University of Oxford

Doctoral Thesis

Structural studies of the neutrophil cell surface receptors GPR97 and CD177 and their interaction with PR3



Céline Zheng-Gérard

Merton College

A thesis submitted for the degree of

Doctor of Philosophy

in the

Department of Biochemistry

Trinity 2022

To papa, mama, and my husband Sam,
for their never-ending support,
and without whom this thesis would not have been possible.

Acknowledgements

I am immensely grateful for the network of colleagues, collaborators, friends and family who have supported me throughout my D.Phil and helped to make the past five years the most enjoyable of my life.

Foremost, I would like to thank my supervisor Elena Seiradake for her continuous support, guidance and advice during this D.Phil. I am deeply appreciative of the training and mentorship she has provided me, as well as for the opportunities to explore projects independently. I am also thankful to my secondary supervisor Mark Sansom for his support and patience with me and the research projects.

I would like to thank our collaborator Hsi-Hsien Lin for bringing this project to the lab, and for his kind advice as well as good ideas, which resulted in a productive collaboration. Others I would like to thank are Ed Lowe for his help with all aspects of crystallography, Kamel El-omari for his help with S-SAD experiments, David Staunton for assistance with SPR and SEC-MALS, and Alan Salama for providing patients' blood serum samples. I would also like to thank Nicole Zitzmann for her kindness and support as my college advisor. I am grateful to Wellcome Trust for their generous funding, and Merton College for their financial support throughout the D.Phil.

I wish to express my warm and sincere thanks to past and present members of the Seiradake group for their help and advice in the lab, but also for the great working environment which made working in the lab so enjoyable: Verity, Onno, Metin, Ditipriya, Fina, Owen, Amy, Miguel, Marilee and Jin. A very special thanks to Maria, our wonderful lab technician, for teaching me cell culture and protein purification, and for maintaining the lab so that our research projects can run as smoothly as possible.

I would like to thank my friends Maria, Antonio, Amy, Nick, Faez and Leon for all the happy memories and late nights playing board games.

Finally, I wish to thank my family, papa, mama and my brothers Christophe and Romain, for their moral support and encouragements since the beginning. Most importantly, I wish to thank my husband Sam for believing in me since high school, and for convincing me to apply to this D.Phil programme. I am forever grateful for your never-ending support, both in and out of the lab, and for constantly inspiring and motivating me to achieve my full potential. Last but not least, I thank you for providing me with a wonderful home and for building a new little family together with Elian.

Abstract

Neutrophils are potent immune cells which act as the first line of defence against invading pathogens and whose activities are tightly regulated to prevent widespread collateral damage to surrounding tissue. On their surface, some receptors mediate processes involved in the regulation of neutrophil activation and potentiation. This thesis focuses on the study of some of these receptors using X-ray crystallography and biophysical methods. The first chapter describes the crystal structure of GPR97, an adhesion G protein-coupled receptor predominantly expressed on the surface of neutrophils. Collaborative work with Prof. Hsi-Hsien Lin shows that GPR97 triggers neutrophil activation upon interaction with Proteinase 3 (PR3) on the surface of neutrophils. More particularly, PR3 appears to be part of a newly identified CD177/GPR97/PAR2/CD16b interactome, which is presented as a critical modulator of neutrophil activation, and directed by a novel GPR97-PAR2 transactivation mechanism.

PR3 is an atypical neutrophil serine protease presented on the surface of neutrophils by its receptor CD177. It is also the main target of autoantibodies, named ANCAs, in the pathogenesis of granulomatosis with polyangiitis (GPA). However, purification and characterisation of PR3 have been hampered by its instability in solution. The second chapter describes the design of PR3 variants with modified surface properties, and highlights the production of a monomeric and highly soluble variant, $PR3^{IWG\text{-}NGT}$. Importantly, $PR3^{IWG\text{-}NGT}$ binds to CD177 and ANCAs, therefore offering promises for clinical applications.

The third chapter describes the structural characterisation of CD177 alone and in complex with PR3. The crystal structure of the full ectodomain of CD177 was solved using S-SAD phasing and reveals the presence of four Ly6/uPAR domains arranged in two subdomains A and B. Biophysical and crystallographic experiments with $PR3^{IWG\text{-}NGT}$ resulted in the first crystal structure of the PR3-CD177 complex, and showed that CD177 binds to PR3 solely through its subdomain A. The model reveals the PR3-CD177 interface is governed by hydrophobic interactions and ultimately provides molecular details into PR3 presentation on the surface of neutrophils.

