 1.1 Potential myostatin inhibiting therapeutics evaluated in animal models of disease.....	20
Table 1.2 Amyloid diseases and associated precursor proteins and protein aggregates.	22
Table 2.1 Summary of antibodies and detection methods used for Western blotting.	44
Table 3.1 Summary of main expression systems trialled for MstnF21.....	62
Table 3.2 Refolding methods used for the myostatin precursor protein.	64
Table 4.1 <i>In silico</i> prediction of primary sequence motifs in MstnPP.....	77
Table 4.2 Identification of proteins and protein complexes from gel filtration of the acid-treated latent complex.	94
Table 6.1 Modelling the three most common rotamers available to C313Y	137
Table 8.1 Measurement of distance between twists for MstnPP fibrils.	157
Table 9.1 Summary of biochemical and biophysical properties of the MstnPP dimer, soluble aggregates, protofibrils and mature amyloid fibrils.	176
Table 12.1 Normalised CDDN deconvolution of MstnGF CD spectra.	208
Table 12.2 Normalised CDDN deconvolution tables.....	213

Abbreviations

Aβ	Amyloid beta protein
AβPP	Amyloid beta precursor protein
ACN	Acetonitrile
βME	Beta-mercaptoethanol
BMP	Bone morphogenic protein
C313Y	Piedmontese mutation (cysteine to tyrosine at position 313)
CD	Circular dichroism
DTT	Dithiothreitol
ECM	Extracellular matrix
EDTA	Ethylenediaminetetra-acetic acid
ER	Endoplasmic reticulum
GDF	Growth and differentiation factor
GndHCl	Guanidine hydrochloride
HCA	Hydrophobic cluster analysis
IAPP	Islet amyloid polypeptide
IDP	Intrinsically disordered protein
IPTG	Isopropyl beta-D-thiogalactoside
kB	Kilobases
kDa	Kilodaltons
LB	Luria broth
LB-amp	Luria broth containing 100 µg/mL ampicillin
MstnGF	Myostatin growth factor
MstnPP	Myostatin precursor protein
MWt	Molecular weight
NF-κB	Nuclear factor kappa B
NMR	Nuclear magnetic resonance
NR	Non-reducing (SDS-PAGE)
PCR	Polymerase chain reaction
PrP	Prion protein
R	Reducing (SDS-PAGE)
ROS	Reactive oxygen species
RP-HPLC	Reverse phase-high performance liquid chromatography

· Abbreviations ·

SAXS	Small angle X-ray scattering
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sIBM	Sporadic inclusion body myositis
TCEP	Tris 2-carboxyethyl phosphine
TEM	Transmission electron microscopy
ThT	Thioflavin T
TGF-β	Transforming growth factor beta
Tm	Melting temperature
UPR	Unfolded protein response
VEGF	Vascular endothelial growth factor
WT	Wild-type

1 Introduction

1.1 Myostatin

Myostatin was first described as Growth and Differentiation Factor-8 (GDF-8) in 1997 following a PCR screen on murine DNA that used degenerate oligonucleotides corresponding to a highly conserved sequence among the transforming growth factor β (TGF- β) family (McPherron, Lawler *et al.* 1997). The TGF- β family member was subsequently shown to play a role in the inhibition of muscle growth and development and its null mutation to be responsible for the naturally-occurring double-muscled phenotype observed in a number of cattle breeds (McPherron and Lee 1997). This function has implicated a role for myostatin in muscular atrophy and placed myostatin as a promising target for the treatment of a number of muscle wastage diseases. Although the structure and regulation of the myostatin growth factor have been studied, little is known about the biochemistry of the precursor protein and the secreted latent complex, two important aspects of myostatin and TGF- β family biology.

1.1.1 Myostatin as a TGF- β family member

The human TGF- β superfamily represents a group of 33 cytokines (Fig. 1.1) that play important roles in a wide variety of processes associated with normal tissue development and homeostasis such as wound healing, mammary gland development, bone metabolism, skin formation and immune function (Keah and Hearn 2005). The superfamily is characterized by dimeric growth factors produced via cleavage of a precursor protein, a cysteine-knot tertiary motif and signaling via type II and type I transmembrane Ser/Thr kinase receptors. Members can be divided into three classes based on sequence identity, receptor utilization and shared inhibitors; the TGF- β s, bone morphogenic proteins (BMPs)/growth and differentiation factors (GDFs) and the activin/inhibins (Cash, Rejon *et al.* 2009).

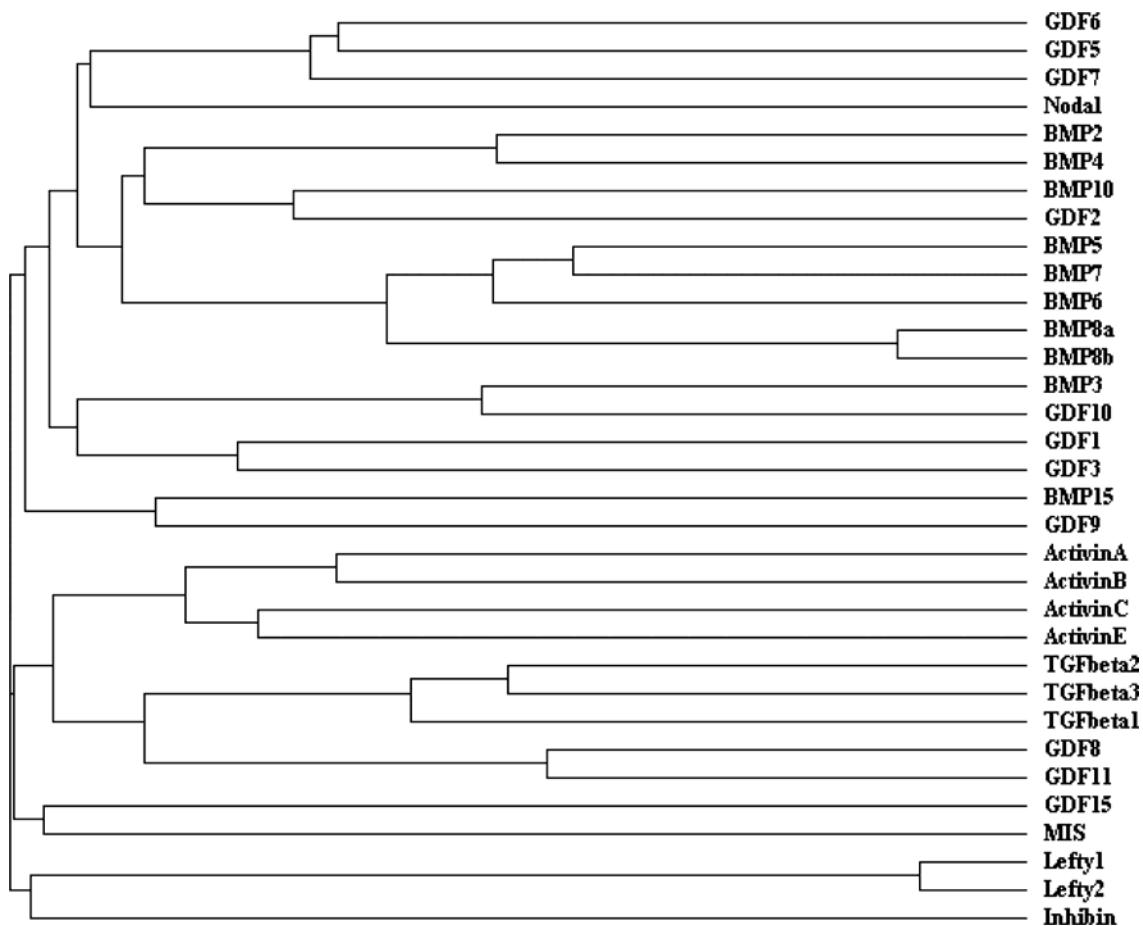


Figure 1.1 Phylogenetic tree of the human TGF- β superfamily members.

Tree constructed using amino acid sequence information from Expasy and ClustalW alignment (Thompson, Higgins *et al.* 1994).

The myostatin primary sequence has all the hallmarks of the TGF- β family (Fig. 1.2) including an N-terminal secretory signal sequence (blue), N-terminal propeptide and C-terminal growth factor regions separated by a dibasic furin convertase RXRR recognition motif (green) and the conserved cysteine residues that contribute to the cysteine knot (red) and dimerisation (red, underlined) in the growth factor domain. An N-glycosylation site is predicted in the propeptide region (orange). The myostatin growth factor domain sequence is identical between humans, mice, rats, pigs, chickens and turkeys (Lee 2004). The propeptide region shows less conservation.

a.

Canine	-MQRLQICVYIYL FVLIVAGPV --DLSENSEQKENVEKEGLCNACMRQNTKSSRIEAI	56
Equine	-MQQLQISVYYIYL FVLILAGPV --DLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAI	56
Bovine	-MQQLQISVYYIYL FMLIVAGPV --DLNENSEQKENVEKEGLCNACLWRRENTTSSRLEAI	56
Sheep	-MQQLQIFVYYIYL FMLIVAGPV --DLNENSEQKENVEKEGLCNACLWRQNNKSSRLEAI	56
Pig	-MQQLQIVYYIYL FMLIVAGPV --DLNENSEQKENVEKEGLCNACMRQNTKSSRIEAI	56
Human	-MQQLQIVYYIYL FMLIVAGPV --DLNENSEQKENVEKEGLCNACMRQNTKSSRIEAI	56
Mouse	MMQLQLQMVYYIYL FMLIAAGPV --DLNEGSEREENVEKEGLCNACAWRQNTRYSRIEAI	57
Chicken	-MQKLAVVYIYL FMLQIAVDPV --ALDGSSQPTENAEKDGLCNACTWRQNTKSSRIEAI	56
Zebrafish	--- MHFTQVLISLSVLIAACGPVGYGDITAHQQPSTATESEQCSTCEFRQHSKLMRLHAI	57
* * * : : .** : . : . * . * : * : : . * : .**		
Canine	KIQILSKLRLETAPNI SDAVRQLPRAAPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Equine	KIQILSKLRLETAPNI KDAIRQLLPKAPPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Bovine	KIQILSKLRLETAPNI KDAIRQLLPKAPPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Sheep	KIQILSKLRLETAPNI KDAIRQLLPKAPPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Pig	KIQILSKLRLETAPNI KDAIRQLLPKAPPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Human	KIQILSKLRLETAPNI KDVIRQLLPKAPPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Mouse	KIQILSKLRLETAPNI KDAIRQLLPRAAPPLRELIDQYDVRQDDSSDGSLEDDDYHATTE	117
Chicken	KIQILSKLRLEQAPNI SDVIRQLLPKAPPPLQELIDQYDVRQDDSSDGSLEDDDYHATTE	116
Zebrafish	KSQILSKLRLKQAPNI SDVVKQLLPKAPPPLQQLLDQYDVRQDDSSDGSLEDDDYHATTE	117
* * * * * : * * * * : * * * * : * * * * : * * * * : * * * * : * * * * :		
Canine	TVIAMPETD LIMMQVEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Equine	TIITMPTESD LIMMQVEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Bovine	TVIAMPTE SDLIAEVQEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Sheep	TIITMPTE SDLIAEVQEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Pig	TIITMPTE SDFLMQVEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Human	TIITMPTE SDFLMQVEGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	176
Mouse	TIITMPTE SDFLMQADGKPKCCFFKSSKIQYNKVVKAQLWIYL RPVKTPTVFQILRL	177
Chicken	TIITMPTE SDFLVQMEGKPKCCFFKSSKIQYNKVVKAQLWIYL RQVKPTTVFQILRL	176
Zebrafish	TIMTMATEPD PIVQVDRKPCKCFSSPKIQA NIRVRAQLWHLRPAEATTVFLQISRL	177
*: * : * : * : : : * * * * * : * * * * : * * * * : * * * * : * * * * :		
Canine	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Equine	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Bovine	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Sheep	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Pig	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Human	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	236
Mouse	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKALDENGH	237
Chicken	IKPMKDGT RTYTGIRSLKLDMNPGTGIWQSIDVKT VQLNWLKQPESNLGIEIKAFDETGRD	236
Zebrafish	M-PVKDGGHR-IRS LKIDVNAGVTSWSQSIDVKQVLTVWLKQPETNRGIEINAYDAKND	235
: *; * * : * : * : . * . * * * * * : * * * * : * * * * : * * * * : * . *		
Canine	LAVTFPGPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Equine	LAVTFPRPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Bovine	LAVTFPEGEDGLTPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Sheep	LAVTFPEGEGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Pig	LAVTFPGPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Human	LAVTFPGPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Mouse	LAVTFPGPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	297
Chicken	LAVTFPGPGEDGLNPFPLEV KVTDTPKRSRRDFGLCDDEHSTESRC CRYPLTVDFEAFGWD	296
Zebrafish	LAVTSTEGEDGLNPFPMEVKR ISEGPKRIRRDSGLCDENSSESRC CRYPLTVDFEAFGWD	295
***** . * * ; * * : * : : * * * * * : * * * * * : * * * * * :		
Canine	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	356
Equine	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	356
Bovine	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGE	356
Sheep	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPKGSAGP CC TPTKMSPINMLYFNGK	356
Pig	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	356
Human	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	356
Mouse	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	357
Chicken	WIIAPKRYKANY CSGEC EFVFLQKYPHTHLVHQANPRGSAGP CC TPTKMSPINMLYFNGK	356
Zebrafish	WIIAPKRYKANY CSGEC DYMYLQKYPHTHLVNKASPRGTAGP CC TPTKMSPINMLYFNGK	355
***** . * * ; * * : * : : * * * * * : * * * * * : * * * * * :		
Canine	EQIYGKIPAMVV DRCGCS	375
Equine	EQIYGKIPAMVV DRCGCS	375
Bovine	GQIYGKIPAMVV DRCGCS	375
Sheep	EQIYGKIPGMVV DRCGCS	375
Pig	EQIYGKIPAMVV DRCGCS	375
Human	EQIYGKIPAMVV DRCGCS	375
Mouse	EQIYGKIPAMVV DRCGCS	376
Chicken	EQIYGKIPAMVV DRCGCS	375
Zebrafish	EQIYGKIPSMVV DRCGCS	374
***** . *****		

b.



Figure 1.2 Conservation of myostatin across species.

(previous page) a. ClustalW alignment (Thompson, Higgins *et al.* 1994). The signal peptide is shown in blue, the predicted glycosylation site in orange, the furin recognition motif in green, and cysteines involved in the cysteine knot are in red. ClustalW similarity key is as follows: *, exact match; :, strong similarity; ., weak similarity.

b. Phylogenetic tree constructed by ClustalW from alignment in a.

1.1.2 Production of the myostatin growth factor

The 375 amino acid human myostatin precursor protein is directed to and translated in the endoplasmic reticulum (ER) (Fig. 1.3a). Oxidative folding in the ER allows formation of the correctly folded precursor dimer, with dimerisation via a disulphide bond at the C-terminus (Fig. 1.3b). The precursor protein is processed, most likely by the *trans*-Golgi localized furin convertase enzyme, at a dibasic RSRR motif (Fig. 1.3c), generating a complex of noncovalently-associated propeptide and growth factor dimer (Fig. 1.3d) which is secreted from the cell (Fig. 1.3e) (Lee 2004; Keah and Hearn 2005). Activity of the growth factor is inhibited in this complex until a second cleavage event disrupts the complex (Fig. 1.3e) (Lee 2004). Myostatin protein may also be exported in the unprocessed precursor form (Figure 1.3g); a pool of extracellular myostatin precursor was identified in mouse skeletal muscle (Anderson, Goldberg *et al.* 2008) suggesting that two mechanisms for the control of myostatin activation exist.

Evidence suggests that the precursor protein and latent complex play a number of important roles. First, the propeptide appears to play a chaperone-like function; folding of the growth factor only occurs in the precursor form (Jin, Dunn *et al.* 2004; Lee 2004; Funkenstein and Rebhan 2007).

Second, formation of the latent complex may be necessary for secretion. This is well-documented for other TGF- β family members such as the inhibins/activins and TGF- β 1 (Gray and Mason 1990) and the binding interface essential for this function is conserved in the myostatin sequence (Walton, Makanji *et al.* 2009). The propeptide regions of TGF- β 1 and TGF- β 2 contain two predicted glycosylation sites that when mutated reduce secretion of the growth factor (Brunner, Lioubin *et al.* 1992); the myostatin propeptide has a similar predicted glycosylation site (Fig. 1.2). Furthermore, the C-terminal region of the myostatin propeptide has been suggested to

play a role in stability of the growth factor (Jiang, Liang *et al.* 2004) which may be important during export.

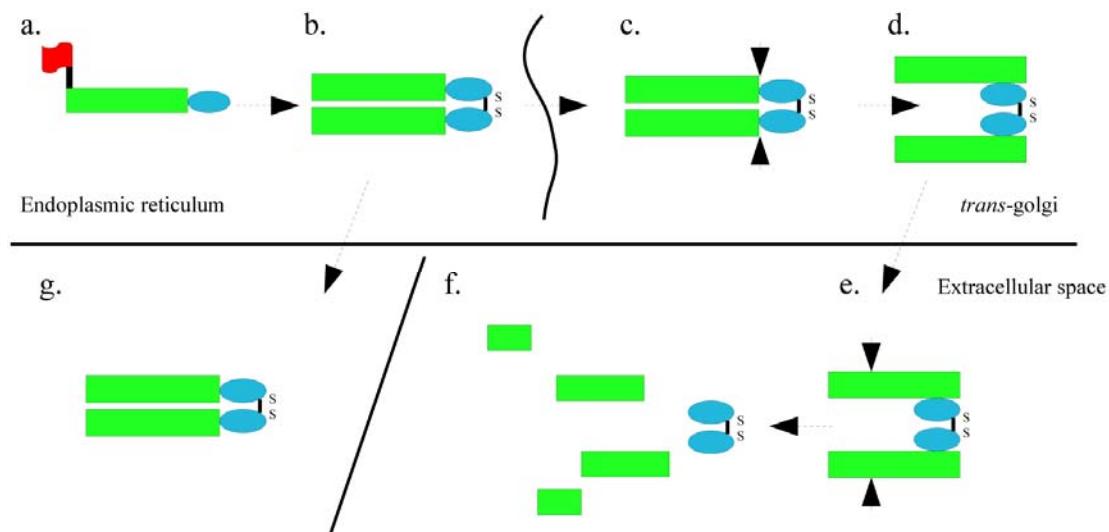


Figure 1.3 The processing of the myostatin precursor protein.

The myostatin precursor protein consists of an N-terminal propeptide region (green) and the C-terminal growth factor domain (blue) (a). The signal sequence (red) directs the precursor protein to the ER for dimerisation (b). The precursor protein is cleaved at a conserved furin RXXR motif (Lee 2004) (c) to generate the latent complex (d) which is secreted to the extracellular space (e). Cleavage by metalloproteases (Wolfman, McPherron *et al.* 2003) disrupts the latent complex to release the active myostatin growth factor dimer (f). Myostatin may also be secreted as a precursor protein (g) (Anderson, Goldberg *et al.* 2008).

Third, both the precursor and latent complex play regulatory functions. Extracellular precursor protein must be cleaved by furin convertase to form the latent complex (Anderson, Goldberg *et al.* 2008). The latent complex, whether originating from extra- or intracellular precursor, must be disrupted for activation of the myostatin growth factor. *In vitro*, this can be achieved using acid or heat treatment (Zimmers, Davies *et al.* 2002; Funkenstein and Rebhan 2007) and *in vivo* is most likely to occur via cleavage N-terminal to Asp 76 of the propeptide region by members of the bone morphogenetic protein-1/tolloid (BMP-1/TLD) family of metalloproteases (Wolfman, McPherron *et al.* 2003; Lee 2008). The inhibitory region of the myostatin propeptide has been characterized as residues 42-115 (Jiang, Liang *et al.* 2004). Recent evidence suggests that the stoichiometry of latent complex formation for some of the bone

morphogenic proteins, including BMP2, BMP4, BMP7, BMP10 and GDF5 is two propeptides bound to one growth factor dimer (Sengle, Charbonneau *et al.* 2008). Inhibition is likely to be due to competition for receptor binding (Walton, Makanji *et al.* 2009) as seen for BMP-7 (Sengle, Charbonneau *et al.* 2008).

Fourth, latent complex formation may assist in localization of the myostatin signal. A number of interactions with proteins in the extracellular matrix (ECM) have been documented for the propeptide domains of TGF- β family members including BMP7 with fibrillin (Gregory, Ono *et al.* 2005), TGF- β s 1-2 and BMPs 2 and 4 with heparin or heparin sulphate (Rider 2006) and the TGF- β s with integrin receptors (Wipff and Hinz 2008). Latent TGF- β complexes are secreted from the cell associated with a latent TGF- β binding protein (LTBP); the interaction is mediated via a disulphide linkage between the propeptide and the LTBP (Rifkin 2005). The LTBPs are multidomain glycoproteins that associate with the ECM to restrict the circulation of TGF- β ligands. While a myostatin propeptide/fibrillin interaction is unlikely (Sengle, Charbonneau *et al.* 2008), interaction of the myostatin precursor with LTBP3 has been implicated (Anderson, Goldberg *et al.* 2008).

1.1.3 Myostatin function

Initially labeled as GDF-8, functional studies in mice indicated a role as a negative regulator of muscle growth and development, hence the name myostatin (McPherron, Lawler *et al.* 1997). A myostatin knockout mouse, generated by deletion of a portion of the gene encoding the C-terminal domain of myostatin, exhibited dramatic and widespread increases in skeletal muscle mass in a dose-dependent manner, with mice heterozygous for the null mutation showing an intermediate phenotype (McPherron, Lawler *et al.* 1997). If myostatin activity is inhibited by the administration of a myostatin-blocking antibody to adult mice, a 13-30% increase in muscle size is observed (Joulia-Ekaza and Cabello 2006). *In vitro* investigations using mouse C2C12 myoblasts showed that cell proliferation is significantly inhibited in the presence of myostatin by comparison to myostatin-free cells (Thomas, Langley *et al.* 2000; Taylor, Bhansin *et al.* 2001) and systemic overexpression of myostatin in adult mice induces profound muscle and fat loss analogous to that seen in human cachexia syndromes (Zimmers, Davies *et al.* 2002).

Inhibition of muscle cell proliferation after myostatin signaling is achieved primarily through the upregulation of p21, a cyclin dependent kinase (Cdk) inhibitor involved in the G1/S transition of the cell cycle. p21 expression results in a decrease in Cdk activity and subsequent activation of the retinoblastoma protein (Rb); Rb represses transcription factors involved in the expression of S-phase specific genes (Thomas, Langley *et al.* 2000). Myostatin signaling has also been shown to augment the degradation of cyclin D1, a cell-cycle co-regulator, through a proteasome-dependent pathway (Yang, Chen *et al.* 2006).

Myostatin signaling inhibits the expression of the myogenic regulatory factors (MRFs) which encode transcription factors regulating muscle differentiation (Patel and Amthor 2005). MyoD expression is down-regulated in C2C12 myoblasts treated with the myostatin growth factor (Langley, Thomas *et al.* 2002) and myogenin is an important target of both endogenous and over-expressed myostatin (Joulia, Bernardi *et al.* 2003). In contrast, *in vivo* research suggests that myostatin promotes the terminal differentiation of embryonic muscle progenitors; targeted gene expression of myostatin specifically to the early muscle progenitor population showed that cell cycle arrest induced by myostatin signaling maintained a balance between proliferation and differentiation (Manceau, Gros *et al.* 2008). Furthermore, *in vivo* myostatin overexpression led to an activation of MyoD expression. It is likely that the environmental context in which myostatin signaling acts is crucially important (Manceau, Gros *et al.* 2008).

Early results suggested that postnatally, myostatin signaling maintains the quiescence of satellite cells, a dormant muscle stem cell population required for repair and regeneration of adult muscle (Patel and Amthor 2005). It was widely accepted that muscle hypertrophy following an inhibition or deletion of myostatin activity was attributed to the proliferation of these stem cells. Recent research however shows that muscle hypertrophy driven by myostatin blockade does not require satellite-cell activity (Amthor, Otto *et al.* 2009). Alternative mechanisms, such as the modulation of protein synthesis and the turnover of structural muscle fibre proteins, are responsible.

Myostatin function may also have an effect on oxidative metabolism; lack of myostatin results in a profound muscle-fibre type conversion towards a glycolytic phenotype, suggesting a deficit in the oxidative metabolism of skeletal muscle (Amthor, Macharia *et al.* 2007; Dumonceaux and Amthor 2009). This effect is not well-understood.

In addition to myostatin, other TGF- β family members may contribute to the regulation of muscle mass (Lee, Reed *et al.* 2005), including TGF- β 1 (Angelis, Borghi *et al.* 1998), and the activins and BMP-11 (Souza, Chen *et al.* 2008).

Naturally occurring myostatin-null phenotypes

A number of mammalian species possess naturally occurring myostatin null mutations that result in a double-muscled phenotype (Fig. 1.4). A large number of cattle breeds such as the Belgian Blue (Fig. 1.4a) have been recognized by breeders for almost 200 years; six different myostatin mutations account for most, if not all, of the double-muscled cattle breeds identified to date (Lee 2007). In 2004 the first case of a naturally occurring mutation in the myostatin gene in a human child was reported (Fig. 1.4b). The child has increased muscle mass, is unusually strong and does not show any negative side effects from the mutation (Schuelke, Wagner *et al.* 2004). Other species include Texel sheep (Fig. 1.4c) and racing whippets (Fig. 1.4d, top image). Myostatin expression follows Mendelian genetics with heterozygous individuals exhibiting an intermediate muscle phenotype (Fig. 1.4d) (Lee 2007).

The majority of myostatin null-mutations result in premature stop codons and an absence of myostatin protein (Kambadur, Sharma *et al.* 1997). In contrast, the myostatin gene of the Piedmontese cattle breed contains a point mutation at nucleotide position 938 with a guanine substituted for an adenine. This mutation results in the replacement of cysteine 313 with a tyrosine and the loss of one of the cysteine residues involved in the TGF- β family cysteine knot structure. Piedmontese myostatin failed to inhibit C2C12 myoblast proliferation on comparison to wildtype myostatin *in vitro* (Berry, Thomas *et al.* 2002) and acted as a competitive inhibitor of wildtype myostatin when added in excess. Early results indicated that the removal of cysteine 313 results in the distortion of the cysteine knot structure, a decrease in stability and a loss of receptor binding (Berry, Thomas *et al.* 2002).

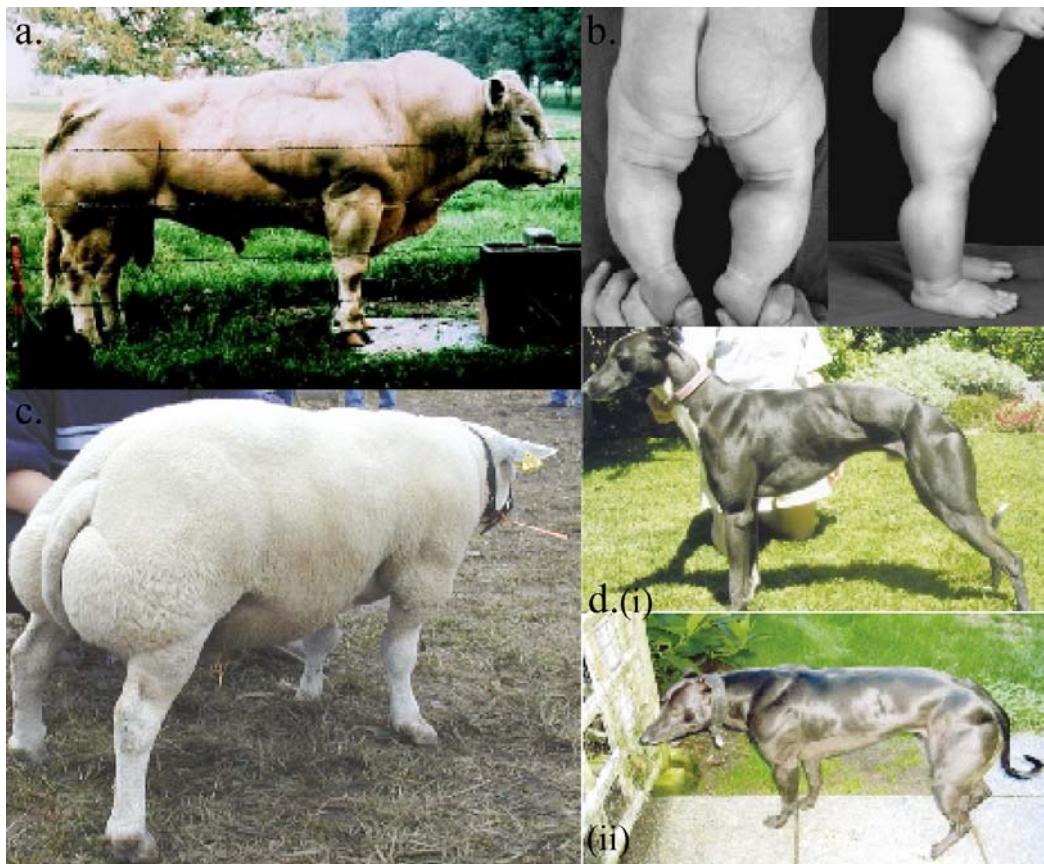


Figure 1.4 Naturally occurring myostatin-null phenotypes produce a double-muscled phenotype.
Myostatin null mutations have been found in a. a number of cattle breeds, such as the Belgian Blue (Lee 2004); b. A human child (Schuelke, Wagner et al. 2004); c. Texel sheep (Lee 2004); d. Racing whippets which exhibit both heterozygous (i) and homozygous (ii) null phenotypes (Lee 2004).

Myostatin in non-muscle tissues

Although predominant in skeletal muscle, myostatin expression has also been reported in other tissues such as adipose (Hirai, Matsumoto *et al.* 2007), the heart (Gaussian and Depre 2005), placenta (Mitchell, Osepchook *et al.* 2006) and the mammary gland (Thomas, Langley *et al.* 2000).

Myostatin appears to act on adipogenic cells directly; myostatin treatment of cultured 3T3-L1 cells blocks differentiation independent of TNF- α or IL-6, known mediators of cachexia in mice (Zimmers, Davies *et al.* 2002). In addition, myostatin inhibits differentiation of the bovine preadipocyte (Hirai, Matsumoto *et al.* 2007). Contrasting evidence suggests that the reduction in body fat seen in myostatin-null mice is due to a metabolic shift where dietary fat utilization is shifted toward muscle rather than

adipose tissue (Lee 2004; Yang and Zhao 2006). Myostatin expression was upregulated in cardiomyocytes following myocardial infarction, suggesting a role for myostatin in cardiac development and pathophysiology (Sharma, Kambadur *et al.* 1999). Evidence that myostatin is synthesized, released and acts within the human placenta by contributing to placental glucose homeostasis was published in 2006 (Mitchell, Osepchook *et al.* 2006).

Myostatin signalling

TGF- β family growth factors signal by binding to type II and type I transmembrane Ser/Thr kinase receptor heterodimer complexes (Fig. 1.5). Binding of the growth factor ligand to a type II receptor homodimer stimulates recruitment of the type I receptor homodimer and heterodimerisation.

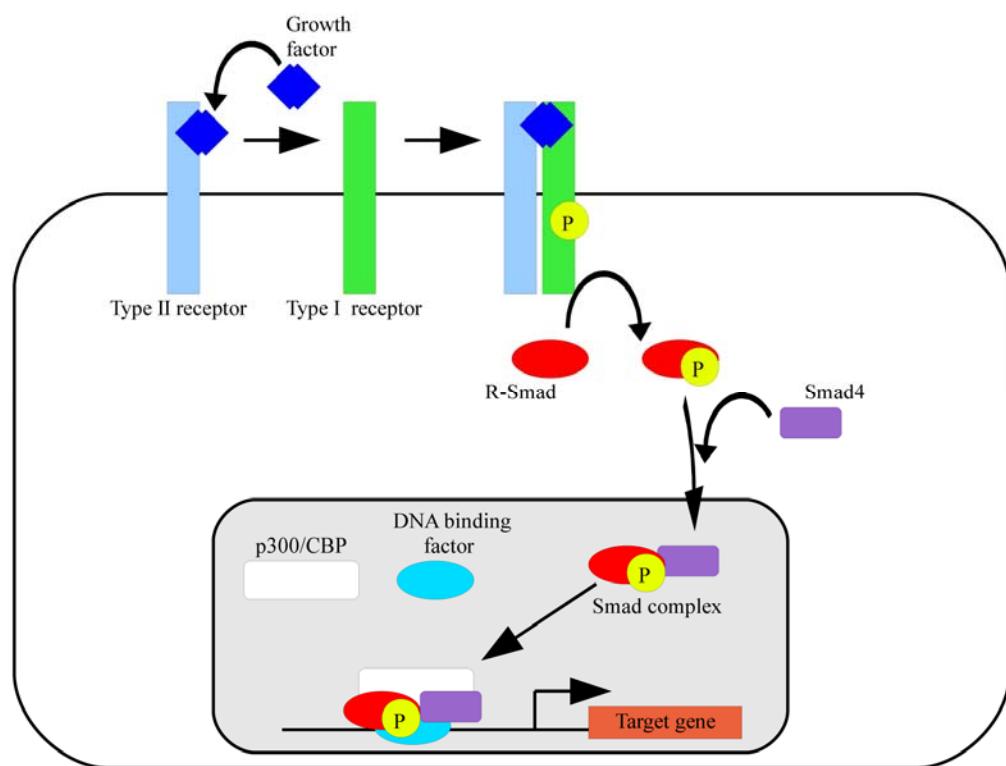


Figure 1.5 Canonical signalling pathway of the TGF- β superfamily.

TGF- β family growth factor dimers signal through type II and type I transmembrane Ser/Thr kinase receptors. Activation of the type I receptor allows intracellular signal propagation via R-Smads and the central Smad 4 protein, which accumulate in the nucleus to control gene expression. I-Smads antagonize signaling. Figure adapted from (Dijke and Hill 2004).

Activation of the type I receptor through phosphorylation by the type II receptor allows intracellular signal propagation via phosphorylation of Receptor-Smad proteins (R-Smads) which form heteromeric complexes with the central Smad 4 protein and accumulate in the nucleus to control gene expression in a cell-type specific and ligand dose-dependent manner. Inhibitory Smads (I-Smads) antagonize signaling through a number of mechanisms such as competition for receptor binding (Dijke and Hill 2004).

Myostatin may also signal through alternative pathways such as the extracellular signal-related kinase 1/2 (Erk1/2) MAPK pathway (Yang, Chen *et al.* 2006) and the PI3K/Akt/GSK-3 β signaling cascade (Yang, Zhang *et al.* 2006). Myostatin has also been shown to increase the expression of genes involved in ubiquitin-mediated proteolysis by reversal of the IGF-1/PI3K/Akt hypertrophy pathway (McFarlane, Plummer *et al.* 2006). These alternative pathways may serve as nodes for crosstalk of the Smads with other major signaling pathways, integrating cellular mechanisms and physiological responses (Moustakas and Heldin 2005).

1.1.4 Structural characteristics

The myostatin growth factor

The X-ray crystal structure of the myostatin growth factor was recently determined (Cash, Rejon *et al.* 2009). High resolution three-dimensional structures have been determined for at least nine other TGF- β superfamily members (Keah and Hearn 2005).

Myostatin displays the traditional TGF- β family architecture which is hand-shaped (Fig. 1.6a) (Cash, Rejon *et al.* 2009) with each monomer consisting of 4 curved β -strands or ‘fingers’, a cysteine knot motif in the ‘palm’ region and a major helix or ‘wrist’. Myostatin has the intricate cysteine knot motif where two disulphide bridges connect two neighbouring chain segments to form a ring structure and a third disulphide bridge penetrates the ring segment to cross-link two additional chain segments (Figure 1.6b) (Muller and Heiring 2002). Each myostatin monomer contains four intramolecular disulphides; three are involved in the cysteine knot (Fig. 1.6c, cysteines 15-74, 43-106 and 47-108). Two myostatin monomers come together palm to palm in an antiparallel direction to generate concave and convex surfaces (Fig.

1.6a) and dimerise via a ninth intermolecular disulphide bond, which provides additional stabilization of hydrophobic interactions that are the major driving force for dimerisation (Venkataraman, Sasisekharan *et al.* 1995). The cysteine knot and dimerisation are considered to be the major determinants of protein stability in the TGF- β superfamily (Muller and Hering 2002), although the dimerisation disulphide is not essential in all family members, such as the vascular endothelial growth factor (VEGF) (Muller and Hering 2002) and GDF-9 (Chang, Brown *et al.* 2002). Members of the TGF- β superfamily exhibit high thermal and structural stability (Brownh, Wakefiel *et al.* 1990), properties attributed to the cysteine knot motif and common to a range of cysteine-knot containing proteins such as collagen type III (Barth, Kyrieleis *et al.* 2003) and peptide toxins (Craik, Daly *et al.* 2000).

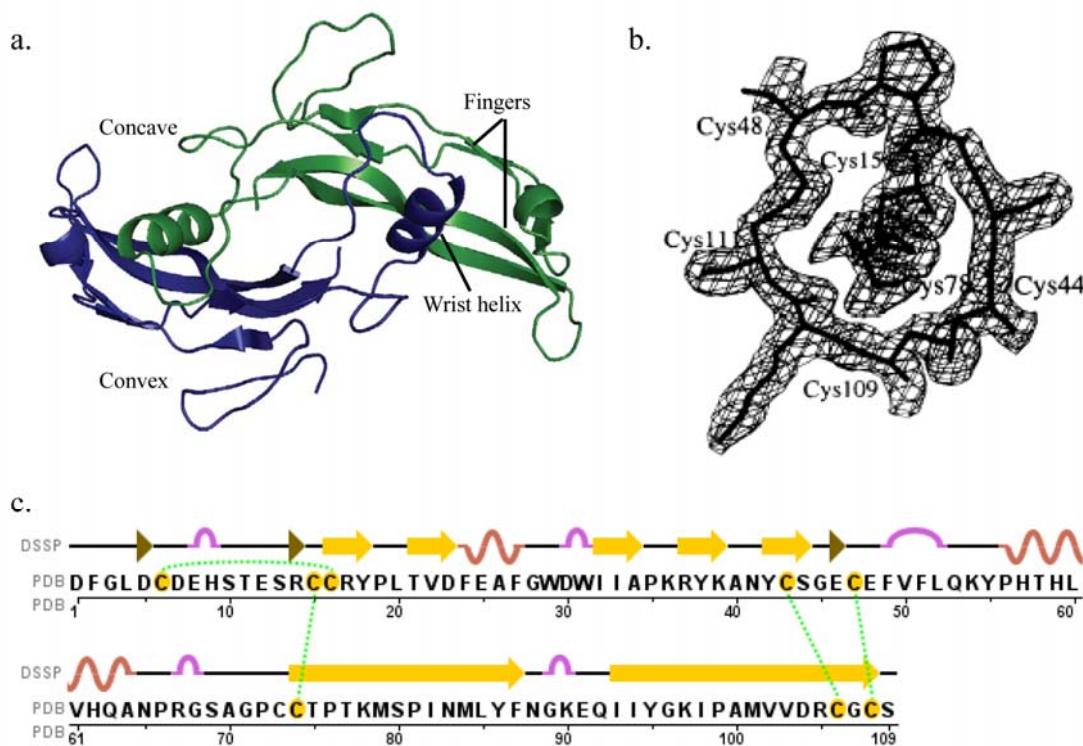


Figure 1.6 Structural characteristics of the myostatin growth factor.

- The X-ray crystal structure of the myostatin growth factor (Cash, Rejon *et al.* 2009).
- The cysteine knot motif of vascular endothelial growth factor (VEGF) (Muller and Hering 2002).
- The myostatin amino acid sequence showing secondary structure, taken from the RSCB. Cysteines involved in intramolecular disulphide bonds (green dotted lines) are shown in yellow. Secondary structure is as follows: yellow arrow, β -sheet; pink wave, α -helix; purple loop, β -turn.

The published structures of the TGF- β growth factors point to three modes of dimerisation (Thompson, Woodruff *et al.* 2003). TGF- β 1, TGF- β 2, BMP-2 and BMP-7 show the common elongated dimer structure where the monomers dimerise in a head-to-tail fashion with the α -helix of one monomer interacting with the concave β -strand surface of the other monomer. TGF- β 3 and GDNF exhibit a more open conformation and the activin/inhibin dimers have a more compact conformation than TGF- β s 1 and 2. In addition, there are at least two modes of receptor binding (Keah and Hearn 2005). Crystallographic analysis has revealed that BMPs and activins bind the ectodomains of type II receptors on their convex surface and type I receptors on their concave surface (Thompson, Woodruff *et al.* 2003). TGF- β s also bind type I receptors on their concave surface but bind type II receptors more distally towards the fingertip region; in this case, the two types of receptors interact with each other, which is necessary for high-affinity ternary complex formation and signaling (Keah and Hearn 2005).

Traditionally, myostatin was thought to belong to the activin class of the TGF- β superfamily due to 40% amino acid sequence identity with activin and binding of the activin type II receptors and the activin inhibitor follistatin (Cash, Rejon *et al.* 2009). The X-ray crystal structure however bears similarities to the TGF- β class as well. Although the myostatin structure is similar to other TGF- β family members, there are significant differences in the N-terminus and region preceding the wrist helix (the prehelix loop) which is situated in the type I receptor binding site. The prehelix loop is most similar to TGF- β ; the activin prehelix loop contains a number of glycine and serine residues and is flexible whereas myostatin has several large hydrophobic residues in this area.

The myostatin precursor, propeptide and latent complex

No structure of the precursor protein, the propeptide region or latent complex for any TGF- β family member has been published. However, circular dichroism (CD), denaturation studies and mutagenesis provide information on secondary structure, stability and latent complex interactions respectively.

Recombinant TGF- β 1 propeptide has a CD spectrum dominated by a minimum at 206 nm and a shoulder at 223 nm (Fig. 1.7a) (McMahon, Dignam *et al.* 1996). Although the authors state that these minima are characteristic of β -sheet structure, a number of

published CD reviews state these represent α -helical structures (Kelly, Jess *et al.* 2005). The latent complex has a spectrum composed primarily of α -helix with some β -sheet. An addition spectrum of the β -sheet predominant TGF- β 1 growth factor and the TGF- β 1 propeptide differs from that of the latent complex with a reduction in intensities of both structural minima suggesting that structural rearrangement accompanies the formation of the latent complex (McMahon, Dignam *et al.* 1996). Circular dichroism of the BMP-2 precursor protein and isolated propeptide showed similar results with minima at 212 and 218 nm for the precursor and a predominance of α -helix for the propeptide domain (Hillger, Herr *et al.* 2005).

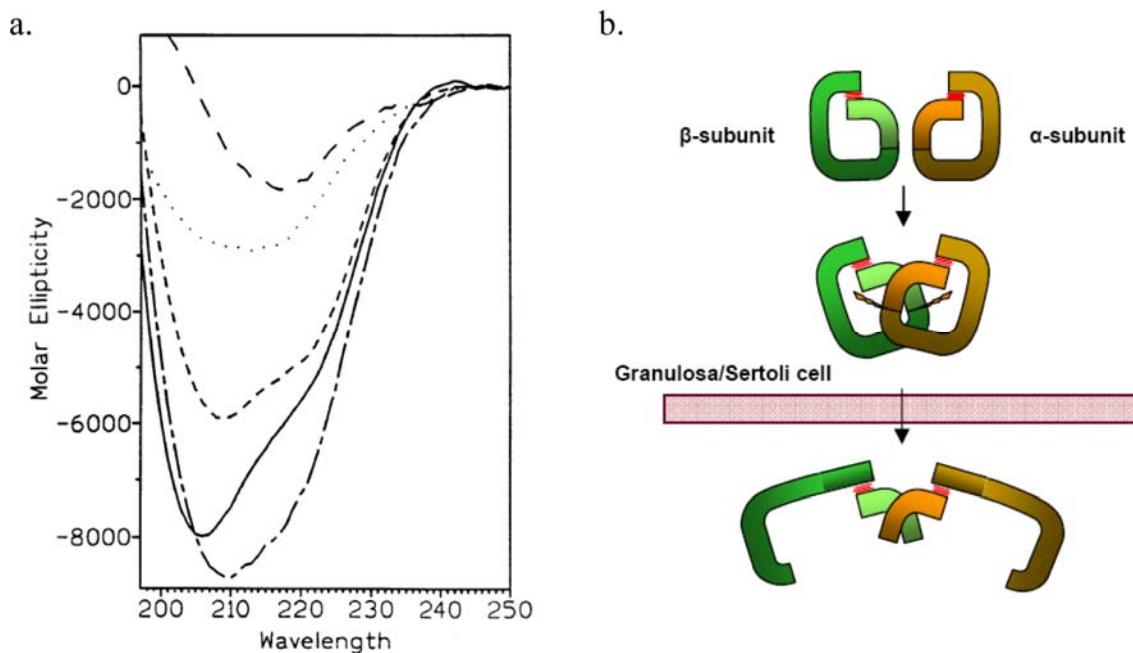


Figure 1.7 A conformational change accompanies TGF- β family latent complex formation.

- a. CD analysis of TGF- β 1 latent complex formation (McMahon, Dignam *et al.* 1996). CD spectra of recombinant TGF- β 1 growth factor (long-dash line), propeptide (solid line), latent complex (short-dash line), summed spectra of TGF- β 1 and propeptide (dash-dot-dash line) and difference spectra between latent complex and summed spectra (dotted line).
- b. A proposed model for precursor dimerisation, latent complex formation and post-secretion structure (Walton, Makanji *et al.* 2009).

Thermodynamic stability of BMP-2 species, measured by urea and guanidine HCl-induced denaturation, indicates reduced structural stability of the propeptide in comparison to the growth factor, and that the growth factor domain stabilizes the structure of the propeptide in the precursor protein (Hillger, Herr *et al.* 2005).

Site-specific mutagenesis of the TGF- β 1 propeptide characterized residues 50-85 as important for interaction with the TGF- β 1 growth factor (Sha, Yang *et al.* 1991). Similarly, mutagenesis of the myostatin propeptide domain characterized the inhibitory region to be residues 42-115, implying that this region contains residues important for interaction with the growth factor in the latent complex (Jiang, Liang *et al.* 2004). The C-terminus of the myostatin propeptide was suggested to confer propeptide stability.

Investigation into the interactions involved in the inhibin/activin latent complex showed that hydrophobic interactions involving the N-terminus of the propeptide were important for growth factor binding in the precursor protein (Walton, Makanji *et al.* 2009); the outer convex surface of the growth factor is postulated to be involved in this interaction. A conserved hydrophobic motif (Hyd-Hyd-X-X-Hyd-X-Hyd) containing the identified residues in the propeptide domains of other TGF- β family members implies a common mechanism for interaction, for example, the TGF- β 1 propeptide is able to form a complex with the growth factor regions of other family members (McMahon, Dignam *et al.* 1996).

A proposed model for precursor dimerisation, latent complex formation and secretion is shown in Figure 1.7b. Hydrophobic interactions between the N-terminal portion of the propeptide and the ‘finger regions’ of the growth factor maintain the molecule in a conformation competent for dimerisation. Cleavage by furin generates the latent complex which consists of a growth factor dimer non-covalently associated to two propeptide domains, and is secreted from the cell (Walton, Makanji *et al.* 2009).

1.1.5 Regulation of myostatin

Myostatin is subject to regulation at all levels, ensuring coordinated and tightly controlled activity. A number of mechanisms are apparent including post-translational modification, competition for receptor binding and feedback control.

Transcriptional regulation

The myostatin promoter contains a number of transcription factor and hormone binding sites. Eight putative E-boxes corresponding to MyoD binding sites have been identified (Du, Chen *et al.* 2005). Stimulation of promoter activity by MyoD suggests that direct up-regulation of myostatin expression is involved in the ability of MyoD to control differentiation (Joulia-Ekaza and Cabello 2006). The down-regulation of MyoD expression by myostatin signaling (Langley, Thomas *et al.* 2002) suggests that feedback mechanisms are in place.

The androgen, progesterone and glucocorticoid response elements were also observed (Du, Chen *et al.* 2005). Glucocorticoid-induced muscle atrophy is a well-known condition observed during treatment of inflammatory disorders in humans (Joulia-Ekaza and Cabello 2006); an increase in myostatin gene expression by glucocorticoids may contribute to muscle atrophy, with a dose-dependent induction of endogenous myostatin transcription in response to dexamethasone reported (Joulia-Ekaza and Cabello 2006). FoxO1 regulatory sites are present in the myostatin promoter region (Allen and Unterman 2007). FoxO1 overexpression leads to an upregulation of myostatin activity and FoxO transcription factors have been suggested to play a central role in the regulation of gene expression during skeletal muscle atrophy.

Post-translational processing and modification

As introduced previously, production of myostatin as a precursor protein constitutes a significant means of inhibition, as the precursor must be cleaved either intra- or extracellularly by furin convertases to generate a latent complex. This complex can be purified from the serum of mammalian cells (Lee and McPherron 2001) with myostatin growth factor activity only detected following acid treatment to disrupt the latent complex (Hill, Davies *et al.* 2002; Zimmers, Davies *et al.* 2002). Transgenic mice expressing high levels of the myostatin propeptide exhibit dramatic increases in skeletal muscle mass comparable to those seen in the myostatin-knockout mouse, indicating that the propeptide inhibits myostatin function *in vivo* (Lee and McPherron 2001). The propeptide domain may compete for receptor-binding; studies of the BMP-7 latent complex suggest that type II receptors displace binding of TGF- β family propeptide domains to the growth factor (Sengle, Ono *et al.* 2008).

Intracellular binding partners

In skeletal muscle, myostatin can be found complexed with hSGT (human small glutamine-rich TPR-containing protein), a protein required for progress through cell division (Winnefeld, Grewinig *et al.* 2006). hSGT interacts with the N-terminal signal peptide region of myostatin, suggesting an inhibition of myostatin translation and/or translocation (Joulia-Ekaza and Cabello 2006).

The sarcomeric protein titin cap binds to mature myostatin, appearing to prevent latent complex formation and consequently inhibit myostatin secretion; overexpression of the titin cap inhibits the antiproliferative influence of myostatin without alteration in myostatin production (Nicholas, Thomas *et al.* 2002; Joulia-Ekaza and Cabello 2006). However, the relationship between the sarcomeric localization of the titin cap protein and the secreted nature of myostatin is unclear.

Extracellular binding partners

Follistatin is a major TGF- β family antagonist (Lee and McPherron 2001); transgenic mice over-expressing follistatin show increased muscling (Lee and McPherron 2001). The crystal structure of myostatin in complex with follistatin provides key information regarding this inhibition (Cash, Rejon *et al.* 2009). Two follistatin molecules wrap around the myostatin growth factor dimer and completely block all four receptor binding sites, as seen in the activin:follistatin complex (Thompson, Lerch *et al.* 2005). The myostatin:follistatin interaction is more intimate than the corresponding activin complex, with a larger amount of buried surface area (Cash, Rejon *et al.* 2009). Comparison of the published structures suggests that follistatin adopts alternate conformations to adapt to the shape of the different ligands (Thompson, Lerch *et al.* 2005; Cash, Rejon *et al.* 2009). The key interaction between myostatin and follistatin involves hydrogen bonding between the ‘fingertip’ motif of myostatin and the N-terminal domain helix of follistatin (Cash, Rejon *et al.* 2009); the latter is suggested to confer specificity to the antagonist. In addition to a blockage of receptor binding, complexed follistatin is known to bind heparin on the cell surface causing cellular uptake and ligand degradation. Follistatin complexed with myostatin has increased affinity for heparin (Cash, Rejon *et al.* 2009), suggesting two modes of myostatin regulation by follistatin.

Two other myostatin-binding proteins have been detected complexed with myostatin in serum; the follistatin-related gene (FLRG) and the growth and differentiation factor-associated serum protein-1 (GASP-1) (Hill, Qiu *et al.* 2003). FLRG is known to bind and inhibit the action of other TGF- β family members *in vitro* and functional studies show that FLRG inhibits myostatin activity in a reporter gene assay (Hill, Davies *et al.* 2002). GASP-1 binds directly to both mature myostatin and its propeptide and contains multiple protease domains, suggesting function is via inhibition of propeptide cleavage, preventing activation of mature myostatin (Hill, Qiu *et al.* 2003).

Myostatin activity is also modulated by decorin, a small leucine-rich ECM-localised proteoglycan previously shown to bind and regulate the activity of TGF- β 1 (Miura, Kishioka *et al.* 2006); decorin binding prevented the growth inhibitory effect of myostatin on C2C12 myoblasts.

Autoregulation by myostatin

Myostatin signaling has been shown to negatively regulate the expression of furin convertase (McFarlane, Langley *et al.* 2005). Myostatin exhibits reduced processing during differentiation; it is possible that this is due to a down-regulation of furin protease activity by an increase in myostatin expression prior to differentiation.

Myostatin upregulates the expression of Smad7, an inhibitory Smad (Forbes, Jackman *et al.* 2006). Smad7 overexpression leads to an inhibition of myostatin gene expression and the inhibition of Smad7 expression by RNA interference increases myostatin promoter activity. These mechanisms may autonomously control myostatin expression in response to extracellular myostatin levels.

1.1.6 Myostatin as a therapeutic target

The role of myostatin in the regulation of muscle development and the striking muscle wastage and double-muscling phenotypes, seen in the overexpression and absence of myostatin respectively, suggest myostatin inhibition as an effective strategy for the treatment of a number of muscle degenerative diseases (Patel and Amthor 2005; Tobin and Celeste 2005; Bradley, Yaworsky *et al.* 2008).

Myostatin function has been implicated in a number of diseases where muscle wastage is a predominant feature. Circulating levels of myostatin are elevated in HIV-infected patients with muscle wasting (Gonzalez-Cadavid, Taylor *et al.* 1998; Lee 2004), in the muscle of patients with muscular atrophy due to disuse (Patel and Amthor 2005) and in an experimental model of cancer cachexia (Costelli, Muscaritoli *et al.* 2008). Mutations in the Titin-Cap protein, shown to bind myostatin and prevent secretion, lead to limb girdle muscular dystrophy type II (Patel and Amthor 2005). Overexpression of the myostatin2 isoform in zebrafish reduces the expression of dystrophin associated protein complex (DAPC) which leads to muscular dystrophy (Amali, Lin *et al.* 2008). In addition, myostatin overexpression has been observed in a cell-culture model of the muscle-wasting disorder sporadic inclusion body myositis (Nogalska, Wojcik *et al.* 2007).

The myostatin pathway may also be a target for the treatment of obesity and type II diabetes. Postnatal inhibition or deletion of myostatin activity prevents fat accumulation in mice fed a high-fat diet (Hamrick, Pennington *et al.* 2006; Yang and Zhao 2006) and the introduction of the myostatin null mutation into two different obese strains of mice resulted in a significant reduction in fat accumulation and the suppression of the development of insulin resistance (Lee 2004).

Myostatin inhibitors

To date, six potential myostatin inhibiting therapeutics have been evaluated in animal models of disease including muscular dystrophy and motor neuron disease (Bradley, Yaworsky *et al.* 2008); these are summarized in Table 1.1. In general, all increased muscle mass by hypertrophy. A truncated form of myostatin (myostatin peptide) also promoted muscle regeneration after injury and increased satellite cell activation (Siriett, Salerno *et al.* 2007).

It has been pointed out that myostatin is highly conserved and functional across healthy individuals. Therefore, the hypothesis that deleting or blocking myostatin is beneficial for skeletal muscle may be over-simplistic, and may result in a negligence to explore its proper function in skeletal muscle, such as the way it modifies oxidative metabolism (Dumonceaux and Amthor 2009).

Myostatin regulatory molecule	Mechanism of action	Animal model	Disease indication
Neutralizing antibodies to myostatin	Bind myostatin and prevent receptor binding	<i>mdx</i>	Duchenne muscular dystrophy (DMD)
		<i>Sgcd</i> ^{-/-}	Limb girdle muscular dystrophy 2F (LGMD2F)
		<i>Sgcg</i> ^{-/-}	LGMD2C
		SOD1 ^{G93A} Transgenic	Amyotrophic lateral sclerosis (ALS)
Myostatin propeptide	Prevents activation of myostatin growth factor	<i>mdx</i>	DMD
		<i>Capn3</i> ^{-/-}	LGMD2A
		<i>Sgca</i> ^{-/-}	LGMD2D
Follistatin	Binds to myostatin	SOD1 ^{G93A} Transgenic	ALS
HDAC inhibitors	Induces follistatin	<i>mdx</i>	DMD
		<i>Sgca</i> ^{-/-}	LGMD2D
Myostatin peptide	Likely prevents Type II receptor binding (ActRIIB)	Aged mice	Sarcopenia
Soluble ActRIIB	Soluble, decoy receptor for myostatin	<i>mstn</i> ^{-/-} Caveolin 3 transgenic mice	LGMD1C

Table 1.1 Potential myostatin inhibiting therapeutics evaluated in animal models of disease.

Table adapted from (Bradley, Yaworsky *et al.* 2008)

The majority of research surrounding myostatin biochemistry and its role in disease and disease treatment focuses on the active myostatin growth factor. Little focus had been placed on the precursor until a role in the amyloid disease sporadic inclusion body myositis, the most common muscle wasting disease of the elderly, was suggested (Wojcik, Engel *et al.* 2005).

1.2 Amyloid

1.2.1 History of amyloid

The term ‘amyloid’ was initially coined by Schleiden and then by Virchow in the mid-19th century to describe the iodine-stained deposits seen in the liver and nervous system at autopsy (Meredith 2005; Rambaran and Serpell 2008). These were mistakenly identified as starch (starch is *amylum* in Latin) until the high nitrogen content of the deposits was established. The tinctorial property of specific binding of amyloid to the dye Congo Red, producing an apple-green birefringence when examined under cross-polarised light, suggested that the amyloid was fibrillar in structure (Divry and Florkin 1927) and electron microscopy confirmed this (Cohen and Calkins 1959). Soon after, several physical techniques including X-ray fibre diffraction indicated that amyloid fibrils are composed of β -sheets (Eanes and Glenner 1968).

The amyloid- β (A β) peptide, which forms the bulk of amyloid plaques in Alzheimer’s disease, was identified and biochemically characterized in 1984 (Glenner and Wong 1984) and is now one of the best-understood amyloid systems. The ‘prion’ hypothesis was put forward two years earlier to describe a protein particle that unlike other amyloid-forming proteins was infective and capable of causing scrapie in sheep (Prusiner 1982); bovine and human prion proteins have since emerged as responsible for Bovine spongiform encephalopathy (BSE) and Creutzfeld-Jacob disease (vCJD) respectively, and suggest transmissibility of prions between species (Rambaran and Serpell 2008). It is now suggested that more than 20 different degenerative human pathologies including Alzheimer’s and Parkinson’s diseases, Type II diabetes and the spongiform encephalopathies (Table 1.2) may have a common cause and pathological mechanism; the misfolding, aggregation and deposition of amyloid-forming proteins, resulting in apoptosis (Soto 2003; Glabe 2006). Many amyloid forming proteins are produced by the proteolytic processing of a precursor protein (Table 1.2) (Rambaran and Serpell 2008).

Amyloid fibrils are also formed by a number of non-disease related proteins and it was recently suggested that all proteins are able to form amyloid given the right conditions (Stefani and Dobson 2003). These observations have led to the hypothesis that amyloid formation is a generic property of the polypeptide chain.

Protein Aggregate	Precursor Protein	Clinical Disease
Amyloid β peptides	Amyloid precursor protein	Alzheimer's disease Down Syndrome Sporadic inclusion body myositis
Prion (PrP^{Sc})	PrP^{C}	Kuru Creutzfeld-Jakob disease/scrapie Bovine spongiform encephalopathy Gerstmann-Straussler-Scheinker disease Fatal familial insomnia
Alpha-synuclein	Alpha-synuclein	Parkinson's disease
Glutamine repeats	Huntington protein	Huntington's disease Spinocerebellar ataxia Machado-Joseph atrophy Dentato-rubro-pallidoluysian atrophy
IAPP/amylin	Pro-IAPP	Type II diabetes
ALys	Lysosome	Hereditary nonneuropathic systemic amyloidosis
Serum amyloid A	(Apo) serum amyloid A	Secondary systemic amyloidosis
ATTR	Transretinatin	Familial amyloid polyneuropathy I Senile systemic amyloidosis Familial cardiac amyloid
AApoA1	Apolipoprotein A1	Familial amyloid polyneuropathy II Familial visceral amyloid
β -microglobulin	$\text{A}\beta 2\text{M}$	Haemodialysis-related amyloidosis Prostatic amyloid

Table 1.2 Amyloid diseases and associated precursor proteins and protein aggregates.

Table adapted from (Stefani and Dobson 2003; Meredith 2005; Rambaran and Serpell 2008).

In addition to amyloid aggregation, degenerative diseases share a striking number of common pathological features, such as evidence of membrane damage, oxidative stress and damage, mitochondrial dysfunction, altered ion and metal homeostasis, aberrant signal transduction, upregulation of autophagy and cell death (Glabe 2006). However, which mechanisms are primary, and the relationship between them is not well-understood. This understanding is important in deciding which of the pathological mechanisms to target for therapeutic development (Glabe 2006).

Endoplasmic reticulum (ER) stress and the unfolded protein response (UPR) are well-documented (Lin and Popko 2009). As the site of oxidative protein folding, the ER is highly sensitive to cell stresses such as perturbations in calcium homeostasis, redox status and elevated secretory protein synthesis rates (Lin and Popko 2009). Cell stress results in the accumulation of unfolded or misfolded proteins in the ER lumen and

activation of the UPR, which functions to restore homeostasis by mechanisms such as the transcriptional induction of chaperones and genes involved in degradation (Schroder 2008). If the UPR fails, the apoptotic pathway is stimulated.

1.2.2 Characteristics of amyloid fibrils

The classical characterization of amyloid fibrils consists of three major structural criteria, regardless of the protein making up the fibres (Stromer and Serpell 2005); staining and birefringence with Congo Red, a cross- β diffraction pattern by X-ray fibre diffraction and straight, unbranching fibrils around 10 nm in diameter visualized by electron microscopy. While these characteristics are still relevant, advances in biophysical techniques such as solid-state NMR and cryo-electron microscopy have allowed a detailed investigation into amyloid structure.

Tinctorial properties

Congo red staining is routinely used as a diagnostic tool to detect amyloid plaques in the brain post-mortem (Frid, Anisimov *et al.* 2007); amyloid fibrils stain red when visualized under a light microscope (Fig. 1.8a (i)) and produce an apple-green birefringence when examined between cross-polarisers (Fig. 1.8a (ii)) (Rambaran and Serpell 2008). Assays can also be carried out in solution, as binding to amyloid induces a characteristic shift in Congo red maximal absorbance from 490 nm to 540 nm (Frid, Anisimov *et al.* 2007). It is generally believed that Congo red binding depends on the cross β -sheet configuration of the fibril by either ligation of the Congo red terminal groups to every fifth peptide chain or via intercalation between peptide chains. Congo red has been implicated in the treatment of amyloid diseases; evidence suggests the dye has the capacity to interfere with processes of misfolding and aggregation (Frid, Anisimov *et al.* 2007).

A second amyloid-specific dye, thioflavin T (ThT), has become widespread for solution studies, since the sensitivity of Congo red can be variable and produce false negatives (Bely and Makovitsky 2006). Furthermore, ThT can be used to detect amyloid fibrillation *in situ* (Groenning, Norrman *et al.* 2007). Binding to amyloid leads to a new excitation maximum at 450 nm and enhanced emission at 482 nm, as opposed to the 385 nm excitation and 445 nm emission of the free dye (Levine-III 1993). The mode of ThT binding to amyloid fibrils is largely unknown and various

binding sites have been proposed; binding within a cavity rather than via specific electrostatic interactions is a recent hypothesis (Groenning, Norrman *et al.* 2007). A drawback is the lack of a strict quantitative relationship between ThT fluorescence and fibril content (Bourhim, Kruzel *et al.* 2007).

Electron microscopy

By transmission electron microscopy (TEM), amyloid fibrils are straight and unbranching (Fig. 1.8b) and have a similar diameter of around 10 nm, irrespective of the precursor protein (Stromer and Serpell 2005). Fibrils usually consist of a number of protofilaments which twist to form rope-like fibrils (Tompa 2009). The environment in which amyloid assembly occurs affects the macromolecular morphology of the fibrils; variations in twists, the number of protofilaments, and the presence of sheet-like/ribbon assemblies and tubes have been documented (Stromer and Serpell 2005). Cryo-EM of A β (11-25) indicates repeating detail within otherwise unadorned fibrils; striations 4.7 Å apart and running perpendicular to the fiber axis can be observed (Stromer and Serpell 2005).

X-ray fibre diffraction

X-ray fibre diffraction shows a cross- β diffraction pattern that arises from the underlying repeating core structure of amyloid fibrils that consists mainly of β -sheet (Stromer and Serpell 2005). There are two predominant reflections (Fig. 1.8c). First, a stronger sharp reflection on the meridian corresponds to the fibre axis at 4.7 Å and is thought to arise from the H-bonding distance between strands in a β -sheet. Weaker, higher angle reflections can also be observed, including a second order of the 4.7 to 4.8 Å spacing at 2.4 Å. The observed meridional spacings can be fitted to a helical repeat distance of 115 Å which represents 24 β -strands (Fig. 1.8d) (Sunde, Serpell *et al.* 1997). The second major reflection is more diffuse, equatorial and at ~10-12 Å (Fig. 1.8c). It is likely that this corresponds to the varying distance between β -strands in a β -sheet, which depends on the side chain content of the peptide involved.

The characteristic amyloid diffraction pattern indicates that the β -strands run perpendicular to the fibre axis. Synchrotron X-ray sources result in reflections out to 2 Å, indicating that individual fibrils have highly ordered internal structures along the fibre axis (Sunde, Serpell *et al.* 1997).

Small angle X-ray scattering

Small-angle X-ray scattering (SAXS) on insulin amyloid fibrils reveals information about the organization of the coherently scattering volume inside the fibril, associated with its repeating unit (Vestergaard, Groenning *et al.* 2007). The 30 nm fibril consists of three intertwining helical protofilaments, each with an individual diameter of about 10 nm and an overall left-handed twist as observed for most amyloid fibrils (Fig. 1.8e). *Ab initio* modeling produces an elliptically-shaped repeating unit with overall dimensions of $690 \times 390 \times 160 \text{ \AA}^3$, and the hydrated volume of the repeating helical unit corresponds to an approximate molecular weight of 7200 kDa or about 1240 insulin monomers. By contrast, SAXS of α -synuclein fibrils indicates a structure in which a single filament appears to make up the mature fibre (Rambaran and Serpell 2008).

Solid-state NMR

The majority of solid-state NMR (ssNMR) work has focused on amyloid fibrils formed by various peptides homologous to regions of A β (Rambaran and Serpell 2008) and shows that the β -sheet may be arranged parallel or anti-parallel within the protofilaments, depending on the properties of the precursor polypeptide. Studies of fibrils formed from full-length A β indicate that the peptide folds into a β -bend structure that then associates with other molecules to form a parallel, in-register β -structure (Petkova, Ishii *et al.* 2002; Luhrs, Ritter *et al.* 2005). Analysis of A β (1-40) shows that residues 12-24 and 30-40 adopt β -strand conformations and form parallel β -sheets through intermolecular H-bonding (Fig. 1.8f (i)) (Petkova, Ishii *et al.* 2002). Residues 25-29 contain a bend of the peptide backbone that brings the two sheets in contact through sidechain-sidechain interactions; the only charged sidechains in the core are D23 and K28 which form saltbridges (Fig. 1.8f (ii)). Approximately the first ten residues are structurally disordered. The 3D structure of A β (1-42) shows a similar pattern, although subtle differences, such as disorder for residues 1-17 and β -sheet formation by residues 18-26 and 31-42, are apparent (Luhrs, Ritter *et al.* 2005). ssNMR data suggest that part of the polypeptide chain remains intrinsically disordered and does not become part of the ordered amyloid core.

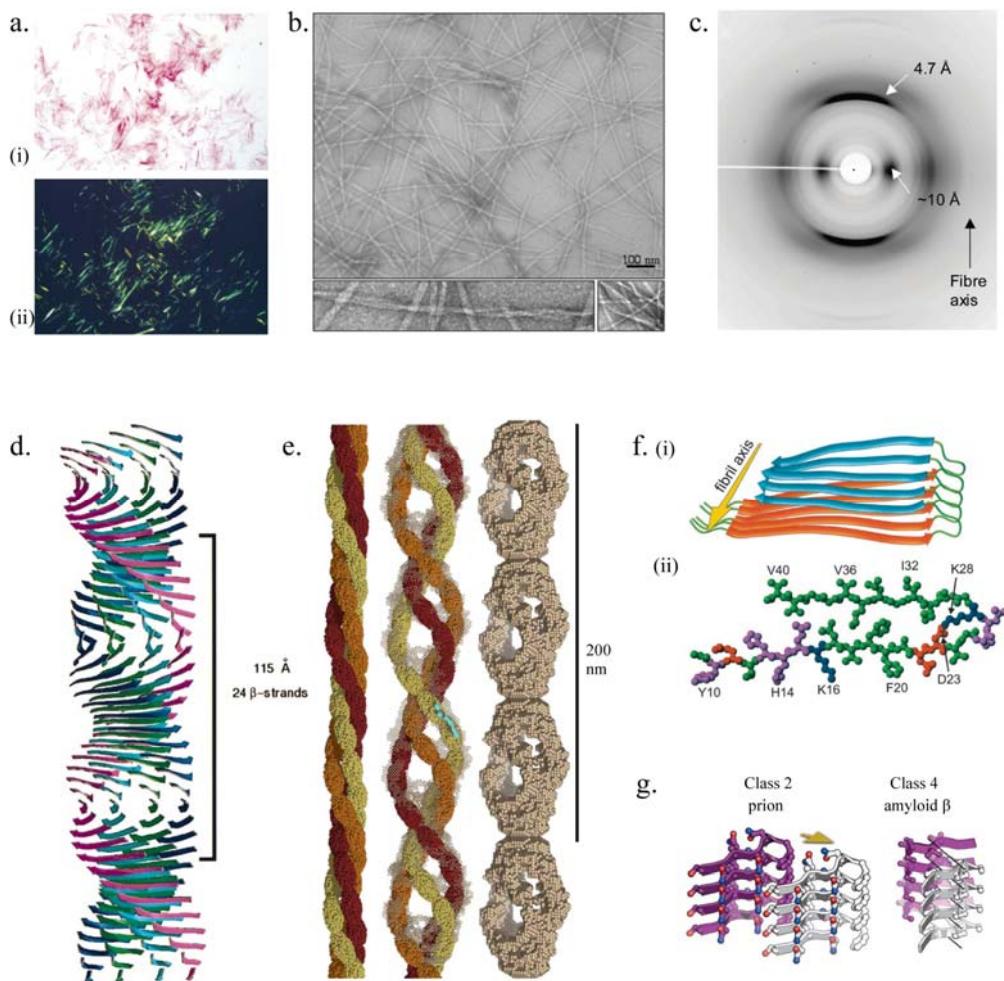


Figure 1.8 The characteristics of amyloid fibrils.

- Congo red staining under a light microscope (i), and apple-green birefringence between cross-polarisers (ii) (Rambaran and Serpell 2008).
- By TEM, amyloid- β fibrils are straight and unbranching (Stromer and Serpell 2005).
- The cross- β pattern of amyloid fibril X-ray fibre diffraction (Rambaran and Serpell 2008).
- The observed meridional spacings from X-ray fibre diffraction can be fitted to a helical repeat distance of 115 Å which represents 24 β-strands (Sunde, Serpell *et al.* 1997)
- Small angle X-ray scattering (SAXS) analysis of the 30 nm insulin fibril (Vestergaard, Groenning *et al.* 2007).
- Solid-state NMR analysis of A β (1-40) shows the formation of parallel β-sheets through intermolecular H-bonding (i) and a bend of the peptide backbone from residues 25-29 (ii) (Petkova, Ishii *et al.* 2002).
- The steric zipper model by X-ray microcrystallography. The fundamental unit of fibrils is a steric zipper formed by two tightly interdigitated β-sheets (Sawaya, Sambashivan *et al.* 2007).

Structural disorder in amyloid may be a special case of ‘fuzziness’ where disorder in the bound state serves different functions such as accommodation of destabilizing residues and mediation of secondary interactions between fibrils. The presence of disorder may be a prominent feature of amyloid fibrils despite the classical theme that amyloid fibrils are highly ordered structures (Tompa 2009).

X-ray microcrystallography

Amyloid-forming segments from a number of proteins, including A β , PrP, tau, IAPP and insulin, form microcrystals that are structurally similar to amyloid fibrils, allowing X-ray microcrystallography and analysis of amyloid structure at atomic resolution (Sawaya, Sambashivan *et al.* 2007) (Fig. 1.8g). A steric zipper model has been proposed, where the fundamental unit of fibrils is a steric zipper formed by two tightly interdigitated β -sheets. Each sheet is formed from extended strands of the segment H-bonding up and down the sheet to identical molecules all perpendicular to the axis of the fibril. The two sheets mate tightly at a completely dry interface where the residue side chains intermesh with close complementarity. In fibrils from full-length proteins, extra spine regions may remain on the periphery of the spine and could even be in a native conformation.

Three different structural characteristics are significant. First, the strands within each sheet can be parallel or antiparallel. Second, the sheets pack with either the same or different surfaces adjacent to one another and third, each sheet can be oriented parallel or antiparallel with respect to one another. Different combinations of these lead to 8 classes of amyloid structure; however, several proteins contain more than one fibril-forming segment suggesting more complicated geometries exist. For example, A β has two fibril forming segments, one stretch forms parallel sheets and the other antiparallel.

Sequence determinants of amyloid formation

A sequence-dependent propensity to form amyloid is apparent (Rambaran and Serpell 2008); the classical example is the poly-glutamine stretch responsible for amyloid formation in the huntingtin protein. Additional sequence motifs have been proposed, in particular, phenylalanine pairs occur in A β and the amyloidosis-related serum amyloid A (Makin, Atkins *et al.* 2005), and motifs of three consecutive alanine residues and AAXK are found throughout the repeat region of α -synuclein.

A number of prediction algorithms calculate the propensity of a polypeptide chain for β -aggregation and amyloid formation from amino acid sequence alone. The PASTA server (Trovato, Chiti *et al.* 2006; Trovato, Seno *et al.* 2007) predicts which portions of the sequence are more likely to stabilize the cross- β core of fibrillar aggregates. Tango (Fernandez-Escamilla, Rousseau *et al.* 2004; Reumers, Schymkowitz *et al.* 2009) uses a statistical mechanics approach to represent competition between major conformational states and predict β -aggregation propensity. Waltz (Reumers, Schymkowitz *et al.* 2009) predicts amyloidogenic regions of proteins by combining terms from amino acid sequence scoring, physical property analysis and homology modeling.

1.2.3 The pathway to amyloid formation

Amyloid formation involves a conformational transition from a natively folded, or unfolded, protein or peptide to the highly ordered cross- β structure of the fibril. One or more intermediate states have been identified resulting in a general model (Fig. 1.9). The natively folded protein or peptide somehow adopts a random coil or α -helical conformation to form a misfolded intermediate. In this form, hydrophobic regions that are normally buried in the native structure may become exposed to the aqueous environment (Soto 2003). The α -helical conformation has been suggested to play a major thermodynamic role in self-association (Abedini and Raleigh 2009); helical intermediates have been observed during amyloid formation by several designed proteins and peptides, as well as in the fibrillization of insulin and the assembly of silk *in vivo* into its cross β -structure.

The misfolded intermediate has a high tendency to aggregate and become stabilized by the formation of an oligomeric β -sheet structure (Soto 2003). This first phase is usually referred to as the lag phase or rate-limiting step, although for some proteins this phase can be extremely fast (Frieden 2007). The conformational change may be responsible for the lag but it is unclear whether this change drives association or is a result of association.

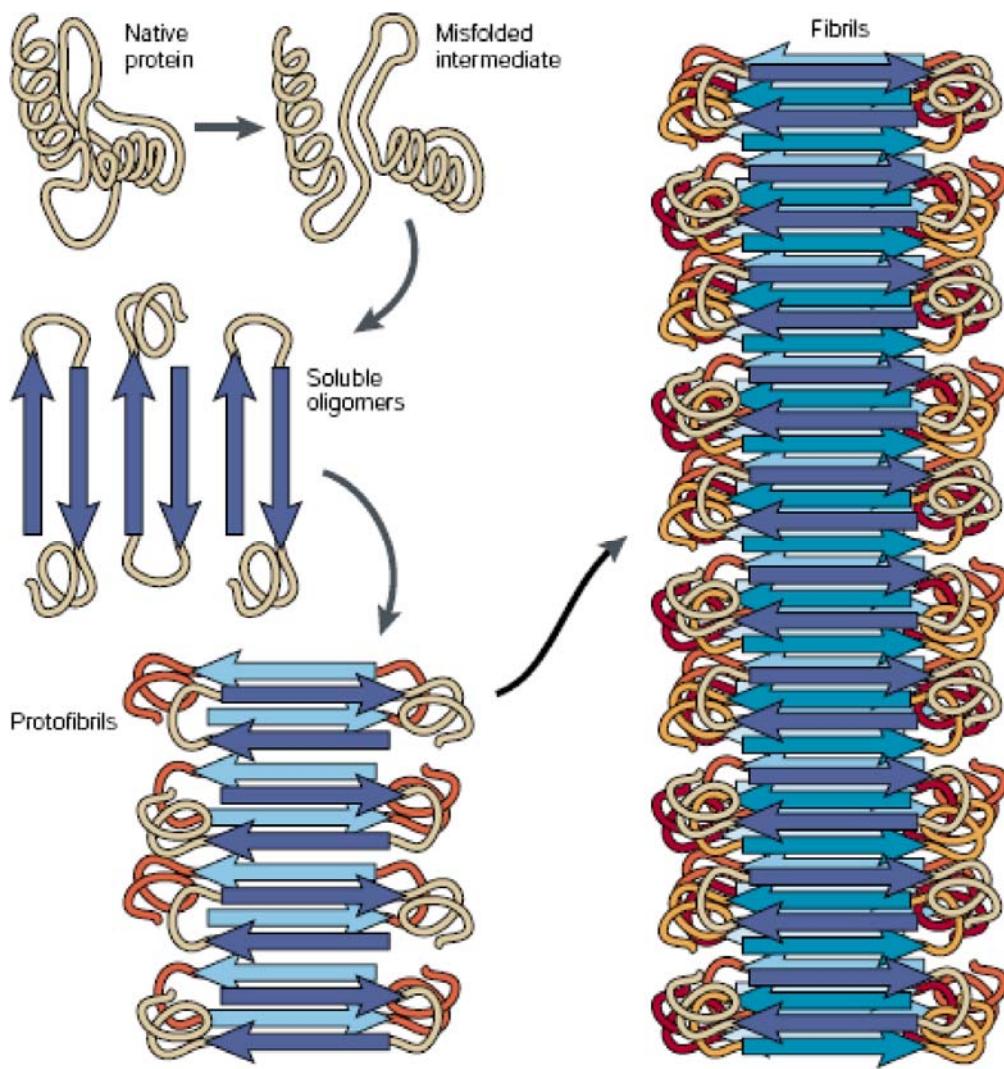


Figure 1.9 The general pathway to amyloid formation.

The natively folded protein or peptide somehow adopts a random coil or α -helical conformation to form a misfolded intermediate that has a high tendency to aggregate and become stabilised by the formation of an oligomeric β -sheet structure. The oligomers assemble both length-wise and laterally giving rise to protofibrils and cross- β amyloid-like fibrils. Figure taken from (Soto 2003).

The second phase is a molecular assembly step whereby the oligomers assemble both length-wise and laterally once critical nucleation has been adopted. This process gives rise to protofibrils and finally to cross- β amyloid-like fibrils (Soto 2003; Frieden 2007). The protofibril is proposed to be the direct precursor of the fibril; major conformational changes occur prior to protofibril formation such that the structure of the protofibril may define that of the mature fibril (Chimon, Shaibat *et al.* 2007).

The exact pathway differs by way of kinetics and the intermediate structures observed, depending on the protein involved and the conditions used to stimulate

amyloid formation (Frieden 2007). There are two major kinetic pathways described for amyloid formation, nucleation-elongation, which is the most commonly described, and isodermic (Frieden 2007). The former involves a long lag phase where monomers associate to form a critical nucleus that serves as a structure for elongation of the cross- β amyloid structure. The latter is a linear mechanism where there is no specific critical concentration of monomer and no nucleus required for elongation.

The A β peptide is perhaps the best-characterised amyloid-forming protein; A β follows a nucleation-dependent pathway where micellar aggregates propagate to form protofibrils (Etienne, Edwin *et al.* 2007). A β protofibrils are in equilibrium with low molecular weight (LMW) misfolded A β suggesting that protofibril formation is reversible. However, the protofibrils give rise to amyloid fibrils, from which dissociation does not readily occur. The process from LMW A β (random coil) to fibrils probably involves a transitory α -helical component (Walsh, Hartley *et al.* 1999). Kinetic studies using ThT fluorescence to monitor β -sheet formation suggest two structural intermediates in the A β fibrillation pathway, an aggregate devoid of amyloid-like β -sheet structure and a toxic spherical intermediate predominated by well-ordered parallel β -sheets (Chimon, Shaibat *et al.* 2007).

Amyloid formation by insulin is nucleation-dependent and has a multistage, hierarchical character (Jansen, Dzwolak *et al.* 2005) where the shape of the nuclei partly determines the fibrillar architecture of the amyloid. Globule-shaped oligomers queue to assemble into elongated protofibrillar forms capable of cooperative transformation into twisted, highly ordered superhelices of mature amyloid. An increase in temperature results in both faster kinetics and a higher degree of morphological variability with a prominence of straight rods, twisted ribbon-like structures and rod bundles.

Analysis of the amyloid formation pathway is often carried out on model proteins; amyloid formation by the 89-residue barstar protein has been studied extensively and is a classical example of how alterations in conditions affect the kinetics of amyloid formation. Barstar amyloid formation can be induced by both trifluoroethanol (TFE) and heat (Kumar and Udgaoonkar 2009) and the protofibrils from each are different in their external dimensions, internal structures and the mechanisms of formation. For example, TFE protofibrils are about half the thickness of those from heat, suggesting

that the TFE-induced structure is a β -sheet monolayer whereas that from heat is a bilayer of β -sheets. The kinetics also differ; in heat, protofibril formation is monophasic without a lag phase, whereas TFE results in sigmoidal kinetics with an initial lag phase.

The exact mechanisms behind the structural transitions that occur are not known. Aggregation can be associated with specific mutations, misprocessing phenomena, aberrant interactions with metal ions, changes in environmental conditions such as pH or temperature, or chemical modification (Stefani and Dobson 2003).

The recent steric zipper model based on X-ray crystal data of amyloid fibrils raises the hypothesis that the process of fibril formation begins with the unmasking of the zipper-forming segments allowing β -sheet stacking and interdigitation (Sawaya, Sambashivan *et al.* 2007). As a large number of intrinsically disordered proteins (IDPs) form amyloid, intrinsic disorder may have a role in transition to the amyloid state (Tompa 2009). For example, the disordered region may be involved in the interaction of the monomer with the end of the amyloid fibril or in interactions between protofibrils such as in the interfibril interface.

A number of amyloid-forming proteins, such as α -synuclein and insulin, are disulphide-bonded (Wang, Liu *et al.* 2009) and this may play a significant role in the early stages of amyloid formation (Yamamoto, Yagi *et al.* 2008; Fei and Perrett 2009; Wang, Liu *et al.* 2009). Amyloid fibril formation by hen egg white lysozyme is inhibited by TCEP in a dose-dependent manner (Wang, Liu *et al.* 2009), but only if added prior to fibrillization, suggesting that the disulphides involved are located within the core of the fibril. A unique cysteine mutant of the yeast prion protein Ure2p can greatly accelerate fibril assembly kinetics under oxidising conditions, likely due to its proximity to a potential amyloid stretch buried entirely within the hydrophobic core of mature fibrils (Fei and Perrett 2009). In contrast, disulphide bonds introduced at other positions in Ure2p cause inhibition of amyloid fibril formation. It has been suggested that disulphide bonds provide opportunities for amyloid interactions by bringing two peptides into proximity or act as conformational constraints to confine the backbone of the polypeptide chain in a manner favourable for β -sheet conversion (Fei and Perrett 2009). A number of other models have been proposed, including the polar zipper model and the domain-swapping model (Ohnishi and Takano 2004).

Structural characterisation of intermediates

Although in-depth structural analysis of mature amyloid fibrils has been carried out, the intrinsic instability, non-crystallinity and heterogeneous nature of oligomeric intermediates makes structural analysis of these species more difficult (Chimon, Shaibat *et al.* 2007). However, as prefibrillar amyloid species have been suggested to play a prominent role in disease pathogenesis, the structural characterisation of any intermediate species is important for rational inhibitor design (Abedini and Raleigh 2009).

