Molecular, Cell and Tissue Studies of Galectin-3:CD98 Interactions in IPF and COVID Pneumonitis

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Mohammed Azim Miah MSc (Hons)

Respiratory Department, University of Leicester

June 2021

Abstract

Mohammed Azim Miah

Molecular, Cell and Tissue Studies of Galectin-3:CD98 Interactions in IPF and COVID Pneumonitis

Galectin-3 critically mediates experimental fibrosis and is implicated in idiopathic pulmonary fibrosis (IPF). It can bind multiple proteins through protein:glycan as well as protein:protein interactions, including self-oligomerisation. It binds to TGF-β receptor (TGF-βR) II and potentiates TGF-β1 signalling. It also interacts with CD98, a mediator of innate inflammatory responses in such cells. I tested the gal-3-fibrosome hypothesis, that galectin-3 mediates IPF pathogenesis by nucleating a macromolecular complex of pro-fibrotic factors including CD98:integrin complex and TGF-βRII at the cell surface, in lung tissue, cell lines, and primary human lung cells.

Co-localisation of galectin-3, β1-integrin and CD98 was supported by proximity ligation assay and immunohistochemistry in human lung tissue, and confocal microscopy and co-immunoprecipitation studies in cells. TGF-β1 stimulation increased expression and co-localisation of all three proteins in A549 cells. CD98 silencing abrogated epithelial wound healing similarly to galectin-3 knockdown. In myofibroblasts TGF-β1 stimulation increased protein levels of β1-integrin and CD98 whilst reducing galectin-3. However, co-localisation of galectin-3 with β1-integrin and CD98 appeared relatively preserved in IPF myofibroblasts. Galectin-3 oligomerisation was most tractable to electron microscopy (EM) analysis by cross-linking ligand-induced oligomers and peptide *N*-glycosidase treatment, but further optimisation is required ahead of cryo-EM studies. I successfully extracted endogenous CD98-containing complexes from cell membranes within nanodiscs, using styrene maleic acid (SMA). CD98 co-immunoprecipitated with galectin-3 in nanodiscs. Both CD98 and galectin-3 were essential for interleukin (IL)-6 and IL-8 responses to SARS-CoV-2 Spike receptor binding domain in lung epithelial and macrophage cell lines.

Together, my data support a gal-3-fibrosome response to disease-relevant stimuli in tissue and cells. In myofibroblasts, they suggest a negative feedback loop that is overridden in IPF. Extensive work has defined promising conditions for future structural studies. CD98-dependent inflammatory responses are observed across cell types as well as diverse injury stimuli. Galectin-3 interactions with CD98:β1-integrin may transduce acute injury responses to pro-fibrotic behaviour in lung tissue, e.g. in acute exacerbations of IPF.

Acknowledgements

The journey of completing my PhD has been an exhilarating rollercoaster. Safe to say, it has been one of the hardest task I have undertaken both mentally and physically. I therefore would like to honour those special few individuals who supported, encouraged, and guided me during my PhD.

Firstly, I would like to take out a moment to display my sincere appreciation to my supervisor Professor Bibek Gooptu ("Sir"). His professionalism, thoughtful planning, comprehensive inspection, and scrutiny impacted my research project to be thesis worth.

I am indebted to my fellow lab member Dr Panayiota Stylianou. I deeply thank her for the time, patience, and skills she has devoted on me. She has regularly been there for me as to when I needed assistance, second opinion and reassurance.

I sincerely thank Dr Omeed Darweesh along with the remainder of the Gooptu team members, and my extended colleagues within the Department of Respiratory Sciences in Clinical Sciences Wing, Glenfield Hospital Leicester. To Fahad, a relationship that began as mere lab colleagues, has now spiralled into life-long brotherhood, thank you for the moments of joy we have shared together.

I want to express my humble gratitude to my close circle of friends and family. To my sole sibling, Nazim, thanks brother. To my fiancé, Tahrin, thank you for your unconditional love through every thick and thin impediment I encountered. Lastly, yet above all others, to my Amma and Abba. Without both of your support and upbringing, I would have perhaps never risen so high. You both are a key reason I strived to be a doctor, to make you proud, and I dedicate this thesis to you both.

Table of Contents

Chapter 1 1

1.1 Interstitial Lung Disease and Pulmonary Fibrosis 2

1.2 IPF Demographics 8

1.3 IPF Risk Factors 9

1.3.1 Inhalational Risk Factors 10

1.3.1.1 Cigarette Smoking 10

1.3.1.2 Occupational and Environmental Risk Factors 11

1.3.2 Infection and the Respiratory Microbiome 12

1.3.3 Gastroesophageal Reflux Disease and Chronic Microaspiration 13

1.3.4 Genetics 14

1.4 Aberrant Wound Healing and Clinical Observations in IPF 18

1.5 Experimental Studies Defining IPF Pathogenesis 20

1.5.1 Histopathology on Human IPF Tissue (Ex Vivo) 20

1.5.2 Explant Model of IPF 22

1.5.3 Animal models of IPF (In Vivo and Ex Vivo) 23

1.5.4 In Vitro Cell Models of IPF 24

1.5.5 Inferences Taken from Model Studies of IPF 28

1.6 The Pro-Fibrotic Cytokine TGF-β and its Role in IPF 32

1.6.1 Integrins Regulate Mechanotransduction and Potentiation of TGF-β Signalling through Cell Matrix Interactions 33

1.6.2 TGF-β Signalling Through the Recruitment of TGF-β-receptors 35

1.6.3 Experimental Models Assessing the Role of TGF-β in an IPF Context 37

1.7 Galectin-3, a Key Player in IPF Pathogenesis? 38

1.7.1 Galectin-3: Synthesis, Domain Architecture, Structural Behaviour and Intermolecular Interactions 39

1.7.2 Galectin-3 Mediates Lung Inflammation and Fibrosis 42

1.7.3 TD139 Inhibits Galectin-3 in IPF Lung 43

1.8 CD98, Structure, Function, and Inflammatory Relevance 44

1.8.1 CD98:β-integrin Extracellular Matrix Interactions 45

1.9 Interactions of CD98hc and β1-integrin with Galectin-3 in IPF 47
1.10 Hypothesis 48
1.11 A Global Pandemic - Covid Pneumonitis and its Severity with Pre-Existing ILD 50
1.11.1 Topical Objective: The Putative Gal-3-Fibrosome Hypothesis Potentially Mediates Covid Pneumonitis 51
1.12 Aims and Objective 53
Chapter 2 54
2.1 Cell Culture 55
2.1.1 Cell Line 55
2.1.1.1 A549 Cell Culture (in vitro) 55
2.1.1.2 Screening A549 Cell Stocks for Mycoplasma Contamination 56
2.1.1.3 THP-1 Cell Culture (in vitro) 56
2.1.1.4 Differentiating THP-1 Cells into M0 Macrophage Phenotype 57
2.1.2 Primary Cell Isolation and Culture (ex vivo) 57
2.1.2.1 Non-fibrotic Control Derived Human Lung Myofibroblasts 57
2.1.2.2 IPF Derived Human Lung Myofibroblasts 58
2.2 Cell Stimulation with the Pro-Fibrotic Cytokine TGF-β1 59
2.3 Lysate Extraction and Western-blotting 59
2.3.1 Whole cell lysate (WCL) Preparation from Human Cells 59
2.3.2 Cytosol and Membrane Lysate Fractionation (MEM-PER Plus Membrane Protein Extraction Kit) 60
2.4 Quantification of Sample Concentrations 62
2.4.1 Quantification of Whole Cell and Membrane-Cytosol Fractionated Lysate Concentrations using Bicinchoninic Acid (BCA) Assay 62
2.4.2 Quantification of Purified Recombinant Protein Lysate Concentrations Using Bicinchoninic Acid (BCA) Assay 63
2.4.3 Quantification of SMALP Protein Lysate and DNA Concentrations Using Nanodrop Spectrophometer 63
2.5 Polyacrylamide Gel Electrophoresis (PAGE) and Protein Visualisation 64
2.5.1 PAGE Conditions 64
2.5.1.1 SDS-PAGE for Whole Cell, Membrane-Cytosol Fractionated and SMALP Lysates 64
2.5.1.2 SDS-PAGE for Recombinant Purified Proteins 65
2.5.1.3 Non-denaturing (native)-PAGE for Recombinant Purified Proteins 66
2.5.1.4 Styrene Maleic Acid non-denaturing (SMA native)-PAGE for SMALP Extracts 69

2.5.2 PAGE-Based Quantitation of Specific Protein Species 69
2.5.2.1 Detection of Proteins in Whole Cell Lysate (Including Membrane/Cytosol Fractions) 69
2.5.2.2 Detection of the SMALP Lysed Proteins 70
2.5.2.3 Detection of the Recombinant Purified Protein Species from SDS or Non-denaturing PAGE 70
2.5.2.4 Western Blot Densitometric Analysis 71
2.5.3 qRT-PCR to Assess mRNA Transcripts of Galectin-3 and CD98 71
2.6 Immunoprecipitation 72
2.6.1 Preparation of Dynabeads™ Protein G 73
2.6.3 Antibody-Lysate-Bead Complex (Protein Elution in a Denaturing Condition) 73
2.6.3.1 Antibody-Lysate-Bead Complex (Protein Elution in a Non-Denaturing (Native) Condition) 74
2.7 Immunostaining Microscopy and Staining Quantification 74
2.7.1.1 Direct Staining Immunofluorescence Microscopy 74
2.7.1.2 Indirect staining immunofluorescence microscopy 75
2.7.1.3 Co-immunostaining of CD98hc and β1-integrin with Galectin-3 in A549, Non-Fibrotic-Control and IPF Human Lung Myofibroblast Cells 75
2.7.2.1 Single Channel Immunostaining Quantification 76
2.7.2.2 Co-Localisation Immunostaining Quantification 76
2.7.2.3 Co-immunostaining of Lung Tissue Using the Duolink® Proximity Ligation Assay Kit 77
2.7.3 Immunohistochemistry on NFC and IPF Human Lung Tissue 79
2.7.3.1 Glycol Methacrylate (GMA) Resin Embedding 80
2.7.3.2 Dako EnVision FLEX + Staining Technique 80
2.8 Wound Healing Assay in A549 Cells Subjected to CD98hc Silencing 81
2.8.1 A549 Cell Culture for CD98 Knockdown 81
2.8.2 Transfection and siRNA Knockdown Optimisation 81
2.8.3 Wound Healing Response of CD98 Depleted A549 Cells in the Presence of Pro-Fibrotic Stimulation 83
2.9 Flow Cytometry 84
2.9.1 Testing the Cell Surface Expression and Distribution of CD98hc in A549 and HLMF Cells 84
2.9.2 Analysis of Flow Cytometry Data 86
2.10 Molecular Biology of Galectin-3 86
2.10.1 Preparation of Recombinant Galectin-3 Glycerol Stocks 86
2.10.2 Recombinant Galectin-3 Sequence Confirmation Through Plasmid Purification 88

2.10.3 Protein Expression of Recombinant Galectin-3 88
2.10.4 Protein Purification of Recombinant Galectin-3 89
2.10.4.1 Preparation of the Divinyl Sulfone Activated Lactosyl-Sepharose Affinity Column 89
2.10.4.2 Packing the Affinity Chromatography Column with Lactosyl-Sepharose 91
2.10.4.3 Galectin-3 Pellet Lysis (Protein Solubilisation) and Protein Elution 91
2.10.4.4 Preparing Galectin-3 for Storage 92
2.10.4.5 Size Exclusion Chromatography 92
2.10.4.6 Size Exclusion Chromatography Multi Angle Light Scattering (SEC-MALS) 93
2.10.4.7 Sucrose Density Gradient Ultracentrifugation 94
2.10.4.8 Galectin-3 Induced Oligomerisation 95
2.10.4.9 Crosslinking Galectin-3:ASF Complexes Within a Sucrose Density Gradient 95
2.10.4.10. Negative Stain Electron Microscopy (EM) 96
2.10.4.11 Exposed Sugar Chain Cleavage using PNGase Treatment (Non-Denaturing) 97
2.11 Covid-19-Related Project Methods 98
2.11.1 SARS-CoV-2 Construct Generation 2.11.1.1 Oligonucleotide Design 98
2.11.1.2 DNA Amplification by Polymerase Chain Reaction (PCR): Strategy 99
2.11.1.3 DNA Amplification by Polymerase Chain Reaction (PCR): Set-Up 101
Primer Annealing Temperature and Extension for Each Construct: 102
2.11.1.4 Agarose Gel Preparation and Visualisation for PCR Product Analysis 102
2.12 Cloning and Plasmid Preparation 103
2.12.1.1 Cloning Strategy 103
2.12.1.2 Preparation of NEB Competent Cells (Large Scale Stock) 105
2.12.1.3 Competent Cell Transformation 105
2.12.1.4 Colony Screening PCR (Pre-Sequencing Sense Check) 106
2.12.1.5 Large Scale Plasmid Isolation (Maxi prep) in Alkaline Conditions 106
2.13 HEK 293 Suspension Cell Culture and DNA Transfection 109
2.13.1 Cell Line 109
2.13.1.1 HEK 293F Cell Culture 109
2.13.2 DNA Transfection 109
2.13.3 Protein Expression and Purification 110
2.13.3.1 FLAG-tag Purification from Cell Pellet and Media 110
2.13.3.2 His-tag Purification from Protein Secreted into Cell Media 111

2.13.4 Enzyme-Linked Immunosorbent Assay (ELISA) 113
2.14 Statistical Analysis 114
Materials 115
2.15 General Chemicals and Reagents: 115
2.16 Cell Culture Media 117
2.17 Buffers and Solutions 117
2.18 Regents and material for Western-blot: 119
2.19 Antibodies 119
2.20 CD98hc Knock-Down and Scratch Test Materials 121
Chapter 3 122
3.1 Introduction 123
3.2 Results 126
3.2.1 Levels of the Putative Galectin-3-Fibrosome Components and Their Co-Localisation Addressed in the Epithelium of Non-Fibrotic-Control and IPF Lung Tissue 126
3.2.2 Level of Putative Gal-3-Fibrosome Components in A549 Cells 129
3.2.2.1 Levels of Galectin-3, β1-integrin and CD98 Increases Following TGF-β1 Stimulation 129
3.2.2.2 Endogenous TGF-βRII Levels were Undetectable 133
3.2.3 Co-Localisation of the Putative Galectin-3-Fibrosome Components in A549 Cells 134
3.2.3.1 Co-localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Increases Following TGF-β1 Treatment 134
3.2.3.2 Co-Localisation Addressed by Co-Immunoprecipitation (co-IP) 141
3.2.3.3 Proximity Ligation Assay of Galectin-3 with CD98 and β1-integrin in Ex Vivo Lung Tissue Model of Early Fibrosis 145
3.2.4 Functional Cross-Talk Between CD98 and an Epithelial Readout of Pro-Fibrotic Behaviour 147
3.2.4.1 CD98 Mediates Wound Healing Responses in A549 Cells 149
3.2.5 Flow Cytometry to Characterise Sub-Cellular Localisation of CD98hc in A549 cells 151
3.2.5.1 Confirming the Specificity of the CD98-44D7 Antibody Used in Flow Cytometry 154
3.3 Discussion 156
Chapter 4 160
4.1 Introduction 161
4.2 Results 163
4.2.1 Characterisation of Ex vivo Derived Human Lung Myofibroblast Phenotype 163
4.2.1.1 Patient Characteristics 163
4.2.1.2 Characterisation of Primary Human Lung Myofibroblasts 165

4.2.2 Level of Putative Galectin-3-Fibrosome Components in Myofibroblasts 168

4.2.2.1 Levels of Galectin-3, CD98hc and β1-integrin in NFC and IPF HLMF are Altered Following TGF-β1 Stimulation 168

4.2.3 Is CD98hc Expressed and Distributed on the Cell Surface of HLMF 174

4.2.3.1 Cell surface Expression of CD98hc in HLMFs 174

4.2.4 Optimising Immunofluorescence Studies of Cell Surface Co-localisation of Galectin-3 with CD98 and β1-integrin in Human Lung Myofibroblasts 177

4.2.4.1 Co-Localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Decreases Following TGF-β1 Treatment in Non-Fibrotic-Control Human Lung Myofibroblasts ... 179

4.2.4.2 Co-Localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Remains Stable and Unchanged Following TGF-β1 Treatment in IPF Human Lung Myofibroblasts 186

4.2.5 Co-Localisation Addressed by Co-immunoprecipitation (co-IP) 193

4.3 Discussion 195

Chapter 5 199

5.1 Introduction 200

5.2 Results 206

5.2.1 Applying Structure Prediction Algorithms to the Galectin-3 Sequence 206

5.2.2 Galectin-3 Purification 210

5.2.2.1 Galectin-3 Plasmid Sequence Confirmation 210

5.2.2.2 Purification of Recombinant Galectin-3 212

5.2.2.3 Maximising Galectin-3 Elution 214

5.2.3 Assessing for Concentration-Dependent Multimerisation of Galectin-3 216

5.2.3.1 Characterisation of Purified Galectin-3 by Size Exclusion Chromatography and Multi Angle Light Scattering (SEC-MALS) 216

5.2.3.2 Sucrose Density Gradient Ultracentrifugation of Purified Galectin-3 220

5.2.3.3 Native-PAGE Analysis of Purified Galectin-3 222

5.2.4 Induced Galectin-3 Oligomerisation. 228

5.2.4.1 Optimisation of Glutaraldehyde Cross-linked Galectin-3 228

5.2.4.2 Size Exclusion Chromatography Characterisation of Cross-linked Species Following Glutaraldehyde Treatment of Monomeric Galectin-3 230

5.2.5 Optimisation of Ligand Induced Galectin-3 Oligomerisation 233

5.2.5.1 Galectin-3 Induced Oligomerisation with the Glycan-Ligand N-Acetyl-D-Lactosamine (LacNAc) 234

5.2.5.1.2 Stochastic Behaviour When Cross-linking Galectin-3:LacNAc 237

5.2.5.1.3 Characterisation of Cross-linked Galectin-3:LacNAc by Size Exclusion Chromatography 239
5.2.6 Optimisation of Galectin-3 Induced Oligomerisation with the Glycoprotein Ligand Asialofetuin 242
5.2.6.1 Size Exclusion Chromatography of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking 244
5.2.6.2 Size Exclusion Chromatography Characterisation of Cross-linked Species Formed due to Galectin-3:ASF Interactions 246
5.2.7 Characterisation of Cross-linked Galectin-3:ASF Oligomeric Species using Sucrose Density Gradient Ultracentrifugation 250
5.2.7.1 Sucrose Density Gradient Ultracentrifugation of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking 250
5.2.7.2 Sucrose Density Gradient Ultracentrifugation of Cross-linked Galectin-3:ASF 253
5.2.8 Negative Stain Electron Microscopy Assessment of Galectin-3 Induced Oligomerisation with the Glycoprotein Ligand Asialofetuin 255
5.2.8.1 Negative Staining Electron Microscopy of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking 255
5.2.8.2 Negative Stain Electron Microscopy Characterisation of Cross-linked Species Formed due to Galectin-3:ASF Interactions 258
5.2.8.3 Peptide N-Glycosidase F Treatment of Cross-Linked Galectin-3:ASF Species 260
5.2.9 Isolation of Membrane Distributed Proteins Within Their Native Lipid Bilayer for Structural Analysis 264
5.2.9.1 Generation of SMALP A549 Cell Lysates Under Stimulated Conditions 264
5.2.9.2 CD98hc Enrichment from A549 SMALP Lysates 268
5.3 Discussion 272
Chapter 6 277
6.1 Introduction 278
6.2 Results 281
6.2.1 Generation of Spike Protein Constructs 281
6.2.1.1 Generation of Spike Protein Constructs and Steps 281
6.2.2 Protein Expression and Purification 284
6.2.2.1 Initial Protein Expression 284
6.2.2.2 Troubleshooting to Confirm TEV-Protease is Functional 289
6.2.2.3 CD5-RBD Expression 291
6.2.3 Optimisation of Purified RBD Concentration for Cellular Stimulations to Release Inflammatory Cytokines in A549 and THP-1 Cells 293

6.2.4 Galectin-3-Fibrosome Regulating RBD Induced Pro-Inflammation 298

6.2.4.1 Inhibiting CD98hc using Cynaropicrin Attenuates RBD Induced IL-6 and IL-8 Release, but not TNF-α Release, in Both A549 and THP-1 Cells..... 298

6.2.4.2 Effects of Silencing CD98hc and Galectin-3 upon RBD-Induced Cytokine Release in A549 cells..... 301

6.2.5 Effects of Dexamethasone Treatment upon RBD-Induced Cytokine Release in A549 and THP-1 cells 306

6.3 Discussion 310

Chapter 7..... 315

General Discussion..... 315

7.1 General Discussion..... 316

7.2 Future Work 325

Appendix 328

References..... 341

List of Figures

Figure 1.1, Current accepted classification of interstitial lung disease……4

Figure 1.2, HRCT scan of IPF ……6

Figure 1.3, Histological Identification of IPF……21

Figure 1.4, A diverse range of cell types may contribute to myofibroblast numbers and behaviours in fibrotic diseases…… 27

Figure 1.5, Cellular level of IPF pathogenesis model……30

Figure 1.6, Integrin dependent mechanotransduction:cell and TGF-β1 signalling……34

Figure 1.7, Schematic representation of the TGF-β1 signalling cascade through the TGF-β-receptors…36

Figure 1.8, Cartoon representation of human galectin family……39

Figure 1.9, Galectin typical CRD arrangement and accommodation of β-galactoside……40

Figure 1.10, Working model of the galectin-3-fibrosome organisation……49

Figure 2.1, pLeics-49 Plasmid Map……98

Figure 2.2, Full length Sars-CoV-2 Structure in a prefusion conformation……99

Figure 3.1, Immunohistochemistry staining focusing on the levels and co-localisation of galectin-3, CD98hc and β1-integrin between NFC and IPF derived lung tissue serial sections……127

Figure 3.2, Quantification of galectin-3, CD98 and β1-integrin levels from the immunohistochemistry-stained NFC and IPF lung tissue……128

Figure 3.3, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in A549 cells by western-blotting……131

Figure 3.4, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin levels from figure 3.4 immunoblotting……132

Figure 3.5, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in A549 cells……136

Figure 3.6, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in A549 cells……133

Figure 3.7, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations……140

Figure 3.8, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations……141

Figure 3.9, Monoclonal antibody pulldown from whole cell and membrane-cytosol fractions of A549 lysed cells……143

Figure 3.10, Proximity ligation assay staining focusing on the co-localisation of galectin-3 with CD98hc (a) and β1-integrin (b) on pre-treated TGF-β1 NFC ex vivo lung tissue defining early fibrosis……..146

Figure 3.11, Western-blot monitoring the expression levels of basal CD98 following siRNA treatment in A549 cells against relevant controls……148

Figure 3.12, Scratch test in CD98 silenced A549 cells and the quantification of the wound healing response……150

Figure 3.13, Flow cytometry FloJo analysed histograms of CD98 staining monitoring signal of CD98hc in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) A549 cells…152

Figure 3.14, Statistical quantification of CD98hc signal analysed from FloJo in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) A549 cells……153

Figure 3.15, Confirming the specificity of CD98-44D7 antibody used for flow-cytometry in A549 cells……155

Figure 4.1, Characterising the human lung myofibroblast (HLMF) from healthy and IPF donors. HLMF characterisation using five positive controls……166

Figure 4.2, Characterising the human lung myofibroblast (HLMF) from healthy and IPF donors. HLMF characterisation using two negative controls……167

Figure 4.3, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in NFC myofibroblast cells by western-blotting……170

Figure 4.4, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin levels from figure 4.3 immunoblotting……171

Figure 4.5, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in IPF myofibroblast cells by western-blotting……172

Figure 4.6, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin levels from figure 4.5 immunoblotting……173

Figure 4.7, Flow cytometry FloJo analysed histograms of CD98 staining monitoring signal of CD98hc in non-permeabilised (cell membrane intact) NFC myofibroblast cells both basally and in response to TGF-β1 stimulation……175

Figure 4.8, Statistical quantification of CD98hc signal following TGF-β1 stimulation, analysed from FloJo in non-permeabilised NFC myofibroblast cells……176

Figure 4.9, Comparing the effect of immunostaining distribution when using different fixation methodologies (Methanol vs 4% (w/v) Paraformaldehyde (PFA))……178

Figure 4.10, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in NFC HLMF……180

Figure 4.11, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in NFC HLMF……182

Figure 4.12, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations in NFC HLMF……184

Figure 4.13, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations in NFC HLMF……185

Figure 4.14, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in IPF HLMF……187

Figure 4.15, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in IPF HLMF……189

Figure 4.16, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations in IPF HLMF……191

Figure 4.17, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations in IPF HLMF……192

Figure 4.18, Anti-galectin-3 pulldown of whole cell NFC HLMF lysates. CD98 co-IP (a) and β1-integrin co-IP (b)……194

Figure 5.1, X-ray crystallographic structure of the NTS and repeat IX-and a structural model defining the CRD with the two new regions identified in the Gal-3[NTS/VII-IX] variant……201

Figure 5.2, SMA mediated membrane protein solubilisation and purification……204

Figure 5.3, SMALP CD98hc purification approach……205

Figure 5.4, Structure prediction using the program Protein Homology/analogY Recognition Engine (Phyre2) and inputting the Human Galectin-3 (LGALS3) Amino Acid Sequence……208

Figure 5.5, Structure prediction using the program Glob-Plot to assess human galectin-3 amino acid sequence……209

Figure 5.6, Sequence alignment of human Galectin-3 (LGALS3)……211

Figure 5.7, Multi-steps in galectin-3 purification……213

Figure 5.8, Step elution titration experiment to maximise galectin-3 elution……215

Figure 5.9, Initial Galectin-3 Characterisation……218

Figure 5.10, SDS-PAGE analysis of size exclusion chromatography elution fractions (from figure 5.9a)…… 219

Figure 5.11, Sucrose density gradient ultracentrifugation fractionation of purified galectin-3……221

Figure 5.12, Native-PAGE analyses of galectin-3……223

Figure 5.13.1, Reversed-polarity native-PAGE of galectin-3……226

Figure 5.13.2, pH>pI native-PAGE of galectin-3 (normal polarity, higher Tris concentrations)……..226

Figure 5.13.3, pH>pI native-PAGE of galectin-3 (normal polarity, with standard Tris concentrations)…227

Figure 5.13.4, pH<pI native-PAGE of galectin-3 (reversed polarity)……227

Figure 5.14, Titration experiment to optimise final glutaraldehyde concentration (v/v) yielding the best galectin-3 oligomeric species……229

Figure 5.15, Size exclusion chromatography analysis of galectin-3 ± cross-linking……..232

Figure 5.16, Titration experiment optimising LacNAc concentration and incubation with 50 µM galectin-3 to yield tractable oligomeric species……236

Figure 5.17, Cross-linking galectin-3 with LacNAc inconsistently forms precipitates……238

Figure 5.18, Size exclusion chromatography analysis of galectin-3 with LacNAc ± cross-linking……..240

Figure 5.19, Titration experiment optimising galectin-3 and ASF concentrations at 1 h incubation to assess best yielded oligomeric species……243

Figure 5.20, Size exclusion chromatography analysis of ASF ± crosslinking……245

Figure 5.21, Size exclusion chromatography analysis of galectin-3 with ASF ± cross-linking……..247

Figure 5.22, Sucrose density gradient ultracentrifugation profile of galectin-3 ± cross-linking……251

Figure 5.23, Sucrose density gradient ultracentrifugation profile of ASF ± cross-linking……252

Figure 5.24, Sucrose density gradient ultracentrifugation profile of galectin-3:ASF ± cross-linking…..254

Figure 5.25, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of galectin-3 and ASF controls in the absence and presence of cross-linking……..257

Figure 5.26, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of incubated galectin-3:ASF in absence and presence of cross-linking……259

Figure 5.27, Size exclusion chromatography analysis of cross-linked galectin-3:ASF subjected to deglycosylation……262

Figure 5.28, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of PNGase F Treated cross-linked galectin-3:ASF……263

Figure 5.29, Assessing the ability of SMA-2000P to form SMALP lysates in TGF-β1 and LPS stimulated A549 cells with CD98hc encapsulated within the nanodisc……..266

Figure 5.30, Assessing the ability of SMA-30010p to form SMALP lysates in TGF-β1 and LPS stimulated A549 cells with CD98hc encapsulated within the nanodisc……267

Figure 5.31, Enriching CD98hc bound proteins by applying anti-CD98hc pulldown on TGF-β1 and LPS stimulated SMALPs and eluting these species in a native condition……270

Figure 5.32, Negative stain-EM analysis on stimulated SMALP lysates subjected to CD98hc immunoprecipitation and native condition sample elution……..271

Figure 6.1, Multi-step approach in DNA cloning of S1, NTD and RBD domains into vector pLeics-49..283

Figure 6.2, Anti-FLAG pulldown of FLAG-Tagged Spike protein constructs. 6.2a (RBD), b (S1) and e (FL spike protein)……286

Figure 6.3, Troubleshooting experiment to confirm TEV-cleavage buffer used is functional……290

Figure 6.4, RBD purification from cell media using Ni-NTA affinity column……292

Figure 6.5, Titration experiment to assess optimal RBD concentration inducing IL-6, IL-8 and TNF-α release in A549 cells……..295

Figure 6.6, Titration experiment to assess optimal RBD concentration inducing IL-6, IL-8 and TNF-α release in THP-1 cells……296

Figure 6.7, RBD titration concentration and stimulation duration does not alter cell proliferation/survival in THP-1 cells, illustrated by MTS assay……297

Figure 6.8, Co-stimulation of RBD with cynaropicrin reduces IL-6 and IL-8 release, though TNF-α release remains stabilised in A549 cells……..300

Figure 6.9, Co-stimulation of RBD with cynaropicrin reduces IL-6 and IL-8 release, though TNF-α release remains stabilised in THP-1 cells……301

Figure 6.10, Western-blot assessing the ability to silence CD98 in THP-1 cells as successfully done in A549 cells……..303

Figure 6.11, CD98 silencing abrogates RBD induced IL-6 and IL-8 release, though TNF-α levels remain stabilised in A549 cells……304

Figure 6.12, Galectin-3 silencing abrogates RBD induced IL-6 and IL-8 release, though TNF-α levels remain stabilised in A549 cells……305

Figure 6.13, Co-stimulation of RBD with dexamethasone reduces IL-6, IL-8 and TNF-α release in A549 cells……308

Figure 6.14, Co-stimulation of RBD with dexamethasone reduces IL-6, IL-8 and TNF-α release in THP-1 cells……309

Figure 7.1, Schematic displaying the protein responses in TGF-β1 pro-fibrotic conditions on the cellular level of different cell types (inferences taken from summary Table 7.1) which can cross-talk and possibly modulate overall tissue responses……320

Figure 7.2, Schematic displaying acute inputs and its associated pro-inflammatory outputs, along with the incorporation of the gal-3-fibrosome hypothesis……324

Figure A1, Negative control staining of NFC lung tissue sections to confirm the specificity of immunohistochemistry staining focusing on the levels and co-localisation of galectin-3, CD98 and β1-integrin between NFC and IPF derived lung tissue serial sections……329

Figure A2, Zoomed projections of the stained panels of figure 3.1, displaying Gal-3-fibrosome proteins localising on the epithelium with increasing amounts in IPF subjects……330

Figure A3, Assessment of mycoplasma contamination in A549 cells using a PCR based approach……331

Figure A4, Full western-blot membrane (uncropped) of figure 3.3……332

Figure A5, Isotype control staining to confirm the specificity of the co-immunofluorescence staining of galectin-3, CD98hc and β1-integrin in A549 cells……..333

Figure A6, Cellular gating strategy for flow cytometry analysis of CD98hc staining signal in A549 cells……..334

Figure A7, Full western-blot membrane (uncropped) of figure 4.3……335

Figure A8, Full western-blot membrane (uncropped) of figure 4.5……336

Figure A9, Cellular gating strategy for flow cytometry analysis of CD98hc staining signal in NFC cells……..337

Figure A10, Gene expression levels of galectin-3 and CD98 from HLMF and lung tissue was assessed in basal and TGF-β1 stimulated conditions using qRT-PCR……..338

Figure A11, Immunofluorescence staining to assess cellular retention of SARS-CoV-2 following CD98 and galectin-3 knockdown……339

Figure A12, SARS-CoV-2 RBD pulldown in A549 whole cell lysates……340

List of Tables

Table 2.1, Native PAGE Stock Solutions……68

Table 2.2, Native PAGE Gel Mixes……68

Table 2.3, Luria Bertani media and agar mix per litre……87

Table 2.4, KOD Hot start DNA polymerase: 50µl total reaction volume……101

Table 2.5, PCR Thermocycler Settings……101

Table 2.6, ELISA standards range for each cytokine being measured (DuoSet Ancillary Reagent kit 2 (DY008)……114

Table 2.7, Chemicals and reagents utilised in this study……115

Table 2.8, Cell culture media used in this study……117

Table 2.9, Whole cell lysate (RIPA) and detergent free membrane (SMA) protein extraction lysis buffer……117

Table 2.10, Membrane cytosol Protein extraction lysis buffer (from MEM-PER™Plus Membrane extraction Kit (89842), ThermoScientific)……118

Table 2.11, Immunocytochemistry Buffers……118

Table 2.12, PAGE running and Western Blot Buffers……118

Table 2.13, Western-blot material……119

Table 2.14, Primary antibodies used in this study……119

Table 2.15, Secondary antibodies used in this study……121

Table 2.16, Reagents for Knock-down and scratch test stimulants……121

Table 4.1, Clinical characteristics of NFC and IPF patients……164

Table 5.1, Initial Galectin-3 Characterisation……218

Table 5.2, SEC-MALS analysis on galectin-3 and ASF cross-linked eluted fractions from both peak 1 and 2……249

Table 6.1, Summary of the expression of each SARS-CoV-2 protein constructs……288

Table 7.1, Summary of individual gal-3-fibrosome protein/ and mRNA gene expression responses in different pro-fibrotic experimental contexts (Chapters 3 and 4). In the alveolar epithelial cell (AEC) model- A549 cells……319

List of Abbreviations

αSMA	Alpha smooth muscle actin
AE	Acute exacerbation
AECs	Alveolar epithelial cells
ASM	Airway smooth muscle cells
BALF	Bronchoalveolar lavage fluid
BME	Beta-Mercaptoethanol
bp	Base pairs
BSA	Bovine Serum Albumin
Ca^{2+}	Calcium
CRD	Carbohydrate recognition domain
CTD	C-terminal domain
DAPI	4’,6-diamidino-2-phenylindole
$_dH_20$	Distilled water
DNA	Deoxyribonucleic acid
DPBS	Dulbecco’s phosphate buffered saline
ECM	Extracellular matrix

EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMT	Epithelial to mesenchymal transition
ER	Endoplasmic reticulum
FEV_1	Forced expiratory volume in the first second
FPLC	Fast protein liquid chromatography
FSP	Fibroblast surface protein
FVC	Forced vital capacity
Gal-3	Galectin-3
GMA	Glycolmethacrylate
GORD	Gastroesophageal reflux disease
h	Hour
HBSS	Hank's balanced salt solution
HLMF	Human lung myofibroblast
HRCT	High resolution computed tomography
IFN	Interferon
IIP	Idiopathic interstitial pneumonia

IL	Interleukin
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis
LISCB	Leicester Institute of Structural and Chemical Biology
LPS	Lipopolysaccharide
MALS	Multi angle light scattering
MMPs	Matrix metalloproteinases
Na^+/K^+	Sodium-potassium ATPase
NEAA	Non-essential amino acids
NFC	Non-fibrotic control
NHS	National Health Services
NO	Nitric oxide
NTD	N-terminal domain
O.D.	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PFA	Paraformaldehyde

PMSF Phenylmethanesulfonylfluoride

PNGase Peptide:N-glycosidase

PVDF Polyvinylidene Fluoride

RIPA Radio-Immuno-Precipitation-Assay

RONS Reactive oxygen and nitrogen species

SDS Sodium dodecyl sulfate

PAGE Polyacrylamide gel electrophoresis

SEC Size exclusion chromatography

SMA Styrene maleic acid

SMALP Styrene maleic acid lipid particles

TBE Tris-borate-EDTA

TGF-β Transforming growth factor-beta

TGF-βR Transforming growth factor-beta receptor

TNF Tumour necrosis factor

UIP Usual interstitial pneumonias

v/v Volume to volume

w/v Weight to volume

Chapter 1

Introduction

1.1 Interstitial Lung Disease and Pulmonary Fibrosis

In the UK, ~4000 new patients are diagnosed with interstitial lung disease (ILD) yearly, with apparent increase observed in its annual incidence rate within the last decade (NHS. 2019). ILD is responsible for a major clinical burden globally (Ortega and Molina. 2019, Sweeney *et al*. 2020). The lung interstitium is the interface between the air-filled alveoli (surface area of ~75 m^2) (Pavelka and Roth. 2010, Stanford Medicine. 2021), and the vast pulmonary capillary bed that accommodates the entire cardiac stroke volume (Panagiotou *et al*. 2017, Stanford Medicine. 2021). It is therefore bounded by alveolar epithelial cells on one side, and pulmonary capillary endothelium on the other (Nova *et al*. 2019). The major cells within healthy interstitium are mesenchymal cells, namely fibroblasts and myofibroblasts (White. 2015, Habiel and Hogaboam. 2017), which synthesise extracellular matrix components (White. 2015). The interstitium has to accommodate the volume changes of the respiratory cycle and represents the barrier that oxygen must cross to keep blood saturation acceptable for life (West. 2009).

During ILD the interstitium is affected by inflammatory changes and/or scarring (Wallace *et al*. 2007, Leslie. 2008). The former are characterised by immune cell infiltrate (Semenzato *et al*. 2000), acute inflammatory cytokine release (Keane. 2008), and oedema (Bourke. 2006). Scarring within the interstitial compartment of the lung is described as pulmonary fibrosis (Lasky and Brody. 2000). It is characterised by surplus deposition of extracellular matrix (ECM) (Burgess *et al*. 2016) and may be distributed within alveolar walls, perilymphatic, peribronchiolar and/or perivascular connective tissue (Scadding. 1974, Andoh *et al*. 1992, Raghu *et al.* 2018). This condition is a result of excessive wound healing response to both acute and chronic lung injury (White and Mantovani. 2013).

Inflammation is a reversible process (Ikegami *et al*. 2005, Kenyon and Last. 2005, Yao *et al.* 2008) whilst fibrosis is generally considered physiologically and pharmacologically irreversible (Kaur *et al.* 2017, Spagnolo *et al*. 2020, British Lung Foundation. 2021).

Both inflammation and fibrosis cause clinical disease, impairing alveolar gas exchange (Plantier *et al.* 2018) whilst increasing lung stiffness and hence the work of breathing (Kim *et al*. 2020). When sufficient lung tissue is affected, ILD can lead to end-stage respiratory failure and premature death (White and Mantovani. 2013).

The term ILD in fact covers a heterogeneous group of diseases currently divided into >200 ILDs (Flaherty *et al*. 2017). ILD may be classified into specific sub-types by various criteria (Figure 1.1), e.g. by one or more of: apparent aetiology, radiological features, histopathological characteristics, molecular markers or behaviour over time (American Thoracic Society. 2001, 2002, Tzouvelekis *et al.* 2005, Raghu *et al.* 2008, Lota and Renzoni. 2012, Antoniou *et al.* 2014, Mancano. 2018, Walsh and Kolb. 2018).

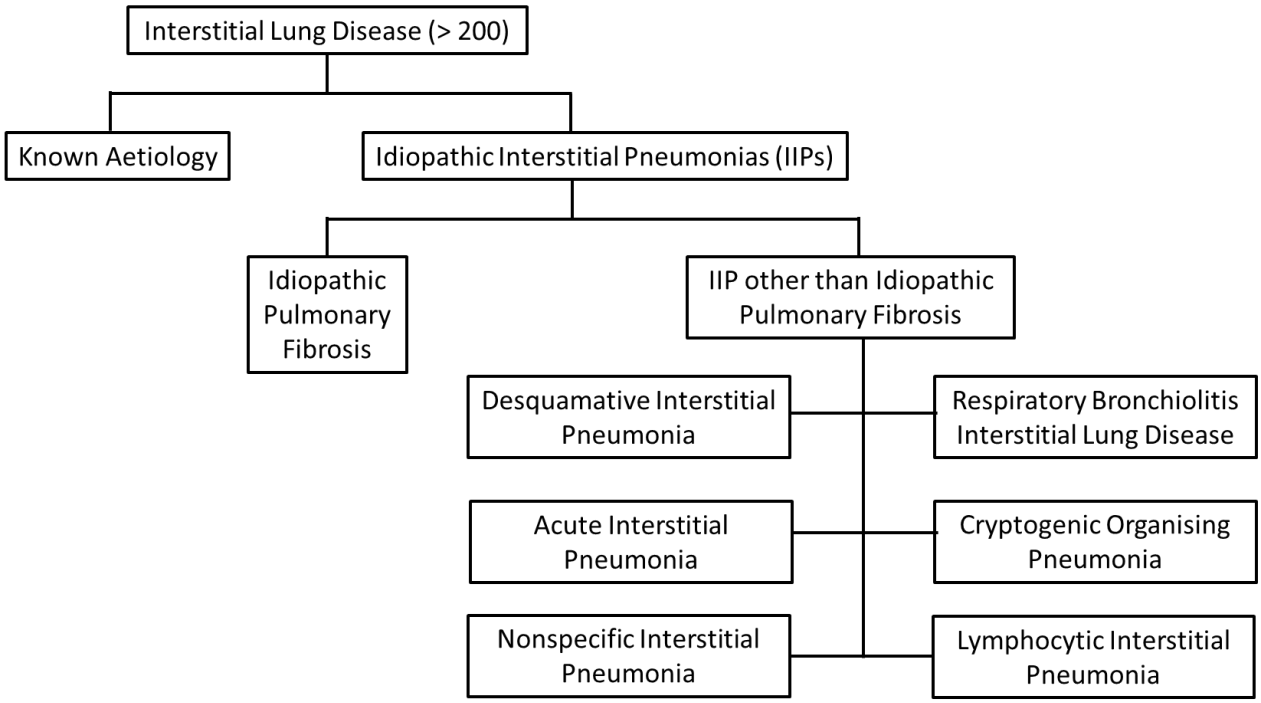


Figure 1.1, Current accepted classification of interstitial lung disease. Interstitial Lung Disease (ILD) consists of disorders of known aetiology (examples include auto-immune, environmental or drug related) as well as disorders of unknown cause (IIPs). The most important distinction between the IIPs is between idiopathic pulmonary fibrosis (IPF) and IIP other than IPF which include nonspecific interstitial pneumonia (a provisional term), desquamative interstitial pneumonia, respiratory bronchiolitis-associated interstitial lung disease, acute interstitial pneumonia, cryptogenic organizing pneumonia, and lymphocytic interstitial pneumonia. (American Thoracic Society. 2001, 2002, Tzouvelekis *et al*. 2005, Raghu *et al*. 2008, Lota and Renzoni. 2012, Antoniou *et al*. 2014, Mancano. 2018, Walsh and Kolb. 2018).

The medical definition of the word idiopathic 'relates to a disease which arises spontaneously or for which the cause is unknown' (MedicineNet. 2020). This term could now be deemed misleading since 'idiopathic' aetiology of pulmonary fibrosis is no longer considered a disease of 'unknown cause' (cryptogenic) (Wells *et al.* 2019). Instead, it refers to disease processes that arise primarily as a result of abnormal responses of cells within the lung parenchyma, excluding those of the adaptive immune system, to common factors (Wells *et al.* 2019). This makes the

current accepted ILD classification somewhat redundant and potentiates revaluation of the classification system criteria (Wells *et al.* 2019). Conversely, inflammatory ILDs are driven by adaptive cell responses to inhaled triggers (e.g. hypersensitivity pneumonitis (Selman *et al.* 2012)) or systemically distributed autoantigens (e.g. rheumatoid arthritis-related and connective tissue disease-related ILDs) (Vij *et al*. 2011, 2013). ILDs that are primarily inflammatory are typically treated with corticosteroids as broad-spectrum anti-inflammatory agents, and/or treatments which reduce adaptive immune cell activity more specifically (Bouros and Antoniou. 2005, Serrano *et al*. 2012). Such treatments are logical and have been used for many years (Wells and Hirani. 2008).

Some of the idiopathic interstitial pneumonias (IIPs) may involve similar immune cell types and give similar patterns but exogenous or systemic stimuli are not identified (Bouros and Antoniou. 2005). In these cases, similar treatments are used with varying consistencies of benefit. In contrast, the prototypic IIP idiopathic pulmonary fibrosis (IPF) is characterised by chronic progressive fibrosis, with histopathological appearances of usual interstitial pneumonia (UIP) (Cavazza *et al*. 2010, Wuyts *et al*. 2014). UIP refers to a morphologic entity defined by patchy interstitial fibrosis with alternating areas of normal lung, temporal heterogenous fibrosis characterised by scattered fibroblastic foci and collagen, and architectural disruption resultant of scarring (honeycomb pattern) (Eunhee and Cagle. 2020). This is lowly associated with an adaptive immune cell inflammatory infiltrate, which could be observed during acute exacerbations (Juarez *et al*. 2015, Butler and Keane. 2018).

Multi-disciplinary consensus is recommended to establish a diagnosis of IPF, informed by tests such as chest radiographs, high resolution computerised tomography (HRCT) scanning and pulmonary function tests (American Thoracic Society. 1999). IPF usually affects both side of the lungs, but lesion distribution can be very heterogeneous and patchy opacities are regularly observed in chest radiographs, commonly in the sub-pleural areas (Isah *et al*. 2016, Zhao *et al*. 2017). However, chest radiographs can appear normal in patients with IPF. Therefore, whilst IPF

may first be suspected following this test, it is typically used as a method for initial diagnostic approach ruling out other non-respiratory diseases such as left heart failure (Walsh *et al*. 2018).

HRCT is a more sensitive imaging technique than plain imaging of the chest and is central in the diagnosis of IPF (Walsh *et al*. 2018). HRCT scans typically display the presence of classical signs of UIP including sub-pleural and bi-basal predominance, patchy reticular opacities, and honeycombing (Figure 1.2) (Devaraj. 2014, Walsh *et al*. 2018).

If UIP is not observed from HRCT scans, then surgical lung biopsy was traditionally recommended with histopathology evaluation in combination with HRCT scans and clinical information to distinguish IPF from other IIPs (Cottin. 2016). However, surgical lung biopsy is associated with substantial mortality and morbidity risks (Hutchinson *et al.* 2015), and so is far less commonly used in diagnosis of IPF currently.

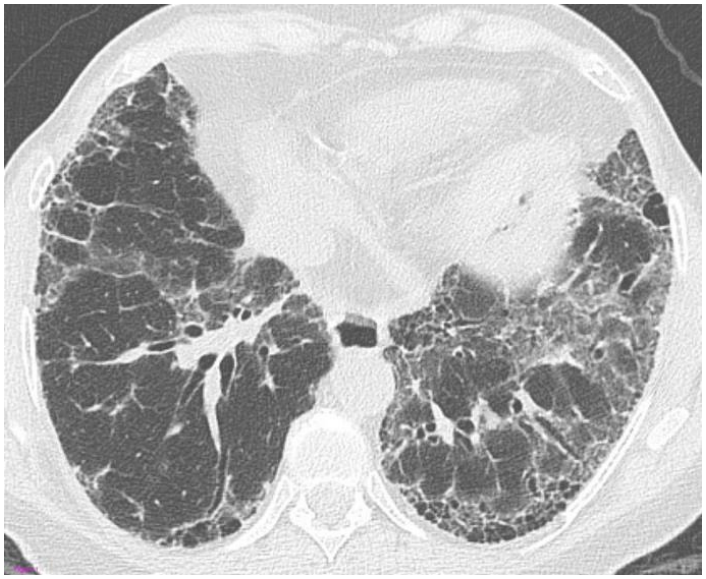


***Figure 1.2 HRCT scan of IPF**:* UIP pattern shown with bi-basilar, peripheral reticular abnormalities, tracheal changes and subpleural basal honeycombing (largely evident in the left lung). In an idiopathic context this indicates a diagnosis of IPF. Image taken from Sverzellati. 2013.

In IPF, steroids and immunomodulatory treatments (prednisolone, azathioprine and N-Acetylcysteine combination treatments) were shown to dramatically increase mortality over placebo (Bouros and Antoniou. 2005, Walter *et al*. 2006, Raghu *et al*. 2012). The only treatment strategy proven to have benefit in terms of survival and progression so far is with the anti-fibrotic small molecules, nintedanib and pirfenidone (Fala. 2015, Raghu *et al*. 2015, Williamson. 2018). Both of these FDA and NICE approved licensed treatments slow but do not stabilise or reverse progression of the disease (Spagnolo *et al*. 2015, Hirani *et al.* 2016).

Nintedanib is a triple tyrosine kinase inhibitor, taken in the form of capsules (Fala. 2015, Rangarajan *et al*. 2016). This has been shown to slow the rate of lung scarring. Nintedanib blocks fibroblast growth factors and indirectly the pro-fibrotic cytokine transforming growth factor (TGF)-β1, thus supressing fibrosis (Rangarajan *et al.* 2016, Case *et al*. 2017).

Pirfenidone is taken in the form capsules or tablets with meals (Williamson. 2018). The mechanism of action is unclear, but it exhibits both anti-inflammatory and anti-fibrotic properties. The end result of pirfenidone use is observed to suppress fibroblast proliferation and the inhibition of pro-fibrotic cytokines including TGF-β1, in addition to reduction in collagen (extracellular matrix) synthesis (Raghu *et al*. 2015). This treatment attenuates forced vital capacity decline in IPF patients (Raghu *et al*. 2015).

Without treatment, median survival of an IPF patient is ~3 years (Fabrellas *et al*. 2018). These current licensed treatments halve the rate of lung function decline and appear to increase survival length approximately two-fold (Margaritopoulous *et al*. 2018, Noor *et al*. 2021).

IPF is the most common form of IIP in populations of European descent, accounting for ~40% of ILD clinic populations (Meltzer and Noble. 2008, Wuyts *et al*. 2014). Despite the recent advances in treatment, overall, there remain major clinical burden and unmet clinical need, that are increasing with incidence (~4% yearly) (Williams and Wilson. 2008, Wolters *et al.* 2018, British Lung Foundation. 2020, British Thoracic Society ILD Registry Annual report. 2020).

Similar features to those observed in IPF may be seen in other ILDs, as a separate aspect of the disease to the inflammatory changes (Travis *et al*. 2013, Cottin *et al*. 2018). Various data support the concept that a common pathway of IPF-like disease, particularly 'progressive fibrotic' ILD (PF-ILD) may occur downstream of distinct aetiologic triggers (Cottin *et al*. 2019). Importantly such characteristics in non-IPF ILDs appear to respond to anti-fibrotic therapies similarly to IPF (Flaherty *et al*. 2019, Maher *et al*. 2019, Acharya *et al*. 2020), although at time of writing their evaluation for clinical use in these contexts within the UK National Health Service (NHS) has not been completed. Findings relating to general pro-fibrotic mechanisms in IPF may therefore have wider relevance to progressive fibrosis in other ILDs (Collins and Raghu. 2019).

1.2 IPF Demographics

IPF is responsible for ~1% of total death (~5% of respiratory related death) within the UK (Wolters *et al.* 2018, British Lung Foundation. 2020). Prevalence is more common in the elderly with typical age of onset ≥ ~65 years (Fell *et al*. 2010, Jo *et al*. 2016). IPF is therefore a global health concern, particularly where there are substantial elderly populations.

IPF progression appears to affect far more men than women (Ley and Collard. 2013, British Lung Foundation. 2020). A study by Han *et al*. 2008, monitored the desaturation levels in serial 6-

minute walk test over a 1-year period of suspected IPF patients. Their research displayed an average desaturation in the 6-minute walk test by 1.37% (female), compared to 2.83% (male). Overall, this suggests a potential survival advantage for females over males, but the mechanisms underlying these differences by sex are unclear. Contrary to this, other studies have not supported this. Commonly measured index of lung function (including the analysis of FVC and D_{lco}) appear to have no bearing between sex differences during the progression of IPF (Caro *et al*. 2016, Kaunisto *et al*. 2019).

IPF is generally viewed as a disease that is more common in patients of European ancestry (Meltzer and Noble. 2008). One study reported mortality rates being higher amongst non-Hispanic whites (Caucasian) (Swigris *et al*. 2012). Out of 251,000 deceased IPF subjects, >87% of the cases were Caucasian (Swigris *et al.* 2012). However, not all studies support this (Lederer *et al*. 2006). A retrospective cohort study by Lederer displayed increased hazard ratio of blacks and Hispanic whites compared to Caucasian subjects.

In the UK, the highest incidence rate of IPF is reported in the north of England (around Yorkshire). The reasons for this are not clear but may relate to demographics (including occupational and environmental risk factors and social groupings) within the different regions (British Lung Foundation. 2020).

1.3 IPF Risk Factors

Several risk factors have been proposed to facilitate the development of IPF although the underlying pro-fibrotic causal mechanisms for such factors remain under-defined. Risk factors include exposure to occupational and damaging environmental factors (Baumgartner *et al*. 2000),

smoking, and viral infections (Baumgartner *et al*. 1997, Thornton *et al*. 2009, Naik *et al*. 2011). Identification of probable risk factors for IPF is important since it flags the increased likelihood of developing the disease. However, further experimental studies based upon epidemiological studies are still required to prove causal links sufficiently to allow the development of patient-specific IPF prevention and population screening strategies. Overall, a small proportion of individuals exposed to one or more of these suggested risk factors develop disease, suggesting a requirement for multifactorial contributions to IPF pathogenesis (Kreuter *et al*. 2015).

1.3.1 Inhalational Risk Factors

1.3.1.1 Cigarette Smoking

Analogous to other respiratory diseases (The Health Consensus of Smoking-50 years Progress. 2014) smoking has consistently been associated with IPF (Bumgartner *et al*. 1997, Bellou *et al*. 2017). Studies have confirmed an increased risk of attaining IPF in younger age group of current smokers than never and ex- smokers (Karkkainnen *et al*. 2017). The probability of developing IPF increases with the pack-year exposure history to cigarettes (Oh *et al*. 2012). However, only an association between smoking and IPF is observed, there are no specific causal correlation(s) defined between IPF diagnosis and the history of cigarettes smoked (Baumgartner *et al*. 1997).

Mechanisms by which smoking could contribute to the pathogenesis of IPF have been proposed. Cigarette smoking contains particulate matters as well as chemicals including highly toxic reactive oxygen and nitrogen species (RONS) (Pryor *et al*. 1993, Zhao and Hopke. 2011). These can disequilibrate oxidant/antioxidant and protease/anti-protease levels in the lungs, resulting in exacerbating inflammation and consequently cellular dysfunction and tissue damage (Wellman and Bloomer. 2009, Liguori *et al*. 2018). *In vitro* studies demonstrate that smoking can cause

fibroblasts to produce excessive levels of matrix metalloproteinases (MMPs) (Kim *et al*. 2004, Ning *et al*. 2006). Under normal healthy conditions, MMPs tightly regulate extracellular matrix turnover; though excessive amounts can disrupt the structural and functional integrity of the alveolar-capillary basement membrane (Atkinson and Senion. 2002, Pardo *et al.* 2002). This activates pro-inflammatory and/or pro-fibrotic mediators including latent-TGF-β1 (Yu and Stamekovic. 2000, Oh *et al*. 2012). *In vivo* studies have reported cigarette smoking induces the promotion of an M2-type macrophage phenotype (Hodge *et al.* 2011). This has indirect consequence in overexpressing residential fibroblasts through the secretion of TGF-β1, thereby increasing collagen (extracellular matrix) secretion (Zhang *et al*. 2018).

1.3.1.2 Occupational and Environmental Risk Factors

Numerous occupational environments sampling a wide range of inhalational dust exposures have been associated with the development of IPF independent of age, sex, region and smoking status (Miyake *et al.* 2005). A positive relationship is observed between IPF patients and their exposure to metal dusts and this has been studied in some detail (Baumgartner *et al*. 2000, Miyake *et al*. 2005). The amplitude of risk seems to be proportional to the duration exposed to such dust (Baumgartner *et al.* 2000). *In vitro* and *in vivo* studies of inorganic dust such as the hard metal cobalt tungsten carbide suggest that the initiation of the pro-fibrotic mechanism may depend upon the formation of reactive oxygen species (Lison *et al.* 1996).

Agricultural employment has also been linked to an increased risk in developing IPF (Nordgren and Bailey. 2016). Workers in this sector are exposed to high levels of airborne dust particles from various sources including pesticides, grains, feed additives, fertilisers and biological aerosols from plant or animal matter (Rumchev *et al*. 2019). Compared to non-exposed controls, the risk of IPF was observed to increase significantly (Baumgarter *et al*. 2000, Awadalla *et al*. 2012).

Exposure to wood dust, adhesives and solvents have also been associated with increased risk of IPF when compared to relevant controls (Hubbard. 2001).

1.3.2 Infection and the Respiratory Microbiome

Previous definitions of acute exacerbations of (AE-)IPF specified exclusion of infectious causes but a more recent consensus position removes this distinction (Ryerson *et al.* 2015). An important reason for this change is due to previous assumptions that normal healthy lung was a sterile environment (Dickson *et al*. 2015), and that any relevant infection could be detected by standard culture techniques, have been superseded by lung microbiome studies. These have provided a far more comprehensive understanding of the lung microbiome in health, IPF and AE-IPF and demonstrate shifts in microbial ecology between each of these states and it seems likely that this has at least some relevance to pathogenesis (Molyneux *et al*. 2014, Salisbury *et al.* 2017, Invernizzi *et al*. 2018). Hospitalisations related to viral infection have been noted in AE-IPF patients, including the presence of parainfluenza, rhinovirus, syncytial virus and cytomegalovirus (Wooton *et al.* 2011, Ushiki *et al*. 2014).

Other viruses reported to be amplified in IPF patients in comparison to healthy controls include Torque-Teno Virus and Epstein-Barr virus (EBV) (Moore and Moore. 2015). High levels of EBV was found to replicate in type II alveolar epithelial cells of IPF patients (Marzouk *et al.* 2005) though these are also found in patients presenting other IIP respiratory diseases (Marzouk *et al*. 2005, Moore and Moore. 2015).

Furthermore, multiple reports suggest increased levels of Hepatitis C Virus (HCV) in IPF patients in comparison to healthy controls (Ueda *et al*. 1992, Arase *et al*. 2008, Moore and Moore. 2015).

However, HCV is generally more prevalent in a number of respiratory diseases and is not specific to IPF (Meliconi *et al*. 1996).

1.3.3 Gastroesophageal Reflux Disease and Chronic Microaspiration

Evidence suggests gastroesophageal reflux disease (GORD) is associated with IPF (Ley and Collard. 2013) with an estimated prevalence of 90% (Lee *et al*. 2010). GORD is a presumed risk factor for chronic microaspiration of both acidic and non-acidic gastric contents leading to repetitive lung injury and resulting in lung injury responses causing pulmonary fibrosis (Raghu and Meyer. 2012).

Conversely, basal lung fibrosis is likely to contribute to disruption of the physiological sphincter at the gastro-oesophageal junction through tractional effects (Ghisa *et al*. 2019)

In vitro studies on human cells subjected to treatment with the bile acid component chenodeoxycholic acid, reported increased production of the pro-fibrotic cytokine TGF-β1 in epithelial cells, combined with increased mesenchymal fibroblast cell proliferation (Perng *et al*. 2007). Furthermore, treatment with GORD medications or anti-acid treatments, typically displayed longer survival time in patients with IPF (Lee *et al*. 2011, 2013) further supporting that GORD and chronic microaspiration may play crucial role in the development of IPF. Contrastingly, a *post hoc* analysis of IPF patients (INPULSIS® trials) treated with anti-acid in combination with nintedanib and placebo at baseline, had worse outcomes than those not receiving the treatments (Raghu *et al*. 2015). Moreover, a second *post hoc* analysis of IPF patients treated with anti-acid in combination with pirfenidone and placebo (CAPACITY studies 004/006 and ASCEND 016 trials) showed no benefit. Indeed, there was a signal for increased pulmonary infection rates in patients

with advanced stages of IPF on such treatment (Kreuter *et al.* 2016). It is difficult to untangle whether the treatment or the underlying severity of the GORD is the actual culprit here.

Therefore, prospective placebo controlled interventional randomised control studies are needed, and are currently underway (TIPAL, ISRCTN13526307) to determine whether preventing acid reflux impacts upon IPF progression. Overall, it is highly accepted that GORD and survival time of IPF is a real risk factor.

1.3.4 Genetics

Genetic variants have been identified as risk factors for the development of IPF and is supported in an increasing number. Familial IPF provides an obvious starting point. It is defined as at least two or more first-degree relatives of primary biological family members having IPF (also regarded as familial pulmonary fibrosis) (Ley and Collard. 2015, Sousa *et al*. 2019). The mean age of onset is often observed at an earlier age (~55 years) than sporadic IPF (~65 years of age) (Marshall *et al*. 2000).

1.3.4.1 Genetic Predisposition Interlinked with Cellular Senescence

IPF is an age-related disease (Cho and Stout-Delgado. 2020), therefore increased focus has been attributed on the impact of the pathogenic role of cellular senescence (Schafer *et al.* 2017, Cho and Stout-Delgado. 2020, Katzen and Beers. 2020). Senescence is a complex cellular programme, but can broadly be subdivided into replicative and cellular senescence (Campisi, J. 1997, Kumari and Jat. 2021). Replicative senescence results from intrinsic cellular events, including telomere shortening, whilst cellular senescence can be driven by several factors including DNA damage

(Campisi, J. 1997, Kumari and Jat. 2021). Assessing telomere length as a risk factor has shown that patients with IPF are more likely to have shortened telomeres than age-matched controls and these shortened telomere length present worse outcomes in IPF subjects (Alder *et al.* 2008, Newton *et al.* 2016). Short telomeres in alveolar epithelial cells are found both in patients with telomere gene mutations and in sporadic IPF with telomere shortening (Alder *et al.* 2008). Increased senescence markers (p16, p21 and p53) are regularly identified in the lung epithelia around fibroblastic foci in IPF biopsies (Schafer *et al.* 2017, Cho and Stout-Delgado. 2020, Katzen and Beers. 2020). The mechanisms of senescent alveolar epithelial cells contribution in fibrotic remodelling remains unclear, however, speculations indicate a potential association in the transition to a senescence-associated secretory phenotype (SASP) (Katzen and Beers. 2020). SASP is associated with production of chemokines and cytokines that are implicated in crosstalk with immune cells as well as (co)factors that are associated with mesenchymal crosstalk and tissue remodelling (Schafer *et al.* 2017). Another possible impact of senescent alveolar epithelial cells includes the loss of epithelial repair and progenitor capacity, and thereby maximising fibrotic cell retention (Katzen and Beers. 2020).

Genetic mutation burden are frequently observed in multiple genes of the telomerase complex. These are most common in the *telomerase reverse transcriptase (TERT)* component of the complex, accounting for ~15% of affected families (Kannengiesser *et al*. 2015). A single mutation (c.2701C > T) located in exon 11 of the *TERT* gene, replacing arginine with tryptophan (Arg901Trp) was identified (Sousa *et al*. 2019). Mutation of this gene is related to premature shortening of telomeres in the peripheral blood and lungs, resulting in short telomere syndrome (Mangaonkar and Patnaik. 2018). It is believed that loss of function of telomerase complex may influence the turnover and disrupt the healing rate of alveolar epithelial cells (pushing them to senescent state) post-exposure to lung injury stimulus, thereby triggering IPF (Sousa *et al*. 2019).

A recent Mendelian randomisation study looked at a panel of seven genetic variants that interact to cause predictable telomere shortening (Duckworth *et al.* 2020). This indicated that a

genetically determined increase in telomere shortening rate was significantly associated with IPF risk and development. One standard deviation telomere shortening gave a > 4-fold increased risk of IPF. Interestingly this association was not apparent for chronic obstructive pulmonary disease (COPD) that is also associated with accelerated senescence.

1.3.4.2 Other Genetic Risk Factors

Polymorphism of the mucin protein gene *Mucin-5B (MUC5B),* is the strongest single genetic risk factor for both sporadic and familial IPF (Seibold *et al*. 2011). *MUC5B* gain-of-function promoter is a common replicated genetic risk factor prognostic in IPF with a 6.8 odds ratio for disease subjects (Yang *et al*. 2015). The association of this *MUC5B* polymorphism is predominantly described in Caucasian populations (Borie *et al*. 2013, Kropski *et al*. 2015, Helling *et al.* 2017). It is observed in 20% of diagnosed IPF patients. Specifically, a single nucleotide polymorphism (SNP) located ~3 kb upstream of the *MUC5B* transcription start site, causing G→T transversion, disrupting the E2 transcription factor (E2F) genes transcriptional binding site by generating two new transcriptional binding sites (HOX9 and PAX2) (Yang *et al*. 2015). This area of the promoter has been reported to bind to NF-κB in airway epithelial cells, which could result in the increase in pro-inflammatory cytokines (Kropski *et al*. 2015, Yang *et al*. 2015). MUC5B in IPF lung is localised in the bronchiolar epithelium (Nakano *et al*. 2016, Helling *et al.* 2017). Currently there are no direct evidence from cellular or animal models to define a mechanistic cause of action, but speculations have been made by David Schwartz research group (Yang *et al.* 2015): MUC5B overexpression results in chronic mucus hypersecretion and accumulation in small airways. This could impair mucocilliary transport, thereby resulting in mucus adhesion in the bronchoalveolar region, consequently inducing and potentiating chronic inflammation and injury responses (Yang *et al*. 2015). An additional speculation to explain this association is that the excessive mucus production in small airways could result in blockage of air supply to the acinar units, thereby preventing their expansion in response to inspiratory pressure changes (Gooptu, B. 2021). The

interface between blocked and expandable acini will be subjected to increased mechanical shear forces, thereby activating pro-fibrotic pathways (Gooptu, B. 2021). Additionally, mucins require extensive chaperoning and refolding to ensure the large number of cysteines they contain form the correct disulphide bonds in the endoplasmic reticulum (ER) prior to secretion (Park *et al*. 2009, Gooptu, B. 2021). The increased mucin levels could overwhelm the ER resulting in ER stress on the redox systems (Bradley *et al.* 2021). Marker of ER stress correlate with the degree of fibrosis irrespective of ILD subtype (Burman *et al*. 2018).

A SNP in the pro-fibrotic cytokine TGF-β-1 is associated with IPF risk, specifically in 869T/C at codon 10 and 915G/C at codon 25. This occurs within the hydrophobic α-helix of the signal sequence, replacing the leucine with proline residue and thus disrupting the α-helical structure of this region. The consequential effect is the upregulation of active TGF-β1 secretion, subsequently increasing the TGF-β1 pro-fibrotic downstream signalling pathway (Son *et al*. 2013, Liu *et al*. 2017).

Genome Wide Association Study (GWAS) by Allen *et al*. 2017 and 2020, support more independent genetic risk factors for IPF. One such genetic risk factor of recent interest is the SNP rs62025270 near *(A-Kinase Anchoring Protein 13) AKAP13* gene. Elevated mRNA expression levels of *AKAP13* (1.42-fold increase) are observed in lung tissue from IPF patients. The minor allele A was detected in these individuals. This is also associated with increased AKAP13 expression and increased susceptibility to IPF. AKAP13 is a Rho nucleotide exchange factor regulating RhoA activation which is known to be involved in pro-fibrotic signalling (Smith *et al*. 2003). This modulates fibroblast turnover and collagen synthesis (Watts *et al*. 2006), consequently resulting in increased extracellular matrix secretion and deposition.

1.4 Aberrant Wound Healing and Clinical Observations in IPF

During a stereotypical wound healing response, tissue injury stimulates acute inflammation stimulating migration of mesenchymal myofibroblast cells to the wound site, where they proliferate in response to inflammatory cytokines (Xue and Jackson. 2015). The myofibroblasts contract and secrete extracellular matrix proteins *in situ* to close and seal the wound (Darby *et al*. 2014). Subsequently myofibroblasts apoptose, whilst epithelial cells proliferate and reconstitute the basement membrane (Madden and Peacock. 1971, Roh and Lyle. 2006, Strieter. 2008, Felice *et al*. 2009, Xue and Jackson. 2015, Kasper and Barth. 2017). Rapid healing of a limited wound may not result in any scar formation (Xue and Jackson. 2015). However, majority of cases in IPF, the extracellular matrix deposition results in the formation of a scar with collagen aligned predominantly along a single axis, and cells excluded (Keane *et al*. 2018).

Clinically-significant pulmonary fibrosis may arise from prolonged or severe lung injury and injury treatments. Examples include the following: patients subjected to mechanical ventilation treatment for acute respiratory distress syndrome (ARDS) (Benitez et al. 2014, Mooney *et al*. 2017); in addition a sub-set of ARDS patients (not subjected to mechanical ventilation treatment) are seen to develop fibroproliferative phenotype (Benitez *et al*. 2014, Burnham *et al.* 2014). Other prolonged lung injury-derived fibrosis includes active connective tissue disease (Cottin. 2013, Walsh *et al*. 2014) or fibrotic hypersensitivity pneumonitis (Sforza and Marinou. 2017). In IPF, aberrant wound repair mechanisms are activated in such a way that fibrosis can progress in the absence of a major inflammatory insult (Selman and Pardo. 2002, Xue and Jackson. 2015). However acute progression is also observed in the context of inflammatory flares known as acute exacerbations of (AE-)IPF (Juarez *et al*. 2015), indicating that acute injury can act as a further driver of downstream fibrosis.

IPF is characterised by loss of type II alveolar epithelial cells (AECs) and their basement membrane (Ding *et al.* 2011, Camelo *et al.* 2013) but also degeneration of capillary endothelium (Ding *et al*.

2011). Loss of basement membrane integrity impairs alveolar re-epithelialisation and may therefore prolong wound healing (Strieter. 2008). Additionally, in response to this, both hyperplastic and proliferating type II alveolar epithelial cells function to repair disrupted basement membrane (Kasper and Barth. 2017). This repair process produces mediators (including TGF-β1), causing exaggerated proliferation, activation and migration of underlying fibroblasts into provisional matrix (Sakai *et al*. 2013). Fibroblastic foci are a pathological hallmark of IPF (Jones *et al*. 2016). They are clusters of apoptotic resistant myofibroblasts around alveolar airspaces, specifically sub-epithelial regions in close proximity to injured or hyperplastic type II alveolar epithelial cells (Eickelberg and Laurent. 2010). The effect of this clustering are enhanced through deposition of extracellular matrix components (Sakai *et al*. 2013). The loss of alveolar units impairs overall gas exchange within the lung (Parimon *et al*. 2020), limiting maximal oxygen uptake and reducing exercise tolerance (Zoz *et al*. 2012). The increased stiffness (reduced compliance) of the scarred lung tissue increases the effort required to take each breath and reduces lung volumes (Plantier *et al*. 2018, Kim *et al*. 2020), including forced vital capacity (FVC), a commonly measured index of lung function.

An early small clinical trial conducted by Raghu *et al*. 1991, suggested prednisolone and azathioprine immunosuppression treatments were beneficial to IPF patients, providing support for the hypothesis that low grade adaptive immune-mediated inflammation played a key role in pathogenesis (Bringardner *et al*. 2008). Microarray dataset analyses comparing IPF with healthy lung tissue revealed an upregulation in pro-inflammatory cytokines, chemokines, antioxidants, and amyloid(s) (all of which are associated with chronic inflammatory disorders) (Kaminski *et al*. 2006, Murray *et al*. 2010). However, the PANTHER-IPF (NCT00650091) study (Raghu *et al*. 2012) has changed the view of adaptive immune response mediation in IPF. IPF patients treated with prednisolone and azathioprine responded negatively in comparison to placebo controls. Far higher mortality was apparent in the treated individuals. Interestingly recent bio-sample analysis showed that excessive mortality was seen in those who had short telomeres but even so there was no evidence of any treatment benefit in the others (Molina. 2019, Newton *et al*. 2019).

Shortly after the PANTHER-IPF study, the TIPAC trial (17464641) assessing the antibiotic co-trimoxazole as an IPF treatment (Shulgina *et al.* 2013) was reported. The primary endpoint of a pre-defined improvement in lung function, based upon encouraging preliminary data was not met. However, amongst secondary endpoints co-trimoxazole treatment appeared to dramatically reduce mortality in IPF patients. Recently the TIPAC-2 (EME-TIPAC) clinical trial, assessed the use of co-trimoxazole in addition to current IPF treatments and displayed no such extra benefit of reduced mortality (Wilson *et al*. 2020).

1.5 Experimental Studies Defining IPF Pathogenesis

Figure 1.5 provides a framework suggesting how observed risk factors could relate to the pathogenesis of IPF. Inhaled stimuli, increased small airway mucus production, shifts in the microbiome, ageing-related processes and heightened injury responses may all contribute. Each of these factors can be tested or complemented by hypothesis-driven experimental studies.

1.5.1 Histopathology on Human IPF Tissue (Ex Vivo)

Histological characterisation of IPF is usually defined by UIP, though this is not completely unique for IPF (Wuyts *et al*. 2014). Histologically various characteristics become apparent following fibrosis, including significant distortion of the lung architecture and scattering of fibroblastic foci containing fibroblasts and myofibroblasts (Jones *et al*. 2016). Additionally, honeycomb patterns are observable (resultant from excessive collagen, causing fibrosis) (Figure 1.3) (Cottin. 2016).

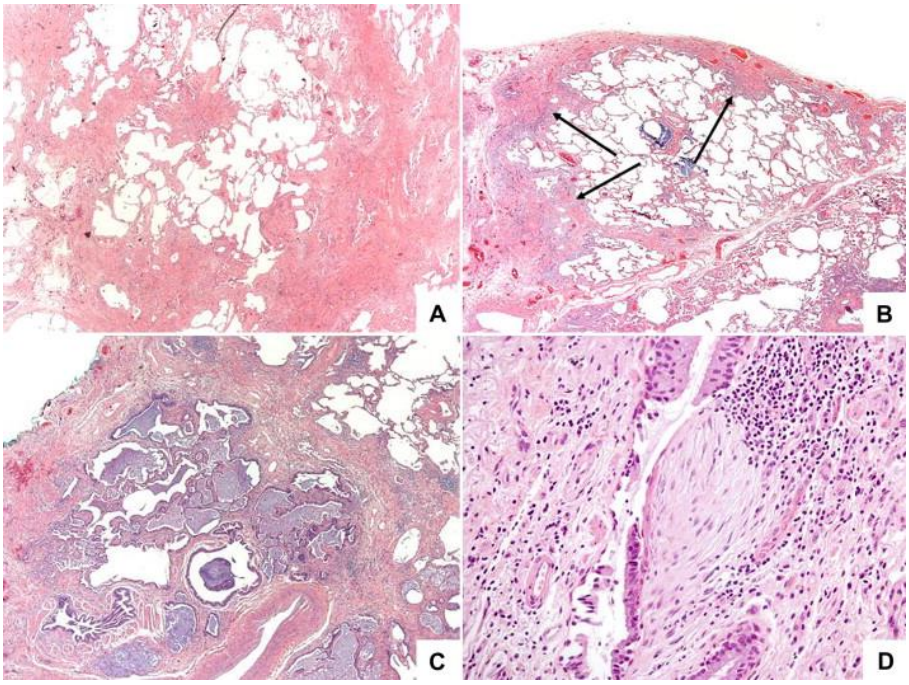


Figure 1.3 Histological Identification of IPF. **Histopathological features of IPF (UIP):** A) Patchy abrupt scarred lung versus normal lung observed at low magnification (diagnostic key feature). B) Fibrosis at periphery of lobule in subpleural-paraseptal regions (shown by the black arrows). C) Honeycombing consists of enlarged airspaces (architectural distortion and the abrupt transition observed compared with residual normal lung, clear in the upper right hand corner. D) Focal area of collagen deposition and fibroblastic foci's (dome shaped) consisting of myofibroblasts cells. The fibroblast foci are covered by bronchiolar epithelium in this image. Image taken from Cavezza *et al*. 2010.

Additional pro-fibrotic markers are probed histologically to assess IPF lung tissue, which have provided further insight into lung tissue remodelling behaviour. E-cadherin levels have been assessed in IPF vs healthy lung tissue supporting EMT as a putative driver in IPF pathogenesis and consequently lung remodelling (Lomas *et al*. 2012). In IPF lung, reduced levels of E-cadherin are evident in the areas denoting basement membranes and the bronchiolar epithelium distribution (Lomas *et al*. 2012, Kulkarni *et al*. 2016). Additionally, elevated levels of α-smooth muscle actin (SMA) are observed in IPF fibroblastic foci, consistent with myofibroblast phenotype (Lomas *et*

al. 2012, Sun *et al*. 2016). Furthermore, increased levels of the pro-fibrotic cytokine TGF-β1 are also observed in IPF tissue versus healthy controls (Lomas *et al*. 2012).

Collagen in a healthy lung has dynamic equilibrium; synthesis and degradation is tightly regulated (Thickett *et al*. 2001). During the fibrotic lung disease state, synthesis outweighs degradation (Hansen *et al*. 2016). This excessive collagen attachment thickens the thin alveolar interstitium leading to the destruction and remodelling of the lung architecture (Hansen *et al*. 2016). The excessive collagen deposited in fibrotic foci and the presence of fibrotic foci are important in the histopathological diagnosis of IPF (Enomoto *et al*. 2013). Collagen replaces functional tissue, proportionately reducing the full capability of the tissue, and its physical characteristics may themselves lead to further dysfunction (Singer and Clarke. 1999, Midwood *et al*. 2004).

1.5.2 Explant Model of IPF

Colleagues from the department of Respiratory Sciences, University of Leicester (Bradding group) have developed and validated a lung tissue explant model to further elucidate IPF pathology (Roach *et al*. 2018). Healthy lung tissue resections were cultured *ex vivo* and stimulated with the pro-fibrotic cytokine TGF-β1. Microarray analysis displayed an upregulation of multiple genes considered to be involved in pro-fibrotic processing. Elevated gene levels were notably observed for extracellular matrix- collagen type I and III accompanied by extracellular matrix remodelling enzymes. Furthermore, subunits of αvβ6 levels were elevated, which are important transmembrane proteins modulating TGF-β1 signalling pathways in animal model of fibrosis (Goodwin *et al*. 2009). Decreased levels of anti-fibrotic genes were observed including interleukin-13 receptor alpha 2.

Immunohistochemistry assessment of tissue protein levels confirmed increased levels of collagen type I and III, α-SMA and fibroblast surface proteins (FSP). Sircol assay further confirmed increased collagen levels secreted into the supernatant of the culture media. Inhibition of a putative target for IPF treatment, calcium activated potassium channel 3.1 ($K_{Ca}3.1$) attenuated upregulation of pro-fibrotic gene responses to TGF-β1 stimulation. Furthermore, immunohistochemistry assessments of these $K_{Ca}3.1$ inhibited tissue sections displayed reductions in levels of collagen type I and III, α-SMA and FSP.

1.5.3 Animal models of IPF (In Vivo and Ex Vivo)

The best-characterised animal model currently for experimental IPF is the bleomycin-treated mouse model (Moeller *et al.* 2008, Jenkins *et al*. 2017). Levels of hydroxyproline are assessed as a biomarker of collagen deposition along with histological staining of tissue resections. The bleomycin-treated mouse model however has important differences from IPF. It involves a florid immune cell-mediated inflammatory phase, and reverses after the insult is stopped (Moeller *et al.* 2008). This alleviates it from one of the critical hallmarks of IPF, which is that it is an irreversible process. However, this is not a universal finding, and fibrotic reversion could take several months (Scotton *et al.* 2013).

The anti-neoplastic drug bleomycin can be administered through several routes including intratracheal, intraperitoneal, subcutaneous, intravenous, inhalational and through osmotic pump (Moeller *et al*. 2008, Moore *et al.* 2013, Peng *et al*. 2013). Bleomycin induces lung injury in response to acute inflammatory responses (Moeller *et al*. 2008, Peng *et al*. 2013), observed by the increase levels of inflammatory cell infiltrates in the bronchoalveolar lavage fluid (BALF) of mice (Peng *et al*. 2013). Hydroxyproline levels are increased in the BALF, indirectly implying increased levels of collagen (Peng *et al*. 2013). *Ex vivo* histological studies of bleomycin treated

lung sections display typical fibrotic lesions (Moeller *et al*. 2008). The lung tissue in early stages of bleomycin administration is characterised by inflammatory cell infiltrate in the interstitium and alveolar spaces, with relatively few myofibroblasts (Peng *et al*. 2013). However, post-5 week bleomycin treatment, heterogenous characteristics become apparent with myofibroblasts interspersed with small partially collapsed alveoli and disrupted bronchoalveolar epithelial cells (Peng *et al*. 2013). Bleomycin treatment increases the gene expression of pro-fibrotic factors, including TGF-β1 and its corresponding receptors (Breen *et al*. 1992, Higashiyama *et al*. 2007, Khalil and O'Connor. 2008). Unsurprisingly, high gene expression levels were also observed for SMAD 1-3 which are the canonical signal transducer of downstream TGF-β1 mediated signalling (Higashiyama *et al*. 2007). TGF-β1 overexpression model has been discussed in more detail below (1.6.3).

Recently, bleomycin-treated sheep model has been developed for preclinical testing of novel anti-fibrotic drugs (Dewage *et al*. 2019). Pirfenidone treatment attenuated bleomycin-induced fibrosis in sheep, therefore this treatment can be used as a standard positive control to assess novel future therapeutics in this animal model.

1.5.4 In Vitro Cell Models of IPF

Aberrant epithelial and mesenchymal responses and cell-cell interactions contribute to IPF pathology *in vivo*. Human *in vitro* models are used to recapitulate these. Although the pathogenic mechanisms of IPF remain to be fully elucidated, evidence that the mesenchymal myofibroblasts may be derived partially from alveolar epithelial cells by the process of EMT is consistently observed in histopathological, animal and explant models.

1.5.4.1 Epithelial Cells

Human lung cell lines including A549 cells have been used to model IPF. The A549 lung epithelial cell line is derived from human broncho-alveolar cancer cells. A cancer cell line, particularly from non-small cell lung cancer, contains high mutation burden. Its behaviour may therefore differ from that of normal alveolar epithelial cells in a variety of ways (Jenkins *et al.* 2012). Nevertheless, A549 cells are experimentally tractable and are widely used to model type II alveolar epithelial cell behaviour. In the correct cell culture conditions they generate intracellular lamellar bodies containing surfactant proteins B and C, a hallmark of type II alveolar epithelial cell phenotype (Cooper *et al*. 2016, Selenium 2019).

TGF-β1 stimulation in these cells displays EMT characteristics since increased levels of the mesenchymal marker vimentin, and reduction in the epithelial marker E-cadherin is observable (Kasai *et al.* 2005, Kim *et al*. 2007). In addition, elevated levels of α-SMA (typical myofibroblast marker) (Kasai *et al*. 2005), and collagen type I and III (Kasai *et al*. 2005) are observed. It is not clear how far EMT proceeds *in vivo* in IPF, and the extent to which this TGF-β response contributes to the observed pool of myofibroblasts that form the fibroblastic foci (O'Connor *et al*. 2013, Bartis *et al*. 2014).

To determine whether lung alveolar epithelial cell injury induces myofibroblast expansion through EMT, A549 cells have been cultured with BALF (obtained from bleomycin treated mice). Expression levels of the epithelial marker E-cadherin is lost, whilst α-SMA, collagen and vimentin increased (Chen *et al*. 2016, Wang *et al*. 2020) indicating mesenchymal phenotype transition. Furthermore, the increase in α-SMA, collagen and vimentin positively, and the decrease in E-cadherin, negatively corresponds to the expression of β-catenin (Chen *et al*. 2016). β-catenin silenced A549 cells have been cultured with bleomycin induced BALF. The expression levels of E-cadherin, α-SMA, vimentin and collagen type I remain unchanged (Chen *et al*. 2016). This

suggests that β-catenin and its downstream signalling is of importance in the end-stage output of the mesenchymal pro-fibrotic signal transduction.

Additional epithelial cell models have been used to investigate IPF pathogenesis, including primary lung epithelial cells. The effects of TGF-β1 and the data observed in the A549 cell model is recapitulated in the primary lung epithelial cell models with similar response outcomes (MacKinnon *et al*. 2012, Gabasa *et al.* 2017).

1.5.4.2 IPF Myofibroblast Cells

Myofibroblasts are differentiated forms of fibroblasts, which as the name suggests are muscle related and display high contractility features with prominent stress fibres and focal adhesions (Desai *et al.* 2014). *Ex vivo* derived myofibroblast primary cells isolated from lung tissue are used to model IPF *in vitro* (Sundarakrishnan *et al*. 2018). Myofibroblast cell isolation and cell culturing is well established (Roach *et al*. 2013). Whilst primary cells are better at replicating *in vivo* lung phenotype, they are limited by donor-based variabilities (Mahmoudi *et al*. 2018) and the cells reach a state of senescence (Hayflick limit) in early passages (Orgunc and Fagagna. 2011).

Myofibroblasts may arise from multiple sources. TGF-β1 induced canonical (SMAD) and non-canonical signalling is critical in lung fibroblast activation, proliferation, and differentiation into matrix secreting myofibroblasts (Desmouliere *et al*. 1993, Sousa *et al*. 2007, Baum *et al.* 2011, Zhou *et al*. 2011). Lung epithelial cells could function as a source of myofibroblasts via the process of EMT in response to TGF-β1 stimulation (Hinz *et al*. 2007, Doerner and Zuraw. 2009, Gabasa *et al*. 2017). Circulating fibrocytes also infiltrate at sites where fibrosis is observed and differentiate into myofibroblasts (Abu El-Asrar *et al*. 2008) (Figure 1.4).

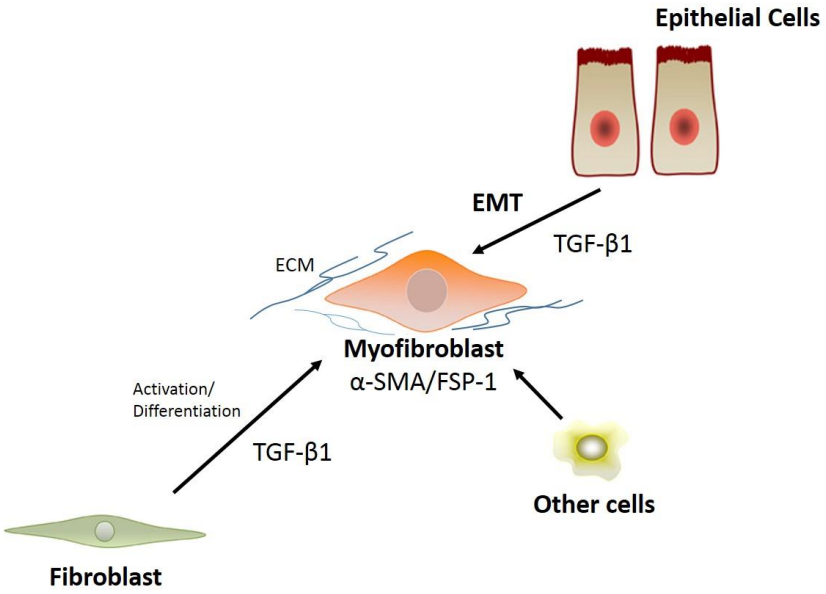


Figure 1.4, A diverse range of cell types may contribute to myofibroblast numbers and behaviours in fibrotic diseases. A variety of precursor cell types have the potential to differentiate into myofibroblasts or adopt similar mesenchymal features. Tissue-resident fibroblasts are activated and differentiated by TGF-β1 stimulation. *In vitro* epithelial cells demonstrate features of epithelial to mesenchymal transition (EMT) following TGF-β1 stimulation. Whilst EMT is the most-studied of these processes, a range of other cell types can show similar behaviour. These include endothelial cells undergoing endothelial-mesenchymal transition (EndMT) upon TGF-β1 stimulation, as well as fibrocytes and pericytes, perivascular adventitial cells and smooth muscle cells differentiating into myofibroblasts. (Desmouliere *et al*. 1993, Hinz *et al*. 2007, Sousa *et al*. 2007, Abu El-Asrar *et al*. 2008, Doerner and Zuraw. 2009, Baum *et al*. 2011, Zhou *et al*. 2011, Gabasa *et al*. 2017).

Myofibroblasts go through permanent reprogrammed physiological and pathological phenotypic changes, characterised by uncontrolled proliferation and survival which promotes fibrogenesis (Moore and Herzog. 2013, Wolters *et al.* 2014). In basal conditions the levels of the mesenchymal marker (α-SMA and stress fibres) and pro-fibrotic mediators (including collagen type I, TGF-β1, K_{Ca}3.1, total SMAD 2/3) are higher in IPF-derived myofibroblasts than that of healthy control (Ramos *et al*. 2001, Bocchino *et al*. 2010, Roach *et al*. 2014). Furthermore, a sub-analysis of α-

SMA probed by immunofluorescence staining displayed increased stress fibre characteristics in IPF derived myofibroblasts (Roach *et al*. 2014).

Alongside the increased levels of collagen, other extracellular matrix components are increased in IPF myofibroblasts, including fibronectin and laminin which are present in the active form of the disease (Faouzi *et al.* 1999, Klingberg *et al.* 2013, LeBleu *et al*. 2014). Furthermore, myofibroblasts can also mediate innate inflammatory responses, accelerating the progression of fibrosis (Linthout *et al.* 2014).

Pro-fibrotic stimulated IPF cells display increased invasive phenotype and resistance to apoptosis (Li *et al*. 2011, Wolters *et al*. 2014). These stimulated IPF myofibroblasts also display increased levels in mesenchymal and pro-fibrotic mediators, e.g. include α-SMA, Collagen type I, SMAD 2/3 (Luo *et al*. 2015).

Donor derived non-fibrotic control and IPF myofibroblasts could include both fast and slow progressors for example with differing proliferation rates. Commonly, IPF-derived myofibroblasts show advanced cellular proliferation in comparison to healthy non-fibrotic controls, however, at late passages this proliferation rate is severely reduced, marked by the reductions in typical myofibroblast markers (Oda *et al*. 1990, Roach *et al*. 2013, Waters *et al*. 2018).

1.5.5 Inferences Taken from Model Studies of IPF

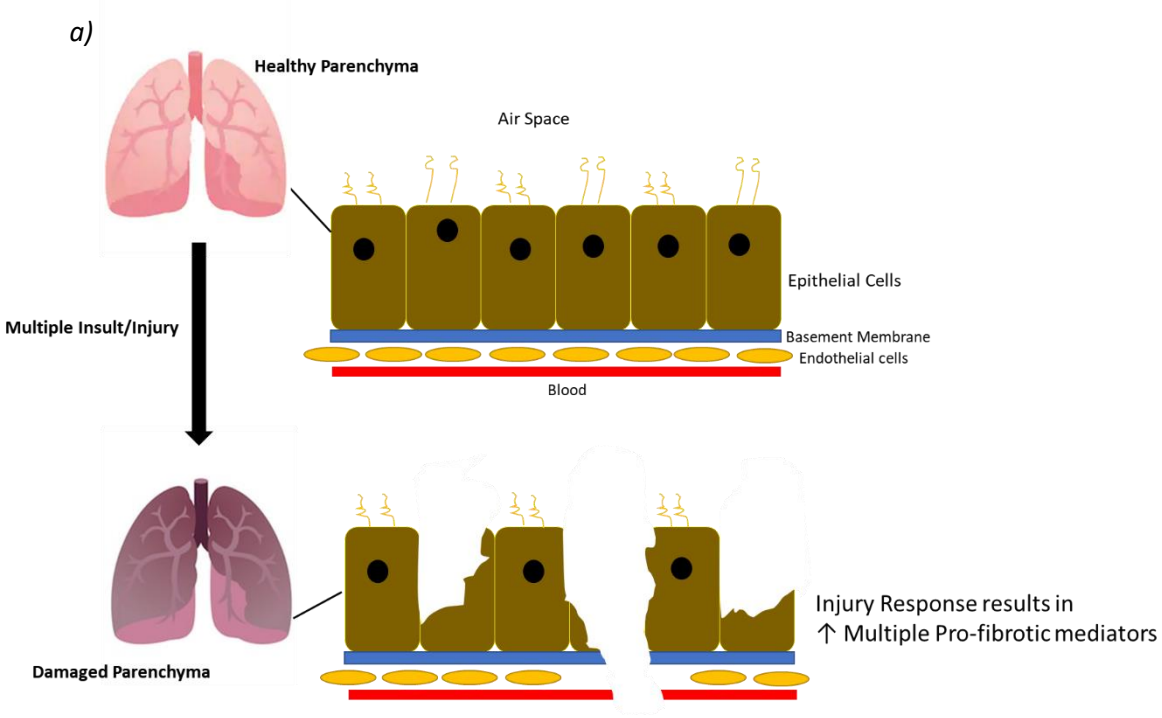
Ex vivo histopathology studies of IPF tissue demonstrate fibroblastic foci within and underneath the regions of the damaged epithelium (schematic, Figure 1.5). IPF is initiated due to the resultant aberrant epithelium and mesenchyme interactions, with little role of inflammation (Selman and

Pardo. 2002, Horowitz *et al*. 2006). Mesenchymal myofibroblasts promote excessive collagen deposition thus favouring fibrogenesis.

Other groups suggest both inflammatory and fibrotic pathogenesis mechanisms work in concert (Horowitz *et al.* 2006). Acute inflammation may precede chronic fibrotic processes, but in the lung these are not sufficient for fibrogenesis (Sheppard. 2001, Gauldie *et al*. 2002). Kolb *et al*. 2001, demonstrated the overexpression of interleukin-1β (IL-1β) in rat lung results in acute inflammatory responses and pulmonary fibrosis (Kolb *et al*. 2001). IL-1β mediated inflammation lead to defects in cross-talk between mesenchymal and epithelial cells causing failure of normal re-epithelialisation/endothelialisation. Subsequent fibrosis is then driven by TGF-β1.

Pro-fibrotic behaviour involves increased activity of pro-fibrotic cytokines, such as TGF-β1, tumour necrosis factor-α (TNF-α) and IL-13 (Fernandez *et al*. 2012, Harris *et al*. 2013). Furthermore, in fibrotic conditions it appears there are deficiencies in anti-fibrotic mediators, such as interferon gamma (IFN-γ) and prostaglandin E2 (PGE2). Overall, in IPF, pro-fibrotic factors outweigh anti-fibrotic factors (Dong *et al.* 2015).

From both clinical and experimental observations, a schema representing pro-fibrotic pathogenic pathway of IPF on the cellular level is displayed in Figure 1.5:


a)
Healthy Parenchyma
Air Space
Epithelial Cells
Basement Membrane
Endothelial cells
Blood
Multiple Insult/Injury
Damaged Parenchyma
Injury Response results in
↑ Multiple Pro-fibrotic mediators

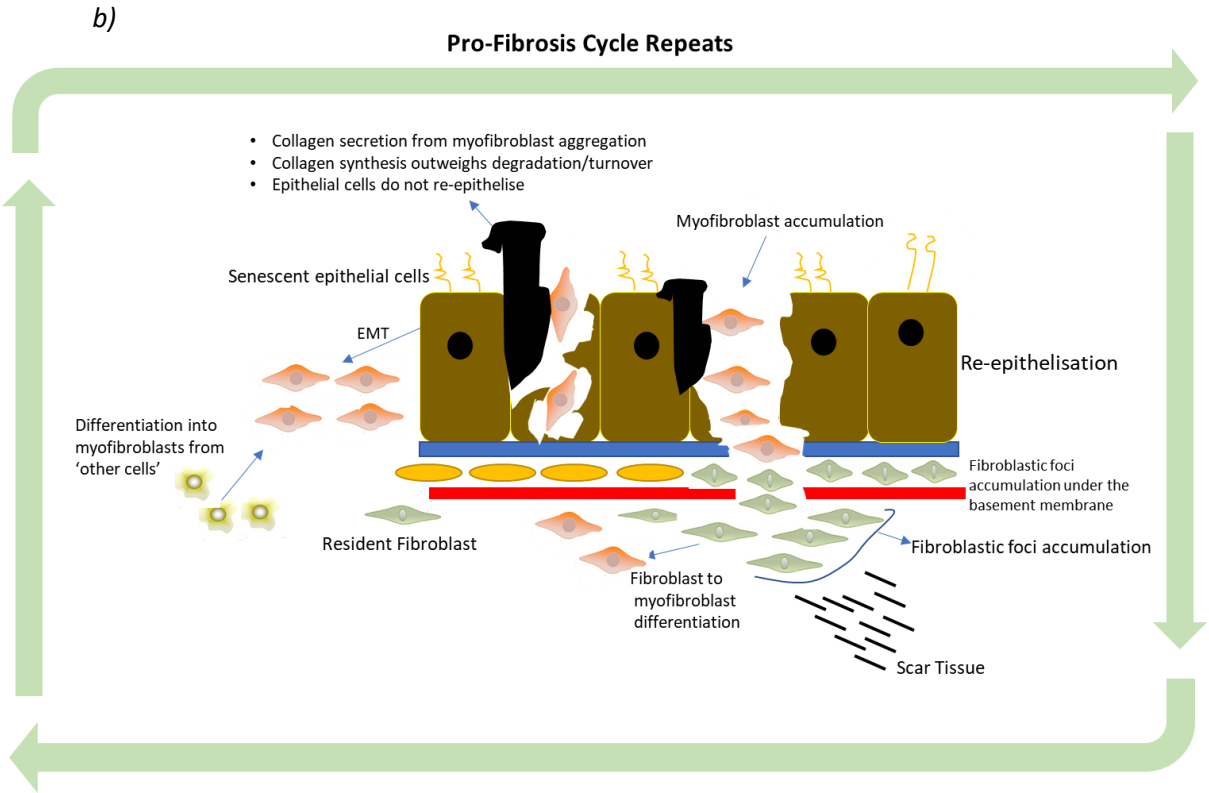

b)
Pro-Fibrosis Cycle Repeats
• Collagen secretion from myofibroblast aggregation
• Collagen synthesis outweighs degradation/turnover
• Epithelial cells do not re-epithelise
Myofibroblast accumulation
Senescent epithelial cells
EMT
Re-epithelisation
Differentiation into
myofibroblasts from
'other cells'
Fibroblastic foci
accumulation under the
basement membrane
Resident Fibroblast
Fibroblastic foci accumulation
Fibroblast to
myofibroblast
differentiation
Scar Tissue

Figure 1.5, Cellular level of IPF pathogenesis model. 1.5a, **In healthy lung/parenchyma** the alveolar epithelial cells, basement membrane and the endothelial cells are intact. The endothelial cells keep the alveolar epithelial cells separated from blood. **Multiple injury** of the lung at multiple sites can arise from risk factors proposed in section 1.3 (including cigarette smoking, viral infection, environmental factors). This causes damage and injury to the endothelial cells, basement membrane, and epithelial cells. In response to injury, the cells are triggered to release multiple pro-fibrotic mediators and growth factors, including the release of TGF-β1 (these can often leak into the airspace).

b, Response to injury leads to the accumulation and infiltration of myofibroblasts and fibroblastic foci within the area of injury from resident fibroblast precursor cells, circulating fibrocytes, epithelial-mesenchymal-transition or differentiation from other cells (myofibroblast differentiation displayed in Figure 1.3). Scar tissue is formed. The myofibroblast secreted collagen is resistant to dynamic collagen turnover and apoptosis is hindered. The collagen then accommodates and incorporates into damaged region of the epithelial cells and form a mesh. The epithelial cells do not re-epithelise, the integrity of basement membrane is irreversible damaged, and myofibroblast apoptosis is hindered (synthesis outweighs turnover), resulting in continuation of pro-fibrotic cycling. (Kolb *et al*. 2001, Selman and Pardo. 2002, Horowitz *et al*. 2006, MacKinnon *et al*. 2012, Jones *et al*. 2016, Roach *et al*. 2018).

1.6 The Pro-Fibrotic Cytokine TGF-β and its Role in IPF

Transforming growth factor-beta (TGF-β) is a pleiotropic polypeptide cytokine which belongs to the transforming growth factor superfamily (Doerner *et al*. 2009). In humans there are three distinct mammalian TGF-β isoforms, (TGF-β1-3) which display both overlapping and unique functions (dependent on the cell type) (Nagaraj *et al*. 2010). TGF-β1 is the most commonly studied isoform in experimental lung fibrosis and is accepted as a key pro-fibrotic mediator in fibrotic disease (Froese *et al*. 2016) and IPF rodent models (Biernacka *et al*. 2011, Saito *et al*. 2018). TGF-β1 is produced by a wide variety of cell types including macrophages and alveolar epithelial cells, though myofibroblasts are the major source of TGF-β1 (Zhang *et al*. 1995). TGF-β1 is assembled as a latent complex in the endoplasmic reticulum and processed in the Golgi before it is secreted and stored as a latent complex in the tissue (Horiguchi *et al*. 2012).

Latent TGF-β can be activated via several mechanisms including the proteolytic cleavage between amino acids 278 and 279, yielding mature TGF-β of approximately 25 kDa (Khalil. 1999, Stoll *et al*. 2004, Jagadeesan and Bayat. 2007). The latent TGF-β1 interacts with a latency associated peptide (LAP), forming a complex termed as small latent complex (SLC) (Nunes *et al.* 1997, Robertson *et al.* 2015). This complex retains within cells until it is bound by another protein, the latent TGF-β-binding protein (LTBP) (Nunes *et al.* 1997, Robertson *et al.* 2015). This forms a large complex termed as the large latent complex (LLC). The LLC is then secreted into the extracellular matrix and is subjected to proteolytic cleavage (Nunes *et al*. 1997, Robertson *et al.* 2015).

Increased levels of active TGF-β1, along with TGF-β1 canonical signalling transducers and transcriptional modulators (SMAD-2/3) are activated in fibrotic rat lung and IPF tissue exposed to cyclic stretch (Froese *et al*. 2016). Furthermore, lung tissue displaying increased stiffness at baseline, when exposed to cyclic stretch, displays increased levels of active TGF-β1 expression.

Early reports displayed Rho associated protein kinase (ROCK) activity to increases matrix stiffness (Huang *et al*. 2012) and the inhibition of ROCK prevents myofibroblast differentiation and bleomycin induced pulmonary fibrosis (Jiang *et al*. 2012, Zhou *et al*. 2013). More recent data has shown ROCK inhibited IPF tissue attenuated levels of active TGF-β1 when exposed to cyclic stretch (Froese *et al*. 2016).

1.6.1 Integrins Regulate Mechanotransduction and Potentiation of TGF-β Signalling through Cell Matrix Interactions

Integrins are a family of cell surface heterodimeric glycoproteins composed of non-covalently bound α and β subunits (Alberts *et al*. 2002). They function primarily as receptors for extracellular matrix ligands and critically mediate bi-directional cross-talk between cells and their immediate environment (Feral *et al*. 2005), via signalling through various biochemical pathways (Jones *et al*. 1999).

Integrins are required for normal epithelial cell development and epithelial cells structurally rest on extracellular matrix (Sheppard. 1996, Merlin *et al*. 2001). Integrins expressed in airway epithelial cells function as mechanosensors detecting changes in cellular activity generated internally or externally through mechanical stress or matrix stiffness (Sheppard. 2003, Ingber. 2006, Tschumperlin *et al*. 2018). This triggered integrin activation can induce changes in the cells (Gkretsi and Stylianopoulos. 2018) (Figure 1.6).

The extracellular matrix environment is a critical regulator of cell behaviour and tissue mechanical properties behaviour (Yue. 2014). Interactions of cells and extracellular matrix regulate a variety of cellular functions including proliferation, differentiation, adhesion, migration and apoptosis (Yue. 2014) (Figure 1.6).

Integrin-cytoskeleton interactions with latent-TGF-β1 in the extracellular matrix function as a source for active TGF-β1. Integrin subtypes containing αv-β(1/3/5/6)-integrin isoforms can bind to latent-TGF-β1 (Brown and Marshal. 2019). These β-integrin isoforms recognise a tri-peptide motif (Arg-Gly-Asp (RGD)) on the latent-TGF-β1 pro-domain, allowing integrin binding (Munger and Sheppard. 2011, Shi *et al*. 2011). In addition to integrins activating latent-TGF-β1, integrins have the capacity to activate downstream TGF-β1-mediated pathways (Boo and Dagnino. 2013) (Figure 1.6).

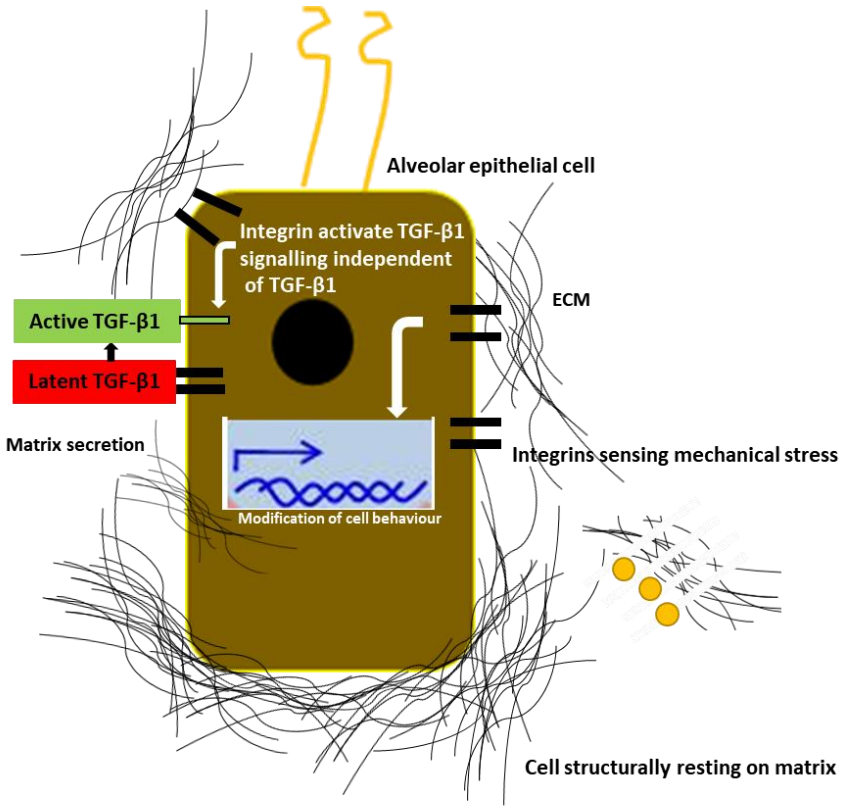


Figure 1.6, Integrin dependent mechanotransduction:cell and TGF-β1 signalling. Schematic diagram representing alveolar epithelial cells resting structurally on extracellular matrix (ECM). Integrins (black bars) are sat on the cell surface and can transduce intracellular signalling from interactions with the ECM or by sensing mechanical forces. Integrins also have the capacity to signal through the TGF-β pathway by binding to latent TGF-β1 (recognising RGD tripeptide) thereby forming active TGF-β1 which signals through the TGF-βRII (green bar) pathway. In response to ECM interactions integrins can activate TGF-βRII independent of TGF-β1 (depicted by the white upper arrow). The alveolar epithelial cells can function to secrete ECM, which has the potential in creating a positive feedback loop. ECM is also subjected to enzymatic remodelling (orange circles depicted as enzymes). (Sheppard. 1996, Merlin *et al*. 2001, Alberts *et al*. 2002, Sheppard. 2003, Feral *et al*. 2005, Ingber. 2006, Munger and Sheppard. 2011, Shi *et al*. 2011, Boo and Dagnino. 2013, Yue. 2014, Gkretsi and Stylianopoulos. 2018, Tschumperlin *et al*. 2018).

Both the inhibition and knockout of αvβ6-integrin in a mice model displayed reduced production of active TGF-β1 (Hahm *et al*. 2007). These set of experiments displayed inhibited accumulation of activated fibroblasts and deposition of interstitial collagen matrix. Furthermore, transcript profiling of tissues displayed reduced disease associated changes in fibrotic and inflammatory mediators including TGF-β1, TGF-βRII and collagen. Another study showed that using a generic αv-integrin small molecule inhibitor attenuates stretch induced TGF-β1 activation and SMAD-2/3 phosphorylation (Froese *et al*. 2010). This suggests a shared regulatory function of of αv-integrin and TGF-β1, and poses integrins as a potential fibrotic therapeutic target.

1.6.2 TGF-β Signalling Through the Recruitment of TGF-β-receptors

TGF-β1 signals via the activation of one of three receptors, type I, type II and type III receptors (TGF-βRI, TGF-βRII and TGF-βRIII) (Huang and Chen. 2012, Hata and Chen. 2016). TGF-βRIII is an accessory receptor, which binds to TGF-β1 recruiting it to bind to TGF-βRII (Boxall *et al*. 2006). This then leads to the hetero-oligomerization with TGF-βRI, allowing the constitutively active kinase domain of TGF-βRII to phosphorylate and activate the kinase domain of TGF-βRI (Boxall *et al*. 2006). This induces conformational changes leading to the downstream canonical intracellular signalling cascade through phosphorylating SMAD proteins, or non-canonical signalling pathways including through ROCK, mitogen activated protein kinases (MAPK), phosphinositide-3-kinases (PI3K) downstream signalling (Figure 1.7) (Zhang. 2009, Cichon *et al*. 2014, Boxall *et al*. 2016, Cervello *et al*. 2017, Finnson *et al.* 2020).

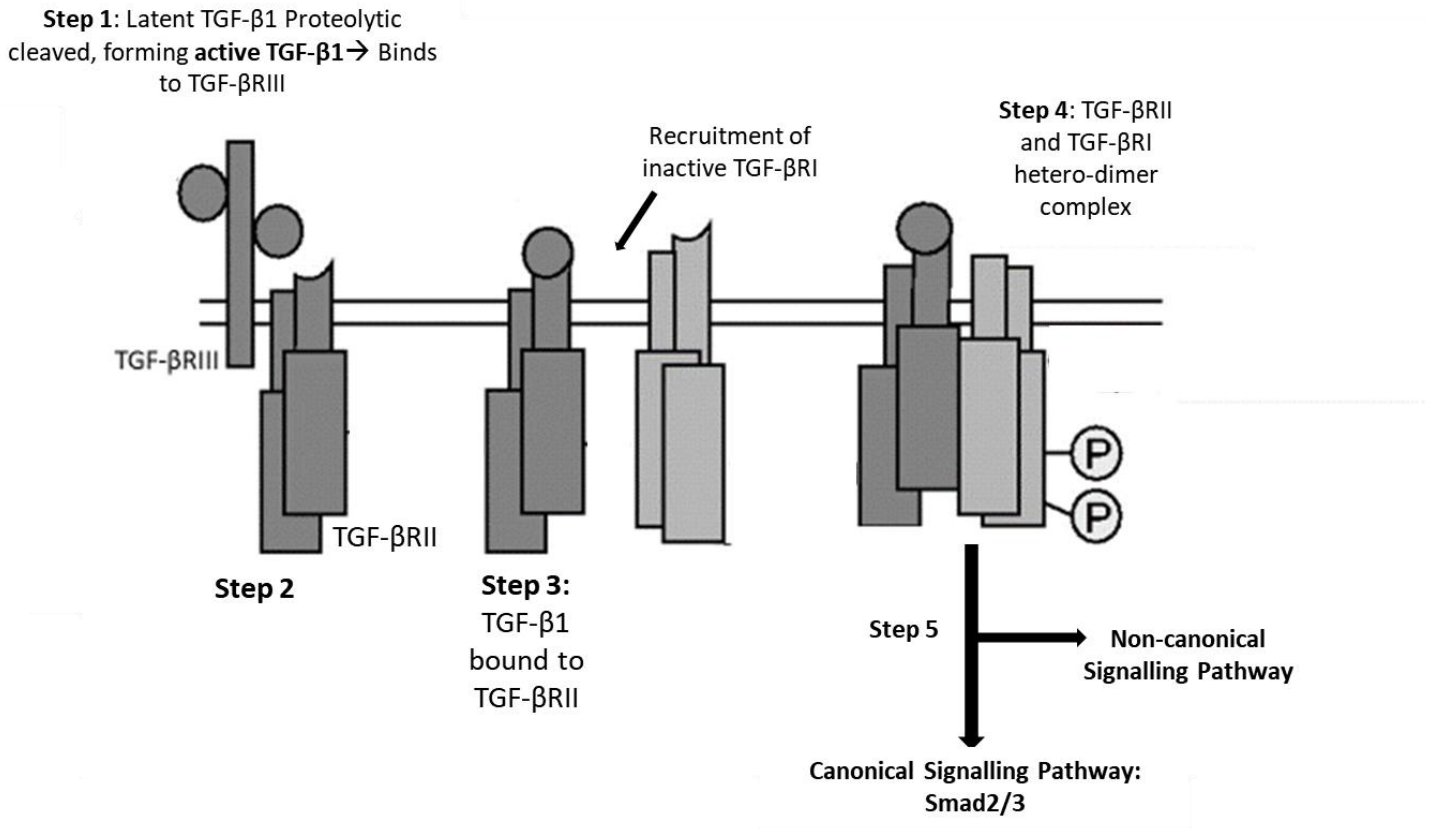


***Figure 1.7, Schematic representation of the TGF-β1 signalling cascade through the TGF-β-receptors*.** In **step 1**: Active TGF-β1 binds to the TGF-βRIII (accessory receptor). This increases the TGF-β1 affinity for TGF-βRII, and thus facilitating binding (**step 2**). Once TGF-β1 is bound to TGF-βRII (**step 3**), it recruits TGF-βRI. Within this heterodimeric complex, TGF-βRI is activated by phosphorylation (**step 4**). The TGF-β-activated TGF-β-Receptor complex transduces the signal into the cytoplasm via canonical SMAD, or non-canonical signalling cascades (**step 5**). (Boxall *et al*. 2006, Zhang. 2009, Huang and Chen. 2012, Cichon *et al*. 2014, Hata and Chen. 2016, Boxall *et al*. 2016, Cervello *et al*. 2017, Finnson *et al*. 2020).

Cross-talk of TGF-β1 and Wnt signalling pathways may be of relevance during the TGF-β1-induced myofibroblast differentiation (Akhmetshina *et al*. 2011). Canonical Wnt signalling is induced whilst downregulating the Wnt antagonist, Dicckopf-1. Conversely, over expression of Dicckopf-1 reduces TGF-β1 mediates fibrosis, demonstrating canonical Wnt signalling is required for TGF-β1 mediated fibrosis (Akhmetshina *et al*. 2011).

1.6.3 Experimental Models Assessing the Role of TGF-β in an IPF Context

Overexpression of TGF-β1 in rodent models drives a pulmonary fibrosis phenotype. TGF-β1 overexpression displayed similar results to bleomycin treatment (Kolb *et al*. 2001, Liu *et al*. 2001, Lee *et al*. 2004, Moore *et al*. 2013). Increased expression of extracellular matrix is observed with decreased epithelial (E-cadherin) and elevated mesenchymal (α-SMA) levels, consequently with the presence of elevated levels of myofibroblast and apoptosed alveolar epithelial cells. Furthermore, in these TGF-β1 overexpressed conditions, levels of protease inhibitors inhibiting enzymatic breakdown of extracellular matrix is elevated (Verrecchia and Mauviel. 2002). However, the progression to fibrosis with TGF-β1 overexpression is observed in the presence of low inflammation (Biernacka *et al*. 2011).

In vitro TGF-β1 overexpression results in overexaggerated, whilst inhibition results in the reduction, of wound healing responses (Crowe *et al*. 2000, MacKinnon *et al*. 2012). Potential molecular mechanisms include TGF-β1-induced reductions in levels of pro-apoptotic nitric oxide (NO) through inhibition of inducible NO synthase (iNOS), and through maintaining levels of the anti-apoptotic protein B-cell leukaemia (Bcl-)2 (Berg *et al*. 2007). In the context of wound healing this may prolong the myofibroblast response inappropriately.

TGF-β1 and its effector pathway have been considered potential targets for anti-fibrotic treatments (Shah *et al*. 1995). However, direct inhibition of TGF-β signalling is not favoured due to its importance in multiple mechanisms in different contexts (Xie *et al*. 2003). Targeting of molecular potentiators of this pathway with specificity for pulmonary fibrosis remains an attractive therapeutic option given its central importance in the disease. Pre-clinical assessments of drugs targeting non-canonical TGF-β1 signalling in experimental models of fibrosis have demonstrated encouraging results with decline in markers of fibrosis (including α-SMA and collagen type I) (Finnson *et al*. 2020). An example of such pre-clinical assessment includes the use

of the FDA and NICE licensed anti-fibrotic drug pirfenidone (Finnson *et al.* 2020). The LOTUSS Trial (NCT01933334) displayed decreased TGF-β1 mediated non-canonical signalling phosphorylation of AKT and ERK (Finnson *et al.* 2020).

Both canonical (SMAD2/3) and non-canonical TGF-β1 mediated signalling have the capacity to modulate senescence which is implicated in IPF (Tominaga and Suzuki. 2019). Modulating of senescence is achieved by several routes including the elevation in cell cycle regulators such as p16, p21, and p53, resulting in cell cycle arrest (Tominaga and Suzuki. 2019). Decreased production of TERT and increased production of ROS resulting in telomere shortening and DNA damage (Tominaga and Suzuki. 2019). Activation of chemokines and extracellular matrix regulators shift cells into SASP (Tominaga and Suzuki. 2019). SASP can then further amplify levels of pro-fibrotic/inflammatory factors including TGF-β1, thereby further amplifying TGF-β1 signalling (Tominaga and Suzuki. 2019).