Overall, this work paves the way for the development of novel therapies targeting the PR3-CD177 complex in GPA or the PR3/CD177/GPR97/PAR2/CD16b interactome in neutrophil-mediated inflammatory diseases.

Contents

List of Figures **xv**

List of Tables **xix**

List of Abbreviations **xxi**

1 Introduction **1**

1.1 The role of neutrophils in inflammation 1
1.2 Proteinase 3 4
1.2.1 Neutrophil serine proteases 4
1.2.2 PR3 expression in neutrophils 4
1.2.3 Biological roles of PR3 5
1.2.4 Role of PR3 in Granulomatosis with polyangiitis 6
1.3 CD177 8
1.3.1 The Ly6/uPAR family 8
1.3.1.1 LU domain architecture 8
1.3.1.2 Biological function 9
1.3.2 Cluster of Differentiation 177 10
1.3.2.1 Human neutrophil antigen 2 10
1.3.2.2 Biological roles of CD177 10
1.4 The adhesion G-protein-coupled receptor GPR97 12
1.4.1 Adhesion G-protein-coupled receptors 12
1.4.1.1 Molecular organisation of adhesion GPCRs 12
1.4.1.2 Dual role of adhesion GPCRs 13
1.4.2 aGPCRs in immunology 17
1.4.2.1 ADGRE subfamily 17
1.4.2.2 BAI1/ADGRB1 18
1.4.2.3 ADGRG subfamily 19
1.4.3 GPR97 19
1.4.3.1 GPR97 neutrophil surface expression in inflammation . . 21
1.4.3.2 GPR97 ligand(s) on the surface of neutrophils 22
1.4.3.3 Identification of GPR97 ligand(s) 24
1.4.3.4 GPR97 interaction with PR3 on the surface of neutrophils 26
1.5 Research aims 29

2 Materials & Methods **31**
2.1 Molecular Cloning . . . 31
2.1.1 Genes . . . 31
2.1.2 Generation of constructs . . . 31
2.1.3 Expression vectors . . . 32
2.1.4 Generation and selection of recombinant plasmid DNA . . . 33
2.2 Protein production . . . 33
2.2.1 HEK293 cell lines and control of N-linked glycosylation . . . 33
2.2.2 HEK293 cell culture maintenance . . . 35
2.2.3 Small-scale transient protein expression . . . 35
2.2.4 Large-scale transient protein expression . . . 36
2.2.4.1 Roller bottle preparation and transfection . . . 36
2.2.4.2 Cell culture medium harvesting . . . 36
2.2.5 Protein purification . . . 37
2.2.5.1 Immobilised metal-ion affinity chromatography . . . 37
2.2.5.2 Size-exclusion chromatography . . . 37
2.2.6 Protocol adaptation for PR3 expression and purification trials . 38
2.2.6.1 PR3 expression in CD177-conditioned medium . . . 38
2.2.6.2 PR3 purification trials . . . 38
2.3 Protein analysis and detection . . . 39
2.3.1 SDS-polyacrylamide gel electrophoresis . . . 39
2.3.2 Western blotting . . . 39
2.3.3 PNGase F deglycosylation assay . . . 40
2.3.4 His-tag pull-down assay . . . 40
2.3.5 PR3 recovery assay . . . 41
2.3.6 PR3 pull-down assay . . . 41
2.4 Functional assays . . . 42
2.4.1 Cell-surface staining assay . . . 42
2.4.2 Enzyme-linked immunosorbent assay . . . 43
2.4.3 ANCA-depletion assay . . . 43
2.5 Biophysical methods . . . 44
2.5.1 Size-exclusion chromatography-multiangle light scattering . . . 44
2.5.2 SEC binding assay . . . 44
2.5.3 Surface Plasmon Resonance . . . 44
2.6 X-ray crystallography . . . 45
2.6.1 Protein crystallisation and crystal optimisation . . . 45
2.6.2 GPR97^{ecto} crystallisation and structure solution . . . 46
2.6.2.1 Crystallisation trials . . . 46
2.6.2.2 Data collection and data processing . . . 47