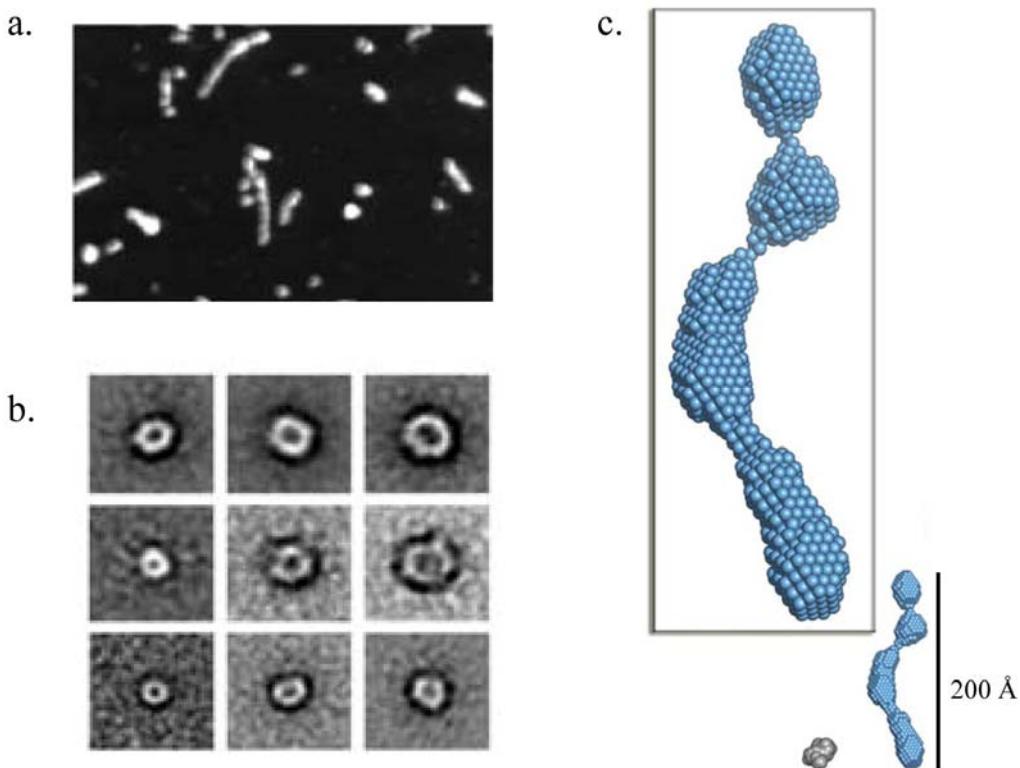


Figure 1.10 Structural characterisation of prefibrillar oligomeric intermediates.

a.-b. The ‘bead-on-a-string’ structure (a) and rings or ‘doughnuts’ (b) by high resolution electron microscopy (Stefani and Dobson 2003).

c. SAXS analysis of the insulin protofibril reveals a helical structure (Vestergaard, Groenning *et al.* 2007).

A β protofibrils are capable of binding Congo red and ThT albeit with lower efficiency, suggesting they possess some of the β -sheet properties of amyloid fibrils (Walsh,

Hartley *et al.* 1999). In contrast, a stable oligomer of the prion protein did not produce ThT fluorescence, suggesting that the cross- β structure had not been obtained or that ThT binding sites were blocked in this species (Walsh, Yau *et al.* 2009).

As observed by high-resolution electron microscopy, oligomeric amyloid species exhibit great variations in morphology. The ‘bead-on-a-string’ structure appears to predominate in the protofibril (Fig. 1.10a). However, a number of other structures including spheres, chains of spheres, and rings or ‘doughnuts’ (Fig. 1.10b) have been documented for the majority of amyloid-forming proteins (Conway, Lee *et al.* 2000; Stefani and Dobson 2003; Juárez, Taboada *et al.* 2009; Walsh, Yau *et al.* 2009); these may represent sequential intermediates in the amyloid-formation pathway (Conway, Lee *et al.* 2000). Structural heterogeneity may result from structural polymorphism within the core region or from a kind of polymorphism in transient interactions of adjacent disordered regions in protofibrils (Tompa 2009).

Prefibrillar intermediates exhibit notable mass differences. A β oligomers range from dimers to dodecamers of the 4 kDa A β monomer (Moore, Rangachari *et al.* 2009) and tetramers, octomers and 20-mers are observed during fibrillization of equine lysozyme (Malisauskas, Ostman *et al.* 2005). Insulin oligomers have an average mass of 32 kDa, corresponding to 5-6 insulin monomers (Vestergaard, Groenning *et al.* 2007) and the hydrodynamic radius of stable prion oligomers is about 30 nm (Walsh, Yau *et al.* 2009).

Circular dichroism and Fourier transform infrared (FTIR) spectroscopy on a number of proteins indicate that while an α -helical transitory component is likely, protofibrillar species have β -sheet structure (Stefani and Dobson 2003; Abedini and Raleigh 2009). ssNMR on a stable oligomeric PrP intermediate indicates extended β -strand formation by the hydrophobic core (Walsh, Yau *et al.* 2009) although the absence of ThT fluorescence suggests that the cross- β structure of the mature fibril has not been adopted. A β (1–40) has well-defined fibril-like parallel β -sheet structures but with only a metastable supramolecular packing before final fibrillization (Chimon, Shaibat *et al.* 2007). However, attenuated total-reflection FTIR analysis of A β suggests that while the β -sheet of fibrillar A β is parallel, the spectra of oligomeric A β has an antiparallel β -sheet contribution (Cerf, Sarroukh *et al.* 2009). This implies that even if fibrils and oligomers are both rich in β -sheets, they do not adopt the same

conformation and is in contrast to the hypothesis that the structure of the protofibril determines that of the mature amyloid (Chimon, Shaibat *et al.* 2007).

Molecular dynamics simulations of the prion protein, lysozyme variants, transthyretin and β_2 -microglobulin identified a novel structure in amyloid formation, the α -sheet (Daggett 2006; Milner-White, Watson *et al.* 2006). In contrast to the β -sheet, the extended α -chain conformation is defined by an alternation of residues in the right-handed and left-handed helical conformations; while each residue is locally defined as helical, the alternation leads to the formation of a sheet. This α - to β -sheet transition may occur by peptide-plane-flipping, a 180° rotation of the –CO-NH- peptide plane that has little effect on the rest of the polypeptide including side-chains (Milner-White, Watson *et al.* 2006).

The low-resolution shape of insulin amyloid oligomers was recently characterized by SAXS analysis and reveals a helical structure with a length of 200 Å and a twist of 270 Å, or three quarters of a full helical turn in the direction of the long axis (Fig. 1.10c) (Vestergaard, Groenning *et al.* 2007). By molecular weight and independent rigid-body modelling, the helical oligomer is an assembly of six monomeric units.

1.2.4 The role of amyloid in disease

Over one hundred years ago, Alois Alzheimer presented the clinical and neuropathological characteristics of the disease that was subsequently named after him; the abnormal protein deposits in the brain that define the neurodegenerative disease are now known to be central to the disease process (Goedert and Spillantini 2006). Amyloid deposition and its localisation varies between diseases (Soto 2003; Stefani and Dobson 2003). While the A β plaques of Alzheimer's disease are extracellular, tangles comprising aggregates of hyperphosphorylated tau protein are located in the cytoplasm of degenerating neurons. In Parkinson's disease, the neuronal cytoplasm contains α -synuclein aggregates called Lewy bodies, and the polyglutamine-rich huntingtin protein forms intranuclear deposits in Huntington's disease (Soto 2003) (Fig. 1.11).

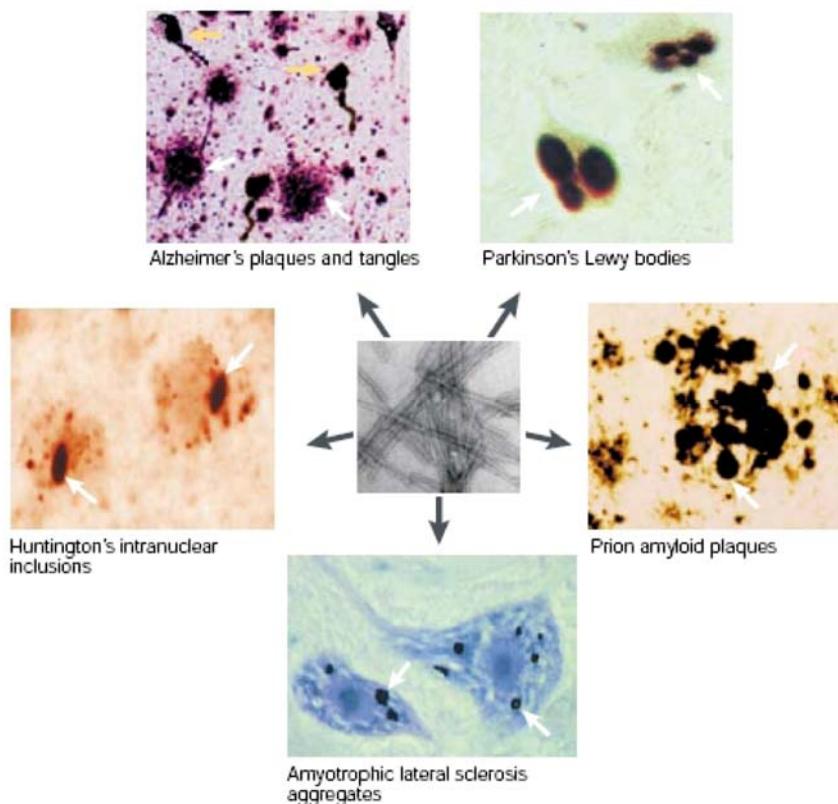


Figure 1.11 Variation in amyloid deposition and localisation between diseases.

Figure taken from (Soto 2003).

Despite observations that amyloid is a typical feature of degenerative diseases and initial research suggesting that the neurotoxicity of A β is dependent on the formation of mature amyloid fibrils (Lorenzo and Yankner 1994), a direct correlation between amyloid formation and disease severity has not been shown (Soto 2003; Malisauskas, Ostman *et al.* 2005). Post-mortem studies show a poor correlation between the load of amyloid deposits in the brain of Alzheimer's disease patients and the severity of clinical symptoms. Furthermore, there is a mixture of different types of aggregates in some diseases and protein deposits appear in clinically normal people. However, mutations in genes that encode the protein components of fibrillar aggregates (such as amyloid- β and α -synuclein) are genetically associated with familial forms of amyloid diseases; these usually have an earlier onset, greater severity than sporadic cases and are also associated with a greater load of protein aggregates (Soto 2003; Meredith 2005). In Huntington's disease, age of onset and disease severity is directly correlated with the length of the polyglutamine tract region in the huntingtin protein (Meredith

2005) and the infectious agent responsible for prion disorders is the misfolded form of the prion protein (Soto 2003).

The popular hypothesis presently is that the cytotoxic species is not mature fibrils but prefibrillar intermediates (Stefani and Dobson 2003). Analysis of the biological activity of the A β protofibril suggested that like mature fibrils, protofibrils were able to affect the normal metabolism of cultured neurons (Walsh, Hartley *et al.* 1999). Subsequent research on A β (Demuro, Mina *et al.* 2005; Manelli, Bulfinch *et al.* 2007) and other disease-associated proteins such as lysozyme (Demuro, Mina *et al.* 2005; Malisauskas, Ostman *et al.* 2005), α -synuclein (Volles and Lansbury 2003), islet amyloid polypeptide (IAPP) and prion (Demuro, Mina *et al.* 2005), suggested that oligomeric intermediates may be the most toxic amyloid species. In addition, a correlation between oligomer size and toxicity is apparent for some proteins such as equine lysozyme (Malisauskas, Ostman *et al.* 2005).

Early onset forms of amyloid diseases are associated with mutations that increase propensity for β -aggregation (Conway, Lee *et al.* 2000; Stefani and Dobson 2003; Kim and Hecht 2008). The acceleration of oligomerization rather than fibrillization is a shared property of both α -synuclein mutations linked to early-onset Parkinson's disease (Conway, Lee *et al.* 2000) and accelerating A β fibrillization decreases both oligomer levels and functional deficits in Alzheimer's mouse models (Cheng, Scearce-Levie *et al.* 2007).

In addition, amyloid intermediates of the non-disease associated proteins HypF-N (bacterial) and bovine phosphatidyl-inositol-3'-kinase (PI3-SH3) are cytotoxic when added to the medium of cultured cells (Bucciantini, Giannoni *et al.* 2002) or injected into rat brains (Baglioni, Casamenti *et al.* 2006). In contrast, monomers and fibrils had no activity. The toxicity of non-disease-associated amyloid aggregates implies a general phenomenon, and that the pathogenicity of protein deposition diseases may be related primarily to the structural nature of the aggregates rather than to the specific sequences of the proteins from which they arise (Bucciantini, Giannoni *et al.* 2002).

Mechanisms of cytotoxicity

Much research is currently focused around understanding how cytotoxic effects are mediated. Different amyloid oligomers share a common structure and are generically toxic to cells; however, some amyloids arise from cytosolic proteins, while others are derived from secretory or extracellular proteins. The most obvious target that is accessible to both the cytosolic and extracellular compartments is the plasma membrane, and a growing body of evidence suggests that membrane permeabilization may represent the common mechanism of pathogenesis of amyloid-related degenerative disease (Glabe 2006).

$\text{A}\beta$, α -synuclein, polyglutamine and IAPP have been shown to form discrete pores or single channels in membranes leading to formation of the ‘channel hypothesis’ (Stefani and Dobson 2003; Glabe 2006), that oligomeric pores may insert irreversibly into membranes as discrete channels. An antibody that recognizes higher molecular-weight prefibrillar oligomers of a wide variety of amyloidogenic peptides and proteins, but not lower molecular weight oligomers or fibrils, also recognizes the pore-forming bacterial toxin α -haemolysin, and human perforin (Cerf, Sarroukh *et al.* 2009). These proteins are known to form oligomeric pores in their toxic conformation with each monomer supplying several β -strands that self-assemble and fold into a β -barrel. A model for β -barrel formation by the $\text{A}\beta$ oligomer has been proposed (Fig. 1.12a) (Cerf, Sarroukh *et al.* 2009).

In contrast, other evidence shows that amyloid oligomers can disrupt membranes without the formation of discrete pores (Glabe 2006). Although the exact mechanisms remain unknown, it is widely agreed that membrane disruption by oligomeric species has an important role in pathogenesis. An increase in membrane permeabilization is associated with increases in calcium concentration and the production of reaction oxygen species (ROS) (Demuro, Mina *et al.* 2005; Glabe 2006). In culture, cell death induced by HypF-N is preceded by entrance into the cytoplasm and an early rise in reactive oxygen species (ROS) and free Ca^{2+} levels. (Bucciantini, Calloni *et al.* 2004).

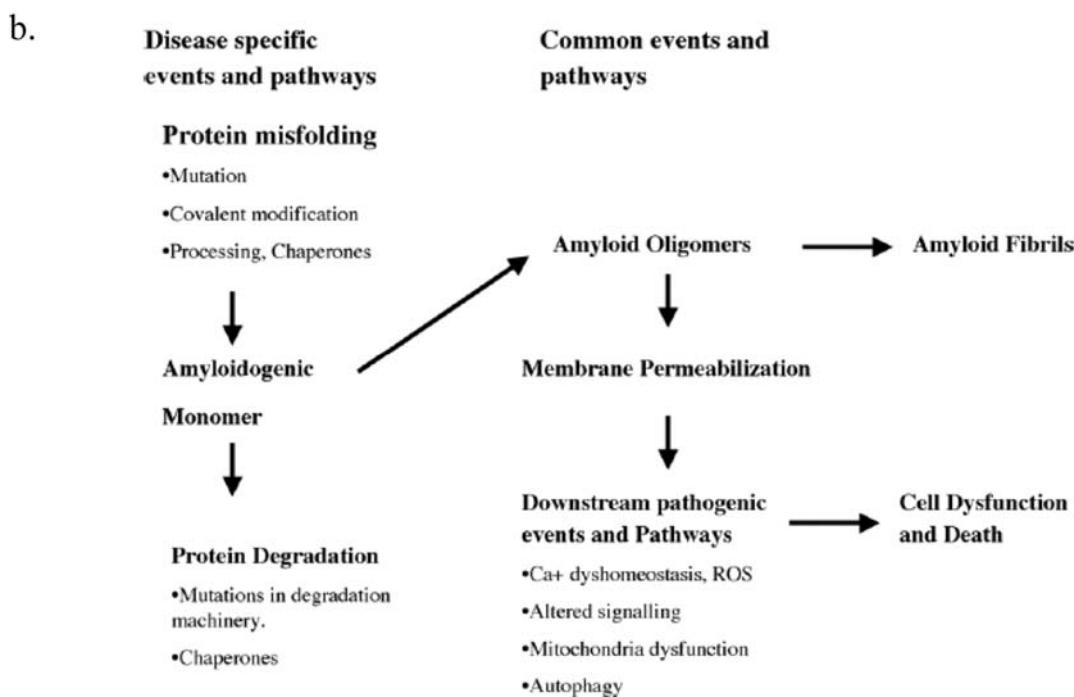
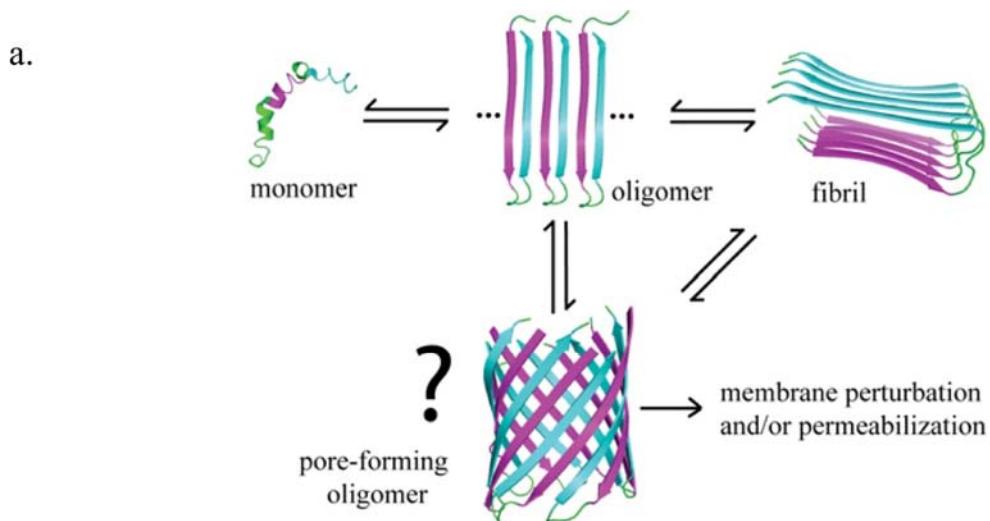


Figure 1.12 Models for the mechanism of amyloid cytotoxicity.

- a. The model for β -barrel formation by the A β oligomer by ssNMR (Cerf, Sarroukh *et al.* 2009).
 b. An overall scheme for the role of amyloid in disease (Glabe 2006).

The dysregulation of Ca^{2+} ions plays a crucial role because of their strong trans-membrane concentration gradient and involvement in cell dysfunction and death (Demuro, Mina *et al.* 2005); a number of studies have suggested a close relationship between Alzheimer's, Parkinsons and the prion diseases and the dysregulation of calcium homeostasis (Stefani and Dobson 2003). ROS may be generated in the very early stages of protein aggregation when protofibrils or soluble oligomers are present, but in the absence of mature amyloid fibrils (Tabner, El-Agnaf *et al.* 2005), and may represent a fundamental molecular mechanism underlying the pathogenesis of oxidative damage. A scheme for the role of amyloid in disease is outlined in Figure 1.12b; membrane permeabilization may also result in altered signaling, mitochondrial dysfunction and autophagy (Glabe 2006).

1.2.5 Functional amyloid

A number of functional amyloid structures exist. The classic example is spider silk, which shares a number of structural characteristics with disease amyloids. Comparison of recombinant *Crysopa flava* egg-stalk silk with amyloid fibrils of the yeast prion protein Sup35p reveals that spider silk contains a cross- β structure by ThT dye binding and X-ray fibre diffraction (Slotta, Hess *et al.* 2007). The latter indicates distinct differences in the overall structure however, with an equatorial reflection at 5.3 Å rather than the amyloid-characteristic 10 Å reflection.

Other functional amyloids include the yeast prions, curli and Pme17 (Rambaran and Serpell 2008). The yeast prions have been suggested to form cytoplasmic fibrillar assemblies associated with a method of heritable information transfer. Curli is a functional fibre on the surface of bacteria such as *E.coli*. Pme17 is a mammalian protein involved in the templating of melanin biosynthesis.

Amyloid has been described as having similar strength to steel (Rambaran and Serpell 2008), a property that has contributed to its use in nanotubular scaffolding for bionanotechnology. Amyloid fibrils can be functionalized by assembling as fusion proteins or used as templates for binding to metals (Rambaran and Serpell 2008). For example, a very short peptide of two phenylalanine residues forms amyloid-like nanotubes that can be functionalized using ionic silver in the centre of the nanotube to make wires ~ 20 nm in diameter.

Amyloid fibrils have also been suggested as a promising means of drug delivery, by acting as stable depots for the long-term, controlled release of an active peptide drug (Maji, Schubert *et al.* 2008).

1.2.6 Sporadic inclusion body myositis and myostatin

Sporadic inclusion body myositis

Sporadic inclusion body myositis (sIBM) is the most common progressive muscle wastage disease of the elderly and leads to pronounced muscle fibre atrophy and weakness (Askanas and Engel 2008). It is the only muscle disease in which accumulation of A β in abnormal muscle fibres appears to play a key pathogenic role (Vattemi, Nogalska *et al.* 2009); the development of sIBM in cellular and animal models involves an increase of the A β precursor protein (A β PP) and the accumulation of A β (Vattemi, Nogalska *et al.* 2009). Other amyloidosis-associated proteins such as α -synuclein and tau are also observed.

Both degeneration and mononuclear-cell inflammation are components of the sIBM pathology, but how each relates to the pathogenesis is unsettled (Askanas and Engel 2008). sIBM muscle fibers bear intriguing similarities to the brains of Alzheimer's and Parkinson's disease patients, including cellular aging, mitochondrial abnormalities, oxidative stress, proteasome inhibition and multiprotein aggregates (Askanas and Engel 2008).

ER-stress and the UPR are evident (Nogalska, Wojcik *et al.* 2007); an upregulation of the ER stress marker NF- κ B induces the transcription of a number of genes involved in inflammation.

A role for myostatin?

A role for myostatin in the pathogenesis of sIBM was suggested when myostatin, or the myostatin precursor protein (MstnPP), was shown to be accumulated within muscle fibers of sIBM patients and colocalised with A β (Wojcik, Engel *et al.* 2005). Double-label gold immune electron microscopy showed an association of both myostatin and A β with 6-10 nm amyloid-like fibrils as well as floccular and amorphous material. Immunoblots showed increased levels of both myostatin and MstnPP and immunoprecipitation studies suggested a physical association of A β and MstnPP. A subsequent study showed similar results in an inclusion body myositis cell

culture model (Wojcik, Nogalska *et al.* 2007). Furthermore, overexpression of A β PP and A β increased the accumulation of MstnPP.

The same group published subsequent evidence that MstnPP expression was increased following ER stress in cultured human muscle fibres (Nogalska, Wojcik *et al.* 2007). Well-differentiated muscle fibres were treated with tunicamycin or thapsigargin, compounds that inhibit protein glycosylation and disrupt ER calcium levels respectively; the induction of ER stress led to an increase in MstnPP mRNA and protein. MstnPP upregulation may have occurred via activation of the ER stress transcription factor NF- κ B (Nogalska, Wojcik *et al.* 2007). MstnPP has an NF- κ B consensus site in its promoter, MstnPP upregulation was not observed in the presence of NF- κ B inhibitors (Nogalska, Wojcik *et al.* 2007), and treatment with the antioxidant resveratrol reduced NF- κ B activation and myostatin expression (Nogalska, D'Agostino *et al.* 2008).

How might myostatin/MstnPP be involved in the pathogenesis of sIBM?

One possibility is that an increase in the level of circulating myostatin growth factor, due to increased expression and processing of MstnPP, contributes to the prominent muscle-fiber atrophy (Zimmers, Davies *et al.* 2002). An increase in myostatin growth factor may also help to explain a diminished potential of satellite cells to proliferate (Karpati and O'Ferrall 2009).

Myostatin may also be acting intracellularly; one suggestion is that post-translational modification of MstnPP, perhaps through association with A β PP/A β , lessens its degradation and traffic, resulting in MstnPP accumulation and an increase in ER stress (Wojcik, Nogalska *et al.* 2007). The authors propose that the most likely scenario is A β PP/A β accumulation preceding that of MstnPP. However, MstnPP may also be detrimental by itself, for example, by negatively influencing A β and other proteins to cause their oligomerization (Wojcik, Engel *et al.* 2005). The upregulation of MstnPP mRNA following ER stress suggests that additional mechanisms may be involved (Nogalska, Wojcik *et al.* 2007). A greater understanding of the role of MstnPP in sIBM may provide insight into causes of this sporadic form of muscle wastage.

Summary

Myostatin is a TGF- β family member with an important role in the regulation of muscle growth and development and is a potential target in the treatment of muscle wastage diseases. A number of naturally-occurring myostatin null mutations exist, such as that observed in the Piedmontese cattle breed. The myostatin precursor protein consists of an N-terminal propeptide domain which plays important roles in the folding, secretion and regulation of the active growth factor, within both the precursor and the secreted latent complex. In addition, aggregated myostatin precursor protein may be involved in sporadic inclusion body myositis, a debilitating muscle wastage disease with a predominant amyloid component. Knowledge of the structure and function of the myostatin precursor protein and latent complex is therefore extremely important for an understanding of the roles of these species in disease. However, little information in this area has been published to date.

This thesis presents biochemical and biophysical analyses of the structure and function of both the WT and Piedmontese mutant myostatin precursors and latent complexes, and addresses the capacity of the myostatin precursor protein to form amyloid *in vitro*. These novel results contribute important information for an understanding of the role of the myostatin precursor protein in muscle wastage disease.

2 Materials and Methods

2.1 Water, chemicals and media

2.1.1 Water and chemicals

All water used for this work was milliQ, purified over two ion-exchange filters and two organic filters in a Barnstead NANOpure II system (Thermoscientific). Chemicals used were the highest grade available.

2.1.2 Media for *E.coli* growth

Luria broth (LB) media (Invitrogen) was made up in milliQ at 25 g/L and autoclaved at 121 °C for 15 minutes before use. To make LB agar, 1 % bacterial agar was added to the broth before autoclaving. Ampicillin was added to LB growth media at a concentration of 100 µg/mL after autoclaving to maintain plasmids in *E.coli*, from a filtered (0.2 µm) 100 mg/mL stock stored at -20 °C. Hereafter, LB media containing ampicillan will be referred to as LB-amp.

2.2 Electrophoresis methods

2.2.1 Agarose gel electrophoresis

Agarose gels were made to 1% agarose (w/v) in TAE buffer (40 mM Tris acetate, 1 mM EDTA). DNA samples were mixed at a 1:1 ratio with sucrose loading buffer (4% sucrose (w/v), 0.1% bromophenol blue) and gels were run at 100 V. Gels were stained in ethidium bromide (5 µg/mL in TAE) and visualized with the GelDoc XR gel documentation system and software (BioRad).

2.2.2 SDS-PAGE

SDS-PAGE was carried out according to Laemmli (Laemmli 1970). Samples were boiled in 5x SDS-PAGE sample buffer (0.125 M Tris-HCl pH 6.8, 35% glycerol, 2.5% SDS, 0.025% bromophenol blue) in the presence or absence of 2 M β-mercaptoethanol (βME) as a reducing agent and run at 150-200 V. All gels for visualisation of MstnPP were 12%; where observation of the myostatin growth factor was required, 15% gels were used. Gels were made in-house. Gels were stained with Coomassie Brilliant Blue R-250 (0.0625% w/v in 40% EtOH, 10% acetic acid),

destained in 20% EtOH, 10% acetic acid, and visualized using the GelDoc XR gel documentation system and software (BioRad).

2.2.3 Western blotting

Protein samples were separated on a 12% SDS-PAGE gel at 150 V and transferred to a nitrocellulose membrane overnight at 15 V (approximately 40-150 mA) in cooled Tris/glycine transfer buffer. For analysis of high molecular weight samples, transfer was continued the following day at 450 mA for one hour. Efficient transfer was confirmed by Ponceau S staining. Membranes were blocked with 5% non-fat milk in TBS-Tween (TBS-T; 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20) for one hour at room temperature and subsequently incubated in primary antibody diluted in 1% non-fat milk TBS-T, for 90 minutes. Following washing with TBS-T (5 x 5 minutes), membranes were incubated in horse radish peroxidase-conjugated secondary antibody diluted in 1 % non-fat milk, TBS-T for 90 minutes. All steps were performed at room temperature. After washing with TBS-T (5 x 5 minutes), proteins were visualised with the SuperSignal West Pico Chemiluminescent Substrate (Pierce) according to manufacturer's instructions.

Antibody concentrations depended on the detection method used. Details of antibodies and dilutions are summarised in Table 2.1.

Detection Method	Rabbit anti-myostatin (Chemicon International) - Polyclonal	Anti-rabbit HRP-conjugate (Jackson ImmunoResearch) - Polyclonal	Mouse anti-His (Sigma) - Clone HIS-1	Anti-mouse HRP-conjugate (Sigma) - Polyclonal
X-ray film/developer	1/10,000	1/50,000	1/5000	1/10,000
Intelligent Dark Box	1/7,500	1/12,500	1/2,500	1/5,000

Table 2.1 Summary of antibodies and detection methods used for Western blotting.

2.3 Cloning

All myostatin constructs were produced from the full-length human myostatin cDNA (1-375) in the pBlueScript vector (from Se-Jin Lee, John Hopkins University School of Medicine, Baltimore, USA).

2.3.1 PCR

The myostatin precursor protein (MstnPP) construct was engineered to omit the N-terminal signal sequence. Because the published data does not agree with the site for the start of the propeptide region, three constructs were made initially, to begin at amino acids 19 (Bentsen, Nielsen *et al.* 2004), 21 (Jin, Dunn *et al.* 2004) and 25 (Lee 2004) (F19, F21 and F25 respectively). Forwards and reverse primers (Sigma-Genosys) contained *BamHI* and *XhoI* restriction sites (underlined below) respectively, for cloning into the pProEX-Htb expression vector (Invitrogen, Appendix 2, Fig. 12.6). The forwards primers were as follows:

F19: ggt gga tcc ggt cca gtg gat cta aat gag

F21: ggt gga tcc gtg gat cta aat gag aac agt

F25: ggt gga tcc gag aac agt gag caa aaa gaa

The same reverse primer was used:

B375: ggt ctc gag tga gca ccc aca gcg gtc tac

PCR was carried out using 2.5 U of Roche PWO polymerase per 50 µL reaction. Each reaction contained 1 ng DNA template, 20 pmol of each primer and 0.5 mM dNTPs in a MgSO₄-containing buffer. PCR incubation involved 35 cycles with denaturation at 95 °C, an annealing temperature of 58 °C and extension at 72 °C. PCR products were confirmed on agarose gel electrophoresis and purified using the Roche PCR purification kit.

2.3.2 Restriction digest

Restriction digests were performed on all cDNA constructs (10 µL per digest) and the pProEX-Htb vector (5 µL per digest), with a total digest volume of 20 µL. 1 U of *BamHI* and *XhoI* (Roche) enzymes and 2 µL of 10 x Roche Buffer B were used, with digests carried out at 37 °C for one hour. cDNA digests were purified using the Roche

PCR purification kit. The entire vector reaction was electrophoresed on a 1% agarose gel and the vector backbone was excised under high-wavelength UV light, with purification by the Roche PCR purification kit, adapted for agarose gel products. Single- and no-enzyme digests were conducted as controls and analysed with agarose gel electrophoresis. DNA concentrations were established using NanoDrop technology (Thermo Scientific).

2.3.3 Ligation

A 3:1 molar ratio of insert to vector backbone was used per ligation. 1 U of T4 DNA ligase in T4 ligase ATP-buffer (Roche) was added in a total volume of 10 µL per reaction. A negative ligation control without insert was also performed. Ligations were incubated overnight at 4 °C.

2.3.4 Transformation

The entire ligation reaction was added to 50 µL competent Top 10 *E.coli* cells and placed on ice for 15 minutes. The incubations were heat-shocked at 42 °C for 1 minute and placed back on ice for 2 minutes before plating on LB-amp plates for overnight incubation at 37 °C.

2.3.5 Colony PCR

Colony PCR was used to determine recombinant colonies using pProEX forward and reverse sequencing primers. In brief, 9 colonies per plate were sampled by swirling a toothpick that had been touched to a colony in 25 µL buffer (50 mM Tris-HCl, pH 8.5, 150 mM NaCl). Toothpicks were subsequently used to scratch a grid on an LB-amp plate for growth of positive colonies. Solutions were boiled for 10 minutes then centrifuged at 17,000 x g, room temperature for 5 minutes and 5 µL from the bottom of each eppendorf was added to 20 µL PCR reaction mix containing 0.5 U Taq polymerase (Roche). Positive colonies were identified using agarose gel electrophoresis (Fig. 2.1) and used to inoculate overnight cultures for plasmid preparations (Roche Plasmid Purification kit).

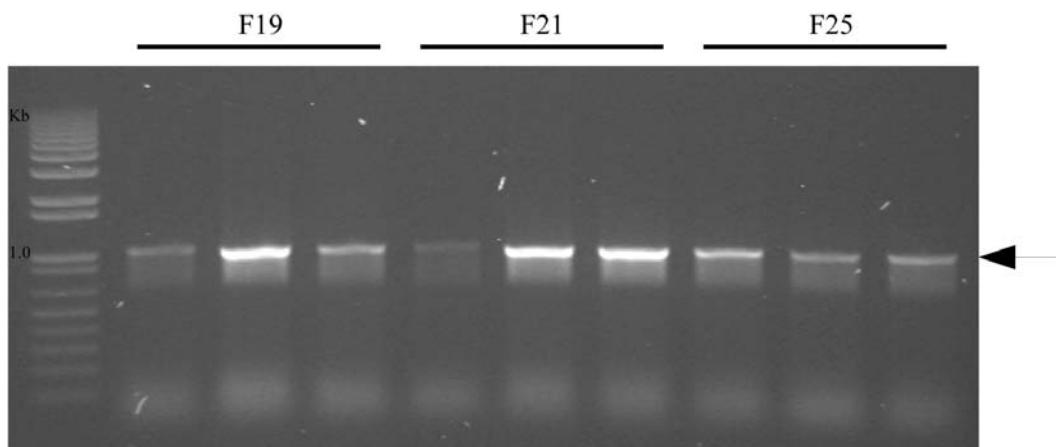


Figure 2.1 Colony PCR of cells carrying pProEX-MstnF19/F21/F25 constructs.

Arrow indicates a positive result. A band just above 1.0 Kb represents myostatin precursor protein cDNA.

2.3.6 Sequencing

Traditional Sanger sequencing of all myostatin constructs was performed by The Allan Wilson Centre Genome Service, Massey University, Palmerston North, using ABI3730 capillary instruments. Primers were pProEX-Htb-specific forward and reverse sequencing primers. Sequences were viewed using 4Peaks software (Mac OS X) and confirmed by NCBI BLAST analysis.

2.3.7 Site-directed mutagenesis

The C313Y mutation was constructed using the splice-overlap-extension (SOE) method of site-directed mutagenesis (Fig. 2.2). Reverse complementary primers (Sigma-Genosys) containing the C313Y mutation (G938A by nucleic acid sequence) were designed and used in conjunction with the F21 and B375 forwards and reverse primers in two separate PCR reactions (Fig. 2.2a, blue and black arrows). The purified products, now containing the mutation (Roche PCR purification kit) were used in a second PCR reaction with F21 and B375 primers (Fig. 2.2b) to generate the full-length mutant construct (Fig. 2.2c). Following restriction digest, ligation, transformation and colony PCR, sequencing ensured the mutation had been created.

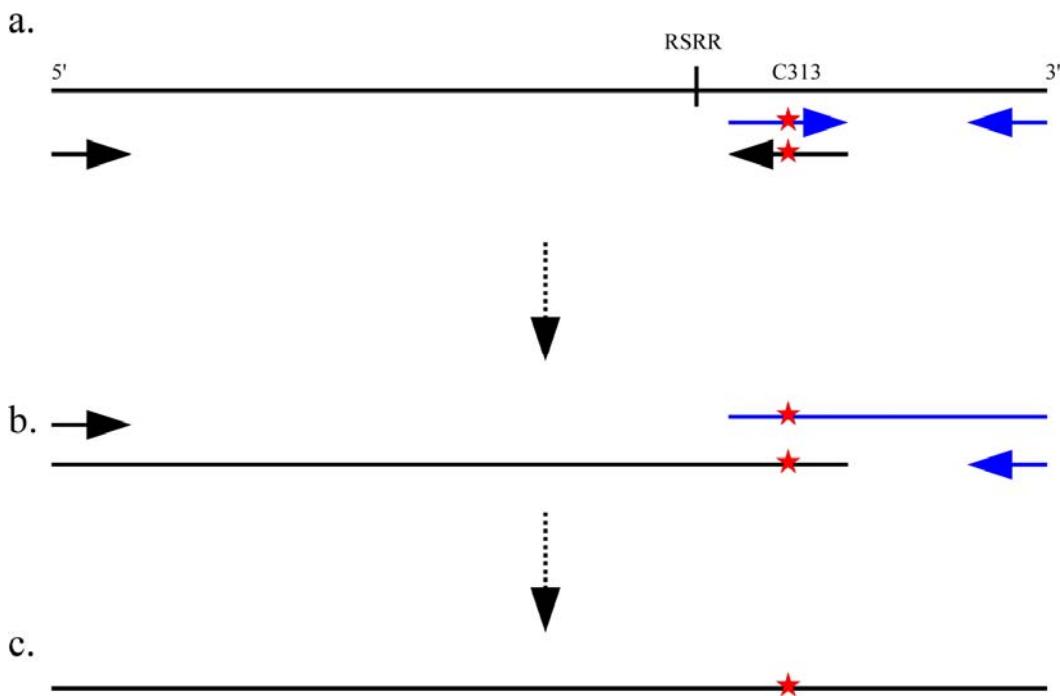


Figure 2.2 Splice-overlap extension (SOE) mutagenesis protocol.

- Reverse complementary primers containing the Piedmontese mutation (red star) were used in conjunction with F21 (black arrow) and B375 (blue arrow) in two separate PCR reactions.
- The purified products of a. were used in a second PCR reaction with F21 and B375.
- Purification of the PCR reaction product from b. generates the full-length mutant construct.

Mutagenesis primers are shown below; the mutation is highlighted in red:

SDM_F	tac tgc tct gga gag ta t gaa ttt gta ttt
SDM_R	aaa tac aaa ttc at a ctc cct aga gca gta

2.4 Protein Expression and Purification

2.4.1 Expression of recombinant MstnPP

E.coli BL21 (DE3) cells were transformed with approximately 50 ng of plasmid DNA using a shortened transformation protocol (Pope and Kent 1996). In brief, cells were incubated with DNA on ice for 5 minutes then plated directly onto LB-amp plates and incubated overnight at 37 °C. Transformed BL21 colonies containing pPro-MstnPP

were grown in a starter culture of 20 mL LB-amp at 37 °C overnight. The starter culture was used to inoculate 2 L LB-amp and cells were grown at 25 °C to an OD₆₀₀ of 0.8. Expression was induced by the addition of 0.1 mM IPTG and cells were incubated overnight (16-18 hours). Cells were collected by centrifugation at 4,000 x g for 30 min at 4 °C and stored at -20 °C.

2.4.2 Inclusion body isolation, solubilisation and refolding of MstnPP

The frozen cell pellet from each 2 L culture was thawed at room temperature and resuspended in a final volume of 70 mL 20 mM Tris-HCl pH 8.5. Cells were lysed by two passages through a French Press at 5,000 psi and the inclusion body pellet was collected by centrifugation at 30,000 x g at 4 °C for 30 min. The inclusion bodies were washed with two 20 mL volumes of wash buffer (50 mM Tris HCl pH 8.5, 0.5 M NaCl, 10% glycerol, 0.5% Triton-X 100) and two washes with the same volume of milliQ H₂O (Fig. 2.3). Pellets were collected after each wash by centrifugation at 30,000 x g for 10 min at 4 °C.

Solubilisation and refolding were carried out using methods published previously for the zebra-fish myostatin precursor protein (Funkenstein and Rebhan 2007) with some modifications. Each inclusion body pellet from 2 L of cell culture was solubilised in 8 mL 6 M guanidine hydrochloride (GndHCl), 50 mM Tris-HCl pH 8.5, 1 mM EDTA, 0.1 mM DTT, at room temperature, with shaking for 5 hours. Insoluble proteins and cell debris were removed by centrifugation at 17,000 x g, 4 °C for 30 min. The supernatant was acidified to pH ~5.0 by adding a few drops of concentrated HCl and dialysed overnight against 6 M GndHCl, 50 mM MES pH 5.0, 1 mM EDTA to remove DTT. After dialysis, the sample was centrifuged again as above. The solubilised protein (8 mL) was rapidly diluted in 150 mL of freshly prepared refolding buffer, consisting of 50 mM Tris-HCl pH 8.5, 1 M NaCl, 0.5 M L-arginine, 5 mM EDTA, 5 mM reduced glutathione and 1 mM oxidised glutathione. The pH was re-adjusted to 8.5 before addition of protein. The protein was left to refold without agitation for 6 days at 4 °C then centrifuged at 30,000 x g, 4 °C for 30 min before overnight dialysis into 20 mM Tris-HCl pH 8.5. Soluble protein was harvested as supernatant after a second centrifugation and filtered (0.45 µM).

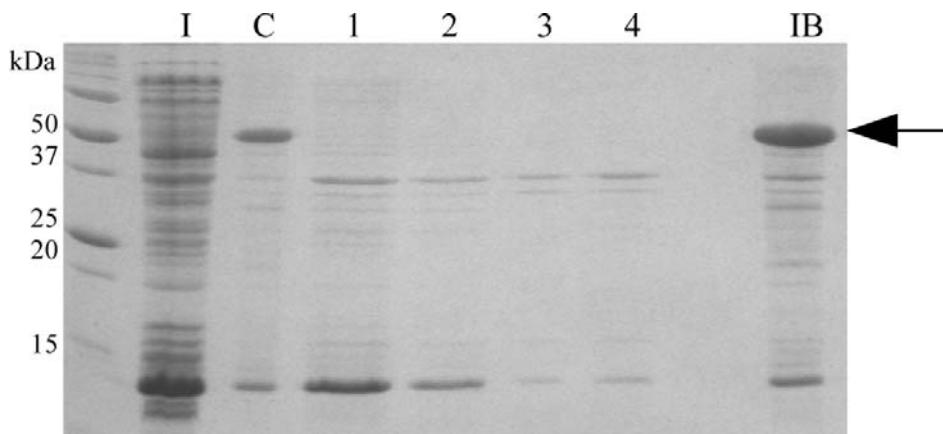


Figure 2.3 Purification of inclusion bodies containing MstnPP.

Lanes are as follows: I, induced bacterial culture; C, cell pellet following lysis; 1-4, inclusion body washes (supernatant fractions); IB, purified inclusion bodies; Arrow, MstnPP.

2.4.3 Periplasmic extraction

Periplasmic extraction was necessary for production of MstnF21 using the pTUM4 chaperone co-expression system (Schlapschy, Grimm *et al.* 2006). In brief, the cell pellet from 2 L of culture was resuspended in 20 mL ice-cold 500 mM sucrose, 1 mM EDTA, 100 mM Tris-HCl, pH 8.0 containing 100 µg/mL lysozyme and incubated on ice for 30 minutes. Spheroblasts were removed by centrifugation at 30,000 x g, 4 °C for 30 minutes and the supernatant was collected as the periplasmic fraction.

2.4.4 Purification of refolded MstnPP dimer and soluble aggregates

Purification of the *in vitro* refolded MstnPP dimer and misfolded soluble aggregates was performed by heparin affinity followed by gel filtration chromatography.

Heparin affinity chromatography

The filtered protein solution was loaded onto a 5 mL HiTrap Heparin HP column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.5, at a flow rate of 1 mL/min using the AKTA Explorer system (GE Healthcare). The flow-through containing unbound protein was discarded and a stepwise NaCl elution was used to elute the correctly folded dimer at 200 mM NaCl (H1) and the majority of soluble aggregates at 600 mM NaCl (H2). Elution was monitored by absorbance at 280 nm and differently folded forms were identified using non-reducing SDS-PAGE.

Size exclusion chromatography

Further purification of the MstnPP dimer and soluble aggregates was achieved using gel filtration chromatography. Gel filtration columns were calibrated using both commercially available standards and standards prepared in-house. Protein standards at 1 mg/mL were applied to a Superdex S200 10/300 HR column equilibrated with 50 mM Tris-HCl pH 8.5, 150 mM NaCl or 50 mM HEPES pH 7.5, 0.15 M NaCl for furin-digested samples. Two separate columns were used for this work and calibrations were performed on both (Appendix 3, Fig. 12.7).

For protein purification, the 200 mM NaCl peak from heparin chromatography was filtered (0.22 µM), concentrated to 350 µL and applied to a Superdex S200 10/300 HR column (GE Healthcare) equilibrated with 50 mM Tris-HCl, pH 8.5, 150 mM NaCl, at a flow rate of 0.4 mL/min. The 600 mM NaCl peak was buffer-exchanged into 50 mM Tris-HCl, pH 8.5, 150 mM NaCl before concentration and gel filtration. Elution was monitored by absorbance at 280 nm and purity was established using reducing and non-reducing SDS-PAGE.

If the dimer was to be cleaved with furin convertase following gel filtration, the heparin fractions were first buffer-exchanged into the furin buffer, 50 mM HEPES pH 7.5, 0.15 M NaCl, before gel filtration using a column equilibrated with the same buffer.

2.4.5 Ni-NTA affinity chromatography

A nickel-conjugated 5 mL HisTrap HP column (GE Healthcare) was used for the purification of the pTUM4-expressed His-F21 construct and initial purification of refolded His-tagged MstnPP. Columns were equilibrated in a 50 mM Tris-HCl, pH 8.5, 500 mM NaCl buffer containing 10 mM imidazole to reduce non-specific binding. A step-wise imidazole gradient from 20 – 500 mM imidazole was used to elute the His-tagged precursor at a flow rate of 1 mL/min. Fractions were analysed by reducing SDS-PAGE.

2.4.6 Furin cleavage

Purified MstnPP dimer, in a HEPES pH 7.5, 150 mM NaCl buffer, was concentrated to 10 mg/mL and furin cleavage buffer (50 mM HEPES pH7.5 + 1 mM CaCl₂) was added to a final volume of 250 µL per 1 mg of protein. Human furin convertase (Sigma, 2 U/µL) was added at a ratio of 1 µL furin to 100 µg protein. The reaction

was incubated undisturbed at 30 °C for 64 hours and subsequently centrifuged at 17,000 x g, 4 °C for 5 minutes to remove precipitated protein.

2.5 Amyloid formation

2.5.1 Generation of insulin amyloid fibrils

Insulin protofibrils and fibrils were produced as described (Groenning, Norrman *et al.* 2007). In brief, different concentrations of insulin (bovine pancreas, Sigma) ranging from 0.5-2 mg/mL (85-340 µM) were prepared in 25 mM HCl (pH 1.6) and placed at 60 °C for an hour (protofibrils) or at least overnight (fibrils) although incubation times were dependent on protein concentration.

2.5.2 Generation of MstnPP amyloid aggregates and fibrils

MstnPP soluble aggregates were sterile-filtered (0.22 µm), concentrated 10-fold and resuspended in dilute HCl solutions at a final concentration of 1 mg/mL, as measured by both Bradford assay and absorbance at 280 nm. Initially, a range of pH solutions were investigated for fibril formation from pH 1.6, as used for insulin amyloid formation (Groenning, Norrman *et al.* 2007) through to pH 6.3. Samples were placed at 60 °C and monitored frequently for amyloid formation by Thioflavin T (ThT) fluorescence, transmission electron microscopy (TEM) and circular dichroism (CD) spectroscopy. Subsequent to this optimisation, MstnPP aggregates were incubated at pH 5.3 (0.005 mM HCl) and 60°C for at least one week. Identical MstnPP amyloid fibrils were formed in solutions containing 0.1% sodium azide.

2.6 Cell culture

2.6.1 C2C12 growth and maintenance

C2C12 mouse myoblasts were cultured in Advanced Dulbecco's Modified Eagle's Medium (Gibco, Invitrogen) containing 4.5 g/L D-glucose and 110 mg/L sodium pyruvate and supplemented with 10% foetal calf serum, 4 mM L-glutamine and penicillin/streptomycin. Cells were incubated in a 5% CO₂ humidified environment at 37 °C and passaged at 80-90% confluency.

2.6.2 C2C12 proliferation assays

Cells were plated in fresh medium in optical bottom 96-well plates (Nunc) at a density of 1,000 cells/well. After 24 hours, media was removed and 100 µL/well fresh medium containing protein was added. Three independent experiments with triplicate wells were used for each condition for WT; two independent experiments were conducted in triplicate for C313Y. Cells incubated in media only and media containing buffer (50 mM HEPES pH7.5 + 1 mM CaCl₂) only were used as negative controls. The myostatin growth factor concentration was 10 µg/mL, estimated from SDS-PAGE after furin digest; equivalent concentrations of the untreated furin digest and MstnPP were used. Activation of the myostatin growth factor was performed by either acid- or heat-treatment. For acid treatment, the furin digested sample was adjusted to pH 2.5 with 1 M HCl, incubated on ice for 1 hour at 4 °C and neutralized to pH 7.5 with 1 M NaOH (Lawrence, Pircher *et al.* 1985; Funkenstein and Rebhan 2007). Heat treatment was carried out by incubation at 80 °C for 5 minutes (Funkenstein and Rebhan 2007). Cells were incubated at 37 °C, 5% CO₂ with protein for 72 hours and cell growth was measured using the WST-1 Cell Proliferation Assay (Roche).

2.6.3 C2C12 cytotoxicity assays

Cells were plated in fresh medium in optical bottom 96-well plates (Nunc) at a density of 2,500 cells/well. After 24 hours, media was removed and 100 µL/well fresh medium containing protein was added. Two independent experiments with triplicate wells were used for each condition. Cells incubated in media containing buffer only (50 mM Tris-HCl pH 8.5, 150 mM NaCl for soluble aggregates and dimer; 0.005 mM HCl, pH 5.3 for protofibrils and fibrils) were used as a control. Concentrations of myostatin samples used in the assay were 25, 10 and 5 µM, comparable to those from previous protein fibril studies (Walsh, Hartley *et al.* 1999; Bucciantini, Giannoni *et al.* 2002; Bucciantini, Calloni *et al.* 2004). Cells were incubated at 37 °C, 5% CO₂ with protein for 24 hours and cytotoxicity was measured using the WST-1 Cytotoxicity Assay (Roche).

2.6.4 WST-1 colorimetric assay for cell proliferation and protein cytotoxicity

The WST-1 reagent (Roche) is a tetrazolium-based salt that can be metabolized by viable cells only to produce a water-soluble formazan product measurable in a

colorimetric assay. After incubation with protein, 10 µL WST-1 was added to each well and the plate was incubated for a further 3 hours. Absorbance at 450 nm was measured in a PowerWave XS 96-well plate reader (BioTek Instruments, Inc) with 630 nm readings taken as reference. Phenol red-free medium was not used as the indicator does not interfere with the WST-1 assay; as phenol red absorbs slightly at 450 nm however, wells containing media only were used as an additional control (Francoeur and Assalian 1996). The Student's t-test was performed using GraphPad Prism (GraphPad Software, Inc).

2.7 Protein analysis

2.7.1 Circular dichroism spectroscopy

CD spectra in the far-UV region (180-260 nm) were obtained on a Chirascan CD spectrometer (Applied Biophysics) using a 0.1 mm path cell and protein concentrations of 1 mg/mL, or a 1 mm path cell and protein concentrations of 0.1 mg/mL, at 4 °C. For all samples, 20 runs were performed with 1 nm readings taken every 2.5 seconds, followed by smoothing and baseline subtraction. Deconvolution of CD spectra was performed using CDDN software (Applied Biophysics) (Appendix 1, Table 12.1 and Appendix 4, Table 12.2a and b).

2.7.2 CD thermal denaturation

For CD thermal denaturation, 1 nm/2.5 second readings were taken at every 5 °C (precursor protein and soluble aggregates) or 10 °C (latent complex) increase in temperature from 5 – 90 °C with a 30 second equilibration time at each temperature and a tolerance level of 0.2 °C.

2.7.3 Thermal shift assays

Purified WT or C313Y MstnPP in 50 mM Tris-HCl, 150 mM NaCl, pH 8.5 was concentrated to 1.0 mg/mL. Sypro Orange (Sigma) was diluted 100x according to manufacturer's instructions in milliQ H₂O and 2 µL was added to 18 µL of protein solution, to give a final dilution factor for the dye of 1000 x. A negative control containing buffer and dye only was also made. The RotorGene 6000 thermocycler (Corbett) was used for analysis with excitation at 470 nm and fluorescence emission measured at 555 nm over increasing temperature from 30 to 95 °C in 1 °C increments;

each run lasted approximately 19 minutes. Differentiation of the data was performed by the Rotor Gene 6000 Series software.

2.7.4 X-ray crystallography

Purified MstnPP dimer was concentrated to 5 mg/mL, centrifuged at 17,000 x g for 10 minutes at 4 °C and filtered (0.22 µm) for crystallisation trials. 1 µL of protein was mixed with 1 µL of mother liquor in sitting drops using 96-well Crystalquick plates (Greiner Bio-One) containing 100 µL well volumes of Crystal Screen and Crystal Screen 2 (Hampton Research) at 22 °C. Purified lyophilized myostatin growth factor (MstnGF) was resuspended in 10 mM acetic acid to a concentration to 3 mg/mL, centrifuged at 17,000 x g for 10 minutes at 4 °C and filtered (0.22 µm). 200 nL of protein was mixed with 200 nL of mother liquor by the Mosquito crystallisation robot (TTP Labtech) at Otago University in a hanging drop format. Plates contained 100 µL wells of Crystal Screen and Crystal Screen 2 (Hampton Research) and Structure Screen I and II (Molecular Dimensions) in a 96-well format. For both proteins, optimisation of conditions was done by the hanging-drop, vapour-diffusion method using a range of protein to mother liquor (v/v) ratios; for the growth factor a ratio of 2:1 was used. Crystals were screened for diffraction in-house using an RAXIS IV++ image plate coupled to a MicroMax-007 X-ray generator (Rigaku) with Osmic Blue optics. Diffraction was monitored using Crystal Clear software (Rigaku). Crystals were imaged and photographed using a Zeiss Axiophot compound light microscope fitted with a Leica DFC320 digital camera, and Leica Application Suite V3.3.0 software.

2.7.5 Protease resistance analysis

MstnPP samples were tested for protease resistance using a myostatin to trypsin (bovine pancreas, Sigma) ratio of 100:1 (MstnPP WT and C313Y dimer and soluble aggregates), or 20:1 (MstnPP protofibrils and fibrils) (w/w). 100:1 solutions were incubated at 4 °C, room temperature or 37 °C. Samples were taken at 0.5, 1, 2, 3, 4 and 18 hours (overnight), immediately denatured to end the reaction and analysed by reducing and non-reducing SDS-PAGE. For additional protease resistance analysis of MstnPP amyloid protofibrils and fibrils, the myostatin to trypsin ratio was 20:1 (w/w) with incubation at 37 °C for 4 and 18 hours and analysis by SDS-PAGE and TEM.

2.7.6 Negative-stained transmission electron microscopy

Dimer and aggregate samples direct from purification (heparin affinity or gel filtration) were buffer-exchanged into water. Amyloid solutions were used directly in dilute HCl. 200-mesh carbon-coated Formvar grids were placed on 30 µL protein solution (0.5 mg/mL) droplets for 45 seconds. Excess sample was drawn off using filter paper and the grids placed on an equal volume of 2% uranyl acetate for 45 seconds. Excess stain was drawn off as above and the grids were air-dried briefly before viewing with a Philips CM10 transmission electron microscope. Fibril width measurement and statistical analysis was performed using iTEM software (Olympus) with a sample size of 10 fibrils from the same grid. Where a single fibril was concerned, at least 5 measurements along the same fibril were taken. For twist distance measurement and analysis, a sample size of at least 5 twists from the same fibril was used.

2.7.7 Thioflavin T binding assays

A solution of Thioflavin T (Sigma) was prepared at 400 µM in water (Bourhim, Kruzel *et al.* 2007) then diluted directly into the protein solution giving a final concentration of 20 µM per assay, as done previously for insulin (Groenning, Norrman *et al.* 2007)). A final protein concentration of 1 mg/mL was used for all samples. Experiments at pH 7.5 were performed after diluting both protein and ThT into 50 mM Tris-HCl pH 7.5, 100 mM NaCl. Fluorescence assays were conducted with the Perkin Elmer LS50B Luminescence Spectrometer with excitation at 450 nm and the emission spectrum measured from 460-530 nm. Error bars are standard errors for four independent measurements. The Student's t-test was performed using GraphPad Prism (GraphPad Software, Inc).

2.7.8 One dimensional NMR

1D NMR experiments were performed on a Bruker Avance 700 MHz spectrometer equipped with a cryoprobe, four rf-channels and gradient pulse capabilities. All spectra were acquired at 25 °C on 300 µL samples containing 5% D₂O in Shigemi NMR tubes. The myostatin growth factor concentration was 0.2 mM in 50 mM Na acetate, pH 4, 100 mM NDSB 201. ¹H spectra were recorded with 32 transients, and 160000 points with a spectral width of 8389.2 Hz. Fourier transformation and final

processing was accomplished using TopSpin 2.1 (Bruker BioSpin GmbH, Rheinstetten, Germany).

2.8 *In silico* analyses

2.8.1 Sequence alignment and tree building

All sequence alignment and tree building was performed with ClustalW (Thompson, Higgins *et al.* 1994) using default settings. Sequence information was obtained from the Protein Knowledgebase (UniProtKB, www.uniprot.org) available via the Expasy website (<http://ca.expasy.org/>).

2.8.2 Primary sequence

The MstnPP amino acid sequence was analysed using the NPS@: Network Protein Sequence Analysis server (Combet, Blanchet *et al.* 2000) available from the Expasy website. Default settings were used, with sequence identity set at 100%. The option to exclude motifs with a high probability of occurrence was selected. The full-length myostatin sequence was analysed for signal peptide motifs using the Signal P 3.0 Server (Bentsen, Nielsen *et al.* 2004) and UniProtKB, the Protein Knowledge Database (<http://www.uniprot.org/uniprot/O14793>).

2.8.3 Secondary structure

Secondary structure prediction for MstnPP was performed using JPred (Cuff, Clamp *et al.* 1998), PredictProtein (Rost, Yachdav *et al.* 2004) and Phyre (Kelley and Sternberg 2009). Jpred and PredictProtein are available via the Expasy website. Default settings were used for all programs.

2.8.4 Intrinsic disorder

Predictions of intrinsic disorder were carried out using PONDR (Romero, Obradovic *et al.* 1997), DisEMBL (<http://dis.embl.de/>), PrDOS (Ishida and Kinoshita 2007), GlobPlot (<http://globplot.embl.de/>) and FoldIndex (Prilusky, Felder *et al.* 2005). Regions of the sequence where three or more of the algorithms predicted a high propensity for disorder (Appendix 4, Fig. 12.8) were selected as having a significant likelihood of disorder. Default settings were used for all algorithms. Raw data are

shown in Appendix 4 (Fig. 12.8). Hydrophobic cluster analysis (HCA) was performed using the drawhca algorithm and default settings (Callebaut, Labesse *et al.* 1997).

2.8.5 Aggregation and amyloid formation propensity

MstnPP amino acids 21-375 were used in β -aggregation propensity prediction algorithms. Tango (Fernandez-Escamilla, Rousseau *et al.* 2004) and Waltz (Reumers, Schymkowitz *et al.* 2009) output was reproduced from www.tango.embl.de; PASTA output was taken from www.protein.cribi.unipd.it/pasta (Trovato, Seno *et al.* 2007). Default settings were used for all programs. Raw data is shown in Appendix 5 (Fig. 12.9).

2.8.6 Structural modelling

The coordinates of the myostatin/follistatin complex structure (pdb reference: 10.2210/pdb3hh2/pdb) were downloaded from the RCSB Protein Databank (<http://www.rcsb.org/pdb/home/home.do>). Modelling of the C313Y mutation in the myostatin structure was performed using Coot (Emsley and Cowtan 2004). Subsequent structural analysis and figure construction was carried out using PyMol (<http://pymol.sourceforge.net/>).

Overview of Results

This PhD presents the structural and functional analysis of the human myostatin precursor protein, with comparisons to the myostatin latent complex and the equivalent Piedmontese proteins. These findings have implications for our understanding of the regulation of muscle growth and diseases of muscle wastage.

Production of the myostatin precursor protein required *in vitro* refolding after production in bacteria as insoluble inclusion bodies. Subsequent purification of the correctly folded dimer and *in vitro* recapitulation of the *in vivo* processing of the precursor protein allowed production of the myostatin latent complex, and the biochemical and biophysical analysis of both species. The methods used for the wildtype protein were then directly transferable for use in production of the Piedmontese myostatin precursor protein. The Piedmontese cattle breed exhibit a double-muscled phenotype due to the production of a form of myostatin that is unable to elicit inhibitory signalling.

The refolding and analysis of the myostatin precursor protein also led to an investigation of the misfolded forms of myostatin refolding produces; a link for misfolded myostatin protein in the muscle wastage disease sporadic inclusion body myositis (sIBM) was recently identified. sIBM is a disease of amyloid formation and the myostatin precursor protein has a high propensity to form amyloid fibrils *in vitro* under destabilising conditions.

This PhD therefore has two major themes. The first is the study of the properties of the correctly folded protein and the information that this provides for the function of the protein *in vivo*. The second is the study of misfolded protein, its ability to form cytotoxic amyloid fibrils and the implications this has for sIBM.

To date, little has been published in either research area. The results presented here are therefore extremely novel and have implications for the mode of action of myostatin in both normal and diseased backgrounds.

The initial objective of this PhD was the structural determination of the human myostatin growth factor dimer by X-ray crystallography. Although the structure of myostatin in complex with follistatin was published before this objective could be completed, meaning a change in focus, the results that were obtained are included as Appendix 1 (Chapter 12).

3 Production of recombinant human myostatin

3.1 Cloning of human myostatin

The full-length myostatin cDNA (amino acids 1-375) was used as a template for cloning. As introduced previously (Chapter 1), the full-length protein contains a signal sequence N-terminal of the precursor protein that can be divided into propeptide and mature growth factor regions. Constructs corresponding to all three putative propeptide start sites (as detailed in Chapter 2), amino acids 19 (Bentsen, Nielsen *et al.* 2004), 21 (Jin, Dunn *et al.* 2004) and 25 (Lee 2004), were made (F19, F21 and F25 respectively). All constructs contained 5' *Bam*HI sites and 3' *Xho*I sites for cloning into the pProEX-Htb expression vector, which allows production of an N-terminally His-tagged protein in *E.coli*.

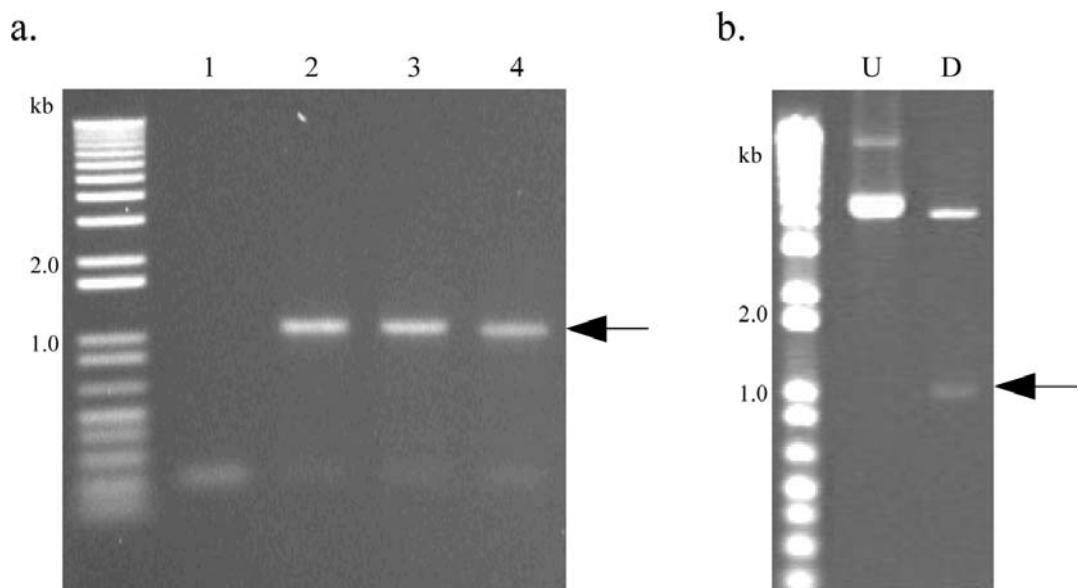


Figure 3.1 Cloning of F19, F21, F25 and F267 myostatin cDNA.

- PCR reactions generating F19 (lane 2), F21 (lane 3) and F25 (lane 4) constructs (arrow). Lane 1 is a negative control with no template DNA.
- Bam*HI/*Xho*I double digest of pProEX-MstnF21 plasmid. U, undigested plasmid; D, digested plasmid. Arrow indicates MstnF21 construct.

After PCR, precursor constructs were just over 1 kB (Figure 3.1a, arrow). Positive colonies were selected using colony PCR (Chapter 2, Figure 2.1) and the presence of myostatin constructs was confirmed by restriction digest of each plasmid (Figure 3.1b). Sequences were verified by DNA sequencing with pProEX-Htb sequencing primers, followed by NCBI BLAST.

3.2 Expression

The expression of all three precursor protein constructs (MstnF19/21/25) was investigated in preliminary expression trials. The expression analysis of MstnF21 is shown in Figure 3.2; no obvious expression differences between the constructs were seen and all were observed as a 50 kDa band within insoluble *E.coli* inclusion bodies (Fig. 3.2, arrow). Although this molecular weight (MWt) is consistent with previous publications (Jin, Dunn *et al.* 2004; Funkenstein and Rebhan 2007), it is significantly larger than the expected 40 kDa size predicted by sequence. The precursor construct starting at amino acid 21 (MstnF21) was selected for use in future studies as this was the start site used in other papers documenting successful production of soluble myostatin precursor protein.

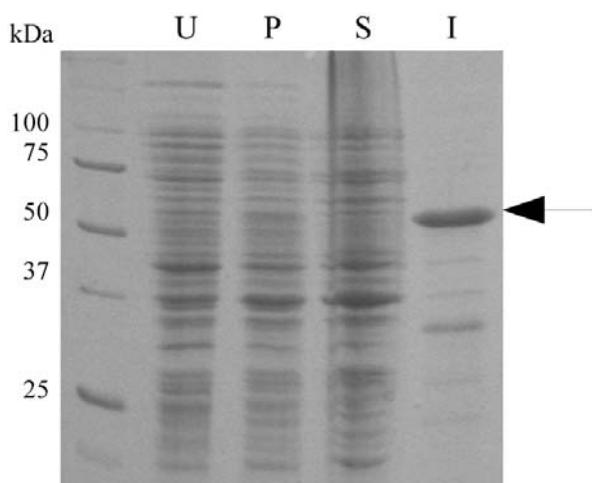


Figure 3.2 Expression analysis of MstnF21 in *E.coli*.

Lanes are as follows: U, uninduced culture; P, post-induction; S, soluble fraction post-lysis; I, insoluble fraction post-lysis. Arrow, Mstn21 protein.

Expression trials

As refolding requires a ‘trial and error’ approach and can be time and materials-consuming, a range of expression trials were carried out initially in an attempt to promote the solubility of the myostatin constructs in *E.coli* (Table 3.1). Alterations to expression temperature, time and IPTG concentration were also investigated with no increase in protein solubility. With the exception of the use of co-expression of the pTUM4 vector, all systems trialled for MstnF21 were unsuccessful and will not be discussed here in any more detail.

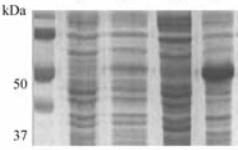
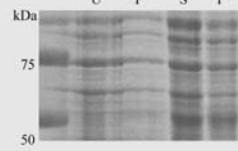
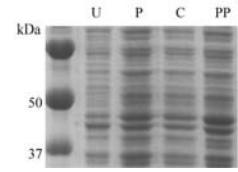
Expression System	Details of System	Results	Summary
SUMO fusion tag	- N-terminal fusion protein tag that promotes solubility		- Insoluble expression only - SUMO-F19/21/25 band at ~70 kDa (*)
ORIGAMI B (<i>E.coli</i> variant)	- Oxidising environment for disulphide bond formation		- Insoluble expression only - His-F19/21/25 band at 50 kDa (*)
EtOH SHOCK	- 6 % EtOH added 1 hour prior to induction - Stress expected to cause upregulation of bacterial chaperones and enhanced folding of recombinant protein		- His-F19/21/25 band at 50 kDa (*) - Some solubility apparent after initial trials (low volume) - Larger volumes 100% insoluble
pTUM4	- pTUM4 vector contains four bacterial chaperones - Coupled with periplasmic expression for formation of native disulphide bonds		- Expression of 50 kDa His-F19/21/25 in periplasm(*) - Protein is soluble - Low yield

Table 3.1 Summary of main expression systems trialled for MstnF21.

Lanes are as follows: U, uninduced; P, post-induction; S, soluble fraction; I, insoluble fraction; C, cytoplasmic fraction; PP, periplasmic fraction; *, myostatin precursor protein.

pTUM4 expression

The pTUM4 expression system was published as a method to promote the correct folding of disulphide-bond containing proteins in the *E.coli* periplasm by co-expression of four established periplasmic bacterial chaperones and folding catalysts

(Schlapschy, Grimm *et al.* 2006). These are DsbA (21 kDa) and DsbC (23.4 kDa), thiol-disulphide oxidoreductases that catalyse the formation of and isomerisation of disulphide bridges, and FkpA (26 kDa) and SurA (45 kDa), peptidyl-prolyl cis/trans-isomerases with chaperone activity (Fig. 3.3a, lane 1). This system requires cloning of the protein of interest in a vector that effects expression in the periplasmic space; the pOmpA vector, which contains the OmpA periplasmic sequence tag, was used in this instance.

Use of the pTUM4 expression system resulted in soluble MstnF21 expression in the periplasmic space, albeit at a low level (Table 3.1, last row *). Protein could be purified by Ni-NTA affinity chromatography with elution at 500 mM imidazole (Figure 3.3a, lane 2, arrow) and gel filtration chromatography (Figure 3.3b, arrow). Elution from the S200 gel filtration column occurred at 13.5 mL representing a size of 110 kDa and thus the precursor dimer (Appendix 3, Fig. 12.7a). However, expression and solubility levels were not consistent and purified protein yields were low. For these reasons, and the successful refolding of MstnF21 as described below, this system was not continued.

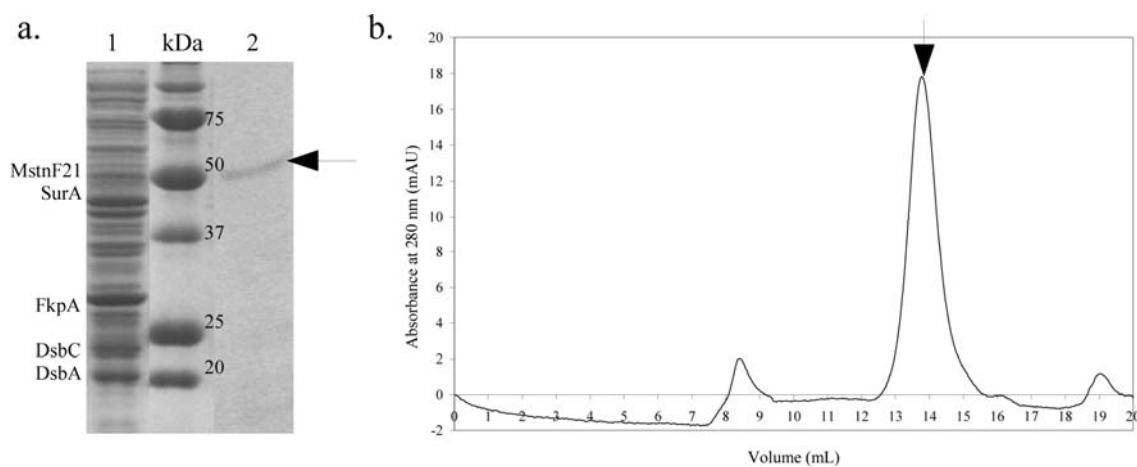


Figure 3.3 The pTUM4 system for production of soluble MstnF21.

- Reducing SDS-PAGE. Lanes as follows: 1, expression of MstnF21 and the bacterial chaperones and folding catalysts of the pTUM4 system (SurA, FkpA, Dsb C and DsbA); 2, Ni-NTA affinity chromatography-purified MstnF21 (arrow).
- Gel filtration purification of the MstnF21 dimer (arrow) which elutes at 13.5 mL.

Refolding

Although some refolding conditions can be predicted, such as the necessity of a redox system where disulphide bonds are involved, exact conditions vary between proteins, making refolding a trial and error process. For myostatin, the majority of conditions tested were taken from papers showing successful refolding of related proteins. Purified inclusion bodies (Fig. 2.3) were used for refolding procedures, summarised in Table 3.2.

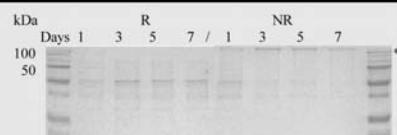
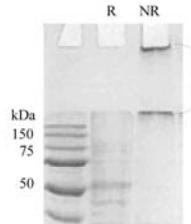
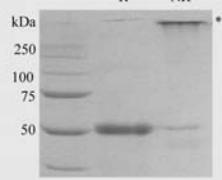
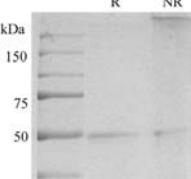
Refolding System	Summary of system	Results
Refolding of porcine precursor	<ul style="list-style-type: none"> Solubilisation in detergent (N-laurylsarcosine) and alkaline pH (CAPS pH 11.0) Refolding in presence of glutathione redox system with GSH:GSSG ratio of 1.3:1 	 <ul style="list-style-type: none"> Not reproducible for human myostatin precursor. Formation of soluble aggregates (*) with disappearance of monomer
Additives:	<ul style="list-style-type: none"> NDSBs BMC 	<ul style="list-style-type: none"> NDSBs proposed to assist in refolding via reduction of aggregation BMC said to have isomerase-like activity Positive synergistic interaction between NDSB 201 and BMC  <ul style="list-style-type: none"> Formation of soluble aggregates (*) Dimer absent
Gentle solubilisation	Based on presence of secondary structure in inclusion bodies – mild solubilisation may preserve inherent structures assisting in refolding process	 <ul style="list-style-type: none"> Formation of soluble aggregates (*) Dimer absent
Gradual removal of reducing agent	Gradual reoxidation proposed to allow formation of native disulphide bonds	 <ul style="list-style-type: none"> Formation of soluble aggregates (*) Dimer absent

Table 3.2 Refolding methods used for the myostatin precursor protein.

SDS-PAGE was conducted in both reducing (R) and non-reducing (NR) conditions as a measure of folding success. Aggregated MstnF21 protein is indicated by an asterisk (*).

The first condition tried was that published for the porcine myostatin precursor (Jin, Dunn *et al.* 2004). In brief, this uses detergent solubilisation (*N*-lauryl sarcosine) and a glutathione redox system for refolding, with a molar ratio of reduced to oxidised glutathione of 1.3:1. Reproduction of this method for the human myostatin precursor protein was not successful, indicated by the absence of MstnF21 dimer on non-reducing (NR) SDS-PAGE and the presence of a large molecular weight complex unable to enter the resolving gel (Table 3.2, *). As MstnF21 monomer is observed as the expected 50 kDa band reducing SDS-PAGE, the likely explanation is that this is an aggregated species greater than 250 kDa, due to non-native disulphide formation. There are thirteen cysteine residues in MstnF21, nine of which are known to participate in disulphide bonds in the native structure. Non-native disulphide formation is therefore a high possibility during the refolding process. Exposed hydrophobic regions may also contribute to aggregation.

Successful refolding of MstnF21 was achieved using a system published in 2007 for the zebrafish myostatin precursor protein (Funkenstein and Rebhan 2007), shown by the presence of the disulphide-bonded dimer on non-reducing (NR) SDS-PAGE (Figure 3.4). The His-tagged MstnF21 monomer runs just above 50 kDa in reducing conditions as expected from previous results, yet three bands are present upon the removal of reducing agent. The predominant species is the MstnF21 dimer at approximately 125 kDa (**), larger than expected by sequence, consistent with previous results. Also present are misfolded monomer (around 50 kDa, *) and the high molecular weight, aggregated species (#). The monomer appears as two bands. Exactly what these represent is unknown, although one explanation is that these are monomeric proteins exhibiting different degrees of disulphide bonding. One of these may represent the native structure and be in equilibrium with the dimer.

The refolding buffer in this method contains a glutathione redox system with a 5:1 molar ratio of reduced to oxidised glutathione, high salt (1 M NaCl) and a high concentration of L-arginine (0.5 M), suggested to prevent aggregation. The exact method was followed initially (Funkenstein and Rebhan 2007) then optimised for the human precursor protein. A key feature of the method is solubilisation with 6 M guanidine HCl (GndHCl) and 0.1 mM DTT, followed by overnight dialysis into a low pH, GndHCl buffer. This keeps disulphide bonds in a reduced state in the absence of reducing agent, which may interfere during refolding. Solubilisation is followed by a

rapid 15-fold dilution into refolding buffer and incubation at 4 °C for 6 days. The original method uses a 12.5-fold dilution and a 14 day refolding period; for MstnF21, the dimer yield was reduced dramatically after day 6. These differences may be attributed to the higher dilution factor or sequence differences between the zebrafish and human myostatin precursor protein that translate into structural and/or biochemical differences.

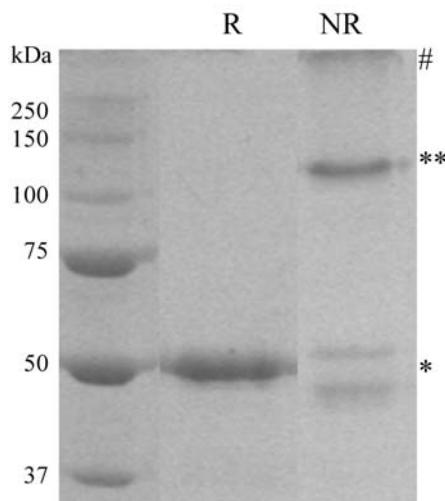


Figure 3.4 Refolding of the myostatin precursor protein dimer.

Reducing (R) vs. non-reducing (NR) SDS-PAGE showing the 125 kDa dimer (**), the 50 kDa monomer (*) and large molecular weight soluble aggregates (#).

It is intriguing that the human myostatin precursor protein was able to fold in a system developed for the zebrafish precursor rather than that of the porcine precursor, since by sequence alignment, the human protein shows higher similarity to the latter (Fig. 3.5). Sequence analysis indicates only three places in the myostatin sequence where the human and zebrafish proteins have similarity not shared by the pig protein (Fig. 3.5, blue). While two are conservative, one involves a charge difference (blue box); a lysine residue in pig compared to glutamic acid in human and zebrafish. One possibility is that this residue plays a key role in the proposed chaperone activity of the propeptide region. Another possibility is that the porcine method is not reproducible, with the porcine precursor protein folding more readily in the zebrafish system.

CLUSTAL 2.0.12 multiple sequence alignment

sp O14793 GDF8_HUMAN	MQKLQLCVIYILFMLIVAGP	V---DLNENSEQKENVEKEGLCNACTWRQN	47
sp O18831 GDF8_PIG	MQKLQIYVYIYLFMLIVAGP	V---DLNENSEQKENVEKEGLCNACMWRQN	47
sp O42222 GDF8_ZEBRAFISH	MHFTQ--VLISLSVLIA	CGPVGYGDITAHQPSTATEESELCSCTCEFRQH	48
	: * * * * :**..	*.: . :.*.: **..*: ***:	***:
sp O14793 GDF8_HUMAN	TKSSRLEAIKIQILSKLRLETAPNISKD	VIRQLLPKAPPRELIDQYDVQ	97
sp O18831 GDF8_PIG	TKSSRLEAIKIQILSKLRLETAPNISKD	AIRQLLPKAPPRELIDQYDVQ	97
sp O42222 GDF8_ZEBRAFISH	SKLMRLHAIKSQILSKLRLKQAPNISRD	VVKQLLPKAPPQQLLDQYDVL	98
	:* * :.***: ****:*****:	*****:*.:.*****:*****:*****:	*****
sp O14793 GDF8_HUMAN	RDDSSDGSLLEDDDYHATTETI	IITMPTESDFLMQVDGKP	147
sp O18831 GDF8_PIG	RDDSSDGSLLEDDDYHATTETI	IITMPTESDFLMQVEGKP	147
sp O42222 GDF8_ZEBRAFISH	GDDSGLDAVEEDDEHATTETIMT	MAPEPDPIVQVDRKP	148
	****:****:****:****:****	KKCCFFSFSPKIQ	****
sp O14793 GDF8_HUMAN	YNKVVKAQLWIYLRPV	EPTTVFVQILRLIKPMKDGT	197
sp O18831 GDF8_PIG	YNKVVKAQLWIYLRPV	TRYTGIRSLKLDNM	197
sp O42222 GDF8_ZEBRAFISH	ANRIVRAQLWVHLRP	EATTVFLQISRLM-PVKDGGRRH-IRSLKIDVN	196
	*:****:****:****:****	IRSLKIDVN	****:****:****
sp O14793 GDF8_HUMAN	PGTGIWQSIDVKTVLQNWLKQ	PESNLGIEIKALDENGHDLAVTF	247
sp O18831 GDF8_PIG	PGTGIWQSIDVKTVLQNWLKQ	PESNLGIEIKALDENGHDLAVTF	247
sp O42222 GDF8_ZEBRAFISH	AGVTWSQSIDVKQVLTVWLKQ	PETNRGIEINAYDAKGNDLAVTSTETGED	246
	:*:****:****:****	***:****:****:****:****	****
sp O14793 GDF8_HUMAN	GLNPFFLEVKTDTPKRSRRD	FGLDCDEHSTESRC	297
sp O18831 GDF8_PIG	GLNPFFLEVKTDTPKRSRRD	FGLDCDEHSTESRC	297
sp O42222 GDF8_ZEBRAFISH	GLLPFMEVKISEGP	CRYPLTVDFEDFGWDW	296
	:*:****:****:****	*****:*****:*****:*****	*****
sp O14793 GDF8_HUMAN	IIAPKRYKANYC	SGECEFVFLQKYPH	347
sp O18831 GDF8_PIG	IIAPKRYKANYC	SGECEFVFLQKYPH	347
sp O42222 GDF8_ZEBRAFISH	IIAPKRYKANYC	SGECDYMYLQKYPH	346
	*****:*****:*****:*****	TLVNVNKASPRGTAGPC	*****
sp O14793 GDF8_HUMAN	INMLYFNGKEQIIY	YKGKIPAMV	375
sp O18831 GDF8_PIG	INMLYFNGKEQIIY	YKGKIPAMV	375
sp O42222 GDF8_ZEBRAFISH	INMLYFNGKEQIIY	YKGKIPSMV	374
	*****:*****:*****:*****	YKGKIPSMV	*****

Figure 3.5 Sequence alignment of human, pig and zebrafish myostatin.

Sequence alignment was performed using ClustalW (Thompson, Higgins *et al.* 1994). Regions where human and zebrafish share identity that pig does not are shown in blue; the most significant difference is indicated by a blue box. The signal peptide is in red. The ClustalW similarity key is as follows: *, exact match; :, strong similarity; ., weak similarity.

There are a number of differences between the refolding methods that may contribute to this result. First, the ratio of reduced to oxidised glutathione in the porcine refolding buffer is approximately 3-fold less than that of the zebrafish buffer. Second, the zebrafish refolding buffer contains L-arginine which is suggested to reduce aggregation. Third, the solubilisation methods, known to have a marked effect on refolding efficiency (Tsumoto, Ejima *et al.* 2003) are significantly different. Solubilisation in the zebrafish method is followed by dialysis into a low pH GndHCl buffer; the porcine protocol uses detergent solubilisation. Detergent solubilisation is

an ‘on-or-off’ system, as detergent micelle formation occurs at a critical concentration (Tsumoto, Ejima *et al.* 2003). Therefore, depending on the concentration of the detergent, the protein is either fully denatured or not. In contrast, urea and GndHCl allow denaturation to occur gradually. These properties are important on dilution of the solubilised protein into refolding buffer; rapid refolding must occur on removal of detergent, compared to the gradual process enabled on removal of urea and GndHCl. The successful refolding of MstnF21 led to sole use of this construct. From here on, the human myostatin precursor protein is referred to as MstnPP.

3.3 Purification of refolded MstnPP

The published zebrafish protocol uses Ni-NTA affinity chromatography followed by size exclusion chromatography for purification of the refolded dimer (Funkenstein and Rebhan 2007). While pure His-tagged MstnPP eluted at 100 mM imidazole off an Ni-NTA column (Fig. 3.6a), subsequent isolation of the dimer from misfolded and aggregated protein using size exclusion chromatography was inefficient (Fig. 3.6b).

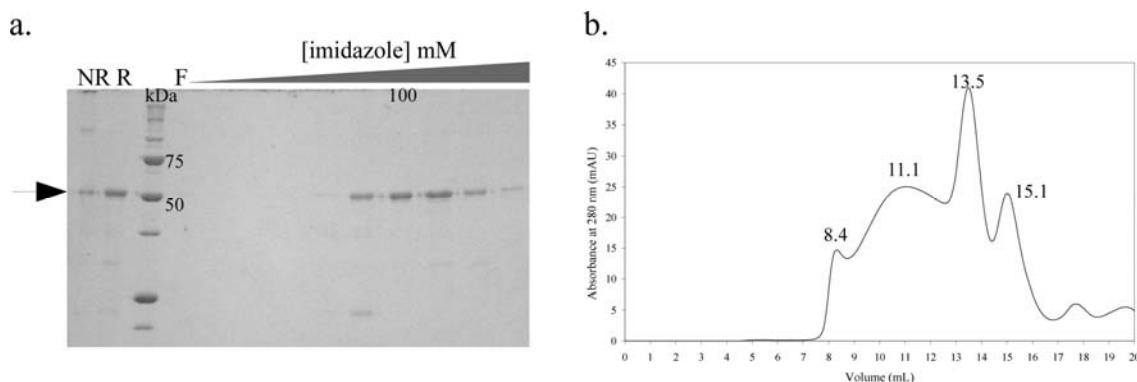


Figure 3.6 Purification of the refolded His-tagged myostatin precursor protein by Ni-NTA chromatography.