1.7 Galectin-3, a Key Player in IPF Pathogenesis?

Whilst no biomarkers of IPF are in routine clinical use, a number have been proposed in literature (Ley *et al*. 2014, Drakopanagiotakis *et al*. 2018). One of these is the protein galectin-3 (Li *et al*. 2014, Dong *et al*. 2018). Galectin-3 is found elevated in multiple organ fibrosis, including the liver (Hara *et al*. 2020), kidney (Okamura *et al*. 2011), heart (Gonzalez *et al*. 2014) and quite prominently IPF lung (Mackinnon *et al*. 2012, Li *et al.* 2014). Elevated levels are found in the BALF and fibroblastic foci of IPF patients relative to controls (Nishi *et al.* 2007, MacKinnon *et al*. 2012, Ho *et al*. 2016). Furthermore, serum galectin-3 is a biomarker for interstitial lung abnormalities (ILAs) that may represent early IPF before the onset of clinical symptoms (Ho *et al*. 2016, Keane 2016). Moreover, galectin-3 appears to play an important role in mediating experimental models of IPF *in vitro* and *in vivo* (MacKinnon *et al*. 2012).

1.7.1 Galectin-3: Synthesis, Domain Architecture, Structural Behaviour and Intermolecular Interactions

Galectin-3 is a β-galactoside binding lectin (carbohydrate-binding protein) (Saraboji *et al*. 2011) that is encoded by the *LGALS3* gene located on chromosome 14 (Raimond *et al*. 1997). It is a member of the lectin family of which twelve human galectins have been identified (Cummings and Liu. 2009). Galectins can be classified into three groups according to their C-terminal-carbohydrate recognition domain (CRD) organisation: (i) the prototypic galectins (display a single CRD), (ii) tandem repeat-type galectins (display two bridged-CRD), and (iii) chimeric type galectin (display CRD fused to another non-lectin (large amino terminal domain)) (Cummings and Liu. 2009, Modenutti *et al.* 2019) (Figure 1.8). The C-terminal domain constitutes ~130-140 amino acid residues, folding into a five (F-sheet, convex side) versus six (S-sheet, concave side) stranded anti-parallel β-sheet sandwich (Figure 1.9) (Cummings *et al*. 2017, Modenutti *et al*. 2019).

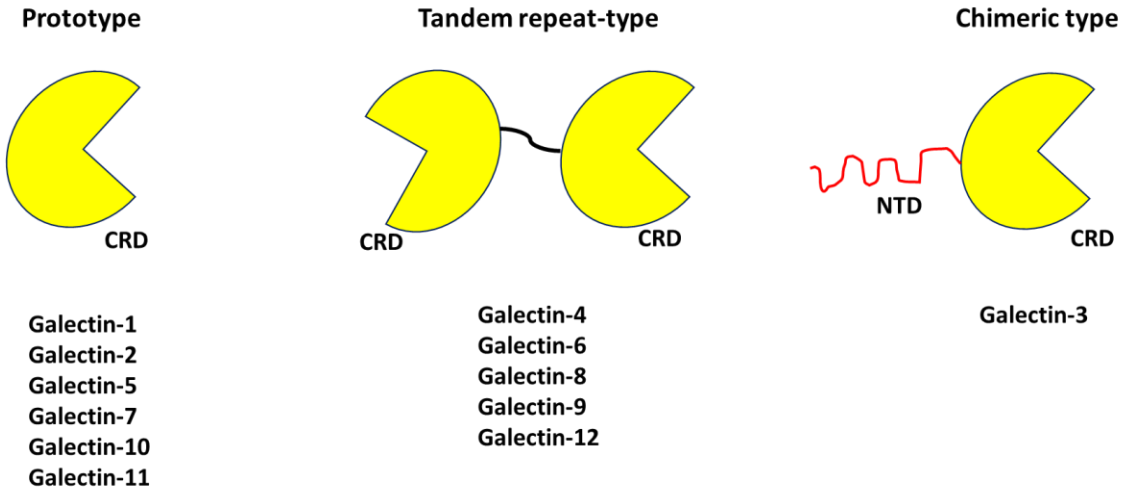


Figure 1.8, Cartoon representation of the human galectin family. The twelve human galectins are classified into three groups according to their C-terminal- carbohydrate recognition domain (CRD) organisation. These include prototype (single CRD), tandem repeat-type (two bridged-CRD), and chimeric type (CRD fused to a large N-terminal domain (NTD)). The CRD is illustrated as a yellow 'pac-man' motif. Galectin-3 is the only chimeric galectin-3 (NTD displayed as a red squiggly line). (Cummings and Liu. 2009, Modenutti *et al*. 2019).

Galectin-3 is synthesised as a cytosolic protein and is found in multiple compartments (Alvarez and Ortega. 2017) including in the nucleus and the cytoplasm, in arrays either side of the cell membrane, and at the cell surface. The extracellular secretion mechanism involves a non-classical secretion pathway avoiding the endoplasmic reticulum (ER) and Golgi complex since it lacks typical signal sequence for ER translocation (Lindstedt *et al*. 1993). Galectin-3 is expressed in various cell types, but high levels are found in macrophage, epithelial and myofibroblast cells (Nishi *et al*. 2007, MacKinnon *et al*. 2012, Ho *et al*. 2016).

Galectin-3 uniquely combines the characteristic galectin carbohydrate recognition domain (Figure 1.9) (Modenutti *et al*. 2019), as a ~15 kDa C-terminal domain in a unique chimeric arrangement with a ~15 kDa N-terminal domain (Menon and Hughes. 2001, Kopitz *et al*. 2014).

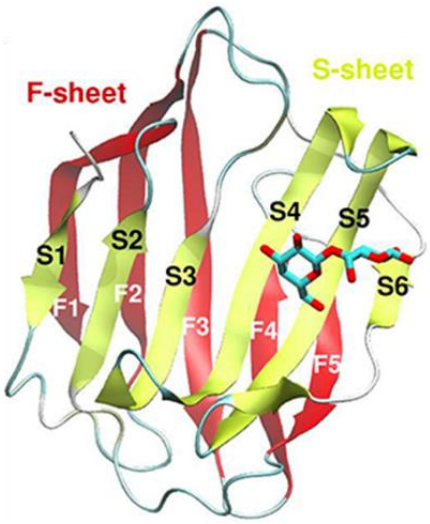


Figure 1.9, Galectin typical CRD arrangement and accommodation of β-galactoside. Ribbon diagram of the crystal structure of human galectin-1 complexed with lactose. Detail of the galectin-1 monomer differentiating the 'S-sheet' (strands S1–S6) in yellow, and the 'F-sheet' (strands F1–F5) in red, forming the β-sandwich. S-sheet S4, S5 and S6 mediate β-galactoside recognition. (Modenutti *et al*. 2019)

Galectin-3 has specificity for N-acetyllactosamine (LacNAc)-modified glycans on glycoproteins (Ahmad *et al*. 2003, Dam *et al.* 2005, Tejler *et al*. 2005, Bumba *et al.* 2018). Many NMR and crystallographic structures have been published with (Sorme *et al*. 2005, Shilova *et al*. 2020) and without ligand binding (Seetheraman *et al*. 1998, Umemoto and Leffler. 2001, Manzoni *et al*. 2018). The N-terminal domain consists of 21 amino-acids known as the N-terminal stretch, which contains two serine residues that are phosphorylation sites, followed by nine Pro/Gly/Tyr-rich repeats (I-IX) (Ibarra *et al*. 2018). The first twelve amino-acids, including phosphorylated Ser6, are critical for both nuclear localisation and extracellular secretion (Gong *et al*. 1999, Takenaka *et al*. 2004, Lin *et al*. 2017). The N-terminal domain appears to adopt an intrinsically disordered state in isolation (Lin *et al*. 2017). Galectin-3 monomers are partially constrained by contacts with the C-terminal domain when in an apo- (non-ligand bound) state. Upon binding to a carbohydrate ligand the N-terminus is released (Halimi *et al*. 2014, Lin *et al*. 2017, Chiu *et al*. 2020). N-terminal release is essential for galectin-3 oligomerisation, which can occur either via homotypic interactions between N-N terminal domains (N-type oligomerisation) or between C-C terminal domains (C-type oligomerisation) (Nieminen *et al*. 2007, Lepur *et al*. 2012). More recently, the oligomeric behaviour has been described as "fuzzy" as the multimerisation behaviour could also be due to homotypic N-C terminal domain interactions (Lin *et al*. 2017). It is currently not known how these oligomerisation mechanisms, observed in cell free *in vitro* conditions relate to galectin-3 oligomerisation behaviour *in vivo*. Nevertheless, the abilities to undergo hetero- and homotypic interactions via protein:glycan and protein:protein interactions respectively provide galectin-3 with extensive potential to nucleate lattice structures incorporating a range of proteins and glycoproteins. Experiments with truncated variants of the N-terminal domain demonstrate its important in both *cis*-binding to cell surface binding partners and *trans*-binding (or cell-bridging) (Kopitz *et al*. 2014).

At the cell surface, galectin-3 can bind to a wide range of proteins with varying degrees of evidence (Loimaranta *et al*. 2018). Notably in the interest of IPF, there is strong evidence for direct binding to TGF-βRII (Partridge *et al.* 2004, Konigshoff. 2012), and extensive evidence for functional interactions with the transmembrane amino-acid transporter channel CD98 (Dong and

Hughes. 1997, Dalton *et al*. 2007, MacKinnon *et al*. 2008). Other proposed binding partners of interest include EGFR, a potent mitogen for collagen producing mesenchymal cells (Markowska *et al*. 2011, Boscher *et al*. 2013).

1.7.2 Galectin-3 Mediates Lung Inflammation and Fibrosis

Early studies of galectin-3 in models of ILD described significant increase in radiation-induced lung inflammation and repair in mice (Kasper *et al*. 1996). It was observed at the cell membrane of newly synthesised type I alveolar epithelial cells, and at the apical surface of type II alveolar epithelial cells in this model. Elevated galectin-3 levels were then reported in the BALF of IPF patients, though corticosteroid treatments abrogated this elevation (Nishi *et al*. 2007). Administration of purified galectin-3 can induce the release of pro-inflammatory cytokines including TNF-α and IL-8 in macrophages (Filer *et al*. 2008), inducing migration and collagen synthesis (Nishi *et al*. 2007). Galectin-3 may therefore mediate inflammatory and cell migration effects, which are of biological importance in IPF pathogenesis.

MacKinnon *et al*. 2012, demonstrated increased levels of galectin-3 in IPF versus non-fibrotic control human lung tissue. Galectin-3 knockdown mouse models showed dramatic attenuation of fibrotic changes induced by TGF-β1 overexpression or bleomycin administration. Similar protection was afforded by treatment with a galectin-3 antagonist designed to block glycan binding, the small molecule TD139 (3,3'-Bis-(4-aryltriazol-1-yl) thio-digalactoside).

A549 lung epithelial cells subjected to galectin-3 silencing were resistant to TGF-β1-induced EMT and showed reduced wound healing responses relative to controls. Moreover, primary lung myofibroblasts from the mouse galectin-3 knockdown model displayed reduced expression of collagen in response to TGF-β1 stimulation (MacKinnon *et al*. 2012).

Taken together, these findings suggest that galectin-3 antagonism represents a promising therapeutic approach for IPF patients.

1.7.3 TD139 Inhibits Galectin-3 in IPF Lung

Currently TD139 (now termed GB0139) is in a large cohort phase II/III clinical trial; based on phase I/II studies displaying promising effect as an anti-fibrotic medication (Hirani *et al*. 2017). It has high (~14 nM K_d) affinity for the carbohydrate recognition domain of galectin-3 (Mackinnon *et al*. 2012, Delaine *et al*. 2016, Hirani *et al*. 2017). TD139 is the first inhaled therapeutic agent to be investigated in an IPF clinical setting (Hirani *et al.* 2020).

In vitro, *in vivo* and *ex vivo* studies using TD139 displays its effect on all key IPF cell types modulating fibroblast activation and reducing effect of pro-fibrotic growth factors that act on myofibroblasts and epithelial cells (Mackinnon *et al*. 2012).

The anti-fibrotic role is central to the inhibition of the recruitment and expansion of galectin-3 secreting cells driving local myofibroblast activation. The phase I/IIa clinical study was a randomised, double-blind, placebo-controlled trial, allowing the assessment of drug safety, tolerability, pharmokinetics and pharmodynamic properties (Hirani *et al*. 2017, 2020).

This study was divided into two-parts (a and b). a) included healthy subjects and demonstrated single doses of TD139 between 0.15 mg and 50 mg were well tolerated. This success then led to b), which assessed multiple ascending dose of TD139 in IPF individuals (Hirani *et al*. 2017, 2020).

Levels of galectin-3 were assessed in serum and BAL macrophages (Hirani *et al.* 2017, 2020). Galectin-3 levels were inhibited in IPF subjects in a dose dependent manner. A 10 mg dose responded to give maximal inhibition. In addition, other pro-fibrotic biomarkers, e.g. PDGF, were

observed to decrease. No patients were discontinued from the treatment arm since drug-related adverse effects were minimal (Hirani *et al*. 2020).

1.8 CD98, Structure, Function, and Inflammatory Relevance

In IPF, mechanotransductive cell:matrix interactions potentiate pro-fibrotic cytokine effects (Duscher *et al.* 2014). CD98 is a putative binding partner of galectin-3 with potential to mediate these pathways. CD98 has been implicated in proliferative (Bulus *et al.* 2012, Cantor and Ginsberg. 2012) and pro-inflammatory pathways (Nguyen and Merlin. 2012). It is also a binding partner of integrins (Kolesnikova *et al*. 2001, Cai *et al.* 2005, Kim and Hahn. 2008).

CD98 is a 125 kDa disulphide linked heterodimeric type II transmembrane glycoprotein (Cantor and Ginsberg. 2012). It is composed of an ~85 kDa glycosylated heavy chain (CD98hc) encoded by the *solute carrier 3 member 2 (SLC3A2)* gene and a variable ~40 kDa non-glycosylated permease light chain encoded by the *solute carrier 7 member (SLC7A)* gene that is almost entirely membrane-embedded and contains the solute channel (Kim *et al*. 2008, MacKinnon *et al*. 2008, Cantor and Ginsberg. 2012). The extracellular domain of CD98hc forms a disulphide bond with the light chain and the complex functions as a transporter with specificity for essential neutral 'L-type' amino acids (Ballina *et al*. 2016).

Initial findings reported the CD98 heterodimer functioned as a lymphocyte activation-related antigen (Haynes *et al*. 1981). This T-cell activation was observed to mediate CD3-mediated signalling independent of exchanges of T-cell proteins (Haynes *et al*. 1981, Cantor *et al*. 2011). CD98 also contributes to the direction of cell-cell interactions and cellular migration mechanisms resulting in activation of lymphocytes (Deves and Boyd. 2000). It regulates lymphocyte function by participating in a post-transcriptional mechanism regulating accumulation of activated lymphocytes prior to cell proliferation (Nguyen and Merlin. 2012, Han *et al*. 2018). CD98 is

expressed in all cell types excluding platelets (Yan *et al*. 2008). Highest expression of CD98 is observed in the gastrointestinal tract and kidney tubules. CD98 is termed as a 'molecular facilitator' in regulating different types of adhesion molecules including lymphocyte function associated antigen-1 (LFA-1) in the gastrointestinal tract, denoting CD98 as a key player in intestinal inflammation (Yan *et al*. 2006, 2008). Pro-inflammatory cytokines including interferon-γ (IFN-γ) increase CD98 expression levels, whilst anti-inflammatory agents such as 5-aminosalicylic acid inhibit its expression levels in a dose dependent manner (Yan *et al.* 2006, 2008).

Experimental treatment of cells with the Gram-negative bacterial membrane component endotoxin or lipopolysaccharide (LPS), as an acute injury model increases mRNA levels of CD98hc (Rittirsch *et al*. 2008, Laroui *et al*. 2014, Canup *et al*. 2019). Furthermore, data from colleagues in our (Gooptu) group (*Manuscript in preparation*) and Laroui *et al*. 2014, have studied the effects of siRNA silencing of CD98hc or inhibition of integrin-dependent CD98 functions using cynaropicrin. These treatments abrogated release of the pro-inflammatory cytokines IL-6, IL-8 and TNF-α in response to multiple experimental stimuli mimicking acute lung injury and acute exacerbations of IPF and intestinal inflammation, *in vitro*, *in vivo* and in *ex vivo* models.

1.8.1 CD98:β-integrin Extracellular Matrix Interactions

CD98 appears to be constitutively associated with β1-integrin and to facilitate β1-integrin dependent cell spreading, adhesion and migration, and protection from apoptosis (Cai *et al*, 2005, Feral *et al*. 2005, Prager *et al*. 2007). However, dependent on the cellular context, the association of CD98 with β3-integrin is of importance in its downstream signalling pathways (Salmani *et al*. 2008). Exogenous CD98 stimulation increases β-integrin affinity for its ligands, along with increased clustering on the cell surface (Cai *et al*. 2005, Kim and Hahn. 2008, Salmani *et al*. 2008, Poettler *et al.* 2013).

In vitro studies initially localised the CD98 interaction with β1-integrin exclusively to the heavy-chain (Fenczik *et al*. 2001, Feral *et al.* 2005). Subsequent studies defined it more closely to the CD98hc transmembrane sequence (Kolesnikova *et al*. 2001) and then, with the aid of chimeric CD98 constructs, to a specific amino acid sequence (WALLL) at residues 82-86 within the transmembrane domain (Henderson *et al*. 2004, Cai *et al.* 2005).

CD98hc enhances integrin-activated adhesion signalling by increasing its affinity to other binding partners corresponding to downstream β-integrin-mediated signalling pathways including FAK, AKT, RAS and PI3K activation (Cai *et al*. 2005, Kim and Hahn. 2008, Salmani *et al.* 2008). Reciprocally, CD98hc requires the presence of β1-integrin to activate intracellular signalling (Rintoul *et al*. 2002).

CD98hc acts as an amplifier of integrin-dependent mechanotransduction. The increase in stiffness sensing provided by CD98hc forms the core of a positive feedback loop promoting further stiffening. Knockout of CD98hc in mice fibroblasts reduces tissue stiffness and increases elasticity (Estrach *et al*. 2014). Activation of integrin-regulated promoters involved in cell survival, proliferation, and cell invasion, namely FAK and AKT, were reduced. Both wild-type and a mutant CD98hc (with retained capacity to couple to integrins) rescued the defect in stiffness sensing in the fibroblasts. Conversely, mutant CD98 variants facilitating amino acid transport without integrin association failed to rescue stiffness sensing. Taken together, this data indicates that CD98hc:integrin interactions modulate the ability of a cell to sense the degree of matrix stiffness, and also suggests integrins bridge the gap between CD98hc and its capacity to sense matrix stiffness (Estrach *et al*. 2014).

CD98hc also contributes to regulation of the integrin-dependent signalling Rho GTPases (Salmani *et al*. 2008). RhoA signalling is involved in matrix stiffness and CD98hc has been implicated in the regulation of RhoA in fibroblasts (Estrach *et al*. 2014). CD98hc triggers integrin derived RhoA signalling allowing cells to apply force on the extracellular matrix (Cantor and Ginsberg. 2012).

1.9 Interactions of CD98hc and β1-integrin with Galectin-3 in IPF

Galectin-3 is known to contribute to acute inflammation via its involvement in neutrophil activation and chemoattraction of monocytes and macrophages (Henderson and Sethi. 2009).

CD98 is highly expressed on both T-cells and macrophages and has been reported as a galectin-3 binding partner (MacKinnon *et al*. 2008, Chen *et al.* 2014). This was defined by affinity chromatography of Jurkat cell lysates with CD98 binding to galectin-3-Sepharose affinity column (Dong and Hughes. 1996, 1997).

The interaction of galectin-3 with CD98 in macrophages promotes increased expression of pro-inflammatory cytokines (Markovic *et al*. 2016, Alvarez. 2017). Galectin-3 binds CD98 to orchestrate the activation and differentiation of bone marrow-derived alternative macrophages (MacKinnon *et al*. 2008).

β1-integrin carries a β-1,6-N-acytyl-D-glycosamine branched glycan, which functions as a ligand for galectin-3 (Feral *et al*. 2005). Reciprocal co-immunoprecipitation confirms galectin-3 associates with β1-integrin in retinal pigment epithelial cells (Priglinger *et al*. 2013). Addition of exogenous galectin-3 enhances β1-integrin lateral mobility (Yang *et al*. 2017). Furthermore, it activates integrin signalling similarly to antibody-mediated cross-linking of CD98 on the cell surface (Henderson *et al.* 2004). Galectin-3 also cooperates with β1-integrin to induce EMT of epithelial cells, with apparent increases in β1-integrin-mediated cell adhesion to the extracellular matrix proteins fibronectin and laminin in alveolar cells and fibroblasts (Margadant *et al*. 2013).

1.10 Hypothesis

Taken together, these data suggest that galectin-3 may mediate fibrosis in the lung and other organs through its architectural, extended lattice-forming capabilities that can facilitate the clustering of pro-fibrotic co-factors, notably CD98, integrins and TGF-βRII. My project explores the hypothesis that galectin-3 nucleates an ordered assembly that stabilises and spatiotemporally clusters a pro-fibrotic macromolecular assembly at the surface of alveolar epithelial cells and myofibroblast cells by bridging integrins and TGF-β1 mediated signalling. Experimental data suggest the involvement of at least four components: galectin-3, CD98hc (heavy chain), β1-integrin and TGF-βRII (Figure 1.10) (Estrach *et al*. 2014, Hashiba *et al*. 2014). The Gooptu group term this putative assembly the 'gal-3-fibrosome'.

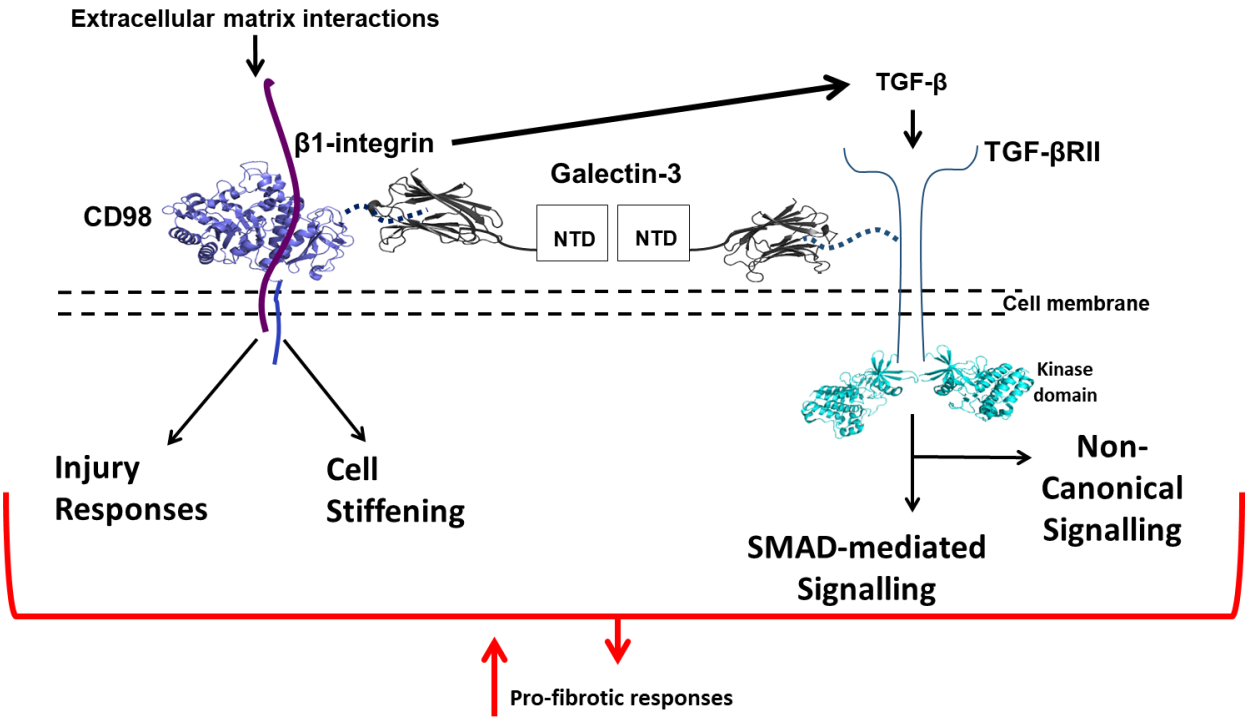


Figure 1.10, Working model of the galectin-3-fibrosome organisation. Schematic view of a galectin-3-fibrosome perpendicular to the cell surface. Simplest form of galectin-3 self-association shown (dimeric). Galectin-3 oligomerises through its N-terminal domain (NTD). Galectin-3 associates with TGF-βRII likely via C-terminal domain:glycan interactions. The C-terminus of a linked molecule of galectin-3 can similarly associate with β1-integrin through a glycan chain motif. β1-integrins consistently found in complex with CD98 heavy chain . In addition to the molecules depicted here, β1-integrin will be complexed to an α-integrin chain, and may interact with fibrillar adhesions via tensin-1, whilst TGF-βRII exists as heterodimers with TGF-βRI. (Miah *et al.* British Thoracic Society Winter Meeting 2019).

1.11 A Global Pandemic - Covid Pneumonitis and its Severity with Pre-Existing ILD

Coronavirus disease-2019 (COVID-19) is caused by infection with severe acute respiratory syndrome-coronavirus (SARS-CoV)-2 (Peterson *et al*. 2020). Its emergence as a novel pathogen resulted in a global pandemic as declared by the World Health Organisation (WHO) in March 2020. Most severe form of the disease, and death in COVID-19 results from severe lung inflammation (COVID pneumonitis) associated with high levels of pro-inflammatory cytokines, notably interleukin-6 (IL-6) (Magro. 2020).

The spike protein of SARS-CoV-2 is responsible for initial contact with the host cell. By analogy with SARS(-CoV-1) (Jia *et al*. 2005, Mutsuyama *et al*. 2010), host cell infection is considered to be initiated when the spike protein binds to a transmembrane form of the protein angiotensin-converting enzyme (ACE-)2 (Hoffmann *et al*. 2020, Sungnak *et al*. 2020). It is then processed by transmembrane protease serine 2 protease (TMPRSS2). Cellular endocytosis can proceed via a clathrin- and caveolin-independent pathway (Yang *et al*. 2020), although conventional endocytosis may also occur (Glebov. 2020).

Since the onset of the COVID-19 pandemic, researchers have investigated whether patients with pre-existing ILD are more susceptible to COVID-19 than those without ILD and have further evaluated the impact of ILD on disease severity in patients with COVID-19 (Esposito *et al*. 2020, Lee *et al*. 2021).

Esposito *et al*. 2020, established in their study increased mortality of COVID-19 positive patients with pre-existing ILD compared to control subjects without ILD (33% versus 13%). This represented an increased odds ratio of death in patients with ILD of 3.2. Increased mortality was

observed post adjustment for age, sex, race, smoking status and any chronic immunosuppression (odds ratio, 4.3).

Lee *et al*. 2021, established in their study patients with ILD were more likely to have severe COVID-19 than patients without ILD (49.3% versus 13.1%), including increased mortality (13.4% versus 2.8%). The risk of severe COVID-19 was significantly higher in patients with pre-existing ILD than in those without ILD (odds ratio, 2.32).

With the aid of HRCT scans, IPF like characteristics (UIP patterns) have been observed in patients with severe or critical COVID-19 when compared to patients with moderate or mild COVID-19 (McGroder *et al.* 2021, Zou *et al.* 2021). Observations of linear opacities, interlobular septal thickening, reticulation, honeycombing were evident. The more severe the clinical classification of COVID-19, the more severe the residual pulmonary fibrotic like characteristics were apparent. However, in most cases, pulmonary fibrotic like patterns were resolved significantly post 90 days after patient discharge (McGroder *et al.* 2021, Zou *et al.* 2021).

1.11.1 Topical Objective: The Putative Gal-3-Fibrosome Hypothesis Potentially Mediates Covid Pneumonitis

Remarkably, the SARS-CoV-2 spike protein has evolved to contain an arginine-glycine-aspartic acid (RGD) tri-peptide amino-acid motif within the receptor binding domain (RBD) (Lan *et al*. 2020). This motif binds integrins including β1-integrin, and is not seen in other known coronaviruses including SARS-CoV-1 (Sigrist *et al*. 2020).

The SARS-CoV-2 spike protein S1 region comprises of an N-terminal domain (NTD) that is characteristic in the β-coronavirus family and adopts a galectin fold with carbohydrate ligand binding capabilities, and the neighbouring receptor binding domain (Caniglia *et al*. 2020). The N-terminal and receptor binding domain therefore have considerable potential to bind to galectin-3 and/or CD98:integrin complex on the cell surface.

We therefore hypothesised that the S1 region of the SARS-CoV-2 interacts with CD98hc:integrin complex as a co-receptor on cells and such interactions (with broader gal-3-fibrosome protein involvement) trigger hyper-inflammatory cytokine responses.

1.12 Aims and Objective

This project's goal is to explore the gal-3-fibrosome hypothesis and get a better understanding of the relationships between galectin-3, CD98 and β1-integrin in IPF using *in vitro* and *ex vivo* models of the disease. I also aimed to understand the multimeric behaviour of galectin-3 and other relevant protein:protein interactions using biochemical and biophysical methods, since this might define a novel strategy for anti-fibrotic therapy. The specific objectives are therefore:

1) To examine cell surface expression levels of galectin-3, CD98hc, β1-integrin and TGF-βRII expression levels under basal and pro-fibrotic conditions in alveolar epithelial cells and in both non-fibrotic control and IPF-derived myofibroblasts

2) To examine co-localisation of galectin-3 with CD98, β1-integrin and TGF-βRII under pro-fibrotic conditions in alveolar epithelial cells and in both non-fibrotic control and IPF-derived myofibroblasts

3) To assess the effects of CD98hc knockdown on pro-fibrotic functional read-outs

4) To express and purify full-length recombinant galectin-3 and to probe galectin-3 oligomeric behaviour using appropriate carbohydrate ligands and investigating a favourable oligomeric state for analysis by electron microscopy

5) To extract endogenously expressed complexes of gal-3-fibrosome components from the cell membrane in a lipid (nanodisc) environment for electron microscopy studies

Lastly the unexpected onset of the COVID pandemic during my PhD research resulted in pursuit of another related objective (see Chapter 6):

6) To assess whether the spike protein SARS-CoV-2 may stimulate observed increases in pro-inflammatory cytokines IL-6 and IL-8 via a galectin-3 and CD98-dependent mechanism.

Chapter 2

Materials and Methods

2.1 Cell Culture

2.1.1 Cell Line

Human A549 cells (derived from human broncho-alveolar adenocarcinoma cells) were used to model Type II alveolar epithelial cell behaviour; this was obtained from ATCC-LCG, UK. Human THP-1 cells (derived from human monocytic leukemia) were used to model macrophage cell type behaviour.

2.1.1.1 A549 Cell Culture (in vitro)

A549 cells were cultured in sterile tissue cell culture 75 cm^2 flasks (Greiner Bio-One Ltd, UK), in Dulbecco's Modified Eagle's Medium (DMEM) + (4.5g/L D-Glucose, L-Glutamine), incubated at 37 °C under an atmosphere of 95% air/5% CO_2. DMEM was supplemented with 10% (v/v) foetal bovine serum (FBS) and 1% (v/v) penicillin/streptomycin (100 IU/ml and 100 µg/ml). The cells were passaged when the flasks reached ~70-90% confluency by aspirating the media and washing the cells twice with 10 ml of Hank's Balanced Salt Solution (HBSS), thereafter the cells were detached from the flask using 5 ml of 0.25% Trypsin-EDTA, incubating for ~5 minutes at 37 °C. The trypsin was neutralised by adding equal amount of cell culture media. Cells were then centrifuged at 220 *x g* for 5 minutes; the supernatant was discarded, and the cell pellet was re-suspended in 1 ml of media. Cells were counted and viability was assessed via the trypan blue method (using a haemocytometer counting chamber). The cells in suspension were then seeded into 75 cm^2 flasks (seeding density of ~$2x10^5$cells in 10 ml of pre-warmed media).

2.1.1.2 Screening A549 Cell Stocks for Mycoplasma Contamination

Mycoplasma, a self-replicating prokaryote characterised by the lack of a cell wall is a common contaminant that can cause adverse effects in cell culture (Nikfarjam and Farzaneh, 2012). These include alterations in growth rate, gene expression and modulation of cytokine production without obvious indication of contamination such as medium turbidity.

To confirm the A549 cells used in my study were mycoplasma-free, we used a commercially available 'PCR Mycoplasma screening detection kit' (Alfa Aesar, UK). This kit is sensitive for over 70 types of Mycoplasma. The A549 cell line, cultured at multiple passages (8, 12, 16, 20), used in my study were demonstrated to be free from Mycoplasma infection (Appendix, Figure A3).

2.1.1.3 THP-1 Cell Culture (in vitro)

THP-1 suspension cells were cultured in non-TC treated 75 cm^2 canted neck flask with a vented cap (Scientific Laboratory Supplies Ltd, UK), in Roswell Park Memorial Institute (RPMI 1640) + (20 mM HEPES, L-Glutamine) media, incubated at 37 °C under an atmosphere of 95% air/5% CO2. RPMI was supplemented with 10% (v/v) heat inactivated FBS and 1% (v/v) penicillin/streptomycin (100 IU/ml and 100 µg/ml). During the incubation step the flasks were placed in a standing upright position. The cells were passaged if the cell count exceeded a confluence of 800,000 cells/ml. Cell count was conducted by pipetting the media with slight agitation of the flask, followed by withdrawing 1 ml of media into an Eppendorf tube. Cells were counted and viability was assessed via the trypan blue method (using a haemocytometer counting chamber). The cell number was adjusted to 500,000 cells/ml with by adding fresh RPMI media. 20 ml of the cell suspension was added into a new 75 cm^2 flask (yielding a total of 10,000,000 cells).

2.1.1.4 Differentiating THP-1 Cells into M0 Macrophage Phenotype

THP-1 cells were differentiated from a resting phase to a macrophage phenotype by using Phorbal-12-myristate 13 acetate (PMA). This was added to freshy passaged cells. PMA (1 mM stock) was diluted 1:5000 to the cell culture media (200 nM final concentration). Once added to the cells, this was cultured for 72 h at 37 °C under an atmosphere of 95% air/5% CO_2. PMA transitions the cells to become adherent. Post 72 h incubation period, the growth media was aspirated and replaced with fresh RPMI media and cultured for a further 48 h at 37 °C under an atmosphere of 95% air/5% CO_2.

2.1.2 Primary Cell Isolation and Culture (ex vivo)

For all primary cell experiments in this study, non-fibrotic control (NFC) and idiopathic pulmonary fibrosis (IPF)-derived human lung myofibroblasts (HLMFs) cells were used.

2.1.2.1 Non-fibrotic Control Derived Human Lung Myofibroblasts

NFC HLMFs were purified *ex vivo* from healthy tissue from patients undergoing lung resection for carcinoma at Glenfield Hospital, University Hospitals of Leicester NHS Trust. No morphological evidence of disease was visually apparent in the tissue samples utilised for myofibroblast isolation. All subjects or donors gave written informed consent. The study was approved by Leicestershire, Northamptonshire and Rutland Research Ethics Committee 2 (approval codes: 07/MRE08/42, 04/Q2502/74 AND 0//H0406/189).

1-2 mm^3 of lung tissue were placed into flat-bottomed 6 well plates. For the tissue to adhere to the plate, the plate was incubated at 37 °C for 1 h in a humidified chamber (95% air/5% CO_2). After adherence, 2 ml of myofibroblast growth media (DMEM+4.5g/l D-Glucose, L-Glutamine) supplemented with 10% (v/v) FBS, 1% (v/v) Penicillin/Streptomycin and 1% (v/v) Non-Essential Amino Acids (NEAA) was gently added to each well. Care was taken to prevent detachment of the lung pieces. Plates were then placed back into the 37 °C incubator and observed (under an inverted light microscope) every 24 h to monitor cell growth, with myofibroblast media replenished every 48 h. When cells reached a confluence state of ~80-90% (~7-9 days) in each well, the tissue was removed, the media was aspirated, and the cells were washed with HBSS. Cells were subsequently detached using 0.25% Trypsin-EDTA. The trypsin was neutralised by adding equal amount of cell culture media. Cells were then centrifuged at 295 *x g* for 8 minutes. The supernatant was discarded, and the cell pellet was re-suspended in 1ml of media. Cells were counted and viability of sample was assessed in a cytometer chamber using trypan blue, before seeding into a 75 cm^2 flask (~$1.5x10^5$ cells in 10 ml of media). This was defined as Passage 1. Subsequently, cells were washed and fed every ~3 days. The cells were characterised, and the myofibroblast identity of the cells was confirmed by demonstration of the typical stellate morphology (observable under an inverted light microscope) and more definitively by immunofluorescence staining. Cells were stained for the following panel of 'positive' biomarkers supportive of myofibroblast identity: alpha-smooth muscle actin (α-SMA), Thy-1, Fibroblast Surface Protein (FSP), vimentin, and collagen. Additional biomarkers supporting alternative phenotypes, 'negative' myofibroblast characterising biomarkers, included CD68 and CD34. The cells were then passaged (as described above) once they reached a confluence of 80-90% and then harvested up to passages 4-5 for downstream experiments.

2.1.2.2 IPF Derived Human Lung Myofibroblasts

IPF HLMFs were derived from patients with suspected symptoms of IPF undergoing lung biopsy for diagnostic purposes at the University of South Carolina, USA, obtained through our

collaborator Carol Feghali-Bostwick. Patients displayed usual interstitial pneumonia (UIP) on histological examination. Written consent was obtained from all IPF subjects, in accordance with the University of South Carolina institutional review board. IPF derived human lung myofibroblasts were cultured in identical conditions as NFC HLMFs (2.1.2.1).

2.2 Cell Stimulation with the Pro-Fibrotic Cytokine TGF-β1

A549, non-fibrotic and IPF derived HLMF cells were grown to ~70% confluency. Growth media was aspirated and cells were synchronised to the same phase of the cell cycle via serum starvation using a low-FBS media (DMEM+4.5g/L D-Glucose, L-Glutamine supplemented with 1% (v/v) Insulin-Transferrin-Selenium (ITS), 1% (v/v) P/S and 1% (v/v) NEAA, for 24 h. Cells were then stimulated by the addition of the pro-fibrotic cytokine recombinant human (rh)TGF-β1 at a final concentration of 10 ng/ml (unless otherwise specified), diluted in the appropriate growth media for the specific cell type.

2.3 Lysate Extraction and Western-blotting

2.3.1 Whole cell lysate (WCL) Preparation from Human Cells

75 cm^2 cell culture flasks of A549, NFC and IPF-derived HLMF cells were untreated or treated with 10 ng/ml of rhTGF-β1 for various time points. Lysates were prepared by aspirating the media and washing the cells twice with ice-cold 1x Dulbecco's Phosphate Buffer Saline (DPBS). DPBS was then aspirated, and cells were lysed in 0.7 ml of Radio-Immuno-Precipitation Assay (RIPA) lysis buffer (sc-24948, Santa-Cruz), supplemented with 1% (v/v) protease inhibitor cocktail (PIC), 1% (v/v) phenylmethanesulfonylfluoride (PMSF) and 1% (v/v) sodium orthovanadate (Na_3VO_4) (proportional to the volume of RIPA). Flasks were placed on ice and the adherent cells were

scraped off using a sterile disposable cell scraper and collated together followed by gentle pipetting, subsequently the solute was transferred into pre-cooled Eppendorf tubes. Cells in Eppendorf tubes were placed on ice for 30 minutes and vortexed every 5 minutes for 15 seconds. The lysate was then centrifuged at 13,800 *x g* for 20 minutes at 4 °C. The resultant supernatant was carefully transferred into a pre-cooled Eppendorf tube and used directly for downstream experiments or stored at -80 °C.

2.3.2 Cytosol and Membrane Lysate Fractionation (MEM-PER Plus Membrane Protein Extraction Kit)

75 cm^2 cell culture flasks of A549, NFC and IPF derived HLMF cells were untreated or treated with 10 ng/ml of rhTGF-β1 for various time points (at around 70-80% confluency). To assess their composition at baseline and in stimulated conditions, cells were scraped off the flask using a sterile disposable cell scraper and transferred to an suitable sized Falcon tube. Subsequently the cells were centrifuged at 220 *x g* or 295 *x g* (A549 or NFC/ IPF myofibroblasts respectively) for 8 minutes. The pellet was re-suspended in 1 ml of wash buffer (provided with the kit). Cell density was determined using the haemocytometer chamber slides via the trypan blue assay (cell density of $5x10^6$ is required). If cell density is lower or higher, then the permeabilisation and solubilisation buffers need to be recalculated for the number of cells respectively. Additional 2 ml of wash buffer was added to the harvested cell suspension followed by centrifugation at 300 *x g* for 5 minutes. The supernatant was removed and discarded. The pellet was then re-suspended in 1.5 ml of the cell wash buffer and transferred into a 1.5 ml Eppendorf tube. This was then centrifuged again at 300 *x g* for 5 minutes. The supernatant was carefully removed and discarded, followed by the addition of 750 µl of the permeabilisation buffer supplemented with 1% (v/v) PIC, 1% (v/v) Na_3VO_4 and 1% (v/v) PMSF, to the pellet and vortexed to obtain a homogenous cell suspension. The cells were incubated with constant rotational motion (4 °C, 10 minutes). Cells were then centrifuged for 15 minutes at 16,000 *x g* at 4 °C. The supernatant was carefully removed and transferred to a new Eppendorf tube, this contains the **cytosolic fraction**. Subsequently 500 µl of

solubilisation buffer supplemented with 1% (v/v) PIC, 1% (v/v) Na_3VO_4, and 1% (v/v) PMSF, was added to the pellet, which was re-suspended with gentle pipetting until a semi-homogenous suspension was formed. The cells were incubated for a further 30 minutes with rotational motion as previously. Cells were again centrifuged at 16,000 *x g* at 4 °C for 15 minutes, and the supernatant containing the **membrane fraction** was carefully removed and transferred to a new Eppendorf tube for immediate experiments or storage at -80 °C until required.

2.3.3 Styrene Maleic Acid Lipid Particle (SMALP) Lysate preparation

A549 cells cultured in 75 cm^2 cell culture flasks were untreated or treated with 10 ng/ml of rhTGF-β1 or 2 ng/ml of LPS for 72 h and 24 h respectively (stimulation durations based upon previous optimisation studies within Bradding and Gooptu group). Cells were detached from the flask using 0.25% Trypsin-EDTA, incubating for ~5 minutes at 37 °C. The trypsin was neutralised by adding equal amount 1x DPBS. Cells were then centrifuged at 220 *x g* for 5 minutes; the supernatant was discarded, and the cell pellet was washed twice with ice-cold 1x DPBS. The sample was re-spun at 220 *x g* and the DPBS was then aspirated, and cells were lysed by pellet resuspension directly into either SMA2000P (kind gift of Dafforn group, University of Birmingham) or SMA30010P (Orbiscope, Netherlands). Per 50µl of pellet size, 500µl of SMA-lysis buffer was used. The SMA was made at a final (v/v) concentration of 2.5% (from 20% (v/v) stock) in a 50mM Tris, 150mM NaCl, 2mM EDTA and 10% (v/v) glycerol buffer, supplemented with 1% (v/v) PIC, 1% (v/v) PMSF and 1% (v/v) Na_3VO_4.

The lysate was injected in and out of a narrow-gauge needle (20 g) five times, to shear DNA, then incubated for 1 h at room temperature. The lysate was diluted 1:5 with extra SMA buffer (2 ml of 50mM Tris, 150mM NaCl, 2mM EDTA and 10% (v/v) glycerol) and spun at 100,000 *x g* for 45 minutes at 4° C (using thick-wall polycarbonate tubes placed in a TLA-100.3 Fixed angle-rotor, spun in a Beckman ultracentrifuge TLX). The supernatant (soluble fraction) was transferred into

a fresh sterile tube (through a Miracloth). The pellet (insoluble fraction) was discarded. Samples were used immediately for subsequent downstream experiments or stored at -80 °C until required.

2.4 Quantification of Sample Concentrations

2.4.1 Quantification of Whole Cell and Membrane-Cytosol Fractionated Lysate Concentrations using Bicinchoninic Acid (BCA) Assay

A commercially available protein assay kit (ThermoFisher, UK) was used to determine the concentration of protein in lysates. Nine bovine serum albumin (BSA) protein standards were prepared with concentrations ranging from 0 to 2 mg/ml using appropriate lysate buffer (using the manufacturers instruction). BCA Working Reagent was prepared by mixing 50 parts of BCA reagent A with 1 part of BCA reagent B (50:1, Reagent A:B) (the solution was vortexed until a homogenous lime green colouration appeared). 25 µl of each standard or unknown sample was pipetted into a flat-bottomed 96-well plate, subsequently 200 µl of the working reagent was added to each well (1:8, Sample: working reagent) and mixed thoroughly (by a plate shaker for 30 seconds). Lysates containing high concentrations of protein give an immediate colour change upon the addition of the working reagent. If a darker colour was observed than obtained with the highest BSA standard concentration (2 mg/ml) then the sample was diluted with the appropriate lysate buffer (typically two-fold dilution). The plate was then wrapped with foil (to avoid light) and placed in a 37 °C incubator for 30 minutes. Thereafter, the plate was cooled to room temperature and the absorbance readings were measured at 562 nm wavelength using the Enspire Perkin-Elmer Multi-plate reader spectrophometer. The absorbance reading of the nine standards was plotted on GraphPad Prism 9 on the Y-axis and the concentration that each absorbance represents was plotted on the X-axis to generate the standard curve. For experimentally-observed values falling within the linear range of the standard curve, the corresponding protein concentrations in assay samples were directly derived from the

absorbance using this curve. For diluted samples the lysate concentration was obtained by multiplying by the appropriate dilution factor.

2.4.2 Quantification of Purified Recombinant Protein Lysate Concentrations Using Bicinchoninic Acid (BCA) Assay

Purified recombinant galectin-3 was dialysed from a 1x PBS 2 mM EDTA buffer that contained β-mercaptoethanol (2.10.4.4), to one without this. β-mercaptoethanol interferes with the underlying Biuret reaction to give inaccurately high readings (by directly causing reduction of Cu (II) which otherwise occurs as a function of protein concentration). Standards were also prepared using this buffer. The purified protein concentration was then assessed as outlined in 2.4.1.

2.4.3 Quantification of SMALP Protein Lysate and DNA Concentrations Using Nanodrop Spectrophometer

Instead of a BCA assay, quantification of the SMALP protein lysate concentration was analysed on a Thermo Scientific™ Nanodrop™ 2000 spectrophometer, since the maleic acid affects the underlying Biuret reaction, thereby giving false readings (as explained in 2.4.2). The 'Protein A_{280}' option was chosen, and the instrument was blank-calibrated with 1.5µl of the buffer that the SMALP lysate is within (50mM Tris, 150mM NaCl, 2mM EDTA and 10% (v/v) glycerol). The concentration of the sample was assessed by taking 1.5µl of sample (duplicate) and the average of this value was used.

Quantification of plasmid concentrations were analysed on the Nanodrop spectrophometer. The 'Nucleic acid' option was chosen. The absorbance at A_{260} was measured. The instrument was blank-calibrated with 1.5µl of dH_2O (the plasmid was resuspended in this) and the sample was

assessed by taking 1.5µl (duplicate) and the average of this value was used. Purity of the plasmid was also defined by the A_{260}/A_{280} ratio.

2.5 Polyacrylamide Gel Electrophoresis (PAGE) and Protein Visualisation

2.5.1 PAGE Conditions

2.5.1.1 SDS-PAGE for Whole Cell, Membrane-Cytosol Fractionated and SMALP Lysates

Proteins were separated proportional to their molecular weight by using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). 4-15% (v/v) gradient Mini-PROTEAN® TGX™ Precast Protein gels (456-1084, Bio-Rad) were used and they were placed in a Bio-Rad gel running cassette unit, which was placed in a Bio-Rad Mini-PROTEAN Tetra cell tank; subsequently 1X Tris-Glycine-SDS (TGS) running buffer was added into the tank until the gel was covered (reaching the max line on the tank). Protein samples were then loaded at desired concentration(s). The concentrations of the lysates loaded were at 20 µg; this required appropriate dilution in dH_20 (or appropriate lysis buffer) such that there is a total volume of 30 µl per sample. Thereafter 10 µl of 4x Laemmli sample buffer (+/- 10% (v/v) β-Mercaptoethanol) was added (1x concentration was achieved) allowing sample visualisation for loading into the wells (and for the disruption of disulphide bonds if β-Mercaptoethanol was added). The samples were +/- heated at 95 °C for 5 minutes (dependant on the denaturation behaviour of the probed protein), and the 40 µl total sample was loaded onto the gel. MagicMark XP Western Protein Standard (LC5602, Invitrogen) was mixed with the protein ladder SeeBlue®Plus2-Prestained (10 µl MagicMark to 5 µl of ladder) this allowed the visualisation of the marker bands on the Western blotting image detector. Western standard and protein samples were then loaded into the gel and the gel was run at a constant voltage of 150 V using Bio-Rad Powerpack™ Basic for ~60-90 minutes at room temperature until the bromophenol blue marker reached the bottom line of the gel.

Post-electrophoresis, the gels were removed from the tank and cassette. The glass plates were detached and for protein visualisation the gel was transferred onto a Bio-Rad polyvinylidene fluoride (PVDF) membrane, which was then placed on a semi-dry transfer machine (Bio-Rad Trans-Blot® Turbo™ Transfer system). After the transfer, the membrane was placed in Tris-buffered saline 0.1% (v/v) Tween20 (TBST) to wash the membrane.

2.5.1.2 SDS-PAGE for Recombinant Purified Proteins

Recombinant purified proteins were separated proportional to their molecular weight by running an SDS-PAGE. 4-12% (v/v) gradient Bis-Tris Gel (NP0329BOX, Invitrogen by Thermo Fisher Scientific) were used and they were placed in a Invitrogen gel running cassette unit, which was placed in a Invitrogen Mini Gel tank; subsequently 1X NuPAGE™ MOPS SDS running buffer was added into the tank until the gel was covered (reaching the max line on the tank). 10µl of protein samples were then loaded, or if the samples were required to be at a specific desired concentration, they were diluted in appropriate buffer until the required concentration (20µg-100µg) was achieved, thereafter 2x Laemmli sample buffer (+/- 10% (v/v) β-Mercaptoethanol) (until 1x concentration was achieved, therefore 10µl of sample buffer was added to 10µl of sample) was added allowing the protein samples to be visualised for loading into each well (and for the disruption of disulphide bonds if β-Mercaptoethanol was added). The samples were not heated. The protein ladder SeeBlue®Plus2-Prestained (5 µl of ladder) was used and placed in the first well of the gel. The gel was run at a constant voltage of 125 V using Bio-Rad Powerpack™ Basic for ~35 minutes at room temperature until the bromophenol blue marker reached the bottom line of the gel.

Post-electrophoresis, the gel was carefully transferred to either Instant*Blue*™ ultrafast protein stain or Western blotting membrane for protein(s) detection.

2.5.1.3 Non-denaturing (native)-PAGE for Recombinant Purified Proteins

Non-denaturing gels were used and run in the same Bio-Rad Mini-PROTEAN tetra cell tank system as SDS gels but using a discontinuous buffer system. The Bio-Rad gel cassette was assembled, and a line was marked ~1 cm below where the comb sits, dictating the level of the amount of resolving gel to be added into the cassette. The resolving gel was prepared at 10% (w/v) (Table 2.2), and then poured until reaching the marker point. Immediately propanol was added on top of the resolving gel (until reaching the top of the cassette). This excludes air which inhibits polymerisation of adjacent acrylamide and keeps the upper border of the gel as straight as possible. Once the resolving gel was polymerised, the propanol was tilted off, and the top layer of the resolving gel was rinsed with dH_2O. Thereafter, 4% (w/v) stacking gel was prepared (Table 2.2) and poured on top of the resolving gel, followed by insertion of the well-forming comb. Once the stacking gel polymerised the comb was removed, and the wells were washed out with dH_2O. The gel was then placed into the running apparatus according to the manufacturer's instructions. Cathodic running buffer (Table 2.1) was placed in the top reservoir (in between the gel cassettes) and the anodic running buffer in the lower reservoir. Varied concentration of samples were prepared, and 2x native sample buffer (table 2.1) was added in a 1:1 dilution to the sample. The sample was then loaded into the gel. When studying recombinant galectin-3, wild-type α1-antitrypsin was useful as a control sample to confirm how the native-PAGE, as opposed to galectin-3 itself, had run. Whilst different proteins run with varying degrees of clarity on native-PAGE, α1-antitrypsin reliably runs very cleanly.

The iso-electric point of galectin-3 is close to the pH of the gels and running buffers. Under normal native-PAGE running conditions it may therefore exist in a range of ionisation states and not migrate well anodally. To address this, native PAGE optimisation was attempted in a range of alternative conditions:

i) Electrode terminals and discontinuous running buffers were reversed (either both changes were done singly, or together) so that the gel is optimised for 'cationic' protein detection. The gels were run for 1 h to 8 h, the dye front could not be used as a reference as it migrates in the opposite direction (upward and out of the gel) under these conditions.

ii) The pH of Tris-based buffers fluctuates with temperature (New England BioLabs, 2021). As running native-PAGE gels tended to heat the buffers, gel tanks were placed in an icebox in the cold-room at 4 °C to control the temperature of the system, giving control over the working pH range of the Tris-buffer. The buffers were prepared and stored at 4 °C (Chapter 5). The gels running variations above were repeated in these cooling conditions. Post-electrophoresis, the gels were removed from the tank and cassette. The glass plates were detached, and the gel was transferred to either Instant*Blue*™ ultrafast protein stain or Western-blotting membrane for the detection of the protein(s).

Table 2.1, Native PAGE Stock Solutions

Acrylamide	30% (w/v) acrylamide: 0.8% (w/v) bis acrylamide
4x Separation gel buffer	2 M Tris-HCl pH 8.8
4x Stacking gel buffer	0.5 M Tris-HCl pH 6.8
Cathode running buffer	53 mM Tris-HCl, 68 mM glycine pH 8.2 *
Anode running buffer	100 mM Tris-HCl pH 7.8 *
Sample loading buffer	10 ml stacking gel buffer, 2 ml glycerol, 0.04% (w/v) bromophenol blue in 20 ml dH_2O
10% APS	10% ammonium persulphate in dH_2O
TEMED	NNN'N'-tetramethyl-ethylenediamine

*Standard pH value given here. Different pH conditions assessed during optimisation for running galectin-3, further elaborated in Chapter 5.

Table 2.2, Native PAGE Gel Mixes

	10% (w/v) Separating gel	**4% (w/v) Stacking gel**
Buffer	2.5 ml	1 ml
Acrylamide	3.33 ml	0.4 ml
Water	4.17 ml	2.6 ml
TEMED	10 µl	
10% APS	50 µl	

2.5.1.4 Styrene Maleic Acid non-denaturing (SMA native)-PAGE for SMALP Extracts

Proteins were separated by their molecular weight by using SMA native-PAGE (SMA lysates have a net negative charge). 4-15% (v/v) gradient Mini-PROTEAN® TGX™ Precast Protein gels were used (2.5.1.1). Subsequently Tris (25mM) - glycine (192mM) running buffer (without SDS) was prepared (no pH adjustment) and added into the tank until the gel was covered (reaching the max line on the tank). 2x native sample buffer was added in a 1:1 dilution to sample. The sample was then loaded into the gel. The first lane was loaded with HiMark™ Pre-stained Protein Standard (which runs well in non-denaturing gel running conditions). The second lane was loaded with MagicMark XP Western Protein Standard (LC5602, Invitrogen) mixed with the protein ladder HiMark™ Pre-stained Protein Standard (10 µl MagicMark to 10 µl of ladder). This allowed the visualisation of the marker bands on the Western blotting image detector. The gel was run at 150 V for 60 minutes at room temperature, and transferred to a PVDF Western blotting membrane as described in 2.5.1.1

2.5.2 PAGE-Based Quantitation of Specific Protein Species

2.5.2.1 Detection of Proteins in Whole Cell Lysate (Including Membrane/Cytosol Fractions)

When probing samples of proteins from cell lysates, Western blot membranes were blocked for 1 h on a roller in 5% (w/v) milk solution in Tris-buffered-saline-Tween20 (TBST) at room temperature. The membranes were then placed into a Falcon tube with the primary antibody diluted in 5% (w/v) milk (in TBST) and left overnight on a rocker at 4 °C. The membranes were then washed for 10 minutes with TBST three times. Post washing, the membranes were incubated for 1 h with the secondary antibody diluted in 5% (w/v) milk (in TBST) and washed for 5 minutes with TBST three times. Membranes were then developed using the Enhanced ChemiLuminescence (ECL) (ThermoFisher, UK) Western blot detection system (equal volumes of solutions A and B were added) for 2 minutes. Bands were detected using the Image Quant LAS

4000. The antibodies and their concentrations used in Western blot analysis are stated in Table 2.13.

2.5.2.2 Detection of the SMALP Lysed Proteins

Gels were handled and Western blotting conducted largely as described in 2.5.2.1. Though to capture the molecular weight ladder Hi-Mark Pre-stained Protein Standard (ThermoFisher, UK) an extra imaging step (digital image capture setting) was done on the Western-blotting developer Image Quant LAS 4000.

2.5.2.3 Detection of the Recombinant Purified Protein Species from SDS or Non-denaturing PAGE

If the gel was transferred into Instant*Blue*™ ultrafast protein stain, it was left on a rocker overnight, and de-stained using dH_2O until the background was clear. Images of the gel were captured on an appropriate scanner and saved as TIF images.

For protein visualisation through Western blotting, gels were immersed in 0.04% (w/v) SDS Bjerrum transfer buffer and incubated for 5 minutes. 0.45 µm porous PVDF membrane weas used. The membrane was initially activated by immersing in methanol for 5 minutes followed by a thorough rinse with 0.04% (w/v) SDS Bjerrum transfer buffer. Thereafter, Whatman paper thickness '2' (immersed in Bjerrum transfer buffer) was used on either side of the membrane sandwich and then placed on a semi-dry transfer machine (Bio-Rad Trans-Blot® Turbo™ Transfer system). After the transfer, the membrane was washed for 20 minutes with 1x Bjerrum transfer buffer containing 20% (v/v) methanol (without SDS) to remove SDS. The membrane was placed in Tris-buffered saline 0.1% (v/v) Tween20 (TBST) to wash the membrane. Membranes were

blocked for 1 h on a roller in 5% (w/v) milk solution in TBST at room temperature. The membranes were then placed into a Falcon tube with the primary antibody diluted in 5% (w/v) milk (in TBST) and left overnight on a rocker at 4 °C. The membranes were then washed for 10 minutes with TBST three times. After the washing, the membranes were incubated for 1 h with the secondary antibody diluted in 5% (w/v) milk (in TBST) and washed for 5 minutes three times. Membranes were then developed using the Pierce Enhanced ChemiLuminescence (ECL) Western blot detection system (equal volume of solution A and B were added) for 2 minutes. Bands were detected using the Image Quant LAS 4000. The antibodies used in Western blot analysis are shown in Table 2.13.

2.5.2.4 Western Blot Densitometric Analysis

Image J (Fiji) was used to quantify the relative intensity of the bands. Once the image was loaded, the initial step was to reduce background noise with a rolling ball radius between the value of 30-50. Rectangular boxes were drawn around each band to be analysed. After bands in each lane were highlighted, the gel was analysed via the 'plot lane' function, which generates a histogram for each lane. A straight line was pencilled at the bottom of each peak enclosing the area under the peak; the area under the enclosed peak was selected using the 'magic wand' tool, subsequently giving the value of the area (arbitrary units). This was also done for quantification of house-keeping protein expression.

2.5.3 qRT-PCR to Assess mRNA Transcripts of Galectin-3 and CD98

qRT-PCR was conducted in collaboration with Dr Katy Roach (Bradding group, Respiratory department, University of Leicester) to assess the mRNA transcripts of galectin 3 and CD98 in myofibroblasts and human lung tissue. RNA extraction using Qiagen RNeasy mini prep kits was

performed for myofibroblasts, and Qiagen RNeasy Fibrosis tissue kit was used for human tissue (as described previously (Roach *et al.* 2013)).

To assess CD98 mRNA levels, CD98 cDNA was first synthesised based on a standard quantity of mRNA using the SuperScript™ VILO cDNA synthesis kit. Taqman® gene expression assays were then used according to the manufacturer's instructions with primer/probes for SLC3A2-FAM(Hs00374243_m1). The house-keeping gene used for this was 18S-VIC.

To assess galectin-3 mRNA levels, RT-PCR was performed using Quantitect primer assay from Qiagen (Hs_LGALS3_1_SG). The house-keeping gene used for this was ACTB (Hs_ACTB_1_SG), with the Agilent Brilliant III SYBR green qRT-PCR mastermix (600886).

Relative expression was calculated using as -2^-(delta CT), where delta CT is the difference between the CT value for the gene of interest and the house-keeping gene. All qRT-PCRs were performed with the QuantStudio 5 Real-Time PCR system.

2.6 Immunoprecipitation

Immunoprecipitation was done on whole cell, membrane-cytosol fractionated and SMALP lysates prepared as mentioned in section 2.3.

2.6.1 Preparation of Dynabeads™ Protein G

Dynabeads™ Protein G superparamagnetic beads (#1003D, ThermoFisher, UK) were re-suspended by 30 seconds of high-speed vortexing. 30 µl of the suspension were drawn up with a cut 200 µl pipette tip and then washed three times with 1x DPBS. DPBS was removed each time by placing the Eppendorf on a magnetic rack and allowing the beads to migrate to one side of the tube; the beads were then re-suspended in 30 µl of DPBS.

2.6.2 Pre-clearing Lysate and Antibody Binding

Prior to immunoprecipitation incubation of a pre-clearing step was undertaken to deplete any proteins that could non-specifically bind to the beads. Cell lysates containing 500-1500 µg of protein were incubated with 10 µl of washed magnetic beads for 60 minutes rotating at 4 °C.

2.6.3 Antibody-Lysate-Bead Complex (Protein Elution in a Denaturing Condition)

After the cell lysate pre-clearing step, the lysates were placed on a magnetic rack, and the supernatant was removed from the dynabeads. Subsequently anti-Galectin-3 or anti-CD98hc antibody was added at the appropriate dilution to the lysate (to give a final antibody concentration of 2 µg/ml) and left to incubate for 4 h at 4° C on a rocker. The remaining 20 µl of washed dynabeads were transferred to the antibody-lysate mixture for overnight incubation at 4° C on a rocker, to allow the beads to bind to the antigen-antibody complex. The antibody-lysate-bead complex were then washed three times with DPBS buffer on a magnetic rack. 30 µl of 4x Laemmli Sample Buffer was added to the beads, the antibody-lysate-bead complex was agitated for 20-30 seconds every 5 minutes for 30 minutes at room temperature. The sample was placed

on the magnetic rack and the sample was then transferred into a new Eppendorf. The samples were then run on a 4-15% (v/v) SDS-PAGE followed by Western blot analysis.

2.6.3.1 Antibody-Lysate-Bead Complex (Protein Elution in a Non-Denaturing (Native) Condition)

Up until the end of the overnight antibody-lysate-bead complex has been washed, the steps between denaturing and non-denaturing immunoprecipitation were the same. Instead of adding 30 µl of SDS containing 4x Laemmli sample buffer, to pull-down co-immunoprecipitated sample of interest, here, 40 µl of 3M $MgCl_2$ was added to the beads to break the bead-antibody interactions. The antibody-lysate-bead complex was agitated for 20-30 seconds every 5 minutes for 2.5 h at room temperature. The tube was placed on a magnetic rack and the sample was then transferred into a new Eppendorf. The samples were then used for subsequent downstream experiments including initial analysis using SDS and SMA-native PAGE or stored at -80° C.

2.7 Immunostaining Microscopy and Staining Quantification

2.7.1.1 Direct Staining Immunofluorescence Microscopy

For the characterisation of NFC and IPF-derived human lung myofibroblasts cells cultured from different donors, the cells were cultured on 8 well glass chamber slides. Once the cells were confluent, culture media was aspirated, and the cells were washed with 1x DPBS. Cells were fixed by adding 200 µl of ice-cold 4% (w/v) paraformaldehyde (PFA) placing the chamber slides for 20 minutes on ice (or in a 4° C fridge). The fixative agent was removed, and the slides were air-dried. A blocking solution consisting of 3% BSA (w/v) diluted in 1x DPBS was added to each well rocking for 30 minutes, eliminating non-specific binding. The blocking solution was then removed and

primary antibodies (diluted in 1% BSA (w/v) DPBS) were added for 90 minutes rocking at room temperature in the dark (Table 2.13). The cells were washed three times with 0.05% (v/v) Tween20 diluted in DPBS (wash buffer). They were then counterstained with 4',6-diamidino-2-phenylindole (DAPI) (32670, SigmaAldrich) with a 2 minute incubation. Finally, cells were washed six times with DPBS, and rectangular coverslip were mounted with drops of Fluoroshield mounting media (F6182, Sigma-Aldrich) added on each well. Images were captured on a fluorescence microscope (ZEN). Matched exposures were used for isotype controls.

2.7.1.2 Indirect staining immunofluorescence microscopy

For antibody specificity and optimisation in all cell types, and the characterisation of NFC and IPF HLMF phenotypes, cells were singly immunostained with or without prior rhTGF-β1 stimulation. The same procedure was used as described in 2.7.1.1, up to the point of adding the primary antibody. After 90 minutes of incubation with the primary antibody, the cells were washed three times with wash buffer and then the secondary antibodies (diluted in 1% BSA (w/v) DPBS) indirectly labelled with Alexa-Fluor 594 (red) or Alexa-Fluor 488 (green) were applied for a further 90 minutes rocking at room temperature (in the dark). The cells were washed again three times with wash buffer, counterstained with DAPI and coverslips were mounted. Images were captured on a fluorescent microscope (ZEN). Matched exposures were used for isotype controls.

2.7.1.3 Co-immunostaining of CD98hc and β1-integrin with Galectin-3 in A549, Non-Fibrotic-Control and IPF Human Lung Myofibroblast Cells

Immunofluorescence studies of A549 cells and *ex vivo* human lung myofibroblasts co-stained for galectin-3 with β1 integrin or with CD98hc were performed to examine the co-localisation of these protein pairs in cells with and without TGF-β1 stimulation. Cells were co-incubated with

pairs of relevant primary antibodies generated in different animal species, and then probed with appropriate secondary antibodies conjugated to different fluorescent reporters (Table 2.14). Images were captured on a confocal laser scanning microscope (Leica TCS SP4, Leica UK Ltd). The intensity of the laser power used was set as follows: 405 Diode UV at 25% intensity in sequential scan-1, Argon Visible at 30% power and 17% intensity in sequential scan-2, and DPSS 561 visible at 30% intensity in sequential scan-3. The microscope was set to bi-directional and the phase correction was set to -29.80; Z stack layers were taken at 0.3 µm increments. Matched exposures were used for the isotype controls.

2.7.2.1 Single Channel Immunostaining Quantification

Cell F Imaging software (Olympus UK Ltd) and Image J (Fiji) were used to quantify immunofluorescence signals. Once the specified channel of an image (red or green) was loaded onto Cell F, the appropriate colour was converted to grey-scale by using the correct separation colour for the channel (red or green). Subsequently the grey value mean tab was selected to define the region of interest (ROI), and the Average Grey Value Mean for the channel was obtained. Fiji was used to quantify the amount of nuclei count; this then allowed the single channel grey value mean to be represented as 'Grey value mean per nucleus'. The values are used in statistical analysis of the acquired data to compare individual protein expression of non-stimulated to stimulated conditions.

2.7.2.2 Co-Localisation Immunostaining Quantification

The original .LIF file stored from the Leica TCS Confocal microscope were loaded onto Image J. Once the maximum distributed stained Z stack was chosen for both the green and red channel, the channels were split, and only the green and red channel were kept open (the DAPI and brightfield channels are closed). The Plug-in JACoP (Just Another Co-Localisation Plugin) was

opened, and the options 'Pearson's Correlation Coefficient' and 'Overlap Coefficient' (Mander's Overlap Coefficient) are selected. Once thresholding had been applied, the image was then analysed and the Pearson's and Mander's values were obtained. The theory behind the use of each coefficient is as follow: Pearson's Correlation Coefficient is not sensitive to differences in mean signal intensities or range, or a zero offset between the two stained components. The result is +1 for perfect correlation, 0 for no correlation and -1 for anti-correlation, therefore the Pearson's coefficient takes into account the correlation of each fluorophore, and not the degree to which the signals co-occur. Mander's overlap coefficient is proportional to the amount of fluorescence of the colocalising pixels or voxels in each stained channel, value ranges from 0 to 1, expressing fraction of intensity in a channel that is located in pixel where there is above zero intensity in the other channel. The Mander's coefficient determines the overlap of two images while taking into account pixel intensity, termed as co-occurrence, it therefore accounts for total amount of fluorophores that overlap with each other. (Aaron *et al*. 2018). These values are used in downstream statistical analysis to compare co-localisation behaviour of different stimuli conditions.

2.7.2.3 Co-immunostaining of Lung Tissue Using the Duolink® Proximity Ligation Assay Kit

These experiments were conducted with the assistance of Dr P. Stylianou (Gooptu group, University of Leicester). These experiments were performed to examine the expression and cellular co-localisation levels of putative gal-3-fibrosome proteins in NFC lung tissue explant model subjected to 10 ng/ml TGF-β1 stimulation. The tissue slices were embedded into paraffin.

Before proceeding with the experiment, reagents provided within the kit as well as non-provided reagents were prepared in advance: wash buffers A and B (provided with the kit) were prepared by dissolving the contents of one pouch into high purity $_{d}H_2O$ to equate a final volume of 1 l (1x concentration). These buffers were stored at 4° C for long term storage.

On the day of starting the staining procedure the following buffers were prepared: wash buffer B was diluted 1:100 with $_dH_2O$ giving a 0.01x final concentration (kept at 4 ℃), blocking reagents were prepared using 10% (v/v) FBS and 0.5% (v/v) Triton X-100 in 1x PBS (kept at 4 ℃) and antibody diluent were prepared using 1% BSA (w/v) and 2% FBS (v/v) in 1x PBS (kept at 4 ℃). Several Duolink® PLA reagents are supplied as concentrated stocks and are diluted just immediately prior to usage; these are described within the protocol. Diluted PLA reagents are not stored for future use. The protocol takes 2 days to complete:

Day 1

100 µl of blocking solution was added to each lung section placed on a glass slide. Slides were incubated at 37 °C under an atmosphere of 95% air/5% CO_2 for 2 h. After this incubation period the blocking solution was gently removed, and primary antibodies (diluted in antibody diluent) were added to the lung section for overnight staining at 4 ℃ (cold room) (Table 2.13).

Day 2

PLUS and MINUS probes (provided within the kit) were prepped by diluting them 1:5 with the antibody diluent buffer. For a 40 µl total reaction volume (added to each lung section/slide), this was achieved by adding 8 µl each, of PLA probe PLUS and MINUS stock, to 24 µl of antibody diluent. The overnight stained primary antibody solution was gently removed, and the slides were washed 2x for 5 minutes in 1x wash buffer A at room temperature. The PLA probe solution was then added to the sections and incubated at 37 °C under an atmosphere of 95% air/5% CO_2 for 2 h. During this incubation period ligation buffer was prepared, however, ligase was added to the ligation buffer immediately before adding the solution to the sample. For a 40 µl reaction, 8 µl of the 5x ligation buffer was added to 32 µl of antibody diluent. The PLA probes were gently removed, and the slides were washed 2x for 5 minutes in 1x wash buffer A at room temperature. During the last wash, ligase was added to the 1x ligation buffer prepared earlier (1 µl ligase to 39

µl of 1x ligation buffer). The ligation solution was added to the sections and incubated at 37 °C under an atmosphere of 95% air/5% CO_2 for 30 minutes. Subsequent steps hereon were conducted in the dark. During this incubation period amplification buffer was prepared, however the polymerase was added to the amplification buffer immediately before adding the solution to the sample. For a 40 µl reaction, 8 µl of the 5x amplification buffer was added to 32 µl of antibody diluent. The ligation buffer was gently removed from the slides and then washed 2x for 5 minutes with 1x Wash buffer A at room temperature. Polymerase was then added to the 1x amplification buffer prepared earlier (0.5 µl polymerase to 39.5 µl of 1x amplification buffer, 1:80 dilution) and the amplification buffer was added to the slides. Slides were then incubated at 37 °C under an atmosphere of 95% air/5% CO_2 for 2 h. The amplification solution was removed followed by 2x 10 minute wash with 0.01x wash buffer B at room temperature. Excess wash buffer was removed, and lung sections were stained with DAPI for 2 minutes for nuclei staining. DAPI was removed with 3x 5 minute wash with 1x PBS. The slides were mounted using Fluoroshield mounting media. Images were captured using Leica TCS Confocal microscope.

2.7.3 Immunohistochemistry on NFC and IPF Human Lung Tissue

These experiments were conducted with the assistance of Dr P. Stylianou (Gooptu group, University of Leicester). Immunohistochemistry involves the staining and detection of proteins in tissue and was performed to examine the expression and cellular co-localisation levels of putative gal-3-fibrosome proteins in serial sections of both NFC and IPF-derived lung tissue. Tissues were processed into glycol methacrylate resin followed by immunostaining using a commercially available kit (Dako EnVision FLEX + staining technique).

2.7.3.1 Glycol Methacrylate (GMA) Resin Embedding

Tissue was initially fixed in ice-cold acetone containing 20 mM iodoacetamide and 2 mM phenylmethylsulfonyl fluoride overnight at -20 °C. The fixative agent was replaced with fresh acetone for 15 minutes at room temperature, and then transferred into methyl benzoate for 15 minutes at room temperature. Tissue was immersed for 6 h in GMA monomer solution consisting of 5% methyl benzoate. During this 6 h incubation time, the GMA solution was changed and replenished three times. The tissue was then placed in polythene capsules with lids containing GMA embedding resin (70 mg Benzoyl peroxide dissolved in 10 ml GMA solution A and 250 µl GMA solution B) at 4 °C for 16 h. The biopsies were then stored at -20°C in an airtight container.

2.7.3.2 Dako EnVision FLEX + Staining Technique

GMA biopsy sections of 2 µm were serially sliced using the LEICA RM2255 microtome and then placed into a water bath consisting of 0.2% (v/v) ammonia for 60 seconds. The slides were left to either dry overnight in a covered slide tray or air dry at room temperature for 2-4 h. A circle was then drawn around the sections using a IHC-pen (ImmEdge). The sections were then incubated in Peroxidase Blocking Reagent for 10 minutes (adding enough to cover the section). Slides were then subjected to thorough washing using EnVision FLEX wash buffer (5 minutes), repeated three times. Primary antibody was diluted in EnVision FLEX antibody Diluent, and then added to the sections for 1 h at room temperature (50 µl per section). The primary antibody was washed away with the EnVision FLEX wash buffer (5 minutes), repeated three times. The secondary antibody (diluted in EnVision FLEX antibody Diluent) EnVision FLEX/HRP was added to each section (50 µl) for a total of 30 minutes at room temperature. The secondary antibody FLEX-HRP was then washed with the EnVision FLEX wash buffer, and then the slide was developed using DAP (3,3'-diaminobenzidine) chromogen (added drop wise onto the section with 10 minute room temperature incubation). The sections were then counterstained with Haematoxylin (immersed

for ~ 1 minute), the slides were then washed with running tap water for 5 minutes. Thereafter, the slides were left to dry (applying supermount on the section) in the dark at room temperature. A cover slip was placed on the section using aqueous mounting medium (DPX), and then proceeded to downstream imaging using the Zeiss Fluorescent microscope (Axiolab 5, Zeiss). The primary and secondary antibodies used with this technique are stated in table (Tables, 2.13 and 2.14.)

2.8 Wound Healing Assay in A549 Cells Subjected to CD98hc Silencing

2.8.1 A549 Cell Culture for CD98 Knockdown

Confluent A549 cells in sterile culture 75 cm^2 flask was serum-starved by treating in low serum media for 24 h. The cells were then passaged (as mentioned in 2.1) and seeded into 6-well plates at a density of $1x10^5$ cells per well in 2 ml of serum-only growth media (DMEM+4.5g/L D-Glucose, L-Glutamine) supplemented with **10% (v/v) FBS**, without antibiotics. The cells were incubated overnight at 37 °C (95% air/5% CO_2) until the cells reached ~50-60% confluency. Cells were not subjected to transfection until they had reached a minimum of ~50% confluency.

2.8.2 Transfection and siRNA Knockdown Optimisation

Cells were seeded (from step 2.8.1) to carry experiments for un-transfected control, mock-control (the transfection reagent only), siRNA target and siRNA-non-targeting-control. In a sterile tissue culture hood 4 µl of Lipofectamine 2000 (transfection reagent) was diluted into 196 µl of serum-free media (DMEM+4.5g/L D-Glucose, L-Glutamine) which excludes antibiotics and FBS. The solution was mixed thoroughly by gentle pipetting followed by a short incubation of 5 minutes at room temperature. 10 µl of the siGENOME Small-Interfering Ribonucleic acid (siRNA)

targeting CD98hc or HUMAN SLC3A2 (#D-003542-01-0010) and the non-targeting siRNA siGENOME-Control (siCON) (#D-001210-03-05) were diluted in 190 µl volumes of serum-free media. The solutions were mixed carefully by gentle pipetting followed by a short incubation of 5 minutes at RT. Subsequently 200 µl of transfection reagent (Lipofectamine 2000) was added to the 200 µl diluted siRNA or siCON respectively. For the Mock Control, the 200 µl of diluted transfection reagent was added to 200 µl serum free media (keeping the conditions consistent). This was mixed by gentle pipetting, since harsh pipetting can result in low cell uptake. The reagents were incubated for 20 minutes at room temperature. During this incubation period, the serum-only cell media in the 6-well plate was aspirated from the wells and washed using 1 ml of serum-free medium followed by the addition of 1 ml of serum-free media in each well; the cells were incubated with this media for 20 minutes at 37 °C (95% air/5% CO_2).

600 µl of serum-free media was next added to either the 'Lipofectamine + siRNA' and 'Lipofectamine + siCON' complex or the Lipofectamine Mock Control, respectively. The solution was mixed thoroughly followed by pipetting into the appropriate well for the desired condition(s). The non-transfected control simply had 2 ml of serum-free media in the well. The cells were left to incubate for 6 h at 37 °C (95% air/5% CO_2). At the end of this period, the reagents were aspirated from the wells and replenished with fresh serum-only media. The cells were incubated at 37 °C (95% air/5% $C0_2$) for various time points (2-7 days). Once each time point was reached, the cells were subjected to RIPA lysis (2.3.1). Post-lysate preparation, BCA assay was conducted for lysate concentration quantification (2.4.1). 20 µg of each lysate was run on 4-15% (v/v) gradient SDS-PAGE, followed by Western blotting analysis using the monoclonal CD98 antibody (CD98 E5 sc-376815) confirming CD98hc knock-down efficiency.

2.8.3 Wound Healing Response of CD98 Depleted A549 Cells in the Presence of Pro-Fibrotic Stimulation

Confluent A549 cells were serum-starved by incubation in ITS for 24 h. Grids were drawn on the outside base of a sterile 6 well plate using a fine permanent non-black marker in a sterile hood. A549 cells were passaged and seeded at a density of $1x10^5$ cells per well and allowed to adhere and reach 50-60% confluency by incubating at 37 °C (95% air/5% CO_2). Cells were transfected with targeting and non-targeting control siRNA constructs (2.8.2). The cells were then allowed to reach a confluency of ~100% (achieved ~48 h post-transfection). Four artificial wounds were scratched in each monolayer using a sterile 200 µl pipette tip in the predetermined grid pattern creating a linear cell free area. The wells were washed twice with serum-free media removing cell debris. The cells were subjected to stimulation with +/- 10 ng/ml of rhTGF-β1, 2 ng/ml lipopolysaccharide (LPS) and 1x DPBS. A positive control (serum-only media) was used in non-transfected cells. Images were captured on the EVOs® *xl core* AMEX-1200 microscope. The objective of the microscope was set to 10x and the phase turret of the microscope was set at PH 4/10 (for phase observations at 4x or 10x); the brightness was set at 50%. Time 0 h images (T_0) were taken and wound location on the grid was recorded. Four images were captured in different wounded areas per condition. The 6-well plates were placed back into the 37 °C (95% air/5% CO_2) incubator. After 24 h, images of the same wounded areas as T_0 were captured. Image analysis was undertaken for the area of the scratch wound using Photoshop CC 2015.5 (Adobe), using the histogram options for analysis across all pixels of the entire image. To work out the area of the wound, the new layer tab option was selected, and the pencil tool was used to draw around the wound. The paint bucket tool was then used to quantify the area of the wound in pixels for comparison over time. The values obtained were used for further statistical analysis.

2.9 Flow Cytometry

2.9.1 Testing the Cell Surface Expression and Distribution of CD98hc in A549 and HLMF Cells

To confirm the certainty of cell surface restricted expression and distribution of CD98hc, A549 and HLMF cells were analysed by flow cytometry. Both cell types were either non-stimulated or subjected to 10 ng/ml TGF-β1 stimulation. Cells were cultured to confluence in a 75 cm^2 flask (as described in 2.1). To confirm antibody specificity of CD98hc via this technique, A549 cells were also subjected to CD98hc silencing and were cultured as aforementioned (2.8.2).

Once ready to harvest, cells were washed twice with HBSS removing growth media. The cells were then gently detached in order to protect surface proteins, using 5 ml of TrypLE Express (-) phenol red (ThermoFisher, UK); cells were incubated at 37 °C for 10 minutes for A549, and 15 minutes for HLMF cells. Post-detachment the TrypLE Express reagent was neutralised by adding equal amount of HBSS and centrifuged at 295 *x g* for 10 minutes. The supernatant was discarded, and the pellet was re-suspended in 1 ml of HBSS. Cell density and viability of a sample using trypan blue quantitation were assessed in a cytometer chamber. $5x10^4$ cells were split from the cell re-suspension into U-bend-bottom Fluorescence-activated cell sorting (FACS) tubes (ThermoFisher Scientific, UK). An additional 500 µl of HBSS was added into the tube to weigh down the cells into the U bend of the FACS tube. The cells were spun at 295 *x g* for 10 minutes at 4°C. The supernatant was discarded very gently.

Cells were then treated in one of two ways. If the membrane expression and distribution of cells was assessed, cells were not subjected to fixation and permeabilisation, and the appropriate primary antibody was added (diluted in 1% BSA in 1x PBS).

Alternatively, if the cytosolic distribution and expression of protein was assessed (as a control to compare membrane restricted distributed proteins against) the cells were fixed and permeabilised using 300 µl of 4% (w/v) PFA with 0.1% (w/v) saponin for 20 minutes on ice. Cells were spun at 295 *x g* for 10 minutes at 4°C and the supernatant was gently discarded. Thereafter, primary antibody was added (diluted in 1% BSA in 1 x PBS with 0.1% (w/v) saponin).

CD98hc primary antibody was added at a concentration of 4 µg/ml (table 2.13) in the appropriate diluent (1% BSA in 1 x PBS +/- 0.1% (w/v) saponin respectively) at 100 µl per FACS tube. The tubes were vortexed gently for 3 seconds and then left to incubate for 1 h on ice (4 °C). With the appropriate diluent, cells were also stained with the appropriate isotype control. Additionally, the cells were also subjected to no primary antibody staining control; this is considered as 'Fluorescence minus one (FMO)' control. FMO controls contains every component and reagent as per the sample being measured, except the specific primary antibody to stain the protein itself. This helps identify and gate cells correctly, therefore identifying positive signals from negative population. Post-incubation the tubes were topped up with 400 µl of appropriate diluent to sediment the cells to the bottom of the U-bend of the tube and centrifuged at 295 *x g* for 10 minutes at 4°C. The supernatant was gently discarded. Secondary fluorescent antibody was added to the tubes treated with primary antibody and FMO controls, specifically 100 µl of F(ab')2 fragment of goat anti-mouse conjugated to APC was used at a working-concentration of 0.005 mg/ml (1:200 dilution) in the appropriate diluent. The cells were lightly vortexed for 3 seconds, and the FACS tubes were left to incubate for 30 minutes on ice (4 °C) covered in foil (dark). The FACS tubes were topped up with 400 µl of appropriate diluent and then centrifuged at 295 *x g* for 10 minutes at 4 °C. To stain for dead cells, 100 µl of SYTOX™ Orange Dead cell stain was used in a 1:200 dilution using the appropriate diluent. The FACS tubes were gently vortexed and incubated on ice (4°C) for 5 minutes. The tubes were topped up with 400 µl 1x PBS, giving total volume of 500 µl. CD98hc expression was captured by flow cytometry (Attune NxT acoustic focusing cytometer, software: Attune NxT Software v3.1.2). Alongside CD98hc staining and the two controls, three further control conditions were assessed per n number for accurate interpretation of signal: primary antibody alone, unstained control (no antibodies) and forced

dead cells with SYTOX™ orange; cells were killed for the final sample by placing them in FACS tubes on a heating block at 100°C for 5 minutes.

2.9.2 Analysis of Flow Cytometry Data

Data was analysed using the software FlowJo (v10.6.2). The data was quantified as a fold difference in geometric mean fluorescent intensity (GMFI) proportional to the isotype or FMO (unstained secondary only) control.

2.10 Molecular Biology of Galectin-3

2.10.1 Preparation of Recombinant Galectin-3 Glycerol Stocks

Full-length galectin-3 construct was kindly donated by Professor Hakon Leffler (Lund University, Sweden). The expression plasmid for galectin-3 is vector pET3C (Ampicillin resistant) in *E.coli* BL21* and arrived in transport culture tubes (stab cultures). All subsequent downstream procedure was carried out using aseptic conditions (sterile bench-space, equipment and reagents) achieved by wiping with industrial methylated spirits (IMS) or carrying procedures under flame to name examples. Luria-Bertani (LB) agar was prepped (Table 2.3) and the components were both dissolved and sterilised by autoclaving. Once touch cool, Ampicillin was added to the LB agar at a concentration of 50 µg/ml, and then poured into a petri dish to set, keeping half of the lid off under flaming conditions to get rid of condensation and maintain sterility. Once set, a fine sterile loop was dipped into the stab culture, and then streaked onto the LB agar plates. The plates were incubated overnight at 37 °C. Post-overnight incubation, 50 ml of LB broth was prepped (Table 2.3) in 250 ml conical flasks, dissolved and sterilised by autoclaving. Once touch cool, ampicillin was added to the LB broth at a final concentration of 50

µg/ml. A single medium to large sized colony was picked from the streaked agar plates, and the LB broth was inoculated and allowed to grow at 37 °C shaking at 250 rpm. The optical density (OD) was measured at 600 nm wavelength (OD_{600}) every 40-60 minutes using a basic spectrophometer (blanked against LB broth) until exponential growth phase was reached (OD_{600} 0.6-0.8). Once OD had reached, 70% of the 250 ml culture was mixed with 30% (v/v) of 50% (v/v) glycerol (100% glycerol diluted with $_dh_20$ in a 1:1 ratio) and then aliquoted into 1 ml cryovials. Vials were plunge frozen in liquid nitrogen and stored at -80 °C.

Table 2.3, Luria Bertani media and agar mix per litre

	LB Broth	**LB Agar**
Agar (1.5% w/v)		15 g
Sodium Chloride (1% w/v)	10 g	10 g
Tryptone (1% w/v)	10 g	10 g
Yeast Extract (0.5 w/v)	5 g	5 g
Water	Make up to 1 Litre Total Volume	
Ampicillin	Added at a final concentration of 50µg/ml after autoclaving	

2.10.2 Recombinant Galectin-3 Sequence Confirmation Through Plasmid Purification

5 ml of autoclaved LB Broth with 50 µg/ml of ampicillin in a 50 ml Falcon tube was inoculated with a sterile loop dipped in a galectin-3 glycerol stock. This was left to incubate at 37 °C shaking at 170 rpm overnight. The overnight culture was centrifuged at 22,000 *x g* for 10 minutes, and the pellet was subjected to downstream processing according to the manufacturer's instructions (QIAprep Spin Miniprep Kit, Qiagen, UK) using the 'Lyse Blue' technique.

The purified plasmid concentration was measured using the Nanodrop spectrophometer (2.4.3). A concentration of 100 ng/µl was prepared, and the plasmid was sent for sequencing to Sources BioSciences. The T7 (forward and reverse) primer was bought from Sources Biosciences directly for the sequencing.

2.10.3 Protein Expression of Recombinant Galectin-3

All procedures were carried out in aseptic conditions. Autoclaved LB Agar was prepared with 50 µg/ml Ampicillin and poured onto petri dishes until set. Once set, a sterile loop was plunged into a galectin-3 glycerol stock and streaked onto the agar plates, and left to incubate overnight at 37 °C.

A single medium to large colony was picked from the overnight agar/Ampicillin plate and was used to inoculate 400 ml of autoclaved-sterile of LB Broth supplemented with 50 µg/ml ampicillin (cultured in 2 litre flask to achieve a 1:5 ratio for appropriate aeration) and left to incubate overnight at 37 °C shaking at 170 rpm. The solution turned cloudy confirming growth. Expression was induced by adding 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) from a 0.2 M stock and incubation was continued at 33.5 °C (can also be cultured at 30 °C or 37 °C) for a further 3 h

shaking at 170 rpm. After this the cells were poured into Beckman centrifuge pots, balanced by weight, and centrifuged at 4200 RCF for 20 minutes at 4 °C (using the Sorval Evolution RC centrifuge with a Sorval SLC-6000 rotor). The media supernatant was discarded and the pellet scraped off with an indented scraper and placed in a 50 ml Falcon, any remaining residual remnants of the pellet can be pipetted with the residual media at the bottom of the centrifuge pot, and then pipetted up and ejected into the 50 ml Falcon. The pellet was used for subsequent downstream experiment or stored at -80 °C until when needed.

2.10.4 Protein Purification of Recombinant Galectin-3

2.10.4.1 Preparation of the Divinyl Sulfone Activated Lactosyl-Sepharose Affinity Column

Since galectin-3 binds to lactose with good affinity, the recombinant protein can be purified by direct affinity chromatography using a lactosyl-sepharose column, without the need for a generic affinity tag. The resin was prepared in-house by the following two-day protocol:

Day 1:

A sintered glass filter with a vacuum was used, this was cleaned extensively with dH_2O. The sepharose bottle was shaken with force such that the resin was homogenously mixed and then 100 ml of the sepharose was poured into the glass filter connected to a running tap. wThe sepharose was washed through with 2 l of $_dH_2O$. Whilst washing, the sepharose was stirred with a 2 ml stripette. During the washing process, the bead cannot be left to dry or stirred with force, (additionally the tap attached to the glass filter cannot be turned off, as this can back-pressure and the force could disrupt the resin). The sepharose was drained and then it was transferred to a suitable size beaker, with the aid of a scraper. Thereafter, 100 ml of 0.5 M of Na_2CO_3 pH 11, and 10ml of divinyl-sulfone (DVS) (Sigma V370) were very carefully added to the sepharose in a fume

hood. The beaker top was covered with cling film and the resin stirred using a magnetic stirrer with a bar in the middle (preventing sepharose grinding) for 70 minutes. An observation in colour change was apparent during this incubation period whereby the colour turns initially to yellow which then turns into blue/grey. This confirms that the sepahrose has been activated with the vinyl group. After 70 minutes of stirring, the resin was efficiently washed by re-pouring the sepharose back into the glass filter and rinsing with 2 l of dH_2O with gentle stirring giving a pink run off solution. 100 ml of vinyl-activated sepharose was then transferred into a beaker and 20% (w/v) (20 g) lactose added into 100 ml of 0.5 M Na_2CO_3 pH 10. This was then gently mixed with the sepharose using a bar magnetic stirrer overnight at room temperature within the fume-hood. This allows the coupling of the lactose to the activated sepharose. The beaker top was wrapped with clingfilm to contain the toxic fumes.

Day 2:

The lactosyl-sepharose coupling product was poured into a vacuumed glass filter and washed with 4 l dH_2O with gentle stirring, then drained and transferred to a beaker. 100 ml of 0.5 M Na_2CO_3 pH 8.5, was added, with 2 ml of β-mercaptoethanol and stirred for 2 h with a mid-bar magnetic stirrer at room temperature. After 2 h, this was poured back into the vacuumed glass filter and washed with 4 l of $_dH_2O$. The resin was then transferred into a fresh beaker and 10 ml of $_dH_2O$ was added to prevent the resin from drying. Sodium Azide (NaN_3) was used as a resin preservative at a final concentration of 0.002% (v/v) from a 2% (v/v) stock (1:100 dilution), therefore, 1 ml of stock was added to the 100 ml of sepharose beads. The resin was then aliquoted into four 25 ml aliquots and topped up to 50 ml with dH_2O in Falcon tubes and stored at 4 °C.

2.10.4.2 Packing the Affinity Chromatography Column with Lactosyl-Sepharose

An empty column was packed with 10 ml of resin by resuspending the 25 ml aliquot with 25 ml dH_2O. 20 ml of this resuspension was drawn and inserted into the empty column (containing 10 ml of resin and 10 ml of dH_2O). The column had screw caps attached at the top and the bottom. The bottom (outlet port) and top (inlet port) screws were detached and the dH_2O from the resin was drained. The resin was then washed by gravitational flow by adding x5 resin volume (50 ml) of 1 x PBS.

2.10.4.3 Galectin-3 Pellet Lysis (Protein Solubilisation) and Protein Elution

Stored pellets from *E. coli* cultures induced to express galectin-3 (2.10.3) were resuspended with 40 ml of lysis buffer (4mM β-Mercaptoethanol, 2mM EDTA and 1xPBS (MEPBS)). If the pellet was taken out from -80 °C frozen stocks, then the pellet was thawed on ice gently, and then subjected to MEPBS buffer resuspension. 10 µl of the pellet resuspended sample was kept to run on an SDS-PAGE (pre-lysed sample). The resuspended pellet was placed in a plastic beaker with ice and then subjected to 10 micron amplitude ultrasonication for 60 seconds x10 with 60-second resting gaps in between each pulsation to prevent the sample from overheating. 10 µl of this sample was kept to run on an SDS-PAGE (post-lysis sample). The lysed sample was placed into the appropriate centrifugation tubes (50 ml polypropylene copolymer tubes, ThermoFisher, UK) and balanced by weighing. The centrifuge tubes were placed in 4 °C pre-cooled JA-30.50 rotor, and spun at 48,380 *x g* for 30 minutes. The supernatant was carefully decanted (kept 10 µl of sample for analysis by SDS-PAGE), this is the post-sonicated soluble fraction. The sample was poured into the lactosyl-sepharose column (~40ml), the bottom (outlet) and top (inlet) caps were unscrewed, the sample was allowed to migrate by gravitational force, and the flow through fraction was collected in a 50 ml Falcon (10 µl of sample was kept to run on an SDS-PAGE). The resin was washed with x5

resin volume (50 ml) of MEPBS buffer and the wash fraction was collected in a 50 ml Falcon (10µl of sample was kept for SDS-PAGE analysis). The protein was then eluted (galectin-3 fraction) off the resin by adding x5 resin volume (50 ml) of 150 mM lactose-MEPBS buffer and the eluent was collected in a 50 ml Falcon (10 µl of sample was kept to run on an SDS-PAGE). Samples were assessed by SDS-PAGE (2.5) to observe and analyse the protein purification procedure.

2.10.4.4 Preparing Galectin-3 for Storage

Lactose and β-mercaptoethanol were removed by two rounds of dialysis into 1.5 l E-PBS (30-fold dilution per round) using a 12-14 kDa molecular weight cut off dialysis tubing (Medicell Membranes Ltd, UK), pre-cleaned in $_{d}H_2O$. Post-dialysis, the sample was kept and concentrated using an Amicon® Ultra-15 Centrifugal Filter Unit 3K (3 kDa) MWCO (Merck Life Science Ltd, UK), sequentially pre-washed with 10 ml of dH_2O followed by 10 ml of E-PBS buffer (in each case spinning at 1500 *x g*, for 15 minutes, 4 °C). The dialysed protein sample was then concentrated 10-fold. The sample was then stored as 0.5 ml aliquots, snap frozen in liquid nitrogen and stored at -80 °C.

2.10.4.5 Size Exclusion Chromatography

Size Exclusion Chromatography (SEC) was used as a characterisation step rather than a polishing step to remove unwanted protein contaminants. Protein samples of galectin-3 alone, and galectin-3 bound to glycoprotein ligands (ASF or LacNAc), with or without glutaraldehyde crosslinking at varied concentrations, were studied.

Un-crosslinked galectin-3 samples were filtered through a 0.22 µm spin filter at a speed of 23,755 *x g* for 2 minutes to exclude any microbial contaminants from the filtrate, and the galectin-3:ligand samples were spun at 23,755 *x g* for 15 minutes at 4°C. The supernatant was used to load onto the AKTA purifier (GE Healthcare, USA).

Samples were passed through a Sephacryl® 400-HR resin with resolution covering a wide mass range (20-8000 kDa). This was packed in a 30 ml column connected to the FPLC apparatus. The column was washed with filtered/degassed dH_2O and pre-equilibrated with 2.5 column volumes of SEC buffer (E-PBS buffer filtered/degassed). 100-1000 µl of sample was loaded onto the column at a flow rate of 0.5 ml/min. The absorbance of the outflow was monitored for a wavelength of 280 nm to detect protein in solution (using a UV-900, GE Healthcare, USA). Fractions were collected in a 96 well tray (using the fraction collector tool, Frac-950, GE Healthcare, USA). Fractions collected with high absorbance peaks were then characterised by methods including SDS-PAGE, Western blotting, SEC-Multi angle light scattering (SEC-MALS) and negative stain electron microscopy (EM). At the end of the SEC run, the column was washed 5x column volume with dH_2O and stored in a 20% (v/v) degassed ethanol buffer.

2.10.4.6 Size Exclusion Chromatography Multi Angle Light Scattering (SEC-MALS)

These experiments were carried out with the expert assistance of Dr Chris Millard and Dr Louise Fairall (Schwabe group, University of Leicester, LISCB).

SEC-MALS was conducted on pooled peak fractions from SEC analysis of samples of galectin-3 and of galectin-3 bound to ASF (stabilised with glutaraldehyde crosslinkage). The same SEC column was coupled to a MALS instrument (Wyatt Technology Ltd, UK) and the column was pre-equilibrated with 2.5x column volume of filtered/degassed E-PBS. MALS instruments were also

washed with E-PBS (Dawn HELEOS II (Laser source) and T-rEX (laser detector) instruments). The AKTA purifier Unicorn software (SEC software) was aligned with the ASTRA (laser detector) software. The following ASTRA light settings were selected: Light scattering signal (red), refractive index (blue), UV 280nm (green). The 'EASI graph' function within the ASTRA software, derives a molecular weight distribution corresponding to the peak was selected to analyse the light scattering data.

2.10.4.7 Sucrose Density Gradient Ultracentrifugation

A 5%-45% (w/v) sucrose density gradient was prepared by making 50 ml stock of 5% (w/v) and 45% (w/v) sucrose in E-PBS buffer. The sucrose was dissolved by spinning on a rocker at room temperature for 2 h. The dissolved sucrose was syringe-filtered through a 0.45 µm membrane filter (Sarstedt, UK).

Gradient Master IP U-bend tubes (4.5 ml) were used. This was placed on a stable tube holder. 5% (w/v) sucrose was dispensed into the tube until it was half full. 45% (w/v) sucrose was then injected (with a fine needle) beneath this, at the base of the U-bend, completely filling the tube. The tube was then carefully transferred to the Gradient Master IP SW60 rotor (BioComp Instruments, Canada), and a 5%-45% (w/v) sucrose gradient option was selected to create a lineate 5%-45% (w/v) sucrose gradient in the tube.

500 µl sucrose solution was slowly removed from the top (5% (w/v) end) of the gradient and gently replaced with the galectin-3 or galectin bound to ASF +/- glutaraldehyde crosslinkage sample (500µl). The loaded gradient was then ultracentrifuged overnight in an ultracentrifuge (Discovery 100SE Ultracentrifuge, Beckman Coulter, USA) using a swinging bucket rotor (TH660, ThermoFisher, UK). The centrifugation was for a total of 18 h, 36,000 rpm, at 4 °C; both the rotor and ultracentrifuge were pre-cooled for 2 h prior to use. This process separates protein species

according to their molecular mass (correlates with sucrose density). After ultracentrifugation, sequential 200 µl fractions were collected across the full gradient and assessed by SDS-PAGE.

2.10.4.8 Galectin-3 Induced Oligomerisation

Galectin-3 was chemically cross-linked with glutaraldehyde. Titration studies were undertaken to optimise the concentration and cross-linking duration. These studies confirmed a final concentration of 0.125% (v/v) glutaraldehyde, with a 20-minute cross-linkage duration, on ice. The cross-linking reaction was quenched by adding 1 M stock of Tris-HCL pH 7.5 to equate a final concentration of 50 mM (room temperature).

Titration experiments established best concentration of the glycoprotein asialofetuin (ASF) binding to galectin-3 (20 µM:100 µM (galectin-3:ASF)).

Cross-linked samples of lone and ligand bound galectin-3 species were analysed via multiple methods including SDS-PAGE, Western blotting, sucrose density ultracentrifugation and SEC. Galectin-3:ASF complex formation (-/+ glutaraldehyde cross-linking) was assessed by SEC-MALS (2.10.4.5 -2.10.4.6).

2.10.4.9 Cross-linking Galectin-3:ASF Complexes Within a Sucrose Density Gradient

To attempt to reduce cross-linking between heterogeneous complexes, a 5%-45% (w/v) sucrose gradient containing 0.125% (v/v) glutaraldehyde (present in both stock solutions) was prepared and run (all other details as per 2.10.4.7). Fractions were assessed by SDS-PAGE.

2.10.4.10. Negative Stain Electron Microscopy (EM)

Negative staining is an established technique in which the background is stained, leaving the actual specimen of interest unstained, and therefore visible. There are multiple stains all of which are heavy metal salts (we chose to use uranyl acetate). The stain can scatter electrons strongly and also adsorb to biological matter (Carol and Harris, 2011). Viewing the sample under negative stain-EM was the first-port of electron microscopy for initial screening purposes and to analyse the homogeneity of sample preparation prior to subjecting the samples for cryo-EM data collections and structural studies.

Samples for negative-stain were initially buffer exchanged via diafiltration from E-PBS into 0.5 M Tris-HCL pH 7.5 buffer, as the heavy metal stain uranyl acetate reacts with phosphate. The sample was diafiltrated to an equivalent of >1000-fold dilution of phosphate.

2% (w/v) uranyl acetate was used and vortexed to dissolve any precipitate. 300 mesh carbon-copper grids (Agar Scientific Ltd, UK) were used. The grids were handled using a fine inverted grid-tweezer and was subjected to glow-discharge by placing on the star grid holder block. Glow-discharge was done using Quorum® Q150 TES for TEM Grids (Quorum Tech, UK). Post glow-discharge, the grids were placed on parafilm, and 10 µl of sample was placed onto the grid for 1 minute. Excess sample was removed from the grid by rotating the grid at a 90° angle and lightly dabbing on a tissue paper. Grids were placed onto parafilm, with the sample facing up, for two rounds of staining. To stain, 10 µl of uranyl acetate was added for 1 minute then blotted as above (repeated the staining twice). Control grids consisting of buffer plus stain, and stain alone, were also prepared. With the expert assistance of Dr Natalie Allcock (EM facility manager, University of Leicester), grids were studied on a JEM-1400 Transmission Electron Microscope (JEOL, Japan) using an accelerating voltage of 120kV. Images were captured on EMSIS Xarosa 20MP CMOS digital camera (ISS Group Services Ltd, UK) and camera calibration was achieved using a 2160 lines/mm cross diffraction grating replica (Agar Scientific Ltd, UK).

2.10.4.11 Exposed Sugar Chain Cleavage using PNGase Treatment (Non-Denaturing)

SEC indicated that very large (>8 MDa) galectin-3:ASF complexes were formed by incubating 20 µM galectin-3 with 100 µM ASF and cross-linking with glutaraldehyde. These did not seem tractable for detailed EM/cryo-EM analysis.

The very large (500-8000 kDa) galectin-3:ASF glutaraldehyde cross-linked complexes, were subjected to treatment with Peptide N-Glycosidase (PNGase) F (New England Biolabs, USA, #P0704S) with manufacturer's instructions scaled up 5-fold. 10 µg of the sample was mixed by gentle pipetting with 10 µl of Glycobuffer 2 (10x) and dH_2O (as required to give total volume of 100 µl). 17.5 µl PNGase-F was then added and mixed in the same way. Optimisation experiments indicated a 4 h incubation at 37 °C was sufficient for maximal de-glycosylation. The sample was then assessed by SEC. Protein-containing fractions of interest were pooled in proportion to the A_{280} absorbance peak, and was buffer exchanged into 1 M Tris-HCL pH 7.5 to be analysed by negative stain EM.

2.11 Covid-19-Related Project Methods

2.11.1 SARS-CoV-2 Construct Generation

2.11.1.1 Oligonucleotide Design

Genes of interest were incorporated into *Kpn*I and *Xho*I restriction sites of pLeics49 vector (University of Leicester, Figure 2.1). Overlapping oligonucleotide primers ~50 bp in length were designed with AT:CG ratio of ~50%, to ensure an annealing temperature ~59-66 °C. Oligonucleotides were synthesised and purified by Eurofins Discovery. The oligonucleotides were dissolved in dH_2O to make a stock concentration of 100 µM (as instructed by the manufacturers), and stored at -20 °C.

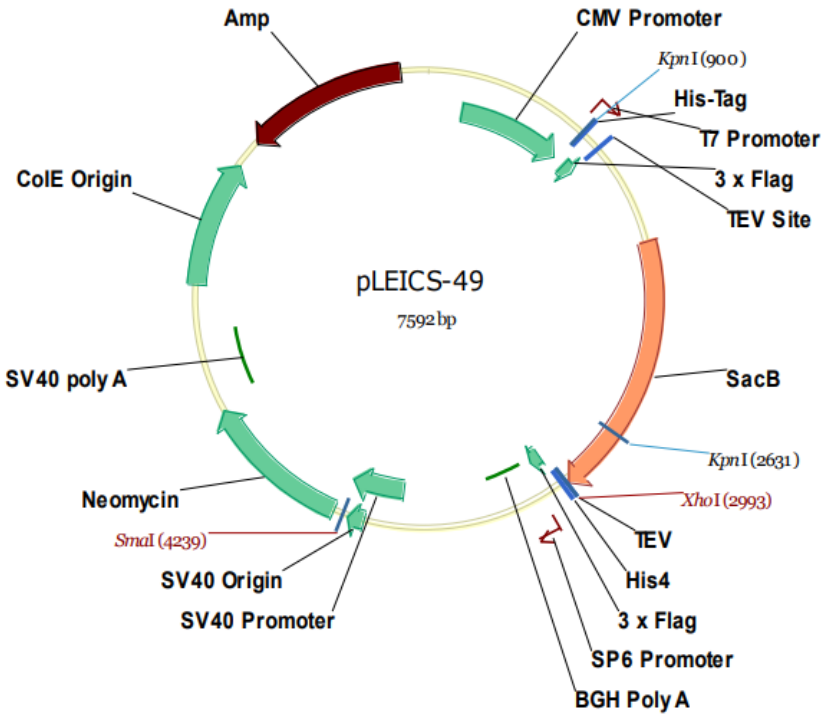


Figure 2.1, pLeics-49 Plasmid Map. pLeics-49 plasmid restriction map, prior to cloning in of Sars-CoV-2 spike construct genes at *KpnI* (900 and 2631) and *XhoI* (2993) site. The location of the Ampicillin resistance gene is also displayed and will remain post-cloning.

Primer Synthesis

NTD Fwd 5': **CTTTTGCCACTGGTAAGCAGTCAGAGTGTTGCAAGTCAAAGTATTATAGC**

NTD Rev 3': **AAAATACAGGTTCTCGAGAAAATTTGATGTTTGGTAGATTCCCTTTTC**

RBD Fwd 5': **CTTTTGCCACTGGTAAGCAGTCAGAGGGTACAACCTACGGAATCTAT**

RBD Rev 3': **AAAATACAGGTTCTCGAGAAAATTGACGCATTTATTCTTCACCAG**

S1 Fwd 5': **ACCCAAGCTTGGTACCATGTTTGTTTTCTTGGTTCTTTTGCCACTGGTAA**

S1 Rev 3': **AAAATACAGGTTCTCGAGTCGCGCTCTTCGCGGGGAATTG**

COVID_2 Fwd 5': **TTTGTTTTCTTGGTTCTTTTGCCACTGGTAAGCAGTC**

COVID_3 Fwd 3': **ACCCAAGCTTGGTACCATGTTTGTTTTCTTGGTTCTTTTGCCAC**

COVID_4 Rev 3': **AAAATACAGGTTCTCGAGTTGTTCATACTTACCGAGTTCTTGAAG**

2.11.1.2 DNA Amplification by Polymerase Chain Reaction (PCR): Strategy

A cDNA construct encoding the full-length SARS-CoV-2 spike protein (SARS-CoV-2-S, wild type sequence with native secretion signal, 2x C-terminal Strep-tagged, kind gift of Krogan group, Gladstone institute, USA) was used as a template (Figure 2.2). The native secretion signalling peptide was retained (at the N-terminus) and a FLAG-tag was inserted in the C-terminus.

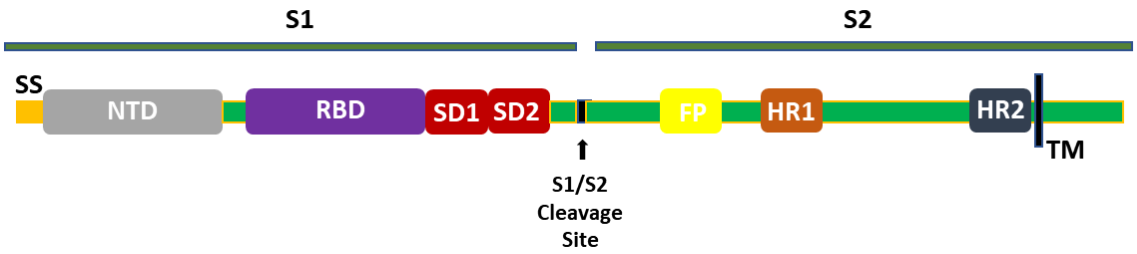


Figure 2.2, Full length Sars-CoV-2 Structure in a prefusion conformation. Schematic representation of the 2019 Full length Sars-CoV-2 primary structure (1D), simplified. The structure is annotated as follows: SS, signal sequence. NTD, N-terminal domain. RBD, receptor binding domain. SD1 & SD2, subdomain 1 & 2. FP, Fusion peptide. HR1 & HR2, Heptad repeat 1 & 2. TM, transmembrane domain. The SS, NTD, RBD and SD1/SD2 makes up the S1 (spike 1) domain. The FP, HR1, HR2 and TM makes up the S2 (spike 2) domain. The arrow displays the protease cleavage site between the S1 and S2 domain.

To Generate S1 Construct, Primers Used:

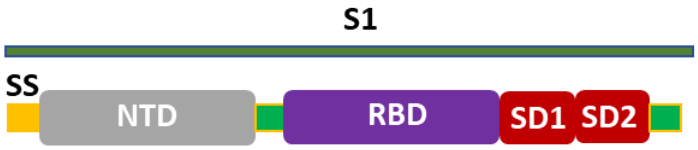
3 PCR reactions:

PCR #1: S1 FRW and S1 REV

PCR #2: COVID_2 and S1 REV

PCR #3: COVID_3 and S1 REV

To generate:



To Generate S1-NTD Construct, Primers Used:

PCR: NTD FRW and NTD REV

To generate:

To Generate S1-RBD Construct, Primers Used:

3 PCR reactions:

PCR #1: RBD FRW and RBD REV

PCR #2: COVID_2 and RBD REV

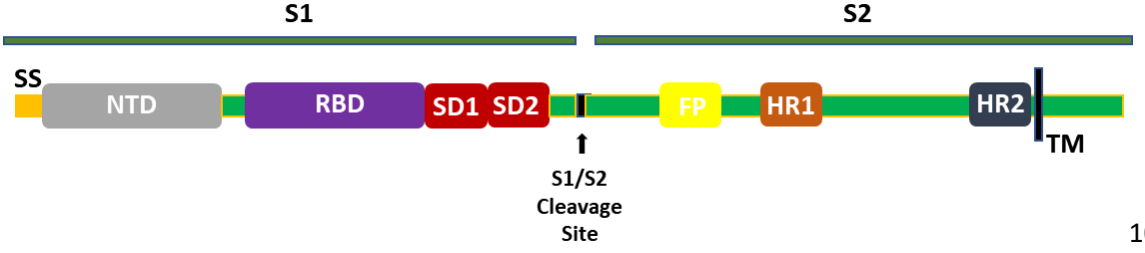
PCR #3: COVID_3 and RBD REV

To generate:

To Generate Full-Length Spike Construct, Primers Used:

PCR: COVID_3 AND COVID_4

To generate:



2.11.1.3 DNA Amplification by Polymerase Chain Reaction (PCR): Set-Up

All amplifications were carried out in a Storm PCR thermocycler. To produce the insert DNA for cloning, KOD hot start DNA polymerase (Sigma Aldrich, USA) was used. Table 2.4 displays the buffer and reagent composition per PCR reaction. Annealing temperature was varied depending on the oligonucleotide sequence composition. PCR was done during 35 cycles. Thermocycler was set as displayed in Table 2.5. PCR was conducted as duplicates, since one PCR product was used for gel electrophoresis to confirm the presence of the PCR products, and the second was used for cloning procedure.

Table 2.4, KOD Hot start DNA polymerase: 50µl total reaction volume

COMPONENT	VOLUME(µl)
10x buffer for KOD hot start polymerase	**5**
25mM MgSO4	**3**
dNTPS	**5**
PCR Grade Water	**30**
Sense 5' Primer(10µM)	**2.5**
Anti-sense 3' Primer (10µM)	**2.5**
Template DNA (10ng/µl)	**1**
KOD Hot Start DNA Polymerase	**1**
	TOTAL= 50

Table 2.5, PCR Thermocycler Settings

Step	**Target Size**			
	<500 bp	500-1000 bp	1000-3000 bp	>3000 bp
Polymerase Activation	95°C for 2 minutes			
Denature	95°C for 10 seconds			
Annealing	10 seconds (at lowest annealing temperature)			
Extension	70°C for 10s/kb	70°C for 15s/kb	70°C for 20s/kb	70°C for 25s/kb
Repeat Steps 2-4	35 cycles			

Primer Annealing Temperature and Extension for Each Construct:

S1-RBD: 59°C for annealing temp, estimated size ~ 0.6kb, for extension step use 70 °C 20s

S1-NTD: 59°C for annealing temp, estimated size ~1kb, for extension step use 70 °C for 20s

S1 Region: 66°C for annealing temp, estimated size ~2kb, for extension step use 70 °C for 40s

FL-Spike: 60°C for annealing temp, estimated size, ~4kb, for extension step use 70 °C for 100s

2.11.1.4 Agarose Gel Preparation and Visualisation for PCR Product Analysis

Agarose gel was prepared by initially preparing 1x TBE buffer (89 mM Tris (base), 89 mM Boric Acid and 2 mM EDTA). 1% (w/v) agarose TBE gel was made by adding 1 g of agarose to 100 ml of 1xTBE, heating and stirring to dissolve it entirely. Ethidium bromide was added to the solution at a final concentration of 0.5 µg/ml (in the fume-hood) and mixed by stirring. The solution was poured in an agarose gel cassette with an appropriate comb insert. Once the gel polymerised, the comb was removed, and the gel was immersed in 1xTBE (running buffer). To determine the size of the amplified DNA fragments, 3.5 µl of Quick Load purple 1kb plus DNA ladder was loaded. The PCR reaction was mixed with 5x NEB loading buffer (New England Biolabs, USA) to a final 1x dilution, and loaded into a neighbouring well. The electrophoresis was run at 80 V for 45 minutes. PCR products were analysed under UV light (302nm), (Syngene U Genius 3 system, USA). If the band of interest at the desired molecular weight was observed, then the PCR amplified material was subjected to clean-up and cloning, and, at key points, sequencing.