2.6.2.3 Structure determination 47
2.6.3 $CD177^{ecto}$ crystallisation and structure solution 48
2.6.3.1 Crystallisation trials 48
2.6.3.2 Heavy-atom soaking 48
2.6.3.3 Data collection and data processing 49
2.6.3.4 Structure determination 50
2.6.4 $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystallisation and structure solution . . 50
2.6.4.1 Crystallisation trials 50
2.6.4.2 Data collection and data processing 51
2.6.5 $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ crystallisation and structure solution . . . 51
2.6.5.1 Crystallisation trials 51
2.6.5.2 Data collection and data processing 51
2.6.5.3 Structure determination 51

3 The role of GPR97 in inflammatory activation of human neutrophils 53
3.1 Summary 53
3.2 $GPR97^{ecto}$ production and crystallisation 54
3.2.1 Bioinformatic analyses 54
3.2.2 Construct screening 57
3.2.3 Protein production and structure determination 59
3.2.3.1 Protein production and crystallisation trials 59
3.2.3.2 Structure determination 62
3.3 Structural analysis of $GPR97^{ecto}$ 65
3.3.1 The N-terminal domain 66
3.3.2 The GAIN domain 67
3.4 GPR97 interaction with PR3 on the surface of neutrophils 69
3.4.1 Mapping the GPR97-PR3 binding interface 69
3.4.2 *In vitro* and *in vivo* analyses of GPR97-PR3 interaction 70
3.4.3 Identification of a new receptor interactome 72
3.4.4 Functional relevance of the CD177/GPR97/PAR2/CD16b interactome 74
3.5 Discussion 76
3.5.1 Functional importance of GPR97 NTD 76
3.5.2 A scaffolding role for subdomain A for the GAIN domain 77
3.5.3 Importance of the free cysteine C245 77
3.5.4 A GPR97-mediated PAR2 transactivation on human neutrophils 78

4 Production and characterisation of PR3 **81**
4.1 Expression and initial purification trials of PR3 . . . 81
4.1.1 Bioinformatic analysis . . . 81
4.1.2 Construct screening . . . 82
4.1.3 Protein purification . . . 84
4.2 Optimisation of the purification protocol . . . 85
4.2.1 The purification buffer . . . 86
4.2.2 Protein storage stability . . . 87
4.2.3 Centrifugation of culture medium . . . 89
4.2.4 PR3 recovery from culture medium . . . 90
4.2.4.1 By washing the cell membranes . . . 90
4.2.4.2 By direct purification from the crude culture medium . 90
4.2.4.3 By expression in CD177-conditioned medium . . . 92
4.2.4.4 By using additives before centrifugation . . . 93
4.2.4.5 Reproducibility between small- and large-scale experiments 95
4.3 PR3 stability during protein expression . . . 97
4.3.1 Monitoring PR3 expression in the culture medium . . . 97
4.3.2 Large-scale purification trials of PR3 . . . 98
4.3.3 Large-scale purification trials of PR3-CD177^{ecto} . . . 99
4.4 PR3 surface mutants . . . 100
4.4.1 Design of PR3 surface mutants . . . 101
4.4.2 Mutant screening . . . 102
4.4.3 Large-scale purification of PR3 mouse mutants . . . 103
4.4.4 Large-scale purification of PR3 double glycosylation mutant . . . 103
4.4.5 Interaction of PR3 variants with CD177 . . . 105
4.4.5.1 PR3 pull-downs . . . 106
4.4.5.2 SEC binding assays . . . 106
4.4.6 Investigation of CD177 binding site . . . 108
4.4.6.1 Design and production of CD177-binding mutants . . . 108
4.4.6.2 Binding assays . . . 109
4.5 Putative dimerisation of PR3 . . . 111
4.5.1 Design and production of PR3 dimer variants . . . 112
4.5.2 Identification of PR3 variants' oligomerisation states . . . 113
4.5.3 SEC binding assays . . . 114
4.6 PR3$^{IWG\text{-}NGT}$ recognition by GPA patients' sera . . . 115
4.6.1 ELISA experiments . . . 115
4.6.2 ANCA depletion experiment . . . 116
4.7 Discussion . . . 116
4.7.1 Advantages of using a recombinantly-produced PR3 . . . 116