- a. Reducing SDS-PAGE of His-MstnPP (arrow) Ni-NTA purification. Lanes are as follows: NR, post-refolding crude sample in non-reducing sample buffer; R, post-refolding crude sample in reducing sample buffer; F, flowthrough; 100, 100 mM imidazole.
- b. S200 gel filtration chromatography of the Ni-NTA purified precursor protein.

This is most likely because the Superose 12 column used in the original protocol has a relatively large column volume (125 mL), which separated the dimer from aggregated

protein very effectively. The Superdex 200 column used for MstnPP, with a notably reduced volume of 24 mL, was not able to separate the dimer at 13.5 mL from soluble aggregates, which eluted from the column in a broad peak from 8 to over 13 mL (Fig. 3.6b).

It was therefore necessary to establish a protocol where effective separation of the correctly folded dimer from aggregated protein could be achieved prior to gel filtration. TGF- β proteins are known to bind heparin *in vivo* (Rider 2006) and heparin affinity chromatography has been used for purification of the myostatin growth factor produced in mammalian cells (Lee and McPherron 2001). As heparin binding is based on charge, and differently folded forms of myostatin are likely to have different charge properties, it is expected that these different forms will bind heparin with different affinities. An optimised heparin separation (Fig. 3.7a) results in the elution of correctly folded precursor dimer at 200 mM NaCl (H1) and the majority of aggregated protein at 600 mM NaCl (H2), although some aggregates are also present in the dimer peak (Fig. 3.7c, lane H1). This initial result suggests that the aggregates may represent two populations that differ in structure.

After heparin separation, gel filtration chromatography with a Superdex 200 column can be used to achieve an effective isolation of the MstnPP dimer (Fig. 3.7b and 3.7c, PD, **). The dimer eluted from the matrix at 13.3 mL, corresponding to a molecular weight of 115,000 Da (Appendix 3, Fig. 12.7b). This retention volume differs from that of the pTUM4 and Ni-NTA purified proteins (Figs. 3.3b and 3.6b respectively) due to a change in gel filtration column (Appendix 3, Fig. 12.7a). The estimated 115 kDa molecular weight is consistent with the high molecular weight of the dimer by non-reducing SDS-PAGE; as gel filtration retention times are affected by protein shape (Erikson 2009), this result suggests an elongated structure for the precursor protein.

Although some aggregates are present in the void volume (8.3 mL, Fig. 3.7 c, PA, #), heparin separation prior to gel filtration removes the broad aggregate peak seen previously (Fig. 3.6b). A Western blot of the purification peaks (Fig. 3.7d) shows that despite the different levels of folding and aggregation, all peaks represent the same protein and confirms that the dimer peak does not contain aggregated protein.

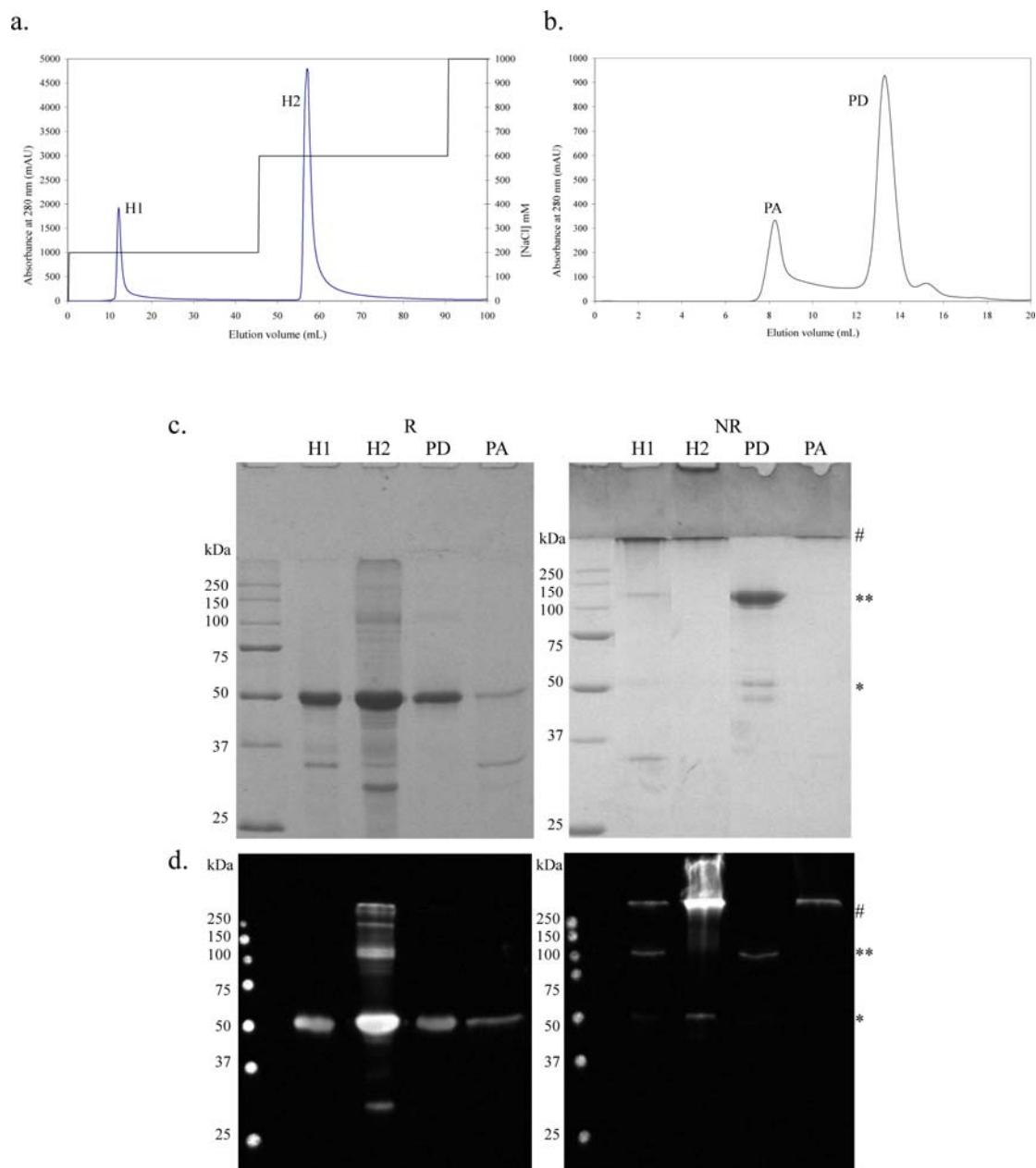


Figure 3.7 Purification of refolded MstnPP by heparin affinity and gel filtration chromatography.

- a. Heparin affinity chromatography.
- b. S200 gel filtration chromatography of heparin peak H1.
- c.-d. The purification process by reducing (R) vs. non-reducing (NR) SDS-PAGE (c) and Western blotting using anti-myostatin antibodies (d).

Labels and lanes are as follows: H1, heparin peak 1; H2, heparin peak 2; PA, purified aggregates; PD, purified dimer. SDS-PAGE/Western bands are as follows: #, MstnPP soluble aggregates; **, MstnPP dimer; *, MstnPP monomer.

3.4 Evidence of refolding success: production of active human MstnGF

The initial indicator for successful refolding of myostatin is the presence of a dimer in non-reducing conditions. However, this is not a guarantee that the native structure has been obtained. To confirm refolding has produced the native structure, activity must be established; this is not possible for MstnPP because the protein has no known measurable activity. *In vivo*, mature myostatin is produced via cleavage of the myostatin precursor protein by furin convertase. Therefore, production of the growth factor and subsequent activity measurements was carried out using commercially available furin convertase enzyme (Jin, Dunn *et al.* 2004; Funkenstein and Rebhan 2007).

3.4.1 Furin digest

Furin digest was carried out according to published methods and the manufacturer's guidelines initially, with subsequent modifications as detailed in Chapter 2.

Analysis of myostatin furin cleavage by reducing (R) SDS-PAGE shows reduction of the precursor band (1) at 50 kDa and concomitant appearance of bands at 37 (2) and 12 (3) kDa (Fig. 3.8a). These represent the propeptide region and the mature growth factor region monomer respectively. In the absence of reducing agent, the growth factor monomer is replaced with a dimer band at 24 kDa (4), indicating successful refolding of the growth factor during the MstnPP refolding protocol. Undigested MstnPP dimer is also visible (5).

To confirm SDS-PAGE results, Western blotting with antibodies against both the myostatin growth factor (anti-MSTN) and the N-terminal His-tag (anti-HIS) was performed. The anti-MSTN antibody recognises a short region at the C-terminus of the growth factor (349-364) and will therefore bind to both the precursor and the growth factor, but not the propeptide region. The anti-HIS antibody will bind to the His-tagged precursor and propeptide, but not the digested growth factor.

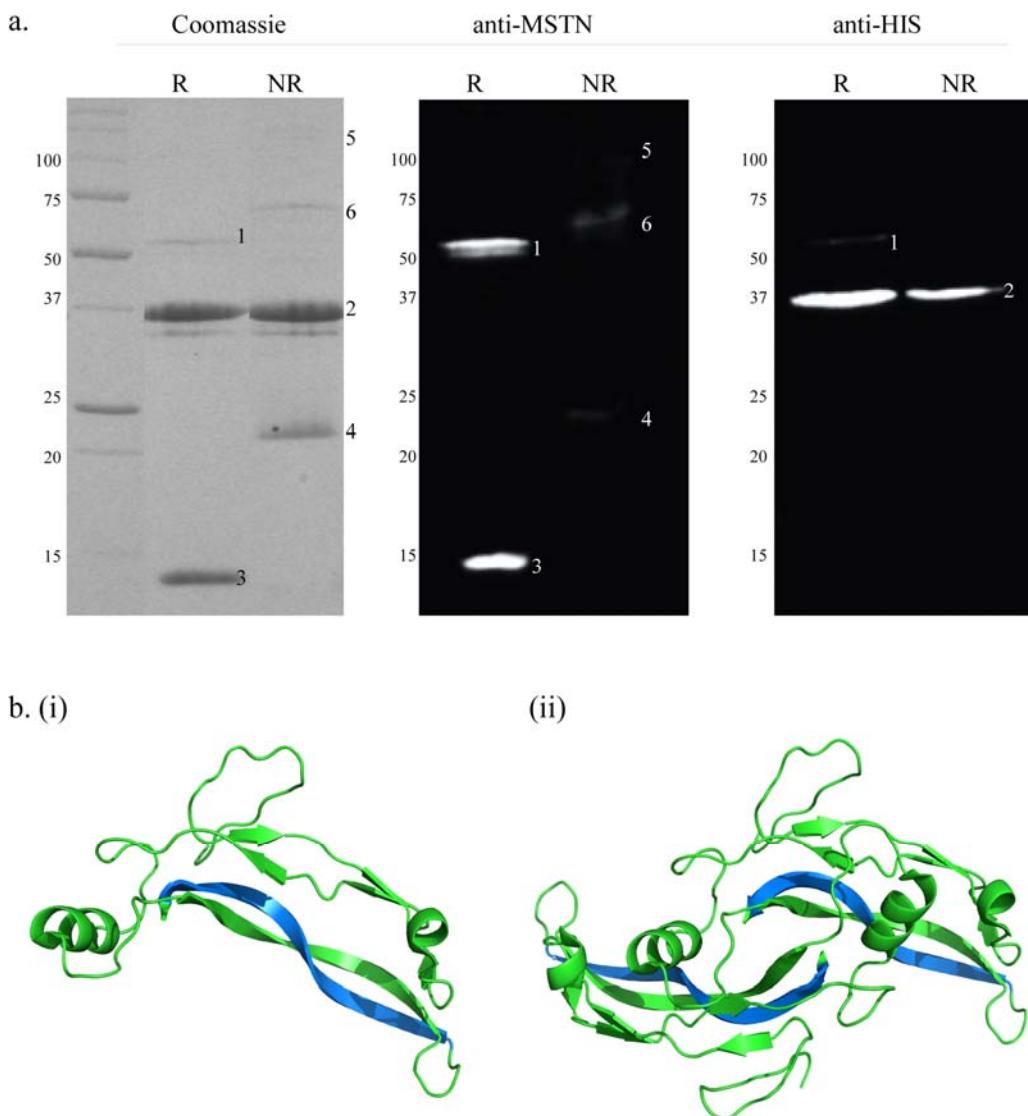


Figure 3.8 Furin digest of MstnPP.

a. SDS-PAGE and Western blotting using antibodies against myostatin (anti-MSTN) and the His-tag (anti-HIS) in reducing (R) and non-reducing (NR) conditions. Bands are as follows: 1, undigested MstnPP monomer; 2, propeptide; 3, growth factor monomer; 4, growth factor dimer; 5, undigested MstnPP dimer; 6, unknown 70 kDa band.

b. The anti-MSTN epitope (blue) modeled in the myostatin growth factor monomer (i) and dimer (ii).

Western blots confirm SDS-PAGE results of successful furin digestion and the production of propeptide and growth factor regions. The anti-MSTN Western shows smearing, and a reduction in band intensities, in the non-reducing (NR) lane. One explanation for the former may be an increase in globularity of the disulphide-linked dimer, causing altered movement through the gel. Transfer onto the nitrocellulose membrane may also allow increased mobility. A reduction in band intensity may be due to reduced accessibility of the antibody to the antigenic region in the dimer. This

suggestion is supported by observation of the myostatin crystal structure (Cash, Rejon *et al.* 2009) which shows that the epitope region (blue) is partially hidden in the dimer (Fig. 3.8b (ii)) compared to the monomer (Fig. 3.8b (i)). As expected, the anti-HIS Western blot shows propeptide bands and a small amount of undigested MstnPP only; an absence of smearing and propeptide dimerisation suggest that the propeptide does not form disulphide bonds.

An intriguing result is the presence of a band at approximately 70 kDa in non-reducing conditions (6), seen in both the Coomassie-stained gel and the anti-MSTN Western blot. The 125 kDa MstnPP dimer is present although faint, suggesting formation of a novel complex. By size, possibilities for the 70 kDa species are the growth factor dimer plus one propeptide ($24 + 37 = 61$ kDa), precursor plus growth factor monomer ($50 + 12 = 62$ kDa) with dimerisation between the two growth factor domains, or two propeptide domains ($37 + 37 = 74$ kDa). Previous research into the latent myostatin complex expressed in a mammalian system presented non-reducing SDS-PAGE and anti-myostatin Western blot data that also shows a band of approximately 70 kDa (Wolfman, McPherron *et al.* 2003). Although antibodies against both the propeptide and C-terminus detected this band, it does not appear in the anti-HIS Western blot shown here (Fig. 3.8a). This may be due to blocking of the His-tag during complex formation, or cleavage of the His-tag, which was often observed in the absence of Tev protease. Alternatively, the 70 kDa band has been suggested to represent misfolded protein (Wolfman, McPherron *et al.* 2003), also a possibility as there was often a proportion of undigested precursor protein remaining after furin digest (Chapter 3), which was not removed by the addition of extra furin convertase.

In either situation, these results show production of the growth factor dimer and successful processing of MstnPP by furin, supporting the conclusion that refolding was successful.

3.4.2 Activity assays

The standard assay for myostatin growth factor activity is an ability to inhibit the proliferation of C2C12 mouse myoblasts (Thomas, Langley *et al.* 2000; Taylor, Bhasin *et al.* 2001; Joulia, Bernardi *et al.* 2003) with cell proliferation routinely

measured using an MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, a tetrazole)-based assay. The yellow MTT salt is reduced to purple formazan in the mitochondria of living cells; absorbance is directly correlated to cell number. Original MTT-assay kits require washing, harvesting and solubilization steps prior to spectrophotometric analysis and the use of phenol-red free media to prevent interference of the indicator with measurements. In contrast, the WST-1 reagent (Roche) is a modified form of MTT that produces a water-soluble formazan product. No extra steps are necessary and phenol-red containing media can be used.

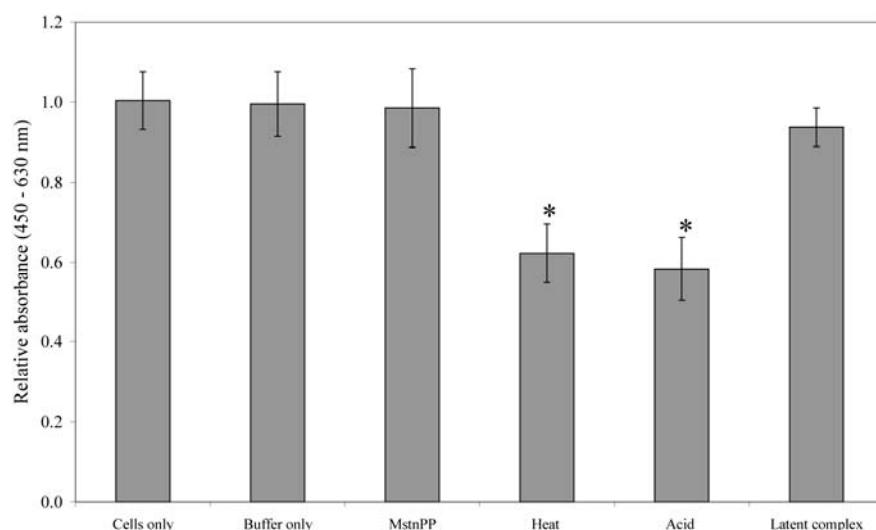


Figure 3.9 Activity of the myostatin growth factor by C2C12 proliferation assays.

Relative absorbance (450-630 nm) is proportional to cell growth using the WST-1 reagent (Roche). Bars are labelled as follows: Heat, heat-treated furin digest; Acid, acid-treated furin digest; MstnPP, undigested MstnPP; Latent complex, untreated furin-digested MstnPP. Cells only and buffer only bars indicate negative controls. Effective concentration of growth factor used was 10 µg/mL within the latent complex; concentration of all other proteins was equivalent to that of the latent complex when the growth factor was 10 µg/mL. * P < 0.05 by Student's T-test.

Following furin cleavage, the latent complex was either heat-treated (80°C for 5 minutes) (Funkenstein and Rebhan 2007) or acid-treated (1 M HCl to pH 2.5 and incubation on ice for 1 hour at 4 °C followed by neutralization with 1 M NaOH to pH 7.5) (Lawrence, Pircher *et al.* 1985), methods shown to be effective previously (Lawrence, Pircher *et al.* 1985), and added to the media of C2C12 myoblasts. The

concentration of the growth factor (10 µg/mL) was estimated by comparing band intensity of furin-digested samples to BSA standards after reducing SDS-PAGE.

Both heat- and acid-treated digests were able to significantly inhibit the proliferation of mouse C2C12 myoblasts when compared to all controls (Fig. 3.9), which were cells only, cells plus the furin digest buffer only, the undigested precursor protein and untreated latent complex. Protein concentrations for the latter two matched that of the total furin digest that represented 10 µg/mL of growth factor. This concentration is greater than that seen in other published reports (Berry, Thomas *et al.* 2002) indicating reduced activity for the refolded myostatin growth factor produced here. Although one possibility is that a proportion of the myostatin growth factor is not correctly folded, the most likely explanation is incomplete dissociation of the latent complex during acid- and heat-treatment, as suggested in Chapter 4 (Fig. 4.10).

4 Biochemical and biophysical analysis of the myostatin precursor protein and latent complex

The TGF- β propeptide region and latent complex play important roles in the production and regulation of the mature protein. The propeptide region of the precursor protein has been suggested to have a chaperone-like function in the correct folding of the growth factor as the myostatin growth factor is unable to fold *in vitro* in its absence (Jin, Dunn *et al.* 2004). After cleavage and latent complex formation the propeptide has a role in export and stability, likely through interactions with the extracellular matrix, and post-secretion, is a potent inhibitor of the growth factor. Therefore, investigation into the structure and function of the precursor, the propeptide and the latent complex is paramount for an understanding of the biology of all TGF- β growth factors, and for myostatin, will provide important information for therapeutic intervention. However, little investigation into the precursor proteins or latent complexes of the TGF- β family has been carried out to date.

This chapter describes the biochemical and biophysical analysis of MstnPP, the processed latent complex, and the propeptide domain within these. The presented results provide novel information that has implications for the understanding of myostatin processing and regulation *in vivo*.

4.1 *In silico* analysis of MstnPP

4.1.1 Primary sequence analysis

Known primary structure motifs were presented in Chapter 1 (Fig. 1.2). Additional motifs were investigated initially using the NPS@: Network Protein Sequence Analysis server (Combet, Blanchet *et al.* 2000) with sequence identity set at 100% (Table 4.1). NPS@ identified the expected TGF- β family signature, as well as a number of phosphorylation sites, a myristylation site, an EF-hand calcium-binding domain and a leucine zipper motif. Initial predictions were tested using additional programs (Table 4.1).

Phosphorylation of TGF- β family proteins has not been documented. Although predictions about the possible role of phosphorylation to myostatin function *in vivo* cannot be made, the majority of these sites are in the propeptide domain. An additional phosphorylation prediction algorithm, KinasePhos 2.0 (Wong, Lee *et al.*

2007) agreed with NPS@ results, suggesting this predicted modification deserves further investigation.

Predicted Motif	Sequence details		Mstn domain	Additional algorithms
TGF-β family signature	298-313	IIAPKRYKANYCSGEC	Growth factor	
Serine phosphorylation sites	51-54	<u>S</u> RIE	Propeptide	Agreement by KinasePhos 2.0
	50-52	<u>S</u> SR	Propeptide	
	105-108	<u>S</u> LED	Propeptide	
	143-145	<u>S</u> SK	Propeptide	
	191-193	<u>S</u> LK	Propeptide	
	264-266	<u>S</u> RR	Furin site	
Threonine phosphorylation site	43-45	<u>T</u> WR	Propeptide	Agreement by KinasePhos 2.0
	123-126	<u>T</u> ESD	Propeptide	
	260-262	<u>T</u> PK	Propeptide	
Tyrosine phosphorylation site	178-186	KPMKD <u>G</u> TRY	Propeptide	Agreement by KinasePhos 2.0
N-myristylation site	37-42	GLCNAC	Propeptide	No agreement by NMT-MYR Predictor
	183-188	GTRYTG	Propeptide	
	201-206	GIWQSI	Propeptide	
EF-hand Ca-binding domain	99-111	DDSSDG <u>S</u> LEDDDY	Propeptide	Agreement by InterPro scan
Leucine zipper motif	216-237	LKQPESNL <u>G</u> I <u>E</u> IKALD ENGHDL	Propeptide	No agreement by 2ZIP algorithm

Table 4.1 *In silico* prediction of primary sequence motifs in MstnPP.

The MstnPP amino acid sequence was analysed using the NPS@: Network Protein Sequence Analysis server (Combet, Blanchet *et al.* 2000) with subsequent analyses performed as detailed.

Myristoylation refers to the covalent attachment of myristate, a 14-carbon saturated fatty acid, and is important for targeting of proteins to membrane locations (Farazi, Waksman *et al.* 2001). This prediction may therefore be relevant for the post-translational export and regulation of myostatin as all motifs are within the propeptide domain. However, an additional myristoylation prediction program, NMT (<http://mendel.imp.ac.at/myristate/SUPLpredictor.htm>), did not predict myristoylation sites in MstnPP. It is therefore unlikely that MstnPP undergoes post-translational myristoylation.

The EF-hand calcium binding domain is responsible for calcium-binding in a range of proteins such as calmodulin and myosin light chains (Grabarek 2006). The motif was also predicted by an InterPro scan (Quevillon, Silventoinen *et al.* 2005), increasing

confidence in the initial result. MstnPP is folded in the ER where calcium plays a number of important roles; however, direct binding of calcium to any TGF- β family propeptide has not been shown, and the existence of EF-hands no longer able to bind calcium is possible.

The leucine zipper motif is located at the C-terminus of the propeptide region. The little structural information available about this region in other family members has not included such a motif and the 2ZIP leucine zipper-specific prediction algorithm (Bornberg-Bauer, Rivals *et al.* 1998) did not recognise this sequence. The presence of a leucine-zipper motif in MstnPP is therefore unlikely.

4.1.2 Secondary structure prediction

Structural information on the TGF- β family precursor proteins is limited to CD analysis of the TGF- β 1 propeptide and latent complex (McMahon, Dignam *et al.* 1996), and the BMP-2 precursor and propeptide (Hillger, Herr *et al.* 2005).

The secondary structure prediction programs PredictProtein (Rost, Yachdav *et al.* 2004), JPred (Cuff, Clamp *et al.* 1998) and Phyre (Kelley and Sternberg 2009), which presents a consensus of predictions from Psi-Pred, SSPro and JNet, were used to analyse the amino acid sequence of the myostatin precursor protein (Fig. 4.1); α -helix is shown in blue and β -sheet in red, with black amino acids predicted to be loop regions or random coil. All programs predict the N-terminus of the propeptide region to contain α -helical structure, up to the putative metalloproteinase cleavage site (green), shown previously to be between Arg 75 and Asp 76 (Wolfman, McPherron *et al.* 2003; Lee 2008). The numbering used here (Fig. 4.1) places these residues as 78 and 79 using a propeptide start site of residue 21, and 98 and 99 for the full-length sequence (numbering in brackets); the published numbering is consistent with a propeptide start site around residue 24, which appears to be favoured by this group (Lee 2004). The rest of the propeptide region is largely β -sheet except for an α -helical segment from amino acids 179-197 (199-217 for the full-length protein). PredictProtein predicts an extra β -sheet in the propeptide region compared to JPred from amino acids 216-221 (236-241). The crystal structure of the growth factor region, which follows the underlined RSRR furin cleavage site, shows a predominance of β -sheet. The only program successful in predicting the wrist helix of the growth factor from amino acids 302-310 (322-330) is Phyre. As Phyre is

traditionally a protein homology recognition engine (Kelley and Sternberg 2009), this success may be due to comparisons to known TGF- β family structures; however, secondary structure predictions using Phyre are performed by independent algorithms prior to homology modeling.

Secondary structure programs are not accurate and predictions cannot be taken as such. However, this information provides initial insight into the unknown structure of the myostatin propeptide region.

1 (21)	JPred PredProt Phyre	VDLNENSEQKENVEKEGLCNACTWRQNTK SSRIEAIKIQILSKLRLETAP VDLNENSEQKENVEKEGLCNACTWRQNTK SSRIEAIKIQILSKLRLETAP VDLNENSEQKENVEKEGLCNACTWR QNTKSSRIEAIKIQILSKLRLETAP
51 (71)	JPred PredProt Phyre	NISKD VIRQLPKAPPRELIDQYDVQR DDSSDGSL EDDDYHATTETIIT NISKD VIRQLPKAPPRELIDQYDVQR DDSSDGSL EDDDYHATETIIT NISKD VIRQLPKAPPRELIDQYDVQR DDSSDGSL EDDDYHATETIIT
101 (121)	JPred PredProt Phyre	MPTESDFLMQVDGKP KCCFFKFSSKIQYNKVVKVKAQLWIYL RPVETPTTVF MPTESDFLMQVDGKP KCCFFKFSSKIQYNKVVKVKAQLWIYL RPVETPTTVF MPTESDFLMQVDGKP KCCFFKFSSKIQYNKVVKVKAQLWIYL RPVETPTTVF
151 (171)	JPred PredProt Phyre	VQILRLIK KPMKDGTTRYTGIRSLKLDMNP <small>G</small> TG IWQSIDVKTVLQNWLKQPE VQILRLIK KPMKDGTTRYTGIRSLKLDMNP <small>G</small> TG IWQSIDVKTVLQNWLKQPE VQILRLIK KPMKDGTTRYTGIRSLKLDMNP <small>G</small> TG IWQSIDVKTVLQNWLKQPE
201 (221)	JPred PredProt Phyre XRCS	SNL GIEIKALDENGHDLAVTFPGPGEDGLNP FLEV KVTDTPKRSRRDFGL SNL GIEIKALDENGHDLAVTFPGPGEDGLNP FLEV KVTDTPKRSRRDFGL SNL GIEIKALDENGHDLAVTFPGPGEDGLNP FLEV KVTDTPKRSRRDFGL DFGL
251 (271)	JPred PredProt Phyre XRCS	DCDEHSTESR CCRYPLTVDFEAFGWDWII APKRYKANYCSGECEFVFLQK DCDEHSTESRC CCRYPLTVDFEAFGWDWII APKRYKANYCSGECEFVFLQK DCDEHSTESRC CCRYPLTVDFEAFGWDWII APKRYKANYCSGECEFVFLQK DCDEHSTESRC CCRYPLTVDFEAFGWDWII APKRYKANYCSGECEFVFLQK
301 (321)	JPred PredProt Phyre XRCS	YPHTHLVHQANPRGSAGPCCTPTKMSP INMLYFNGKEQIIYGKIPAMVVD YPHTHLVHQANPRGSAGPCCTPTKMSP INMLYFNGKEQIIYGKIPAMVVD YPHTHLVHQANPRGSAGPCCTPTKMSP INMLYFNGKEQIIYGKIPAMVVD YPHTHLVHQANPRGSAGPC CTPTKMSPINMLYFNGKEQIIYGKIPAMVVD
351 (371)	JPred PredProt Phyre XRCS	RCGCS RCGCS RCGCS RCGCS

Figure 4.1 Secondary structure predictions for MstnPP.

Analysis of the MstnPP amino acid sequence (21-375) with Jpred (Cuff, Clamp *et al.* 1998), PredictProtein (PredProt) (Rost, Yachdav *et al.* 2004) and Phyre (Kelley and Sternberg 2009) algorithms. XRCS represents the known secondary structure for the myostatin growth factor, taken from the X-ray crystal structure of myostatin (Cash, Rejon *et al.* 2009). α -helix is shown in blue and β -sheet in red. The putative metalloproteinase cleavage site is shown in green. Numbering is from 1-355 (21-275); brackets denote numbering for the full-length myostatin amino acid sequence when the signal peptide is present.

4.2 Biophysical analysis of MstnPP

4.2.1 Circular dichroism spectroscopy

Circular dichroism spectroscopy (CD) was used to assess the secondary structures present in the refolded MstnPP dimer (Fig. 4.2). The spectrum is indicative of a mixture of α -helix and β -sheet and is therefore in agreement with the crystal structure of the mature myostatin growth factor (Cash, Rejon *et al.* 2009) and the mature regions of other TGF- β family members (Schreuder, Liesum *et al.* 2005). It also closely resembles the CD spectra of the BMP-2 precursor protein (Hillger, Herr *et al.* 2005), strongly suggesting that the native MstnPP structure has been obtained.

There are two minima at 217 nm and 209 nm, which are characteristic of β -sheet and α -helical structures respectively (Kelly, Jess *et al.* 2005). The 190 nm maximum supports the presence of α -helix. The absolute negative value at 200 nm implies a degree of intrinsic disorder exists in the structure (Walsh, Hartley *et al.* 1999; Obradovic, Peng *et al.* 2003; Receveur-Brechot, Bourhis *et al.* 2006), which may be functionally significant due to the different roles suggested for the propeptide region of the myostatin precursor, requiring flexibility in the N-terminal domain. The type-I β turn also exhibits a negative value at 200 nm and may also contribute to the spectrum of MstnPP. To our knowledge, this is the first structural analysis for the myostatin precursor protein.

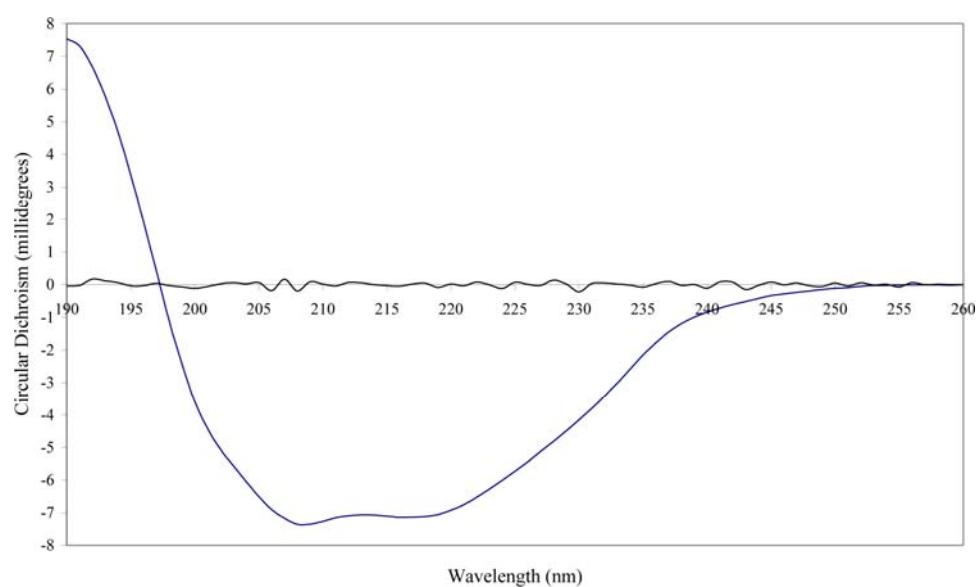


Figure 4.2 Circular dichroism spectrum of MstnPP.

A number of algorithms exist for the deconvolution of CD spectra. However, the accuracy of these is highly dependent on the reference data set used. The spectra for MstnPP was analysed using the CDDN algorithm to have 16% α -helix, 25% antiparallel β -sheet, 10% parallel β -sheet, 17% β -turn and 32% random coil (Appendix 4, Table 12.2); these values are comparable to the conclusions drawn by visual analysis.

It is important to acknowledge that CD analysis gives an averaged measure of the secondary structures present only. The MstnPP species used for CD was purified dimer and the solution is expected to be homogenous. However, the presence of undigested MstnPP after incubation with furin convertase (Chapter 3, Section 3.5.1) indicates that a proportion of the purified dimer solution may be incorrectly folded; if so, these structures will also contribute to the spectrum.

4.2.2 X-ray crystallography

Determination of the crystal structure of MstnPP will provide details about interactions between the propeptide and growth factor regions within the precursor protein. Crystallisation screens were set up using the purified MstnPP dimer and small crystals ($<0.05\text{ }\mu\text{m}$) were obtained in a number of different conditions (Fig. 4.3).

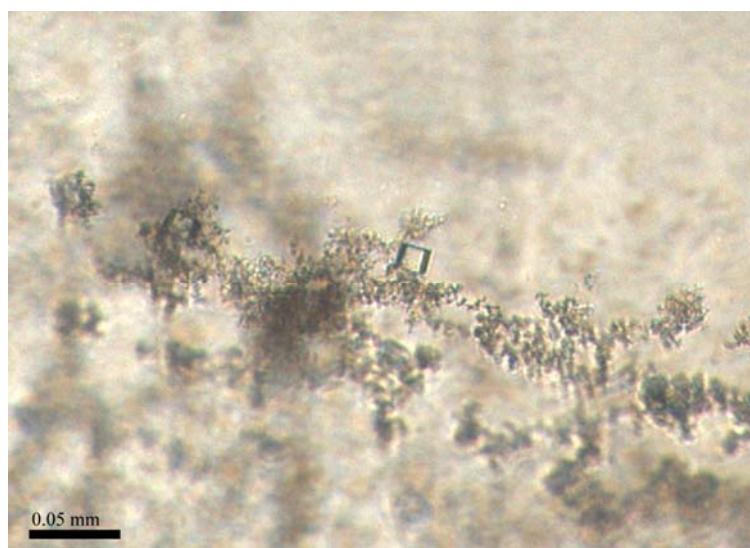


Figure 4.3 Preliminary crystals of MstnPP.

Crystals were obtained in Hampton Research crystallisation screen II condition 22 (0.1 M MES pH 6.5, 12 % (w/v) PEG 20,000).

These could be enlarged by optimisation of some conditions, such as a 2:1 (v/v) ratio of protein to mother liquor, but the resulting crystals do not diffract. The presence of disorder may contribute to this. For example, ordered regions, such as the growth factor domain, may play a role in nucleation, with subsequent growth limited by flexibility in the propeptide domain. This explanation also provides a reason for the small crystal size. The production of diffraction-quality crystals was not achieved within the time frame of this PhD.

4.3 Stability of MstnPP

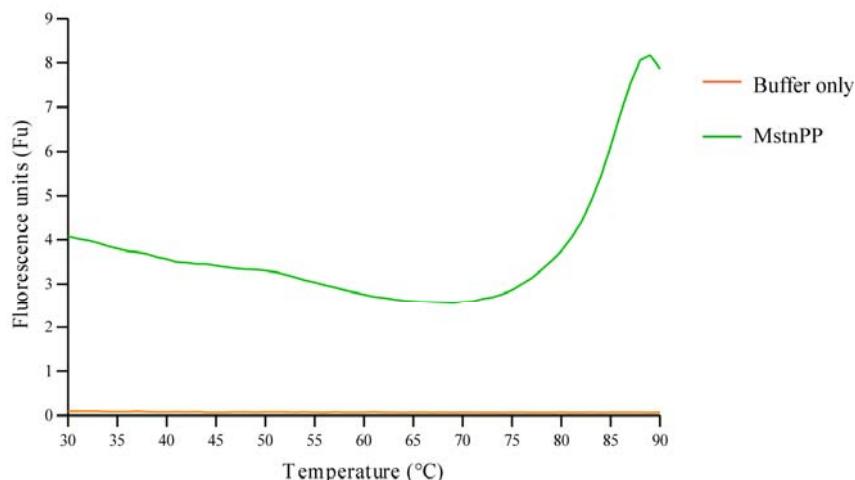
Stability measurements not only provide information about specific protein characteristics but for a refolded protein, give insight into whether or not the native protein fold has been obtained. This is especially useful for a protein such as the myostatin precursor protein for which an activity assay is not possible.

4.3.1 Fluorescence-based thermal shift assay (Sypro Orange)

An estimate of melting temperature (T_m) can be obtained using the fluoro-dye Sypro Orange. The environment-sensitive dye fluoresces upon binding to hydrophobic regions; as protein structure breaks down, fluorescence increases and the midpoint of the unfolding transition is taken as an estimation of T_m (Geerlof, Brown *et al.* 2006). Using a Rotorgene 6000 thermocycler with excitation at 470 nm and fluorescence emission measured at 555 nm, temperature was increased from 30 to 90 °C (Fig. 4.4a). Differentiation of raw data was used to calculate the T_m (Fig. 4.4b).

The thermal shift assay estimates the T_m of WT MstnPP to be 86 °C. This high T_m value is expected as TGF- β proteins are known to be very thermally stable and activation of TGF- β growth factors from the latent complex is routinely achieved by heating at 80 °C (Brownh, Wakefiel *et al.* 1990; Funkenstein and Rebhan 2007). As far as we know, these are the first thermostability measurements published for any TGF- β precursor protein.

a.



b.

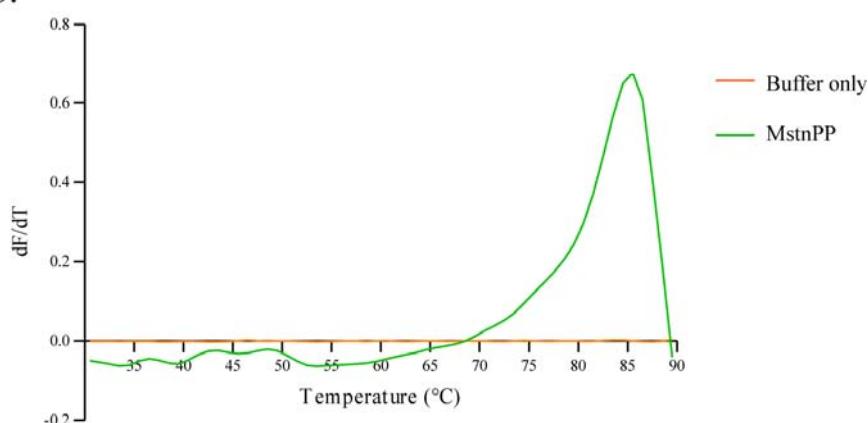


Figure 4.4 Thermal shift analysis of MstnPP by Sypro Orange binding and fluorescence.

- a. Raw data; Sypro Orange fluorescence as a function of temperature.
- b. Differentiation of the curve in a. produces a peak T 86 °C which represents the Tm.
- MstnPP, green; Buffer-only control, orange.

Melting temperature determined using Sypro Orange fluorescence is an estimate only, as it is unknown whether the curve represents the entire protein or a single domain. In the analysis of myostatin and other cysteine-knot containing proteins, the 90 °C limit of the instrument used may prevent detection of the denaturation of the growth factor region. However, these results show high thermal stability for refolded MstnPP and support previous evidence that the native MstnPP structure has been obtained.

4.3.2 CD thermal denaturation

Thermal stability was also assessed by monitoring changes in secondary structure using circular dichroism (CD) spectroscopy. Increasing temperature from 5 to 90°C results in a decrease in intensity of both α -helical (209 nm) and β -sheet (217 nm) minima (Fig. 4.5). However, above 45 °C, an erratic absorption pattern can be observed between 190 and 200 nm (Fig. 4.5a), suggesting alternative behaviours such as aggregation. Irreversible heat-induced aggregation and/or misfolding during CD thermal denaturation has been documented previously (Benjwal, Verma *et al.* 2006). After 80 °C a dramatic change can be observed where the mixed spectrum is replaced by one rich in β -sheet (Fig. 2.5b). These results are reproducible between samples and therefore characteristic of the myostatin precursor protein.

In the two-state unfolding model, a protein moves from a completely folded state to a completely unfolded state (Mayne and Englander 2000), and calculation of the melting temperature (T_m) can be obtained by plotting circular dichroism against temperature at each wavelength of interest to generate a sigmoidal curve. The differential of the curve is used to determine the point of inflection, which represents the T_m . The normalised circular dichroism (percentage of total decrease) versus temperature plot was generated for the 209 and 217 nm minima of MstnPP (Fig. 4.5c). However, neither dataset has a purely sigmoidal shape; both plots are ambiguous and may contain more than one transition. Multi-state unfolding models have been proposed for a number of proteins where a native protein is unfolded via one or more stable intermediates (Mayne and Englander 2000). Both the 209 and 217 nm data have regions that resemble multiple inflection points; however, these may also reflect inconsistencies in the data.

The β -sheet minimum appears to decrease significantly at a lower temperature than the α -helix, suggesting reduced thermal stability of the former. For both plots, the points lose the shape of a smooth curve at higher temperatures; aggregation, known to interfere with thermodynamic analysis of unfolding (Benjwal, Verma *et al.* 2006), may provide an explanation for this.

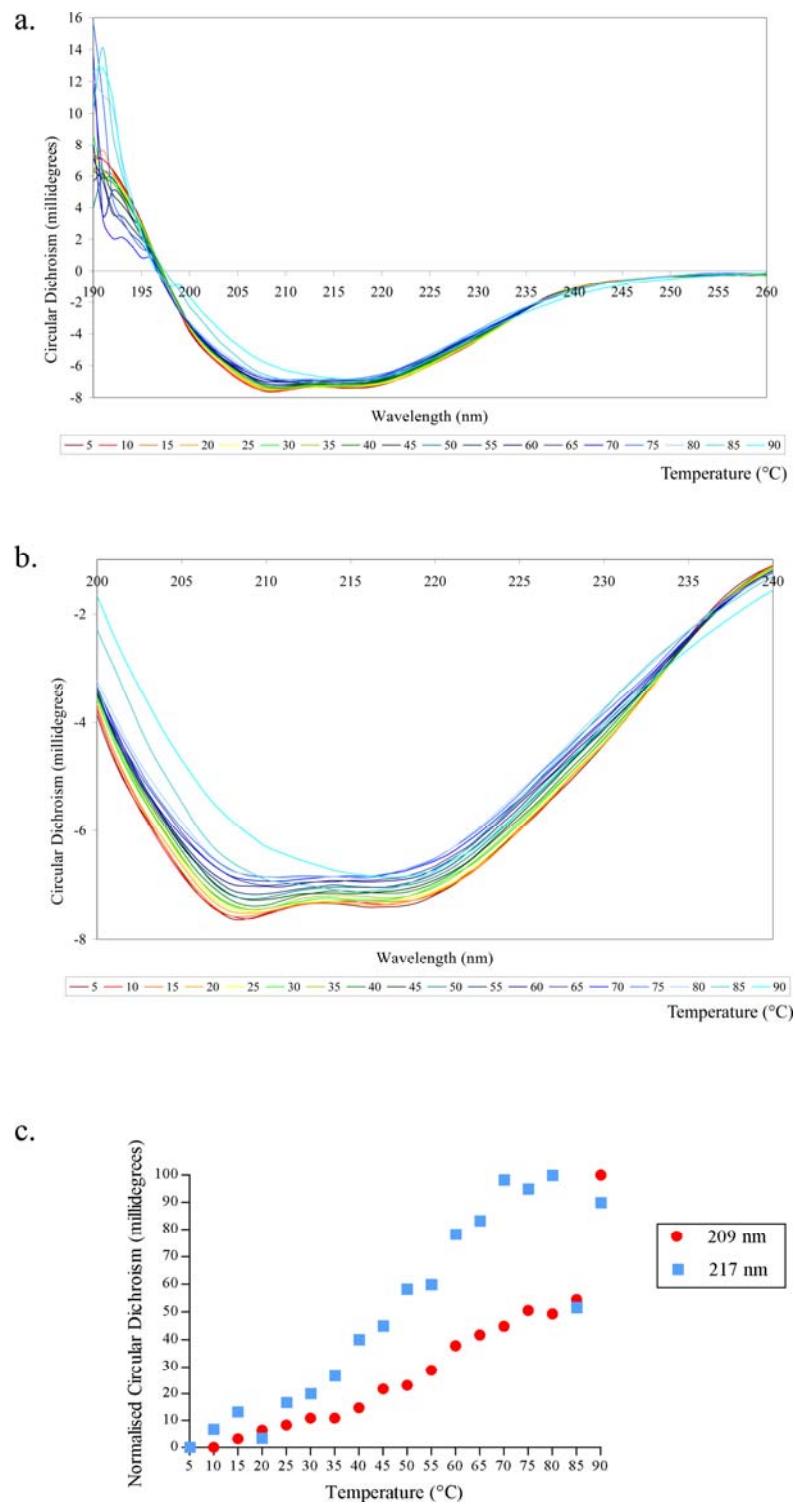


Figure 4.5 CD thermal denaturation of MstnPP.

- CD profile of MstnPP from 190-260 nm from 5-90 °C.
- CD profile as in a. from 200-240 nm.
- Plot of normalised CD vs. temperature for the 209 nm (red) and 217 nm (blue) minima.

The shape and characteristics of the spectra after 80 °C suggest a transition to a β -sheet predominant structure, which then decreases in intensity, suggesting further denaturation. Aggregation has been suggested; another possibility is that this represents the stable cysteine-knot of the growth factor. In either case, the transition indicates that rather than complete unfolding of MstnPP occurring, an alternative structure becomes predominant. The association of β -sheet structure with aggregation is suggestive of amyloid formation, a result that has been observed for a number of proteins such as myoglobin (Benjwal, Verma *et al.* 2006) and ubiquitin (Yang, Larios *et al.* 2003).

The CD thermal denaturation results are quite different to the 86 °C Tm value estimated from Sypro Orange analysis. One explanation for this is that while Sypro Orange analysis results represent melting of the entire myostatin precursor protein, CD data shows the behaviour of all secondary structures present, the data of which may overlap to make interpretation difficult. The fluorescence-estimated Tm is comparable to the temperature of the observed β -sheet transition, suggesting that 86 °C may represent the midpoint of a structural change, or the kinetics of aggregation, rather than denaturation alone.

4.3.3 Limited proteolysis

Limited proteolysis provides information about the structural stability of a protein. If a protein is readily digested, this may be an indication that the polypeptide chain is in a largely unfolded or disordered state (Heiring and Muller 2001). Refolded MstnPP was subjected to proteolysis using a 100:1 ratio (w/w) of MstnPP to trypsin and incubation at 37 °C overnight (18 hours), with samples taken immediately after the addition of trypsin and at 0.5, 1, 2, 3, 4 and 18 hours. MstnPP is quite stable up to 1 hour of digestion at 37°C (Fig. 4.6a), supporting previous evidence for successful refolding of MstnPP. A number of degradation products can be seen at 4 hours and bands are largely absent from the SDS-PAGE gel after overnight incubation indicating complete digestion. However, three faint bands can be seen at 18 hours (Fig. 4.6a and b). Bands are present at 10, 15 and 18 kDa in reducing (R) conditions and 10, 18 and 26 kDa in non-reducing (NR) conditions, suggesting that some regions of the polypeptide chain are resistant to proteolytic digest. The 26 kDa band may be a dimer of the 15 kDa peptide.

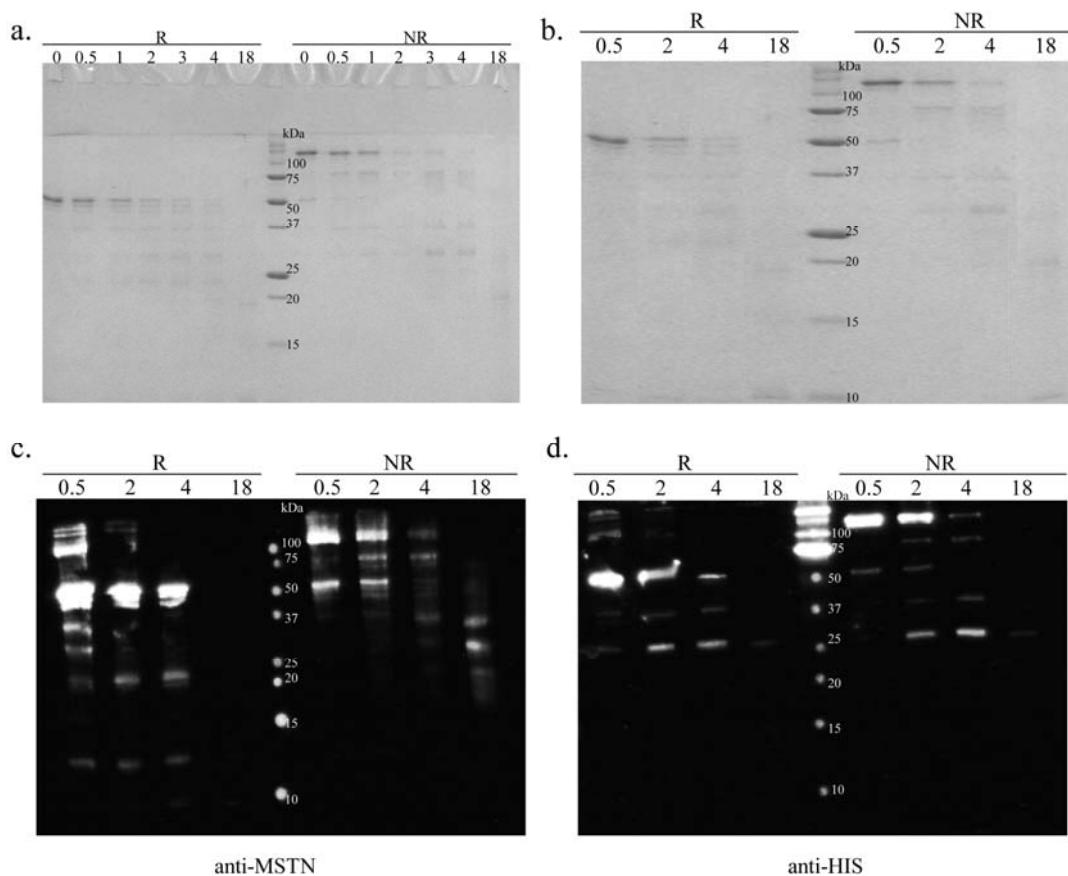


Figure 4.6 Limited proteolysis of MstnPP.

a. and b. Reducing (R) vs. non-reducing (NR) SDS-PAGE.

a. Samples were taken at 0, 0.5, 1, 2, 3, 4 and 18 hours.

b. The 0.5, 2, 4 and 18 hour timepoints were selected for Western analysis (c. and d.)

c. and d. Western blots of gel in b. using antibodies against myostatin (anti-MSTN, c) and the His-tag (anti-HIS, d).

A 100:1 myostatin to trypsin (w/w) ratio was used at 37 °C.

One possibility is structural stability due to the presence of the cysteine knot motif. To investigate this, bands at 0.5, 2, 4 and 18 hours of trypsin digestion (Fig. 4.6b) were analysed by Western blotting using antibodies to the C- (anti-MSTN, Fig 4.6c) and N- (anti-HIS, Fig. 4.6d) termini of His-tagged MstnPP. Overall, the results are ambiguous on comparison with SDS-PAGE and therefore require further investigation. However, some hypotheses can be presented.

A large number of bands are visible throughout the digestion, predominantly the MstnPP monomer and dimer bands at 50 and 100 kDa respectively. Some dimer can

be observed in reducing conditions at 0.5 hours; this is likely to be non-covalently associated dimer as the dimer interface is known to be hydrophobic (Cash, Rejon *et al.* 2009). Other major bands are at 22, 15 and 12.5 kDa using anti-MSTN, and 37 and 26 kDa with anti-HIS. The non-overlapping nature of these band sizes suggests that trypsin cleaves between the N- and C-termini early in digestion, possibly at the dibasic furin recognition site.

There are discrepancies between the Coomassie-stained gel and the Western blot results at 18 hours. The C-terminal antibody recognises 10 and 12.5 kDa bands under reducing conditions, with the removal of reducing agent producing a smear that contains discrete bands at 20, 30 and 35 kDa. The smearing effect may be due to disulphide-bonded oligomers that have a range of sizes. The 12.5 and 35 kDa bands are not visible on the Coomassie-stained gel; however, Western-detection is more sensitive than Coomassie-staining. The 20 and 30 kDa bands may represent the 18 and 26 kDa bands observed by SDS-PAGE. Alternatively, along with the 15 kDa band noted in the Coomassie-stained gel, the latter have neither antigenic region. The 10 kDa MSTN-positive band may represent the stable cysteine-knot region.

Detection of His-tag containing bands produces one species only, at 26 kDa in both reducing and non-reducing conditions. This species is an N-terminal fragment that is either devoid of lysine and arginine residues or has structure that hides these. Analysis of the myostatin amino acid sequence suggests that the latter is likely as there is no place in the primary sequence where an absence of basic residues would produce a 26 kDa peptide. The 26 kDa band may also be due to non-specific binding of the antibody. However, as this band is not visible in all non-reducing lanes, this is a less likely possibility.

4.4 The latent complex

The myostatin latent complex is formed following furin cleavage of the precursor protein. Consisting of propeptide domains non-covalently associated with the growth factor dimer, it is important for growth factor regulation in the majority of TGF- β family members and therefore an attractive target for structural investigation. Structural and biophysical analyses of the myostatin latent complex, produced by digestion of the precursor with furin convertase, was carried out and compared to the undigested myostatin precursor protein.

4.4.1 Circular dichroism spectroscopy

The CD spectra of the latent complex and MstnPP were compared (Fig. 4.7). For direct comparison, circular dichroism in millidegrees was converted into ellipticity ($\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$). The spectra look almost identical; the only difference is the intensity of the peaks. Deconvolution using CDDN suggests an increase in α -helix with concomitant decreases in β -sheet and random coil (Appendix 4, Table 12.2).

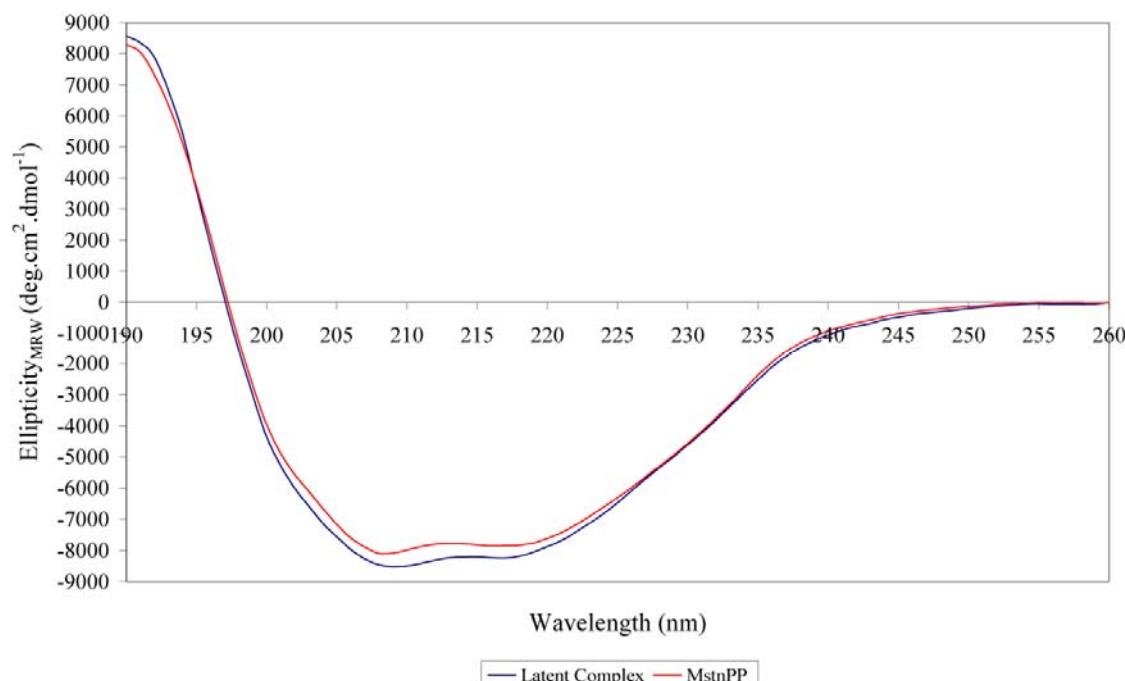


Figure 4.7 CD comparison of myostatin latent complex and MstnPP.

These results may indicate a subtle structural difference occurs on complex formation but may also be due to differences in concentration. The latter is not unexpected, as the same extinction coefficient was used in concentration calculations for both species and that of the latent complex may be affected by differences in the environment of key aromatic residues. As CD follows secondary structure only, subtle structural rearrangements are unlikely to produce significant spectral differences.

4.4.2 CD thermal denaturation

To gain information about the stability of the latent complex, CD thermal denaturation of furin-digested MstnPP was performed (Fig. 4.8). As the complex consists of non-covalently associated propeptide and growth factor regions, this analysis and comparison to the undigested precursor (Fig. 4.5) may provide information about the two regions in isolation. Consistent with denaturation of the precursor, the data show erratic behaviour in the region from 190-200 nm for temperatures above 50 °C (Fig. 4.8a).

Similar to precursor results, the α -helical minimum at 209 nm appears to disappear at high temperatures (Fig. 4.8b); however, this transition occurs at the lower temperature of 70 °C for the latent complex. This may be due to the non-covalent association of the propeptide and growth factor regions and represent latent complex dissociation. β -aggregation may also be a possibility for the latent complex, raising the hypothesis that this transition is reflective of amyloid formation. If this is the case, the lower transition temperature for the latent complex may suggest that separation of the two regions facilitates this change.

In comparison to thermal denaturation of MstnPP, the reduction of both 209 and 217 nm minima occur at a similar temperature and the profiles show a more sigmoidal shape (Fig. 4.8c), supporting the hypothesis that the two regions behave in a different manner when linked non-covalently. As suggested previously for the Sypro Orange analysis, the inflection points of these curves may represent the kinetics of aggregation in addition to unfolding.

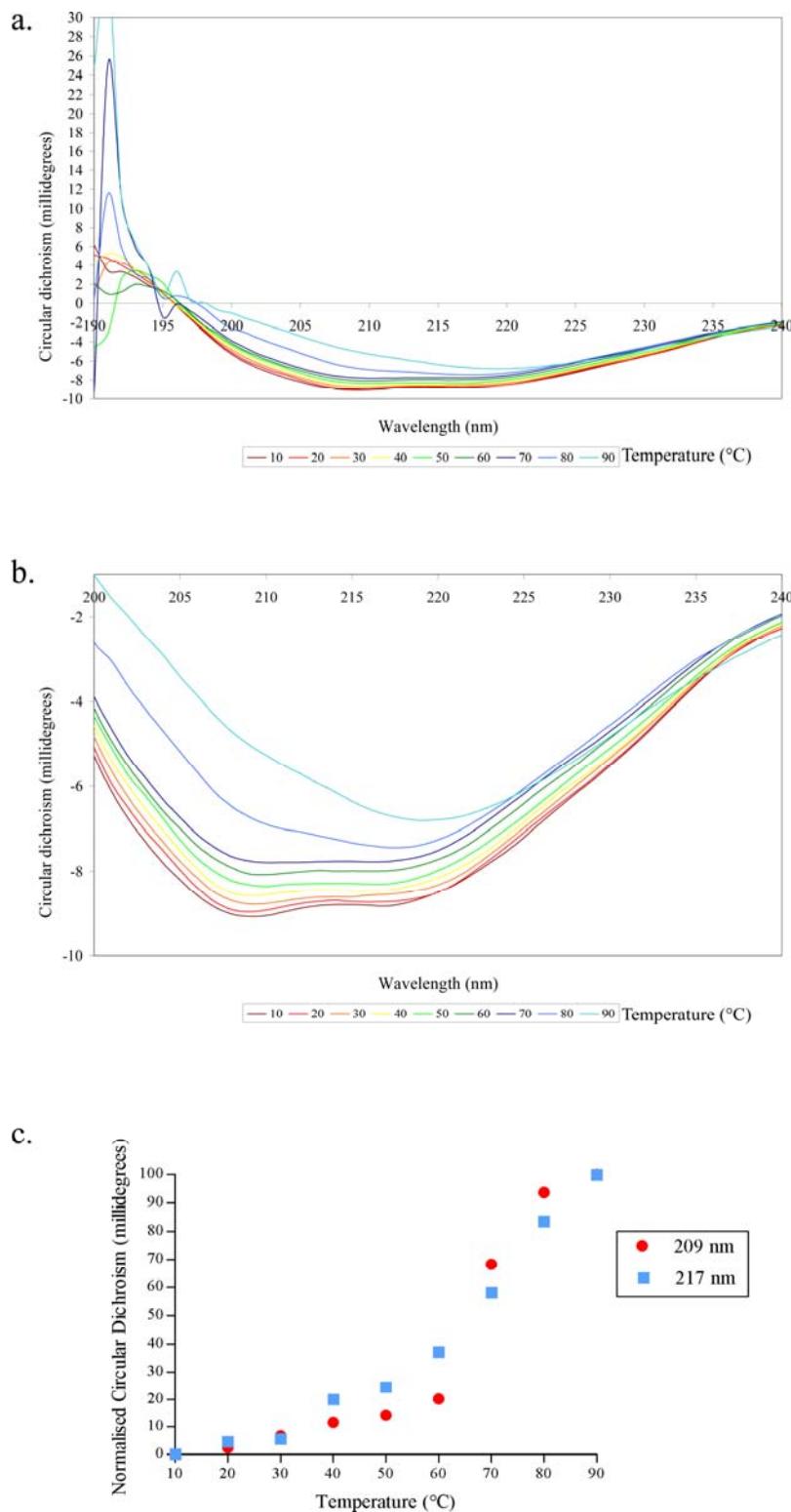


Figure 4.8 CD thermal denaturation of the furin-digested myostatin latent complex.

- CD profile from 190-260 nm from 5-90 °C.
- CD profile as in a. from 200-240 nm.
- Plot of normalised CD vs. temperature for the 209 nm (red) and 217 nm (blue) minima.

4.4.3 Gel filtration chromatography

The furin-digested complex was analysed by calibrated gel filtration chromatography (Appendix 3, Fig. 12.7c) both immediately following digestion (Fig. 4.9) and after acid treatment to cause dissociation of the complex (Fig. 4.10).

There are two major peaks in the elution profile of the latent complex (Figure 4.9a), at 13.7 (1) and 15.4 mL (2). The nature of each peak was visualised using reducing (R) vs. non-reducing (NR) SDS-PAGE (Figure 4.9b). Although both peaks contain propeptide (i) and growth factor monomer (ii), the peaks are clearly distinct and therefore represent two different complexes. The 13.7 mL elution volume of the first peak represents a molecular weight of 100 kDa, 15 kDa less than that of undigested MstnPP (Fig. 3.7). As the digested complex is expected to consist of a growth factor dimer (iii) flanked by two propeptide domains, and retention time on gel filtration is affected by protein shape, the difference in apparent molecular mass may be due to a more compact conformation of the complex. Peak 1 also contains undigested MstnPP monomer (iv) and dimer (v), and consistent with previous results (Fig. 3.8) and published research (Wolfman, McPherron *et al.* 2003), a band at about 70 kDa (vi).

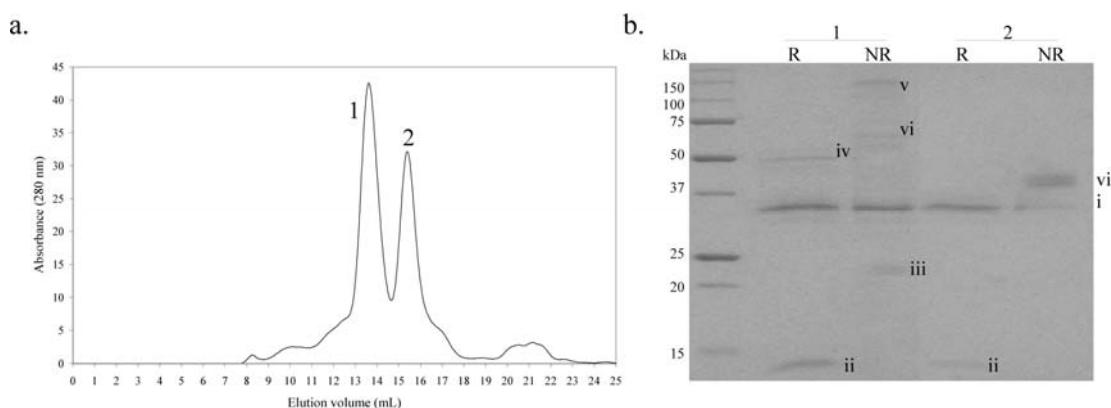


Figure 4.9 Gel filtration analysis of the furin-digested latent complex.

a. S200 gel filtration chromatogram. Peak 1, 13.7 mL; Peak 2, 15.4 mL

b. Reducing (R) vs. non-reducing (NR) SDS-PAGE of peaks 1 and 2. Bands are as follows: i, propeptide; ii, myostatin growth factor monomer; iii, myostatin growth factor dimer; iv, undigested precursor monomer; v, undigested precursor dimer; vi, unknown 70 kDa band; vii, unknown 45 kDa doublet.

At 15.4 mL, the second peak has an estimated molecular weight of 47 kDa. Although reducing SDS-PAGE indicates that this peak contains both propeptide and growth factor, non-reducing conditions produce a doublet (vii) at about 45 kDa with the concomitant disappearance of the growth factor band and partial reduction in the intensity of the propeptide band. One explanation is that this species is a complex of the growth factor dimer plus one propeptide, suggesting either partial dissociation or that the latent complex also exists in a ‘one-propeptide’ state. The absence of this complex in reducing conditions may indicate formation of a disulphide bond between the growth factor dimer and the propeptide; strong, yet non-covalent association is also possible.

Gel filtration of the acid-treated latent complex (Fig. 4.10a) produces a similar first peak at 13.6 mL although the second peak is broad and appears to contain three smaller peaks at 15.5, 16.1 and 16.7 mL. An additional small peak is at 20.4 mL.

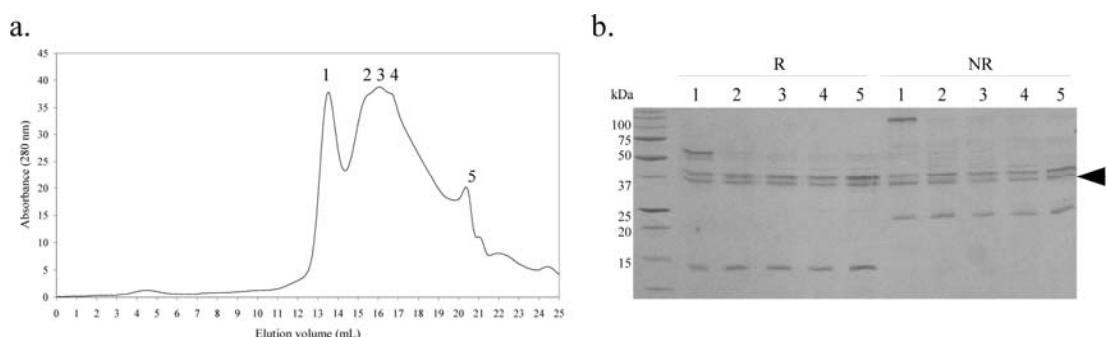


Figure 4.10 Gel filtration analysis of the acid-treated latent complex.

- S200 gel filtration chromatogram. Peak 1, 13.6 mL; Peak 2, 15.5 mL; Peak 3, 16.1 mL; Peak 4, 16.7 mL; Peak 5, 20.4 mL.
- Reducing (R) vs. non-reducing (NR) SDS-PAGE of peaks 1-5. Arrow indicates acid-cleaved propeptide band.

SDS-PAGE shows that the first peak contains undigested precursor, propeptide and the growth factor as observed for the untreated complex (Fig. 4.10b). Fractions sampled across peaks 2 to 4, and peak 5, appear largely similar in both reducing and

non-reducing conditions. However, this is most likely due to the poor resolution of the individual peaks.

Acid treatment is expected to cause dissociation of the latent complex and produce distinct peaks for the growth factor and the propeptide regions. It is likely that some dissociation has occurred as acid treatment has been used previously to produce active growth factor dimer (Fig. 3.9). The appearance of a band at about 35 kDa (Fig. 4.10, arrow) is probably due to degradation of the propeptide via acid-induced cleavage (Asp-X/X-Asp) (Smith 2002) at the metalloproteinase Arg 78-Asp 79 site, and has been observed previously (Wolfman, McPherron *et al.* 2003). Table 4.2 summarises the calculated MWt of each peaks and probable identity of the major species that each peak represents. The 35 kDa band is most likely represented by the 16.7 mL peak; a MWt of 29 kDa is consistent with acid-induced cleavage at Asp 79.

Peak	Retention volume (mL)	Calculated MWt (kDa)	Possible protein species
1	13.6	105	MstnPP Latent complex
2	15.5	47	One propeptide + growth factor dimer
3	16.1	36	Propeptide
4	16.7	29	Growth factor dimer Degraded propeptide
5	20.4	6	Propeptide degradation product

Table 4.2 Identification of proteins and protein complexes from gel filtration of the acid-treated latent complex.

Retention volume (mL) and estimated molecular weight (MWt, kDa) of peaks in Figure 4.10 with possible protein species that each represents.

The undigested 37 kDa propeptide band is still visible and the ratio of the 37 kDa propeptide to the 35 kDa degradation product is approximately 1:1. This suggests that the most likely explanation for these ambiguous results is partial dissociation, a phenomenon that also explains the reduced activity observed for the acid-activated growth factor in the C2C12 proliferation assay (Fig. 3.9).

The 45 kDa doublet observed for peak 2 of the associated complex (Fig. 4.9) is not visible here, suggesting that dissociation of this complex has occurred.

Taken together, these results suggest that the latent complex may have a more compact structure than the undigested precursor. Furthermore, the latent complex may consist of either one or two propeptides bound to the growth factor dimer.

5 Discussion: The structure and function of the human myostatin precursor and the latent complex

To our knowledge, this is the first biochemical and biophysical data shown for the human myostatin precursor protein and latent complex. The characteristics of these species both individually, and on comparison with each other, raise a number of interesting hypotheses with implications for function *in vivo*.

5.1 The structure and properties of human MstnPP

5.1.1 Summary

MstnPP consists of a propeptide domain and a growth factor domain, the latter of which has been structurally determined by X-ray crystallography (Cash, Rejon *et al.* 2009). Secondary structure predictions that agree best with the determined structure were produced using Phyre, the algorithms of which predict the propeptide region of MstnPP to have an α -helical N-terminus and a C-terminus consisting primarily of β -sheet and loop regions. An additional α -helix lies in the middle of the β -sheet domain. Consistent with these predictions, the secondary structure profile of MstnPP is a mixture of α -helix, β -sheet and random coil. As the MstnGF is predominantly β -sheet, the α -helical component is likely to stem primarily from the propeptide. This hypothesis is supported by CD analysis of the TGF- β 1 (McMahon, Dignam *et al.* 1996) and BMP-2 (Hillger, Herr *et al.* 2005) propeptide regions. Although both authors interpreted the CD spectra as β -sheet predominant, absorption at the wavelengths shown (208 and 223 nm) is characteristic of α -helical structure (Kelly, Jess *et al.* 2005).

MstnPP can be successfully crystallised, supporting all evidence suggesting that the recombinant precursor protein is at least partially structured and is likely to contain the native fold. However, the crystals are small and do not diffract. The MstnPP structure may contain disorder, and as the native growth factor is known to be structured, this is likely to come from the propeptide region.

The melting temperature of MstnPP by fluorophore estimation is 86 °C, a thermal stability in line with the TGF- β family. However, CD thermal denaturation suggests

that MstnPP does not follow a two-state unfolding model and shows a β -sheet transition at high temperatures that may indicate either β -aggregation or the stable cysteine knot core. By secondary structure predictions, the α -helical minimum is likely to reflect the propeptide region primarily and the β -sheet minimum both growth factor and propeptide. The β -sheet minimum appears less thermally stable than that of the α -helix; if the high-temperature transition represents the stable cysteine knot core, the early β -sheet unfolding may reflect the β -sheet structures of the propeptide region. However, if β -aggregation is occurring at temperatures above 45°C, this is likely to preclude analysis of the unfolding of the native structure. Chemical denaturation of the BMP-2 precursor protein suggested that the α -helical propeptide region was the least thermodynamically stable (Hillger, Herr *et al.* 2005); however, structural stability may not directly relate to thermal stability in analysis of TGF- β proteins (Muller and Hering 2002). Chemical denaturation of the myostatin precursor protein requires investigation.

The differentiated Sypro Orange fluorescence profile has a peak of 86 °C. While this is usually taken to represent the melting temperature, the similarity of this temperature to that for the β -sheet transition observed during CD thermal denaturation suggests that the fluorescence-estimated T_m represents this transition, whether due to aggregation or cysteine knot stability. Sypro Orange fluoresces when bound to hydrophobic regions, hence both the proposed situations are possible. Aggregation is likely to follow the exposure of hydrophobic regions induced by high temperatures, and a predominant absorption by the cysteine knot assumes unfolding of the propeptide and remaining growth factor regions. It is worth noting that the differentiated fluorescence data (Fig. 4 4b) is not symmetrical and has a lag phase prior to the peak, which may represent the initial stages of denaturation observed by CD.

In either situation, unfolding of the MstnPP produces complex results when analysed over increasing temperature. It is not clear whether this is due to the propeptide or growth factor regions; however, previous analysis of TGF- β family growth factors in isolation have not indicated such results (McMahon, Dignam *et al.* 1996), suggesting that the propeptide region is responsible.

Limited proteolysis of MstnPP supports the suggestion that the propeptide has inherent structure. Two propeptide bands appear stable by Western analysis, one just below 37 kDa up to 4 hours digestion and one after overnight digest at 26 kDa. The 37 kDa band is likely to represent the entire propeptide domain as furin digest results in a propeptide band of this size (Fig. 3.8). The relatively high resistance of this region to trypsin implies that the propeptide is structured in a way that hides many of the potential trypsin cleavage sites in the sequence. The 26 kDa band that remains after 18 hours must contain the N-terminal HIS-tag, indicating that the C-terminus of the propeptide is more flexible by comparison. If this hypothesis is investigated using amino acid sequence, predicted secondary structure (Fig. 4.1) and MWt, the most likely regions for cleavage are around the third α -helix or in the β -sheet that follows this helix. These regions may be more exposed in the tertiary structure and therefore more susceptible to trypsin cleavage.

5.1.2 MstnPP as an intrinsically disordered protein (IDP)

A number of results suggest that the myostatin precursor protein has regions of intrinsic disorder, which may be functional. Intrinsically disordered proteins (IDPs) are ones which fulfil essential biological functions while lacking well-defined secondary and tertiary structures and therefore defy the structure–function paradigm (Iakoucheva, Brown *et al.* 2002). The number of IDPs known to be involved in cell-signalling and regulation is growing rapidly and the inherent disorder is often functional, enabling high specificity binding to a partner that subsequently induces folding of the unstructured region.

Hallmarks of intrinsic disorder

Intrinsic disorder is often indicated during the protein purification process; three major characteristics are aberrant electrophoretic mobility, high apparent molecular mass and resistance to denaturing conditions such as heat and acid (Receveur-Brechot, Bourhis *et al.* 2006).

IDPs have been suggested to bind less SDS than globular proteins due to a unique amino acid composition and have an apparent molecular mass 1.2–1.8 times higher than expected by SDS-PAGE (Receveur-Brechot, Bourhis *et al.* 2006). The myostatin

precursor exhibits aberrant electrophoretic mobility on reducing and non-reducing SDS-PAGE; although by sequence, the predicted molecular weights of the His-tagged MstnPP monomer and dimer are approximately 40 and 80 kDa respectively, the monomer runs as a 50 kDa band (Fig. 3.2, lane I) and the dimer appears at approximately 125 kDa. After furin cleavage the myostatin growth factor runs as expected at 12 kDa (Fig. 3.8), yet the propeptide region runs at 37 kDa despite its theoretical MWt of 30 kDa, suggesting disorder is localised to the propeptide primarily. Previous research has shown that glycosylation of the myostatin propeptide domain expressed in mammalian cells may be responsible for aberrant mobility on SDS-PAGE (Jiang, Liang *et al.* 2004). However, the myostatin protein produced for this thesis was expressed in *E. coli*; although evidence of glycosylation in *E. coli* exists (McBride, Yu *et al.* 2000), this is rare. Furthermore, the inclusion body localisation of MstnPP is expected to prevent glycosylation in this instance.

By gel filtration, the MstnPP dimer has a retention volume of 13.3 mL (Fig. 3.7b), representing a molecular weight higher than would be expected for a globular protein, indicating a relatively open or elongated structure. The absence of diffraction-quality MstnPP crystals supports both the presence of disordered regions and an open, more flexible structure, as diffraction requires uniform, ordered packing within a crystal. Resistance to heat and acid is hard to interpret as the myostatin growth factor domain has high stability in denaturing conditions (Brownh, Wakefiel *et al.* 1990). Fluorophore analysis and CD thermal denaturation suggest a structural transition at 86 °C, rather than a complete loss of structure. Acid resistance was not tested directly; however, hydrochloric acid treatment of the furin-digested complex, performed for activation of the growth factor dimer, did not cause precipitation of the propeptide, indicating acid-stability.

Structural characteristics

A number of tools are used to analyse the global structure and shape of an IDP, including small angle X-ray scattering (SAXS) and nuclear magnetic resonance (NMR) spectroscopy (Receveur-Brechot, Bourhis *et al.* 2006). Although these methods are not within the scope of this PhD, some information can be obtained from the CD spectrum of MstnPP.

If the MstnPP CD spectrum (Fig. 4.2) is compared to that of a typical IDP, there are notable differences. The CD spectrum of the latter is typified by a largely negative ellipticity at 200 nm, negligible ellipticity at 222 nm and ellipticity close to zero at 185 nm (Receveur-Brechot, Bourhis *et al.* 2006). The only feature of the MstnPP CD spectrum that is similar is overall negative ellipticity at 200 nm. As CD is an averaging technique, the highly ordered growth factor regions of MstnPP, as well as those predicted to be ordered in the propeptide, are likely to predominate the spectrum. Therefore, if it exists, intrinsic disorder is not expected to be the major structural aspect, but localised in discrete domains.

Pinpointing intrinsically disordered regions of MstnPP

Which regions of MstnPP are intrinsically disordered? A number of *in silico* algorithms predict intrinsic disorder in proteins via analysis of the amino acid chain. Five were selected for the analysis of MstnPP on the basis of their predominance in the literature, year of publication and means by which disorder is calculated (He, Wang *et al.* 2009). These are PONDR (Romero, Obradovic *et al.* 1997), DisEMBL (<http://dis.embl.de/>), PrDOS (Ishida and Kinoshita 2007), GlobPlot (<http://globplot.embl.de/>) and FoldIndex (Prilusky, Felder *et al.* 2005).

The sequence of MstnPP from 1-355, which corresponds to amino acids 21-375 (numbering in brackets), contains four regions with a significant likelihood of disorder, 1-10, 73-99, 221-222 and 227-249 (Fig. 5.1, D, green and underlined, and Appendix 5, Fig. 12.8). These regions were selected using the criteria that at least 3 out of the 5 algorithms predicted a high propensity for disorder. Interestingly, all regions are localised to the propeptide region, except for the furin cleavage site (underlined in black) and four amino acids at the N-terminus of the growth factor region (267-269) which are in a loop region of the crystal structure (Cash, Rejon *et al.* 2009). The predicted secondary structure from Chapter 4 is shown below the disorder prediction (S), with α -helices in blue and β -sheet in red. Two of the four predicted disordered regions also contain predicted secondary structure; this is minimal and is primarily β -sheet.

1 (21)	D	VDLNENSEQK ENVEKEGLCN ACTWRQNTKS SRIEAIKIQI LSKLRLETAP
	S	VDLNENSEQK ENVEKEGLCN ACTWR QNTKS SRIEAIKIQI LSKLRLETAP
51 (71)	D	NISKDVI RQL LPKA PPLREL IDQYDVQRDD SSDGSL EDDD YHATTEIIT
	S	NISKDVI RQL LPKA PPLREL IDQYDVQRDD SSDGSL EDDD YHAT TETIIT
101 (121)	D	MPTESDFL MQ VDGKP KCCFF KFSSKI QYNK VVKA QLW IYL RPVETPTTVF
	S	MPTESDFL M MQ VDGKP KCCFF KFSSKI QYNK VVKA QLW IYL RPVETPTT VF
151 (171)	D	VQILRLIKPM KDGTRYTGIR SLKLD MNPGT GIWQS IDV KT VLQNWLKQPE
	S	VQILRLI KPM KDGTRYT GIR SLKLD MNPGT GIWQS IDV KT VLQNWLKQPE
201 (221)	D	SNLGIEIKAL DENGHD LAVT FFPGPGE DGLN PFLEV KVTDT PKRSRRDFG L
	S	SNLGIEIKAL DENGHD LAVT FFPGPGE DGLN PFLEV KVTDT PKRSRRDFG L
251 (271)	D	DCDEHSTESR CCRYPL TVDF EAFGWD WIIA PKRYKANYCS GECEFVFLQ K
	S	DCDEHSTESR CCRYPL TVDF EAFGWD WIIA PKRYKANYCS GECEFVFLQ K
301 (321)	D	YPHTHLVHQ A NPRGSAGPCC TPTKM SPINM LYFNGKE QII YGKIPAMVVD
	S	YPHTHLVHQ A NPRGSAGPCC TPTKM SPINM LYFNGKE QII YGKIPAMVVD
351 (371)	D	RCGCS
	S	RCGCS

Figure 5.1 Intrinsic disorder versus secondary structure in MstnPP.

Disorder, D; Secondary structure, S. Regions of predicted intrinsic disorder (consensus of PONDR, DisEMBL, PrDOS, GlobPlot and FoldIndex) are in green and underlined. Predicted (Phyre) and known (Cash, Rejon *et al.* 2009) secondary structure shows α -helix in blue and β -sheet in red. Numbering is from 1-355 (21-275); brackets denote numbering for the full-length myostatin amino acid sequence when the signal peptide is present.

Also used in the study of disorder is hydrophobic cluster analysis (HCA), a two-dimensional, helical representation of a protein sequence that allows intuitive visualization of clusters of hydrophobic amino acids. In general, these correspond to secondary structure elements in globular proteins (Gaboriaud, Bissery *et al.* 1987); in contrast, non-globular regions are generally poor in hydrophobic residues and rich in polar residues (Karlin, Ferron *et al.* 2003). HCA of MstnPP using the drawhca algorithm (Callebaut, Labesse *et al.* 1997) highlights a predominance of charged residues in the first two predicted regions of disorder (Fig. 5.2, 1 and 2). There is a proline-rich stretch near the third (3), which is also an indicator of disorder (Karlin, Ferron *et al.* 2003). The last region (4) shows some propensity for secondary structure although is also relatively charge-rich and contains the furin recognition motif RSRR. Trypsin proteolysis indicates that the most stable region in the propeptide is the N-terminus. Intriguingly, the MWt of this region indicates cleavage just prior to the last two regions of predicted disorder.

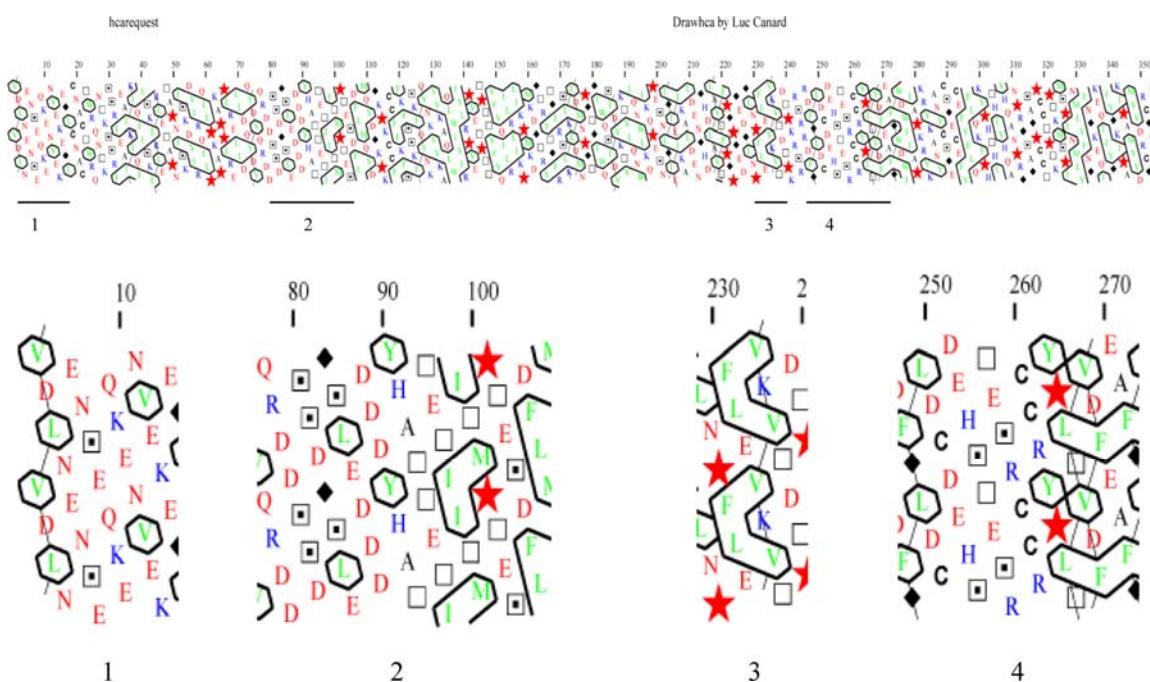


Figure 5.2 Hydrophobic cluster analysis for MstnPP.