2.12 Cloning and Plasmid Preparation

2.12.1.1 Cloning Strategy

PCR-amplified Material Clean Up

Following PCR and agarose gel analysis, the PCR-amplified product of the NTD, RBD, S1 and Full-length constructs were purified by phenol chloroform extraction. For this the sample was diluted two-fold in dH_2O to give a 100 µl total volume. 100 µl of phenol chloroform was then added to the sample (1:1 dilution). The sample was mixed by high speed vortexing for 10 s. The sample was centrifuged at 25,000 *x g* for 4 minutes allowing two phase-separated layers to form. The upper aqueous layer, containing the DNA product was carefully transferred into a fresh Eppendorf tube without disturbing the lower interphase and organic phase phenol chloroform layer containing RNA and lipids, which was then discarded. The purified DNA was then washed and precipitated by adding 200 µl of 100% ethanol and 10 µl of 5 M NaCl. The sample was mixed by inverting the tube five times. The sample was then centrifuged at 25,000 *x g* for 20 minutes. Post-centrifugation the top layer (supernatant) was removed (without disturbing the pellet), and another 200 µl of 100% ethanol was added. The sample was centrifuged for a further 5 minutes at 25,000 *x g*. Post-centrifugation, the supernatant was completely removed and the pellet was left to air-dry with the Eppendorf lid kept open (~30 minutes) on the bench-top at room temperature. 25 µl of dH_20 was then added to resuspend the pellet. The concentration was measured using the Nanodrop spectrophometer (2.4.3).

Double Restriction Enzyme Digestion

Following phenol chloroform clean-up, 3.5 µg of the PCR product was digested using *KpnI* and *Xho*I restriction endonucleases at 3 U each. 3 µl of 10x CutSMART Buffer was used and the

reaction was topped up (with dH_2O) to give a total reaction volume of 30 µl. The sample was then incubated for 3 h at 37°C.

Isolation of DNA Fragments by Agarose Gel Electrophoresis

Restriction digest products were assessed by 1% (w/v) agarose gel electrophoresis. Bands of interest (at the expected molecular weight for intended products) were identified by UV fluorescence of intercalated ethidium bromide, and were cut out in a minimal volume of gel using a sterile razor and placed into an Eppendorf tube. QIAquick (QIAGEN Ltd. West Sussex, UK) gel extraction kit (according to manufacturer's instructions) was then used to extract the DNA fragment. Purified DNA products were eluted in 30 µl of sterile dH_2O. The eluted DNA is subjected to phenol chloroform clean-up. Concentration was determined measuring the A_{260} value using the Nanodrop spectrophometer (2.4.3).

Digestion of Original Plasmid to Isolate the Vector of Interest

3.5 µg of plasmid was used for digestion. Restriction endonucleases (*KpnI* and *XhoI*) were used at 3 U each. 3 µl of 10x CutSMART Buffer with 1.2 µl of CIP (1 µl per 25 µl reaction) was used and the reaction was topped up (with dH_2O) to equate a total reaction volume of 30 µl. Digestion was completed by incubating for 3 h at 37 °C. This digested vector was subjected to phenol-chloroform clean up.

Construct and Vector Bonding by Ligation

T4 ligase was used to ligate the gene construct of interest into the vector backbone. 50 ng of the DNA construct was added to 150 ng of the vector linearised by *KpnI* and *XhoI* restriction digestion.

Subsequently, 1 µl of 10x Ligation buffer and 0.5 µl of T4 ligase were added together with dH_2O as required to give a final reaction volume of 10 µl. The sample was incubated for 2 h at room temperature.

2.12.1.2 Preparation of NEB Competent Cells (Large Scale Stock)

Under strict aseptic conditions, 50 µl of Stable Competent *E. coli* (High Efficiency) cells (New England Biolabs, USA) was gently thawed on ice, and added directly to 500 ml of LB broth, containing ampicillin (50 µg/ml). The cell culture was incubated at 37 °C, shaking at 170 rpm until the OD_{600} reached a value of between 0.34-0.4. OD_{600} measurements were checked every 30 minutes, by blanking the spectrophometer with LB Broth media, and pipetting out 1 ml of the bacterial culture. Once the desired OD_{600} had reached, the cells were centrifuged at 1715 *x g* for 10 minutes at 4°C. The media was removed. The remainder of the experiment was done in 4°C. The pellet was resuspended by adding 100 ml of ice cold sterile-filtered 0.1M $MgCl_2$ buffer. This was transferred and split into 2x 50 ml Falcon tubes. The Falcon's were placed on ice to incubate for 30 minutes. Thereafter, cells were centrifuged at 1715 *x g* for 10 minutes at 4°C. The supernatant was discarded and the pellet was resuspended with 25 ml of ice cold filter-sterilised 0.1 M CaCl2, 14% (v/v) glycerol buffer. This was then divided into 500 µl aliquots into cryovials, snap frozen in liquid nitrogen, and stored at -80°C.

2.12.1.3 Competent Cell Transformation

100 µl of NEB competent cells was gently thawed on ice and mixed with the total 10 µl ligated sample and left to incubate on ice for 20 minutes. The tubes were then placed on a heat block, at 42 °C for 45 seconds to heat shock cells, and immediately after placed on ice for a further 5 minutes. The sample was topped up with an additional 500 µl of LB Broth and the sample was left to incubate at 37 °C for 1 h. The sample was then centrifuged at 1715 *x g* for 1 minute. The

500 µl of LB media was removed, and the pellet was resuspended. This was then pipetted onto an (50 µg/ml) Ampicillin-resistant agar plate, and streaked using a sterile spreader. The plates were left to incubate overnight at 37 °C. Vector only construct was used as a control.

2.12.1.4 Colony Screening PCR (Pre-Sequencing Sense Check)

Post-overnight incubation, if there was at least a 1:5 ratio difference between the grown colonies of the vector control and ligated plates, then a pre-sequencing sense check was done (otherwise ligation and transformation was repeated). At random, ten-fifteen colonies were picked from the ligated plate with a sterile loop and used to inoculate 5 ml of (50 µg/ml) Ampicillin treated LB Broth (in a 50 ml Falcon). This was left to incubate overnight at 37 °C, shaking at 170 rpm. Post-overnight culture the culture was centrifuged at 1715 *x g* for 10 minutes. The media supernatant was discarded, and the pellet was used for plasmid purification using the QIAprep Spin Miniprep Kit (Qiagen Ltd. West Sussex, UK). The purified plasmid was subjected to double restriction digestion test (as per 2.12.1, with three-fold decrease in all enzyme and buffer quantities). Samples characterised by bands at the correct molecular weight by agarose gel electrophoresis were checked by sequencing.

2.12.1.5 Large Scale Plasmid Isolation (Maxi prep) in Alkaline Conditions

Transformation was repeated on the purified plasmid (mini-prep) that were sequenced positively, using 10 ng of plasmid with 100 µl of NEB competent cells (2.12.1.3). Post-overnight incubation, colonies were picked with a sterile loop to inoculate 800 ml LB Broth containing 50 µg/ml ampicillin, in a 2 l conical flask. The flask was left to incubate overnight at 37 °C, shaking at

175rpm. The culture was then centrifuged at 1715 *x g* for 20 minutes at 4 °C, and the pellet was transferred into a 50 ml Falcon tube and stored at -80 °C if necessary.

The pellet was resuspended in 12 ml of 10 mM Tris-HCL pH 8, 10 mM EDTA buffer. Thereafter, 4 ml of 10 mM Tris-HCL pH 8.0, 10 mM EDTA with 50 mg Lysozyme buffer, was added to the Falcon tube and mixed by inverting five times. This was then equally split into 2x 50 ml Falcon tubes and left to incubate at room temperature for five minutes. For each tube, 15 ml of 0.2 M NaOH 1% (w/v) SDS buffer was added, and mixed by inverting five times. After incubation at room temperature for 5 minutes, 11 ml of 3 M potassium acetate pH 4.8 was added per tube, mixed by inverting five times, and again left to incubate at room temperature for 5 minutes. These treated lysates were then centrifuged at 1715 *x g* for 10 minutes at 4°C.

A Miracloth (22-25 µm pore size) was placed on top of a fresh 50 ml Falcon tube, and the supernatant was filtered though. Isopropanol was added proportional to 0.6x volume of the supernatant, then the tube was inverted five times, and left to incubate at room temperature for 5 minutes. This was then centrifuged at 1715 *x g* for 10 minutes at 4 °C. The supernatant was carefully discarded, and the pellet was resuspended with 7.5 ml of 10 mM Tris-HCL pH 8.0, 1 mM EDTA buffer. The Falcon was vortexed vigorously until the pellet completely dissolved. An equal volume of 5 M LiCl was added to the resuspension, and the tube was inverted five times and left to Incubate on ice for 5 minutes. This was then centrifuged at 1715 *x g* for 10 minutes at 4°C.

The supernatant was transferred to a new 50 ml Falcon tube, and the pellet was discarded. Proportional to the volume of supernatant, 0.6x volume of isopropanol was added and the tube was inverted five times, this was then left to incubate at room temperature for 5 minutes. The sample was then centrifuged at 1715 *x g* for 10 minutes at 4°C. The supernatant was discarded. The pellet was resuspended in 10ml of Tris-HCL pH 8.0, 1 mM EDTA buffer by vigorous shaking until completely dissolved. To the Falcon, 50 µl of 10 mg/ml of heat-treated RNase A was added

and mixed by inverting the tube ten times. This was left to incubate for 15 minutes at room temperature. Proportional to the volume, 1/3 volume of 30% PEG 6000, 2.5 M NaCl buffer was added. The solution should turn cloudy. If the solution does not turn cloudy, then more buffer can be slowly added in (but cannot equate to more than ½ volume). This was then left to Incubate for 30 minutes on ice followed by centrifugation at 1715 *x g* for 10 minutes at 4 °C

The pellet was resuspended in 10 ml, of 10 mM Tris-Hcl pH 8.0, 1 mM EDTA buffer and dissolved completely by vortexing. To this resuspension, 2 ml of chloroform was added (to extract PEG), this was vortexed briefly (20 seconds) followed by centrifugation at 1715 *x g* for 2 minutes at 4 °C. The aqueous (upper) layer was retained, and carefully transferred into a clean 50 ml Falcon tube.

To precipitate the DNA, relative to the DNA solution volume, 1/10 volume of 5 M NaCl and 3x volume 100% ethanol was added and mixed by inverting the tubes five times, followed by subsequent centrifugation at 1715 *x g* for 10 minutes at 4 °C. The supernatant was discarded, and the pellet was then washed with 20 ml of 100% ethanol and centrifuged at 1715 *x g* for 2 minutes at 4 °C. The supernatant was carefully removed, and the pellet was left to air-dry overnight, by placing the Falcon tube inverted on a piece of clean paper. Post overnight drying, the pellet was resuspended in 2 ml of syringe filtered (0.22 µm, Sarstedt, UK) 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA buffer, then further filter sterilised. The concentration of the purified plasmid was measured by testing the A_{260} value using the Nanodrop Spectrophometer (2.4.3) and stored at -20°C until required for downstream experiments.

2.13 HEK 293 Suspension Cell Culture and DNA Transfection

2.13.1 Cell Line

Human Embryonic Kidney (HEK)293F suspension-adapted cells were kindly donated by Schwabe group (University of Leicester, LISCB). Cells were received at a seeded density of $5x10^5$ cells per ml (60 ml total volume).

2.13.1.1 HEK 293F Cell Culture

For growth and maintenance, HEK 293F cells were cultured in a sterile tissue cell culture hood in a Corning 250 mL Polycarbonate Erlenmeyer Flask with a vent cap attached, using FreeStyle 293 Expression medium + GlutaMAX, without antibiotic or serum supplementation in a total of 60 ml volume. The cells were cultured and incubated in an orbital shaker at 37 °C under an atmosphere of 95% air/5% CO_2, shaking at 150 rpm and passaged every third day (usual duration to reach confluence). For cell count, the Erlenmyer flask was agitated by hand with medium-force, and the culture was then resuspended by stripette-mixing multiple times. 1 ml of medium was aspirated. The cell count and viability was assessed by Trypan blue assay, and the cells were then seeded into a fresh Erlenmyer flask at a density of $5x10^5$ cells/ml, topped up to a total volume of 60 ml using FreeStyle 293 Expression Media.

2.13.2 DNA Transfection

24 h prior to transfection confluent HEK 293F cells were passaged and seeded at a cell density of $5x10^5$ cells/ml for small- (60 ml test expression) or large-scale (300 ml) transfections. This corresponds to a predicted cell density of ~$1x10^6$ cells/ml after 24 h. For the small-scale transfections initial cultures were undertaken in a Corning 250 ml Polycarbonate Erlenmeyer

Flask with a cent cap attached. Large-scale preps were done in a 1 l Roller Bottle with 431132 Easy Grip Vented Roller Bottle Cap (Corning, USA).

To transfect, 1 µg/ml of sterilised DNA was added into 6 ml (for small-scale expression) or 30 ml (large-scale) of 1x DPBS (1:10 dilution to the original cell culture media) and shaken vigorously for 20 seconds. Subsequently, 120 µl (small-scale) or 600 µl (large-scale) of 1 mg/ml filter-sterilised polyethylenimine (PEI) transfection reagent was added to the PBS-DNA solution and shaken vigorously for 20 seconds. This was then left to incubate for 20 minutes at room temperature in the tissue-culture hood.

The DNA-PEI mixture was added to the 60 ml or 300 ml of cells cultured for 24 h. The transfected cells were then incubated in an orbital shaker for a further 72 h at 37 °C, shaking at 150 rpm, under an atmosphere of 95% air/5% CO_2.

2.13.3 Protein Expression and Purification

2.13.3.1 FLAG-tag Purification from Cell Pellet and Media

Proteins from the transfected cells were harvested post-72 h cell culture, by centrifuging cells at 1715 *x g* for 20 minutes at 4 °C. The media supernatant was transferred into a clean tube prior to immuno-affinity purification. The cell pellet was resuspended in 1 ml of lysis buffer (50 mM Tris/HCL pH 7.5, 150 mM potassium acetate (KAc), 10% (v/v) glycerol, 0.3% (v/v) Triton-X-100 and 1 tablet of EDTA-free protease inhibitor). The pellet was then placed on ice in a beaker, and lysed by sonicating at 10 micron amplitude, 3x for 15 seconds (with 15 second resting gaps in between each pulsation to prevent the sample from overheating). The lysed sample was centrifuged at 48,380 *x g*, 30 minutes at 4 °C in a pre-cooled JA-30.50 rotor. The supernatant was carefully decanted whilst the insoluble pellet was discarded. Anti-FLAG resin was prepared, by gently

thawing the resin on ice from its -20 °C storage. Once thawed, the tube was vortexed to mix the resin until homogenous. The resin was kept in a 50% (v/v) suspension in 100% stock glycerol. Using a truncated 200 µl pipette tip, 40 µl of suspension was drawn out of the stock tube into an Eppendorf tube. This was centrifuged for 30 seconds at 6000 *x g*. The resin was placed on ice for ~5 minutes to settle. The glycerol (upper layer) was removed, and the resin was washed with 1 ml of lysis buffer by inverting the Eppendorf five times. This was then re-centrifuged for 30 seconds at 6000 *x* g. The resin was placed on ice for a further ~5 minutes to settle, and the buffer (top layer) was removed. This was re-centrifuged for a further 30 seconds at 6000 *x g* and then placed on ice for 2 minutes to allow it to settle. Residual buffer was carefully removed without disturbing the resin. The lysed cell pellet was added directly to the resin, or the resin added directly to the media supernatant. This mixture was incubated for 2 h at 4 °C with vigorous agitation. It was then centrifuged for 1 minute at 6000 *x g*, and placed on ice for 5 minutes. After the supernatant was removed, the beads were washed with 1 ml of lysis buffer, followed by 4x wash with 1ml of wash buffer (50mM Tris-HCL pH 7.5, 150mM KAc, 5% (v/v) glycerol), and 4x wash with cleavage buffer (50mM Tris-HCL pH 7.5, 50mM KAc, 5% (v/v) glycerol and 9.5 mM DTT). Between each wash, the tube was centrifuged at 6000 *x g*, and placed on ice for 2 minutes to let the resin settle. 10 µl of sample was kept to run on an SDS-PAGE (Pre-TEV Sample). Once settled, the cleavage buffer was removed, and the resin was resuspended in 50 µl of cleavage buffer with 1 µl of TEV (1.5 µg/µl) followed by overnight incubation at 4 °C on a rocker. Post-overnight incubation, the tube was centrifuged at 6000 *x* g for 5 minutes and then placed on ice, allowing time for the resin to settle. The top supernatant layer contains eluted sample of interest (10 µl of this and the resin were kept to run and analyse on an SDS-PAGE), and was immediately used for subsequent experiments or stored at -80 °C.

2.13.3.2 His-tag Purification from Protein Secreted into Cell Media

Proteins from the transfected cells were harvested post-72 h cell culture, by centrifuging cells at 1715 *x g* for 15 minutes at 4 °C. The pellet was discarded, but the media fraction was transferred

into a new centrifuge pot, and centrifuged for a further 20 minutes at 1715 *x g* at 4 °C, to remove any residual pelleted fraction. The media was then transferred to a clean sterile bottle large enough to accommodate the 300 ml of media volume. 30 ml (10% (v/v) of the starting media volume) of equilibration/binding buffer (20 mM Tris/HCL pH 8.0, 0.5 M NaCl, 20 mM imidazole + 10% (v/v) glycerol) was added and mixed by rotating for 10 minutes (room temperature) to adjust the ionic strength and pH, and termed this the clarified sample. A 5 ml pre-packed Ni-NTA chromatography column (HisPur Ni-NTA Chromatography Cartridge #90099) was equilibrated with 50 ml (10x cartridge volume) of equilibration/binding buffer (applied through the column manually at low pressure by using a syringe). The ~300 ml of media-buffer mix was then injected over the column at low pressure, and the flow-through saved until the protein was successfully purified. The column was then washed with 10x column volume (50 ml) of wash buffer (50 mM Tris/HCL pH 8.0, 300 mM NaCl, 20 mM Imidazole, 10% (v/v) glycerol) and the wash outflow saved until successful protein purification was confirmed. The protein was then step-eluted at 0.5 M imidazole using 10x column volume of elution buffer (20 mM Tris-HCL pH 8.0, 0.5 M NaCl, 0.5 M Imidazole and 10% (v/v) glycerol). The column was then stored by washing with 5x column volume of dH_2O and storing in 20% (v/v) ethanol (70%). For re-use, the column was regenerated by washing with 10x column volume of dH_2O and subsequently with 10x column volume of equilibration/binding buffer.

The 50 ml of eluate was placed into size 12-14 kDa molecular weight cut off dialysis tubing (Medicell Membranes Ltd, UK) cleaned with dH_2O. It was dialysed against 1 l of Tris-buffered-saline with 5% (v/v) glycerol for 4-5 h, and then into identical fresh conditions overnight, at 4 °C. The sample was then concentrated as required (described in 2.10.4.4), aliquoted and used immediately for downstream experiments or stored at -80 °C.

2.13.4 Enzyme-Linked Immunosorbent Assay (ELISA)

These set of experiments were done with the assistance of PDRA Dr P. Stylianou and Dr O. Darweesh (Gooptu Group). A549 and THP-1 cells were seeded at a cell density of $8x10^4$ and $5x10^5$ cells in 1 ml of media respectively (loaded in 24-well plates) and cultured for 24 h. Both A549 and THP-1 cells were then subjected to various stimulations, including lone RBD stimulation at a final concentration of 2 µg/ml (for A549 cells) and 10 µg/ml (for THP-1 cells). Another variation in treatment of both A549 and THP-1 cells included pre-stimulating cells with 2 µg/ml (for A549 cells) or 10 µg/ml (for THP-1 cells) of RBD for 1 h, followed by the secondary treatment of either 5 µM cynaropicrin or 1 µM dexamethasone for a total of 24 h treatment. The cells supernatant was then harvested and levels of the pro-inflammatory cytokines IL-6, IL-8 and TNF-α was quantified using the below described ELISA kit.

96-well high-binding microplates (DuoSet Ancillary Reagent kit 2 (DY008, R&D, UK)) were coated with 100 µl/well of capture antibody diluted in 1xPBS (as stated by the manufacturers) incubated overnight at room temperature. Plates were washed three times with wash buffer (0.05% (v/v) Tween-20 in 1xPBS). Non-specific binding was prevented by adding 1% (w/v) BSA in 1xPBS (blocking buffer) to each well, and then incubated for 1 h at room temperature. Plates were washed again three times with wash buffer. Seven standards were prepared according to manufacturer's detection range and recommendation (see table 2.6) diluted in reagent diluent (1% (w/v) BSA in 1xPBS); 100 µl was added to each well as duplicates. 100 µl of harvested supernatants were added to each well (as duplicates) and plates were incubated for 120 minutes at room temperature. After three washes with wash buffer, streptavidin antibody conjugated to Horseradish Peroxidase (HRP) diluted in reagent diluent (1:40) were added at 100 µl per well. Plates were incubated for 20 minutes at room temperature, and then washed three times with wash buffer. The plates were developed using substrate solution for ELISA (DuoSet Ancillary Reagent kit 2 (DY008), R&D, UK) added at 100 µl per well. Samples were incubated for 20 minutes in the dark and stopped with 50 µl of stop solution (DuoSet Ancillary Reagent kit 2 (DY008), UK).

Plates were read at 450 nm and 570 nm wavelengths using the Enspire Perkin-Elmer Multi-plate reader spectrophometer. The 570 nm (background/noise) wavelength absorbance reading was subtracted from the 450 nm wavelength absorbance reading giving the true absorbance reading. Sample concentrations were determined by generating an interpolated standard curve with the seven standards absorbance readings using GraphPad prism 7 software.

ELISA kit	Cytokine	Species	Range	Supplier/Country
IL-6 DuoSet	IL-6	Human	9.38 - 600 pg/ml	R&D Systems, UK
CXCL8/IL-8 DuoSet	IL-8	Human	31.20 - 2,000 pg/ml	R&D Systems, UK
TNF-α DuoSet	TNF-α	Human	15.60 - 1,000 pg/ml	R&D Systems, UK

Table 2.6, ELISA standards range for each cytokine being measured (DuoSet Ancillary Reagent kit 2 (DY008)

2.14 Statistical Analysis

All data are presented as the mean ± standard deviation. Where appropriate, parametric data were analysed by analysis of variance (ANOVA) across groups, (multiple) paired t-test between related groups and unpaired t-test between unrelated groups. Data analysis was performed using GraphPad Prism software version 9 (GraphPad Software, Inc., La Jolla, CA, USA). The threshold for statistical significance was set to $p < 0.05$.

Materials

2.15 General Chemicals and Reagents:

Table 2.7, Chemicals and reagents utilised in this study

Reagent	**Supplier (+Product code)**
Albumin Standard (BSA)	ThermoScientific (23209)
Ampicillin	Merck (A0166)
AMPS	Merck (A3672)
Agar	Merck (05040)
ASF fom fetal calf serum	Merck (A4781)
β-Mercaptoethanol	Sigma (M-7522)
Bovine serum albumin (BSA)	Sigma (A2153)
Bromophenol blue	Merck (B0126)
Cell counting chamber 0.1mm depth	Hawksley
Chloroform	Merck (C2432)
Cynoropicrin	Carbosynth (35730-78-0)
DAPI	Sigma (32670)
DMSO	Sigma (D2650)
DPBS (-$CaCl_2$, -$MgCl_2$)	Giboco® (14190-144)
Divinyl Sulfone	Merck (V3700)
DynaBeads™ Protein-G	Invitrogen (10003D)
DynaMag™-2	LifeTechnologies (123211)
EDTA	Merck (E6511)
FBS	ThermoScientific (A4766801)
Fluoroshield Mount media	Sigma (1002602639)
Glutaraldehyde	Merck (G6257)
Hydrochloric acid	Fisher (7647-01-0)
Imidazole	Merck (I5513)
Isopropanol	Merck (I9516)
Methanol	Acros Organics (67-56-1)
Pierce® ECL Western Blot Substrate	ThermoScientific (32106)
Supersignal® West Pico Chemiluminescent Substrate	ThermoScientific (34087)
Supersignal® West Femto Chemiluminescent substrate	ThermoScientific (34094)
Gibco™ MEM Non-Essential Amino Acids Solution (100X)	FisherScientific (11570486)
Gibco™ Heat Inactivated Fetal Bovine Serum	FisherScientific (10500064)
Gibco™ Trypsin-EDTA (0.25%), Phenol red	FisherScientific (11570626)
Glycerol	Merck (G5516)
Glycine	Sigma (101285852)

ITS	Gibco® (41400-045)
KpnI	NEB (R3142-S)
LAcNAc	Merck (A7791)
Lactose	Merck (17814)
Laemmli Sample buffer 4x	Bio-Rad (1610747)
Lipopolysaccharide	Sigma (L5668)
Lithium chloride	Merck (L9650)
Lysozyme	ThermoScientific (89833)
MagicMark™ XP Western Protein Standard	ThermoScientific (LC5602)
Magnesium chloride	Merck (M8266)
Mem-PER™Plus Membrane protein extraction kit	ThermoScientific (89842)
PEG 6000	Merck (528877)
PEI	Merck (408727)
Penicillin/ Streptomycin (stock 10,000 Units/mL Penicillin and 10,000 µg/ml Streptomycin) (100X)	Gibco® (15140-122)
PFA	Sigma (P6148)
Phenol chloroform	ThermoScientific (15593031)
Polyethylene glycol sorbitan monolaurate (Tween-20)	Sigma (P1379)
SeeBlue® Prestained Protein Ladder	Invitrogen (LC5925)
Protease inhibitor cocktail (PIC)	Sigma
PMSF	Bio basic
QIAprep Spin Miniprep Kit	Qiagen (27104)
QIAquick Gel Extraction Kit	Qiagen (28115)
rhTGF-b1	R&D systems (240-B-002)
RNase	ThermoScientific (12091021)
Sepharose 4B	Merck (4B200)
SMA 30010P	Orbiscope
Sodium Azide	Merck (S2002)
Sodium Carbonate	Merck (222321)
Sodium Chloride	Merck (S3014)
Sodium dodecyl sulfate (SDS)	Sigma (L4390)
Sodium Orthovanadate	Merck (S6508)
Sucrose	Merck (57501)
Supersignal® West Pico Chemiluminescent Substrate	ThermoScientific (34087)
Supersignal® West Femto Chemiluminescent substrate	ThermoScientific (34094)
10x TBS (Tris-buffered saline)	Bio-Rad (1706435)
TEMED	ThermoScientific (17919)
Tris-Base	ThermoScientific (BP1521)
Triton-X00	Merck (900293)
10x TGS (Tris/Glycine/SDS)	Bio-Rad (161-0772)

Trypan Blue (0.4%)	Sigma (T8154)
Tryptone	Merck (T7293)
TrypLE	ThermoScientific (12604021)
Tween-20	Merck (900564)
T4 DNA Ligase Kit	NEB (M020S)
XhoI	NEB (R0146-S)
Yeast extract	Merck (Y1625)

2.16 Cell Culture Media

Table 2.8 Cell culture media used in this study

Media	**Supplier**
Dulbecco's modified Eagle's medium (DMEM) + 4.5 µg/L D-Glucose and L-GlutaMin	Gibco® FisherScientific (41965-039)
FreeStyle™ 293 Expression Medium+Glutamax	Gibco® FisherScientific (12338018)
Hank's Balanced Salt Solution (HBSS) -$CaCl_2$ and $MgCl_2$	Gibco® FisherScientific (11530476)
Trypsin-EDTA (0.25%), Phenol red	Gibco® FisherScientific (11570626)

DMEM was supplemented appropriately for the cell-type to be cultured (as mentioned in methods).

2.17 Buffers and Solutions

Table 2.9, Whole cell lysate (RIPA) and detergent free membrane (SMA) protein extraction lysis buffer

Buffer	**Composition**
Radio-Immunoprecipitation Assay (RIPA) buffer	0.01 mol/L Tris-HCL pH 8.0, 0.15 mol/L NaCl, 0.1% (w/v) SDS, 1% (v/v) NP-40, 0.5% (w/v) Na deoxycholate, 50mmol/L NaF, 30mmol/L sodium pyrophosphate, 1% (v/v) $Na3VO_4$, PIC and PMSF
SMA 2000P or SMA 30010P	Unknown (aqueous SMA polymers, confidential ingredients by manufacturers), 1% (v/v) $Na3VO_4$, PIC and PMSF

Table 2.10, Membrane cytosol Protein extraction lysis buffer (from MEM-PER™Plus Membrane extraction Kit (89842), ThermoScientific)

Buffer	**Composition**
Wash Buffer	Unknown (1862784)
Solubilisation Buffer	Unknown (1862783)
Perm Buffer	Unknown (89842X)

Table 2.11, Immunocytochemistry Buffers

Buffer	**Composition**
PBS-3% BSA Blocking Buffer	600mg of BSA dissolved in 20ml of 1xDPBS (-$MgCl_2$ and$CaCl_2$)
PBS-1% BSA Antibody diluent	8ml of PBS-3% BSA diluted in 16ml of 1xDPBS (-$MgCl_2$ and$CaCl_2$)
PBS-Tween20 Wash Buffer	0.05% (v/v)n Tween20 (25µl of Tween20 in 50ml of DPBS)

Table 2.12, PAGE running and Western Blot Buffers

Buffer	**Composition**
SDS-Page Running Buffer (1x)	25mM Tris-HCl, 192mM Glycine, 0.1% (w/v) SDS pH 8.3
Native Running buffer (SMALP)	25mM Tris-HCl, 192mM Glycine
1× Tris-Saline-Tween 20 (TST) (Wash Buffer and Antibody diluent)	20mM Tris-HCl, 500mM NaCl (pH 7.5) 0.1% (v/v) Tween-20
Blocking Buffer	5% (w/v) BSA or Milk (Marvel) powder in 1X TST buffer with 0.1% (v/v) Tween-20
4× Sample Buffer (denaturing)	60 mM Tris-HCl pH 6.8, 10 % Glycerol (v/v), 2% SDS (w/v), 0.01% bromophenol blue (w/v), 10% β-Mercaptoethanol (v/v)
2x Sample Buffer (non denaturing)	10 ml stacking gel buffer, 2 ml glycerol, 0.04% (w/v) bromophenol blue in 20 ml dH_2O

All other buffer composition is stated within the written methodology.

2.18 Regents and material for Western-blot:

Table 2.13 Western-blot material

Product	Product Code
Mini Protean®TGX™ Gels (4-15%)	456-1084
10x TGS	161-0772
10x TBS	1706435
4x Laemli Sample Buffer	1610747
Trans-Blot®Turbo™ Transfer Pack (mini PVDF membrane)	1704156

2.19 Antibodies

Table 2.14, Primary antibodies used in this study

Antibody	Host species	Clonality	Working Concentration	Application	Supplier	Isotype control	Supplier
Anti-Galectin3 sc-23938	Rat	Monoclonal	0.125 µg/ml 4.00 µg/ml	Immunoblotting Immunocytochemistry	SantaCruz	Rat IgG2a 559073	BD Pharmigen
Anti-β1-integrin 610468	Mouse	Monoclonal	0.005 µg/ml	Immunoblotting	BD Biosciences		
Anti-β1-integrin ab24693	Mouse	Monoclonal	3.00 µg/ml	Immunocytochemistry	AbCam	Mouse IgG1 X0931	Dako
Anti-CD98 E5 sc-376815	Mouse	Monoclonal	0.125 µg/ml	Immunoblotting	SantaCruz		
Anti-CD98 ab23495	Mouse	Monoclonal	3.00 µg/ml 4.00 µg/ml	Immunocytochemistry Flow cytometry	AbCam	Mouse IgG1 X0931	Dako

Anti-Na^+/K^+ ATPase NB300-146	Mouse	Monoclonal	**(1:5000 Dilution)**	Immunoblotting	Novus Biologicals		
Anti-β-actin sc-47778	Mouse	Monoclonal	0.04 µg/ml	Immunoblotting	SantaCruz		
Anti-α-SMA-FITC F3777	Mouse	Monoclonal	10 µg/ml	Immunocytochemistry	Sigma-Aldrich	IgG2a-FITC X0933	Dako
FSP F4771	Mouse	Monoclonal	4.00 µg/ml	Immunocytochemistry	Sigma-Aldrich	Mouse IgM X0942	Dako
Thy-1 CP28	Mouse	Monoclonal	3.00 µg/ml	Immunocytochemistry	Merck	Mouse IgG1 X0931	Dako
Vimentin M7020	Mouse	Monoclonal	1.50 µg/ml	Immunocytochemistry	Dako	Mouse IgG2a X0943	Dako
Collagen 550346	Rabbit	Polyclonal	20.00 µg/ml	Immunocytochemistry	Millipore	Rabbit IgG P021402-2	Dako
CD68 M0718	Mouse	Monoclonal	6.4µg/ml	Immunocytochemistry	Dako	Mouse IgG1 X0931	Dako
CD34-FITC	Mouse	Monoclonal	0.5µg/ml	Immunocytochemistry	Catlag		

Table 2.15, Secondary antibodies used in this study

Antibody	Dilution	Application	Supplier
Goat Anti-Rat IgG, HRP-linked antibody	1:1000	Western Blotting	SantaCruz(sc-2006)
Polyclonal Goat Anti-mouse IgG, HRP-linked antibody	1:5000	Western Blotting	Dako (P0447)
Goat Anti-Rat IgG, Alexa Flour 594	1:100	Immunocytochemistry	ThermoFisher Scientific (A21209)
Goat Anti-Mouse IgG, Alexa Fluor 594	1:100	Immunocytochemistry	ThermoFisher Scientific (A11005)
Goat Anti-Rat IgG, Alexa Fluor 488	1:200	Immunocytochemistry	ThermoFisher Scientific (A11006)
Goat Anti-Mouse IgG, Alexa Fluor 488	1:200	Immunocytochemistry	ThermoFisher Scientific (A11001)
Rabbit Anti-Mouse IgG FITC	1:10	Immunocytochemistry	Dako (F0313)
Goat Anti-Rabbit IgG, FITC	1:20	Immunocytochemistry	Sigma (F0382)
F(ab')2-Goat anti-Mouse IgG (H+L) Secondary Antibody, APC	1:200	Flow-cytometry	ThermoFisher Scientific (17401082)

2.20 CD98hc Knock-Down and Scratch Test Materials

Table 2.16, Reagents for Knock-down and scratch test stimulants

Product	Product Code
siRNA HUMAN SLC3A2	#D-003542-01-0010
SiRNA siGENOME Control	#D-001210-03-05
Lipofectamine™2000	#11668019
Lipopolysaccharide from *Escherichia Coli* O127:B8	L5668
rhTGF-β1	R&D systems (240-B-002)

Chapter 3

Examination of the Gal-3-Fibrosome Hypothesis in *Ex Vivo* Lung Tissue and *In Vitro* Human Type II Alveolar Epithelial Cell Model (A549 Cell Line)

3.1 Introduction

Work underpinning this project supported roles of galectin-3 in mediating lung fibrosis through its function in both alveolar epithelial cells and myofibroblasts (Mackinnon *et al*. 2012). I focused on these cells in turn. This chapter describes work using *ex vivo* lung tissue from patients with established IPF and in an experimental model of early lung fibrosis, and then more detailed studies in the A549 broncho-alveolar epithelial cell line.

The alveolar epithelium is comprised of two main cell types: type I and type II alveolar epithelial cells (Ward and Nicholas. 1984, Wong and Johnson. 2013). ~95% of the alveolar surface constitutes of type I alveolar epithelial cells and serves as the gaseous exchange barrier in the alveolus, whilst the remaining ~5% is covered by type II alveolar epithelial cells. Type II alveolar epithelial cells operate as 'caretakers' of the alveolar compartment, thereby maintaining lung homeostasis (Ward and Nicholas. 1984, Parimon *et al*. 2020). These cells function as progenitor cells to damaging insults exposed on both type I and II alveolar epithelial cells (Matthew *et al*. 2019). In addition, type II alveolar epithelial cells produce surfactant proteins, thereby enabling optimised gas exchange by lowering the surface tension within the alveoli (Camelo *et al*. 2013).

Type II alveolar epithelial cells are abundant secretors of galectin-3 into the extracellular matrix (Nishi *et al*. 2007, MacKinnon *et al*. 2012, Ho *et al*. 2016). Injury to alveolar epithelial cells and its (over)-activation is a major contributor in IPF pathogenesis *in vivo* (Camelo *et al*. 2013). Ultrastructural studies demonstrate the loss of both type I and type II alveolar epithelial cells with degenerated basement membrane (Ding *et al*, 2011). The loss of basement membrane integrity promotes inflammation (Strieter. 2008). Progenitor type II alveolar epithelial cells become metabolically active and function to repair the wounded basement membrane by re-epithelisation (Liebler *et al*. 2014). Consequently, several multiple pro-fibrotic mediators are increased including TGF-β1 (Camelo *et al*. 2013, Parimon *et al*. 2020). In line with re-

epithelisation, this can additionally cause exaggerated proliferation, activation and migration of underlying fibroblasts into injury sites, resulting in the formation of fibroblastic foci (Sakai *et al*. 2013). Furthermore, increased levels of TGF-β1 in these activated type II alveolar epithelial cells could induce epithelial to mesenchymal transition (EMT) of neighbouring epithelium resulting in increased expression of the mesenchymal myofibroblasts (Camelo *et al*. 2013). The myofibroblasts occupy the alveolar airspace, promoting excessive deposition of extracellular matrix, increasing lung stiffness and work of breathing whilst limiting diffusion of oxygen into alveolar capillaries (Selman and Pardo. 2002, Horowitz *et al*. 2006, Sakai *et al*. 2013).

Unfortunately for experimental studies, the use of primary type II alveolar epithelial cells is practically hampered by unreliable supply, and the technical challenges of performing isolation and culture procedures *in vitro* (Fuchs *et al*. 2003). More fundamentally, type II alveolar epithelial cells lose their specific phenotype over 1–2 weeks when cultured *in vitro*, de-differentiating to develop features characteristic of type I alveolar epithelial cells (Fuchs *et al*. 2003, Mao *et al*. 2015).

The A549 cell line has been extensively used to model type II alveolar epithelial cell behaviour *in vitro*. It was established in 1972 from a broncho-alveolar cell (BAC, lepidic) subtype of adenocarcinoma (Giard *et al*. 1973, Foster *et al*. 1998, Swain *et al*. 2010). A549 cells express multiple characteristic 'caretaking' features of type II alveolar-epithelium including the production of surfactant proteins, which in the lungs reduce surface tension and prevent alveolar collapse during ventilation (Nova *et al*. 2020). When cultured *in vitro* these cells are adherent (ATCC, UK), and display type II alveolar epithelial cell phenotype including multilamellar body expression (Lieber *et al*. 1976, Cooper *et al*. 2016, Selenius *et al*. 2019).

A549 cells are experimentally tractable, and a considerable body of literature describes their behaviour in different contexts (ATCC, UK). Nevertheless, interpretation of data from A549 cells

is limited by their atypical cellular origin, and the potential breadth of phenotype, covering type I alveolar epithelial cell and bronchial epithelial features (Swain *et al*. 2010, Fessart *et al*. 2013, Schefller *et al*. 2015, Cooper *et al*. 2016, Selenius *et al*. 2019).

I used A549 cells to model for type II alveolar epithelial cell behaviour in the absence and presence of TGF-β1 stimulation. I assessed the expression levels of galectin-3 and the putative fibrosome components CD98 and β1-integrin, and their co-localisation behaviour under basal and pro-fibrotic (TGF-β1 stimulated) conditions. Multiple methods were used including: western-blotting, co-immunoprecipitation, confocal immunofluorescence (on cells and lung tissue) microscopy and immunohistochemistry. The cell surface restricted distribution and co-localisation of the fibrosome-proteins was confirmed using western blotting and flow-cytometry. To examine the importance of functional crosstalk between acute and chronic exacerbation, wound healing responses of CD98hc silenced A549 cells was also assessed under pro-fibrotic conditions (TGF-β1 stimulation).

3.2 Results

3.2.1 Levels of the Putative Galectin-3-Fibrosome Components and Their Co-Localisation Addressed in the Epithelium of Non-Fibrotic-Control and IPF Lung Tissue

I first addressed the levels and distribution patterns of the putative gal-3-fibrosome proteins in non-fibrotic and IPF lung tissue, working with Dr Panayiota Stylianou (PDRA, Gooptu Group, University of Leicester). To these ends we used the monoclonal antibodies against galectin-3, CD98hc and β1-integrin and the Dako EnVision staining kit for immunohistochemistry of tissue samples, gratefully received from patient donors who had given informed consent for such work. Representative images of the immunostaining are shown in Figure 3.1. Antibody-free and IgG isotype controls (rat IgG2a for galectin-3 studies, mouse IgG1 for CD98, β1-intgerin, and E-cadherin studies) demonstrated no staining (Appendix, Figure A1).

In addition to looking at each protein individually in unrelated sections, the data presented here focussed on serial sections of tissues to assess their potential co-localisation at a tissue:cellular level (Figures 3.1a-b). To help understand this in terms of epithelial and non-epithelial cells, the epithelial marker E-cadherin was also included in these studies. The images are oriented as similarly as possible. Quantitation of the data indicated that the proteins assessed demonstrated increased expression in IPF lung tissue compared with non-fibrotic control lung tissue (Figure 3.2). Additional zoomed projections of these micrographs focussing more closely on the lung epithelia are displayed in Appendix, Figure A2. These studies overall supported similar cellular expression profiles for galectin-3, CD98hc and β1-integrin, most strongly within E-cadherin positive parenchymal cells, supporting an initial focus on lung epithelia.

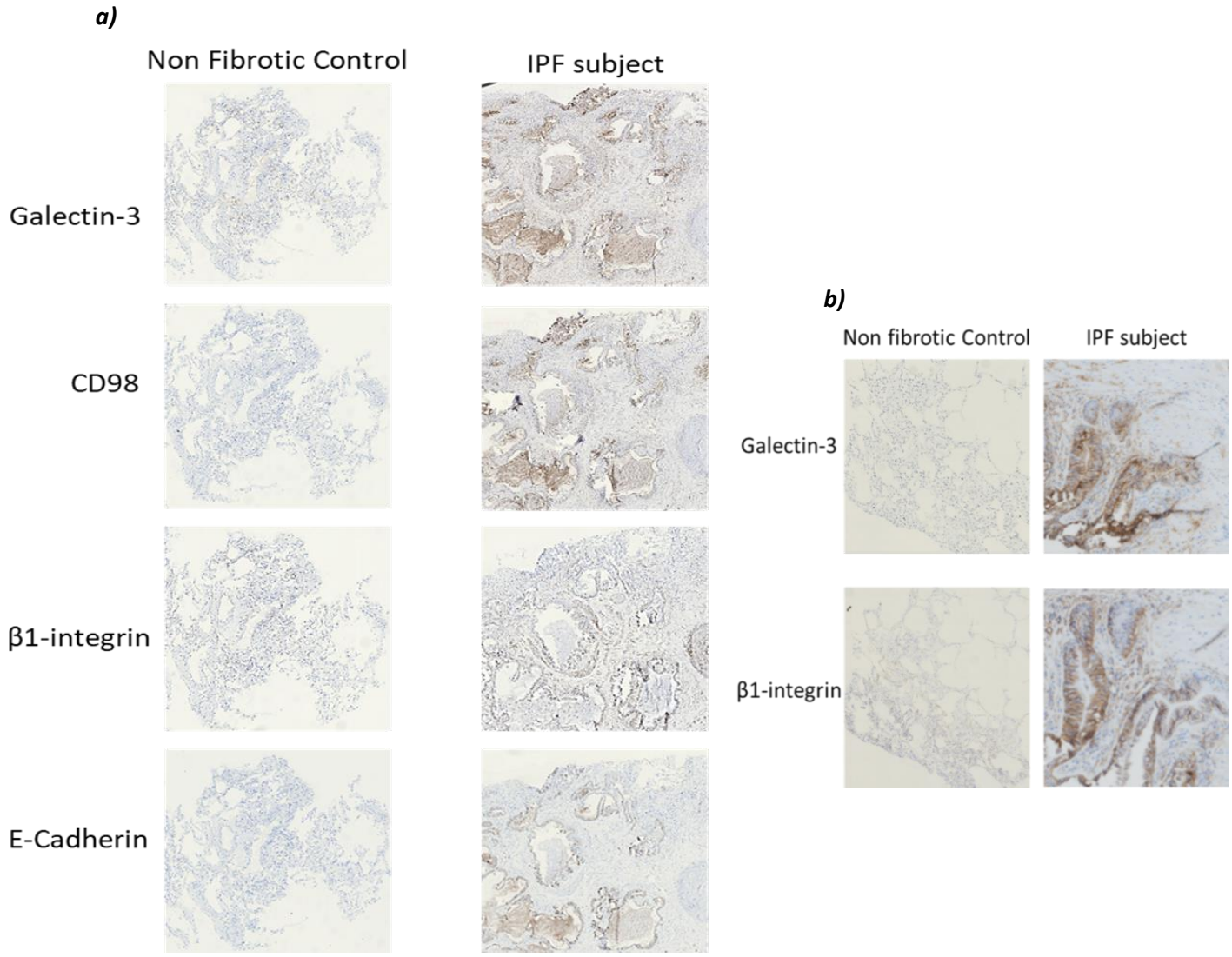


Figure 3.1, Gal-3-fibrosome proteins localise on the epithelium with increasing amounts in IPF subjects. Immunohistochemistry staining focusing on the levels and co-localisation of galectin-3, CD98hc and β1-integrin between NFC and IPF derived lung tissue serial sections. In both panel ***a)*** and ***b),*** NFC (on the left hand column) and IPF (on the right hand column) lung tissue samples were serially sectioned and stained. Tissue sections were singularly stained to probe for the gal-3-fibrosome proteins with monoclonal primary antibodies; 4 µg/ml galectin-3, 5 µg/ml CD98, 2.5 µg/ml β1-integrin. 5 µg/ml of E-cadherin monoclonal antibody was applied to highlight epithelial cells. 1 µg/ml of anti-rat-HRP secondary antibody was used to probe for galectin-3 (brown-stain) and 50 µl of secondary anti-mouse-HRP antibody (Dako EnVision kit) was used to stain for CD98, β1-integrin and E-cadherin (brown-stain). The nuclei were stained by immersing the slides into hematoxylin (blue stain), (n=4).

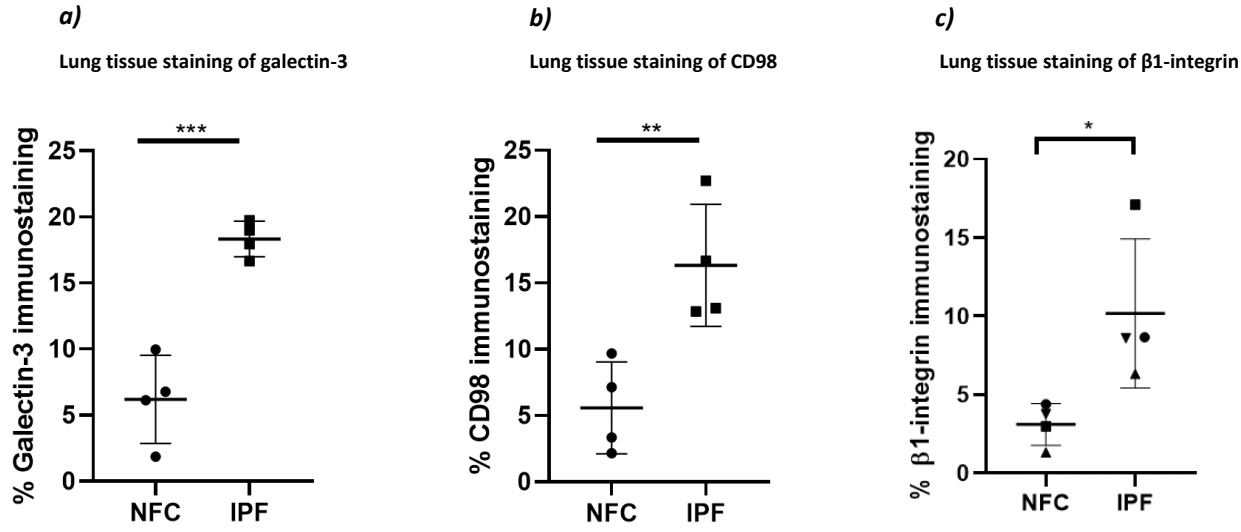


Figure 3.2, Quantification of galectin-3, CD98 and β1-integrin levels from the immunohistochemistry-stained NFC and IPF lung tissue. Quantification for the staining distribution of galectin-3 ***(a),*** CD98 ***(b),*** and β1-integrin ***(c),*** was done to compare their levels between NFC and IPF lung tissue serial-sections. This was achieved using a custom-made macro plug-in (with the help of Dr Kees Straatman, Senior Experimental Officer: Manager of the Advanced Imaging Facility, University of Leicester) using the software Image J. The macro analyses any brown (protein) staining signal, and reports this as a % of staining in the field of view at the magnification that the images were taken (Y-axis). The X-axis inscribes the appropriate lung tissue (NFC or IPF). The data was plotted using graph-pad prism 7. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by different stained tissue slices from different donors) are represented by the symbols, (n=4). Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from unpaired t-test (*p<0.05, p**<0.01, ***p<0.001).

3.2.2 Level of Putative Gal-3-Fibrosome Components in A549 Cells

I examined the expression levels of the putative galectin-3-fibrosome components: galectin-3, CD98hc and β1-integrin in the A549 cells. Expression of each protein was assessed by western blot analysis of the membrane and cytosolic fractions of lysed cells, and by immunofluorescence confocal microscopy in the absence or presence of the pro-fibrotic cytokine TGF-β1 (10 ng/ml, 24 h and 72 h stimulations). In addition, co-localisation levels of the putative gal-3-fibrosome components were probed by co-immunostaining confocal microscopy and co-immunoprecipitation.

3.2.2.1 Levels of Galectin-3, β1-integrin and CD98 Increases Following TGF-β1 Stimulation

Fractionated lysates (2.3.2) were analysed via western blotting. The blots were immunoblotted with the appropriate antibodies (Table 2.13). Single bands were detected for galectin-3, CD98 and β1-integrin at the predicted molecular weights of ~31 kDa, ~85 kDa and ~130 kDa respectively, confirming the specificity of the mAb antibodies used in my study. The protein levels of galectin-3 increased progressively when stimulated with TGF-β1 (comparing 24 h and 72 h TGF-β1 stimulation to their matched non-stimulated controls) in both the membrane and cytosolic fractions (Figure 3.3). Likewise, a progressive increase in the protein levels of CD98 and β1-integrin (both protein distribution are cell membrane restricted, absence of bands in the cytosol fraction) was observed. Membranes were probed for β-actin primarily as a cytosolic protein loading control and for the Na^+/K^+ ATPase subunit ATP1A1 as a membrane protein loading control. β-actin was detected in membrane as well as cytosolic fractions. This may reflect binding of actin by membrane proteins e.g. within fibrillar adhesions. Conversely ATP1A1 was only detected in the membrane but not cytosolic fractions, supporting the effectiveness of the fractionation technique and any potential membrane cytosol contamination. The expression levels of both housekeeping proteins are unchanged following TGF-β1 stimulation.

The levels of each protein normalised with reference to the ATP1A1 loading control, increased progressively when stimulated with TGF-β1 (comparing findings after 24 h and 72 h with non-stimulated controls) (Figures 3.4a-c). Levels of galectin-3 and CD98hc increased in the membrane fraction following TGF-β1 stimulation at both 24 h and 72 h (Figures 3.4a-b). Levels of β1-integrin in the membrane fraction increased following 72 h but not 24 h TGF-β1 stimulation (Figure 3.4c). Full uncropped western-blot scans of all probed proteins are provided in Appendix (Figure A4).

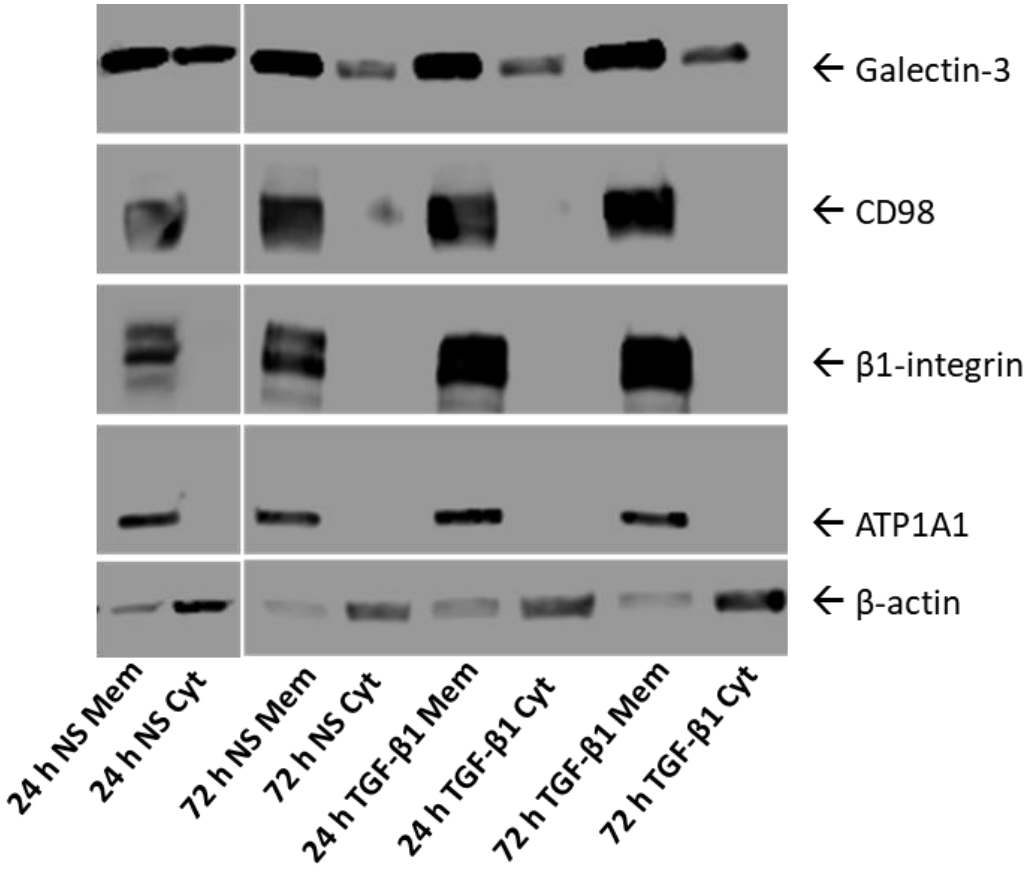
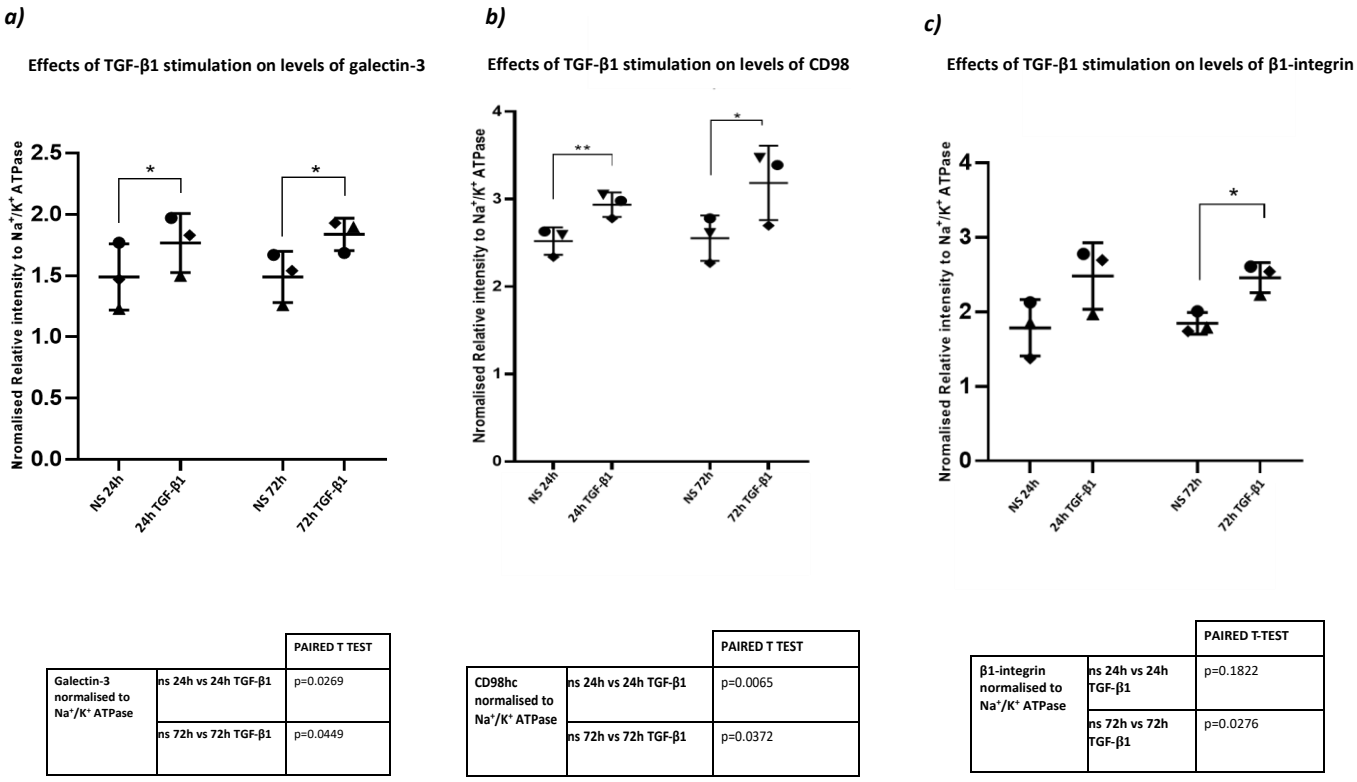


Figure 3.3, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in A549 cells by western-blotting. Exponentially growing A549 cells was used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3, 0.125 µg/ml CD98, 0.005 µg/ml β1-integrin, 0.04 µg/ml β-actin and 1:5000 Na^+/K^+ ATPase (ATP1A1). Lane 1-8 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective protein being represented. The blots were immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3).



		PAIRED T TEST
Galectin-3 normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.0269
	ns 72h vs 72h TGF-β1	p=0.0449

		PAIRED T TEST
CD98hc normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.0065
	ns 72h vs 72h TGF-β1	p=0.0372

		PAIRED T-TEST
β1-integrin normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.1822
	ns 72h vs 72h TGF-β1	p=0.0276

Figure 3.4, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin membrane levels from figure 3.4 immunoblotting. Quantification of the levels of galectin-3 ***(a),*** CD98 ***(b),*** and β1-integrin ***(c),*** was conducted from the western blot images using the software Image J and plotted on GraphPad Prism 7. The blot densitometric intensities were normalised against the membrane specific housekeeping protein ATP1A1 (Na^+/K^+ ATPase) and has been labelled on the Y-axis. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table below each scatterplot represents the full p value of each protein expressed as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

3.2.2.2 Endogenous TGF-βRII Levels were Undetectable

Unfortunately, TGF-βRII expression data are not shown as it was not detected by its respective primary antibody in any experimental context in A549 fractionated cell lysates and co-immunostaining according to manufacturer's instructions. To assess whether the proteins were lost or denatured during the cellular fractionation process, whole cell lysate samples were probed by western-blotting. We loaded varying concentrations of sample (20, 40, 60, 80 and 100 µg) and titrating antibody concentrations across and beyond manufacturers' recommendation, but the protein could not be detected in any case. These findings could theoretically indicate that this protein is not found at appreciable levels in A549 cells. However, other data from our and other studies indicate this is unlikely to be the case (Cheng *et al*. 2011). Instead, it supports the interpretation that these primary antibodies were not effective probes for western-blot analysis; whether it is the result of batch to batch variation or fundamental ineffectiveness in this application (despite manufacturers' product specifications to the contrary).

3.2.3 Co-Localisation of the Putative Galectin-3-Fibrosome Components in A549 Cells

Confocal microscopy was undertaken to assess co-localisation interactions of galectin-3 with CD98hc (Figure 3.5) and β1-integrin (Figure 3.6) following pro-fibrotic stimulation with 10 ng/ml TGF-β1. Images were captured using the Leica TCS SP5 confocal microscope. Co-localisation was probed between cell surface distributed proteins, by fixing the cells using 4% paraformaldehyde (PFA) instead of methanol (causing minimal permeabilisation and better immunostaining distribution) **(as in Figure 4.9)**. In these studies, one protein is stained with a probe signalling through a green channel (Alexa-fluor 488) and one with a probe signalling through a red channel (Alexa-fluor 594). Co-localisation is identified within the resolution of the technique, the microscope and the magnification in pixels where merging of the green and red channel data show signals from both (depicted as yellow pixel colouring).

3.2.3.1 Co-localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Increases Following TGF-β1 Treatment

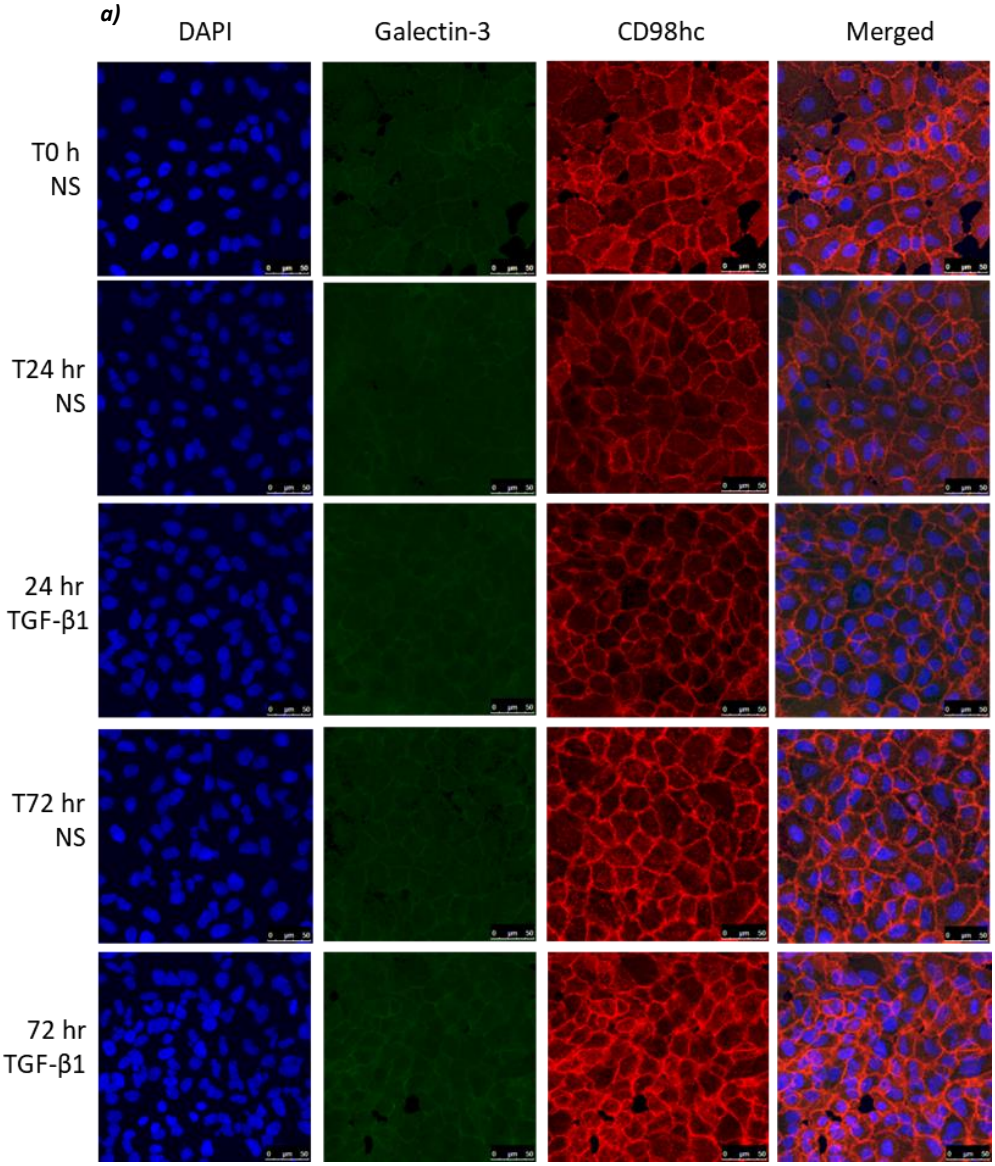
Rat IgG2a (for galectin-3) and mouse IgG1 (for both CD98hc and β1-integrin) isotypes were used as negative controls and did not produce any staining (Appendix, Figure A5). Following TGF-β1 stimulation, increases in immunofluorescence signals for galectin-3, CD98 and β1-integrin individually (Figures 3.5a and 3.6a) and for pairwise co-localisation (merged projections, yellow staining) appeared convincing once relative staining intensities were considered. These findings were quantified using Image J and 'Just Another Co-localisation Plug-in' (JACoP, Fiji suite).

Single channel immunostaining quantification from the galectin-3 and CD98hc co-immunostaining pairing demonstrated a significant increase of both proteins at 24 h TGF-β1 stimulation (not at 72 h TGF-β1 stimulation) (Figures 3.5b-c). Interestingly co-localisation quantification reported by Pearson's linear correlation indicated a significant increase at 72 h as a result of TGF-β1 stimulation (Figure 3.7a), though not at the 24 h timepoint, whilst Mander's

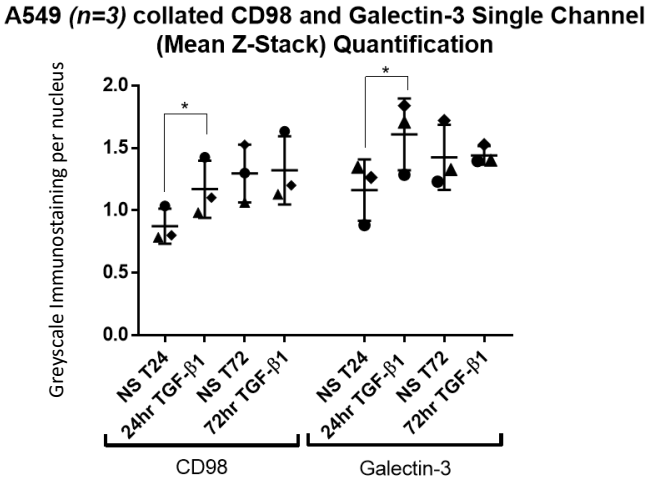
overlap coefficient did not report any change, at either timepoint (Figure 3.7b). These findings have to be taken in the context of baseline high signal for both measures, potentially limiting the sensitivities of these methods to detecting change.

Single channel immunostaining quantification from the galectin-3 and β1-integrin co-immunostaining pairing resulted a significant increase in β1-integrin at 24 h TGF-β1 stimulation (not at 72 h TGF-β1 stimulation) (Figures 3.6b-c). However, the galectin-3 response as measured by immunofluorescence appeared to demonstrate variability as the changes observed for galectin-3 did not reach statistical significance overall in these experiments. This was in contrast to the galectin-3 findings when assessed in the co-immunofluorescence experiments described above. This is potentially due to staining galectin-3 in different channels between the two protein pairs. The staining channel of galectin-3 was changed since β1-integrin did not stain at all in the red channel (Alexa-fluor 594) like CD98hc. Nevertheless, co-localisation assessment by Pearson's linear correlation and Mander's overlap coefficient showed a strikingly similar response profile to galectin-3:CD98 co-immunofluorescence. Once again co-localisation assessed by Pearson's linear correlation quantification was significantly increased by TGF-β1 stimulation but at 72 h, not at 24 h, whilst Mander's overlap coefficient did not demonstrate any change to the same significance (Figure 3.8a-b). Overall, whilst qualitatively the data matches the expected response profiles of each protein's response to TGF-β1 stimulation, the statistical analysis of single channel quantification is hard to interpret, and is best to consider this approach as a semi-quantitative method.

a)



b)

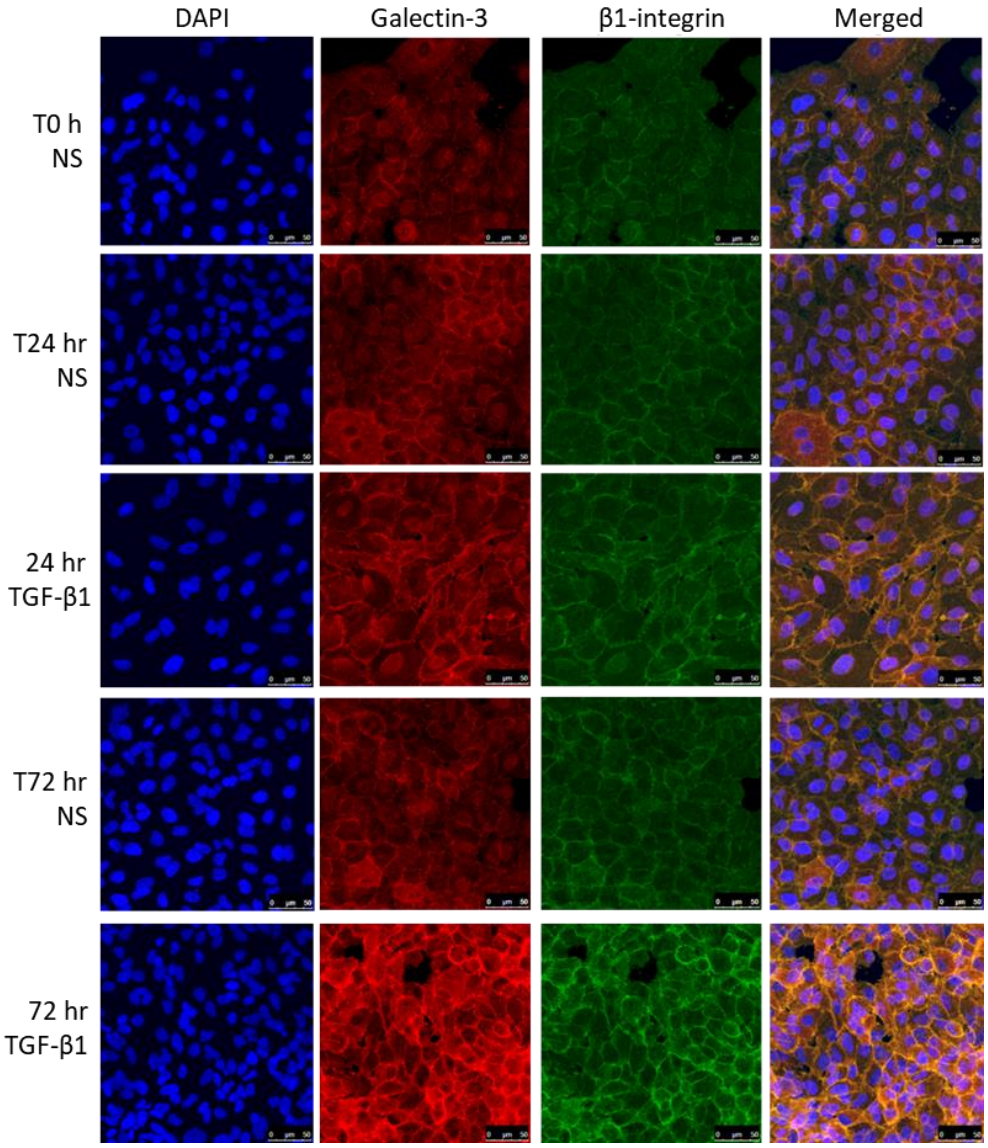


c)

		PAIRED T TEST
CD98hc Single Channel Immunostain	ns 24h vs 24h TGF-β1	p=0.0328
	ns 72h vs 72h TGF-β1	p=0.9074
Galectin-3 Single Channel Immunostain	ns 24h vs 24h TGF-β1	p=0.0211
	ns 72h vs 72h TGF-β1	p=0.8961

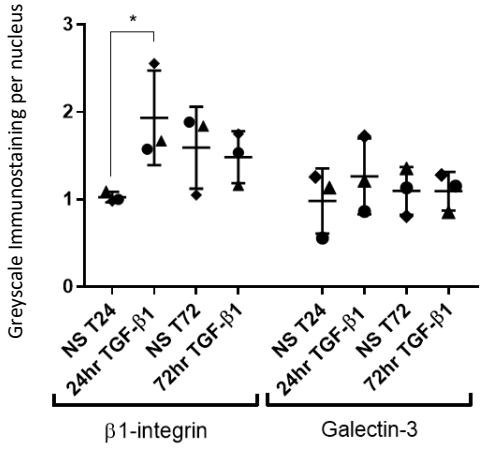
Figure 3.5, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in A549 cells. 3.5a, Cells were either untreated (T0 h NS, 24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 μg/ml galectin-3 and 5 μg/ml CD98hc using goat-anti-rat Alexa-fluor 488 (green) to stain galectin-3 and goat-anti-mouse Alexa-fluor 594 (red) to stain CD98hc. The nuclei were stained with DAPI. Each channel is shown singularly and with a merged projection. This is depicted as yellow stain but given the relative intensities of green (fainter) and red (more intense), pixels in common project as slightly more orange than the red single channel appearances. Representative data from n=3. ***3.5b-c,*** Single channel quantification of galectin-3 and CD98hc from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (n=*3*). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of CD98 and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table in figure ***3.5c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

a)



b)

A549 *(n=3)* collated β1-integrin and Galectin-3 Single Channel (Mean Z-Stack) Quantification



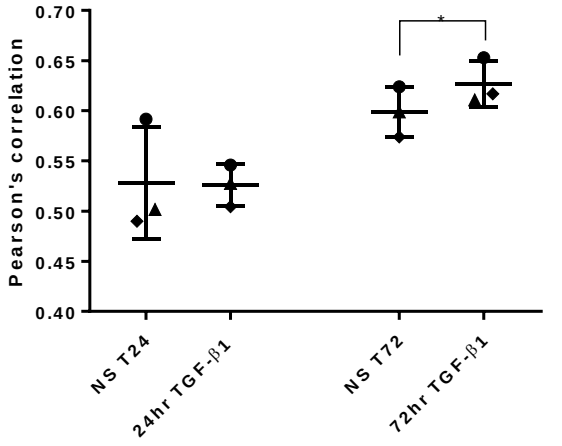
c)

		PAIRED T TEST
β1-integrin Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.039
	ns 72h vs 72h TGF-β1	p=0.8172
Galectin-3 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.1294
	ns 72h vs 72h TGF-β1	p=0.9936

Figure 3.6, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in A549 cells. 3.6a, Cells were either untreated (T0 h NS, 24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 µg/ml galectin-3 and 3 µg/ml β1-integrin using goat-anti-rat Alexa-fluor 594 (red) to stain galectin-3 and goat-anti-mouse Alexa-fluor 488 (green) to stain β1-integrin. The nuclei were stained with DAPI. Each channel is shown singularly and with a merged projection. This is depicted as yellow but given the relative intensities of green (fainter) and red (more intense), pixels in common project as slightly more orange than the red single channel appearances. Representative data from n=3. ***3.6b-c,*** Single channel quantification of galectin-3 and β-integrin from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (n=3). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of CD98 and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table in figure ***3.6c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

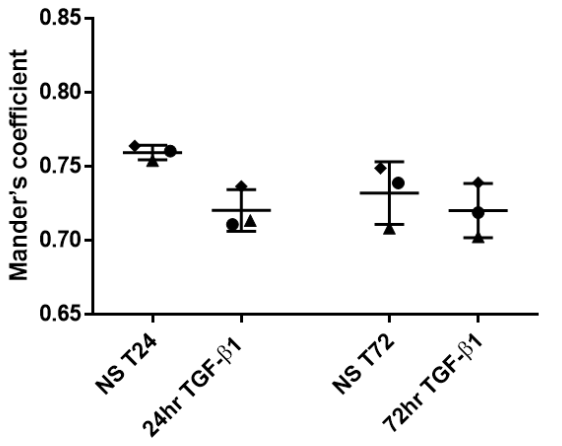
Co-localisation quantification of galectin-3 and CD98hc co-immunostaining following TGF-β1 stimulations

a)



		PAIRED T TEST
CD98 and Galectin-3 Pearson's correlation co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.9470
	ns 72h vs 72h TGF-β1	p=0.0325

b)

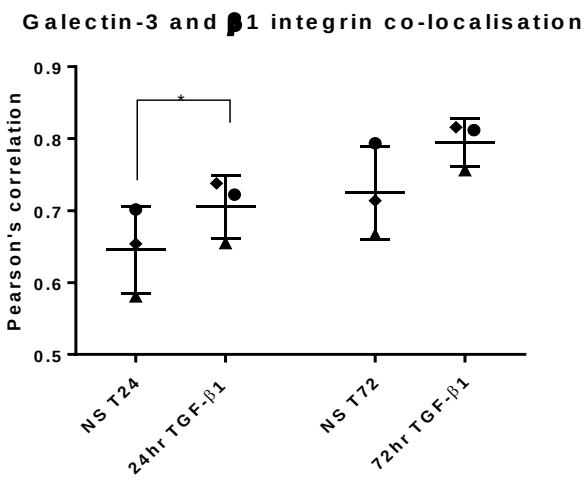


		PAIRED T TEST
CD98 and Galectin-3 Mander's coefficient co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.5817
	ns 72h vs 72h TGF-β1	p=0.1069

Figure 3.7, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations. 3.7a, Pearson's quantification of co-localisation and ***3.7b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded significance represented for paired t-tests with asterisks are displayed on the graphs. Full p-values for this are given in the lower panels (*p<0.05, p**<0.01, ***p<0.001).

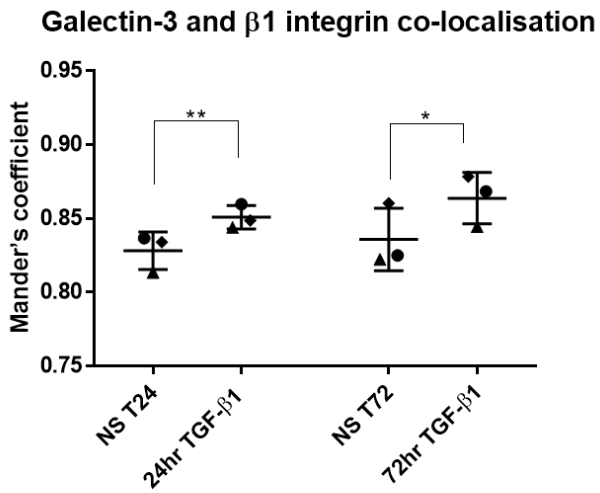
Co-localisation quantification of galectin-3 and β1-integrin co-immunostaining following TGF-β1 stimulations

a)



		PAIRED T TEST
β1-integrin and Galectin-3 Pearson's correlation co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.0441
	ns 72h vs 72h TGF-β1	p=0.0686

b)



		PAIRED T TEST
β1-integrin and Galectin-3 Mander's coefficient co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.0065
	ns 72h vs 72h TGF-β1	p=0.0194

Figure 3.8, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations. 3.8a, Pearson's quantification of co-localisation and ***3.8b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded significance represented for paired t-tests with asterisks are displayed on the graphs. Full p-values for this are given in the lower panels (*p<0.05, p**<0.01, ***p<0.001).

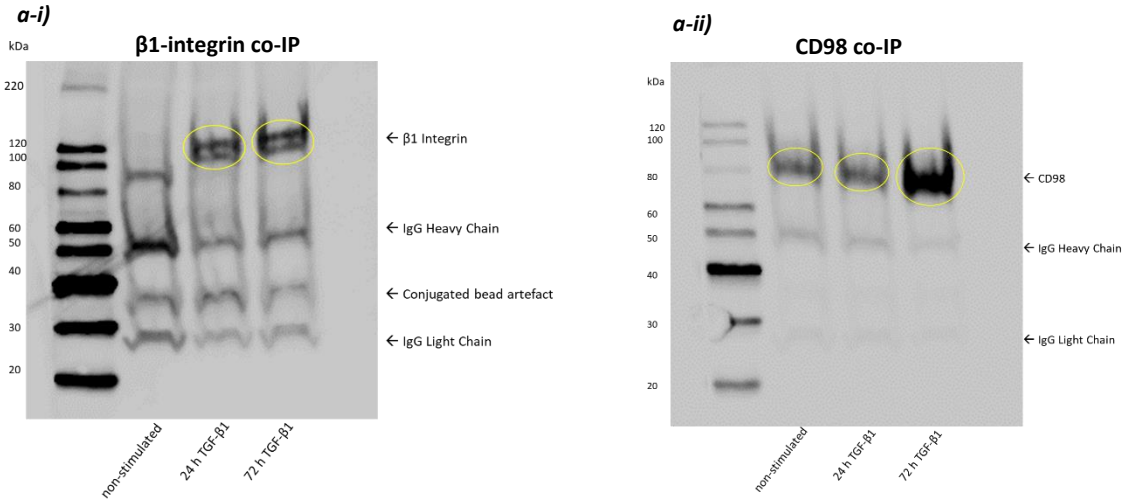
3.2.3.2 Co-Localisation Addressed by Co-Immunoprecipitation (co-IP)

I next addressed co-localisation of galectin-3 with β1-integrin and CD98hc by co-immunoprecipitation analysis. Galectin-3 or CD98hc were immunoprecipitated from A549 whole cell lysates in the absence or presence of TGF-β1 stimulation, using the appropriate monoclonal antibodies (table 2.13). Consistent results were observed with triplicate experiments. TGF-β1 stimulation increased the co-immunoprecipitation of both CD98hc and β1-integrin with galectin-3 pulldown (Figures 3.9a-i and 3.9a-ii). Reciprocally galectin-3 and β1-integrin co-immunoprecipitated increasingly with CD98hc pulldown following TGF-β1 stimulation when compared to relevant non-stimulated control (Figures 3.9b-i and 3.9b-ii).

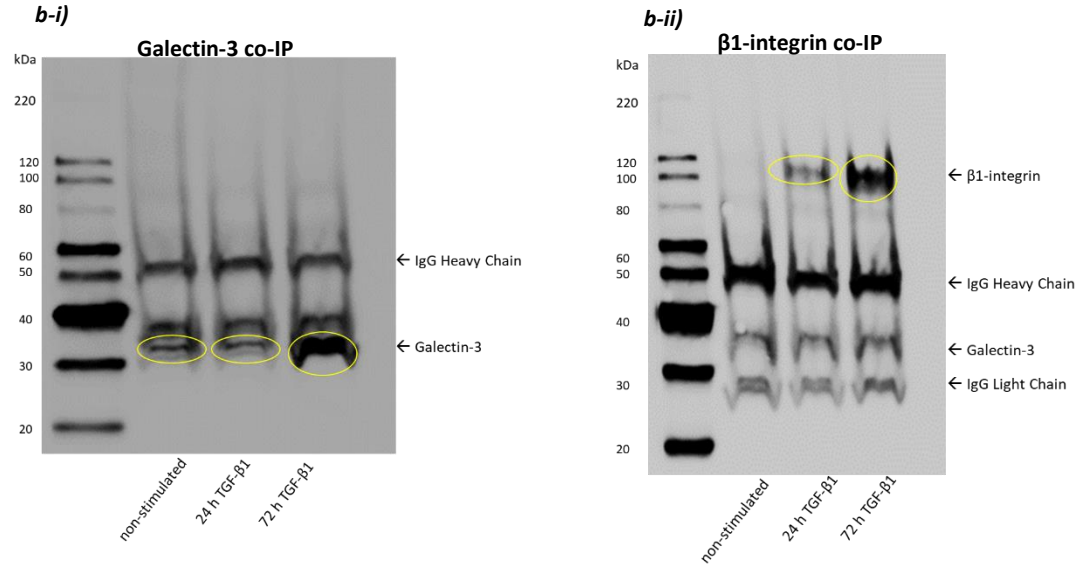
In addition, I conducted preliminary experiments to assess the co-localisation of galectin-3 with β1-integrin on the membrane and cytosolic fractionated lysates in non-stimulated (basal) conditions (Figure 3.9c). This was done using the anti-galectin-3 mAb for pulldown as above, followed by probing for β1-integrin. Since β1-integrin is expected to be exclusively a membrane localised protein (shown in Figure 3.3), this initial result was very encouraging as it suggests that β1 integrin co-immunoprecipitated with galectin-3 in the membrane but not the cytosolic fraction (Figure 3.9c).

Anti-galectin-3 pulldown in both whole cell (Figure 3.9a) and membrane-cytosol fractionated (Figure 3.9c) lysates in all assessed conditions displayed a band at ~31 kDa when probing for β1-integrin (this molecular weight coincidently corresponds to monomeric galectin-3 mass). This has been labelled as conjugated bead artefact. Galectin-3 molecular weight bands were observed in the absence of probing directly for galectin-3 using its appropriate and specific mAb antibody with its corresponding rat-HRP-secondary antibody to display such bands. This suggests that the initial galectin-3 pulldown antibody contains overlapping parts to the primary mAb of β1-integrin, or the secondary mouse-HRP-antibody in some manner can cross-bind with remnants of the rat-galectin-3 species that will most likely remain present in the eluted loaded sample.

Galectin-3 pulldown



CD98 pulldown



Galectin-3 pulldown on fractionated lysates

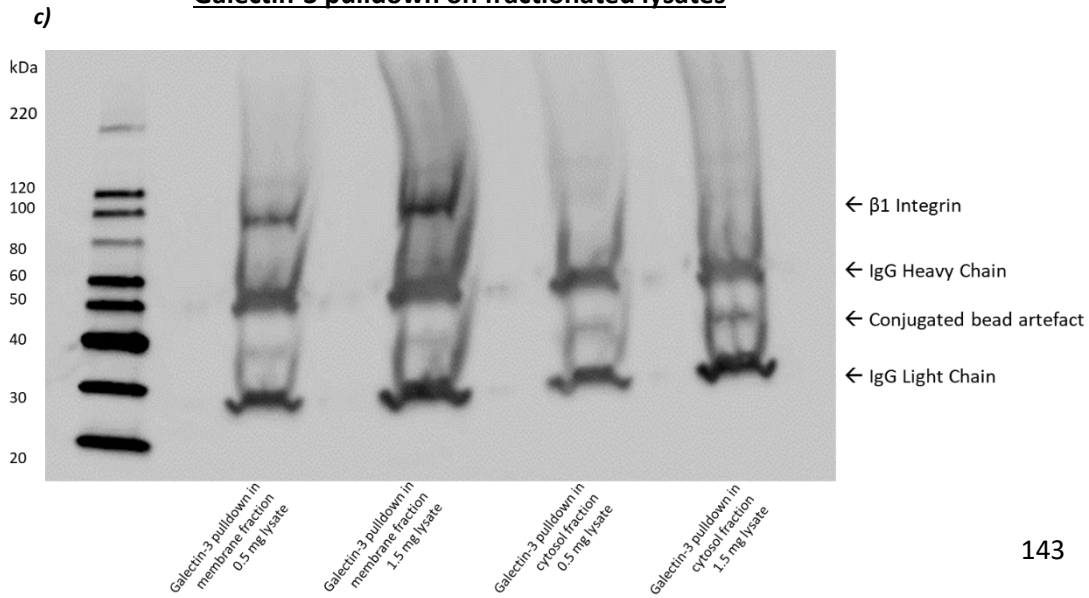


Figure 3.9, Monoclonal antibody pulldown from whole cell and membrane-cytosol fractions of A549 lysed cells. All samples were loaded on a 4-20% (v/v) SDS-PAGE to analyse material co-immunoprecipitated from the lysates. ***a-b) Whole cell lysate co-IP,*** was analysed using antibodies immunoprecipitating against galectin-3 (***a***) and CD98 (***b***), probed by western blotting for proteins as indicated above the blots. The samples indicated in blots ***a-b*** were loaded in identical order**:** Lane 2 represents non-stimulated cell lysates, Lanes 3 and 4 represents lysates from cells stimulated with 10 ng/ml TGF-β1 for 24 and 72 h respectively. All probed proteins displayed bands at the expected molecular weight of β1-integrin (~130 kDa), CD98 (~85 kDa) and galectin-3 (~31 kDa). ***c) Initial membrane and cytosol lysate fraction immunoprecipitation of galectin-3.*** Lanes 2 and 3 show anti-galectin-3 immunoprecipitation of the membrane fraction lysate loaded at 0.5 mg and 1.5 mg total lysate respectively, probed for β1-integrin co-IP (bands observed appropriately at the correct molecular weight at ~130 kDa). Lanes 4 and 5 represents anti-galectin-3 pulldown of the cytosolic fraction lysate at 0.5 mg and 1.5 mg total lysate respectively. No bands were observed on probing for β1-integrin in the cytosol fraction, supporting effective fractionation and membrane-restricted expression. In all blots the yellow ring highlights the bands of the co-IP protein. The bands at ~50 kDa and ~30 kDa in all blots represent the immunoglobulin heavy and light chains from the pull-down antibody that are identified by the secondary blot antibody. The marker used on all blots was MagicMarkXP (in Lane 1). All blots were immersed with EZ-ECL chemiluminescent detection kit and images were taken at an exposure time ranging from 2-10 minutes. (n=3).

3.2.3.3 Proximity Ligation Assay of Galectin-3 with CD98 and β1-integrin in Ex Vivo Lung Tissue Model of Early Fibrosis

We assessed whether the co-localisation of galectin-3 with CD98 and β1-integrin was also supported by higher resolution findings in the *ex vivo* model of early lung fibrosis (Roach *et al*, 2018). I worked with Dr Panayiota Stylianou (PDRA, Gooptu group, University of Leicester), to perform proximity ligation assays (PLA) on TGF-β1 pre-treated non-fibrotic control lung tissue.

In these studies, a PLA kit (Duolink® PLA Technology, Sigma Aldrich, UK) was used to define co-localisation between the two protein pairs of galectin-3 with CD98 and β1-integrin. These were probed by their respective monoclonal antibodies raised in two separate species. From the PLA kit, secondary antibodies coupled with oligonucleotides (PLA probes) bind to the primary antibodies. If the probed proteins are in a distance of ≤40 nm then connector oligos are able to link the PLA probes by ligation. This circular closed DNA amplifies in presence of DNA polymerase, and detection oligos coupled to fluorochromes can then hybridise to the repeating sequences in the amplicons. Co-localisation can then be identified as fluorescently labelled red spots under immunofluorescence microscope.

Our data indicated that galectin-3 with CD98 (Figure 3.10a) or β1-integrin (Figure 3.10b) are found within 40 nm of each other. The levels of these co-localising pairs increased with TGF-β1 pre-treatment.

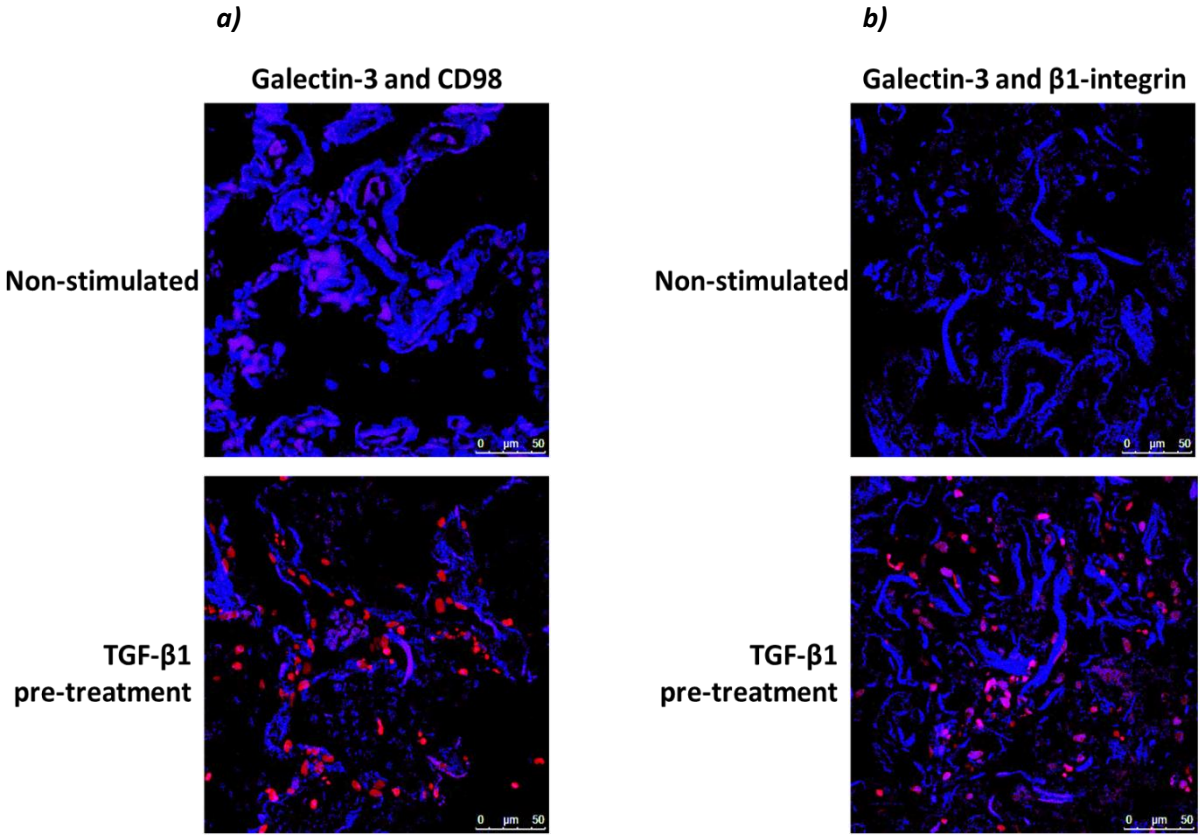


***Figure 3.10, Proximity ligation assay staining focusing on the co-localisation of galectin-3 with CD98hc (a) and β1-integrin (b) on pre-treated TGF-β1 NFC ex vivo lung tissue defining early fibrosis*.** PLA signals were assessed in non-stimulated or 7-day 10 ng/ml TGF-β1 pre-treated NFC lung tissue sections. Dual protein of interest were probed with with 4 µg/ml galectin-3 and- 5µg/ml CD98hc or 3 µg/ml β1-integrin (dually probed proteins antibodies were raised in different species), and stained fluorescently with the Duolink® PLA kit. PLA defines co-localisation of protein pairs to ≤40 nm resolution which is depicted as red spots. The nuclei were stained with DAPI. Images were captured using the Leica TCS SP5 confocal microscope.

3.2.4 Functional Cross-Talk Between CD98 and an Epithelial Readout of Pro-Fibrotic Behaviour

Our group has shown that the putative galectin-3 ligand CD98 mediates pro-inflammatory cytokine responses to acute injury stimuli (Wang *et al*. *Manuscript in preparation*). The co-localisation data supports close physical interactions between CD98 and galectin-3, consistent with the gal-3-fibrosome hypothesis, such that CD98 may be implicated at the junction of pro-inflammatory and pro-fibrotic behaviour.

IPF is characterised as a disease of exaggerated wound healing responses. Scratch tests of epithelial wound healing have been used to demonstrate the role of galectin-3 in this response (Mackinnon *et al*. 2012) in the context of tissue evidence of its role in mediating lung fibrosis *in vivo*. To test whether CD98 was similarly required I therefore undertook a wound healing assay without and with selective silencing of this protein.

The scratch testing assay itself was set up to establish the degree of wound closure within 24 h after the wound was established. Prior to that the cells were grown to confluency (48-72 h). CD98 silencing was conducted at the outset. I therefore first assessed the duration of CD98 silencing by siRNA-mediated knockdown. Timepoints from 24-120 h were studied (Figure 3.11, data from 48-72 h). The findings demonstrated strong basal expression of CD98 (Figure 3.11) and consistent, near-total knockdown with CD98 siRNA relative to relevant controls over the entire time course.

This confirmed that the siRNA approach used in this study could efficiently silence CD98hc and that the CD98hc mAb used in this study for multiple techniques (table 2.13) was specific to CD98hc. Equal sample concentration loading was verified by immunoblotting for β-actin (Figure 3.11).

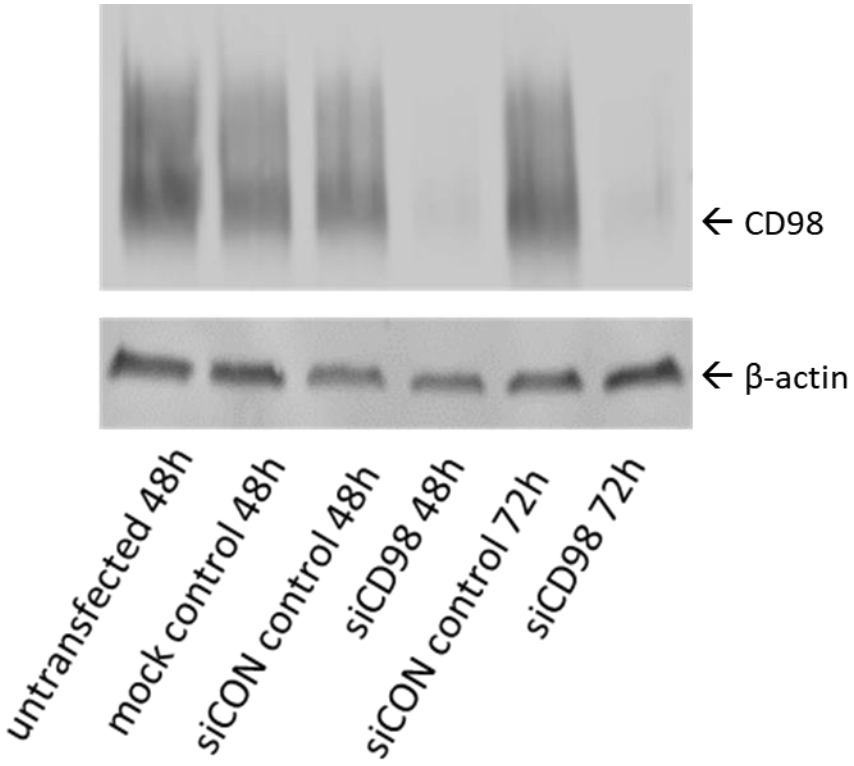
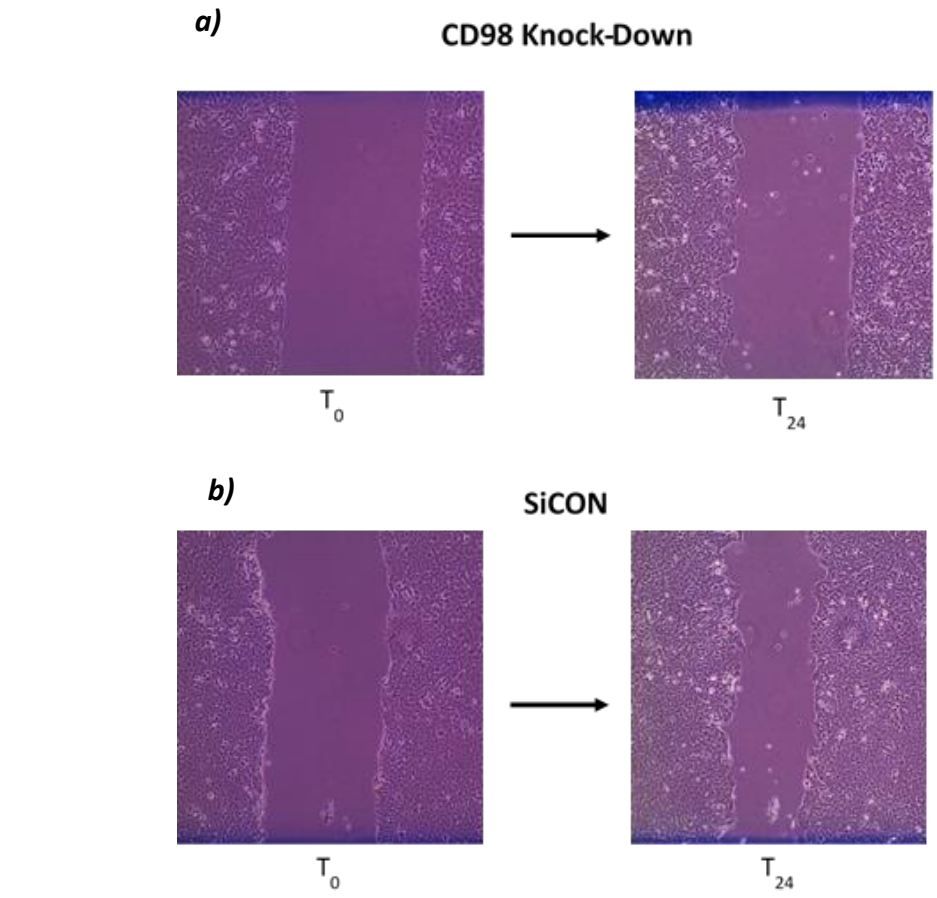


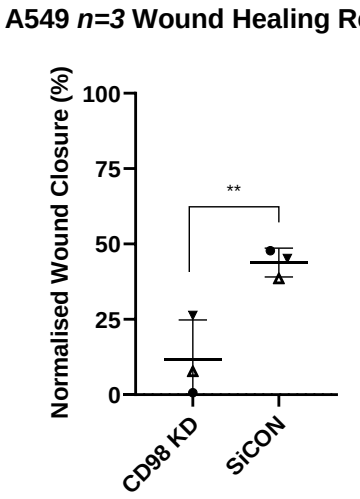
Figure 3.11, Western-blot monitoring the expression levels of basal CD98 following siRNA treatment in A549 cells against relevant controls. Exponentially growing A549 cells was used. Whole cell lysate was prepared by using RIPA Lysis Buffer. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml CD98 and 0.04 µg/ml β-actin. Lane 1-6 has been labelled with the appropriate conditions (on the X-axis), with siCON denoting siRNA scramble control and siCD98 denoting siRNA for CD98. The Y-axis has been labelled with the respective protein being represented. The blots were immersed with EZ-ECL chemiluminescent detection kit and the images were taken at exposure time of 2 minute.

3.2.4.1 CD98 Mediates Wound Healing Responses in A549 Cells

Three individual experimental repeats displayed consistent results demonstrating that when CD98hc is knocked down in A549 cells there is an abrogation of the wound healing response. This is illustrated in Figure 3.12a. Quantification of the wound closure area at 24 h post-scratch confirmed a statistically significant difference between cells treated with siRNA targeting CD98hc expression and siRNA control (Figure 3.12b). Wound closure values were normalised to 10% (v/v) FBS treated positive control media (Figure 3.12c). This indicates that CD98hc is required for wound-healing responses in A549 cells, with potential relevance to alveolar epithelial cell behaviour in IPF.



c)



	PAIRED T TEST
CD98 KD vs SiCON	p=0.0024

Figure 3.12, Scratch test in CD98 silenced A549 cells and the quantification of the wound healing response. ***3.12a,*** CD98 depleted A549 cells and ***3.12b,*** siRNA control (siCON): T_0 image represents the initial scratch, T_{24} image represents the same location of scratch post- 24 h of cell culture. ***3.12c***, The X-axis represents the different conditions: normalised to the positive control (DMEM + FBS which reaches 100% wound repair). The Y-axis represents the wound closure as normalised wound closure area (%). The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table to the right of the scatterplot represents the full p value of each protein expressed as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

3.2.5 Flow Cytometry to Characterise Sub-Cellular Localisation of CD98hc in A549 cells

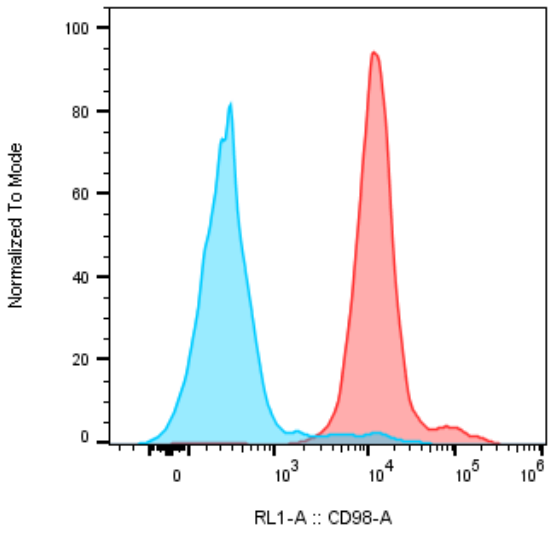
My previous experiments demonstrated that CD98hc is predominantly membrane-restricted (Figure 3.3). To assess if the expression of CD98hc is restricted to the cell surface membrane in my alveolar epithelial cell model I used flow-cytometry with the assistance of PhD student Michael Biddle. I focused on CD98 for demonstrating cell surface distribution since it is less widely expressed than β1-integrin and galectin-3, but closely associated with each.

I compared the protein expression levels of CD98hc in permeabilised against non-permeabilised cells. The CD98-E5 antibody (Santa-Cruz Biotechnology, U.S.A) that I used for western blotting, co-immunofluorescence and (co)-immunohistochemistry analysis did not display any signal when used for flow cytometry, irrespective of antibody concentration. The monoclonal antibody CD98-44D7 (Abcam, Cambridge, UK) reported robust and consistent signal in flow-cytometry and so this was used for flow-cytometry.

Cells were assessed without or following treatment with the absence and presence of 4% PFA with 0.1% (w/v) saponin for fixation and permeabilisation. Triplicate repeats displayed consistent results. Histograms of CD98hc signal distribution was very similar between permeabilised (Figure 3.13a) and non-permeabilised cells (Figure 3.13b) (pink peak at ~Log 10^4). As expected, the isotype control in both conditions show almost zero signal (blue peak). Further, once the mean value of CD98hc signal was quantified (subtracted from either the isotype control (mouse IgG) (Figure 3.14a) or fluorescence minus one (FMO) control) (Figure 3.14b)), there were no statistically significant differences in the signal intensity of both conditions. Taken together these data confirmed that CD98hc protein expression is restricted to the cell surface membrane in A549 cells. A549 cellular gating strategy displayed as dot plots is provided in Appendix (Figure A6).

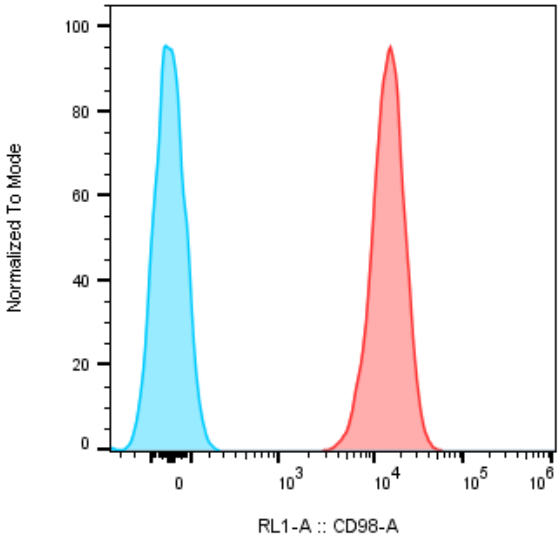
a)

Permeabilised



b)

Non-Permeabilised



c)

Superimposed histogram of Permeabilised vs Non-Permeabilised

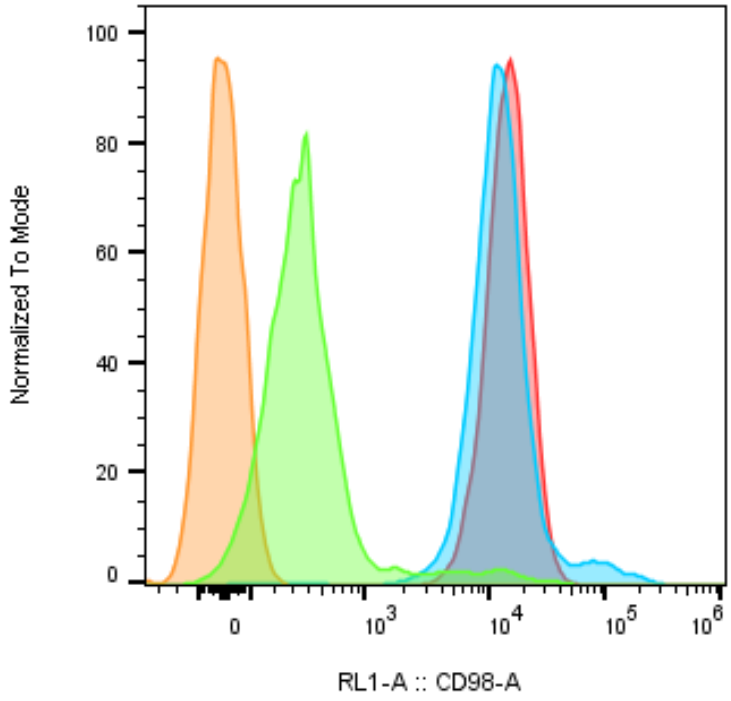
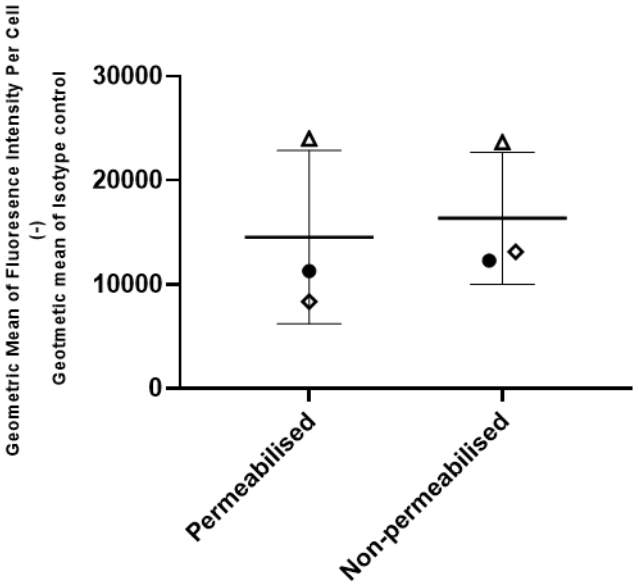


Figure 3.13, Flow cytometry FloJo analysed histograms of CD98 staining monitoring signal of CD98hc in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) A549 cells. Figure 3.13a (permeabilised) and 3.13b (non-permeabilised): pink peak represents CD98hc staining, and the blue peak represents the isotype control staining (Mouse IgG). ***3.13c***, overlay of Figures *3.13a* and *3.13b*; the pink and blue peak represents CD98hc staining. Isotype control is represented as green peak for permeabilised and orange peak for non-permeabilised cells. Peak signal (Y-axis) on the histograms represent arbitrary normalised modal value; the X axis represents fluorescence intensity (in the Red Laser 1 (RL1) (the secondary antibody used is APC), which fluoresces in the red channel) given as a logarithmic scale. n=3

a)

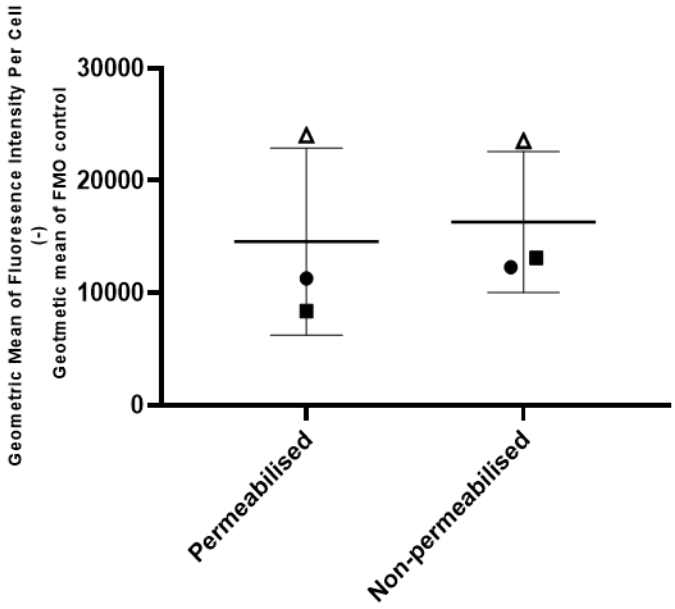
Effect of CD98 expression levels on permeabilised VS non-permeabilised cells



	PAIRED T TEST
Permeabilised vs non-permeabilised	p=0.3620

b)

Effect of CD98 expression levels on permeabilised VS non-permeabilised cells



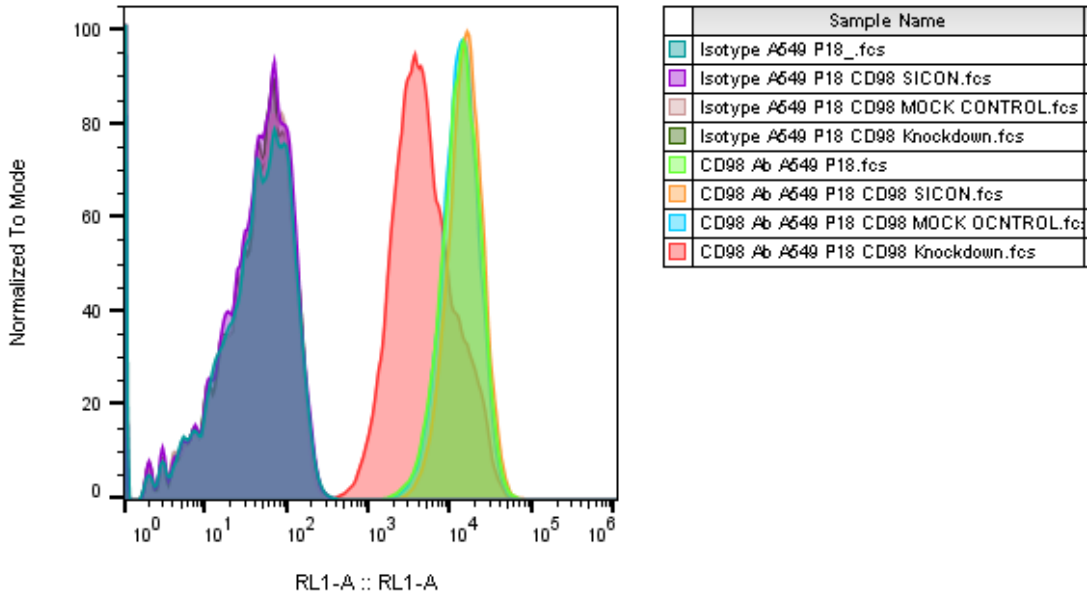
	PAIRED T TEST
Permeabilised vs non-permeabilised	p=0.3779

Figure 3.14, Statistical quantification of CD98hc signal analysed from FloJo in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) A549 cells. The graphs represents data as geometric mean normalised to either isotype control (3.14a) or to FMO control (3.14b). The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The table to the right of the scatterplots represent the full p value of each protein expressed as paired t-test.

3.2.5.1 Confirming the Specificity of the CD98-44D7 Antibody Used in Flow Cytometry

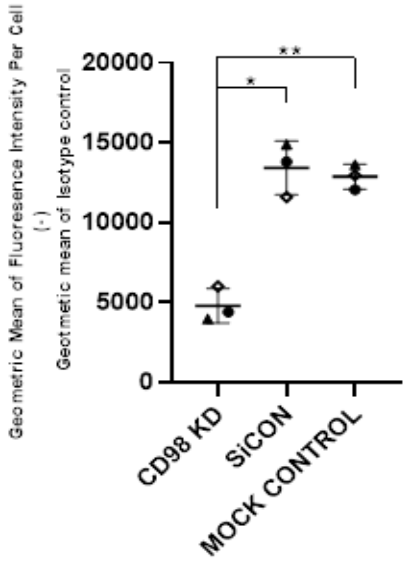
For all previous experimental context in my study, the antibody CD98-E5 from SantaCruz Biotechnology has worked well with high specificity for CD98hc. However, when used in a flow-cytometry context this did not report any signal. CD98-44D7 antibody from abcam displayed strong signalling in flow-cytometry. Since expression levels of CD98 are the same between permeabilised and non-permeabilised cells, the subsequent experiment was conducted in non-permeabilised cells. Signal levels of CD98 was reduced upon CD98 knockdown (pink peak signal compared to all relevant control peak signals, Figures 3.15a-b) confirming the specificity of the CD98-44D7 mAb, used in this experimental technique.

a)



b)

CD98 KD in A549 cell-Antibody Specificity (non permeablised)



	PAIRED T TEST	ONE WAY ANOVA
CD98 KD vs SiCON	p=0.0322	p=0.0003
CD98 KD vs Mock Control	p=0.0099	p=0.0004
SiCON vs Mock Control	p=0.6276	P=0.8151

Figure 3.15, Confirming the specificity of CD98-44D7 antibody used for flow-cytometry in A549 cells. 3.15a, FloJo analysed histogram of CD98 staining in non-permeabilised cells, The teal, purple, grey and dark green peaks represent the isotype control (mouse IgG) staining. The green, orange, blue and pink peaks represent CD98 staining. A key-table colour which describes each condition has been given (top-right hand corner). On the histograms, the Y-axis represents arbitrary normalised modal value; the X axis represents fluorescence intensity (in the Red Laser 1 (RL1) (the secondary antibody used is APC), which fluoresces in the red channel) given as a logarithmic scale. ***3.15b, Quantification of CD98 knockdown vs relevant controls in non-permeabilised cells.*** The graph is presented as a scatter-plot. The X-axis represents the different conditions The Y-axis represents the fluorescence intensities geometric mean normalised to isotype control. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table to the right of the scatterplot represents the full p value of protein expression as multiple paired t-test and one-way ANOVA (*p<0.05, p**<0.01, ***p<0.001).

3.3 Discussion

Understanding the molecular pathways and the proteins involved in fibrosis is important for improving our approaches to anti-fibrotic therapies. Due to galectin-3's binding capability we hypothesised that galectin-3 nucleates a pro-fibrotic macromolecular assembly at the cell surface. We termed this as 'galectin-3-fibrosome' with galectin-3 spatiotemporally clustering integrin dependent pathways with TGF-β1 mediated signalling. Experimental data suggest the involvement of at least four components, namely galectin-3, CD98hc, β1-integrin and TGF-βRII (Estrach *et al*. 2014, Hashiba *et al*. 2014).