4.7.2 A challenging protein production and purification 117
4.7.2.1 The recombinantly-produced PR3 variants are functionally relevant 117
4.7.2.2 PR3 is an aggregation-prone protein 118
4.7.2.3 PR3 is lost upon prolonged interaction with the cell culture medium 119
4.7.3 Using surface mutants to improve the protein surface properties 119
4.7.4 PR3 dimerisation and/or oligomerisation 120
4.7.5 The PR3-CD177 binding site 121

5 Structural studies of the PR3-CD177 complex 123
5.1 $CD177^{ecto}$ production and crystallisation 124
5.1.1 Bioinformatic analysis 124
5.1.2 Construct screening 125
5.1.3 Protein production and crystallisation trials 126
5.1.4 Structure determination 128
5.1.4.1 Data processing 128
5.1.4.2 Molecular replacement 128
5.1.4.3 Heavy-atom soaking 130
5.1.4.4 S-SAD phasing 132
5.1.5 $CD177^{ecto}$ structure analysis 133
5.2 PR3 interaction with CD177 137
5.2.1 $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystallisation trials 137
5.2.1.1 Co-purification of $PR3^{IWG\text{-}NGT}$ with $CD177^{ecto}$ 137
5.2.1.2 1:1 mixture of $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$ 138
5.2.1.3 SEC-purified $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ complex 139
5.2.2 Design of shorter PR3-binding CD177 constructs 141
5.2.2.1 SPR binding assays 141
5.2.2.2 $CD177^{subA}$ production 142
5.2.2.3 $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ crystallisation trials 143
5.2.3 $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ structure determination 143
5.2.4 Improvement of $CD177^{ecto}$ model 145
5.2.5 Analysis of the PR3-CD177 binding interface 146
5.3 Discussion 147
5.3.1 $CD177^{ecto}$ reveals a tandem arrangement of LU domains 147
5.3.2 Flexibility between subdomains A and B 148
5.3.3 PR3-CD177 binding interface 149

6 Discussion 151
6.1 Identification of new signalling interactome in human neutrophils 151
6.2 The role of CD177-PR3 in pathogenesis of GPA 152

Appendices

A Supplementary Materials & Methods **157**

B Constructs and primers used in this work **163**

C GPR97 sequence alignments **167**

D PR3 sequence alignment **171**

References **173**

List of Figures

1.1 Overview of neutrophil development and function 2
1.2 Structural comparison of LU domains with 3FTs 9
1.3 Representation of the nine distinct aGPCR subfamilies and their unique N-terminal domain architecture 13
1.4 Disinhibinition models of aGPCR activation 16
1.5 GPR97 expression during inflammatory disorders 22
1.6 GPR97 ligand screening 23
1.7 Identification of CD177 and PR3 as potential GPR97 ligands 24
1.8 Identification of PR3 as GPR97 ligand 26
1.9 A weak GPR97-PR3 direct interaction 27
1.10 Neutrophil activation upon GPR97-PR3 interaction 27
1.11 GPR97 effect on PR3 proteolytic activity 28

2.1 MCS of the pHLsec vector 32
2.2 Control of N-linked glycosylation in HEK293 cells 34

3.1 Predicted GPR97 domain architecture 55
3.2 Comparison of the crystal structures of other aGPCRs 56
3.3 Expression trials of GPR97^{ecto} constructs 58
3.4 Large-scale production of GPR97^{ecto} 60
3.5 GPR97^{ecto} crystallisation trials 61
3.6 Structural building of the NTD 63
3.7 Structural analysis of the NTD 66
3.8 Structural analysis of GPR97^{ecto} 68
3.9 Mapping the GPR97-PR3 binding interface 70
3.10 SPR analysis of GPR97-PR3 interaction 71
3.11 Identification of a new receptor interactome 74
3.12 Decrease in neutrophil activation upon inhibition of PAR2 75
3.13 Simplified schematic of the CD177/GPR97/PAR2/CD16b interactome on the surface of neutrophils 79

4.1 The conserved serine protease fold of PR3 82
4.2 Expression trials of PR3 constructs 83
4.3 Initial large-scale purification trials of PR3 84
4.4 Monitoring proPR3 and PR3 during purification 86