The drawhca algorithm (Callebaut, Labesse *et al.* 1997) was used. Symbols denote certain amino acids: red star, proline; black diamond, glycine; open square, threonine; open square with dot inside, serine. Blue and red lettering show positively- and negatively-charged amino acids respectively; green indicates hydrophobic residues. Cysteines are shown in bold black. Predicted structures are contained within black lines.

Taken together, the analyses presented so far support the hypothesis that MstnPP has regions of intrinsic disorder, localised primarily to the propeptide region.

Intrinsic disorder and amyloid

The CD spectrum of MstnPP shows erratic behaviour between 190 and 200 nm and a β -sheet transition at higher temperatures. One explanation for this unexpected profile is β -aggregation and the formation of amyloid structures. In line with this hypothesis, IDPs have been shown to possess a high propensity for amyloid formation (Tompa 2009). During *in vitro* refolding of MstnPP, large molecular weight soluble aggregates form spontaneously (Fig. 3.4), indicating that the myostatin precursor protein has a high propensity for aggregation.

Formation of the myostatin latent complex

A predominant feature of IDP biochemistry is coupled folding and binding, where an unstructured peptide or protein domain folds into an ordered structure on partner binding (Ward, Sodhi *et al.* 2004). Molecular recognition involving IDPs has a number of important functional advantages, for example, disordered regions can bind their targets with high specificity and low affinity. In addition, intrinsic disorder promotes binding diversity by enabling proteins to interact with numerous partners (Iakoucheva, Brown *et al.* 2002).

One hypothesis is that MstnPP contains intrinsically disordered regions that are also capable of folding into secondary structures such as β -sheet after furin cleavage and latent complex formation. In this situation, the disorder may be functional, allowing flexibility prior to furin cleavage and post-cleavage, high-specificity binding to the growth factor domain for effective inhibition.

A number of lines of evidence support structural differences between the myostatin precursor protein and latent complex. First, the CD spectrum of the latent complex has increased intensity compared to the precursor protein (Fig. 4.7). The shape is similar, suggesting structural rearrangements are likely to be subtle, without a major impact on the CD spectrum. The myostatin latent complex has a spectrum similar to that of the TGF- β 1 latent complex (McMahon, Dignam *et al.* 1996), although the β -sheet minimum is more profound for myostatin. CD analysis of the TGF- β 1 latent complex and comparison to the CD spectra of TGF- β 1 growth factor and propeptide regions alone, as well as an addition spectra of the latter two, suggested that a conformational change takes place on latent complex formation (McMahon, Dignam *et al.* 1996). Although this work did not present the CD spectrum of the TGF- β 1 precursor protein, the BMP-2 precursor protein and propeptide CD spectra (Hillger, Herr *et al.* 2005) are similar to those of MstnPP and TGF- β 1 respectively. The agreement that exists between myostatin, TGF- β 1 and BMP-2, may indicate that a similar phenomenon involving intrinsic disorder and conformational change exists for all TGF- β family members during latent complex formation.

Second, CD thermal denaturation spectra show differences between the latent complex and MstnPP, especially for the 209 nm minimum, which is likely to represent the propeptide domain primarily. The propeptide may therefore undergo the largest structural alterations. Although a quantitative comparison cannot be made due

to ambiguity, a significant β -sheet transition, that occurs at a reduced temperature in the latent complex, is apparent. This transition may represent denaturation of the propeptide domain, β -aggregation, or dissociation of the latent complex. It is also possible that all these situations occur and influence each other. For example, if the transition reflects β -aggregation, separation of the propeptide and growth factor regions may facilitate the transition.

The third characteristic of the latent complex that suggests a difference in structure compared to the precursor protein is a difference in retention time by gel filtration. While the MstnPP dimer elutes at 13.3 mL (Fig. 4.7b), corresponding to a molecular weight of 115 kDa, the equivalent peak after furin digest elutes at 13.7 mL (Fig. 4.9a), corresponding to a 100 kDa species and suggesting a more compact structure. However, 100 kDa is still significantly larger than the 80 kDa theoretical MWt by amino acid number, indicating that the complex maintains an elongated structure. It must be noted that retention time by gel filtration is also influenced by protein concentration; in general, a decrease in concentration results in an increased retention time. As the absorbance at 280 nm is over ten-fold lower for the latent complex compared to the precursor protein, estimated molecular weight differences may also be explained by this.

5.2 Implications for MstnPP and latent complex function *in vivo*

5.2.1 β -Aggregation by MstnPP

CD thermal denaturation profiles show a marked transition to a β -sheet rich spectrum at high temperatures, suggestive of β -aggregation, a characteristic of amyloid formation (Benjwal, Verma *et al.* 2006). What implications does this have for the myostatin precursor protein and latent complex *in vivo*? For the native protein, β -aggregation is not expected to occur, as both the MstnPP and latent complex spectra showed stability at physiological temperature. However, recent studies have suggested a role for aggregated MstnPP in the amyloid disease sporadic inclusion body myositis (sIBM) (Wojcik, Engel *et al.* 2005; Wojcik, Nogalska *et al.* 2007; Askanas and Engel 2008). An investigation into amyloid formation by MstnPP is presented in Chapters 8 and 9.

5.2.2 The propeptide as a chaperone

The myostatin precursor protein is the translated form of myostatin and evidence suggests that the propeptide region in this form plays a role in the correct folding of the myostatin growth factor (Jin, Dunn *et al.* 2004; Funkenstein and Rebhan 2007). Refolding of the human myostatin precursor was successful in a protocol designed for zebrafish precursor and not in that used for the porcine protein (Chapter 3) despite a higher sequence similarity between human MstnPP and the latter. If sequence differences are important, residue 144 (Fig. 5.3, circled in yellow), which is lysine in pig and glutamic acid in human and zebrafish, may contribute to refolding differences and therefore have a role in chaperone activity. Glu 144 is within a region of the MstnPP that is relatively scarce in both predicted secondary structure and disorder (Fig. 5.3). If this residue is involved in folding of the growth factor, the loop-location coupled with an absence of disorder would impart flexibility, yet stability. Consistent with this hypothesis, Glu 144 is in the C-terminus of the myostatin propeptide domain, suggested previously to be important for stabilisation (Jiang, Liang *et al.* 2004).

51 (71)	NISKDVIRQL LPKAPPLREL IDQYDVQRDD SSDGSLEDDD YHATTEIIT NISKDVIRQL LPKAPPLREL IDQYDVQRDD SSDGSLEDDD YHATTEIIT
101 (121)	MPTE SDFLMQ VDGKPCKCCFF KFSSKIQYNK VVKAQLWIYL RPVE TPTTVF MPTE SDFLMQ VDGKPCKCCFF KFSSKIQYNK VVKAQLWIYL RPVE TPTTVF
151 (171)	VQILRLIKPM KDGTRYTGIR SLKLDMNPGT GIWQSIDVKT VLQNWLKQPE VQILRLIKPM KDGTRYTGIR SLKLDMNPGT GIWQSIDVKT VLQNWLKQPE
201 (221)	SNLGIEIKAL DENGHDLAVT FPGPGEDGLN PFLEVKTDT PKRSRRDFGL SNLGIEIKAL DENGHDLAVT FPGPGEDGLN PFLEVKTDT PKRSRRDFGL

Figure 5.3 The propeptide as a chaperone.

Section of sequence analysis taken from Figure 5.1. Disorder, green; α -helix, blue; β -sheet, red. The glutamic acid residue that may be involved in propeptide-assisted growth factor folding is circled in yellow.

A second hypothesis is that the predicted disorder may play a role in the chaperone-like activity of the propeptide domain by imparting flexibility and allowing regions to

undergo numerous different noncovalent interactions with the growth factor as it folds. The challenge of TGF- β folding lies predominantly in the structure of the cysteine knot and therefore, the formation of native disulphides. Flexible regions of the myostatin propeptide may be able to interact transiently with specific growth factor regions during the folding process to enable the formation of one disulphide over another, or to promote the function of disulphide isomerases in the endoplasmic reticulum. Furthermore, the disorder-prone regions contain a high proportion of charged residues which may be able to stabilise charges in the growth factor domain during folding. As the growth factor also has a high proportion of hydrophobic residues, another possibility is that charged regions of the propeptide shield these hydrophobic regions from the environment to prevent aggregation during folding.

Primary sequence analysis shows that the myostatin propeptide contains an EF-hand calcium-binding domain (Table 4.1). This has not been documented for any other TGF- β family member and NPS@ searches on TGF- β 1-3, BMP-2, inhibin- α , activin A and GDF-11 do not show this motif. The reasons for this motif in the myostatin sequence, if any, are unclear. Calcium plays an important role in ER processes; since chaperone-assisted folding occurs in the ER, one possibility is that the propeptide uses calcium as a cofactor for chaperone activity or that calcium binding may be involved in the recruitment of other chaperones.

5.2.3 Export of myostatin

As a member of the TGF- β family, the myostatin precursor is most likely processed in the trans-Golgi by furin convertase and exported as a latent complex; complex formation has been suggested to be important for export (Gray and Mason 1990; McMahon, Dignam *et al.* 1996; Jiang, Liang *et al.* 2004; Walton, Makanji *et al.* 2009).

The N-terminus of the myostatin propeptide contains one conserved *N*-glycosylation motif (NISK, Fig. 1.2). As glycosylation plays an important role in the export and subsequent bioactivity of TGF- β 1 and TGF- β 2 (Brunner, Lioubin *et al.* 1992), it may perform a similar function for myostatin; there are no glycosylation motifs in the growth factor domain supporting a necessity for the propeptide.

Evidence suggests that the myostatin precursor protein is also secreted. In contrast to the latent complex which is secreted into the serum, MstnPP is found extracellularly in skeletal muscle, suggesting a local role (Anderson, Goldberg *et al.* 2008). Therefore, while formation of the latent complex may not be necessary for the export of myostatin, the propeptide domain probably is.

Residues 99-266 of the propeptide have been suggested to be involved in stabilization of myostatin (Jiang, Liang *et al.* 2004), which may be necessary for export of both the complex and precursor protein. Although evidence supports some level of structural rearrangement on latent complex formation, a proportion of the propeptide/growth factor interactions involved in stabilization may be conserved between the precursor and the latent complex. Residues 99-266 are predicted to be β -sheet-predominant and contain less random coil relative to the N-terminus; an increase in structure may provide the proposed stability. However, limited proteolysis results suggest decreased structural stability of this region relative to the N-terminus (Fig. 4.6). Therefore, further investigation is necessary.

5.2.4 Inhibition by the myostatin propeptide

The role of the propeptide in binding and inhibition of growth factor activity is well-documented (Hill, Davies *et al.* 2002; Jiang, Liang *et al.* 2004; Yang and Zhao 2006). One possibility is that the propeptide blocks the receptor-binding site of the growth factor in a manner similar to follistatin (Cash, Rejon *et al.* 2009) and/or BMP-7 (Sengle, Ono *et al.* 2008); however, no structural evidence to support this has been published.

Deletion mutagenesis has shown residues 42-115 of the propeptide to be responsible for inhibition (Jiang, Liang *et al.* 2004). This region is predicted to consist of two N-terminal α -helices and a C-terminal β -sheet (Fig. 4.1), separated by a region of predicted intrinsic disorder (Fig. 5.1). Interestingly, the disordered region contains the putative metalloproteinase cleavage site involved in post-secretion activation of the mature growth factor (Wolfman, McPherron *et al.* 2003; Lee 2008). The predicted intrinsic disorder may therefore be functional, allowing protrusion of the sequence for cleavage and growth factor activation.

After cleavage, structuring of disordered regions may result in enhanced interactions of the propeptide with the growth factor dimer, for high affinity and specificity binding and inhibition.

How might cleavage result in release of the myostatin growth factor? *In vitro*, heat and acid treatment can be used to activate latent myostatin (Zimmers, Davies et al. 2002; Funkenstein and Rebhan 2007). Acid treatment is most likely to have a similar effect as *in vivo* metalloproteinases, with cleavage at one of the many aspartate residues present in the inhibitory domain (Smith 2002). Heat-induced release probably stems from denaturation of the propeptide, especially if some regions of propeptide are protruding and relatively flexible.

5.3 The structure of the latent complex

5.3.1 Structural modelling

The results presented in this thesis, combined with other studies of the interactions between the propeptide, or follistatin, and the mature growth factor for both myostatin (Jiang, Liang *et al.* 2004) and other TGF- β family members (McMahon, Dignam *et al.* 1996; Walton, Makanji *et al.* 2009), provide insight into the structural characteristics of the human myostatin latent complex.

The crystal structure of myostatin in complex with follistatin (Cash, Rejon *et al.* 2009) shows the mechanism by which follistatin prevents receptor binding by the myostatin growth factor (Fig. 5.3a-c); two molecules of follistatin wrap completely around one myostatin growth factor dimer to block the putative Type I and Type II receptor binding sites (Fig. 5.3a). One follistatin molecule consists of four domains, an N-terminal domain (ND) and three follistatin domains (FSD1-3) (Fig. 5.3b and c); while ND interacts to block the Type I receptor binding site located on the concave surface of the growth factor, the Type II receptor binding site on the convex surface is blocked by FSD1 and FSD2.

Site-directed mutagenesis of inhibin and activin growth factors indicates a number of hydrophobic residues likely to be important for the interaction of the TGF- β growth factors with their respective propeptide domains in the latent complex (Walton, Makanji *et al.* 2009). In myostatin these residues are F290, I298, P301, I348 and M350, and map to the FSD1-blocked region of the Type II receptor binding site (Fig. 5.4d, red). Interactions with the propeptide in this region may act in a similar manner

to follistatin to block Type II receptor binding. The corresponding myostatin propeptide residues possibly involved in the myostatin latent complex interaction are I53, I60 and L64, which map to predicted α -helical secondary structure (Fig. 4.1).

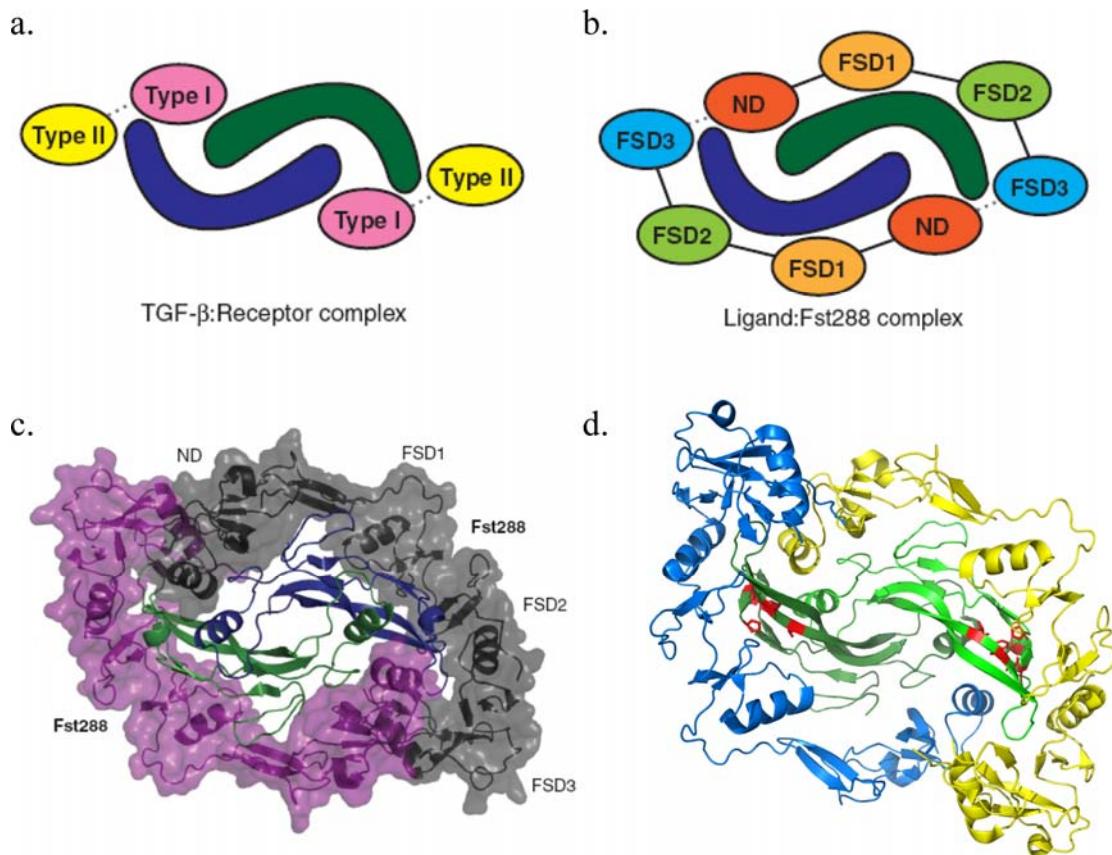


Figure 5.4 Interactions of myostatin with follistatin, TGF- β family receptors and the myostatin propeptide.

- a. Regions involved in receptor binding by TGF- β .
 - b. Schematic of ligand antagonism by follistatin.
 - c. The crystal structure of the myostatin:follistatin complex showing interactions in b.
 - d. The myostatin:follistatin structure with residues proposed to be important for propeptide binding by the inhibin/activin growth factors (Walton, Makanji *et al.* 2009) highlighted in red.
- a.-c. were taken from (Cash, Rejon *et al.* 2009) and d. was constructed using PyMol.

These results are consistent with those of deletion mutagenesis in myostatin that implicated residues 42-115 as the region of the propeptide important for inhibition (Walton, Makanji *et al.* 2009), and site-specific mutagenesis in the TGF- β 1 propeptide that characterized residues 50-85 as important for interaction with the

TGF- β 1 growth factor residues (Sha, Yang *et al.* 1991). In addition, velocity sedimentation experiments of the BMP-7 latent complex or free growth factor, together with the BMP Type II receptor, show that the BMP-7 propeptide domain competes with the growth factor for Type II receptor binding (Sengle, Ono *et al.* 2008).

An alignment of myostatin propeptide residues 40-120 with FSD1 of follistatin (Fig. 5.5a) indicates that the predicted hydrophobic residues (arrows and underlined) and α -helical secondary structure in the myostatin propeptide align to an α -helical region in FSD1. The myostatin/follistatin crystal structure (Cash, Rejon *et al.* 2009) suggests that the interaction between FSD1 and myostatin is primarily via two loop regions (Fig. 5.5a, underlined and Fig. 5.5b (i) and (ii)). The first loop aligns to I53 of the myostatin propeptide and is in close contact with P301 of the myostatin growth factor ((i), solid arrow); however, the second loop does not appear to contact the putative propeptide hydrophobic interaction domain of the growth factor (Walton, Makanji *et al.* 2009) ((i) and (ii), dashed arrows). The FSD1 α -helix does not appear to have a prominent involvement. It is likely that the binding mechanism of both FSD1 and the myostatin propeptide shows some conservation; the FSD1/myostatin binding site contains hydrophobic interactions and involves at least one of the hydrophobic residues that may interact with the propeptide within the latent complex (Walton, Makanji *et al.* 2009). However, alternative binding mechanisms for FSD1 exist.

In deletion mutagenesis studies, the C-terminus of the myostatin propeptide, residues 120-262, were not necessary for inhibition and therefore suggested to have a role in stabilization (Jiang, Liang *et al.* 2004). However, it has been proposed that the receptor-binding mode of TGF- β 1, which myostatin appears to resemble most closely, may involve initial binding to the type II receptor with subsequent recruitment of the type I receptor (Keah and Hearn 2005; Groppe, Hinck *et al.* 2008). If this is also the mode used by the myostatin growth factor, blocking of the type II receptor binding site by residues 42-115 would prevent observation of any downstream inhibition effected by the C-terminus of the propeptide. Although the C-terminus alone was not sufficient to inhibit myostatin growth factor activity (Jiang, Liang *et al.* 2004), this region may require the presence of residues 42-115 for growth factor interaction.

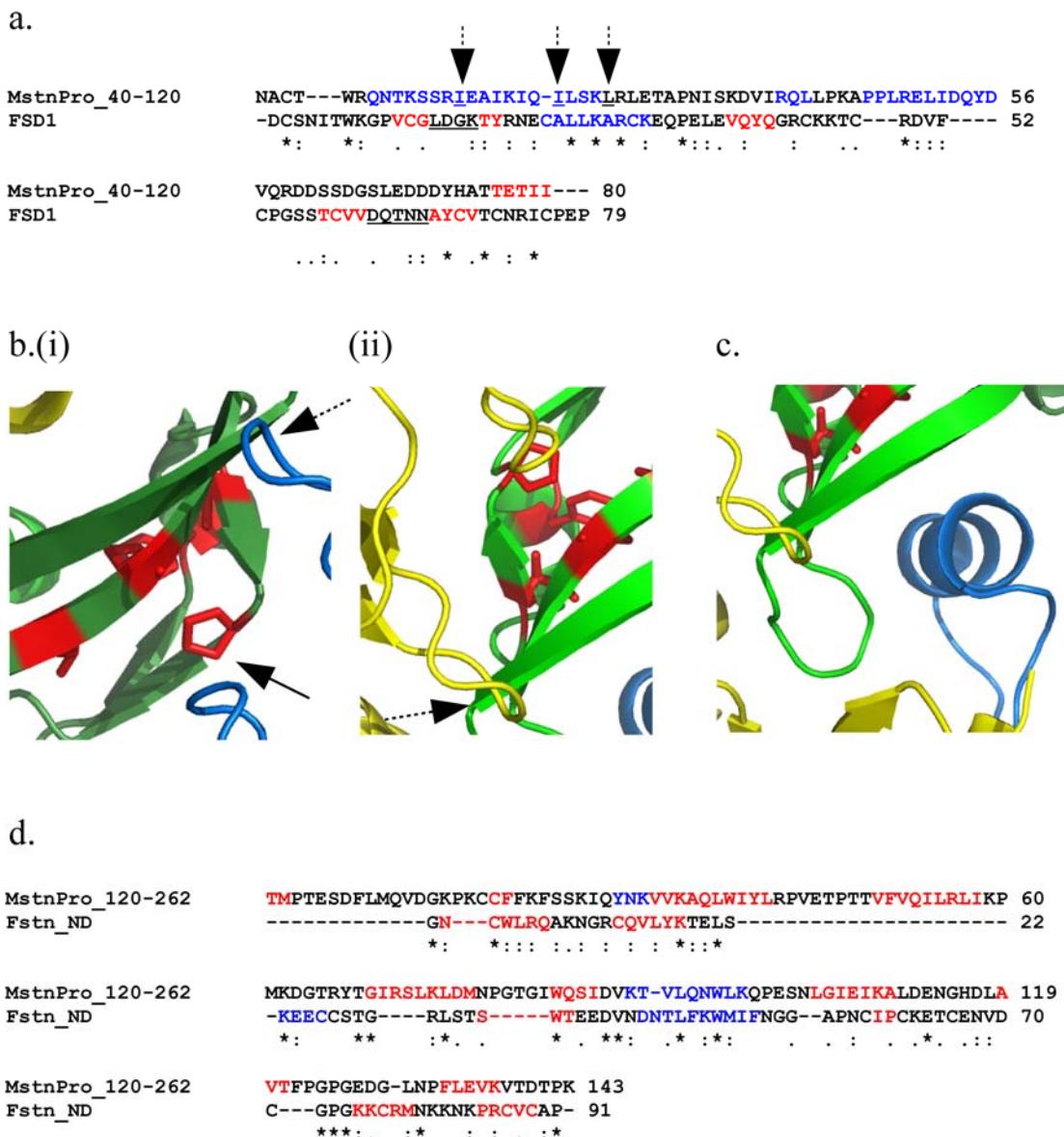


Figure 5.5 Myostatin interaction domains of follistatin versus the myostatin propeptide.

- a. Sequence alignment of myostatin propeptide residues 40-120 and FSD1 of follistatin. Residues proposed to interact with the myostatin growth factor are indicated by an arrow and underlined. Loop regions of FSD1 that appear to be involved in growth factor interactions are underlined.
- b. Zoomed-in view from Figure 5.4d of FSD1 loop 1 ((i), solid arrow) and loop 2 ((i) and (ii), dashed arrows) interactions with myostatin. (i) and (ii) represent the two myostatin monomers.
- c. Zoomed-in view from Figure 5.4d of ND helix and growth factor interaction.
- d. Sequence alignment of myostatin propeptide residues 120-262 and ND of follistatin . Sequence alignments were performed with ClustalW (Thompson, Higgins *et al.* 1994) and show α -helix in blue, β -sheet in red. ClustalW similarity key is as follows: *, exact match; :, strong similarity; ., weak similarity.

The N-terminal domain of follistatin (ND) plugs the type I receptor binding site of myostatin, primarily via contacts between the ND helix and the fingertip region of myostatin (Fig. 5.5c). If residues 120-262 of the myostatin propeptide are compared to the sequence of ND (Fig. 5.4 b), the helical regions align, indicating that a similar mode of interaction may occur between the myostatin growth factor and the C-terminus of the myostatin propeptide. For example, the latter may also function to block the Type I receptor binding site and play a role in growth factor inhibition.

If interactions between the propeptide and both type I and type II receptor binding sites of the growth factor are involved in the myostatin latent complex, two propeptides may completely wrap around one growth factor dimer, in a manner similar to that shown by follistatin (Cash, Rejon *et al.* 2009).

5.3.2 The role of cysteine residues in the propeptide domain

The myostatin propeptide domain contains 4 cysteine residues, Cys 39, Cys 42, Cys 137 and Cys 138. However, whether these are involved in intra- or intermolecular disulphides, or in interactions with other proteins *in vivo*, is not known.

Disulphide bonding between the propeptide and growth factor within the latent complex has been discounted previously (Lee and McPherron 2001; Thies, Chen *et al.* 2001; Jiang, Liang *et al.* 2004) and observation of the propeptide at the same MW_t on both reducing and non-reducing SDS-PAGE (Fig. 3.8) suggests an absence of disulphide bonding between propeptide domains (Jiang, Liang *et al.* 2004). However, after furin digest of the myostatin precursor protein, bands at 70 and 47 kDa were observed (Figs. 3.8 and 4.9 respectively) in non-reducing conditions, with a concomitant reduction in intensity of the propeptide band and disappearance of the growth factor for the latter. The 70 kDa band has been observed previously (Wolfman, McPherron *et al.* 2003) but was suggested to represent misfolded protein. While this cannot be discounted, the results presented in this thesis suggest the formation of novel complexes between the myostatin propeptide and growth factor dimer, which may involve disulphide bonding. The 70 kDa band may reflect dimerisation of two free propeptide domains. Intermolecular disulphides between TGF- β family propeptides have been shown previously; Cys223 and Cys225 of the TGF- β 1 propeptide region are necessary for latent complex formation and growth

factor inhibition (Brunner, Marquardt *et al.* 1989). Alternatively, both complexes may represent covalent interaction between the propeptide and growth factor dimer. This hypothesis has been discounted previously for the myostatin latent complex (Lee and McPherron 2001; Thies, Chen *et al.* 2001; Jiang, Liang *et al.* 2004) and all 9 cysteines of the growth factor are involved in either intramolecular disulphide bonds or dimerisation in the isolated growth factor (Cash, Rejon *et al.* 2009). However, the disulphide-bonds of the myostatin growth factor may differ within the latent complex, the structure of which has not been determined. Furthermore, the 70 and 47 kDa bands may indicate the formation of novel complexes within which disulphide bonding interactions differ to those of the traditionally-observed latent complex. An alternative possibility is that the 70 and 47 kDa bands reflect strong yet non-covalent association.

Intra-propeptide disulphide bonds are also possible. For example, covalent interactions between Cys 39 and Cys 137, or Cys 42 and Cys 138, may contribute to structural stability, in a similar manner to the Cys 272-Cys 281 disulphide of the myostatin growth factor domain (Cash, Rejon *et al.* 2009).

An interaction between the latent TGF- β binding protein, LTBP-3, and the secreted myostatin precursor has been shown (Anderson, Goldberg *et al.* 2008). Unlike the LTBP/TGF- β propeptide interaction that involves a disulphide bond between Cys 33 of TGF- β and the cysteine repeat domain of LTBP, the myostatin/LTBP-3 interaction appears to be non-covalent (Anderson, Goldberg *et al.* 2008). An interaction between the processed myostatin latent complex and LTBP-3 has not been investigated; if it occurs, the mechanism of this interaction may differ from that of the precursor and involve disulphide bonding.

5.3.3 One or two propeptides per growth factor?

The above modeling and discussion propose that two propeptides may wrap around one growth factor in a manner similar to that shown by follistatin. Previous research on the latent myostatin complex suggests that two propeptide domains associate with one disulphide-linked growth factor dimer (Thies, Chen *et al.* 2001) and studies

carried out for other TGF- β family members indicate the same 2:1 ratio, two propeptides to one growth factor dimer (Sengle, Charbonneau *et al.* 2008).

Gel filtration results suggest that after furin cleavage two complexes may exist, a large molecular weight complex composed of two propeptide regions per growth factor dimer and a novel, smaller structure consisting of one propeptide plus the growth factor dimer (Fig. 4.9).

The most likely explanation is partial dissociation of the latent complex during acid treatment; this also explains the lower levels of myoblast proliferation inhibition seen for the acid-treated complex (Berry, Thomas *et al.* 2002). If this is the case, a 2:1 propeptide to growth factor dimer ratio in the myostatin latent complex is supported.

5.3.4 Model: Formation of the myostatin latent complex from a partially structured precursor

What might the structure of the myostatin precursor protein look like, and what rearrangements, if any, might occur during latent complex formation? A model, based on the results and discussion presented in this thesis, has been constructed (Fig. 5.3), with both a schematic diagram (Fig. 5.3a) and the corresponding more detailed structural model (Fig. 5.3b) shown. The structural model is based around the crystal structure of the myostatin growth factor dimer (Cash, Rejon *et al.* 2009), and the secondary structures predicted by Phyre algorithms (Fig. 4.1); however, it must be noted that the structure of the propeptide as depicted in Figure 5.3 is speculative. Only the α -helical secondary structures important for this discussion are shown; although β -sheet and loop interactions may also be important, no evidence exists to support this possibility at this stage.

The precursor protein (Fig. 5.3a(i) and b(i)) is likely to dimerise in the ER with formation of the cysteine-knot core of the growth factor, with the furin cleavage sites (black arrows) exposed. The propeptide region is predicted to contain regions of intrinsic disorder (black stars); while disorder may block interactions between the C-terminal helix of the propeptide (red star (a) and square (b)) and growth factor, conserved hydrophobic interactions between the N-terminal helix of the propeptide and the fingertip regions of the growth factor may be possible (green stars (a) and ovals (b)).

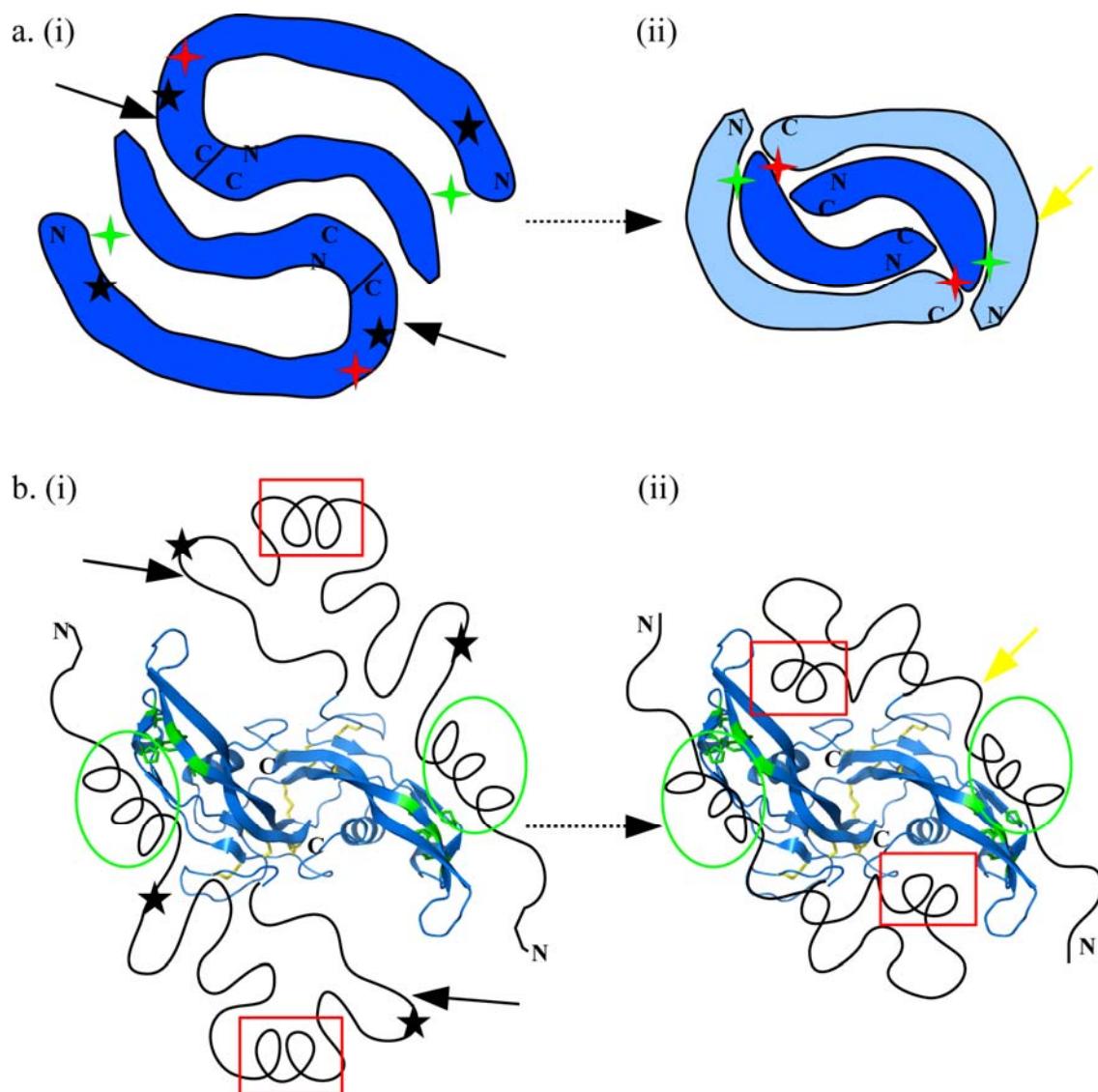


Figure 5.6 Model for the structural characteristics of myostatin latent complex formation.

a. Schematic model.

b. Structural model using the myostatin growth factor crystal structure (Cash, Rejon *et al.* 2009).

(i) The myostatin precursor protein and (ii) the myostatin latent complex.

Key is as follows: Black stars, regions of predicted intrinsic disorder; Green stars (a) and ovals (b), interaction between the N-terminus of the propeptide and the growth factor; Red stars (a) and squares (b), interaction between a putative helical region in the propeptide and the concave surface of the growth factor; Black arrows, furin cleavage site; Yellow arrows, metalloprotease cleavage site.

Furin cleavage results in a release of the growth factor from the propeptide domain and may allow structural rearrangement of the propeptide C-terminus for enhanced interaction between the C-terminal helix and the type I receptor-binding interface of the growth factor, similar to that observed in the follistatin complex (Fig. 5.3a(ii) and b(ii), green). Therefore, this implies induced structuring of the intrinsically disordered regions of the precursor protein.

Overall, this model states that two propeptides may wrap around one growth factor dimer to block both type I and type II receptor binding sites. The interactions shown here are focused around the fingertip region of the myostatin growth factor. Although other interactions have not been shown in the literature to date, these may also exist to stabilise and/or enhance complex formation. For example, disulphide bonds between propeptide domains are a possibility. The approximate site of metalloproteinase cleavage is indicated (yellow arrow). The mechanism by which cleavage results in dissociation is unclear; however, one possibility is that cleavage disrupts these stabilising interactions. Although no conclusions about the nature of the complex after furin cleavage can be drawn, conformational rearrangement of the propeptide region is possible and this may result in the formation of a more compact inhibitory structure than that of MstnPP.

Summary

Little is known about the structure of the myostatin precursor protein and latent complex, *in vivo* or *in vitro*. These chapters present *in vitro* biophysical analysis and novel results for both species. The myostatin propeptide region is likely to contain a significant amount of structure, as well as some functional intrinsic disorder. Furthermore, MstnPP may be prone to β -aggregation and amyloid formation above physiological temperature. The latent complex exhibits structural differences by comparison to MstnPP, suggesting a conformational change on latent complex formation, and a model for this transition and the structure of the latent complex has been proposed.

6 Analysis of C313Y, the Piedmontese myostatin mutation

Myostatin null mutations occur in a number of species including sheep, racing whippets, humans and a number of cattle breeds (Kambadur, Sharma *et al.* 1997; Lee 2004; Lee 2007). Piedmontese cattle produce a mutant myostatin protein that has a point mutation, G938A, in the codon for one of the cysteine residues of the cysteine knot, resulting in the replacement of cysteine at position 313 with a tyrosine (C313Y). Mutated protein is translated, processed and exported from the cell. Although detected circulating in the serum of Piedmontese cattle (Berry, Thomas *et al.* 2002), signalling does not appear to occur. Understanding the mechanisms behind this will provide insight into the signaling of wild-type myostatin and possibilities for therapeutic intervention. The structural analysis of the C313Y mutant protein, and comparison to the wildtype protein is an important aspect of this understanding.

The C313Y mutation was produced by site-directed PCR mutagenesis in the human precursor and refolded for biochemical and biophysical comparison to the WT precursor protein.

6.1 Production of the C313Y precursor protein (C313Y MstnPP)

6.1.1 Site-directed mutagenesis

The splice overlap extension (SOE) method of site-directed mutagenesis was used to produce the C313Y myostatin construct (Chapter 2). In brief, reverse complementary primers spanning the region to be mutated were used with the F21 and B375 primers designed for production of the WT construct. This produced two sections of the cDNA with overlapping regions containing the G938A mutation, that were used in a second PCR reaction to produce full-length myostatin mutant precursor cDNA beginning at amino acid 21. This construct is referred to as C313Y MstnPP and was cloned into the pProEX-Htb vector for expression as done for WT myostatin. Sequencing was carried out to ensure that mutagenesis had been successful (Figure 6.1).

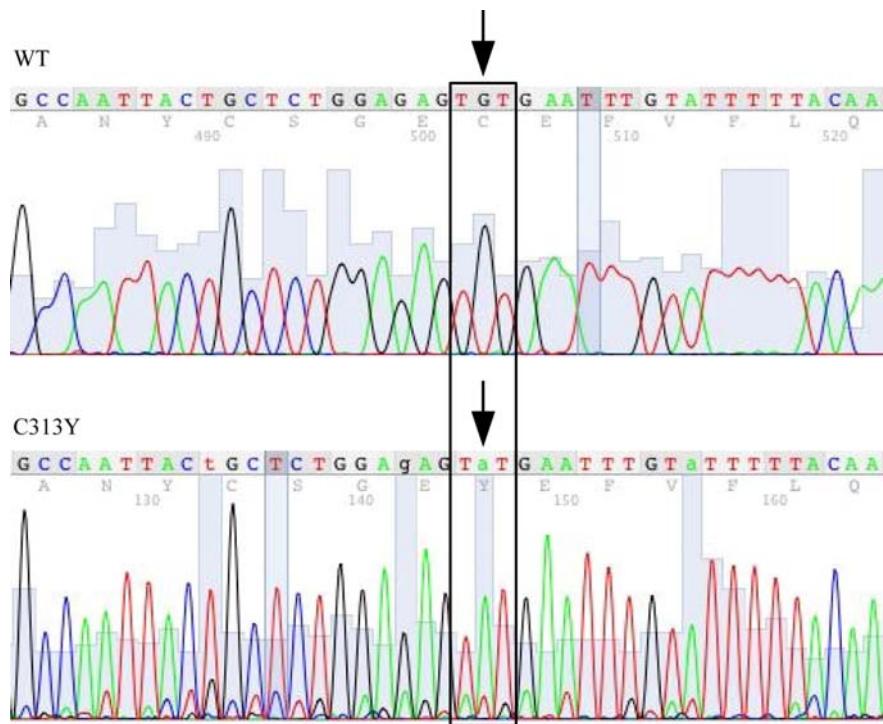


Figure 6.1 Nucleotide sequencing of C313Y cDNA.

Comparison of WT and C313Y sequences shows G938A nucleotide substitution.

6.1.2 Expression

The C313Y MstnPP construct was expressed, refolded and purified using identical procedures to those employed for the WT protein and results at each stage were compared.

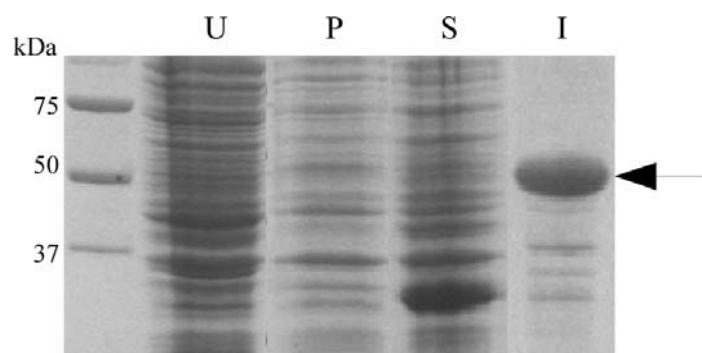


Figure 6.2 Expression of the myostatin C313Y precursor protein.

Lanes are as follows: U, uninduced; P, post-induction; S, soluble fraction; I, insoluble fraction. C313Y protein is indicated by an arrow.

There were no notable differences in expression (Figure 6.2). Post-expression analysis by reducing SDS-PAGE revealed little difference by way of expression level (lane P) or proportions of soluble versus insoluble protein (lanes S and I respectively). All C313Y MstnPP was contained within insoluble bacterial inclusion bodies (arrow). In addition, no differences in other steps involved in expression, such as transformation efficiency and handling of the pellets, were observed.

6.1.3 Refolding

The first noticeable difference between the WT and mutant precursor proteins was apparent following refolding. Although refolding of the C313Y MstnPP dimer was achieved, efficiency was reduced markedly in comparison to WT MstnPP, as seen using reducing (R) versus non-reducing (NR) SDS-PAGE (Figure 6.3). This result is consistent with the hypothesis that the cysteine knot may be important for folding (Muller and Heiring 2002).

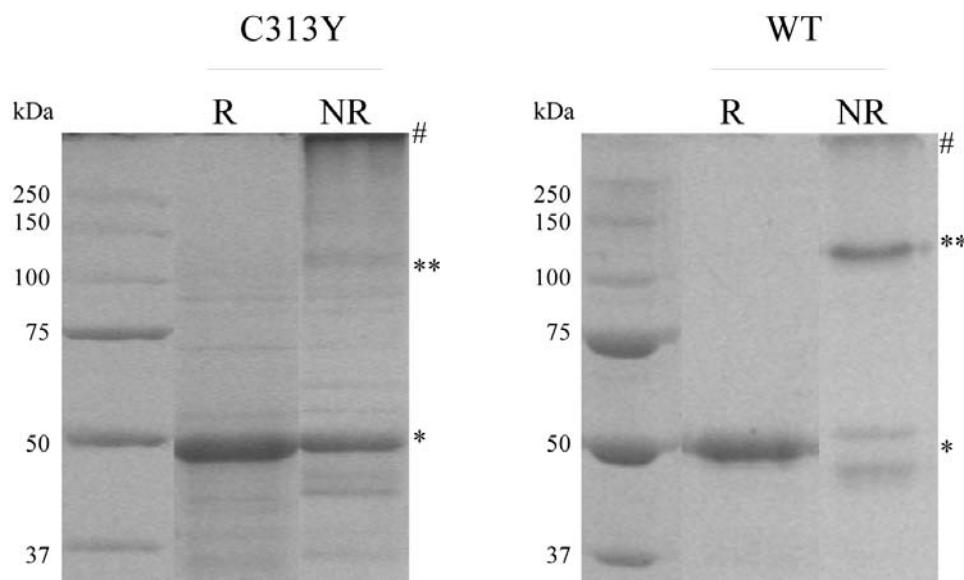


Figure 6.3 Refolding of the C313Y and WT precursor proteins.

Reducing (R) and non-reducing (NR) SDS-PAGE. Bands are labeled as follows: *, monomer; **, dimer; #, soluble aggregates. Intervening gel lanes have been removed.

The C313Y dimer band is very faint compared to the WT dimer (** bands) with concomitant increases in the levels of monomeric (*) and aggregated (#) C313Y protein, indicating that reduction in dimer intensity is not due to reduced myostatin

overall. Intriguingly, in non-reducing conditions, the C313Y monomer does not display the mobility differences that the WT monomer does. A possible explanation for this is that removal of one disulphide bond reduces the extent of differently folded forms produced during refolding.

6.1.4 Purification

The purification of refolded C313Y MstnPP yielded further differences. Heparin affinity chromatography of the mutant protein (Fig. 6.4a) appears almost identical to that of the WT (Fig. 3.7a).

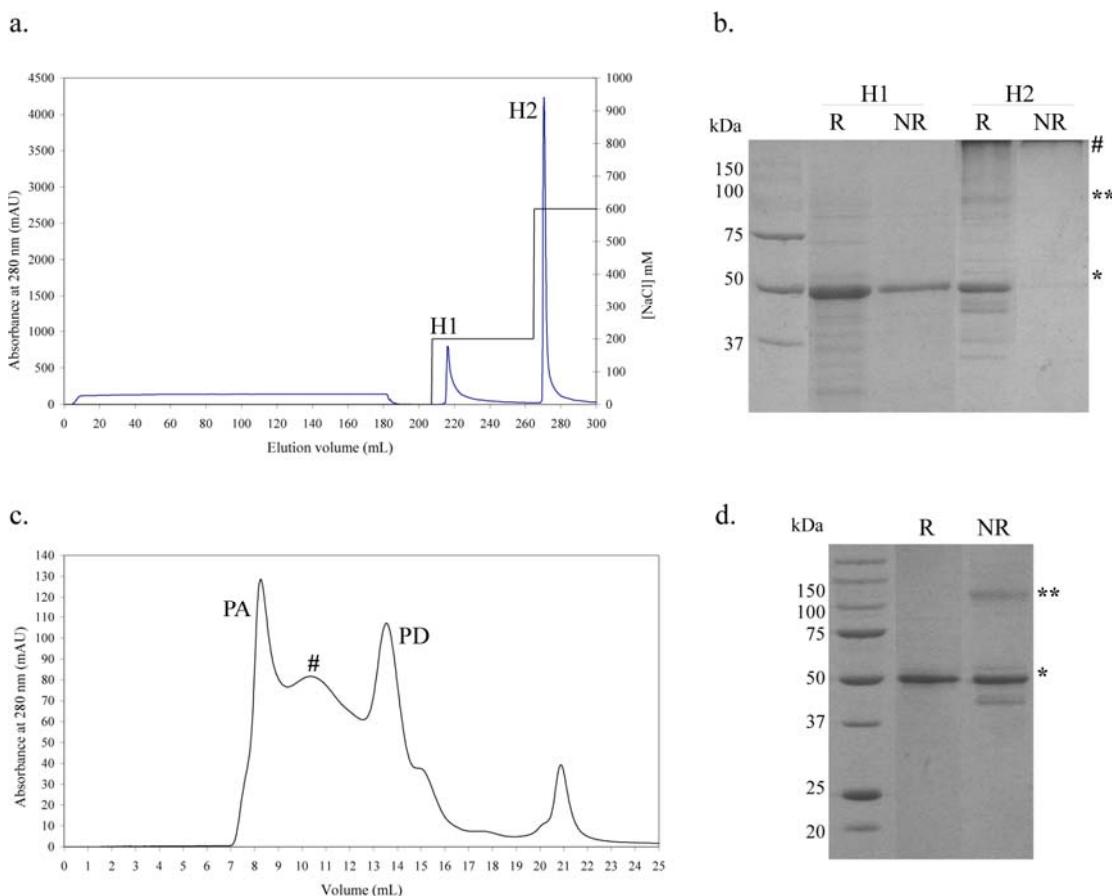


Figure 6.4 Purification of the refolded C313Y precursor protein.

- Heparin affinity chromatography.
- Reducing (R) vs. non-reducing (NR) SDS-PAGE of H1 and H2 from heparin affinity chromatography. Bands are labelled as follows: *, monomer; **, dimer; #, soluble aggregates.
- S200 gel filtration chromatography of heparin peak H1. PA, purified aggregates; PD, purified dimer; # aggregates of mixed molecular weight.
- Reducing (R) vs. non-reducing (NR) SDS-PAGE of peak PD from gel filtration chromatography. Bands are monomer (*) and dimer (**).

It is comprised of two large peaks (H1 and H2), the second of which (H2) is predominant. These peaks contain the same species of protein seen for the WT purification (Fig. 6.4b) with the first consisting of C313Y dimer (**) and some aggregated protein (#) and the second consisting entirely of aggregate. Consistent with previous results, the refolded C313Y dimer population is markedly reduced compared to the WT. The ratio of peak H2 to peak H1 is greater for the C313Y MstnPP purification; at 5.4:1 this is almost double that of the WT, suggesting that aggregation is enhanced during refolding of the mutant. Purification of H1 using gel filtration (Fig. 6.4c) results in isolation of the C313Y dimer (Fig. 6.4c, PD and 6.4d,***) from aggregated protein (PA). The proportion of dimer remains reduced compared to the WT, with the majority of protein present in monomeric form (*). The C313Y heparin separation does not remove aggregated species as effectively as that of the WT, shown by the presence of a broad peak similar to that seen for the WT protein without heparin separation (Fig. 3.6) at 10.5 mL (Fig. 6.4c, #). This volume represents an average MW_t of 380 kDa and most likely contains C313Y MstnPP oligomers of varying sizes; 380 kDa reflects the formation of an octomer. The mutant dimer peak resolved by gel filtration has a retention time of 13.6 mL and estimated molecular weight of 105 kDa, reduced by comparison to the WT, which may be explained by a more compact structure or concentration differences. The predominance of a monomer in this peak by NR SDS-PAGE (Fig. 6.4d,*) suggests that dimerisation is non-covalent; the C313Y mutation may hinder the potential for covalent dimerisation.

6.2 Verification of refolding success

6.2.1 Production of the C313Y growth factor

The C313Y growth factor was produced by the WT furin digest protocol (Chapter 2 and Fig. 6.5). Reducing (R) SDS-PAGE highlights the production of both the growth factor monomer (1) and the propeptide region (2). There is also some undigested C313Y MstnPP remaining (3). The C313Y growth factor dimer is not visible by non-reducing (NR) SDS-PAGE, even in Western blots, consistent with previous results (Figs. 6.3 and 6.4) that indicate covalent dimerisation may be limited in the mutant. However, epitope inaccessibility may also contribute to this result, as suggested for the WT growth factor dimer (Fig. 3.8). As a faint amount of undigested C313Y

MstnPP dimer can be observed (4), dimerisation may be facilitated in the precursor protein.

A C313Y growth factor monomer band can be observed in NR conditions (5) of both SDS-PAGE and anti-MSTN Western analysis. This runs faster than the reduced form of the protein indicating disulphide-bond formation. Anti-HIS analysis detects propeptide only; undigested C313Y MstnPP is not visible, despite its detection by the anti-MSTN antibody (3). The most likely explanations are a difference in loading between gels and different antibody sensitivities.

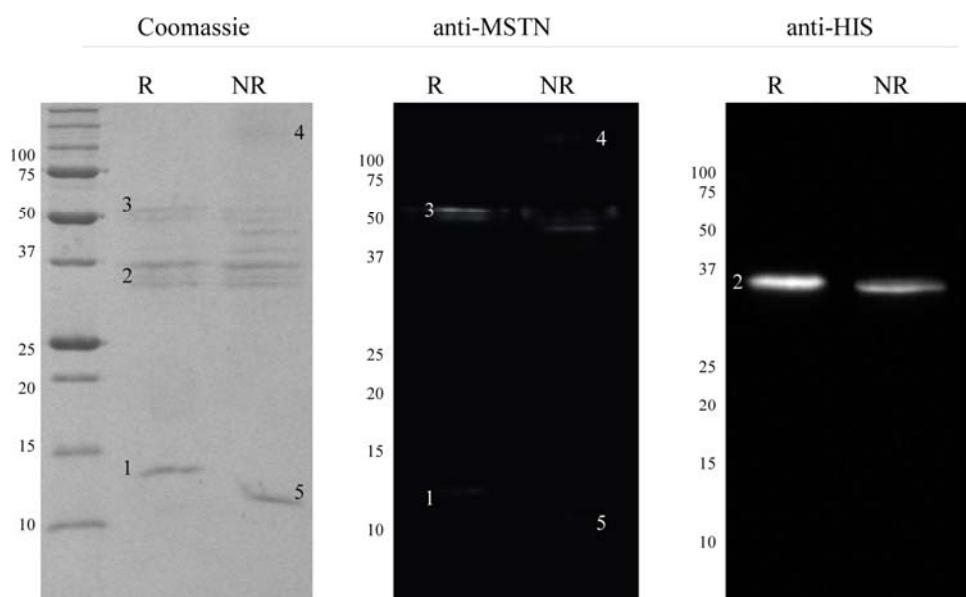


Figure 6.5 Furin digest of the C313Y precursor protein.

SDS-PAGE and Western blotting using antibodies against myostatin (anti-MSTN) and the His-tag (anti-HIS) in reducing (R) and non-reducing (NR) conditions. Bands are labeled as follows: 1, C313Y growth factor monomer; 2, C313Y propeptide; 3, undigested C313Y MstnPP monomer; 4, undigested C313Y MstnPP dimer; 5, Disulphide-bonded C313Y growth factor monomer.

Like WT MstnPP, refolded C313Y MstnPP can be digested *in vitro* by furin convertase to produce the growth factor and propeptide domains; results shown thus far indicate that the native C313Y precursor protein has been produced.

6.2.2 Activity assays

In order to establish inactivity of the C313Y growth factor, C2C12 proliferation assays were performed (Figure 6.6). As acid-treatment was the most effective method for activation of the WT growth factor (Fig. 3.9), this method was selected for C313Y. Assays were also conducted for the C313Y precursor protein and latent complex using concentrations equivalent to those used for the WT protein.

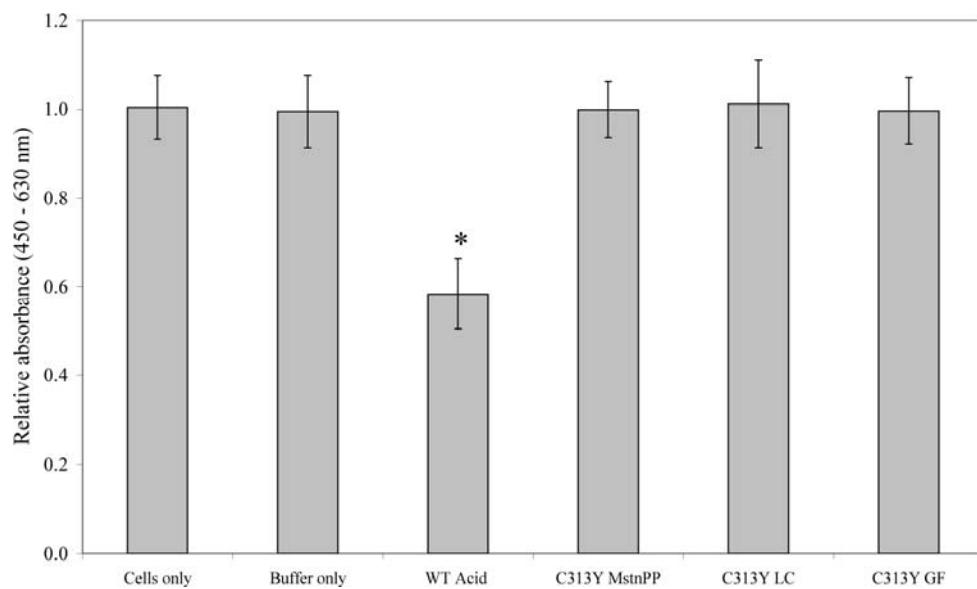


Figure 6.6 Activity of the mutant myostatin growth factor by C2C12 proliferation assays

Relative absorbance (450-630 nm) is proportional to cell growth using the WST-1 reagent (Roche). Bars are labeled as follows: C313Y GF, the C313Y growth factor; C313Y LC, untreated latent complex, the furin-digested C313Y precursor protein; C313Y MstnPP, undigested C313Y precursor protein; WT Acid, acid-treated furin digest of WT precursor protein (Fig. 3.8). Cells only and buffer only bars indicate negative controls. Concentration of WT and C313Y growth factors used was 10 µg/mL, within the latent complex; concentration of all other proteins was equivalent to that of the latent complex when the growth factor was 10 µg/mL. * P < 0.05 by Student's t-test.

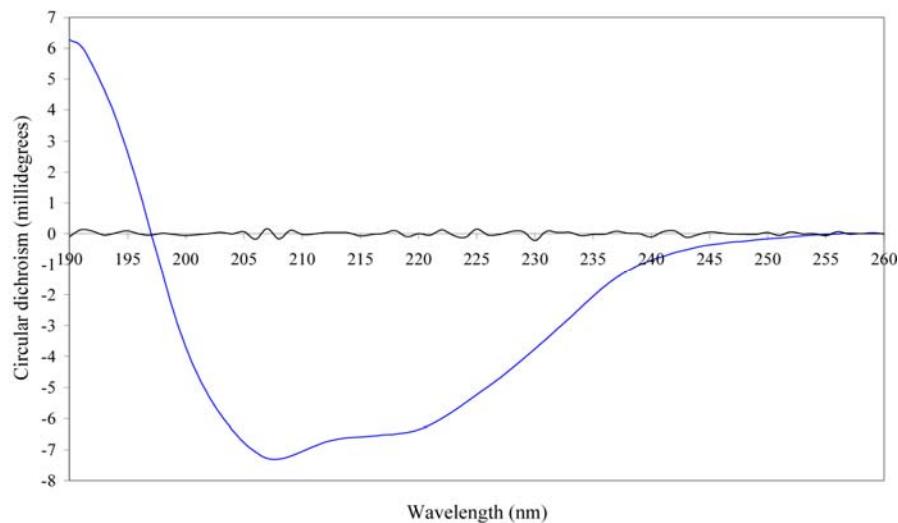
The acid-activated mutant growth factor (C313Y GF) had no effect on the proliferation of C2C12 myoblasts, indicating that the C313Y mutation had removed myostatin activity, consistent with previous results (Berry, Thomas *et al.* 2002). The mutant precursor and latent complex showed a similar result. The absence of activity is not definite evidence that production of native C313Y growth factor has been achieved as incorrectly folded protein is also expected to be inactive in this assay. However, biochemical and biophysical analyses complement these results.

6.3 Biochemical analysis of C313Y MstnPP and comparison to WT MstnPP

6.3.1 Secondary structure

The secondary structure of refolded C313Y MstnPP dimer was assessed (Figure 6.7a) and compared to that of the WT (Figure 6.7b) using circular dichroism spectroscopy.

a.



b.

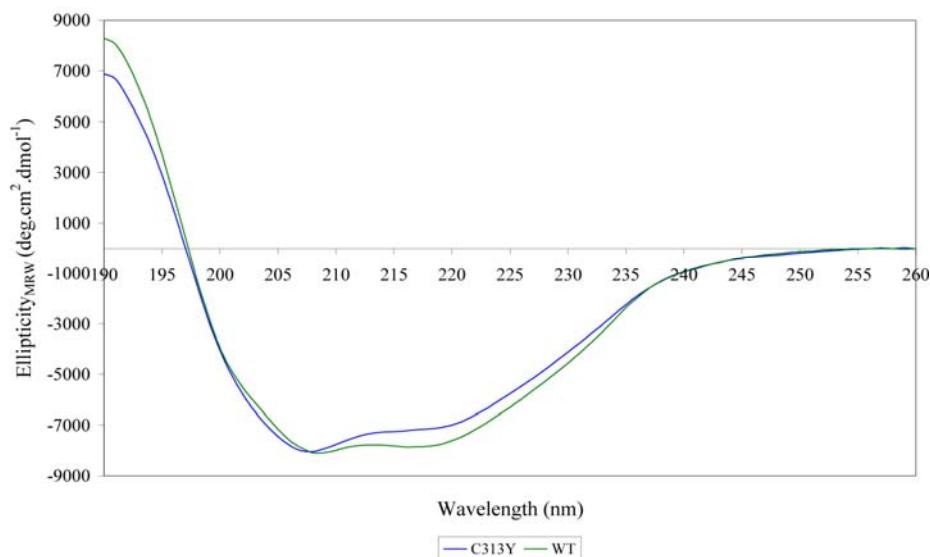


Figure 6.7 Circular dichroism of the C313Y precursor protein.

a. CD spectrum of C313Y.

b. Comparison between the CD spectra of C313Y (blue) and WT (green) precursor proteins.

As seen for WT MstnPP, the C313Y precursor protein spectrum contains contributions from both α -helix at 209 nm and β -sheet at 217 nm, as well as disorder with an absolute negative value at 200 nm. The spectra are largely similar except for a decrease in intensity of the C313Y β -sheet absorption intensity, suggesting that the C313Y mutation compromises the integrity of some β -sheet structures. The 190 nm maximum, usually correlated with α -helical absorption, is stronger in the WT protein; this may imply that the WT protein also has a higher proportion of α -helix. Deconvolution using the CDDN algorithm predicts a slight increase in α -helix, a decrease in β -sheet structures and comparable levels of random coil (Appendix 4, Table 12.2), consistent with the observations described.

6.3.2 Thermal stability

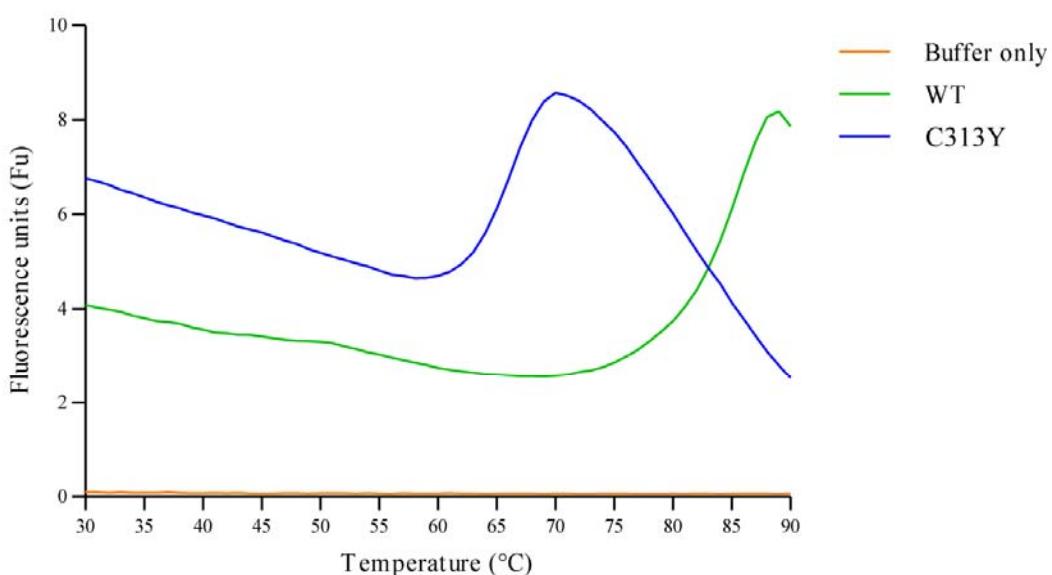
The thermal stability of C313Y and WT precursor proteins were compared using Sypro Orange fluorescence and CD thermal denaturation.

Thermal shift assay

Fluorescence measurements estimate that the C313Y precursor protein has a melting temperature of 66 °C (Fig. 6.8), 20°C lower than that of the WT, indicating that the mutant protein has decreased thermal stability compared to the WT. This result is proportionately consistent with the reduced thermal stability observed for an equivalent VEGF cysteine knot mutant (Muller and Hering 2002).

A possible explanation is that the C313Y mutation renders the protein chain more flexible and thus more prone to temperature-induced movement and denaturation.

a.



b.

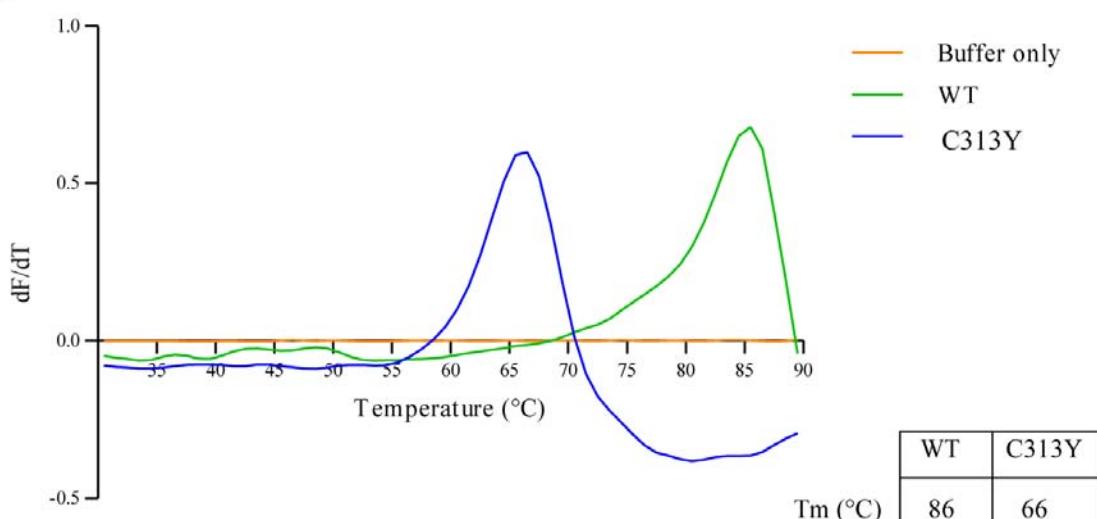


Figure 6.8 Thermal shift analysis of C313Y and WT precursor proteins by Sypro Orange binding and fluorescence.

- Raw data; Sypro Orange fluorescence as a function of temperature.
 - Differentiation of the curves in a. produces peaks which represent T_m values of 86 and 66 °C for the WT and C313Y proteins respectively.
- C313Y MstnPP, blue; WT MstnPP, green; Buffer-only control, orange.

CD thermal denaturation

CD thermal denaturation was used to investigate the behaviour in fluorescence thermal denaturation further (Figure 6.9). Like WT MstnPP, erratic absorption between 190 and 200 nm can be observed (Fig. 6.9a) and both α -helical and β -sheet minima decrease in intensity as temperature increases (Fig. 6.9b).

Plots of the normalised circular dichroism as a function of temperature for both 209 and 217 nm minima (Fig. 6.9c) exhibit a similar pattern to the WT protein; although some sigmoidal shape is apparent, the data are ambiguous and may contain more than one transition. In addition, the 217 nm minimum shows a noticeable reduction in intensity at a lower temperature than that of 209 nm.

The WT thermal denaturation profile showed a β -sheet transition around 85 °C, suggested to be due to β -aggregation or to represent the stable cysteine knot motif. This is not visible for C313Y. However, a sharp decrease in normalised circular dichroism can be observed between 65 and 70 °C (Fig. 6.9c); a similar, although more dramatic, decrease can be observed for the WT at 85 °C (Fig. 4.5c). As WT Sypro Orange analysis gave a transition temperature of 86 °C for WT MstnPP, corresponding to the observed CD transition, C313Y MstnPP may undergo a similar, yet less intense, transition at around 65 °C, close to the Tm value estimated by Sypro Orange (Fig. 6.8). If this is the case, and the transition reflects aggregation, these results suggest that the mutant precursor has increased propensity for β -aggregation. If this transition is indicative of cysteine knot stability, the hypothesis that the C313Y mutation compromises the stability of the cysteine knot is supported.

In either situation, clear differences between the thermal denaturation of WT and C313Y MstnPP suggest that the C313Y mutation decreases the thermal stability of MstnPP.

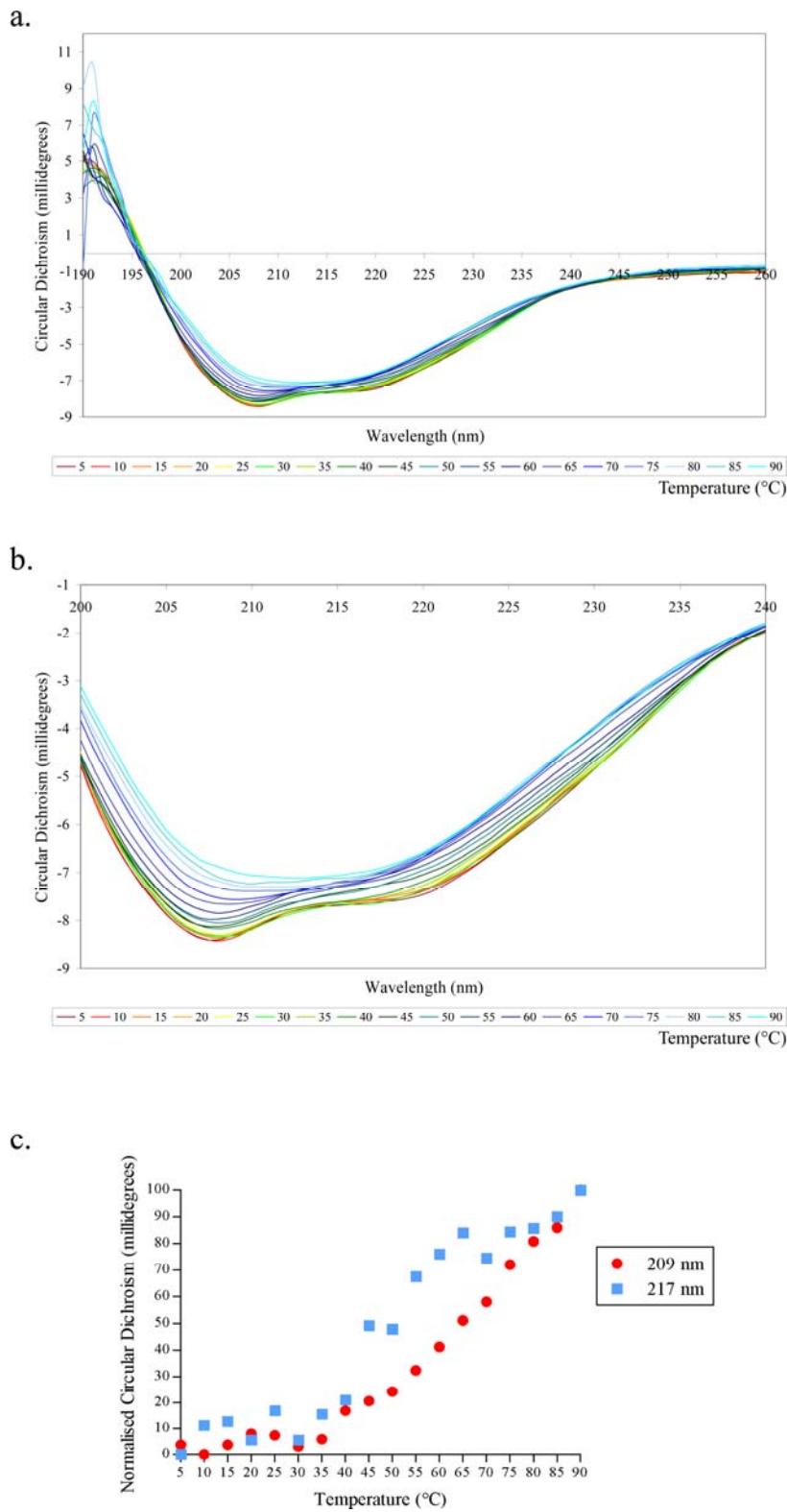


Figure 6.9 CD thermal denaturation of the C313Y precursor protein.

- CD profile from 190-260 nm from 5-90 °C.
- CD profile as in a. from 200-240 nm.
- Plot of normalised CD vs. temperature for the 209 (red) and 217 (blue) minima.

6.3.3 Protease resistance

The resistance of the C313Y precursor to limited proteolysis was compared to the WT at 37 °C using reducing (R) and non-reducing (NR) SDS-PAGE (Figure 6.10). Overall, the product pattern is very similar, indicating an absence of major structural differences between the WT and mutant proteins. However, the stability of C313Y appears to be decreased relative to WT. After one hour at 37 °C, full-length C313Y monomer (*) and dimer (**) have decreased significantly compared to WT.

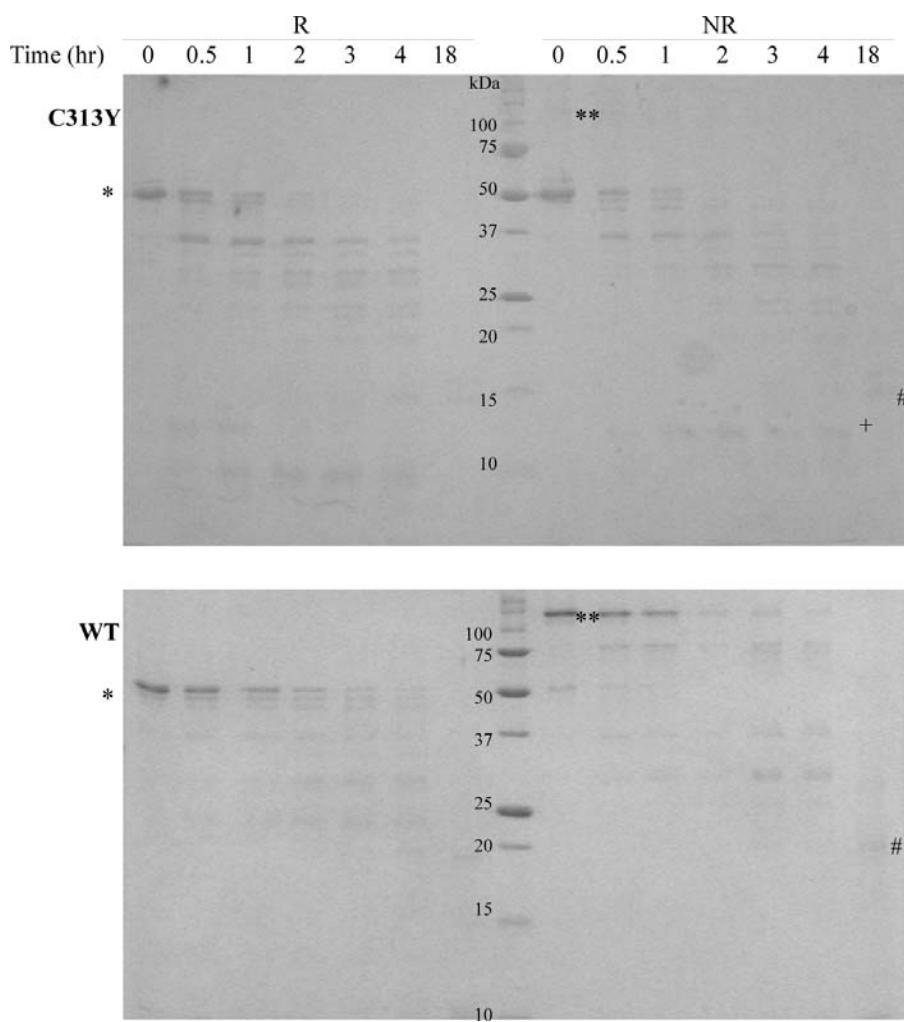


Figure 6.10 Limited proteolysis of C313Y and WT precursor proteins.

Samples were taken at 0, 0.5, 1, 2, 3, 4 and 18 hours and analysed using reducing (R) and non-reducing (NR) conditions. A 100:1 myostatin to trypsin (w/w) ratio was used at 37 °C.

The initial degradation product for C313Y MstnPP is 37 kDa. As suggested for WT MstnPP, this may represent the propeptide region indicating that cleavage around or near to the furin site has occurred. A 12.5 kDa (+) band is prominent from 0.5 hours in both reducing and non-reducing lanes and a band just below 10 kDa can be also observed in reducing conditions. The 10 and 12.5 kDa bands are not clearly visible in the WT SDS-PAGE gel but are present by Western analysis (Fig. 4.6).

After overnight digest, only a 15 kDa band remains in both R and NR C313Y lanes (#) compared to the 10, 15 and 18 kDa (R) and 10, 18 and 26 kDa (NR) bands of the WT (Fig. 4.6). Neither antibody in Western analysis of the WT digest detected the 15 kDa band; the C313Y 15 kDa band may represent a similar product that lacks both the His-tag and C-terminal antigens.

The primary conclusion from this analysis is that WT MstnPP has greater protease resistance and therefore structural stability than the C313Y mutant protein. The C313Y mutation, which disrupts the cysteine knot structure, may increase structural flexibility and/or reduce the integrity of stable structural elements.

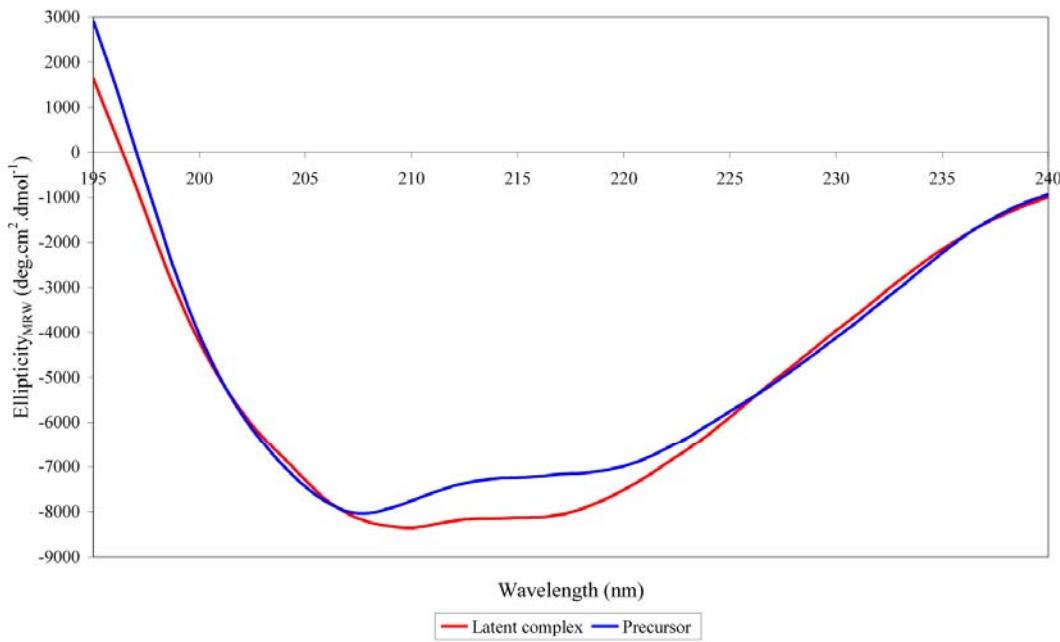
6.4 The C313Y latent complex

Analysis of the C313Y latent complex and comparison to the WT may provide additional information for an understanding of the biological consequences of the C313Y mutation.

6.4.1 Circular dichroism spectroscopy

The CD spectrum of the C313Y latent complex was compared to both the C313Y precursor protein (Fig. 6.11a) and the WT latent complex (Fig. 6.11b). In contrast to the highly similar WT precursor and latent complex spectra (Fig. 4.7), there is a noticeable difference in spectrum shape between the C313Y precursor and latent complex; the latent complex has increased intensity at both minima, especially for β -sheet at 217 nm. The concomitant reduction in intensity of the latent complex 190 nm peak supports this. CDDN deconvolution suggests increases in both α -helix and β -sheet regions (Appendix 4, Table 12.2) and a decrease in random coil, supporting an increase in structure for the latent complex.

a.



b.

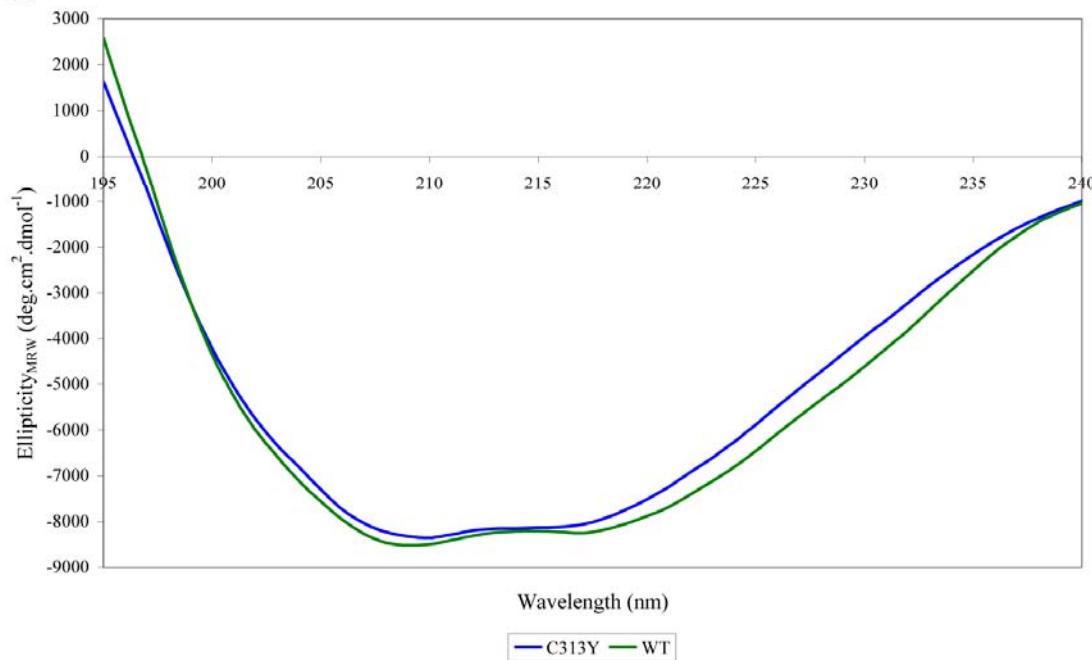


Figure 6.11 Circular dichroism of the C313Y latent complex.

- Comparison of the C313Y latent complex (red) and precursor protein (blue) spectra.
- Comparison of the C313Y (blue) and WT (green) latent complex spectra.

The WT and C313Y latent complexes have comparable CD spectra, suggesting formation of a similar complex (Fig. 6.11b). There are two differences worth noting however that may indicate subtle structural differences between the two complexes. First, the WT spectrum has slightly increased intensity at all wavelengths and second, the peak from 215-220 nm is narrowed in the mutant. While the former could be due to concentration differences, this is unlikely as the protein concentrations were consistent between samples. Spectrum deconvolution indicates a higher proportion of ordered structure in the WT protein (Appendix 4, Table 12.2). These results suggest that while structured regions may be increased in the C313Y latent complex compared to the C313Y precursor, the mutation has compromised some structural aspects relative to the WT protein.

6.4.2 CD thermal denaturation

Thermal denaturation of furin-digested C313Y MstnPP was performed (Fig. 6.12). As observed for all other CD thermal denaturation analyses, the region from 190-200 nm shows erratic behaviour (Fig. 6.12a). The WT latent complex shows a dramatic decrease for the 209 nm minimum at 80 °C (Fig. 4.8); this same phenomenon occurs at 70 °C for the C313Y latent complex (Fig. 6.12b). This result may reflect a number of possible scenarios. First, it may indicate that the α -helical component of the digested precursor is less thermally stable in the mutant compared to the WT. Second, this transition may also represent latent complex dissociation, suggesting that the entire C313Y latent complex is less stable than the WT. Third, if β -aggregation is responsible for the spectral shift, the C313Y mutation may increase the propensity of MstnPP for aggregation, a phenomenon that is most likely due to decreased thermal stability.

Neither dataset fits a two-state model of unfolding (Fig. 6.12c). However, denaturation occurs at similar temperatures for both 209 and 217 nm minima, as observed for the WT complex, supporting the hypothesis that complex formation is associated with structural alterations.