With Dr Panayiota Stylianou I initially studied the co-localisation of galectin-3 with other putative gal-3-fibrosome components at the whole cell level, looking at *ex vivo* IPF and non-fibrotic control human lung tissue. The findings indicated that staining for all these proteins increased in IPF tissue relative to control, with evidence of higher co-expression in alveolar epithelial cells relative to mesenchymal cells. I therefore next studied co-localisation of these proteins in the A549 lung epithelial cell model in the absence and presence of the pro-fibrotic cytokine TGF-β1.

Fractionation studies confirmed that CD98hc and β1-integrin were found predominantly in a membrane-restricted distribution. Densitometric analysis of western blot data, and immunofluorescence studies demonstrated that galectin-3, CD98 and β1-integrin levels increased with TGF-β1 stimulation in the membrane fractions.

Unsurprisingly, in the fractionation studies β1-integrin and CD98hc levels increased in the membrane fraction in a similar manner. This is consistent with reports repetitively stating CD98hc is constitutively coupled with β1-integrin as a heterodimer (Feral *et al*. 2005). Furthermore, data from our group has shown the importance of CD98hc:β1-integrin mediating acute exacerbation (Wang *et al*. *Manuscript in preparation*). Flow cytometry studies demonstrated that CD98hc was

entirely expressed at the cell surface membrane only, in agreement with existing reports stating CD98hc expression and distribution on the cell membrane (Kim *et al*. 2008, Ballina *et al*. 2016). Co-immunostaining confocal microscopy on A549 cells along with PLA studies on NFC lung tissue confirmed increased co-localisation behaviour of galectin-3 with CD98 and β1-integrin on epithelial cells following TGF-β1 stimulation.

Since PLA studies define co-localisation to ≤ 40 nm, these are planned to be conducted on IPF lung tissue as of when they become readily available. This will allow a stronger resolution assessment of basal co-localisation levels of gal-3-fibrosome proteins in IPF vs NFC lung tissue, and aid with a better biological understanding of such physical interactions.

Co-localisation of these proteins was also supported by immunoprecipitation experiments on A549 whole cell lysates subjected to TGF-β1 stimulation. These immunoprecipitation studies are a starting point for initiating mass spectrometry studies, which could allow us to identify additional components forming the gal-3-fibrosome complex.

Using these optimised immunoprecipitation conditions, I attempted anti-galectin-3 immunoprecipitation on basal non-stimulated fractionated lysates. This initial finding confirmed that galectin-3 and β1-integrin interactions are cell membrane restricted. This initial data has led to further development within the group, and the experiment has been continued on by PDRA Dr Panayiota Stylianou (Gooptu Group, University of Leicester). Her data shows galectin-3 pulldown co-immunoprecipitates with CD98hc on the membrane fraction (restricted). Reciprocated pulldown with CD98hc confirms both galectin-3 and β1-integrin co-immunoprecipitated (on the membrane fraction only) increasingly following TGF-β1 stimulation, in agreement with all my experimental data.

Co-localisation between galectin-3 with CD98hc and β1-integrin (which increases in a pro-fibrotic context) is therefore observed repeatedly by multiple experimental readouts and occurs at epithelial cell surface. The findings are consistent with the gal-3-fibrosome hypothesis and suggest the following: the galectin-3 C-terminal domain interacts with a branched glycan motif (β-1,6-N-acytyl-D-glycosamine) of the glycoprotein β1-integrin and its complexed binding partner CD98hc (Feral *et al*. 2005). During TGF-β1 mediated fibrosis, there is an increased production of β1-integrin:CD98hc. Since β1-integrin has the capacity to bind to the RGD tripeptide sequence, this will predictably concentrate latent TGF-β1 at the epithelial cell surface, close to TGF-βRII (Munger and Sheppard. 2011). This can be activated by stretch and/or αvβ6-integrin or any β3-integrin bound to CD98. This will set up a positive feedback loop in which TGF-β1 stimulation causes an upregulation of β1-integrin:CD98hc which can increase local TGF-β1 levels near TGF-βRII. This increased TGF-β1 retention to TGF-βRII could be probed indirectly by measuring downstream canonical TGF-β1 signalling pathway (i.e. measuring SMAD phosphorylation). However, indirect assessment of TGF-β1 has limitations since non-canonical signalling pathways can also be activated, thereby potentially confounding the overall amount TGF-β1 being measured as inaccurate. The increased levels of galectin-3 in close proximity will potentiate the ligand binding effects by extending its duration by bridging both β1-integrin and TGF-βRII through its CRD interactions. Unfortunately, I was unable to assess the involvement of TGF-βRII stabilisation within a gal-3-fibrosome directly, due to lack of an effective anti-TGF-βRII mAb for detection of endogenous protein.

TGF-β1 induces the expression of α-SMA and collagen (Hu *et al*. 2012, Mackinnon *et al.* 2012). TGF-β1 is both an initiator and a driver of tissue stiffness since it stimulates the production of collagen and other matrix proteins via its action on nearby myofibroblasts (MacKinnon et al, 2012). These will promote integrin-dependent latent TGF-β1 activation. This could be a two-fold hit, where TGF-β1 stimulation causes pro-fibrotic signalling via the TGF-βRII signalling pathway, and the increased collagen expression (extracellular matrix) can then interact with β1-integrin:CD98hc to further amplify downstream pro-fibrotic signalling.

Taken together these findings support the general concept of a gal-3-fibrosome, nucleated by galectin-3, clustering a range of pro-inflammatory and pro-fibrotic elements including CD98hc and β1-integrin. Because of the close proximity of CD98hc to galectin-3, I performed simple scratch test studies that demonstrated that CD98 was required for epithelial wound healing responses similarly to galectin-3 (Mackinnon *et al*. 2012). In line with its importance in inflammation (Wang *et al*. *Manuscript in preparation*) CD98 is now seen to be functionally important in pro-fibrotic pathway. And so, CD98 may potentially bridge pro-inflammatory and pro-fibrotic responses *in vivo,* thereby facilitating IPF pathogenesis.

Chapter 4

Co-localisation of Galectin-3 with CD98 and β1-integrin in IPF and Non-Fibrotic Control Human Lung Myofibroblasts

4.1 Introduction

Human lung myofibroblast (HLMF) cells were initially described as modified fibroblast cells with increased contractility features causing increased secretion of collagen to form fibres within the extracellular space (Gabbiani *et al*. 1971). Myofibroblast cells are now understood to play a key role in IPF pathogenesis (Moore and Herzog. 2012, Kendall and Bostwick *et al*. 2014). They mediate tissue contraction and exaggerated extracellular matrix deposition in IPF (Zhang *et al*. 1994), leading to lung stiffness, reduced gas transfer, and ultimately end stage respiratory failure (McKleroy *et al*. 2013).

In vivo, myofibroblasts may arise from multiple sources including residential parenchymal fibroblast precursor cells, lung epithelial cells (type II alveolar epithelial cells) via the process of EMT in response to TGF-β1 stimulation, and from circulating fibrocytes recruited to fibrotic sites (Abu El-Asrar *et al*. 2008, Doerner and Zuraw. 2009, Zhou *et al*. 2011). The relative contribution of each of these mechanisms for IPF *in vivo* remains to be elucidated (Gabasa *et al*. 2017).

Galectin-3 regulates myofibroblast activation and proliferation resulting in increased extracellular matrix production (Henderson *et al*. 2006, Taniguchi *et al*. 2012). Myofibroblast cells derived from galectin-3 null mice display reduced extracellular matrix (collagen) expression under TGF-β1 stimulation (MacKinnon *et al*. 2012). Integrins similarly modulate myofibroblast differentiation (Lygoe *et al*. 2004). Gene expression analysis of myofibroblasts isolated from bleomycin-treated mouse lung demonstrates high expression of CD98 (Akamatsu *et al*. 2013). The integrin:CD98 heterodimer receptors can mediate both cell-cell and cell-matrix interactions. (Schnittert *et al*. 2018). Taken together, I hypothesised galectin-3 and CD98:integrin complex interact functionally to mediate the myofibroblast pro-fibrotic response and co-localise at the myofibroblast surface in a gal-3-fibrosome analogous to observations in A549 cells (Chapter 3).

I used *ex vivo* human lung myofibroblasts derived from non-fibrotic control (NFC) tissue, and IPF patient lung tissue (Table 4.1). Samples were obtained with informed consent and the work was conducted with ethical approval from the local research ethics committee. I isolated NFC-myofibroblasts from healthy lung resection margin tissue from patients undergoing surgery at Glenfield Hospital, University Hospitals Leicester, NHS Trust (Leicestershire, UK), for possible or confirmed cancer. IPF myofibroblasts were isolated from patients undergoing lung biopsy for diagnostic purposes at the University of South Carolina. Both sources of myofibroblasts were grown and cultured in identical conditions (2.1.2), using a well-established protocol developed by colleagues within the Department of Respiratory Sciences, University of Leicester (Roach *et al*. 2013).

Due to the structural similarities and overlap of myofibroblasts to fibroblasts and airway smooth muscle (ASM) cells, I characterised the myofibroblasts using several markers (discussed in greater depth in 4.2.1.2). These markers are reduced or absent in fibroblasts and ASM cells, and so as a panel they allow robust discrimination of myofibroblasts from these cells (Roach *et al*. 2013).

NFC- and IPF- myofibroblast cells were used and studied at low passages (P4-5) to maintain myofibroblast behaviour, in the absence or presence of TGF-β1 stimulation. I assessed the expression levels of galectin-3 and the putative fibrosome components, and its proximity (interactions) under pro-fibrotic conditions (TGF-β1 stimulation) primarily using western blotting and immunofluorescence confocal microscopy.

4.2 Results

4.2.1 Characterisation of Ex vivo Derived Human Lung Myofibroblast Phenotype

4.2.1.1 Patient Characteristics

Tissue samples from three NFC and three IPF patients were used as the source for purified HLMF cells. The patient clinical characteristics can be found in Table 4.1. Some of the demographic features are not typical of real world IPF such as having donors subjected to prednisolone treatment. However, such cases will be overrepresented in biopsy samples due to pre-test diagnostic uncertainties. Furthermore, biopsy-proven IPF is a frequent inclusion criterion for translation and clinical studies in IPF.

Noteworthy features of patient demographics in my study included reduced lung function of the IPF donors; the FEV_1 and FVC are both lower than the non-fibrotic controls. The IPF patients also displayed reduced efficiency of gas transfer; the diffusing lung capacity for carbon monoxide (DLCO) is below 50%.

Characteristic	IPF (n=3)	NFC (n=3)
Sex (no. of subjects)		
Male	2	3
Female	1	-
Age (years)		
Mean	63.9 ±0.7	76 ±2.52
Range	62-65	71-79
Length of symptoms (years)		
Mean	12.67 ±4.67	N/A
Range	4-20	N/A
Smoking (no. with >10 pack/years)	0	2
FEV_1 (mean % predicted ± SE)	62 ±9	73 ±12.52
FVC (mean % predicted ± SE)	51.5 ±2.5	89.07 ±8.08
DLCO (mean % predicted ± SE)	44 ±3	-
PA mean (mean % predicted ± SE)	22.5 ±1.5	-
Treatments		
Prednisolone	1	N/A
Cyclophosphamide	1	

Table 4.1, Clinical characteristics of NFC and IPF patients. This table demonstrates key clinical characteristics of the patient donors who provided their tissue for the culture of human lung myofibroblasts. FEV_1= Forced expiratory volume in first second. FVC= Forced vital capacity. DLCO= Diffusing capacity of carbon monoxide. PA= Pulmonary artery.

4.2.1.2 Characterisation of Primary Human Lung Myofibroblasts

NFC HLMFs were purified *ex vivo* from healthy tissue of resected lung. IPF-derived myofibroblasts were received at passage 2. Both NFC and IPF cultured myofibroblasts viewed under an inverted light microscope in the initial instance displayed the typical morphology of myofibroblast-like cells (spindle or stellate shape). Subsequently, NFC- and IPF- derived myofibroblasts were phenotypically characterised (at passages 2-4). Five positive and two negative biomarkers for myofibroblast phenotype were assessed by immunofluorescence. All HLMF preparations within my study successfully demonstrated immunostaining characteristics of myofibroblast cells (Figure 4.1-4.2).

Cells stained positively for a panel of myofibroblast markers (Figure 4.1) including α-smooth muscle actin (α-SMA) (Baum and Duffy. 2012, Moiseenko *et al*. 2017), collagen type I (Pan. 2008, Baum and Duffy. 2011, Bourke *et al*. 2011) and vimentin (Wang *et al*. 2006). These markers are lowly expressed in ASM cells, therefore additional markers were probed. Thy-1 is highly expressed in fibroblasts (Johnson *et al*. 2016), which stained strongly. In addition, fibroblast surface protein (FSP), as the name suggests, stains for fibroblasts, and indeed stained strongly in my cell cultures. However, FSP is also weakly expressed in macrophages (Kahounova *et al*. 2017, Moiseenko *et al*. 2017). Therefore, to rule out the presence of macrophages in my cell culture, I probed a macrophage/monocyte specific marker (CD68), this displayed absence of any staining (Figure 4.2). Since myofibroblasts cells are of mesenchymal origin, a specific hematopoietic stem cell marker (CD34) was probed, which stained negatively (Figure 4.2). However, a positive control was not assessed to probe for both CD68 and CD34, and therefore there is a slim chance this could suggest an antibody issue. The matched isotype controls stained negatively, confirming any positive staining to be antibody specific. These features were consistently observed in all purified cells, indicating high purity of the myofibroblast cell preparations.

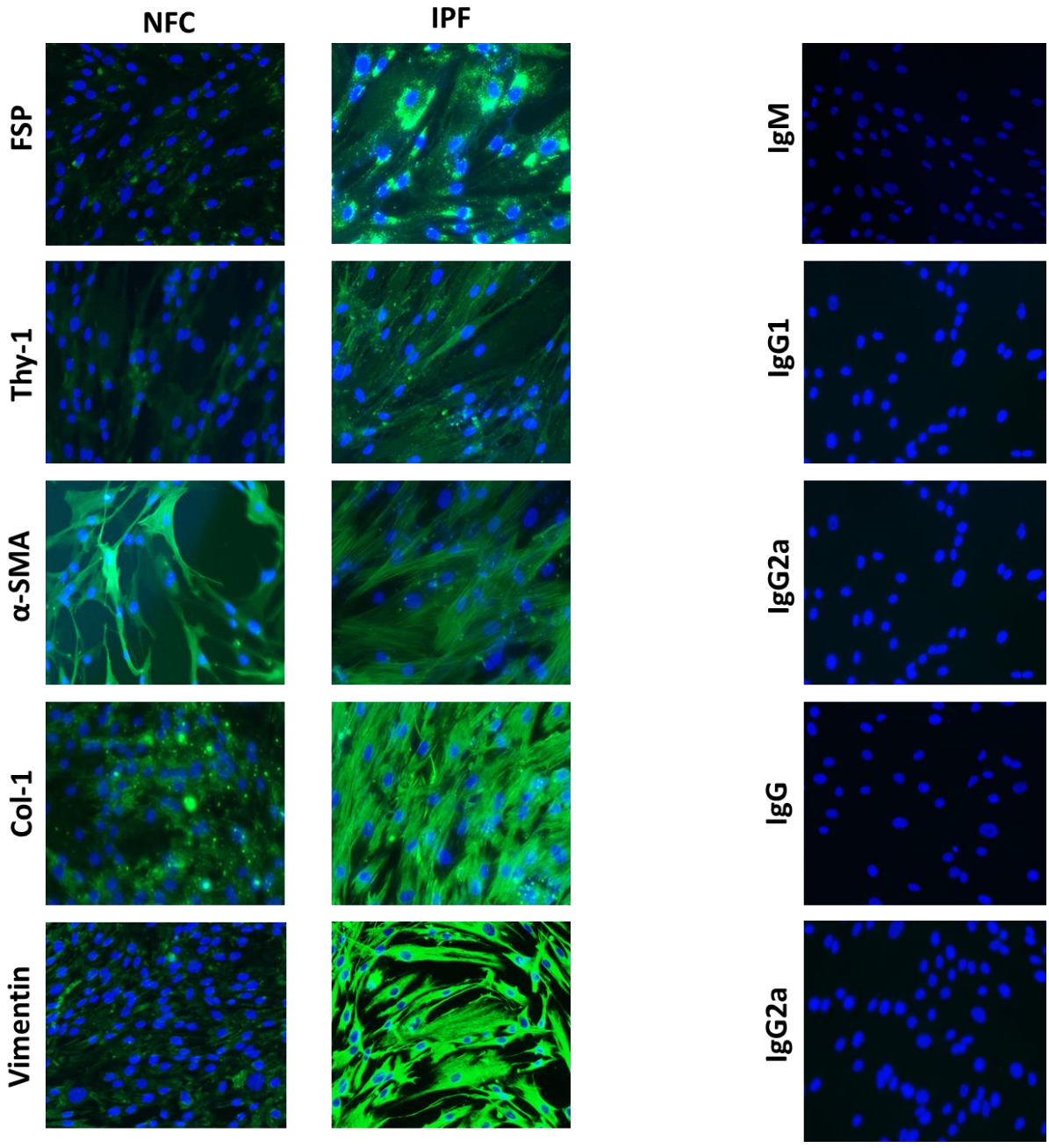


Figure 4.1, Characterising the human lung myofibroblast (HLMF) from both healthy and IPF donors. HLMF characterisation using five positive controls: a) Fibroblast surface protein (FSP) and the mouse isotype control IgM, ***b)*** Thy-1 and the mouse isotype control IgG1, ***c)*** α-SMA and the isotype control IgG2a, ***d)*** Collagen type I and the rat isotype control IgG, and ***e)*** Vimentin and the mouse isotype control IgG2a. Collagen type I was stained using secondary anti-rat FITC (green) and the rest of the panel were stained using secondary anti-mouse-FITC (green). Nuclei were stained using DAPI.

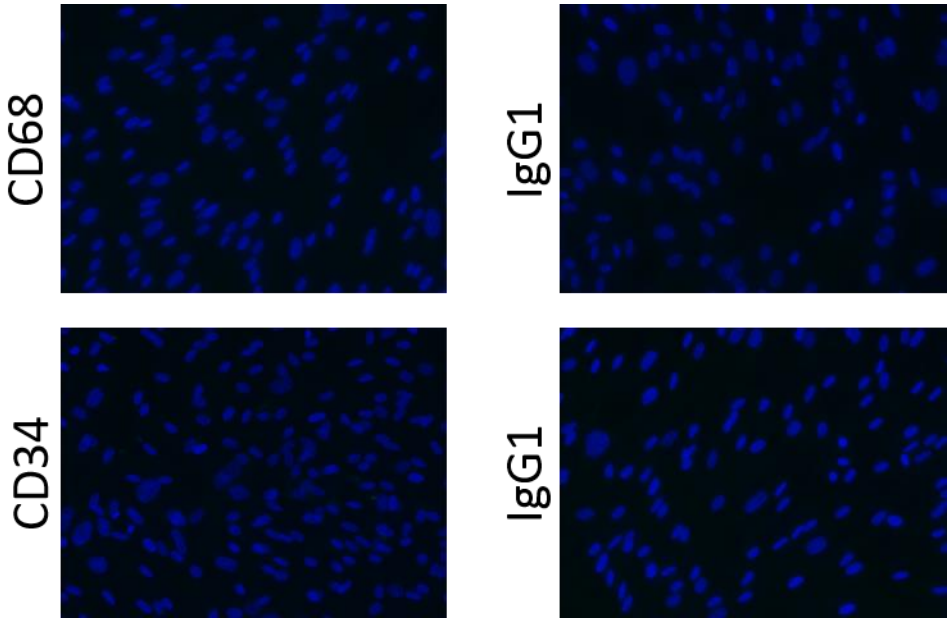


Figure 4.2, Characterising the human lung myofibroblast (HLMF) from healthy and IPF donors. HLMF characterisation using two negative controls: a) CD68 and the mouse isotype control IgG1 and ***b)*** CD34 and the mouse isotype control IgG1. The panel was stained using secondary anti-mouse-FITC (green). Nuclei were stained using DAPI.

4.2.2 Level of Putative Galectin-3-Fibrosome Components in Myofibroblasts

I examined the expression levels of the putative galectin-3-fibrosome components: galectin-3, CD98hc and β1-integrin in the lung mesenchymal cells of both NFC- and IPF- derived myofibroblasts. Expression of each protein was assessed by western blot analysis of SDS-PAGE of the membrane and cytosolic fractions of lysed cells, and by immunofluorescence confocal microscopy in the absence or presence of the pro-fibrotic cytokine TGF-β1 (10 ng/ml, 24 h and 72 h stimulation). In addition, levels and co-localisation of the putative gal-3-fibrosome components were assessed by co-immunostaining confocal microscopy.

4.2.2.1 Levels of Galectin-3, CD98hc and β1-integrin in NFC and IPF HLMF are Altered Following TGF-β1 Stimulation

NFC and IPF HLMF lysates were prepared and fractionated in the same way as A549 cells. They were assessed by SDS-PAGE and western blot using appropriate antibodies (Table 2.13). Bands of appropriate molecular weight were detected for galectin-3 (~31 kDa), CD98hc (~85 kDa) and β1-integrin (~130 kDa) (Figures 4.3 and 4.5). Membranes were probed for β-actin primarily as a cytosolic protein loading control and for the Na^+/K^+-ATPase subunit ATP1A1 as a membrane protein loading control. β-actin was detected in membrane as well as cytosolic fractions.

In both NFC (Figure 4.3) and IPF (Figure 4.5) myofibroblasts, β1-integrin and CD98hc were detected only in the cell membrane fraction. Their levels, normalised with reference to the ATP1A1 loading control, increased progressively when stimulated with TGF-β1 (comparing findings after 24 h and 72 h with non-stimulated controls). This was clearly observed after 72 h in NFC HLMFs (Figure 4.4) and after both 24 h and 72 h in IPF HLMFs (Figure 4.6). These findings were consistent with the membrane fraction distribution and changes in their expression levels

observed for A549 cells (Figure 3.3). Unsurprisingly the error bars within these datasets are quite spread, agreeing with donor variable response differences. Surprisingly, in both NFC and IPF myofibroblasts, the protein levels of galectin-3 in both the membrane and cytosolic fractions decreased progressively when stimulated with TGF-β1. This finding is the opposite of galectin-3 responses observed in A549 cells, indicating cell-specific variation in galectin-3 responses to TGF-β1 stimulation. Full uncropped western-blot scans of all probed proteins are provided in Appendix (Figures A7 and A8).

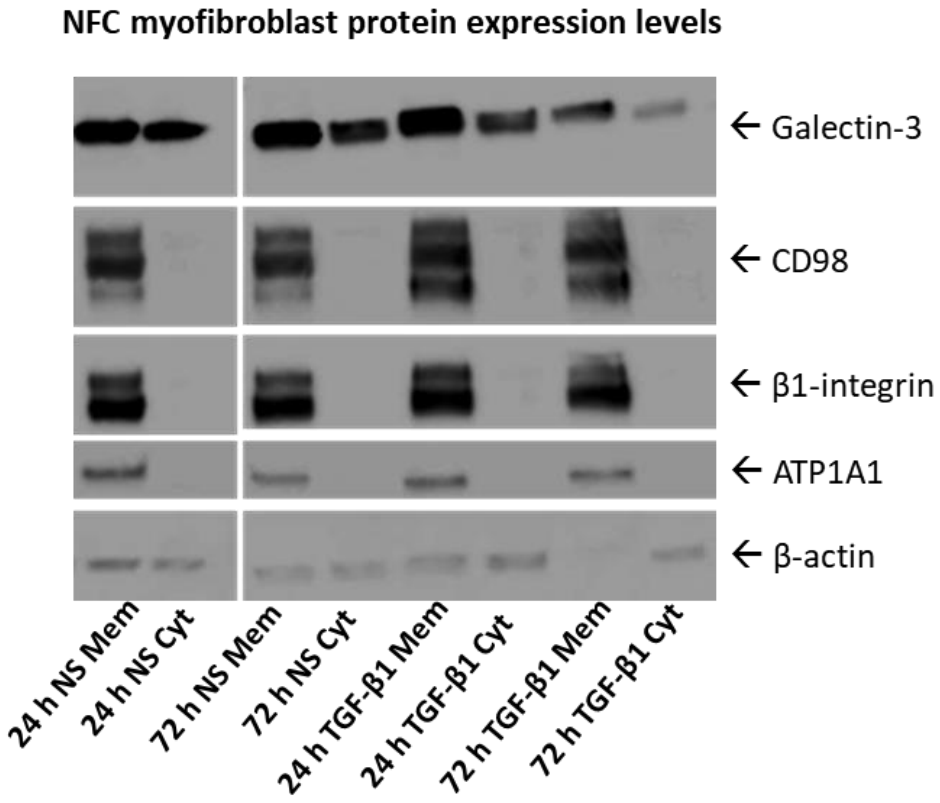
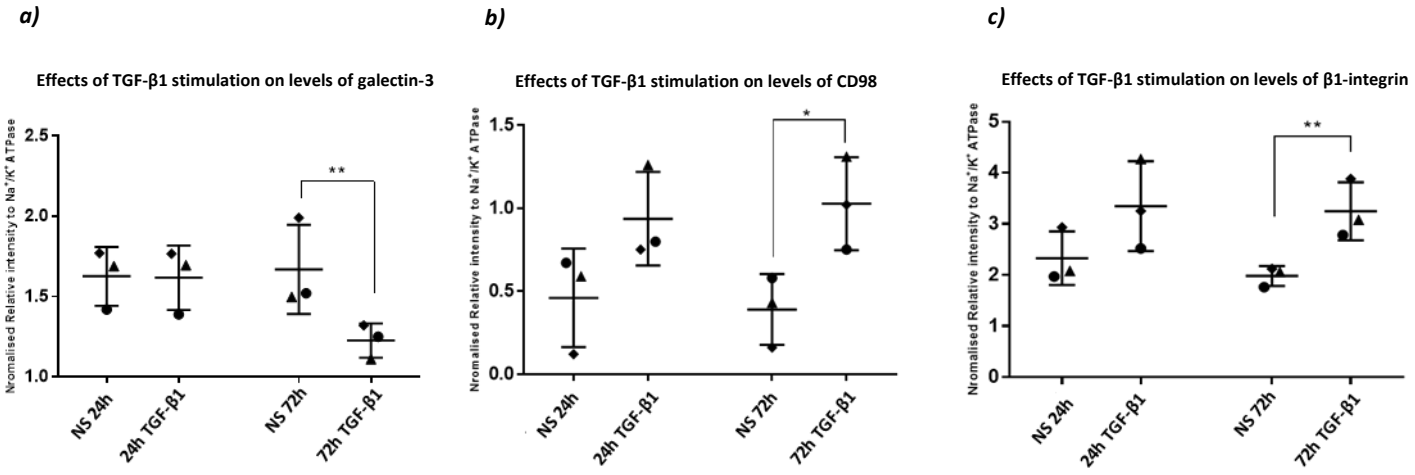


Figure 4.3, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in NFC myofibroblast cells by western-blotting. *Ex-vivo* HLMF cells were used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3, 0.005 µg/ml β1-integrin, 0.125 µg/ml CD98hc, 0.004 µg/ml β-actin and 0.5 µg/ml Na^+/K^+-ATPase (ATP1A1). Lane 1-10 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective protein being represented. The blot was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3)



		PAIRED T TEST
Galectin-3 normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.4573
	ns 72h vs 72h TGF-β1	p=0.0049

		PAIRED T TEST
CD98 normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.0663
	ns 72h vs 72h TGF-β1	p=0.0310

		PAIRED T TEST
β1-integrin normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.1723
	ns 72h vs 72h TGF-β1	p=0.0032

Figure 4.4, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin membrane levels from figure 4.3 immunoblotting. Quantification of the levels of galectin-3 ***(a),*** CD98 ***(b),*** and β1-integrin ***(c),*** was conducted from the western blot images using the software Image J and plotted on GraphPad Prism 7. The blot densitometric intensities were normalised against the membrane specific housekeeping protein ATP1A1 (Na^+/K^+ ATPase) and has been labelled on the Y-axis. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table below each scatterplot represents the full p value of each protein expressed as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

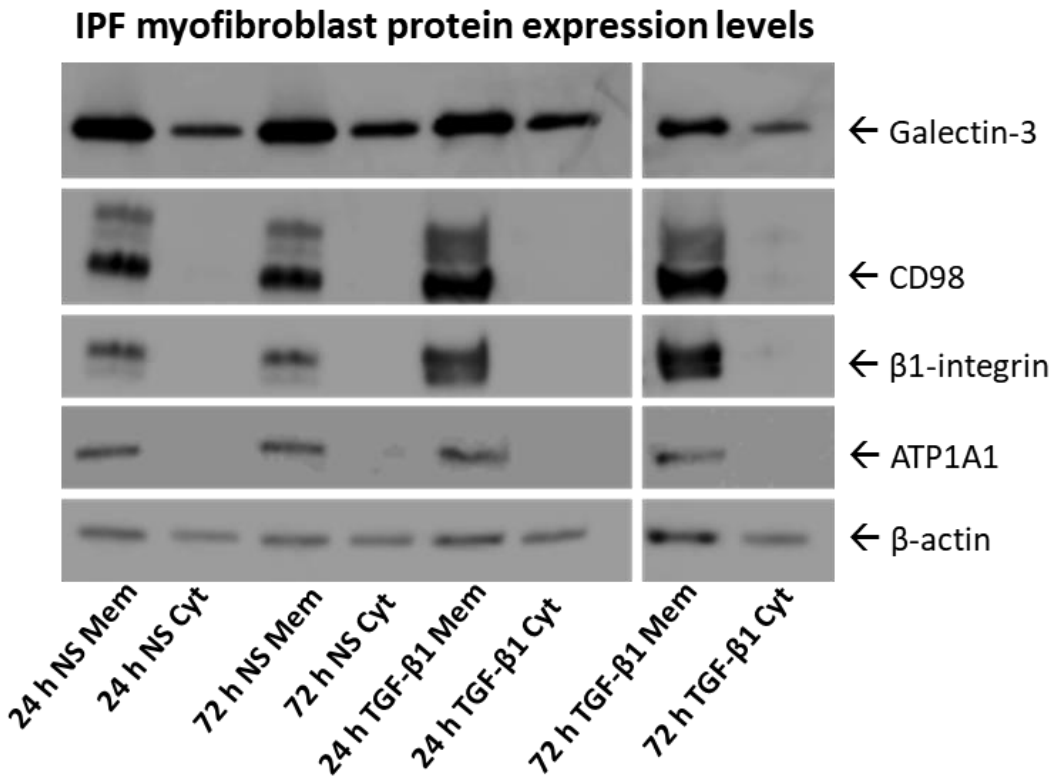
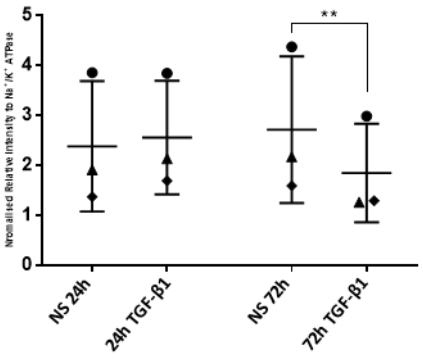


Figure 4.5, Monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in IPF myofibroblast cells by western-blotting. *Ex-vivo* HLMF cells were used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3, 0.005 µg/ml β1-integrin, 0.125 µg/ml CD98hc, 0.004 µg/ml β-actin and 0.5 µg/ml Na^+/K^+-ATPase (ATP1A1). Lane 1-10 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective protein being represented. The blot was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3)

a)

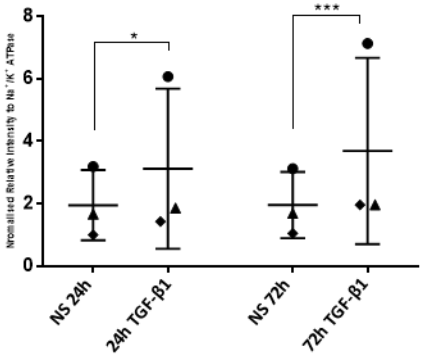
Effects of TGF-β1 stimulation on levels of galectin-3



		PAIRED T TEST
Galectin-3 normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.2505
	ns 72h vs 72h TGF-β1	p=0.0098

b)

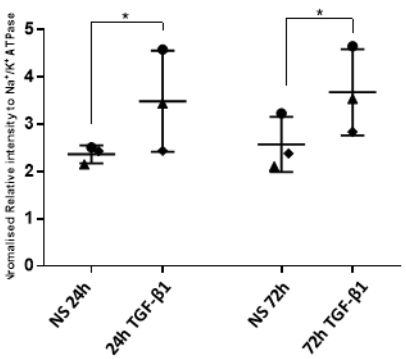
Effects of TGF-β1 stimulation on levels of CD98



		PAIRED T TEST
CD98hc normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.0231
	ns 72h vs 72h TGF-β1	p=0.0004

c)

Effects of TGF-β1 stimulation on levels of β1-integrin



		PAIRED T TEST
β1-integrin normalised to Na^+/K^+ ATPase	ns 24h vs 24h TGF-β1	p=0.0201
	ns 72h vs 72h TGF-β1	p=0.0179

Figure 4.6, Densitometric scatter-plot and quantification of Galectin-3, CD98 and β1-integrin membrane levels from figure 4.5 immunoblotting. Quantification of the levels of galectin-3 ***(a),*** CD98 ***(b),*** and β1-integrin ***(c),*** was conducted from the western blot images using the software Image J and plotted on GraphPad Prism 7. The blot densitometric intensities were normalised against the membrane specific housekeeping protein ATP1A1 (Na^+/K^+ ATPase) and has been labelled on the Y-axis. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table below each scatterplot represents the full p value of each protein expressed as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

4.2.3 Is CD98hc Expressed and Distributed on the Cell Surface of HLMF

In A549 cells I demonstrated CD98hc is distributed exclusively on the cell surface by comparing expression levels of CD98hc between permeabilised (whole cell sampling) against non-permeabilised (predominantly cell surface sampling) cells. In addition, siRNA-mediated CD98hc silencing confirmed the specificity of the CD98-44D7 antibody used for the flow-cytometry technique, observed by the reduced signalling levels (3.2.5.1). For HLMFs, western blotting data (4.2.2) suggest that CD98hc is membrane-restricted in NFC- and IPF- myofibroblasts. However, this does not discriminate between intracellular membranes such as mitochondrial membrane and the cell surface. Here I re-used the flow cytometry technique in non-permeabilised (cell surface intact) NFC HLMF cells to assess if CD98hc expression levels are assessable on the cell surface of myofibroblasts.

4.2.3.1 Cell surface Expression of CD98hc in HLMFs

Using flow cytometry in non-permeabilised NFC HLMF cells, I demonstrated that CD98hc expression levels were restricted to the cell surface of myofibroblasts (Figure 4.7). Cell surface CD98 expression levels increased following 24 h TGF-β1 stimulation (Figures 4.7 and 4.8). NFC HLMF cellular gating strategy displayed as dot plots is provided in Appendix (Figure A9).

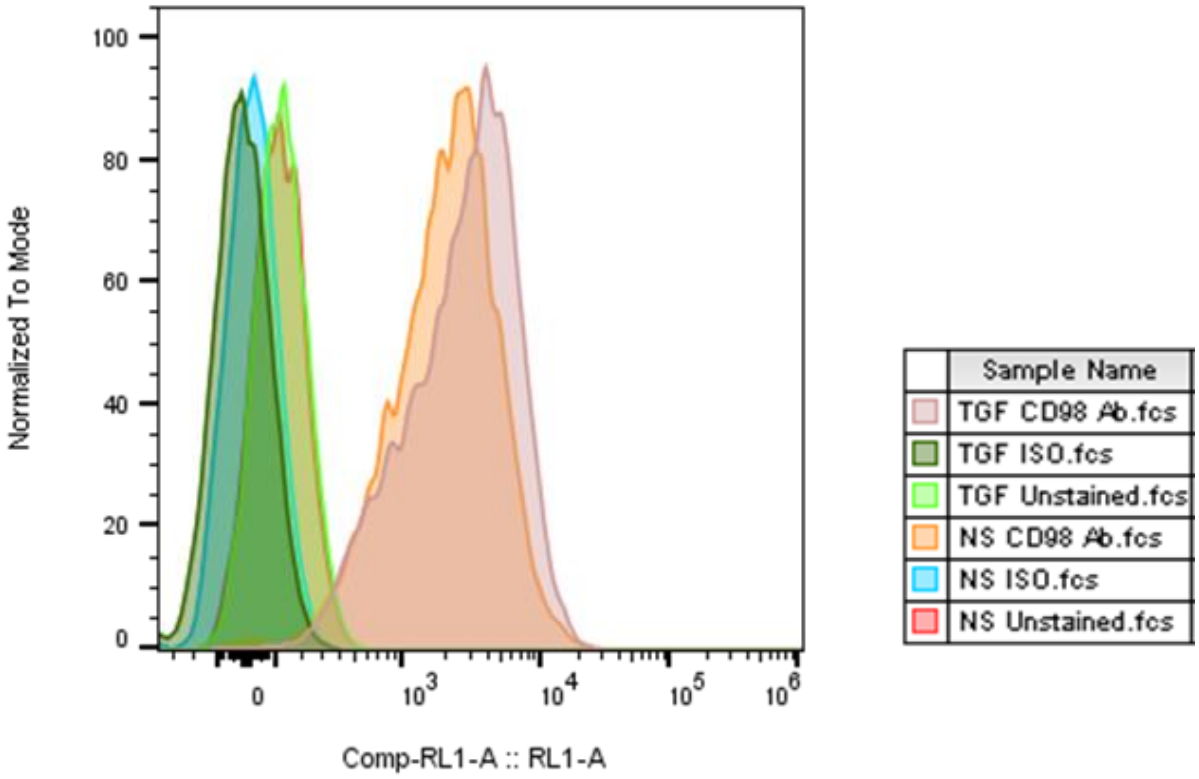
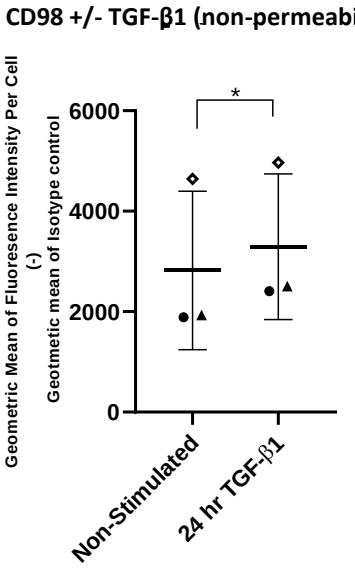


Figure 4.7, Flow cytometry FloJo analysed histograms of CD98 staining monitoring signal of CD98hc in non-permeabilised (cell membrane intact) NFC myofibroblast cells both basally and in response to TGF-β1 stimulation. The orange peak represents the staining signal of CD98 in non-stimulated conditions, and the grey peak represents the staining signal of CD98 following 24 h TGF-β1 stimulation. The blue and green peaks represent the isotype control staining signal (Mouse IgG) in non-stimulated and 24 h TGF-β1 stimulated conditions respectively. The pink and yellow peaks represent the unstained primary antibody control (fluorescence minus one or FMO) staining signal in non-stimulated and 24 h TGF-β1 stimulated conditions respectively. On the histogram, the Y-axis represents arbitrary normalised modal value; the X axis represents fluorescence intensity (in the Red Laser 1 (RL1) (the secondary antibody used is APC), which fluoresces in the red channel) given as a logarithmic scale. n=3

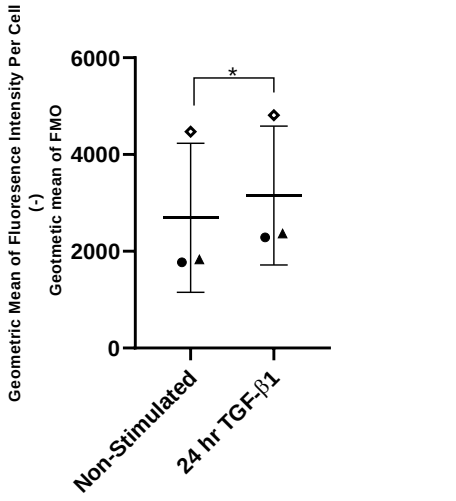
a)



b)

	PAIRED T TEST
Non-stimulated vs 24h TGF-β1	p=0.0245

c)



d)

	PAIRED T TEST
Non-stimulated vs 24h TGF-β1	p=0.0176

Figure 4.8, Statistical quantification of CD98hc signal following TGF-β1 stimulation, analysed from FloJo in non-permeabilised NFC myofibroblast cells. 4.8a, graph of geometric mean normalised to isotype control and 4.8c, graph of geometric mean normalised to FMO (unstained primary antibody) control, the graphs are presented as scatter-plots. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The X-axis represents the different conditions. The Y-axis represents the fluorescence intensities geometric mean. The asterisk plotted on the graphs represents significance from paired t-test. The table in figure ***4.8b and 4.8d,*** represents the full p value as paired t-test. (*p<0.05, p**<0.01, ***p<0.001).

4.2.4 Optimising Immunofluorescence Studies of Cell Surface Co-localisation of Galectin-3 with CD98 and β1-integrin in Human Lung Myofibroblasts

I wanted to assess the co-localisation interactions of cell surface galectin-3 with CD98hc (Figure 4.10) and β1-integrin (Figure 4.11) following pro-fibrotic stimulation with 10 ng/ml TGF-β1. Images were captured using the Leica TCS SP5 confocal microscope. As mentioned in 3.2.3, in these studies, one protein is stained with a probe signalling through a green channel (Alexa-fluor 488) and one with a probe signalling through a red channel (Alexa-fluor 594). Co-localisation is identified within the resolution of the technique, microscope and magnification in pixels where merging of the green and red channel data show signals from both (depicted by yellow pixel colouring).

Cell surface protein staining was enhanced by fixing cells using 4% (w/v) paraformaldehyde (PFA) instead of methanol in order to minimise cell membrane permeabilisation and rupture. Figure 4.9 compares galectin-3 and β1-integrin co-immunostaining with paraformaldehyde versus methanol fixation (in basal conditions). Visually it is apparent that the distribution of galectin-3 staining (in the red channel) when fixed with paraformaldehyde is more defined (restricting the galectin-3 staining on the cell surface). In contrast when cells are fixed with methanol the staining is greater and more diffuse and heavily saturated. A rational explanation for this would be methanol fixation allows the staining of both cell surface and cytosolic galectin-3. The staining intensity of β1-integrin (in the green channel) visually appears to look similar between both PFA and methanol fixative methods. This is consistent with the cell surface restricted expression of this protein. Furthermore, the nuclear staining (DAPI, blue channel) intensity visually appears to be similar between the two different fixative methods, confirming that PFA fixation did not hinder DAPI's ability to stain for nuclei. With all HLMF co-immunostaining data, consistent results were observed in triplicate experiments using and treating cells at low passages (P4 to P5). Considerable variability was observed with similar studies using HLMFs at higher passage (passage 7, not shown), possibly due to loss of myofibroblast differentiation at passage >6.

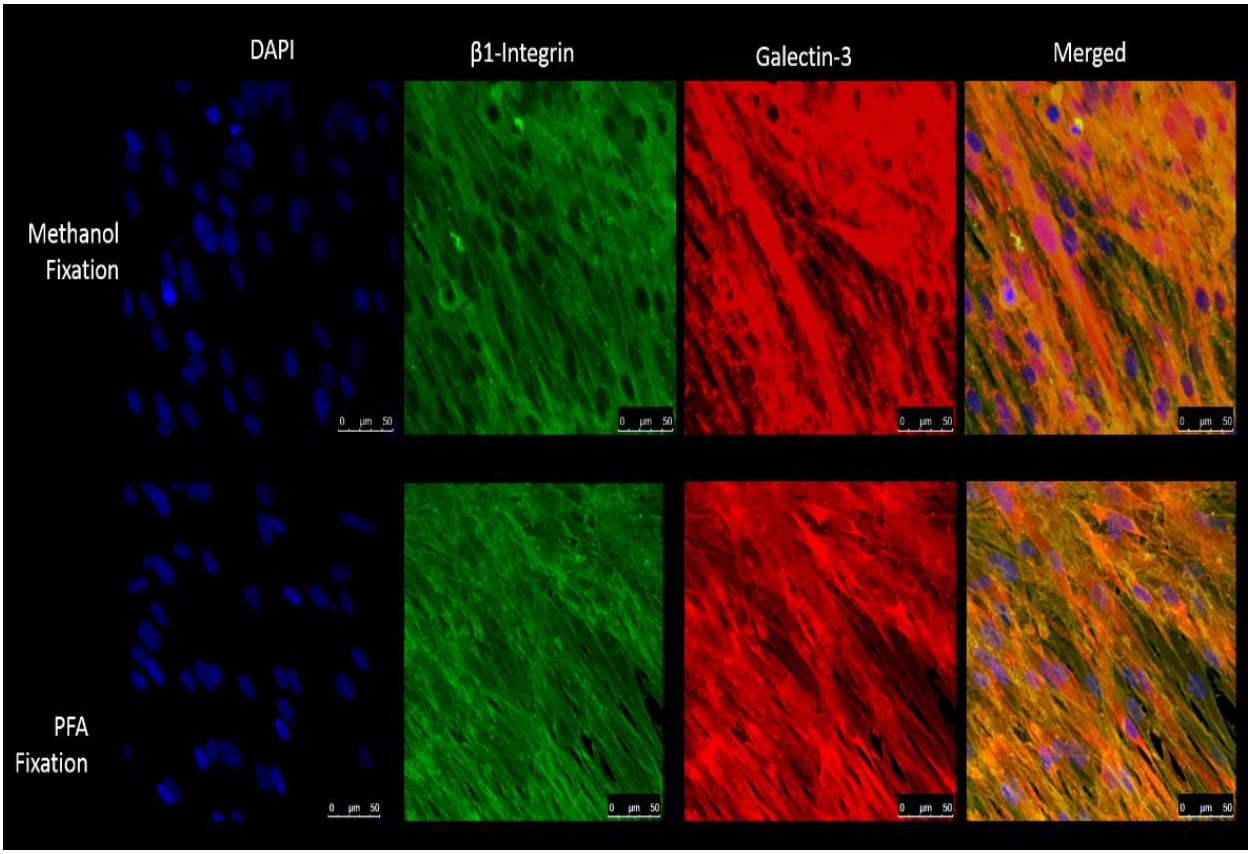


Figure 4.9, Comparing the effect of immunostaining distribution when using different fixation methodologies (Methanol vs 4% (w/v) Paraformaldehyde (PFA)). Cells were co-stained with 4 µg/ml galectin-3 and 3 µg/ml β1-integrin using goat-anti-rat Alexa-fluor 594 (red) to stain galectin-3 and goat-anti-mouse Alexa-fluor 488 (green) to stain β1-integrin. The nuclei were stained using DAPI. From each fixation condition, each channel was shown singularly and with a merged projection. All repeats were done on non-stimulated NFC HLMF cells at passage 4. All conditions between the two samples were the same, from passage number to the concentration of the antibodies used as well as microscope settings. Only difference in experimental condition is the fixation solution used. The merged image depicted yellow staining representing co-localisation (red and green overlap). A 50 µm scale-bar is drawn on the bottom right hand corner of each image.

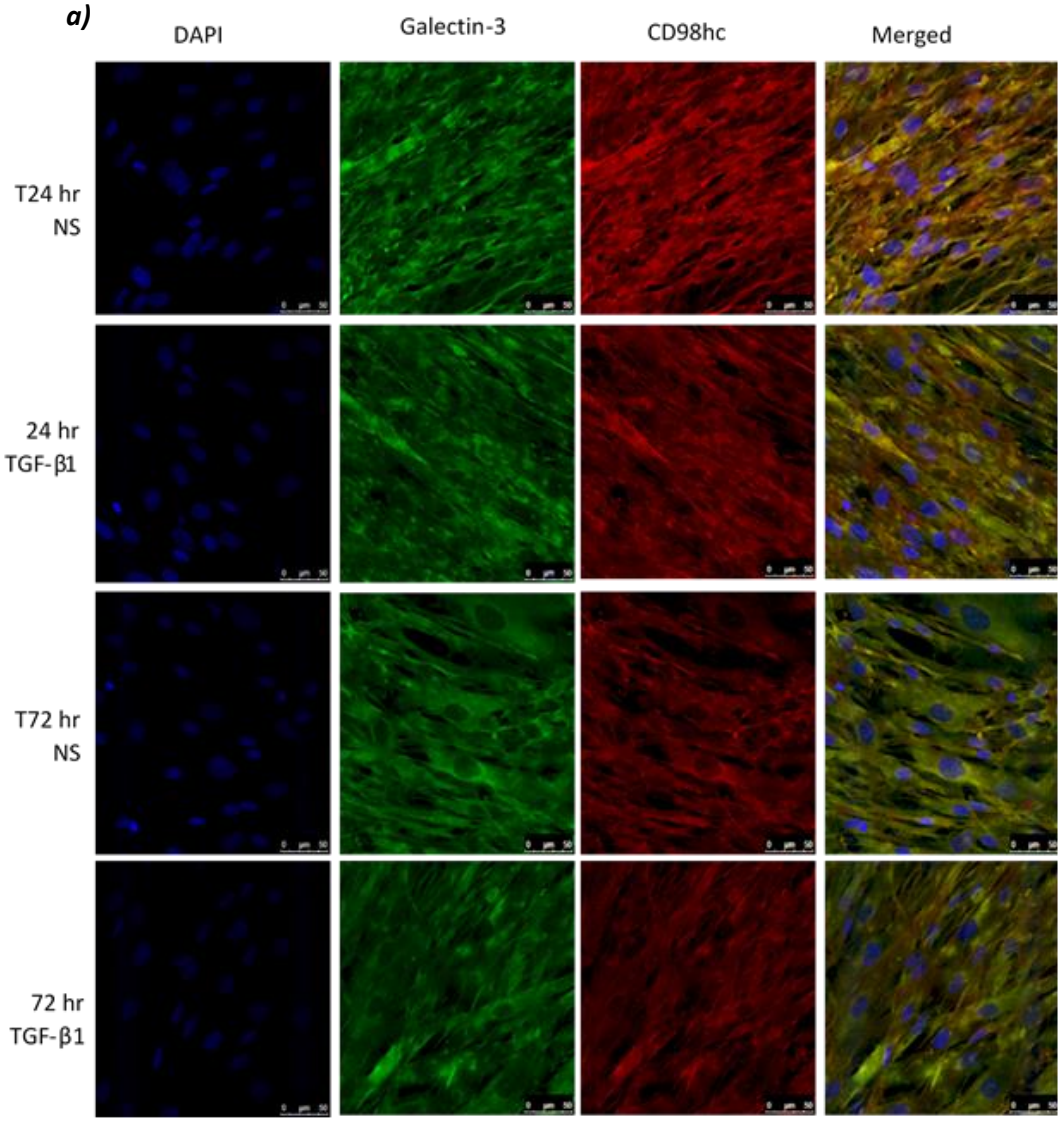
4.2.4.1 Co-Localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Decreases Following TGF-β1 Treatment in Non-Fibrotic-Control Human Lung Myofibroblasts

Direct observation from immunofluorescence studies supported the findings from western blot analyses of NFC HLMF membrane and cytosol lysate fractions (4.2.2). Following TGF-β1 stimulation, galectin-3 signals appeared to decrease (Figures 4.10 and 4.11, decreased staining in green and red channels respectively). However, the staining signals for CD98hc (Figure 4.10, red channel) and β1-integrin (Figure 4.11 green channel) protein expression appeared to increase. The reduction in the cell surface staining of galectin-3 concomitantly corresponded to the reduced co-localisation between both galectin-3 with CD98hc and β1-integrin protein pairs demonstrated by apparent reduction in the yellow merged signal intensity.

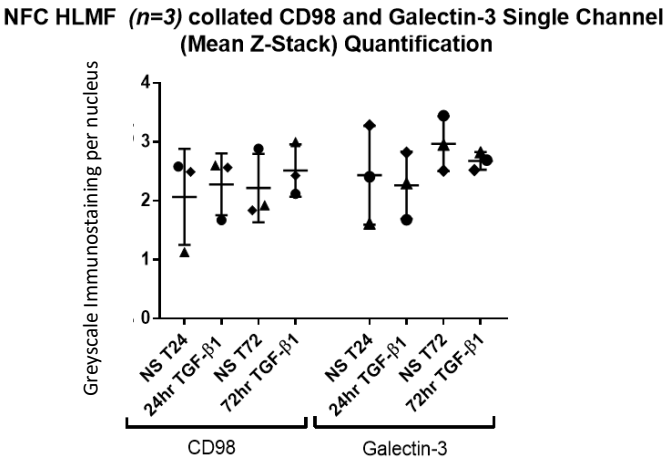
Quantification of the immunofluorescence findings for individual proteins alone, undertaken using 'Just Another Co-localisation Plug-in' (JACoP) (Bolt and Cordelieres, 2006) in Image J (Schindelin *et al*, 2012), was less clear-cut than the densitometric western blot analysis. Single channel immunostaining quantification of galectin-3 with CD98hc (Figures 4.10b-c) and β1-integrin (Figures 4.11b-c) co-immunostaining pairing did not demonstrate significant decreases in galectin-3 signal, or increase in the signals for CD98hc or β1-integrin following TGF-β1 stimulations.

On the other hand, assessing co-localisation with the Pearson's linear correlation and Mander's overlap coefficient quantification did support the visual impression at the microscope. Overall significant decreases in co-localisation of galectin-3 with CD98hc and β1-integrin were observed following TGF-β1 stimulation (Figures 4.12 and 4.13).

a)



b)

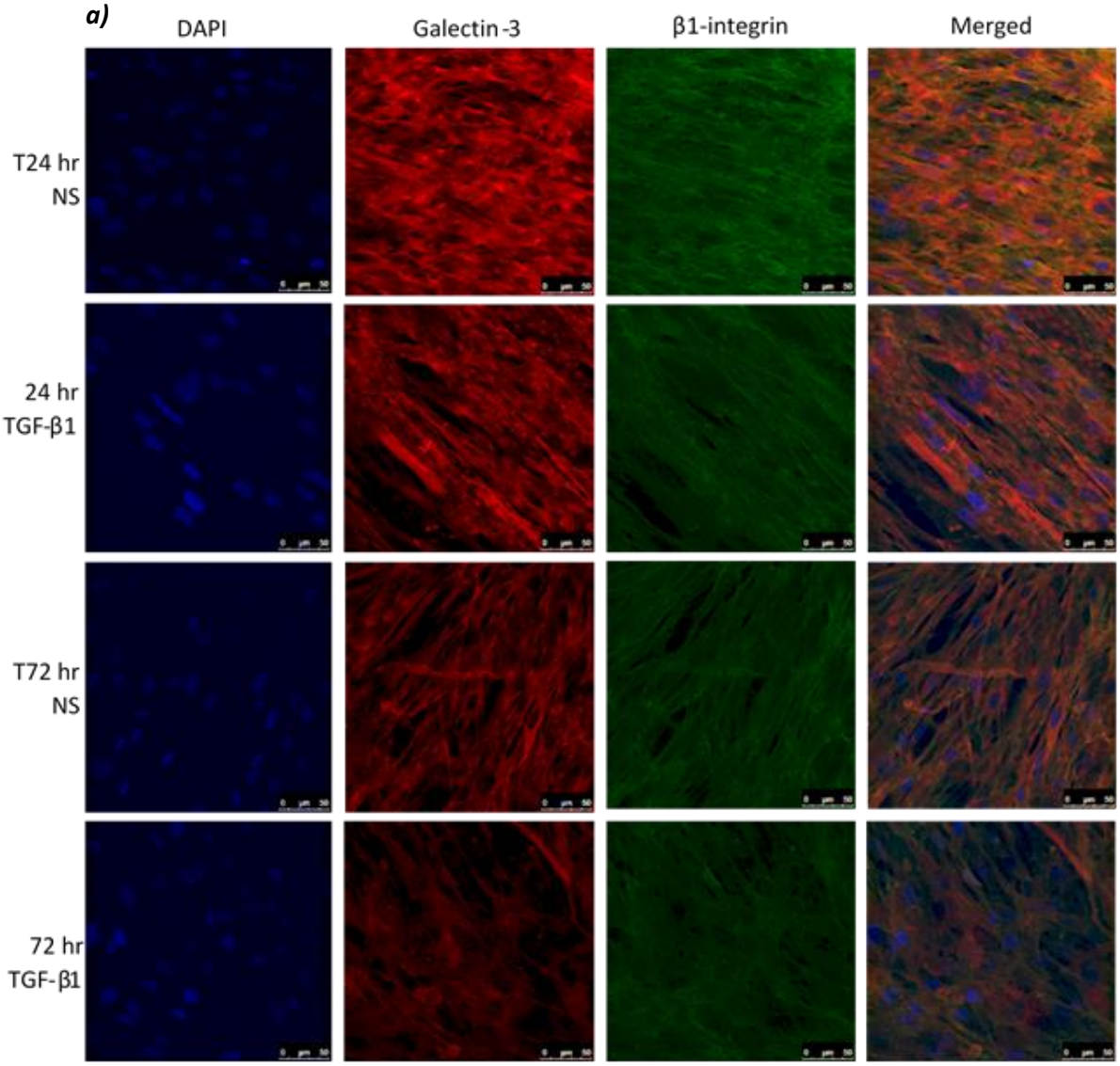


c)

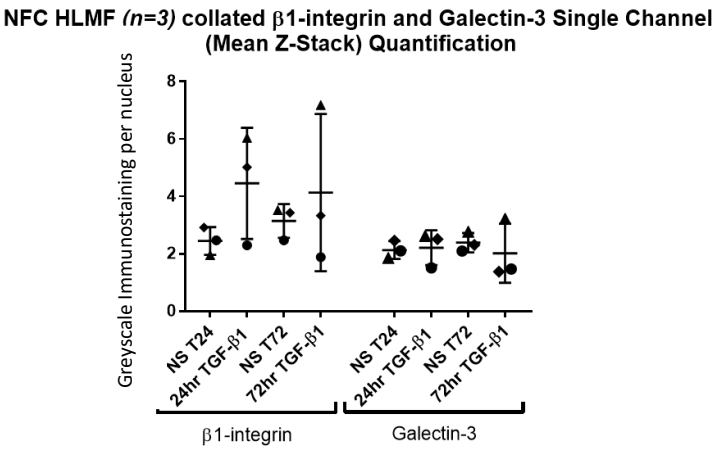
		PAIRED T TEST
CD98hc Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.7863
	ns 72h vs 72h TGF-β1	p=0.6423
Galectin-3 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.7322
	ns 72h vs 72h TGF-β1	p=0.3419

Figure 4.10, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in NFC HLMF. 4.10a, Cells were either untreated (24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 µg/ml galectin-3 and 5 µg/ml CD98hc using goat-anti-rat Alexa-fluor 488 (green) to stain galectin-3 and goat-anti-mouse Alexa-fluor 594 (red) to stain CD98hc. The nuclei were stained with DAPI. Each channel was shown singularly and with a merged projection. (n=3). ***4.10b-c,*** Single channel quantification of galectin-3 and CD98hc from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (n=3). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of CD98 and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The table in figure ***4.10c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

a)



b)



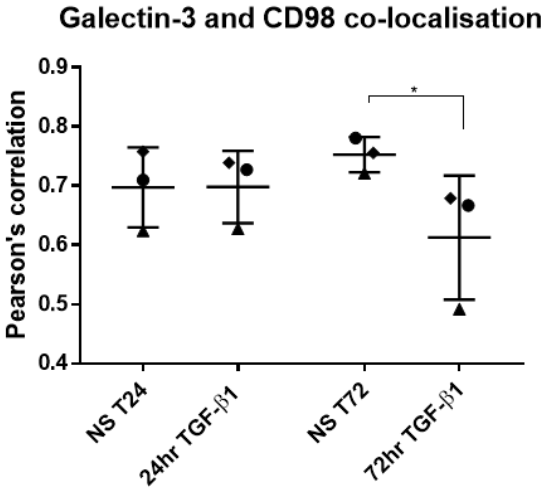
c)

		PAIRED T TEST
β1-integrin Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.2443
	ns 72h vs 72h TGF-β1	p=0.5376
Galectin-3 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.8615
	ns 72h vs 72h TGF-β1	p=0.4725

Figure 4.11, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in NFC HLMF. 4.11a, Cells were either untreated (24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 µg/ml galectin-3 and 3 µg/ml β1-integrin using goat-anti-rat Alexa-fluor 594 (red) to stain galectin-3 and goat-anti-mouse Alexa-fluor 488 (green) to stain β1-integrin. The nuclei was stained with DAPI. Each channel was shown singularly and with a merged projection. (n=3). ***4.11b-c,*** Single channel quantification of galectin-3 and β1-integrin from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (*n=3*). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of β1-integrin and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The table in figure ***4.11c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

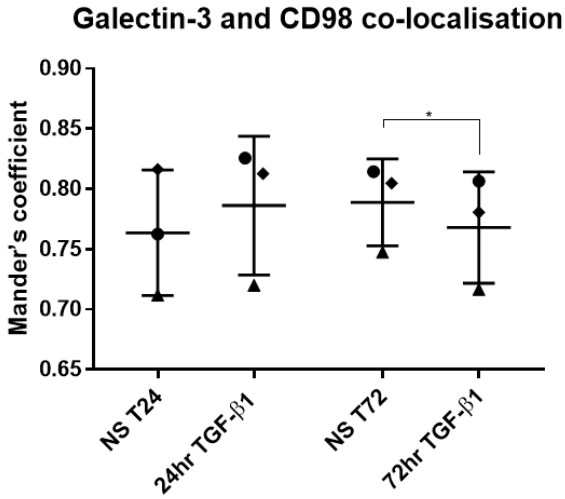
Co-localisation quantification of galectin-3 and CD98hc co-immunostaining following TGF-β1 stimulations

a)



		PAIRED T TEST
CD98 and Galectin-3 Pearson's correlation co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.9340
	ns 72h vs 72h TGF-β1	p=0.0108

b)

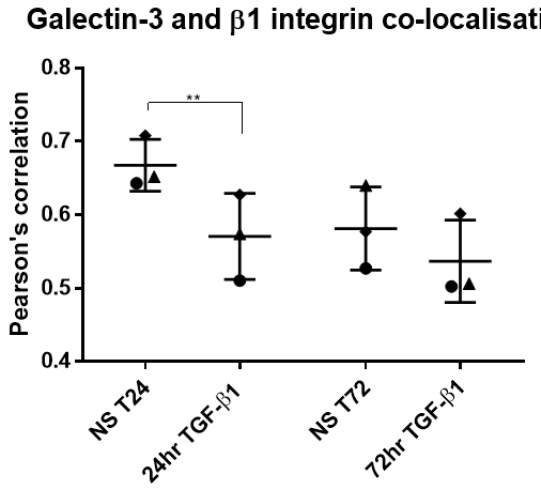


		PAIRED T TEST
CD98 and Galectin-3 Mander's coefficient co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.4441
	ns 72h vs 72h TGF-β1	p=0.0409

Figure 4.12, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations in NFC HLMF. 4.12a, Pearson's quantification of co-localisation and ***4.12b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded significance represented for paired t-tests with asterisks are displayed on the graphs. Full p-values for this analysis are given in the lower panels (*p<0.05, p**<0.01, ***p<0.001).

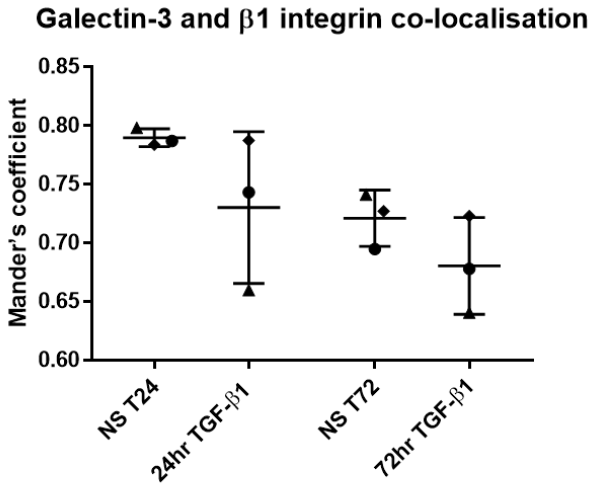
Co-localisation quantification of galectin-3 and β1-integrin co-immunostaining following TGF-β1 stimulations

a)



		PAIRED T TEST
β1-integrin and Galectin-3 Pearson's correlation co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.0368
	ns 72h vs 72h TGF-β1	p=0.4340

b)



		PAIRED T TEST
β1-integrin and Galectin-3 Mander's coefficient co-localisation quantification	ns 24h vs 24h TGF-β1	p=0.2891
	ns 72h vs 72h TGF-β1	p=0.3059

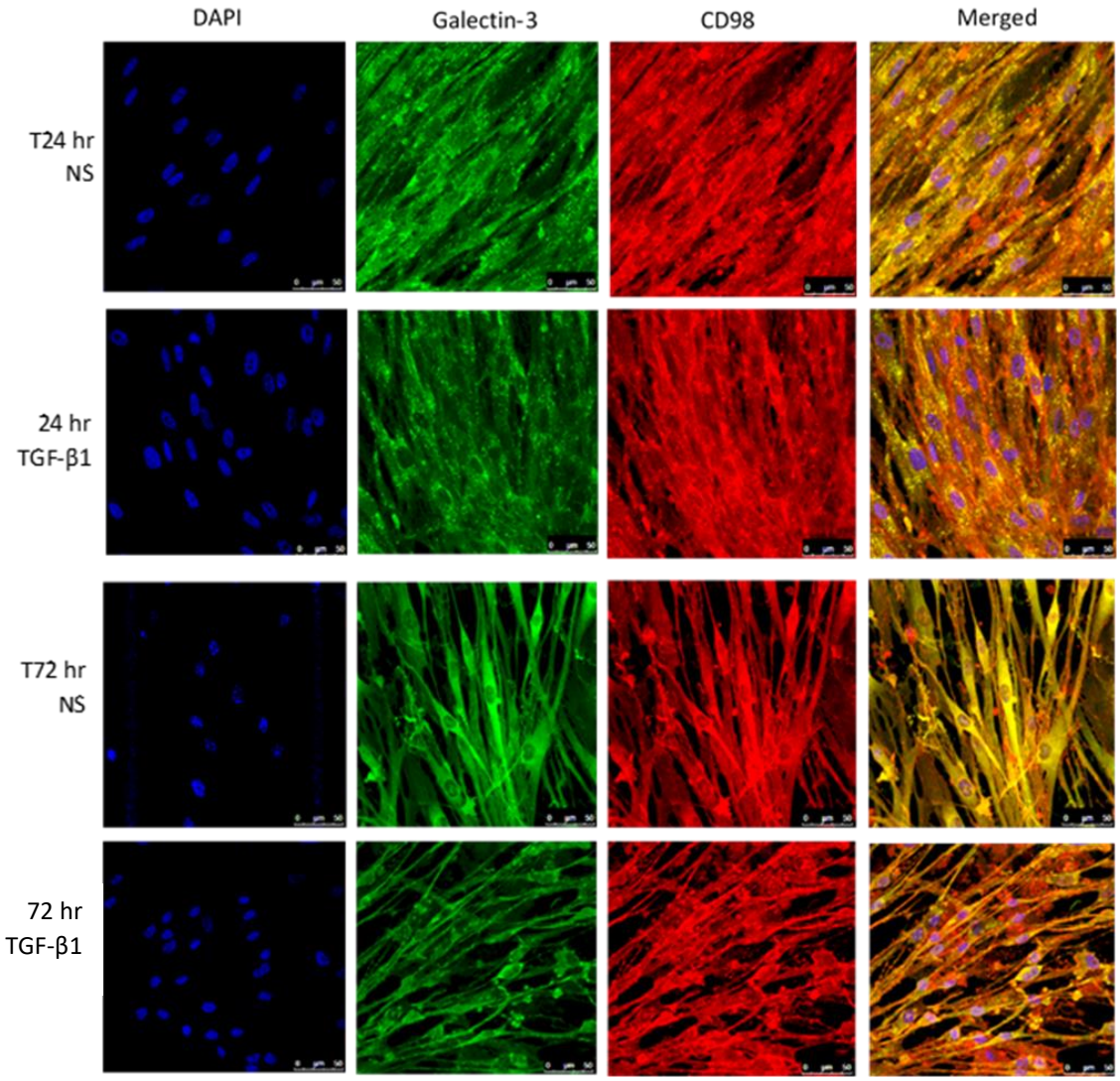
Figure 4.13, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations in NFC HLMF. 4.13a, Pearson's quantification of co-localisation and ***4.13b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded significance represented for paired t-tests with asterisks are displayed on the graphs. Full p-values for this analysis are given in the lower panels (*p<0.05, p**<0.01, ***p<0.001).

4.2.4.2 Co-Localisation Between the Protein Pairing of Galectin-3 with CD98hc and β1-integrin Remains Stable and Unchanged Following TGF-β1 Treatment in IPF Human Lung Myofibroblasts

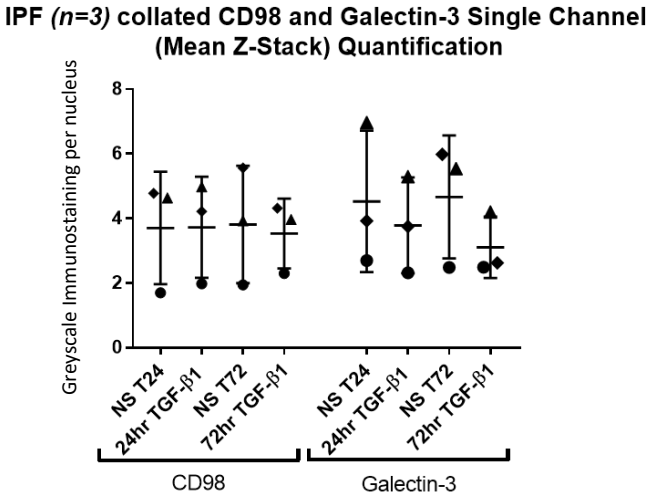
Visually, single channel immunofluorescence data reporting galectin-3, CD98hc or β1-integrin in dual stained IPF myofibroblast cells (Figures 4.14a and 4.15a) appeared to display similar profiles to those observed in NFC HLMFs, consistent with the findings from western blot analysis in these cells (4.2.2). However, this was not further supported by quantitative analysis of the immunofluorescence. Instead, none of these proteins demonstrated any significant changes in response to TGF-β1 stimulation detectable from the study of cells from three donors (Figures 4.14b and 4.15b). Interestingly, the co-localisation of galectin-3 with CD98hc (Figure 4.16) and β1-integrin (Figure 4.17) protein pairs also remained unchanged when assessed by Pearson's and Mander's coefficients in IPF HLMFs, unlike NFC HLMFs.

Intriguingly the staining distribution pattern appears punctate in IPF HLMFs when compared to NFC HLMFs. Based on speculation, this could suggest the presence of focal adhesions. This is in agreement with the increased contractility features of IPF myofibroblasts over healthy NFC myofibroblasts.

a)



b)

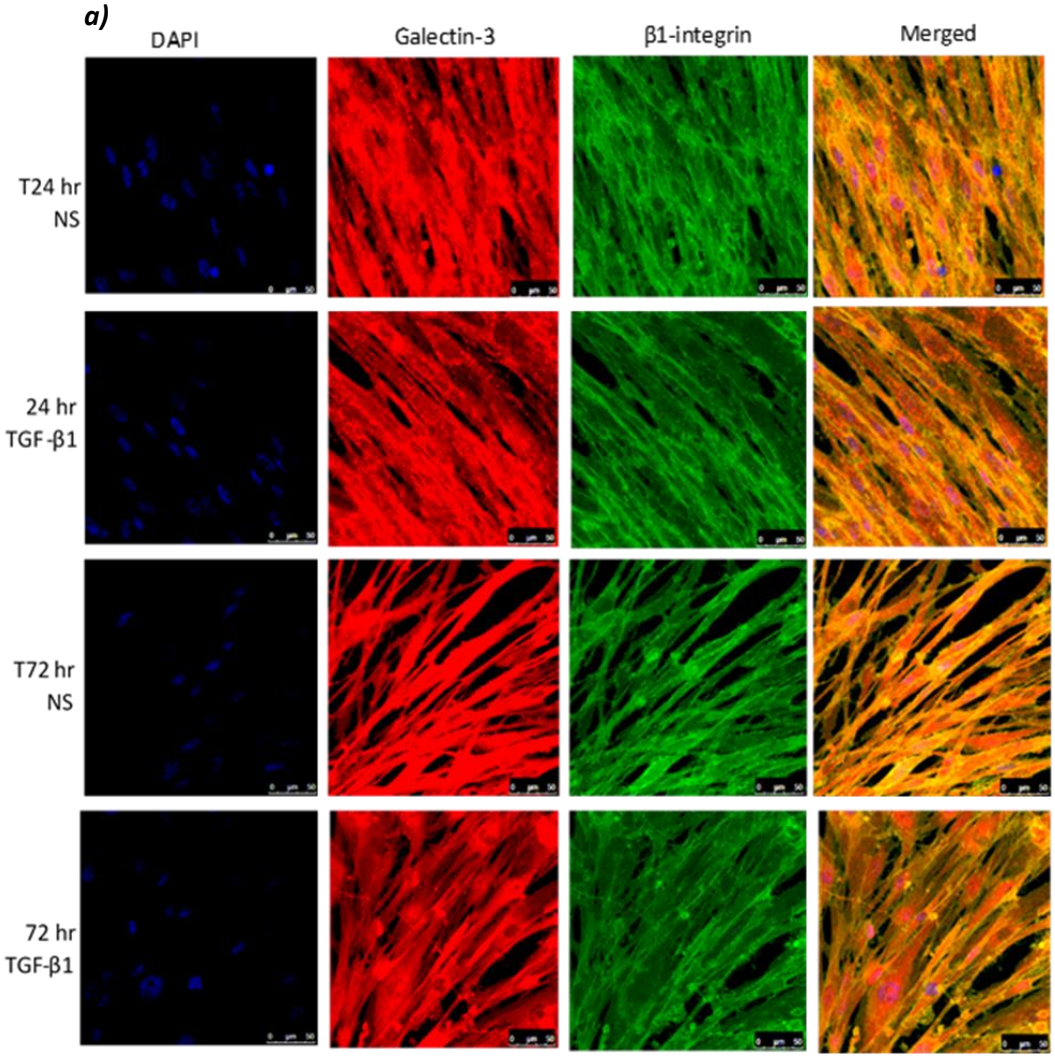


c)

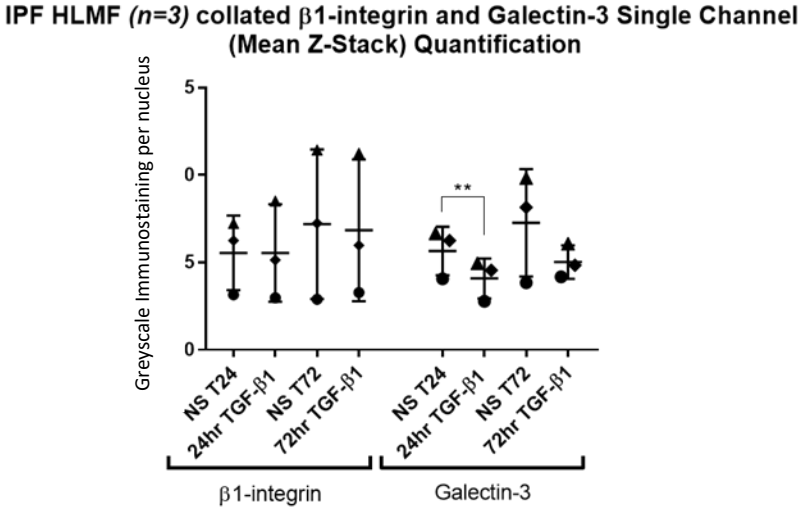
		PAIRED T TEST
CD98 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.9503
	ns 72h vs 72h TGF-β1	p=0.6221
Galectin-3 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.2542
	ns 72h vs 72h TGF-β1	p=0.2505

Figure 4.14, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with CD98hc following TGF-β1 stimulations in IPF HLMF. 4.14a, Cells were either untreated (24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 µg/ml galectin-3 and 5 µg/ml CD98hc using goat-anti-rat Alexa-fluor 488 (green) to stain galectin-3 and goat-anti-mouse Alexa-fluor 594 (red) to stain CD98hc. The nuclei were stained with DAPI. Each channel was shown singularly and with a merged projection. (n=3). ***4.14b-c,*** Single channel quantification of galectin-3 and CD98hc from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (*n=3*). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of CD98 and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The table in figure ***4.14c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

a)



b)



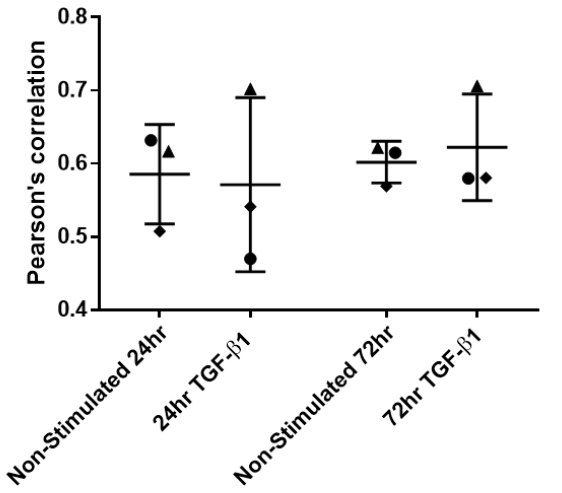
c)

		PAIRED T TEST
β1-integrin Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.9976
	ns 72h vs 72h TGF-β1	p=0.5425
Galectin-3 Single Channel Immunostain	**ns 24h vs 24h TGF-β1**	p=0.0080
	ns 72h vs 72h TGF-β1	p=0.2251

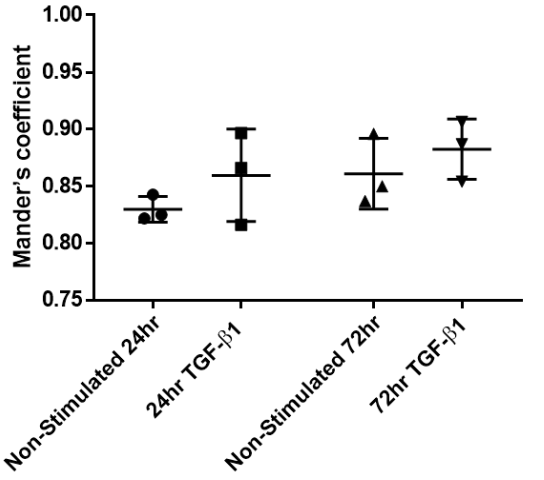
Figure 4.15, Co-Immunofluorescence staining focusing on the co-localisation of galectin-3 with β1-integrin following TGF-β1 stimulations in IPF HLMF. 4.15a, Cells were either untreated (24 h NS and 72 h NS) or treated with 10 ng/ml TGF-β1 stimulation for 24 h and 72 h. Cells were co-stained with 4 µg/ml galectin-3 and 3 µg/ml β1-integrin using goat-anti-rat Alexa-fluor 594 (red) to stain galectin-3 and goat-anti-mouse Alexa-fluor 488 (green) to stain β1-integrin. The nuclei were stained with DAPI. Each channel was shown singularly and with a merged projection. (n=3). ***4.15b-c,*** Single channel quantification of galectin-3 and β1-integrin from the co-immunostaining pairing. Quantification of the single channel staining and nuclei count was done using image J (*n=3*). The Y-axis represents arbitrary proportion of staining when the single channel staining is normalised to per nucleus of the particular field view image. The X-axis represents the treatment condition of the cells. On the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed and the separation of β1-integrin and Galectin-3 is highlighted. The graph is presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test. The table in figure ***4.15c,*** represents the full p value as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

Co-localisation quantification of galectin-3 and CD98hc co-immunostaining following TGF-β1 stimulations

a)



b)



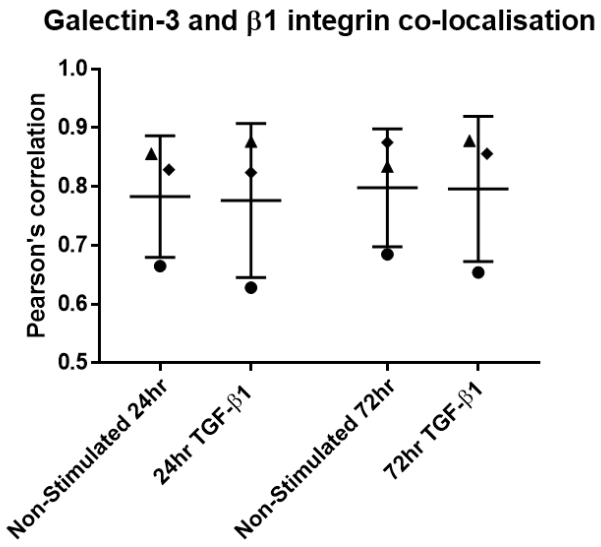
		PAIRED T TEST
CD98 and Galectin-3 Pearson's correlation co-localisation quantification	**ns 24h vs 24h TGF-β1**	p=0.8665
	ns 72h vs 72h TGF-β1	p=0.6169

		PAIRED T TEST
CD98 and Galectin-3 Mander's coefficient co-localisation quantification	**ns 24h vs 24h TGF-β1**	p=0.2408
	ns 72h vs 72h TGF-β1	p=0.4699

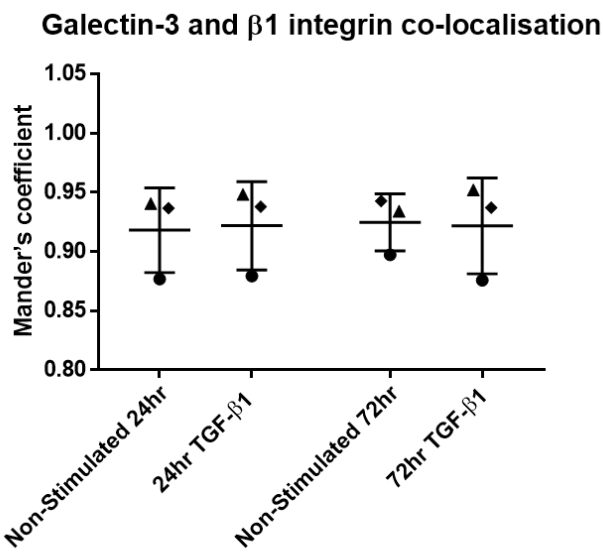
Figure 4.16, Quantification of the galectin-3 and CD98 co-immunostaining pairing subjected to TGF-β1 stimulations in IPF HLMF. 4.16a, Pearson's quantification of co-localisation and ***4.16b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded full p-values for this analysis are given in the lower panels as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

Co-localisation quantification of galectin-3 and β1-integrin co-immunostaining following TGF-β1 stimulations

a)



b)



		PAIRED T TEST
β1-integrin and Galectin-3 Pearson's correlation co-localisation quantification	**ns 24h vs 24h TGF-β1**	p=0.7198
	ns 72h vs 72h TGF-β1	p=0.9397

		PAIRED T TEST
β1-integrin and Galectin-3 Mander's coefficient co-localisation quantification	**ns 24h vs 24h TGF-β1**	p=0.1991
	ns 72h vs 72h TGF-β1	p=0.8158

Figure 4.17, Quantification of the galectin-3 and β1-integrin co-immunostaining pairing subjected to TGF-β1 stimulations in IPF HLMF. 4.17a, Pearson's quantification of co-localisation and ***4.17b,*** Mander's quantification of co-localisation was done using image J (n=3). Y-axis represents the extent of co-localisation (co-localisation is quantified on a 0-1 scale) and on the X-axis the appropriate non-stimulated (NS) and TGF-β1 stimulated condition has been inscribed. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by donor number) are represented by different symbols. Mean values and standard deviations are plotted. Thresholded full p-values for this analysis are given in the lower panels as paired t-test (*p<0.05, p**<0.01, ***p<0.001).

4.2.5 Co-Localisation Addressed by Co-immunoprecipitation (co-IP)

I attempted to further assess the apparent reductions in co-localisation behaviour of galectin-3 with CD98 and β1-integrin by co-immunoprecipitation in NFC HLMF cells to test and optimise the protocol. Compared to A549 cells, to observe a detectable signal in the NFC HLMFs, triple concentration of lysate was required. I did not have access to sufficient IPF lung tissue to study these readouts in IPF HLMFs.

Galectin-3 was immunoprecipitated from the cell lysate in the absence or presence of TGF-β1 stimulation, using the appropriate monoclonal antibody. The co-immunoprecipitation of β1-integrin and CD98 was assessed by SDS-PAGE and western blot analysis. At 24 h TGF-β1 stimulation both CD98 (Figure 4.18a) and β1-integrin (Figure 4.18b) appear to slightly increase, however by 72 h TGF-β1 stimulation, both CD98 (Figure 4.18a) and β1-integrin (Figure 4.18b) co-immunoprecipitation with galectin-3 appears to have decreased when compared to the non-stimulated control. Interestingly when observing the β1-integrin co-immunoprecipitation blot (Figure 4.18b), galectin-3 levels decreased following TGF-β1 stimulation, consistent with both the western blotting (4.2.2) and (co)-immunostaining (4.2.4) data. This work needs further optimisation and a method needs to be implemented allowing subsequent densitometric quantification normalised to an appropriate control.

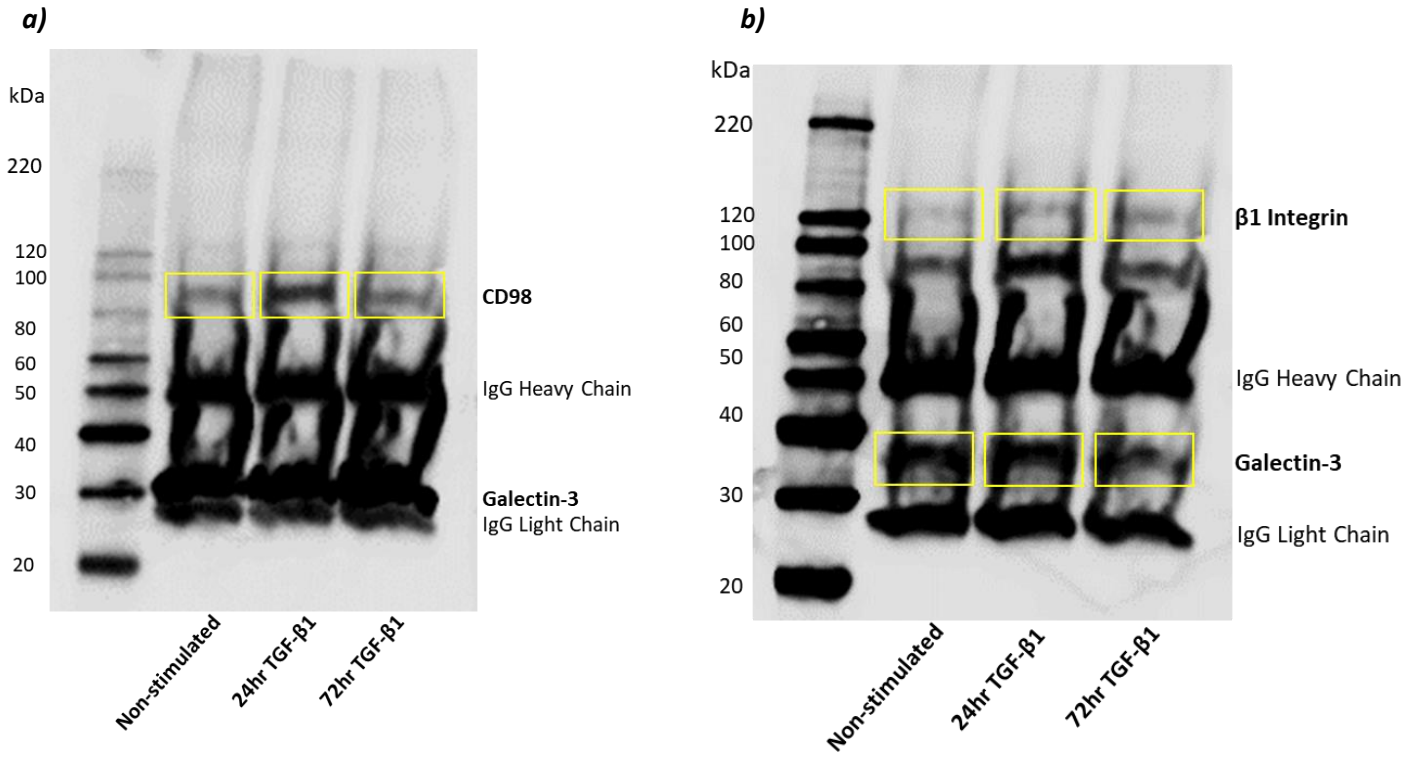


Figure 4.18, Anti-galectin-3 pulldown of whole cell NFC HLMF lysates. CD98 co-IP (a) and β1-integrin co-IP (b). Lane 1 represents non-stimulated cell lysates, Lanes 2 and 3 represents 24 h and 72 h 10 ng/ml TGF-β1 stimulated cell lysates. At the ~85 kDa molecular weight, CD98hc co-immunoprecipitation is observed. At ~120 kDa molecular weight β1-integrin co-immunoprecipitation is observed. 1.5 mg of lysate was subjected to anti-galectin-3 immunoprecipitation. The band at ~32kDa represents galectin-3. The bands at ~50kDa and ~25kDa represents the immunoglobulin heavy and light chain. All blots were immersed with EZ-ECL chemiluminescent detection kit and the image was taken at an exposure time of between 2-10 minutes. (n=2).

4.3 Discussion

I hypothesised that expression and co-localisation behaviour of the putative gal-3-fibrosome proteins including galectin-3, CD98 and β1-integrin, would be conserved between the A549 lung epithelial cell model and *ex vivo* derived HLMFs.

Understanding pro-fibrotic molecular pathways is key for developing novel anti-fibrotic therapies. Currently primary human alveolar epithelial cells cannot be reliably used in tissue culture and so A549 cells were used to model their behaviour. However, the mesenchymal human lung myofibroblasts can be purified *ex vivo* and so I took the opportunity to probe the gal-3-fibrosome hypothesis in these cells from NFC and IPF conditions in the absence and presence of TGF-β1 stimulation.

Before any downstream experiments, the *ex vivo* derived myofibroblast cells from each donor were critically characterised by immunofluorescence biomarker staining, due to the structural similarities and overlap of myofibroblast to fibroblasts and airway smooth muscle (ASM) cells. Unsurprisingly, the cell culture media used are the same for myofibroblasts, fibroblasts and ASM cells (Singh and Hall. 2007, Sigma. UK). The positive markers assessed are not entirely specific to myofibroblast cells as they are present in other cell types, however the culture media is selective for these three overlapping cell types and therefore it restricts the comparison between the three. Taken together, the purity of myofibroblast identity was confirmed for all donors used in my study, in its current best ability by using a panel of biomarkers. This was vital for my data integrity as analysing the expression of the gal-3-fibrosome proteins in basal and in response to TGF-β1 stimulated conditions is that of *ex vivo* derived mesenchymal matrix secreting myofibroblast cells.

As in A549 cells, flow-cytometry confirmed CD98 expression was restricted to the cell surface in HLMFs. Supporting our initial hypothesis, western blot analysis strongly indicated that TGF-β1 stimulations caused an increase in cell surface CD98 and β1-integrin levels in both NFC and IPF HLMFs. Unsurprisingly, β1-integrin and CD98 increased similarly, consistent with reports repetitively stating CD98hc is constitutively coupled with β1-integrin as a heterodimer on the cell surface (Feral *et al*. 2005). However, in contrast to the A549 cells, following TGF-β1 stimulation galectin-3 levels decreased in both NFC and IPF HLMFs. This was surprising since galectin-3 is reported to mediate pro-fibrotic responses to TGF-β1 in fibroblasts (Henderson *et al*. 2006, Mackinnon *et al.* 2012).

In contrast, Professor Gisli Jenkins's group (work done at University of Nottingham) report galectin-3 mRNA transcripts increase following TGF-β1 stimulation from IPF fibroblasts (Parmer *et al*. poster presentations and personal communication, *Manuscript in preparation*). To address the differences in our findings, it may be relevant that this work was done at slightly higher passage values (> P6), where we have observed loss of original phenotype in a range of readouts. It would also be helpful to know how their cells were phenotypically characterised.

In case the differences in my protein expression data and their mRNA data were due to regulation of translation in the mesenchymal cells following TGF-β1 stimulation, I worked with Dr Katy Roach (Bradding group, University of Leicester) to analyse the LGALS3 mRNA transcripts for galectin-3 following TGF-β1 stimulation in both NFC and IPF HLMFs. In the cell samples we studied, LGALS3 (galectin-3) mRNA decreased, consistent with the protein expression levels (Appendix, Figure A10b).

We then conducted further experiments in whole lung tissue from the explant model. Here too galectin-3 (LGALS3) mRNA levels decreased, and CD98 (SLC3A2) mRNA transcripts increased following TGF-β1 stimulation (Appendix, Figure A10c and A10d). Dr. Roach has already

established that β1-integrin (ITGB1) mRNA transcripts are upregulated in response to TGF-β1 in lung tissue (Roach *et al.* 2018).

Unlike the western blot quantitation, the single channel immunofluorescence quantitation was less clear-cut. However, in agreement with the reduced protein levels of galectin-3, overlap analysis of immunofluorescence co-localisation indicated that galectin-3 co-localisation with CD98 and β1-integrin decreased in NFC HLMFs following TGF-β1 stimulation. Interestingly this was not observed in IPF HLMFs. This reduced co-localisation behaviour in NFC HLMFs following TGF-β1 stimulation was further supported by initial (n=2) co-immunoprecipitation experiments. As and when sufficient IPF tissue are available for comparison, co-immunoprecipitation experiments should be conducted to confirm the co-immunofluorescence findings.

A reasonable explanation for this decreased co-localisation behaviour could be that in our gal-3-fibrosome model, galectin-3 itself may not be relevant for a pro-fibrotic status unlike β1-integrin and CD98hc via TGF-β1 stimulated fibrosis in healthy NFC myofibroblast. Another explanation for this could be the operation of a negative feedback loop in the healthy NFC myofibroblasts as a means to limit fibrotic responses, which is disabled or overridden in IPF.

Interestingly when observing the co-immunostaining micrographs of galectin-3 with CD98hc and β1-integrin pairings, stronger staining is observed in IPF than NFC HLMFs in basal conditions. In particular, basal levels of galectin-3 staining is the most clear-cut. Images of both NFC and IPF HLMF cells were captured on the same confocal microscope under identical settings (e.g. exposure and magnification). However, increased number of samples are likely required for a definitive conclusion based upon standard threshold of statistical significance. Although, basal levels of galectin-3 mRNA transcripts are higher in IPF than NFC myofibroblasts (Appendix, Figure A10a), supporting the validity of this observation.

If confirmed a finding that gal-3-fibrosome proteins are found in high amounts in basal conditions in IPF myofibroblasts, this could suggest potential to mediate pro-fibrotic responses autonomous of further TGF-β1 stimulation.

As with A549 cells, the absence of an effective TGF-βRII mAb in our hands meant I could not directly draw conclusions about this protein's expression and co-localisation behaviour.

Taken together, my data for *ex vivo*-isolated NFC and IPF myofibroblast cells are complex in the expression and co-localisation responses of galectin-3, CD98 and β1-integrin to TGF-β1 stimulation than was observed in A549 epithelial cells. Therefore, it is difficult to draw definitive conclusions from this model of study, however speculations can be made. Expression levels of galectin-3 on the protein level and its subsequent decreased co-localisation behaviour with CD98hc and β1-integrin in NFC myofibroblasts modified the initial 'positive feedback loop' model of gal-3-fibrosome behaviour in this context. In the context of lung tissue this reduction in myofibroblast galectin-3 production might be offset by paracrine supplementation with galectin-3 secreted from epithelial cells. However, in IPF myofibroblasts following TGF-β1 stimulation the expression levels of galectin-3 decreased, yet the co-localisation with CD98hc and β1-integrin remained stable, consistent with a concept of co-stabilisation of gal-3-fibrosome proteins within the complex in IPF. Basal levels of individual fibrosome proteins seem to be more in IPF derived myofibroblasts, suggesting that they might have an important role in IPF myofibroblasts (disease state), potentially autonomous of further TGF-β1 binding.

Chapter 5

Electron Microscopy Studies Probing the Galectin-3 Oligomeric Behaviour and the Gal-3-Fibrosome Hypothesis

5.1 Introduction

Galectin-3 does not itself directly mediate any signalling or enzymatic activity. Its key characteristic is its ability to participate in homo- and heterotypic protein:protein and protein:glycan interactions via N- and C-terminal domain interactions. At the cell surface level putative binding partners include multiple proteins implicated in fibrosis. These include the type II receptor for the key pro-fibrotic cytokine transforming growth factor (TGF-)β, TGF-βRII, and CD98hc, that appears constitutively complexed to β1-integrin (Chapters 3 and 4) (MacKinnon *et al*. 2008, 2012). Other binding candidates with relevance in pro-inflammation and fibrosis have been suggested, including CD11b/CD18 (forming the CD3 heterodimer) (McCubbery *et al*. 2018, Sciacchitano *et al*. 2018). Though, we are not assessing these potential interactions at the current moment.

Galectin-3 binds glycan chains on glycoproteins modified by glycosylation (Santos *et al*. 2016, Nielson *et al*. 2018) through its C-terminal carbohydrate recognition domain (CRD). This domain consists of ~135 amino acids (Zou *et al*. 1995, Kamili *et al*. 2016) and its structure has been solved in the apo- state (Seetheraman *et al*. 1998, Umemoto and Leffler. 2001, Zhuang *et al*. 2008, Manzoni *et al*. 2018) and bound to many ligands, particularly by NMR (Sorme *et al*. 2005, Shilova *et al*. 2020). The CRD lacks any α-structure and folds to a β-sandwich structure into which the glycan ligand intercalates (Varki *et al*. 2009) (Figure 1.5). In contrast, no high-resolution structures of the full-length N-terminal domain have yet been solved, agreeing with the N-terminal domain adopting an intrinsically disordered state in isolation (Lin *et al*. 2017). The N-terminal domain consists of 21 amino-acids identified as the N-terminal stretch (NTS), which contains two serine residues that are phosphorylation sites, followed by nine Proline/Glycine/Tyrosine-rich collagen like repeats (I-IX) (Ibarra *et al*. 2018).

To address the flexibility issue, a mutant galectin-3 variant with truncated N-terminal domain constituting of the NTS with the collagen-like repeats VII-IX (galectin-3[NTS/VII-IX]) was characterised crystallographically in the apo-state by another group whilst I was pursuing this

PhD. A 2.2 Å resolution X-ray crystallographic structure was determined (Flores-Ibarra *et al*. 2018). In conjunction with the C-terminal domain, two parts of the N-terminal domain were structurally characterised: namely the NTS and the terminal region of the collagen-like repeat (IX). These were observed to form a β-hairpin motif, extending the F-sheet with the addition of F-1 (Phe5-His8 residues) and F0 (Gly13-Asn16 residues) strand (Figure 5.1). Repeat (IX) was visualised as a free N-terminal tail (Figure 5.1c).

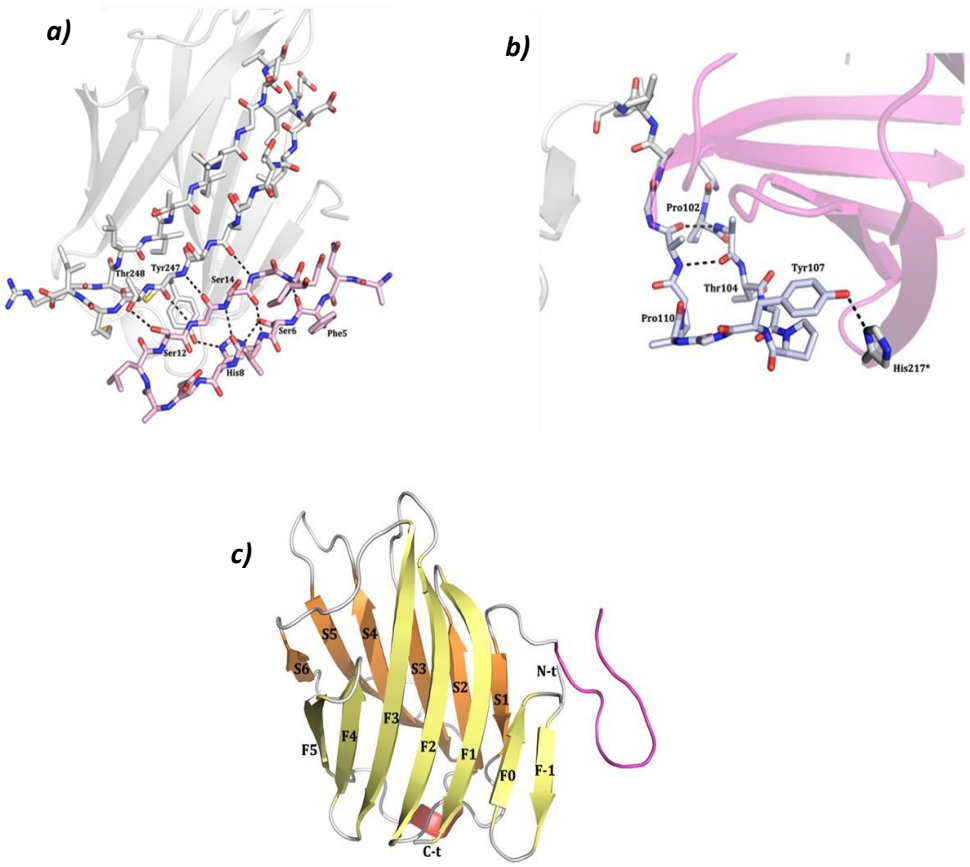


Figure 5.1, X-ray crystallographic structure of the NTS and repeat IX-and a structural model defining the CRD with the two new regions identified in the Gal-3[NTS/VII-IX] variant. ***(a)*** Specific set of contacts between the C-terminal region of the CRD and NTS residues, typical of a β-sheet. F-1 and F0 represents the newly identified β-sheet, taking the F-sheet from F1-F5, now to F-1 to F5. ***(b)*** Close-up view of the collagen-like repeat IX in the N-terminal domain region (Tyr101-Val116). Two hydrogen bonds between Ala103 and Ala111 stabilise this section. One extra hydrogen bond with a symmetry-related molecule between Tyr107 and His217 allowed to define this segment. ***(c)*** Schematic cartoon structure defining the secondary structure of the CRD and the two regions of the N-terminal domain defined by this study. (Flores-Ibarra *et al*. 2018), PDB ID: 1A3K

Prior to the assessment of the truncated variant of galectin-3[NTS/VII-IX], only the CRD had been structurally solved. Whilst the use of the truncated variant has led to further understanding of the intrinsically disordered N-terminal domain and has given the possibility to model further annotations of the overall structure, many queries remain unanswered. These include, to further understand the structural effect of the (total) nine collagen like repeats (I-IX) and its overall interactions with the CRD. Furthermore, alongside the apo- state, it remains unclear how the NTS and the repeats behave in a CRD:ligand bound state. Crucially it does not directly inform on the oligomerisation mechanism (Nieminen *et al*. 2007, Halimi *et al*. 2012, Lepur *et al*. 2012).

Protein:protein interactions of galectin-3 are mediated by the N-terminal domain. This has capacity to self-associate (N-type oligomerisation) (Nieminen *et al*. 2007, Lepur *et al*. 2012) or bind to other proteins (Halimi *et al*. 2014). Such interactions of galectin-3 as a whole are dependent upon dissociation of non-covalent interactions of the N-terminal domain with the carbohydrate recognition domain (Halimi *et al*. 2014, Lin *et al*. 2017, Flores-Ibbara *et al*. 2018, Chiu *et al*. 2020) in a conformational change that is facilitated by glycan binding. Oligomerisation of galectin-3 is described at high concentrations in the absence of glycan ligands (Lepur *et al*. 2012), but in their presence (e.g. with the disaccharide lactose *N*-acetyllactosamine) it occurs far more readily (Ahmad *et al*. 2004, Dam *et al*. 2005, Lepur *et al*. 2012, Bumba *et al*. 2018). However, the carbohydrate recognition domain can also self-associate (C-type oligomerisation) (Lepur *et al*. 2012, Halimi *et al*. 2014) and so it is possible that the importance of the N-terminal domain rests in its regulation of these interactions.

I wanted to study the oligomerisation behaviour of galectin-3 in both non-cellular and cellular context; as targeting this could provide a novel small molecule treatment strategy for fibrosis in the lung and other organs.

Our collaborator (Professor Hakon Leffler, Lund University, Sweden) kindly donated human galectin-3 construct. The expression plasmid for galectin-3 was the pET3C (Ampicillin resistant) vector which was transfected into *E.coli* BL21*. I worked with Professor Russell Wallis (University

of Leicester, LISCB) to successfully make my own stock of lactosyl-Sepharose resin, and used this for affinity column chromatography in a single step purification protocol (Pa *et al*. 2015).
To understand how galectin-3 oligomerisation behaviour may nucleate lattice assemblies at the cell surface (and oligomeric conditions in health and disease), purified galectin-3 oligomerisation was studied *in vitro* in the absence and presence of ligands. Oligomerisation was assessed though multiple readouts: non-denaturing (native)-PAGE, cross-linkage followed by SDS-PAGE, SEC and SEC-MALS, negative stain-EM and sucrose gradient fractionation.

Styrene maleic acid (SMA) is a synthetic polymer composed of styrene (hydrophobic moieties) and maleic acid (reactive, functional moieties) monomers (Smith *et al*. 2019, Orbiscope. 2021). SMA co-polymers interact with cell phospholipid membranes to extract membrane proteins within SMA lipid particle (SMALP) nanodiscs (Pollock *et al*. 2019, Smith *et al*. 2019). The protein species can then be studied in a relatively physiological lipid environment by structural biology methods e.g. crystallography (Broecker *et al*. 2017) or cryo-electron microscopy (cryo-EM) (Postis *et al*. 2015, Parmar *et al*. 2018). This contrasts with traditional membrane crystallography that is conducted in aqueous solution in the presence of detergent. Species can be extracted from cells in the presence or absence of stimulation to understand how responses are mediated.

To probe for proteins associated with the gal-3-fibrosome, I chose to explore the ability of SMALPs to extract species containing CD98. Galectin-3 is widely distributed between cellular compartments, whilst β1-integrin is found widely on the cell surface and may have a wide range of distinct interactions not involving CD98, and endogenous TGF-βRII is not strongly identified by the antibodies used in my study. CD98 therefore seemed the proposed constituent most likely to produce observations relevant to the gal-3-fibrosome (e.g. cell surface interactions).

Figure 5.2 schematically displays how the proteins are encapsulated by the SMA nanodiscs during the cell lysate preparation to form SMALPs.

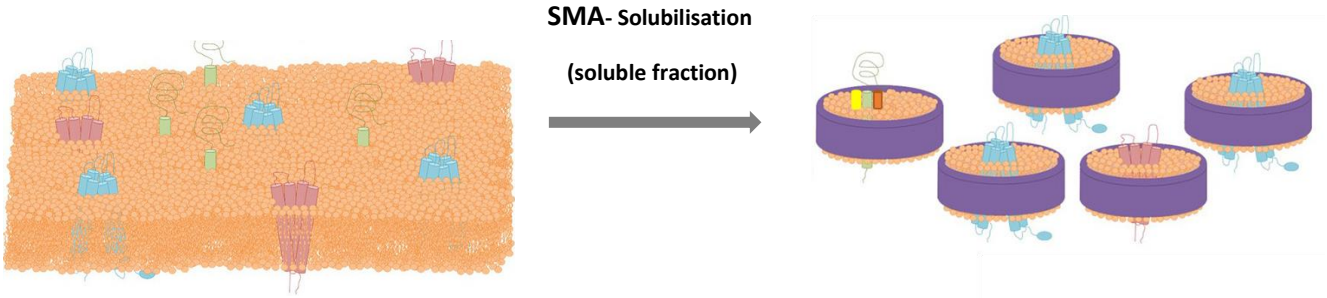


Figure 5.2, SMA mediated membrane protein solubilisation and purification. SMA is added to cell lysate preparations where the polymer inserts into the membrane and punches holes forming small discs of proteins within their lipid bilayer encircled by the polymer (~15 nm in size); this is termed as SMALP (SMA Lipid particles) nanoparticles. The SMALPs contains proteins and lipid bilayer surrounded by the polymer. The soluble fractions are separated from the insoluble material by a high-speed ultracentrifugation step (2.3.3). The lysate includes all solubilised proteins, and at this point is crude. The subsequent particles are stable, soluble and amenable to downstream experiments. (Figure taken from Pollock *et al*. 2018, and amended)

The SMA encapsulated soluble proteins can be further purified from the total cell lysate by methods such as affinity chromatography (Pollock *et al*. 2018), resulting in the extraction purification of membrane protein of interest, without removing the proteins from their lipid bilayer environment (Pollock *et al*. 2018, Hall *et al*. 2020).

As I was studying endogenous protein without tags for affinity chromatography, my strategy, (outlined in Figure 5.3) was to perform immunoprecipitation using protein G magnetic beads. The monoclonal anti-CD98hc antibody (CD98-E5) was used (Table 2.13) to bind exposed CD98hc antigen embedded and exposed in the nanodiscs. Once the antigen-antibody complex is coupled to the protein G magnetic beads, the only sample that will be bound to the beads is that of the CD98hc encapsulated nanodiscs and co-encapsulated proteins. The bead-bound complex can then be eluted off (using a native elution approach) and subsequently probed and analysed for

downstream experiments including SDS-PAGE, non-denaturing (native)-PAGE and negative stain-EM for sample preparation analysis to work up for future cryo-EM studies.

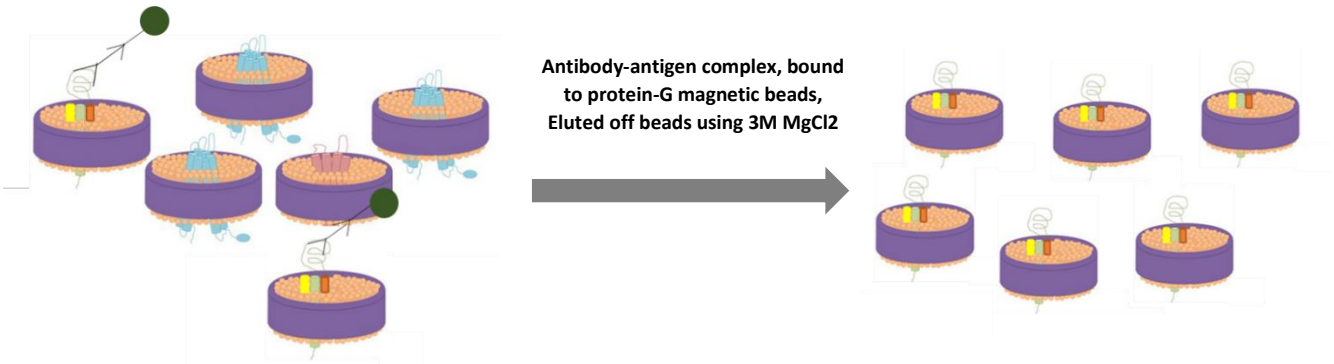


Figure 5.3, SMALP CD98hc purification approach. SMALP solubilised lysates were purified specifically for CD98 by adding the appropriate CD98hc antibody, allowing it to bind to the antigen. The Fc portion of the antibody is then bound to the protein G magnetic beads. The sample is eluted off using magnesium chloride keeping the sample in its native state.

5.2 Results

5.2.1 Applying Structure Prediction Algorithms to the Galectin-3 Sequence

The structure of around ~50% of galectin-3 (N-terminal domain) is minimally characterised. Conversely the remaining C-terminal, carbohydrate recognition domain has been studied in high resolution structural detail on multiple occasions. To assess the potential for the N-terminal domain to adopt secondary structure in the monomeric state, the full-length protein sequence (human 'LGALS3' extracted from UniProtKB) was subjected to secondary structure prediction.

The carbohydrate recognition domain was included as a control of known structure. Protein Homology/analogY Recognition Engine (Phyre^2) and Glob-plot was used. Structure prediction suggested with high confidence that the N-terminal domain is intrinsically disordered in a monomeric state. Figure 5.4 shows Phyre^2 predictions (predicted disorder depicted as '**?**', heatmap colouring for strength of confidence, red=high).

Predictions for the carbohydrate recognition domain were made with high confidence and indicated 30% β-structure, with 6% transmembrane helix and 1% α-helix. This corresponds to ~0 % α- and transmembrane helix in contrast to ~50% β-strand in the solved structures, allowing a general cross-check of algorithm accuracy. The predictions were also cross-referenced with Glob-Plot analysis of disorder propensity (Figure 5.5). This generates a 'running total' plot of predicted disorder propensity as the sequence progresses from N- to C-terminus. The profile displayed shows a fairly linear increase in this index across the N-terminal domain, which then falls steadily as the carbohydrate recognition domain (indicated by cross-hatched background due to matching with published structure) is taken into account. The final value approaches 50% of the

peak value, broadly consistent with a prediction of 50% disorder (N-terminal domain) and 50% ordered structure (C-terminal carbohydrate recognition domain).

Figure 5.4, Structure prediction using the program Protein Homology/analogY Recognition Engine (Phyre^2) and inputting the Human Galectin-3 (LGALS3) Amino Acid Sequence. The human galectin-3 amino acid sequence was obtained from UniProtKB from FASTA (canonical tab). The prediction was generated by using the Phyre^2 website and loading the amino acid sequence onto the prediction tool. The key on the bottom right hand corner of the figure dictates the motif of the predicted sequence, and the confidence of each prediction (red being the highest (numeric value 9), and purple being the lowest (numeric value 0) confidence).

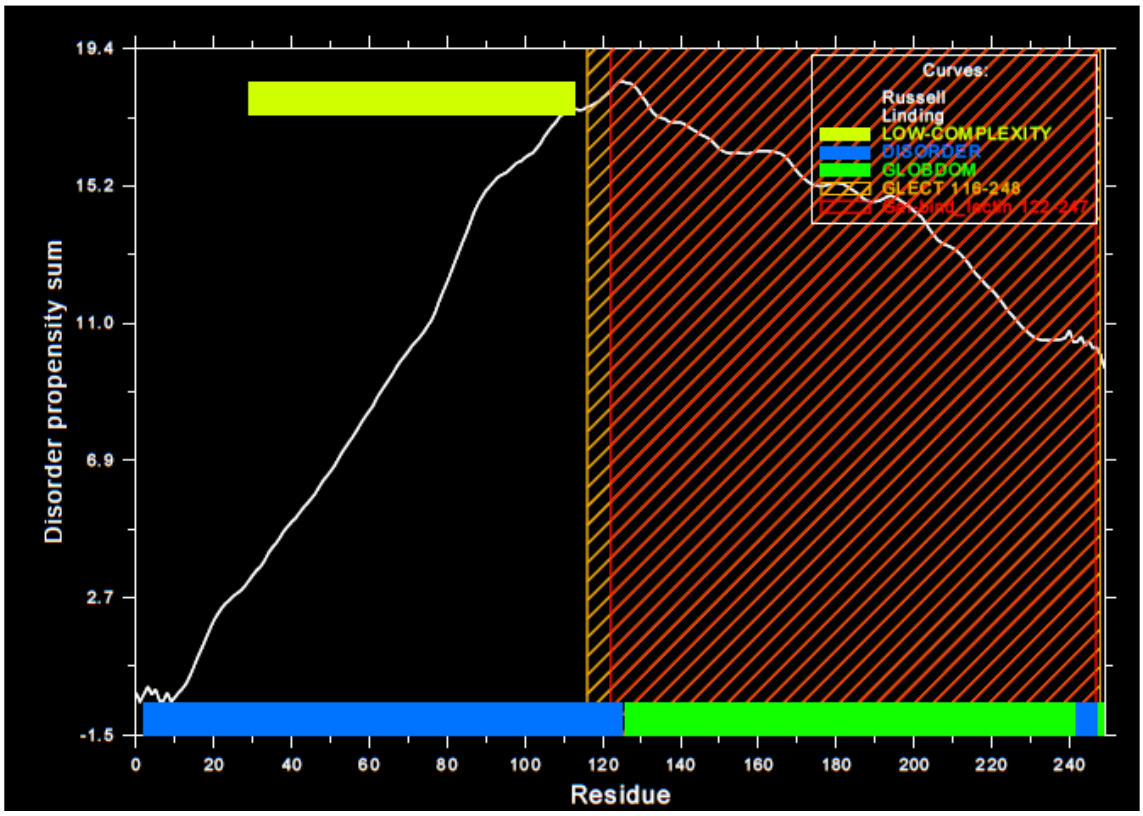


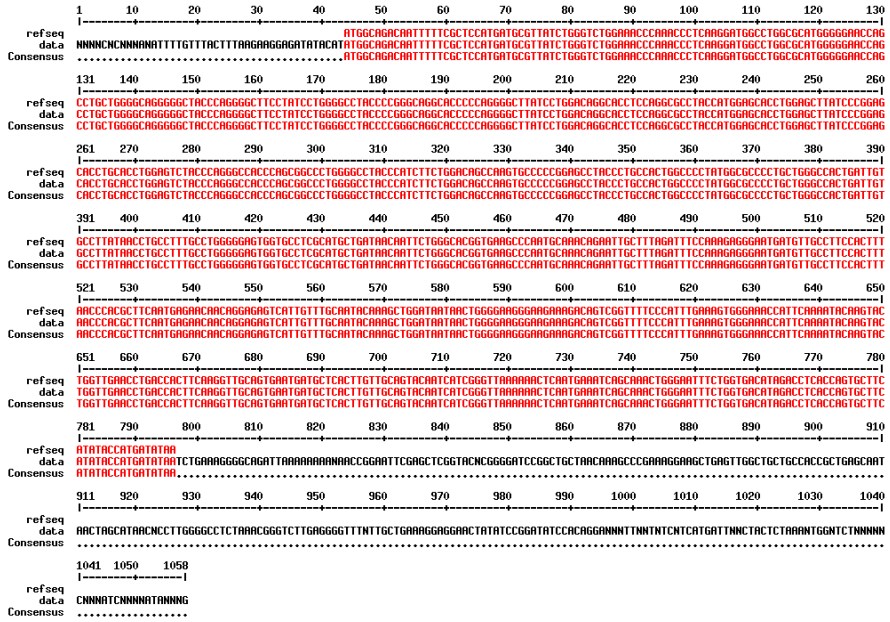
Figure 5.5, Structure prediction using the program Glob-Plot to assess human galectin-3 amino acid sequence. The human galectin-3 amino acid sequence was obtained from UniProtKB from FASTA (canonical tab). The prediction was generated by using the GlobPlot website and loading the amino acid sequence onto the prediction tool. The key on the top right hand corner of the figure dictates the motif of the predicted sequence. The X-Axis represents the amino acid residue number. The Y axis displays the propensity for a given amino acid to be in 'random coil' or in 'regular secondary structure'.