4.5 PR3 stability tests . . . 88
4.6 PR3 stability during centrifugation of culture medium . . . 89
4.7 PR3 recovery from washing cell membranes . . . 91
4.8 PR3 expression in CD177-conditioned medium . . . 93
4.9 Use of additives to improve PR3 solubilisation . . . 94
4.10 Large-scale purification of PR3 in DDM . . . 96
4.11 Evolution of PR3 solubility during protein expression . . . 97
4.12 Large-scale purification of PR3 from early harvest . . . 99
4.13 Large-scale purification of PR3 from early harvest in presence of CD177 100
4.14 Design of PR3 surface mutants . . . 101
4.15 Centrifugation trial of PR3 variants . . . 102
4.16 Large-scale purification of PR3 mouse mutants . . . 104
4.17 Large-scale purification of PR3 double glycosylation mutant . . . 105
4.18 PR3 pull-down experiment . . . 106
4.19 SEC binding assays of PR3 variants with $CD177^{ecto}$ . . . 107
4.20 Large-scale purification of PR3 single glycosylation mutants . . . 108
4.21 Evolution of $PR3^{IWG\text{-}NGT}$ and $PR3^{TFF\text{-}NGT}$ solubility during protein expression . . . 109
4.22 SEC binding assays of PR3 single glycosylation mutants with $CD177^{ecto}$ 110
4.23 SPR analysis of $PR3^{IWG\text{-}NGT}$ and $PR3^{TFF\text{-}NGT\ IWG\text{-}NGT}$ interaction with $CD177^{ecto}$ . . . 111
4.24 PR3 putative dimerisation interface . . . 112
4.25 Large-scale purification of PR3 synthetic and dimer mutant . . . 113
4.26 PR3 variants SEC-MALS experiment . . . 114
4.27 SEC binding assays of PR3 synthetic and dimer mutant . . . 114
4.28 $PR3^{IWG\text{-}NGT}$ recognition by GPA patients' sera . . . 115
4.29 Summary and comparison of all purified PR3 variants . . . 120

5.1 Predicted CD177 domain architecture . . . 124
5.2 Expression trials of $CD177^{ecto}$ constructs . . . 126
5.3 Large-scale production of $CD177^{ecto}$ . . . 127
5.4 Anomalous signal in function of resolution . . . 131
5.5 Substructure determination of $CD177^{ecto}$ sulphur atoms . . . 132
5.6 Initial electron density map obtained from S-SAD phasing . . . 133
5.7 $CD177^{ecto}$ crystal structure . . . 134
5.8 Network of hydrophobic interactions stabilising each CD177 subdomain 135
5.9 Comparison with AlphaFold model . . . 135
5.10 Large-scale purification of PR3-CD177 complex . . . 137
5.11 Crystal hits obtained from 1:1 mixture of $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$. 138
5.12 Production of pure $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ complex . . . 140

5.13 SPR analysis of $PR3^{IWG\text{-}NGT}$ interaction with CD177 constructs 142
5.14 Large-scale purification of $CD177^{subA}$ 142
5.15 $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ crystals 144
5.16 Crystal structure of $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ complex 147

C.1 Predicted GPS for all aGPCRs 168
C.2 GPR97 hydrophobic residues involved in NTD-GAIN stabilisation . . . 169
C.3 Predicted subdomain A of GAIN domains for all aGPCRs 170

D.1 Sequence alignment of serine proteases 172

List of Tables

1.1 Summary of known extracellular ligands 15

2.1 Summary of genes used for sub-cloning 31
2.2 Summary of reagents tested in the PR3 recovery assay 42
2.3 Summary of all heavy-atom derivatives used for crystal soaking 49

3.1 Summary of GPR97^{ecto} crystallisation trials 62
3.2 Data collection and refinement statistics for GPR97^{ecto} 65

4.1 Total number of published structures of all NSPs 118

5.1 Summary of CD177^{ecto} crystallisation trials 128
5.2 Comparison between several native datasets 129
5.3 Comparison between several heavy-atom derivatised datasets 131
5.4 Data collection and initial refinement statistics for CD177^{ecto} 136
5.5 Challenging indexing during data processing 139
5.6 Summary of PR3$^{IWG\text{-}NGT}$-CD177^{ecto} crystallisation trials 141
5.7 Summary of PR3$^{IWG\text{-}NGT}$-CD177^{subA} crystallisation trials 143
5.8 Comparison between several PR3$^{IWG\text{-}NGT}$-CD177^{subA} datasets 143
5.9 Data collection and initial refinement statistics for PR3$^{IWG\text{-}NGT}$-CD177^{subA} 145
5.10 Improved refinement statistics for CD177^{ecto} 146