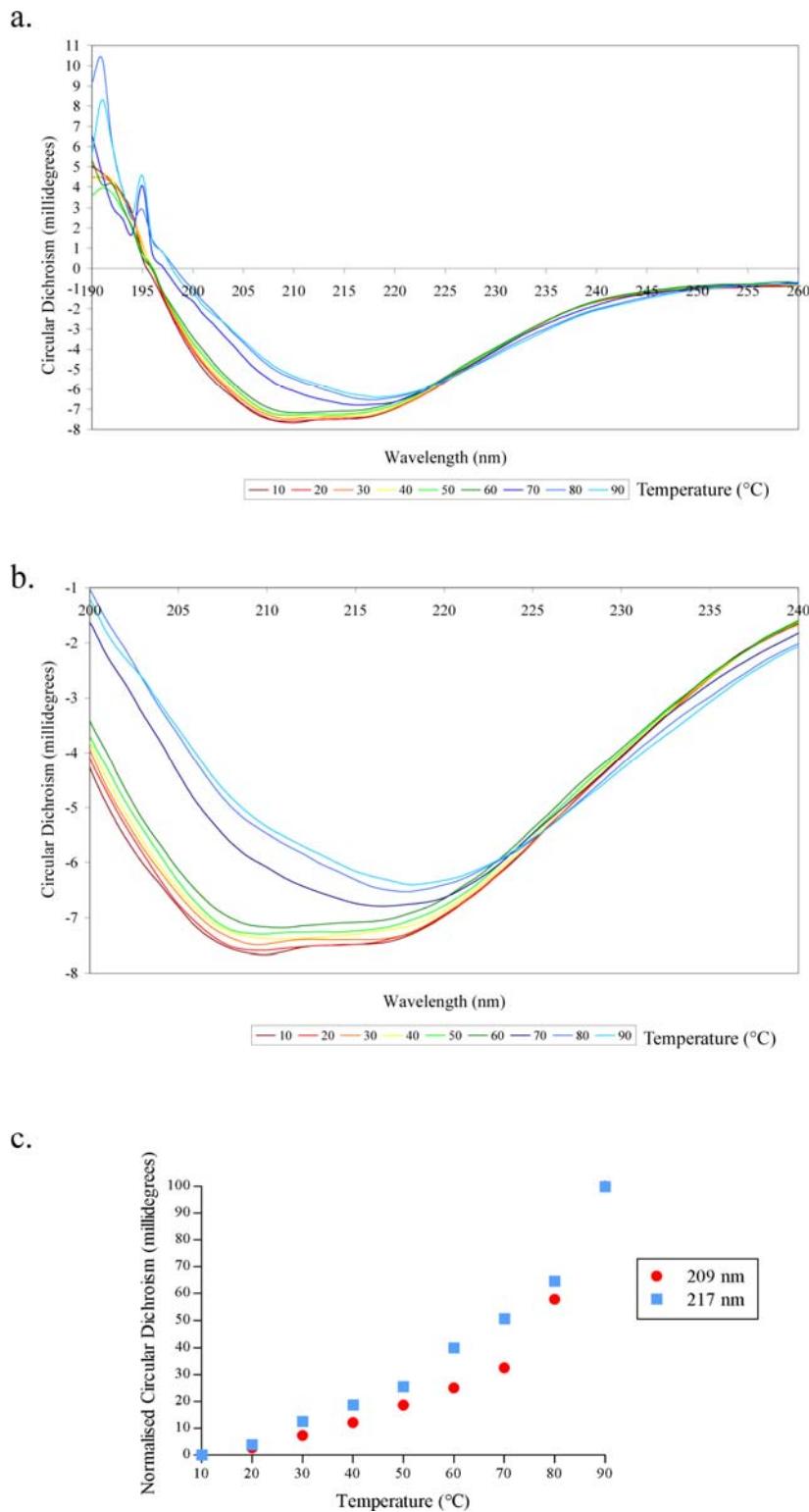


Figure 6.12 CD thermal denaturation of the C313Y latent complex.

- CD profile from 190-260 nm from 10-90 °C.
- CD profile as in a. from 200-240 nm.
- Plot of normalised CD vs. temperature for the 209 (red) and 217 (blue) nm minima.

6.4.3 Gel filtration analysis

Furin-digested C313Y was analysed using gel filtration chromatography after acid treatment (Fig. 6.15). There are two peaks in the chromatogram (Fig. 6.15a), the second of which is predominant. Retention times are 13.6 (1) and 15.3 (2) mL, corresponding to MWts of 105 and 52 kDa respectively. Reducing (R) and non-reducing (NR) SDS-PAGE analysis of the peaks (Fig. 6.15b) shows that while the first peak contains mainly undigested precursor protein monomer (i) and dimer (ii), there is also some propeptide present (iii). The growth factor is not visible. Peak two contains propeptide and growth factor monomer (iv). Consistent with previous results, only the growth factor monomer is present under non-reducing conditions (v) and an increase in mobility suggests the presence of intramolecular disulphide-bonds.

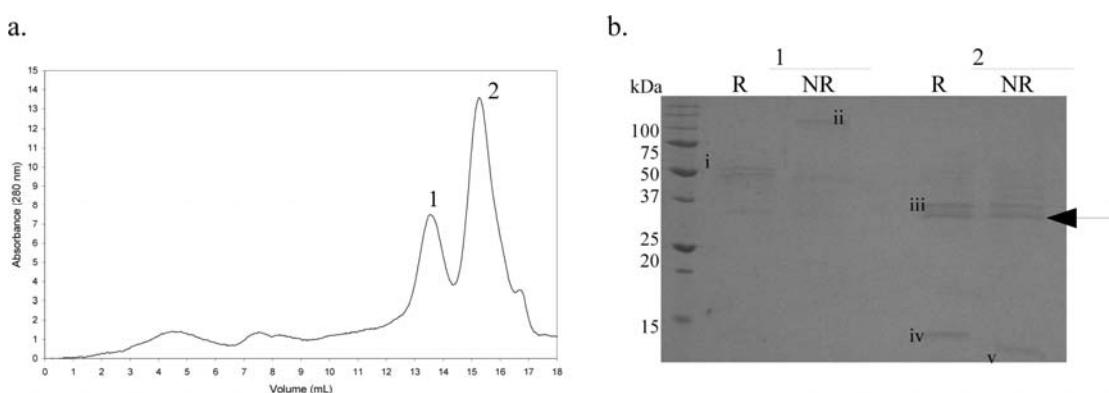


Figure 6.13 Gel filtration analysis of acid-treated C313Y latent complex.

- a. Gel filtration chromatography. Peak 1, 13.6 mL; Peak 2, 15.3 mL.
- b. Reducing (R) vs. non-reducing (NR) SDS-PAGE of peaks 1 and 2. Bands are as follows: i, undigested C313Y precursor monomer; ii, undigested C313Y precursor dimer; iii, C313Y propeptide; iv, reduced C313Y growth factor monomer; v, disulphide-bonded C313Y growth factor monomer.

These results are similar to those obtained for the WT latent complex. However, peak 2 of C313Y is symmetrical, in contrast to the multi-peak nature in the WT profile (Fig. 4.10). The 35 kDa propeptide degradation product can be observed (arrow). Analysis of the acid-treated WT latent complex estimated a MWt of 29 kDa for this band, corresponding to a peak eluting at 16.7 mL. As the small shoulder to the right of peak 2 has a retention time of 16.7 mL, this is likely to represent acid-degraded C313Y propeptide. The ratio of the 37 to 35 kDa bands is approximately 1:1, as observed for the WT. Dissociation was therefore incomplete.

6.5 Structural modelling of the C313Y mutation

The C313Y mutation was modelled in the published myostatin crystal structure (Fig. 7.2) with the use of Coot (Emsley and Cowtan 2004) and Pymol (<http://pymol.sourceforge.net/>). Disulphide bonds are shown in yellow and the C313Y mutation is shown in red, in one monomer only. C313Y is located within the concave region of the myostatin dimer structure (Fig. 1.6) which is likely to contain the Type I receptor-binding site (Fig. 5.4a) and is in close proximity to the intermolecular disulphide (Fig. 7.2, arrow).

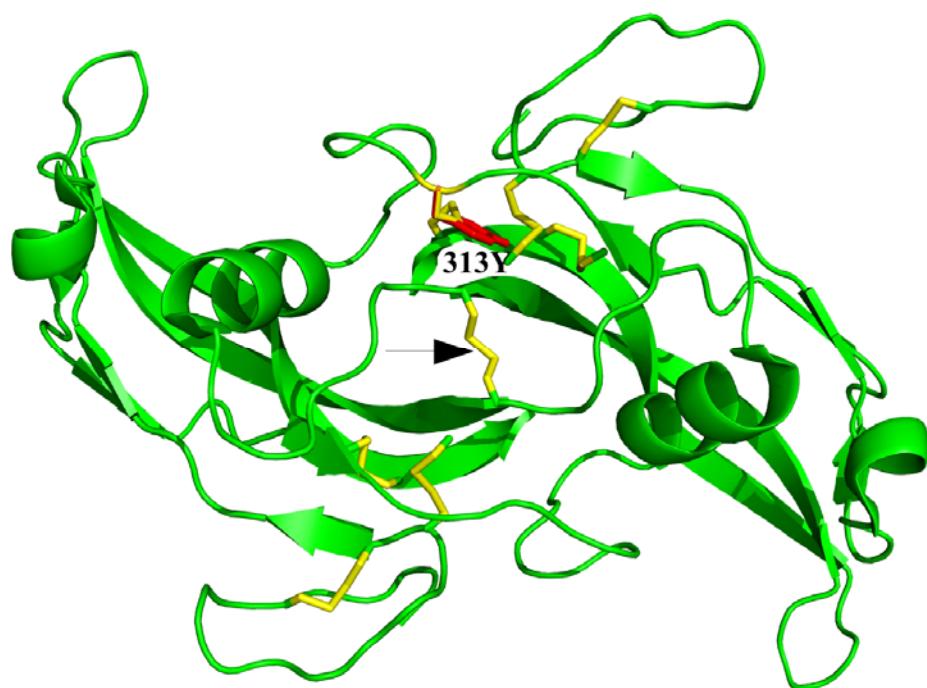


Figure 6.14 Modelling of the C313Y mutation in the myostatin growth factor crystal structure.
Native disulphides are shown in yellow; the modeled C313Y mutation is shown in red for one monomer only. The dimerisation disulphide is indicated by an arrow.

The three most likely rotamers (Fig. 6.15 and Table 6.1) were visualized in Coot. The native cysteine residue is shown in yellow and the substituted tyrosine in red, with intramolecular disulphide bonds in brown and the intermolecular disulphide in blue.

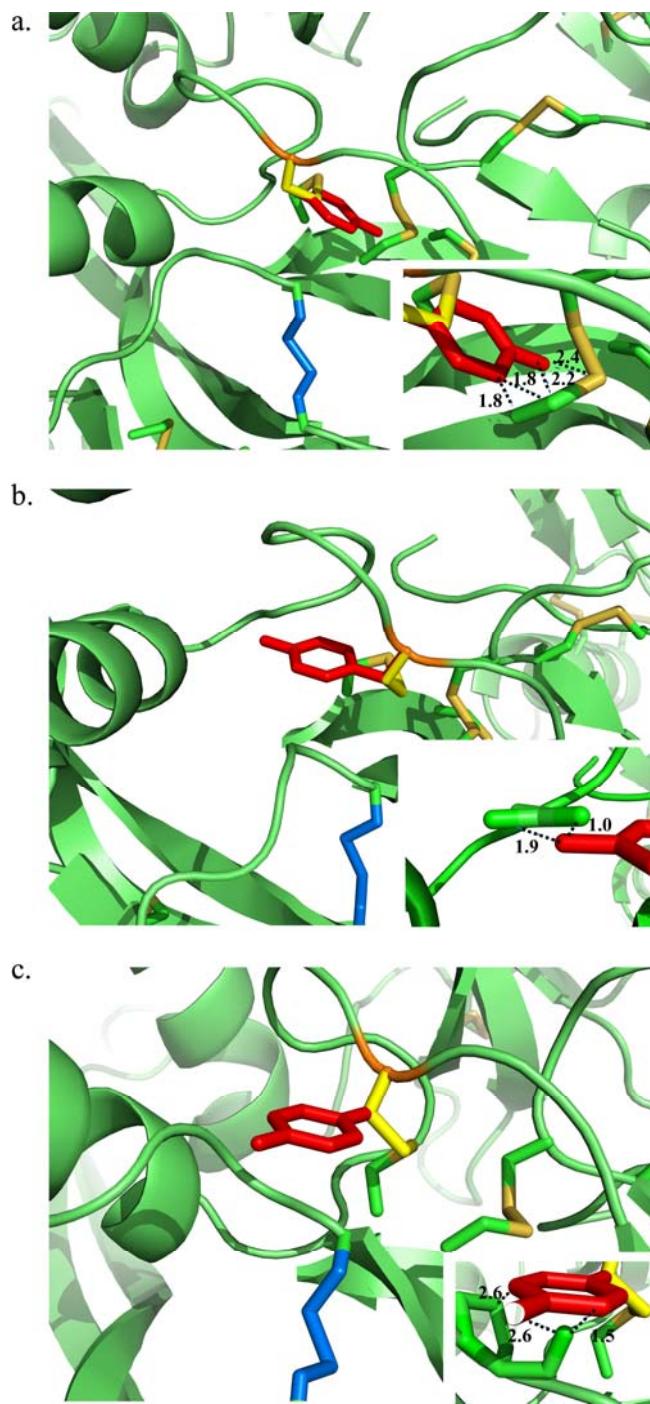


Figure 6.15 Rotamers available to C313Y.

As summarised in Table 6.1. Insets for each are a zoomed-in view of probable steric clashes.

a. Rotamer 1; Chi 1, -65°.

b. Rotamer 2; Chi 1, 176°.

c. Rotamer 3; Chi 1, 63°.

The native C313 residue is shown in yellow with the mutant tyrosine overlaid in red. The intermolecular dimerisation disulphide is shown in blue. Bonds that represent steric clashes are shown by black dotted lines with bond distances in Angstroms (Å).

All structures shown are unlikely due to steric clashes with other contacts (Fig 7.3, insets); distances are all 2.6 Å or less. While rotamer 1 (Fig. 6.15a) is in close contact with the C309-C372 disulphide bond, rotamer 2 (Fig. 7.3b) contacts Q319 and rotamer 3 (Fig. 7.3c), P338. The large tyrosine residue is therefore likely to cause structural changes for its accommodation. Which rotamer is most likely? Rotamer 1 would require alterations to the location of the disulphide bond so is the least likely scenario. Although it is not possible to make any conclusions, rotamer 2 appears to be the most viable option. Rotamer 3 makes contact with a proline which is likely to be more fixed in its conformation than the glutamine contacting rotamer 2. Rotamer 2 is also more commonly found in protein structures of the Coot database (Table 6.1).

Rotamer	Coot reference	Chi 1 value (°)	Frequency in conformation (%)
1 (Fig. 6.15a)	m-85	-65	43
2 (Fig. 6.15b)	t80	176	34
3 (Fig. 6.15c)	p90	63	13

Table 6.1 Modelling the three most common rotamers available to C313Y .

7 Discussion: Analysis of C313Y, the Piedmontese myostatin mutation

Analysis of the Piedmontese mutation in the human myostatin precursor, comparison to the WT equivalent and structural modelling suggest that the C313Y mutation has effects on the structure and stability of both the precursor and the latent complex. These results have implications for the structure of the Piedmontese growth factor and the function of myostatin *in vivo*.

7.1 The human myostatin precursor as a model system

The C313Y mutation is found in the Piedmontese cattle breed but for this thesis has been studied in the context of the human precursor protein. Is this an ideal model for investigation?

The growth factor regions of human and bovine myostatin share 98% sequence identity and contain the cysteine residue affected by the C313Y mutation (Fig. 7.1, red square). All cysteine residues involved in disulphide bonds are conserved (black stars), suggesting that the core cysteine knot structure of the human and bovine growth factors is the same. Analysis of the C313Y mutation within the human growth factor should therefore provide a fair representation of the Piedmontese situation.

The human and bovine propeptide regions share 93% identity at the amino acid level, not including the signal sequence, and most differences are conservative. As little is known about propeptide/growth factor interactions, the similarity between human and bovine proteins in this regard cannot be determined. However, from the results and model presented in Chapters 4 and 5, some structural comparisons can be made.

The least well conserved region is the N-terminal inhibitory domain; the C-terminus is almost identical, consistent with the proposed role of this region in myostatin stability. Phyre algorithms predict the bovine propeptide to contain most of the propeptide domain secondary structures, with almost identical C-terminal predictions. Importantly, the three hydrophobic residues suggested to be important for interaction of the propeptide with the growth factor (Walton, Makanji *et al.* 2009) are conserved (open stars). The bovine sequence is also highly similar to the human in regions of predicted disorder (black lines) indicating any functional disorder will be conserved.

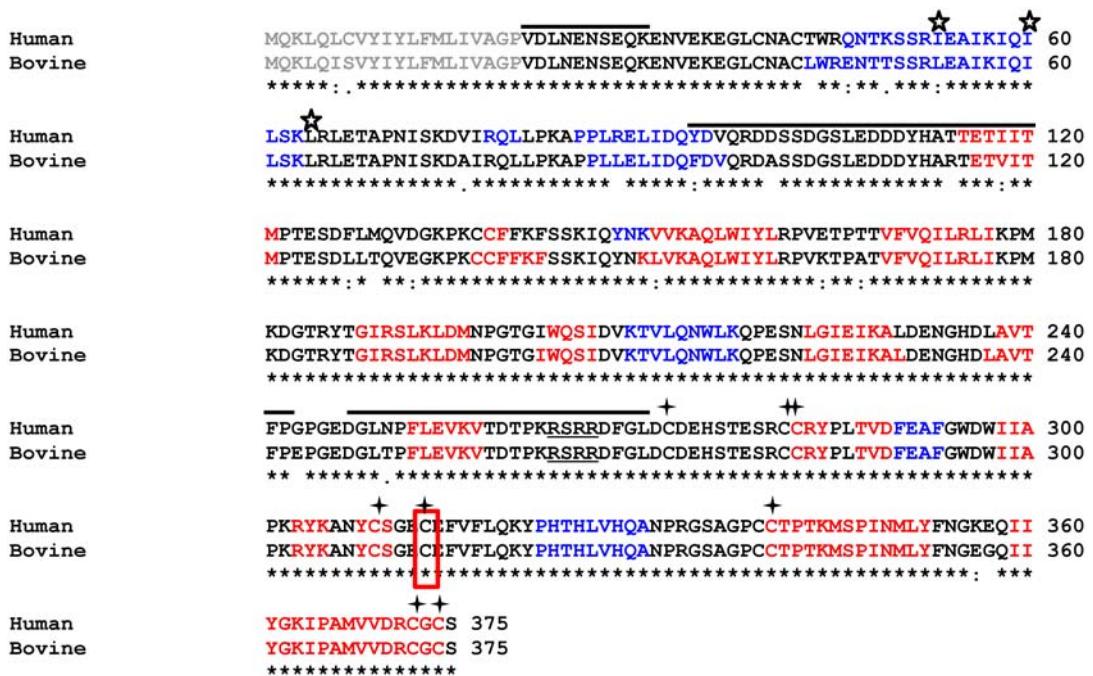


Figure 7.1 Sequence alignment of human and bovine myostatin precursor proteins.

Sequence alignment was performed using ClustalW (Thompson, Higgins *et al.* 1994). Key is as follows: Blue, predicted (Phyre) or known (crystal structure) helix; red, predicted or known β -sheet; black lines, predicted intrinsic disorder; open five-pointed stars, residues of the propeptide proposed to be involved in hydrophobic interactions with the growth factor; black 4-pointed stars, conserved cysteine residues involved in disulphide-bonds; underlined, the furin cleavage site; grey, signal peptide. ClustalW similarity key is as follows: *, exact match; :, strong similarity; ., weak similarity.

In conclusion, any differences between the bovine and human myostatin sequences at the amino acid level are subtle and the human protein is a sufficient model system to study the effects of the C313Y mutation on protein structure and function.

7.2 Biochemical consequences of C313Y

7.2.1 Dimerisation

The myostatin growth factor dimerises via an intermolecular disulphide bond and the majority of TGF- β family members require covalent dimerisation for activity (Brunner, Lioubin *et al.* 1992; Vitt and Hsueh 2001) and stabilisation (Daopin, Piez *et al.* 1992). A number of lines of evidence suggest that the potential for covalent dimerisation is hindered in the C313Y mutant; an inability to dimerise covalently may be one reason behind the inactivity of the C313Y growth factor *in vivo*.

First, after refolding of the mutant protein, there is proportionately less dimer by non-reducing SDS-PAGE than observed for the WT, with the majority of refolded protein present in a monomeric form (Fig. 6.3).

Second, structural modelling shows that the C313Y mutation site is close to the dimerisation disulphide bond (Fig. 6.14). In order to accommodate the large tyrosine residue, the structure in this region may be distorted, causing an increase in the distance between the two cysteines involved in the intermolecular disulphide is increased and perturbation of covalent interaction.

Third, the retention time of the refolded C313Y protein is indicative of a 105 kDa dimer by gel filtration, despite the predominance of the C313Y monomer by non-reducing SDS-PAGE (Fig. 6.4), suggesting dimerisation via non-covalent interactions at the hydrophobic dimerisation interface (Cash, Rejon *et al.* 2009).

Last, the C313Y growth factor dimer is absent following furin digestion (Fig. 6.5) by both SDS-PAGE and the more sensitive technique of Western blotting. The possible inaccessibility of the epitope recognized by anti-MSTN and low protein concentration used in this analysis may be responsible for this absence. However, the disulphide-bonded monomer is clearly visible.

After both refolding and gel filtration, a small amount of dimer can be observed by SDS-PAGE. Flexibility in the structure may allow the formation of the dimerisation disulphide, with formation and dissociation in equilibrium. An alternative explanation is that although the C313Y mutation interferes with disulphide formation in the majority of molecules, the tyrosine residue resides in different rotamers between molecules (Fig. 6.15), allowing covalent dimerisation in some. The absence of the growth factor dimer post-furin digest (Fig. 6.5) may indicate that the propeptide region facilitates the observed dimerisation and removal of this region completely inhibits disulphide bond formation.

If an inability to dimerise covalently is responsible for an absence of signalling and some covalent dimerisation still occurs *in vivo*, a low level of signalling may be possible, with additional mechanisms responsible for the double-muscled phenotype. It has been suggested that the C313Y mutant acts as a competitive inhibitor *in vivo* (Berry, Thomas *et al.* 2002); this inhibition may overcome a low level of signalling resulting from functional dimerisation. Competitive inhibition *in vitro* was assessed by the addition of the Piedmontese growth factor to WT C2C12 proliferation assays

(Berry, Thomas *et al.* 2002). As Piedmontese cattle are homozygous for the C313Y mutation (Kambadur, Sharma *et al.* 1997), WT inhibition will not be observed *in vivo*.

An absence of the dimerisation disulphide may result in a more open structure which prevents receptor binding. In contrast to this hypothesis, the estimated molecular weight of the C313Y dimer by gel filtration is 10 kDa less than the WT precursor protein, suggesting a more compact conformation than the WT, which may also inhibit receptor-binding. An increase in retention time may also reflect the lower concentration of C313Y protein used. However, this result is consistent with structural analysis of the equivalent mutation in the cysteine-knot protein VEGF (vascular endothelial growth factor), that showed a reduction in diameter of 2 Å by comparison to the WT (Fig. 7.2) (Muller and Heiring 2002) and tighter packing in the dimerisation interface.

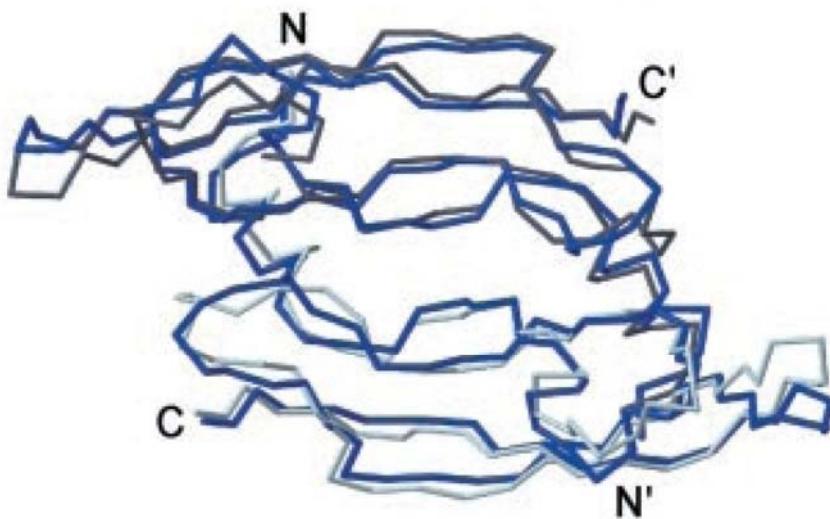


Figure 7.2 Molecular structure of VEGF mutant equivalent to C313Y.

WT VEGF is in grey, mutant VEGF in blue. Figure taken from (Muller and Heiring 2002).

7.2.2 Stability and flexibility

Structural differences

The C313Y mutation results in disruption of the C313-C374 disulphide (Fig. 6.14). This is one of the disulphide bonds of the cysteine knot core structure, suggested to play a major role in the high thermal and structural stability of TGF- β family proteins (Muller and Heiring 2002), indicating that these properties may be compromised in the Piedmontese mutant. Residue 313 lies in a loop region of the growth factor; the native C313-C374 disulphide bond may stabilise this loop so that bond disruption results in increased loop movement. This may contribute to an increase in structural flexibility.

The Piedmontese mutation replaces a cysteine residue with a large tyrosine; structural modelling suggests that perturbations will be necessary for the accommodation of tyrosine at position 313 (Fig. 6.15) and these perturbations may have an effect on dimerisation. Structural alterations are also likely to impact the greater surrounding area, involving both side- and main-chain rearrangement. In addition, global changes are possible, affecting the overall shape of the myostatin growth factor.

Circular dichroism indicates a reduction in β -sheet for the Piedmontese mutant compared to the WT protein (Fig. 6.7). As the local structure of residue 313 is β -sheet predominant (Fig. 6.14), these β -sheet secondary structures may be partially stabilized by the C313-C374 disulphide bond and disruption may interfere with β -sheet formation and/or structural stability. The latent complex of Piedmontese myostatin shows moderate differences in secondary structure compared to the precursor protein, with enhanced intensity for both α -helical and β -sheet minima, especially the latter (Fig. 6.12a). As these differences were not observed in a comparison of the WT precursor and latent complexes (Fig. 4.7), one possibility is that structural differences in the C313Y growth factor influence the structure of the latent complex. An increase in secondary structure for the C313Y latent complex suggests that interactions formed in the latent complex stabilise the increased flexibility seen in the mutant precursor protein. Consistent with this, the overall shape of the WT and C313Y latent complexes is similar (Fig. 6.12b); however, subtle differences suggest the C313Y mutation effects structural alterations in the latent complex also.

The construction of disulphide-bond cysteine mutants in the cysteine-knot protein VEGF (Muller and Heiring 2002) produced results that support those shown for the C313Y myostatin mutant and suggest additional conformational changes may exist. Although VEGF and myostatin share little sequence homology, the structure of the cysteine knot is well-conserved and the VEGF mutant ΔIII (C61A) is equivalent to C313Y; like C313, C61 is located within a loop region of VEGF on the concave surface (Fig. 7.3a, blue). The crystal structure of ΔIII shows dramatic structural changes. First, the global structure showed displacement of the outer ‘finger’ of VEGF (Fig. 7.3a, arrow), the site of Type II receptor binding, by about 3.2 Å parallel to the two-fold axis, a direct consequence of rearrangements in side-chain orientations in the dimer interface. These changes are quite distal to the site of the mutation and have implications for type II receptor binding by the VEGF ΔIII mutant; similar global alterations may exist for the Piedmontese myostatin structure.

The local structure showed a conformational change in the loop segment containing residue 61 (Fig. 7.3b) and a novel disulphide bond had been introduced in the loop; this is not possible for myostatin due to the absence of equivalent cysteine residues. As mentioned previously, tighter packing in the dimerisation interface was observed, accompanied by a number of new interactions, such as salt-bridges and a reduction of 1.7 Å in the distance between β-strands lining the interface.

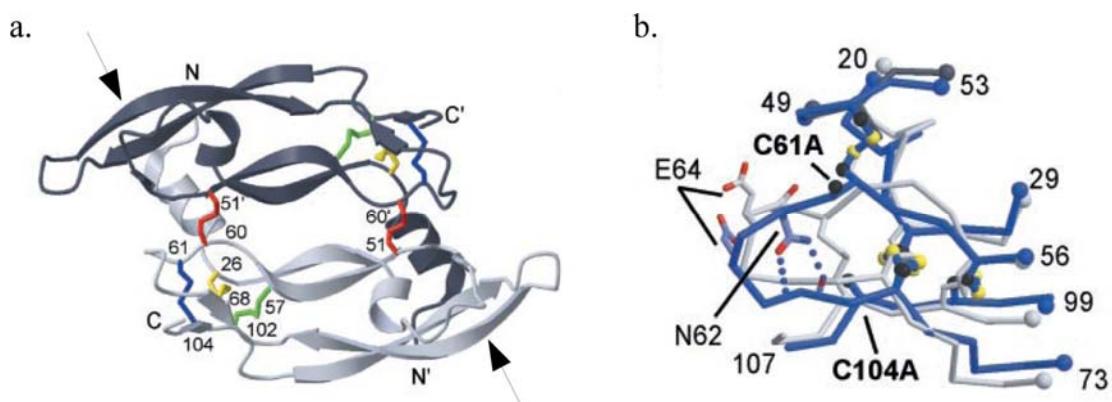


Figure 7.3 Location of the VEGF ΔIII mutant and associated structural changes.

a. WT VEGF structure showing disulphide bond locations. The ΔIII mutant contains C to A substitutions in the C61-C104 bond (blue). Arrows indicate the outer β-sheet fingers of VEGF and the type II receptor binding site.

b. Local structural changes observed for VEGF ΔIII. WT, grey; ΔIII, blue.

Figures taken from (Muller and Heiring 2002).

The VEGF Δ III mutant contained C to A substitutions at both residues involved in the C61 disulphide bond (C61A-C104A). Therefore, the large structural changes observed may also be due to the C104A mutation. It is worth noting then, that cysteine 374 is left in an unpaired state by the C313Y mutation. This cysteine residue is located in the C-terminal β -sheet of the growth factor, one amino acid away from residue C372, which is involved in a disulphide bond with C309. A second VEGF mutant, Δ II, is without the equivalent disulphide bond to C309-C372, and shows moderate structural changes only on comparison to Δ III. Although the C309-C372 disulphide of C313Y myostatin is likely to have a stabilizing effect in the region of C374, it is also possible that increased flexibility of the C-terminus may hinder formation of the C309-C372 disulphide bond. If this is the case, the observed structural and biochemical changes may also reflect the absence of this other disulphide.

Thermal stability

The C313Y precursor protein and latent complex exhibit altered thermal stability (Figs. 6.8–6.10, Table 6.1) by comparison to the WT. Evidence suggests that a reduction is most likely, consistent with the reduced thermal stability of VEGF cysteine knot mutants (Muller and Heiring 2002). The results shown here can be explained by disruption of the core cysteine knot motif, an increase in structural flexibility and/or the absence of a stabilising intermolecular disulphide bond (Daopin, Piez *et al.* 1992).

CD thermal denaturation for all species indicates a β -sheet transition at high temperatures, suggested to represent either retention of the stable cysteine knot, amyloid-like β -aggregation or, when the furin-digested precursor is studied, dissociation of the latent complex. Both Sypro Orange fluorescence analysis and CD thermal denaturation indicate that this transition, independent of its nature, occurs at a lower temperature for the mutant in comparison to the WT, with Sypro Orange analysis calculating a difference of 20 °C (Fig. 6.8). If the structural transition represents the stable cysteine knot, this may indicate that the absence of the C313-C374 disulphide bond compromises the integrity of the cysteine knot. If the β -sheet change is reminiscent of β -aggregation, the hypothesis that the Piedmontese mutation alters the structure in a way that increases the propensity for aggregation is raised, likely through decreased thermal stability and an increase in flexibility.

Proteolysis

The C313Y precursor protein exhibits greater sensitivity to trypsin proteolysis by comparison to the WT precursor protein (Fig. 6.10). Degradation is visible earlier for the mutant precursor, implying that some regions may be more flexible or disordered in this species. This supports the suggestion that the loop-location of residue 313, coupled with disruption of the C313-C374 disulphide bond, increases flexibility in this region.

Analysis of possible trypsin cleavage sites (Lys and Arg residues) on one monomer of the myostatin growth factor shows a lysine located on the loop containing residue 313 (Fig. 7.4). Therefore, increased flexibility of this loop may increase cleavage at this site. Although it is also possible that the C313Y mutation causes structural alterations that place other trypsin-cleavage sites in a more accessible position, such as the global changes observed for the VEGF Δ III mutant (Muller and Heiring 2002), information about this is not available from the results obtained here.

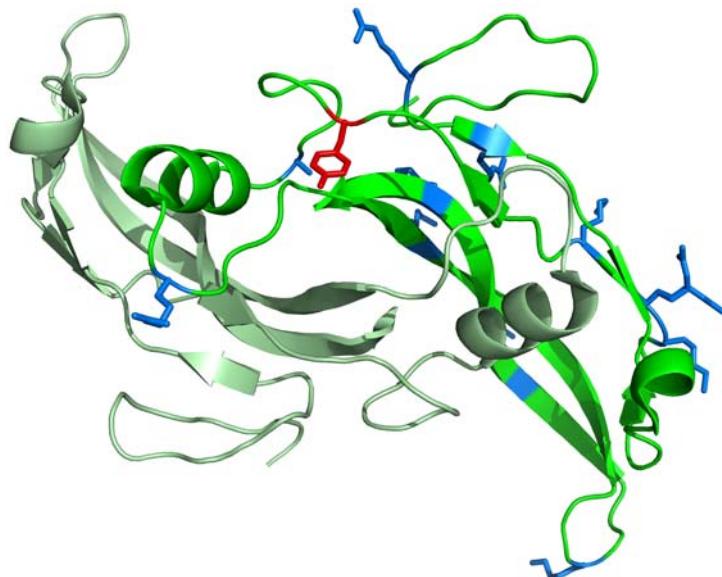


Figure 7.4 Trypsin cleavage sites of the C313Y growth factor.

Arginine and lysine residues are shown in blue; the C313Y mutation is in red.

7.2.3 Processing, secretion and biological activity

Analysis of the expression and processing of myostatin from Piedmontese skeletal muscle showed that the mutation does not lead to unstable mRNA but may influence

processing or stability, with less mature myostatin protein relative to WT, despite high levels of precursor protein (Berry, Thomas *et al.* 2002). In contrast, no change in serum circulatory levels of Piedmontese myostatin could be observed by comparison to WT, suggesting that secretion is not affected.

In vitro refolding of the Piedmontese C313Y myostatin precursor was associated with increased aggregation and a subsequent reduction in refolding efficiency (Fig. 6.3 and 6.4). If folding of the Piedmontese precursor is similar *in vivo*, with an increase in aggregation, this is likely to be accompanied by ubiquitination and proteasomal degradation (Schroder 2008). Previous research showed high levels of Piedmontese precursor with a number of differently-sized bands detected by Western analysis (Berry, Thomas *et al.* 2002), which may represent ubiquitinated protein; the myostatin antibody is unlikely to distinguish between differently folded forms of the precursor protein.

Processing of the myostatin precursor is not inhibited by the C313Y mutation. The mutant precursor dimer can be cleaved *in vitro* by furin convertase with efficiency comparable to the WT (Fig.6.5). This indicates that a mutant latent complex is formed for export extracellular circulation, and that reduced processing of the correctly folded C313Y precursor is not involved in the double-muscling phenotype. These results are in contrast to *in vivo* observations (Berry, Thomas *et al.* 2002); however, this difference may be due to the presence of aggregated C313Y precursor protein, as structural differences may inhibit furin-cleavage. Furin digestion of aggregated C313Y MstnPP was not attempted.

CD thermal denaturation suggests that dissociation of the latent complex may occur more readily in the mutant. After modelling of growth factor/propeptide interactions for the WT latent complex, it was proposed that interactions may form between the C-terminus of the propeptide and the concave surface of the growth factor (Fig. 5.6). Increased flexibility in this region of the C313Y growth factor may weaken the proposed interactions, contributing to a decrease in latent complex stability. If this is the case, the mutant growth factor will be activated to a greater extent than the WT *in vivo*. This is not expected to translate into an increase in myostatin signalling due to the presence of the C313Y mutation; however, it may help to explain the unaltered levels of circulating growth factor Piedmontese cattle despite decreases in processing and stability (Berry, Thomas *et al.* 2002).

In contrast, gel filtration analysis after acid treatment of the C313Y latent complex indicated that dissociation of the mutant complex may be reduced by comparison to the WT. This may indicate that different mechanisms are involved for heat and acid-induced dissociation.

C2C12 proliferation assays performed using acid-treated furin-cleaved C313Y latent complex show that the mutant growth factor is inactive (Fig. 6.6) as expected (Berry, Thomas *et al.* 2002). Similar results have been observed for the equivalent TGF- β 1 disulphide mutant (Brunner, Lioubin *et al.* 1992). It has been suggested that the Piedmontese growth factor exhibits dominant negative inhibition of WT myostatin, as proliferation assays conducted using increased molar ratios of mutant protein over the wild-type protein rescued myoblasts from growth inhibition (Berry, Thomas *et al.* 2002). Furthermore, the mutant growth factor alone caused increases in cell growth over and above control cells. An increase in cell growth was not observed in the assays done here, even at a concentration of 10 μ g/mL, which effected a 20% increase in cell growth in the previous reports (Berry, Thomas *et al.* 2002). This may be due to the incomplete dissociation observed for the latent complex (Fig. 6.4).

7.3 Implications for myostatin function *in vivo*

A popular hypothesis is that the Piedmontese mutation inhibits binding of the growth factor to its receptors in some way as *in vivo*, levels of the circulating growth factor are unaltered in comparison to the WT protein (Berry, Thomas *et al.* 2002).

The Piedmontese protein exhibits reduced dimerisation, and a decrease in thermal and proteolytic stability. These characteristics are most likely due to disruption of the cysteine knot core motif, an increase in flexibility of the loop containing residue 313, and/or the presence of tyrosine at position 313 forcing structural perturbations on the concave surface. The C313Y mutation is localized on the concave surface, close to the type I receptor-binding site (Fig. 7.5). Therefore, the mechanisms behind an inhibition of receptor-binding may include structural disruptions and increased flexibility in the type I receptor-binding site, a hypothesis supported by the crystal structure of the equivalent VEGF mutant (Muller and Heiring 2002).

In addition, altered thermal stability for the Piedmontese latent complex may indicate that the C313Y mutation lowers the affinity of interactions between the propeptide and growth factor domains, some of which have been proposed to occur at the type I receptor binding site.

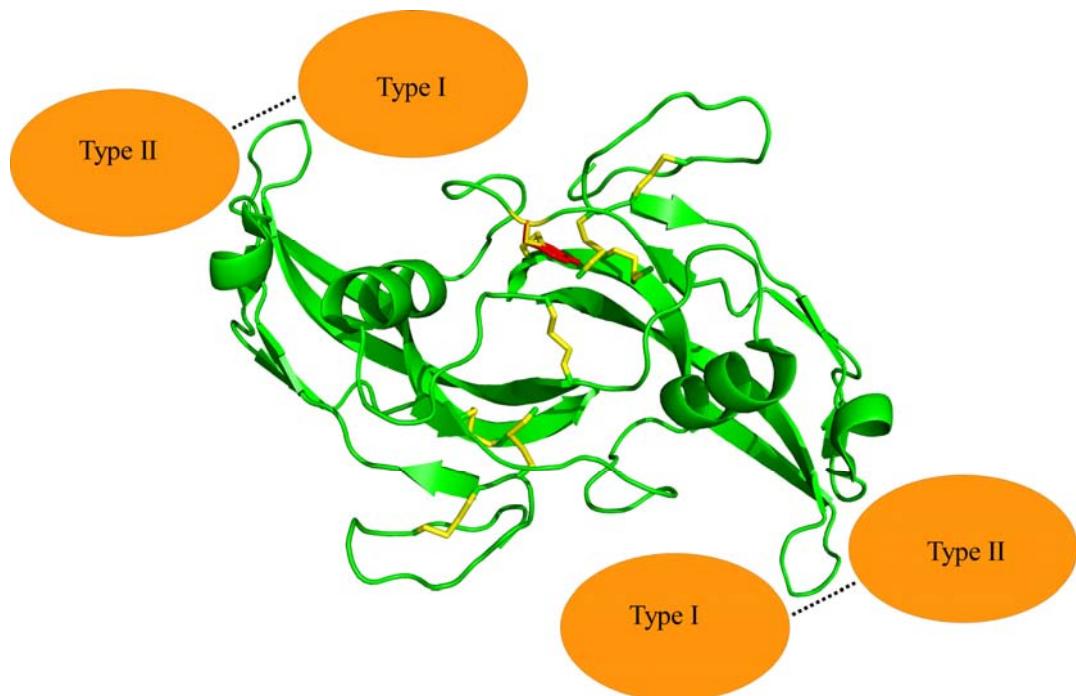


Figure 7.5 Putative type I and type II receptor interactions for the myostatin growth factor.

The TGF- β receptor-binding sites modeled for the myostatin growth factor. Receptors shown in orange, with dotted lines indicating interaction between receptors. C313Y mutation is in red and disulphide bonds are in yellow.

Modelling of the WT latent complex included interactions between the N-terminus of the propeptide and the finger regions of the growth factor. These interactions are quite distal from the site of the C313Y mutation; however, the equivalent VEGF mutant showed significant structural differences in this region, which contains the type II receptor binding site (Muller and Hering 2002). If the C313Y mutation causes similar conformational changes in Piedmontese myostatin, type II receptor binding may also be affected.

Increased proteolysis in the presence of trypsin may indicate that the Piedmontese growth factor is more susceptible to degradation *in vivo*. However, this was not apparent in Western blots of circulating Piedmontese and WT growth factors (Berry, Thomas *et al.* 2002). For this thesis, trypsin digests were performed at 37°C, suggesting that increased proteolysis may be due to an increase in mutant flexibility. As basal body temperature is around 37°C, increased flexibility may also occur *in vivo*.

Summary

The results presented in Chapters 6 and 7 suggest that differences in structure and stability may underlie the inability of the C313Y Piedmontese myostatin growth factor dimer to signal and elicit an inhibition of muscle growth *in vivo*. These results are consistent with observations presented for other cysteine-knot mutant proteins and *in vivo* Piedmontese studies. The predominant hypothesis is that structural perturbations inhibit the ability of Piedmontese myostatin to bind to its type I receptor. Global changes in conformation may also have an impact on type II receptor binding.

8 Amyloid formation by MstnPP

The results presented so far focus on analysis of the native myostatin protein, whether as a precursor or latent complex, for the WT and Piedmontese mutant. However, thermal denaturation of these species suggests the occurrence of amyloid-like β -aggregation at high temperatures. Furthermore, during *in vitro* refolding of the myostatin precursor protein, large molecular weight soluble aggregates form in addition to the MstnPP dimer, that bear a striking resemblance morphologically to the prefibrillar amyloid aggregates of lysozyme, amyloid β and HypF-N (Walsh, Hartley *et al.* 1999; Malisauskas, Ostman *et al.* 2005; Calloni, Lendel *et al.* 2008). As MstnPP has been implicated in the amyloid disease sporadic inclusion body myositis (sIBM) (Wojcik, Engel *et al.* 2005; Askanas and Engel 2008), the propensity of MstnPP for amyloid formation was investigated.

8.1 *In silico* predictions

A number of *in silico* programs exist that have been shown to accurately predict the regions of a polypeptide chain that are prone to β -sheet aggregation and amyloid formation (He, Wang *et al.* 2009). Those that appeared most popular in the literature were used to analyse the MstnPP amino acid sequence (Fig. 8.1a); Tango (Fernandez-Escamilla, Rousseau *et al.* 2004) and PASTA (Trovato, Chiti *et al.* 2006; Trovato, Seno *et al.* 2007) were selected for the analysis of β -aggregation propensity and Waltz for amyloidogenic regions (Reumers, Schymkowitz *et al.* 2009). Original output from the algorithms is shown in Appendix 5, Figure 12.9. Both Tango and PASTA predicted a high propensity (red) for β -aggregation for residues 154-177 within the propeptide region; residues 139-153 were also predicted by PASTA to have β -aggregation tendencies, albeit at a comparatively lower level (blue). Tango also identified growth factor domain sequences, 314-321 and 347-355. Waltz predicted the same propeptide amyloidogenic regions, 153-162 and 168-177 as well as the very C-terminus of the protein, 348-364. The regions where the three predictions overlap are underlined. Mapping of aggregation-prone regions of the myostatin growth factor onto the crystal structure (Cash, Rejon *et al.* 2009) shows that residues 348-364 localise to the β -sheet ‘fingers’ of the growth factor (Fig. 8.1b).

a.

```

21VDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSKLRLE
TAPNISKDVIRQLLPKAPPRLRELIDQYDVRDDSSDGSLEDDDYHATTETIITMPTES
DFLMQVDGKPCKCFFKFSSKIQYNKVVKAQLWIYLRPVETPTVFVQILRLICKPMK
DGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDL
AVTFPGPGEDGLNPLEVKTDTPKRSRRDGLCDEHSTESRCCRYPLTVDFEAF
GWDWIIAPKRYKANYCSECEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPIN
MLYFNGKEQIYGKIPAMVVDRCGCS 375

```

b.

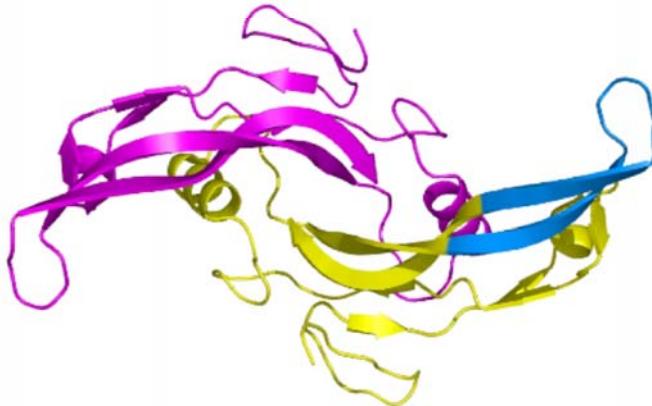


Figure 8.1 *In silico* predictions of propensity for β -sheet aggregation and amyloid formation.

β -aggregation was measured by Tango and PASTA; amyloid formation by Waltz.

- a. MstnPP sequence 21-375 showing regions of elevated propensity at two levels; red regions have higher propensity than those in blue. Regions where Waltz, Tango and PASTA predictions overlap are underlined. The RSRR furin proteolysis sequence that separates the propeptide from the growth factor domain is shown in bold.
- b. Mapping of the 348-364 sequence (blue) onto the myostatin growth factor crystal structure. The region has been shown on one monomer only.

8.2 The amyloid-like properties of MstnPP soluble aggregates

8.2.1 Transmission electron microscopy

High molecular weight, misfolded aggregates were formed during refolding of the MstnPP dimer (Fig. 3.4). Transmission electron microscopy (TEM, Fig. 8.2a (i)) reveals that the soluble aggregates exhibit a morphology and size similar to that documented for amyloid protofibrils from a number of other proteins such as lysozyme (Goda, Takano *et al.* 2000; Malisauskas, Ostman *et al.* 2005) and HypF-N (Calloni, Lendel *et al.* 2008) as well as the insulin protofibrils generated as a positive

control (Fig. 8.2b) using a well-established protocol (Jansen, Dzwolak *et al.* 2005; Groenning, Norrman *et al.* 2007; Mauro, Craparo *et al.* 2007). The insulin sample contains both prefibrillar aggregates (white arrow) and fibrils (black arrow). The myostatin aggregates have the typical ‘bead-on-a-string’ morphology, appearing as elongated oligomers of associated spheres (Fig. 8.2a (ii)) with a diameter of 15.6 ± 1.6 nm (standard deviation, SD). A range of MstnPP structures, including rings (Fig 8.2a (iii)) can be seen, and lateral association of the elongated aggregates is evident (Fig. 8.2a (iv)).

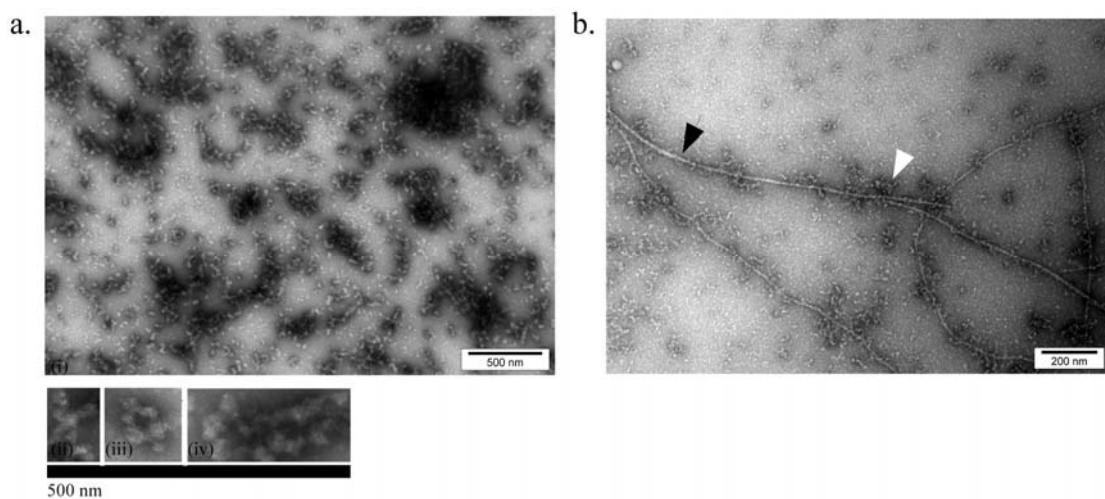


Figure 8.2 Characterisation of MstnPP soluble aggregates by negative-stain transmission electron microscopy.

- a. (i) MstnPP soluble aggregates; (ii - iv) zoomed-in views of (i) showing overlapping of elongated aggregates (ii), ring-like oligomers (iii) and lateral association (iv).
- b. Insulin positive control containing both protofibrils (white arrow) and mature fibrils (black arrow).

8.2.2 Thioflavin T assays

A defining characteristic of amyloid fibrils and protofibrils is their ability to bind the fluorescent dye thioflavin T (ThT) (Walsh, Hartley *et al.* 1999; Bourhim, Kruzel *et al.* 2007; Groenning, Norrman *et al.* 2007). ThT binding assays were carried out on the MstnPP soluble aggregates (Fig. 8.3, green) and compared to a solution containing a mixture of insulin protofibrils and fibrils as a positive control (red), the correctly refolded MstnPP dimer (yellow) and a blank solution (blue). The MstnPP soluble

aggregates bound ThT with intensity comparable to insulin fibrils whereas the correctly folded MstnPP dimer did not, results which suggest that a population of MstnPP can aggregate spontaneously to form amyloid-like oligomers.

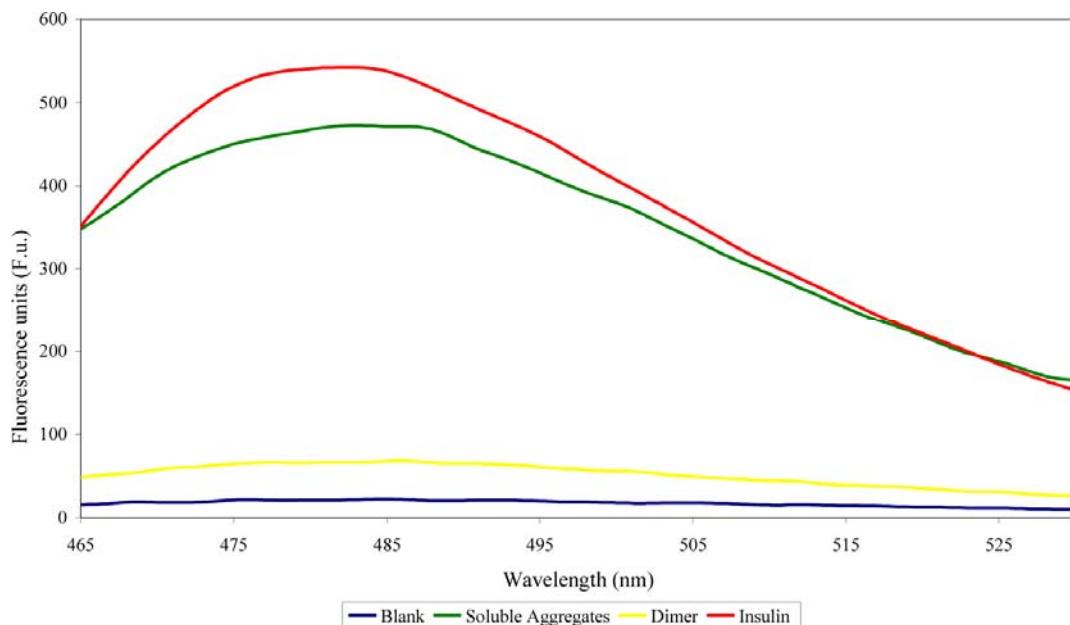


Figure 8.3 Thioflavin T fluorescence of MstnPP soluble aggregates.

MstnPP soluble aggregates (green) were compared to a fibril-containing insulin positive control (red), the refolded MstnPP dimer (yellow) and a buffer-only blank (blue).

8.3 Formation of MstnPP amyloid fibrils at acidic pH and elevated temperature

8.3.1 ThT and TEM

The presence of MstnPP amyloid-like oligomers suggests that amyloid fibril formation is also possible. MstnPP aggregates were concentrated and resuspended in dilute HCl solutions (pH range 1.6 to 6.3) and incubated at 60 °C, conditions including those under which insulin forms amyloid fibrils (Groenning, Norrman *et al.* 2007; Mauro, Craparo *et al.* 2007). Solutions were monitored by ThT binding and TEM.

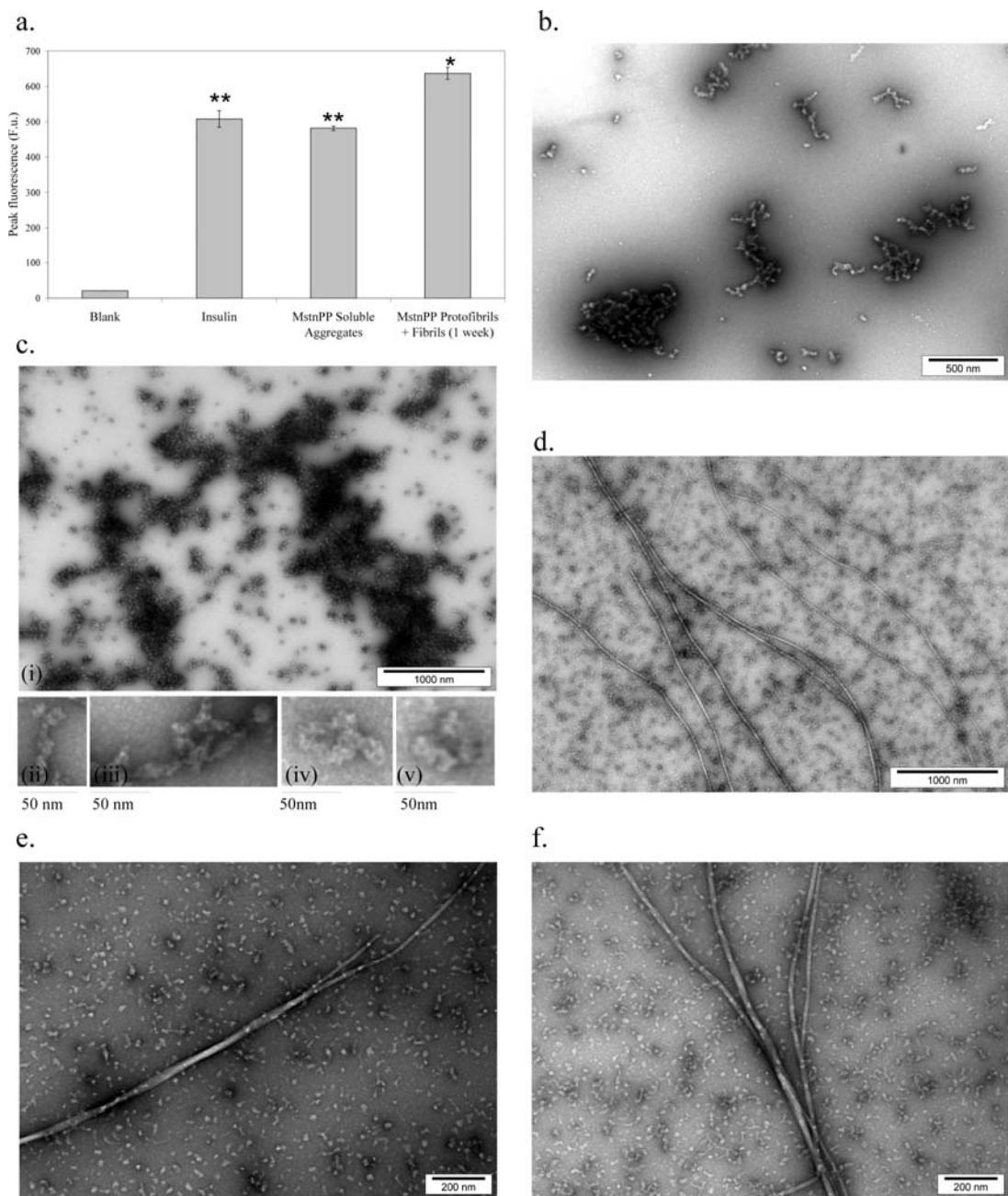


Figure 8.4 Characterisation of MstnPP prefibrillar oligomers and amyloid fibrils by ThT binding and TEM.

- ThT binding after incubation at 60 °C and pH 5.3 for one week compared to insulin and MstnPP soluble aggregates. Each result represents the average of four independent experiments and error bars represent SEM. ** P < 0.005 and * P < 0.01 by paired Student's t-test. Blank is 0.005 mM HCl, pH 5.3 and ThT.
- Overnight incubation at 60 °C in pH 5.3.
- Three days incubation showing (i) large, dense three-dimensional arrays; (ii-v) zoomed in examples of lateral association and longitudinal fusion (ii and iii) and pore-like structures (iv and v).
- Characteristic amyloid fibrils after one week.
- f. Twisting of two (e) or more (f) fibrils around each other.

After one week at 60 °C, ThT fluorescence had increased significantly in the pH 5.3 solution (Fig. 8.4a), which contained a mixture of prefibrillar aggregates and amyloid fibrils, observed by TEM (Fig. 8.4 d-f).

Since ThT binding may be affected by changes in pH (Sabaté, Lascu *et al.* 2008), the fluorescence assays were also performed by diluting protein solutions into a pH 7.5 buffer. In these assays, fluorescence values for the soluble aggregates (SA) were similar for both pH 8.5 and pH 7.5, but an increase in intensity for a mixture of protofibrils and fibrils (PF/F) was observed at pH 7.5 compared to pH 5.3 (Fig. 8.5). These results support the conclusion that the increase in ThT fluorescence previously observed for both soluble aggregates and fibril-containing samples (Figs. 8.3 and 8.4a) is significant.

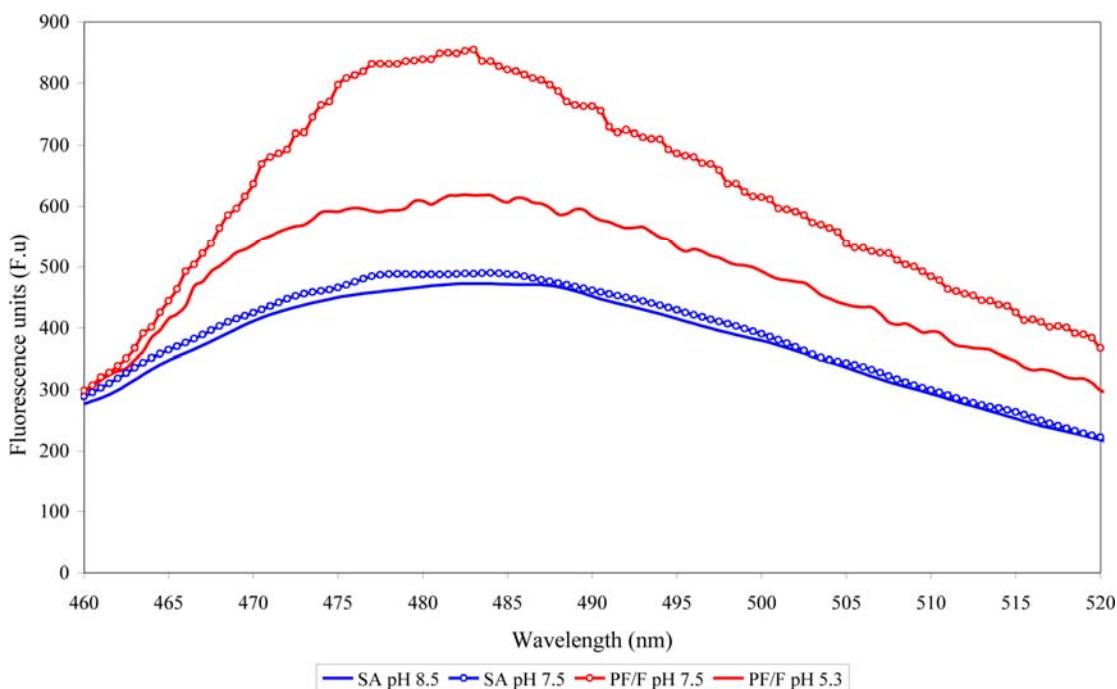


Figure 8.5 ThT fluorescence at pH 7.5.

The ThT fluorescence of soluble aggregates (SA) and a solution containing both protofibrils and fibrils (PF/F) at pH 7.5.

Myostatin fibril formation occurs via a number of stages. For a 3 mg/mL solution, overnight incubation at 60 °C in pH 5.3 produces oligomers that are compact and retain the beaded morphology (Fig. 8.4b) as observed for amyloid-β (Aβ) and

lysozyme fibrils (Walsh, Hartley *et al.* 1999; Malisauskas, Ostman *et al.* 2005). By three days, extensive arrays of aggregation were present; often three-dimensional, these appeared as areas of high electron density under the TEM (Fig. 8.4c (i)). Increased elongation and in some places lateral association and longitudinal fusion of oligomers can be observed (Fig. 8.4c (ii) and (iii)). This granular to smooth transition has been shown during A β fibril formation (Walsh, Hartley *et al.* 1999). Smaller pore-like oligomers, previously documented on the amyloid formation pathway of a number of proteins (Lashuel 2005; Juárez, Taboada *et al.* 2009), are also apparent (Fig. 8.4c (iv) and (v)). By one week, fibrils that show the characteristic morphology of amyloid had appeared, with a diameter of 15.4 ± 0.7 nm (SD) and a straight, unbranching structure (Fig. 8.4d). The fibrils are extremely long, in excess of 5 μm . At two weeks, higher order amyloid structures can be observed, such as two (Fig. 8.4e) or more (Fig. 8.4f) protofilaments twisting around each other to give diameters of 24.7 ± 0.9 nm (SD) and 40.7 ± 0.9 nm (SD) respectively.

Twisting is periodic for individual fibrils. Although the distance between twists is similar for fibrils from the same preparation, this distance varies notably between different preparations (Table 8.1). Fibrils 1 and 2 are from the same sample and have similar twist distances. Fibril widths of 21.6 ± 0.8 (SD) and 18.5 ± 0.6 (SD) nm respectively suggest twisting of two protofilaments. The reduced diameter of fibril 2 suggests a more stretched structure than fibril 1. Fibrils 3 and 4 have twist-distances at least 3-fold greater than 1 and 2 despite similar widths. Fibril 3 is likely to contain two protofilaments with a mean width of 26.3 ± 0.7 nm (SD); the width of fibril 4 is 42.5 ± 0.9 nm (SD), indicating three protofilaments twisted around each other. Fibril dimensions are known to be influenced by factors such as the number of protofilaments per fibril and internal interactions (Sunde, Serpell *et al.* 1997; Jimenez, Nettleton *et al.* 2002; Stromer and Serpell 2005). The inherent kinetics of different preparations may also account for some of the differences seen here (Jansen, Dzwolak *et al.* 2005; Kumar and Udgaonkar 2009).

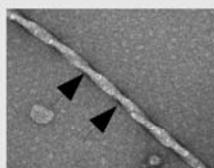
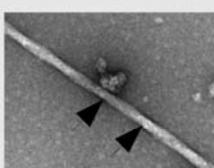
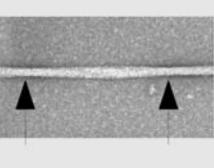
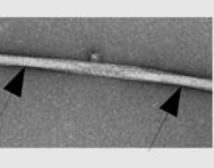
	1	2	3	4
TEM				
Distance between twists (nm)	90.4 ± 6.3	100.4 ± 4.5	304.2 ± 2.9	431.5 ± 16.6

Table 8.1 Measurement of distance between twists for MstnPP fibrils.

iTEM software (Olympus) was used for measurement and statistical analysis. Arrows indicate representative points between which measurements were taken.

8.3.2 Circular dichroism spectroscopy

Amyloid formation is accompanied by the adoption of a β -sheet rich secondary structure regardless of the structure of the native protein (Marshall and Serpell 2009). CD spectroscopy was used to study the structural changes that occur during amyloid formation (Fig. 8.6). The spectrum of correctly refolded MstnPP (blue) has been shown previously (Fig. 4.2) and is indicative of a mixture of α -helix, β -sheet and random coil.

Quite a different profile is observed for the MstnPP soluble aggregates (green) with a very strong minimum at 208 nm indicating secondary structure that is primarily α -helical in nature. There is broadening of the spectrum between 215-225 nm, which may arise from contributions of β -sheet and α -helix to the spectra. The CD spectra of protofibrils of other amyloid-forming proteins are characterised by β -sheet absorption at 215-218 nm (Picotti, Franceschi *et al.* 2007; Calloni, Lendel *et al.* 2008; Rezaei-Ghaleh, Zweckstetter *et al.* 2009), indicating that the MstnPP soluble aggregates represent the α -helical oligomeric intermediate observed for the majority of proteins on the pathway to protofibril formation (Abedini and Raleigh 2009).

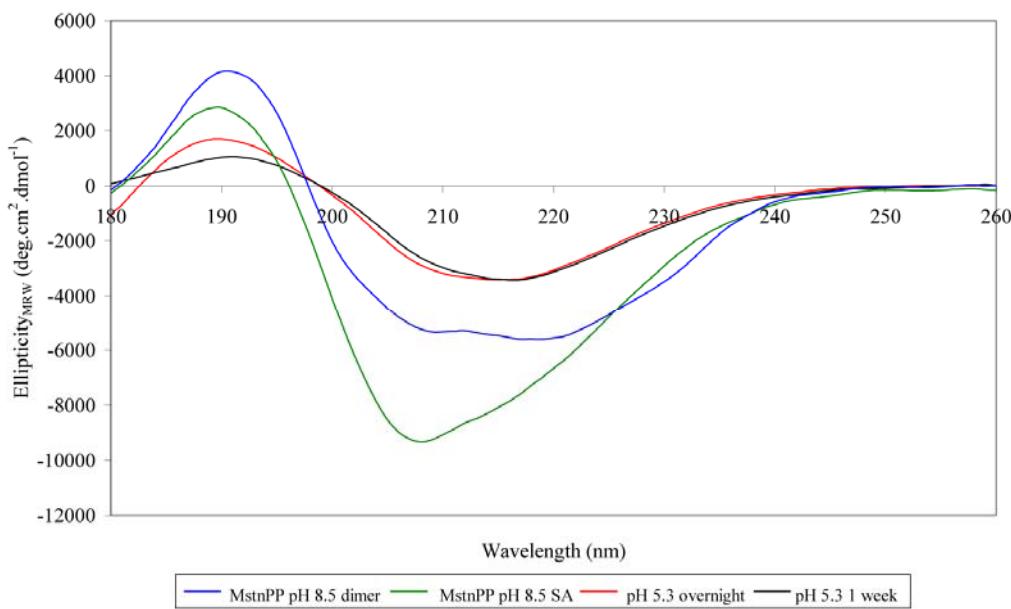


Figure 8.6 Circular dichroism spectra showing the structural transitions that occur in amyloid fibril formation.

Correctly folded MstnPP dimer (blue); MstnPP soluble aggregates (SA) before acidification (green); prefibrillar aggregates after overnight incubation in pH 5.3 at 60°C (red); mixture of prefibrillar aggregates and fibrils after one week incubation (black).

Incubation of the aggregates overnight at pH 5.3 and 60 °C produces a CD spectrum dominated by β -sheet (red) with the α -helical minimum at 208 nm completely replaced by a minimum at 218 nm. TEM images of the same sample reveals there are no fibrils at this stage but the aggregates are more compact and are present in dense arrays (Fig. 8.4b and c), which may represent the start of β -sheet stacking as amyloid protofibrils. The ellipticity for the spectrum is comparatively low, similar to that observed in the fibrillogenesis of A β (Walsh, Hartley *et al.* 1999). This reduction of signal is most likely due to differential absorption flattening, a phenomenon often seen in the CD spectra of samples containing suspensions of solid-phase material (Calloni, Lendel *et al.* 2008) and causing both a decrease in intensity and red-shift of all minima. After one week of incubation the spectrum is similar (black trace) except for flattening of the 190 nm peak. At this stage, fibrils are present as observed by TEM analysis of an aliquot of the same sample (Fig. 8.4d). The further flattening of the spectrum over time is likely to be due to both an increase in β -sheet structure as well as differential absorption flattening (Kelly, Jess *et al.* 2005; Calloni, Lendel *et al.*

2008). CD shows that the soluble aggregates differ dramatically in secondary structure compared to native MstnPP and that the aggregates may represent an α -helical-containing intermediate in the transition to protofibrils and fibrils rich in β -sheet structure. Deconvolution of the CD spectra using CDDN agrees with the observations made above (Appendix 4, Table 12.2b).

8.3.3 Resistance to proteolytic digest by trypsin

Resistance to proteolytic digestion is another defining characteristic for the presence of highly stable amyloid fibrils (Bocharova, Breydo *et al.* 2005; Hartley, Zhao *et al.* 2008). Comparative trypsin digestions were performed for the MstnPP fibril-containing sample, soluble aggregates and native dimer (Fig. 8.7).

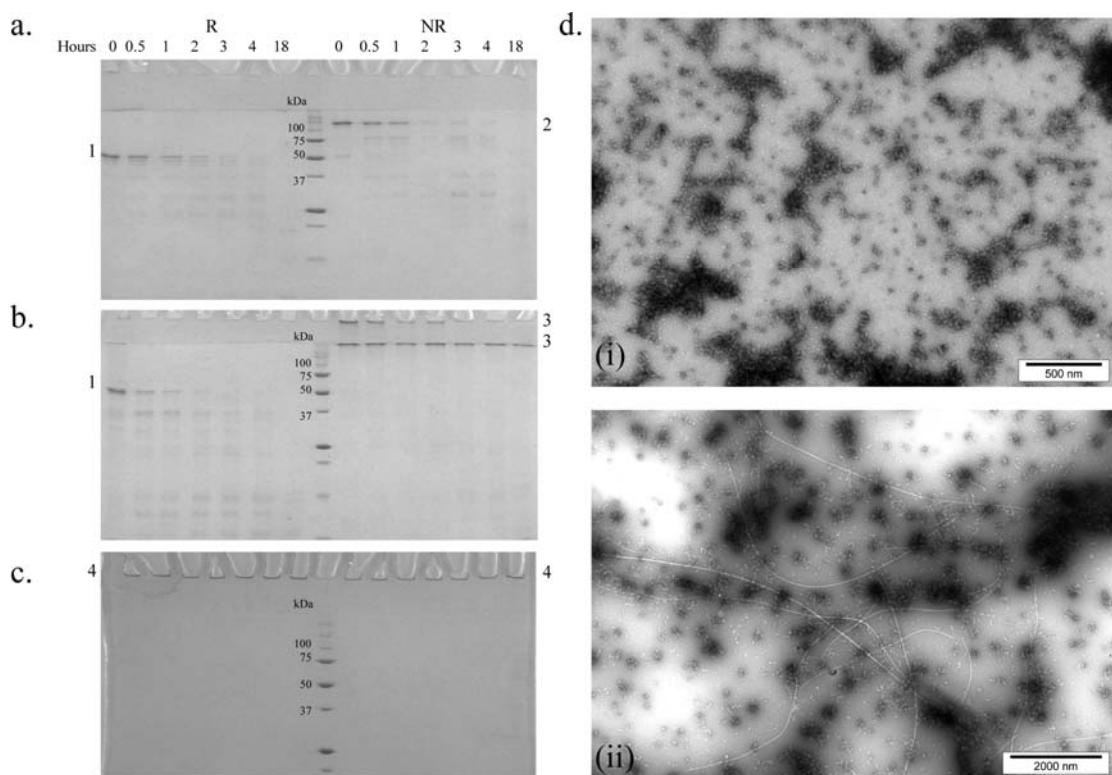


Figure 8.7 MstnPP resistance to trypsin digestion by SDS-PAGE analysis and TEM.

a.-c. Reducing (R) vs. non-reducing (NR) SDS-PAGE of dimer (a), soluble aggregates (b) and a MstnPP protofibril/fibril mixture (c). Major bands are indicated: 1. MstnPP monomer; 2. MstnPP dimer; 3. MstnPP soluble aggregates; 4. prefibrillar aggregates plus fibrils.

d. TEM shows resistance of (i) protofibrils and (ii) amyloid fibrils to trypsin.

Trypsin digests were performed with a MstnPP:trypsin ratio of 100:1 (a and b) or 20:1 (c and d) for a total of 18 hours at 37°C.

Digestion was carried out with a MstnPP to trypsin ratio of 100:1 (w/w) at 37 °C with samples taken after 0.5, 1, 2, 3, 4 and 18 hours. The precursor dimer and aggregates have comparable trypsin susceptibility and are fully digested after an overnight incubation (18 hours, Fig. 8.7a and b respectively). The susceptibility of the soluble aggregates to trypsin digest suggests an open, flexible structure, supporting the conclusion that this species represents a prefibrillar intermediate rather than protofibrils. Prior to digestion the fibril-containing sample is not able to enter the top of the 4% stacking gel owing to its extremely large size. Following incubation with trypsin this property is maintained even after overnight incubation (18h) and a 5-fold increase in trypsin concentration (MstnPP:trypsin 20:1) (Fig 6c). Inability to enter the stacking gel is not a definitive indication of proteolysis resistance since partial hydrolysis may result in products smaller than fibrils and/or protofibrils yet still large enough to be retained in the stacking gel. To address this possibility, TEM analysis revealed that both MstnPP protofibrils (Fig. 8.7d (i)) and mature fibrils (Fig. 8.7d (ii)) had unchanged morphology after trypsin incubation.

8.3.4 SDS resistance and DTT

Amyloid structures are known to be SDS-resistant (Cerf, Sarroukh *et al.* 2009) and disulphide bonding has been suggested to play a role in the amyloid formation pathway for some proteins (Yamamoto, Yagi *et al.* 2008; Fei and Perrett 2009; Wang, Liu *et al.* 2009). The MstnPP sequence contains 13 cysteine residues; therefore, the effects of SDS and DTT on the MstnPP aggregates were investigated.

The inability of both protofibrils and fibrils to enter the 4% SDS-PAGE stacking gel (Figure 8.7c) suggests resistance to solubilisation in SDS. The soluble aggregates are able to enter the stacking gel yet not the 12% resolving gel (Fig. 8.7b). However, this is only the case in the absence of reducing agent; when reducing SDS-PAGE buffer is used, the soluble aggregates are represented as a 50 kDa monomer band (Fig. 8.7b, R), suggesting that disulphide-bonding is a factor in aggregate stabilisation. Overnight incubation and β -sheet formation remove sensitivity to reducing agent, raising the hypothesis that disulphides involved in structural stability of the amyloid species are buried and inaccessible at this stage.

To study the effects of reduction on MstnPP amyloid formation in more detail, soluble aggregate samples were incubated with different concentrations of DTT in both purification (pH 8.5, 4°C) and amyloid-promoting (pH 5.3, 60 °C) conditions and visualised under TEM. The MstnPP dimer (Fig. 8.8a,***) can be reduced to the monomeric form (*) with 1 mM DTT; 10 mM DTT results in complete reduction. Addition of 1 mM DTT to MstnPP soluble aggregates at pH 8.5 appears to partially suppress the formation of prefibrillar aggregates (Fig. 8.8b). Rather than a high concentration of elongated species with a ‘bead-on-a-string’ morphology, the grid is predominated by single sphere-like particles or beads (Fig. 8.8b, inset), suggesting that the interaction is partly mediated by disulphide-bonding. Increasing the concentration of DTT to 10 mM has a similar effect (Fig. 8.8c). The dense arrays are largely removed but some elongation is still apparent. Incubation of soluble aggregates with DTT in amyloid-promoting conditions resulted in precipitation of MstnPP at both DTT concentrations, possibly due to exposure of hydrophobic regions previously hidden due to disulphide interactions. Precipitation could not be studied by TEM due to the degenerative effects this has on the instrument. However, in more dilute areas of the grid (Fig. 8.8d), protofibrillar aggregates that displayed striking ring-like morphologies (Fig. 8.8d, inset) could be observed, suggesting DTT-independent aggregation or the aggregation of disulphide-bonded structures that escaped reduction.

ThT binding of soluble aggregates (SA) in the presence of 10 mM DTT was unchanged, suggesting that disulphide bond formation has no effect on the site of ThT binding (Fig. 8.8e). However, DTT reduction alters the CD spectra of soluble aggregates (Fig. 8.8f); the α -helical minimum at 208 nm is increased, there is a broadening and increase in intensity of the β -sheet minimum from 215-220 nm, the intensity of the maximum at 190 nm is increased and the spectrum has shifted to the right. These changes indicate an increase in α -helical structure and suggest that disulphide-bonding plays a role in the maintenance of β -sheet structures in the soluble aggregates. All results shown in Figure 8.8 represent partial reduction only; reducing SDS-PAGE (Fig. 8.7) renders the soluble aggregates to a monomeric form with a β -mercaptoethanol concentration of 2 M. However, taken together, the results presented here support the hypothesis that disulphide-bonding plays a role in amyloid formation by MstnPP.

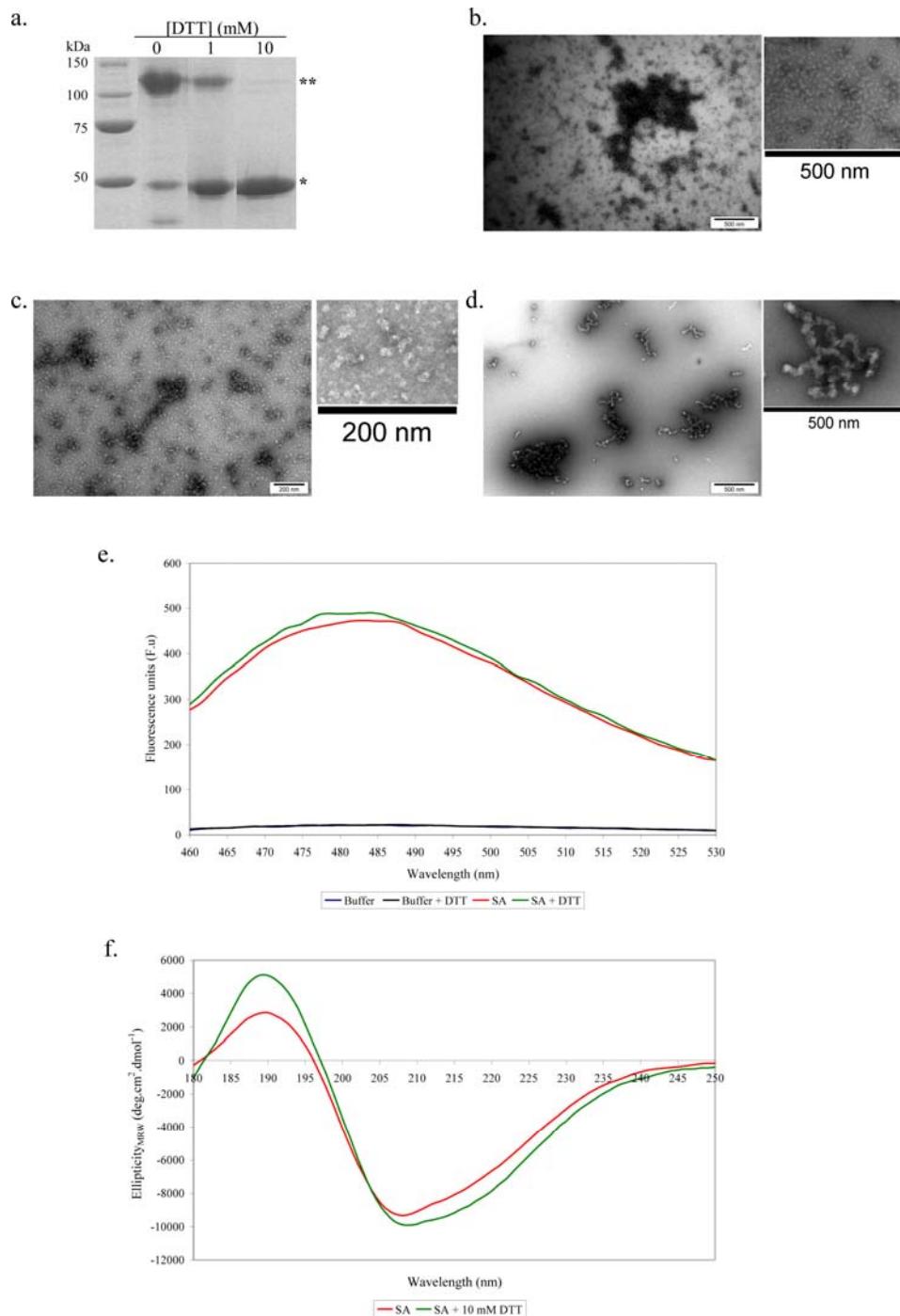


Figure 8.8 The effects of DTT on amyloid formation shown by TEM, ThT binding and CD spectroscopy.

- Non-reducing SDS-PAGE showing DTT reduction of MstnPP dimer (**) to the monomeric form (*). Intervening gel lanes have been removed.
- Addition of 1mM (b) and 10 mM (c) DTT to soluble aggregates at pH 8.5, 4 °C.
- Addition of 10 mM DTT to soluble aggregates at pH 5.3, 60 °C.
Insets for b.-d. show zoomed-in views of main image.
- ThT binding of soluble aggregates (SA) in the presence (green) and absence (red) of 10 mM DTT compared to buffer-only (blue) and buffer-only + DTT (black) blanks.
- CD spectra of soluble aggregates (SA) in the presence (green) and absence (red) of 10 mM DTT.

8.4 Alternative conditions for amyloid formation

8.4.1 Physiological temperature

The formation of amyloid fibrils by MstnPP at 60 °C and pH 5.3 is comparable to conditions routinely used to promote rapid amyloid formation in *in vitro* model systems for amyloid diseases. However, analysis of secondary structure from 10 to 65 °C immediately after suspension in pH 5.3 using CD indicates that the β-sheet transition may occur at physiological temperature; the spectrum had changed significantly by 45 °C with a loss of α-helix beginning from 30 °C and the appearance of a characteristic amyloid β-sheet-rich spectrum at 65 °C (Fig. 8.9a). Interestingly, the erratic absorption observed for correctly folded MstnPP and the latent complex between 190 and 200 nm (Fig. 4.5) is not visible; this may be due to structural differences. Subsequently, MstnPP aggregates were incubated at 37 °C at pH 5.3 and monitored with ThT binding, CD spectroscopy and TEM (Fig. 8.9b,c, and d respectively).

After an overnight incubation, ThT binding had increased only slightly (Fig. 8.9b), although CD indicated that β-sheet aggregation has occurred (Fig. 8.9c). Under TEM, the 37 °C protofibrils have a similar morphology to the 60 °C samples, with elongation, clustering and ring-like structures apparent (Fig. 8.9d). However, aggregation is less extensive. After one week of incubation at 37 °C, ThT binding had increased significantly (Fig. 8.9b) and CD showed a definite β-sheet transition (Fig. 8.9c). TEM images show increased aggregation, lateral association and a granular to smooth morphology transition (Fig. 8.9e). Although fibril formation did not occur at 37 °C over the time period observed for formation at 60 °C, it is likely that the kinetics of fibril formation will be slower at the reduced temperature. However, these results show that the properties of protofibrils, proposed to be direct precursors of mature amyloid fibrils, have been adopted at 37 °C; amyloid formation is likely to occur over an extended period of time.

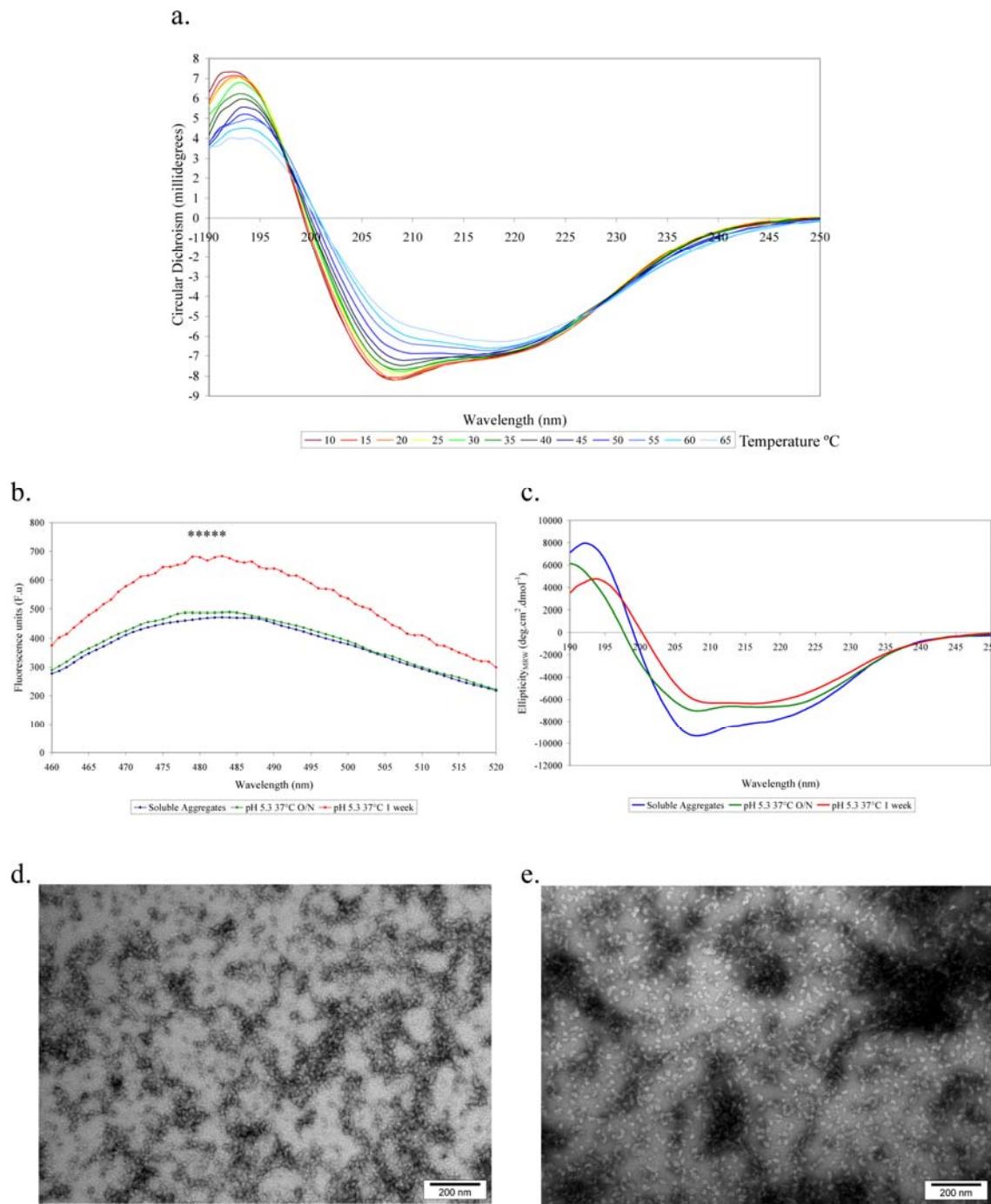


Figure 8.9 Characterisation of MstnPP amyloid formation at 37°C by CD, ThT binding and TEM.