5.2.2 Galectin-3 Purification

5.2.2.1 Galectin-3 Plasmid Sequence Confirmation

The sequence of the donated cDNA construct encoding galectin-3 was assessed. Colonies of cells successfully transfected with the construct were inoculated into 5 ml of LB Broth for small scale plasmid preparation (Mini Prep, 2.10.2). Plasmid sequencing was performed by Sources Bio-Sciences UK (Nottingham, UK). T7 forward and reverse complementary primers were used. The observed sequence was aligned with the reference sequence of human galectin-3 (extracted from RefSeq LGALS3 using base pair positions 40-792) using the 'MultiAlin tool'. There was a 100% match (Figure 5.6) between the reference and data sequence observed by both the forward and reverse primers. The extra data sequences correspond to the vector sequence (pET3c).

T7 Primer-Forward



T7 Primer-Reverse (Reverse complement)

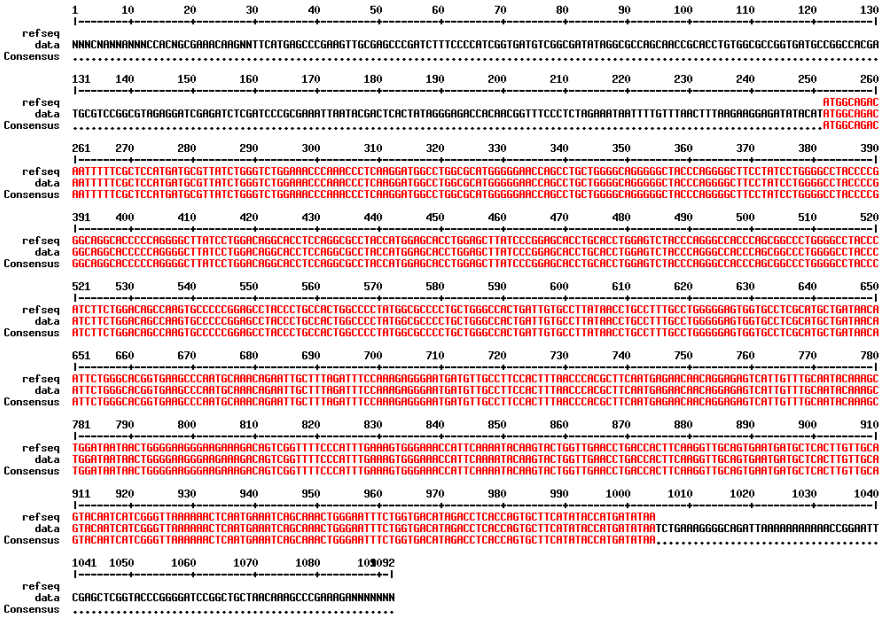


Figure 5.6, Sequence alignment of human Galectin-3 (LGALS3). LGALS3 sequence obtained from RefSeq. The Reference Sequence was aligned against our data sequence using the 'MultiAlin tool'. LGALS3 sequence was used from base pair positions 40-792. Extended data sequence is that of the vector (pET3c). For the forward sequence T7 Forward primer was used. For the reverse sequence T7 Reverse primer (reverse complement) was used. All sequence between the Reference and Data aligned with a 100% match.

5.2.2.2 Purification of Recombinant Galectin-3

Galectin-3 was successfully purified by affinity chromatography using the lactosyl-sepharose column (Figure 5.7). Minimal galectin-3 (31 kDa) was observed in the flow through or buffer wash. A single intense band of protein at the expected ~31 kDa weight for galectin-3 eluted with the step elution condition of 150 mM lactose.

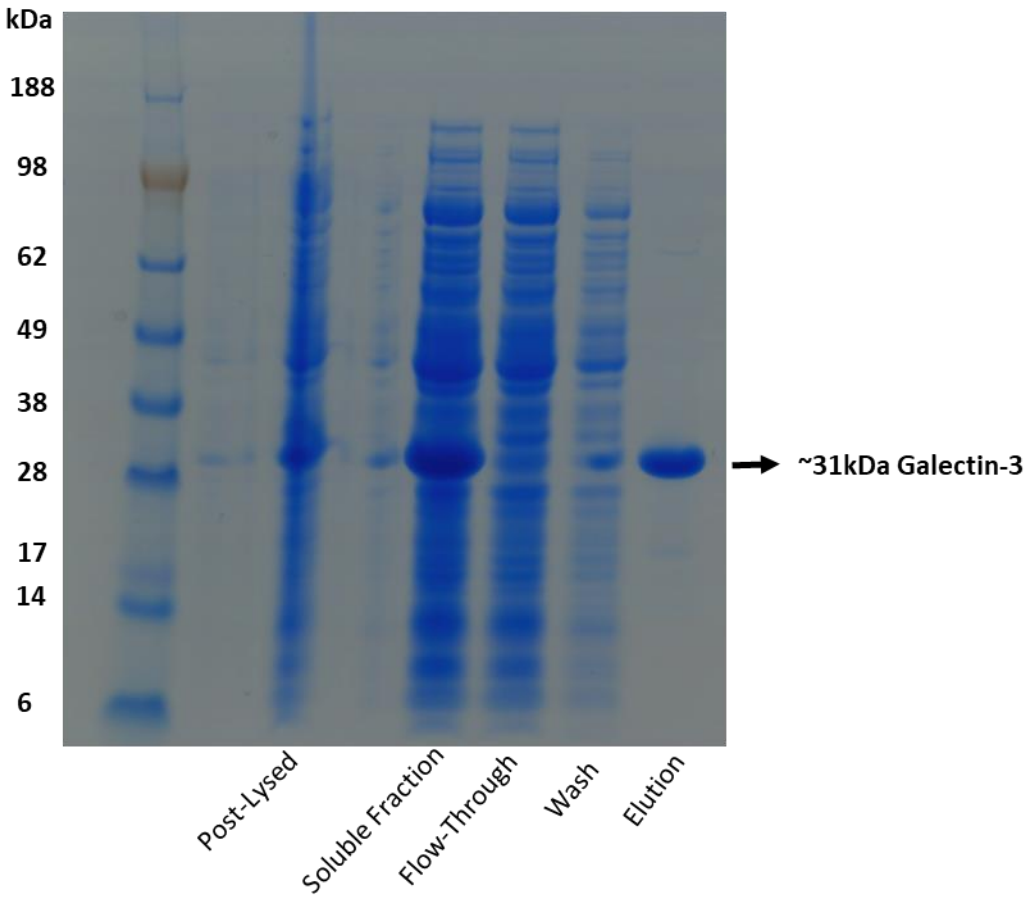


Figure 5.7, Multi-steps in galectin-3 purification. Recombinant galectin-3 purified post-culture. The galectin-3 sample was harvested post 3 h 0.5M IPTG induction. 10 µl of each sample was loaded with 5µl of 2x Laemmli sample buffer on a 4-12% (v/v) SDS-PAGE resolving gel. The gel was stained overnight in Expedeon instant blue. Lanes are labelled with the appropriate sample conditions (X-axis). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Galectin-3 molecular weight at ~31 kDa has been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.2.3 Maximising Galectin-3 Elution

150 mM lactose is a standard concentration used to elute galectin-3 (Massa *et al*. 1993, Cederfur *et al*. 2008). However, to ensure I was using the optimal conditions for elution, I tested multiple step elution conditions at small scale, varying lactose concentrations (10-350 mM, Figure 5.8). In my hands lower concentrations of lactose gave similar yields, but no conditions gave a better yield than 150 mM. I therefore continued with using 150 mM lactose as the concentration for galectin-3 elution.

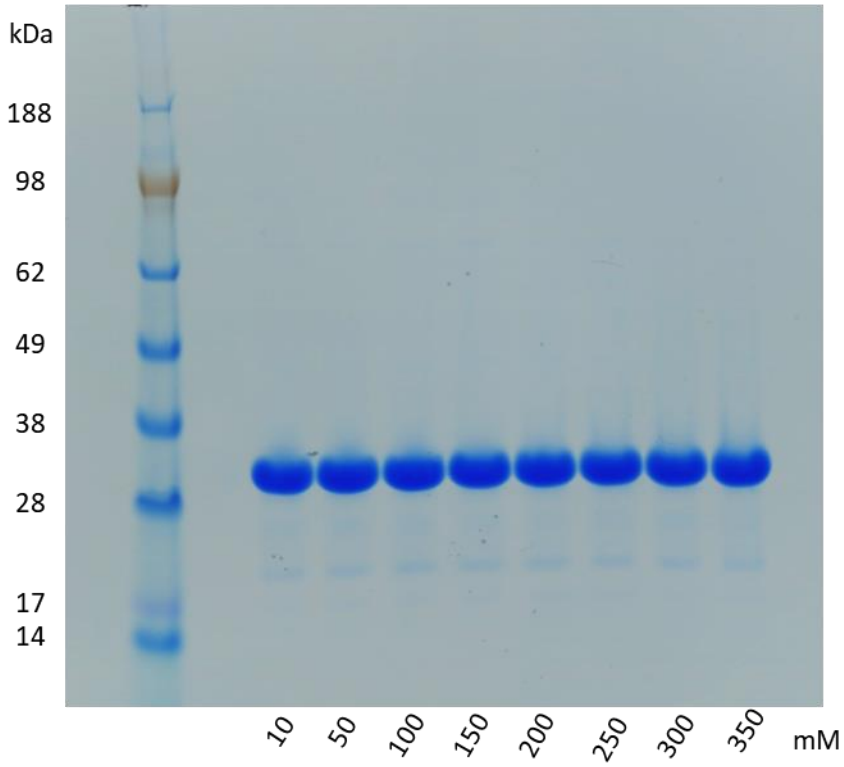


***Figure 5.8, Step elution titration experiment to maximise galectin-3 elution*.** Galectin-3 purified post-culture. The galectin-3 sample was harvested post 3 h 0.5 M IPTG induction. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. Lane 1-8 has been labelled with the appropriate final (w/v) concentration of Lactose in the Lactose-MEPBS elution buffer (on the X-axis). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Galectin-3 monomeric molecular weight is at ~31 kDa. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.3 Assessing for Concentration-Dependent Multimerisation of Galectin-3

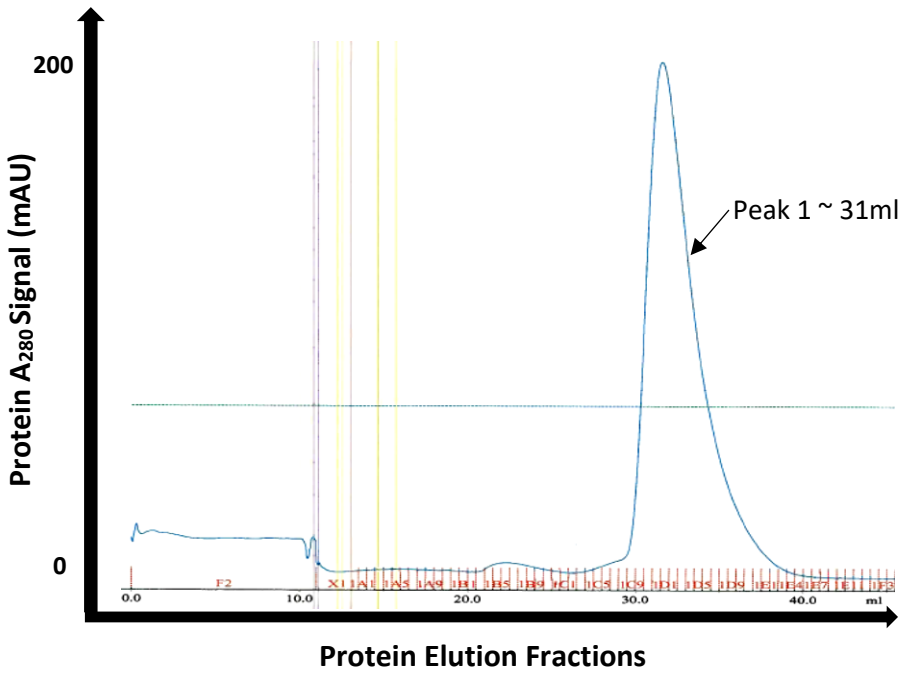
Concentration-dependent multimerisation of galectin-3 has been reported to occur in the absence of specific carbohydrate ligand binding (Ahmad *et al*. 2004, Lepur *et al*. 2012). Multimerisation will be mediated by non-covalent intermolecular interactions and so cannot be directly detected by SDS-PAGE. I therefore evaluated a number of alternative approaches to try and detect any galectin-3 multimers generated in this way.

5.2.3.1 Characterisation of Purified Galectin-3 by Size Exclusion Chromatography and Multi Angle Light Scattering (SEC-MALS)

After concentrating the purified galectin-3 sample to ≥100 µM using centrifugal filter units with 3 kDa molecular weight cut-off (Merck Life Sciences Ltd, UK), the sample was loaded onto an S400-sephacryl column. Size exclusion chromatography (SEC) was performed to assess for evidence of concentration-induced multimerisation (Figure 5.9). The sample was eluted off in a 96 well fraction collector (labelled A1-A12, B1-B12, until H12). The sample was injected into the resin at 11 ml elution fraction (X-axis). The maximum protein signal (A_{280}) peak was observed at an elution fraction of ~31 ml or collection fractions C11, C12 and D1. Fractions contributing to the peak (peak fraction C11) were assessed by multi angle light scattering (MALS) and the data were collected. MALS measures the light scattered by a sample into a plurality of angles and determines the average size of the molecules in solution. These were consistent with a species of 24.1 (±1.2%) kDa molecular weight, indicating that the purified galectin-3 (>100 µM) behaved entirely as a monomer (Table 5.1). This is slightly lower than the quoted molecular mass of galectin-3 (31 kDa) and so SDS-PAGE was undertaken to confirm that no degradation had occurred during the gel filtration (Figure 5.10). Lanes 1 and 2 were loaded with sample aliquots taken before loading onto the S400-sephacryl column, and the remaining lanes were loaded with the fractions corresponding to the peak. All bands ran consistently with an apparent molecular

weight of ~31 kDa than ~24 kDa (migrating more slowly than the 28 kDa molecular weight marker). The potential difference in size observed for galectin-3 between the MALS and SDS-PAGE data could be explained by how light is scattered within the MALS instrument. Therefore, it is used as a relative and not an absolute tool for assessing molecular weight (Nobbmann *et al.* 2007).

a)



b)

	Mw (kDa)
S75 Galectin FrC11from S400 030519	24.1 (±1.2%)
Average	24.1
Standard deviation	n/a
% Standard deviation	n/a
Minimum	24.1
Maximum	24.1

Figure 5.9, Initial Galectin-3 Characterisation. 5.9a, SEC elution profile of purified recombinant galectin-3. 200µM of Galectin-3 was loaded onto the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). The A_{280} signal peaks at ~31 ml (fractions C12 and D1 correspond to this peak). The X-axis represents the protein elution fraction(s), The Y-Axis represents the A_{280} protein signal intensity peaking at 200 mAU (arbitrary, milli-absorbance unit). Fractions were collected in a collection tub that has 12 rows (i.e A1-A12, B1-B12) with 8 columns (96 wells in total). ***5.9b,*** Table 5.1 displaying SEC-MALS analysis of fraction C11 (peak fraction from elution profile). The sample eluted at fraction C11 (corresponding to Figure 5.9a, elution profile peak and Figure 5.10, strongest band intensity on the SDS-Gel) was loaded on a S75 column for SEC-MALS analysis. The average predicted molecular weight from the ASTRA software has been screenshot and shown with a % error (24.1 kDa ± 1.2%), which displays monomeric galectin-3 species.

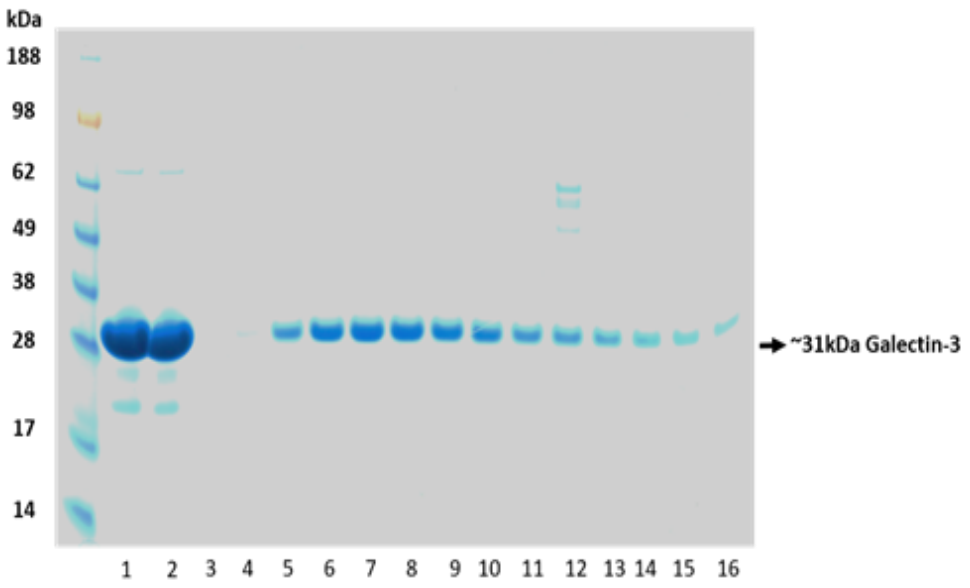


Figure 5.10, SDS-PAGE analysis of size exclusion chromatography elution fractions (from figure 5.9a). 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. The lanes were loaded as follows: Lanes 1 and 2, pre and post filter sample of galectin-3 (200 µM) before loading onto the gel-filtration column. Lanes 3-16, fractions C8-D9 (peak intensity of corresponding fractions from elution profile is C11/C12/D1, lanes 6, 7 and 8). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric Galectin-3 molecular weight at ~31 kDa has been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.3.2 Sucrose Density Gradient Ultracentrifugation of Purified Galectin-3

Next, the sample was assessed by 5%-45% (w/v) sucrose density gradient ultracentrifugation. Purified galectin-3 of ≥100 μM concentration was ultracentrifuged overnight and the distribution of galectin-3 species across the sucrose gradient was analysed by SDS-PAGE (Figure 5.11). 24 fractions were collected in total (from the top 5% (w/v) to the bottom 45% (w/v) sucrose). Every second fraction was loaded onto the gel. Consistent with the SEC-MALS analysis (Figures 5.9 a and b), galectin-3 was found in a single distribution suggesting the purified product was conformationally homogeneous in the ultracentrifugation conditions.

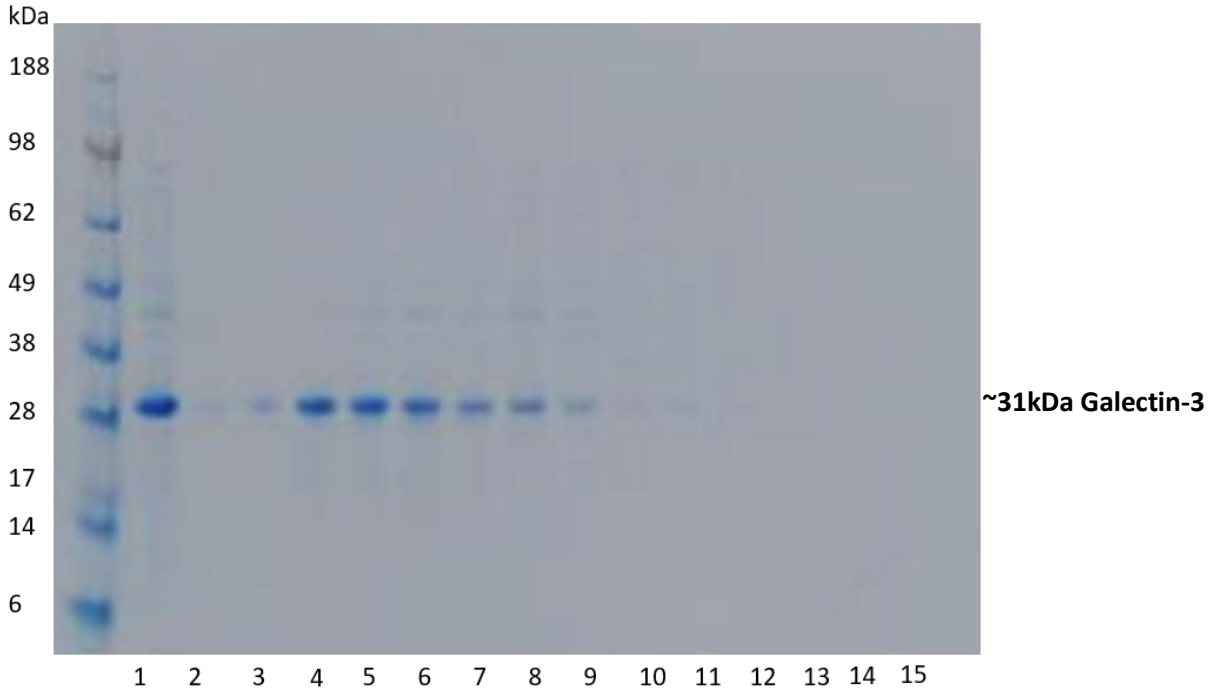


Figure 5.11, Sucrose density gradient ultracentrifugation fractionation of purified galectin-3. 4-12% (v/v) resolving gel displaying the fractions collected from a 5%-45% (w/v) sucrose density ultracentrifugation gradient. 200 µl fractions were collected from top to bottom of the centrifuged tube. The fraction collector was arranged in rows of 24 fractions. Lane 1 is loaded with galectin-3 sample aliquoted prior to ultracentrifugation (loading control). Fractions 1-9 and 13-17 were loaded (Lanes 2-15). 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer. The gel was stained overnight in Expedeon instant blue. The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric Galectin-3 molecular weight at ~31 kDa has been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.3.3 Native-PAGE Analysis of Purified Galectin-3

5.2.3.3.1 Galectin-3 Electrophoresis in Standard Native-PAGE Running Conditions

Non-denaturing (native-)PAGE can be used as a powerful tool for assessing different oligomeric states within a sample, e.g. polymerisation of members of the serpin superfamily of proteins, such as α_1-antitrypsin, which is typically well-followed by this method (Irving *et al*. 2011, Lomas *et al*. 2016). Samples in an SDS-PAGE system usually migrate as linearised species saturated with negative charge (SDS), with all non-covalent interactions denatured. Differences in anodal migration through the polyacrylamide gel matrix are therefore entirely due to protein mass. In contrast protein migration in native-PAGE is also influenced by conformation and the net charge in their native state. Larger and/or more extended species migrate more slowly. Since native α_1-antitrypsin, unlike many proteins, runs as a distinct band on native-PAGE, recombinant α_1-antitrypsin was used as a gel running control (Figures 5.12-5.13). Lanes 2-5 were loaded with increasing quantities of galectin-3. Initially, conditions typically used to assess α_1-antitrypsin including a discontinuous buffer running system were used (regarded as 'standard' native-PAGE) and the gel was run for 2 h at room temperature (Figure 5.12a). Whilst recombinant α_1-antitrypsin ran well as a characteristic doublet of compact bands, galectin-3 samples migrated very slowly. The majority of the protein failed to enter the separation gel. Some protein was clearly retained in the stacking gel and at the interface with the loading well space (Figure 5.12a). Similar results were observed in galectin-3 migration when the gel was run for 4 h instead of 2 h, though less α_1-antitrypsin was visible due to migrating out of the gel (Figure 5.12b). When the gel was run for 6 h, as expected α_1-antitrypsin completely migrated off the gel, whilst increased levels of galectin-3 did enter the gel in multiple more diffuse bands. Again however, most of the protein failed to enter the gel and was retained in the stacking gel and interface with the loading well space (Figure 5.12c). This suggests that running the gel for increased time has limited benefits in the resolution of galectin-3 migration. Western-blot analysis of each gel confirmed the presence of α_1-antitrypsin or galectin-3 in all observed bands and smears.

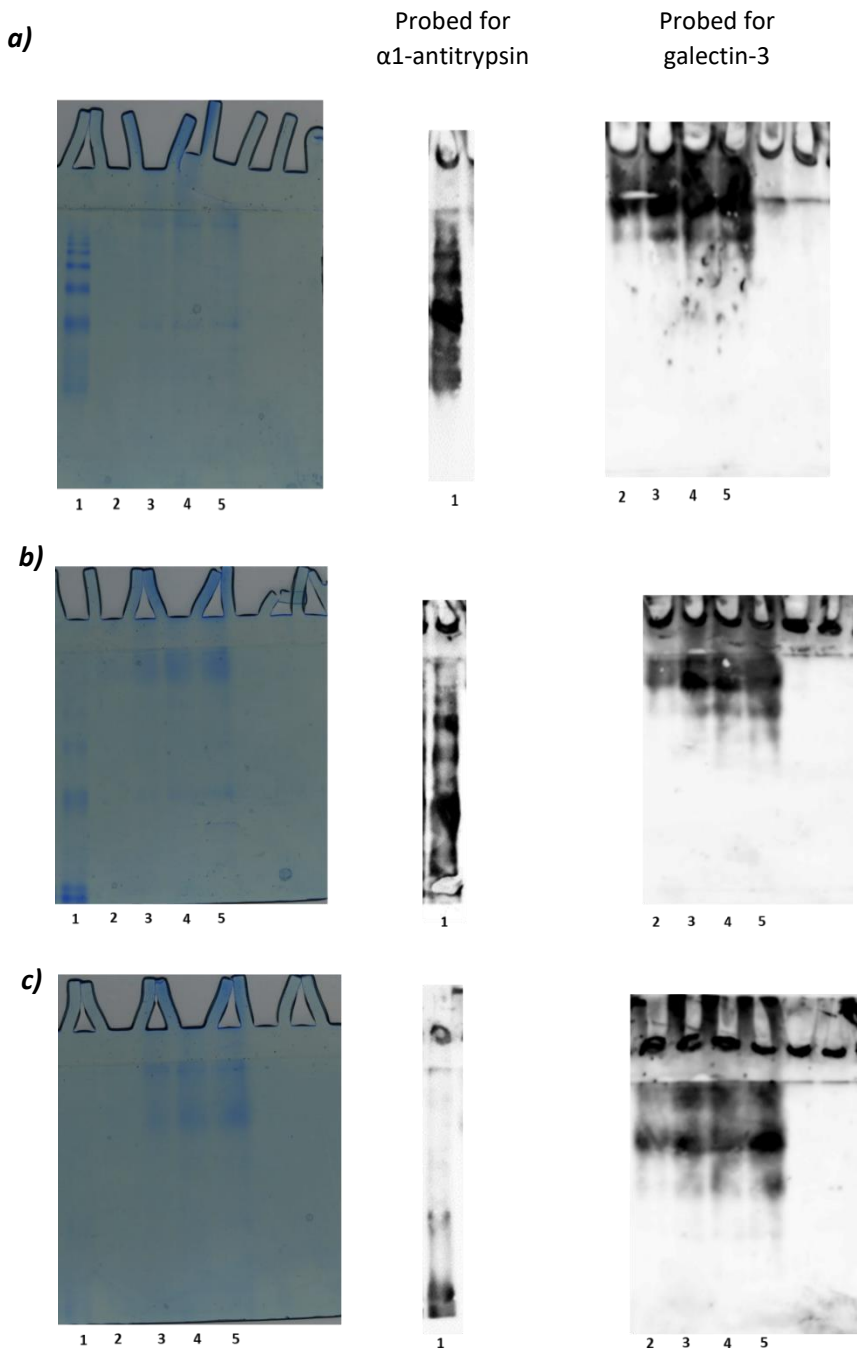


Figure 5.12, Native-PAGE analyses of galectin-3. 10% (v/v) Native-gel were run in a discontinuous buffer method at 150V for 2 h, 4 h and 6 h (***a, b, c*** respectively) at room temperature. Samples were loaded in a 1:1 ratio with 2x native sample buffer. The lanes were loaded as follows: Lane 1, 7.5 µg of α1-antitrypsin, Lanes 2-5 20 µg, 40 µg, 80 µg and 120 µg of galectin-3. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel-scanner. For the western blot the same gel was duplicated and transferred to a methanol activated 0.45 µm pore size PVDF membrane. The membrane was incubated with 5% (w/v) milk in 1x TBS-Tween-20 and immunoblotted with 1:2000 α1-antitrypsin and 0.125 µg/ml galectin-3. Lanes 1-5 were loaded (as stated for the native gel loading). The blots were immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times on the ImageQuant™LAS400.

5.2.3.3.2 Galectin-3 Electrophoresis in Adjusted Native-PAGE Running Conditions

Anodal migration of proteins on native-PAGE depends upon a net negative charge, which in turn depends upon the protein's iso-electric point (pI) and the buffer pH. pH>pI results in net negative charge, pH<pI results in net positive charge. Where pH and pI are similar, a range of ionisation states may be observed. To assess whether these considerations might explain the migration pattern observed with standard native-PAGE conditions I assessed the pI of galectin-3 using the ProtParam tool (Gasteiger *et al*. 2005) accessed through the ExPASy platform (Swiss Institute of Bioinformatics) to analyse its amino acid sequence. This gave a pI of 8.58. The pH for the upper buffer in the standard running conditions (as used in 5.2.3.3.1) is 8.2 so this would result in a slight net positive charge, and/or a range of charge states.

Strongly cationic proteins can be studied on native-PAGE by reversing the electrode polarity to observe cathodal electrophoresis. I therefore assessed what effect this would have on migration of galectin-3 in native-PAGE at a 2 h gel running duration (Figure 5.13.1). As expected, the α_1-antitrypsin loading control completely migrated upward and out of the gel (Lane 1). However, galectin-3 (Lanes 2-5, loaded with increasing concentrations) was still retained within the stacking portion of the gel and the interface with the loading well space without any migration either in or out of the gel.

Whilst running native-PAGE in these conditions, I noticed the gel running system was heating up. This was a potential confounding factor as the pH of Tris-based buffers is temperature sensitive. Subsequent experiments were therefore conducted by placing the gel-tanks in an icebox and running the gels in a 4 °C temperature-controlled cold room.

The similarities of migration of galectin-3 with and without reversal of electrode polarity suggested that the pI and pH were close enough to cause a distribution of slightly net positive and negative charge states. I therefore tried to promote a clear charge state by changing the pH of the running buffers. I first used the standard electrode polarity with a cathodic, upper running buffer pH of 9.1 and an anodic, lower running buffer pH of 8.8. As these conditions are outside the optimal buffering capacity of Tris, I doubled the Tris concentration to mitigate this. As before, α_1-antitrypsin loading control (Lane 1) ran well (Figure 5.13.2). Galectin-3 did now enter the resolving gel (Lanes 2-4), but still only migrated slowly through this portion. The same samples ran similarly with the standard Tris concentrations with the altered pH settings (Figure 5.13.3).

To complete the survey of buffer and electrode polarity conditions, I altered the pH of upper and lower buffers so that they were considerably lower than the pI of galectin-3 (7.6 and 7.2 respectively). These conditions should provide galectin-3 with a net strongly positive charge (i.e. highly cationic). The gel was then run with reversed polarity electrodes (Figure 5.13.4). α1-antitrypsin (Lane 1) migrated upward and out of the gel. Again, most galectin-3 (Lanes 2-4) entered the resolving portion of the gel, however the proteins were retained very close to the stacking gel. Overall, these findings suggested that the protein charge state distribution was only one factor determining slow migration of galectin-3 by native-PAGE. These findings alone do not exclude the loaded galectin-3 samples being oligomeric as opposed to monomeric. However, our other readouts (as summarised in earlier sections) did not provide strong evidence for this, and none of the conditions generated a robust report of the relatively large amounts of galectin-3 loaded. This is evident from the consistently poor signals for 20-80 µg galectin-3, compared with the 7.5 µg of α1-antitrypsin control protein. I therefore moved to try and determine optimal readouts of galectin-3 oligomerisation by studying conditions which generate irreversible intermolecular linkages (covalent cross-linkage) or induce galectin-3 to oligomerise at lower concentrations (ligand-induced).

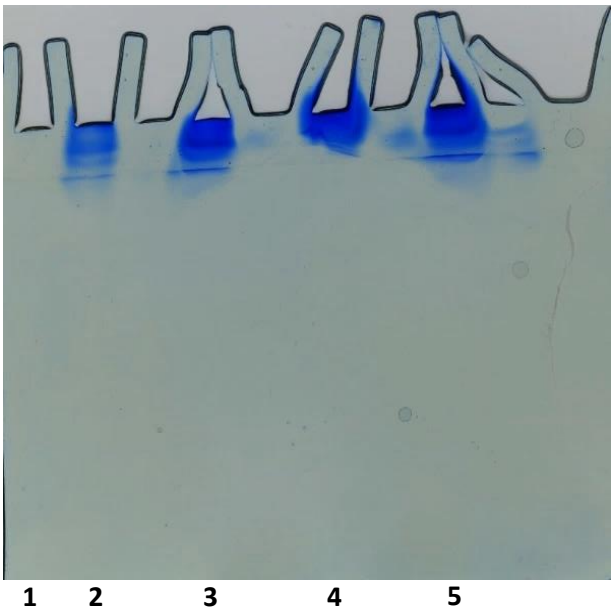


Figure 5.13.1, Reversed-polarity native-PAGE of galectin-3. 10% (v/v) native-gel. The gel was run in a discontinuous buffer method with the cathodic and anodic terminals reversed, at 150 V for 2 h at room temperature. Samples were loaded in a 1:1 ratio with 2x Native sample buffer. The lanes were loaded as follows: Lane 1, 7.5 µg of α1-anti-trypsin. Lanes 2-5 are loaded with 20 µg, 40 µg, 60 µg, and 80 µg of galectin-3. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

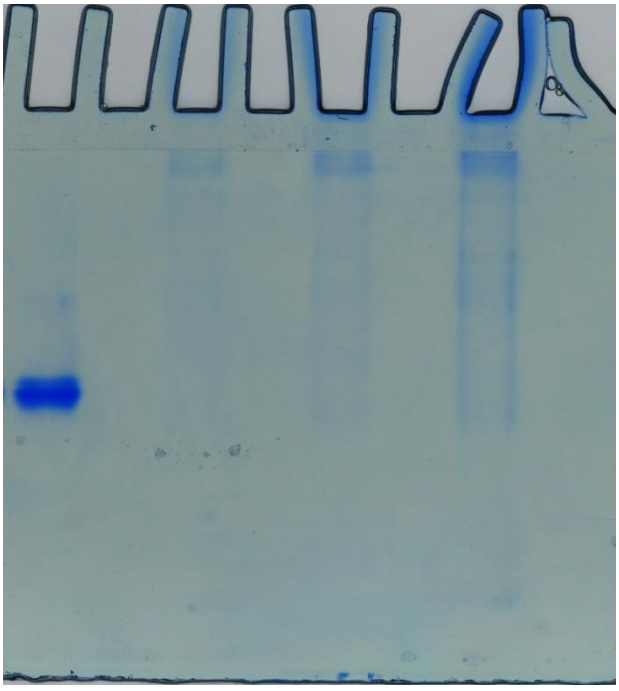


Figure 5.13.2, pH>pI native-PAGE of galectin-3 (normal polarity, higher Tris concentrations). 10% (v/v) Native gel. The gel was placed in a discontinuous buffer system with adjusted pH and doubled Tris concentrations. Cathode buffer Tris concentration was doubled (100 mM Tris, with 68 mM glycine) and the pH was adjusted to 9.1. Anode buffer Tris concentration was doubled (200 mM Tris) and the pH was adjusted to 8.8. The gel was run at 150 V for 2 h at 4 °C. Samples were loaded in a 1:1 ratio with 2x Native sample buffer. The lanes were loaded as follows: Lane 1, 7.5 µg of α1-antitrypsin. Lanes 2-4 were loaded with 20 µg, 40 µg and 80 µg of galectin-3. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

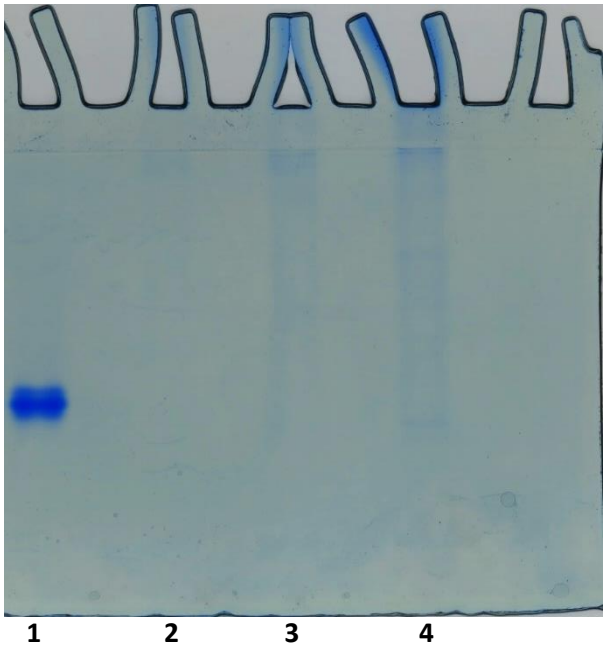


Figure 5.13.3, pH>pI native-PAGE of galectin-3 (normal polarity, with standard Tris concentrations). 10% (v/v) Native-gel. The gel was placed in a discontinuous buffer system with adjusted Tris pH. Cathode buffer Tris pH was adjusted to 9.1. Anode buffer Tris pH was adjusted to 8.8. The gel was run at 150 V for 2 h at 4 °C. Samples were loaded in a 1:1 ratio with 2x Native sample buffer. The lanes were loaded as follows: Lane 1, 7.5 µg of α1-antitrypsin. Lanes 2-4 were loaded with 20 µg, 40 µg and 80 µg of galectin-3. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

Figure 5.13.4, pH<pI native-PAGE of galectin-3 (reversed polarity). 10% (v/v) Native-gel. The gel was run in a discontinuous buffer method with the cathodic and anodic terminals reversed at 150 V for 2 h at 4 °C. Buffer pH adjusted: Cathode buffer pH was adjusted to 7.6. Anode buffer pH was adjusted to 7.2. Samples were loaded in a 1:1 ratio with 2x Native sample buffer. The lanes were loaded as follows: Lane 1, 7.5 µg of α1-anti-trypsin. Lanes 2-4 were loaded with 20 µg, 40 µg and 80 µg of galectin-3. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.4 Induced Galectin-3 Oligomerisation.

To assess whether it was possible to stabilise tractable galectin-3 oligomeric species I undertook glutaraldehyde-mediated cross-linking studies of galectin-3.

5.2.4.1 Optimisation of Glutaraldehyde Cross-linked Galectin-3

I initially optimised conditions for glutaraldehyde-mediated cross-linkage of monomeric galectin-3 by performing time course and concentration titration experiments. Cross-linking reaction was performed for durations of 1, 10, 20, and 30 minutes. Cross-linking increased with time up to 20 minutes, reported by increasing quantities of multimeric bands, but beyond this point no further increase was observed. I therefore standardised the duration of cross-linking to 20 minutes. Glutaraldehyde concentrations above 0.125% (v/v) resulted in formation of insoluble aggregates in suspension (Figure 5.14a) whilst at 0.125% (v/v) the sample-mix remained transparent. SDS-PAGE (Figure 5.14b) demonstrated an increase in covalently-linked dimer and trimer species up to this concentration whilst at higher concentrations the formation of visible aggregates coincided with loss of monomeric and oligomeric species entering the gel. I therefore proceeded to assess the profile of cross-linked species generated from 200 µM monomeric galectin-3 with 0.125% (v/v) glutaraldehyde treatment for 20 minutes.

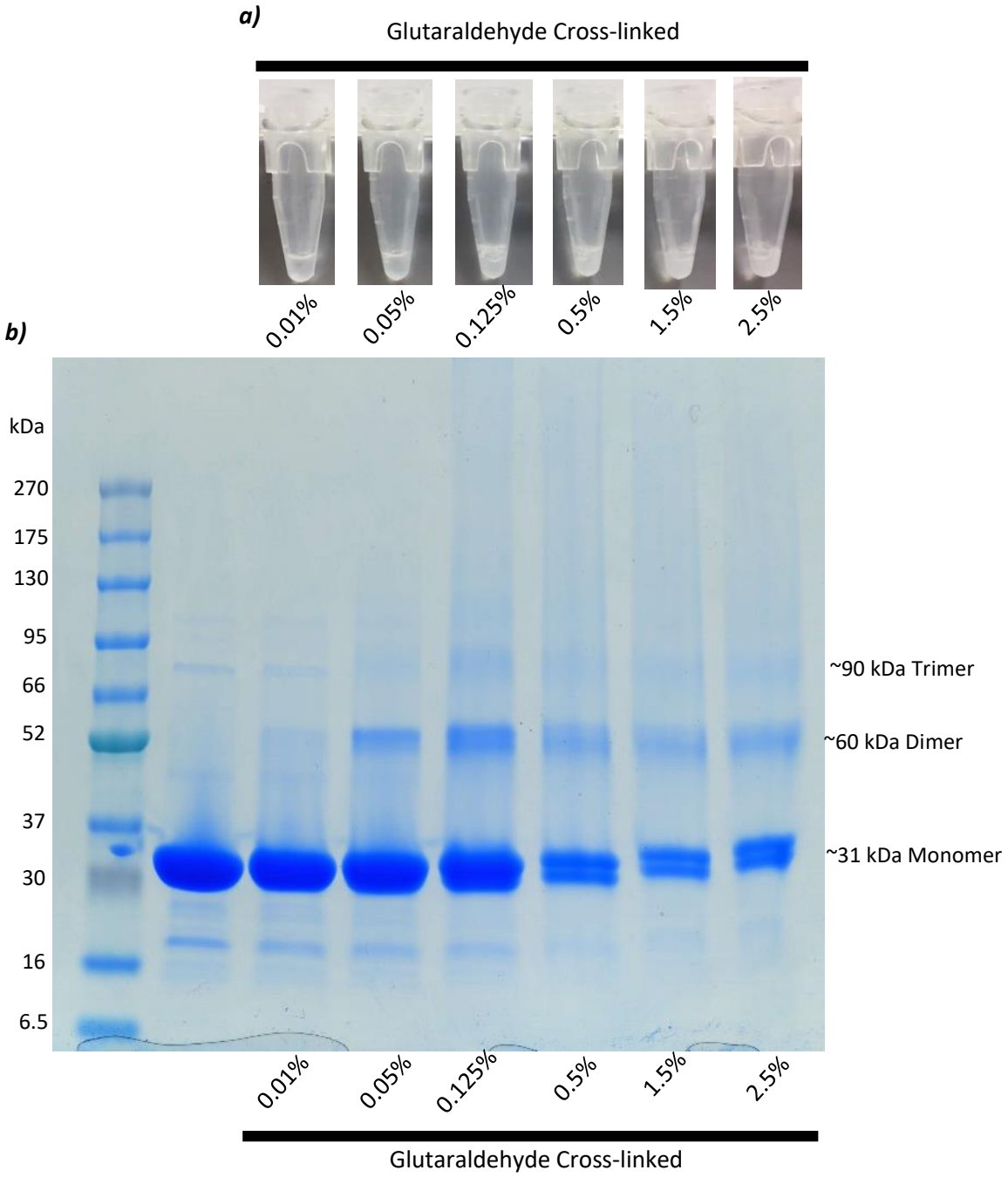


Figure 5.14, Titration experiment to optimise final glutaraldehyde concentration (v/v) yielding the best galectin-3 oligomeric species. 5.14a, Photographic images of the Eppendorf tubes in which 200µm purified galectin-3 was cross-linked with varying concentration of glutaraldehyde % (v/v). The concentration of glutaraldehyde is labelled in its final (v/v) concentration and indicated below the Eppendorf image. ***5.14b,*** 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. The lanes were loaded as follow: Lane 1, purified galectin-3 (200 µM) non-cross-linked (control). Lanes 2-7 has been labelled with the appropriate final % (v/v) concentration of glutaraldehyde (on the X-axis) used to cross-link 200 µM galectin-3. The Y-axis represents the molecular weight marker (GRS Protein Marker MultiColour PLUS). Galectin-3 monomeric and oligomeric molecular weights have been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.4.2 Size Exclusion Chromatography Characterisation of Cross-linked Species Following Glutaraldehyde Treatment of Monomeric Galectin-3

To covalently retain multimeric species, I optimised to cross-link galectin-3 for 20 minutes with a final glutaraldehyde concentration of 0.125% (v/v). SEC was performed to assess for evidence of glutaraldehyde cross-linked galectin-3 oligomers against a control sample of non-cross-linked galectin-3 (Figure 5.15 a-b).

For both cross-linked and non-cross-linked SEC profiles, the maximum protein signal (A_{280} peak) was observed at an elution fraction volume of ~31 ml or collection fractions C12, D1 and D2. In both cross-linked and non-cross-linked samples, a small hump (slight rise in A_{280}) was visible at ~22 ml elution fraction volume, though the size of this appeared larger in the cross-linked sample trace. The peak elution fractions of the cross-linked sample suggested that the eluted material was monomeric galectin-3.

SDS-PAGE (Figure 5.15 c and d) for these eluted fractions were undertaken for two reasons. Firstly, to confirm whether the samples injected into the column were successfully cross-linked, and secondly whether higher ordered molecular weight species were eluted off the column without a distinguishable A_{280} signal trace. SDS-PAGE analysis of sample aliquoted following cross-linking treatment before loading onto the S400-sephacryl column confirmed successful cross-linking with monomer, dimer and trimer bands visible. SDS-PAGE of fractions eluting in association with the small hump observed at ~22 ml (Fig. 5.15c) did not display appreciable bands by Coomassie-based staining whilst fractions eluting in association with the major peak (Figure 5.15c) demonstrated only monomeric material.

The non-cross-linked galectin-3-eluted samples were run on a gel to confirm if there were any obvious difference between the cross-linked and non-cross-linked eluted material (Figure 5.15e). The first two lanes after the marker were loaded with pre and post spin filtered sample of galectin-3 before loading onto the S400-sephacryl column (control). The remaining lanes were loaded with fractions corresponding to the small hump observed at elution fraction volume ~22 ml and eluted fractions corresponding to the peak signal observed at ~31 ml elution fraction volume. No bands were observed for fractions corresponding to the small hump at ~22 ml elution fraction volume. As expected, the peak eluted fractions displayed bands at ~31 kDa (monomeric) galectin-3 mass. Noticeably the band intensity of samples taken from non-cross-linked peak elution fractions (A_{280} signal of 200 mAU) are more intense in comparison to peak elution fractions of the cross-linked samples (A_{280} signal of 40 mAU), in agreement with the intensity of the peak profile. This suggests that the levels of monomeric galectin-3 were reduced following glutaraldehyde cross-linking treatment.

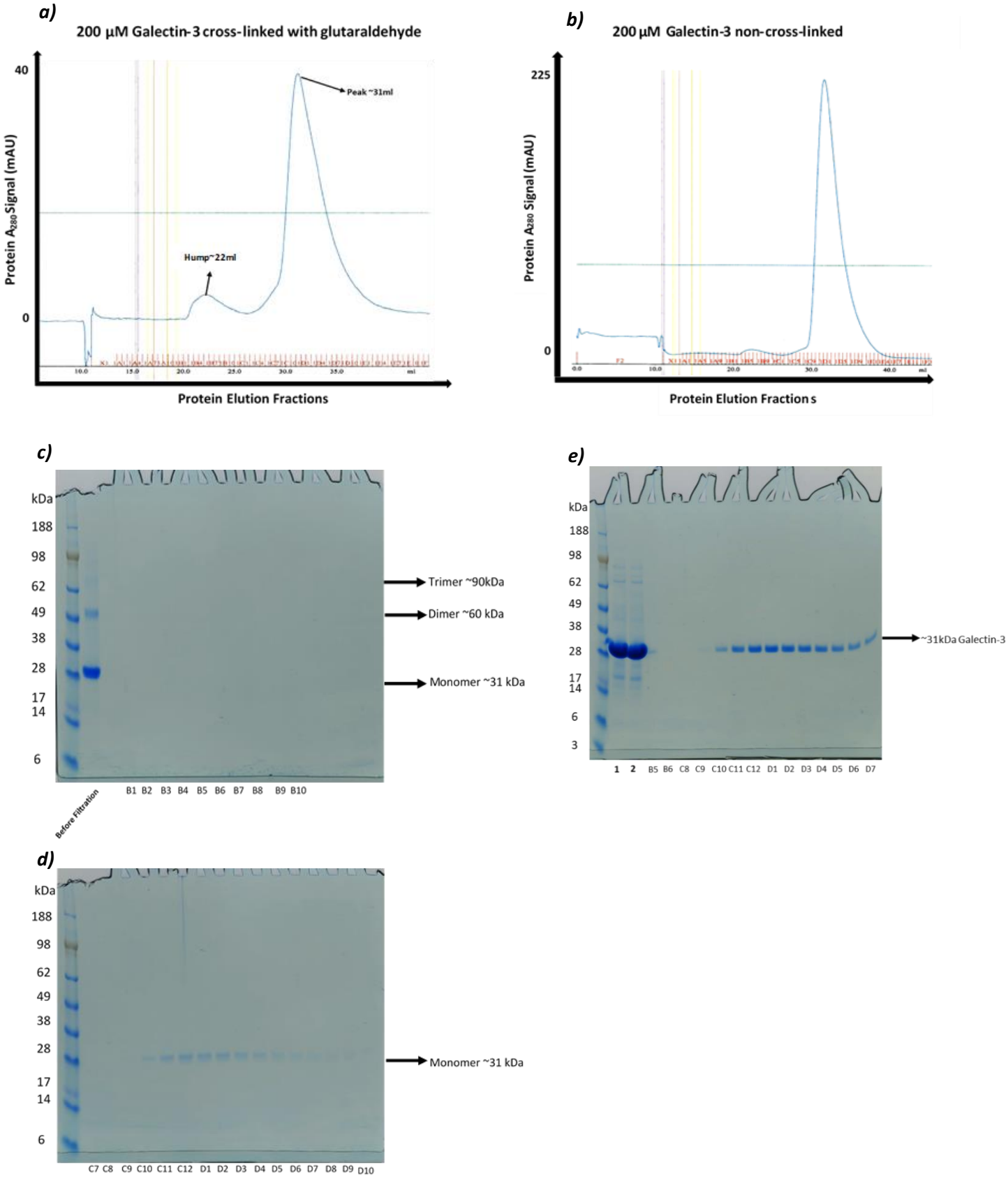

a)
200 µM Galectin-3 cross-linked with glutaraldehyde
Protein A280 Signal (mAU)
40
0
Peak ~31ml
Hump~22ml
Protein Elution Fractions
b)
200 µM Galectin-3 non-cross-linked
Protein A280 Signal (mAU)
225
0
Protein Elution Fractions
c)
kDa
188
98
62
49
38
28
17
14
6
Trimer ~90kDa
Dimer ~60 kDa
Monomer ~31 kDa
Before Filtration
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10
d)
kDa
188
98
62
49
38
28
17
14
6
Monomer ~31 kDa
C7 C8 C9 C10 C11 C12 D1 D2 D3 D4 D5 D6 D7 D8 D9 D10
e)
kDa
188
98
62
49
38
28
17
14
6
3
~31kDa Galectin-3
1 2 B5 B6 C8 C9 C10 C11 C12 D1 D2 D3 D4 D5 D6 D7

Figure 5.15, Size exclusion chromatography analysis of galectin-3 ± cross-linking. 5.15a-b, 200 µM of galectin-3 cross-linked ***(a),*** or non-cross-linked ***(b),*** and loaded onto the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). For the cross-linked sample ***(a)***, the A_{280} signal shows a small hump ~22 ml and a large peak at ~31 ml (representing protein peak), corresponding fractions C12 and D1 gives the highest intensity. For the non-cross-linked sample ***(b),*** the A_{280} peak has been shown ~31 ml (representing protein peak), corresponding fractions C12, D1 and D2 gives the highest intensity. The X-axis represents the protein elution fraction(s). The Y-Axis represents the A_{280} protein signal intensity (arbitrary, milli-absorbance unit (mAU)). SDS-PAGE of the SEC fractions of glutaraldehyde cross-linked (***5.15c-d)*** and non-cross-linked ***(5.15e)*** 200 µM galectin-3. 10 µl of each sample were loaded with 5 µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follow: ***5.15c*** Lane 1, 200 µM galectin-3 cross-linked sample, before loading on the gel-filtration column. Lanes 2-10 correspond to the eluted fractions that show a small hump seen at ~22 ml elution fraction. ***5.15d***, Lanes 1-16 were loaded with eluted fractions that correspond to the peak at ~31 ml (Fractions C11, C12 and D1 display the highest intensity in the peak). ***5.15e,*** Lanes 1 and 2, with pre and post filter sample of galectin-3 (200 µM) before loading onto the gel-filtration column. Lanes 3-4 correspond to the eluted fractions which displaying a small hump (at ~22 ml elution fraction). Lanes 5-16 is loaded with eluted fractions that correspond to the SEC peak at ~31 ml (Fractions C11, D1 and D2 display the highest intensity in the peak The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric and oligomeric galectin-3 species are indicated. The gel was stained overnight in Expedeon instant blue. Before image capture, the gel was washed twice in $_dH_2O$, and then scanned using an appropriate gel-compatible scanner.

5.2.5 Optimisation of Ligand Induced Galectin-3 Oligomerisation

Overall, work undertaken with galectin-3 at high concentration indicated that concentration-induced oligomerisation was unlikely to occur sufficiently in the conditions studied to produce useful samples for electron microscopy analysis. Galectin-3 oligomerisation is inducible at lower protein concentrations by binding of glycan ligands (Ahmad *et al*. 2004, Dam *et al*. 2005, Lepur *et al*. 2012, Bumba *et al*. 2018). I therefore went on to assess the feasibility of studying glycan-induced oligomerisation states by electron microscopy (EM) and cryo-EM to generate datasets for single particle analysis and reconstruction of oligomeric states.

5.2.5.1 Galectin-3 Induced Oligomerisation with the Glycan-Ligand N-Acetyl-D-Lactosamine (LacNAc)

To optimise the yield of LacNAc-induced galectin-3 oligomers, concentration and time-point titration experiments were performed. The starting concentrations of galectin-3 and LacNAc assessed was narrowed down from published data (Lepur *et al*. 2012).

Samples of 50 µM galectin-3 treated with LacNAc at 200 µM or 500 µM were incubated at 4 °C for 1 h, 6 h or overnight and then cross-linked for 20 minutes with 0.125% (v/v) glutaraldehyde (Figure 5.16a, lanes 4-9). Similar appearances were observed (proportional to monomeric, dimeric and trimeric masses of galectin-3) in the galectin-3 and LacNAc samples cross-linked with glutaraldehyde, regardless of the concentration of LacNAc or the sample incubation period.

Higher LacNAc concentrations of 750 µM and 1000 µM were assessed with 50 µM of galectin-3 (incubated at 4°C for 1 h and overnight) (Figure 5.16b). Again, similar bands were observed (proportional to monomeric, dimeric and trimeric masses of galectin-3) in the galectin-3 and

LacNAc samples cross-linked between these higher concentration of LacNAc and the sample incubation period. Interestingly, the dimeric band (at ~60 kDa) visually appears slightly intense in the galectin-3 bound (1000 µM) LacNAc cross-linked samples than any lower concentration, without any noticeable difference between the 1 h and overnight incubation.

As expected, individual LacNAc (monomeric mass of ~0.4 kDa) was not observable in both the non-cross-linked and cross-linked samples. Though contrary to expectations, higher multimeric bands of LacNAc were not visible in galectin-3 incubated cross-linked samples.

The samples were assessed against all relevant controls to confirm whether LacNAc induced different behaviour of galectin-3 oligomers in comparison to samples with the absence of LacNAc.

Since the dimeric band of galectin-3 (60 kDa) appeared stronger for 50 µM galectin-3 bound to 1000 µM LacNAc with no obvious difference between the incubation time of 1 h and overnight, this sample was chosen for further assessment.

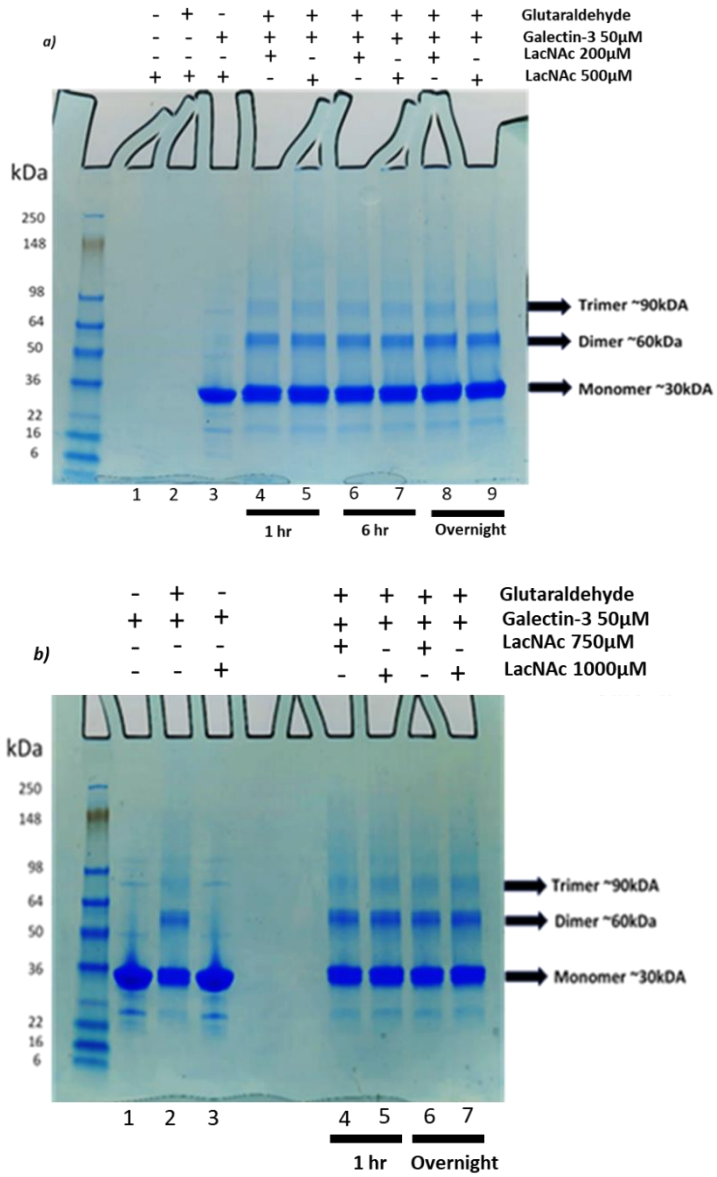


Figure 5.16, Titration experiment optimising LacNAc concentration and incubation with 50 µM galectin-3 to yield tractable oligomeric species. 10µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. ***5.16a,*** Assessing varying concentration of LacNAc incubated with galectin-3 in three different time points. The lane were loaded as follows: Lane 1 and 2 500 µM of LacNAc non-cross-linked (control) and cross-linked (control) respectively. Lane 3, 50 µM of galectin-3 and 500 µM LacNAC non-cross-linked (overnight incubation) (control). Lanes 4-9, 50 µM of galectin-3 and either with 200 µM or 500 µM LacNAc (1 h, 6 h and overnight incubation) followed by cross-linking. ***5.16b***, Assessing higher concentrations of LacNAac with varying incubation timepoints. The lanes were loaded as follows: Lane 1 and 2, 50 µM of galectin-3 (non-cross-linked and cross-linked controls). Lane 3, 50 µM of galectin-3 and 1000 µM of LacNAc non-cross-linked (overnight incubation) (control). Lane 4, 1000 µM of LacNAc cross-linked (control). Lanes 5-8, 50 µM of galectin-3 with either 750 µM or 1000 µM LacNAac (1 h or overnight incubation) followed by cross-linking. All conditions stated in both gels are displayed by the annotation on top of the gels (top X-axis). The bottom X-axis is labelled with the lane numbers, and the incubation time of LacNAc bound to galectin-3. The Y-axis on the left hand represents the marker SeeBlue™ Plus2 Pre-stained Protein Standard, and on the right hand side the monomeric and oligomeric weights have been marked. All cross-linking was done for 20 minutes at a final (v/v) concentration of 0.125% glutaraldehyde and quenched. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.5.1.2 Stochastic Behaviour When Cross-linking Galectin-3:LacNAc

Samples of 50 µM galectin-3 treated with 1000 µM LacNAc for 1 h incubation at 4°C and then cross-linked with glutaraldehyde displayed stochastic behaviour. The sample would occasionally and unreliably form precipitates once crosslinked (Figure 5.17). The precipitates were separated from the supernatant by transferring into a new Eppendorf tube. The precipitate was then resuspended with 1 ml of 1x PBS buffer and assessed by SDS-PAGE (Figure 5.17b). The resuspended pellet displayed (faint) band of monomeric galectin-3 at ~31 kDa, however a smeary continuous band is observable at higher molecular weight. This suggests that some monomeric galectin-3 was present in the precipitate and the higher multimers formed were very large and undefinable proportional to their molecular weights. Furthermore, the smear runs off from the top of the gel, potentially in agreement with the presence of very large multimers which were not able to enter the resolving portion of the gel. Due to its aggregate nature, it is not justifiable to load this resuspended pellet sample onto the gel filtration column and analyse its size exclusion chromatography profile since it risked clogging the equipment.

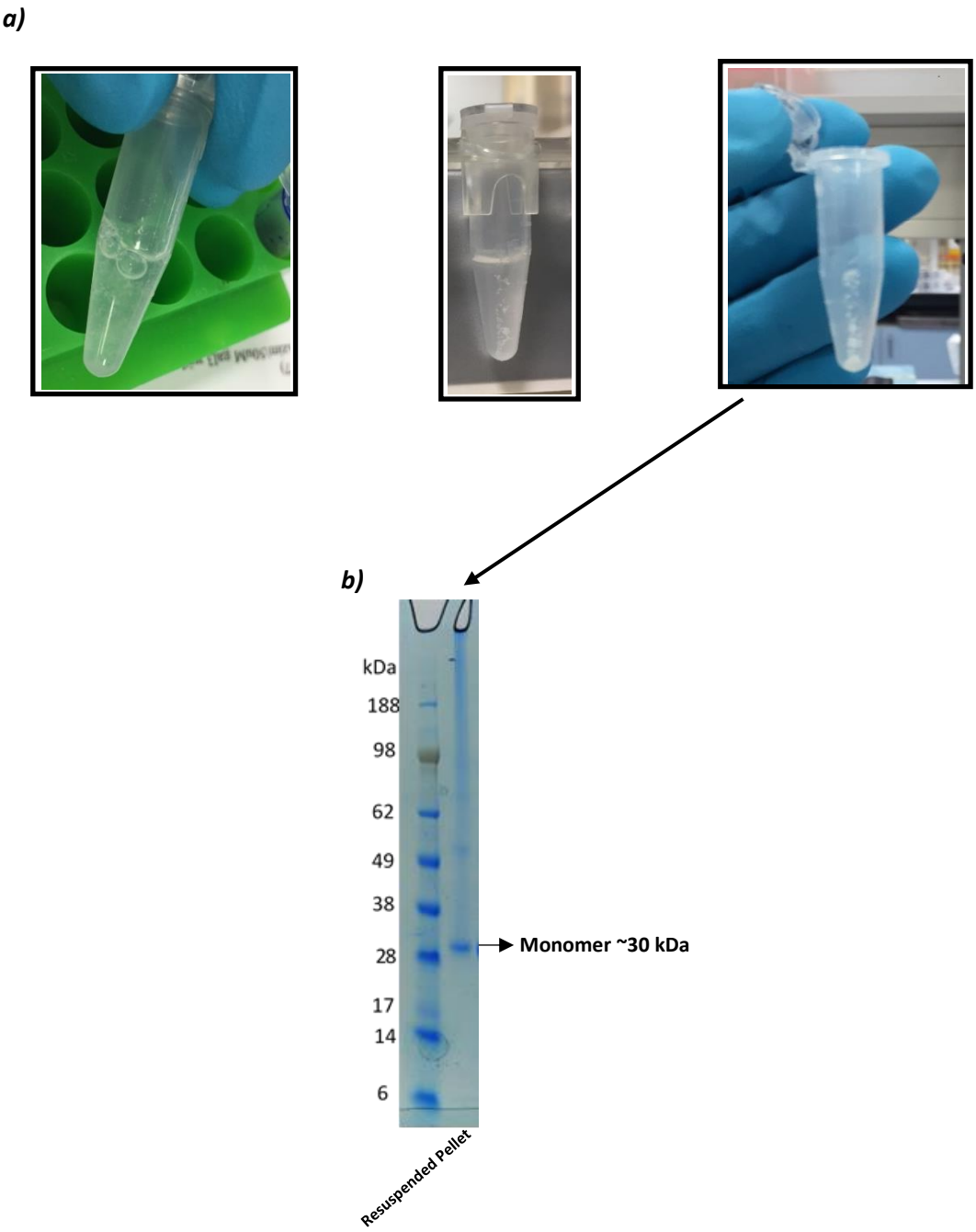
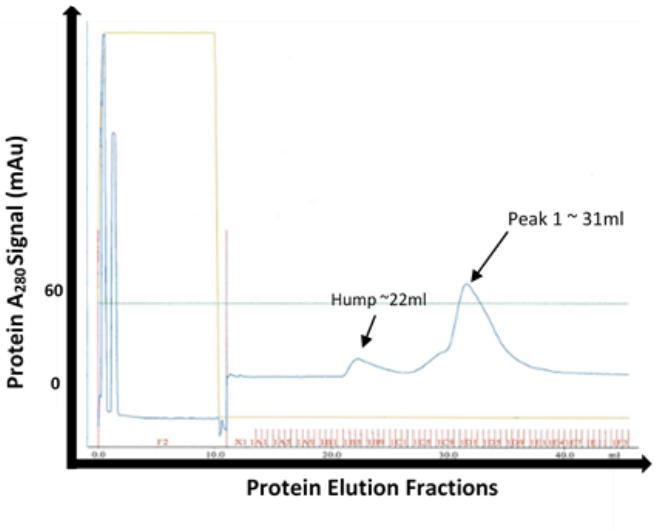


Figure 5.17, Cross-linking galectin-3 with LacNAc inconsistently forms precipitates. 5.17a) Photographic images of the Eppendorf tubes in which 50 µm of galectin-3 was incubated with 1000 µM LacNAc followed by cross-linking with glutaraldehyde at a final concentration of 0.125% (v/v) for 20 minutes. The three photographs from left to right are as follows: The left image and the middle image displays galectin-3 bound LacNAc cross-linked on different occasions, and the same phenomena of precipitation is observed. The right image displays the precipitate from the middle image without the supernatant/solution. The supernatant was transferred into a fresh Eppendorf tube. This precipitate was resuspended with 1 ml of 1x PBS buffer, and then subjected for SDS-PAGE analysis. ***5.17b)*** 10 µl of the resuspended pellet sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Galectin-3 monomeric mass has been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

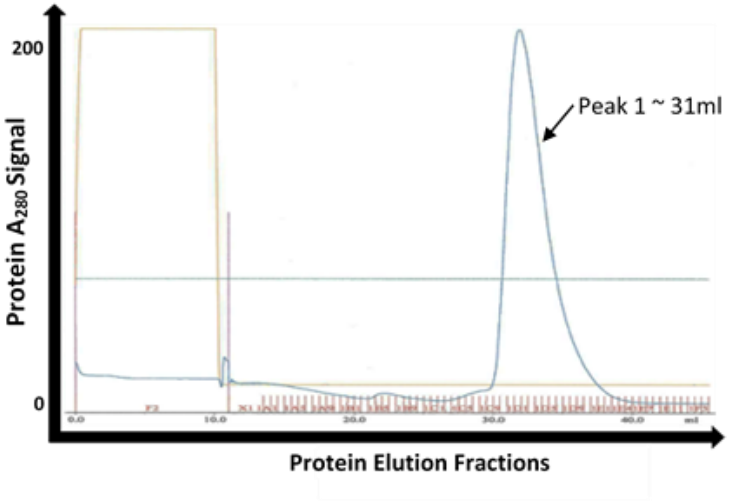
5.2.5.1.3 Characterisation of Cross-linked Galectin-3:LacNAc by Size Exclusion Chromatography

The supernatant was separated from the precipitates produced following cross-linking of galectin-3 with LacNAc and then loaded onto a S400-sephacryl column. SEC was performed to characterise galectin-3 oligomers in the supernatant (Figure 5.18a). SDS-PAGE confirmed the sample loaded onto the column contained species of monomers, dimers and trimers (Figure 5.18c). This SEC profile was assessed against the relevant monomeric control material treated with galectin-3 and LacNAc non-crosslinked (Figure 5.18b). A clear difference in the A_{280} signal intensity was apparent between the cross-linked and non-crosslinked samples. A larger peak signal intensity at ~31 ml elution fraction volume was observed in the non-cross-linked sample (arbitrary unit, 200 mAU) in comparison to the crosslinked sample (arbitrary unit, 60 mAU). This was in agreement with the stronger band intensity of the fractions from the non-cross-linked sample, which displays species of monomeric galectin-3 (Figure 5.18e). However, the overall trace profile is similar between the non-cross-linked and cross-linked sample, suggesting that the higher molecular weight species did not display a tractable A_{280} signal. SDS-PAGE analysis (Figure 5.18c-d) showed no evidence of covalently cross-linked oligomers in the elution peak samples. Interestingly, minimal monomeric material was present in the eluted material of the cross-linked sample. In contrast the initial sample loaded seemed to show high levels of monomeric, dimeric and trimeric species. This could be likely explained by the dilution factor during the collection of fractions, making the fainter higher molecular weight bands in the initial instance now negligible to stain and probe by SDS-PAGE Coomassie-staining.

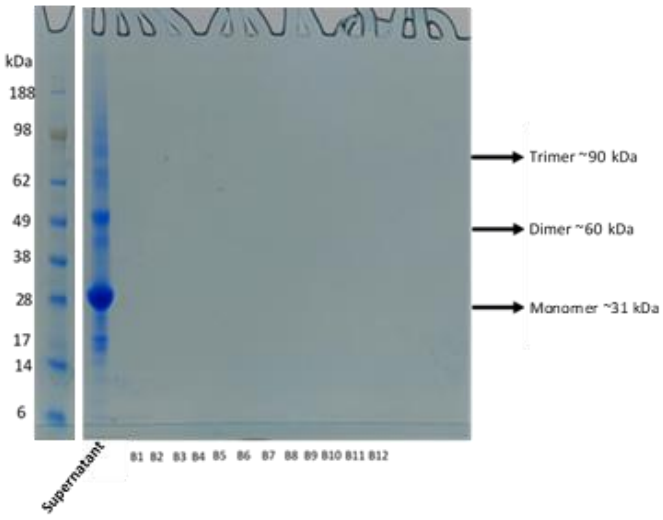
a)


50 µM Galectin-3 and 1000 µM of LacNAc cross-linked
Protein A280 Signal (mAu)
60
0
Peak 1 ~ 31ml
Hump ~22ml
Protein Elution Fractions

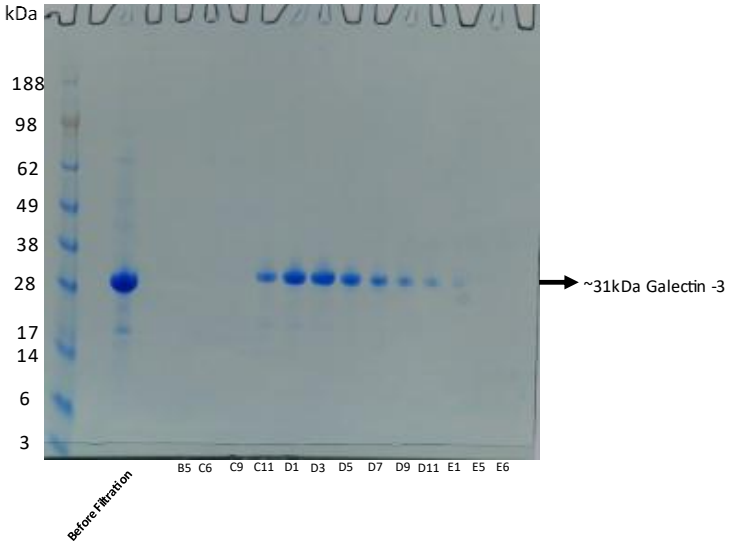
b)


50 µM Galectin-3 and 1000 µM of LacNAc non-cross-linked
Protein A280 Signal
200
0
Peak 1 ~ 31ml
Protein Elution Fractions

c)


kDa
188
98
62
49
38
28
17
14
6
Trimer ~90 kDa
Dimer ~60 kDa
Monomer ~31 kDa
Supernatant
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12

e)


kDa
188
98
62
49
38
28
17
14
6
3
~31kDa Galectin -3
Before Filtration
B5 C6 C9 C11 D1 D3 D5 D7 D9 D11 E1 E5 E6

d)

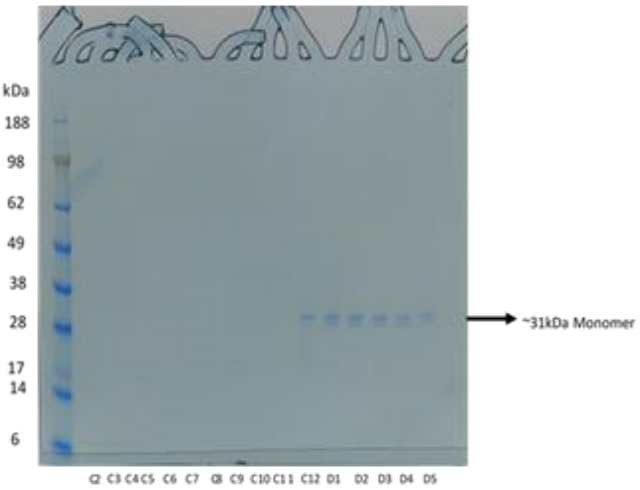

kDa
188
98
62
49
38
28
17
14
6
~31kDa Monomer
C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 D1 D2 D3 D4 D5

Figure 5.18, Size exclusion chromatography analysis of galectin-3 with LacNAc ± cross-linking. 5.18a-b, 50 µm of galectin-3 incubated with 1000 µM LacNAc cross-linked ***(a)*** or non-cross-linked ***(b)*** was loaded onto the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). For the cross-linked sample ***(a)***, the A_{280} signal shows a small hump at ~22 ml and a larger peak at ~31 ml (representing protein peak) corresponding fractions are C12, D1 and D2. For the non-cross-linked sample ***(b)*** The A_{280} shows a large peak at ~31 ml (representing protein peak) corresponding fractions are C12, D1 and D2. The X-axis represents the protein elution fraction(s). The Y-Axis represents the A_{280} protein signal intensity (arbitrary, milli-absorbance unit (mAU)). SDS-PAGE of the SEC fractions of glutaraldehyde cross-linked (***5.18c-d)*** and non-cross-linked ***(5.18e)*** 50 µM galectin-3 with 1000 µM LacNAc. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.18c***, Lane 1, supernatant of cross-linked 50 µM galectin-3 incubated with 1000 µM LacNAc for 1 h before loading onto the gel-filtration column. Lanes 2-13, correspond to the eluted fractions that show a small hump seen at ~22 ml elution fraction. ***5.18d***, Lanes 1-16 is loaded with eluted fractions that correspond to the peak at ~31 ml (Fractions C12, D1 and D2 display the highest intensity in the peak). ***5.18e***, Lane 1, 50 µM of galectin-3 incubated with 1000 µM non-cross-linked sample before loading onto the gel-filtration column. Lanes 2-14 is loaded with eluted fractions that correspond to the SEC peak at ~31 ml (Fractions C12, D1 and D2 display the highest intensity in the peak). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric and oligomeric galectin-3 molecular weight has been marked (kDa). Before image capture, the gel was washed twice in $_dH_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.6 Optimisation of Galectin-3 Induced Oligomerisation with the Glycoprotein Ligand Asialofetuin

Since cross-linked galectin-3 with LacNAc was proving to be unreliable in sample preparation and did not display any tractable multimeric samples of interest, I decided to assess and induce galectin-3 multimerisation with a different ligand. Based on literature (Dam *et al*, 2005, Lepur *et al*. 2012) I chose a high affinity galectin-3 binding partner, asialofetuin (ASF), a ~45 kDa glycoprotein containing nine branched LacNAc chains.

I first characterised the interaction of galectin-3 with ASF by performing time-course and concentration titration experiments. The starting concentrations of galectin-3 and ASF assessed was narrowed down from published data (Dam *et al*. 2005, Lepur *et al*. 2012, Bocker and Elling. 2017).

Galectin-3 was incubated with ASF at 4°C. Cross-linked multimers were apparent by 30 minutes and increased to 1 h but beyond this the cross-linkage profile detected by SDS-PAGE did not change. I therefore standardised the duration of galectin-3 and ASF incubation to 1 h for binding prior to cross-linking. The samples were assessed against all relevant controls to confirm whether ASF induced different behaviour of galectin-3 oligomers in comparison to samples with the absence of ASF (Figure 5.19).

Titration experiments were conducted varying concentrations of galectin-3 and ASF in turn, prior to cross-linking with glutaraldehyde treatment. The strongest cross-linked oligomer signal was achieved when 20 µM galectin-3 was incubated with 100 µM ASF for 1 h at 4°C. These conditions were therefore used for subsequent experiments.

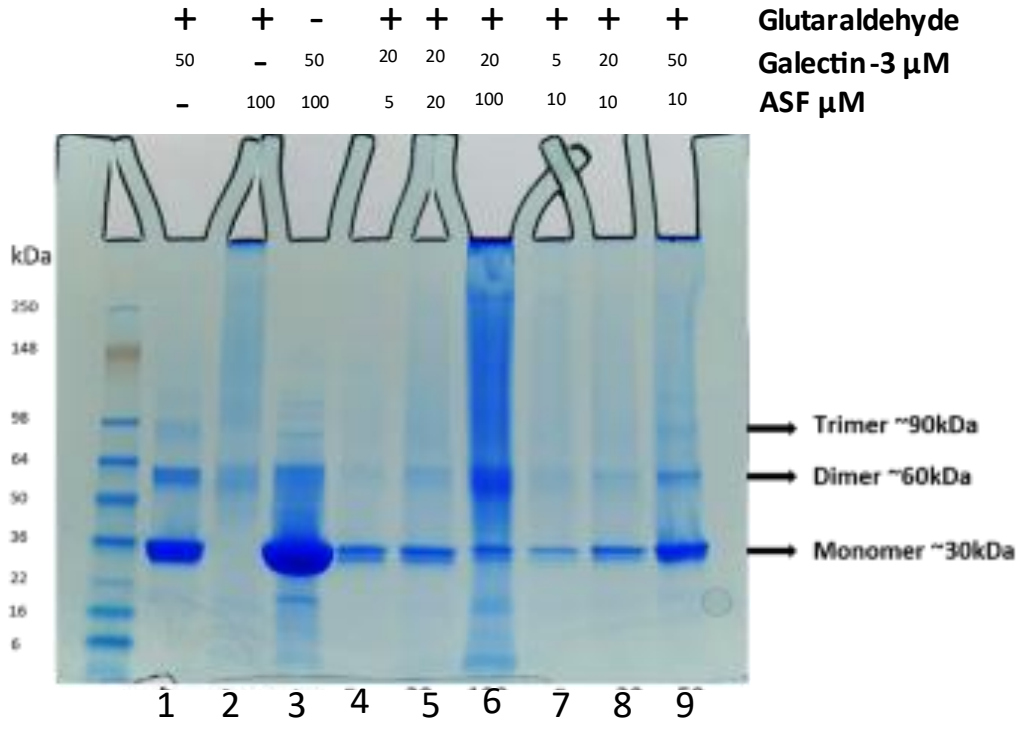


Figure 5.19, Titration experiment optimising galectin-3 and ASF concentrations at 1 h incubation to assess best yielded oligomeric species. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. The lanes were loaded as follows: Lane 1, 50 µM of galectin-3 cross-linked (control). Lane 2, 100 µM of ASF cross-linked (control). Lane 3, 50 µM of galectin-3 and 100 µM of ASF non-cross-linked (control). Lanes 4-6 is loaded with fixed concentration of galectin-3 (20 µM) and varying concentration of ASF (5 µM, 20 µM and 100 µM), subjected to cross-linking. Lanes 7-9 is loaded with fixed concertation of ASF (10 µM) and varying concentration of galectin-3 (5 µM, 20 µM and 50 µM). All conditions loaded into each well are labelled on top of the gel. The Y-axis represents the molecular weight marker SeeBlue™ Plus2 Pre-stained Protein Standard. Monomeric and oligomeric galectin-3 molecular weights have been marked. All cross-linking was done for 20 minutes at a final (v/v) concentration of 0.125% glutaraldehyde. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.6.1 Size Exclusion Chromatography of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking

I wished to understand how any size exclusion chromatography profiles of cross-linked galectin-3:ASF complexes would be independent of galectin-3:galectin-3 cross-linkage, or ASF:ASF cross-linkage. Galectin-3 cross-linkage SEC profiles are already reported in section 5.2.4.2. Overall, this displayed a peak A_{280} signal at a similar elution fraction volume to non-cross-linked galectin-3 control. Furthermore, whilst SDS-PAGE analysis confirmed that oligomeric sample was loaded onto the S400-sephacryl column, the fractions eluted off the column corresponded to bands agreeing with monomeric species.

ASF was assessed without (Figure 5.20 a,c) and with (Figure 5.20 b,d) the same glutaraldehyde treatment, by the same technique. As expected, SEC profile of the non-cross-linked samples reported a single elution peak at ~27 ml elution fraction volume. In addition, SDS-PAGE analysis of this peak fractions displayed purely monomeric migration patterns of ASF. Cross-linked ASF samples displayed a SEC profile similar to non-cross-linked control with a single elution peak at ~27 ml elution fraction volume. However, SDS-PAGE analysis of this peak eluted fractions displayed glutaraldehyde-linked dimer and higher molecular weight oligomers formed due to the cross-linking treatment.

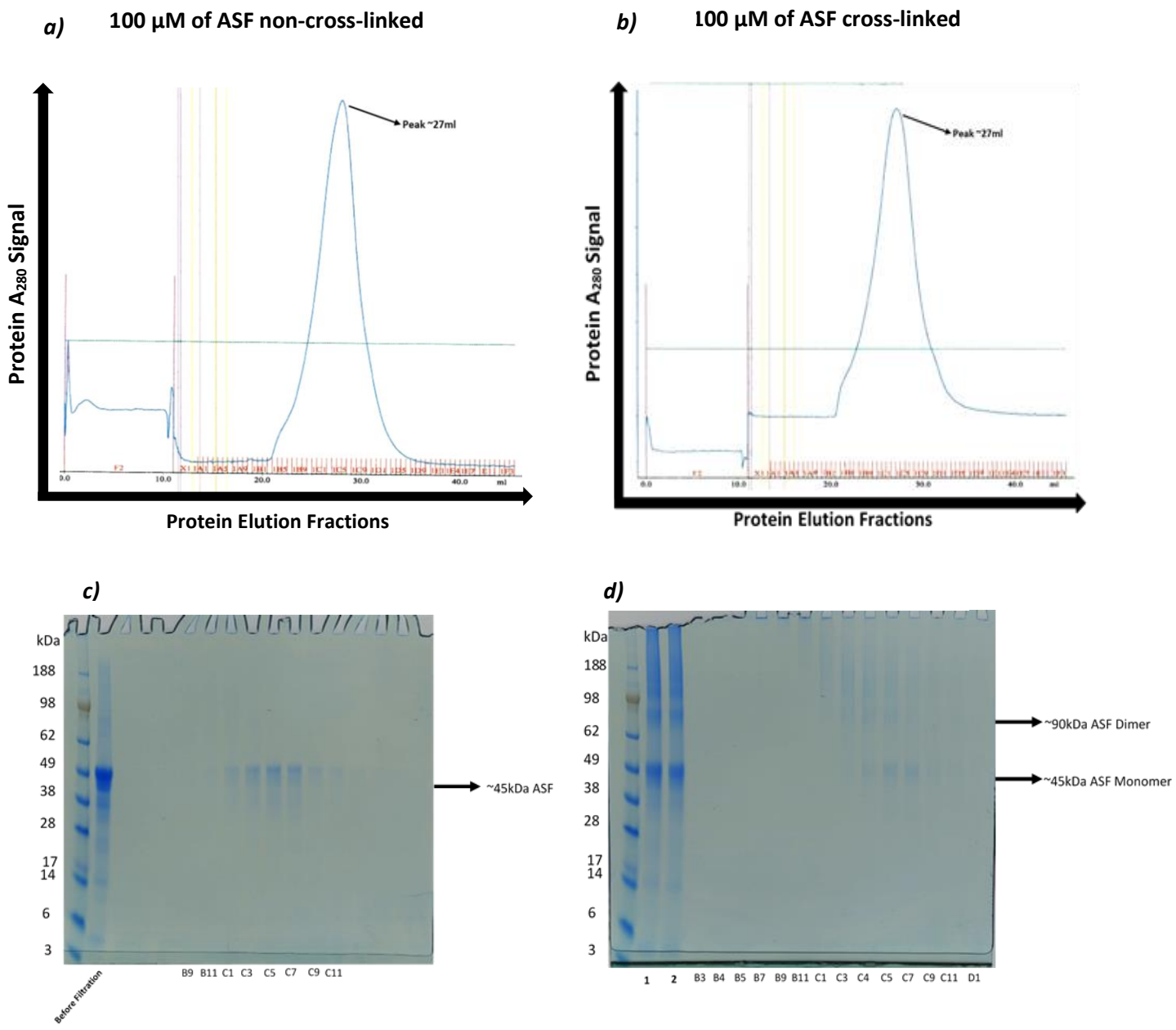


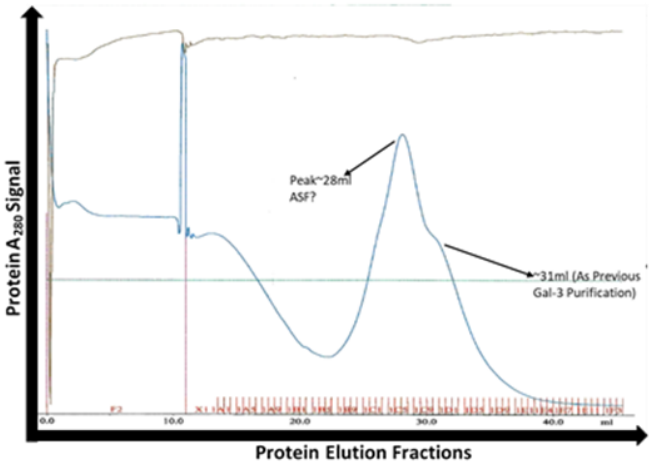
Figure 5.20, Size exclusion chromatography analysis of ASF ± crosslinking. 5.20a-b, 100 µM of ASF non-cross-linked ***(a)*** or cross-linked ***(b)*** was loaded on the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). The A_{280} signal displays a large peak in both profiles at ~27 ml (representing protein peak) corresponding fractions are C4, C5 and C6 for ***(a)*** non-cross-linked and ***(b)*** C3, C4 and C5 for crosslinked. The X-axis represents the protein elution fraction(s), The Y-Axis represents the A_{280} protein signal intensity (arbitrary, milli-absorbance unit (mAU)). SDS-PAGE of the SEC fractions of non-cross-linked (***5.20c)*** and glutaraldehyde cross-linked ***(5.20d)*** 100 µM ASF. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.20c*** Lane 1, 100 µM of non-cross-linked ASF before loading onto the gel-filtration column. Lanes 2-9, is loaded with eluted fractions that correspond to the peak at ~27 ml (Fractions C4, C5 and C6 display the highest intensity in the peak). ***5.20d***, Lanes 1 and 2, pre and post filter sample of 100 µM ASF cross-linked before loading onto the gel-filtration column. Lanes 3-16 is loaded with eluted fractions that correspond to the peak at ~27 ml (Fractions C3, C4 and C5 display the highest intensity in the peak). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric and oligomeric ASF molecular weight has been marked (kDa). Before image capture, the gel was washed twice in $_dH_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.6.2 Size Exclusion Chromatography Characterisation of Cross-linked Species Formed due to Galectin-3:ASF Interactions

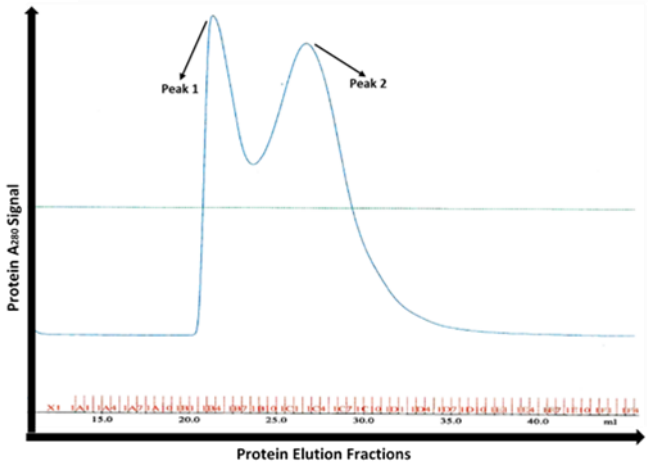
Cross-linked galectin-3:ASF oligomers were characterised by size exclusion chromatography using a S400-sephacryl column, combined with multi-angle light scattering (MALS, i.e. SEC-MALS in combination).

Non-cross-linked sample ran with a dominant peak at ~27 ml elution fraction volume which corresponded to a molecular weight of 45 kDa, consistent with monomeric ASF, and a shoulder at ~31 ml elution fraction volume consistent with the profile for monomeric galectin-3 (Figure 5.21a and 5.21c-d). However, cross-linking the sample shifted the profile into two distinct peaks (Figures 5.21b and 5.21e-f). SEC-MALS analysis (Table 5.2) indicated the fraction corresponding to the first peak (peak 1) was associated with a molecular mass of >66 MDa. This falls outside the elution range of the S400 sephacryl resin, though importantly it signifies the production of large multimeric species. Whilst the fraction corresponding to the second peak (peak 2) was associated with a molecular mass of 628 kDa. Consistently, no species were evident for peak 1 fractions in the SDS-PAGE conditions we used (Figure 5.21e). A small pellet formed when galectin-3:ASF was cross-linked, which also did not yield any protein species small enough to enter the gel (Figure 5.21e). In contrast, bands consistent with monomeric ASF and higher molecular weight oligomers were apparent for samples from peak 2 fractions (Figure 5.21f), suggesting the presence of non-covalent and covalently linked subunits within the species reported by this peak. Furthermore, monomeric galectin-3 bands are barely visible in peak 2 fractions, suggesting that cross-linking galectin-3 with ASF shifted these into higher ordered multimers, which is in agreement with the SEC elution profile of the ~31 ml shoulder (displayed in the non-cross-linked control) is diminished. The galectin-3:ASF cross-linked SEC profile peak 2 overlaps with the ASF cross-linked control peak. However, whilst eluted fractions between these two display overlapping bands at similar eluted fraction volume, there are additional larger molecular weighted species present in the peak 2 fractions of galectin-3:ASF cross-link sample, observed in earlier eluted fractions.

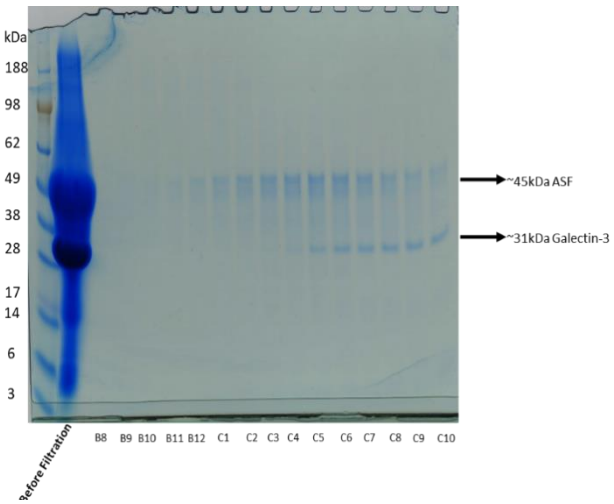
a) 20 µM Galectin-3 and 100 µM of ASF non-cross-linked



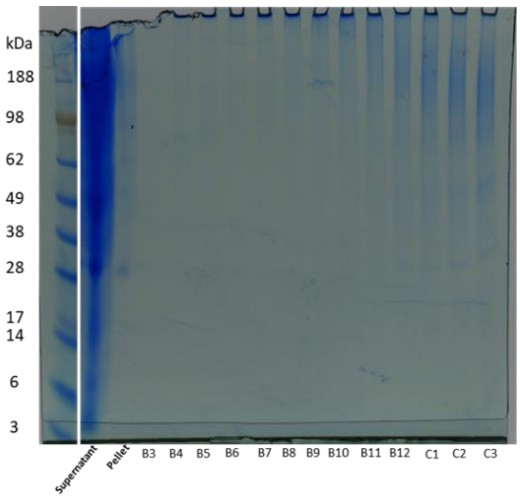
b) 20µM Galectin-3 and 100 µM of ASF glutaraldehyde cross-linked



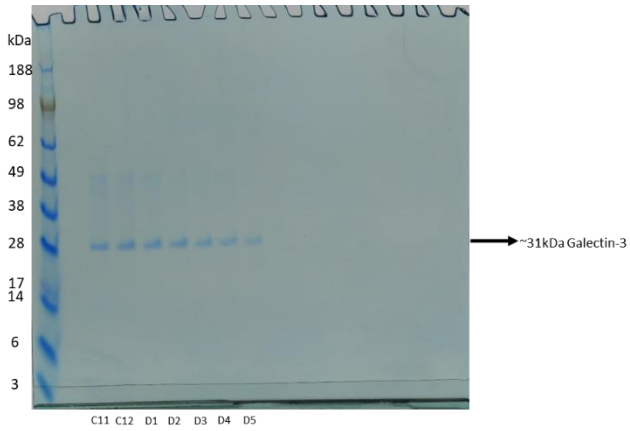
c)



e)



d)



f)

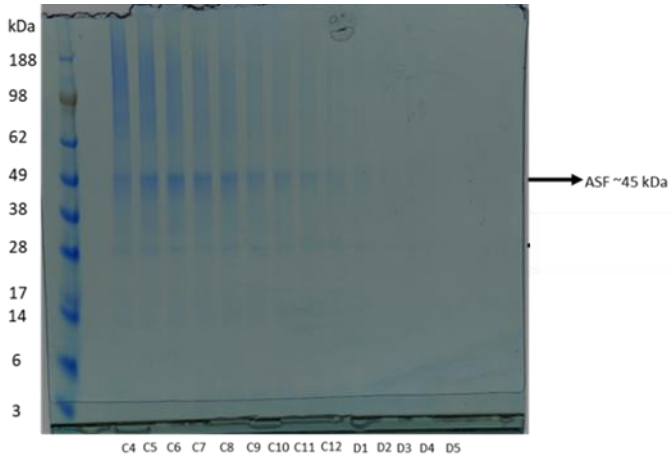


Figure 5.21, Size exclusion chromatography analysis of galectin-3 with ASF ± cross-linking. 5.21a-b, 20 µm of galectin-3 incubated with 100 µM ASF non-cross-linked ***(a)*** or cross-linked ***(b)*** was loaded onto the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). For the non-cross-linked sample ***(a)***, two merged A_{280} peaks are visible. The first peak is at ~28 ml with fractions C4 and C5 corresponding to the peaks highest intensity. The second shoulder peak is at ~31 ml, with fractions C10, C11 and C12 corresponding to the peaks highest intensity. For the cross-linked sample ***(b),*** two distinct A_{280} peaks are visible. The first peak is at ~22 ml with fractions B3, B4 and B5 corresponding to the peaks highest intensity. The second peak is at ~28 ml with fractions C3, C4 and C5 corresponding to the peaks highest intensity. The X-axis represents the protein elution fraction(s). The Y-Axis represents the A_{280} protein signal intensity (arbitrary, milli-absorbance unit (mAU)). SDS-PAGE of the SEC fractions of non-cross-linked (***5.21c-d)*** and glutaraldehyde cross-linked ***(5.21e-f)*** 20 µM galectin-3 with 100 µM ASF. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.21c,*** Lane 1, 20 µM of galectin-3 and 100 µM of ASF non-cross-linked (control), before loading on the gel-filtration column. Lanes 2-16, correspond to the eluted fractions of Peak 1 at ~28 ml (fractions C3 and C4 display the highest intensity in the peak). The gel was then loaded gradually with fractions corresponding to the second shoulder peak at ~31 ml. ***5.21d***, Lanes 1-7 is a continuation of eluted fractions corresponding to shoulder peak at ~31 ml (fractions C11 and C12 display the highest intensity in the peak). ***5.21e***, Lane 1, supernatant of 20 µM galectin-3 and 100 µM ASF cross-linked control, before loading on the gel-filtration column. Lane 2, pellet resuspension (tiny pellet formed when 20 µM galectin-3 and 100 µM ASF incubated samples were cross-linked). Lanes 3-15 correspond to the eluted fractions of peak 1 at ~22 ml (Fractions B3, B4 and B5 display the highest intensity in the peak). The gel was then loaded gradually with fractions corresponding to peak 2 at ~28 ml. ***5.21f,*** Lanes 1-14 is a continuation of eluted fractions corresponding to Peak 2, (fractions C3, C4 and C5 display the highest intensity in the peak). The Y-axis represents the molecular weight marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Monomeric and oligomeric galectin-3 molecular weight has been marked (kDa). Before image capture, the gel was washed twice in $_dH_2O$, and then scanned using an appropriate gel compatible scanner.

Peak 1 (Fraction B4)

	Mw (kDa)
S400 Fraction B4 Gal3 and ASF XLINKED 19th August 2019	66423.7 (±0.6%)
Average	66423.7
Standard deviation	n/a
% Standard deviation	n/a
Minimum	66423.7
Maximum	66423.7

Peak 2 (Fraction C3)

	Mw (kDa)
S400 Fraction C3 Gal3 and ASF XLINKED 19th August 2019	628.0 (±0.7%)
Average	628.0
Standard deviation	n/a
% Standard deviation	n/a
Minimum	628.0
Maximum	628.0

Table 5.2, SEC-MALS analysis on galectin-3 and ASF cross-linked eluted fractions from both peak 1 and 2, The sample eluted at fraction B4 (corresponding to ***Figure 5.21b*** elution profile peak 1) was loaded on a S400 column for SEC-MALS analysis. The average predicted molecular weight from the ASTRA software has been screenshot and shown with a % error (66423.7 kDa ± 0.6%). The sample eluted at fraction C3 (corresponding to ***Figure 5.21b*** elution profile peak 2, and ***Figure 5.21f*** strongest band intensity on the SDS-Gel) was loaded on a S400 column for SEC-MALS analysis. The average predicted molecular weight from the ASTRA software has been screenshot and shown with a % error (628.0 kDa ± 0.7%).

5.2.7 Characterisation of Cross-linked Galectin-3:ASF Oligomeric Species using Sucrose Density Gradient Ultracentrifugation

To fractionate cross-linked galectin-3:ASF species more precisely by molecular weight, this sample was assessed by 5%-45% (w/v) sucrose density gradient ultracentrifugation. The aim was to improve sample homogeneity for EM structural studies. Relevant controls were also assessed and are described first. In all cases, after overnight ultracentrifugation, I analysed the distribution of the species across the sucrose gradient using SDS-PAGE. In total 24 fractions (of sequential 200 µl aliquots) were collected from the top (5% (w/v)) to the bottom (45% (w/v)) sucrose gradient.

5.2.7.1 Sucrose Density Gradient Ultracentrifugation of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking

In the absence of cross-linkage (Figure 5.22a) galectin-3 migrated as a monomer on SDS-PAGE in a single distribution consistent with homogeneously monomeric behaviour in these conditions.

Cross-linked galectin-3 migrated predominantly as a monomer but showed evidence of trace amounts of dimeric and trimeric galectin-3 species at higher sucrose density (Figure 5.22b). No higher molecular weight bands were visible, suggesting the sample did not multimerise past the trimeric level, or that these species were too large to enter the gel.

As expected, non-cross-linked ASF (Figure 5.23a) migrated mainly in a single distribution consistent with entirely monomeric behaviour. However, faint bands at lower molecular weight than ASF is visible. Potentially the high-speed force applied during the ultracentrifugation spin could have caused slight degradation of ASF through the shedding of glycans. Cross-linkage of ASF resulted in fractionation across a range of oligomeric states (Figure 5.23b), with a ‘staircase’ pattern being apparent from the low to high sucrose gradient distributed fractions.

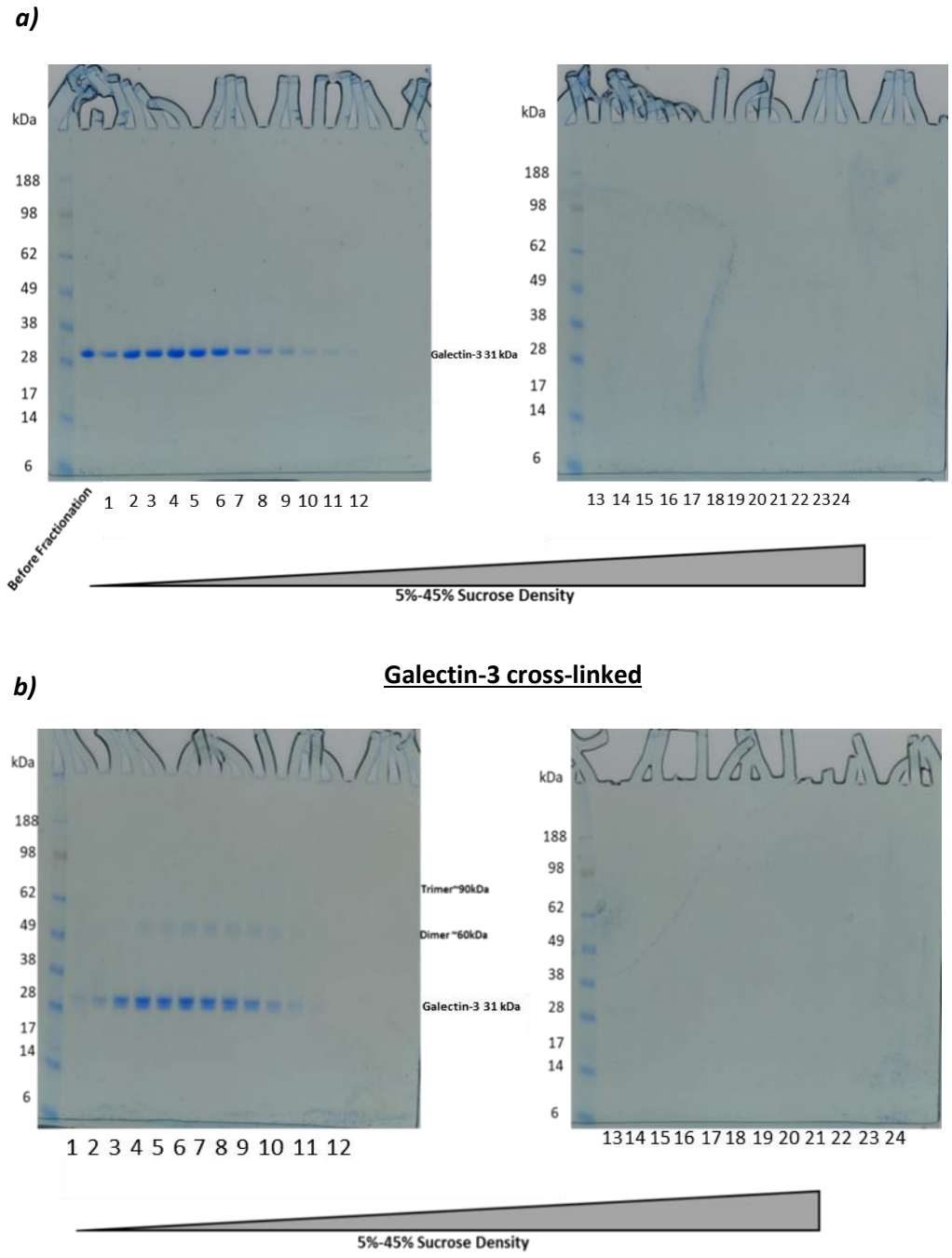


Figure 5.22, Sucrose density gradient ultracentrifugation profile of galectin-3 ± cross-linking. Sucrose gradient analysis of non-cross-linked galectin-3 (***5.22a***) and cross-linked galectin-3 (***5.22b***). Sequential 200 µl fractions were collected across the 5%-45% (w/v) sucrose density gradient. A total of 24 fractions were collected. 10 µl of each sample was loaded with 5µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.22a,*** Lane 1 pre-ultracentrifuged non-cross-linked galectin-3 (control). ***5.22b,*** Lane 1 pre-ultracentrifuged cross-linked galectin-3 (control). The remainder of the lanes in both of the non-cross-linked and cross-linked gels were loaded with fractions 1-24 (distributed between the two gels, and labelled on the X-axis according to the fraction loaded). The gels were stained overnight in Expedeon instant blue. The Y-axis denotes the molecular mass values of the marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Galectin-3 monomeric and oligomeric molecular weights (kDa) have been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

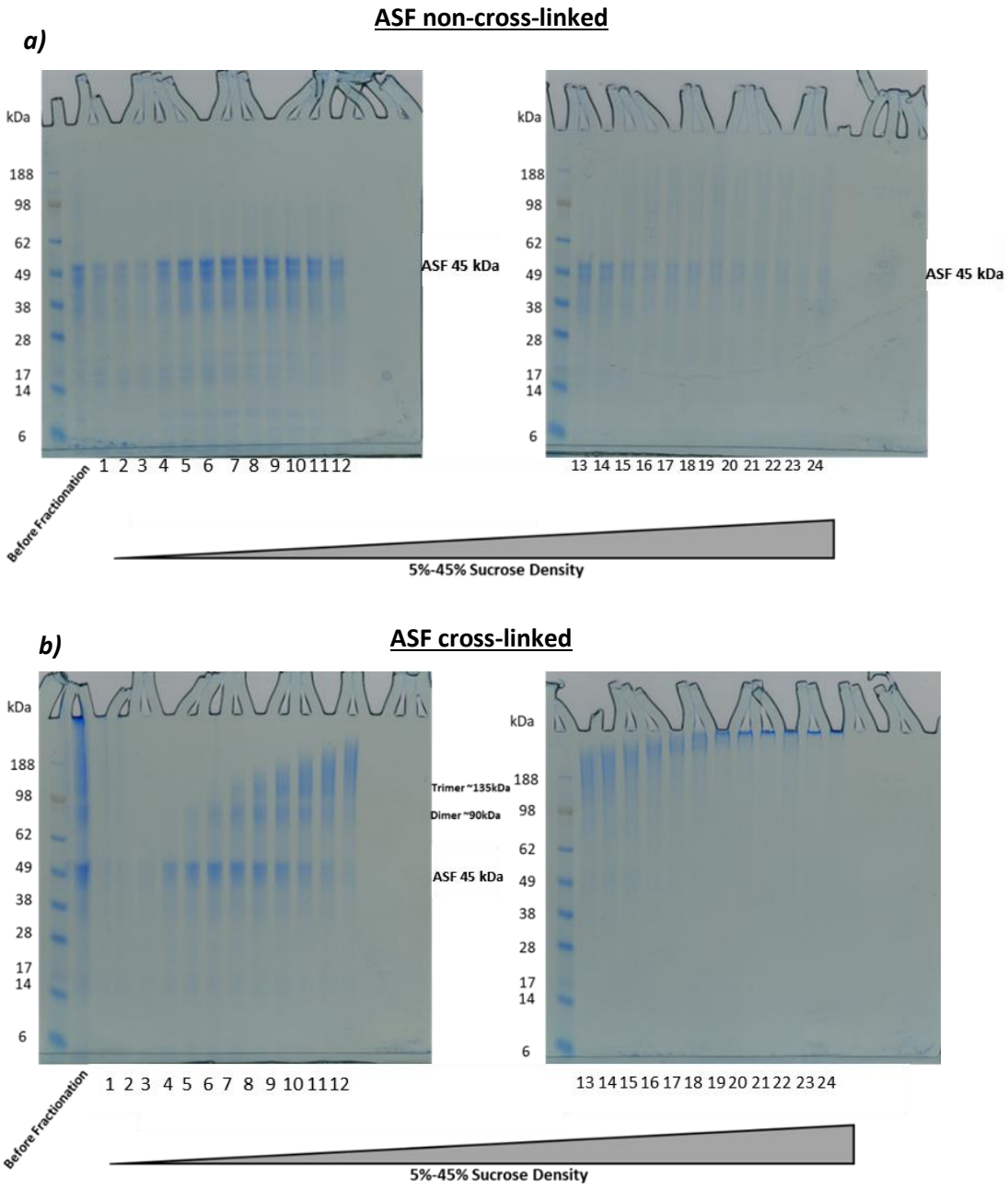


Figure 5.23, Sucrose density gradient ultracentrifugation profile of ASF ± cross-linking. Sucrose gradient analysis of non-cross-linked ASF (***5.23a***) and cross-linked ASF (***5.23b***). Sequential 200 µl fractions were collected across the 5%-45% (w/v) sucrose density gradient. A total of 24 fractions were collected. 10 µl of each sample was loaded with 5µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.23a,*** Lane 1 pre-ultracentrifuged non-cross-linked ASF (control). ***5.23b,*** Lane 1 pre-ultracentrifuged cross-linked ASF (control). The remainder of the lanes in both of the non-cross-linked and cross-linked gels were loaded with fractions 1-24 (distributed between the two gels, and labelled on the X-axis according to the fraction loaded). The gels were stained overnight in Expedeon instant blue. The Y-axis denotes molecular mass values of marker (SeeBlue™ Plus2 Pre-stained Protein Standard). ASF monomeric and oligomeric molecular weights (kDa) have been marked. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.7.2 Sucrose Density Gradient Ultracentrifugation of Cross-linked Galectin-3:ASF

Cross-linked galectin-3:ASF oligomers were similarly characterised using 5%-45% (w/v) sucrose density gradient ultracentrifugation.

Non-cross-linked (control) co-incubated material was distributed in a profile indicating both galectin-3 and ASF were behaving as monomers in these conditions (Figure 5.24a).

Cross-linked co-incubated material fractionated with evidence of oligomeric behaviour (Figure 5.24b). Overall, the SDS-PAGE appearances for the cross-linked co-incubated material seem interpretable in terms of cross-linked ASF species (as in Figure 5.23b, with the similar 'staircase' pattern across the fractionated species), and monomeric galectin-3 (Figure 5.22a). Since the cross-linked galectin-3:ASF samples did not demonstrate the presence of unique galectin-3:ASF oligomers in the sucrose gradient fractions, these approaches were not directly taken forward for further downstream applications.

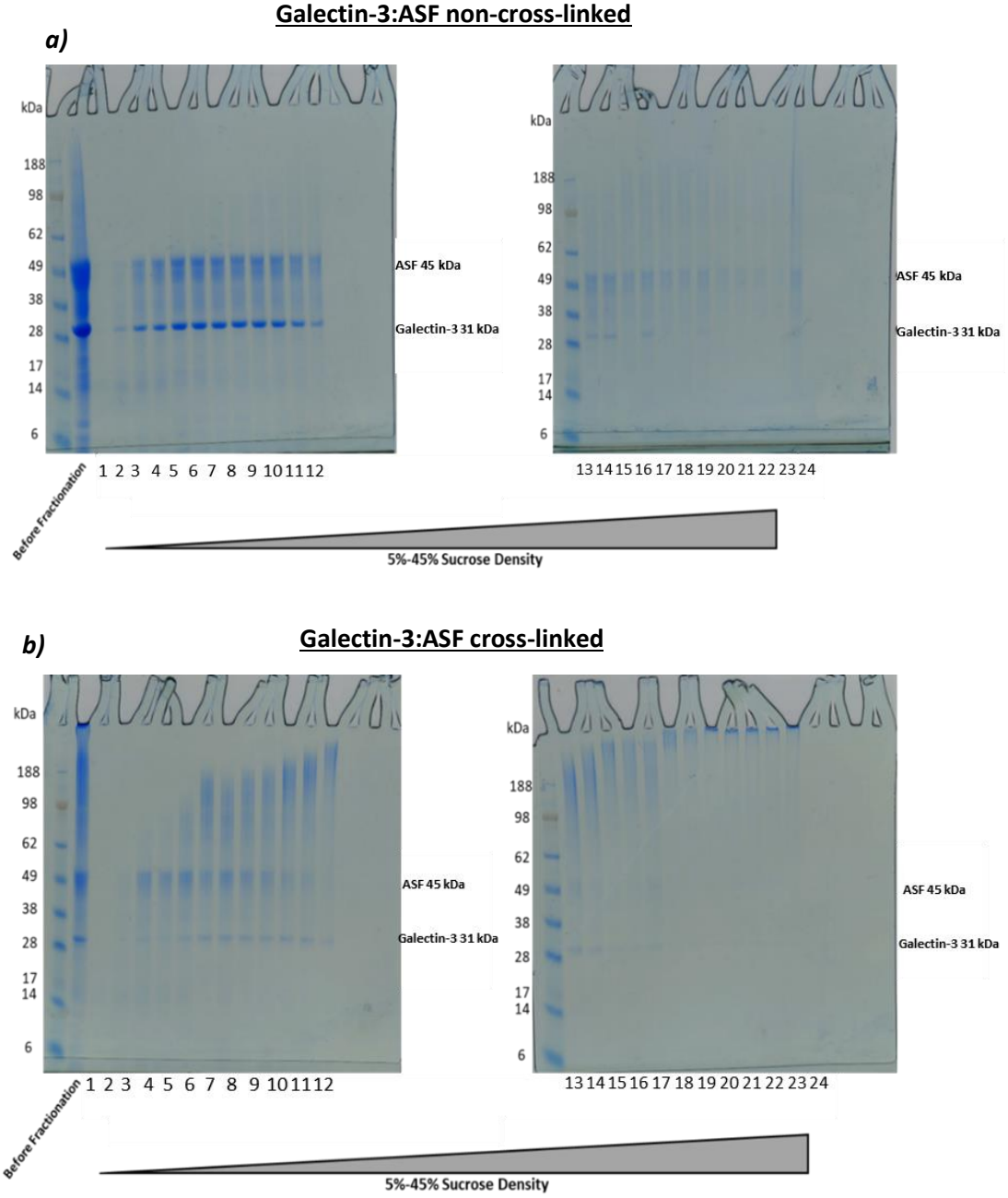


Figure 5.24, Sucrose density gradient ultracentrifugation profile of galectin-3:ASF ± cross-linking. Sucrose gradient analysis of non-cross-linked galectin-3:ASF (***5.24a)*** and cross-linked galectin-3:ASF (***5.24b***). 200 µl fractions were collected from top to bottom of the ultracentrifuged tube. A total of 24 fractions were collected. 10 µl of each sample was loaded with 5 µl of 2x Laemmli sample buffer onto a 4-15% (v/v) resolving gel. The lanes were loaded as follows: ***5.24a***, Lane 1 pre-ultracentrifuged non-cross-linked galectin-3:ASF. ***5.24b*** Lane 1 pre-ultracentrifuged cross-linked galectin-3:ASF. The remainder of the lanes in both of the non-cross-linked and cross-linked gels were loaded with fractions 1-24 (distributed between the two gels, and labelled on the X-axis according to the fraction loaded). The gels were stained overnight in Expedeon instant blue. The Y-axis denotes molecular mass value of marker (SeeBlue™ Plus2 Pre-stained Protein Standard). Galectin-3 and ASF molecular weight (kDa) has been marked. Before image capture, the gel was washed twice in $_dH_2O$, and then scanned using an appropriate gel compatible scanner.