B.1 Summary of GPR97^{ecto} and CD177^{ecto} constructs used in this work . . 164
B.2 Summary of PR3 constructs used in this work 166

List of Abbreviations

Materials and Methods

A_{280} . . . . . **A**bsorbance at **280** nm

AAV . . . . . **A**NCA-**a**ssociated **v**asculitis

ANCA . . . . . **A**nti**n**eutrophil **c**ytoplasmic **a**ntibodies

ASU . . . . . **As**ymmetric **u**nit

$CC_{1/2}$ . . . . . **C**orrelation **c**oefficient between random half datasets

CD177-CM . . **CD177**-**c**onditioned **m**edium

CHAPS . . . . . 3-[(3-**ch**olamidopropyl)dimethyl**a**mmonio]-1-**p**ropane**s**ulfonate

cmc . . . . . **c**ritical **m**icelle **c**oncentration

DDM . . . . . n-**D**o**d**ecyl-β-D-**M**altopyranoside

DLS . . . . . **D**iamond **L**ight **S**ource

DMEM . . . . . **D**ulbecco's **M**odified **E**agle **M**edium

DPBS . . . . . **D**ulbecco's **PBS**

EDTA . . . . . **E**thylene**d**iamine**t**etra**a**cetic acid

ELISA . . . . . **E**nzyme-**l**inked **i**mmuno**s**orbent **a**ssay

ER . . . . . **E**ndoplasmic **r**eticulum

FBS . . . . . **F**etal **B**ovine **S**erum

fMLF . . . . . N-**f**ormyl-**M**et-**L**eu-**P**he

geoMFI . . . . . **geo**metric **m**ean **f**luorescence **i**ntensity

GPA . . . . . **G**ranulomatosis with **P**oly**a**ngiitis

HEK . . . . . **H**uman **E**mbryonic **K**idney

His_6 . . . . . Hexahistidine

HRP . . . . . **H**orse **r**adish **p**eroxidase

IMAC . . . . . **I**mmobilised **m**etal ion **a**ffinity **c**hromatography

LB . . . . . **L**ysogeny **B**roth

LMNG . . . . . **L**auryl Maltose Neopentyl Glycol

mAb . . . . . **M**onoclonal **an**tibody

mAu . . . . . **M**illi-**a**bsorbance **u**nits

MCS . . . . . **M**ultiple **c**loning **s**ites

MWCO . . . . . **M**olecular **w**eight **c**ut-**o**ff

NDSB . . . . . **N**on-**d**etergent **s**ulfo**b**etaines

NEB . . . . . **N**ew **E**ngland **B**iolabs
Ni-IDA . . . . . **Ni**ckel-**i**mino**d**iacetic **a**cid
OG . . . . . n-**O**ctyl-β-D-**G**lucoside
pAb . . . . . **P**olyclonal **an**tibody
PBS . . . . . **P**hosphate **b**uffer **s**aline
PCR . . . . . **P**olymerase **C**hain **R**eaction
PEI . . . . . **P**oly**e**thylen**i**mine
PLA . . . . . **P**roximity **l**igation **a**ssay
PMSF . . . . . **P**henyl**m**ethane**s**ulfonyl **f**luoride
PTM . . . . . **P**ost-**t**ranslational **m**odification
RMSD . . . . . **r**oot **m**ean **s**quare **d**eviation
ROS . . . . . **R**eactive **o**xygen **s**pecies
r.t. . . . . . **r**oom **t**emperature
RU . . . . . **R**esponse **u**nit
SDS . . . . . **S**odium **d**odecyl **s**ulfate
SDS-PAGE . . **SDS**-**p**oly**a**crylamide **g**el **e**lectrophoresis
SEC . . . . . **S**ize-**e**xclusion **c**hromatography
SEC-MALS . . **SEC**-**M**ulti**a**ngle **l**ight **s**cattering
SPR . . . . . **S**urface]lasmon **r**esonance
S-SAD . . . . . **S**ulphur-**s**ingle wavelength **a**nomalous **d**ispersion