- a. CD thermal denaturation of MstnPP soluble aggregates from 10 to 65°C.
- b.-c. ThT binding and fluorescence (b) and CD spectra (c) of soluble aggregates (blue) after overnight (green) and week-long (red) incubations. *****, P < 0.001 by Student's t-test using averaged values over peak.
- d. and e. TEM of soluble aggregates after overnight (d) and week-long (e) incubations.

8.4.2 Alternative pH, alternative pathway

A number of amyloid-forming proteins such as barstar (Kumar and Udgaonkar 2009) and insulin (Jansen, Dzwolak *et al.* 2005) exhibit differences in the morphology of structural intermediates and the kinetics of amyloid formation when the conditions used for amyloid formation are altered. The effects of temperature on the kinetics of amyloid formation by MstnPP have been shown (Fig. 8.9); alterations in pH produced further differences. Incubation at 60 °C in pH 1.6 (as used for insulin), 4.3, and 6.3 did not lead to fibrillization of MstnPP. The protein remained in solution as amorphous aggregates (Fig. 8.10a), an intriguing result as the same buffer was used for all protein suspensions. However, incubation in pH 3.3 allowed fibril formation with structural intermediates that displayed dramatic morphological differences (Fig. 8.10b-d).

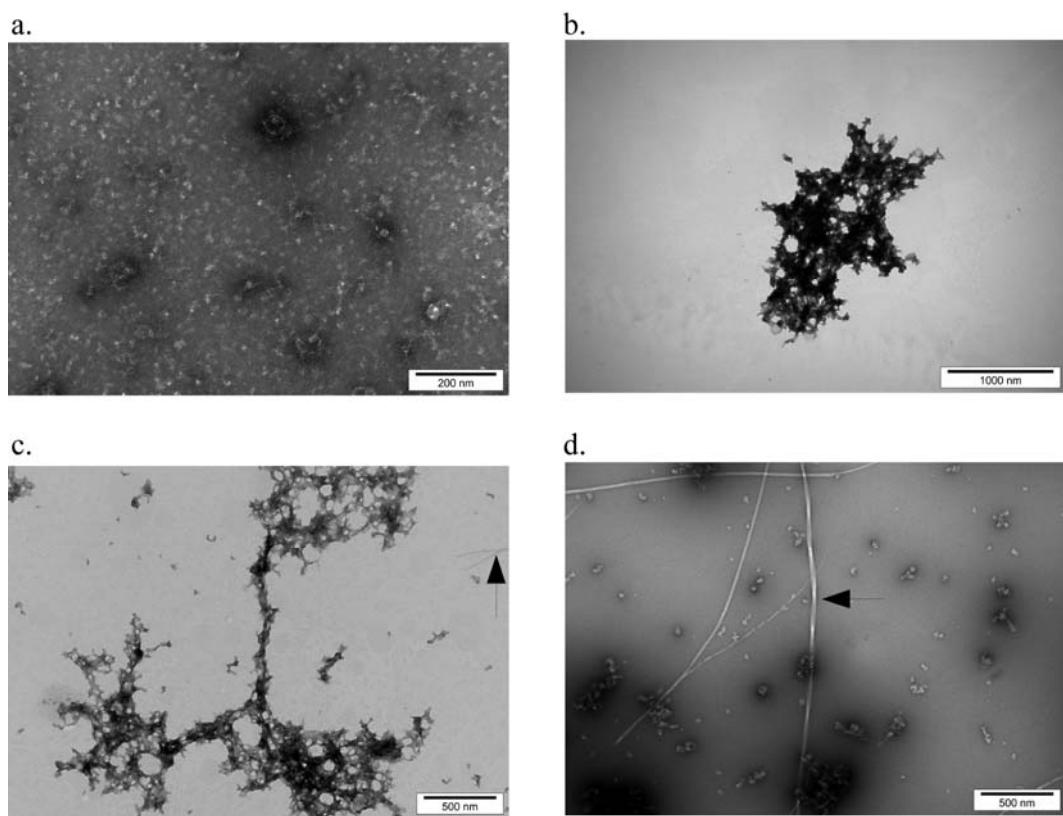


Figure 8.10 Characterisation of MstnPP amyloid formation at alternative pH values by TEM.

- Incubation at pH 1.6, 4.3 and 6.3 does not result in aggregation or fibril formation.
- pH 3.3 produces morphologically different oligomers after overnight (b) and 72 hours (c) incubation, with fibrils after one week (d).

Overnight incubation yielded large aggregates that have a squashed morphology (Fig. 8.10b). The electron-dense regions seen in the pH 5.3 sample are not apparent, suggesting that the pH 3.3 arrays are not as three-dimensional. The individual beads within the clusters have a smooth, fluid-like appearance and lateral association is a predominant feature. After 72 hours incubation, short fibrils had appeared amongst the aggregates (Fig. 8.10c, arrow), which had retained the alternative morphology. This time-scale is reduced in comparison to the pH 5.3 incubation, indicating that changes in pH affect the kinetics of amyloid formation.

After one week, longer, unbranching, mature fibrils were present (Fig. 8.10d) and twisting could be observed (Fig. 8.10d, arrow). At 16.7 ± 1.1 nm (SD) in diameter, the pH 3.3 fibrils are slightly wider than those from pH 5.3. The twist distance of the fibril shown here (Fig. 8.10d, arrow) is 289.7 ± 7.35 nm (SD). Intriguingly, the initial oligomer conformation had disappeared and the oligomeric species surrounding the fibrils looked very similar to those found at pH 5.3, suggesting that the alternative oligomers were a transient intermediate only. These results imply that there are different pathways for amyloid formation by MstnPP, influenced by the environment. The mature fibrils appear quite similar, except for the slight increase in diameter at pH 3.3. It has been suggested that alternative structures for oligomeric species result in differences in fibril morphology also; the techniques used for this thesis are not of a resolution to investigate this.

8.5 Formation of amyloid by the refolded MstnPP dimer

In contrast to the soluble aggregates (Fig. 8.9a), thermal denaturation of the refolded MstnPP dimer suggested an occurrence of aggregation at high temperatures. The WT MstnPP dimer was investigated for an ability to form amyloid by TEM.

TEM analysis of the MstnPP dimer (Fig. 8.11a (i)) suggests that while the majority of the protein is present as small oligomers that likely represent the dimer (Fig. 8.11a (ii)), elongated species are also present (Figure 8.11a (iii) and (iv)). MstnPP is observed as a 125 kDa band on non-reducing SDS-PAGE immediately after purification (Fig. 3.7); however, spontaneous aggregation of this species over time at 4 °C has been noticed. This is not a consistent phenomenon, suggesting that contamination of the dimer species with soluble aggregates is responsible and that

purity of the refolded dimer is paramount if aggregation is to be avoided. Analysis of the sample shown here by non-reducing SDS-PAGE confirms the presence of higher molecular weight species (Fig. 8.11b,*) in addition to the dimer (**); however, these are smaller than the soluble aggregates present after refolding and can enter the gel. Sizes of 225 and 250 kDa suggest the presence of tetramers and pentamers respectively.

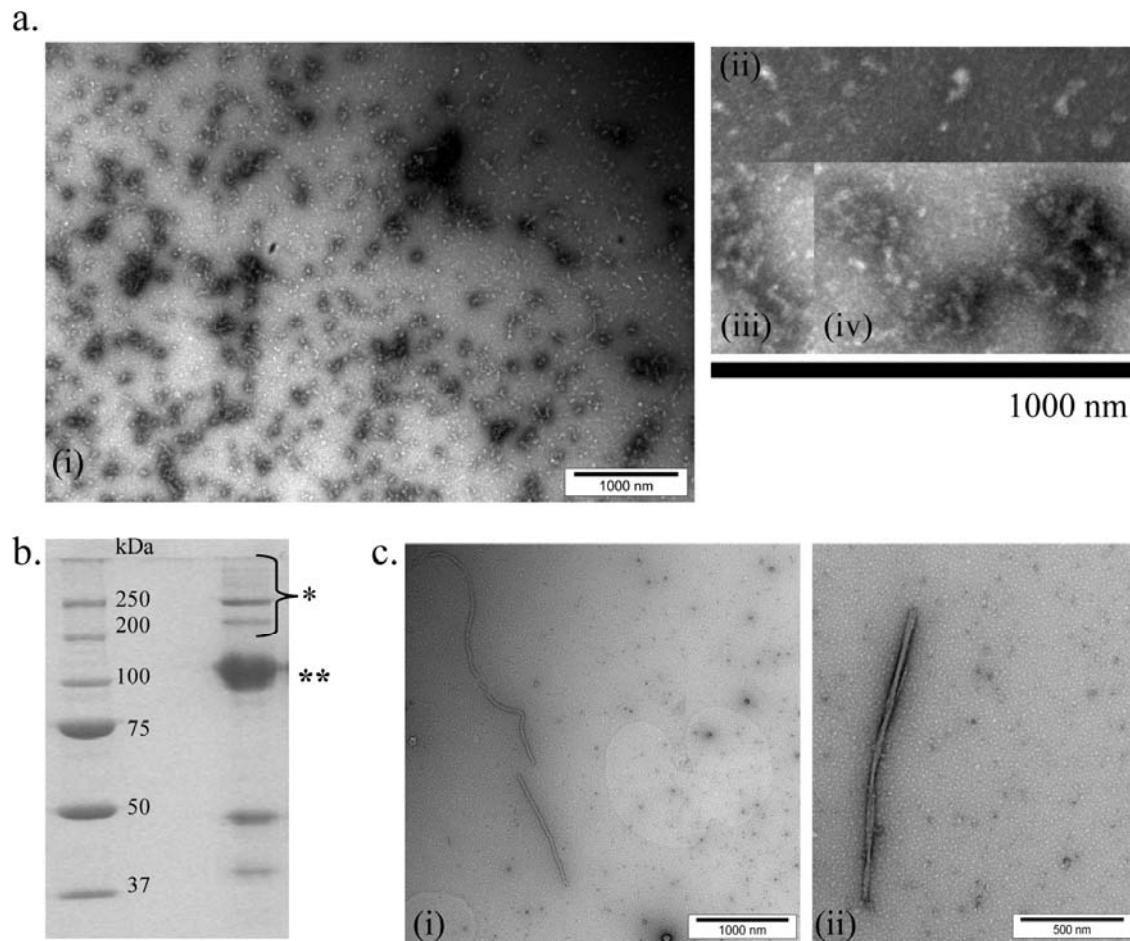


Figure 8.11 Characterisation of amyloid formation by the refolded MstnPP dimer by TEM.

- a. (i) TEM of the MstnPP dimer in pH 8.5. Insets show zoomed-in views of small oligomers, likely to represent the dimer (ii), and elongated species (iii and iv).
- b. Non-reducing SDS-PAGE of purified dimer (***) and high MWt oligomers (*).
- c. MstnPP dimer amyloid fibrils. (i) single protofilament, and (ii) two protofilaments twisted around each other.

Consistent with the suggestion of β -aggregation during CD thermal denaturation (Fig. 4.5), incubation of the MstnPP dimer at 60 °C and pH 5.3 produces amyloid fibrils (Fig. 8.11c) that appear both as single protofilaments (Fig. 8.11c (i)) and as two protofilaments twisted around each other (Fig. 8.11c (ii)). The diameter of dimer fibrils is 13.6 ± 0.7 nm (SD), reduced in comparison to those from aggregated protein. The twisted fibril in Figure 8.11c (ii) has a diameter of 34.4 ± 1.5 nm (SD) and a twist-distance of 350.2 ± 2.3 nm (SD). Although detailed analysis of the kinetics of amyloid formation has not been carried out, the MstnPP dimer appeared to form mature amyloid at a faster rate than the soluble aggregates, at an equivalent protein concentration. The regions involved in amyloid formation may be more accessible in the dimer; as the soluble aggregate solution is heterogeneous, some species may require partial unfolding for amyloid formation.

8.6 Biological activity of MstnPP aggregates

The mediator of amyloid pathogenesis is proposed to be oligomeric prefibrillar intermediates (Stefani and Dobson 2003). Oligomeric species in the amyloid formation pathway of a number of proteins are cytotoxic when added to the media of cultured cells (Walsh, Hartley *et al.* 1999; Bucciantini, Giannoni *et al.* 2002; Bucciantini, Calloni *et al.* 2004). The effect of MstnPP soluble aggregates, protofibrils and fibrils on the viability of C2C12 mouse myoblasts, the standard model cell-line used for the analysis of myostatin activity (Thomas, Langley *et al.* 2000; Joulia, Bernardi *et al.* 2003), was investigated by monitoring the absorbance of formazan (Fig. 8.12) produced after addition of WST-1 (Roche).

25 μ M of soluble aggregates (SA) and at least 10 μ M of protofibrils (PF) decreased cell viability significantly compared to buffer (B) of the same pH and the correctly folded MstnPP dimer. Lower concentrations of both SA and PF had a reduced effect. Although not significant by Student's t-test, this may still be a biologically relevant result. A solution containing a mixture of MstnPP protofibrils and fibrils (PF/F) had a negative yet not significant effect on C2C12 cell survival, in line with the suggestion that mature fibrils may be an inert species (Stefani and Dobson 2003). These results indicate that oligomeric species in the amyloid formation pathway of MstnPP affect the normal functioning of C2C12 myoblasts.

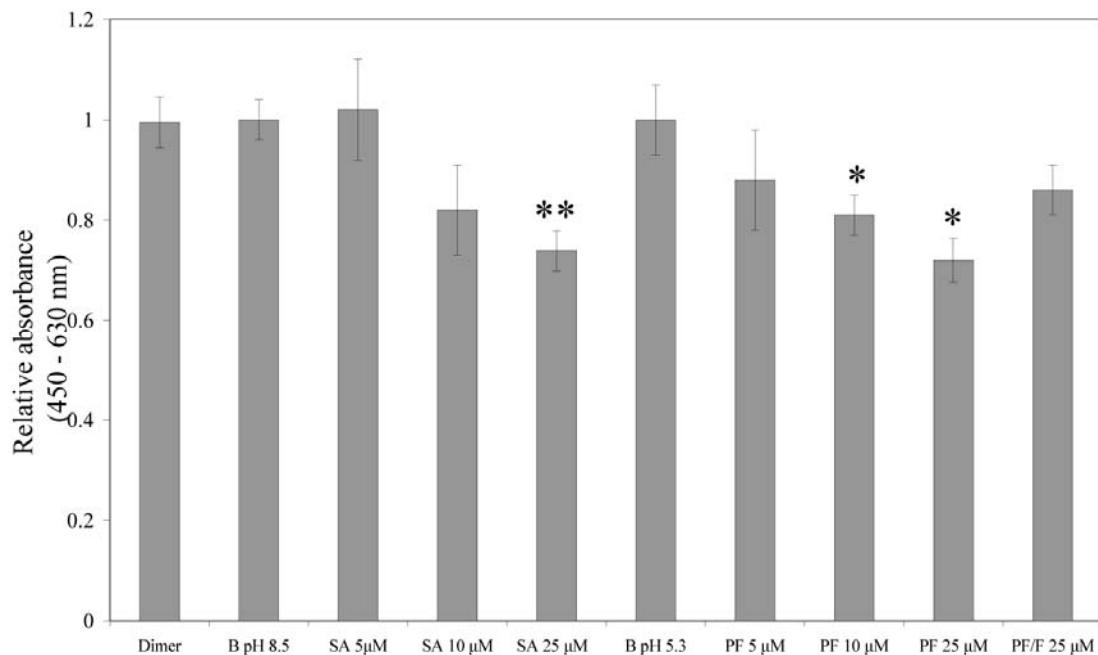


Figure 8.12 MstnPP amyloid C2C12 cytotoxicity assay

WST-1 reagent was used to monitor the reduction of formazan. The difference in absorbance at 450 and 630 nm directly correlates to cell density after incubation with increasing concentrations of MstnPP soluble aggregates (SA, 5 – 25 μ M), protofibrils (PF, 5 – 25 μ M), 25 μ M MstnPP dimer and 25 μ M PF/ F (fibrils). Concentrations are expressed as monomer equivalents. Cells incubated in media containing buffer (B) only (50 mM Tris-HCl pH 8.5, 150 mM NaCl for soluble aggregates and dimer; 0.005 mM HCl, pH 5.3 for protofibrils and fibrils) were used as a control. Error bars represent the standard error of the mean for triplicate samples from two independent experiments. Statistical significance was calculated using a paired Student's t-test where ** $P < 0.01$ and * $P < 0.05$.

9 Discussion: Amyloid formation by MstnPP

9.1 *In vivo implications of MstnPP amyloid formation and sporadic inclusion body myositis*

Amyloid formation is a predominant feature of sporadic inclusion body myositis (sIBM) (Askanas and Engel 2008) and involves the Alzheimer's disease protein A β (Wojcik, Engel *et al.* 2005; Vattemi, Nogalska *et al.* 2009). A role for MstnPP in sIBM was proposed after co-localisation and direct association with A β was observed in diseased cells (Wojcik, Engel *et al.* 2005), and because ER stress causes an upregulation of myostatin expression (Nogalska, Wojcik *et al.* 2007). The results presented here show that misfolded human MstnPP aggregates share the morphological characteristics of amyloid oligomers and the ability to bind the amyloid-specific dye thioflavin T. The α -helical predominant secondary structure of these aggregates suggests they are likely to represent an intermediate species that occurs in the transition between native protein and β -sheet-rich protofibrils. When subjected to the mildly denaturing conditions of pH 5.3 and 60 °C, the aggregates form amyloid protofibrils and after one week, are able to form long, linear and unbranching amyloid fibrils. At 37 °C, β -sheet rich protofibrils form over an increased time period. These *in vitro* results show that MstnPP is capable of amyloid protofibril and fibril formation, supporting a role for misfolding of the myostatin precursor in the pathogenesis of sIBM. Furthermore, MstnPP aggregates and protofibrils have a cytotoxic effect on mouse myoblasts when added to the culture medium.

9.1.1 Implications for the folding of MstnPP *in vivo*

After *in vitro* refolding of the myostatin precursor protein, misfolded yet soluble aggregates that form amyloid protofibrils and fibrils at a lowered pH, can be observed. Although elevated temperature has been used to produce mature MstnPP fibrils quickly in this study, a physiological temperature allows protofibril formation over an extended time period; by comparison, fibril formation is expected to occur over many years *in vivo* and sIBM is considered to be a disease of the elderly (Garlepp and Mastaglia 2007; Karpati and O'Ferrall 2009).

Growth factors such as myostatin and most other secreted proteins are folded in the ER. A major obstacle to folding of native TGF- β family proteins is the formation of the correct disulphide-bonded cysteine pairs (Heiring and Muller 2001; Muller and Heiring 2002); therefore, the correct redox environment and assistance from disulphide bond chaperones is critical. The myostatin propeptide has been proposed to have a chaperone-like function in the correct folding of the growth factor domain (Jin, Dunn *et al.* 2004; Funkenstein and Rebhan 2007). During ER-stressed situations, such as an altered redox environment, protein misfolding can be extensive (Pratico 2008; Schroder 2008). Although the unfolded protein response (UPR) results in the elimination of misfolded proteins and/or apoptosis if necessary (Paschen 2003; Lai, Teodoro *et al.* 2007; Schroder 2008), if the production of misfolded proteins overwhelms the UPR, protein aggregation may lead to amyloid formation (Bucciantini, Giannoni *et al.* 2002; Bucciantini, Calloni *et al.* 2004).

It is hypothesised that in an ER-stressed situation, newly-translated MstnPP may have the potential to misfold and form similar aggregates to those now observed *in vitro*. Although in most situations the UPR will function to return the ER to homeostasis thereby eliminating misfolded myostatin, overwhelming stress may promote the aggregation and structural rearrangement of misfolded forms of myostatin, allowing amyloid association and protofibril formation. This hypothesis is made more complex by the observation that the correctly folded MstnPP dimer is also able to form amyloid aggregates and fibrils (Fig. 8.11). These results imply that in an ER-stressed situation, MstnPP that has already undergone chaperone-assisted folding may also be capable of partial denaturation and structural rearrangements that allow amyloid formation.

9.1.2 Implications for sporadic inclusion body myositis

A role for myostatin and/or MstnPP in the pathogenesis of sIBM was suggested in 2005 when it was shown in 12 sIBM biopsies that myostatin/MstnPP was accumulated within muscle fibers and colocalised with A β (Wojcik, Engel *et al.* 2005). A subsequent study showed similar results in an inclusion body myositis culture model (Wojcik, Nogalska *et al.* 2007), and that overexpression of A β PP/A β

promoted the accumulation and aggresomal targeting of MstnPP. Further research suggested that MstnPP expression, at both the mRNA and protein level, was increased following ER stress in cultured human muscle fibres (Nogalska, Wojcik *et al.* 2007). This upregulation is most likely mediated by NF-κB; MstnPP has an NF-κB consensus site in its promoter and upregulation was not observed in the presence of NF-κB inhibitors (Nogalska, Wojcik *et al.* 2007). Furthermore, treatment with the anti-oxidant resveratrol reduced both NF-κB activation and myostatin expression (Nogalska, D'Agostino *et al.* 2008).

How might myostatin/MstnPP be involved in the pathogenesis of sIBM?

Increased activity of the myostatin growth factor

One possibility is that an increase in the level of circulating myostatin growth factor, due to increased expression and processing of MstnPP, contributes to the prominent muscle-fiber atrophy (Zimmers, Davies *et al.* 2002) as well as the diminished potential of satellite cells to proliferate (Karpati and O'Ferrall 2009). However, the observation that MstnPP is present as amorphous aggregates (Wojcik, Engel *et al.* 2005) and the possible aggresomal localisation (Wojcik, Nogalska *et al.* 2007) may inhibit processing by furin convertase. Aggresomes have been proposed to have a cytoprotective function, serving as recruitment centres to facilitate the degradation of toxic proteins (Taylor, Tanaka *et al.* 2003). Aggregated MstnPP may be targeted for degradation by the ubiquitin pathway, which is upregulated by the UPR (Schroder 2008), supporting a reduction in active growth factor levels.

Furin cleavage of soluble aggregates was not attempted in this thesis. However, after furin digestion, a proportion of undigested MstnPP was often observed despite the addition of extra furin (Chapter 3), which may represent misfolded precursor protein. It has been suggested that putative disorder in the region of the polypeptide chain immediately N-terminal to the furin cleavage site assists in access of furin to the polypeptide chain (Chapter 5); if this is the case, and aggregation of MstnPP results in an altered structure in this area, furin cleavage may be inhibited.

MstnPP aggregation

Post-translational modification of MstnPP, perhaps through association with A β /A β PP, may lessen its degradation and traffic, resulting in MstnPP accumulation

and an increase in ER stress (Wojcik, Nogalska *et al.* 2007). One scenario is that A β PP/A β accumulation precedes that of MstnPP at the protein level, as A β PP overexpression does not increase MstnPP mRNA. Another suggestion is that MstnPP negatively influences A β and other proteins to cause their oligomerization (Wojcik, Engel *et al.* 2005).

Is it possible that MstnPP is acting independently of A β in sIBM? The results presented in Chapter 8 suggest that MstnPP misfolding and amyloid aggregation may have a negative effect on cell viability without an effect on, or stimulation from, A β . The milieu of different events that occur in amyloid disease suggest that no aspect is completely in isolation and each has an influence on the others. Protein misfolding is both promoted by and contributes to ER stress and the production of reactive oxygen species (ROS) that act negatively on both ER and mitochondrial function (Schroder 2008). The exact cause of sIBM is unknown although ER stress is a common theme underlying degenerative disease (Paschen 2003). A likely scenario, and the suggestion favoured here, is that ER stress causes the misfolding and amyloid-like aggregation of newly-translated and/or chaperone-folded myostatin. Cytotoxicity assays in cultured mouse myoblasts show that aggregated myostatin has a negative effect on cell viability (Fig. 8.12); therefore, misfolded myostatin may contribute to sIBM via a cytotoxic mechanism.

9.1.3 MstnPP cytotoxicity

The cytotoxicity exhibited by MstnPP aggregates and protofibrils in C2C12 cells supports a mechanism in which aggregation by MstnPP in sIBM contributes to muscle degeneration; this may occur independently of atrophy that may result from increased myostatin growth factor signalling. The *in vitro* assays investigate the effect of extracellular protein on cell viability. In sIBM tissue and cell culture models, MstnPP aggregates were localised to the cytoplasm, aggresomes and nuclear regions (Wojcik, Engel *et al.* 2005; Wojcik, Nogalska *et al.* 2007). However, an absence of extracellular MstnPP was not shown. Since the myostatin precursor protein is secreted (Anderson, Goldberg *et al.* 2008), post-secretion aggregation or secretion of aggregated/misfolded MstnPP may contribute to muscle fiber atrophy in sIBM. Extracellular cytotoxicity studied for other aggregated proteins involves mechanisms

such as alterations to endo- and exocytosis (Walsh, Hartley *et al.* 1999), alterations to Ca^{2+} homeostasis and the generation of reactive oxygen species (Bucciantini, Calloni *et al.* 2004). The presence of pore-like structures for MstnPP under TEM suggests that the cytotoxic mechanism may include membrane permeabilization (Lashuel 2005), the favoured hypothesis for amyloid oligomer cytotoxicity. A number of models place membrane permeabilization upstream of processes such as Ca^{2+} dyshomeostasis and ROS production (Glabe 2006).

Alternatively, MstnPP aggregates may exert toxic effects from an intracellular locality, with the observed cytotoxicity in cell culture due to internalisation of protein (Bucciantini, Calloni *et al.* 2004). Localisation in aggresomes suggests that the myostatin aggregates are targeted for degradation; if cell homeostasis is perturbed and degradation is not carried out, enhanced aggregation by MstnPP may add to cell stress and contribute to disease in this way. Although localisation to the ER was not shown, it was also not discounted; MstnPP aggregation within the ER may have a similar effect. The observation of MstnPP aggregates in the nucleus (Wojcik, Engel *et al.* 2005) is intriguing as a nuclear role for MstnPP or any TGF- β family member has not been shown to date. These results, coupled with the cytoplasmic staining shown for MstnPP (Wojcik, Nogalska *et al.* 2007), suggest that trafficking of the MstnPP protein may be altered in sIBM, a scenario that may be due to, or promote, MstnPP aggregation.

In any case, both extra- and intracellular cytotoxicity may occur via membrane permeabilization. The disruption of nuclear, mitochondrial, ER and Golgi membranes may occur in addition to that of the cell.

ER stress-induced upregulation of MstnPP

ER stress induces an upregulation of MstnPP at both the mRNA and protein levels, most likely via activation of NF- κ B (Nogalska, Wojcik *et al.* 2007). However, it is not clear why, in a situation where ER homeostasis is perturbed, a protein requiring oxidative folding and Ca^{2+} -regulated processing is upregulated.

As the goal of the UPR is to halt cell processes until homeostasis has been restored, an upregulation of MstnPP may occur to increase myostatin growth factor levels. Signalling by the myostatin growth factor may assist via upregulation of p21 and halting of the cell cycle (Thomas, Langley *et al.* 2000), locally inhibiting cell growth and/or development until homeostasis is restored. ER stress responses include both

protective and apoptotic mechanisms (Schroder 2008); an upregulation of myostatin is most likely protective. In a cultured sIBM cell model, NF- κ B activation of MstnPP and the ER stress markers Grp78 and Herp was not accompanied by pro-apoptotic cleavage of procaspase-3, suggesting that the protective pathway was active (Nogalska, Wojcik *et al.* 2007). As there is currently no evidence to support these suggestions, further investigation is required to understand the mechanisms involved here.

9.2 Model for amyloid formation

A number of both disease and non-disease proteins are capable of forming amyloid given the right conditions, and the general pathway (Fig. 1.9) appears to be conserved. However, the mechanisms behind the structural transitions responsible for the formation of amyloid protofibrils and fibrils remain unknown, and while the core structure of the amyloid fibril is inherently the same (Sunde, Serpell *et al.* 1997), the proteins from which they appear are structurally unrelated. Furthermore, the details of the amyloid formation pathway appear to differ between proteins.

9.2.1 Summary of results

The major results from Chapter 8 are summarised in Table 9.1. While the native dimer does not exhibit any amyloid characteristics, the soluble aggregates present after refolding appear to be intermediate species between the dimer and protofibrils. Mature fibrils and protofibrils have quite similar characteristics overall, consistent with the hypothesis that protofibrils are the direct precursors of fibrils and that fibril structure is determined by that of the protofibril (Chimon, Shaibat *et al.* 2007). The only notable differences are morphological and the reduced activity of mature fibrils in cell culture; however, this latter observation requires further investigation.

	Native Dimer	Soluble Prefibrillar Aggregates	Protofibrils	Mature fibrils
TEM morphology	Spherical	Elongated, bead-on-a-string, rings	Dense arrays and lateral association	Long, unbranching fibrils
ThT binding	None	Yes	Increased	Increased
Secondary structure	Mixture α -helix, β -sheet and random coil	Predominantly α -helix and significant disorder	β -sheet	β -sheet
Proteolytic resistance	No	No	Yes	Yes
β -ME/DTT resistance	No	Yes	Yes	Yes
SDS resistance	No	Yes, unless reduced	Yes	Yes
Biological activity	No	Yes	Yes	Reduced

Table 9.1 Summary of biochemical and biophysical properties of the MstnPP dimer, soluble aggregates, protofibrils and mature amyloid fibrils.

9.2.2 The role of non-native disulphides

MstnPP soluble aggregates can be reduced on SDS-PAGE with a high concentration of β -mercaptoethanol (Fig. 8.6) and treatment with standard conditions for the reduction of proteins in solution (10 mM DTT) is able to fully reduce the MstnPP dimer (Fig. 8.8a). However, this has a reduced effect for the soluble aggregates (Fig. 8.8 b.–e.). While TEM analysis indicates that some disruption of aggregation occurs and CD suggests an increase in α -helical structure, the soluble aggregates are still able to bind ThT. Protofibrils and mature fibrils show complete resistance to reduction by reducing SDS-PAGE. These results suggest that non-native disulphide-bonding may play a role in the formation and stabilization of prefibrillar MstnPP oligomers. Similar hypotheses have been proposed for a number of other amyloid-forming proteins such as lysozyme (Wang, Liu *et al.* 2009), insulin (Jimenez, Nettleton *et al.* 2002), the yeast prion protein Ure2p (Fei and Perrett 2009) and β_2 -microglobulin (Smith and Radford 2001).

The WT myostatin precursor protein sequence contains 13 cysteine residues (Fig. 9.1a, green). The 9 growth factor cysteines are all involved in disulphide-bonds in the isolated native structure (Cash, Rejon *et al.* 2009); little is known about the 4

cysteines of the propeptide. Although disulphide interactions between the propeptide and growth factor regions are unlikely in the native structure, in a situation where misfolding occurs, each cysteine residue may be able to interact with any other cysteine residue in the polypeptide chain. Some of these non-native disulphides may promote amyloid formation by bringing two regions of the chain in close enough proximity for β -aggregation to occur, a phenomenon documented previously for the yeast prion protein Ure2p (Fei and Perrett 2009). Putative examples of such disulphides are shown in Figure 9.1b. The regions predicted to have propensity for β -aggregation and amyloid formation are shown in blue (lower propensity) and red (higher propensity) (from Fig. 8.1). For example, cysteines 1 and 3, or 2 and 4, may be able to form a disulphide bond which enhances interactions between aromatic phenylalanine residues (yellow stars). Cysteines 3 and 4 are surrounded by phenylalanine and lysine residues, both implicated as key indicators of amyloid formation (Makin, Atkins *et al.* 2005).

a.

²¹VDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSKLRLE
 TAPNISKDVIRQLLPKAPPLRELIQYDVQRDDSSDGSLDEDDDYHATTETIITMPTES
 DFLMQVDGKPK**CCFFKFSSKIQYNKVVK**AQLWIYL**RPVETPTTVFVQILR**LIKPM
 MKDGTRYTGIRSLKLDNMNP GTGIWQSIDVKTVLQNWLKPESNLGIEIKALDENGH
 DLA VTFPGPGEDGLNP FLEV KVTDTPKRSRRDFGLDCDEHSTESR**CCRYPLTVDFE**
 AFGWDWIIAPKRYKANY**CSGECEFVFLQKYPHTHLVHQANPRGSAGP****CC**PTKMS
PINMLYFNGKEQIIYGKIPAMVVDRCGCS ₃₇₅

b.



Figure 9.1 The putative role of disulphide bonding in amyloid formation.

- MstnPP sequence from 21-375 (from Fig. 8.1).
- Schematic diagram showing examples of possible disulphide bonds between propeptide cysteines (black lines). Yellow stars represent aromatic residues within motifs suggested to be important for amyloid formation.

For a. and b., regions of predicted disorder are shown in blue (moderate propensity) and red (high propensity). Cysteine residues are indicated in green. The furin cleavage site is in black.

The hypothesis presented here is that non-native disulphide bonding promotes amyloid formation in the myostatin protein. Does protein misfolding occur prior to or because of these disulphides? In the case of the correctly folded dimer, partial unfolding and reduction would be necessary to allow new disulphides to form; alternatively, amyloid formation by the MstnPP dimer may occur independent of disulphide-bonding.

Partial reduction of the soluble prefibrillar aggregates results in a decrease in β -sheet structure and concomitant increase in α -helical structure, suggesting that disulphide-bonding promotes β -sheet formation and that an equilibrium may exist between the α -helical aggregates and β -sheet oligomers. Similarly, addition of TCEP to lysozyme inhibits amyloid fibrillation by prevention of the α -to- β transition (Wang, Liu *et al.* 2009). Reduced MstnPP aggregates can still bind thioflavin T, an intriguing result as ThT binding is usually associated with the formation of β -sheet structure. One explanation is that a proportion of the α -helical predominant oligomers contain an extended β -sheet core, a suggestion supported by the CD spectra of soluble aggregates both in the presence and absence of DTT (Fig. 8.8f).

TEM analysis shows partial DTT-induced disruption of elongation and association, suggesting that the formation of non-native disulphides may also occur intermolecularly, assisting in the growth of protofibrils both lengthwise and laterally.

9.2.3 Overall model

Together, the results presented here suggest that the pathway for amyloid formation by the myostatin precursor protein follows the general phenomena observed for other amyloid-forming proteins; the formation of misfolded intermediates rich in α -helical structures, which over time, in conditions of partial denaturation, form β -sheet rich protofibrils and mature fibrils. As suggested above, non-native disulphides may play a key role in the transition from the α -helical intermediate to β -sheet rich oligomers. Although a number of routes to amyloid formation are likely, shown by the different structures observed when alternative incubation conditions are used (Fig. 8.8 and 8.9), a simplified model for the distinct pathway of amyloid formation followed by the myostatin precursor protein is presented (Fig. 9.2).

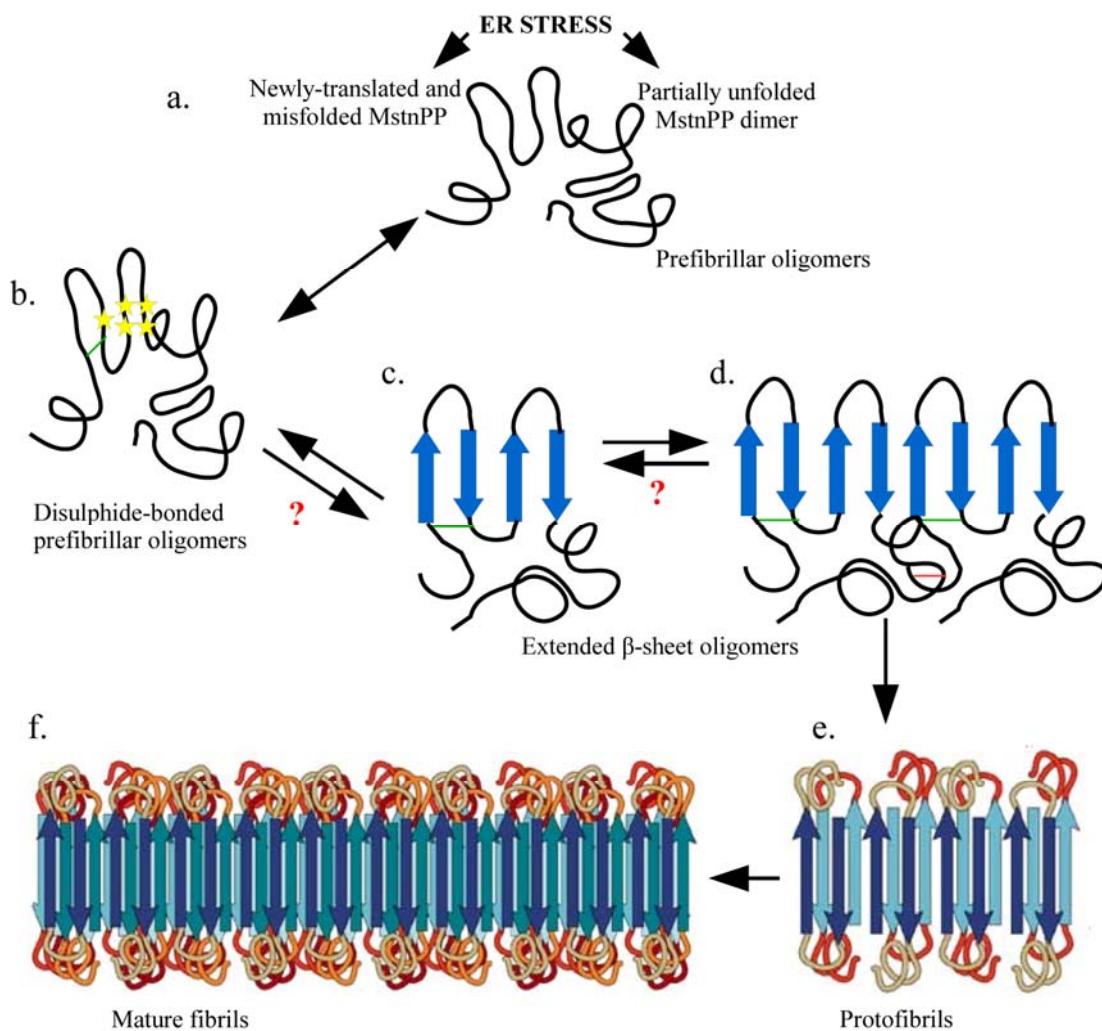


Figure 9.2 Model of the amyloid formation pathway for MstnPP.

ER stress promotes the upregulation, misfolding and accumulation of MstnPP (a). A number of different oligomeric forms are likely to be present. Non-native disulphide bonds (green) promote interactions between key aromatic residues (yellow stars) and the formation of prefibrillar oligomers (b), that may be stabilized by disulphide bonding to allow β-aggregation and the formation of extended β-sheet oligomers (c). Intermolecular disulphides (red) may promote elongation of β-sheet oligomers (d). Protofibril (e) and mature fibril (f) formation occurs via lateral association and strengthening of core interactions.

Red question marks indicate where equilibria may or may not be occurring.

ER stress is likely to be a key instigator of MstnPP misfolding. For example, consequences of ER stress such as inactivity of disulphide isomerases, an altered redox environment or the production of ROS, may result in newly-synthesised MstnPP that is incorrectly folded; alternatively, stress may cause the partial denaturation and subsequent misfolding of the native MstnPP dimer (Fig. 9.2a). These misfolded forms, rich in α -helical structures and random coil, may be highly flexible. Some oligomeric forms may also contain a proportion of β -sheet structure with the whole population quite dynamic in nature. Whether non-native disulphides (Fig. 9.2b, green) are responsible for misfolding or occur subsequently is not clear. As the whole process is likely to be dynamic, an equilibrium may exist between reduced and disulphide-containing oligomers ($a \leftrightarrow b$). Non-native disulphides may also isomerise, with a certain subset promoting the amyloid pathway. The observation that reduction promotes α -helical structures suggests that the formation of a flexible α -helix-rich protein occurs prior to disulphide-formation, which stabilises structures capable of β -aggregation. In addition, the precipitation of the majority of MstnPP when DTT is added at pH 5.3, 60 °C, suggests that disulphide bonding has a stabilising effect during amyloid formation, possibly through the shielding of hydrophobic regions.

It is likely that a number of structural intermediates exist between the soluble aggregates (Fig. 9.2a) and protofibrils (Fig. 9.2e); TEM images show a number of different morphological forms present in the aggregate solution (Fig. 8.2 and 8.3). These may represent different oligomeric intermediates, one of which may be the precursor of the β -sheet oligomer (Fig. 9.2b). At this stage, β -aggregation may be predominantly in the extended β -sheet form rather than the cross- β structure proposed for protofibrils and fibrils, as ThT binding, associated predominantly with cross- β structure (Walsh, Yau *et al.* 2009), increases on protofibril formation. Extended β -sheet oligomers (Fig. 9.2c) may then interact to form elongated structures (Fig. 9.2d) and intermolecular disulphide bonding (red) may contribute to this. Lateral association and further interactions within the core of the structure lead to the formation of protofibrils (Fig. 9.2e) and mature fibrils (Fig. 9.2f). At this stage, even high concentrations of reducing agent have no effect on the integrity of the amyloid structures, implying that the disulphides involved in β -sheet aggregation are located in the core of the structure, or that the interactions that have formed no longer require stabilisation by disulphide bonding.

At which stage is the point of no return? Equilibria are likely to exist between early intermediate species, especially if disulphide bonding is involved. However, the inability to reduce MstnPP protofibrils, even at high concentrations of β ME (Fig. 8.7) suggests that the formation of protofibrils may be irreversible, possibly due to formation of the amyloid cross- β core structure.

The kinetics of MstnPP are likely to follow a nucleation-dependent pathway. Although kinetic studies were not carried out for this thesis, a significant lag-time for fibril formation, dependent on protein concentration, was observed. After overnight incubation, a β -sheet rich secondary structure is apparent by CD; however, the formation of mature fibrils occurs over at least a week. The β -sheet rich secondary structure may represent either extended β -sheet or the characteristic cross- β structure of protofibrils and mature fibrils; the lag phase is most likely to involve formation of the β -aggregated species that directly precedes the latter. The correctly folded MstnPP dimer appeared to have faster kinetics for fibril formation than the misfolded aggregates, suggesting that structure has a significant role in amyloid formation. For example, the soluble aggregates may require comparatively more unfolding than the dimer to allow amyloid formation and/or amyloidogenic regions may be more accessible in the dimer.

9.3 Implications of the model for disease

The most cytotoxic species of MstnPP amyloid aggregation appears after incubation in amyloid-promoting conditions and is therefore either the β -sheet containing oligomer or the protofibril, or both. The mature fibril may have reduced or no cytotoxicity. These results are consistent with those observed for other amyloid-forming proteins and support the hypothesis that targeting prefibrillar oligomeric species is most promising in disease intervention (Bucciantini, Giannoni *et al.* 2002; Stefani and Dobson 2003). If amyloid aggregation by the myostatin precursor protein plays a significant role in the pathogenesis of sIBM, the results presented here suggest that inhibition of non-native disulphide formation will be an effective method for preventing MstnPP β -aggregation and amyloid formation, as suggested for lysozyme, transthyretin and β_2 -microglobulin (Yamamoto, Yagi *et al.* 2008; Wang, Liu *et al.*

2009). However, MstnPP relies on a number of disulphides for activity and distinguishing between native and non-native interactions would be paramount.

Summary

The myostatin precursor protein is able to form cytotoxic amyloid aggregates and mature fibrils *in vitro*, most likely by a nucleation-dependent mechanism enhanced by increasing protein concentration. This finding has important implications for the suggestion that MstnPP has a role in the amyloid disease sporadic inclusion body myositis. ER-stress effects such as an altered redox environment, coupled with a stress-induced upregulation of MstnPP, may lead to accumulation of misfolded MstnPP species and cytotoxic amyloid aggregation *in vivo*.

10 Summary and Conclusions

This PhD thesis is focused predominantly around the structure and function of the human myostatin precursor protein and the implications that these findings have for disease. The major discussion points and conclusions are summarised below.

The native human myostatin precursor protein can be refolded *in vitro* after insoluble expression in *E.coli*

Myostatin is a secreted protein, with a structure that contains at least four disulphide bonds. Production of the native growth factor in isolation has only been successful in mammalian culture. However, the porcine and zebrafish precursor proteins have been produced within insoluble bacterial inclusion bodies, and a similar protocol to the latter can be used for the human myostatin precursor protein.

Refolding success was shown by a number of different means. First, the disulphide-linked MstnPP dimer is observed by non-reducing SDS-PAGE after refolding. Second, *in vitro* furin cleavage of MstnPP yields the latent complex, consisting of propeptide and growth factor dimer domains.

Third, the complex can be dissociated by acid or heat treatment and addition of the dissociated complex to the media of C2C12 mouse myoblasts results in the inhibition of proliferation, showing that the myostatin growth factor dimer is active and likely to be correctly folded.

Fourth, analysis of MstnPP secondary structure shows a mixture of α -helix, β -sheet and random coil, consistent with propeptide secondary structure predictions, the published myostatin growth factor crystal structure (Cash, Rejon *et al.* 2009) and the CD spectra of other TGF- β family propeptide and precursor proteins (McMahon, Dignam *et al.* 1996; Sengle, Ono *et al.* 2008).

Fifth, refolded MstnPP is highly stable. The protein shows elevated thermal stability, consistent with the reported high thermal stability of the TGF- β superfamily (Brownh, Wakefiel *et al.* 1990), and proteolytic digestion indicates two highly stable regions, the cysteine knot and the N-terminus of the propeptide domain.

Sixth, the ability to grow crystals of MstnPP supports the formation of a structured protein.

The human myostatin precursor protein may contain functional intrinsic disorder

At least one region in the myostatin propeptide domain may contain intrinsic disorder, which may be functional, assisting in the proposed functions of the propeptide domain. Intrinsic disorder in MstnPP is indicated primarily by aberrant electrophoretic mobility on reducing SDS-PAGE, a high apparent molecular mass by gel filtration and the possibility of β -aggregation at elevated temperatures.

In silico algorithms predict two regions of intrinsic disorder, one in the inhibitory N-terminal domain and the other at the very C-terminus of the propeptide. Limited trypsin proteolysis followed by Western analysis suggests that the C-terminus is less proteolytically stable and therefore more likely to contain intrinsic disorder.

The myostatin propeptide has been suggested to have a chaperone-like function in folding of the myostatin growth factor domain. Intrinsic disorder may assist in this function by imparting flexibility and allowing the formation and possible interchange of numerous non-covalent interactions, or by shielding the hydrophobic growth factor to prevent aggregation during folding, through interactions with the ER environment.

Formation of the human myostatin latent complex may involve structural rearrangement

The human myostatin latent complex exhibits a number of structural and biochemical characteristics that are in contrast to those of the precursor protein, suggesting a structural rearrangement on complex formation. This is consistent with the hypothesis presented above, as regions of intrinsic disorder often undergo induced structuring on partner binding (Iakoucheva, Brown *et al.* 2002), as well as results shown for the TGF- β 1 latent complex (McMahon, Dignam *et al.* 1996).

First, CD spectroscopy shows an increase in intensity at both α -helical and β -sheet minima for the latent complex. Second, CD thermal denaturation behaviour is altered in comparison to the precursor, especially in the α -helical region, which is likely to represent the propeptide domain predominantly. Third, the complex has an increased retention time by gel filtration suggesting the formation of a more compact structure. These results have implications for both stabilisation and inhibitory functions of the propeptide within the latent complex. Structuring of intrinsically disordered regions may provide increased structural stability for the export of myostatin, as well as an

increase in growth factor affinity via non-covalent interactions and a greater surface area for binding, leading to enhanced inhibition of growth factor receptor-binding.

The structure of the human myostatin latent complex

What might the structure of the human myostatin latent complex look like? Modelling of the complex using the published follistatin/myostatin complex structure and suggested interaction domains between the inhibin/activin propeptide and growth factor regions allows a preliminary model to be hypothesised. In this model, two propeptide molecules wrap around one growth factor dimer to block both type I and type II receptor binding domains. The N-terminal α -helix of the propeptide interacts with the fingers of the myostatin growth factor via conserved hydrophobic domains in both structures to inhibit type II receptor-binding. The C-terminal domain of the propeptide blocks the type I receptor-binding domain through interaction via a second α -helix, in a mechanism similar to that of follistatin's N-terminal domain helix.

The Piedmontese mutant myostatin has reduced dimerisation capacity

The Piedmontese myostatin growth factor contains a mutation of cysteine to tyrosine at position 313, which disrupts the formation of one of the disulphide bonds involved in the cysteine knot motif. Previous *in vivo* studies showed a possible reduction in processing and stability, yet the mutation did not appear to prevent growth factor secretion or circulation (Berry, Thomas *et al.* 2002). What is the mechanism behind the inability of C313Y mutant myostatin to elicit signalling?

Biochemical and biophysical analyses show that the mutant precursor protein has reduced dimerisation capacity. While the C313Y precursor has a retention time characteristic of a dimer by gel filtration, non-reducing SDS-PAGE results indicate that the protein is predominantly monomeric, a phenomenon continued after furin cleavage and latent complex formation that suggests non-covalent dimerisation. Modelling of the C313Y mutation in the myostatin crystal structure shows a close proximity to the intermolecular disulphide bond and a strong likelihood of structural perturbations in this area, factors which may have an inhibitory effect on covalent dimerisation. If the intermolecular disulphide is necessary to maintain a structure competent to bind its receptors, an inability to dimerise covalently may be one reason for lack of activity of the Piedmontese myostatin growth factor.

The equivalent disulphide mutant has been constructed in the cysteine-knot growth factor VEGF and exhibits a more compact structure, with a diameter about 2 Å less than the WT and tighter packing in the dimerisation interface (Muller and Heiring 2002). The inability of C313Y myostatin to dimerise covalently may result in a more open, or a more compact structure, the latter supported by an increase in gel filtration retention time.

The Piedmontese mutant myostatin may have structural perturbations and increased flexibility in the Type I receptor-binding domain

The cysteine knot, disrupted by the C313Y mutation, is regarded as a key determinant of stability in the TGF- β family (Muller and Heiring 2002). Sypro Orange analysis and CD thermal denaturation reveal that the Piedmontese mutant has decreased thermal stability compared to the WT protein, and limited trypsin proteolysis suggests an increase in structural flexibility at physiological temperature that likely translates to a decrease in structural stability. Structural modelling shows that residue 313 is located on the concave surface of the myostatin growth factor, in close proximity to the type I receptor binding domain and is contained within a loop segment of the cysteine knot structure. Therefore, disruption of the cysteine knot structural motif coupled with structural rearrangement caused by introduction of the tyrosine side-chain, and/or an increase in flexibility of the loop region containing residue 313, may inhibit binding of the type I receptor, providing a second explanation for the inability of Piedmontese myostatin to inhibit muscle growth *in vivo*. Analysis of the C313Y latent complex by CD spectroscopy and thermal denaturation suggests that propeptide/growth factor interactions within the complex may also be compromised. If this is the case, WT interactions may involve the type I receptor binding site, supporting the above hypothesis. The equivalent VEGF disulphide mutant shows global structural changes, in particular, rearrangement in the type II receptor binding site (Muller and Heiring 2002). If similar changes occur in the Piedmontese myostatin structure, the C313Y mutation may also inhibit type II receptor binding.

The myostatin precursor protein can form cytotoxic amyloid aggregates and fibrils *in vitro* with implications for sporadic inclusion body myositis

A role for the myostatin precursor protein in the amyloid disease sporadic inclusion body myositis (sIBM) has been suggested (Askanas and Engel 2008). After refolding,

a population of SDS-resistant soluble aggregates is present in addition to the native MstnPP dimer. This species is likely to represent a prefibrillar oligomer on the amyloid formation pathway due to its morphology by TEM, the ability to bind ThT, and an α -helical predominant secondary structure. Incubation in the partially denaturing conditions of pH 5.3 and 60 °C causes an overnight transition to a protofibrillar species and after one week, mature fibrils, with β -sheet rich secondary structures and enhanced ThT fluorescence.

Amyloid formation by MstnPP exhibits the characteristics of a nucleation-dependent pathway, with a prominent lag-phase that is shortened with increasing protein concentration. The correctly-folded MstnPP dimer has a shorter lag-phase than the aggregated form, suggesting that protein structure is important for nucleation. Within the lag-phase, a number of oligomeric intermediates of differing morphology and size can be observed by TEM, including ‘bead-on-a-string’, spherical and ring-shaped species. It is proposed that one of these α -helical predominant intermediates is viable for β -aggregation and amyloid core formation. Subsequent growth and lateral association of β -aggregated intermediates leads to protofibril formation and eventually mature amyloid fibrils.

As suggested for a number of other proteins (Fei and Perrett 2009; Wang, Liu *et al.* 2009), non-native disulphide bonds may contribute to amyloid formation by MstnPP; MstnPP amyloid species show varying resistance to reducing agents. Low concentrations of DTT cause partial dissociation of elongated prefibrillar oligomers and an increase in α -helix, suggesting that disulphide bonding within the α -helical oligomers promotes β -sheet aggregation, and between molecules, may contribute to oligomer elongation and lateral association.

The amyloid aggregates of human MstnPP are cytotoxic *in vitro*. Oligomeric MstnPP species affect the viability of C2C12 mouse myoblasts when added to the culture medium, and TEM analysis shows that these species exhibit a range of morphologies, including ring-shaped molecules that have been implicated in membrane permeabilization. A stressed environment may promote the formation of toxic MstnPP amyloid aggregates *in vivo*, contributing to the pathogenesis of sIBM. It is hypothesised that ER stress either prevents the correct folding of newly-translated MstnPP, or partially unfolds the native MstnPP dimer, or both.

Overall summary and conclusions: the big picture

As a potent inhibitor of muscle growth, the myostatin growth factor is a promising therapeutic target in the treatment of muscle wastage diseases. While research surrounding the structure, function and regulation of the isolated growth factor is steady, the biochemistry of the myostatin precursor protein and latent complex is poorly understood. As key players in the processing and regulation of the myostatin growth factor, details of the structure and function of these species will provide important insight into the roles and regulation of myostatin in disease.

The research and discussion presented here contribute such information. *In vitro* biochemical and biophysical analyses of the refolded myostatin precursor and latent complex show that the propeptide domain is likely to contain a significant degree of structure, as well as regions of intrinsic disorder, which may enable its diverse roles to be carried out. A number of growth factor/propeptide interactions may exist within the precursor protein, and conformational rearrangements of the propeptide domain on latent complex formation may both increase the surface area available for growth factor binding and enhance existing interactions for inhibition.

Comparative analysis between WT and C313Y mutant myostatin precursor proteins and latent complexes suggests a number of mechanisms that may underlie the inability of the C313Y growth factor to elicit signalling *in vivo*. These include an inhibition of covalent dimerisation, disruption of the cysteine knot motif, enhanced flexibility and structural perturbations, all of which are expected to significantly alter the structure of the concave surface of the growth factor where the type I receptor is likely to bind.

MstnPP has the ability to misfold and form cytotoxic amyloid aggregates and mature fibrils in partially denaturing conditions. Importantly, protofibrils, the direct precursor to the amyloid fibril, are able to form at physiological temperature. As MstnPP has been shown to be upregulated in the amyloid disease sporadic inclusion body myositis, these results have implications for amyloid formation by the myostatin precursor protein *in vivo*.

Overall, this thesis presents novel results for the structure and function of both folded and misfolded forms of the WT myostatin precursor protein, suggesting that both the myostatin growth factor and precursor protein may be important targets in diseases of muscle wastage.

11 Future work

This PhD thesis has produced a number of novel findings with implications for the structure and function of the human myostatin precursor protein and latent complex, Piedmontese mutant myostatin and the role of the myostatin precursor protein in the amyloid disease sporadic inclusion body myositis. These results set the scene for further investigation in all areas. This chapter presents a discussion of the main research foci that can follow on from this thesis.

11.1 Structure and function of MstnPP and latent complex

11.1.1 Structural analysis

Crystallisation

Key targets for crystallisation are the myostatin precursor protein, the propeptide domain in isolation and the latent complex. Preliminary crystals of the MstnPP dimer have been obtained in a number of different conditions, but these are small and do not diffract. An initial conclusion from these results was that disorder in the propeptide domain prevents growth of an ordered crystal lattice. However, subsequent biophysical and *in silico* analyses suggest that while the propeptide may have regions of intrinsic disorder, inherent structure also exists, indicating that a partial structure may be obtainable. Additional screening and optimisation will include techniques such as limited proteolysis to remove regions of disorder, alternative crystallisation and additive screens, and seeding using the microcrystals previously obtained. In addition, alternative constructs containing amino acid substitutions that alter the regions predicted to be the most disordered, may be more amenable to structure determination.

Although the precursor protein is expressed in an insoluble form in *E.coli*, previous studies indicate soluble expression of the propeptide domain alone (Jiang, Liang *et al.* 2004). As the only structural information for the propeptide published to date is presented in this thesis, and the propeptide is a key inhibitor of the myostatin growth factor *in vivo*, the three-dimensional structure of the propeptide is an important and attractive target. Structure determination will likely require a method other than

molecular replacement due to an absence of published related structures. This work is currently underway.

The latent complex plays an important role in regulation of the myostatin growth factor; a high-resolution structure of the latent complex will therefore provide valuable information for the mechanism of inhibition by the propeptide. Purification of the latent complex can be achieved by gel filtration following furin digest of the precursor protein and structure determination can make use of the recently determined myostatin/follistatin complex structure (Cash, Rejon *et al.* 2009).

Other biophysical analyses

Other biophysical techniques can be performed to obtain structural information about the myostatin precursor protein, propeptide and latent complex. Small angle x-ray scattering (SAXS) analysis can be used to obtain information about the shape and size of these species in solution. Synchrotron SAXS radiation, available at the Australian Synchrotron, Melbourne, Australia, produces precise scattering patterns that could be used in combination with electron microscopy and high-resolution X-ray diffraction data, such as the published myostatin growth factor crystal structure (Cash, Rejon *et al.* 2009), to gain relatively high-resolution information. The spatial range of SAXS is 1-200 nm.

NMR is used to obtain information about the solution structure of a protein (Wuthrich 1990) and is a further possibility for the propeptide region due to its small size (less than 30 kDa). The propeptide domain is predicted to contain regions of intrinsic disorder which may interfere with crystal growth to preclude structural determination by X-ray crystallography. As NMR is performed on a protein in solution, it is not limited to crystallisable proteins. Furthermore, NMR can provide information about the dynamics and flexibility of a protein structure (Ishima and Torchia 2000; Berjanskii and Wishart 2006); it is therefore an ideal tool for investigation into these features of the myostatin propeptide domain.

Structural changes on complex formation

It has been hypothesised that structural rearrangement within the propeptide domain occurs on formation of the myostatin latent complex and that this may involve induced folding of intrinsically disordered regions. A number of techniques can be used to study this in more detail (Receveur-Brechot, Bourhis *et al.* 2006). ANS (1-

anilino-8-naphthalenesulfonic acid) can access hydrophobic cavities and strongly fluoresces when bound to a non-polar surface; therefore, changes in binding pre- and post-furin digest may indicate changes in the structural environment of the compound. The myostatin propeptide is expected to be highly amenable to this analysis, as both the N- and C-terminal regions of predicted interaction and disorder contain a number of tyrosine and tryptophan residues.

Fluorescence resonance energy transfer (FRET) is a technique that allows specific regions of the protein, expected to undergo conformational change, to be followed. Fluorescent donor and acceptor chromophores are used to label two regions of the polypeptide chain; if these are in close enough proximity (less than 10 nm) energy will be transferred from donor to acceptor, observed as fluorescence. Structural changes that separate the two regions will therefore prevent this transfer and fluorescence will not be observed.

In addition, limited proteolysis of the latent complex can be compared to that of the precursor protein. Subsequent analysis of the resulting peptides by mass spectrometry is expected to identify differences in exposed regions between the molecules.

Intrinsic disorder

The hypothesis that MstnPP contains regions of intrinsic disorder has been presented. A number of methods for investigation of intrinsic disorder in proteins have been described (Reeveur-Brechot, Bourhis *et al.* 2006) and some of these methods may be suitable for MstnPP. For example, an assessment of the global tertiary structure can be made using limited proteolysis followed by mass spectrometry, as limited proteolytic sites are typically found at flexible loop regions. Protein dimensions and shape can be studied using SAXS and dynamic light scattering. The local tertiary structure can be investigated using techniques such as dynamic quenching of fluorescence. This allows the microenvironment of tryptophan and tyrosine residues to be probed, by measuring fluorescence intensity after the addition of compounds such as iodide or succinimide. Fluorescence quenching occurs when these molecules are in physical contact with an excited chromophore, leading to the sharing of excitation energy and a reduction in fluorescence intensity (Reeveur-Brechot, Bourhis *et al.* 2006).

11.1.2 Stability

CD thermal denaturation of the precursor protein and latent complex produced ambiguous results, with spectral shifts to β -sheet rich minima suggesting β -aggregation at high temperatures, observation of the highly stable cysteine knot core or dissociation of the latent complex when this was studied. Recent reports suggest using turbidity or 90° light scattering measurements simultaneously with CD thermal denaturation (Benjwal, Verma *et al.* 2006). These techniques allow aggregation during thermal unfolding to be monitored in real time and will therefore help to determine if the transitions observed during CD thermal denaturation of the myostatin precursor protein and/or the latent complex are representative of aggregation.

To gain further information about the behaviour of MstnPP and the latent complex at high temperatures, fluorescence may be a more appropriate technique. The myostatin precursor protein primary sequence contains 6 tryptophans that, with contributions from 11 tyrosine residues, are expected to produce a strong emission spectrum. In addition, urea and/or guanidine-HCl induced denaturation, monitored by CD and fluorescence can be used as an alternative to temperature, to gain information about myostatin stability characteristics.

11.1.3 Functional analysis

The role of propeptide cysteine residues

The myostatin propeptide domain contains four cysteine residues (Cys39, Cys42, Cys137 and Cys138). All cysteines of the isolated growth factor are involved in either intra- or intermolecular disulphides (Cash, Rejon *et al.* 2009); although the nature of growth factor disulphide bonding within the latent complex is unknown, evidence suggests covalent association between the propeptide and growth factor is unlikely (Thies, Chen *et al.* 2001; Jiang, Liang *et al.* 2004).

Disulphide bonds may form within the propeptide region to influence the structure of this domain, or between the propeptides of different precursor monomers. The TGF- β 1 propeptide region contains three cysteine residues. While Cys33 forms a disulphide bond with the TGF- β 1 growth factor, Cys223 and Cys225 form intermolecular disulphides that are necessary for latent complex formation and growth factor inhibition (Brunner, Marquardt *et al.* 1989).

An investigation into the presence of additional propeptide disulphide bonds can be started with measurement of the total number of free and bonded cysteines in the precursor protein using 4,4'-dithiodipyridine (DTDP) (Reiner, Kada *et al.* 2002). In brief, DTDP reacts with free sulphydryl groups to form a product that can be measured spectrophotometrically at 324 nm. Absorbance is proportional to the number of cysteine residues, which can be calculated mathematically (Walsh, Yau *et al.* 2009). Performing this reaction in both reducing and non-reducing conditions is expected to give a measure of the number of MstnPP cysteines involved in disulphide bonds in the native structure. However, this analysis will not discriminate between intra- and intermolecular bonds. To address this, subsequent MS/MS mass spectrometry in non-reducing conditions can be used to analyse the MstnPP peptides joined by disulphide bonds.

Site-directed mutagenesis of propeptide region cysteine residues may provide insight into the role of cysteine residues at different stages of myostatin processing and regulation. For example, analysis of the folding of MstnPP cysteine mutants using the refolding and purification protocols presented in this thesis will give an indication of the role each cysteine plays in the proposed chaperone activity of the propeptide. Initially, cysteine to alanine substitutions are logical.

Disulphide bonds may also play a role in propeptide structure and stability. The denaturation of cysteine mutants, monitored by CD thermal and chemical-denaturation, Sypro Orange binding and fluorescence, can be compared to that of the WT species.

Mapping of the propeptide/growth factor binding domains

Mutagenesis of inhibin/activin propeptide and growth factor regions suggests a conserved hydrophobic interaction motif within the TGF- β family; for myostatin, this is between the N-terminal helix of the myostatin propeptide and the ‘finger region’ of the growth factor in the vicinity of the type II receptor-binding domain. To investigate the proposed interaction, mutagenesis of these regions in the myostatin growth factor and propeptide can be performed, followed by methods such as gel filtration or isothermal calorimetry. Binding studies can be coupled with effects on processing and secretion, for example, through expression of myostatin mutants in mammalian cells such as CHO, and subsequent analysis via Western blotting and the C2C12 proliferation assay.

A similar approach can be used to study putative interacting domains in other regions of the myostatin propeptide and growth factor. For example, the C-terminal propeptide helix of the propeptide may be involved in interactions with the type I receptor binding domain of the growth factor.

A role for calcium?

Primary structure analysis predicted the presence of an EF-hand Ca^{2+} -binding domain in the myostatin propeptide domain. Although Ca^{2+} -binding has not been documented for other members of the TGF- β family, myostatin is folded in the ER and is expressed in the muscle, localizations that depend on calcium for homeostasis. An investigation into the effects of calcium on the structure and function of the myostatin precursor protein, propeptide and latent complex may therefore provide interesting insights into myostatin regulation. Examples of key experiments are the addition of calcium to the refolding and purification processes, as well as CD, CD thermal denaturation, and proteolytic analyses. The removal of calcium should also be addressed, via addition of EDTA.

11.2 Structure of C313Y and implications for function

11.2.1 Effects of the C313Y mutation

Dimerisation

In vitro, the C313Y mutation affects the ability of Piedmontese myostatin to dimerise. The mechanism behind this may include structural perturbation of the dimerisation interface due to presence of tyrosine at this position. A key experiment is mutation of cysteine 374, the other member of the C313-C374 disulphide bond, while leaving C313 in the native form, followed by a study of the behaviour of this construct during refolding, structural analysis and denaturation. If the presence of tyrosine at position 313 is primarily responsible for reduced dimerisation, replacement of C374 with alanine should not prevent dimerisation. Production of a myostatin construct with alanine at position 313 instead of tyrosine would also probe this hypothesis. Although the smaller residue would not be expected to have such dramatic effects on protein structure, the equivalent VEGF disulphide mutant contained C to A substitutions and showed dramatic structural alterations (Muller and Heiring 2002).

Receptor binding

A major hypothesis presented is that type I receptor-binding is inhibited for Piedmontese myostatin due to the close proximity of the C313Y mutation to the type I receptor-binding site; structural alterations shown for the equivalent VEGF disulphide mutant (Muller and Heiring 2002) suggest that type II receptor binding may also be affected. The binding of the Piedmontese protein to its receptors can be studied using techniques such as isothermal calorimetry and surface plasmon resonance; however, these investigations rely on the purification of sufficient quantities of mutant protein. Similar studies should also be performed using the mutants described above.

Thermodynamic stability

The analysis of VEGF disulphide bond mutants produced an unexpected result; while thermal and structural stability was decreased by comparison to the WT protein, increases in thermodynamic stability were observed (Muller and Heiring 2002). Although the removal of disulphide bonds is expected to increase conformational entropy, this may provide one explanation for the production of the inactive Piedmontese protein *in vivo*. The thermodynamic stability of C313Y myostatin needs to be analysed for comparison to both WT protein and VEGF mutant data (Muller and Heiring 2002). This can be carried out via calculations of the standard enthalpies, entropies and free energies of the C313Y protein using urea, guanidine-HCl and/or pH-induced denaturation at 25 °C.

11.2.2 Crystallisation of C313Y

X-ray crystallography and structure determination of the C313Y growth factor, precursor protein or latent complex would provide invaluable insight into the structural changes caused by the Piedmontese mutation. Furthermore, as the latent complex is known to form, and formation of the WT latent complex may involve interactions in the type I and II receptor-binding sites, comparison of these structures may provide information about receptor-binding by C313Y.

11.3 Amyloid

11.3.1 MstnPP amyloid structure

The structure of MstnPP amyloid oligomers, protofibrils and mature fibrils was studied primarily using TEM, ThT binding and CD spectroscopy for this thesis. However, a number of additional techniques provide higher-resolution information about amyloid structure. X-ray fibre diffraction is used routinely to gain information about the distance between strands in the β -sheets of the amyloid core structure (Sunde, Serpell *et al.* 1997). SAXS can provide relatively high-resolution information about the shape and repeating units of amyloid species (Rambaran and Serpell 2008) especially when used in conjunction with techniques such as cryo-electron microscopy. In addition, both solution and solid-state NMR have been used for the structural analysis of amyloid, with the latter providing the ability to work on solid samples (Rambaran and Serpell 2008).

11.3.2 The MstnPP amyloid formation pathway

Which regions of MstnPP are involved in amyloid formation?

The myostatin precursor protein is large by comparison to other amyloid forming peptides; however, there are only two regions predicted to be prone to β -aggregation and amyloid formation within the primary sequence. Initially it will be important to ascertain which domain of MstnPP, propeptide or growth factor, or both, is responsible for amyloid aggregation. This can be done by studying amyloid formation of isolated domains in the partially denaturing conditions used previously, by TEM, CD and ThT binding. Subsequently, the minimal region of MstnPP that forms amyloid can be probed using different MstnPP constructs, such as peptides 139-177, 314-321 and 347-355. Site-directed mutagenesis can then be used to determine the residues that play major roles in the aggregation and amyloid-formation processes.

The kinetics of MstnPP amyloid formation

MstnPP amyloid formation has been suggested to follow a nucleation-dependent pathway that is affected by environmental conditions. Furthermore, the correctly folded MstnPP dimer appears to form amyloid at a faster rate than aggregated protein and the C313Y mutant may have an increased propensity for β -aggregation at high

temperatures. These observations suggest that protein structure has an effect on the kinetics of amyloid formation by MstnPP. A detailed analysis of the kinetics of MstnPP amyloid formation can be carried out using dynamic light scattering for changes in oligomer size, CD for secondary structure transitions, and ThT and Congo Red binding for the onset of cross- β structure (Kumar, Mohanty *et al.* 2007; Mauro, Craparo *et al.* 2007).

The effects of disulphide-bonding on MstnPP amyloid formation

Results presented in this thesis suggest that disulphide bonding is involved in the amyloid formation process of MstnPP. In order to study this in further detail, MstnPP amyloid formation can be studied in the presence and absence of reducing agent added at different stages. TCEP is recommended for use here rather than DTT or β ME as it exhibits higher stability and stronger reducing activity at low concentrations, and has been used for the study of disulphide-dependent amyloid formation by lysozyme (Wang, Liu *et al.* 2009). Similar studies on MstnPP constructs carrying mutations in one or more cysteine residues will provide insight into the specific cysteine residues involved in this phenomenon.

11.3.3 Implications for sporadic inclusion body myositis (sIBM)

Studies in sIBM tissue

The results presented in this thesis are *in vitro*. While these provide an important initial glimpse into amyloid formation by MstnPP and imply that a similar process may occur *in vivo*, it has been suggested that while most proteins are capable of amyloid-formation given the right conditions *in vitro*, this does not necessarily translate into a disease-scenario (Stefani and Dobson 2003).

To study amyloid formation in a more *in vivo* context, a mammalian cell model of sporadic inclusion body myositis can be produced (Nogalska, Wojcik *et al.* 2007; Wojcik, Nogalska *et al.* 2007). Mouse C2C12 myoblasts are available for use as a starting culture with induction of differentiation into muscle fibres followed by the addition of chemicals that induce ER stress, such as tunicamycin and thapsigargin (Nogalska, Wojcik *et al.* 2007). This is expected to produce a mouse muscle fibre model of sIBM. The nature of MstnPP in this model can then be investigated.

The study of MstnPP aggregates from sIBM patient tissue will provide information about the state of MstnPP amyloid aggregation in a disease-scenario; tissue has been sourced and experimental design is underway.

The caveat of both investigations lies in the separation of myostatin amyloid forms from those of amyloid- β and other proteins known to form amyloid in sIBM. Anti-myostatin antibodies may be useful in pulling down MstnPP from diseased tissue and cells; however, associated aggregates of other proteins may also be retrieved. The commercially available anti-myostatin antibody has been shown to recognise prefibrillar forms of MstnPP amyloid but the possibility that it will not be effective for amyloid fibrils, due to core interactions involving the antigenic region, cannot be discounted. However, it will be possible to isolate oligomeric forms of MstnPP using this method.

The mechanism of MstnPP amyloid cytotoxicity

Addition of MstnPP aggregates and protofibrils to the media of cultured C2C12 mouse myoblasts has a cytotoxic effect. However, the mechanism for this cytotoxicity is unknown. The aggregated protein may be internalised or may exert its toxicity at an extracellular level; membrane permeabilisation is a possibility in both situations. The aggregated protein may also have an indirect effect through the generation of free radicals such as hydrogen peroxide (Bucciantini, Calloni *et al.* 2004).

Internalisation of MstnPP can be studied using fluorescently-labelled protein by confocal microscopy. For example, the effect of aggregates on the permeabilisation of model membranes can be investigated using the self-quenching fluorescent dye calcein (Rabzelj, Viero *et al.* 2008; Walsh, Yau *et al.* 2009).

The mechanism of cell death associated with MstnPP aggregates can be studied by addressing the apoptotic or necrotic status of cells through assessing aspects such as caspase activation, depolarisation of mitochondria and phosphatidylserine externalisation. Caspase activation can be monitored using caspase substrates (Walsh, Cullen *et al.* 2008) such as Ac-DEVD-AFC (Roche), which generates a fluorescent product, AFC, upon cleavage by caspase-3 at the DEVD recognition site. Depolarisation of mitochondrial membrane potential can be investigated by measuring cellular uptake of the fluorescent cationic dye, JC-1 (Sathishkumar, Gao *et al.* 2010) and its segregation into the mitochondrial fraction. Phosphatidylserine externalisation is one of the earliest changes in apoptosis; phosphatidylserine usually

resides in the inner layer of phospholipids (Sathasivam and Shaw 2005) and externalisation can be measured using Annexin V-488, which produces a fluorescent signal on binding to phosphatidylserine.

Intracellular ROS production can be monitored using confocal microscopy and the ROS-sensitive fluorescent probe CM-H2 (Bucciantini, Calloni *et al.* 2004), which fluoresces when oxidised. The measurement of A β -generated H₂O₂ has been carried out previously using electron spin resonance (ESR) spectroscopy in conjunction with a spin-trapping technique (Tabner, El-Agnaf *et al.* 2005; Allsop, Mayes *et al.* 2008). Any hydrogen peroxide formed can be converted to hydroxyl radicals upon addition of Fe(II), which are then trapped by DMPO (5-dimethyl-1-pyrroline N-oxide) to form a hydroxyl radical adduct which has a characteristic ESR spectrum. The use of this technique is likely to require collaboration.

Summary

The study of the structure and function of both native and misfolded myostatin precursor protein, the latent complex and the C313Y Piedmontese mutant, contributes key information to understanding the biochemistry of myostatin and its role in disease. The novel results presented in this thesis contribute important information in these areas of human health and disease, and set the scene for ongoing investigations into a number of different aspects of myostatin structure and function.

12 Appendices

12.1 Appendix 1: The human myostatin growth factor

The initial objective of this PhD was determination of the X-ray crystal structure of the human myostatin growth factor. The mature growth factor (109 amino acids in length, residues 267-375) was cloned into the pProEX-Htb vector with the same reverse primer used for the precursor protein but different forwards primer (F267). A number of expression and refolding systems were trialled; however, soluble myostatin growth factor dimer was unable to be produced in isolation. To this end, the growth factor was produced by furin cleavage of MstnPP and purified by RP-HPLC. Although crystals were produced, these did not diffract and the crystal structure of myostatin in complex with follistatin was published before optimisation of preliminary crystals was achieved. In addition, a number of results were ambiguous and suggest that the purification protocol used for the myostatin growth factor dimer caused partial denaturation of the growth factor tertiary structure.

The main results and discussion surrounding this work will be presented here; although this objective was not completed, the results are interesting and warrant mention.

12.1.1 Expression

In all expression systems trialled, the mature myostatin growth factor (MstnGF) was produced within insoluble bacterial inclusion bodies (Fig. 12.1a, arrow). While the use of a SUMO solubility tag produced soluble His-SUMO-MstnGF protein that could be highly purified by its N-terminal His-tag using Ni-NTA affinity chromatography (Fig. 12.1b, lane P), removal of the SUMO tag with SUMO protease (lane D) resulted in precipitation of MstnGF (lane I) suggesting that the native structure had not been produced. The SUMO protein may have prevented precipitation through shielding of MstnGF hydrophobic regions.

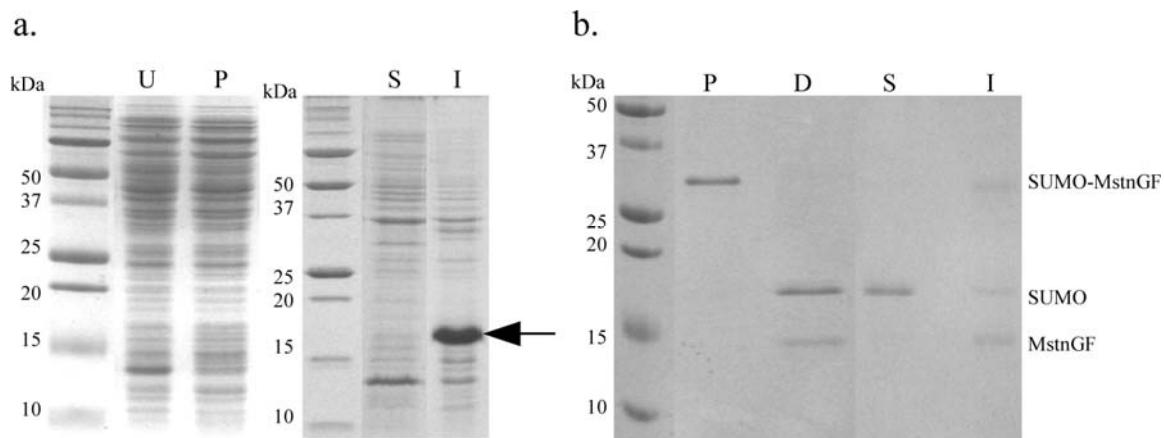


Figure 12.1 Expression analysis of the myostatin growth factor by reducing SDS-PAGE.

a. Insoluble expression (arrow) in *E.coli*. Lanes are as follows: U, uninduced; P, post-induction; S, soluble fraction; I, insoluble fraction.

b. The SUMO-tagged growth factor. Lanes are as follows: P, purified SUMO-MstnGF; D, SUMO-protease digested SUMO plus MstnGF; S, soluble fraction post-digest; I, insoluble fraction post-digest.

12.1.2 Refolding

All refolding conditions trialled for the MstnGF were unsuccessful. These included rapid dilution into a redox system (Jin, Dunn *et al.* 2004), dialysis into a glycerol-containing buffer (Thomas, Langley *et al.* 2000) and an optimized protocol published for the BMP-2 growth factor (Vallejo and Rinus 2004).

The most promising results were obtained via a refolding screen (Appendix 6, Fig. 12.10), made in-house and adapted from a combination of a published screen (Vincentelli, Canaan *et al.* 2004) recommended after personal communication (Christopher Squires, Auckland University) and the commercially available iFOLD screen (Novagen). This screen relies on the increased absorbance of precipitated protein at 340 nm and can therefore be carried out using a 96-well plate and a plate-reader. However, solubility is not always an indication of correct folding; the production of soluble aggregates has been shown previously in this thesis (Chapters 3, 4, 6 and 8). Representative results are shown in Figure 12.2a. The three numbers in each square represent absorbance at 340 nm prior to the addition of protein, immediately after the addition of protein and after 24 hours incubation at 4 °C, from top to bottom.

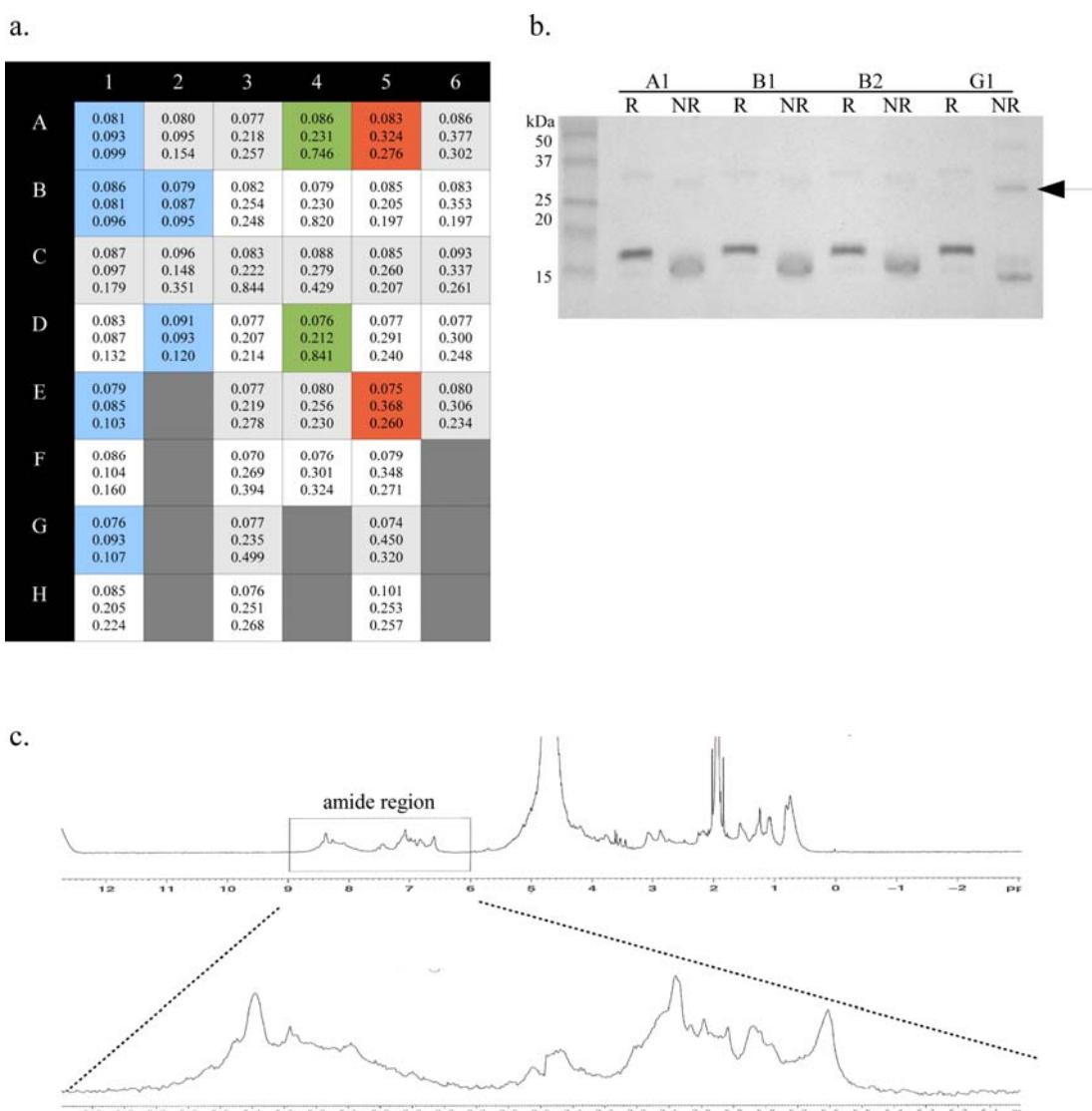


Figure 12.2 The MstnGF refolding screen.

- a. Absorbance at 340 nm was used to monitor protein precipitation. Top number, absorbance at 340 nm for solutions before the addition of solubilised protein; middle number, absorbance immediately after the addition of protein; bottom number, absorbance 24 hours post-incubation at 4 °C. Colours indicate an absorbance change within the 0.05 unit threshold (blue), an increased absorbance over 24 hours (green) and an initial increase followed by decreased absorbance after 24 hours (red).
- b. Reducing (R) vs. non-reducing (NR) 15 % SDS-PAGE for selected wells within the absorbance threshold. Well conditions are detailed in Appendix 6. Possible dimer in G1 is indicated by an arrow.
- c. 1D NMR analysis of MstnGF refolded in condition G1.

The majority of results belonged to one of three categories: absorbance change was less than the recommended 0.05 unit threshold (blue), absorbance increased throughout the incubation (green), or absorbance increased initially then decreased after 24 hours incubation (red). The latter of these suggests that a proportion of the protein precipitates initially, but refolding or soluble aggregation occurs during incubation due to the conditions used. Results where absorbance change was within the 0.05 unit threshold were selected for further analysis. Although a number of conditions satisfied this criteria, analysis by non-reducing (NR) SDS-PAGE did not always indicate the MstnGF dimer, suggesting refolding was unsuccessful (Fig. 12.2b); condition G1 (NDSB (non-detergent sulfobetaine) 201) was the exception (Fig. 12.2b, arrow).

To confirm these results, purified MstnGF was analysed with 1D NMR (Fig. 12.2c) due to an absence of access to circular dichroism instrumentation at the time. A detailed explanation of one-dimensional NMR will not be given here. In brief, the resulting spectrum presents the chemical shifts present for a protein; the characteristic shift for the amide bond is 5-10 ppm if a protein is structured but 6-9 ppm if unfolded. The 1D NMR spectrum for MstnGF spans the 6-9 ppm region, indicating that MstnGF is unfolded.