5.2.8 Negative Stain Electron Microscopy Assessment of Galectin-3 Induced Oligomerisation with the Glycoprotein Ligand Asialofetuin

I assessed SEC fractions of incubation products of interest by negative stain-electron microscopy (EM). Negative stain-EM was used as an initial screening method to observe the homogeneity and quality of sample preparations, prior to cryo-EM analysis.

Samples obtained from SEC elution peaks were buffer-exchanged from EDTA-PBS into Tris-HCL pH 7.5 as the heavy metal stain used (uranyl acetate) reacts with phosphate. Images were captured on the JEOL TEM with Natalie Allcock (EM Facility Manager, University of Leicester).

Prior to undertaking EM studies of elution fractions, I assessed control grids treated with stain alone, and buffer with stain. These consistently displayed blank images, indicating that any species observed under the microscope for each condition assessed were not buffer or stain artefacts.

5.2.8.1 Negative Staining Electron Microscopy of Galectin-3 and ASF Controls in Absence and Presence of Cross-Linking

The best evidence of ligand-induced oligomerisation would come from identification of species not observed with only galectin-3 or ASF. I performed negative stain-EM on samples corresponding to the peak eluted fraction displayed in the SEC profile of both galectin-3 (Figure 5.25 a,c) and ASF (Figure 5.25 b,d) without and with the same glutaraldehyde cross-linking treatment. In the absence of cross-linkage no species could be visualised by this technique in the galectin-3 control (Fig. 5.25a). This is consistent with data from the earlier characterisation methods indicating a lack of oligomer species eluting from the column, since the 31 kDa monomer of galectin-3 is smaller than can typically be identified above background by negative stain-EM.

In contrast, the cross-linked galectin-3 sample demonstrated the presence of some large species (Figure 5.25c), consistent with successful cross-linking treatment of galectin-3. However, as they were not detected by signals within either the SEC profile or the subsequent SDS-PAGE, these species are likely to be aggregates. Furthermore, only a few species were detectable within the field of view limiting assessment of any common features in the absence of more extensive studies. Since this set of experiments was designed specifically to focus upon ligand-induced oligomerisation, this was not pursued further.

Negative stain-EM of both ASF non-cross-linked (Figure 5.25b) and cross-linked (Figure 5.25d) SEC peak fraction demonstrates the presence of species with similar appearance, apparently at higher concentration in the cross-linked sample. This overall agrees with similarities between the SEC profiles for these samples (Figure 5.20a and 5.20b). However overall, no consistent homogeneous features are observable in these fractions.

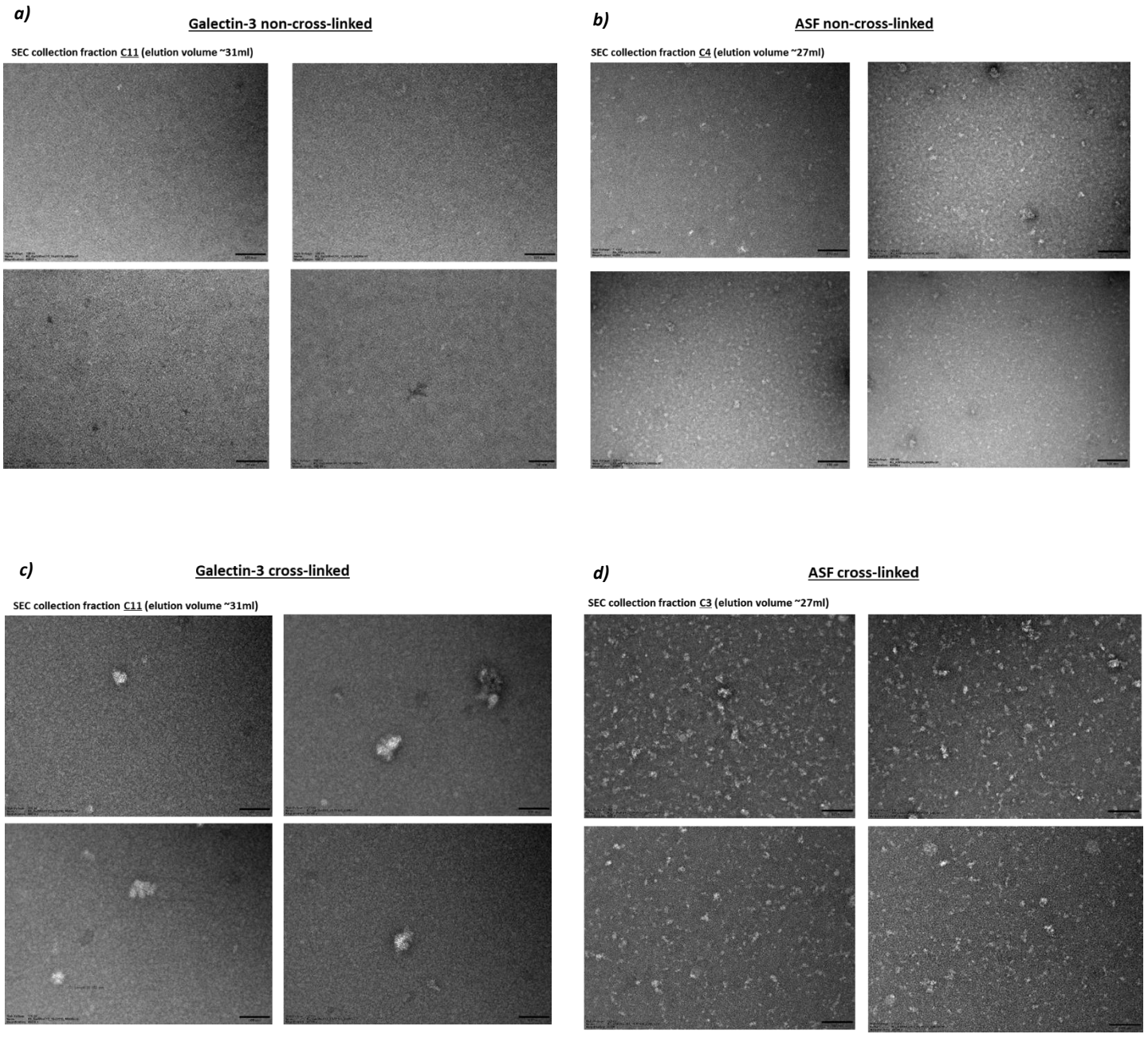


Figure 5.25, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of galectin-3 and ASF controls in the absence and presence of cross-linking. Negative stain-EM was conducted on the SEC peak fraction for all conditions. The micrographs displayed represent the following samples: ***5.25a,*** Negative stain-EM was applied to the peak eluted sample of non-cross-linked galectin-3 (fraction C11). ***5.25b,*** Negative stain-EM was applied to the peak eluted sample of non-cross-linked ASF (fraction C4). ***5.25c,*** Negative stain-EM was applied to the peak eluted sample of cross-linked galectin-3 (fraction C11). ***5.25d,*** Negative stain-EM was applied to the peak eluted sample of cross-linked ASF (fraction C3). Glow-discharged 300 mesh copper grids were used and the sample was dropped onto the grid. Grids were dually-stained with UranylAcetate. Projections were captured on JEOL 1400 Transmission Electron Microscope with EMSIS Xarosa 20MP CMOS digital camera. Micrographs shown are at 60K magnification. Scale bar of 100 nm is displayed at the bottom right hand corner.

5.2.8.2 Negative Stain Electron Microscopy Characterisation of Cross-linked Species Formed due to Galectin-3:ASF Interactions

Samples from the SEC peak for the non-cross-linked galectin-3:ASF sample and the two peaks (Peaks 1 and 2) from cross-linked galectin-3:ASF were assessed by negative stain-EM.

Negative stain-EM analysis of non-cross-linked (control) galectin-3:ASF peak elution fraction displayed similar motifs to the non-cross-linked ASF SEC fraction (Figure 5.26a). This agrees with the similarities in their A_{280} SEC trace profiles and so likely represents dissociated monomeric ASF.

Negative stain-EM analysis of cross-linked galectin-3:ASF of both SEC Peak 1 and Peak 2 samples displayed different features (Figure 5.26b), consistent with the MALS data (indicating species of sizes >66 MDa for Peak 1, and 628 kDa for Peak 2). The Peak 1 sample contained far larger species in comparison to the Peak 2 sample. The species in the Peak 2 sample were well distributed and observed at higher concentrations than Peak 1 species across the EM grid. These findings supported the potential of using cross-linked samples of ligand-induced oligomers for further EM and cryo-EM studies. However, the shapes and sizes of the species within the field of view indicated substantial heterogeneity that would likely confound structural analyses of a larger dataset collection. Since there are also many stoichiometries of galectin-3 and ASF that could be represented in these species, this did not seem a tractable option in the time available.

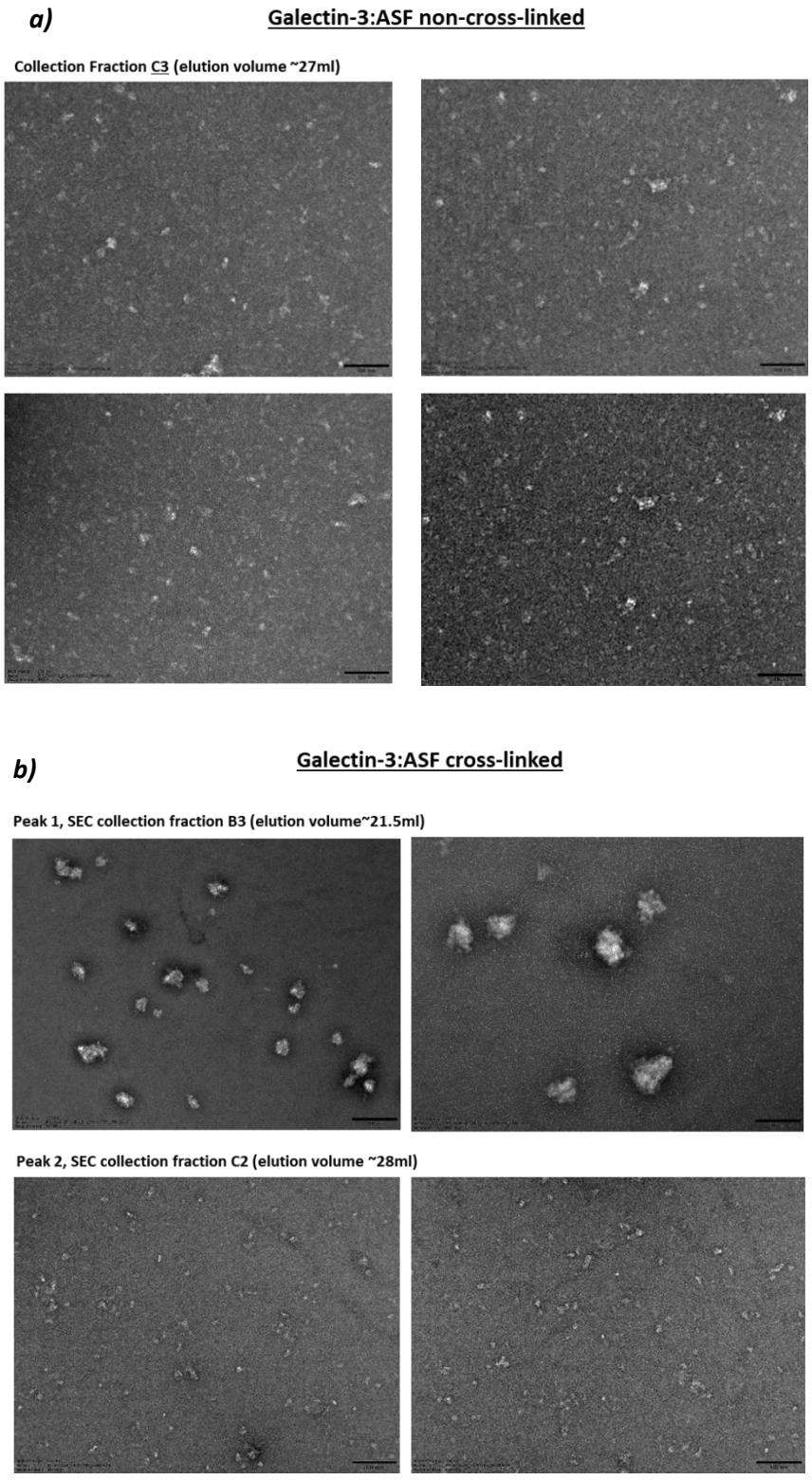


Figure 5.26, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of incubated galectin-3:ASF in absence and presence of cross-linking. Negative stain-EM was conducted on the SEC peak fraction of all conditions. The micrographs displayed represent the following samples: ***5.26a,*** Negative stain-EM was applied to the peak eluted sample of non-cross-linked galectin-3:ASF (fraction C3). ***5.26b,*** Negative stain-EM was applied to the peak eluted sample of cross-linked galectin-3:ASF Peak 1 (fraction B3) and Peak 2 (fraction C2). Glow-discharged 300 mesh copper grids were used and the sample was dropped onto the grid. Grids were dually-stained with UranylAcetate. Projections were captured on JEOL 1400 Transmission Electron Microscope with EMSIS Xarosa 20MP CMOS digital camera. Micrographs shown are at 60K magnification. Scale bar of 100 nm is displayed at the bottom right-hand corner.

5.2.8.3 Peptide N-Glycosidase F Treatment of Cross-Linked Galectin-3:ASF Species

I went on to try digesting the very large cross-linked galectin-3:ASF species seen in the SEC Peak 1 sample using Peptide N-Glycosidase (PNGase) F to see if tractable oligomeric subunits could be liberated. PNGase F is an enzyme used to remove oligosaccharides from glycoprotein by cleaving between the innermost N-acetylglucosamine (GlcNAc) and asparagine residues of complex oligosaccharides (NEB, 2021). As ASF will interact with galectin-3 via its N-linked glycans, the aim was to break up the very large complexes, although the covalent cross-linkage induced by glutaraldehyde treatment would be unaffected.

Following PNGase F treatment, the sample was subjected to size exclusion chromatography. Two peaks are visible in the A_{280} trace of the SEC profile (Figure 5.27). The first peak (Peak A) is highlighted at an early stage of the elution fraction volume, this displays a large A_{280} signal intensity (presence of bound complex) and was carried forward for further analysis. Peak B at later elution fraction volume displayed a low A_{280} signal intensity. This peak corresponds to species of lower molecular weight (confirming the cleavage of free sugar chains that were not cross-linked following cross-linking treatment).

Peak A fractions were pooled and dialysed into Tris-HCL buffer, pH 7.5, and then the sample was analysed by negative stain-EM (Figure 5.28). The species observed were far smaller than those seen for Peak 1 samples prior to PNGase F treatment, indicating the technical success of this strategy. The smaller size may have aided identification of certain morphologies observable in the 2D projections e.g. a 'boomerang' or 'croissant' morphology (highlighted in yellow circles), and a 'paisley' morphology (highlighted in purple circles). However, in consultation with the managers of the University of Leicester EM (Natalie Allcock) and Midlands cryo-EM (Dr Christos Savva) facilities, the heterogeneity was still considered too great to proceed with further EM or cryo-EM studies with this sample.

This approach has proved to be promising, since the multimeric species were further defined through PNGase treatment, however further optimisation is still required for downstream experiments of cryo-EM. Methods to define a better resolution of the multimeric species in order to attain good sample for single image processing are being considered within the Gooptu group, and are being carried on by a current PhD student within the lab and has been described in greater details in the discussion of this chapter.

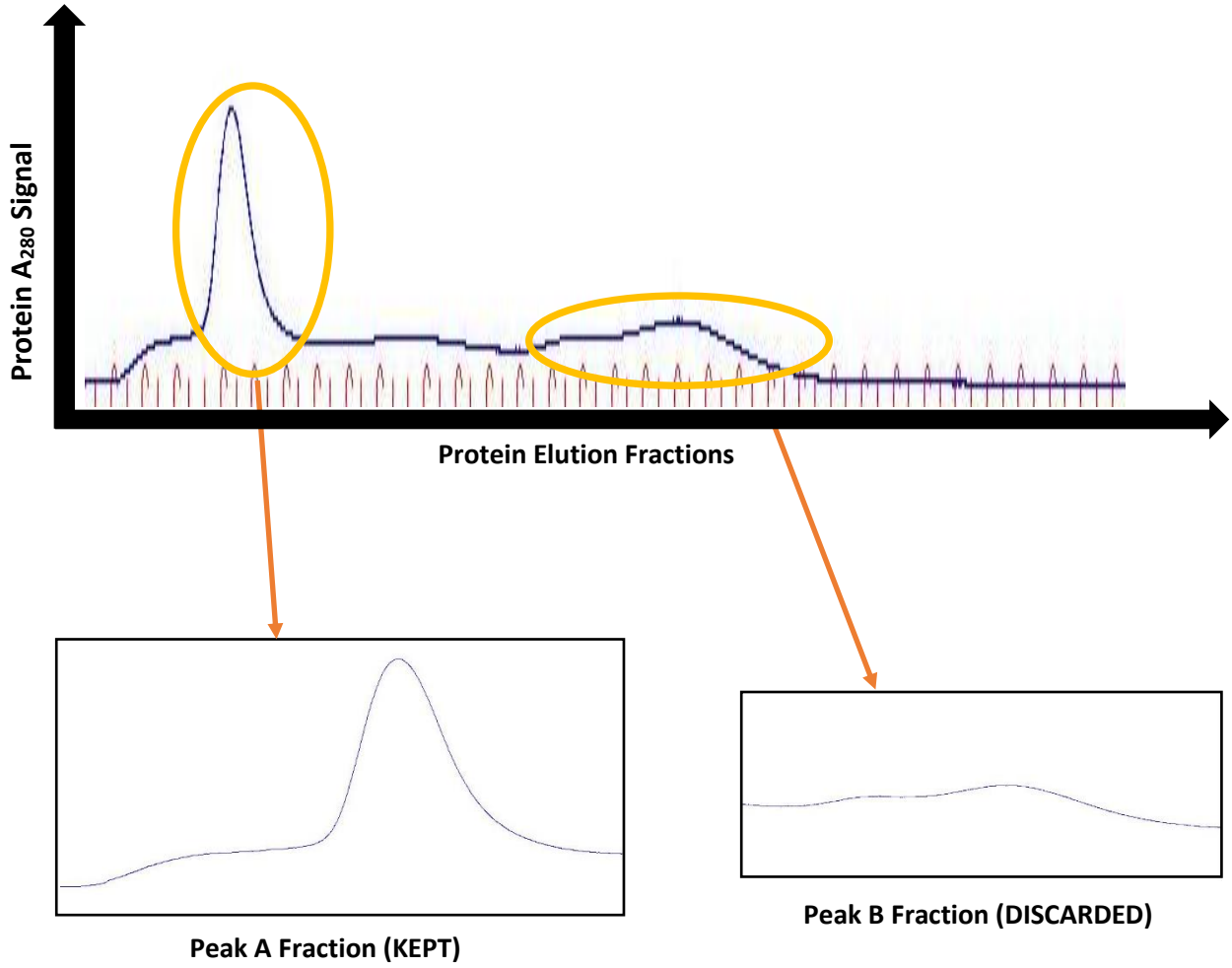


Figure 5.27, Size exclusion chromatography analysis of cross-linked galectin-3:ASF subjected to deglycosylation. Peak 1 SEC elution fractions of cross-linked 20 µM of galectin-3 with 100 µM of ASF was pooled and collated. This was subjected to non-denaturing PNGase F treatment and loaded on the S400 Sephacryl column (elution range from 20 kDa-8.00 MDa). Two A_{280} peak signal is present. Peak A is at ~22 ml (corresponding to covalently bound protein complexes). Peak B is at ~32 ml (corresponding to cleaved sugar chains). The X-axis represents the protein elution fraction(s), The Y-Axis represents the A_{280} protein signal intensity (arbitrary, milli-absorbance unit (mAU)).

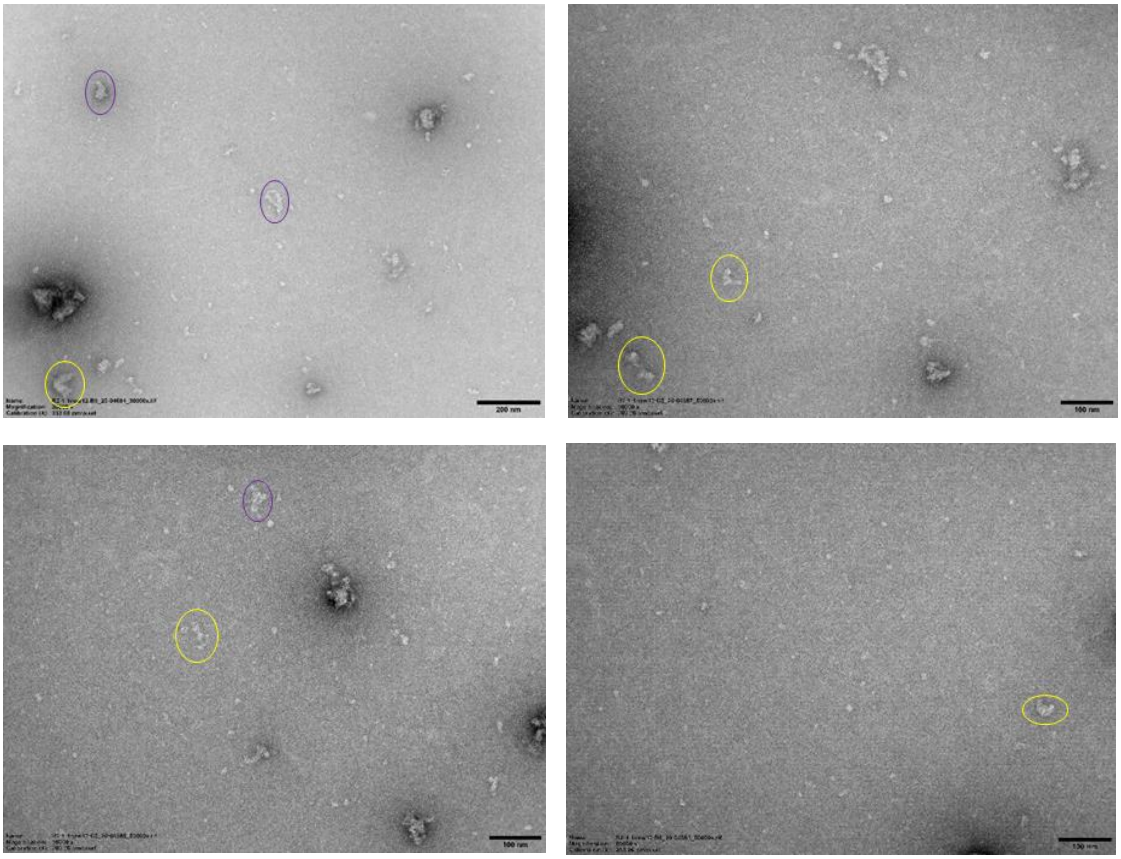


Figure 5.28, Negative Stain-EM analysis on size exclusion chromatography eluted peak fraction of PNGase F Treated cross-linked galectin-3:ASF. Eluted fractions from Peak A of PNGase F treated galectin-3:ASF crosslinked sample was pooled. Glow-discharged 300 mesh copper grids were used and the sample was dropped onto the grid. Grids were dually-stained with UranylAcetate. Projections were captured on JEOL 1400 Transmission Electron Microscope with EMSIS Xarosa 20MP CMOS digital camera. The yellow rings highlight a 'croissant' or 'boomerang' motif. The purple rings highlight a 'paisley-design' motif. Micrographs are shown at 60K magnification. Scale bare of 100 nm is represented at the bottom right-hand corner.

5.2.9 Isolation of Membrane Distributed Proteins Within Their Native Lipid Bilayer for Structural Analysis

5.2.9.1 Generation of SMALP A549 Cell Lysates Under Stimulated Conditions

Galectin-3 is widely distributed between cellular compartments. Therefore, to probe for proteins associated with the galectin-3-fibrosome on the cell surface, I chose to focus on the ability of a SMALP approach to extract species containing CD98.

Based on my earlier studies (Chapter 3) I chose to study this in A549 cells, stimulated with LPS (2 ng/ml, 24 h stimulation) or the pro-fibrotic cytokine TGF-β1 (10 ng/ml, 72 h stimulation) to maximise putative gal-3-fibrosome protein expression at the cell surface. I produced cell surface extracted proteins in their native state by preparing lysates of cells with styrene maleic acid (SMA). SMA lysates were analysed by SDS-PAGE and western blotted. Two different SMA's were used and optimised. One was kindly donated by Dafforn Group, University of Birmingham, using their licensed SMA-2000P. This was used in initial studies with their assistance. The other was a commercially available product, SMA 30010p (Orbiscope, Netherlands).

Lysates from both stimulated conditions demonstrated good signals for CD98hc (~85kDa) in the soluble fraction, supporting successful extraction of this cell membrane protein (Figure 5.29a). I next assessed whether CD98hc was extracted in complex with other proteins by running samples from the same lysate soluble fractions on a non-denaturing (SMA native) PAGE. Whilst SMA does not denature protein as SDS does, they are similar in conferring negative charge on the proteins they interact with, so SMA embedded proteins migrate relatively cleanly on non-denaturing PAGE. Strong CD98hc-positive bands were observed in samples extracted from both LPS-stimulated and TGF-β1-stimulated conditions (Figure 5.29b). Molecular markers indicated the molecular weight of the band was close to 240 kDa. In non-reducing conditions we would expect CD98 to run as a covalently bound dimer with light chain (~40 kDa) partner molecules, giving a

total molecular weight of ~125 kDa. The SMA native-PAGE findings therefore indicated that the extracted CD98hc was in complex with one or more other proteins. Similar findings were observed when the commercially available SMA (SMA-30010p) was used (Figure 5.30a and b), with the non-denaturing analysis indicating the complex had a molecular mass of greater than 220 kDa. Interestingly, treatment with the reducing agent dithiothreitol (DTT) did not alter the migration pattern. This indicates that either the SMA and protein complex stabilise the CD98 light chain, or that it may not be present within the prepared samples any differently to those without, therefore confirming DTT is not required for sample migration. Subsequent experiments were conducted using the commercially available SMA-30010p.

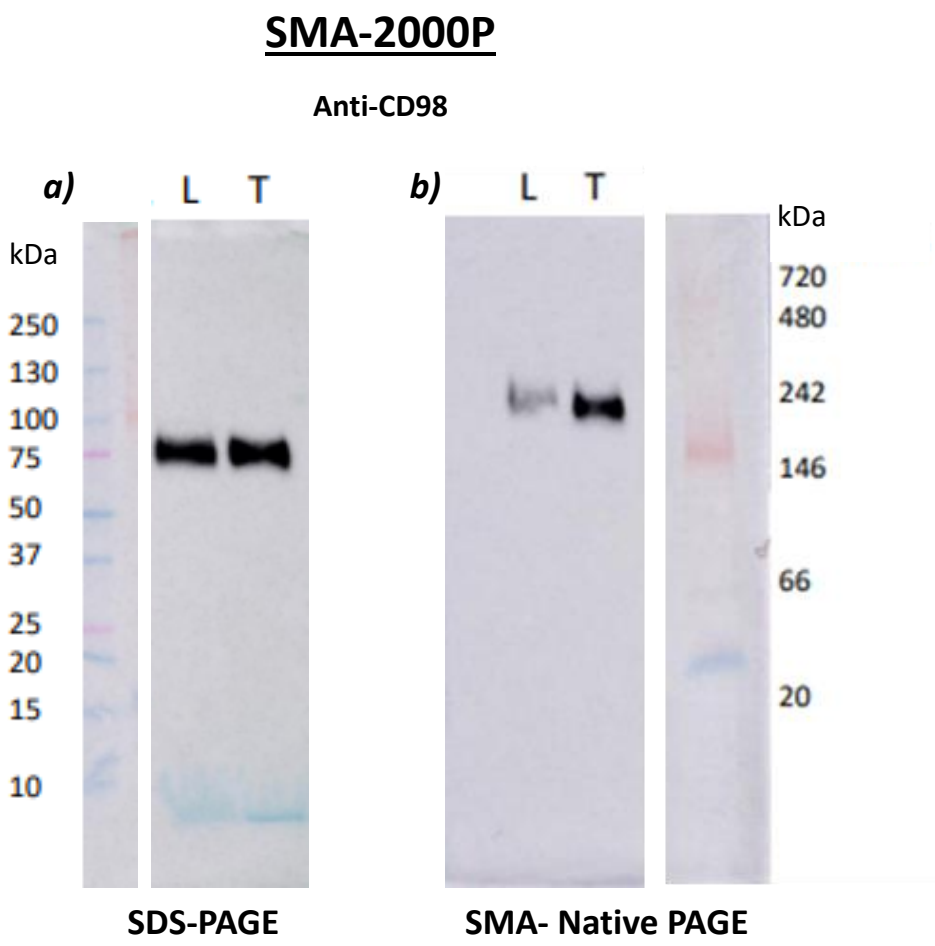


Figure 5.29, Assessing the ability of SMA-2000P to form SMALP lysates in TGF-β1 and LPS stimulated A549 cells with CD98hc encapsulated within the nanodisc. SMALP Cell lysates (cell surface fraction in detergent free medium) was prepared using our collaborators licensed SMA-2000P. For the SDS-PAGE gel ***(5.29a)*** 20 µl of lysate were loaded with 4x Laemmli sample buffer on a 4-15% (v/v) resolving gel (SDS-running buffer). For the SMA-Native-PAGE gel ***(5.29b)*** 20 µl of lysate were loaded with 2 x Native sample buffer on a 4-15% (v/v) resolving gel (SDS free running buffer). The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted 0.125 µg/ml CD98hc. The loaded lanes have been labelled with the appropriate conditions (on top of the gel): T= 72 h TGF-β1 stimulation, L= 24 h LPS stimulation. The Y-axis has been labelled with the respective marker: for SDS gel, Precision Plus Protein™ Dual colour standards and for SMA-Native gel, NativeMark™ protein standards was used. The western blot membrane was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 2-5 minutes. (n=3)

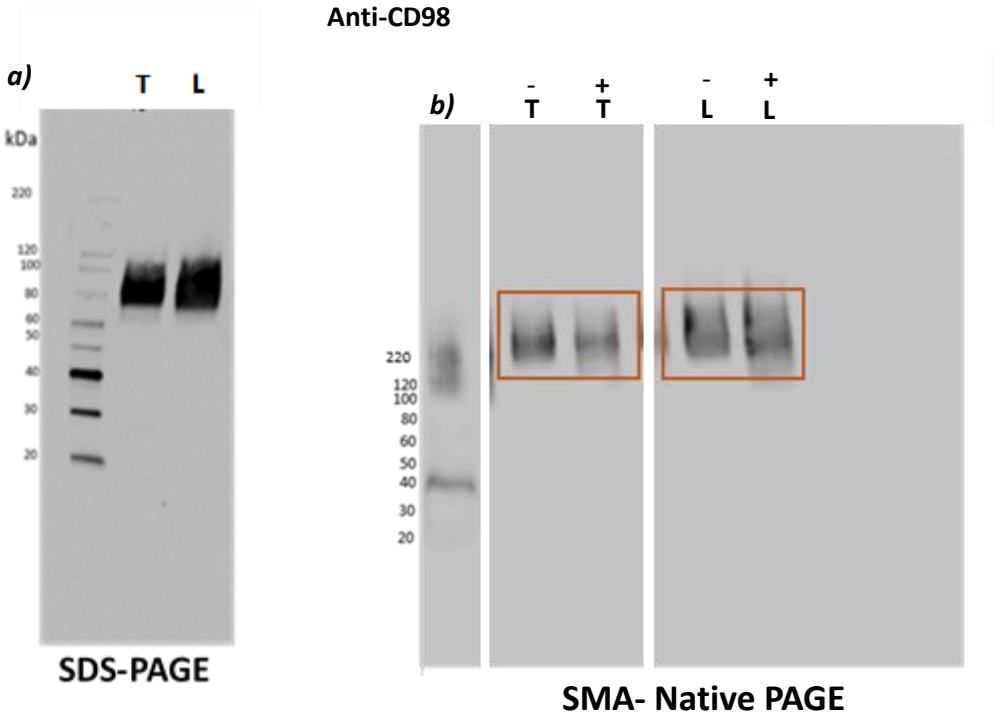


Figure 5.30, Assessing the ability of SMA-30010p to form SMALP lysates in TGF-β1 and LPS stimulated A549 cells with CD98hc encapsulated within the nanodisc. SMALP Cell lysates (cell surface fraction in detergent free medium) was prepared using commercially available SMA-30010P. For the SDS-PAGE gel ***(5.30a)*** 20 µl of lysate were loaded with 4x Laemmli sample buffer on a 4-15% (v/v) resolving gel (SDS-running buffer). For the SMA-Native-PAGE gel ***(5.30b)*** 20 µl of lysate were loaded with 2 x Native sample buffer on a 4-15% (v/v) resolving gel (SDS free running buffer). The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted 0.125 µg/ml CD98hc. The loaded lanes have been labelled with the appropriate conditions (on top of the gel): T= 72 h TGF-β1 stimulation, L= 24 h LPS stimulation, '- or +' = with or without DTT. The Y-axis has been labelled with the respective marker: for SDS gel, Precision Plus Protein™ Dual colour standards and for SMA-Native gel, NativeMark™ protein standards was used. The western blot membrane was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 2-5 minutes. (n=3)

5.2.9.2 CD98hc Enrichment from A549 SMALP Lysates

I immunoprecipitated SMALP lysates from A549 cells stimulated with TGF-β1 and LPS. This was to further clarify and specifically pulldown proteins of interest rather than all cell surface proteins encapsulated by the nano-discs. CD98hc and proteins encapsulated in the nano-discs were immunoprecipitated using a CD98hc monoclonal antibody. After overnight incubation of the antigen-antibody-bead complex, the CD98hc bead-bound sample was eluted off the beads using a salt elution approach (3 M $MgCl_2$ elution).

The immunoprecipitated sample was assessed initially by SDS-PAGE probed by western blot (Figure 5.31a). Immunoprecipitants from both TGF-β1- and LPS-stimulated SMALP lysates displayed bands at ~85 kDa (CD98hc molecular weight), supporting the presence of CD98hc as expected. SMA native-PAGE (Figure 5.31b) again supported purification of complexes containing CD98hc. These appeared to have a molecular mass between 238 kDa and 268 kDa, consistent with those previously observed. Both the SDS and SMA native-PAGE blots were next washed and re-assessed with a monoclonal galectin-3 antibody to examine if this co-immunoprecipitated with the anti-CD98hc immunoprecipitation. Whilst in the SMA native-PAGE western-blot, galectin-3 bands were not detectable in both TGF-β1 and LPS stimulated lysates, the SDS gel displayed a faint band of galectin-3 at its predicted molecular weight in the TGF-β1 stimulated immunoprecipitated sample (Figure 5.31c).

The immunoprecipitated SMALP extracts of the CD98hc complex were next characterised by negative stain-EM (Figure 5.32). Particles of ~22 nm diameter were observed, with no gross differences detectable between samples from cells stimulated with either TGF-β1 or LPS. In addition to blank findings in buffer and stain controls, the particles became less crowded on the grid in step with sample dilutions.

Based on these encouraging findings, fresh samples of TGF-β1 and LPS stimulated CD98hc immunoprecipitated SMALP lysates were prepared and loaded onto a cryo-EM grid prior to flash cooling and cryo-EM imaging (working with Dr Christos Savva, Midlands Cryo-EM Facility Manager). The samples were assessed on a Thermo Fisher Titan Krios G3 equipped with a Gatan K3 direct electron detector. Unfortunately, the particles could not be convincingly visualised. Regrettably, due to the global COVID-19 pandemic I was unable to pursue further time on this line of work. Further work has consequently been taken on within the Gooptu group (Charlie Hitchman).

SMA-30010P

Anti-CD98 (co)immunoprecipitation

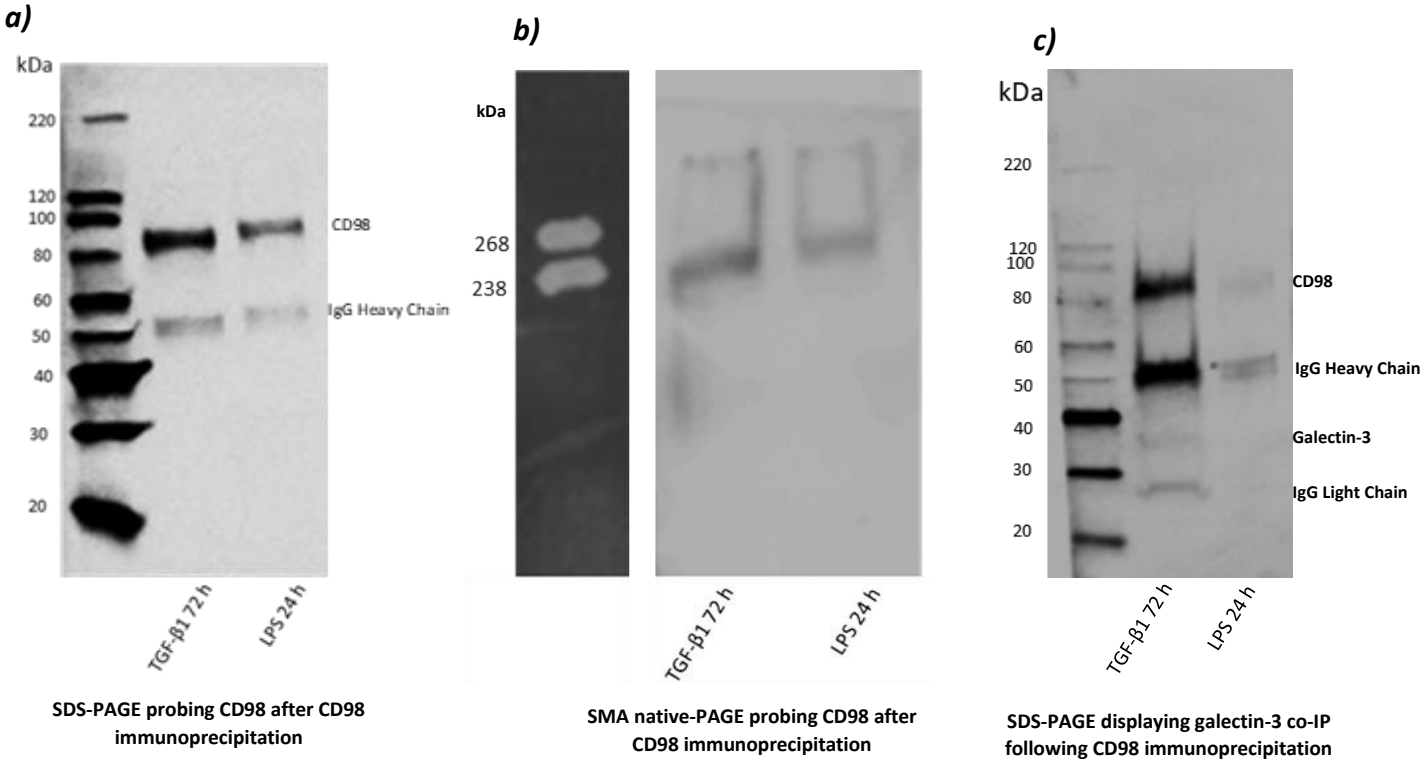
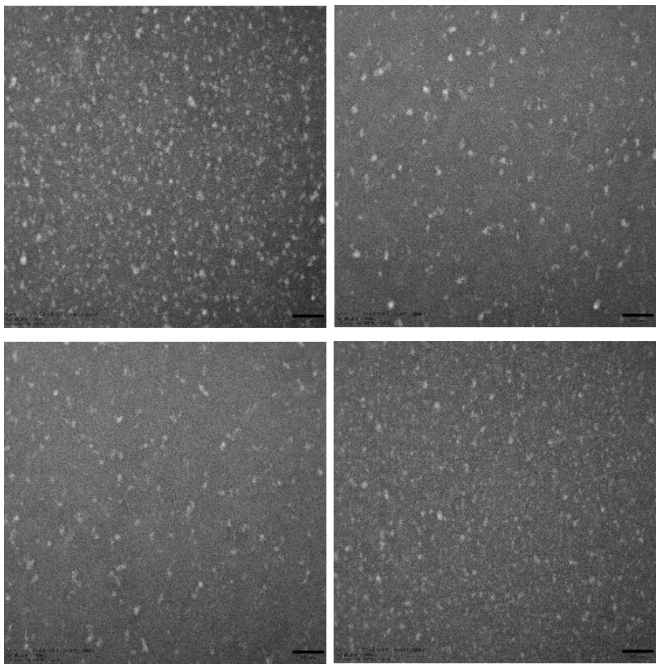


Figure 5.31, Enriching CD98hc bound proteins by applying anti-CD98hc pulldown on TGF-β1 and LPS stimulated SMALPs and eluting these species in a native condition. The lanes were loaded for both gels with samples as labelled. The Y-axis indicates masses of molecular weight markers ***5.31a, SDS-PAGE,*** between ~80-100 kDa molecular weight (CD98hc), the bands represent immunoprecipitation. Bands visible at ~50 kDa represents the Immunoglobulin heavy chain. ***5.31b, SMA-Native PAGE***, single bands are observed that migrate between ~238-268 kDa molecular weight markers. ***5.31c, SDS-PAGE,*** between ~30-40 kDa represents molecular weight of galectin-3 and therefore its successful co-immunoprecipitation following an anti-CD98 pulldown. Bands visible at ~80-100kDa represent CD98hc as it was initially probed on this blot, ~50 kDa and ~25 kDa represents the Immunoglobulin heavy chain and light chain, respectively. The markers on the blots of ***5.31a and c***, is MagicMarkXP and ***5.31b***, is HiMArk prestain protein standard (digitization image was taken to capture the band weight). All blots were immersed with EZ-ECL chemiluminescent detection kit and the image was taken at an exposure time of between 2-10 minutes (n=3).

TGF-β1 stimulated SMALP lysate, CD98hc immunoprecipitation



LPS stimulated SMALP lysate, CD98hc immunoprecipitation

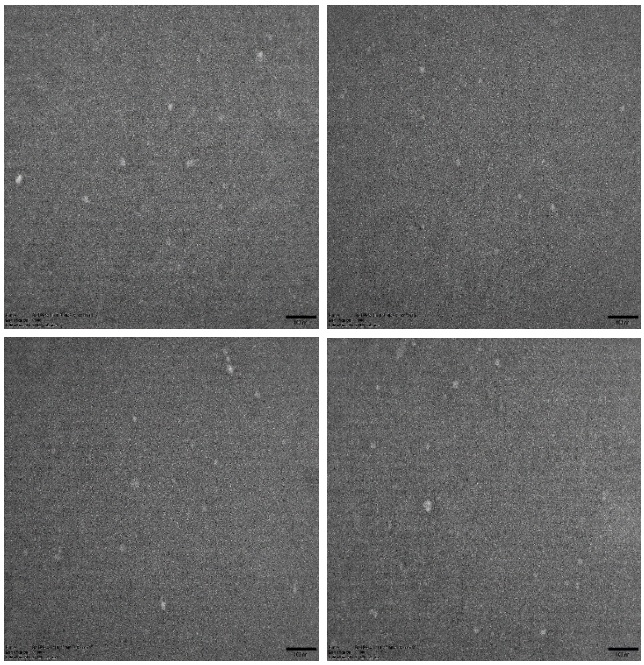


Figure 5.32, Negative stain-EM analysis on stimulated SMALP lysates subjected to CD98hc immunoprecipitation and native condition sample elution. Negative stain EM was done on anti-CD98hc pulldown ***(a)*** 72 h TGF-β1- and ***(b)*** 24 h LPS-treated SMALP lysates eluted with 3 M $MgCl_2$ (native elution). Glow-discharged 300 mesh copper grids were used and the sample was dropped onto the grid. Grids were dually-stained with UranylAcetate. Projections were captured on a JEOL 1400 Transmission Electron Microscope with an EMSIS Xarosa 20MP CMOS digital camera. Micrographs shown were obtained at 60K magnification. Scale bars represent 100 nm (lower right corner).

5.3 Discussion

My research data presented in chapters 3 and 4 broadly support the existence of an important functional interaction between galectin-3 and the putative fibrosome proteins. They do not however, prove direct physical interactions are important in the cell models of disease studied nor do they define any high-resolution information on any such interactions thereby trying to disrupt them. To define galectin-3 oligomeric behaviour and therefore the potential of how these proteins could interact physically, I addressed this using recombinant galectin-3.

Recombinant galectin-3 was successfully cultured and purified in my hands using the Lactosyl-sehparose affinity column I made. The protein yield was high (~50 ml of ~0.5-1 mg/ml, pre-concentrating). My initial interest was to prepare stable oligomeric samples of galectin-3 for structural characterisation using cryo-EM. A major challenge I faced was to find an assay capable of reporting oligomerisation events in cell-free conditions as even 'non-perturbing' approaches require significant changes to the conditions in which oligomerisation is induced, e.g. dilution, potentially reversing it.

Published data indicate that, in the absence of an appropriate carbohydrate ligand, galectin-3-self-oligomerisation is observed at high concentration (≥ 100μM) (Ahmad *et al*. 2004). At the outset an important question was whether in fact galectin-3 purified as a homogeneous monomer or in observable equilibrium with oligomeric species. In addition to dilution, individual biochemical methods may affect oligomeric stability by changes in buffer conditions, molecular crowding, and biophysical effects of the assay itself, e.g. gel electrophoresis. It was therefore important to compare findings from a number of different assays. All of these indicated the sample behaved in a conformationally homogeneous way, consistent with an entirely monomeric product despite purifying at >100 μM concentration. Assuming that the published data reporting oligomerisation in these concentrations (Ahmad *et al*. 2004, Johannes *et al*. 2018) are valid, the

most likely explanation is that at lower concentrations the oligomers dissociate back to the monomeric state. Unfortunately, I did not have access to equipment such as dynamic light scattering to confirm and reproduce the observations of oligomeric behaviour in the same assays using my recombinant galectin-3 product. It also did not seem feasible or sensible to increase the concentrations of the initial sample by the orders of magnitude necessary to keep galectin-3 concentration >100 µM when diluted in the assays. This would be likely to induce non-specific protein aggregation so complicate interpretation of any apparently positive results.

As part of these studies, I specifically tried to find conditions to optimise characterisation of galectin-3 by a native-PAGE approach. The standard native-PAGE Tris-running buffer is commonly used at a pH of 8.2, very close to the pI of galectin-3. Under these conditions galectin-3 migrates as a number of diffuse bands. This may indicate multiple residues populating a range of charge states due to the similarity of pH and pI. In fact, this would predict a majority of galectin-3 migrating out of the gel due to a mildly cationic state overall. However, strategies to overcome these issues, and concern about the running temperature of the system were only partially successful in optimising galectin-3 migration and hence interpretability on native-PAGE. All assessed altered native-PAGE running conditions whilst improved galectin-3 migration from the stacking gel portion into the resolving gel portion, overall the samples did not migrate much further into the resolving gel.

Ligand interactions can induce galectin-3 oligomerisation in low concentrations (Ahmad *et al*. 2004, Dam *et al*. 2005). I therefore decided to use such ligands to support my chances in preparing homogeneous oligomeric samples for cryo-EM. I initially studied the disaccharide N-acetyl-lactosamine (LacNAc) (Bumba *et al*. 2018). Non-cross-linked samples behaved much like galectin-3 monomers, with no evidence of residual oligomers upon dilution in the assays used. However, I also used glutaraldehyde to covalently cross-link amines on nearby amino-acids via the Mannich reaction and/or reductive amination (Dammink *et al*. 1995), providing the chance to observe evidence of any oligomerisation states even in more dilute assay conditions.

Cross-linking of galectin-3:LacNAc precipitated out the sample in an inconsistent, stochastic manner. This prevented identification of optimal conditions to scale up oligomer preparation, and so this line of investigation did not proceed further. However, the finding does itself strongly support the formation of LacNAc-induced oligomers of galectin-3 as precipitation is not seen with cross-linkage of galectin-3 at the same concentration.

Since the glycan ligand LacNAc was not inducing galectin-3 oligomeric samples suitable for my purposes, I next decided to use the glycoprotein ASF (~45kDa). This is known to bind with high affinity to galectin-3 (Lepur *et al*. 2012, Bocker and Elling. 2017) and to induce galectin-3 oligomerisation (Dam *et al*. 2005).

The non-cross-linked experimental reaction mix of galectin-3 and ASF consistently behaved as a simple combination of the two monomeric states when assessed by all techniques used. Interestingly, cross-linking galectin-3:ASF avoided the precipitation issue as seen with galectin-3:LacNAc cross-linkage. However, ASF introduces complexity of its own since it is a multivalent glycoprotein, inevitably cross-linking will create increased likelihood of further interactions.

ASF has nine branched-glycan chains (LacNAc). As each LacNAc interaction may induce galectin-3 to oligomerise, there is potential for large molecular structures to form. A model in which one ASF (45 kDa) binds nine individual galectin-3 molecules (31 kDa each) which can then each recruit further galectin-3 molecules to pentameries, generating a ~1440 kDa (1.44 MDa) species, but further galectin-3:ASF interactions could escalate the size of the assembly.

In practice I did observe species in the megadalton scale (Peak 1, >66 MDa) when galectin-3 interacted with ASF species as well as a ~600 kDa species (Peak 2) that was attributable to cross-linked ASF. Negative stain-EM studies indicated the Peak 1 species were very electron dense and

heterogeneous. These characteristics are adverse for structural studies by cryo-EM, but they were greatly improved by PNGase treatment. However, in the truncated time period available for this work due to the COVID pandemic, further optimisation to try and generate a sample where such work was practical was not feasible.

It is interesting that PNGase treatment appears to dissociate the large species into smaller assembly states. This indicates that LacNAc:galectin-3 interactions may substantially stabilise Peak 1 species even when diluted in the assays used. On the other hand in the absence of covalent cross-linking, stable ASF:galectin-3 species were not detectable in the same assays. This apparent contradiction might be explained if the presence of covalent cross-linkers orientates the ASF and galectin-3 molecules in a way that enhances LacNAc:galectin-3 interactions bridging between the cross-linked subunit assemblies.

Running these PNGase treated cross-linked samples through a sucrose density gradient using ultracentrifugation (with appropriate non-treated controls) might help characterise and minimise the heterogeneity issue. It would allow examination of whether oligomeric species fractionate uniquely within the gradient at a tractable molecular weight, and might generate homogenous sample for EM and cryo-EM analysis. Failing this, if the heterogenous sample of galectin-3:ASF cannot be further cleaned to make it homogenous, mass spectrometry could be used as a future experiment, helping to identify different stoichiometries of the galectin-3:ASF oligomeric species within the sample.

Working in collaboration with Dr Naomi Pollock from Professor Dafforn's Group, University of Birmingham, I was able to use a relatively new technique to isolate membrane proteins in their native environment by using SMA nanodiscs to form SMA-lipid particles (SMALPs). The advantage of this over the preparation of recombinant protein(s) is that it is physiologically more relevant (since for example proteins could be isolated from cells artificially stimulated or directly from

diseased cells thereby observing a certain disease characteristic). SMA nanodiscs are ~15 nm in diameter (Esmaili and Overduin. 2018, Pollock *et al*. 2018). SMALP strongly appears to successfully extract protein complexes containing CD98. SMA native-PAGE displayed a large molecular weight (~240 kDa) species. SMALP lysates were successfully enriched for CD98hc containing species using anti-CD98 mAb for immunoprecipitation. This pulled down species contains galectin-3 as well as CD98hc. CD98hc seems constitutively bound to β1-integrin on the cell surface as a heterodimer (Feral *et al*. 2005). The observed mass likely includes CD98hc (85 kDa) encapsulated within the nanodisc with the presence of CD98 light chain (molecular mass of ~40 kDa) and β1-integrin (molecular mass of 120kDa) giving a total theoretical mass of ~245 kDa. Negative stain-EM analysis on these CD98hc immunoprecipitated SMALP lysates further support sample homogeneity. However, further optimisation is required for cryo-EM sample preparation.

Like other elements of the work in this chapter, these lines of investigation remain promising. Currently this work is being continued on by Charlie Hitchman (PhD student, Gooptu group, University of Leicester). He is attempting to immunoprecipitate TGF-β1 and LPS stimulated SMALP lysates using the less cell surface distributed protein, galectin-3 (core protein bridging our galectin-3-fibrosome model). He will initially conduct identical downstream experiments as I have done, and if the results are successful, he will accordingly move onto small-scale cryo-EM screening of small molecule libraries.

Chapter 6

The Roles of Galectin-3 and CD98 in Inflammatory Cytokine Release by Host Cells in Response to Initial Interaction with the SARS-CoV-2 Spike Protein

6.1 Introduction

SARS-CoV-2 cell infection is initiated when the spike protein binds to ACE-2, followed by processing by TMPRSS2, allowing cellular endocytosis through both a dependent and independent clathrin- and caveolin pathway (Glebov. 2020, Hoffmann *et al*. 2020, Sungnak *et al*. 2020, Yang *et al*. 2020).

Despite the potential severity of COVID pneumonitis, ACE-2 is expressed only at low levels in alveolar epithelial cells, raising the possibility of interactions via co- or alternative receptors. Nevertheless, histopathologic studies and single-cell sequencing do confirm a sub-population of alveolar epithelial cells expressing ACE2 and TMPRSS2, and tissue EM shows viral replication within cells (Bourgonje *et al*. 2020, Yan *et al*. 2020). Interestingly the alveolar epithelial cell model-A549 cells, used in my studies to this point, does not express ACE-2 at detectable levels (The Human Protein Atlas. 2021), and is therefore protected from infection by SARS-CoV-2.

ACE-2 is abundantly expressed in activated type 1 (M1) macrophages under exacerbation stimuli (LPS) conditions (Song *et al*. 2020). Under these exacerbation conditions, M1 macrophages display a pro-inflammatory phenotype, observed by the resultant augmentation in multiple pro-inflammatory cytokines including IL-6, IL-8 and tumour necrosis factor (TNF)-α (Orrechioni *et al*. 2019).

Levels of the pro-inflammatory cytokines IL-6, IL-8 and TNF-α are elevated in serum from COVID-19 patients compared to healthy donors (Valle *et al*. 2020). These cytokine levels are comparable to those observed in other inflammatory acute respiratory distress syndrome (ARDS) states (Swaroopa *et al*. 2016, Lin *et al*. 2018). Therapeutic targeting of the IL-6 signalling pathway with the immunosuppressive drug tocilizumab reduces mortality in moderate to high severity COVID pneumonitis (REMAP-CAP study (REMAP-CAP investigators. 2021) and RECOVERY trial (The

RECOVERY Collaborative Group. 2021)). This indicates that in COVID pneumonitis the IL-6 response is more deleterious than helpful to the host, and likely contributes directly to the pneumonitis. The far weaker effects (Young *et al*. 2021) or negative findings (Jones *et al*. 2020) with antiviral treatments in these patients may suggest that active infection is less important than the inflammatory response in this phase of the disease (typically >2 weeks after initial infection).

Previous data in our group demonstrated that the transmembrane protein interactions of CD98:β1-integrin complex are critical mediators of IL-6 and IL-8 responses to diverse injury stimuli *in vivo* in mice*, ex vivo* in human lung tissue and *in vitro (*Wang *et al. Manuscript in preparation)*. This response pathway results in tissue neutrophilia *in vivo*. Ongoing studies are exploring the potential for such stimulation to enhance interactions of CD98 with galectin-3 and TGF-βRII in line with the gal-3-fibrosome hypothesis.

We hypothesised that the S1 region of the SARS-CoV-2 interacts with CD98hc:integrin complex as a co-receptor on cells and such interactions (with broader gal-3-fibrosome protein involvement) trigger hyper-inflammatory cytokine responses. This hypothesis was based on SARS-CoV-2 spike protein S1 region comprising of an N-terminal domain (NTD) which adopts a galectin fold with carbohydrate ligand binding capabilities and a receptor binding domain containing an RGD tripeptide sequence allowing β1-integrin binding (Caniglia *et al*. 2020).

To probe these responses *in vitro* and assess potential for therapeutic targeting, the Gooptu group were successfully awarded an MRC-funded confidence in concept grant via Leicester Drug Discovery and Diagnostics (LD3). I worked with Dr Oksana Gonchar (PDRA, Schwabe Group- LISCB, University of Leicester) to generate recombinant full-length (FL) spike protein followed by the S1 region and its construct domains specifically: the N-terminal and receptor binding domain. We wished to stimulate cells with these and conduct functional assays. I successfully produced

purified recombinant SARS-CoV-2 receptor binding domain proteins. I worked with Dr Panayiota Stylianou and Dr Omeed Darweesh (PDRAs, Gooptu Group, University of Leicester) to assess the functional output of receptor binding domain stimulations using two cell models- A549 and THP-1 cells. The THP-1 human monocytic cell line was used and differentiated with 72 h 200 nM phorbol-12-myristate-13-acetate (PMA) treatment into non-activated (M0) macrophages. Both A549 and M0 macrophage cells were stimulated with the purified receptor binding domain protein and functionally characterised (assessment of inflammatory cytokine levels) and their dependence of these on CD98hc:β1-integrin and galectin-3 was assessed using siRNA mediated knockdown and small molecule inhibitors.

6.2 Results

6.2.1 Generation of Spike Protein Constructs

I generated different domains of the spike protein constructs on multiple backgrounds defining optimal expression system by in-house design of primers, PCR amplification and T4 DNA ligase cloning strategy. I used a parent construct: wild-type SARS-CoV-2-S (FL-spike) strep-tagged, with native secretion sequence (donated, from Krogan group, USA). I cloned this (with the presence of the native secretion sequence at the N-terminal domain) into the pLeics49 plasmid and incorporated a FLAG-tag and a TEV cleavage site.

6.2.1.1 Generation of Spike Protein Constructs and Steps

From the parent Strep-tagged full length (FL) spike protein construct, I PCR amplified the receptor binding domain, N-terminal domain, full S1 region as well as the full length protein (using specific primers). In the latter I incorporated a step deleting the transmembrane domain at the end of the C-terminus. The molecular weight of each construct matched the expected value of, receptor binding domain ~0.6 kbp, N-terminal domain ~ 1 kbp, S1 ~2 kbp (Figure 6.1a), and FL-SARS-CoV-2 ~3.3 kbp. After phenol chloroform extraction and digestion with *KpnI* and *XhoI* restriction enzymes, I loaded the samples on an agarose gel and purified the bands of interest (Figure 6.1b). The inserts were cloned into pLeics49 (~5.5 kbp) vector. I then transformed the cloned sample, and inoculated colonies into 5 ml of LB Broth for small scale plasmid preparation (Mini Prep, 2.10.2). These plasmids were screened, multiple colonies showed products of consistent size for plasmids containing the receptor binding domain (~6.1 kbp) (Figure 6.1c), N-terminal domain (~6.5 kbp) (Figure 6.1d) and S1 domain (~ 7.5 kbp) (Figure 6.1e). These plasmid samples were initially checked by conducting a restriction digestion test. This is illustrated in Figure 6.1f for S1 plasmid preparations. In this case, of the thirteen digested samples only one sample showed

products at the predicted cut-points (the bands observed at ~5.5 kbp (pLeics49 vector) and ~2kbp (S1 insert). Samples that showed the expected profiles were sequenced by Sources Bio-Sciences (Nottingham, UK). The positive sequenced plasmids were amplified by large scale plasmid preparation (Maxi-prep) ahead of test expression experiments.

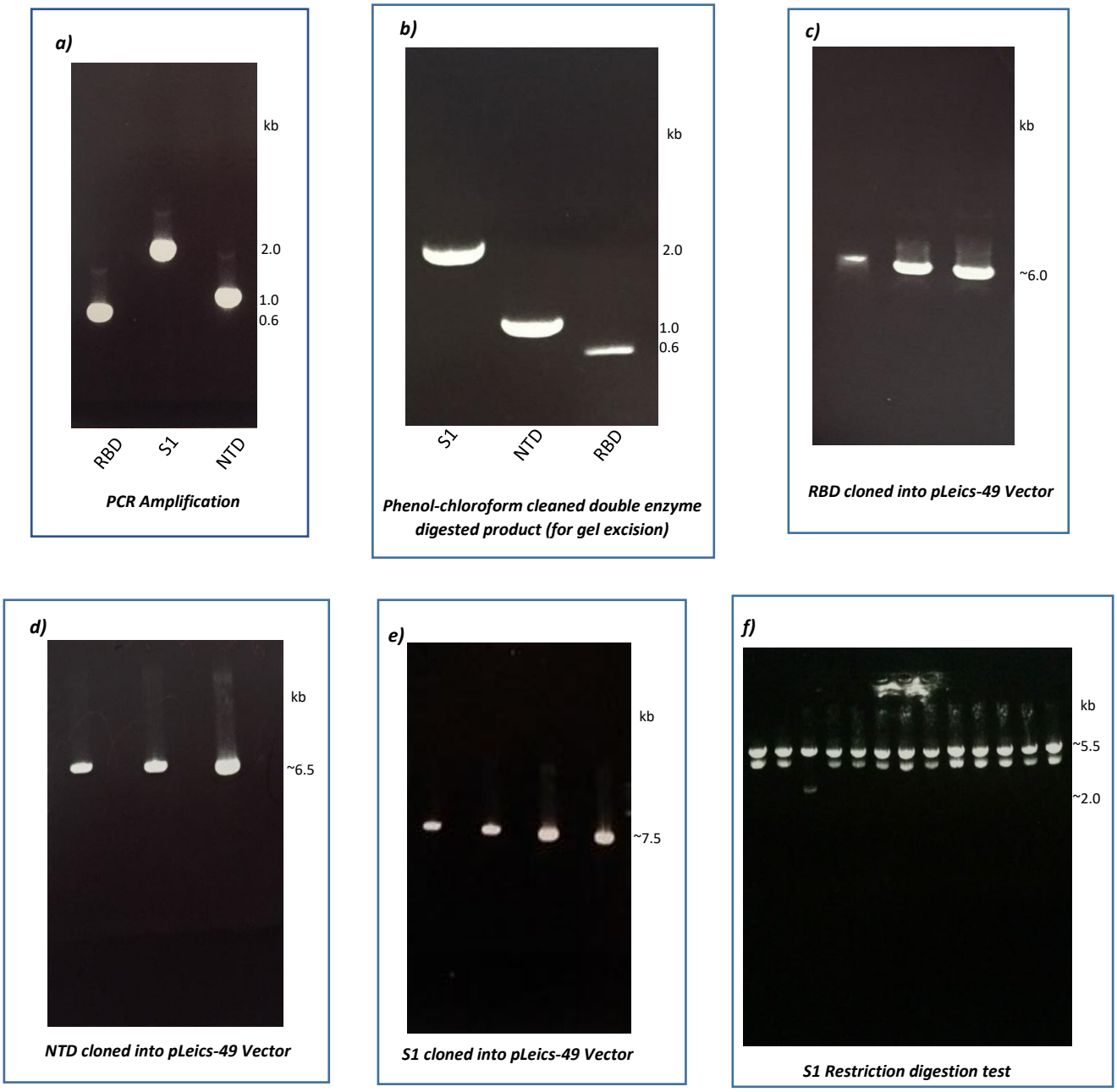


Figure 6.1, Multi-step approach in DNA cloning of S1, NTD and RBD domains into vector pLeics-49. a) PCR Amplification, The wild-type FL spike protein was used as the parent construct, and the inserts of S1, RBD and NTD was PCR amplified using the appropriate primers, visualised by the corresponding theoretical base pairs of each insert. ***b) Double endonuclease digested gel for gel excision/purification*****,** double digested phenol chloroformed PCR products, inserted into the gel for construct extraction using the QIAgen kit. Bands corresponding to each insert was sliced out. ***c-e) NEB competent transformed cells post-cloning,*** Post-transformation, colonies were picked and cultured small-scale followed by small scale plasmid prep. The positive bands at the correct bp size of RBD, NTD and S1 domain were re-transformed. ***f) Restriction digestion test,*** Re-transformed colonies were sense-checked before shipping out for sequencing. The double bands at the correct bp size represent correctly cloned inserts. Those samples were then shipped for sequencing. All samples were loaded on a 1x TBE 1% (w/v) agarose gel for ~45 minute at 80 V. 10 µl of sample was loaded with 2 µl of 6x loading dye. The marker 1kb DNA hyper ladder was used and loaded in lane 1 of each gel. Appropriate conditions have been labelled on the X-axis. The Y-axis has been labelled with the respective bp size. Image of the gel was captured using a UV trans-illuminator imager.

6.2.2 Protein Expression and Purification

For initial protein expression analysis, I transfected the constructs into suspension HEK293F cells (small scale). Protein expression was visualised by SDS-PAGE. If no protein bands of appropriate molecular weight could be visualised by Coomassie–based staining I undertook western blot analysis, probing with an anti-FLAG antibody.

6.2.2.1 Initial Protein Expression

HEK293F cells transiently transfected with the different constructs were cultured and assessed for expression of the recombinant FLAG-tagged proteins. Proteins express constitutively when the plasmid is transfected in this culture system. Anti-FLAG pulldown was performed on both the cell pellet lysate and media to capture these products. When examining the lanes corresponding to pre-TEV (input sample) and post-TEV cleavage of the cell pellet lysates, low levels of protein expression were observed for receptor binding domain (~32 kDa MW) and S1 (~92 kDa MW) constructs by Coomassie-stained SDS-PAGE gel (Figures 6.2a-b). No protein bands were visible in the elution sample lane, suggesting TEV cleavage was unsuccessful. Furthermore, the expression of the receptor binding and S1 constructs were observed only in the cell pellet but not the media. Coomassie-stained SDS-PAGE gel displayed no bands at the predicted molecular weight of the N-terminal domain (~50 kDa). To address this sensitivity issue, the FLAG-tagged N-terminal domain was then probed by western blotting. Expression of the N-terminal domain was observed in anti-FLAG pulldown samples from cell pellet lysate, prior to TEV cleavage (input sample) (Figure 6.2c). No bands were visible in the elution sample for the N-terminal domain, and it appears that the post-TEV treated beads demonstrated the presence of bands at ~50 kDa (Figure 6.2d), suggesting unsuccessful TEV-cleavage. Stronger staining was apparent for the receptor binding domain (MW ~32 kDa) when assessed by western blotting in the cell pellet lysate (supporting its clear visualisation by Coomassie-stained SDS-PAGE gel analysis) (Figure 6.2c), whilst the media demonstrated much weaker sample yield. The full length spike protein also expressed in the cell

pellet lysate and was not secreted into the media. This was barely detected by Coomassie-staining (Figure 6.2e) but western blotting analysis demonstrated presence of a FLAG-containing protein of appropriate molecular weight (>200 kDa) (Figure 6.2f). Interestingly, the Coomassie-stained SDS-PAGE gel of the full length SARS-CoV-2 construct displayed bands at ~98 kDa, suggesting that the protein is potentially proteolytically cleaved. This supports other reports demonstrating the full length SARS-CoV-2 contains a furin cleavage site between the S1 and S2 regions which is susceptible to proteolytic cleavage (Xia *et al*, 2020). All assessed construct findings are summarised in Table 6.1.

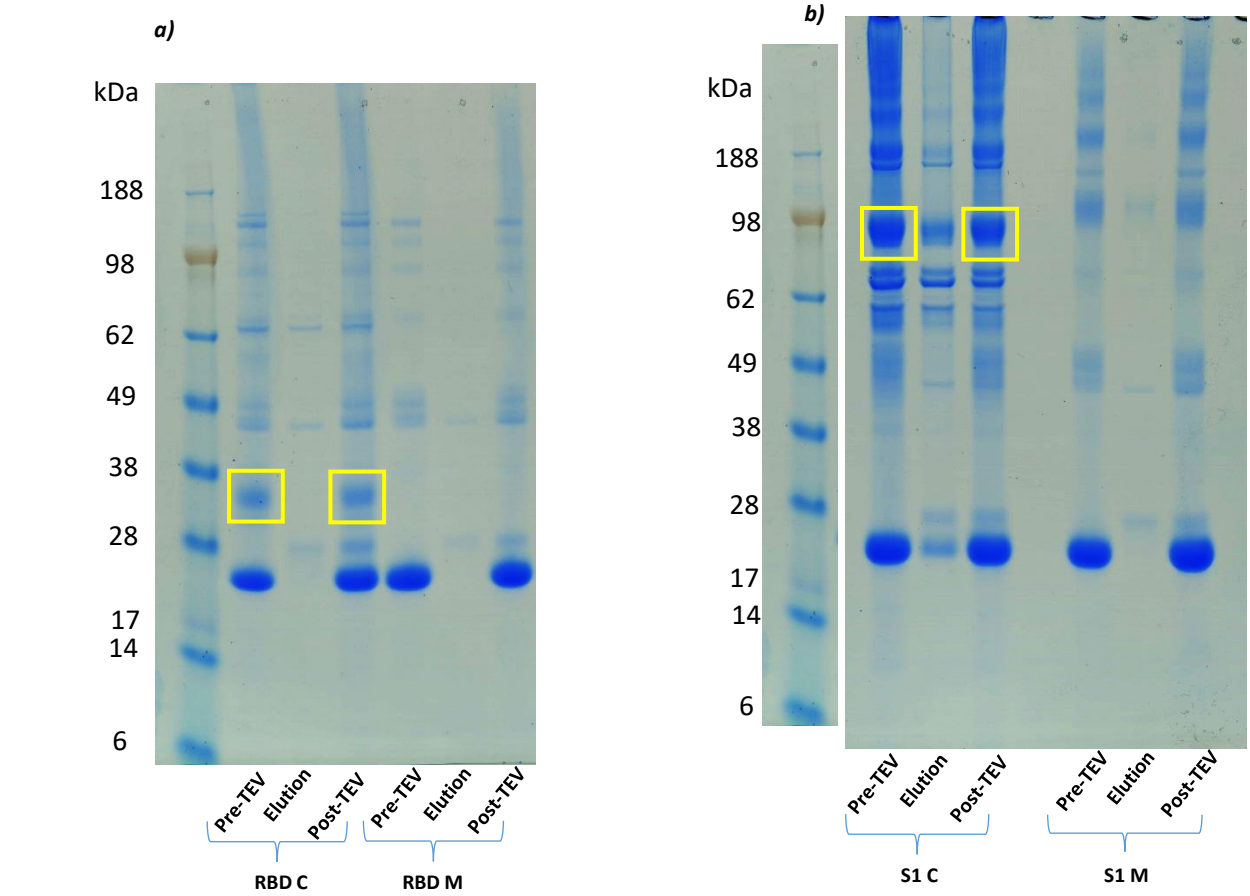

a)
kDa
188
98
62
49
38
28
17
14
6
Pre-TEV
Elution
Post-TEV
Pre-TEV
Elution
Post-TEV
RBD C
RBD M
b)
kDa
188
98
62
49
38
28
17
14
6
Pre-TEV
Elution
Post-TEV
Pre-TEV
Elution
Post-TEV
S1 C
S1 M
c)
NTD
RBD
kDa
C
M
C
M
220
120
100
80
60
50
40
30
20
d)
NTD
kDa
220
120
100
80
60
50
40
30
20
Elution
Post-TEV
e)
kDa
188
98
62
49
38
28
17
14
6
Pre-TEV
Elution
Post-TEV
FL-spike Cell pellet lysate
f)
kDa
220
120
100
80
60
40
30
20
FL

Figure 6.2, Anti-FLAG pulldown of FLAG-Tagged Spike protein constructs. 6.2a (RBD), b (S1) and e (FL spike protein), multi-step protein purification assessed by SDS-PAGE*,* 10 µl of each sample was loaded with 6x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. All conditions are labelled on the X-axis (C= Cell pellet lysate, M= Media lysate). The Y-axis has been labelled with the respective marker SeeBlue™ Plus 2 Pre-Stained protein standard. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate scanner. ***6.2c-d (NTD and RBD), 6.2f (FL spike protein),*** Western blotting to probe low sensitivity of protein expression***.*** 20 µl of cell pellet lysates (*figure 6.2 c and f*) or media lysates (*figure 6.2c*), were loaded with 6x Laemmli sample buffer on a 4-15% (v/v) resolving gel. ***6.2d***, demonstrates the N-terminal domain FLAG immunoprecipitated sample subjected to elution, and its remainder beads following TEV treatment. All blots were incubated with 5% (w/v) milk in 1x TBS-Tween-20 and immunoblotted with 1 µg/ml of anti-FLAG (**DYKDDDDK**) antibody. C= Cell pellet lysate, M= Media Lysate. The Y-axis has been labelled with the respective marker being represented (MagicMark™ XP Western Protein Standard). The blot was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 2-5 minutes.

Construct	Cell Pellet Lysate	Cell Media Lysate	Eluted Following TEV treatment	Bound to Post-TEV beads
S1	**Yes**	**No**	**No**	**Yes**
RBD	**Yes**	**Very Low**	**No**	**Yes**
NTD	**Yes**	**No**	**No**	**Yes**
FL Spike	**Yes**	**No**	**No**	**Yes**

Table 6.1, Summary of the expression of each SARS-CoV-2 protein constructs. All observations made in Figures 6.2a-f, are summarised in this table. No construct eluted into the elution buffer, with all samples being retained on the FLAG-resin post TEV treatment.

6.2.2.2 Troubleshooting to Confirm TEV-Protease is Functional

Although all constructs were being expressed successfully into the cell pellet lysate, I was unable to elute this off from the FLAG-resin using the well-established TEV-cleavage procedure. As a trouble-shoot experiment I wanted to confirm whether the TEV protease used in my study was working efficiently.

I purified a dually expressed construct (MEIR/HDAC) (donated by Professor John Schwabe, LISCB, University of Leicester) since it consistently expresses well in HEK293F suspension cells. High levels of protein was eluted into the elution buffer (~55 kDa) (Figure 6.3) with minimal left behind on the post-TEV cleaved beads. This confirms that the TEV cleavage protease and buffers were optimum and the technique itself was being handled correctly.

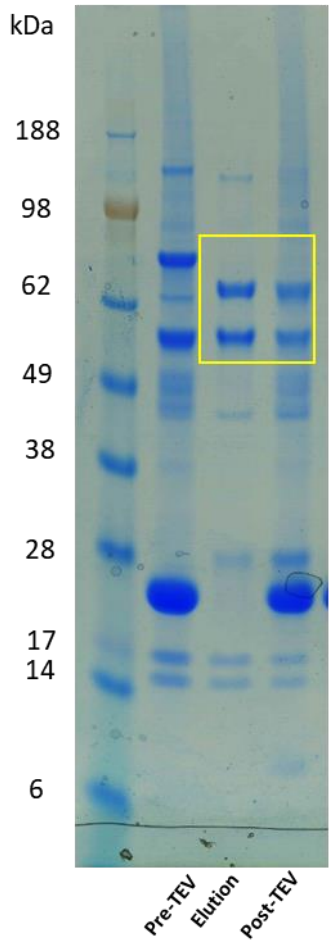


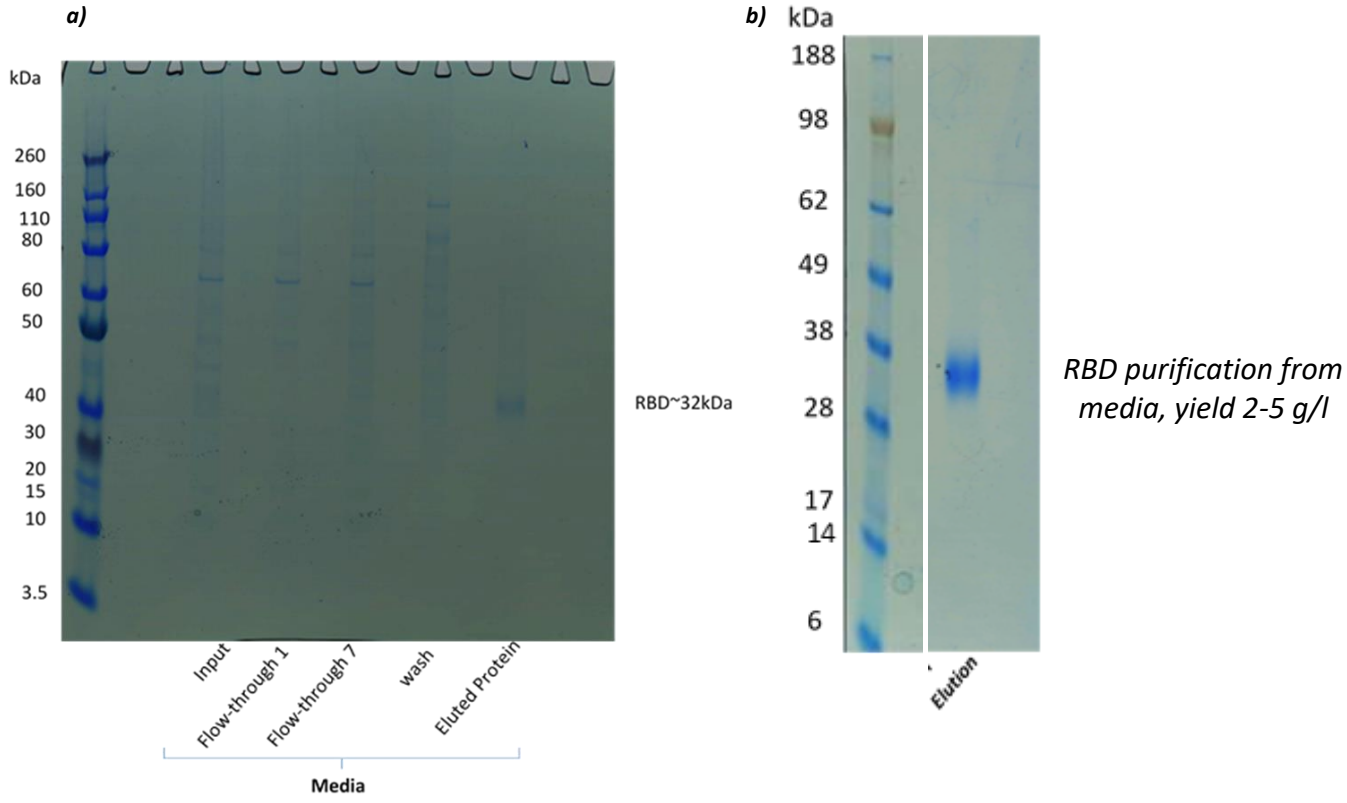
Figure 6.3, Troubleshooting experiment to confirm TEV-cleavage buffer used is functional. Anti-FLAG pulldown of FLAG-Tagged HEK293F suspension cells expressed MEIR/HDAC proteins assessed by SDS-PAGE, 10 µl of each sample was loaded with 6x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. All conditions are labelled on the X-axis. The Y-axis has been labelled with the respective marker SeeBlue™ Plus 2 Pre-Stained protein standard. Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate scanner.

6.2.2.3 CD5-RBD Expression

Although I had successfully managed to clone and express all the constructs of interest, the expressed proteins failed to secrete into the media from HEK293F suspension cells. Whilst the constructs were present in the cell pellet lysate, I was unable to cleave the FLAG-tag via the TEV cleavage site to generate soluble proteins.

I was kindly gifted a receptor binding domain construct with a CD5-derived secretion signal fused at its N-terminal domain and cloned into pcDNA3.1 vector from Professor Nicholas Brindle (LISCB, University of Leicester). I transfected this construct into suspension HEK293F cells. It was successfully expressed and purified by affinity chromatography using a Ni-NTA column (Figure 6.4a). The eluted purified protein was dialysed and concentrated (~five-fold) and a good yield (2-5 g/l of media) was indicated by concentration estimate and band intensity in Figure 6.4b.

Given our experiences with the different construct design strategies, for future work with the N-terminal domain, S1 region, and full-length SARS-CoV-2 spike protein constructs the next logical step would be to express these constructs similarly to this receptor binding domain construct. i.e. to clone these constructs into the pcDNA3.1 vector with the same CD5-derived secretion signal.



RBD purification from media, yield 2-5 g/l

Figure 6.4, RBD purification from cell media using Ni-NTA affinity column. 6.4a, Multi-steps in the production of purified recombinant RBD (CD5 and His-tagged). Gel was run at 150 V for 45 minutes. 10 µl of each sample was loaded with 6x Laemmli sample buffer on a 4-12% (v/v) resolving gel. The gel was stained overnight in Expedeon instant blue. Lanes have been labelled with the appropriate conditions (on the X-axis). Input= initial starting material prior to injecting through Ni-NTA column. Flow-through 1 and 7 represents flowthrough sample which filled 7 Falcons, sample was taken from Falcon 1, and Falcon 7. Wash=Column washed with wash buffer. Eluted protein= Column injected with elution buffer. The Y-axis has been labelled with the respective marker molecular weight being represented (SeeBlue^{TM} Plus 2 Pre-Stained protein standard). Before image capture, the gel was washed twice in $_{d}H_2O$, and then scanned using an appropriate scanner. ***6.4b)*** 5x concentrated (buffer dialysed) RBD preparation. Eluted protein sample shown in figure 6.4a was 5x concentrated using a 3k MWCO 0.5 ml spin filter, and analysed through the above mentioned PAGE and staining technique.

6.2.3 Optimisation of Purified RBD Concentration for Cellular Stimulations to Release Inflammatory Cytokines in A549 and THP-1 Cells

In the interests of time, we proceeded to study interactions of the purified SARS-CoV-2 spike protein receptor binding domain with lung epithelial cells and/or macrophages. We focused on cytokine outputs (IL-6, IL-8, TNF-α) from our cell models (work with Dr Panayiota Stylianou and Dr Omeed Darweesh, PDRAs, Gooptu Group, University of Leicester). Secretion levels of these cytokines in response to the receptor binding domain were assessed relative to responses of LPS stimulation that we had previously demonstrated to be CD98-dependent. To assess submaximal cytokine release in both A549 and THP-1 cells, we optimised the concentration of SARS-CoV-2 receptor binding domain and the duration of stimulation (8 h and 24 h). Since cytokine release levels can differ in different cell types, optimal receptor binding domain concentrations were defined for both A549 and THP-1 macrophages. In addition, the optimised concentrations of SARS-CoV-2 receptor binding domain used had to fall within the linear range of the ELISA standard curve for each cytokine assessed.

Along with receptor binding domain induced cytokine expression, we assessed both positive and negative controls. The supernatant of non-stimulated conditions in the absence and presence of the same concentration of a general serum protein, bovine serum albumin, were selected as negative controls. The supernatant of LPS stimulated cells, a stimulus which releases IL-6, IL-8 and TNF-α significantly at a concentration of 2 ng/ml in multiple cell types (Gooptu group. *manuscript in preparation*) was selected as a positive control.

In A549 cells, treatment with the receptor binding domain induced IL-6 (Figure 6.5a) and IL-8 (Figure 6.5b) release in a concentration dependent manner qualitatively over both 8 h and 24 h. Interestingly the IL-6 and IL-8 cytokine responses to receptor binding domain stimulation peaked following interaction with a receptor binding domain concentration of 2 μg/ml for 24 h and so

this concentration was used to stimulate A549 cells for 24 h. The levels of TNF-α (Figure 6.5c) was similar between receptor binding domain stimulated concentrations of 0.2 μg/ml – 20 μg/ml (in both 8 h and 24 h stimulation period).

In M0-macrophage-THP-1 cells, the levels of IL-6 (Figure 6.6a), IL-8 (Figure 6.6b) and TNF-α (Figure 6.6c), increased linearly with increasing stimulation time and concentration of receptor binding domain. Receptor binding domain concentrations of 10 μg/ml and 20 μg/ml resulted in similarly strong secretion signals for all three cytokines at 24 h duration. MTS assay (Figure 6.7) demonstrated that cell viability was unaffected with the increased concentration of the receptor binding domain used. Since stimulation with 20 μg/ml receptor binding domain concentration displayed similar cytokine release to 10 μg/ml, the latter was used as the standard concentration for subsequent experiments in these cells. Typically, M1-macrophage phenotype cells display increased levels of IL-6 and IL-8, whilst M2-macrophage phenotype cells display increased levels of IL-10 (Orecchioni *et al*. 2019). Therefore, our results indicate that the M0-macrophage-THP-1 cells were polarised to M1-macrophage cells under receptor binding domain stimulated conditions, since increased levels of the pro-inflammatory cytokines IL-6 and IL-8 were consistently observed.

These set of experiments were predominantly conducted to observe whether the receptor binding domain induces apparent biological cytokine response differences to non-treated conditions. The individual data points qualitatively displayed increased release in cytokine patterns (of IL-6 and IL-8 in A549 cells, and all three cytokines in THP-1 cells), however, the statistical analyses of these experiments were not significant ($p < 0.05$). This is likely explained by the large data variability from each individual repeats, and the low n-numbers used for the titration experiment. This was indirectly addressed in the subsequent experiments (6.2.4 and 6.2.5) where the optimised chosen stimulated conditions were statistically significant compared to the untreated-(and all relevant) controls.

RBD concentration and stimulation length titration in A549 cells

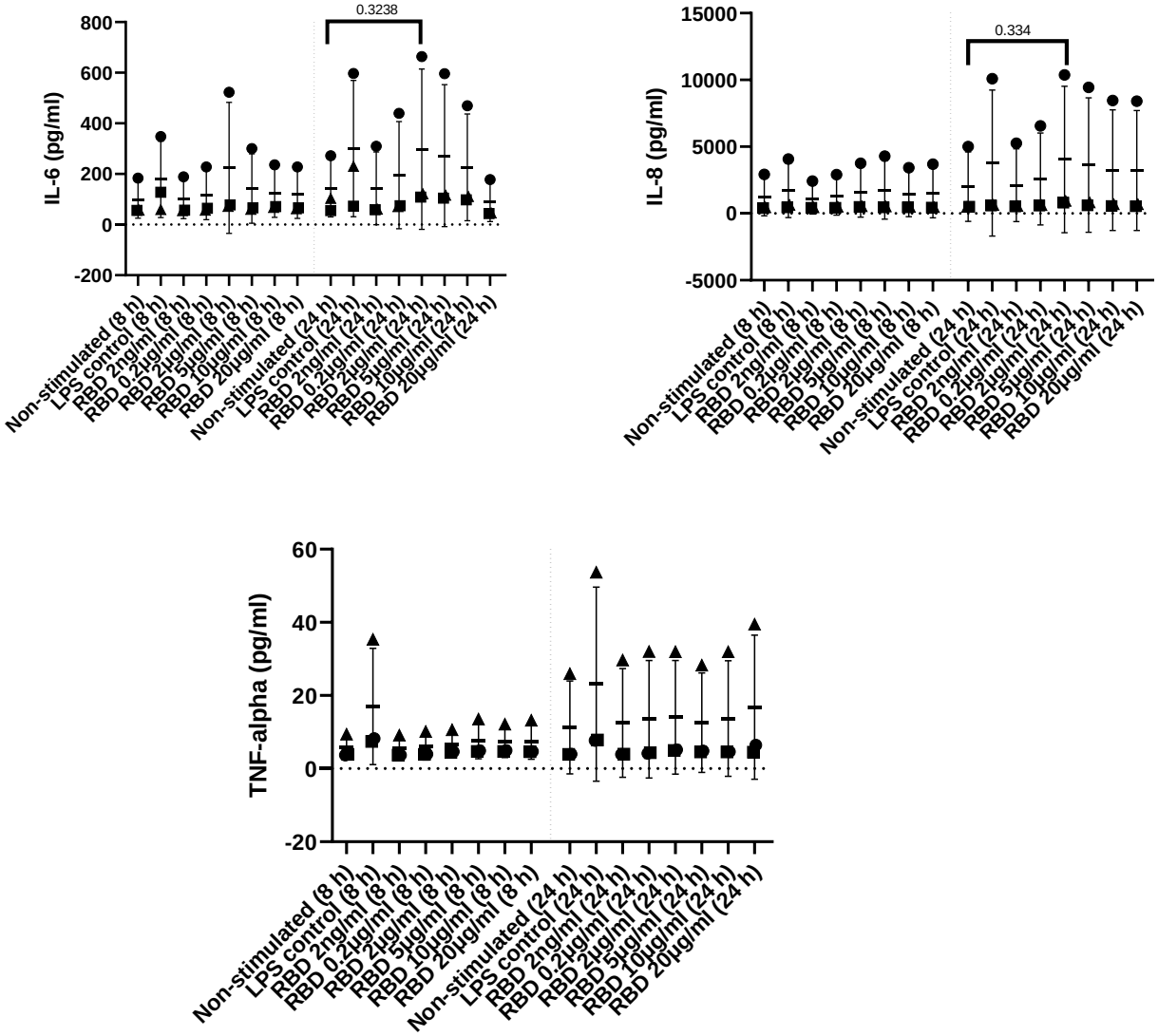


Figure 6.5, Titration experiment to assess optimal RBD concentration inducing IL-6, IL-8 and TNF-α release in A549 cells. A549 cells were non-stimulated or stimulated with 2 ng/ml of LPS (controls). Cells were treated with various concentrations of RBD for 8 h (left section of graphs) or 24 h (right section of graphs). Cell supernatants were harvested at the specified times, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. Statistical analysis were obtained from multiple paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3)

RBD concentration and stimulation length titration in THP-1 cells

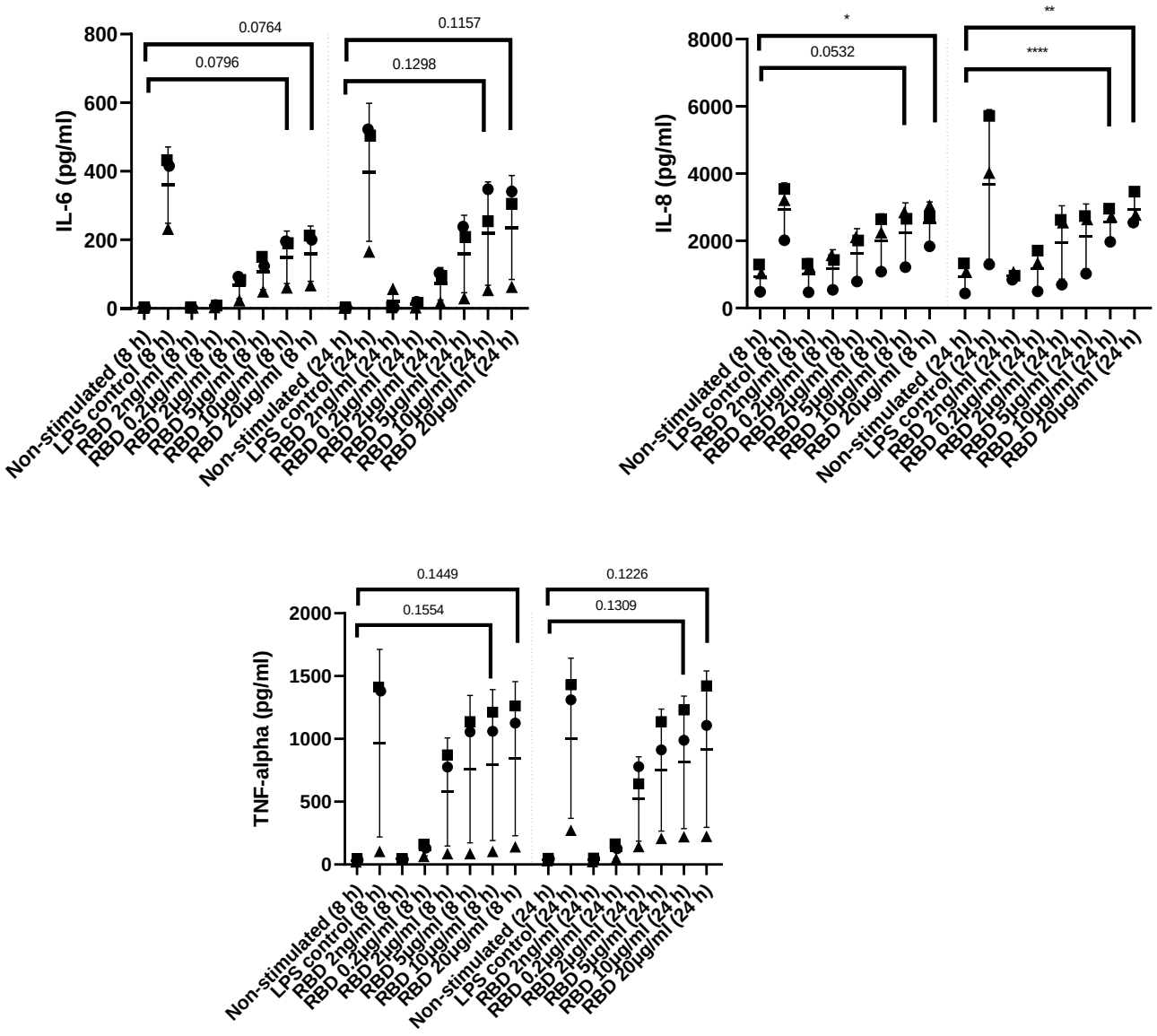


Figure 6.6, Titration experiment to assess optimal RBD concentration inducing IL-6, IL-8 and TNF-α release in THP-1 cells. THP-1 cells were non-stimulated or stimulated with 2 ng/ml of LPS (controls). Cells were treated with various concentrations of RBD for 8 h (left section of graphs) or 24 h (right section of graphs). Cell supernatants were harvested at the specified times, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. Statistical analysis were obtained from multiple paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3).

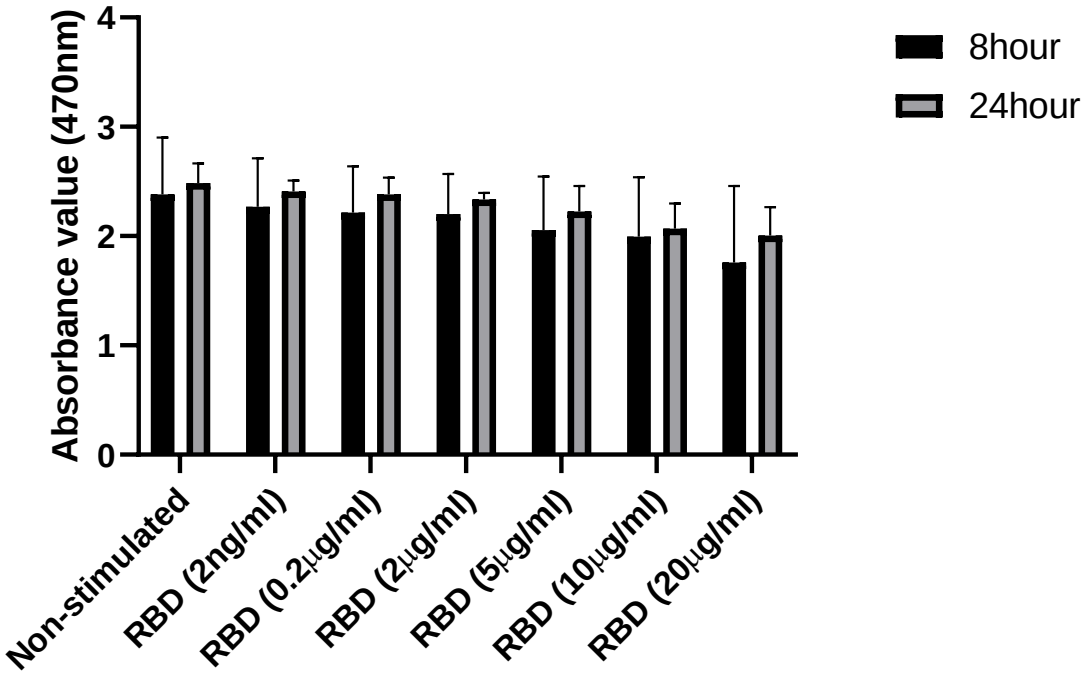


Figure 6.7, RBD titration concentration and stimulation duration does not alter cell proliferation/survival in THP-1 cells, illustrated by MTS assay. THP-1 cells were non-stimulated or stimulated with various concentrations of RBD for 8 h (black bars) or 24 h (grey bars). After the stimulation period, cell viability was assessed by MTS assay. RBD concentrations at 10 μg/ml and 20 μg/ml did not affect THP-1 cell proliferation/survival rates when compared to controls. X-Axis displays the appropriate conditions. The Y-Axis displays the level of the cell signal, depicted from the UV absorbance reading at 470 nm. Data are displayed as their mean (black and grey bars) and ± standard deviation. Multiple paired t-test were conducted. (n=3).

6.2.4 Galectin-3-Fibrosome Regulating RBD Induced Pro-Inflammation

6.2.4.1 Inhibiting CD98hc using Cynaropicrin Attenuates RBD Induced IL-6 and IL-8 Release, but not TNF-α Release, in Both A549 and THP-1 Cells

Our group has demonstrated that CD98:β1-integrin complex critically mediates inflammatory responses in alveolar epithelial cells (Wang *et al*. *Manuscript in preparation*). SARS-CoV-2 receptor binding domain has potential to bind to CD98hc:β1-integrin since it carries the tripeptide RGD amino acid sequence, a motif recognised and bound by β1-integrin (Sigrist *et al*. 2020).

Cynaropicrin is a sesquiterpene inhibitor of integrin-dependent functions of CD98 (Cho *et al*. 2004). Our group's earlier work demonstrated that treatment with 5 μM of cynaropicrin or CD98 silencing abrogates IL-6 and IL-8, but not TNF-α responses to injury stimuli without compromising cell viability.

We assessed whether the spike protein receptor binding domain stimulated cytokine responses in a CD98-dependent manner in both cell types. A549 and THP-1 cells were pre-treated with the receptor binding domain (2 μg/ml or 10 μg/ml respectively) for 1 h, followed by the secondary treatment of 5 μM cynaropicrin for a total of 24 h. The cell supernatant was then harvested and levels of IL-6, IL-8 and TNF-α were quantified by ELISA.

To verify the level of cytokine release was specifically proportional to cells pre-stimulated with receptor binding domain followed by cynaropicrin, and not as a result of buffer or solution effects, five control conditions were assessed (Lanes 1-5, Figures 6.8 and 6.9). The controls assessed are briefly described as follows:

1) *Unstimulated control*.

2) *Vehicle control*, this signified the reagent (DMSO) cynaropicrin was reconstituted in.

3) *Vehicle + TBS supplemented with 5% (v/v) glycerol buffer* (our purified RBD was dialysed into TBS buffer supplemented with 5% (v/v) glycerol).

4) *Cynaropicrin + TBS supplemented with 5% (v/v) glycerol buffer*.

5) *RBD + vehicle (DMSO) control*.

Comparing the signal observed in control number '5' to the dual treatment condition (RBD + cynaropicrin) was of most importance. This signifies the levels of the cytokine release in response to receptor binding domain stimulation with and without the presence of the CD98hc inhibitor, cynaropicrin.

In the initial instance, receptor binding domain induces increased release of IL-6, IL-8 and TNF-α in both A549 (Figure 6.8) and THP-1 (Figure 6.9) cells. This is observed by comparing the 'RBD + vehicle control' lane to the untreated lane. Levels of IL-6 were considerably lower in THP-1 cells (Figure 6.9a) than A549 cells (Figure 6.8a). An explanation for this could be that THP-1 cells are not professional inflammatory macrophages and initiate their cytokine responses from a M0 (resting) macrophage phenotype. As expected, cynaropicrin treatment did not attenuate RBD-induced TNF-α release in both A549 (Figure 6.8c) and THP-1 (Figure 6.9c) cells. This implies that RBD-induced TNF-α pro-inflammatory signalling is not dependent on CD98hc *in vitro*. In contrast cynaropicrin inhibition of CD98hc significantly attenuated RBD-induced IL-6 and IL-8 release in both A549 (Figure 6.8a-b) and THP-1 (Figure 6.9a-b) cells. This implies RBD-induced IL-6 and IL-8 pro-inflammatory signalling is CD98hc dependent for both A549 and THP-1 cells *in vitro*.

RBD induced cytokine release in A549 cells

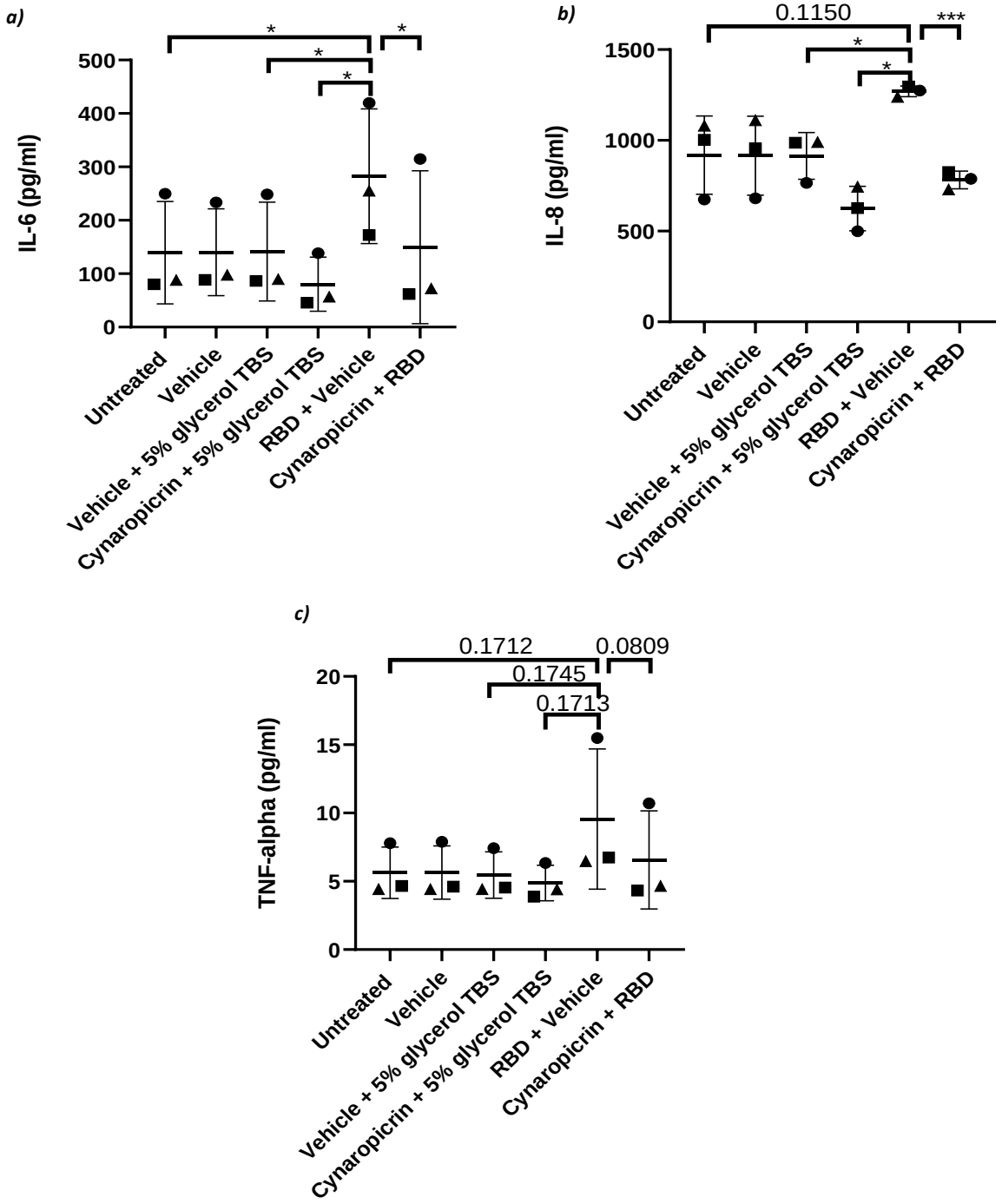


Figure 6.8, Co-stimulation of RBD with cynaropicrin reduces IL-6 and IL-8 release, though TNF-α release remains stabilised in A549 cells. A549 cells were untreated or (co)-treated with relevant controls and RBD with Cynaropicrin. Cells were treated with 2 µg/ml of RBD with 5 µm of Cynaropicrin for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3).

RBD induced cytokine release in THP-1 cells

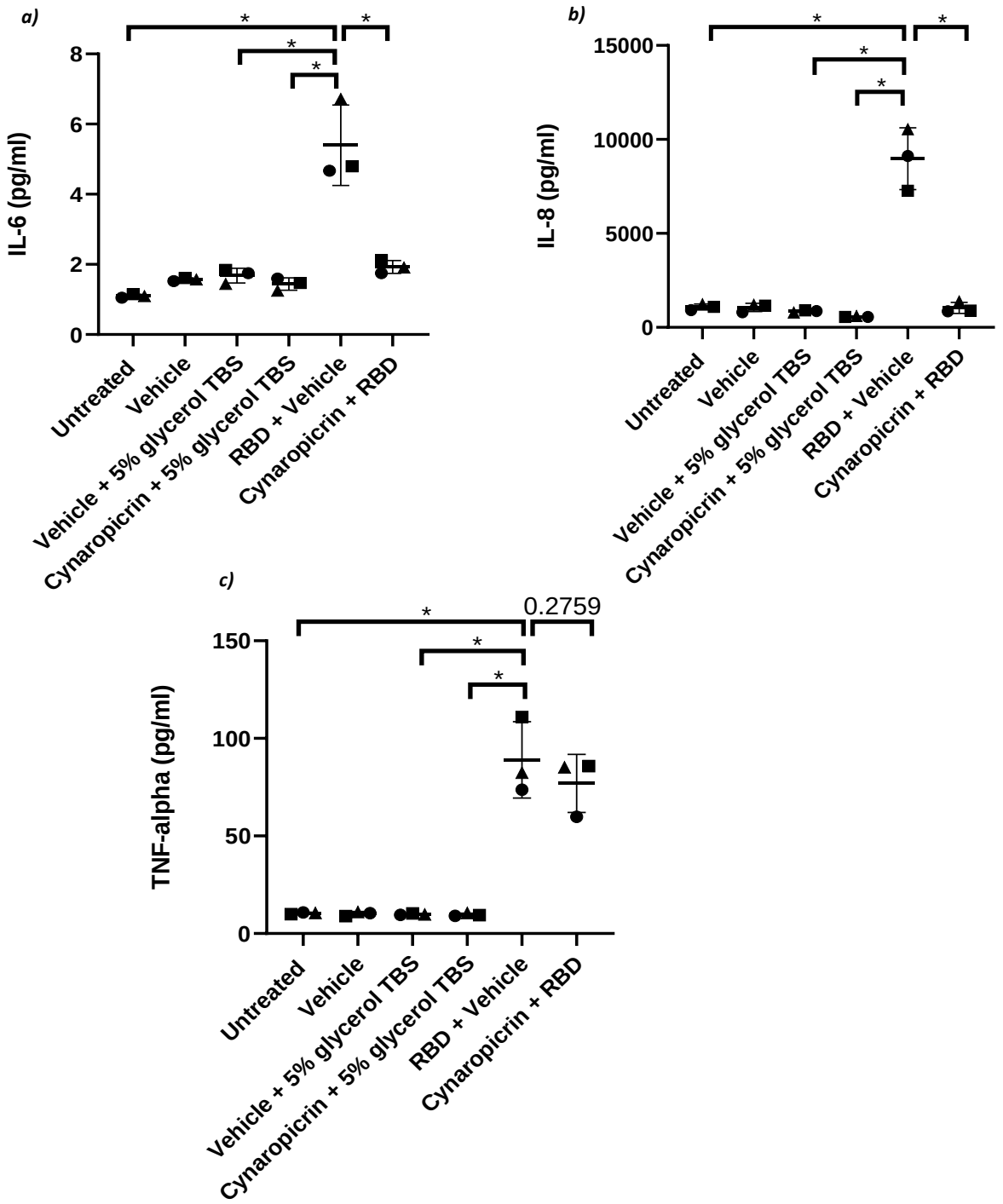


Figure 6.9, Co-stimulation of RBD with cynaropicrin reduces IL-6 and IL-8 release, though TNF-α release remains stabilised in THP-1 cells. THP-1 cells were untreated or (co)-treated with relevant controls and RBD with Cynaropicrin. Cells were treated with 2 µg/ml of RBD with 5 µm of Cynaropicrin for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3).

6.2.4.2 Effects of Silencing CD98hc and Galectin-3 upon RBD-Induced Cytokine Release in A549 cells

Silencing of CD98 and galectin-3 using the lipofectamine 2000 transfection reagent was achieved as previously in A549 cells (2.8.2). This method was not effective in THP-1 cells despite doubling the concentration of siRNA used (Figure 6.10), consistent with published experience in this cell type (Mae *et al*. 2014, Jing *et al*. 2014, ThermoFisher Scientific. 2021).

Silencing of CD98 abrogated IL-6 and IL-8 responses to stimulation by the SARS-CoV-2 receptor binding domain (Figure 6.11), consistent with the data from cynaropicrin treatment. The TNF-α response to such stimulation was lower and CD98 silencing was associated with less evidence of a significant effect on this (n=4). Silencing of galectin-3 had very similar effects (Figure 6.12). Levels of IL-6 were considerably lower than that observed in 6.2.4.1. A rational explanation for this could be based around the cell culture conditions used to achieve the knockdown of both galectin-3 and CD98. The cell culture media does not contain any FBS, which is an important cell growth supplement and is present under normal non-knockdown conditions. Together these data imply that receptor binding domain-induced IL-6 and IL-8 pro-inflammatory signalling is both CD98hc and galectin-3 dependent in A549 cells *in vitro*. The findings are consistent with co-stabilisation of these proteins at the cell surface, and their critical involvement in mediating IL-6 and IL-8 responses to SARS-CoV-2 spike receptor binding domain in A549 lung epithelial cells.

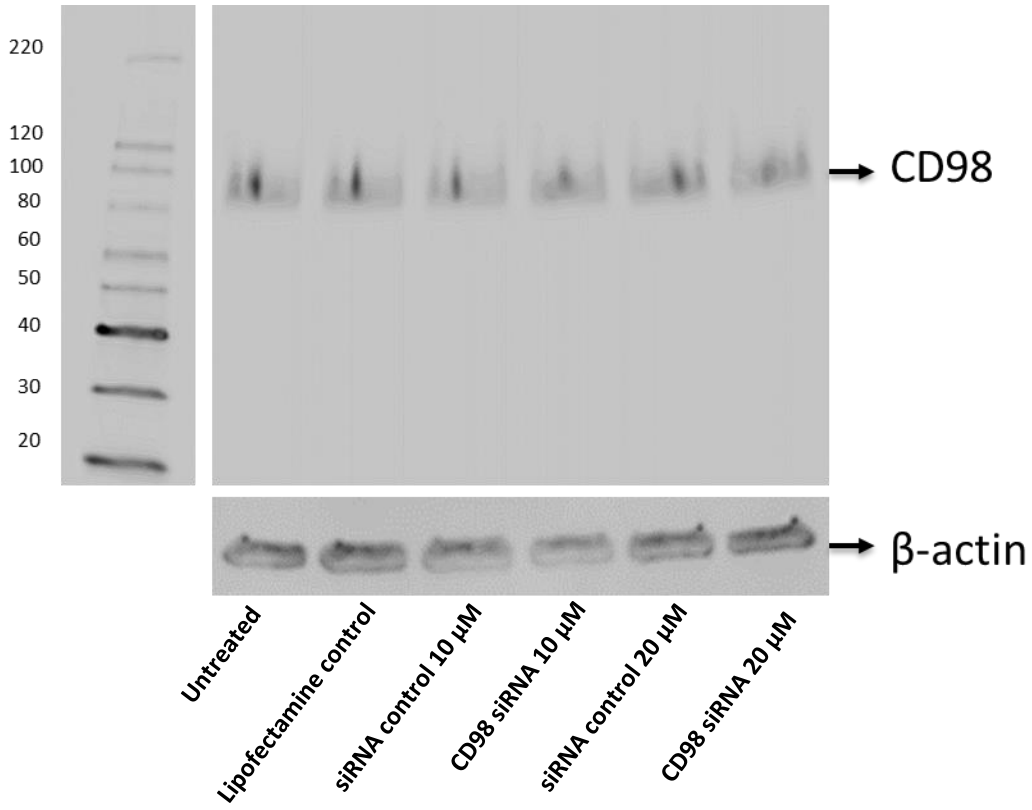


Figure 6.10, Western-blot assessing the ability to silence CD98 in THP-1 cells as successfully done in A549 cells. THP-1 whole cell lysate was prepared by using RIPA Lysis Buffer. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml CD98 and 0.04 µg/ml β-actin. Lane 1-6 has been labelled with the appropriate conditions (on the X-axis). The Y-axis has been labelled with the respective marker being represented. β-actin was used as a housekeeping protein. The blots were immersed with EZ-ECL chemiluminescent detection kit and the images were taken at exposure time of 5 minutes. (n=3).

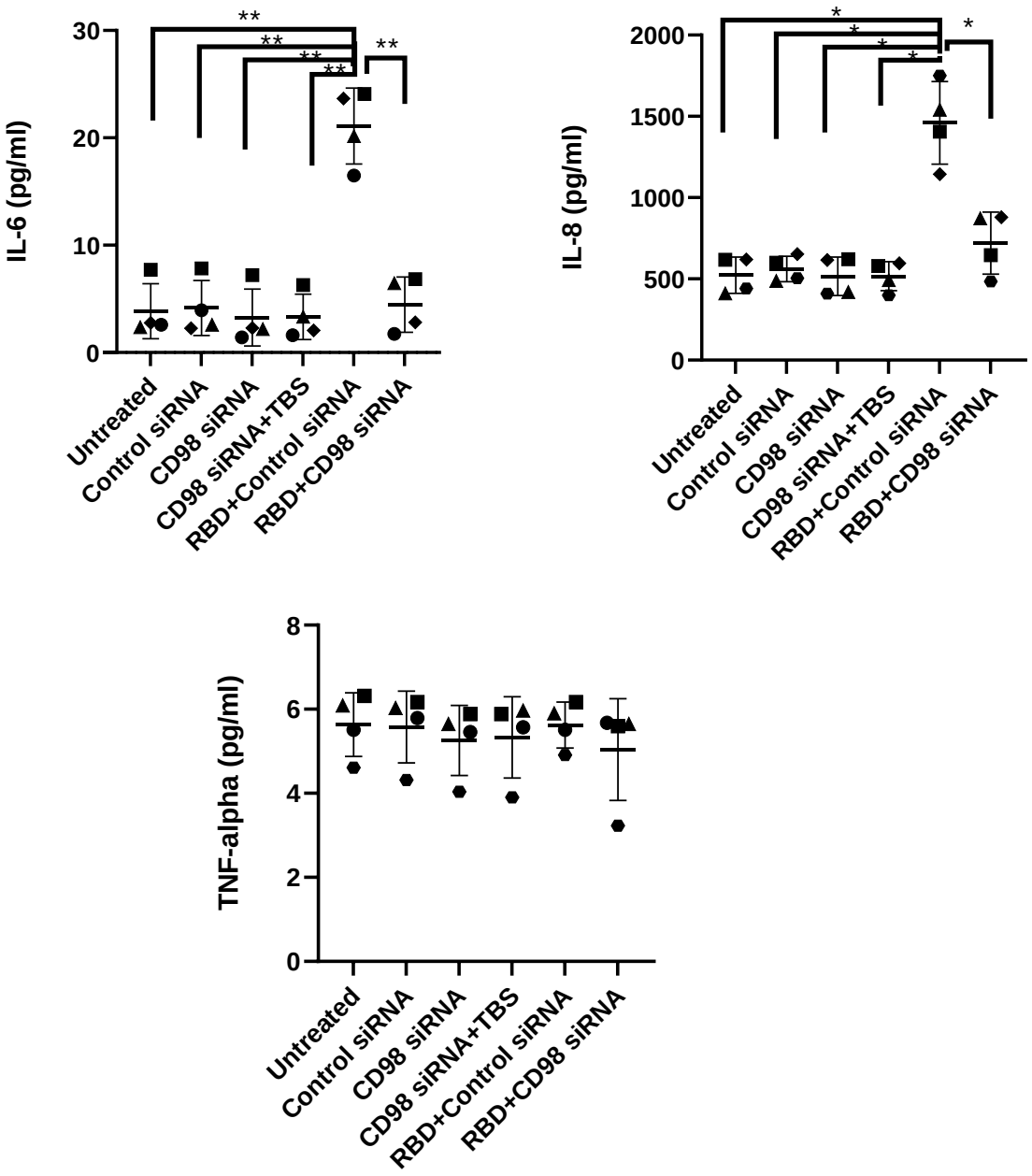


Figure 6.11, CD98 silencing abrogates RBD induced IL-6 and IL-8 release, though TNF-α levels remain stabilised in A549 cells. A549 cells were subjected to siRNA mediated CD98 silencing. All appropriate controls were assessed (untreated and siRNA control). Cells were stimulated with RBD (2 µg/ml) for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=4).