Protein and domain names

3FT . . . . . **3**-**f**inger **t**oxin
7TM . . . . . **7** trans**m**embrane
A1AT α**1**-**a**nti**t**rypsin
aGPCR . . . . . **a**dhesion **G** **p**rotein-**c**oupled **r**eceptor
BSA . . . . . **B**ovine **s**erum **a**lbumin
CD177 . . . . . **C**luster of **D**ifferenciation **177**
CG . . . . . **C**athepsin **G**
CTF . . . . . **C**-**t**erminal **f**ragment
ECD . . . . . **E**xtra**c**ellular **d**omain
EGF . . . . . **E**pidermal **g**rowth **f**actor
EPCR . . . . . **E**ndothelial **p**rotein **C** **r**eceptor
GAIN . . . . . **G**PCR **a**utoproteolysis-**in**ducing
GnTI . . . . . **G**lc**N**Ac-**t**ransferase **I**
GPCR . . . . . **G** **p**rotein-**c**oupled **r**eceptor

GPI . . . . . **G**lycosyl**p**hosphatidyl**i**nositol
GPS . . . . . **G**PCR **p**roteolysis **s**ite
HNE . . . . . **H**uman **n**eutrophil **e**lastase
HormR/Horm **Horm**one **r**eceptor domain
LU domain . . **L**y6/**u**PAR domain
Ly6/uPAR . . . **Ly**mphocyte antigen-**6**/**u**rokinase **p**lasminogen **a**ctivator **r**eceptor
nPR3 . . . . . **n**eutrophil **Pr**oteinase **3**
NSP . . . . . **N**eutrophil **s**erine **p**rotease
NTD . . . . . **N**-**t**erminal **d**omain
NTF . . . . . **N**-**t**erminal **f**ragment
PAR . . . . . **P**roteinase-**a**ctivated **r**eceptor
PR3 . . . . . **Pr**oteinase **3**
SEA . . . . . **S**perm protein, **E**nterokinase and **A**grin domain
TNF-α . . . . . **T**umour **n**ecrosis **f**actor-α

Chapter 1
Introduction

1.1 The role of neutrophils in inflammation

Neutrophils were first described in 1865 by Max Johann Sigismund Schultze as one of the four different types of white blood cells along with lymphocytes, monocytes and eosinophils [1]. However, their role and function during inflammation and innate immunity were only understood and defined in 1892 by Elie Metchnikoff, who later shared the 1908 Nobel Prize with Paul Ehrlich for their founding work on immunity. Today, neutrophils are known to be the most important and abundant type of circulating immune cells (40-80%) as they form the first line of defence of the innate immune system against invading pathogens. They are typically characterised by a multi-lobular nucleus and by a high cytoplasmic granular content, which is why they are also referred to as polymorphonuclear leukocytes and neutrophilic granulocytes [2].

Upon infection and injury, neutrophils are the first immune cells to cross the blood vessel wall and enter the inflamed tissue where they act as phagocytic machines. Pathogen engulfment triggers neutrophils to release their arsenal of reactive oxygen species (ROS) and cytotoxic granule proteins into the phagosome where they work to destroy the pathogen. However, while much of the microbial killing occurs intracellularly, an extracellular release of hydrolytic enzymes and ROS also occurs, which inevitably leads to collateral damage to surrounding healthy tissue. In addition, some neutrophils are