12.1.3 The production of myostatin growth factor from MstnPP

Furin cleavage of MstnPP produces the growth factor dimer in complex with the myostatin propeptide domain (Chapter 3, Fig. 3.8). In order to study the growth factor dimer in isolation, purification by reverse-phase high performance liquid chromatography (RP-HPLC) was employed, as this has been used routinely to purify other TGF- β family growth factors (Daopin, Piez *et al.* 1992; Mueller, Gottermeier *et al.* 2005; Schreuder, Liesum *et al.* 2005). Although considered to be a harsh method, the stable cysteine knot of the growth factor region is expected to be resistant to acetonitrile and/or acidic denaturation (Lawrence, Pircher *et al.* 1985). For subsequent TGF- β family crystallographic analysis, RP-HPLC is followed by lyophilization, storage at -20 °C and resuspension (Scheufler, Sebald *et al.* 1999; Mueller, Gottermeier *et al.* 2005; Schreuder, Liesum *et al.* 2005; Harrington, Morris-Triggs *et al.* 2006).

Purification of the myostatin growth factor dimer can be achieved using a 0-80% acetonitrile gradient and a Jupiter 5u C4 300A column (Phenomenex) (Fig. 12.3). RP-HPLC produces two peaks (Fig. 12.3a). The mature dimer (1) elutes at 45% ACN and the propeptide region (2) at 50% ACN, indicating that the propeptide is more hydrophobic than the correctly folded MstnGF, most likely due to denaturation. By SDS-PAGE the growth factor appears quite pure and a dimer can be observed upon removal of reducing agent (Fig. 12.3b, arrows). A small amount of growth factor is also visible in peak 2; this may represent incorrectly folded protein that has a higher degree of random coil and is more hydrophobic than that of peak 1.

Peak 1 is quite broad by comparison to peak 2, suggesting heterogeneity, although evidence of this is not clear by SDS-PAGE. Differently folded forms of MstnGF may be present, one of which is likely to represent the correctly folded protein. Growth factor fractions were lyophilized overnight following separation to remove acetonitrile, and stored at -80°C.

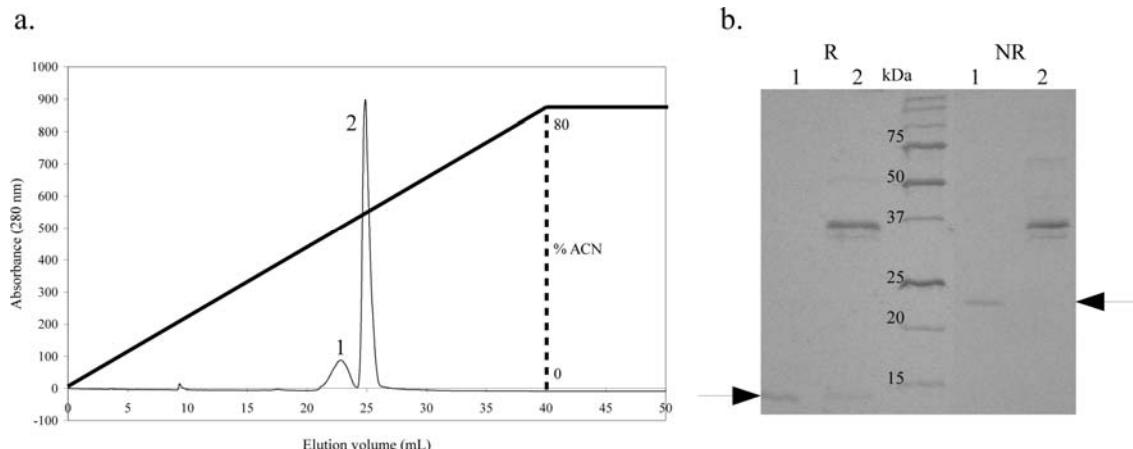


Figure 12.3 Purification of the myostatin growth factor post-furin digest by RP-HPLC.

- RP-HPLC chromatogram from 0-80% acetonitrile (ACN).
- Reducing (R) vs. non-reducing (NR) SDS-PAGE of RP-HPLC peaks 1 and 2. Arrows indicate MstnGF monomer (left) and dimer (right).

12.1.4 Structural and biochemical analysis

X-ray crystallography

Structural analysis of the myostatin growth factor dimer by X-ray crystallography was attempted. Crystal screens (Hampton and Molecular Dimensions) of lyophilized myostatin growth factor dimer, resuspended in 10 mM acetic acid, were set up in hanging drops, in a 96-well format, using the Mosquito crystallisation robot at Otago University. These initial screens provided a number of positive hits in conditions similar to those used for structural determination of other TGF- β family members. Scaling up of these conditions and the use of a 2:1 protein to mother liquor ratio produced large needles in two conditions (Fig. 12.4a and b). However, neither of these gave a diffraction pattern. Although optimisation of the initial conditions was started, the crystal structure of the MstnGF dimer/follistatin complex was published before diffracting crystals were produced (Cash, Rejon *et al.* 2009).

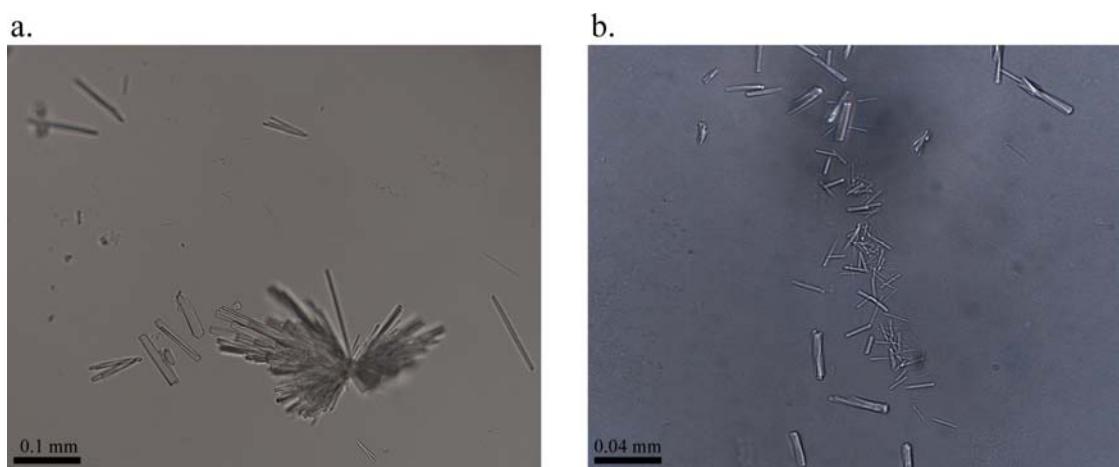


Figure 12.4 Preliminary crystals of the myostatin growth factor.

a. 30% MPD, 0.1 M Na acetate pH 4.6, 0.2 M NaCl.

b. 20% isopropanol, 0.1 M Na acetate pH 4.6, 0.2 M CaCl₂ dihydrate

A protein to mother liquor ratio of 2:1 (v/v) and a hanging-drop format were used.

Circular Dichroism spectroscopy

CD was used to gain insight into the structural characteristics of MstnGF in the absence of X-ray data. The CD spectrum of MstnGF after furin cleavage, purification by RP-HPLC, lyophilization and resuspension in 10 mM acetic acid, is predominated by a minimum at 195 nm which is indicative of disorder (Figure 12.5a, red); a second minimum can be observed at 218 nm.

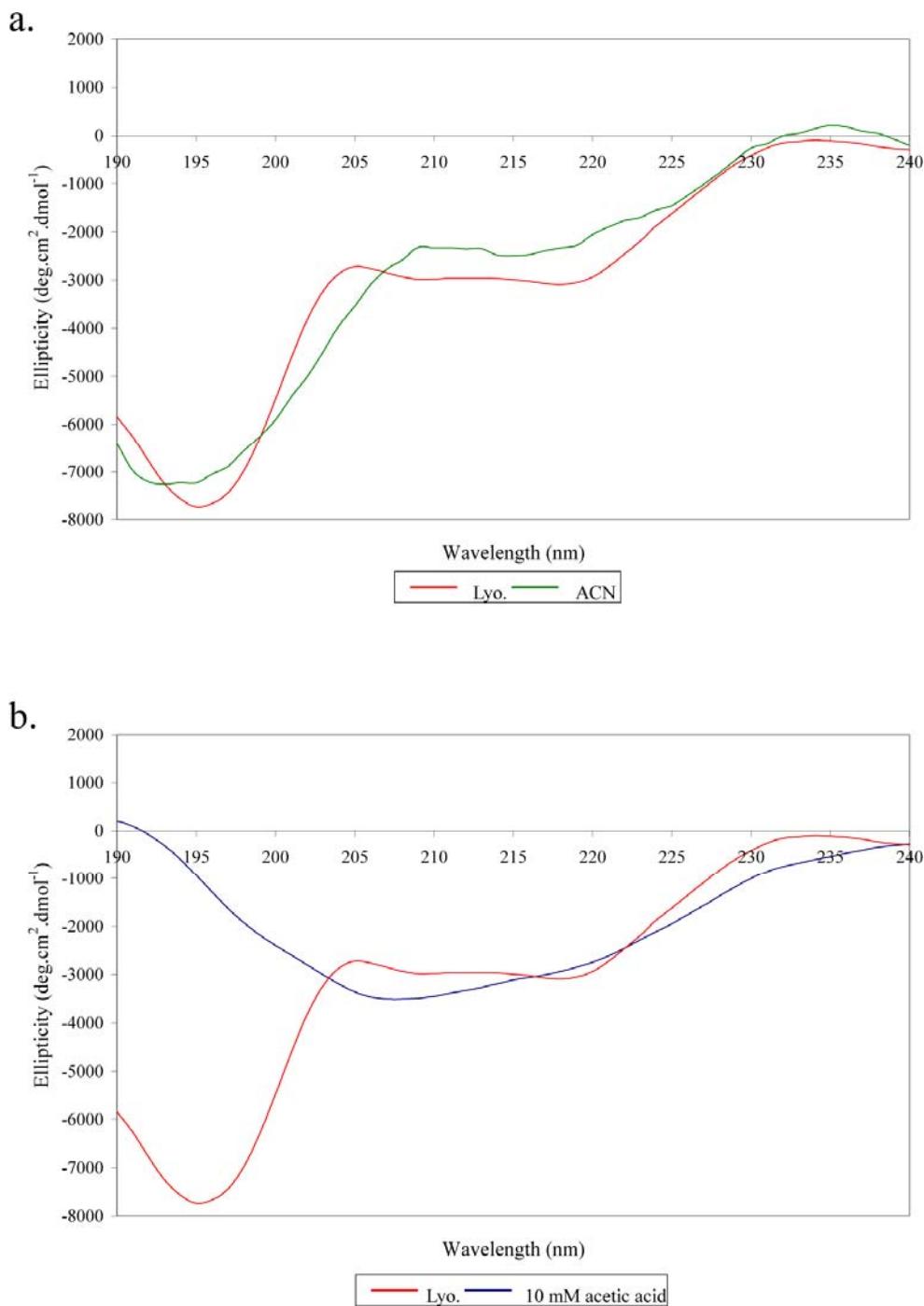


Figure 12.5 CD spectra of the myostatin growth factor in different purification conditions.

- Comparison of the lyophilized protein resuspended in 10 mM acetic acid (Lyo., red) with non-lyophilized protein in 45% acetonitrile (ACN, green).
- Comparison of the lyophilized spectrum in a. with protein buffer-exchanged from 45% ACN into 10 mM acetic acid (blue).

CDDN deconvolution predicts a structure predominated by β -sheet and random coil (Table 12.1). However, the CD spectra of other TGF- β family members contain a β -sheet minimum only (McMahon, Dignam *et al.* 1996), supporting the suggestion that the broad growth factor peak previously observed in RP-HPLC results is a heterogenous solution containing both correctly folded and unfolded protein. The CD spectra may also indicate that steps in the purification process, such as lyophilization, have partially denatured the MstnGF protein. In order to address this possibility, the CD spectrum of the RP-HPLC purified growth factor in 45 % acetonitrile (ACN) without lyophilization was also measured (Figure 12.5a, green). The shape of this spectrum is similar to that of the lyophilized protein but the β -sheet minimum is reduced in intensity and there is an increase in disorder; CDDN results are in agreement (Table 12.1). Interestingly, this suggests that the presence of ACN is more denaturing to the myostatin growth factor than the lyophilization process.

The removal of ACN and subsequent resuspension in acetic acid may allow some retrieval of the native structure. To investigate this, the RP-HPLC purified protein was buffer-exchanged into 10 mM acetic acid; this sample produces a CD spectrum that lacks the large 195 nm minimum seen for the previous samples (Fig. 12.5b, blue). However, the predominant minimum in the spectrum is at 208 nm, indicative of α -helix. Although the spectrum appears to most closely resemble that of a folded protein, CDDN deconvolution produces structure estimates intermediate between those for the ACN and lyophilized proteins, indicating the presence of residual ACN.

The CDDN deconvolution data are similar to the secondary structure profile of myostatin taken from the crystal structure (Table 12.1, last column), for the lyophilized protein in particular, suggesting that the complete removal of ACN is a critical step in production of the native myostatin protein.

SECONDARY STRUCTURE	Lyophilized (red)	45% Acetonitrile (green)	10 mM acetic acid (blue)	Crystal structure (Cash, Rejon <i>et al.</i> 2009)
α-helix	12	19	14	14
β-sheet	45	33	39	43
β-turn	16	17	16	11
Random Coil	27	31	30	32

Table 12.1 Normalised CDDN deconvolution of MstnGF CD spectra.

The secondary structures of myostatin, taken from the crystal structure (Cash, Rejon *et al.* 2009) are shown in the last column for comparison.

12.1.6 Discussion

The crystallisation and structural determination of other TGF- β family growth factors has involved RP-HPLC for purification followed by lyophilization and resuspension in acidic conditions, hence the protocol chosen for this thesis. It is intriguing that these methods do not produce similar results for the human myostatin growth factor; CD indicates a large degree of disorder present following purification by RP-HPLC. However, one possibility is that purification was comparably harsh for other TGF- β family members and the optimal crystallisation conditions provided both a means for refolding of the native structure, as well as purification away from any incorrectly folded protein. The absence of diffraction for the myostatin crystals may have been due to a heterogeneous solution, such that the presence of partially denatured forms inhibited formation of an ordered crystal lattice. If this is the case, optimisation of the initial conditions and further crystallisation screens would be expected to produce diffraction-quality crystals. An alternative explanation is that the myostatin growth factor has structural and stability characteristics that are in contrast to other members of the TGF- β family. For example, while TGF- β family members have been routinely refolded in isolation, this has not been achieved for the myostatin growth factor to date. The recently published X-ray crystal structure does not indicate any major structural differences (Cash, Rejon *et al.* 2009); however, the presence of four

disulphide bonds in the myostatin growth factor monomer, compared to the three of the majority of other TGF- β family members, may provide a reason for refolding difficulties. Furthermore, the myostatin crystal structure was published in complex form; this may indicate difficulty in crystallization of the isolated myostatin growth factor.

Summary

The active myostatin growth factor was successfully produced via refolding of the human myostatin precursor protein and furin cleavage (Chapter 3). However, purification of the growth factor post-cleavage using traditional methods of RP-HPLC and lyophilization affect the integrity of the myostatin growth factor structure, leading to the presence of misfolded forms of the protein. While this may suggest that the native cysteine knot structure had not been obtained, all structural and biochemical analyses of the refolded precursor protein suggest otherwise, and the acid-activated growth factor is able to inhibit the growth of mouse C2C12 myoblasts. Alternative methods such as Ni-NTA affinity chromatography and gel filtration will be required to purify the protein successfully.

12.2 Appendix 2: pProEX-Htb Vector Map

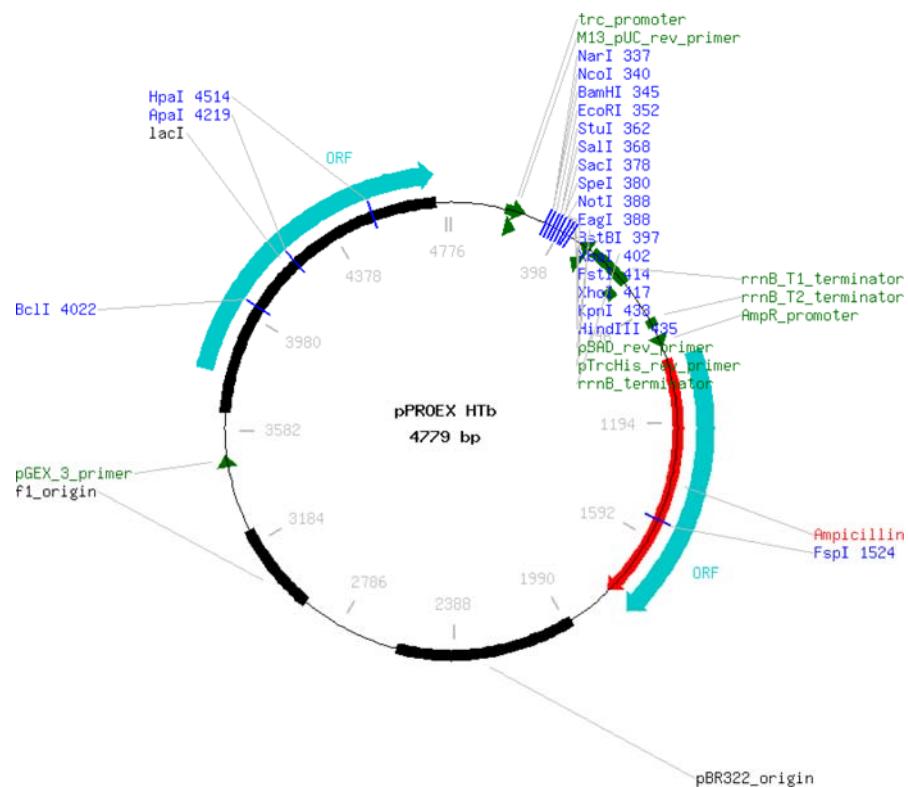


Figure 12.6 The pProEX-Htb vector map (Invitrogen).

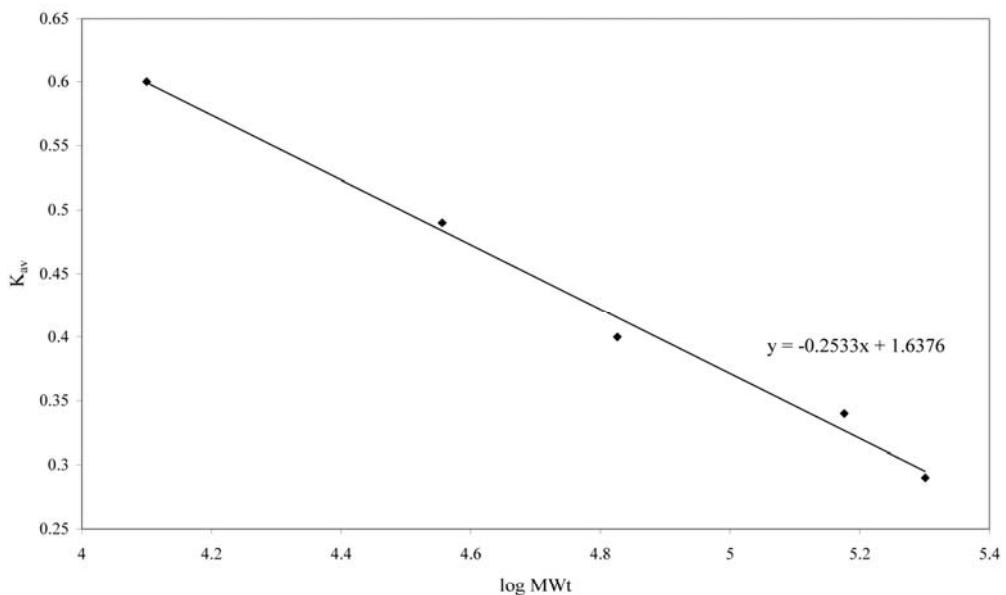
Myostatin constructs were cloned using the *Bam*HI and *Xho*I restriction sites.

12.3 Appendix 3: Gel filtration calibration curves

a. $K_{av} = (V_r - V_o) / (V_c - V_o)$

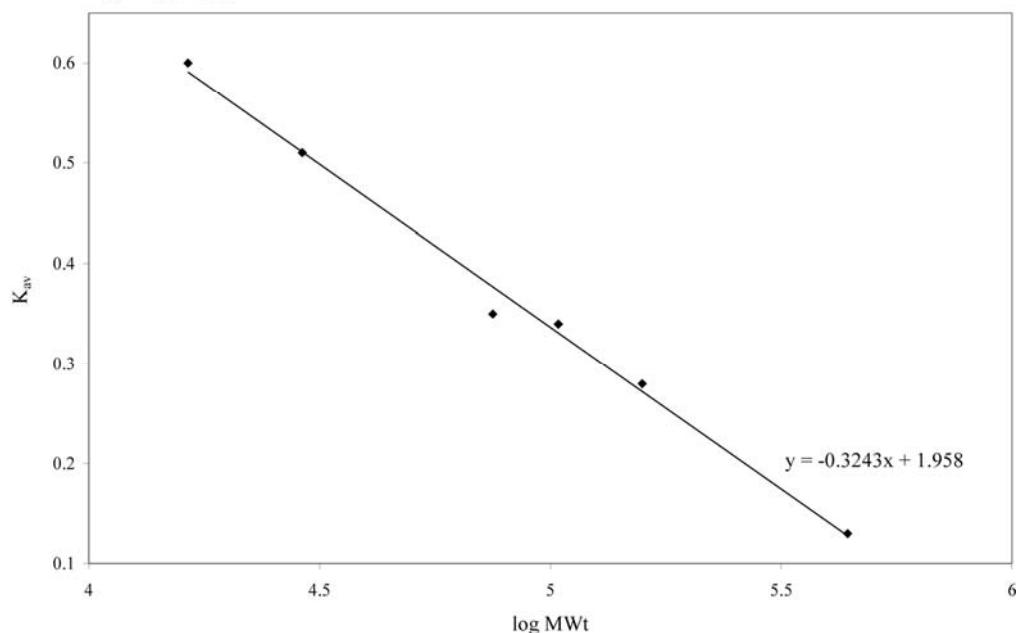
Superdex 200 1. (50 mM Tris-HCl pH 8.5, 150 mM NaCl)

$$V_c = 24 \text{ mL}$$
$$V_o = 8.29 \text{ mL}$$



b. **Superdex 200 2.** (50 mM Tris-HCl pH 8.5, 150 mM NaCl)

$$V_c = 25 \text{ mL}$$
$$V_o = 8.1 \text{ mL}$$



c.

Superdex S200 2. (50 mM HEPES pH 7.5, 150 mM NaCl)

$$V_c = 25 \text{ mL}$$
$$V_o = 8.0 \text{ mL}$$

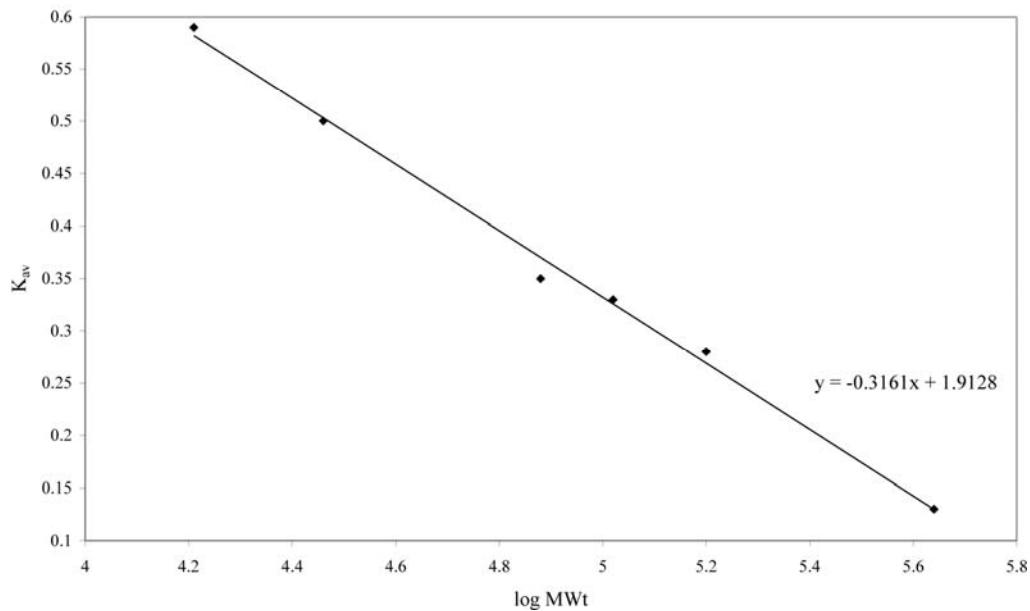


Figure 12.7 Superdex S200 calibration curves.

Two Superdex S200 columns were used in this thesis and calibration curves for both are presented. Superdex S200 1. refers to the column used in Figures 3.3 and 3.6; Superdex S200 2. is relevant from Figure 3.7.

- Superdex S200 calibration curve 1, 50 mM Tris-HCl, pH 8.5, 150 mM NaCl.
- Superdex S200 calibration curve 2, 50 mM Tris-HCl, pH 8.5, 150 mM NaCl.
- Superdex S200 calibration curve 2, 50 mM HEPES, pH 7.5, 150 mM NaCl.

12.4 Appendix 4: CDDN Deconvolution tables

a.

SECONDARY STRUCTURE	WT PP	C313Y PP	WT LC	C313Y LC
α-helix	16	18	23	20
β-sheet	35	31	34	32
β-turn	17	18	18	18
Random Coil	32	33	25	30

b.

SECONDARY STRUCTURE	MstnPP SA	Protofibrils	Amyloid
α-helix	21	13	13
β-sheet	22	38	39
β-turn	18	15	16
Random Coil	39	34	32

Table 12.2 Normalised CDDN deconvolution tables.

- a. Predicted secondary structures for correctly folded forms of WT and C313Y precursor proteins (PP) and latent complexes (LC).
- b. Predicted secondary structures for MstnPP soluble aggregates (SA), protofibrils and mature amyloid.

12.5 Appendix 5: *In silico* raw output

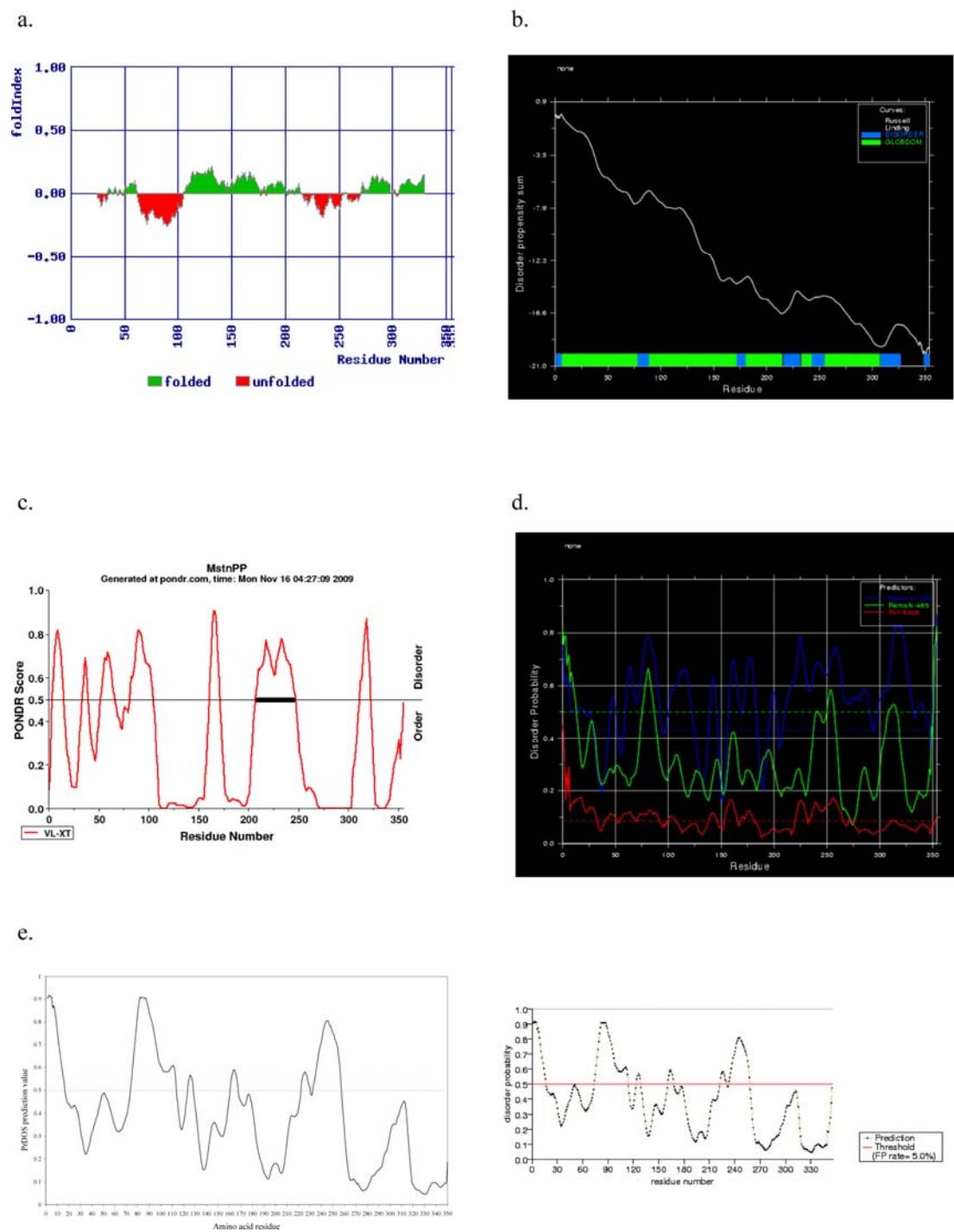


Figure 12.8 MstnPP disorder predictions: raw data.

a. FoldIndex (Prilusky, Felder *et al.* 2005); b. GlobPlot (<http://globplot.embl.de/>); c. PONDR (Romero, Obradovic *et al.* 1997); d. DisEMBL (<http://dis.embl.de/>); e. PrDOS (Ishida and Kinoshita 2007).

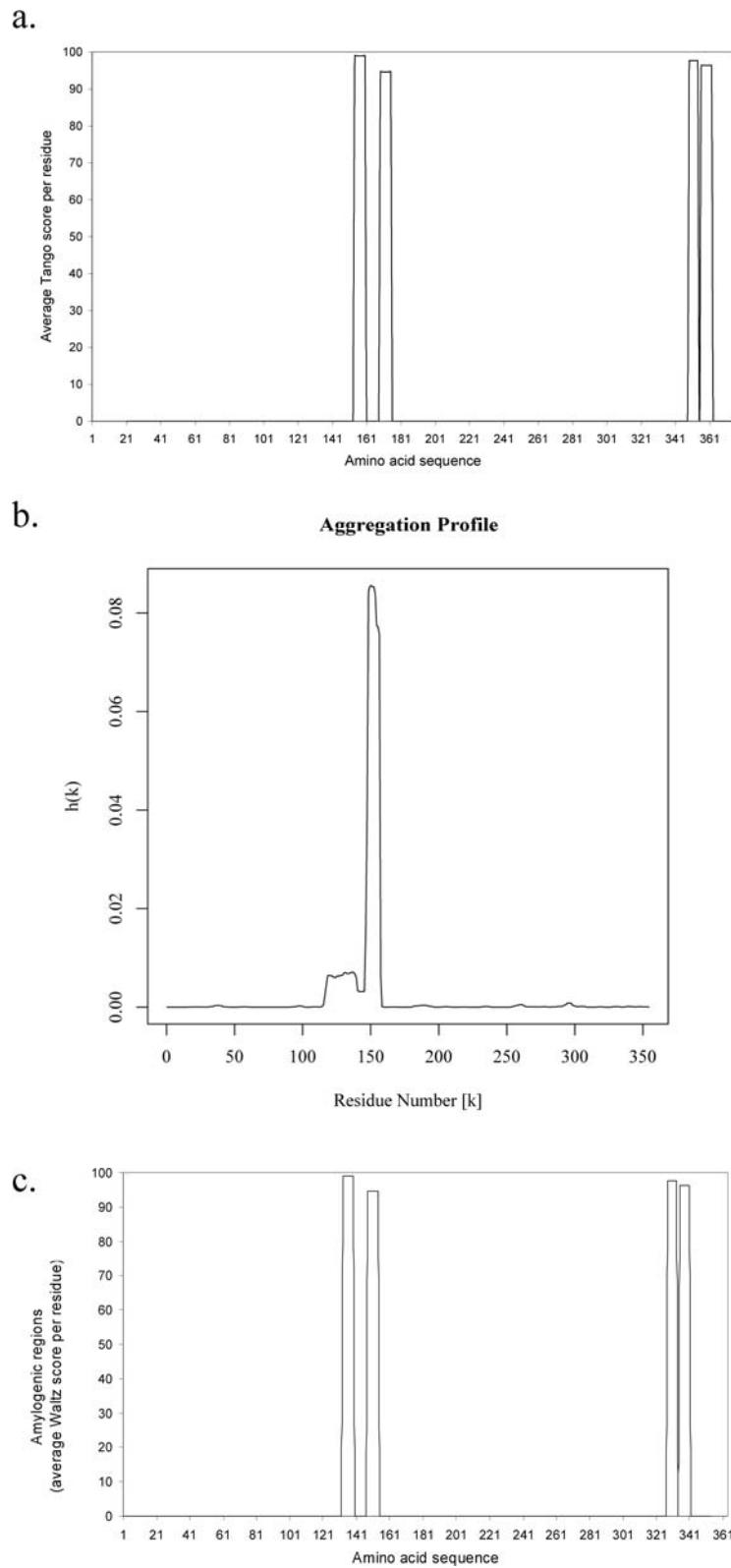


Figure 12.9 MstnPP β -aggregation and amyloid formation propensity predictions: raw data.

a. Tango (Fernandez-Escamilla, Rousseau *et al.* 2004); b. PASTA (Trovato, Seno *et al.* 2007); c. Waltz (Reumers, Schymkowitz *et al.* 2009).

12.6 Appendix 6: MstnGF Refolding screen

	1	2	3	4	5	6	7	8	9	10	11	12
A	pH 4 TCEP KCl TCEP NDSB 256	pH 4 KCl TCEP NDSB 256	pH 5 EDTA	pH 5 PEG 400 TCEP	pH 6 PEG 400 TCEP NDSB 201	pH 6 Glyc. TCEP	pH 7 PEG 400 TCEP NDSB 256 Gluc.	pH 7 NDSB 201	pH 8 EDTA NDSB 256 Gluc.	pH 8 0.2 M NaCl TCEP Gluc.	pH8 ** 5 mM GSSG	pH9 0.1 M NaCl Gluc.
B	TCEP NDSB 256	TCEP NDSB 201 Gluc.	0.1 M NaCl	TCEP NDSB 256	TCEP NDSB 195 Gluc.	EDTA	PEG 4000 TCEP	Glyc. TCEP	PEG 400 NDSB 195	** 10 mM GSSG	TCEP NDSB 195	
C	TCEP 0.1 M NaCl	KCl NDSB 195	PEG 4000 TCEP NDSB 201	Glyc. TCEP NDSB 195	Glyc. NDSB 256	0.1 M NaCl TCEP NDSB 195	KCl	NDSB 195 Gluc.	KCl TCEP NDSB 201	TCEP KCl NDSB 201 Gluc.	pH9 TCEP KCl NDSB 201 Gluc.	PEG 400
D	Glyc.	PEG 4000 Gluc.	0.2 M NaCl NDSB 195	TCEP	0.2 M NaCl TCEP Gluc.	0.2 M NaCl NDSB 201	EDTA TCEP NDSB 195	Gluc.	NDSB 201	TCEP		
E	PEG 400 Gluc.		0.2 M NaCl	KCl Gluc.	TCEP NDSB 201	PEG 4000	Glyc. TCEP	MgCl ₂ CaCl ₂		** 5 mM GSH	0.1 M NaCl NDSB 256	EDTA TCEP
F	EDTA		NDSB 256	MgCl ₂ CaCl ₂	PEG 400 NDSB 256		PEG 4000 NDSB 256		PEG 4000 NDSB 195	** 5 mM GSH 2 mM GSSG	TCEP NDSB 195	NDSB 195
G	NDSB 201		EDTA TCEP NDSB 201		KCl		0.1 M NaCl NDSB 201		TCEP Gluc.	** 5 mM GSH 5 mM GSSG	0.2 M NaCl TCEP NDSB 256	MgCl ₂ CaCl ₂
H	0.2 M NaCl NDSB 195 TCEP		TCEP Glyc.		NDSB 256 Gluc.		TCEP		0.1 M NaCl TCEP NDSB 256	** 2 mM GSH 5 mM GSSG	Glyc. NDSB 201	PEG 4000 TCEP Gluc.

Figure 12.10 MstnGF refolding screen.

Base buffers and chemical concentrations are as follows: pH4, 50 mM Na acetate; pH 5 and pH 6, 50 mM MES; pH 7 and pH 8, 50 mM Tris-HCl; pH 9, 50 mM CHES. **, base buffer of 50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA; TCEP, 1 mM; NDSB 196/201/256, 100 mM; Glucose (Gluc.), 500 mM; Glycerol (Glyc.), 20 % (v/v); PEG 400/4000, 0.05 % (w/v); KCl, 100 mM; EDTA, 1 mM; MgCl₂ and CaCl₂, 0.25 mM.

13 References

- Abedini, A. and D. P. Raleigh (2009). A Role for Helical Intermediates in Amyloid Formation by Natively Unfolded Polypeptides? *Physical Biology*. **6**: e15005.
- Allen, D. L. and T. G. Unterman (2007). Regulation of Myostatin Expression and Myoblast Differentiation by Foxo and Smad Transcription Factors. *American Journal of Physiology - Cell Physiology*. **292**: 188-199.
- Allsop, D., J. Mayes, S. Moore, A. Masad and B. J. Tabner (2008). Metal-Dependent Generation of Reactive Oxygen Species from Amyloid Proteins Implicated in Neurodegenerative Disease. *Biochemical Society Transactions*. **36**: 1293-1298.
- Amali, A., C. Lin, Y. Chen, W. Wang, H. Gong, R. Rekha, J. Lu, T. Chen and J. Wu (2008). Overexpression of Myostatin2 in Zebrafish Reduces the Expression of Dystrophin Associated Protein Complex (DAPC) Which Leads to Muscle Dystrophy. *Journal of Biomedical Science*. **15**: 595-604.
- Amthor, H., A. Otto, A. Vulin, A. Rochat, J. Dumonceaux, L. Garcia, E. Mouisel, C. Hourde, R. Macharia, M. Friedrichs, F. Relaix, P. S. Zammit, A. Matsakas, K. Patel and T. Partridge (2009). Muscle Hypertrophy Driven by Myostatin Blockade Does Not Require Stem/Precursor-Cell Activity. *Proceedings of the National Academy of Sciences*. **106**: 7479-7484.
- Anderson, S. B., A. L. Goldberg and M. Whitman (2008). Identification of a Novel Pool of Extracellular Pro-Myostatin in Skeletal Muscle. *The Journal of Biological Chemistry*. **283**: 7027-7035.
- Angelis, L. D., S. Borghi, R. Melchionna, L. Berghella, M. Baccarani-Contri, F. Parise, S. Ferrari and G. Cossu (1998). Inhibition of Myogenesis by Transforming Growth Factor- β is Density-Dependent and Related to the Translocation of Transcription Factor MEF2 to the Cytoplasm. *Proceedings of the National Academy of Sciences*. **95**: 12358-12363.

Askanas, V. and W. K. Engel (2008). Inclusion-Body Myositis: Muscle Fiber Molecular Pathology and Possible Pathogenic Significance of its Similarity to Alzheimer's and Parkinson's Disease Brains. *Acta Neuropathology*. **116**: 583-595.

Baglioni, S., F. Casamenti, M. Bucciantini, L. M. Luheshi, N. Taddei, F. Chiti, C. M. Dobson and M. Stefani (2006). Prefibrillar Amyloid Aggregates Could Be Generic Toxins in Higher Organisms. *The Journal of Neuroscience*. **26**: 8160-8167.

Barth, D., O. Kyrieleis, S. Frank, C. Renner and L. Moroder (2003). The Role of Cysteine Knots in Collagen Folding and Stability, Part II. Conformational Properties of (Pro-Hyp-Gly)_n Model Trimers with N- and C-Terminal Collagen Type III Cysteine Knots. *Chemistry*. **9**: 3703-3714.

Bely, M. and J. Makovitsky (2006). Sensitivity and Specificity of Congo Red Staining According to Romhanyi. Comparison with Puchtler's or Bennhold's Methods. *Acta Histochemistry*. **108**: 175-180.

Benjwal, S., S. Verma, K. H. Rohm and O. Gursky (2006). Monitoring Protein Aggregation During Thermal Unfolding in Circular Dichroism Experiments. *Protein Science*. **15**: 635-639.

Bentsen, J., H. Nielsen, G. v. Heijne and S. Brunak (2004). Improved Prediction of Signal Peptides: SignalP 3.0. *Journal of Molecular Biology*. **16**: 783-795.

Berjanskii, M. and D. S. Wishart (2006). NMR: Prediction of Protein Flexibility. *Nature Protocols*. **1**: 683-688.

Berry, C., M. Thomas, B. Langley, M. Sharma and R. Kambadur (2002). Single Cysteine to Tyrosine Transition Inactivates the Growth Inhibitory Function of Piedmontese Myostatin. *American Journal of Physiology*. **283**: 135-141.

Bornberg-Bauer, E., E. Rivals and M. Vingron (1998). Computational Approaches to Identify Leucine Zippers. *Nucleic Acids Research*. **26**: 2740-2746.

Bourhim, M., M. Kruzel, T. Srikrishnan and T. Nicotera (2007). Linear Quantitation of A β Aggregation Using Thioflavin T: Reduction in Fibril Formation by Colostrinin. *Journal of Neuroscience Methods*. **160**: 264-268.

Bradley, L., P. J. Yaworsky and F. S. Walsh (2008). Myostatin as a Therapeutic Target for Musculoskeletal Disease. *Cellular and Molecular Life Sciences*. **65**: 2119-2124.

Brownh, P. D., L. M. Wakefiel, A. D. Levinson and M. B. Sporn (1990). Physicochemical Activation of Recombinant Latent Transforming Growth Factor- β s 1, 2, and 3. *Growth Factors*. **3**: 35-43.

Brunner, A. M., M. N. Lioubin, H. Marquardt, A. R. Malacko, W. C. Wang, R. A. Shapiro, M. Neubauer, J. Cook, L. Madisen and A. F. Purchio (1992). Site-Directed Mutagenesis of Glycosylation Sites in the Transforming Growth Factor β 1 (TGF β 1) and TGF β 2 (414) Precursors and of Cysteine Residues Within Mature TGF β 1: Effects on Secretion and Activity. *Molecular Endocrinology*. **6**: 1691-1700.

Brunner, A. M., H. Marquardt, A. R. Malacko, M. N. Lioubin and A. F. Purchio (1989). Site-Directed Mutagenesis of Cysteine Residues in the Pro Region of the Transforming Growth Factor β 1 Precursor. *The Journal of Biological Chemistry*. **264**: 13660-13664.

Bucciantini, M., G. Calloni, F. Chiti, L. Formigli, D. Nosi, C. M. Dobson and M. Stefani (2004). Prefibrillar Amyloid Protein Aggregates Share Common Features of Cytotoxicity. *The Journal of Biological Chemistry*. **279**: 31374-31382.

Bucciantini, M., E. Giannoni, F. Chiti, F. Baroni, L. Formigli, J. Zurdo, N. Taddel, G. Ramponi, C. M. Dobson and M. Stefani (2002). Inherent Toxicity of Aggregates Implies a Common Mechanism for Protein Misfolding Diseases. *Nature* **416**: 507-511.

Callebaut, I., G. Labesse, P. Durand, A. Poupon, L. Canard, J. Chomilier, B. Henrissat and J. Mornon (1997). Deciphering Protein Sequence Information Through

Hydrophobic Cluster Analysis (HCA): Current Status and Perspectives. *Cellular and Molecular Life Sciences.* **53**: 621-645.

Calloni, G., C. Lendel, S. Campioni, S. Giannini, A. Gliootti, A. Relini, M. Vendruscolo, C. M. Dobson, X. Salvatella and F. Chiti (2008). Structure and Dynamics of a Partially Folded Protein are Decoupled from Its Mechanism of Aggregation. *Journal of the American Chemical Society.* **130**: 13040-13050.

Cash, J. N., C. A. Rejon, A. C. McPherron, D. J. Bernard and T. B. Thompson (2009). The Structure of Myostatin:Follistatin 288: Insights into Receptor Utilization and Heparin Binding. *The EMBO Journal.* **28**: 2662-2676.

Cerf, E., R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y. F. Dufrene, V. Narayanaswami, E. Goormaghtigh, J.-M. Ruysschaert and V. Raussens (2009). Antiparallel β -Sheet: A Signature Structure of the Oligomeric Amyloid β -Peptide. *Biochemical Journal.* **421**: 415-423.

Chang, H., C. W. Brown and M. M. Matzuk (2002). Genetic Analysis of the Mammalian Transforming Growth Factor- β Superfamily. *Endocrine Reviews.* **23**: 787-823.

Cheng, I. H., K. Scearce-Levie, J. Legleiter, J. J. Palop, H. Gerstein, N. Bien-Ly, J. Puolivali, S. Lesne, K. H. Ashe, P. J. Muchowski and L. Mucke (2007). Accelerating Amyloid- β Fibrillization Reduces Oligomer Levels and Functional Deficits in Alzheimer Disease Mouse Models. *The Journal of Biological Chemistry.* **282**: 23818-23828.

Chimon, S., M. A. Shaibat, C. R. Jones, D. C. Calero, B. Aizezi and Y. Ishii (2007). Evidence of Fibril-Like β -Sheet Structures in a Neurotoxic Intermediate of Alzheimer's β -Amyloid. *Nature Structural and Molecular Biology.* **14**: 1157-1164.

Cohen, A. S. and E. Calkins (1959). Electron Microscopic Observation on a Fibrous Component in Amyloid of Diverse Origins. *Nature.* **183**: 1202-1203.

Combet, C., C. Blanchet, C. Geourjon and G. Deleage (2000). NPS@: Network Protein Sequence Analysis. *Trends in Biochemical Sciences*. **291**: 147-150.

Conway, K. A., S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson and P. T. Lansbury (2000). Acceleration of Oligomerization, Not Fibrillization, Is a Shared Property of Both α -Synuclein Mutations Linked to Early-Onset Parkinson's Disease: Implications for Pathogenesis and Therapy. *Proceedings of the National Academy of Sciences*. **97**: 571-576.

Costelli, P., M. Muscaritoli, A. Bonetto, F. Penna, P. Reffo, M. Bossola, G. Bonelli, G. B. Doglietto, F. M. Baccino and F. R. Fanelli (2008). Muscle Myostatin Signalling is Enhanced in Experimental Cancer Cachexia. *European Journal of Clinical Investigation*. **38**: 531-538.

Craik, D. J., N. L. Daly and C. Waine (2000). The Cysteine Knot Motif in Toxins and Implications for Drug Design. *Toxicon*. **39**: 43-60.

Cuff, J., M. Clamp, A. Siddiqui, M. Finlay and G. Barton (1998). JPred: A Consensus Secondary Structure Prediction Server. *Bioinformatics*. **14**: 892-893.

Daggett, V. (2006). α -Sheet: The Toxic Conformer in Amyloid Diseases? *Accounts of Chemical Research*. **59**: 594-602.

Daopin, S., K. Piez, Y. Ogawa and D. Davies (1992). Crystal Structure of Transforming Growth Factor- β 2: An Unusual Fold for the Superfamily. *Science*. **257**: 369-373.

Demuro, A., E. Mina, R. Kayed, S. C. Milton, I. Parker and C. G. Glabe (2005). Calcium Dysregulation and Membrane Disruption as a Ubiquitous Neurotoxic Mechanism of Soluble Amyloid Oligomers. *The Journal of Biological Chemistry*. **280**: 17294-17300.

Dijke, P. and C. Hill (2004). New Insights into TGF- β Signalling. *Trends in Biochemical Sciences*. **29**: 265-273.

Divry, P. and M. Florkin (1927). Sur Les Proprietes Optiques De L'amyloid. *Societe de Biologie*. **97**: 180-181.

Du, R., Y. F. Chen, X. R. An, X. Y. Yang, Y. Ma, L. Zhang, X. L. Yuan, L. M. Chen and J. Qin (2005). Cloning and Sequence Analysis of Myostatin Promoter in Sheep. *DNA Sequence*. **16**: 412-417.

Dumonceaux, J. and H. Amthor (2009). Current Advances in the Development of Therapies for Neuromuscular Disorders based on Myostatin Signalling. *Neuromuscular Disorders*. **19**: 797-799.

Eanes, E. D. and G. G. Glenner (1968). X-Ray Diffraction Studies on Amyloid Filaments. *Journal of Histochemistry and Cytochemistry*. **16**: 673-677.

Emsley, P. and K. Cowtan (2004). Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D*. **60**: 2126-2132.

Erikson, H. P. (2009). Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online*. **11**: 32-51.

Etienne, M. A., N. J. Edwin, J. P. Aucoin, P. S. Russo, R. L. McCarley and R. P. Hammer (2007). β -Amyloid Protein Aggregation. *Methods in Molecular Biology*. **386**: 203-225.

Farazi, T. A., G. Waksman and J. I. Gordon (2001). The Biology and Enzymology of Protein N-Myristoylation. *The Journal of Biological Chemistry*. **276**: 39501-39504.

Fei, L. and S. Perrett (2009). Disulphide Bond Formation Significantly Accelerates the Assembly of Ure2p Fibrils Due to Proximity of a Potential Amyloid Stretch. *The Journal of Biological Chemistry*. **284**: 11134-11141.

- Fernandez-Escamilla, A. M., F. Rousseau, J. Schymkowitz and L. Serrano (2004). Prediction of Sequence-Dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nature Biotechnology*. **22**: 1302-1306.
- Forbes, D., M. Jackman, A. Bishop, M. Thomas, R. Kambadur and M. Sharma (2006). Myostatin Auto-Regulates Its Expression by Feedback Loop through Smad7 Dependent Mechanism. *Journal of Cellular Physiology*. **206**: 264-272.
- Francoeur, A. M. and A. Assalian (1996). MICROCAT: A Novel Cell Proliferation and Cytotoxicity Assay Based on WST-1. *Biochemica*. **3**: 19-25.
- Frid, P., D. V. Anisimov and N. Popovic (2007). Congo Red and Protein Aggregation in Neurodegenerative Diseases. *Brain Research Reviews*. **53**: 135-160.
- Frieden, C. (2007). Protein Aggregation Processes: In Search of the Mechanism. *Protein Science*. **16**: 2334-2344.
- Funkenstein, B. and Y. Rebhan (2007). Expression, Purification, Renaturation and Activation of Fish Myostatin Expressed in Escherichia coli: Facilitation of Refolding and Activity Inhibition by Myostatin Prodomain. *Protein Expression and Purification*. **10**: 54-65.
- Gaboriaud, C., V. Bissery, T. Benchetrit and J. Mornon (1987). Hydrophobic Cluster Analysis: An Efficient New Way to Compare and Analyse Hydrophobic Amino Acid Sequences. *FEBS Letters*. **224**: 149-155.
- Garlepp, M. J. and F. L. Mastaglia (2007). Inclusion Body Myositis: New Insights Into Pathogenesis. *Current Opinion in Rheumatology*. **20**: 662-668.
- Gaussian, V. and C. Depre (2005). Myostatin, the Cardiac Chalone of Insulin-Like Growth Factor-1. *Cardiovascular Research*. **68**: 347-349.
- Geerlof, A., J. Brown, B. Coutard, M. P. Egloff, F. J. Enguita, M. J. Fogg, R. J. C. Gilbert, M. R. Groves, A. Haouz, J. E. Nettleship, P. Nordlund, R. J. Owens, M. Ruff,

S. Sainsbury, D. I. Svergun and M. Wilmanns (2006). The Impact of Protein Characterization in Structural Proteomics. *Acta Crystallographica Section D*. **62**: 1125-1136.

Glaebe, C. G. (2006). Common mechanisms of amyloid oligomer pathogenesis in degenerative disease. *Neurobiology of Aging*. **27**: 570-575.

Glenner, G. G. and C. W. Wong (1984). Alzheimer's Disease: Initial Report of the Purification and Characterisation of a Novel Cerebrovascular Amyloid Protein. *Biochemical and Biophysical Research Communications*. **120**: 885-890.

Goda, S., K. Takano, Y. Yamagata, R. Nagata, H. Akutsu, S. Maki, K. Namba and K. Yutani (2000). Amyloid Protofilament Formation of Hen Egg Lysozyme in Highly Concentrated Ethanol Solution. *Protein Science*. **9**: 369-375.

Goedert, M. and M. G. Spillantini (2006). A Century of Alzheimer's Disease. *Science*. **314**: 777-781.

Gonzalez-Cadavid, N., W. Taylor, K. Yarasheski, I. Sinha-Hikim, K. Ma, S. Ezzat, R. Shen, R. Lalani, S. Asa, M. Mamita, G. Nair, S. Arver and S. Bhasin (1998). Organization of the Human Myostatin Gene and Expression in Healthy Men and HIV-Infected Men with Muscle Wasting. *Proceedings of the National Academy of Sciences*. **95**: 14938-14943.

Grabarek, Z. (2006). Structural Basis for Diversity of the EF-Hand Calcium-Binding Proteins. *Journal of Molecular Biology*. **359**: 509-525.

Gray, A. M. and A. J. Mason (1990). Requirement for Activin A and Transforming Growth Factor- β 1 Pro-Regions in Homodimer Assembly. *Science*. **247**: 1328-1330.

Gregory, K. E., R. N. Ono, N. L. Charbonneau, C. L. Kuo, D. R. Keene, H. P. Bachinger and L. Y. Sakai (2005). The Prodomain of BMP-7 Targets the BMP-7 Complex to the Extracellular Matrix. *The Journal of Biological Chemistry*. **280**: 27970-27980.

Groenning, M., M. Norrman, J. M. Flink, M. v. d. Weert, J. T. Bukrinsky, G. Schluckebier and S. Frokjaer (2007). Binding Mode of Thioflavin T in Insulin Amyloid Fibrils. *Journal of Structural Biology*. **159**: 483-497.

Groppe, J., C. S. Hinck, P. Samavarchi-Tehrani, C. Zubieta, J. P. Schuermann, A. B. Taylor, P. M. Schwarz, J. L. Wrana and A. P. Hinck (2008). Cooperative Assembly of TGF- β Superfamily Signaling Complexes is Mediated by Two Disparate Mechanisms and Distinct Modes of Receptor Binding. *Molecular Cell*. **29**: 157-168.

Hamrick, M. W., C. Pennington, C. N. Webb and C. M. Isales (2006). Resistance to Body Fat Gain in 'Double-Muscled' Mice Fed a High-Fat Diet. *International Journal of Obesity*. **30**: 868-870.

Harrington, A. E., S. A. Morris-Triggs, B. T. Ruotolo, C. V. Robinson, S. Ohnuma and M. Hyvonen (2006). Structural Basis for the Inhibition of Activin Signalling by Follistatin. *The EMBO Journal*. **25**: 1035-1045.

He, B., K. Wang, Y. Liu, B. Xue, V. N. Uversky and A. K. Dunker (2009). Predicting Intrinsic Disorder in Proteins: An Overview. *Cell Research*. **19**: 929-949.

Heiring, C. and Y. A. Muller (2001). Folding Screening Assayed by Proteolysis: Application to Various Cysteine Deletion Mutants of Vascular Endothelial Growth Factor. *Protein Engineering*. **14**: 183-188.

Hill, J. J., M. V. Davies, A. A. Pearson, J. H. Wang, R. M. Hewick, N. M. Wolfman and Y. Qiu (2002). The Myostatin Propeptide and the Follistatin-Related Gene Are Inhibitory Binding Proteins of Myostatin in Normal Serum. *The Journal of Biological Chemistry*. **277**: 40735-40741.

Hill, J. J., Y. Qiu, R. M. Hewick and N. M. Wolfman (2003). Regulation of Myostatin In Vivo by Growth and Differentiation Factor-Associated Serum Protein-1: A Novel Protein with Protease Inhibitor and Follistatin Domains. *Molecular Endocrinology*. **17**: 1144-1154.

Hillger, F., G. Herr, R. Rudolph and E. Schwarz (2005). Biophysical Comparison of BMP-2, ProBMP2, and the Free Pro-Peptide Reveals Stabilization of the Pro-Peptide by the Mature Growth Factor. *The Journal of Biological Chemistry*. **280**: 14974-14980.

Hirai, S., H. Matsumoto, N. Hino, H. Kawachi, T. Matsui and H. Yano (2007). Myostatin Inhibits Differentiation of the Bovine Preadipocyte. *Domestic Animal Endocrinology*. **32**: 1-14.

Iakoucheva, L. M., C. J. Brown, J. D. Lawson, Z. Obradovic and A. K. Dunker (2002). Intrinsic Disorder in Cell-Signaling and Cancer-Associated Proteins. *Journal of Molecular Biology*. **323**: 573-584.

Ishida, T. and K. Kinoshita (2007). PrDOS: Prediction of Disordered Protein Regions From Amino Acid Sequence. *Nucleic Acids Research*. **35**: 460-464.

Ishima, R. and D. A. Torchia (2000). Protein Dynamics from NMR. *Nature Structural Biology*. **7**: 740-743.

Jansen, R., W. Dzwolak and R. Winter (2005). Amyloidogenic Self-Assembly of Insulin Aggregates Probed by High Resolution Atomic Force Microscopy. *Biophysical Journal*. **88**: 1344-1353.

Jiang, M., L. Liang, S. Wang, T. Ratovitski, J. Holmstrom, C. Barker and R. Stotish (2004). Characterization and Identification of the Inhibitory Domain of GDF-8 Propeptide. *Biochemical and Biophysical Research Communications*. **315**: 525-531.

Jimenez, J. L., E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson and H. R. Saibil (2002). The Protofilament Structure of Insulin Amyloid Fibrils. *Proceedings of the National Academy of Sciences*. **99**.

Jin, H. J., M. A. Dunn, D. Borthakur and Y. S. Kim (2004). Refolding and Purification of Unprocessed Porcine Myostatin Expressed in Escherichia coli. *Protein Expression and Purification*. **35**: 1-10.

Joulia-Ekaza, D. and G. Cabello (2006). Myostatin Regulation of Muscle Development: Molecular Basis, Natural Mutations, Physiopathological Aspects. *Experimental Cell Research.* **312**: 2401-2414.

Joulia, D., H. Bernardi, V. Garandel, F. Rabenoelina, B. Vernus and G. Cabello (2003). Mechanisms Involved in the Inhibition of Myoblast Proliferation and Differentiation by Myostatin. *Experimental Cell Research.* **286**: 263-275.

Juárez, J., P. Taboada and V. Mosquera (2009). Existence of Different Structural Intermediates on the Fibrillation Pathway of Human Serum Albumin. *Biophysical Journal.* **96**: 2353-2370.

Kambadur, R., M. Sharma, T. P. L. Smith and J. J. Bass (1997). Mutations in Myostatin (GDF8) in Double-Muscled Belgian Blue and Piedmontese Cattle. *Genome Research.* **7**: 910-915.

Karlin, D., F. Ferron, B. Canard and S. Longhi (2003). Structural Disorder and Modular Organization in Paramyxovirinae N and P. *Journal of General Virology.* **84**: 3239-3252.

Karpati, G. and E. K. O'Ferrall (2009). Sporadic Inclusion Body Myositis: Pathogenic Considerations. *Annals Neurology.* **65**: 7-11.

Keah, H. and M. Hearn (2005). A Molecular Recognition Paradigm: Promiscuity Associated with the Ligand-Receptor Interactions of the Activin Members of the TGF- β Superfamily. *Journal of Molecular Recognition.* **18**: 385-403.

Kelley, L. and M. Sternberg (2009). Protein Structure Prediction on the Web: A Case Study Using the Phyre Server. *Nature Protocols.* **4**: 363-371.

Kelly, S. M., T. J. Jess and N. C. Price (2005). How to Study Proteins by Circular Dichroism. *Biochimica et Biophysica Acta.* **1751**: 119-139.

Kim, W. and M. Hecht (2008). Mutations Enhance the Aggregation Propensity of the Alzheimer's A β Peptide. *Journal of Molecular Biology*. **377**: 565-574.

Kumar, S., S. K. Mohanty and J. B. Udgaonkar (2007). Mechanism of Formation of Amyloid Protofibrils of Barstar from Soluble Oligomers: Evidence for Multiple Steps and Lateral Association Coupled to Conformational Conversion. *Journal of Molecular Biology*. **367**: 1186-1204.

Kumar, S. and J. B. Udgaonkar (2009). Structurally Distinct Amyloid Protofibrils Form on Separate Pathways of Aggregation of a Small Protein. *Biochemistry*. **48**: 6441-6449.

Laemmli, U. K. (1970). Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4. *Nature*. **227**: 680-685.

Lai, E., T. Teodoro and A. Volchuk (2007). Endoplasmic Reticulum Stress: Signaling the Unfolded Protein Response. *Physiology*. **22**: 193-201.

Langley, B., M. Thomas, A. Bishop, M. Sharma, S. Gilmour and R. Kambadur (2002). Myostatin Inhibits Myoblast Differentiation by Down-Regulating MyoD Expression. *The Journal of Biological Chemistry*. **277**: 49831-49840.

Lashuel, H. A. (2005). Membrane Permeabilization: A Common Mechanism in Protein-Misfolding Diseases. *Science of Aging Knowledge Environment*. **2005**: pe28.

Lawrence, D. A., R. Pircher and P. Jullien (1985). Conversion of a High Molecular Weight Latent β -TGF from Chicken Embryo Fibroblasts into a Low Molecular Weight Active β -TGF Under Acidic Conditions. *Biochemical and Biophysical Research Communications*. **133**: 1026-1034.

Lee, S. J. (2004). Regulation of Muscle Mass by Myostatin. *Annual Reviews in Cell and Developmental Biology*. **20**: 61-86.

Lee, S. J. (2007). Sprinting Without Myostatin: A Genetic Determinant of Athletic Prowess. *Trends in Genetics*. **23**: 475-477.

Lee, S. J. (2008). Genetic Analysis of the Role of Proteolysis in the Activation of Latent Myostatin. *PLoS One*. **3**: e1628.

Lee, S. J. and A. C. McPherron (2001). Regulation of Myostatin Activity and Muscle Growth. *Proceedings of the National Academy of Sciences*. **98**: 9306-9311.

Lee, S. J., L. A. Reed, M. V. Davies, S. Girgenrath, M. E. P. Goad, K. N. Tomkinson, J. F. Wright, C. Barker, G. Ehrmantraut, J. Holmstrom, B. Trowell, B. Gertz, M.-S. Jiang, S. M. Sebald, M. Matzuk, E. Li, L. Liang, E. Quattlebaum, R. L. Stotish and N. M. Wolfman (2005). Regulation of Muscle Growth by Multiple Ligands Signaling through Activin Type II Receptors. *Proceedings of the National Academy of Sciences*. **102**: 18117-18122.

Levine-III, H. (1993). Thioflavine T Interaction with Synthetic Alzheimer's Disease β -Amyloid Peptides: Detection of Amyloid Aggregation in Solution. *Protein Science*. **2**: 404-410.

Lin, W. and B. Popko (2009). Endoplasmic Reticulum Stress in Disorders of Myelinating Cells. *Nature Neuroscience*. **12**: 379-385.

Lorenzo, A. and B. A. Yankner (1994). β -Amyloid Neurotoxicity Requires Fibril Formation and is Inhibited by Congo Red. *Proceedings of the National Academy of Sciences*. **91**: 12243-12247.

Luhrs, T., C. Ritter, M. Adrian, D. Riek-Lohr, B. Bohrmann, H. Dobeli, D. Schubert and R. Riek (2005). 3D Structure of Alzheimer's Amyloid- β (1-42) fibrils. *Proceedings of the National Academy of Sciences*. **102**: 17342-17347.

Maji, S., D. Schubert, C. Rivier, S. Lee, J. Rivier and R. Riek (2008). Amyloid as a Depot for the Formulation of Long-Acting Drugs. *PLoS Biology*. **6**: e17.

Makin, O. S., E. Atkins, P. Sikorski, J. Johansson and L. C. Serpell (2005). Molecular Basis for Amyloid Fibril Formation and Stability. *Proceedings of the National Academy of Sciences*. **102**: 315-320.

Malisauskas, M., J. Ostman, A. Darinskas, V. Zamotin, E. Liutkevicius, E. Lundgren and L. A. Morozova-Roche (2005). Does the Cytotoxic Effect of Transient Amyloid Oligomers From Common Equine Lysozyme In Vitro Imply Innate Amyloid Toxicity? *The Journal of Biological Chemistry*. **280**: 6269-6275.

Manceau, M., J. Gros, K. Thome, A. McPherron, B. Paterson and C. Marcelle (2008). Myostatin Promotes the Terminal Differentiation of Embryonic Muscle Progenitors. *Genes and Development*. **22**: 668-681.

Manelli, A., L. Bulfinch, P. Sullivan and M. LaDu (2007). Aβ42 Neurotoxicity in Primary Co-Cultures: Effect of ApoE Isoform and Abeta Conformation. *Neurobiology of Aging*. **28**: 1139-1147.

Marshall, K. E. and L. C. Serpell (2009). Structural Integrity of β-Sheet Assembly. *Biochemical Society Transactions*. **37**: 671-676.

Mauro, M., E. F. Craparo, A. Podesta, D. Bulone, R. Carrotta, V. Martorana, G. Tiana and P. L. S. Biagio (2007). Kinetics of Different Processes in Human Insulin Amyloid Formation. *Journal of Molecular Biology*. **366**: 258-274.

Mayne, L. and S. W. Englander (2000). Two-State vs. Multistate Protein Unfolding Studied by Optical Melting and Hydrogen Exchange. *Protein Science*. **9**: 1873-1877.

McBride, J. W., X. J. Yu and D. H. Walker (2000). Glycosylation of Homologous Immunodominant Proteins of *Ehrlichia chaffeensis* and *Ehrlichia canis*. *Infection and Immunity*. **68**: 13-18.

McFarlane, C., B. Langley, M. Thomas, A. Hennebry, E. Plummer, G. Nicholas, C. McMahon, M. Sharma and R. Kambadur (2005). Proteolytic Processing of Myostatin is Auto-Regulated During Myogenesis. *Developmental Biology*. **283**: 58-69.

McFarlane, C., E. Plummer, M. Thomas, A. Hennebry, M. Ashby, N. Ling, H. Smith, M. Sharma and R. Kambadur (2006). Myostatin Induces Cachexia by Activating the Ubiquitin Proteolytic System through an NF- κ B-Independent, Foxo1-Dependent Mechanism. *Journal of Cellular Physiology*. **209**: 501-514.

McMahon, G. A., J. D. Dignam and L. E. Gentry (1996). Structural Characterization of the Latent Complex Between Transforming Growth Factor β 1 and β 1-Latency-Associated Peptide. *Biochemical Journal*. **313**: 343-351.

McPherron, A. C., A. M. Lawler and S. J. Lee (1997). Regulation of Skeletal Muscle Mass in Mice by a New TGF- β Superfamily Member. *Nature*. **387**: 83-90.

McPherron, A. C. and S. J. Lee (1997). Double Muscling in Cattle Due to Mutations in the Myostatin Gene. *Proceedings of the National Academy of Sciences*. **94**: 12457-12461.

Meredith, S. C. (2005). Protein Denaturation and Aggregation. Cellular Responses to Denatured and Aggregated Proteins. *Annals of the New York Academy of Sciences*. **1066**: 181-221.

Milner-White, E. J., J. D. Watson, G. Qi and S. Hayward (2006). Amyloid Formation May Involve α - to β Sheet Interconversion Via Peptide Plane Flipping. *Structure*. **14**: 1369-1376.

Mitchell, M. D., C. C. Osepchook, K. C. Leung, C. D. McMahon and J. J. Bass (2006). Myostatin is a Human Placental Product That Regulates Glucose Uptake. *Journal of Clinical Endocrinology and Metabolism*. **91**: 1343-1437.

Miura, T., Y. Kishioka, J. I. Wakamatsu, A. Hattori, A. Hennebry, C. J. Berry, M. Sharma, R. Kambadur and T. Nishimura (2006). Decorin Binds Myostatin and Modulates Its Activity to Muscle Cells. *Biochemical and Biophysical Research Communications*. **340**: 675-680.

Moore, B. D., V. Rangachari, W. M. Tay, N. M. Milkovic and T. L. Rosenberry (2009). Biophysical Analyses of Synthetic Amyloid- β (1-42) Aggregates Before and After Covalent Cross-Linking. Implications for Deducing the Structure of Endogenous Amyloid- β Oligomers. *Biochemistry*. **48**: 11796-11806.

Moustakas, A. and C. H. Heldin (2005). Non-Smad TGF- β Signals. *Journal of Cell Science*. **118**: 3573-3584.

Mueller, T. D., M. Gottermeier, W. Sebald and J. Nickel (2005). Crystallization and Preliminary X-Ray Diffraction Analysis of Human Growth and Differentiation Factor 5 (GDF-5). *Acta Crystallographica Section F*. **61**: 134-136.

Muller, Y. and C. Heiring (2002). The Cysteine Knot Promotes Folding and Not Thermodynamic Stability in Vascular Endothelial Growth Factor. *Journal of Biological Chemistry*. **277**: 43410-43416.

Nicholas, G., M. Thomas, B. Langley, W. Somers, K. Patel, C. F. Kemp, M. Sharma and R. Kambadur (2002). Titin-Cap Associates with, and Regulates Secretion of, Myostatin. *Journal of Cellular Physiology*. **193**: 120-131.

Nogalska, A., C. D'Agostino, W. K. Engel and V. Askanas (2008). Resveratrol, a Polyphenol Found in Red Wine, Reduces NF κ B-Activation and Myostatin in Endoplasmic-Reticulum-Stress (ERS)-Provoked Cultured Human Muscle Fibers (CHMFs): Relevance to Treatment of Sporadic Inclusion Body Myositis (s-IBM). *Annals of Neurology*. **64**: S9.

Nogalska, A., S. Wojcik, W. K. Engel, J. McFerrin and V. Askanas (2007). Endoplasmic Reticulum Stress Induces Myostatin Precursor Protein and NF- κ B in Cultured Human Muscle Fibers: Relevance to Inclusion Body Myositis. *Experimental Neurology*. **240**: 610-618.

Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac, C. J. Brown and A. K. Dunker (2003). Predicting Intrinsic Disorder from Amino Acid Sequence. *Proteins: Structure, Function, and Genetics*. **53**: 566-572.

- Ohnishi, S. and K. Takano (2004). Amyloid Fibrils from the Viewpoint of Protein Folding. *Cellular and Molecular Life Sciences*. **61**: 511-524.
- Paschen, W. (2003). Endoplasmic Reticulum: A Primary Target in Various Acute Disorders and Degenerative Diseases of the Brain. *Cell Calcium*. **34**: 365-383.
- Patel, K. and H. Amthor (2005). The Function of Myostatin and Strategies of Myostatin Blockade - New Hope for Therapies Aimed at Promoting Growth of Skeletal Muscle. *Neuromuscular Disorders*. **15**: 117-126.
- Petkova, A. T., Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio and R. Tycko (2002). A Structural Model for Alzheimer's β -Amyloid Fibrils Based on Experimental Constraints from Solid State NMR. *Proceedings of the National Academy of Sciences*. **99**: 16742-16747.
- Picotti, P., G. D. Franceschi, E. Frare, B. Spolaore, M. Zambonin, F. Chiti, P. P. d. Laureto and A. Fontana (2007). Amyloid Fibril Formation and Disaggregation of Fragment 1-29 of Apomyoglobin: Insights into the Effect of pH on Protein Fibrillogenesis. *Journal of Molecular Biology*. **367**: 1237-1245.
- Pope, B. and H. M. Kent (1996). High Efficiency 5 min Transformation of Escherichia coli. *Nucleic Acids Research*. **24**: 536-537.
- Pratico, D. (2008). Evidence of Oxidative Stress in Alzheimer's Disease Brain and Antioxidant Therapy. *Annals of the New York Academy of Sciences*. **1147**: 70-78.
- Prilusky, J., C. E. Felder, T. Zeev-Ben-Mordehai, E. H. Rydberg, O. Man, J. S. Beckmann, I. Silman and J. L. Sussman (2005). FoldIndex: A Simple Tool to Predict Whether a Given Protein Sequence is Intrinsically Unfolded. *Bioinformatics*. **21**: 3435-3438.
- Prusiner, S. B. (1982). Novel Proteinaceous Infectious Particles Cause Scrapie. *Science*. **216**: 136-144.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Moolder, R. Apweiler and R. Lopez (2005). InterProScan: Protein Domains Identifier. *Nucleic Acids Research*. **33**: 116-120.

Rabzelj, S., G. Viero, I. Gutierrez-Aguirre, V. Turk, M. D. Serra, G. Anderluh and E. Zerovnik (2008). Interaction with Model Membranes and Pore Formation by Human Stefin B - Studying the Native and Prefibrillar States. *FEBS Journal*. **275**: 2455-2466.

Rambaran, R. N. and L. C. Serpell (2008). Amyloid Fibrils. Abnormal Protein Assembly. *Prion*. **2**: 112-117.

Receveur-Brechot, V., J. M. Bourhis, V. N. Uversky, B. Canard and S. Longhi (2006). Assessing Protein Disorder and Induced Folding. *Proteins: Structure, Function, and Bioinformatics*. **62**: 24-45.

Reiner, C. K., G. Kada and H. J. Gruber (2002). Quick Measurement of Protein Sulphydryls with Ellman's reagent and with 4,4'-Dithiopyridine. *Analytical and Bioanalytical Chemistry*. **373**: 266-276.

Reumers, J., J. Schymkowitz and F. Rousseau (2009). Using Structural Bioinformatics to Investigate the Impact of Non Synonymous SNPs and Disease Mutations: Scope and Limitations. *BMC Bioinformatics*. **10**: S9.

Rezaei-Ghaleh, N., M. Zweckstetter, D. Morshedi, A. Ebrahim-Habibi and M. Nemat-Gorgani (2009). Amyloidogenic Potential of α -Chymotrypsin in Different Conformational States. *Biopolymers*. **91**: 28-36.

Rider, C. (2006). Heparin/Heparin Sulphate Binding in the TGF- β Cytokine Superfamily. *Biochemical Society Transactions*. **34**: 458-460.

Rifkin, D. B. (2005). Latent Transforming Growth Factor-beta (TGF- β) Binding Proteins: Orchestrators of TGF- β Availability. *The Journal of Biological Chemistry*. **280**: 7409-7412.

- Romero, P., C. Obradovic, C. Kissinger, J. Villafranca and A. Dunker (1997). Identifying Disordered Regions in Proteins from Amino Acid Sequences. *Proceedings of the IEEE International Conference on Neural Networks*, Houston, Texas.
- Rost, B., G. Yachdav and J. Lui (2004). The PredictProtein Server. *Nucleic Acids Research*. **32**: 321-326.
- Sabaté, R., I. Lascu and S. J. Saupe (2008). On the Binding of Thioflavin-T to HET-s Amyloid fibrils Assembled at pH 2. *Journal of Structural Biology*. **162**: 387-396.
- Sathasivam, S. and P. J. Shaw (2005). Apoptosis in Amyotrophic Lateral Sclerosis - What is the Evidence? *The Lancet Neurology*. **4**: 500-509.
- Sathishkumar, K., X. Gao, A. Raghavamenon, S. Murthy, P. Kadowitz and R. Uppu (2010). Determination of Glutathione, Mitochondrial Transmembrane Potential, and Cytotoxicity in H9c2 Cardiomyoblasts Exposed to Reactive Oxygen and Nitrogen Species. *Methods in Molecular Biology*. **610**: 51-61.
- Sawaya, M. R., S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen, C. Riek and D. Eisenberg (2007). Atomic Structures of Amyloid Cross- β Spines Reveal Varied Steric Zippers. *Nature*. **447**: 453-457.
- Scheufler, C., W. Sebold and M. Hülsmeyer (1999). Crystal Structure of Human Bone Morphogenic Protein-2 at 2.7 Å Resolution. *Journal of Molecular Biology*. **287**: 103-115.
- Schlapschy, M., S. Grimm and A. Skerra (2006). A System for Concomitant Overexpression of Four Periplasmic Folding Catalysts to Improve Secretory Protein Production in Escherichia coli. *Protein Engineering Design and Selection*. **19**: 385-390.
- Schreuder, H., A. Liesum, J. Pohl, M. Kruse and M. Koyama (2005). Crystal Structure of Recombinant Human Growth and Differentiation Factor 5: Evidence for

Interaction of the Type I and Type II Receptor-Binding Sites. *Biochemical and Biophysical Research Communications*. **329**: 1076-1086.

Schroder, M. (2008). Endoplasmic Reticulum Stress Responses. *Cellular and Molecular Life Sciences*. **65**: 862-894.

Schuelke, M., K. R. Wagner, L. E. Stolz, C. Hubner, T. Riebel, W. Komen, T. Braun, J. F. Tobin and S. J. Lee (2004). Myostatin Mutation Associated with Gross Muscle Hypertrophy in a Child. *New England Journal of Medicine*. **350**: 2682-2688.

Sengle, G., N. L. Charbonneau, R. N. Ono, T. Sasaki, J. Alvarez, D. R. Keene, H. P. Bachinger and L. Y. Sakai (2008). Targeting of Bone Morphogenetic Protein Growth Factor Complexes to Fibrillin. *Journal of Biological Chemistry*. **283**: 13874-13888.

Sengle, G., R. N. Ono, K. M. Lyons, H. P. Bächinger and L. Y. Sakai (2008). A New Model for Growth Factor Activation: Type II Receptors Compete with the Prodomain for BMP-7. *Journal of Molecular Biology*. **381**: 1025-1039.

Sha, X., L. Yang and L. Gentry (1991). Identification and Analysis of Discrete Functional Domains in the Pro Region of Pre-Pro-Transforming Growth Factor β 1. *The Journal of Cell Biology*. **114**: 827-839.

Sharma, M., R. Kambadur, K. G. Matthews, W. G. Somers, G. P. Devlin, J. V. Conaglen, P. J. Fowke and J. J. Bass (1999). Myostatin, a Transforming Growth Factor- β Superfamily Member, is Expressed in Heart Muscle and is Upregulated in Cardiomyocytes After Infarct. *Journal of Cellular Physiology*. **180**: 1-9.

Siriett, V., M. S. Salerno, C. Berry, G. Nicholas, R. Bower, R. Kambadur and M. Sharma (2007). Antagonism of Myostatin Enhances Muscle Regeneration During Sarcopenia. *Molecular Therapy*. **15**: 1463-1470.

Slotta, U., S. Hess, K. Spiess, T. Stromer, L. Serpell and T. Schiebel (2007). Spider Silk and Amyloid Fibrils: A Structural Comparison. *Macromolecular Bioscience*. **7**: 183-188.

Smith, B. J. (2002). Chemical Cleavage of Proteins at Aspartyl-X Peptide Bonds. *The Protein Protocols Handbook*, Humana Press. **4**: 499-502.

Smith, D. P. and S. E. Radford (2001). Role of the Single Disulphide Bond of β 2-Microglobulin in Amyloidosis In Vitro. *Protein Science*. **10**: 1775-1784.

Soto, C. (2003). Unfolding the Role of Protein Misfolding in Neurodegenerative Diseases. *Nature Reviews Neuroscience*. **4**: 49-60.

Souza, T. A., X. Chen, Y. Guo, P. Sava, J. Zhang, J. J. Hill, P. J. Yaworsky and Y. Qiu (2008). Proteomic Identification and Functional Validation of Activins and BMP-11 as Candidate Novel Muscle Mass Regulators. *Molecular Endocrinology*. **22**: 2689-2702.

Stefani, M. and C. M. Dobson (2003). Protein Aggregation and Aggregate Toxicity: New Insights into Protein Folding, Misfolding Diseases and Biological Evolution. *Journal of Molecular Medicine*. **81**: 678-699.

Stromer, T. and L. C. Serpell (2005). Structure and Morphology of the Alzheimer's Amyloid Fibril. *Microscopy Research and Technique*. **67**: 210-217.

Sunde, M., L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys and C. C. F. Blake (1997). Common Core Structure of Amyloid Fibrils by Synchrotron X-Ray Diffraction. *Journal of Molecular Biology*. **273**: 729-739.

Tabner, B. J., O. M. A. El-Agnaf, M. J. German, N. J. Fullwood and D. Allsop (2005). Protein Aggregation, Metals and Oxidative Stress in Neurodegenerative Diseases. *Biochemical Society Transactions*. **33**: 1082-1086.

Tabner, B. J., O. M. A. El-Agnaf, S. Turnbull, M. J. German, K. E. Paleologou, Y. Hayashi, L. J. Cooper, N. J. Fullwood and D. Allsop (2005). Hydrogen Peroxide is Generated During the Very Early Stages of Aggregation of the Amyloid Peptide Implicated in Alzheimer Disease and Familial British Dementia. *The Journal of Biological Chemistry*. **280**: 35789-35792.

Taylor, J. P., F. Tanaka, K. Robitschek, C. M. Sandoval, A. Taye, S. Markovic-Plese and K. H. Fischbeck (2003). Aggresomes Protect Cells by Enhancing the Degradation of Toxic Polyglutamine-Containing Protein. *Human Molecular Genetics*. **1**: 749-757.

Taylor, W. E., S. Bhasin, J. Artaza, F. Byhower, M. Azam, D. H. Willard, F. C. Kull and N. Gonzalez-Cadavid (2001). Myostatin Inhibits Cell Proliferation and Protein Synthesis in C2C12 Muscle Cells. *American Journal of Physiology - Endocrinology and Metabolism*. **280**: 221-228.

Thies, R. S., T. Chen, M. V. Davies, K. N. Tomkinson, A. A. Pearson, Q. A. Shakey and N. M. Wolfman (2001). GDF-8 Propeptide Binds to GDF-8 and Antagonizes Biological Activity by Inhibiting GDF-8 Receptor Binding. *Growth Factors*. **18**: 251-259.

Thomas, M., B. Langley, C. Berry, M. Sharma, S. Kirk, J. Bass and R. Kambadur (2000). Myostatin, a Negative Regulator of Muscle Growth, Functions by Inhibiting Myoblast Proliferation. *The Journal of Biological Chemistry*. **275**: 40235-40243.

Thompson, J., D. Higgins and T. Gibson (1994). CLUSTALW: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research*. **22**: 4673-4680.

Thompson, T. B., T. F. Lerch, R. W. Cook, T. K. Woodruff and T. S. Jardetzsky (2005). The Structure of the Follistatin:Activin Complex Reveals Antagonism of Both Type I and Type II Receptor Binding. *Developmental Cell*. **9**: 535-543.

Thompson, T. B., T. K. Woodruff and T. S. Jardetzsky (2003). Structures of an ActRIIB:Activin Complex Reveal a Novel Binding Mode for TGF- β Ligand-Receptor Interactions. *The EMBO Journal*. **22**: 1555-1566.

Tobin, J. F. and A. J. Celeste (2005). Myostatin, a Negative Regulator of Muscle Mass: Implications for Muscle Degenerative Diseases. *Current Opinion in Pharmacology*. **5**: 328-332.

Tompa, P. (2009). Structural Disorder in Amyloid Fibrils: Its Implication in Dynamic Interactions of Proteins. *FEBS Journal*. **276**: 5406-5415.

Trovato, A., F. Chiti, A. Maritan and F. Seno (2006). Insight into the Structure of Amyloid Fibrils from the Analysis of Globular Proteins. *PLoS Computational Biology*. **2**: e170.

Trovato, A., F. Seno and S. C. E. Tosatto (2007). The PASTA Server for Protein Aggregation Prediction. *Protein Engineering Design and Selection*. **20**: 521-523.

Tsumoto, K., D. Ejima, I. Kumagai and T. Arakawa (2003). Practical Considerations in Refolding Proteins from Inclusion Bodies. *Protein Expression and Purification*. **28**: 1-8.

Vallejo, L. F. and U. Rinus (2004). Optimized Procedure for Renaturation of Recombinant Human Bone Morphogenic Protein-2 at High Protein Concentration. *Biotechnology and Bioengineering*. **85**: 601-609.

Vattemi, G., A. Nogalska, W. K. Engel, C. D'Agostino, F. Checler and V. Askanas (2009). Amyloid- β 42 is Preferentially Accumulated in Muscle Fibers of Patients with Sporadic Inclusion Body Myositis. *Acta Neuropathology*. **117**: 569-574.

Venkataraman, G., V. Sasisekharan, C. L. Cooney, R. Langer and R. Sasisekharan (1995). Complex Flexibility of the Transforming Growth Factor- β Superfamily. *Proceedings of the National Academy of Sciences*. **92**: 5406-5410.

Vestergaard, B., M. Groenning, M. Roessle, J. S. Kastrup, M. v. d. Weert, J. M. Flink, S. Frokjaer, M. Gajhede and D. I. Svergun (2007). A Helical Structural Nucleus is the Primary Elongating Unit of Insulin Amyloid Fibrils. *PLoS Biology*. **5**: 1089-1097.

Vincentelli, R., S. Canaan, V. Campanacci, C. Valencia, D. Maurin, F. Frassineti, L. Scappucini-Calvo, Y. Bourne, C. Cambillau and C. Bignon (2004). High-Throughput Automated Refolding Screening of Inclusion Bodies. *Protein Science*. **13**: 2782-2792.

Vitt, U. A. and A. J. W. Hsueh (2001). Stage-Dependent Role of Growth Differentiation Factor-9 in Ovarian Follicle Development. *Molecular and Cellular Endocrinology*. **183**: 171-177.

Volles, M. and P. Lansbury (2003). Zeroing in on the Pathogenic Form of α -Synuclein and its Mechanism of Neurotoxicity in Parkinson's Disease. *Biochemistry*. **42**: 7871-7878.

Walsh, D. M., D. M. Hartley, Y. Kusumoto, Y. Fezoui, M. M. Condron, A. Lomakin, G. B. Benedek, D. J. Selkoe and D. B. Teplow (1999). Amyloid β -Protein Fibrillogenesis. Structure and Biological Activity of Protofibrillar Intermediates. *The Journal of Biological Chemistry*. **274**: 25945-25952.