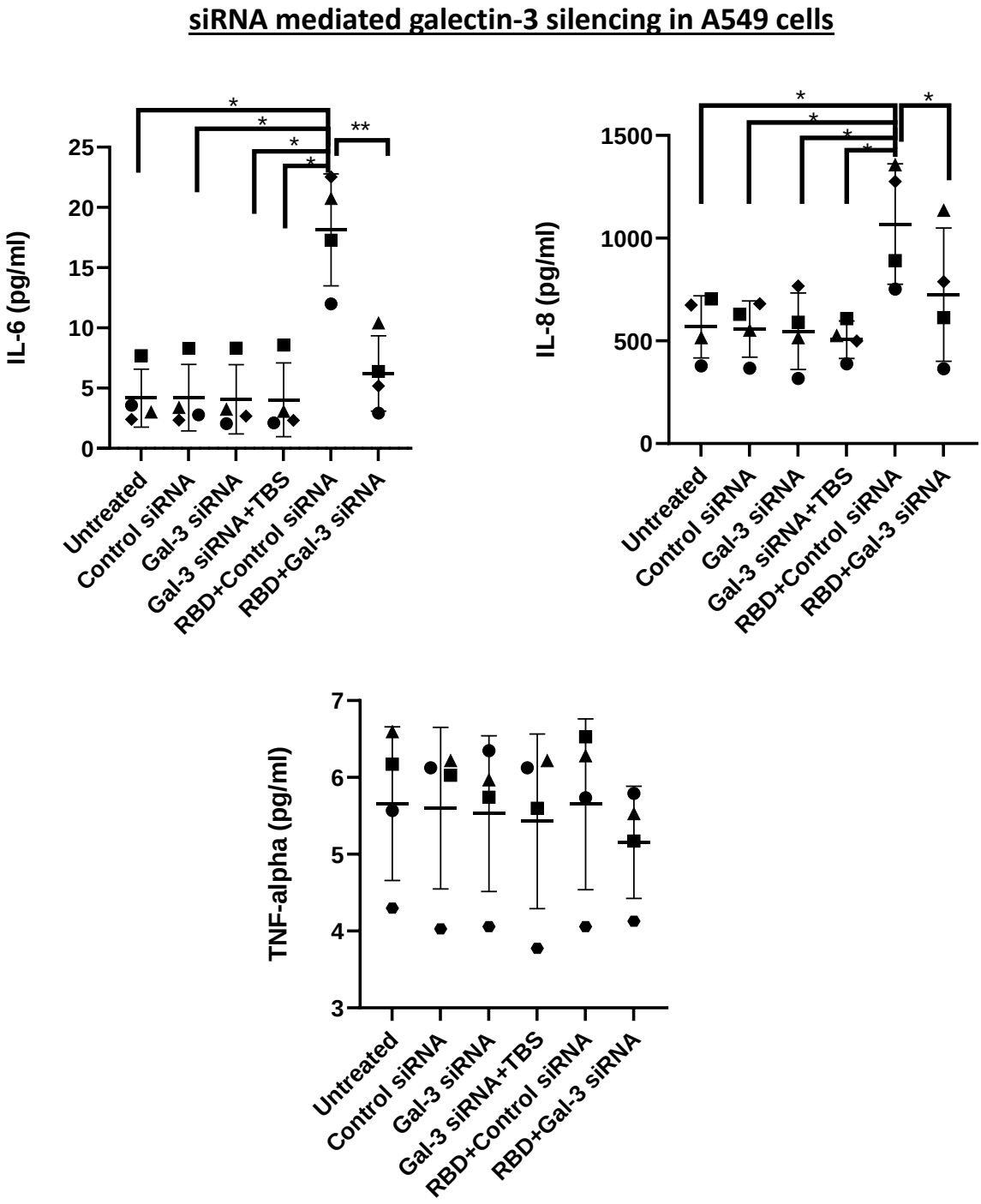


Figure 6.12, Galectin-3 silencing abrogates RBD induced IL-6 and IL-8 release, though TNF-α levels remain stabilised in A549 cells. A549 cells were subjected to siRNA mediated galectin-3 silencing. All appropriate controls were assessed (untreated and siRNA control). Cells were stimulated with RBD (2 µg/ml) for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (***a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=4).

6.2.5 Effects of Dexamethasone Treatment upon RBD-Induced Cytokine Release in A549 and THP-1 cells

Dexamethasone is an anti-inflammatory corticosteroid (Coutinho and Chapman. 2011), exerting its effect by binding to glucocorticoid receptors (Mayeux *et al*. 1983, Meijsing *et al*. 2007). COVID-19 patients treated with dexamethasone show lower 28-day mortality amongst those who were receiving either invasive mechanical ventilation or oxygen alone (Horby *et al*. 2020, RECOVERY trial. 2021). No benefit is seen in individuals self-ventilating on air.

We wanted to assess the capacity of this corticosteroid in reducing pro-inflammatory IL-6, IL-8 and TNF-α cytokine release. Both A549 and THP-1 cells were treated with dexamethasone in the presence of SARS-CoV-2 receptor binding domain stimulation. Physiologically- and pharmacologically-relevant levels of dexamethasone administration are between 10^{-5}-10^{-7} M concentration (Kream *et al.* 1983. Hawrylowicz *et al.* 2002). We decided to use the midpoint of this range, 10^{-6}M (1 μM). To assess the anti-inflammatory characteristics of dexamethasone, we pre-treated both A549 and THP-1 cells with receptor binding domain (2 μg/ml or 10 μg/ml respectively) for 1 h, followed by the secondary treatment of 1 μM dexamethasone for a total of 24 h. The cell supernatant was then harvested and levels of IL-6, IL-8 and TNF-α were quantified by ELISA.

To verify the level of cytokine release was specifically proportional to cells pre-stimulated with RBD followed by dexamethasone, and not as a result of buffer or solution effects, five controls were assessed:

1) *Unstimulated control*.

2) *Vehicle control*, this signified the reagent (ethanol) dexamethasone was reconstituted in.

3) *Vehicle + TBS supplemented with 5% (v/v) glycerol buffer* (our purified RBD was dialysed into TBS buffer supplemented with 5% (v/v) glycerol).

4) *Dexamethasone + TBS supplemented with 5% (v/v) glycerol buffer*.

5) *RBD + vehicle (ethanol) control*.

Comparing the signal observed in control number '5' to the dual treatment condition (RBD + dexamethasone) was of most importance. This comparison highlights changes in the levels of cytokine release in response to receptor binding domain stimulation due to dexamethasone treatment.

SARS-CoV-2 receptor binding domain induced significant release of IL-6, IL-8 and TNF-α in both A549 (Figure 6.13) and THP-1 (Figure 6.14) cells. This is observed by comparing the 'RBD + vehicle control lane' to the 'untreated' lane. In agreement with the clinical use of dexamethasone in COVID-19 patients, our data displayed dexamethasone significantly attenuating RBD-induced IL-6, IL-8 and TNF-α release in both A549 (Figure 6.13a-c) and THP-1 (Figure 6.14a-c) cells. This implies that the anti-inflammatory inhibition mechanism by which dexamethasone functions has profound inhibition in the release of IL-6, IL-8 and TNF-α when assessing cells *in vitro* subjected to receptor binding domain stimulation.

RBD induced cytokine release in A549 cells

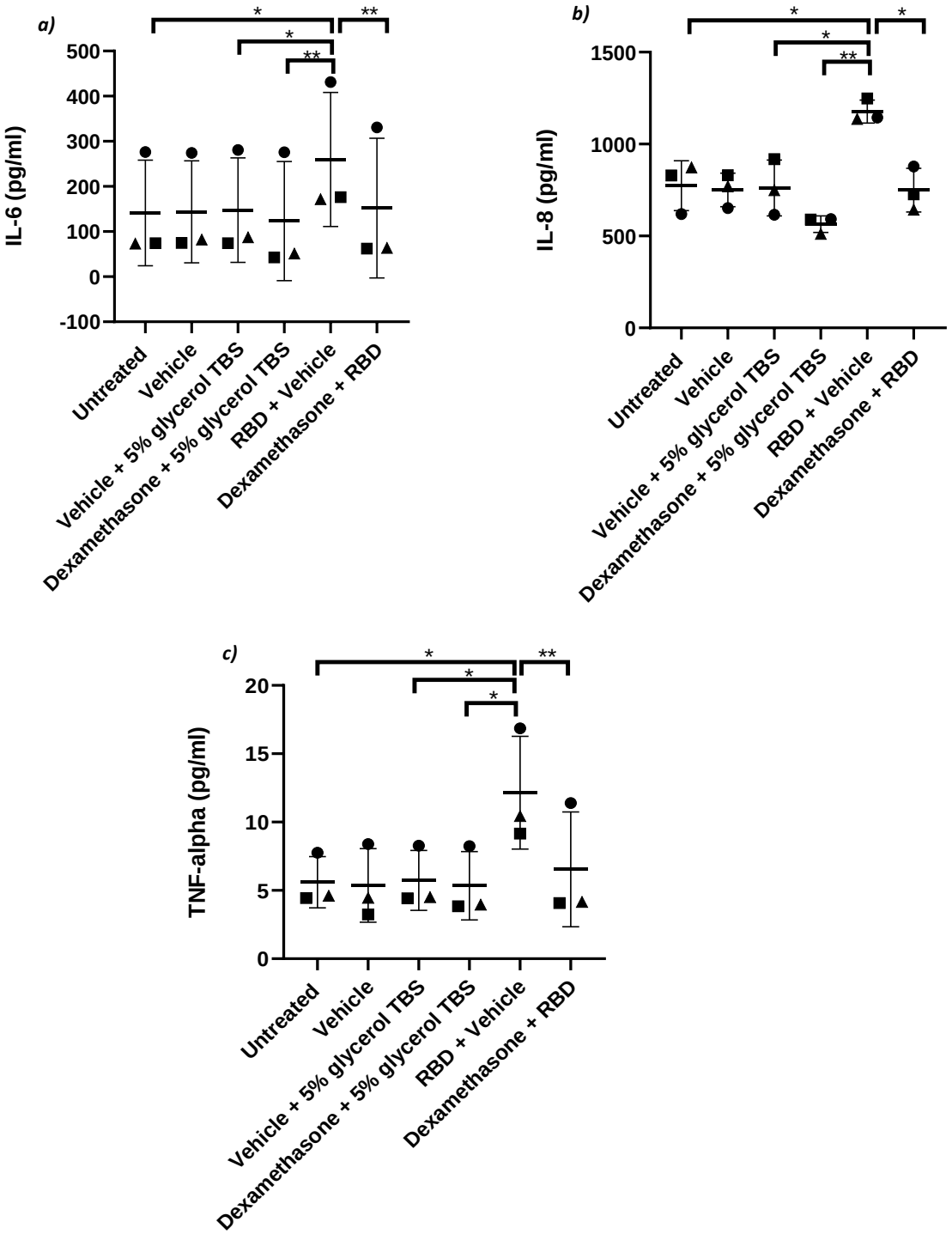


Figure 6.13, Co-stimulation of RBD with dexamethasone reduces IL-6, IL-8 and TNF-α release in A549 cells*.** A549 cells were untreated or (co)-treated with relevant controls and RBD with Dexamethasone. Cells were treated with 2 µg/ml of RBD with 1 µm of Dexamethasone for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3).

RBD induced cytokine release in THP-1 cells

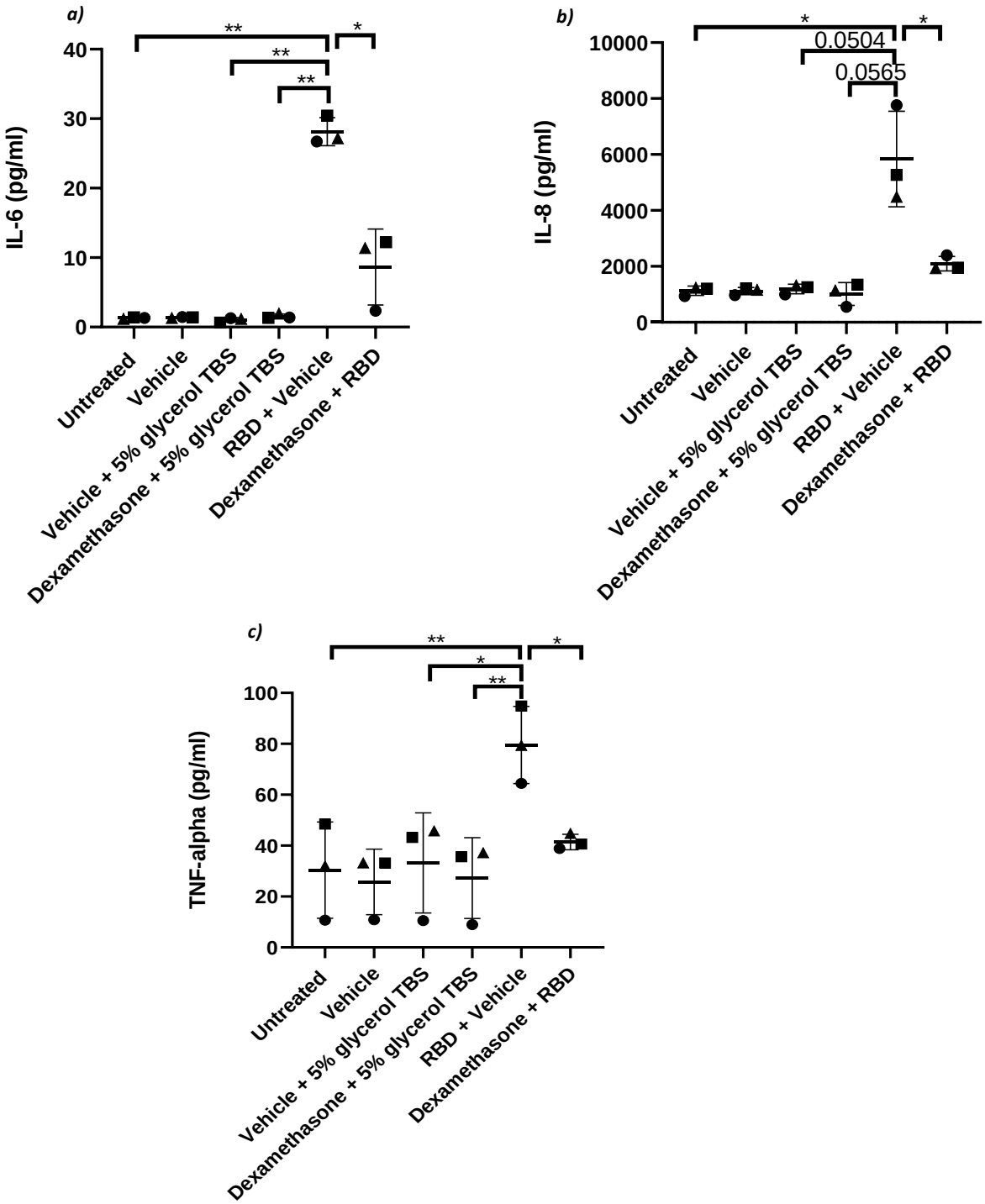


Figure 6.14, Co-stimulation of RBD with dexamethasone reduces IL-6, IL-8 and TNF-α release in THP-1 cells*.** THP-1 cells were untreated or (co)-treated with relevant controls and RBD with Dexamethasone. Cells were treated with 2 µg/ml of RBD with 1 µM of Dexamethasone for 24 h. Cell supernatants were harvested post stimulation, and IL-6 (a***) and IL-8 (***b***) and TNF-α (***c***) levels were quantified using the appropriate ELISA kit. The graphs are presented as a scatter-plot. Data from different sets of biological repeats (varying by passage number) are represented by different symbols. The X-Axis displays the appropriate conditions. The Y-Axis displays the level of the probed cytokine release concentration as pg/ml. Mean values and standard deviations are plotted. The asterisk plotted on the graph represents significance from paired t-test (*p<0.05, p**<0.01, ***p<0.001). (n=3).

6.3 Discussion

Cloning of FLAG-tagged S1, receptor binding domain, N-terminal domain and full length-SARS-CoV-2 constructs in the pLeics-49 vector was successful with the endogenous secretion signal. However, the protein constructs were not efficiently secreted into the media and the full-length protein was vulnerable to protease-mediated cleavage. Other investigators who have achieved success with structural studies of the full length protein have used a Pro-Pro double mutant that renders it resistant to cleavage. This may be useful for our group to incorporate for future studies with the full-length protein.

HEK293F suspension cells was chosen for protein expression over bacterial cell culture for several reasons. These include the absence of specific co-factors, chaperones, and post-translational modifications (spike protein glycosylation is important for protein function) that may cause loss of function, misfolding, and disruption of protein-protein interactions of surface receptors and secreted proteins in bacterial cell culture (Portolano *et al*. 2014). Using mammalian cell expression system addresses these drawbacks as it provides a eukaryotic expression environment.

Although cloning was achieved as planned, the constructs were poorly expressed in HEK293F suspension cells. Amanat *et al*. 2020, generated S1-constructs that secreted into the media with the 'native' endogenous secretion signal sequence (MFVFLVLLPLVSSQ). In my hands this was not an effective strategy. Much of the protein in the lysate was retained within the cell pellet (Figures 6.2a-f). This would not have been a problem if expression were higher, and the yield was similar to expressed proteins in a bacterial cell cultured pellet, since it will generate enough protein to conduct downstream *in vitro* cellular stimulations. However, alongside the issue of cell pellet retained protein expression, I was unable to elute these proteins from the FLAG-immunoprecipitation beads by TEV cleavage. Therefore, lanes of an SDS-PAGE gel loaded with

this sample displayed the expected molecular weight of the protein expressed. I confirmed this was not an experimental handling error, or an issue with the reagents such as the TEV cleavage buffer by expressing and successfully eluting a positive control protein provided by Professor John Schwabe (LISCB, University of Leicester). The design of each construct may explain why the FLAG-tags (at C-termini of each construct) were not cleaved by TEV. The TEV-cleavage sites at the C-termini end of each construct were not spaced between the end of the construct sequence and the TEV cleavage site itself. Addition of a linker sequence (such as 'GSGS...') between the end of each construct sequence and the beginning of TEV site could aid in TEV recognition and cleavage (Correa *et al*. 2014).

Expression of the SARS-CoV-2 receptor binding domain construct with the CD5-derived secretion signal in the pcDNA3.1 vector was successful, allowing its use in cell culture experiments. Future strategies for further optimising production of these proteins include dual-tagging this construct, by cloning in a FLAG tag beyond the TEV-cleavage site in addition to the current 6x His-tag within pCDNA3.1. This would allow us to compare the yield between anti-FLAG immunoprecipitation versus His-tag affinity chromatography (optimising maximal protein generation).

We also plan to clone in the other spike protein constructs (N-terminal domain and S1 region) into the same system (pcDNA3.1 vector; CD5 derived signal sequence; His-tag parent construct) with additionally adding TEV cleavage site and FLAG tag to make the constructs dual-tagged. This approach will be undertaken by PhD student Charlie Hitchman (Gooptu group). To achieve this, a synthetic SARS-CoV-2 S1 region gDNA construct has been brought and will be cloned into the digested parent construct. Once dual-tags are inserted, the N-terminal domain and receptor binding domain can then be amplified from the S1 cloned construct and inserted into the vector again.

We used the purified protein to stimulate cells and conduct downstream experiments to characterise functional outputs, such as measuring inflammatory cytokine responses, as has been described here for the receptor binding domain.

Significant cytokine responses of IL-6, IL-8 and TNF-α release were in response to SARS-CoV-2 receptor binding domain stimulation apparent in both A549 and THP-1 cells. These findings may model parenchymal epithelial (A549), and macrophage (THP-1) responses to the initial encounter with the virus. Furthermore, under such stimulated conditions, the increase of these cytokines observed from resting (M0) macrophage cells (THP-1), denoted the differentiation into M1 (pro-inflammatory) macrophages (Orrechioni *et al*. 2019). Stimulating both the A549 and THP-1 cells using the full-length SARS-CoV-2 or just the S1 region (consisting of the N-terminal domain and receptor binding domain) potentially would result in further amplified cytokine storm. This speculation is made since the N-terminal domain consists of galectin-3 like characteristics, which therefore could result in increased retention of the receptor binding domain interactions to β1-integrin, and binding to TGF-βRII, thereby clustering the fibrosome proteins (as discussed in chapters 3 and 4).

Our data suggests that IL-6 and IL-8 release are modulated by CD98hc and galectin-3 under pathological conditions. IL-6 and IL-8 release was abrogated using the CD98hc inhibitor cynaropicrin, or silencing CD98hc and galectin-3 using our established siRNA mediated protocol. However, TNF-α levels remained low and unchanged, suggesting this cytokine pathway is not directly involved in transmitting the consequences of spike protein receptor binding, and CD98 or galectin-3 silencing do not affect this. This pattern of IL-6 and IL-8 reductions with unchanged levels of TNF-α was also observed in the previous studies with cynaropicrin or CD98 knockdown in both A549 and TT1 alveolar epithelial cell models after stimulation with LPS or mechanical stressors.

Dexamethasone suppresses inflammation by two main mechanisms, namely, transactivation and transrepression (King *et al*. 2012). During transactivation, the glucocorticoid diffuses into the cytoplasm, thereafter, binds to homo-dimeric glucocorticoid receptor-α (GRα), forming a GC-GRα complex (Barnes. 2006). The GC-GRα complex translocate into the nucleus and binds to glucocorticoid response element on promoter regions of multiple target anti-inflammatory genes (Barnes. 2006). This is assisted by in parallel recruitment of different transcriptional co-activators of these genes. These co-activators have histone acetyl transferase activity, facilitating the transcription of the anti-inflammatory proteins (Newton. 2000, Barnes. 2006). Transrepression mechanism is thought to inhibit (or repress) the expression of pro-inflammatory cytokines by interacting with multiple transcription factors (Hubner *et al*. 2015, Weikum *et al*. 2017). In contrast to transactivation, transrepression is mediated by the binding of glucocorticoid to monomeric GRα (Hubner *et al.* 2015). The GC-GRα complex recruit multiple co-activators but in this instance these function to either hinder the activity of histone acetylation or recruit histone deacetylases to reverse histone acetylation. This subsequently decreases the transcription of pro-inflammatory cytokines (Barnes. 2006).

Dexamethasone treatment has been demonstrated to improve outcomes in COVID-19 pneumonitis requiring respiratory support through supplemental oxygen and/or mechanical support of ventilatory function. In our cell assays, stimulation by the spike protein receptor binding domain subjected to subsequent dexamethasone treatment reduced levels of IL-6, IL-8 and TNF-α. Dexamethasone may therefore inhibit pathways mediated by CD98hc and/or galectin-3 as part of a broader mechanism of action.

Taken together these findings support the hypothesis that SARS-CoV-2 interactions with lung epithelial cells are inherently pro-inflammatory at the point of initial spike protein:cell surface protein encounters, even without or prior to infection. This concept is further supported by recent observations of Dr Panayiota Stylianou and Charlie Hitchman (PDRA and PhD student, Gooptu group, University of Leicester) in collaboration with Dr Richard Haigh (PDRA, Barer group,

University of Leicester) working with the SARS-CoV-2 virus. In line with other studies, A549 cells were not infected by the virus, suggesting all the interactions leading to cytokine amplification to be occurring on the cell surfaces. Calu-3 cells, a bronchial epithelial cell line, are readily infected due to high expression of ACE-2 focused in the apical plasma membrane (Ren *et al*, 2006). Calu-3 cells expressed galectin-3 and CD98 similarly to A549 cells. Expression was readily knocked down by targeted siRNA treatments. Interestingly, the potential for Calu-3 cell infection by SARS-CoV-2, was unaffected by CD98 or galectin-3 knockdown. Moreover, A549 cells did not show any obvious reduction in interactions with the receptor binding domain when these proteins were knocked down (Appendix, Figure A11). However, knockdown of these proteins does abrogate IL-6 and IL-8 responses to spike receptor binding domain construct (Charlie Hitchman and Dr Omeed Darweesh, Gooptu group, University of Leicester). Together these findings support a broadly conserved galectin-3 and CD98 dependent mechanism mediating cytokine responses to spike protein receptor binding domain across quite different cell types. This appears distinct from ACE-2 mediated mechanisms of viral infection.

Currently within the Group, my colleagues are working with Galecto Biotech, Denmark. They are testing small molecule galectin-3 inhibitors to assess the effect upon receptor binding domain induced pro-inflammatory cytokine release. Initial data demonstrate abrogation of pro-inflammatory cytokine release in response to receptor binding domain stimulation like siRNA mediated galectin-3 silencing. Further biological repeats are still needed to conclude these set of experiments.

The successful generation of all spike protein-constructs, and the repeat of functional assays of cytokine assessment in both A549 and THP-1 cells with these proteins introduce exciting future experiments to assess how effects of the receptor binding domain interactions may be modulated by the presence of the galectin-like domain and other spike regions.

Chapter 7

General Discussion

7.1 General Discussion

The pathogenesis of IPF remains to be fully elucidated. IPF is responsible for ~1% of total death within the UK, with a median survival of ~3 years from point of diagnosis (Zoz *et al*. 2010, Kanaji *et al*. 2016, British Lung Foundation. 2018, Wolters *et al.* 2018). Current FDA and NICE approved treatments reduce the progression of scarring in IPF, but do not reverse or permanently stabilise scarring. This signifies the importance as to why we need to further understand and elucidate the molecular pathway(s) of this disease to improve anti-fibrotic therapies.

Whilst no routine clinical biomarkers exist in the diagnosis of IPF, several have been proposed to be increased in IPF patients. One such biomarker is the protein galectin-3. Galectin-3 may mediate fibrosis in the lung through its extended lattice-forming capabilities that can facilitate the clustering of pro-fibrotic factors. My project explores the hypothesis that galectin-3 nucleates an ordered assembly that stabilises and spatiotemporally clusters a pro-fibrotic macromolecular assembly at the surface of alveolar epithelial and myofibroblast cells constituting of at least four components: galectin-3, CD98hc (heavy chain), β1-integrin and TGF-βRII (Estrach *et al*. 2014, Hashiba *et al*. 2014). This latticed assembly could potentially facilitate integrin-dependent TGF-β1 mediated signalling to boost chronic disease responses cross-talking with acute inflammatory signalling responses, and we termed this as the gal-3-fibrosome. Defining this may identify novel therapeutic strategies.

Galectin-3, CD98 and β1-integrin proteins were all expressed at increased levels in IPF lung tissue relative to NFC lung tissue (Chapter 3). These proteins co-localised in the same cellular patterns, predominantly in parenchymal epithelial cells as evidenced by E-cadherin staining. Less intense staining was seen in the parenchymal background, which may suggest their distribution in mesenchymal cells also. It is important to note that alveolar/airway resident immune cells (e.g.

airway and alveolar macrophages) may not be represented due to the nature of sample acquisition and treatment.

In A549 cells, these individual increased protein responses were recapitulated by pro-fibrotic (TGF-β1) stimulation. Flow-cytometry confirmed that detectable CD98 protein is restricted to the cell surface membrane. Co-localisation of galectin-3 with CD98 and β1-integrin was strongly supported by both co-immunoprecipitation and co-immunofluorescence analyses in response to the pro-fibrotic cytokine TGF-β1 stimulation.

CD98 therefore may play a role in both acute inflammatory responses (Wang et al. *Manuscript in* preparation) and due its close interaction with galectin-3, in pro-fibrotic, exaggerated wound healing responses. Wound healing in A549 scratch test indicates that CD98 is required for this response. This finding supports the hypothesis that galectin-3:CD98 interactions may couple acute and longer term fibrotic consequences of exacerbation, and that this interaction may have potential as a therapeutic target. TD139 (now known as GB0139) in phase 3 clinical studies may function like this. There is scope for other therapeutics primarily targeting CD98 or the specific interaction with galectin-3 to be developed.

To assess how far these findings were conserved in lung parenchymal myofibroblasts, *ex vivo* derived NFC and IPF myofibroblasts were studied in Chapter 4. In both NFC and IPF myofibroblasts the increased protein level expressions of CD98 and β1-integrin were similar following TGF-β1 stimulation to what had been observed in A549 cells. However, the similarities were less clear cut for galectin-3 protein expression. In fact, levels of galectin-3 in these cells decreased both in the membrane and cytosol fractions. This decreased protein level of galectin-3 was supported by the decreased mRNA gene expression levels of galectin-3 following TGF-β1 stimulation from both NFC and IPF derived myofibroblasts.

Co-localisation of galectin-3 with both CD98 and β1-integrin in NFC-derived myofibroblasts appeared to decrease following treatment with TGF-β1, consistent with a reduction in galectin-3 production by the myofibroblasts. This may represent an inherent negative feedback loop in these 'healthy' cells, that normally limits gal-3-fibrosome formation. In contrast, IPF-derived myofibroblasts displayed evidence of a relative preservation of such co-localisation in response to TGF-β1.

TGF-β1 stimulated NFC lung tissue (explant model, Roach *et al*. 2018) demonstrated increased mRNA expression of CD98:β1-integrin, and decreased mRNA expression of galectin-3, relative to the non-stimulated NFC lung tissue. However, proximity ligation assay in these tissues demonstrated increased co-localisation of galectin-3 with β1-integrin and CD98 (following TGF-β1 stimulation) within a distance of ≤ 40 nm. Findings of these individual proteins (and gene) expression levels and their co-localisation in different cell types and lung tissue are summarised in table 7.1.

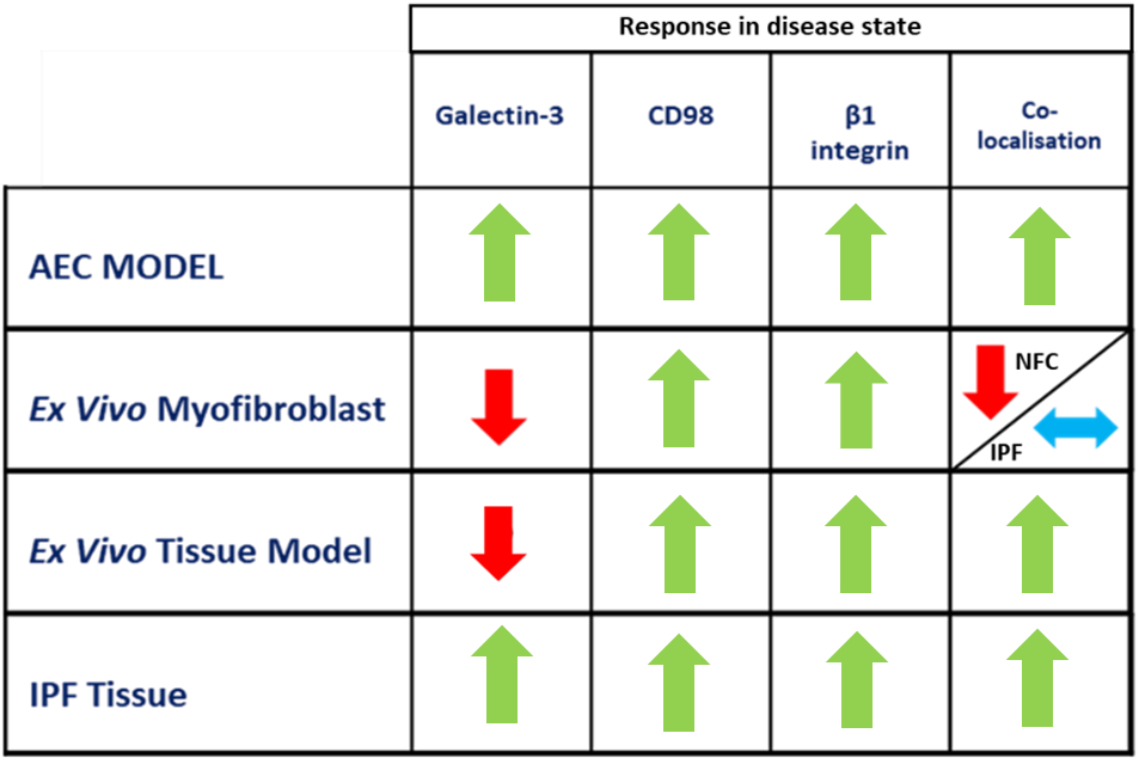


Table 7.1, Summary of individual gal-3-fibrosome protein/ and mRNA gene expression responses in different pro-fibrotic experimental contexts (Chapters 3 and 4). **In the alveolar epithelial cell (AEC) model- A549 cells**, stimulation with the pro-fibrotic cytokine TGF-β1 increased protein levels of galectin-3, CD98 and β1-integrin, and its subsequent co-localisation. **In both NFC- and IPF- *ex vivo* derived myofibroblast**, stimulation with the pro-fibrotic cytokine TGF-β1 decreased protein levels of galectin-3, but increased levels of CD98 and β1-intgerin similar to A549 cells. In agreement with the reduction of galectin-3, co-localisation went down in NFC, but remained unchanged in IPF derived myofibroblasts. **In an *ex vivo* pre-treated TGF-β1 tissue model,** mRNA levels of CD98:β1-integrin went up whilst galectin-3 went down. Co-localisation assessments on these tissue by PLA displayed overall increased co-localisation between the proteins. **Immunohistochemistry assessment of IPF vs NFC lung** tissue displayed increased levels of all three proteins in IPF tissue compared to NFC tissue. All proteins co-localised around E-cadherin. Overall, this supports the gal-3-fibrosome hypothesis. The 'Green vertical up' arrow represents increased levels. The 'red vertical down' arrow represents decreased levels. The 'blue horizontal' arrow represents stable levels.

Given the situation in lung tissue where epithelial cells and myofibroblasts can cross-talk, increased galectin-3 secreted from alveolar epithelial cells might compensate for the lack of upregulation in myofibroblasts (Figure 7.1), therefore explaining the overall increased levels in the co-localisation of these proteins on the tissue. Future work could explore this in co-culture studies. However, my findings suggest that in established IPF, myofibroblasts may autonomously overcome this limitation on TGF-β1-induced pro-fibrotic behaviour. The studies I have undertaken do not clearly allow speculation about the underlying mechanism.

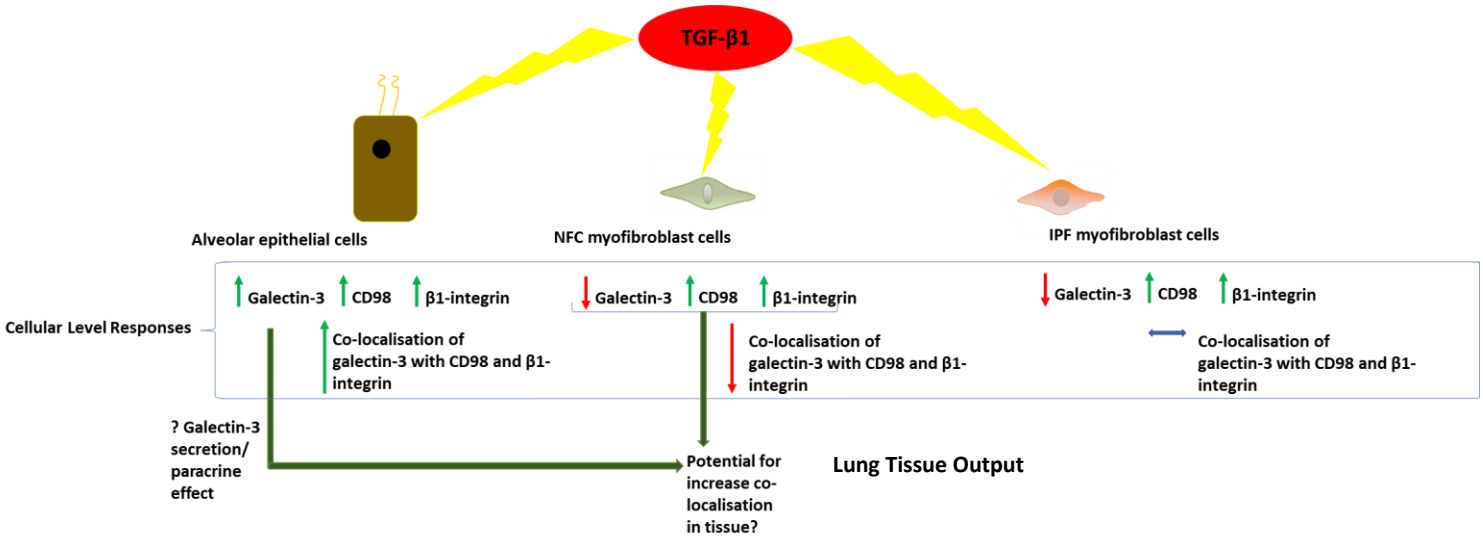


Figure 7.1, Schematic displaying the protein responses in TGF-β1 pro-fibrotic conditions on the cellular level of different cell types (inferences taken from summary Table 7.1) which can cross-talk and possibly modulate overall tissue responses. For the cells assessed in my thesis (Chapters 3 and 4), the responses of individual proteins have been shown labelled for alveolar epithelial (A549) cells, and both NFC- and IPF- derived myofibroblasts. In a healthy individual, the alveolar epithelial cells and NFC myofibroblast are present in the lung (at the very least). In an IPF patient, all three cell types will be present in the lung. In lung tissue, epithelial cells and myofibroblasts could potentially cross-talk/ and exert a paracrine effect: increased galectin-3 secreted from alveolar epithelial cells might offset for the lack of upregulation of galectin-3 in myofibroblasts. The 'Green vertical up' arrow represents increased levels. The 'red vertical down' arrow represents decreased levels. The 'blue horizontal' arrow represents stable levels.

Taken together these findings support the existence of physical, and important functional, interactions between putative gal-3-fibrosome constituents. However, they do not themselves prove that direct physical interactions are important in the cell models of disease studied, or define high resolution information on any such interactions to use in trying to disrupt them. I therefore worked on several approaches to define this (Chapter 5). Although I was unable to progress any of the strategies to the point of obtaining high resolution structural data from a sample, the work does clarify important technical points and allows planning of future work building upon this. Unfortunately, from the moment the COVID pandemic arrived in the UK, I was unable to spend sustained periods on such work but instead surveyed several approaches to see if any led straightforwardly to specific structural studies or conclusions.

I first studied recombinant galectin-3 oligomerisation in conditions reported to induce concentration-dependent or ligand-induced oligomerisation. In doing this I worked to optimise non-denaturing (native)-PAGE of monomeric galectin-3. Ultimately best results were achieved when adjusting buffer pH to take account of the galectin-3 pI. But even with these modifications galectin-3 migrated slowly as a diffuse band, perhaps in part reflecting its high degree of conformational flexibility. The failure of multiple approaches to capture anything other than monomeric behaviour indicates both concentration- and ligand-induced oligomerisation is highly reversible and dependent on concentrations of galectin-3 and any ligand.

Galectin-3 monomers are constrained by contacts with the C-terminal domain when in an apo-state. Binding to a carbohydrate ligand releases the N-terminus (Halimi *et al*. 2014, Lin *et al*. 2017, Chiu *et al*. 2020). This is essential for galectin-3 oligomerisation, which can occur either via homotypic interactions between N-N (N-type oligomerisation) or C-C (C-type oligomerisation) terminal domains (Nieminen *et al*. 2007, Lepur *et al*. 2012). More recently, the oligomeric behaviour has been described as “fuzzy” as the multimerisation behaviour could be due to additional homotypic N-C terminal domain interactions (Lin *et al*. 2017).

These indicate the galectin-3 oligomerisation may result in a heterogeneous range of N-terminal domain interactions that can interconvert dynamically. This is consistent with our findings and might make structural studies focusing on the naturally occurring range of galectin-3 oligomers alone more challenging. Heterogeneity and dynamic conformational change limit acquisition of high-resolution structural data. NMR approaches can work well to characterise the dynamics but there is less experience in translating such data to develop therapeutic strategies.

Cross-linking studies confirmed that oligomerisation was occurring but generated enormous covalently-stabilised complexes that were very electron-dense and unsuitable for cryo-EM. Slightly unexpectedly, use of PNGase produced smaller species. Whilst this displayed common repetitive motifs, species remained largely heterogenous and substantial further optimisation is still required to generate samples tractable by (cryo-)EM studies for structure determination.

My work focusing upon the production of samples in which gal-3-fibrosome components are involved in homotypic or heterotypic protein:protein interactions for targeting has been inconclusive in terms of defining its structural basis. However, it has progressed our understanding of how to undertake such studies in ongoing and future work.

SMALP approaches to isolate endogenously-regulated putative gal-3-fibrosome proteins in complex from cell membranes have shown some initial promise for CD98. However, sample preparations were too dilute for specific cryo-EM analysis and require validation. Further work-up is currently underway by a PhD student within the group (Charlie Hitchman). He is looking into overexpression systems and other approaches.

The challenges of progressing my envisaged PhD project due to COVID-19 were offset to some degree by opportunities we engaged with as a group to research how the gal-3-fibrosome might

be involved in the pathogenesis of COVID pneumonitis (Chapter 6). In addition to noting the presence of a galectin-3 like domain adjacent to the receptor binding domain of SARS-CoV-2, we hypothesised that the receptor binding domain itself might engage with the gal-3-fibrosome via CD98:β1-integrin interactions due to the presence of an RGD motif (well recognised by β1-integrin) in the receptor binding domain.

My focus was on producing recombinant spike protein constructs to be used in ELISA studies with Dr Panayiota Stylianou and Dr Omeed Darweesh (PDRA, Gooptu group, University of Leicester). Studies in A549 cells directly extended the lung epithelial studies done previously for CD98-dependent responses to injury triggers (Wang *et al*, *Manuscript in preparation*). In addition, we looked at a macrophage cell model (PMA treated THP-1 cells). We found that the receptor binding domain stimulated IL-6, IL-8 and TNF-α responses (similar magnitude to experimental LPS conditions) in A549 and THP-1 cells. IL-6 and IL-8 levels were abrogated by knockdown of either CD98 or galectin-3, to a similar extent to treatment with dexamethasone (which additionally abrogated levels of TNF-α)– a key evidence-based treatment for COVID-pneumonitis.

A549 cells do not detectably express ACE2 and do not get infected by SARS-CoV-2, raising the possibility that the pro-inflammatory cytokine response is distinct from the cell infection mechanism. Support for this comes from the absence of any effect of galectin-3 or CD98 knockdown upon COVID infectivity in Calu-3 cells that do express ACE2 and are a good model of infection.

These findings were consistent with previous work in A549 and TT1 alveolar epithelial cell model (Wang *et al*, *Manuscriot in preparation*) with LPS, mechanical stretch, and matrix stiffness for which the mediating intracellular mechanism (involving a positive feedback loop of CD98, TRPV4, $Ca2^{+}$ influx, and NFκB activation) has been dissected (Figure 7.2). Ongoing work by Dr Omeed

Darweesh, in partnership with Galecto Biotech (the developers of GB0139), is now exploring the potential for compounds related to GB0139 to inhibit these cytokine responses.

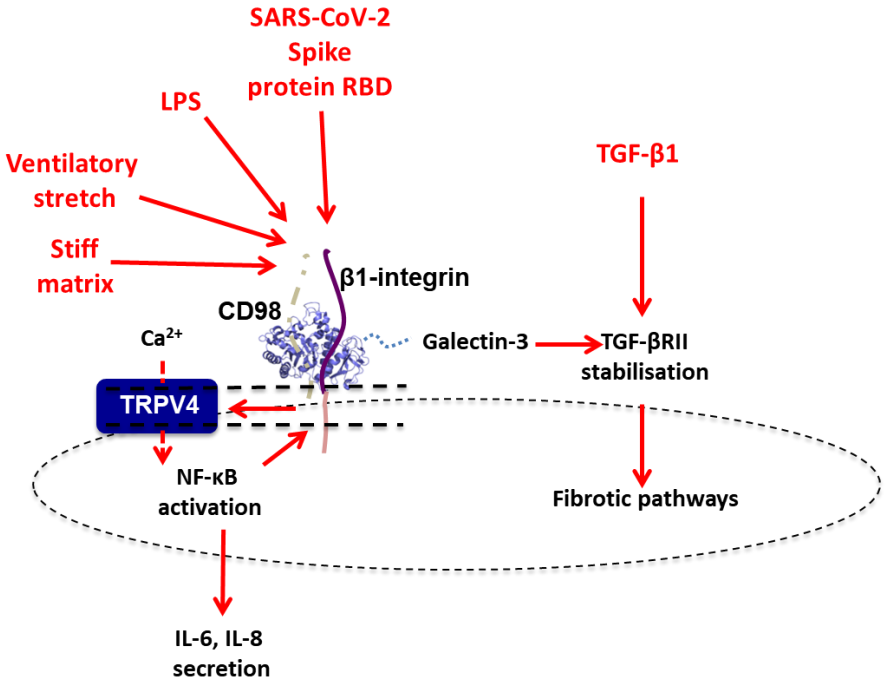


Figure 7.2, Schematic displaying acute inputs and its associated pro-inflammatory outputs, along with the incorporation of the gal-3-fibrosome hypothesis. A representation of pro-inflammatory response inputs which comprises multiple stimuli including LPS, ventilatory stretch and matrix stiffness (Wang *et al*, *Manuscript in preparation*). This is all mediated by an intracellular positive-feedback mechanism which includes CD98, TRPV4 -(and Ca^{2+} influx), and NF-κB activation. CD98 is bound to our gal-3-fibrosome hypothesis, displaying the likelihood of cross-talk between pro-fibrotic and pro-inflammatory pathways. Now in addition, SARS-CoV-2 spike protein RBD interactions has been integrated into this schematic as an acute input mediating pro-inflammatory responses. The common pro-inflammatory cytokines which were elevated between all these acute inputs include IL-6 and IL-8.

Immunoprecipitation of receptor binding domain in A549 cells stimulated by receptor binding domain treatment co-immunoprecipitated both CD98 and galectin-3 (Appendix, Figure A12), supporting a close physical interaction between these proteins. Further studies to test the hypothesis that inflammatory and infective interactions at the cell surface are different are planned. These involve co-immunoprecipitation studies in receptor binding domain-stimulated Calu-3 cells followed by immunoprecipitation with anti-CD98 mAb and probing for the receptor binding domain and ACE2, and reciprocally immunoprecipitating with anti-ACE2 and probing for both CD98 and the receptor binding domain. We would predict that the receptor binding domain, but no ACE2, co-immunoprecipitates with CD98 and vice versa.

Taken together, my work has advanced our understanding of the gal-3-fibrosome in multiple cell type and tissue contexts. It is broadly consistent with the starting model of galectin-3 interacting closely with CD98:integrin complex with potential for cross-talk between acute inflammatory and chronic fibrotic responses, that appear to be key to both IPF progression and COVID-pneumonitis.

7.2 Future Work

Based on the findings from my work, I propose subsequent follow-on experimental strategies:

- Since antibodies were unable to detect endogenous TGF-βRII, a plasmid containing an exogenous TGF-βRII cDNA construct should be transiently transfected into the cell models used in Chapters 3 and 4. This would allow assessment of the co-localisation of TGF-βRII with galectin-3 and CD98hc:integrin complex (in both basal and TGF-β1 stimulated conditions).

- Supplementary co-immunoprecipitation experiments conducted on both NFC and IPF-derived myofibroblasts to confirm co-immunofluorescence findings of reduced (NFC) and stabilised (IPF) co-localisation of galectin-3 with CD98:integrin following TGF-β1 stimulation.

- *In vitro* co-culture studies of alveolar epithelial (A549) cells with NFC derived-myofibroblast cells to address potential paracrine effect of the epithelial cells supplementing the myofibroblasts in the context of TGF-β1 stimulation.

- Repeating studies of putative gal-3-fibrosome protein levels under basal conditions in low passage IPF myofibroblasts compared to NFC to see if the preliminary observations of elevated levels and co-localisation are confirmed.

- Further sample optimisation with a view to cryo-EM data collection for structural analysis of recombinant cross-linked galectin-3:ASF sample treated with PNGase. The use of sucrose density gradient ultracentrifugation as a means to separate different size species may help address this.

- Endogenous gal-3-fibrosome isolation using CD98 immunoprecipitated SMALP lysates require further optimisation for cryo-EM dataset collection, since negative stain-EM displayed homogenous sample preparation. The preparation of co-immunoprecipitation samples and mass-spectrometry readouts needs optimisation as a pathway to identify additional gal-3-fibrosome binding partners (including initial capture target protein).

- Cross-linking samples studied by mass spectrometry to see if this boosts signal for analysis.

- Immunoprecipitate SMALP lysates using anti-galectin-3 mAb. In the initial instance, I would propose repeating the same experiments as done for CD98 immunoprecipitated

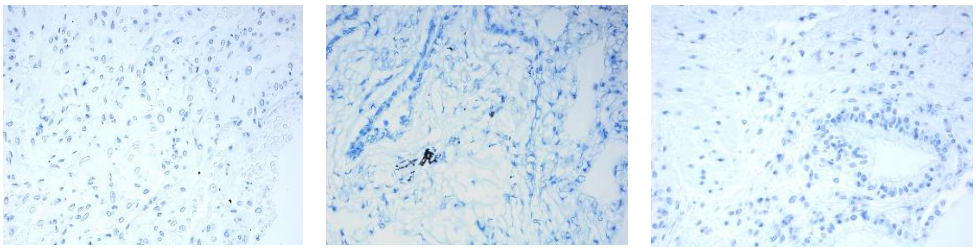
SMALP lysates. If the results were successful these samples could be assessed by cryo-EM screening.

- Generate all remaining SARS-CoV-2 constructs including the full length, S1 region and N-terminal domain into pCDNA3.1 vector as was achieved for the receptor binding domain construct. It may be beneficial to insert dual-tags, i.e. incorporating FLAG and 6x HIS- tag. This will allow to compare the yield between the two pulldown strategies. After-successful cloning, I would express the proteins and repeat similar functional assays as done for the receptor binding domain in the three different cell models (A549, THP-1 and Calu-3 cells) to observe the levels of the pro-inflammatory cytokine outputs.

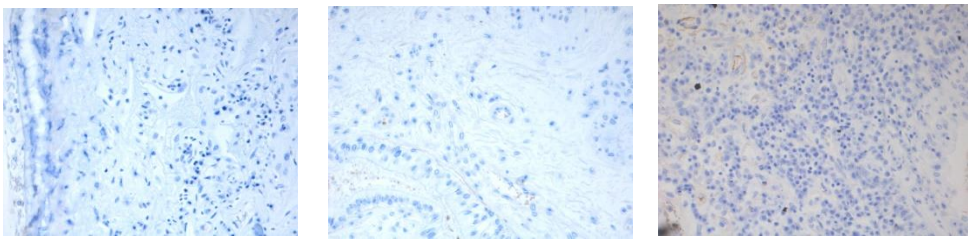
- Other members of the Gooptu group are now collaborating with Galecto Biotech, Denmark, testing novel small molecule galectin-3 inhibitors, to assess the effect of receptor binding domain induced cytokine responses.

- In addition, whether inflammatory and infective interactions are different are required to be assessed. This could be addressed by co-immunoprecipitation studies; the receptor binding domain stimulated Calu-3 cells followed by anti-CD98 mAb immunoprecipitation could be probed for the receptor binding domain and ACE2. Reciprocally immunoprecipitating with anti-ACE2 mAb, followed by probing for CD98 and receptor binding domain could be conducted. If interactions differ, we predict that the receptor binding domain, but not ACE2 will co-immunoprecipitate with CD98 (and vice versa).

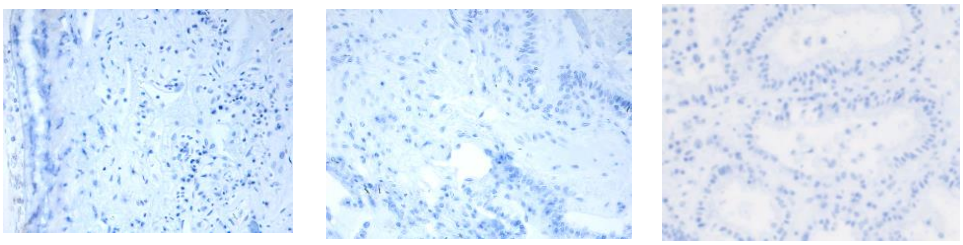
Appendix



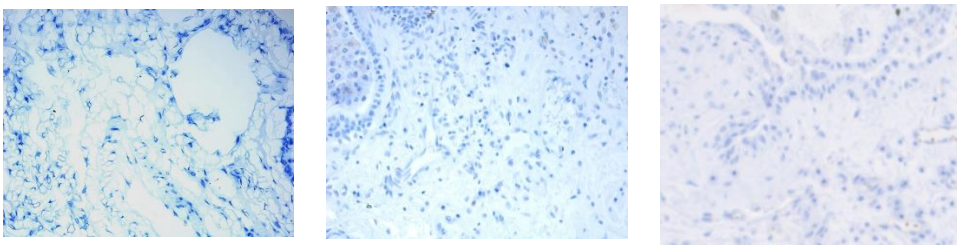
a) Negative control: Secondary antibody only, anti-rat-HRP 1 µg/ml (used to stain for galectin-3)



b) Negative control: Rat IgG2a Isotype Control for galectin-3



c) Negative control: Secondary antibody only, anti-mouse-HRP 50 µl (Dako EnVision Kit) (used to stain for CD98, β1-integrin and E-cadherin)



d) Negative control: Mouse IgG1 Isotype control for CD98, β1-integrin and E-cadherin

Figure A1, Negative control staining of NFC lung tissue sections to confirm the specificity of immunohistochemistry staining focusing on the levels and co-localisation of galectin-3, CD98 and β1-integrin between NFC and IPF derived lung tissue serial sections. Tissue sections were singularly stained as follows: **a)** 1 µg/ml of 'anti-rat-HRP secondary antibody only' was assessed as this was the concentration used to stain and bind to the anti-rat monoclonal galectin-3 antibody. **b)** The galectin-3 isotype control (rat IgG2a) was assessed at a concentration of 4 µg/ml (matched galectin-3 concentration) with 1 µg/ml of anti-rat-HRP secondary antibody. **c)** 50 µl of 'secondary anti-mouse-HRP antibody (Dako EnVision kit) only' was assessed as this was used in the amount stated to stain and bind to the anti-mouse monoclonal antibodies of CD98, β1-integrin and E-cadherin. **d)** The isotype control of CD98, β1-integrin and E-cadherin (mouse IgG1), was assessed at a concentration of 5 µg/ml with 50 µl of secondary anti-mouse-HRP (Dako EnVision kit). The nuclei were stained by immersing the slides into hematoxylin (blue stain).

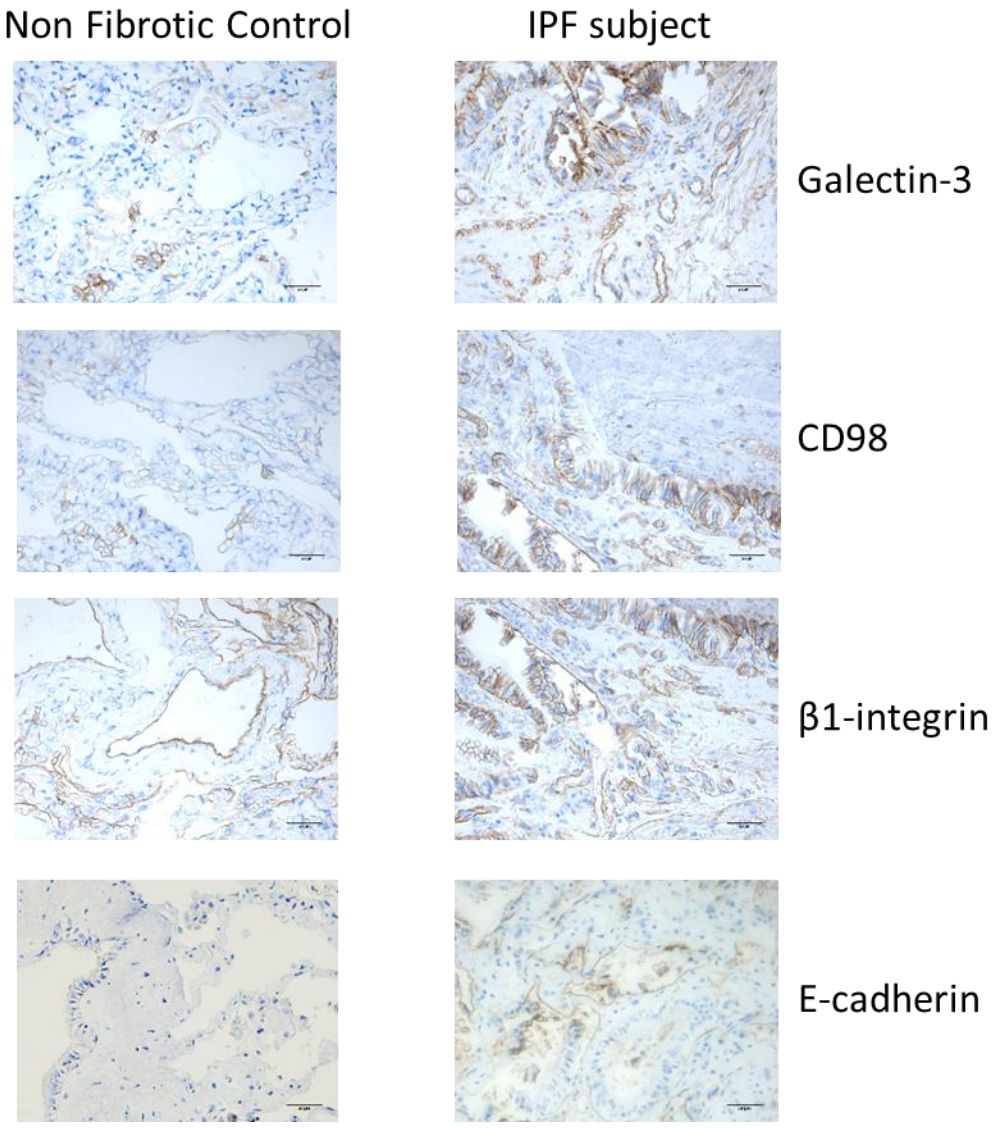


Figure A2, Zoomed projections of the stained panels of figure 3.1, displaying Gal-3-fibrosome proteins localising on the epithelium with increasing amounts in IPF subjects. Immunohistochemistry staining focusing on the levels and co-localisation of galectin-3, CD98hc and β1-integrin between NFC and IPF derived lung tissue serial sections. NFC (on the left hand column) and IPF (on the right hand column) lung tissue samples were serially sectioned and stained. Tissue sections were singularly stained to probe for the gal-3-fibrosome proteins with monoclonal primary antibodies; 4 µg/ml galectin-3, 5 µg/ml CD98, 2.5 µg/ml β1-integrin. 5 µg/ml of E-cadherin monoclonal antibody was applied to highlight epithelial cells. 1 µg/ml of anti-rat-HRP secondary antibody was used to probe for galectin-3 (brown-stain) and 50 µl of secondary anti-mouse-HRP antibody (Dako EnVision kit) was used to stain for CD98, β1-integrin and E-cadherin (brown-stain). The nuclei were stained by immersing the slides into hematoxylin (blue stain). Images were taken in a x20 zoom.

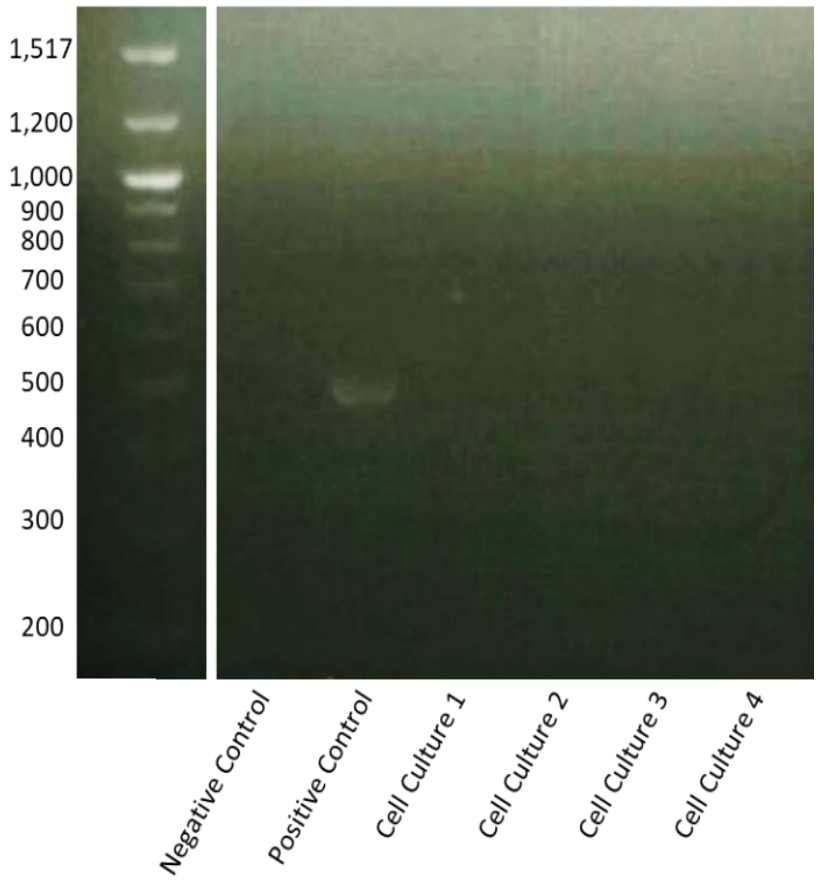


Figure A3, Assessment of Mycoplasma contamination in A549 cells using a PCR based approach. Supernatant from exponentially growing A549 cells was used at differing passages. PCR was run as per manufactures instructions (from the kit). The amplified products were run on an 1% (w/v) 1xTBE agarose gel electrophoresis. The presence of PCR products at approximately 450-550bp indicates the presence of mycoplasma contamination in the cell culture. The Y-axis represents a 100bp DNA ladder ranging from 200-1517bp. Lanes 1-6 have been labelled with the appropriate conditions loaded on the gel (on the X-axis), with cell culture 1, 2, 3 and 4 representing the A549 passage number in an increasing order (passages 8, 12, 16 and 20). The negative control consisted of a non-template-control to confirm the lack of contamination in the reaction. The mycoplasma positive control ensured that no PCR inhibitors interfered with the reaction (band seen at ~500bp marker). The bands were visualised due to ethidium bromide on a UV transilluminator and the gel image was captured.

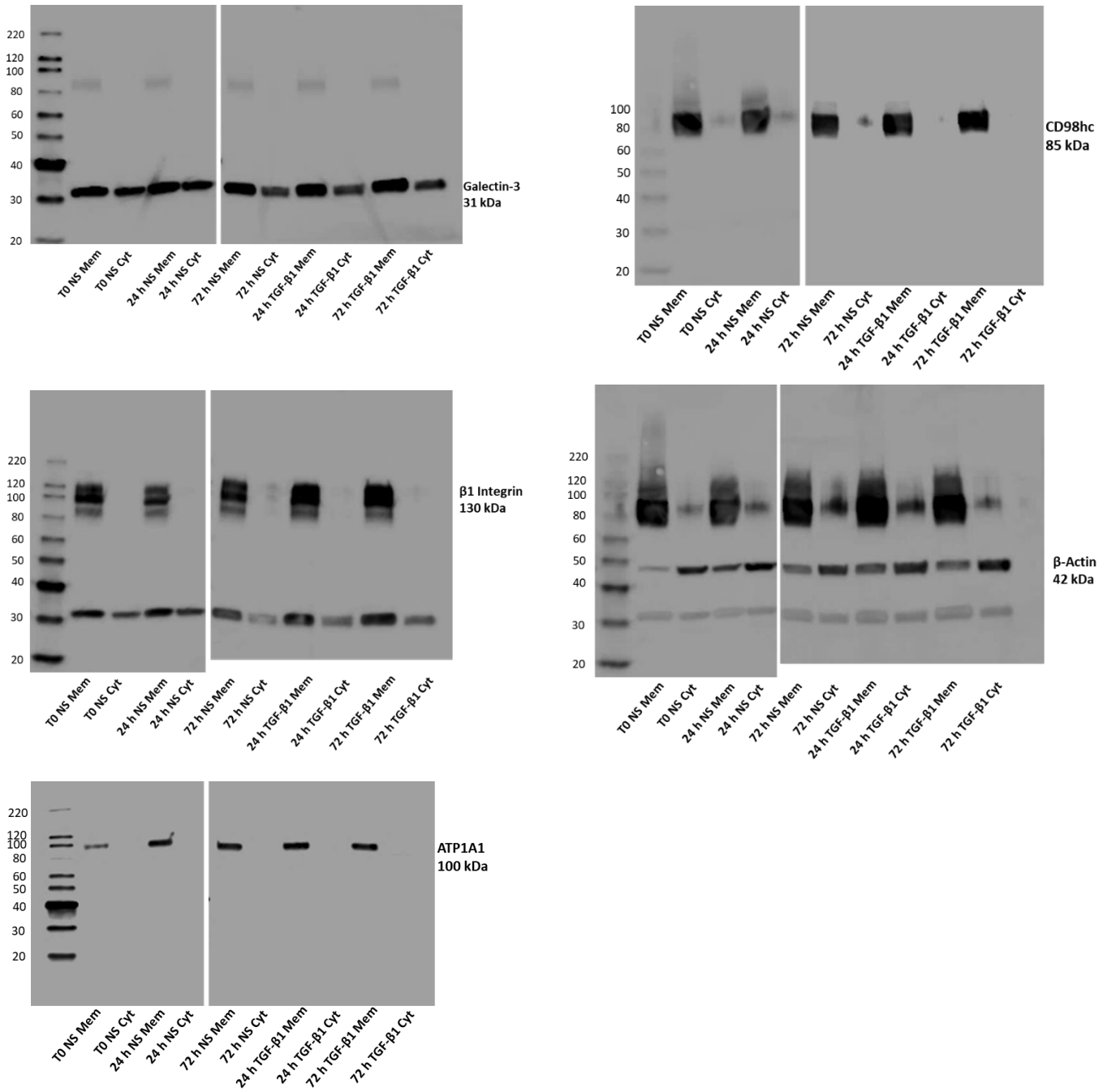
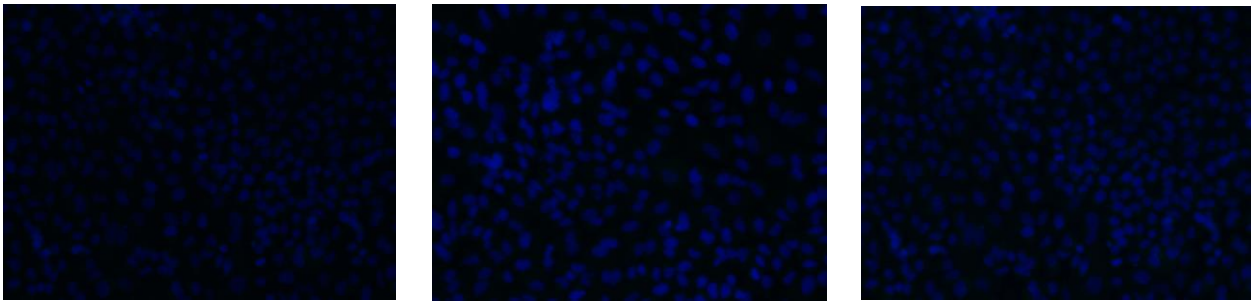
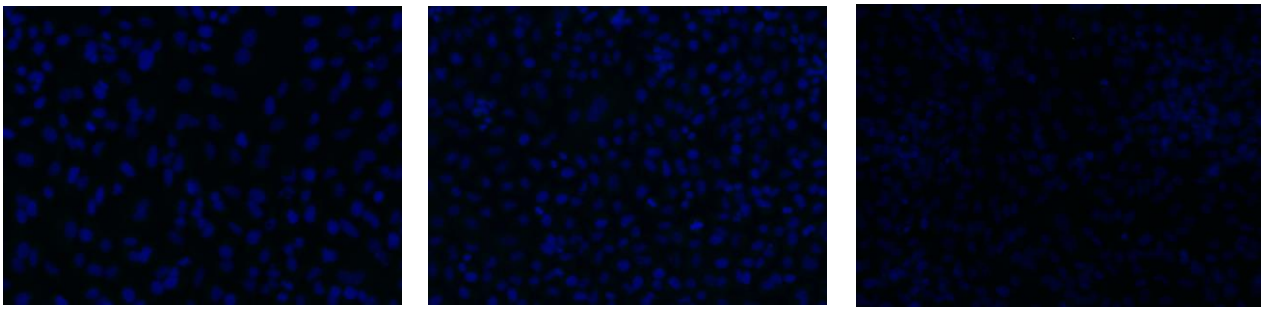


Figure A4, Full western-blot membrane (uncropped) of Figure 3.3- monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in A549 cells by western-blotting. Exponentially growing A549 cells was used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3 (31 kDa), 0.125 µg/ml CD98hc (85 kDa), 0.005 µg/ml β1-integrin (125 kDa), 0.04 µg/ml β-actin (42 kDa) and 1:5000 Na^+/K^+ ATPase (ATP1A1) (100 kDa). Lane 1-10 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective marker being represented (MagicMark™ XP Western Protein Standard). The blots were immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3).



a) Negative control: Mouse IgG1a Isotype Control for CD98 and β1-integrin



b) Negative control: Rat IgG2a Isotype Control for galectin-3

Figure A5, Isotype control staining to confirm the specificity of the co-immunofluorescence staining of galectin-3, CD98hc and β1-integrin in A549 cells. Isotype controls were chosen to be assessed in 72 h TGF-β1 stimulated cells since highest protein expression for the gal-3-fibrosome proteins in A549 cells were statistically highest at this stimulated condition, and hence this will increase the chances of observing any non-specific staining. **a),** Cells were singularly stained with 5 µg/ml of mouse-IgG1 and then stained with goat-anti-mouse Alexa-fluor 594 (red) (as done to stain for CD98) or Alexa fluor 488 (green) (as done to stain for β1-integrin). **b)**, Cells were singularly stained with 4 µg/ml of rat-IgG2a and then stained with anti-rat Alexa fluor 488 (green) (as done to stain for galectin-3). The nuclei were stained with DAPI.

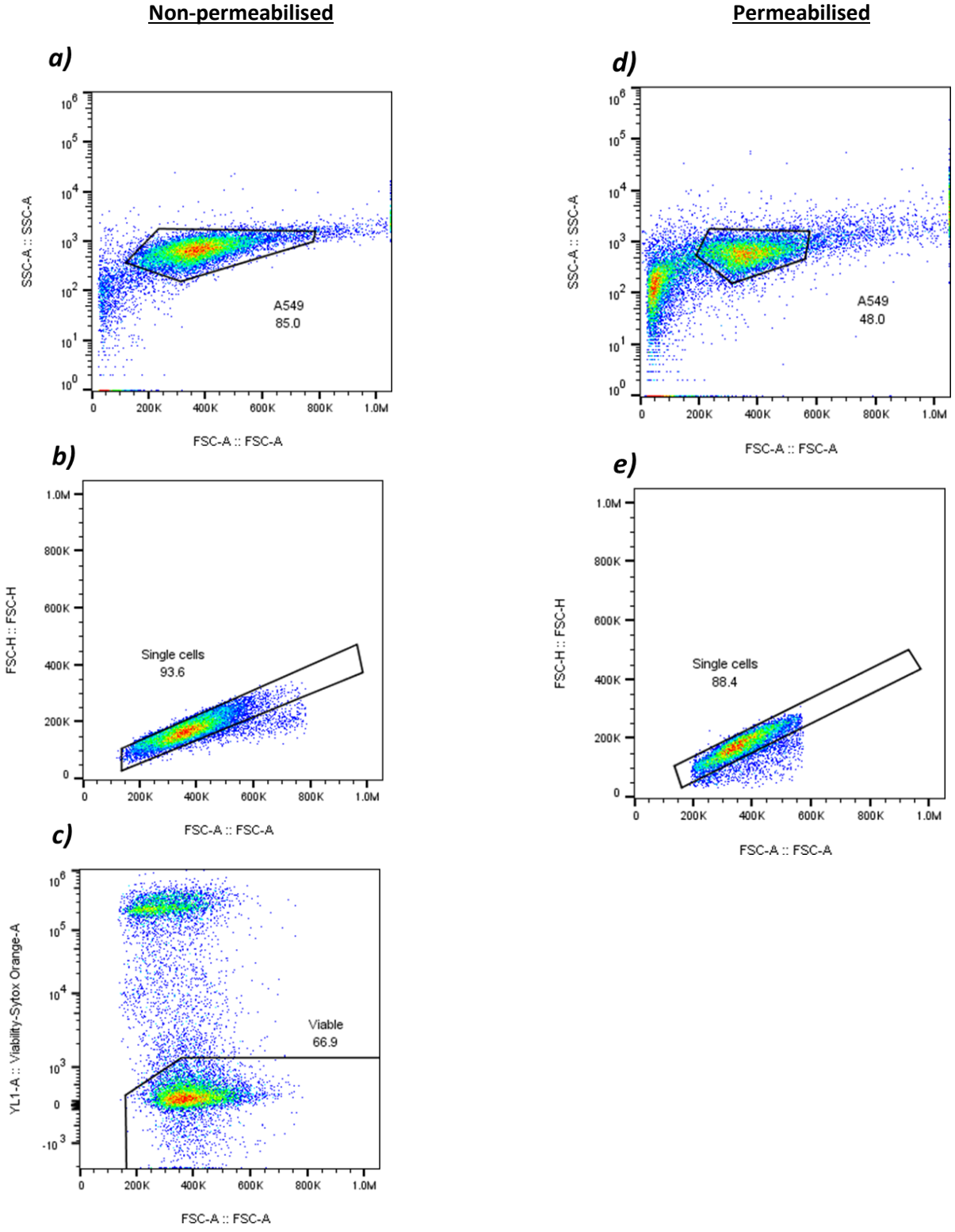


Figure A6, Cellular gating strategy for flow cytometry analysis of CD98hc staining signal in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) A549 cells. Figure A5 a-c (non-permeabilised cells) and A5 d-f (permeabilised cells): A stringent cell sorting strategy was devised to detect pure A549 cells. ***a*** and ***d)*** Cell size (forward scattering area, FSC-A) and granularity (side scatter area, SSC-A) were used to identify the main cell fraction (excluding debris). ***b*** and ***e)*** Cells were gated to exclude cell doublets as depicted (- H = signal heights). ***c)*** A further gating strategy (for extracellular non-permeabilised analysis) was used to exclude dead cells and therefore only examining viable cells. This was achieved by subjecting A549 to forced cell death (further elaborated in materials and method).

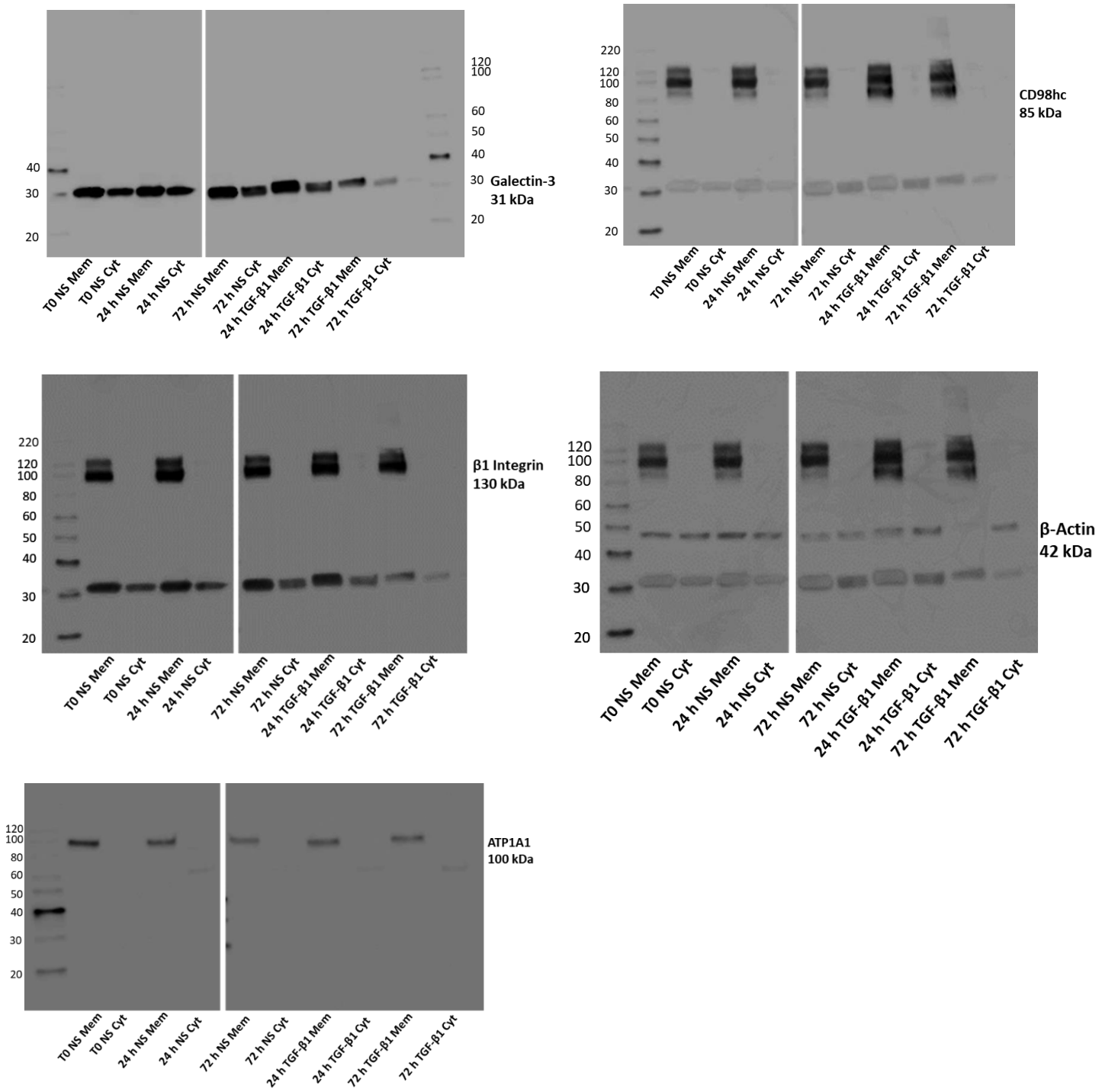


Figure A7, Full western-blot membrane (uncropped) of Figure 4.3-monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in NFC myofibroblast cells by western-blotting. *Ex-vivo* HLMF cells were used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3 (31 kDa), 0.005 µg/ml β1-integrin (125 kDa), 0.125 µg/ml CD98hc (85 kDa), 0.004 µg/ml β-actin (42 kDa) and 0.5 µg/ml Na^+/K^+-ATPase (ATP1A1) (100 kDa). Lane 1-10 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective marker being represented (MagicMark™ XP Western Protein Standard). The blot was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3.)

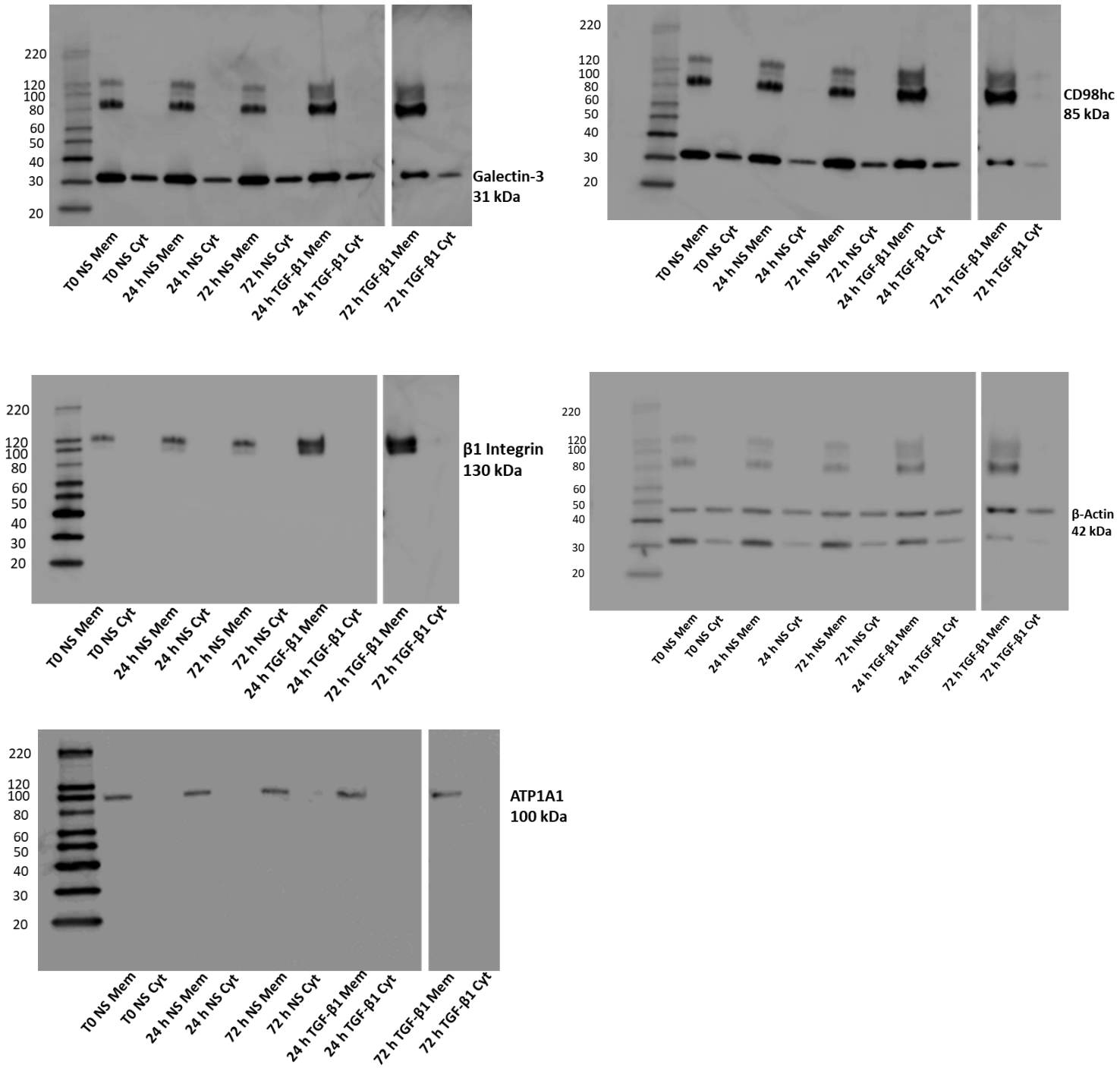


Figure A8, Full western-blot membrane (uncropped) of Figure 4.5-monitoring expression levels of proteins associated with the galectin-3-fibrosome following stimulation with the pro-fibrotic cytokine TGF-β1 in IPF myofibroblast cells by western-blotting. *Ex-vivo* HLMF cells were used. Cell lysates of the cytosol and membrane fractionation were prepared using the Mem-PER kit. 20 µg of proteins were loaded with 4x Laemmli sample buffer on a 4-15% resolving SDS-gel. The blot was incubated with 5% milk (w/v) in 1x TBS-Tween-20 and immunoblotted with 0.125 µg/ml galectin-3 (31 kDa), 0.005 µg/ml β1-integrin (125 kDa), 0.125 µg/ml CD98hc (85 kDa), 0.004 µg/ml β-actin (42 kDa) and 0.5 µg/ml Na^+/K^+-ATPase (ATP1A1) (100 kDa). Lane 1-10 has been labelled with the appropriate conditions (on the X-axis), with membrane fraction being abbreviated as Mem, cytosolic fraction abbreviated as Cyt, and non-stimulated conditions abbreviated as NS. The Y-axis has been labelled with the respective marker being represented (MagicMark™ XP Western Protein Standard). The blot was immersed with EZ-ECL chemiluminescent detection kit and the images were taken at various exposure times ranging from 1-10 minutes. (n=3).

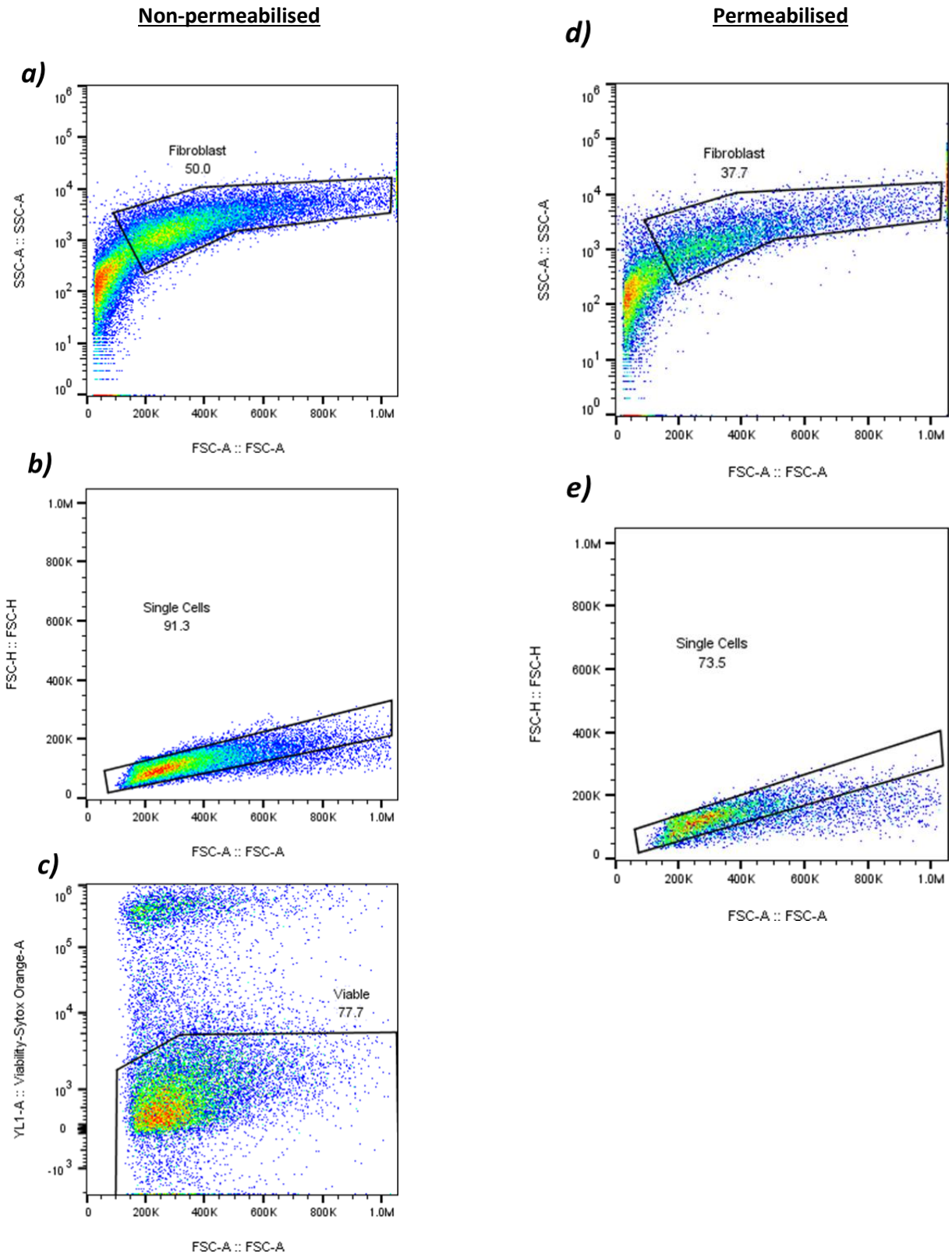


Figure A9, Cellular gating strategy for flow cytometry analysis of CD98hc staining signal in permeabilised (cell membrane ruptured) vs non-permeabilised (cell membrane intact) NFC myofibroblast cells. Figure A8 a-c (non-permeabilised cells) and A5 d-f (permeabilised cells): A stringent cell sorting strategy was devised to detect pure NFC myofibroblast cells. ***a*** and ***d)*** Cell size (forward scattering area, FSC-A) and granularity (side scatter area, SSC-A) were used to identify the main cell fraction (excluding debris). ***b*** and ***e)*** Cells were gated to exclude cell doublets as depicted (- H = signal heights). ***c)*** A further gating strategy (for extracellular non-permeabilised analysis) was used to exclude dead cells and therefore only examining viable cells. This was achieved by subjecting NFC myofibroblasts to forced cell death (further elaborated in materials and method).

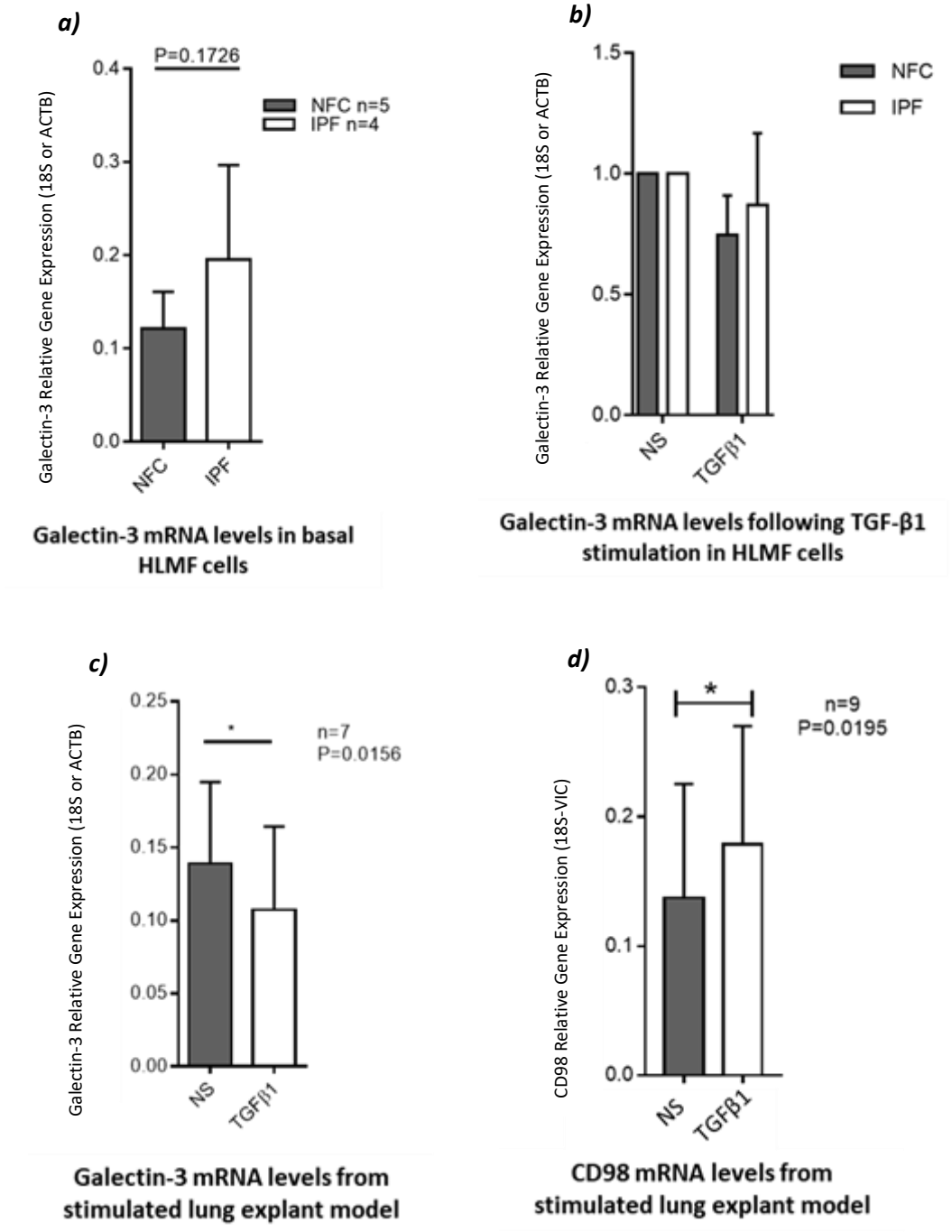


Figure A10, Gene expression levels of galectin-3 and CD98 from HLMF and lung tissue was assessed in basal and TGF-β1 stimulated conditions using qRT-PCR. Gene expression levels of galectin-3 was assessed in NFC and IPF HLMF cells both basally and in response to 24hr TGF-β1 stimulation (***a*** and ***b***). Gene expression levels of galectin-3 and CD98 following 7-day TGF-β1 treatment was assessed from lung fibrogenesis explant model/tissue (***c*** and ***d***). The relative mRNA expression was calculated using the 2^-(ΔCt) method, where delta CT is the difference between the CT value for the galectin-3 gene and its house-keeping gene (18S ACTB) or CD98 gene and its house-keeping gene (18S-VIC).

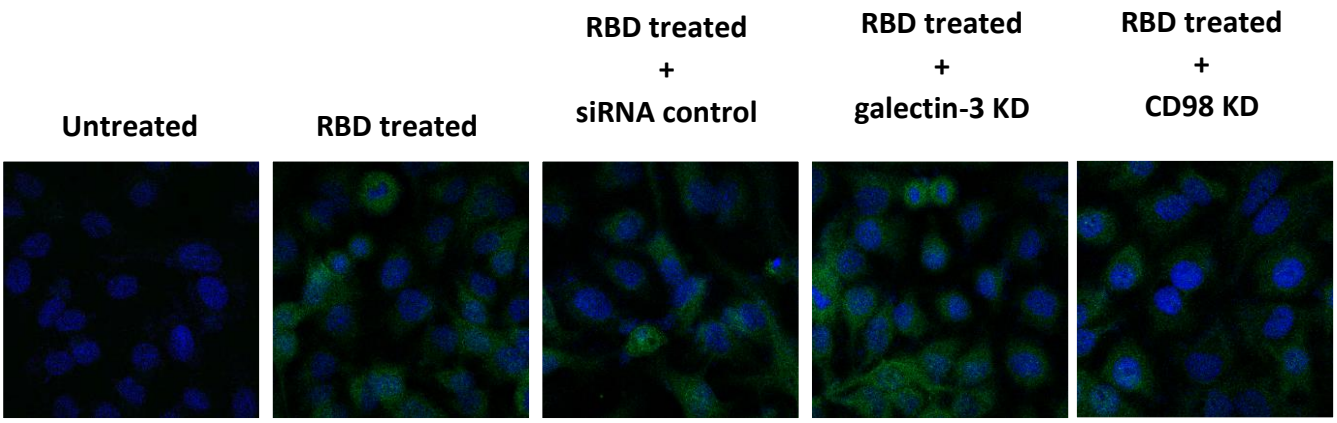


Figure A11, Immunofluorescence staining to assess cellular retention of SARS-CoV-2 following CD98 and galectin-3 knockdown: A549 cells were subjected to siRNA mediated CD98 or galectin-3 silencing using lipofectamine 2000. Cells were treated with the SARS-CoV-2 RBD at a concentration of 2 µg/ml for 24 h, then fixed with 4% PFA. Results show similar expression levels across all conditions. Cells were stained with 5 µg/ml of SARS-CoV-2 (PA5-81795) using secondary anti-rabbit FITC (green). Nuclei were stained using DAPI.

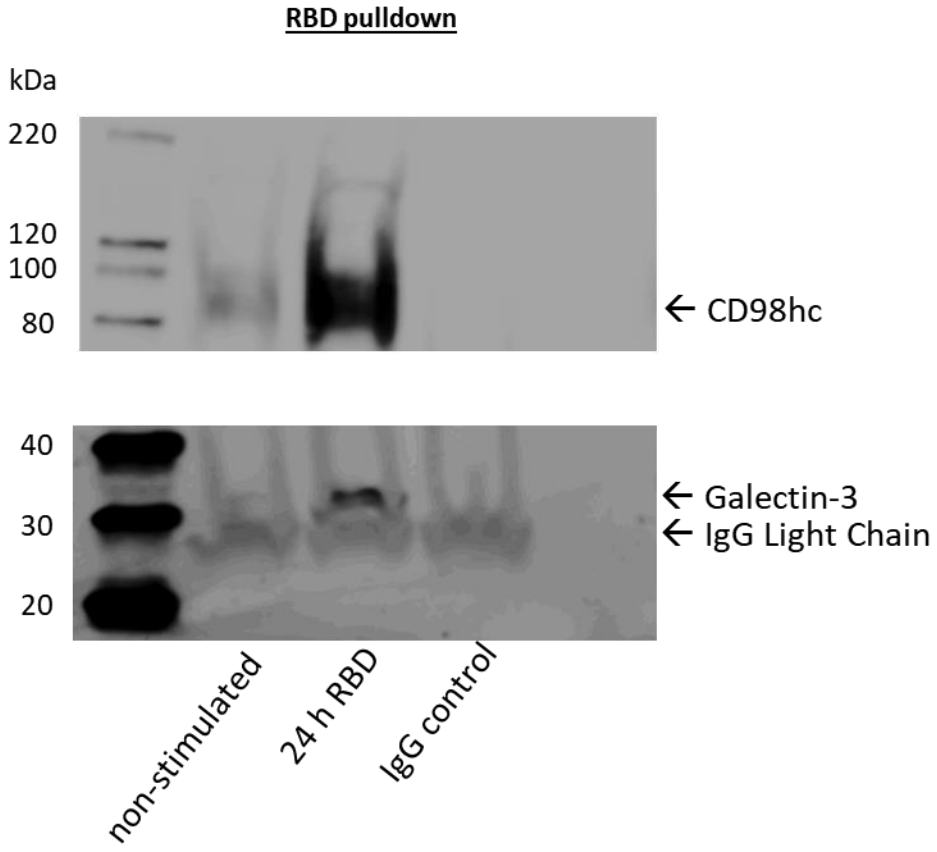


Figure A12, SARS-CoV-2 RBD pulldown in A549 whole cell lysates. This experiment was done by Dr Omeed Darweesh (PDRA, Gooptu Group, University of Leicester). Samples were loaded on a 4-20% (v/v) SDS-PAGE to analyse material co-immunoprecipitated from 1 mg A549 total cell lysates. Antibody immunoprecipitating against SARS-CoV-2 RBD were probed by western blotting for both CD98hc (~85 kDa) and galectin-3 (~31 kDa). Lane 2 represents non-stimulated cell lysates. Lane 3 represents lysates from cells treated with the SARS-CoV-2 RBD at a concentration of 2 µg/ml for 24 h. Lane 4 represents IgG control pull-down in cell lysates subjected to the same treatment as lane 3 (SARS-CoV-2 RBD at a concentration of 2 µg/ml for 24 h). Bands at ~25 kDa represent the immunoglobulin light chain from the pull-down antibody that are identified by the secondary blot antibody. The marker used was MagicMarkXP (in Lane 1). The blot was immersed with EZ-ECL chemiluminescent detection kit and image was taken at an exposure time of 5 minutes. (n=1).

References

- '"Antifibrotic therapy for fibrotic lung disease beyond idiopathic pulmonary fibrosis." Bridget F. Collins and Ganesh Raghu. Eur Respir Rev 2019; 28: 190022', (2019) *European respory review : an official journal of the European Respiratory Society,* 28(154), pp. 10.1183/16000617.5022-2019. Print 2019 Dec 31.
- Abu El-Asrar, A.M., Struyf, S., Van Damme, J. and Geboes, K. (2008) 'Circulating fibrocytes contribute to the myofibroblast population in proliferative vitreoretinopathy epiretinal membranes', *The British journal of ophthalmology,* 92(5), pp. 699-704.
- Acharya, N., Sharma, S.K., Mishra, D., Dhooria, S., Dhir, V. and Jain, S. (2020) 'Efficacy and safety of pirfenidone in systemic sclerosis-related interstitial lung disease-a randomised controlled trial', *Rheumatology international,* 40(5), pp. 703-710.
- Ahmad, N., Gabius, H.J., Andre, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F. and Brewer, C.F. (2004) 'Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes', *The Journal of biological chemistry,* 279(12), pp. 10841-10847.
- Akamatsu, T., Arai, Y., Kosugi, I., Kawasaki, H., Meguro, S., Sakao, M., Shibata, K., Suda, T., Chida, K. and Iwashita, T. (2013) 'Direct isolation of myofibroblasts and fibroblasts from bleomycin-injured lungs reveals their functional similarities and differences', *Fibrogenesis & tissue repair,* 6(1), pp. 15-15.
- Akhmetshina, A., Palumbo, K., Dees, C., Bergmann, C., Venalis, P., Zerr, P., Horn, A., Kireva, T., Beyer, C., Zwerina, J., Schneider, H., Sadowski, A., Riener, M.O., MacDougald, O.A., Distler, O., Schett, G. and Distler, J.H. (2012) 'Activation of canonical Wnt signalling is required for TGF-beta-mediated fibrosis', *Nature communications,* 3, pp. 735.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002) 'Molecular Biology of the Cell. 4th edition', *Garland Science.* Available at: https://www.ncbi.nlm.nih.gov/books/NBK21054/ [Accessed 1st March 2021].
- Alder, J.K., Chen, J.J., Lancaster, L., Danoff, S., Su, S.C., Cogan, J.D., Vulto, I., Xie, M., Qi, X., Tuder, R.M., Phillips, J.A., Lansdorp, P.M., Loyd, J.E. and Armanios, M.Y. (2008) 'Short telomeres are a risk factor for idiopathic pulmonary fibrosis', Proceedings of the National Academy of Sciences of the United States of America, 105(35), pp. 13051-13056.

- Allen, R.J., Guillen-Guio, B., Oldham, J.M., Ma, S.F., Dressen, A., Paynton, M.L., Kraven, L.M., Obeidat, M., Li, X., Ng, M., Braybrooke, R., Molina-Molina, M., Hobbs, B.D., Putman, R.K., Sakornsakolpat, P., Booth, H.L., Fahy, W.A., Hart, S.P., Hill, M.R., Hirani, N., Hubbard, R.B., McAnulty, R.J., Millar, A.B., Navaratnam, V., Oballa, E., Parfrey, H., Saini, G., Whyte, M.K.B., Zhang, Y., Kaminski, N., Adegunsoye, A., Strek, M.E., Neighbors, M., Sheng, X.R., Gudmundsson, G., Gudnason, V., Hatabu, H., Lederer, D.J., Manichaikul, A., Newell, J.D., O'Connor, G.T., Ortega, V.E., Xu, H., Fingerlin, T.E., Bosse, Y., Hao, K., Joubert, P., Nickle, D.C., Sin, D.D., Timens, W., Furniss, D., Morris, A.P., Zondervan, K.T., Hall, I.P., Sayers, I., Tobin, M.D., Maher, T.M., Cho, M.H., Hunninghake, G.M., Schwartz, D.A., Yaspan, B.L., Molyneaux, P.L., Flores, C., Noth, I., Jenkins, R.G. and Wain, L.V. (2020) 'Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis', *American journal of respiratory and critical care medicine,* 201(5), pp. 564-574.
- Allen, R.J., Porte, J., Braybrooke, R., Flores, C., Fingerlin, T.E., Oldham, J.M., Guillen-Guio, B., Ma, S.F., Okamoto, T., John, A.E., Obeidat, M., Yang, I.V., Henry, A., Hubbard, R.B., Navaratnam, V., Saini, G., Thompson, N., Booth, H.L., Hart, S.P., Hill, M.R., Hirani, N., Maher, T.M., McAnulty, R.J., Millar, A.B., Molyneaux, P.L., Parfrey, H., Rassl, D.M., Whyte, M.K.B., Fahy, W.A., Marshall, R.P., Oballa, E., Bosse, Y., Nickle, D.C., Sin, D.D., Timens, W., Shrine, N., Sayers, I., Hall, I.P., Noth, I., Schwartz, D.A., Tobin, M.D., Wain, L.V. and Jenkins, R.G. (2017) 'Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study', *The Lancet.Respiratory medicine,* 5(11), pp. 869-880.
- Aluwihare, P., Mu, Z., Zhao, Z., Yu, D., Weinreb, P.H., Horan, G.S., Violette, S.M. and Munger, J.S. (2009) 'Mice that lack activity of alphavbeta6- and alphavbeta8-integrins reproduce the abnormalities of Tgfb1- and Tgfb3-null mice', *Journal of cell science,* 122(Pt 2), pp. 227-232.
- 'American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS)', (2000) *American journal of respiratory and critical care medicine,* 161(2 Pt 1), pp. 646-664.

- Andoh, Y., Shimura, S., Aikawa, T., Sasaki, H. and Takishima, T. (1992) 'Perivascular fibrosis of muscular pulmonary arteries in chronic obstructive pulmonary disease', *Chest,* 102(6), pp. 1645-1650.
- Antoniou, K.M., Margaritopoulos, G.A., Tomassetti, S., Bonella, F., Costabel, U. and Poletti, V. (2014) 'Interstitial lung disease', *European respiratory review : an official journal of the European Respiratory Society,* 23(131), pp. 40-54.
- Arase, Y., Suzuki, F., Suzuki, Y., Akuta, N., Kobayashi, M., Kawamura, Y., Yatsuji, H., Sezaki, H., Hosaka, T., Hirakawa, M., Saito, S., Ikeda, K. and Kumada, H. (2008) 'Hepatitis C virus enhances incidence of idiopathic pulmonary fibrosis', *World journal of gastroenterology,* 14(38), pp. 5880-5886.
- Arodin Selenius, L., Wallenberg Lundgren, M., Jawad, R., Danielsson, O. and Bjornstedt, M. (2019) 'The Cell Culture Medium Affects Growth, Phenotype Expression and the Response to Selenium Cytotoxicity in A549 and HepG2 Cells', Antioxidants (Basel, Switzerland), 8(5), pp. 10.3390/antiox8050130.
- Atkinson, J.J. and Senior, R.M. (2003) 'Matrix metalloproteinase-9 in lung remodeling', *American journal of respiratory cell and molecular biology,* 28(1), pp. 12-24.
- Awadalla, N.J., Hegazy, A., Elmetwally, R.A. and Wahby, I. (2012) 'Occupational and environmental risk factors for idiopathic pulmonary fibrosis in Egypt: a multicenter case-control study', *The international journal of occupational and environmental medicine,* 3(3), pp. 107-116.
- B Moore, B., Lawson, W.E., Oury, T.D., Sisson, T.H., Raghavendran, K. and Hogaboam, C.M. (2013) 'Animal models of fibrotic lung disease', *American journal of respiratory cell and molecular biology,* 49(2), pp. 167-179.
- B Moore, B., Lawson, W.E., Oury, T.D., Sisson, T.H., Raghavendran, K. and Hogaboam, C.M. (2013) 'Animal models of fibrotic lung disease', American journal of respiratory cell and molecular biology, 49(2), pp. 167-179.
- Balestro, E., Calabrese, F., Turato, G., Lunardi, F., Bazzan, E., Marulli, G., Biondini, D., Rossi, E., Sanduzzi, A., Rea, F., Rigobello, C., Gregori, D., Baraldo, S., Spagnolo, P., Cosio, M.G.

and Saetta, M. (2016) 'Immune Inflammation and Disease Progression in Idiopathic Pulmonary Fibrosis', *PloS one,* 11(5), pp. e0154516.

- Bartis, D., Mise, N., Mahida, R.Y., Eickelberg, O. and Thickett, D.R. (2014) 'Epithelial-mesenchymal transition in lung development and disease: does it exist and is it important?', *Thorax,* 69(8), pp. 760-765.
- Baum, J. and Duffy, H.S. (2011) 'Fibroblasts and myofibroblasts: what are we talking about?', *Journal of cardiovascular pharmacology,* 57(4), pp. 376-379.
- Baumgartner, K.B., Samet, J.M., Coultas, D.B., Stidley, C.A., Hunt, W.C., Colby, T.V. and Waldron, J.A. (2000) 'Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. Collaborating Centers', *American Journal of Epidemiology,* 152(4), pp. 307-315.
- Baumgartner, K.B., Samet, J.M., Stidley, C.A., Colby, T.V. and Waldron, J.A. (1997) 'Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis', *American journal of respiratory and critical care medicine,* 155(1), pp. 242-248.
- Bellou, V., Belbasis, L. and Evangelou, E. (2021) 'Tobacco smoking and risk for pulmonary fibrosis: A prospective cohort study in UK Biobank', *Chest,* .
- Berg, D.T., Gupta, A., Richardson, M.A., O'Brien, L.A., Calnek, D. and Grinnell, B.W. (2007) 'Negative regulation of inducible nitric-oxide synthase expression mediated through transforming growth factor-beta-dependent modulation of transcription factor TCF11', *The Journal of biological chemistry,* 282(51), pp. 36837-36844.
- Biernacka, A., Dobaczewski, M. and Frangogiannis, N.G. (2011a) 'TGF-beta signaling in fibrosis', *Growth factors (Chur, Switzerland),* 29(5), pp. 196-202.
- Bocchino, M., Agnese, S., Fagone, E., Svegliati, S., Grieco, D., Vancheri, C., Gabrielli, A., Sanduzzi, A. and Avvedimento, E.V. (2010) 'Reactive oxygen species are required for maintenance and differentiation of primary lung fibroblasts in idiopathic pulmonary fibrosis', *PloS one,* 5(11), pp. e14003.
- Bocker, S. and Elling, L. (2017) 'Binding characteristics of galectin-3 fusion proteins', *Glycobiology,* 27(5), pp. 457-468.

- Bolte, S. and Cordelieres, F.P. (2006) 'A guided tour into subcellular colocalization analysis in light microscopy', *Journal of microscopy,* 224(Pt 3), pp. 213-232.
- Boo, S. and Dagnino, L. (2013) 'Integrins as Modulators of Transforming Growth Factor Beta Signaling in Dermal Fibroblasts During Skin Regeneration After Injury', *Advances in wound care,* 2(5), pp. 238-246.
- Borie, R., Crestani, B., Dieude, P., Nunes, H., Allanore, Y., Kannengiesser, C., Airo, P., Matucci-Cerinic, M., Wallaert, B., Israel-Biet, D., Cadranel, J., Cottin, V., Gazal, S., Peljto, A.L., Varga, J., Schwartz, D.A., Valeyre, D. and Grandchamp, B. (2013) 'The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population', *PloS one,* 8(8), pp. e70621.
- Boscher, C. and Nabi, I.R. (2013) 'Galectin-3- and phospho-caveolin-1-dependent outside-in integrin signaling mediates the EGF motogenic response in mammary cancer cells', *Molecular biology of the cell,* 24(13), pp. 2134-2145.
- Bourke, J.E., Li, X., Foster, S.R., Wee, E., Dagher, H., Ziogas, J., Harris, T., Bonacci, J.V. and Stewart, A.G. (2011) 'Collagen remodelling by airway smooth muscle is resistant to steroids and beta(2)-agonists', *The European respiratory journal,* 37(1), pp. 173-182.
- Bourke, S.J. (2006) 'Interstitial lung disease: progress and problems', *Postgraduate medical journal,* 82(970), pp. 494-499.
- Bouros, D. and Antoniou, K.M. (2005) 'Current and future therapeutic approaches in idiopathic pulmonary fibrosis', *The European respiratory journal,* 26(4), pp. 693-702.
- Boxall, C., Holgate, S.T. and Davies, D.E. (2006) 'The contribution of transforming growth factor-beta and epidermal growth factor signalling to airway remodelling in chronic asthma', *The European respiratory journal,* 27(1), pp. 208-229.
- Bradley, K.L., Stokes, C.A., Marciniak, S.J., Parker, L.C. and Condliffe, A.M. (2021) 'Role of unfolded proteins in lung disease', Thorax, 76(1), pp. 92-99.
- Breen, E., Shull, S., Burne, S., Absher, M., Kelley, J., Phan, S. and Cutroneo, K.R. (1992) 'Bleomycin regulation of transforming growth factor-beta mRNA in rat lung fibroblasts', *American journal of respiratory cell and molecular biology,* 6(2), pp. 146-152.

- Bringardner, B.D., Baran, C.P., Eubank, T.D. and Marsh, C.B. (2008) 'The role of inflammation in the pathogenesis of idiopathic pulmonary fibrosis', *Antioxidants & redox signaling,* 10(2), pp. 287-301.
- British Lung Foundation. (2020). 'What is pulmonary fibrosis'. [Online] Available at: https://www.blf.org.uk/support-for-you/pulmonary-fibrosis/treatment [Accessed 22nd March 2021].
- British Thoracic Society. (2020). BTS ILD Registry Annual Report 2020. [Online] Available at: file:///C:/Users/Azim%201/Downloads/BTS%20ILD%20Registry%20Annual%20Report%202020.pdf [Accessed 20th March 2021].
- Broecker, J., Eger, B.T. and Ernst, O.P. (2017) 'Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs', *Structure (London, England : 1993),* 25(2), pp. 384-392.
- Brown, N.F. and Marshall, J.F. (2019) 'Integrin-Mediated TGFbeta Activation Modulates the Tumour Microenvironment', *Cancers,* 11(9), pp. 10.3390/cancers11091221.
- Bulus, N., Feral, C., Pozzi, A. and Zent, R. (2012) 'CD98 increases renal epithelial cell proliferation by activating MAPKs', *PloS one,* 7(6), pp. e40026.
- Bumba, L., Laaf, D., Spiwok, V., Elling, L., Kren, V. and Bojarova, P. (2018) 'Poly-N-Acetyllactosamine Neo-Glycoproteins as Nanomolar Ligands of Human Galectin-3: Binding Kinetics and Modeling', *International journal of molecular sciences,* 19(2), pp. 10.3390/ijms19020372.
- Burgess, J.K., Mauad, T., Tjin, G., Karlsson, J.C. and Westergren-Thorsson, G. (2016) 'The extracellular matrix - the under-recognized element in lung disease?', *The Journal of pathology,* 240(4), pp. 397-409.
- Burman, A., Tanjore, H. and Blackwell, T.S. (2018) 'Endoplasmic reticulum stress in pulmonary fibrosis', *Matrix biology : journal of the International Society for Matrix Biology,* 68-69, pp. 355-365.

- Burnham, E.L., Janssen, W.J., Riches, D.W., Moss, M. and Downey, G.P. (2014) 'The fibroproliferative response in acute respiratory distress syndrome: mechanisms and clinical significance', *The European respiratory journal,* 43(1), pp. 276-285.
- Butler, M.W and Keane, M.P., (2018) 'The Role of Immunity and Inflammation in IPF Pathogenesis', *National Public Health Emergency Collection,* (15), pp 97-131.
- Cabrera-Benitez, N.E., Laffey, J.G., Parotto, M., Spieth, P.M., Villar, J., Zhang, H. and Slutsky, A.S. (2014) 'Mechanical ventilation-associated lung fibrosis in acute respiratory distress syndrome: a significant contributor to poor outcome', *Anesthesiology,* 121(1), pp. 189-198.
- Cai, S., Bulus, N., Fonseca-Siesser, P.M., Chen, D., Hanks, S.K., Pozzi, A. and Zent, R. (2005a) 'CD98 modulates integrin beta1 function in polarized epithelial cells', *Journal of cell science,* 118(Pt 5), pp. 889-899.
- Camelo, A., Dunmore, R., Sleeman, M.A. and Clarke, D.L. (2014) 'The epithelium in idiopathic pulmonary fibrosis: breaking the barrier', *Frontiers in pharmacology,* 4, pp. 173.
- Campisi, J. (1997) 'The biology of replicative senescence', European journal of cancer (Oxford, England : 1990), 33(5), pp. 703-709.
- Cantor, J., Slepak, M., Ege, N., Chang, J.T. and Ginsberg, M.H. (2011) 'Loss of T cell CD98 H chain specifically ablates T cell clonal expansion and protects from autoimmunity', *Journal of immunology (Baltimore, Md.: 1950),* 187(2), pp. 851-860.
- Cantor, J.M. and Ginsberg, M.H. (2012) 'CD98 at the crossroads of adaptive immunity and cancer', *Journal of cell science,* 125(Pt 6), pp. 1373-1382.
- Caro, F.M., Alberti, M.L., Fernandez, M.E. and Paulin, F. (2017a) 'The Influence of Sex on Prognosis of Patients With Idiopathic Pulmonary Fibrosis in a Retrospective Cohort', *Archivos de Bronconeumologia,* 53(11), pp. 649-650.
- Cavazza, A., Rossi, G., Carbonelli, C., Spaggiari, L., Paci, M. and Roggeri, A. (2010a) 'The role of histology in idiopathic pulmonary fibrosis: an update', *Respiratory medicine,* 104 Suppl 1, pp. 11.

- Cederfur, C., Salomonsson, E., Nilsson, J., Halim, A., Oberg, C.T., Larson, G., Nilsson, U.J. and Leffler, H. (2008) 'Different affinity of galectins for human serum glycoproteins: galectin-3 binds many protease inhibitors and acute phase proteins', *Glycobiology,* 18(5), pp. 384-394.
- Cervello, M., Augello, G., Cusimano, A., Emma, M.R., Balasus, D., Azzolina, A., McCubrey, J.A. and Montalto, G. (2017) 'Pivotal roles of glycogen synthase-3 in hepatocellular carcinoma', *Advances in biological regulation,* 65, pp. 59-76.
- Charles Patrick Davis, MD, PhD, MedicineNet, (2021). Medical Definition of Idiopathic. [Online] Available at: https://www.medicinenet.com/idiopathic/definition.htm [Accessed 17th September 2021].
- Chen, K.J., Li, Q., Wen, C.M., Duan, Z.X., Zhang, J.Y., Xu, C. and Wang, J.M. (2016) 'Bleomycin (BLM) Induces Epithelial-to-Mesenchymal Transition in Cultured A549 Cells via the TGF-beta/Smad Signaling Pathway', *Journal of Cancer,* 7(11), pp. 1557-1564.
- Chen, L.D., Zhang, Z.Y., Wei, X.J., Cai, Y.Q., Yao, W.Z., Wang, M.H., Huang, Q.F. and Zhang, X.B. (2020) 'Association between cytokine profiles and lung injury in COVID-19 pneumonia', *Respiratory research,* 21(1), pp. 201-2.
- Chen, S.S., Sun, L.W., Brickner, H. and Sun, P.Q. (2015) 'Downregulating galectin-3 inhibits proinflammatory cytokine production by human monocyte-derived dendritic cells via RNA interference', *Cellular immunology,* 294(1), pp. 44-53.
- Chen, X., Shi, C., Meng, X., Zhang, K., Li, X., Wang, C., Xiang, Z., Hu, K. and Han, X. (2016) 'Inhibition of Wnt/beta-catenin signaling suppresses bleomycin-induced pulmonary fibrosis by attenuating the expression of TGF-beta1 and FGF-2', *Experimental and molecular pathology,* 101(1), pp. 22-30.
- Chiu, Y.P., Sun, Y.C., Qiu, D.C., Lin, Y.H., Chen, Y.Q., Kuo, J.C. and Huang, J.R. (2020) 'Liquid-liquid phase separation and extracellular multivalent interactions in the tale of galectin-3', *Nature communications,* 11(1), pp. 1229-3.
- Cho, J.Y., Kim, A.R., Joo, H.G., Kim, B.H., Rhee, M.H., Yoo, E.S., Katz, D.R., Chain, B.M. and Jung, J.H. (2004) 'Cynaropicrin, a sesquiterpene lactone, as a new strong regulator of CD29

and CD98 functions', *Biochemical and biophysical research communications,* 313(4), pp. 954-961.