Walsh, J. G., S. P. Cullen, C. Sheridan, A. U. Luthi, C. Gerner and S. J. Martin (2008). Executioner Caspase-3 and Caspase-7 are Functionally Distinct Proteases. *Proceedings of the National Academy of Sciences*. **105**: 12815-12819.

Walsh, P., J. Yau, K. Simonetti and S. Sharpe (2009). Morphology and Secondary Structure of Stable β -Oligomers Formed by Amyloid Peptide PrP(106-126). *Biochemistry*. **48**: 5779-5781.

Walton, K. L., Y. Makanji, M. C. Wilce, K. L. Chan, D. M. Robertson and C. A. Harrison (2009). A Common Biosynthetic Pathway Governs the Dimerisation and Secretion of Inhibin and Related TGF- β Ligands. *The Journal of Biological Chemistry*. **284**: 9311-9320.

Wang, S. S.-S., K.-N. Liu and Y.-C. Lu (2009). Amyloid Fibrillation of Hen Egg-White Lysozyme is Inhibited by TCEP. *Biochemical and Biophysical Research Communications*. **381**: 639-642.

Ward, J., J. Sodhi, L. McGuffin, B. Buxton and D. Jones (2004). Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology*. **337**: 635-645.

Winnefeld, M., A. Grewinig, M. Schnolzer, H. Spring, T. A. Knoch, E. C. Gan, J. Rommelaere and C. Cziepluch (2006). Human SGT Interacts with Bag-6/Bat-3/Scythe and Cells with Reduced Levels of Either Protein Display Persistence of Few Misaligned Chromosomes and Mitotic Arrest. *Experimental Cell Research.* **312:** 2500-2514.

Wipff, P. J. and B. Hinz (2008). Integrins and the Activation of Latent Transforming Growth Factor β 1 - An Intimate Relationship. *European Journal of Cell Biology.* **87:** 601-615.

Wojcik, S., W. K. Engel, J. McFerrin and V. Askanas (2005). Myostatin is Increased and Complexes with Amyloid- β within Sporadic Inclusion-Body Myositis Muscle Fibers. *Acta Neuropathology.* **110:** 173-177.

Wojcik, S., A. Nogalska, J. McFerrin, W. K. Engel, G. Oledzka and V. Askanas (2007). Myostatin Precursor Protein is Increased and Associates with Amyloid- β Precursor Protein in Inclusion Body Myositis Model. *Neuropathology and Applied Neurobiology.* **33:** 238-242.

Wolfman, N. M., A. C. McPherron, W. N. Pappano, M. V. Davies, K. Song, K. N. Tomkinson, J. F. Wright, L. Zhao, S. M. Sebald, D. S. Greenspan and S.-J. Lee (2003). Activation of Latent Myostatin by the BMP-1/Tolloid Family of Metalloproteinases. *Proceedings of the National Academy of Sciences.* **100:** 15842-15846.

Wong, Y., T. Lee, H. Liang, C. Huang, Y. Yang, C. Chu, H. Huang, M. Ko and J. Hwang (2007). KinasePhos 2.0: A Web Server for Identifying Kinase-Specific Phosphorylation Sites Based on Sequences and Coupling Patterns. *Nucleic Acids Research.* **35:** 588-94.

Wuthrich, K. (1990). Protein Structure Determination in Solution by NMR Spectroscopy. *The Journal of Biological Chemistry.* **265:** 22059-22062.

Yamamoto, K., H. Yagi, D. Ozawa, K. Sasahara, H. Naiki and Y. Goto (2008). Thiol Compounds Significantly Inhibit the Formation of Amyloid Fibrils by β_2 -Microglobulin at Neutral pH. *Journal of Molecular Biology*. **376**: 258-268.

Yang, J. Z. and B. P. Zhao (2006). Postnatal Expression of Myostatin Propeptide cDNA Maintained High Muscle Growth and Normal Adipose Tissue Mass in Transgenic Mice Fed a High-Fat Diet. *Molecular Reproduction and Development*. **73**: 462-469.

Yang, W., Y. Chen, X. Zhang, X. Y. Wang, N. Yang and D. H. Zhu (2006). Extracellular Signal-Regulated Kinase 1/2 Mitogen-Activated Protein Kinase Is Involved in Myostatin-Regulated Differentiation Repression. *Cancer Research*. **66**: 1320-1326.

Yang, W., Y. Zhang, Y. Li, Z. Wu and D. Zhu (2006). Myostatin Induces Cyclin D1 Degradation to Cause Cell Cycle Arrest through a Phosphatidylinositol 3-Kinase/AST/GSK-3 β Pathway and Is Antagonized by IGF-1. *The Journal of Biological Chemistry*. **282**: 3799-3808.

Yang, W. Y., E. Larios and M. Gruebele (2003). On the Extended beta-Conformation Propensity of Polypeptides at High Temperature. *Journal of the American Chemical Society*. **125**: 16220-16227.

Zimmers, T. A., M. V. Davies, L. G. Koniaris, P. Haynes, A. F. Esquela, K. N. Tomkinson, A. C. McPherron, N. M. Wolfman and S. J. Lee (2002). Induction of Cachexia in Mice by Systemically Administered Myostatin. *Science*. **296**: 1486-1488.

Cytotoxic Aggregation and Amyloid Formation by the Myostatin Precursor Protein

Carlene S. Starck, Andrew J. Sutherland-Smith*

Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand

Abstract

Myostatin, a negative regulator of muscle growth, has been implicated in sporadic inclusion body myositis (sIBM). sIBM is the most common age-related muscle-wastage disease with a pathogenesis similar to that of amyloid disorders such as Alzheimer's and Parkinson's diseases. Myostatin precursor protein (MstnPP) has been shown to associate with large molecular weight filamentous inclusions containing the Alzheimer's amyloid beta peptide in sIBM tissue, and MstnPP is upregulated following ER stress. The mechanism for how MstnPP contributes to disease pathogenesis is unknown. Here, we show for the first time that MstnPP is capable of forming amyloid fibrils *in vitro*. When MstnPP-containing *Escherichia coli* inclusion bodies are refolded and purified, a proportion of MstnPP spontaneously misfolds into amyloid-like aggregates as characterised by electron microscopy and binding of the amyloid-specific dye thioflavin T. When subjected to a slightly acidic pH and elevated temperature, the aggregates form straight and unbranched amyloid fibrils 15 nm in diameter and also exhibit higher order amyloid structures. Circular dichroism spectroscopy reveals that the amyloid fibrils are dominated by β -sheet and that their formation occurs via a conformational change that occurs at a physiologically relevant temperature. Importantly, MstnPP aggregates and protofibrils have a negative effect on the viability of myoblasts. These novel results show that the myostatin precursor protein is capable of forming amyloid structures *in vitro* with implications for a role in sIBM pathogenesis.

Citation: Starck CS, Sutherland-Smith AJ (2010) Cytotoxic Aggregation and Amyloid Formation by the Myostatin Precursor Protein. PLoS ONE 5(2): e9170. doi:10.1371/journal.pone.0009170

Editor: Ashley M. Buckle, Monash University, Australia

Received November 1, 2009; **Accepted** January 19, 2010; **Published** February 11, 2010

Copyright: © 2010 Starck, Sutherland-Smith. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by Tertiary Education Commission Top Achiever Doctoral and New Zealand Neuromuscular Alliance Henry Kelsey Scholarships to C.S.S. and grants from the Massey University Research Fund and the Palmerston North Medical Research Foundation to A.J.S.-S.. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: A.J.Sutherland-Smith@massey.ac.nz

Introduction

Myostatin is a member of the transforming growth factor-beta (TGF- β) superfamily of growth and differentiation factors and is a primary regulator of muscle growth both pre- and postnatally, primarily via inhibition of myoblast proliferation and differentiation [1–4]. Like other members of the family, myostatin is translated as a precursor protein (MstnPP) that consists of an N-terminal signal sequence, a regulatory propeptide domain (residues 21–266) and a growth factor domain (residues 267–374) which dimerises at the C-terminus via an inter-molecular disulfide bond [5–7]. The mature growth factor dimer is cleaved from the propeptide region by furin convertase proteolysis in the endoplasmic reticulum (ER) at a conserved RSRR sequence [2,7]. The propeptide region plays at least two important functions. First, as a chaperone in the ER to assist in the folding of the growth factor region [8,9] that contains the intricate cysteine-knot motif characteristic of the family [10,11]. Second, the N-terminal propeptide plays a regulatory role after cleavage, remaining non-covalently associated with the mature dimer to form a latent complex which is exported from the cell [7]. Myostatin remains latent until a second cleavage event immediately N-terminal to aspartate 76 of the propeptide region, most probably by metalloproteinases, that disrupts the association [12,13]. It is possible that furin cleavage of MstnPP also occurs post-secretion with a pool of extracellular MstnPP identified in skeletal muscle

[14]. The mature growth factor dimer is structurally similar to other members of the TGF- β family [11]; however the structural characteristics of MstnPP or the propeptide region remain undetermined.

Signalling by the myostatin growth factor via activin and TGF- β receptors [2] ultimately results in cell-cycle arrest through the upregulation of genes involved in cell-cycle withdrawal such as p21 and p53 and the downregulation of myogenic regulatory factors such as MyoD and myogenin [15–18]. Although postnatally this action maintains the quiescence of muscle satellite cells, the prenatal role is more complex and depends on the environmental context during development, with signalling by myostatin ensuring that a balance between proliferation and differentiation is maintained [19].

Myostatin-null mutations have been identified in dogs, cattle and sheep, resulting in a double-muscled phenotype [1,20,21] and an exceptionally muscular and strong human lacking functional myostatin protein was also recently identified [22,23]. Myostatin overexpression in mice induces profound muscle and fat loss analogous to that seen in human cachexia syndromes [4] and ectopically expressed myostatin rapidly lowers muscle mass in rats [24]. Myostatin signalling can have negative consequences in a diseased background such as the muscular dystrophies [25] and may contribute to cachexia associated with many chronic disease states [4] including HIV [26] and cancer [27]. For these reasons, since its discovery in 1997 [3], the processed myostatin growth

factor dimer has been suggested to hold exciting potential for inhibitory targeting in a wide range of muscle wastage diseases [28,29].

Less focus had been placed on the involvement of MstnPP or the propeptide region in disease until a role for MstnPP in the pathogenesis of sporadic inclusion body myositis (sIBM) was proposed [30,31]. sIBM is the most common progressive muscle wastage disease associated with aging where progressive muscle loss leads to severe atrophy and weakness. Although the pathogenesis is unknown, it is likely that oxidative damage contributes to aging of the muscle fibers [30]. Endoplasmic reticulum (ER) stress and the unfolded protein response (UPR) have been demonstrated in sIBM muscle fibers [30,32]. Inflammation and amyloid formation appear to be predominant features but whether these are causally related and which is the primary cause of sIBM, remain matters of debate [30,33]. The presence of fibrillar inclusions in some diseased tissue suggests that sIBM may be an amyloid disease, where a prominent feature is protein aggregation and the subsequent formation and deposition of large amyloid fibrils analogous to those observed for neurodegenerative disorders such as Alzheimer's, Parkinson's and Huntington's diseases, the spongiform encephalopathies, the systemic amyloidoses and type II diabetes [34,35]. Despite extensive research, the mechanisms behind amyloid formation and how they contribute to disease remain poorly understood. It is well-documented that fibril formation occurs via one or more oligomeric intermediate stages [36–40] and it appears that an intermediate oligomeric species, rather than the mature fibril, is the mediator of cyto- and neurotoxicity [41–43]. Amyloid fibrils, characterised by extensive β -sheet secondary structure, are deposited in tissues and are resistant to degradation but may actually be chemically inert [35].

The filamentous inclusions of sIBM tissue contain a number of molecules normally alien to muscle fibers, predominantly the amyloid beta (A β) protein of Alzheimer's disease [44]. Biopsies from sIBM muscle fibers revealed a close association of the myostatin precursor protein (MstnPP) with A β aggregates and increased levels of both MstnPP and the processed myostatin growth factor dimer [31]. Furthermore, when muscle cells are placed in conditions that result in endoplasmic reticulum (ER) stress, MstnPP is upregulated and an sIBM-like pathology is observed [45]. ER stress may lead to the misfolding and aggregation of MstnPP. Cultured human muscle fibers (CHMFs) overexpressing A β precursor protein showed an increase in MstnPP protein levels and subsequent experimental inhibition of the proteasome caused co-accumulation of MstnPP and A β PP/A β within aggresomes [46]. CHMFs subjected to tunicamycin or thapsigargin, compounds that inhibit protein glycosylation and disrupt ER calcium levels respectively inducing ER stress, also showed an increase in MstnPP expression (in addition to ER stress marker proteins GRP78 and Herp), possibly due to the activation of NF- κ B [45]. In contrast to proteasome inhibition, where MstnPP mRNA levels were decreased, NF- κ B activation increases transcription of MstnPP mRNA in a stressed situation. These differences may be explained by the observation that the nature and duration of ER stress may be important in determining whether a cell activates a protective pathway or enters apoptosis [47,48]. An additional mechanism for how myostatin/MstnPP contributes to sIBM is focused on increased activity by the myostatin growth factor. As myostatin is a procachectic growth factor postnatally, increased production and processing of MstnPP may result in increased levels of circulating myostatin growth factor and signalling that leads to atrophy [45,46,49], though there is little evidence supporting this hypothesis. A number of peptide hormones, such as those in secretory granules of the endocrine system, have been found to aggregate into amyloids in secretory

pathways in a regulated, functional manner [50,51], so there is a possibility that myostatin aggregates are protective in the sIBM diseased state.

The role of MstnPP in sIBM has not been firmly established. Our research shows for the first time that MstnPP is able to form amyloid fibrils independently *in vitro* via a pathway that produces cytotoxic intermediates. When MstnPP is refolded and purified from *E. coli* inclusion bodies, a population of MstnPP spontaneously misfolds into amyloid protofibril-like aggregates that can be visualised using electron microscopy and bind the amyloid-specific dye thioflavin T. When subjected to a slightly acidic pH and elevated temperature, the protofibril-like aggregates form straight and unbranched amyloid fibrils that are 15 nm in diameter and exhibit higher order amyloid structures such as twists and bundles. The fibrils are resistant to trypsin digest after incubation overnight at 37°C. Circular dichroism spectroscopy reveals that MstnPP undergoes a number of structural transitions through aggregation to amyloid protofibril and fibril formation that initiate at a physiologically relevant temperature. Significantly, MstnPP aggregates and protofibrils have a negative effect on the viability of C2C12 mouse myoblasts when added to the culture medium. These results support a role for MstnPP in the pathogenesis of sIBM and present the hypothesis that this role may be within a degenerative context.

Results

MstnPP Is Predicted to Be Prone to Aggregation and Disorder *In Silico*

In silico programs have been shown to accurately predict the regions of a polypeptide chain that are prone to β -sheet aggregation and amyloid formation. The MstnPP amino acid sequence was analysed for β -aggregation propensity (Fig. 1A) using Tango [52] (Fig. S1A) and PASTA [53,54] (Fig. S1B) and for amyloidogenic regions using Waltz [55] (Fig. S1C). Both Tango and PASTA predicted a high propensity for β -aggregation for residues within the pro-peptide region, 154–177 (Fig. 1A); residues 139–153 were also predicted by PASTA to have β -aggregation tendencies, albeit at a comparatively lower level. Tango also identified growth factor domain sequences, 314–321 and 347–355. Waltz predicted the same propeptide amyloidogenic regions, 153–162 and 168–177 as well as the very C-terminus of the protein, 348–364. Mapping of aggregation-prone regions of the myostatin growth factor onto the published crystal structure [11] shows that residues 348–364 localise to the β -sheet 'fingers' of the growth factor (Fig. 1B).

In Vitro Refolding of MstnPP Results in the Production of Large Soluble Aggregates

Human MstnPP is found in inclusion bodies when overexpressed in *E. coli* like other TGF- β family members [56,57] and can be refolded *in vitro* using a modification of methods described for zebra-fish myostatin [8]. After refolding, in addition to the native disulfide-bonded MstnPP dimer, two other species were apparent by non-reducing SDS-PAGE; misfolded monomer and a very high molecular weight species that has been characterised previously as soluble aggregates [8,9] (Fig. 2A), which are frequently observed during protein refolding [58]. Native MstnPP was purified by heparin affinity chromatography that separated the correctly folded MstnPP dimer from the majority of the aggregates followed by the removal of remaining aggregates by gel filtration. During heparin chromatography, the correctly folded dimer eluted at 200 mM NaCl (H1) and the majority of aggregates as a large peak at 600 mM NaCl (H2) (Fig. 2B and Fig. S2A).

A

21VDLNENSEQKENVEKEGLCNACTWRQNTKSSRERAIKIQILSKRLR
 ETAPNISKDVRQLLPKAPPLRELIQDYDVQRDDSSGSLEDDDDYHATTETIITMPT
 ESDFLMQVDGKPKCCF**FKFSSKIQYNKVVK**AQLWYL^YLRP**VETPTTV**FVQILRLIK
 MKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNLQKQPESNLGEIKALDEN
 HDLAVTFPGPGEDGLNPFLLEVKTDTPKRS**R**RD**F**GLC**D**E**H**TESRC**C**RYPLTV
 DFEAFGWWDWIAPKRYKANYCSGECEFVFLQKYPHTHLVHQANPRGSAGPCCTP
 TKMS**PINMLYFNGKEQIIYGKIP**AMVVDRCGCS₃₇₅

B

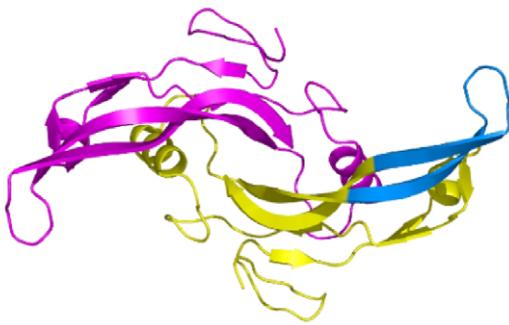


Figure 1. *In silico* predictions of propensity for β -sheet aggregation (Tango and PASTA) and amyloid formation (Waltz). A. MstnPP sequence 21–375 showing regions of elevated propensity at two levels; red regions have higher propensity than those in blue. Regions where Waltz, Tango and PASTA predictions overlap are underlined. The RSRR furin proteolysis sequence that separates the propeptide from the growth factor domain is shown in bold. B. Mapping the 348–364 sequence (blue) onto the myostatin growth factor crystal structure. The region has been shown on one monomer only.

doi:10.1371/journal.pone.0009170.g001

Western blotting identifies both peaks as myostatin (Fig. S2B) but the two species interact differently with the heparin column suggesting a difference in folding. Further purification of both peaks was achieved using gel filtration (Fig. 2C). The aggregates are too large to move from the well into an SDS-PAGE gel (above 250 kDa, Fig. 2A), elute in the void volume during S200 gel filtration (PA) (Fig. 2C) and remain soluble even after ultracentrifugation at 214,000 $\times g$ (data not shown). Since the MstnPP monomer is 43 kDa the soluble aggregates are most likely to be an oligomer of at least 5 monomeric units.

MstnPP Soluble Aggregates Have Characteristics of Amyloid-Like Protofibrils

Transmission electron microscopy (TEM, Fig. 3A and B) revealed that MstnPP soluble aggregates exhibit a morphology and size similar to that documented for amyloid protofibrils from a number of other proteins such as lysozyme [59,60] and HypF-N [61] as well as the insulin protofibrils generated as a positive control (Fig. 3C) using a well-established protocol [39,62,63]. A range of structures, including spheres and oligomers of associated spheres (diameter 15.6 \pm 1.6 nm SD), can be seen, and examples of increased elongation of the aggregates are apparent (Fig. 3B).

A defining characteristic of amyloid fibrils and protofibrils is their ability to bind the fluorescent dye thioflavin T (ThT) [63–65]. When bound to fibrils, ThT can be excited at 450 nm to fluoresce at 485 nm [66]. One limitation of this method is the lack of a strict quantitative relationship [64] with amyloid fibrils often producing a stronger emission than protofibrils by comparison [65]. ThT

binding assays were carried out on the MstnPP aggregates and compared to a solution containing a mixture of insulin protofibrils and fibrils as a positive control. The MstnPP soluble aggregates bound ThT with intensity comparable to insulin fibrils whereas the correctly folded MstnPP dimer did not (Fig. 3D). These results suggest that a population of MstnPP can aggregate spontaneously to form amyloid-like protofibrils.

Formation of MstnPP Amyloid Fibrils at Acidic pH and Elevated Temperature

The presence of MstnPP amyloid-like protofibrils suggests that amyloid fibril formation is also possible. MstnPP aggregates were concentrated and resuspended in dilute HCl solutions (pH range 1.6 to 6.3) and incubated at either 37° or 60°C, conditions including those under which insulin forms amyloid fibrils [62,63]. Solutions were monitored by ThT binding and TEM (Fig. 4). After one week at 60°C, ThT fluorescence had increased significantly in the pH 5.3 solution (Fig. 4A), which contained a mixture of prefibrillar aggregates and amyloid fibrils observed by TEM (Fig. 4B–4F) in both the presence and absence of 0.1% sodium azide. The fluorescence assays were also performed by diluting protein solutions into a pH 7.5 buffer since ThT binding may be affected by changes in pH [67]. In these assays, fluorescence values for the soluble aggregates were largely unchanged but an increase in intensity for the pH 5.3 samples was observed (data not shown), consistent with the lower pH results.

Myostatin fibril formation occurs via a number of stages. For a 3 mg/mL solution, overnight incubation at 60°C in pH 5.3 produces oligomers that are compact and have a ‘bead-on-a-string’ morphology (Fig. 4B) as observed for A β and lysozyme fibrils [60,65]. By three days, extensive arrays of aggregation were present; these are often three-dimensional, appearing as areas of high electron density under the TEM (Fig. 4C*i*). Increased elongation and in some places lateral association and longitudinal fusion of oligomers can be observed (Fig. 4C*ii* and 4C*iii*). This granular to smooth transition has been shown during A β fibril formation [65]. Smaller pore-like oligomers, previously documented on the amyloid formation pathway of a number of proteins [40,43], are also apparent (Fig. 4C*iv* and 4C*v*). By one week, fibrils that show the characteristic morphology of amyloid had appeared, with a diameter of 15.4 \pm 0.7 nm (SD) and a straight, unbranching structure (Fig. 4D). The fibrils are extremely long, in excess of 5 μ m. At two weeks, higher order amyloid structures can be observed, such as two (Fig. 4E) or more (Fig. 4F) fibrils twisting around each other.

Secondary Structure Analysis of MstnPP Fibril Formation by Circular Dichroism Spectroscopy

Amyloid formation is accompanied by the adoption of a β -sheet rich secondary structure regardless of the structure of the native protein [68]. Circular dichroism (CD) spectroscopy was used to study the structural changes that occur during amyloid formation (Fig. 5). Correctly refolded MstnPP has a CD spectrum indicative of a mixture of α -helix and β -sheet, in agreement with the crystal structure of the mature myostatin growth factor [11] and the mature regions of other TGF- β family members [57]. There are two minima, the stronger at 218 nm and a second at 208 nm which are characteristic of β -sheet and α -helical structures respectively [69]. The absolute negative value at 200 nm implies a degree of intrinsic disorder exists in the structure [65,70], which may be functionally significant due to the different roles suggested for the propeptide region of the myostatin precursor, requiring flexibility in the N-terminal domain. To our knowledge, this is the

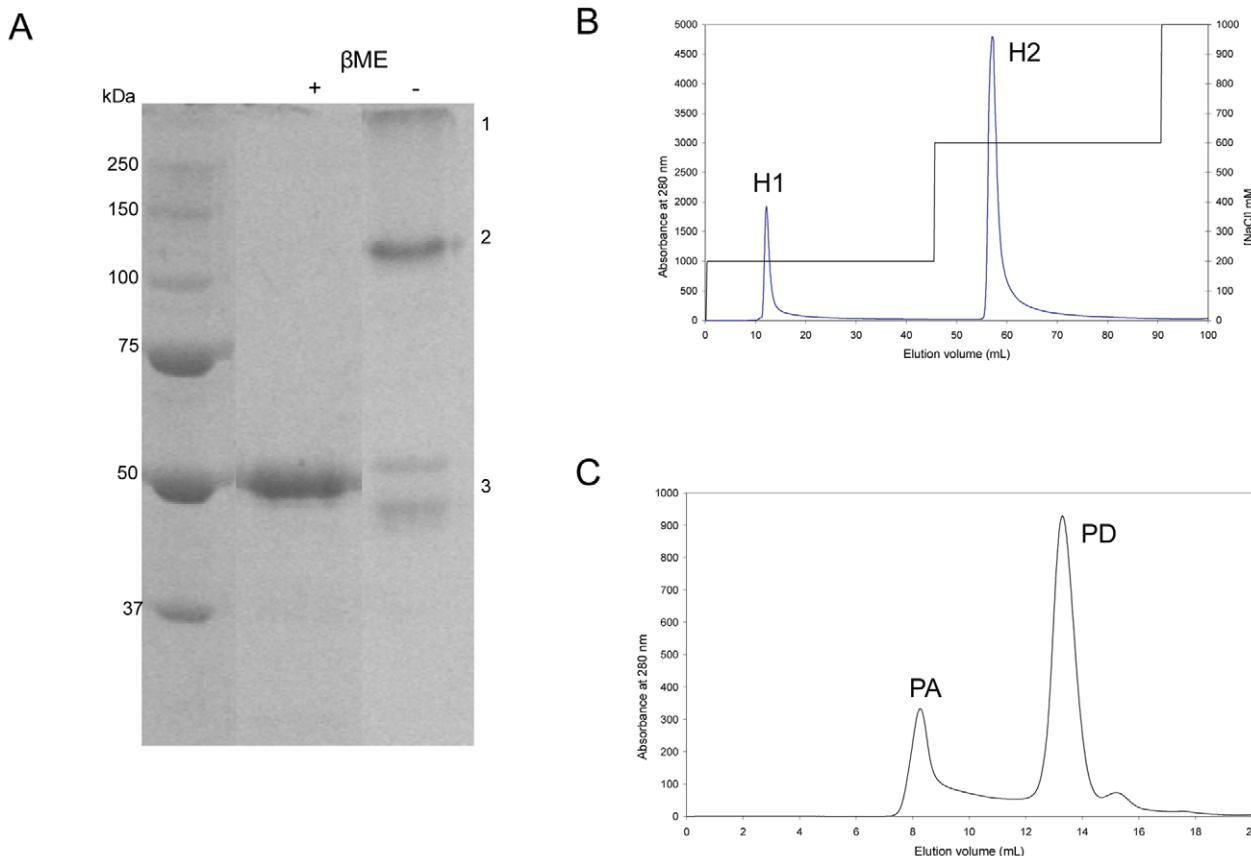


Figure 2. Refolding and purification of human MstnPP. A. Reducing (+ β-mercaptoethanol (βME)) vs non-reducing (−βME) SDS-PAGE (12%) after refolding. β-ME concentration is 2 M. Bands are as indicated: 1. soluble aggregates; 2. disulfide-bonded dimer; 3. monomer either fully reduced (+β-ME) or misfolded at various stages of reduction (−β-ME). Figure assembled from one gel where intervening lanes were removed. B. Heparin affinity chromatography separates the majority of soluble aggregates (H2) from correctly-folded dimer (H1). C. S200 gel filtration chromatography; the MstnPP dimer elutes at 13.5 mL (PD) and the soluble aggregates in the void volume at 8 mL (PA).

doi:10.1371/journal.pone.0009170.g002

first structural analysis for any TGF- β family precursor protein. Quite a different profile is observed for the MstnPP soluble aggregates with the very strong minimum at 208 nm indicating that the secondary structure is primarily α -helical in nature. Broadening of the spectrum between 215–225 nm may arise from contributions of β -sheet and α -helix to the spectra. As the CD spectra of protofibrils of other amyloid-forming proteins is characterised by β -sheet absorption at 215–218 nm [61,71,72], it is likely that the MstnPP soluble aggregates represent an oligomeric intermediate state prior to protofibril formation rather than protofibrils. Incubation of the aggregates overnight at pH 5.3 and 60°C produces a CD spectrum dominated by β -sheet where the α -helix minimum at 208 nm is completely replaced by a minimum at 218 nm. TEM images of the same sample (Fig. 4) reveals there are no fibrils at this stage but the aggregates are more compact and have begun to aggregate into large arrays, which may represent the start of β -sheet stacking as amyloid protofibrils. The ellipticity for the spectrum is comparatively low, similar to that observed in the fibrilllogenesis of A β [65]. This reduction of signal is most likely due to differential absorption flattening, a phenomenon often seen in the CD spectra of samples containing suspensions of solid-phase material [61] and causing both a decrease in intensity and red-shift of all minima. After one week of incubation the spectrum is similar except for flattening of the 190 nm peak. At this stage, fibrils are present as observed by TEM analysis of an aliquot of the same sample (Fig. 4). The further

flattening of the spectrum over time is likely to be due to both an increase in β -sheet structure as well as differential absorption flattening [61,69]. CD shows that the soluble aggregates differ dramatically in secondary structure compared to native MstnPP and that the aggregates may represent an α -helical-containing intermediate in the transition to protofibrils and fibrils rich in β -sheet structure.

MstnPP Protofibrils and Fibrils Are Resistant to Proteolytic Digest by Trypsin

Resistance to proteolytic digestion is another defining characteristic for the presence of amyloid fibrils [73,74]. Comparative trypsin digestions were performed for the MstnPP fibril-containing sample, soluble aggregates and native dimer (Fig. 6). Digestion of the MstnPP dimer and aggregates was carried out with a MstnPP to trypsin ratio of 100:1 (w/w) at 4°C, room temperature (approximately 22°C) and 37°C with samples taken after 0.5, 1, 2, 3, 4 and 18 hours. The precursor dimer and aggregates have comparable trypsin susceptibility at 37°C and are fully digested after an overnight incubation (18 hours, Fig. 6A and B respectively). The susceptibility of the soluble aggregates to trypsin digest suggests an open, flexible structure, consistent with CD results and supporting the conclusion that this species represents a prefibrillar intermediate rather than protofibrils. Prior to digestion the fibril-containing sample is not able to enter the top of the 4%

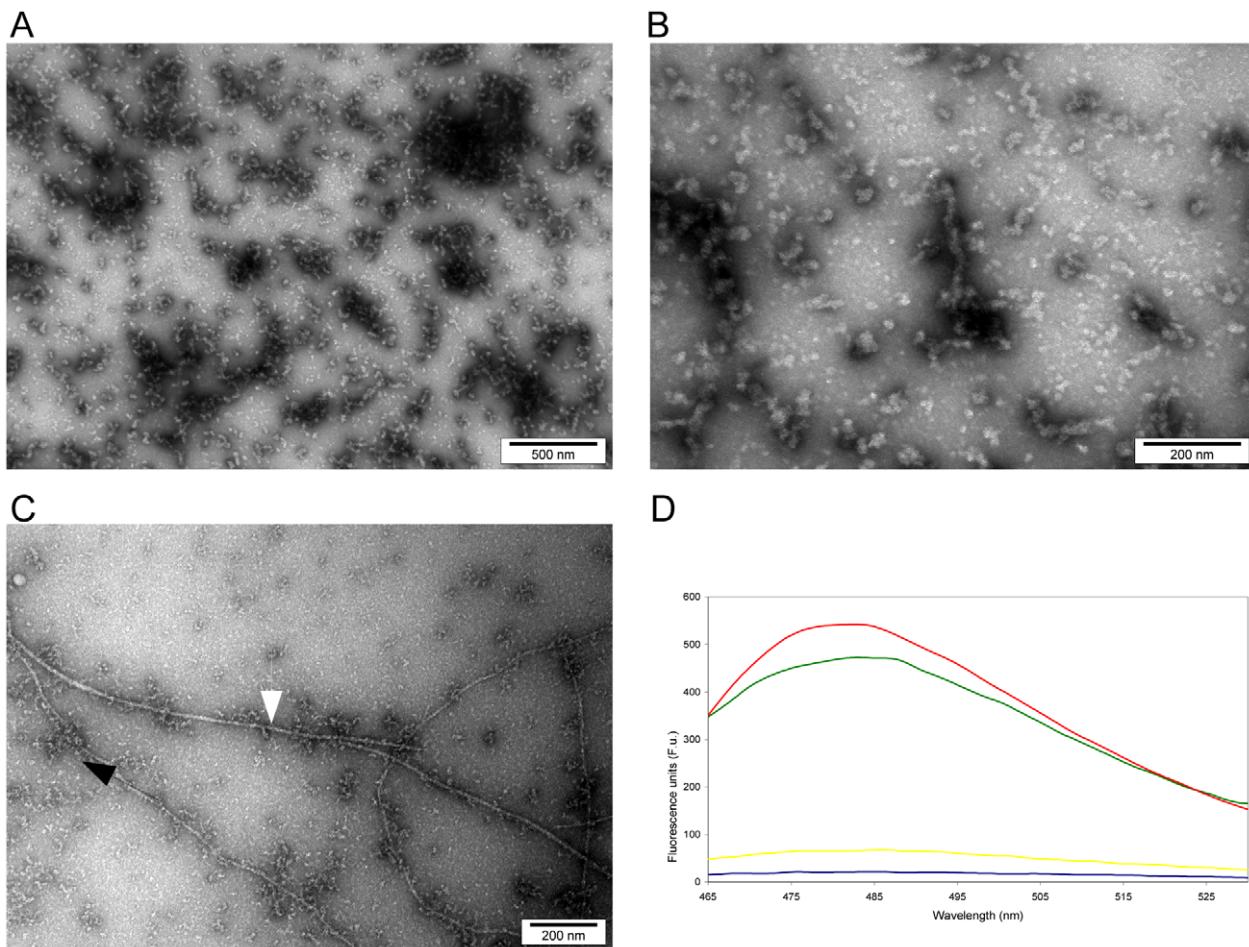


Figure 3. Characterisation of MstnPP soluble aggregates by negative-stain transmission electron microscopy (TEM) and ThT binding. (A–B) TEM of MstnPP soluble aggregates and C. insulin positive control containing both protofibrils (black arrow) and mature fibrils (white arrow). D. ThT binding of MstnPP soluble aggregates (green), compared to an insulin positive control (red), the MstnPP dimer (yellow) and a buffer (50 mM Tris-HCl pH 8.5, 150 mM NaCl) blank (blue).

doi:10.1371/journal.pone.0009170.g003

stacking gel owing to its extremely large size. Following incubation with trypsin this property is maintained even after overnight incubation (18 h) at 37°C and a 5-fold increase in trypsin concentration (MstnPP:trypsin 20:1) (Fig. 6C). Inability to enter the stacking gel is not a definitive indication of proteolysis resistance since partial hydrolysis may result in products smaller than fibrils and/or protofibrils yet still large enough to be retained in the stacking gel. To address this possibility, TEM analysis revealed that both MstnPP protofibrils (Fig. 6Di) and mature fibrils (Fig. 6Dii) had unchanged morphology after trypsin incubation.

MstnPP Aggregates Exhibit Amyloid Characteristics after Incubation at 37°C

The formation of amyloid fibrils by MstnPP at 60°C and pH 5.3 is comparable to conditions routinely used to promote rapid amyloid formation in *in vitro* model systems for amyloid diseases. To investigate the temperature range at which structural changes occur, MstnPP aggregates were analysed by CD from 10 to 65°C immediately after suspension in pH 5.3 (Fig. S3). The spectrum had changed significantly by 45°C with a loss of α -helix beginning from 30°C and the appearance of a characteristic amyloid β -sheet-rich spectrum at 65°C. Subsequently, MstnPP aggregates were incubated at 37°C at pH 5.3 and monitored with

ThT binding, CD spectroscopy and TEM (Fig. 7). After overnight incubation ThT binding increased only slightly (Fig. 7A) and CD indicated that β -sheet aggregation has occurred (Fig. 7B). Under TEM, the 37°C protofibrils have a similar morphology to the 60°C samples, however aggregation is less extensive (Fig. 7C); elongation, clustering and ring-like structures are apparent. After one week of incubation at 37°C, ThT binding increased significantly (Fig. 7A) and CD showed a definite β -sheet transition (Fig. 7B). TEM images show increased aggregation, lateral association and a granular to smooth morphology transition (Fig. 7D). Although fibril formation did not occur at 37°C in the time period observed for incubation at 60°C, it is likely that the kinetics of fibril formation will be slower at the lower temperature. However, these results show that the properties of protofibrils, proposed to be direct precursors of mature amyloid fibrils, have been adopted at 37°C; amyloid formation is likely to occur over an extended period of time.

MstnPP Aggregates Have a Negative Effect on C2C12 Myoblast Viability

Oligomeric species in the amyloid formation pathway of a number of proteins are cytotoxic when added to the media of cultured cells [41,42,65]. The effect of MstnPP soluble aggregates,

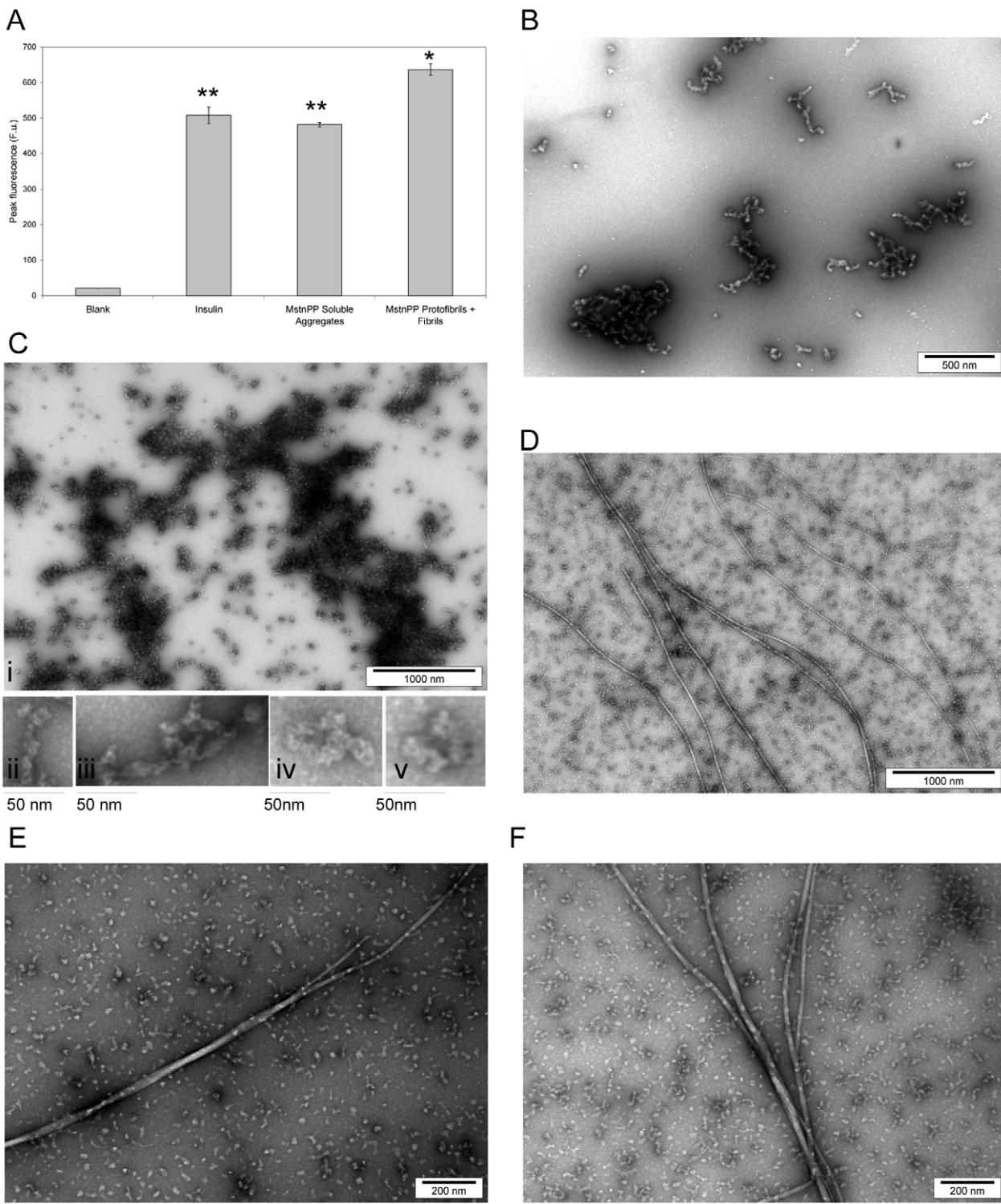


Figure 4. Characterisation of MstnPP amyloid fibrils by ThT binding and TEM. A. ThT binding after incubation at 60°C and pH 5.3 compared to insulin and MstnPP soluble aggregates. Each result represents the average of four independent experiments and error bars represent SEM. ** P<0.005 and * P<0.01 by paired Student's t-test. Blank is 0.005 mM HCl, pH 5.3 and ThT. (B–F) TEM showing different stages in the formation of prefibrillar structures and fibrils by human MstnPP. B. Overnight incubation at 60°C in pH 5.3 produces 'beaded' prefibrillar structures; C. (i) By three days large, dense three-dimensional arrays show up as areas of high electron density, (ii and iii) and pore-like structures (iv and v) from (i); D. Characteristic amyloid fibrils after one week; E–F. Formation of higher order structures with the twisting of two (E) or more (F) fibrils around each other.

doi:10.1371/journal.pone.0009170.g004

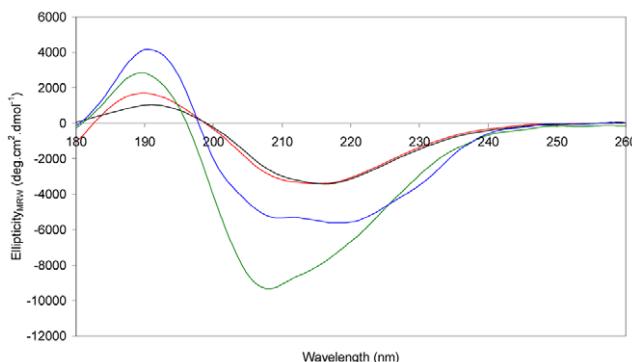


Figure 5. Circular dichroism spectra showing the structural transitions that occur in amyloid fibril formation. Correctly folded MstnPP dimer (blue); MstnPP soluble aggregates before acidification (green); prefibrillar aggregates after overnight incubation in pH 5.3 at 60°C (red); mixture of prefibrillar aggregates and fibrils after one week incubation (black).

doi:10.1371/journal.pone.0009170.g005

protofibrils and fibrils on the viability of C2C12 mouse myoblasts was investigated by monitoring the absorbance of formazan (Fig. 8) produced after addition of WST-1 (Roche). Only viable cells reduce WST-1 to produce formazan with the ratio of absorbances at 450 and 630 nm correlated to the number of viable cells in the culture. C2C12 mouse myoblasts are the standard model cell-line

used for the analysis of myostatin activity [16,18]. Although human MstnPP is used here, sequence identity between the mouse and human myostatin precursors is 96%; the murine cell-line is therefore suitable for initial studies. 25 μM of soluble aggregates and at least 10 μM of protofibrils decreased cell viability significantly compared to buffer only and correctly folded MstnPP dimer controls, at the same pH. A solution containing a mixture of MstnPP protofibrils and fibrils, as well as lower concentrations of soluble aggregates and protofibrils, had a reduced effect. These results indicate that oligomeric species in the amyloid formation pathway of MstnPP affect the normal functioning of C2C12 myoblasts.

Discussion

Amyloid formation is a predominant feature of sIBM [30] with the Alzheimer's disease protein Aβ associated with sIBM amyloid structures [31,44]. A role for MstnPP in sIBM was proposed after co-localisation and direct association with Aβ was observed in diseased cells [31], and also because ER stress causes an upregulation of myostatin expression [45]. We now show that a human MstnPP misfolded species shares the morphological characteristics of amyloid protofibrils and the ability to bind the amyloid-specific dye thioflavin T. The α-helical secondary structure of these aggregates suggests they are most likely to represent an intermediate oligomeric species that occurs in the transition between native protein and β-sheet-rich protofibrils. When subjected to the mildly denaturing conditions of

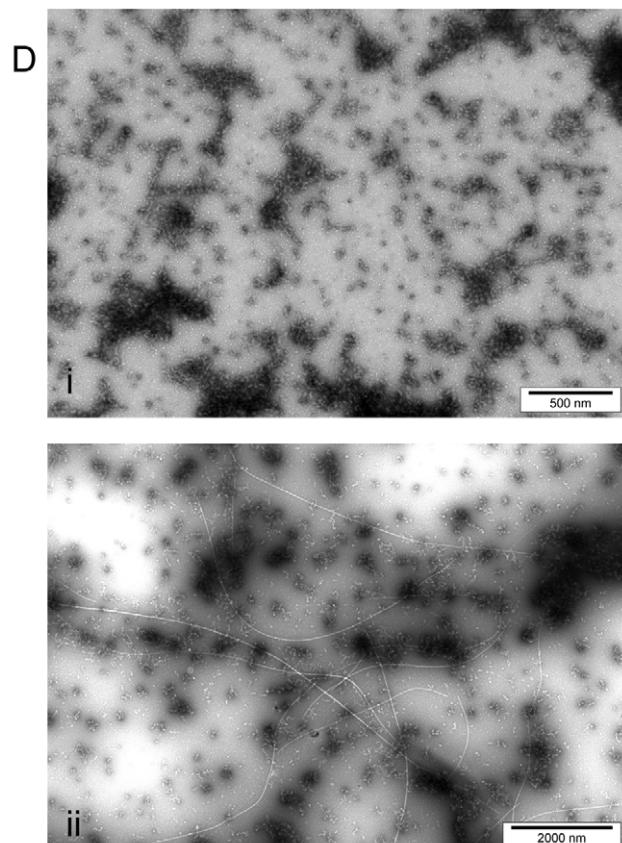
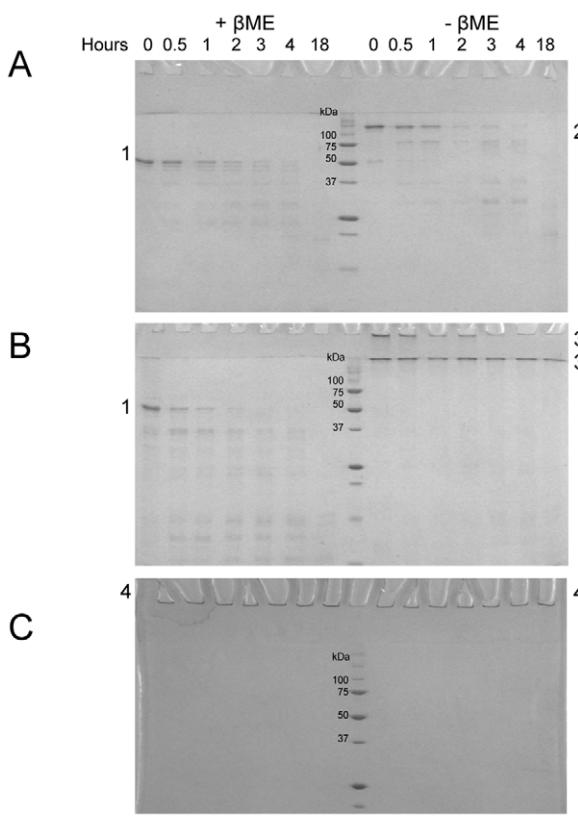


Figure 6. SDS-PAGE and TEM analysis showing MstnPP resistance to trypsin digestion. A. dimer; and B. soluble aggregates. Major bands are indicated: 1. MstnPP monomer; 2. MstnPP dimer; 3. MstnPP soluble aggregates. C. MstnPP fibrils: 4. prefibrillar aggregates plus fibrils. Samples were analysed in sample buffer +/− βME where βME concentration is 2 M. D. TEM shows resistance of (i) protofibrils and (ii) amyloid fibrils to trypsin. Trypsin digests were performed with a MstnPP:trypsin ratio of 100:1 (A and B) or 20:1 (C and D) for 18 hours at 37°C.

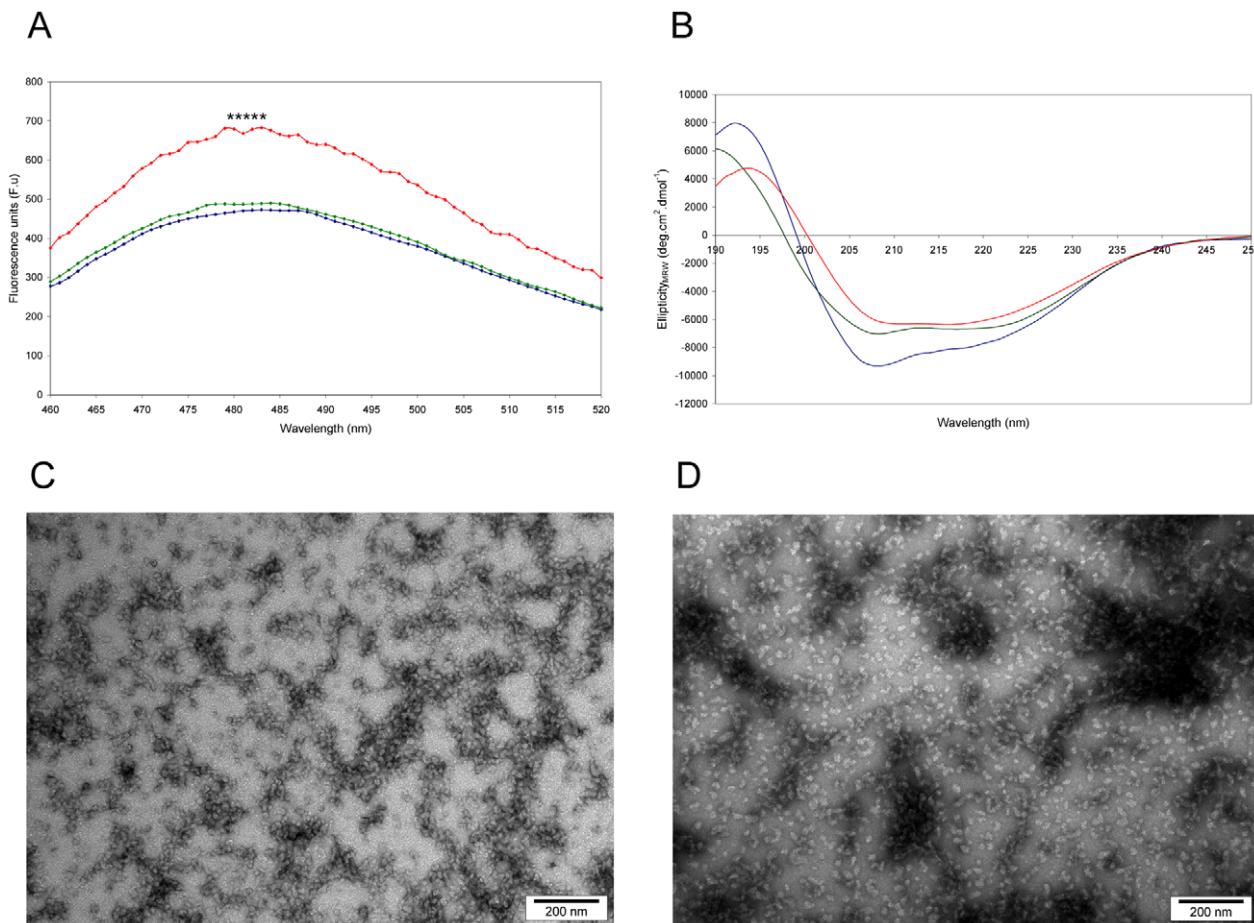


Figure 7. ThT binding, CD spectroscopy and TEM show the formation of prefibrillar structures and protofibrils by MstnPP aggregates at 37°C and pH 5.3. A. ThT binding assays before incubation (blue), after overnight incubation (green) and after one week incubation (red). ***** P<0.001 by Student's t-test using averaged values over peak. B. CD spectra of MstnPP soluble aggregates (blue), aggregates after overnight incubation at 37°C in pH 5.3 (green) and after one week incubation (red). C. TEM shows elongation and clustering after overnight incubation; and D. a granular to smooth transition and enhanced lateral association after one week.

doi:10.1371/journal.pone.0009170.g007

pH 5.3 and 60°C, the aggregates form amyloid protofibrils and after one week are able to form long, linear and unbranching amyloid fibrils. At 37°C, protofibrils and mature fibrils form over an increased time period. These *in vitro* results show that MstnPP is capable of amyloid protofibril and fibril formation, supporting a role for misfolding of the myostatin precursor in the pathogenesis of sIBM. MstnPP aggregates and protofibrils have a cytotoxic effect on mouse myoblasts when added to the culture medium.

CD spectroscopy shows there is a transition from the correctly folded protein structure to a structure consisting primarily of β -sheet. For a number of peptides and proteins this transition is from a loop or disordered region through an α -helical intermediate stage [36,75]. Analysis of the differently folded forms of MstnPP suggests a shift towards a predominance of α -helix on misfolding. Whether this α -helical shift represents the intermediate stage described for other amyloid species is not known but is an attractive possibility. TEM images show a number of different morphological forms present in the aggregate solution. These may represent different oligomeric intermediates at different stages of the fibril formation process, some of which may be the direct precursors of protofibrils. Overnight incubation of MstnPP aggregates in pH 5.3 at 60°C produces a transition from the α -helical structure to one dominated by β -sheet. This result is

consistent with the major models proposed for amyloid formation, which describe an unfolded state in equilibrium with a partly unfolded α -helix-containing intermediate that accelerates fibril formation [36,37,76]. Aggregation of the intermediate into oligomers precedes structural rearrangement into β -sheet-rich protofibrils. A detailed structure of protofibrils has not been determined; they are defined primarily by morphology, the ability to bind dyes such as ThT and a structural shift to a predominance of β -sheet. Amyloid fibrils have been well-studied structurally by NMR and X-ray fiber diffraction [77].

For growth factors such as myostatin, and a number of other secreted proteins, perhaps the biggest obstacle to refolding is the formation of the correct disulfide-bonded cysteine pairs in the ER [78]. Myostatin has the characteristic TGF- β cysteine knot motif in its mature domain [2,10,11] and so likely requires both an appropriate redox environment and the presence of disulfide-bond chaperones for correct folding in the ER. The myostatin propeptide that lies N-terminal to the mature domain has been proposed to have a chaperone-like function in the correct folding of the mature domain [8,9]. During ER-stressed situations, such as an altered redox environment, misfolding can be extensive [48,79] with the unfolded protein response (UPR) resulting in the elimination of misfolded proteins and/or apoptosis if necessary [48,80,81]. If the

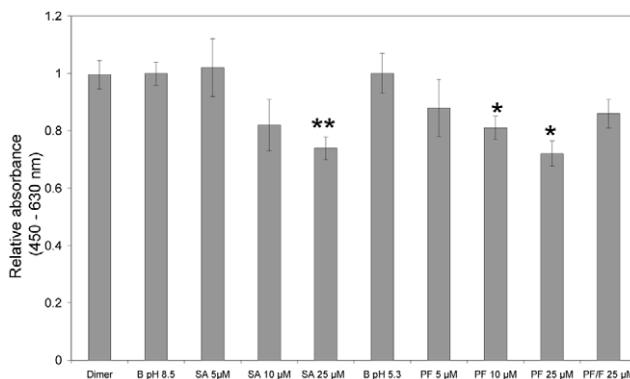


Figure 8. MstnPP aggregates and protofibrils (PF) are cytotoxic to C2C12 mouse myoblasts. C2C12 cytotoxicity assay using the WST-1 reagent where the difference in absorbance at 450 and 630 nm directly correlates to cell density after incubation with increasing concentrations of MstnPP soluble aggregates (SA, 5–25 μM), protofibrils (PF, 5–25 μM), 25 μM MstnPP dimer and 25 μM PF/F (fibrils). Concentrations are expressed as monomer equivalents. Cells incubated in media containing buffer (B) only (50 mM Tris-HCl pH 8.5, 150 mM NaCl for soluble aggregates and dimer; 0.005 mM HCl, pH 5.3 for protofibrils and fibrils) were used as a control. Error bars represent the standard error of the mean for triplicate samples from two independent experiments. Statistical significance was calculated using a paired Student's t-test where ** P<0.01 and * P<0.05.

doi:10.1371/journal.pone.0009170.g008

production of misfolded proteins overwhelms the UPR, protein aggregation may result in the generation of cytotoxic amyloid protofibrils [41,42] and mature amyloid fibrils, contributing to amyloid diseases, including sIBM [32]. After refolding of MstnPP, misfolded yet soluble aggregates that form amyloid protofibrils and fibrils at a lowered pH can be observed. Although elevated temperature has been used to produce mature MstnPP fibrils quickly in this study, a physiological temperature allows protofibril and mature fibril formation over an extended time period; by comparison, fibril formation is expected to occur over many years *in vivo*.

sIBM is characterised by amyloid aggregation of the Aβ protein and severe muscle atrophy [30,33]. Previous studies have shown that MstnPP associates with Aβ fibrils in diseased cells [31] and that ER stress results in upregulation of MstnPP in a cultured human sIBM-model [45]. Increased activity of the myostatin growth factor following MstnPP upregulation may contribute to atrophy in sIBM, however another study showed MstnPP accumulated with Aβ/AβPP within aggresomes following ER stress by proteasome inhibition in CHMFs [46]. This accumulation in aggresomes may prevent processing of MstnPP and decrease levels of circulating growth factor, as mature myostatin growth factor is absent in immunoblots of CHMFs transiently expressing AβPP [46].

The *in vitro* formation of MstnPP aggregates and amyloid species suggests a similar phenomenon could occur *in vivo* having a negative effect on cell viability and/or assisting in Aβ fibrillogenesis. The cytotoxicity exhibited by MstnPP aggregates and protofibrils in C2C12 cells supports a mechanism in which aggregation by MstnPP in sIBM contributes to muscle degeneration. This may occur independently of atrophy resulting from increased myostatin growth factor signalling. Even though MstnPP aggregates were localised to the cytoplasm, aggresomes and nuclear regions of sIBM tissue in previous studies [31,46], an absence of extracellular MstnPP was not shown. Since MstnPP is secreted, possibly for local inhibition of muscle growth [14], post-

secretion aggregation or secretion of aggregated/misfolded MstnPP may contribute to muscle fiber atrophy in sIBM via a cytotoxic mechanism that may involve alterations to endo- and exocytosis [65], alterations to Ca²⁺ homeostasis and the generation of reactive oxygen species [41]. The presence of pore-like structures under TEM suggests that the cytotoxic mechanism may include permeabilization of extracellular and/or intracellular membranes [43].

The results presented here show that human MstnPP is able to form cytotoxic amyloid-like aggregates and amyloid fibrils *in vitro*, raising the possibility that amyloid formation by MstnPP *in vivo* may contribute to the pathogenesis of sIBM.

Materials and Methods

MstnPP β-Aggregation Propensity by *In Silico* Analysis

MstnPP amino acids 21–375 were used in β-aggregation propensity prediction algorithms. Tango [52] and Waltz [55] output was reproduced from www.tango.embl.de; PASTA output was taken from www.protein.cribi.unipd.it/pasta [54]. Default settings were used for all programs.

Expression of Recombinant MstnPP

Full-length human myostatin cDNA was used as a template for PCR cloning with primers encompassing cDNA corresponding to human MstnPP amino acid 21 (C-terminal to the signal peptidase cleavage site) to the C-terminal amino acid 375. PCR fragments were cloned into a modified pET vector via *Bam*H and *Xba*I restriction sites. Positive colonies were selected by colony PCR. The pET-MstnPP plasmid was isolated, sequenced and used to transform *E. coli* BL21(DE3) cells for protein expression. Transformed BL21 cells containing pET-MstnPP were grown in a starter culture of 20 mL LB in the presence of 100 μg/mL ampicillin (LB amp) at 37°C overnight. The starter culture was used to inoculate 2 L LB amp and cells were grown at 25°C to an OD₆₀₀ of 0.8. Expression was induced by the addition of 0.1 mM IPTG and cells were incubated overnight (16–18 hours). Cells were collected by centrifugation at 4,000×g for 30 min at 4°C and stored at –20°C.

Inclusion Body Isolation and Solubilisation and Refolding of Recombinant MstnPP

The frozen cell pellet from each 2 L culture was thawed at room temperature and resuspended in a final volume of 70 mL 20 mM Tris-HCl pH 8.5. Cells were lysed by two passages through a French Press at 5,000 psi and the inclusion body pellet was collected by centrifugation at 30,000×g at 4°C for 30 min. The inclusion bodies were washed with two 20 mL volumes of wash buffer (50 mM Tris HCl pH 8.5, 0.5 M NaCl, 10% glycerol, 0.5% Triton-X 100) and two washes with the same volume of milliQ H₂O. Pellets were collected after each wash by centrifugation at 30,000×g for 10 min at 4°C. Solubilisation and refolding were carried out using methods published previously for the zebrafish myostatin precursor protein [8] with some modifications. Each inclusion body pellet from 2 L of cell culture was solubilised in 8 mL 6 M guanidine hydrochloride (GndHCl), 50 mM Tris-HCl pH 8.5, 1 mM EDTA, 0.1 mM DTT, at room temperature, with shaking for 5 hours. Insoluble proteins and cell debris were removed by centrifugation at 17,000×g, 4°C for 30 min. The supernatant was acidified to pH ~5.0 by adding a few drops of concentrated HCl and dialysed overnight against 6 M GndHCl, 50 mM MES pH 5.0, 1 mM EDTA to remove DTT. After dialysis, the sample was centrifuged again as above. The solubilised protein (8 mL) was rapidly diluted in 150 mL of freshly

prepared refolding buffer, consisting of 50 mM Tris-HCl pH 8.5, 1 M NaCl, 0.5 M L-arginine, 5 mM EDTA, 5 mM reduced glutathione and 1 mM oxidised glutathione. The pH was re-adjusted to 8.5 before addition of protein. The protein was left to refold without agitation for 6 days at 4°C then centrifuged at 30,000 ×g, 4°C for 30 min before overnight dialysis into 20 mM Tris-HCl pH 8.5. Soluble protein was harvested as supernatant after a second centrifugation and filtered (0.45 µM).

Purification of Refolded MstnPP Dimer and Soluble Aggregates

Purification of the *in vitro* refolded MstnPP dimer and misfolded soluble aggregates was performed by heparin affinity followed by gel filtration chromatography. The filtered protein solution was loaded onto a 5 mL HiTrap Heparin HP column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.5, at a flow rate of 1 mL/min. The flow-through containing unbound protein was discarded and a stepwise NaCl elution was used to elute the correctly folded dimer at 200 mM NaCl (H1) and the majority of soluble aggregates at 600 mM NaCl (H2). Elution was monitored by absorbance at 280 nm and differently folded forms were identified using non-reduced SDS-PAGE. Further purification of the MstnPP dimer was achieved using gel filtration chromatography. The 200 mM NaCl peak was filtered (0.22 µM), concentrated to 350 µL and applied to a Superdex S200 10/300 HR column (GE Healthcare) equilibrated with 50 mM Tris-HCl, pH 8.5, 150 mM NaCl, at a flow rate of 0.5 mL/min. The dimer eluted at approximately 13.5 mL (PD) and was separated from any remaining soluble aggregates that eluted in the void volume (PA).

Western Blotting

Protein samples were separated on a 12% SDS-PAGE gel at 150 V and transferred to a nitrocellulose membrane overnight at 15 V (approximately 40–150 mA) in cooled Tris/glycine transfer buffer. The next day transfer of high molecular weight samples was ensured at 450 mA for one hour. Efficient transfer was confirmed by Ponceau S staining. Membranes were blocked with 5% non-fat milk in TBS-Tween (TBS-T; 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20) for one hour at room temperature and subsequently incubated in rabbit anti-myostatin polyclonal primary antibody (Chemicon International), diluted 1/10,000 in 1% non-fat milk TBS-T, for 90 minutes. Following washing with TBS-T, membranes were incubated in a 1/50,000 dilution (1% non-fat milk, TBS-T) of horse radish peroxidase-conjugated anti-rabbit antibody (Jackson ImmunoResearch) for 90 minutes. All steps were performed at room temperature. After washing with TBS-T, proteins were visualised with the Super-Signal West Pico Chemiluminescent Substrate (Pierce) according to manufacturer's instructions.

Generation of MstnPP Amyloid Fibrils

MstnPP soluble aggregates were sterile-filtered (0.22 µm), concentrated 10-fold and resuspended in dilute HCl solutions at a final concentration of 1 mg/mL, as measured by both Bradford assay and absorbance at 280 nm. Initially, a range of pH solutions were investigated for fibril formation from pH 1.6, as used for insulin amyloid formation [63] through to pH 6.3. Samples were placed at either 37°C or 60°C and monitored frequently for amyloid formation by Thioflavin T (ThT) fluorescence, TEM and CD spectroscopy. Subsequent to this optimisation, MstnPP aggregates were incubated at pH 5.3 (0.005 mM HCl) and 60°C for at least one week. Identical MstnPP amyloid fibrils were formed in equivalent solutions containing 0.1% sodium azide.

Generation of Insulin Protomembrane and Fibrils

Insulin protomembrane and fibrils were produced as described [63]. In brief, different concentrations of insulin (bovine pancreas, Sigma) ranging from 0.5–2 mg/mL (85–340 µM) were prepared in 25 mM HCl (pH 1.6) and placed at 60°C for an hour (protofibrils) or at least overnight (fibrils) although incubation times were dependent on protein concentration.

Negative-Stained Transmission Electron Microscopy

Dimer and aggregate samples direct from purification (heparin affinity or gel filtration) were buffer-exchanged into water. Amyloid solutions were used directly in dilute HCl. 200-mesh carbon-coated Formvar grids were placed on 30 µL protein solution (0.5 mg/mL) droplets for 45 seconds. Excess sample was drawn off using filter paper and the grids were placed on an equal volume of 2% uranyl acetate for 45 seconds. Excess stain was drawn off as above and the grids were air-dried briefly before viewing with a Philips CM10 transmission electron microscope. Fibril measurement and statistical analysis was performed using iTEM software (Olympus) with a sample size of 10 fibrils from the same grid.

Thioflavin T Binding Assays

A solution of Thioflavin T (Sigma) was prepared at 400 µM in water [64] then diluted directly into the protein solution giving a final concentration of 20 µM per assay (as done previously for insulin [63]). A final protein concentration of 1 mg/mL was used for all samples. Experiments at pH 7.5 were performed after diluting both protein and ThT into 50 mM Tris-HCl pH 7.5, 100 mM NaCl. Fluorescence assays were conducted with the Perkin Elmer LS50B Luminescence Spectrometer with excitation at 450 nm and the emission spectrum measured from 460–530 nm. Error bars are standard errors for four independent measurements. The Student's t-test was performed using GraphPad Prism (GraphPad Software, Inc).

Circular Dichroism Spectroscopy

CD spectra in the far-UV region (180–260 nm) were obtained on a Chirascan CD spectrometer (Applied Biophysics) using a 0.1 mm path cell and protein concentrations of 1 mg/mL at 4°C. For all samples, 20 runs were performed with 1 nm readings taken every 2.5 seconds, followed by smoothing and baseline subtraction. For CD thermal denaturation, 1 nm/2.5 second readings were taken at every 5°C increase in temperature from 10–65°C with a 30 second equilibration time at each temperature and a tolerance level of 0.2°C.

Protease Resistance Analysis

MstnPP samples were tested for protease resistance using a myostatin to trypsin (bovine pancreas, Sigma) ratio of 100:1 or 20:1 (w/w). 100:1 solutions were incubated at 4°C, room temperature or 37°C. Samples were taken at 0.5, 1, 2, 3, 4 and 18 hours (overnight), immediately denatured to end the reaction and analysed by reducing and non-reducing SDS-PAGE. For subsequent protease resistance analysis of MstnPP amyloid fibrils, the myostatin to trypsin ratio was 20:1 (w/w) with incubation at 37°C for 4 and 18 hours and analysis by SDS-PAGE and TEM.

Cell Culture and Incubation with Protein Aggregates

C2C12 mouse myoblasts were cultured in Advanced Dulbecco's Modified Eagle's Medium (Gibco, Invitrogen) containing 4.5 g/L D-glucose and 110 mg/L sodium pyruvate and supplemented with 10% foetal calf serum, 4 mM L-glutamine and penicillin/

streptomycin. Cells were incubated in a 5% CO₂ humidified environment at 37°C. Cells were plated in fresh medium in optical bottom 96-well plates (Nunc) at a density of 2,500 cells/well for cytotoxicity assays. After 24 hours, media was removed and 100 µL/well fresh medium containing protein was added. Two independent experiments with triplicate wells were used for each condition. Cells incubated in media containing buffer (B) only (50 mM Tris-HCl pH 8.5, 150 mM NaCl for soluble aggregates and dimer; 0.005 mM HCl, pH 5.3 for protofibrils and fibrils) were used as a control. Concentrations of myostatin samples used in the assay were comparable to those from previous protein fibril studies [41,42,65].

WST-1 Colorimetric Assay for Cytotoxicity

The WST-1 reagent (Roche) is a tetrazolium-based salt that can be metabolized by viable cells only to produce a water-soluble formazan product measurable in a colorimetric assay. After a 24-hour incubation with protein, 10 µL WST-1 was added to each well and the plate was incubated for a further 3 hours. Absorbance at 450 nm was measured in a PowerWave XS 96-well plate reader (BioTek Instruments, Inc) with 630 nm readings taken as reference. Phenol red-free media was not used as the indicator does not interfere with the WST-1 assay; as phenol red absorbs slightly at 450 nm however, wells containing media only were used as an additional control [82].

Supporting Information

Figure S1 *In silico* predictions of propensity for β-sheet aggregation (A and B) and regions responsible for amyloid formation (C) by MstnPP. (A) Tango; (B) PASTA; and (C) Waltz. The MstnPP amino acid sequence from residues 21–375 were used for the calculations at a theoretical pH of 7. Default settings were used for all algorithms.

References

1. Joulia-Ekaza D, Cabello G (2006) Myostatin regulation of muscle development: molecular basis, natural mutations, physiopathological aspects. *Exp Cell Res* 312: 2401–2414.
2. Lee SJ (2004) Regulation of muscle mass by myostatin. *Ann Rev Cell Dev Biol* 20: 61–86.
3. McPherron AC, Lawler AM, Lee SJ (1997) Regulation of skeletal muscle mass in mice by a new TGF-β superfamily member. *Nature* 387: 83–90.
4. Zimmers TA, Davies MV, Koniaris LG, Haynes P, Esquela AF, et al. (2002) Induction of cachexia in mice by systemically administered myostatin. *Science* 296: 1486–1488.
5. Hill JJ, Davies MV, Pearson AA, Wang JH, Hewick RM, et al. (2002) The myostatin propeptide and the follistatin-related gene are inhibitory binding proteins of myostatin in normal serum. *J Biol Chem* 277: 40735–40741.
6. Jiang M, Liang L, Wang S, Ratovitski T, Holmstrom J, et al. (2004) Characterization and identification of the inhibitory domain of GDF-8 propeptide. *Biochem Biophys Res Commun* 315: 525–531.
7. Lee SJ, McPherron AC (2001) Regulation of myostatin activity and muscle growth. *Proc Natl Acad Sci* 98: 9306–9311.
8. Funkenstein B, Rebban Y (2007) Expression, purification, renaturation and activation of fish myostatin expressed in *Escherichia coli*: facilitation of refolding and activity inhibition by myostatin prodomain. *Protein Expr Purif* 10: 54–65.
9. Jin HJ, Dunn MA, Borthakur D, Kim YS (2004) Refolding and purification of unprocessed porcine myostatin expressed in *Escherichia coli*. *Protein Expr Purif* 35: 1–10.
10. Berry C, Thomas M, Langley B, Sharma M, Kambadur R (2002) Single cysteine to tyrosine transition inactivates the growth inhibitory function of Piedmontese myostatin. *Am J Physiol* 283: 135–141.
11. Cash JN, Rejon CA, McPherron AC, Bernard DJ, Thompson TB (2009) The structure of myostatin:follistatin 288: insights into receptor utilization and heparin binding. *EMBO J* 28: 2662–2676.
12. Lee SJ (2008) Genetic analysis of the role of proteolysis in the activation of latent myostatin. *PLoS ONE* 3: e1628.
13. Wolfman NM, McPherron AC, Pappano WN, Davies MV, Song K, et al. (2003) Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. *Proc Natl Acad Sci* 100: 15842–15846.
14. Anderson SB, Goldberg AL, Whitman M (2008) Identification of a novel pool of extracellular pro-myostatin in skeletal muscle. *J Biol Chem* 283: 7027–7035.
15. Amthor H, Otto A, Macharia R, McKinnell I, Patel K (2006) Myostatin imposes reversible quiescence on embryonic muscle precursors. *Dev Dyn* 235: 672–680.
16. Joulia D, Bernardi H, Garandel V, Rabenoelina F, Vernus B, et al. (2003) Mechanisms involved in the inhibition of myoblast proliferation and differentiation by myostatin. *Exp Cell Res* 286: 263–275.
17. Langley B, Thomas M, Bishop A, Sharma M, Gilmore S, et al. (2002) Myostatin inhibits myoblast differentiation by down-regulating MyoD expression. *J Biol Chem* 277: 49831–49840.
18. Thomas M, Langley B, Berry C, Sharma M, Kirk S, et al. (2000) Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *J Biol Chem* 275: 40235–40243.
19. Manceau M, Gros J, Thome K, McPherron A, Paterson B, et al. (2008) Myostatin promotes the terminal differentiation of embryonic muscle progenitors. *Genes Dev* 22: 668–681.
20. Lee SJ (2007) Sprinting without myostatin: a genetic determinant of athletic prowess. *Trends Genet* 23: 475–477.
21. Kambadur R, Sharma M, Smith TPL, Bass JJ (1997) Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. *Genome Res* 7: 910–915.
22. McNally EM (2004) Powerful genes - myostatin regulation of human muscle mass. *N Engl J Med* 350: 2642–2644.
23. Schuelke M, Wagner KR, Stolz LE, Hubner C, Riebel T, et al. (2004) Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med* 350: 2682–2688.
24. Durieux AC, Amirouche A, Banzet S, Koulmann N, Bonnefoy R, et al. (2007) Ectopic expression of myostatin induces atrophy of adult skeletal muscle by decreasing muscle gene expression. *Endocrinology* 148: 3140–3147.
25. Wagner K, McPherron A, Winik N, Lee S (2002) Loss of myostatin attenuates severity of muscular dystrophy in *mdx* mice. *Ann Neurol* 52.
26. Gonzalez-Cadavid N, Taylor W, Yarasheski K, Sinha-Hikim I, Ma K, et al. (1998) Organization of the human myostatin gene and expression in healthy men and HIV-infected men with muscle wasting. *Proc Natl Acad Sci* 95: 14938–14943.

Found at: doi:10.1371/journal.pone.0009170.s001 (0.24 MB DOC)

Figure S2 Western blot and SDS-PAGE of MstnPP purification procedure. (A) 12% reducing (+ βME) vs non-reducing (− βME) SDS-PAGE and (B) subsequent Western blot of purification procedure. Lanes are as follows: H1, heparin peak 1; H2, heparin peak 2; PD, gel filtration purified dimer; PA, gel filtration purified aggregates. Major bands are indicated: 1. soluble aggregates; 2. dimer; 3. monomer. β-ME concentration in + β-ME lanes is 2 M. Found at: doi:10.1371/journal.pone.0009170.s002 (1.42 MB DOC)

Figure S3 CD thermal denaturation of MstnPP soluble aggregates at pH 5.3 shows a gradual transition from an α-helical dominated spectrum at 10°C to one rich in β-sheet at 65°C. Significant loss of the α-helical minimum begins at 40°C.

Found at: doi:10.1371/journal.pone.0009170.s003 (0.21 MB DOC)

Acknowledgments

The authors would like to thank Se-Jin Lee (John Hopkins University, USA) for the generous gift of plasmid containing full-length human myostatin cDNA. We thank Juliet Gerrard and Jackie Healy (University of Canterbury, New Zealand) for technical assistance with Thioflavin T binding assays and Doug Hopcroft (Manawatu Microscopy and Imaging Centre, Massey University) for assistance with electron microscopy.

Author Contributions

Conceived and designed the experiments: CSS AJS-S. Performed the experiments: CSS. Analyzed the data: CSS AJS-S. Wrote the paper: CSS AJS-S.

27. Costelli P, Muscaritoli M, Bonetto A, Penna F, Reffo P, et al. (2008) Muscle myostatin signalling is enhanced in experimental cancer cachexia. *Eur J Clin Invest* 38: 531–538.
28. Patel K, Amthor H (2005) The function of myostatin and strategies of myostatin blockade - new hope for therapies aimed at promoting growth of skeletal muscle. *Neuromuscul Disord* 15: 117–126.
29. Tobin JF, Celeste AJ (2005) Myostatin, a negative regulator of muscle mass: implications for muscle degenerative diseases. *Curr Opin Pharmacol* 5: 328–332.
30. Askanas V, Engel WK (2008) Inclusion-body myositis: muscle fiber molecular pathology and possible pathogenic significance of its similarity to Alzheimer's and Parkinson's disease brains. *Acta Neuropathol* 116: 583–595.
31. Wojcik S, Engel WK, McFerrin J, Askanas V (2005) Myostatin is increased and complexes with amyloid- β within sporadic inclusion-body myositis muscle fibers. *Acta Neuropathol* 110: 173–177.
32. Vattemi G, Engel WK, McFerrin J, Askanas V (2004) Endoplasmic reticulum stress and unfolded protein response in inclusion body myositis muscle. *Am J Pathol* 164: 1–7.
33. Karpati G, O'Ferrall EK (2009) Sporadic inclusion body myositis: pathogenic considerations. *Ann Neurol* 65: 7–11.
34. Meredith SC (2005) Protein denaturation and aggregation. Cellular responses to denatured and aggregated proteins. *Ann NY Acad Sci* 1066: 181–221.
35. Stefani M, Dobson CM (2003) Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med* 81: 678–699.
36. Abedini A, Raleigh DP (2009) A role for helical intermediates in amyloid formation by natively unfolded ploypeptides? *Phys Biol* 6: 15005.
37. Kumar S, Mohanty SK, Udgaoarkar JB (2007) Mechanism of formation of amyloid protofibrils of barstar from soluble oligomers: evidence for multiple steps and lateral association coupled to conformational conversion. *J Mol Biol* 367: 1186–1204.
38. Svane AS, Jahn K, Deva T, Malmendal A, Otzen DE, et al. (2008) Early stages of amyloid fibril formation studied by liquid-state NMR: the peptide hormone glucagon. *Biophys J* 95: 366–377.
39. Jansen R, Dzwolak W, Winter R (2005) Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy. *Biophys J* 88: 1344–1353.
40. Juárez J, Taboada P, Mosquera V (2009) Existence of different structural intermediates on the fibrillation pathway of human serum albumin. *Biophys J* 96: 2353–2370.
41. Bucciantini M, Calloni G, Chiti F, Formigli L, Nosi D, et al. (2004) Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *J Biol Chem* 279: 31374–31382.
42. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, et al. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416: 507–511.
43. Glabe CG (2006) Common mechanisms of amyloid oligomer pathogenesis in degenerative disease. *Neurobiology of Aging* 27: 570–575.
44. Vattemi G, Nogalska A, Engel WK, D'Agostino C, Checler F, et al. (2009) Amyloid- β 42 is preferentially accumulated in muscle fibers of patients with sporadic inclusion body myositis. *Acta Neuropathol* 117: 569–574.
45. Nogalska A, Wojcik S, Engel WK, McFerrin J, Askanas V (2007) Endoplasmic reticulum stress induces myostatin precursor protein and NF- κ B in cultured human muscle fibers: relevance to inclusion body myositis. *Exp Neurol* 240: 610–618.
46. Wojcik S, Nogalska A, McFerrin J, Engel WK, Oledzka G, et al. (2007) Myostatin precursor protein is increased and associates with amyloid- β precursor protein in inclusion body myositis model. *Neuropath Appl Neuro* 33: 238–242.
47. Lin W, Popko B (2009) Endoplasmic reticulum stress in disorders of myelinating cells. *Nat Neurosci* 12: 379–385.
48. Schroder M (2008) Endoplasmic reticulum stress responses. *Cell Mol Life Sci* 65: 862–894.
49. Garlepp MJ, Mastaglia FL (2007) Inclusion body myositis: new insights into pathogenesis. *Curr Opin Rheumatol* 20: 662–668.
50. Badtke MP, Hammer ND, Chapman MR (2009) Functional amyloids signal their arrival. *Sci Signal* 2: pe43.
51. Maji SK, Perrin MH, Sawaya MR, Jessberger S, Vadodaria K, et al. (2009) Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325: 328–332.
52. Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22: 1302–1306.
53. Trovato A, Chiti F, Maritan A, Seno F (2006) Insight into the structure of amyloid fibrils from the analysis of globular proteins. *PLoS Comput Biol* 2: e170.
54. Trovato A, Seno F, Tosato SCE (2007) The PASTA server for protein aggregation prediction. *Protein Eng Des Sel* 20: 521–523.
55. Reumers J, Schymkowitz J, Rousseau F (2009) Using structural bioinformatics to investigate the impact of non synonymous SNPs and disease mutations: scope and limitations. *BMC Bioinformatics* 10: S9.
56. Mittl PRE, Priestle JP, Cox DA, McMaster G, Cerletti N, et al. (1996) The crystal structure of TGF- β 3 and comparison to TGF- β 2: implications for receptor binding. *Protein Sci* 5: 1261–1271.
57. Schreuder H, Liesum A, Pohl J, Kruse M, Koyama M (2005) Crystal structure of recombinant human growth and differentiation factor 5: evidence for interaction of the type I and type II receptor-binding sites. *Biochem Biophys Res Comm* 329: 1076–1086.
58. Tsumoto K, Ejima D, Kumagai I, Arakawa T (2003) Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif* 28: 1–8.
59. Goda S, Takano K, Yamagata Y, Nagata R, Akutsu H, et al. (2000) Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein Sci* 9: 369–375.
60. Malisauskas M, Ostman J, Darinskas A, Zamotin V, Liutkevicius E, et al. (2005) Does the cytotoxic effect of transient amyloid oligomers from common equine lysozyme *in vitro* imply innate amyloid toxicity? *J Biol Chem* 280: 6269–6275.
61. Calloni G, Lendel C, Campioni S, Giannini S, Gliozzi A, et al. (2008) Structure and dynamics of a partially folded protein are decoupled from its mechanism of aggregation. *J Am Chem Soc* 130: 13040–13050.
62. Mauro M, Craparo EF, Podesta A, Bulone D, Carrotta R, et al. (2007) Kinetics of different processes in human insulin amyloid formation. *J Mol Biol* 366: 258–274.
63. Groenning M, Normann M, Flink JM, Weert Mvd, Bukrinsky JT, et al. (2007) Binding mode of Thioflavin T in insulin amyloid fibrils. *J Struct Biol* 159: 483–497.
64. Bourhim M, Kruzel M, Srikrishnan T, Nicotera T (2007) Linear quantitation of $\text{A}\beta$ aggregation using Thioflavin T: reduction in fibril formation by colostrinin. *J Neurosci Methods* 160: 264–268.
65. Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, et al. (1999) Amyloid β -protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J Biol Chem* 274: 25945–25952.
66. Levine-III H (1993) Thioflavine T interaction with synthetic Alzheimer's disease β -amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci* 2: 404–410.
67. Sabaté R, Lascu I, Saupe SJ (2008) On the binding of Thioflavin-T to HET-s amyloid fibrils assembled at pH 2. *J Struct Biol* 162: 387–396.
68. Marshall KE, Serpell LC (2009) Structural integrity of β -sheet assembly. *Biochem Soc T* 37: 671–676.
69. Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta* 1751: 119–139.
70. Receveur-Brechet V, Bourhis JM, Uversky VN, Canard B, Longhi S (2006) Assessing protein disorder and induced folding. *Proteins Struct Funct Bioinform* 62: 24–45.
71. Rezaei-Ghaleh N, Zweckstetter M, Moshedi D, Ebrahim-Habibi A, Nemat-Gorgani M (2009) Amyloidogenic potential of α -chymotrypsin in different conformational states. *Biopolymers* 91: 28–36.
72. Picotti P, Franceschi GD, Frare E, Spolaore B, Zambonin M, et al. (2007) Amyloid fibril formation and disaggregation of fragment 1–29 of apomyoglobin: insights into the effect of pH on protein fibrillogenesis. *J Mol Biol* 367: 1237–1245.
73. Bocharova OV, Breydo L, Parfenov AS, Salnikov VV, Baskakov IV (2005) In vitro conversion of full-length mammalian prion protein produces amyloid form with physical properties of PrPSc. *J Mol Biol* 346: 645–659.
74. Hartley DM, Zhao C, Speier AC, Woodard GA, Li S, et al. (2008) Transglutaminase induces protofibril-like amyloid- β protein assemblies that are protease-resistant and inhibit long-term potentiation. *J Biol Chem* 283: 16790–16800.
75. Soto C, Castaño EM, Frangione B, Inestrosa NC (1995) The α -helical to β -strand transition in the amino-terminal fragment of the amyloid β -peptide modulates amyloid formation. *J Biol Chem* 270: 3063–3067.
76. Frieden C (2007) Protein aggregation processes: In search of the mechanism. *Protein Sci* 16: 2334–2344.
77. Makin OS, Serpell LC (2005) Structures for amyloid fibrils. *FEBS J* 272: 5950–5961.
78. Heiring C, Muller YA (2001) Folding screening assayed by proteolysis: application to various cystine deletion mutants of vascular endothelial growth factor. *Protein Eng* 14: 183–188.
79. Pratico D (2008) Evidence of oxidative stress in Alzheimer's disease brain and antioxidant therapy. *Ann NY Acad Sci* 1147: 70–78.
80. Lai E, Teodoro T, Volchuk A (2007) Endoplasmic reticulum stress: signaling the unfolded protein response. *Physiology* 22: 193–201.
81. Paschen W (2003) Endoplasmic reticulum: a primary target in various acute disorders and degenerative diseases of the brain. *Cell Calcium* 34: 365–383.
82. Francoeur AM, Assalian A (1996) MICROCAT: A novel cell proliferation and cytotoxicity assay based on WST-1. *Biochemica* 3: 19–25.