- Cho, S.J. and Stout-Delgado, H.W. (2020) 'Aging and Lung Disease', Annual Review of Physiology, 82, pp. 433-459.
- Cichon, M.A. and Radisky, D.C. (2014) 'Extracellular matrix as a contextual determinant of transforming growth factor-beta signaling in epithelial-mesenchymal transition and in cancer', *Cell adhesion & migration,* 8(6), pp. 588-594.
- Cottin, V. (2013) 'Significance of connective tissue diseases features in pulmonary fibrosis', *European respiratory review : an official journal of the European Respiratory Society,* 22(129), pp. 273-280.
- Cottin, V. (2016) 'Idiopathic interstitial pneumonias with connective tissue diseases features: A review', *Respirology (Carlton, Vic.),* 21(2), pp. 245-258.
- Cottin, V., Hirani, N.A., Hotchkin, D.L., Nambiar, A.M., Ogura, T., Otaola, M., Skowasch, D., Park, J.S., Poonyagariyagorn, H.K., Wuyts, W. and Wells, A.U. (2018) 'Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases', *European respiratory review : an official journal of the European Respiratory Society,* 27(150), pp. 10.1183/16000617.0076-2018. Print 2018 Dec 31.
- Cottin, V., Wollin, L., Fischer, A., Quaresma, M., Stowasser, S. and Harari, S. (2019) 'Fibrosing interstitial lung diseases: knowns and unknowns', *European respiratory review : an official journal of the European Respiratory Society,* 28(151), pp. 10.1183/16000617.0100-2018. Print 2019 Mar 31.
- Coutinho, A.E. and Chapman, K.E. (2011b) 'The anti-inflammatory and immunosuppressive effects of glucocorticoids, recent developments and mechanistic insights', *Molecular and cellular endocrinology,* 335(1), pp. 2-13.
- Crowe, M.J., Doetschman, T. and Greenhalgh, D.G. (2000) 'Delayed wound healing in immunodeficient TGF-beta 1 knockout mice', *The Journal of investigative dermatology,* 115(1), pp. 3-11.

- Dalton, P., Christian, H.C., Redman, C.W., Sargent, I.L. and Boyd, C.A. (2007) 'Membrane trafficking of CD98 and its ligand galectin 3 in BeWo cells--implication for placental cell fusion', *The FEBS journal,* 274(11), pp. 2715-2727.
- Dam, T.K., Gabius, H.J., Andre, S., Kaltner, H., Lensch, M. and Brewer, C.F. (2005) 'Galectins bind to the multivalent glycoprotein asialofetuin with enhanced affinities and a gradient of decreasing binding constants', *Biochemistry,* 44(37), pp. 12564-12571.
- Darby, I.A., Laverdet, B., Bonte, F. and Desmouliere, A. (2014) 'Fibroblasts and myofibroblasts in wound healing', *Clinical, cosmetic and investigational dermatology,* 7, pp. 301-311.
- De Felice, B., Garbi, C., Santoriello, M., Santillo, A. and Wilson, R.R. (2009) 'Differential apoptosis markers in human keloids and hypertrophic scars fibroblasts', *Molecular and cellular biochemistry,* 327(1-2), pp. 191-201.
- de la Ballina, L R, Cano-Crespo, S., Gonzalez-Munoz, E., Bial, S., Estrach, S., Cailleteau, L., Tissot, F., Daniel, H., Zorzano, A., Ginsberg, M.H., Palacin, M. and Feral, C.C. (2016) 'Amino Acid Transport Associated to Cluster of Differentiation 98 Heavy Chain (CD98hc) Is at the Cross-road of Oxidative Stress and Amino Acid Availability', *The Journal of biological chemistry,* 291(18), pp. 9700-9711.
- Del Valle, D.M., Kim-Schulze, S., Huang, H.H., Beckmann, N.D., Nirenberg, S., Wang, B., Lavin, Y., Swartz, T.H., Madduri, D., Stock, A., Marron, T.U., Xie, H., Patel, M., Tuballes, K., Van Oekelen, O., Rahman, A., Kovatch, P., Aberg, J.A., Schadt, E., Jagannath, S., Mazumdar, M., Charney, A.W., Firpo-Betancourt, A., Mendu, D.R., Jhang, J., Reich, D., Sigel, K., Cordon-Cardo, C., Feldmann, M., Parekh, S., Merad, M. and Gnjatic, S. (2020) 'An inflammatory cytokine signature predicts COVID-19 severity and survival', *Nature medicine,* 26(10), pp. 1636-1643.
- Delaine, T., Collins, P., MacKinnon, A., Sharma, G., Stegmayr, J., Rajput, V.K., Mandal, S., Cumpstey, I., Larumbe, A., Salameh, B.A., Kahl-Knutsson, B., van Hattum, H., van Scherpenzeel, M., Pieters, R.J., Sethi, T., Schambye, H., Oredsson, S., Leffler, H., Blanchard, H. and Nilsson, U.J. (2016) 'Galectin-3-Binding Glycomimetics that Strongly Reduce Bleomycin-Induced Lung Fibrosis and Modulate Intracellular Glycan

Recognition', *Chembiochem : a European journal of chemical biology,* 17(18), pp. 1759-1770.

- Desai, V.D., Hsia, H.C. and Schwarzbauer, J.E. (2014) 'Reversible modulation of myofibroblast differentiation in adipose-derived mesenchymal stem cells', PloS one, 9(1), pp. e86865.
- Desmouliere, A., Geinoz, A., Gabbiani, F. and Gabbiani, G. (1993) 'Transforming growth factor-beta 1 induces alpha-smooth muscle actin expression in granulation tissue myofibroblasts and in quiescent and growing cultured fibroblasts', *The Journal of cell biology,* 122(1), pp. 103-111.
- Devaraj, A. (2014) 'Imaging: how to recognise idiopathic pulmonary fibrosis', *European respiratory review : an official journal of the European Respiratory Society,* 23(132), pp. 215-219.
- Deves, R. and Boyd, C.A. (2000) 'Surface antigen CD98(4F2): not a single membrane protein, but a family of proteins with multiple functions', *The Journal of membrane biology,* 173(3), pp. 165-177.
- Dewage, S.N.V., Organ, L., Koumoundouros, E., Derseh, H.B., Perera, K.U.E., Samuel, C.S., Stent, A.W. and Snibson, K.J. (2019) 'The efficacy of pirfenidone in a sheep model of pulmonary fibrosis', Experimental lung research, 45(9-10), pp. 310-322.
- Diaz-Alvarez, L. and Ortega, E. (2017) 'The Many Roles of Galectin-3, a Multifaceted Molecule, in Innate Immune Responses against Pathogens', *Mediators of inflammation,* 2017, pp. 9247574.
- Dickson, R.P., Erb-Downward, J.R., Martinez, F.J. and Huffnagle, G.B. (2016) 'The Microbiome and the Respiratory Tract', *Annual Review of Physiology,* 78, pp. 481-504.
- Ding, Q., Luckhardt, T., Hecker, L., Zhou, Y., Liu, G., Antony, V.B., deAndrade, J. and Thannickal, V.J. (2011) 'New insights into the pathogenesis and treatment of idiopathic pulmonary fibrosis', *Drugs,* 71(8), pp. 981-1001.
- Doerner, A.M. and Zuraw, B.L. (2009) 'TGF-beta1 induced epithelial to mesenchymal transition (EMT) in human bronchial epithelial cells is enhanced by IL-1beta but not abrogated by corticosteroids', *Respiratory research,* 10, pp. 100.

- Dong, L.H., Jiang, Y.Y., Liu, Y.J., Cui, S., Xia, C.C., Qu, C., Jiang, X., Qu, Y.Q., Chang, P.Y. and Liu, F. (2015) 'The anti-fibrotic effects of mesenchymal stem cells on irradiated lungs via stimulating endogenous secretion of HGF and PGE2', *Scientific reports,* 5, pp. 8713.
- Dong, R., Zhang, M., Hu, Q., Zheng, S., Soh, A., Zheng, Y. and Yuan, H. (2018) 'Galectin-3 as a novel biomarker for disease diagnosis and a target for therapy (Review)', *International journal of molecular medicine,* 41(2), pp. 599-614.
- Dong, S. and Hughes, R.C. (1996) 'Galectin-3 stimulates uptake of extracellular Ca2+ in human Jurkat T-cells', *FEBS letters,* 395(2-3), pp. 165-169.
- Dong, S. and Hughes, R.C. (1997) 'Macrophage surface glycoproteins binding to galectin-3 (Mac-2-antigen)', *Glycoconjugate journal,* 14(2), pp. 267-274.
- Drakopanagiotakis, F., Wujak, L., Wygrecka, M. and Markart, P. (2018) 'Biomarkers in idiopathic pulmonary fibrosis', *Matrix biology : journal of the International Society for Matrix Biology,* 68-69, pp. 404-421.
- du Bois, R. and King, T.E. (2007) 'Challenges in pulmonary fibrosis x 5: the NSIP/UIP debate', *Thorax,* 62(11), pp. 1008-1012.
- Duckworth, A., Gibbons, M.A., Allen, R.J., Almond, H., Beaumont, R.N., Wood, A.R., Lunnon, K., Lindsay, M.A., Wain, L.V., Tyrrell, J. and Scotton, C.J. (2021) 'Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a mendelian randomisation study', The Lancet.Respiratory medicine, 9(3), pp. 285-294.
- Duscher, D., Maan, Z.N., Wong, V.W., Rennert, R.C., Januszyk, M., Rodrigues, M., Hu, M., Whitmore, A.J., Whittam, A.J., Longaker, M.T. and Gurtner, G.C. (2014) 'Mechanotransduction and fibrosis', *Journal of Biomechanics,* 47(9), pp. 1997-2005.
- Eickelberg, O. and Laurent, G.J. (2010) 'The quest for the initial lesion in idiopathic pulmonary fibrosis: gene expression differences in IPF fibroblasts', *American Journal of Respiratory Cell and Molecular Biology : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association,* 42(1), pp. 1-2.
- Elsebai, M.F., Mocan, A. and Atanasov, A.G. (2016) 'Cynaropicrin: A Comprehensive Research Review and Therapeutic Potential As an Anti-Hepatitis C Virus Agent', *Frontiers in pharmacology,* 7, pp. 472.

- Endeman, H., Meijvis, S.C., Rijkers, G.T., van Velzen-Blad, H., van Moorsel, C.H., Grutters, J.C. and Biesma, D.H. (2011) 'Systemic cytokine response in patients with community-acquired pneumonia', *The European respiratory journal,* 37(6), pp. 1431-1438.
- Enomoto, N., Suda, T., Kono, M., Kaida, Y., Hashimoto, D., Fujisawa, T., Inui, N., Nakamura, Y., Imokawa, S., Funai, K. and Chida, K. (2013) 'Amount of elastic fibers predicts prognosis of idiopathic pulmonary fibrosis', *Respiratory medicine,* 107(10), pp. 1608-1616.
- Esmaili, M. and Overduin, M. (2018) 'Membrane biology visualized in nanometer-sized discs formed by styrene maleic acid polymers', *Biochimica et biophysica acta.Biomembranes,* 1860(2), pp. 257-263.
- Esposito, A.J., Menon AA AUIDCID:, 0., Ghosh, A.J., Putman, R.K., Fredenburgh, L.E., El-Chemaly, S.Y., Goldberg, H.J., Baron, R.M., Hunninghake, G.M. and Doyle, T.J. (2020) 'Increased Odds of Death for Patients with Interstitial Lung Disease and COVID-19: A Case-Control Study', American journal of respiratory and critical care medicine, 202(12), pp. 1710-1713.
- Estrach, S., Lee, S.A., Boulter, E., Pisano, S., Errante, A., Tissot, F.S., Cailleteau, L., Pons, C., Ginsberg, M.H. and Feral, C.C. (2014) 'CD98hc (SLC3A2) loss protects against ras-driven tumorigenesis by modulating integrin-mediated mechanotransduction', *Cancer research,* 74(23), pp. 6878-6889.
- Eunhee, S. Yi and Cagle, P.T. (2020). Pathology of Usual Interstitial Pneumonia. [Online] Available at: https://emedicine.medscape.com/article/2078722-overview [Accessed 18th September 2021].
- Fala, L. (2015) 'Ofev (Nintedanib): First Tyrosine Kinase Inhibitor Approved for the Treatment of Patients with Idiopathic Pulmonary Fibrosis', *American health & drug benefits,* 8(Spec Feature), pp. 101-104.
- Faouzi, S., Le Bail, B., Neaud, V., Boussarie, L., Saric, J., Bioulac-Sage, P., Balabaud, C. and Rosenbaum, J. (1999) 'Myofibroblasts are responsible for collagen synthesis in the stroma of human hepatocellular carcinoma: an in vivo and in vitro study', *Journal of hepatology,* 30(2), pp. 275-284.

- Fell, C.D., Martinez, F.J., Liu, L.X., Murray, S., Han, M.K., Kazerooni, E.A., Gross, B.H., Myers, J., Travis, W.D., Colby, T.V., Toews, G.B. and Flaherty, K.R. (2010) 'Clinical predictors of a diagnosis of idiopathic pulmonary fibrosis', *American journal of respiratory and critical care medicine,* 181(8), pp. 832-837.
- Fenczik, C.A., Zent, R., Dellos, M., Calderwood, D.A., Satriano, J., Kelly, C. and Ginsberg, M.H. (2001) 'Distinct domains of CD98hc regulate integrins and amino acid transport', *The Journal of biological chemistry,* 276(12), pp. 8746-8752.
- Feral, C.C., Nishiya, N., Fenczik, C.A., Stuhlmann, H., Slepak, M. and Ginsberg, M.H. (2005b) 'CD98hc (SLC3A2) mediates integrin signaling', *Proceedings of the National Academy of Sciences of the United States of America,* 102(2), pp. 355-360.
- Fernandez Fabrellas, E., Peris Sanchez, R., Sabater Abad, C. and Juan Samper, G. (2018) 'Prognosis and Follow-Up of Idiopathic Pulmonary Fibrosis', *Medical sciences (Basel, Switzerland),* 6(2), pp. 10.3390/medsci6020051.
- Fernandez, I.E. and Eickelberg, O. (2012) 'The impact of TGF-beta on lung fibrosis: from targeting to biomarkers', *Proceedings of the American Thoracic Society,* 9(3), pp. 111-116.
- Fernández-Fabrellas, F., Sánchez, R.P., Abad, C.S., Samper, G.J., (2018) 'Prognosis and Follow-Up of Idiopathic Pulmonary Fibrosis', *Medical Sciences,* 6(2), pp. 51
- Filer, A., Bik, M., Parsonage, G.N., Fitton, J., Trebilcock, E., Howlett, K., Cook, M., Raza, K., Simmons, D.L., Thomas, A.M., Salmon, M., Scheel-Toellner, D., Lord, J.M., Rabinovich, G.A. and Buckley, C.D. (2009) 'Galectin 3 induces a distinctive pattern of cytokine and chemokine production in rheumatoid synovial fibroblasts via selective signaling pathways', *Arthritis and Rheumatism,* 60(6), pp. 1604-1614.
- Finnson, K.W., Almadani, Y. and Philip, A. (2020) 'Non-canonical (non-SMAD2/3) TGF-beta signaling in fibrosis: Mechanisms and targets', *Seminars in cell & developmental biology,* 101, pp. 115-122.
- Fisher-Wellman, K. and Bloomer, R.J. (2009) 'Acute exercise and oxidative stress: a 30 year history', *Dynamic medicine : DM,* 8, pp. 1-1.

- Fisher-Wellman, K., Bell, H.K. and Bloomer, R.J. (2009) 'Oxidative stress and antioxidant defense mechanisms linked to exercise during cardiopulmonary and metabolic disorders', *Oxidative medicine and cellular longevity,* 2(1), pp. 43-51.
- Flaherty, K.R., Brown, K.K., Wells, A.U., Clerisme-Beaty, E., Collard, H.R., Cottin, V., Devaraj, A., Inoue, Y., Le Maulf, F., Richeldi, L., Schmidt, H., Walsh, S., Mezzanotte, W. and Schlenker-Herceg, R. (2017) 'Design of the PF-ILD trial: a double-blind, randomised, placebo-controlled phase III trial of nintedanib in patients with progressive fibrosing interstitial lung disease', *BMJ open respiratory research,* 4(1), pp. 000212. eCollection 2017.
- Flaherty, K.R., Wells, A.U., Cottin, V., Devaraj, A., Walsh, S.L.F., Inoue, Y., Richeldi, L., Kolb, M., Tetzlaff, K., Stowasser, S., Coeck, C., Clerisme-Beaty, E., Rosenstock, B., Quaresma, M., Haeufel, T., Goeldner, R.G., Schlenker-Herceg, R., Brown, K.K. and INBUILD Trial Investigators (2019) 'Nintedanib in Progressive Fibrosing Interstitial Lung Diseases', *The New England journal of medicine,* 381(18), pp. 1718-1727.
- Flores-Ibarra, A., Vertesy, S., Medrano, F.J., Gabius, H.J. and Romero, A. (2018) 'Crystallization of a human galectin-3 variant with two ordered segments in the shortened N-terminal tail', *Scientific reports,* 8(1), pp. 9835-x.
- Foster, K.A., Oster, C.G., Mayer, M.M., Avery, M.L. and Audus, K.L. (1998) 'Characterization of the A549 cell line as a type II pulmonary epithelial cell model for drug metabolism', *Experimental cell research,* 243(2), pp. 359-366.
- Froese, A.R., Shimbori, C., Bellaye, P.S., Inman, M., Obex, S., Fatima, S., Jenkins, G., Gauldie, J., Ask, K. and Kolb, M. (2016) 'Stretch-induced Activation of Transforming Growth Factor-beta1 in Pulmonary Fibrosis', *American journal of respiratory and critical care medicine,* 194(1), pp. 84-96.
- Fuchs, S., Hollins, A.J., Laue, M., Schaefer, U.F., Roemer, K., Gumbleton, M. and Lehr, C.M. (2003) 'Differentiation of human alveolar epithelial cells in primary culture: morphological characterization and synthesis of caveolin-1 and surfactant protein-C', *Cell and tissue research,* 311(1), pp. 31-45.

- Gabasa, M., Duch, P., Jorba, I., Gimenez, A., Lugo, R., Pavelescu, I., Rodriguez-Pascual, F., Molina-Molina, M., Xaubet, A., Pereda, J. and Alcaraz, J. (2017) 'Epithelial contribution to the profibrotic stiff microenvironment and myofibroblast population in lung fibrosis', *Molecular biology of the cell,* 28(26), pp. 3741-3755.
- Gabbiani, G., Ryan, G.B. and Majne, G. (1971) 'Presence of modified fibroblasts in granulation tissue and their possible role in wound contraction', *Experientia,* 27(5), pp. 549-550.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A (2005) 'Protein Identification and Analysis Tools on the ExPASy Server': *The Proteomics Protocols Handbook, Humana Press*, pp. 571-607 Full text - Copyright Humana Press.
- Gauldie, J., Kolb, M. and Sime, P.J. (2002) 'A new direction in the pathogenesis of idiopathic pulmonary fibrosis?', *Respiratory research,* 3, pp. 1.
- George, P.M., Barratt, S.L., Condliffe, R., Desai, S.R., Devaraj, A., Forrest, I., Gibbons, M.A., Hart, N., Jenkins, R.G., McAuley, D.F., Patel, B.V., Thwaite, E. and Spencer, L.G. (2020) 'Respiratory follow-up of patients with COVID-19 pneumonia', *Thorax,* 75(11), pp. 1009-1016.
- Ghisa, M., Marinelli, C., Savarino, V. and Savarino, E. (2019) 'Idiopathic pulmonary fibrosis and GERD: links and risks', *Therapeutics and clinical risk management,* 15, pp. 1081-1093.
- Gkretsi, V. and Stylianopoulos, T. (2018) 'Cell Adhesion and Matrix Stiffness: Coordinating Cancer Cell Invasion and Metastasis', *Frontiers in oncology,* 8, pp. 145.
- Glebov, O.O. (2020) 'Understanding SARS-CoV-2 endocytosis for COVID-19 drug repurposing', *The FEBS journal,* 287(17), pp. 3664-3671.
- Gong, H.C., Honjo, Y., Nangia-Makker, P., Hogan, V., Mazurak, N., Bresalier, R.S. and Raz, A. (1999) 'The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells', *Cancer research,* 59(24), pp. 6239-6245.
- Gonzalez, G.E., Cassaglia, P., Noli Truant, S., Fernandez, M.M., Wilensky, L., Volberg, V., Malchiodi, E.L., Morales, C. and Gelpi, R.J. (2014) 'Galectin-3 is essential for early wound healing and ventricular remodeling after myocardial infarction in mice', *International journal of cardiology,* 176(3), pp. 1423-1425.

- Goodwin, A. and Jenkins, G. (2009) 'Role of integrin-mediated TGFbeta activation in the pathogenesis of pulmonary fibrosis', *Biochemical Society transactions,* 37(Pt 4), pp. 849-854.
- Gooptu, B. *In print* 'Surfactant protein C mutations and familial pulmonary fibrosis: stuck in a loop on the scenic route', European Respiratory Journal.
- Habiel, D.M. and Hogaboam, C.M. (2017) 'Heterogeneity of Fibroblasts and Myofibroblasts in Pulmonary Fibrosis', *Current pathobiology reports,* 5(2), pp. 101-110.
- Hahm, K., Lukashev, M.E., Luo, Y., Yang, W.J., Dolinski, B.M., Weinreb, P.H., Simon, K.J., Chun Wang, L., Leone, D.R., Lobb, R.R., McCrann, D.J., Allaire, N.E., Horan, G.S., Fogo, A., Kalluri, R., Shield, C.F., Sheppard, D., Gardner, H.A. and Violette, S.M. (2007) 'Alphav beta6 integrin regulates renal fibrosis and inflammation in Alport mouse', The American journal of pathology, 170(1), pp. 110-125.
- Hajari Case, A. and Johnson, P. (2017) 'Clinical use of nintedanib in patients with idiopathic pulmonary fibrosis', *BMJ open respiratory research,* 4(1), pp. 000192. eCollection 2017.
- Halimi, H., Rigato, A., Byrne, D., Ferracci, G., Sebban-Kreuzer, C., ElAntak, L. and Guerlesquin, F. (2014) 'Glycan dependence of Galectin-3 self-association properties', *PloS one,* 9(11), pp. e111836.
- Hall, S.C.L., Clifton, L.A., Tognoloni, C., Morrison, K.A., Knowles, T.J., Kinane, C.J., Dafforn, T.R., Edler, K.J. and Arnold, T. (2020) 'Adsorption of a styrene maleic acid (SMA) copolymer-stabilized phospholipid nanodisc on a solid-supported planar lipid bilayer', *Journal of colloid and interface science,* 574, pp. 272-284.
- Han, M.K., Baker, M., Zhang, Y., Yang, C., Zhang, M., Garg, P., Viennois, E. and Merlin, D. (2018) 'Overexpression of CD98 in intestinal epithelium dysregulates miRNAs and their targeted proteins along the ileal villus-crypt axis', *Scientific reports,* 8(1), pp. 16220-9.
- Han, M.K., Murray, S., Fell, C.D., Flaherty, K.R., Toews, G.B., Myers, J., Colby, T.V., Travis, W.D., Kazerooni, E.A., Gross, B.H. and Martinez, F.J. (2008) 'Sex differences in physiological progression of idiopathic pulmonary fibrosis', *The European respiratory journal,* 31(6), pp. 1183-1188.

- Hansen, N.U., Karsdal, M.A., Brockbank, S., Cruwys, S., Ronnow, S. and Leeming, D.J. (2016) 'Tissue turnover of collagen type I, III and elastin is elevated in the PCLS model of IPF and can be restored back to vehicle levels using a phosphodiesterase inhibitor', *Respiratory research,* 17(1), pp. 8.
- Hara, A., Niwa, M., Noguchi, K., Kanayama, T., Niwa, A., Matsuo, M., Hatano, Y. and Tomita, H. (2020) 'Galectin-3 as a Next-Generation Biomarker for Detecting Early Stage of Various Diseases', *Biomolecules,* 10(3), pp. 10.3390/biom10030389.
- Harris, W.T., Kelly, D.R., Zhou, Y., Wang, D., MacEwen, M., Hagood, J.S., Clancy, J.P., Ambalavanan, N. and Sorscher, E.J. (2013) 'Myofibroblast differentiation and enhanced TGF-B signaling in cystic fibrosis lung disease', *PloS one,* 8(8), pp. e70196.
- Hashiba, K., Sano, M., Nio-Kobayashi, J., Hojo, T., Skarzynski, D.J. and Okuda, K. (2014) 'Galectin-3 contributes to luteolysis by binding to Beta 1 integrin in the bovine corpus luteum', *Biology of reproduction,* 91(1), pp. 2.
- Hata, A. and Chen, Y.G. (2016) 'TGF-beta Signaling from Receptors to Smads', *Cold Spring Harbor perspectives in biology,* 8(9), pp. 10.1101/cshperspect.a022061
- Hawrylowicz, C., Richards, D., Loke, T.K., Corrigan, C. and Lee, T. (2002) 'A defect in corticosteroid-induced IL-10 production in T lymphocytes from corticosteroid-resistant asthmatic patients', The Journal of allergy and clinical immunology, 109(2), pp. 369-370.
- Haynes, B.F., Hemler, M.E., Mann, D.L., Eisenbarth, G.S., Shelhamer, J., Mostowsk, Thomas, C.A., Strominger, J.L. and Fauci, A.S. (1981) 'Characterization of a monoclonal antibody (4F2) that binds to human monocytes and to a subset of activated lymphocytes', *Journal of immunology (Baltimore, Md.: 1950),* 126(4), pp. 1409-1414.
- Helling, B.A., Gerber, A.N., Kadiyala, V., Sasse, S.K., Pedersen, B.S., Sparks, L., Nakano, Y., Okamoto, T., Evans, C.M., Yang, I.V. and Schwartz, D.A. (2017) 'Regulation of MUC5B Expression in Idiopathic Pulmonary Fibrosis', *American journal of respiratory cell and molecular biology,* 57(1), pp. 91-99.
- Henderson, N.C. and Sethi, T. (2009) 'The regulation of inflammation by galectin-3', *Immunological reviews,* 230(1), pp. 160-171.

- Henderson, N.C., Collis, E.A., Mackinnon, A.C., Simpson, K.J., Haslett, C., Zent, R., Ginsberg, M. and Sethi, T. (2004) 'CD98hc (SLC3A2) interaction with beta 1 integrins is required for transformation', *The Journal of biological chemistry,* 279(52), pp. 54731-54741.
- Henderson, N.C., Mackinnon, A.C., Farnworth, S.L., Poirier, F., Russo, F.P., Iredale, J.P., Haslett, C., Simpson, K.J. and Sethi, T. (2006) 'Galectin-3 regulates myofibroblast activation and hepatic fibrosis', *Proceedings of the National Academy of Sciences of the United States of America,* 103(13), pp. 5060-5065.
- Higashiyama, H., Yoshimoto, D., Okamoto, Y., Kikkawa, H., Asano, S. and Kinoshita, M. (2007) 'Receptor-activated Smad localisation in bleomycin-induced pulmonary fibrosis', *Journal of clinical pathology,* 60(3), pp. 283-289.
- Hinz, B., Phan, S.H., Thannickal, V.J., Galli, A., Bochaton-Piallat, M-L., Gabbiani, G. (2007). ' The Myofibroblast', *The American Journal of PATHOLOGY',* 170(60), pp.1807-1816.
- Hirani, N., MacKinnon, A.C., Nicol, L., Ford, P., Schambye, H., Pedersen, A., Nilsson, U.J., Leffler, H., Sethi, T., Tantawi, S., Gavelle, L., Slack, R.J., Mills, R., Karmakar, U., Humphries, D., Zetterberg, F., Keeling, L., Paul, L., Molyneaux, P.L., Li, F., Funston, W., Forrest, I.A., Simpson, A.J., Gibbons, M.A. and Maher, T.M. (2020) 'Target-inhibition of Galectin-3 by Inhaled TD139 in Patients with Idiopathic Pulmonary Fibrosis', *The European respiratory journal,* .
- Hirani, N., Nicol, L., MacKinnon, A.C., Ford, P., Schambye, H., Pedersen., Nilsson, U., Leffler, H., Thomas, T., Gibbons, O.K.M., Simpson, J., Maher, T. (2016). 'TD139, A Novel Inhaled Galectin-3 Inhibitor for The Treatment of Idiopathic Pulmonary Fibrosis (IPF). Results from The First in (IPF) Patients Study', *QJM,* 109, pp. 16.
- Ho, J.E., Gao, W., Levy, D., Santhanakrishnan, R., Araki, T., Rosas, I.O., Hatabu, H., Latourelle, J.C., Nishino, M., Dupuis, J., Washko, G.R., O'Connor, G.T. and Hunninghake, G.M. (2016) 'Galectin-3 Is Associated with Restrictive Lung Disease and Interstitial Lung Abnormalities', *American journal of respiratory and critical care medicine,* 194(1), pp. 77-83.

- Hodge, S., Matthews, G., Mukaro, V., Ahern, J., Shivam, A., Hodge, G., Holmes, M., Jersmann, H. and Reynolds, P.N. (2011) 'Cigarette smoke-induced changes to alveolar macrophage phenotype and function are improved by treatment with procysteine', *American journal of respiratory cell and molecular biology,* 44(5), pp. 673-681.
- Horiguchi, M., Ota, M. and Rifkin, D.B. (2012) 'Matrix control of transforming growth factor-beta function', *Journal of Biochemistry,* 152(4), pp. 321-329.
- Horowitz, J.C. and Thannickal, V.J. (2006b) 'Epithelial-mesenchymal interactions in pulmonary fibrosis', *Seminars in respiratory and critical care medicine,* 27(6), pp. 600-612.
- Huang, F. and Chen, Y.G. (2012) 'Regulation of TGF-beta receptor activity', *Cell & bioscience,* 2, pp. 9.
- Huang, X., Yang, N., Fiore, V.F., Barker, T.H., Sun, Y., Morris, S.W., Ding, Q., Thannickal, V.J. and Zhou, Y. (2012) 'Matrix stiffness-induced myofibroblast differentiation is mediated by intrinsic mechanotransduction', *American journal of respiratory cell and molecular biology,* 47(3), pp. 340-348.
- Hubbard, R. (2001) 'Occupational dust exposure and the aetiology of cryptogenic fibrosing alveolitis', *The European respiratory journal.Supplement,* 32, pp. 119s-121s.
- Hutchinson, J.P., Fogarty, A.W., McKeever, T.M. and Hubbard, R.B. (2016) 'In-Hospital Mortality after Surgical Lung Biopsy for Interstitial Lung Disease in the United States. 2000 to 2011', American journal of respiratory and critical care medicine, 193(10), pp. 1161-1167.
- Idiopathic Pulmonary Fibrosis Clinical Research Network, Raghu, G., Anstrom, K.J., King, T.E., Lasky, J.A. and Martinez, F.J. (2012) 'Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis', *The New England journal of medicine,* 366(21), pp. 1968-1977.
- Ikegami, M., Whitsett, J.A., Martis, P.C. and Weaver, T.E. (2005) 'Reversibility of lung inflammation caused by SP-B deficiency', *American journal of physiology.Lung cellular and molecular physiology,* 289(6), pp. 962.

- Ingber, D.E. (2006) 'Cellular mechanotransduction: putting all the pieces together again', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology,* 20(7), pp. 811-827.
- Invernizzi, R. and Molyneaux, P.L. (2019) 'The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis', *European respiratory review : an official journal of the European Respiratory Society,* 28(152), pp. 10.1183/16000617.0045-2019. Print 2019 Jun 30.
- Irving, J.A., Ekeowa, U.I., Belorgey, D., Haq, I., Gooptu, B., Miranda, E., Perez, J., Roussel, B.D., Ordonez, A., Dalton, L.E., Thomas, S.E., Marciniak, S.J., Parfrey, H., Chilvers, E.R., Teckman, J.H., Alam, S., Mahadeva, R., Rashid, S.T., Vallier, L. and Lomas, D.A. (2011) 'The serpinopathies studying serpin polymerization in vivo', *Methods in enzymology,* 501, pp. 421-466.
- Isah, M.D., Abbas, A., Abba, A.A. and Umar, M. (2016) 'Idiopathic pulmonary fibrosis misdiagnosed as sputum-negative pulmonary tuberculosis', *Annals of African medicine,* 15(4), pp. 204-206.
- Jenkins, G., Blanchard, A., Borok, Z., Bradding, P., Ehrhardt, C., Fisher, A., Hirani, N., Johnson, S., Konigshoff, M., Maher, T.M., Millar, A., Parfrey, H., Scotton, C., Tetley, T., Thickett, D., Wolters, P. and ECIPF workshop (2012) 'In search of the fibrotic epithelial cell: opportunities for a collaborative network', Thorax, 67(2), pp. 179-182.
- Jenkins, R.G., Moore, B.B., Chambers, R.C., Eickelberg, O., Konigshoff, M., Kolb, M., Laurent, G.J., Nanthakumar, C.B., Olman, M.A., Pardo, A., Selman, M., Sheppard, D., Sime, P.J., Tager, A.M., Tatler, A.L., Thannickal, V.J., White, E.S. and ATS Assembly on Respiratory Cell and Molecular Biology (2017) 'An Official American Thoracic Society Workshop Report: Use of Animal Models for the Preclinical Assessment of Potential Therapies for Pulmonary Fibrosis', *American journal of respiratory cell and molecular biology,* 56(5), pp. 667-679.
- Jia, H.P., Look, D.C., Shi, L., Hickey, M., Pewe, L., Netland, J., Farzan, M., Wohlford-Lenane, C., Perlman, S. and McCray, P.B. (2005) 'ACE2 receptor expression and severe acute

respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia', *Journal of virology,* 79(23), pp. 14614-14621.

- Jiang, C., Huang, H., Liu, J., Wang, Y., Lu, Z. and Xu, Z. (2012) 'Fasudil, a Rho-kinase inhibitor, attenuates bleomycin-induced pulmonary fibrosis in mice', *International journal of molecular sciences,* 13(7), pp. 8293-8307.
- Jing, G., Bei, C., Wen, P., Yun-qiao, L., Wei, Z. (2014) 'Comparison of transfection efficiency of THP-1 monocytes by different methods', *Journal of Clinical Rehabilitative Tissue Engineering Research,* 18(38), pp. 6105-6109.
- Jo, H.E., Randhawa, S., Corte, T.J. and Moodley, Y. (2016) 'Idiopathic Pulmonary Fibrosis and the Elderly: Diagnosis and Management Considerations', *Drugs & aging,* 33(5), pp. 321-334.
- Johannes, L., Jacob, R. and Leffler, H. (2018a) 'Galectins at a glance', *Journal of cell science,* 131(9), pp. 10.1242/jcs.208884.
- Johnson, P., Beswick, E.J., Chao, C., Powell, D.W., Hellmich, M.R. and Pinchuk, I.V. (2016) 'Isolation of CD 90+ Fibroblast/Myofibroblasts from Human Frozen Gastrointestinal Specimens', *Journal of visualized experiments : JoVE,* (107):e53691. doi(107), pp. e53691.
- Jones, M.G., Fabre, A., Schneider, P., Cinetto, F., Sgalla, G., Mavrogordato, M., Jogai, S., Alzetani, A., Marshall, B.G., O'Reilly, K.M., Warner, J.A., Lackie, P.M., Davies, D.E., Hansell, D.M., Nicholson, A.G., Sinclair, I., Brown, K.K. and Richeldi, L. (2016) 'Three-dimensional characterization of fibroblast foci in idiopathic pulmonary fibrosis', *JCI insight,* 1(5), pp. 10.1172/jci.insight.86375.
- Jones, R., Nelson, M., Bracchi, M., Asboe, D. and Boffito, M. (2020) 'COVID-19 in patients with HIV', The lancet.HIV, 7(6), pp. e383-9. Epub 2020 May 14.
- Juarez, M.M., Chan, A.L., Norris, A.G., Morrissey, B.M. and Albertson, T.E. (2015b) 'Acute exacerbation of idiopathic pulmonary fibrosis-a review of current and novel pharmacotherapies', *Journal of thoracic disease,* 7(3), pp. 499-519.
- Kabir-Salmani, M., Fukuda, M.N., Kanai-Azuma, M., Ahmed, N., Shiokawa, S., Akimoto, Y., Sakai, K., Nagamori, S., Kanai, Y., Sugihara, K. and Iwashita, M. (2008a) 'The membrane-

spanning domain of CD98 heavy chain promotes alpha(v)beta3 integrin signals in human extravillous trophoblasts', *Molecular endocrinology (Baltimore, Md.),* 22(3), pp. 707-715.

- Kahounova, Z., Kurfurstova, D., Bouchal, J., Kharaishvili, G., Navratil, J., Remsik, J., Simeckova, S., Student, V., Kozubik, A. and Soucek, K. (2018) 'The fibroblast surface markers FAP, anti-fibroblast, and FSP are expressed by cells of epithelial origin and may be altered during epithelial-to-mesenchymal transition', *Cytometry.Part A : the journal of the International Society for Analytical Cytology,* 93(9), pp. 941-951.
- Kamili, N.A., Arthur, C.M., Gerner-Smidt, C., Tafesse, E., Blenda, A., Dias-Baruffi, M. and Stowell, S.R. (2016) 'Key regulators of galectin-glycan interactions', *Proteomics,* 16(24), pp. 3111-3125.
- Kaminski, N. and Rosas, I.O. (2006) 'Gene expression profiling as a window into idiopathic pulmonary fibrosis pathogenesis: can we identify the right target genes?', *Proceedings of the American Thoracic Society,* 3(4), pp. 339-344.
- Kannengiesser, C., Borie, R., Menard, C., Reocreux, M., Nitschke, P., Gazal, S., Mal, H., Taille, C., Cadranel, J., Nunes, H., Valeyre, D., Cordier, J.F., Callebaut, I., Boileau, C., Cottin, V., Grandchamp, B., Revy, P. and Crestani, B. (2015) 'Heterozygous RTEL1 mutations are associated with familial pulmonary fibrosis', *The European respiratory journal,* 46(2), pp. 474-485.
- Karkkainen, M., Kettunen, H.P., Nurmi, H., Selander, T., Purokivi, M. and Kaarteenaho, R. (2017) 'Effect of smoking and comorbidities on survival in idiopathic pulmonary fibrosis', *Respiratory research,* 18(1), pp. 160-6.
- Kasai, H., Allen, J.T., Mason, R.M., Kamimura, T. and Zhang, Z. (2005) 'TGF-beta1 induces human alveolar epithelial to mesenchymal cell transition (EMT)', *Respiratory research,* 6, pp. 56-56.
- Kasper, M. and Barth, K. (2017) 'Potential contribution of alveolar epithelial type I cells to pulmonary fibrosis', *Bioscience reports,* 37(6), pp. 10.1042/BSR20171301. Print 2017 Dec 22.

- Kasper, M. and Hughes, R.C. (1996) 'Immunocytochemical evidence for a modulation of galectin 3 (Mac-2), a carbohydrate binding protein, in pulmonary fibrosis', *The Journal of pathology,* 179(3), pp. 309-316.
- Katzen, J. and Beers, M.F. (2020) 'Contributions of alveolar epithelial cell quality control to pulmonary fibrosis', The Journal of clinical investigation, 130(10), pp. 5088-5099.
- Kaunisto, J., Salomaa, E.R., Hodgson, U., Kaarteenaho, R., Kankaanranta, H., Koli, K., Vahlberg, T. and Myllarniemi, M. (2019) 'Demographics and survival of patients with idiopathic pulmonary fibrosis in the FinnishIPF registry', *ERJ open research,* 5(3), pp. 10.1183/23120541.00170-2018. eCollection 2019 Jul.
- Kaur, A., Mathai, S.K. and Schwartz, D.A. (2017) 'Genetics in Idiopathic Pulmonary Fibrosis Pathogenesis, Prognosis, and Treatment', *Frontiers in medicine,* 4, pp. 154.
- Keane, M.P. (2008) 'The role of chemokines and cytokines in lung fibrosis', *European Respiratory review,* 17, pp 151-156.
- Keane, M.P. (2016) 'Galectin-3: Distant Biomarker or Relevant Target?', *American journal of respiratory and critical care medicine,* 194(1), pp. 7-9.
- Keane, T.J., Horejs, C.M. and Stevens, M.M. (2018) 'Scarring vs. functional healing: Matrix-based strategies to regulate tissue repair', *Advanced Drug Delivery Reviews,* 129, pp. 407-419.
- Kendall, R.T. and Feghali-Bostwick, C.A. (2014) 'Fibroblasts in fibrosis: novel roles and mediators', *Frontiers in pharmacology,* 5, pp. 123.
- Kenyon, N.J. and Last, J.A. (2005) 'Reversible and irreversible airway inflammation and fibrosis in mice exposed to inhaled ovalbumin', *Inflammation research : official journal of the European Histamine Research Society ...[et al.],* 54(2), pp. 57-65.
- Khalil, N. (1999) 'TGF-beta: from latent to active', *Microbes and Infection,* 1(15), pp. 1255-1263.
- Khalil, N., O'Conner, R. (2008) 'Transforming Growth Factor-β in Cancer Therapy', *Basic and Clinical Biology,* (1), pp 589-594.

- Kim, H., Liu, X., Kohyama, T., Kobayashi, T., Conner, H., Abe, S., Fang, Q., Wen, F.Q. and Rennard, S.I. (2004) 'Cigarette smoke stimulates MMP-1 production by human lung fibroblasts through the ERK1/2 pathway', *Copd,* 1(1), pp. 13-23.
- Kim, S., Lim, J.H. and Woo, C.H. (2020) 'Therapeutic potential of targeting kinase inhibition in patients with idiopathic pulmonary fibrosis', *Yeungnam University journal of medicine,* 37(4), pp. 269-276.
- Kim, S.M. and Hahn, J.H. (2008) 'CD98 activation increases surface expression and clusteringof beta1 integrins in MCF-7 cells through FAK/Src- and cytoskeleton-independent mechanisms', *Experimental & molecular medicine,* 40(3), pp. 261-270.
- Klingberg, F., Hinz, B. and White, E.S. (2013) 'The myofibroblast matrix: implications for tissue repair and fibrosis', *The Journal of pathology,* 229(2), pp. 298-309.
- Kolb, M., Bonniaud, P., Galt, T., Sime, P.J., Kelly, M.M., Margetts, P.J. and Gauldie, J. (2002) 'Differences in the fibrogenic response after transfer of active transforming growth factor-beta1 gene to lungs of "fibrosis-prone" and "fibrosis-resistant" mouse strains', *American journal of respiratory cell and molecular biology,* 27(2), pp. 141-150.
- Kolb, M., Margetts, P.J., Anthony, D.C., Pitossi, F. and Gauldie, J. (2001) 'Transient expression of IL-1beta induces acute lung injury and chronic repair leading to pulmonary fibrosis', *The Journal of clinical investigation,* 107(12), pp. 1529-1536.
- Kolesnikova, T.V., Mannion, B.A., Berditchevski, F. and Hemler, M.E. (2001a) 'Beta1 integrins show specific association with CD98 protein in low density membranes', *BMC biochemistry,* 2, pp. 10-10. Epub 2001 Oct 15.
- Konigshoff, M. and Rojas, M. (2012) 'Galectin-3: the bridge over troubled waters', *American journal of respiratory and critical care medicine,* 185(5), pp. 473-475.
- Kopitz, J., Vertesy, S., Andre, S., Fiedler, S., Schnolzer, M. and Gabius, H.J. (2014) 'Human chimera-type galectin-3: defining the critical tail length for high-affinity glycoprotein/cell surface binding and functional competition with galectin-1 in neuroblastoma cell growth regulation', *Biochimie,* 104, pp. 90-99.

- Kream, J., Mulay, S., Fukushima, D.K. and Solomon, S. (1983) 'Determination of plasma dexamethasone in the mother and the newborn after administration of the hormone in a clinical trial', The Journal of clinical endocrinology and metabolism, 56(1), pp. 127-133
- Kreuter, M., Bonella, F., Wijsenbeek, M., Maher, T.M. and Spagnolo, P. (2015) 'Pharmacological Treatment of Idiopathic Pulmonary Fibrosis: Current Approaches, Unsolved Issues, and Future Perspectives', *BioMed research international,* 2015, pp. 329481.
- Kreuter, M., Wuyts, W., Renzoni, E., Koschel, D., Maher, T.M., Kolb, M., Weycker, D., Spagnolo, P., Kirchgaessler, K.U., Herth, F.J. and Costabel, U. (2016) 'Antacid therapy and disease outcomes in idiopathic pulmonary fibrosis: a pooled analysis', *The Lancet.Respiratory medicine,* 4(5), pp. 381-389.
- Kropski, J.A., Blackwell, T.S. and Loyd, J.E. (2015) 'The genetic basis of idiopathic pulmonary fibrosis', *The European respiratory journal,* 45(6), pp. 1717-1727.
- Kulkarni, T., de Andrade, J., Zhou, Y., Luckhardt, T. and Thannickal, V.J. (2016) 'Alveolar epithelial disintegrity in pulmonary fibrosis', *American journal of physiology.Lung cellular and molecular physiology,* 311(2), pp. 185.
- Kumari, R. and Jat, P. (2021) 'Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype', Frontiers in cell and developmental biology, 9, pp. 645593.
- Laroui, H., Geem, D., Xiao, B., Viennois, E., Rakhya, P., Denning, T. and Merlin, D. (2014) 'Targeting intestinal inflammation with CD98 siRNA/PEI-loaded nanoparticles', *Molecular therapy : the journal of the American Society of Gene Therapy,* 22(1), pp. 69-80.
- Lasky, J.A. and Brody, A.R. (2000) 'Interstitial fibrosis and growth factors', *Environmental health perspectives,* 108 Suppl 4, pp. 751-762.
- LeBleu, V.S., Taduri, G., O'Connell, J., Teng, Y., Cooke, V.G., Woda, C., Sugimoto, H. and Kalluri, R. (2013) 'Origin and function of myofibroblasts in kidney fibrosis', *Nature medicine,* 19(8), pp. 1047-1053.
- Lederer, D.J., Arcasoy, S.M., Barr, R.G., Wilt, J.S., Bagiella, E., D'Ovidio, F., Sonett, J.R. and Kawut, S.M. (2006) 'Racial and ethnic disparities in idiopathic pulmonary fibrosis: A

UNOS/OPTN database analysis', American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons, 6(10), pp. 2436-2442.

- Lee, C.G., Cho, S.J., Kang, M.J., Chapoval, S.P., Lee, P.J., Noble, P.W., Yehualaeshet, T., Lu, B., Flavell, R.A., Milbrandt, J., Homer, R.J. and Elias, J.A. (2004) 'Early growth response gene 1-mediated apoptosis is essential for transforming growth factor beta1-induced pulmonary fibrosis', *The Journal of experimental medicine,* 200(3), pp. 377-389.
- Lee, H., Choi, H., Yang, B., Lee, S.K., Park, T.S., Park, D.W., Moon, J.Y., Kim, T.H., Sohn, J.W., Yoon, H.J. and Kim, S.H. (2021) 'Interstitial lung disease increases susceptibility to and severity of COVID-19', The European respiratory journal, .
- Lee, J.S., Collard, H.R., Anstrom, K.J., Martinez, F.J., Noth, I., Roberts, R.S., Yow, E., Raghu, G. and IPFnet Investigators (2013) 'Anti-acid treatment and disease progression in idiopathic pulmonary fibrosis: an analysis of data from three randomised controlled trials', *The Lancet.Respiratory medicine,* 1(5), pp. 369-376.
- Lee, J.S., Ryu, J.H., Elicker, B.M., Lydell, C.P., Jones, K.D., Wolters, P.J., King, T.E. and Collard, H.R. (2011) 'Gastroesophageal reflux therapy is associated with longer survival in patients with idiopathic pulmonary fibrosis', *American journal of respiratory and critical care medicine,* 184(12), pp. 1390-1394.
- Lepur, A., Salomonsson, E., Nilsson, U.J. and Leffler, H. (2012) 'Ligand induced galectin-3 protein self-association', *The Journal of biological chemistry,* 287(26), pp. 21751-21756.
- Leslie, K.O. (2009) 'My approach to interstitial lung disease using clinical, radiological and histopathological patterns', *Journal of clinical pathology,* 62(5), pp. 387-401.
- Ley, B. and Collard, H.R. (2013) 'Epidemiology of idiopathic pulmonary fibrosis', *Clinical epidemiology,* 5, pp. 483-492.
- Ley, B. and Collard, H.R. (2015) 'House of Cards? Testing Fundamental Assumptions in Idiopathic Pulmonary Fibrosis Epidemiology', *American journal of respiratory and critical care medicine,* 192(10), pp. 1147-1148.

- Ley, B., Brown, K.K. and Collard, H.R. (2014) 'Molecular biomarkers in idiopathic pulmonary fibrosis', *American journal of physiology.Lung cellular and molecular physiology,* 307(9), pp. 681.
- Li, L.C., Li, J. and Gao, J. (2014) 'Functions of galectin-3 and its role in fibrotic diseases', *The Journal of pharmacology and experimental therapeutics,* 351(2), pp. 336-343.
- Li, Y., Jiang, D., Liang, J., Meltzer, E.B., Gray, A., Miura, R., Wogensen, L., Yamaguchi, Y. and Noble, P.W. (2011) 'Severe lung fibrosis requires an invasive fibroblast phenotype regulated by hyaluronan and CD44', *The Journal of experimental medicine,* 208(7), pp. 1459-1471.
- Liang, C., Li, X., Zhang, L., Cui, D., Quan, X. and Yang, W. (2015) 'The anti-fibrotic effects of microRNA-153 by targeting TGFBR-2 in pulmonary fibrosis', *Experimental and molecular pathology,* 99(2), pp. 279-285.
- Liguori, I., Russo, G., Curcio, F., Bulli, G., Aran, L., Della-Morte, D., Gargiulo, G., Testa, G., Cacciatore, F., Bonaduce, D. and Abete, P. (2018) 'Oxidative stress, aging, and diseases', *Clinical interventions in aging,* 13, pp. 757-772.
- Lin, S., Wu, H., Wang, C., Xiao, Z. and Xu, F. (2018) 'Regulatory T Cells and Acute Lung Injury: Cytokines, Uncontrolled Inflammation, and Therapeutic Implications', *Frontiers in immunology,* 9, pp. 1545.
- Lin, Y.H., Qiu, D.C., Chang, W.H., Yeh, Y.Q., Jeng, U.S., Liu, F.T. and Huang, J.R. (2017) 'The intrinsically disordered N-terminal domain of galectin-3 dynamically mediates multisite self-association of the protein through fuzzy interactions', *The Journal of biological chemistry,* 292(43), pp. 17845-17856.
- Lindstedt, R., Apodaca, G., Barondes, S.H., Mostov, K.E. and Leffler, H. (1993) 'Apical secretion of a cytosolic protein by Madin-Darby canine kidney cells. Evidence for polarized release of an endogenous lectin by a nonclassical secretory pathway', *The Journal of biological chemistry,* 268(16), pp. 11750-11757.
- Lison, D., Lauwerys, R., Demedts, M. and Nemery, B. (1996) 'Experimental research into the pathogenesis of cobalt/hard metal lung disease', *The European respiratory journal,* 9(5), pp. 1024-1028.

- Liu, J.Y., Sime, P.J., Wu, T., Warshamana, G.S., Pociask, D., Tsai, S.Y. and Brody, A.R. (2001) 'Transforming growth factor-beta(1) overexpression in tumor necrosis factor-alpha receptor knockout mice induces fibroproliferative lung disease', *American journal of respiratory cell and molecular biology,* 25(1), pp. 3-7.
- Loimaranta, V., Hepojoki, J., Laaksoaho, O. and Pulliainen, A.T. (2018) 'Galectin-3-binding protein: A multitask glycoprotein with innate immunity functions in viral and bacterial infections', *Journal of leukocyte biology,* 104(4), pp. 777-786.
- Lomas, N.J., Watts, K.L., Akram, K.M., Forsyth, N.R. and Spiteri, M.A. (2012) 'Idiopathic pulmonary fibrosis: immunohistochemical analysis provides fresh insights into lung tissue remodelling with implications for novel prognostic markers', *International journal of clinical and experimental pathology,* 5(1), pp. 58-71.
- Lota, H.K. and Renzoni, E.A. (2012) 'Circulating biomarkers of interstitial lung disease in systemic sclerosis', *International journal of rheumatology,* 2012, pp. 121439.
- Luo, Y., Xu, W., Chen, H., Warburton, D., Dong, R., Qian, B., Selman, M., Gauldie, J., Kolb, M. and Shi, W. (2015) 'A novel profibrotic mechanism mediated by TGFbeta-stimulated collagen prolyl hydroxylase expression in fibrotic lung mesenchymal cells', *The Journal of pathology,* 236(3), pp. 384-394.
- Lygoe, K.A., Norman, J.T., Marshall, J.F. and Lewis, M.P. (2004) 'AlphaV integrins play an important role in myofibroblast differentiation', *Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society,* 12(4), pp. 461-470.
- Lyseng-Williamson, K.A. (2018) 'Pirfenidone tablets in idiopathic pulmonary fibrosis: a profile of their use', *Drugs & therapy perspectives : for rational drug selection and use,* 34(1), pp. 8-15.
- MacKinnon, A.C., Farnworth, S.L., Hodkinson, P.S., Henderson, N.C., Atkinson, K.M., Leffler, H., Nilsson, U.J., Haslett, C., Forbes, S.J. and Sethi, T. (2008) 'Regulation of alternative macrophage activation by galectin-3', *Journal of immunology (Baltimore, Md.: 1950),* 180(4), pp. 2650-2658.

- Madden, J.W. and Peacock, E.E. (1971) 'Studies on the biology of collagen during wound healing. 3. Dynamic metabolism of scar collagen and remodeling of dermal wounds', *Annals of Surgery,* 174(3), pp. 511-520.
- Maess, M.B., Wittig, B. and Lorkowski, S. (2014) 'Highly efficient transfection of human THP-1 macrophages by nucleofection', *Journal of visualized experiments : JoVE,* (91):e51960. doi(91), pp. e51960.
- Maher, T.M., Corte, T.J., Fischer, A., Kreuter, M., Lederer, D.J., Molina-Molina, M., Axmann, J., Kirchgaessler, K.U., Samara, K., Gilberg, F. and Cottin, V. (2020) 'Pirfenidone in patients with unclassifiable progressive fibrosing interstitial lung disease: a double-blind, randomised, placebo-controlled, phase 2 trial', *The Lancet.Respiratory medicine,* 8(2), pp. 147-157.
- Mahmoudi, S., Mancini, E., Moore, A., Xu, L., Jahanbani, F., Hebestreit, K., Srinivasan, R., Li, X., Devarajan, K., Prélot, L., Ang, C.E., Shibuya, Y., Benayoun, B.A., Chang, A.L.S., Wernig,, M., Wysocka, J., Longaker, M.T., Snyder, M.P., Brunet, A. (2018) 'Old fibroblasts secrete inflammatory cytokines that drive variability in reprogramming efficiency and may affect wound healing between old individuals', *Cold Spring Harbor Laboratory.*
- Mancano, A.D. (2018) 'Classification of idiopathic interstitial pneumonias', *Radiologia brasileira JID305000,* 51(5), pp. V-VI.
- Mangaonkar, A.A. and Patnaik, M.M. (2018) 'Short Telomere Syndromes in Clinical Practice: Bridging Bench and Bedside', *Mayo Clinic proceedings,* 93(7), pp. 904-916.
- Manzoni, F., Wallerstein, J., Schrader, T.E., Ostermann, A., Coates, L., Akke, M., Blakeley, M.P., Oksanen, E. and Logan, D.T. (2018) 'Elucidation of Hydrogen Bonding Patterns in Ligand-Free, Lactose- and Glycerol-Bound Galectin-3C by Neutron Crystallography to Guide Drug Design', *Journal of medicinal chemistry,* 61(10), pp. 4412-4420.
- Margadant, C., van den Bout, I., van Boxtel, A.L., Thijssen, V.L. and Sonnenberg, A. (2012) 'Epigenetic regulation of galectin-3 expression by beta1 integrins promotes cell adhesion and migration', *The Journal of biological chemistry,* 287(53), pp. 44684-44693
- Margaritopoulos, G.A., Trachalaki, A., Wells, A.U., Vasarmidi, E., Bibaki, E., Papastratigakis, G., Detorakis, S., Tzanakis, N. and Antoniou, K.M. (2018) 'Pirfenidone

improves survival in IPF: results from a real-life study', *BMC pulmonary medicine,* 18(1), pp. 177-z.

- Markowska, A.I., Jefferies, K.C. and Panjwani, N. (2011) 'Galectin-3 protein modulates cell surface expression and activation of vascular endothelial growth factor receptor 2 in human endothelial cells', *The Journal of biological chemistry,* 286(34), pp. 29913-29921.
- Marshall, R.P., Puddicombe, A., Cookson, W.O. and Laurent, G.J. (2000b) 'Adult familial cryptogenic fibrosing alveolitis in the United Kingdom', *Thorax,* 55(2), pp. 143-146.
- Martinez, F.O. and Gordon, S. (2014) 'The M1 and M2 paradigm of macrophage activation: time for reassessment', *F1000prime reports,* 6, pp. 13-13. eCollection 2014.
- Marzouk, K., Corate, L., Saleh, S. and Sharma, O.P. (2005) 'Epstein-Barr-virus-induced interstitial lung disease', *Current opinion in pulmonary medicine,* 11(5), pp. 456-460.
- Massa, S.M., Cooper, D.N., Leffler, H. and Barondes, S.H. (1993) 'L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity', *Biochemistry,* 32(1), pp. 260-267.
- Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M. and Taguchi, F. (2010) 'Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2', *Journal of virology,* 84(24), pp. 12658-12664.
- McCubbrey, A.L., Barthel, L., Mohning, M.P., Redente, E.F., Mould, K.J., Thomas, S.M., Leach, S.M., Danhorn, T., Gibbings, S.L., Jakubzick, C.V., Henson, P.M. and Janssen, W.J. (2018) 'Deletion of c-FLIP from CD11b(hi) Macrophages Prevents Development of Bleomycin-induced Lung Fibrosis', *American journal of respiratory cell and molecular biology,* 58(1), pp. 66-78.
- McGroder, C.F., Zhang, D., Choudhury, M.A., Salvatore, M.M., D'Souza, B.M., Hoffman, E.A., Wei, Y., Baldwin, M.R. and Garcia, C.K. (2021) 'Pulmonary fibrosis 4 months after COVID-19 is associated with severity of illness and blood leucocyte telomere length', Thorax, .
- Meliconi, R., Andreone, P., Fasano, L., Galli, S., Pacilli, A., Miniero, R., Fabbri, M., Solforosi, L. and Bernardi, M. (1996) 'Incidence of hepatitis C virus infection in Italian patients with idiopathic pulmonary fibrosis', *Thorax,* 51(3), pp. 315-317.

- Meltzer, E.B. and Noble, P.W. (2008) 'Idiopathic pulmonary fibrosis', *Orphanet journal of rare diseases,* 3, pp. 8.
- Menon, R.P. and Hughes, R.C. (1999) 'Determinants in the N-terminal domains of galectin-3 for secretion by a novel pathway circumventing the endoplasmic reticulum-Golgi complex', *European journal of biochemistry,* 264(2), pp. 569-576.
- Merlin, D., Sitaraman, S., Liu, X., Eastburn, K., Sun, J., Kucharzik, T., Lewis, B. and Madara, J.L. (2001) 'CD98-mediated links between amino acid transport and beta 1 integrin distribution in polarized columnar epithelia', *The Journal of biological chemistry,* 276(42), pp. 39282-39289.
- Midwood, K.S., Williams, L.V. and Schwarzbauer, J.E. (2004) 'Tissue repair and the dynamics of the extracellular matrix', *The international journal of biochemistry & cell biology,* 36(6), pp. 1031-1037.
- Miyake, Y., Sasaki, S., Yokoyama, T., Chida, K., Azuma, A., Suda, T., Kudoh, S., Sakamoto, N., Okamoto, K., Kobashi, G., Washio, M., Inaba, Y. and Tanaka, H. (2005) 'Occupational and environmental factors and idiopathic pulmonary fibrosis in Japan', *The Annals of Occupational Hygiene,* 49(3), pp. 259-265.
- Modenutti, C.P., Capurro, J.I.B., Di Lella, S. and Marti, M.A. (2019) 'The Structural Biology of Galectin-Ligand Recognition: Current Advances in Modeling Tools, Protein Engineering, and Inhibitor Design', *Frontiers in chemistry,* 7, pp. 823.
- Moeller, A., Ask, K., Warburton, D., Gauldie, J. and Kolb, M. (2008) 'The bleomycin animal model: a useful tool to investigate treatment options for idiopathic pulmonary fibrosis?', *The international journal of biochemistry & cell biology,* 40(3), pp. 362-382.
- Moeller, A., Ask, K., Warburton, D., Gauldie, J. and Kolb, M. (2008) 'The bleomycin animal model: a useful tool to investigate treatment options for idiopathic pulmonary fibrosis?', The international journal of biochemistry & cell biology, 40(3), pp. 362-382.
- Moiseenko, A., Kheirollahi, V., Chao, C.M., Ahmadvand, N., Quantius, J., Wilhelm, J., Herold, S., Ahlbrecht, K., Morty, R.E., Rizvanov, A.A., Minoo, P., El Agha, E. and Bellusci, S. (2017) 'Origin and characterization of alpha smooth muscle actin-positive cells during murine lung development', *Stem cells (Dayton, Ohio),* 35(6), pp. 1566-1578.

- Molina-Molina, M. (2019) 'Telomere Shortening Is behind the Harm of Immunosuppressive Therapy in Idiopathic Pulmonary Fibrosis', *American journal of respiratory and critical care medicine,* 200(3), pp. 274-275.
- Molyneaux, P.L., Cox, M.J., Willis-Owen, S.A., Mallia, P., Russell, K.E., Russell, A.M., Murphy, E., Johnston, S.L., Schwartz, D.A., Wells, A.U., Cookson, W.O., Maher, T.M. and Moffatt, M.F. (2014) 'The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis', *American journal of respiratory and critical care medicine,* 190(8), pp. 906-913.
- Mooney, J.J., Raimundo, K., Chang, E. and Broder, M.S. (2017) 'Mechanical ventilation in idiopathic pulmonary fibrosis: a nationwide analysis of ventilator use, outcomes, and resource burden', *BMC pulmonary medicine,* 17(1), pp. 84-2.
- Moore, B.B. and Moore, T.A. (2015) 'Viruses in Idiopathic Pulmonary Fibrosis. Etiology and Exacerbation', *Annals of the American Thoracic Society,* 12 Suppl 2, pp. 186.
- Moore, M.W. and Herzog, E.L. (2013) 'Regulation and Relevance of Myofibroblast Responses in Idiopathic Pulmonary Fibrosis', *Current pathobiology reports,* 1(3), pp. 199-208.
- Munger, J.S. and Sheppard, D. (2011) 'Cross talk among TGF-beta signaling pathways, integrins, and the extracellular matrix', *Cold Spring Harbor perspectives in biology,* 3(11), pp. a005017.
- Murray, L.A., Rosada, R., Moreira, A.P., Joshi, A., Kramer, M.S., Hesson, D.P., Argentieri, R.L., Mathai, S., Gulati, M., Herzog, E.L. and Hogaboam, C.M. (2010) 'Serum amyloid P therapeutically attenuates murine bleomycin-induced pulmonary fibrosis via its effects on macrophages', *PloS one,* 5(3), pp. e9683.
- Nagaraj, N.S. and Datta, P.K. (2010) 'Targeting the transforming growth factor-beta signaling pathway in human cancer', *Expert opinion on investigational drugs,* 19(1), pp. 77-91.
- Naik, P.K. and Moore, B.B. (2010) 'Viral infection and aging as cofactors for the development of pulmonary fibrosis', *Expert review of respiratory medicine,* 4(6), pp. 759-771.

- Nakano, Y., Yang, I.V., Walts, A.D., Watson, A.M., Helling, B.A., Fletcher, A.A., Lara, A.R., Schwarz, M.I., Evans, C.M. and Schwartz, D.A. (2016) 'MUC5B Promoter Variant rs35705950 Affects MUC5B Expression in the Distal Airways in Idiopathic Pulmonary Fibrosis', *American journal of respiratory and critical care medicine,* 193(4), pp. 464-466.
- National Health Services England. (2019.) Interstitial lung disease adult. [Online] Available at: https://www.england.nhs.uk/wp-content/uploads/2018/08/Interstitial-lung-disease-service-adult.pdf [Accessed 25th March. 2021].
- New England Biolabs. (2021). PNGase F. [Online] Available at: **https://international.neb.com/products/p0704-pngase-f#Product%20Information** [Accessed 7th May 2021].
- Newton, C.A., Batra, K., Torrealba, J., Kozlitina, J., Glazer, C.S., Aravena, C., Meyer, K., Raghu, G., Collard, H.R. and Garcia, C.K. (2016) 'Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive', The European respiratory journal, 48(6), pp. 1710-1720.
- Newton, C.A., Zhang, D., Oldham, J.M., Kozlitina, J., Ma, S.F., Martinez, F.J., Raghu, G., Noth, I. and Garcia, C.K. (2019) 'Telomere Length and Use of Immunosuppressive Medications in Idiopathic Pulmonary Fibrosis', *American journal of respiratory and critical care medicine,* 200(3), pp. 336-347.
- Nguyen, H.T. and Merlin, D. (2012a) 'Homeostatic and innate immune responses: role of the transmembrane glycoprotein CD98', *Cellular and molecular life sciences : CMLS,* 69(18), pp. 3015-3026.
- Nguyen, H.T. and Merlin, D. (2012b) 'Homeostatic and innate immune responses: role of the transmembrane glycoprotein CD98', *Cellular and molecular life sciences : CMLS,* 69(18), pp. 3015-3026.
- Nielsen, M.I., Stegmayr, J., Grant, O.C., Yang, Z., Nilsson, U.J., Boos, I., Carlsson, M.C., Woods, R.J., Unverzagt, C., Leffler, H. and Wandall, H.H. (2018) 'Galectin binding to cells and glycoproteins with genetically modified glycosylation reveals galectin-glycan specificities in a natural context', *The Journal of biological chemistry,* 293(52), pp. 20249-20262.

- Nieminen, J., Kuno, A., Hirabayashi, J. and Sato, S. (2007) 'Visualization of galectin-3 oligomerization on the surface of neutrophils and endothelial cells using fluorescence resonance energy transfer', *The Journal of biological chemistry,* 282(2), pp. 1374-1383.
- Nikfarjam, L. and Farzaneh, P. (2012) 'Prevention and detection of Mycoplasma contamination in cell culture', *Cell journal,* 13(4), pp. 203-212.
- Ning, W., Dong, Y., Sun, J., Li, C., Matthay, M.A., Feghali-Bostwick, C.A. and Choi, A.M. (2007) 'Cigarette smoke stimulates matrix metalloproteinase-2 activity via EGR-1 in human lung fibroblasts', *American journal of respiratory cell and molecular biology,* 36(4), pp. 480-490.
- Nishi, Y., Sano, H., Kawashima, T., Okada, T., Kuroda, T., Kikkawa, K., Kawashima, S., Tanabe, M., Goto, T., Matsuzawa, Y., Matsumura, R., Tomioka, H., Liu, F.T. and Shirai, K. (2007) 'Role of galectin-3 in human pulmonary fibrosis', *Allergology international : official journal of the Japanese Society of Allergology,* 56(1), pp. 57-65.
- Nobbmann, U., Connah, M., Fish, B., Gee, C., Mulot, S., Chen, J., Zhou, L., Lu, Y., Shen, F., Yi, J. and Harding, S.E. (2007) 'Dynamic light scattering as a relative tool for assessing the molecular integrity and stability of monoclonal antibodies', Biotechnology & genetic engineering reviews, 24, pp. 117-128.
- Noor, S., Nawaz, S. and Chaudhuri, N. (2021) 'Real-World Study Analysing Progression and Survival of Patients with Idiopathic Pulmonary Fibrosis with Preserved Lung Function on Antifibrotic Treatment', *Advances in Therapy,* 38(1), pp. 268-277.
- Nordgren, T.M. and Bailey, K.L. (2016) 'Pulmonary health effects of agriculture', *Current opinion in pulmonary medicine,* 22(2), pp. 144-149.
- Nova, Z., Skovierova, H. and Calkovska, A. (2019) 'Alveolar-Capillary Membrane-Related Pulmonary Cells as a Target in Endotoxin-Induced Acute Lung Injury', *International journal of molecular sciences,* 20(4), pp. 10.3390/ijms20040831.
- Nova, Z., Skovierova, H., Strnadel, J., Halasova, E. and Calkovska, A. (2020) 'Short-Term versus Long-Term Culture of A549 Cells for Evaluating the Effects of Lipopolysaccharide on Oxidative Stress, Surfactant Proteins and Cathelicidin LL-37', *International journal of molecular sciences,* 21(3), pp. 10.3390/ijms21031148.

- Nunes, I., Gleizes, P.E., Metz, C.N. and Rifkin, D.B. (1997) 'Latent transforming growth factor-beta binding protein domains involved in activation and transglutaminase-dependent cross-linking of latent transforming growth factor-beta', The Journal of cell biology, 136(5), pp. 1151-1163.
- O'Connor, J.W. and Gomez, E.W. (2013) 'Cell adhesion and shape regulate TGF-beta1-induced epithelial-myofibroblast transition via MRTF-A signaling', *PloS one,* 8(12), pp. e83188.
- Oda, D., Gown, A.M., Vande Berg, J.S. and Stern, R. (1990) 'Instability of the myofibroblast phenotype in culture', *Experimental and molecular pathology,* 52(2), pp. 221-234.
- Ogrunc, M. and d'Adda di Fagagna, F. (2011) 'Never-ageing cellular senescence', *European journal of cancer (Oxford, England : 1990),* 47(11), pp. 1616-1622.
- Oh, C.K., Murray, L.A. and Molfino, N.A. (2012) 'Smoking and idiopathic pulmonary fibrosis', *Pulmonary medicine,* 2012, pp. 808260.
- Okamura, D.M., Pasichnyk, K., Lopez-Guisa, J.M., Collins, S., Hsu, D.K., Liu, F.T. and Eddy, A.A. (2011) 'Galectin-3 preserves renal tubules and modulates extracellular matrix remodeling in progressive fibrosis', *American journal of physiology.Renal physiology,* 300(1), pp. 245.
- Orbiscope. (2021). SMALP. [Online] Available at: https://www.orbiscope.com/order-your-smalp/ [Accessed 4th May 2021].
- Orecchioni, M., Ghosheh, Y., Pramod, A.B. and Ley, K. (2019) 'Macrophage Polarization: Different Gene Signatures in M1(LPS+) vs. Classically and M2(LPS-) vs. Alternatively Activated Macrophages', *Frontiers in immunology,* 10, pp. 1084.
- Osterreicher, C.H., Penz-Osterreicher, M., Grivennikov, S.I., Guma, M., Koltsova, E.K., Datz, C.,Sasik, R., Hardiman, G., Karin, M. and Brenner, D.A. (2011) 'Fibroblast-specific protein 1 identifies an inflammatory subpopulation of macrophages in the liver', *Proceedings of the National Academy of Sciences of the United States of America,* 108(1), pp. 308-313.
- Panagiotou, M., Church, A.C., Johnson, M.K. and Peacock, A.J. (2017) 'Pulmonary vascular and cardiac impairment in interstitial lung disease', *European respiratory review : an*

official journal of the European Respiratory Society, 26(143), pp. 10.1183/16000617.0053-2016. Print 2017 Jan.

- Pardo, A., Ruiz, V., Arreola, J.L., Ramirez, R., Cisneros-Lira, J., Gaxiola, M., Barrios, R., Kala, S.V., Lieberman, M.W. and Selman, M. (2003) 'Bleomycin-induced pulmonary fibrosis is attenuated in gamma-glutamyl transpeptidase-deficient mice', *American journal of respiratory and critical care medicine,* 167(6), pp. 925-932.
- Parimon, T., Yao, C., Stripp, B.R., Noble, P.W. and Chen, P. (2020) 'Alveolar Epithelial Type II Cells as Drivers of Lung Fibrosis in Idiopathic Pulmonary Fibrosis', *International journal of molecular sciences,* 21(7), pp. 10.3390/ijms21072269.
- Park, S.W., Zhen, G., Verhaeghe, C., Nakagami, Y., Nguyenvu, L.T., Barczak, A.J., Killeen, N. and Erle, D.J. (2009) 'The protein disulfide isomerase AGR2 is essential for production of intestinal mucus', *Proceedings of the National Academy of Sciences of the United States of America,* 106(17), pp. 6950-6955.
- Parmar, M., Rawson, S., Scarff, C.A., Goldman, A., Dafforn, T.R., Muench, S.P. and Postis, V.L.G. (2018) 'Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure', *Biochimica et biophysica acta.Biomembranes,* 1860(2), pp. 378-383.
- Partridge, E.A., Le Roy, C., Di Guglielmo, G.M., Pawling, J., Cheung, P., Granovsky, M., Nabi, I.R., Wrana, J.L. and Dennis, J.W. (2004) 'Regulation of cytokine receptors by Golgi N-glycan processing and endocytosis', *Science (New York, N.Y.),* 306(5693), pp. 120-124.
- Pavelka, M., Roth, J. (2010) ‘Alveoli: Gas Exchange and Host Defesne. In: Functional Ultrastructure’, *Springer, Vienna,* pp 248-249
- Peng, R., Sridhar, S., Tyagi, G., Phillips, J.E., Garrido, R., Harris, P., Burns, L., Renteria, L., Woods, J., Chen, L., Allard, J., Ravindran, P., Bitter, H., Liang, Z., Hogaboam, C.M., Kitson, C., Budd, D.C., Fine, J.S., Bauer, C.M. and Stevenson, C.S. (2013) 'Bleomycin induces molecular changes directly relevant to idiopathic pulmonary fibrosis: a model for "active" disease', *PloS one,* 8(4), pp. e59348.
- Perng, D.W., Chang, K.T., Su, K.C., Wu, Y.C., Wu, M.T., Hsu, W.H., Tsai, C.M. and Lee, Y.C. (2007) 'Exposure of airway epithelium to bile acids associated with gastroesophageal

reflux symptoms: a relation to transforming growth factor-beta1 production and fibroblast proliferation', *Chest,* 132(5), pp. 1548-1556.

- Phan, S.H. (2008) 'Biology of fibroblasts and myofibroblasts', *Proceedings of the American Thoracic Society,* 5(3), pp. 334-337.
- Plantier, L., Cazes, A., Dinh-Xuan, A.T., Bancal, C., Marchand-Adam, S. and Crestani, B. (2018) 'Physiology of the lung in idiopathic pulmonary fibrosis', *European respiratory review : an official journal of the European Respiratory Society,* 27(147), pp. 10.1183/16000617.0062-2017. Print 2018 Mar 31.
- Pollock, N.L., Lee, S.C., Patel, J.H., Gulamhussein, A.A. and Rothnie, A.J. (2018) 'Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer', *Biochimica et biophysica acta.Biomembranes,* 1860(4), pp. 809-817.
- Postis, V., Rawson, S., Mitchell, J.K., Lee, S.C., Parslow, R.A., Dafforn, T.R., Baldwin, S.A. and Muench, S.P. (2015) 'The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy', *Biochimica et biophysica acta,* 1848(2), pp. 496-501.
- Prager, G.W., Feral, C.C., Kim, C., Han, J. and Ginsberg, M.H. (2007) 'CD98hc (SLC3A2) interaction with the integrin beta subunit cytoplasmic domain mediates adhesive signaling', *The Journal of biological chemistry,* 282(33), pp. 24477-24484.
- Priglinger, C.S., Szober, C.M., Priglinger, S.G., Merl, J., Euler, K.N., Kernt, M., Gondi, G., Behler, J., Geerlof, A., Kampik, A., Ueffing, M. and Hauck, S.M. (2013) 'Galectin-3 induces clustering of CD147 and integrin-beta1 transmembrane glycoprotein receptors on the RPE cell surface', *PloS one,* 8(7), pp. e70011.
- Pryor, W.A. and Stone, K. (Invalid date) 'Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite', *Annals of the New York Academy of Sciences,* 686, pp. 12-8.
- Raghu, G. and Meyer, K.C. (2012) 'Silent gastro-oesophageal reflux and microaspiration in IPF: mounting evidence for anti-reflux therapy?', *The European respiratory journal,* 39(2), pp. 242-245.

- Raghu, G., Crestani, B., Bailes, Z., Schlenker-Herce, R., Costabel, U. (2015) 'Effect of anti-acid medication on reduction in FVC decline with nintedanib' *EUROPEAN RESPIRATORY journal,* 46 (59).
- Raghu, G., Depaso, W.J., Cain, K., Hammar, S.P., Wetzel, C.E., Dreis, D.F., Hutchinson, J., Pardee, N.E. and Winterbauer, R.H. (1991) 'Azathioprine combined with prednisone in the treatment of idiopathic pulmonary fibrosis: a prospective double-blind, randomized, placebo-controlled clinical trial', *The American Review of Respiratory Disease,* 144(2), pp. 291-296.
- Raghu, G., Remy-Jardin, M., Myers, J.L., Richeldi, L., Ryerson, C.J., Lederer, D.J., Behr, J., Cottin, V., Danoff, S.K., Morell, F., Flaherty, K.R., Wells, A., Martinez, F.J., Azuma, A., Bice, T.J., Bouros, D., Brown, K.K., Collard, H.R., Duggal, A., Galvin, L., Inoue, Y., Jenkins, R.G., Johkoh, T., Kazerooni, E.A., Kitaichi, M., Knight, S.L., Mansour, G., Nicholson, A.G., Pipavath, S.N.J., Buendia-Roldan, I., Selman, M., Travis, W.D., Walsh, S., Wilson, K.C. and American Thoracic Society, European Respiratory Society (2018) 'Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline', *American journal of respiratory and critical care medicine,* 198(5), pp. e44-e68.
- Raghu. G., Nicholson, A.G., Lynch, D. (2008) 'The classification, natural history and radiological/histological appearance of idiopathic pulmonary fibrosis and the other idiopathic interstitial pneumonias', *EUROPEAN RESPIRATORY review,* (17), pp 108-115.
- Raimond, J., Zimonjic, D.B., Mignon, C., Mattei, M., Popescu, N.C., Monsigny, M. and Legrand, A. (1997) 'Mapping of the galectin-3 gene (LGALS3) to human chromosome 14 at region 14q21-22', *Mammalian genome : official journal of the International Mammalian Genome Society,* 8(9), pp. 706-707.
- Ramos, C., Montano, M., Garcia-Alvarez, J., Ruiz, V., Uhal, B.D., Selman, M. and Pardo, A. (2001) 'Fibroblasts from idiopathic pulmonary fibrosis and normal lungs differ in growth rate, apoptosis, and tissue inhibitor of metalloproteinases expression', *American journal of respiratory cell and molecular biology,* 24(5), pp. 591-598.
- Rangarajan, S., Kurundkar, A., Kurundkar, D., Bernard, K., Sanders, Y.Y., Ding, Q., Antony, V.B., Zhang, J., Zmijewski, J. and Thannickal, V.J. (2016) 'Novel Mechanisms for the

Antifibrotic Action of Nintedanib', American journal of respiratory cell and molecular biology, 54(1), pp. 51-59.

- RECOVERY Collaborative Group, Horby, P., Lim, W.S., Emberson, J.R., Mafham, M., Bell, J.L., Linsell, L., Staplin, N., Brightling, C., Ustianowski, A., Elmahi, E., Prudon, B., Green, C., Felton, T., Chadwick, D., Rege, K., Fegan, C., Chappell, L.C., Faust, S.N., Jaki, T., Jeffery, K., Montgomery, A., Rowan, K., Juszczak, E., Baillie, J.K., Haynes, R. and Landray, M.J. (2021) 'Dexamethasone in Hospitalized Patients with Covid-19', The New England journal of medicine, 384(8), pp. 693-704.
- REMAP-CAP Investigators, Gordon, A.C., Mouncey, P.R., Al-Beidh, F., Rowan, K.M., Nichol, A.D., Arabi, Y.M., Annane, D., Beane, A., van Bentum-Puijk, W., Berry, L.R., Bhimani, Z., Bonten, M.J.M., Bradbury, C.A., Brunkhorst, F.M., Buzgau, A., Cheng, A.C., Detry, M.A., Duffy, E.J., Estcourt, L.J., Fitzgerald, M., Goossens, H., Haniffa, R., Higgins, A.M., Hills, T.E., Horvat, C.M., Lamontagne, F., Lawler, P.R., Leavis, H.L., Linstrum, K.M., Litton, E., Lorenzi, E., Marshall, J.C., Mayr, F.B., McAuley, D.F., McGlothlin, A., McGuinness, S.P., McVerry, B.J., Montgomery, S.K., Morpeth, S.C., Murthy, S., Orr, K., Parke, R.L., Parker, J.C., Patanwala, A.E., Pettila, V., Rademaker, E., Santos, M.S., Saunders, C.T., Seymour, C.W., Shankar-Hari, M., Sligl, W.I., Turgeon, A.F., Turner, A.M., van de Veerdonk, F L, Zarychanski, R., Green, C., Lewis, R.J., Angus, D.C., McArthur, C.J., Berry, S., Webb, S.A. and Derde, L.P.G. (2021) 'Interleukin-6 Receptor Antagonists in Critically Ill Patients with Covid-19', *The New England journal of medicine,* 384(16), pp. 1491-1502.
- Ren, X., Glende, J., Al-Falah, M., de Vries, V., Schwegmann-Wessels, C., Qu, X., Tan, L., Tschernig, T., Deng, H., Naim, H.Y. and Herrler, G. (2006) 'Analysis of ACE2 in polarized epithelial cells: surface expression and function as receptor for severe acute respiratory syndrome-associated coronavirus', *The Journal of general virology,* 87(Pt 6), pp. 1691-1695.
- Riario Sforza, G.G. and Marinou, A. (2017) 'Hypersensitivity pneumonitis: a complex lung disease', *Clinical and molecular allergy : CMA,* 15, pp. 6-7. eCollection 2017.
- Riario Sforza, G.G. and Marinou, A. (2017) 'Hypersensitivity pneumonitis: a complex lung disease', *Clinical and molecular allergy : CMA,* 15, pp. 6-7. eCollection 2017.

- Rintoul, R.C., Buttery, R.C., Mackinnon, A.C., Wong, W.S., Mosher, D., Haslett, C. and Sethi, T. (2002) 'Cross-linking CD98 promotes integrin-like signaling and anchorage-independent growth', *Molecular biology of the cell,* 13(8), pp. 2841-2852.
- Rittirsch, D., Flierl, M.A., Day, D.E., Nadeau, B.A., McGuire, S.R., Hoesel, L.M., Ipaktchi, K., Zetoune, F.S., Sarma, J.V., Leng, L., Huber-Lang, M.S., Neff, T.A., Bucala, R. and Ward, P.A. (2008) 'Acute lung injury induced by lipopolysaccharide is independent of complement activation', *Journal of immunology (Baltimore, Md.: 1950),* 180(11), pp. 7664-7672.
- Rivera-Ortega, P. and Molina-Molina, M. (2019) 'Interstitial Lung Diseases in Developing Countries', *Annals of global health,* 85(1), pp. 10.5334/aogh.2414.
- Roach, K.M., Duffy, S.M., Coward, W., Feghali-Bostwick, C., Wulff, H. and Bradding, P. (2013a) 'The K+ channel KCa3.1 as a novel target for idiopathic pulmonary fibrosis', *PloS one,* 8(12), pp. e85244.
- Roach, K.M., Feghali-Bostwick, C., Wulff, H., Amrani, Y. and Bradding, P. (2015) 'Human lung myofibroblast TGFbeta1-dependent Smad2/3 signalling is Ca(2+)-dependent and regulated by KCa3.1 K(+) channels', *Fibrogenesis & tissue repair,* 8, pp. 5-0. eCollection 2015.
- Roach, K.M., Sutcliffe, A., Matthews, L., Elliott, G., Newby, C., Amrani, Y. and Bradding, P. (2018) 'A model of human lung fibrogenesis for the assessment of anti-fibrotic strategies in idiopathic pulmonary fibrosis', *Scientific reports,* 8(1), pp. 342-9.
- Roach, K.M., Wulff, H., Feghali-Bostwick, C., Amrani, Y. and Bradding, P. (2014b) 'Increased constitutive alphaSMA and Smad2/3 expression in idiopathic pulmonary fibrosis myofibroblasts is KCa3.1-dependent', *Respiratory research,* 15, pp. 155-5.
- Robertson, I.B., Horiguchi, M., Zilberberg, L., Dabovic, B., Hadjiolova, K. and Rifkin, D.B. (2015) 'Latent TGF-beta-binding proteins', Matrix biology : journal of the International Society for Matrix Biology, 47, pp. 44-53.
- Roh, C. and Lyle, S. (2006) 'Cutaneous stem cells and wound healing', *Pediatric research,* 59(4 Pt 2), pp. 100R-3R.

- Rojas-Serrano, J., Gonzalez-Velasquez, E., Mejia, M., Sanchez-Rodriguez, A. and Carrillo, G. (2012) 'Interstitial lung disease related to rheumatoid arthritis: evolution after treatment', *Reumatologia clinica,* 8(2), pp. 68-71.
- Rumchev, K., Gilbey, S., Mead-Hunter, R., Selvey, L., Netto, K. and Mullins, B. (2019) 'Agricultural Dust Exposures and Health and Safety Practices among Western Australian Wheatbelt Farmers during Harvest', *International journal of environmental research and public health,* 16(24), pp. 10.3390/ijerph16245009.
- Ryerson, C.J., Cottin, V., Brown, K.K. and Collard, H.R. (2015) 'Acute exacerbation of idiopathic pulmonary fibrosis: shifting the paradigm', *The European respiratory journal,* 46(2), pp. 512-520.
- Saito, A., Horie, M. and Nagase, T. (2018) 'TGF-beta Signaling in Lung Health and Disease', *International journal of molecular sciences,* 19(8), pp. 10.3390/ijms19082460.
- Sakai, N. and Tager, A.M. (2013) 'Fibrosis of two: Epithelial cell-fibroblast interactions in pulmonary fibrosis', *Biochimica et biophysica acta,* 1832(7), pp. 911-921.
- Salisbury, M.L., Han, M.K., Dickson, R.P. and Molyneaux, P.L. (2017) 'Microbiome in interstitial lung disease: from pathogenesis to treatment target', *Current opinion in pulmonary medicine,* 23(5), pp. 404-410.
- Santos, S.N., Junqueira, M.S., Francisco, G., Vilanova, M., Magalhaes, A., Dias Baruffi, M., Chammas, R., Harris, A.L., Reis, C.A. and Bernardes, E.S. (2016) 'O-glycan sialylation alters galectin-3 subcellular localization and decreases chemotherapy sensitivity in gastric cancer', *Oncotarget,* 7(50), pp. 83570-83587.
- Saraboji, K., Hakansson, M., Genheden, S., Diehl, C., Qvist, J., Weininger, U., Nilsson, U.J., Leffler, H., Ryde, U., Akke, M. and Logan, D.T. (2012) 'The carbohydrate-binding site in galectin-3 is preorganized to recognize a sugarlike framework of oxygens: ultra-high-resolution structures and water dynamics', *Biochemistry,* 51(1), pp. 296-306.
- Scadding, J.G. (1974) 'Diffuse pulmonary alveolar fibrosis', *Thorax,* 29(3), pp. 271-281.
- Schafer, M.J., White, T.A., Iijima, K., Haak, A.J., Ligresti, G., Atkinson, E.J., Oberg, A.L., Birch, J., Salmonowicz, H., Zhu, Y., Mazula, D.L., Brooks, R.W., Fuhrmann-Stroissnigg, H., Pirtskhalava, T., Prakash, Y.S., Tchkonia, T., Robbins, P.D., Aubry, M.C., Passos, J.F.,

Kirkland, J.L., Tschumperlin, D.J., Kita, H. and LeBrasseur, N.K. (2017) 'Cellular senescence mediates fibrotic pulmonary disease', Nature communications, 8, pp. 14532.

- Scheffler, S., Dieken, H., Krischenowski, O. and Aufderheide, M. (2015) 'Cytotoxic Evaluation of e-Liquid Aerosol using Different Lung-Derived Cell Models', International journal of environmental research and public health, 12(10), pp. 12466-12474.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012) 'Fiji: an open-source platform for biological-image analysis', *Nature methods,* 9(7), pp. 676-682.
- Schnittert, J., Bansal, R., Storm, G. and Prakash, J. (2018) 'Integrins in wound healing, fibrosis and tumor stroma: High potential targets for therapeutics and drug delivery', *Advanced Drug Delivery Reviews,* 129, pp. 37-53.
- Sciacchitano, S., Lavra, L., Morgante, A., Ulivieri, A., Magi, F., De Francesco, G.P., Bellotti, C., Salehi, L.B. and Ricci, A. (2018) 'Galectin-3: One Molecule for an Alphabet of Diseases, from A to Z', *International journal of molecular sciences,* 19(2), pp. 10.3390/ijms19020379.
- Scotton, C.J., Hayes, B., Alexander, R., Datta, A., Forty, E.J., Mercer, P.F., Blanchard, A. and Chambers, R.C. (2013) 'Ex vivo micro-computed tomography analysis of bleomycin-induced lung fibrosis for preclinical drug evaluation', The European respiratory journal, 42(6), pp. 1633-1645.
- Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S.H. and Rini, J.M. (1998) 'X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-A resolution', *The Journal of biological chemistry,* 273(21), pp. 13047-13052.
- Seibold, M.A., Wise, A.L., Speer, M.C., Steele, M.P., Brown, K.K., Loyd, J.E., Fingerlin, T.E., Zhang, W., Gudmundsson, G., Groshong, S.D., Evans, C.M., Garantziotis, S., Adler, K.B., Dickey, B.F., du Bois, R.M., Yang, I.V., Herron, A., Kervitsky, D., Talbert, J.L., Markin, C., Park, J., Crews, A.L., Slifer, S.H., Auerbach, S., Roy, M.G., Lin, J., Hennessy, C.E., Schwarz, M.I. and Schwartz, D.A. (2011) 'A common MUC5B promoter polymorphism and pulmonary fibrosis', *The New England journal of medicine,* 364(16), pp. 1503-1512.

- Selman, M. and Pardo, A. (2002) 'Idiopathic pulmonary fibrosis: an epithelial/fibroblastic cross-talk disorder', *Respiratory research,* 3, pp. 3.
- Selman, M., Pardo, A. and King, T.E. (2012) 'Hypersensitivity pneumonitis: insights in diagnosis and pathobiology', *American journal of respiratory and critical care medicine,* 186(4), pp. 314-324.
- Semenzato, G., Adami, F., Maschio, N. and Agostini, C. (2000) 'Immune mechanisms in interstitial lung diseases', *Allergy,* 55(12), pp. 1103-1120.
- Shah, M., Foreman, D.M. and Ferguson, M.W. (1995) 'Neutralisation of TGF-beta 1 and TGF-beta 2 or exogenous addition of TGF-beta 3 to cutaneous rat wounds reduces scarring', *Journal of cell science,* 108 (Pt 3)(Pt 3), pp. 985-1002.
- Sheppard, D. (2001) 'Pulmonary fibrosis: a cellular overreaction or a failure of communication?', *The Journal of clinical investigation,* 107(12), pp. 1501-1502
- Sheppard, D. (2003) 'Functions of pulmonary epithelial integrins: from development to disease', *Physiological Reviews,* 83(3), pp. 673-686.
- Shi, M., Zhu, J., Wang, R., Chen, X., Mi, L., Walz, T. and Springer, T.A. (2011) 'Latent TGF-beta structure and activation', *Nature,* 474(7351), pp. 343-349.
- Shilova, A., Lebrette, H., Aurelius, O., Nan, J., Welin, M., Kovacic, R., Ghosh, S., Safari, C., Friel, R.J., Milas, M., Matej, Z., Hogbom, M., Branden, G., Kloos, M., Shoeman, R.L., Doak, B., Ursby, T., Hakansson, M., Logan, D.T. and Mueller, U. (2020) 'Current status and future opportunities for serial crystallography at MAX IV Laboratory', *Journal of synchrotron radiation,* 27(Pt 5), pp. 1095-1102.
- Shulgina, L., Cahn, A.P., Chilvers, E.R., Parfrey, H., Clark, A.B., Wilson, E.C., Twentyman, O.P., Davison, A.G., Curtin, J.J., Crawford, M.B. and Wilson, A.M. (2013) 'Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: a randomised controlled trial', *Thorax,* 68(2), pp. 155-162.
- Sigrist, C.J. and Le Mercier, P. (2020) 'A potential role for integrins in host cell entry by SARS-CoV-2', *Antiviral Research,* 177, pp. 104759.
- Simovic Markovic, B., Nikolic, A., Gazdic, M., Bojic, S., Vucicevic, L., Kosic, M., Mitrovic, S., Milosavljevic, M., Besra, G., Trajkovic, V., Arsenijevic, N., Lukic, M.L. and Volarevic, V.

(2016) 'Galectin-3 Plays an Important Pro-inflammatory Role in the Induction Phase of Acute Colitis by Promoting Activation of NLRP3 Inflammasome and Production of IL-1beta in Macrophages', *Journal of Crohn's & colitis,* 10(5), pp. 593-606.

- Singer, A.J. and Clark, R.A. (1999) 'Cutaneous wound healing', *The New England journal of medicine,* 341(10), pp. 738-746.
- Singh, S.R. and Hall, I.P. (2008) 'Airway myofibroblasts and their relationship with airway myocytes and fibroblasts', *Proceedings of the American Thoracic Society,* 5(1), pp. 127-132.
- Smith, A.J., Wright, K.E., Muench, S.P., Schumann, S., Whitehouse, A., Porter, K.E. and Colyer, J. (2019) 'Styrene maleic acid recovers proteins from mammalian cells and tissues while avoiding significant cell death', *Scientific reports,* 9(1), pp. 16408-1.
- Smith, P.G., Roy, C., Zhang, Y.N. and Chauduri, S. (2003) 'Mechanical stress increases RhoA activation in airway smooth muscle cells', *American journal of respiratory cell and molecular biology,* 28(4), pp. 436-442.
- Son, J.Y., Kim, S.Y., Cho, S.H., Shim, H.S., Jung, J.Y., Kim, E.Y., Lim, J.E., Park, B.H., Kang, Y.A., Kim, Y.S., Kim, S.K., Chang, J. and Park, M.S. (2013) 'TGF-beta1 T869C polymorphism may affect susceptibility to idiopathic pulmonary fibrosis and disease severity', *Lung,* 191(2), pp. 199-205.
- Song, X., Yu, W. and Guo, F. (2018) 'Pirfenidone suppresses bleomycin-induced pulmonary fibrosis and periostin expression in rats', *Experimental and therapeutic medicine,* 16(3), pp. 1800-1806.
- Sorme, P., Arnoux, P., Kahl-Knutsson, B., Leffler, H., Rini, J.M. and Nilsson, U.J. (2005) 'Structural and thermodynamic studies on cation-Pi interactions in lectin-ligand complexes: high-affinity galectin-3 inhibitors through fine-tuning of an arginine-arene interaction', *Journal of the American Chemical Society,* 127(6), pp. 1737-1743.
- Sousa, A.M., Liu, T., Guevara, O., Stevens, J., Fanburg, B.L., Gaestel, M., Toksoz, D. and Kayyali, U.S. (2007) 'Smooth muscle alpha-actin expression and myofibroblast differentiation by TGFbeta are dependent upon MK2', *Journal of cellular biochemistry,* 100(6), pp. 1581-1592.

- Sousa, S.R., Caetano Mota, P., Melo, N., Bastos, H.N., Padrao, E., Pereira, J.M., Cunha, R., Souto Moura, C., Guimaraes, S. and Morais, A. (2018) 'Heterozygous TERT gene mutation associated with familial idiopathic pulmonary fibrosis', *Respiratory medicine case reports,* 26, pp. 118-122.
- Spagnolo, P., Balestro, E., Aliberti, S., Cocconcelli, E., Biondini, D., Casa, G.D., Sverzellati, N. and Maher, T.M. (2020) 'Pulmonary fibrosis secondary to COVID-19: a call to arms?', *The Lancet.Respiratory medicine,* 8(8), pp. 750-752.
- Stanford Medicine. (2021). 'Understanding ILD'. [Online] Available at: http://med.stanford.edu/ild/patient-resources/understanding-ild.html#:~:text=The%20interstitium%20refers%20to%20the,a%20thin%20sheet%20of%20blood [Accessed 21st March 2021].
- Stoll, C., Mengsteab, S., Stoll, D., Riediger, D., Gressner, A.M. and Weiskirchen, R. (2004) 'Analysis of polymorphic TGFB1 codons 10, 25, and 263 in a German patient group with non-syndromic cleft lip, alveolus, and palate compared with healthy adults', *BMC medical genetics,* 5, pp. 15-15.
- Strieter, R.M. (2008) 'What differentiates normal lung repair and fibrosis? Inflammation, resolution of repair, and fibrosis', *Proceedings of the American Thoracic Society,* 5(3), pp. 305-310.
- Sun, K.H., Chang, Y., Reed, N.I. and Sheppard, D. (2016) 'alpha-Smooth muscle actin is an inconsistent marker of fibroblasts responsible for force-dependent TGFbeta activation or collagen production across multiple models of organ fibrosis', *American journal of physiology.Lung cellular and molecular physiology,* 310(9), pp. 824.
- Sundarakrishnan, A., Chen, Y., Black, L.D., Aldridge, B.B. and Kaplan, D.L. (2018) 'Engineered cell and tissue models of pulmonary fibrosis', *Advanced Drug Delivery Reviews,* 129, pp. 78-94.
- Sverzellati, N. (2013) 'Highlights of HRCT imaging in IPF', *Respiratory research,* 14 Suppl 1, pp. S3-S3. Epub 2013 Apr 16.

- Swain, R.J., Kemp, S.J., Goldstraw, P., Tetley, T.D. and Stevens, M.M. (2010) 'Assessment of cell line models of primary human cells by Raman spectral phenotyping', *Biophysical journal,* 98(8), pp. 1703-1711.
- Swaroopa, D., Bhaskar, K., Mahathi, T., Katkam, S., Raju, Y.S., Chandra, N. and Kutala, V.K. (2016) 'Association of serum interleukin-6, interleukin-8, and Acute Physiology and Chronic Health Evaluation II score with clinical outcome in patients with acute respiratory distress syndrome', *Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine,* 20(9), pp. 518-525.
- Sweeney, D.J., Khor, Y.H. and Goh, N.S.L. (2020) 'The unmet care needs of progressive fibrosing interstitial lung disease', *Respirology (Carlton, Vic.),* 25(12), pp. 1231-1232.
- Swigris, J.J., Olson, A.L., Huie, T.J., Fernandez-Perez, E.R., Solomon, J., Sprunger, D. and Brown, K.K. (2012) 'Ethnic and racial differences in the presence of idiopathic pulmonary fibrosis at death', *Respiratory medicine,* 106(4), pp. 588-593.
- Takenaka, Y., Fukumori, T., Yoshii, T., Oka, N., Inohara, H., Kim, H.R., Bresalier, R.S. and Raz, A. (2004) 'Nuclear export of phosphorylated galectin-3 regulates its antiapoptotic activity in response to chemotherapeutic drugs', *Molecular and cellular biology,* 24(10), pp. 4395-4406.
- Taniguchi, T., Asano, Y., Akamata, K., Noda, S., Masui, Y., Yamada, D., Takahashi, T., Ichimura, Y., Toyama, T., Tamaki, Z., Tada, Y., Sugaya, M., Kadono, T. and Sato, S. (2012) 'Serum levels of galectin-3: possible association with fibrosis, aberrant angiogenesis, and immune activation in patients with systemic sclerosis', *The Journal of rheumatology,* 39(3), pp. 539-544.
- Teixeira, S. and Kuhn, L.C. (1991) 'Post-transcriptional regulation of the transferrin receptor and 4F2 antigen heavy chain mRNA during growth activation of spleen cells', *European journal of biochemistry,* 202(3), pp. 819-826.
- Tejler, J., Leffler, H. and Nilsson, U.J. (2005) 'Synthesis of O-galactosyl aldoximes as potent LacNAc-mimetic galectin-3 inhibitors', *Bioorganic & medicinal chemistry letters,* 15(9), pp. 2343-2345.

- The Human Protein Atlas. (2021). ACE2. [Online] Available at: **https://www.proteinatlas.org/ENSG00000130234-ACE2** [Accessed 21st April 2021].
- ThermoFisher Scientific. (2021). Transfecting Plasmid DNA into THP-1 Cells Using Lipofectamine LTX Reagent. [Online] Available at: https://www.thermofisher.com/uk/en/home/references/protocols/cell-culture/transfection-protocol/transfecting-plasmid-dna-into-thp-1-cells-using-lipofectamine-ltx-reagent.html [Accessed 11th April 2021].
- Thickett, D.R., Poole, A.R. and Millar, A.B. (2001) 'The balance between collagen synthesis and degradation in diffuse lung disease', *Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG,* 18(1), pp. 27-33.
- Thornton, C., Maher, T.M., Hansell, D., Nicholson, A.G. and Wells, A.U. (2009) 'Pulmonary fibrosis associated with psychotropic drug therapy: a case report', *Journal of medical case reports,* 3, pp. 126.
- Tominaga, K. and Suzuki, H.I. (2019) 'TGF-beta Signaling in Cellular Senescence and Aging-Related Pathology', International journal of molecular sciences, 20(20), pp. 10.3390/ijms20205002.
- Travis, W.D., Costabel, U., Hansell, D.M., King, T.E., Lynch, D.A., Nicholson, A.G., Ryerson, C.J., Ryu, J.H., Selman, M., Wells, A.U., Behr, J., Bouros, D., Brown, K.K., Colby, T.V., Collard, H.R., Cordeiro, C.R., Cottin, V., Crestani, B., Drent, M., Dudden, R.F., Egan, J., Flaherty, K., Hogaboam, C., Inoue, Y., Johkoh, T., Kim, D.S., Kitaichi, M., Loyd, J., Martinez, F.J., Myers, J., Protzko, S., Raghu, G., Richeldi, L., Sverzellati, N., Swigris, J., Valeyre, D. and ATS/ERS Committee on Idiopathic Interstitial Pneumonias (2013) 'An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias', *American journal of respiratory and critical care medicine,* 188(6), pp. 733-748.
- Tschumperlin, D.J., Ligresti, G., Hilscher, M.B. and Shah, V.H. (2018) 'Mechanosensing and fibrosis', *The Journal of clinical investigation,* 128(1), pp. 74-84.
- Tzouvelekis, A., Kouliatsis, G., Anevlavis, S. and Bouros, D. (2005) 'Serum biomarkers in interstitial lung diseases', *Respiratory research,* 6, pp. 78-78.

- Ueda, T., Ohta, K., Suzuki, N., Yamaguchi, M., Hirai, K., Horiuchi, T., Watanabe, J., Miyamoto, T. and Ito, K. (1992) 'Idiopathic pulmonary fibrosis and high prevalence of serum antibodies to hepatitis C virus', *The American Review of Respiratory Disease,* 146(1), pp. 266-268.
- Umemoto, K. and Leffler, H. (2001) 'Assignment of 1H, 15N and 13C resonances of the carbohydrate recognition domain of human galectin-3', *Journal of Biomolecular NMR,* 20(1), pp. 91-92.
- Ushiki, A., Yamazaki, Y., Hama, M., Yasuo, M., Hanaoka, M. and Kubo, K. (2014) 'Viral infections in patients with an acute exacerbation of idiopathic interstitial pneumonia', *Respiratory investigation,* 52(1), pp. 65-70.
- Van Linthout, S., Miteva, K. and Tschope, C. (2014) 'Crosstalk between fibroblasts and inflammatory cells', *Cardiovascular research,* 102(2), pp. 258-269.
- Varki, A., Cummings*, R.D.,* Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., Schnaar, R.L., Seeberger*, P.H. (2017) 'Essentials of Glycobiology, Second Edition', Cold Spring Harbor Laboratory Press, Chapter 33.*
- Verrecchia, F. and Mauviel, A. (2002) 'Transforming growth factor-beta signaling through the Smad pathway: role in extracellular matrix gene expression and regulation', *The Journal of investigative dermatology,* 118(2), pp. 211-215.
- Vij, R., Noth, I. and Strek, M.E. (2011) 'Autoimmune-featured interstitial lung disease: a distinct entity', *Chest,* 140(5), pp. 1292-1299.
- Vogelmeier, C. (2003) 'IPF or NSIP? That is the question', *European Respiratory journal',* 22, pp 191-192.
- Wallace, W.A., Fitch, P.M., Simpson, A.J. and Howie, S.E. (2007) 'Inflammation-associated remodelling and fibrosis in the lung - a process and an end point', *International journal of experimental pathology,* 88(2), pp. 103-110.
- Walsh, S.L., Sverzellati, N., Devaraj, A., Keir, G.J., Wells, A.U. and Hansell, D.M. (2014) 'Connective tissue disease related fibrotic lung disease: high resolution computed

tomographic and pulmonary function indices as prognostic determinants', *Thorax,* 69(3), pp. 216-222.

- Walsh, S.L.F. and Kolb, M. (2018) 'Radiological diagnosis of interstitial lung disease: is it all about pattern recognition?', *The European respiratory journal,* 52(2), pp. 10.1183/13993003.01321-2018. Print 2018 Aug.
- Walter, N., Collard, H.R. and King, T.E. (2006) 'Current perspectives on the treatment of idiopathic pulmonary fibrosis', *Proceedings of the American Thoracic Society,* 3(4), pp. 330-338.
- Wang, R., Li, Q. and Tang, D.D. (2006) 'Role of vimentin in smooth muscle force development', *American journal of physiology.Cell physiology,* 291(3), pp. 483.
- Ward, H.E. and Nicholas, T.E. (1984) 'Alveolar type I and type II cells', *Australian and New Zealand Journal of Medicine,* 14(5 Suppl 3), pp. 731-734.
- Waters, D.W., Blokland, K.E.C., Pathinayake, P.S., Burgess, J.K., Mutsaers, S.E., Prele, C.M., Schuliga, M., Grainge, C.L. and Knight, D.A. (2018) 'Fibroblast senescence in the pathology of idiopathic pulmonary fibrosis', *American journal of physiology.Lung cellular and molecular physiology,* 315(2), pp. L162-L172.
- Watts, K.L., Cottrell, E., Hoban, P.R. and Spiteri, M.A. (2006) 'RhoA signaling modulates cyclin D1 expression in human lung fibroblasts; implications for idiopathic pulmonary fibrosis', *Respiratory research,* 7, pp. 88-88.
- Wells, A.U., Brown, K.K., Flaherty, K.R., Kolb, M., Thannickal, V.J., IPF Consensus Working Group and IPF Consensus Working Group: (2019) 'Pulmonary fibrosis: "idiopathic" is not "cryptogenic"', *The European respiratory journal,* 53(3), pp. 10.1183/13993003.00400-2019. Print 2019 Mar.
- Wells, A.U., Hirani, N. (2008) 'Interstitial lung disease guideline', *Thorax,* (63), pp 1-58.
- Weng, C.M., Li, Q., Chen, K.J., Xu, C.X., Deng, M.S., Li, T., Zhang, D.D., Duan, Z.X., Chen, Z.Q., Li, G.H., Chen, J. and Wang, J.M. (2020) 'Bleomycin induces epithelial-to-mesenchymal transition via bFGF/PI3K/ESRP1 signaling in pulmonary fibrosis', *Bioscience reports,* 40(1), pp. 10.1042/BSR20190756.

- West, J.B. (2009) 'Comparative physiology of the pulmonary blood-gas barrier: the unique avian solution', *American journal of physiology.Regulatory, integrative and comparative physiology,* 297(6), pp. 1625.
- White, E.S. (2015) 'Lung extracellular matrix and fibroblast function', *Annals of the American Thoracic Society,* 12 Suppl 1, pp. 30.
- White, E.S. and Mantovani, A.R. (2013) 'Inflammation, wound repair, and fibrosis: reassessing the spectrum of tissue injury and resolution', *The Journal of pathology,* 229(2), pp. 141-144.
- Williams, T.J. and Wilson, J.W. (2008) 'Challenges in pulmonary fibrosis: 7--Novel therapies and lung transplantation', *Thorax,* 63(3), pp. 277-284.
- Wilson, A.M., Clark, A.B., Cahn, T., Chilvers, E.R., Fraser, W., Hammond, M., Livermore, D.M., Maher, T.M., Parfrey, H., Swart, A.M., Stirling, S., Thickett, D.R., Whyte, M. and EME-TIPAC team (2020) 'Effect of Co-trimoxazole (Trimethoprim-Sulfamethoxazole) vs Placebo on Death, Lung Transplant, or Hospital Admission in Patients With Moderate and Severe Idiopathic Pulmonary Fibrosis: The EME-TIPAC Randomized Clinical Trial', *Jama,* 324(22), pp. 2282-2291.
- Wolters, P.J., Blackwell, T.S., Eickelberg, O., Loyd, J.E., Kaminski, N., Jenkins, G., Maher, T.M., Molina-Molina, M., Noble, P.W., Raghu, G., Richeldi, L., Schwarz, M.I., Selman, M., Wuyts, W.A. and Schwartz, D.A. (2018) 'Time for a change: is idiopathic pulmonary fibrosis still idiopathic and only fibrotic?', *The Lancet.Respiratory medicine,* 6(2), pp. 154-160.
- Wolters, P.J., Collard, H.R. and Jones, K.D. (2014) 'Pathogenesis of idiopathic pulmonary fibrosis', *Annual review of pathology,* 9, pp. 157-179.
- Wootton, S.C., Kim, D.S., Kondoh, Y., Chen, E., Lee, J.S., Song, J.W., Huh, J.W., Taniguchi, H., Chiu, C., Boushey, H., Lancaster, L.H., Wolters, P.J., DeRisi, J., Ganem, D. and Collard, H.R. (2011) 'Viral infection in acute exacerbation of idiopathic pulmonary fibrosis', *American journal of respiratory and critical care medicine,* 183(12), pp. 1698-1702.
- Wuyts, W.A., Cavazza, A., Rossi, G., Bonella, F., Sverzellati, N. and Spagnolo, P. (2014) 'Differential diagnosis of usual interstitial pneumonia: when is it truly

idiopathic?', *European respiratory review : an official journal of the European Respiratory Society,* 23(133), pp. 308-319.

- Wynn, T.A. and Ramalingam, T.R. (2012) 'Mechanisms of fibrosis: therapeutic translation for fibrotic disease', *Nature medicine,* 18(7), pp. 1028-1040.
- Xia, S., Lan, Q., Su, S., Wang, X., Xu, W., Liu, Z., Zhu, Y., Wang, Q., Lu, L. and Jiang, S. (2020) 'The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin', *Signal transduction and targeted therapy,* 5(1), pp. 92.
- Xie, L., Law, B.K., Aakre, M.E., Edgerton, M., Shyr, Y., Bhowmick, N.A. and Moses, H.L. (2003) 'Transforming growth factor beta-regulated gene expression in a mouse mammary gland epithelial cell line', *Breast cancer research : BCR,* 5(6), pp. 187.
- Xu, C.C., Wu, L.M., Sun, W., Zhang, N., Chen, W.S. and Fu, X.N. (2011) 'Effects of TGF-beta signaling blockade on human A549 lung adenocarcinoma cell lines', *Molecular medicine reports,* 4(5), pp. 1007-1015.
- Xue, M. and Jackson, C.J. (2015) 'Extracellular Matrix Reorganization During Wound Healing and Its Impact on Abnormal Scarring', *Advances in wound care,* 4(3), pp. 119-136.
- Yan, Y., Dalmasso, G., Sitaraman, S. and Merlin, D. (2007) 'Characterization of the human intestinal CD98 promoter and its regulation by interferon-gamma', *American journal of physiology.Gastrointestinal and liver physiology,* 292(2), pp. 535.
- Yan, Y., Vasudevan, S., Nguyen, H.T. and Merlin, D. (2008) 'Intestinal epithelial CD98: an oligomeric and multifunctional protein', *Biochimica et biophysica acta,* 1780(10), pp. 1087-1092.
- Yang, E.H., Rode, J., Howlader, M.A., Eckermann, M., Santos, J.T., Hernandez Armada, D., Zheng, R., Zou, C. and Cairo, C.W. (2017) 'Galectin-3 alters the lateral mobility and clustering of beta1-integrin receptors', *PloS one,* 12(10), pp. e0184378.
- Yang, I.V., Fingerlin, T.E., Evans, C.M., Schwarz, M.I. and Schwartz, D.A. (2015) 'MUC5B and Idiopathic Pulmonary Fibrosis', *Annals of the American Thoracic Society,* 12 Suppl 2, pp. 193.

- Yao, H., de Boer, W.I. and Rahman, I. (2008) 'Targeting lung inflammation: novel therapies for the treatment of COPD', *Current respiratory medicine reviews,* 4(1), pp. 57-68.
- Young, B., Tan, T.T. and Leo, Y.S. (2021) 'The place for remdesivir in COVID-19 treatment', The Lancet.Infectious diseases, 21(1), pp. 20-21
- Yu, Q. and Stamenkovic, I. (2000) 'Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis', *Genes & development,* 14(2), pp. 163-176.
- Yue, B. (2014) 'Biology of the extracellular matrix: an overview', *Journal of glaucoma,* 23(8 Suppl 1), pp. 20.
- Zhang, K., Rekhter, M.D., Gordon, D. and Phan, S.H. (1994) 'Myofibroblasts and their role in lung collagen gene expression during pulmonary fibrosis. A combined immunohistochemical and in situ hybridization study', *The American journal of pathology,* 145(1), pp. 114-125.
- Zhang, L., Wang, Y., Wu, G., Xiong, W., Gu, W. and Wang, C.Y. (2018) 'Macrophages: friend or foe in idiopathic pulmonary fibrosis?', *Respiratory research,* 19(1), pp. 170-2.
- Zhang, Y.E. (2009) 'Non-Smad pathways in TGF-beta signaling', *Cell research,* 19(1), pp. 128-139.
- Zhao, J., Hopke, P.K., (2011) 'Concentration of Reactive Oxygen Species (ROS) in Mainstream and Sidestream Cigarette Smoke' *Aerosol Science and Technology,* (46), pp 191-197.
- Zhao, Y.D., Yin, L., Archer, S., Lu, C., Zhao, G., Yao, Y., Wu, L., Hsin, M., Waddell, T.K., Keshavjee, S., Granton, J. and de Perrot, M. (2017) 'Metabolic heterogeneity of idiopathic pulmonary fibrosis: a metabolomic study', *BMJ open respiratory research,* 4(1), pp. 000183. eCollection 2017.
- Zhou, X., Wu, W., Hu, H., Milosevic, J., Konishi, K., Kaminski, N. and Wenzel, S.E. (2011) 'Genomic differences distinguish the myofibroblast phenotype of distal lung fibroblasts from airway fibroblasts', *American journal of respiratory cell and molecular biology,* 45(6), pp. 1256-1262.

- Zhou, Y., Huang, X., Hecker, L., Kurundkar, D., Kurundkar, A., Liu, H., Jin, T.H., Desai, L., Bernard, K. and Thannickal, V.J. (2013) 'Inhibition of mechanosensitive signaling in myofibroblasts ameliorates experimental pulmonary fibrosis', *The Journal of clinical investigation,* 123(3), pp. 1096-1108.
- Zou, J., Glinsky, V.V., Landon, L.A., Matthews, L. and Deutscher, S.L. (2005) 'Peptides specific to the galectin-3 carbohydrate recognition domain inhibit metastasis-associated cancer cell adhesion', *Carcinogenesis,* 26(2), pp. 309-318.
- Zou, J.N., Sun, L., Wang, B.R., Zou, Y., Xu, S., Ding, Y.J., Shen, L.J., Huang, W.C., Jiang, X.J. and Chen, S.M. (2021) 'The characteristics and evolution of pulmonary fibrosis in COVID-19 patients as assessed by AI-assisted chest HRCT', PloS one, 16(3), pp. e0248957.
- Zoz, D.F., Lawson, W.E. and Blackwell, T.S. (2011) 'Idiopathic pulmonary fibrosis: a disorder of epithelial cell dysfunction', *The American Journal of the Medical Sciences,* 341(6), pp. 435-438.