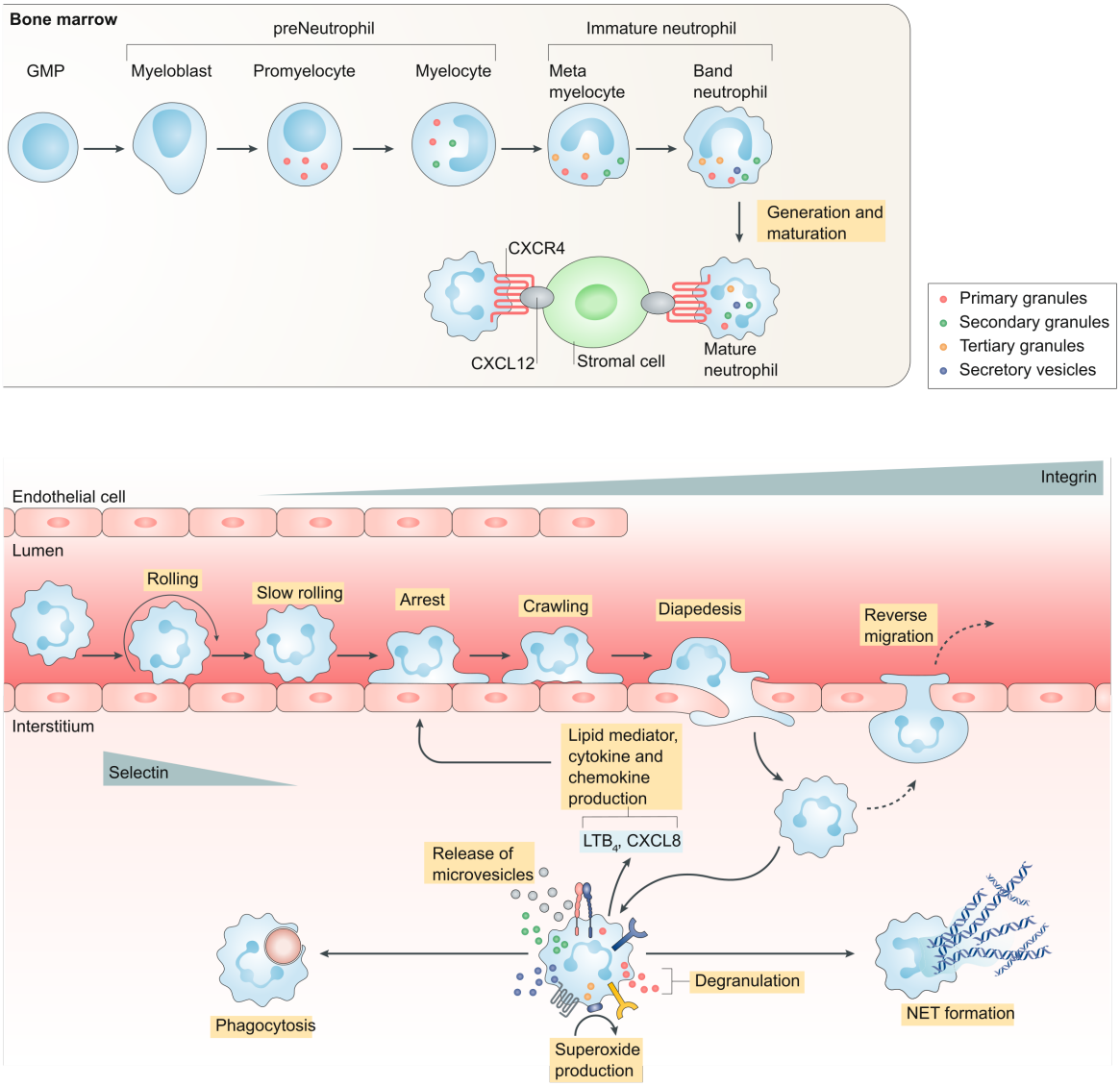


Figure 1.1: Overview of neutrophil development and function. Generation and maturation of neutrophils occur in the bone marrow. Upon microbial or inflammatory stimuli, neutrophils undergo a special recruitment cascade and eventually migrate into the tissues. At the site of inflammation, neutrophils exert their different effector functions, leading to the elimination of the invading microorganisms or promoting the inflammatory response. Abbreviations: CXCL, CXC-chemokine ligand; CXCR, CXC-chemokine receptor; GMP, granulo-monocytic progenitor; LTB_4, leukotriene B_4. Image taken and adapted from [3].

also capable of releasing neutrophil extracellular traps (NETs) at extracellular pathogens in order to immobilise them (Figure 1.1).

Neutrophils are therefore short-lived, potent immune cells whose activity needs to be tightly regulated in order to prevent a widespread damage to host tissue. In fact, failure of regulatory mechanisms often leads to detrimental inflammatory diseases such as organ failure associated with sepsis [4] and systemic inflammatory response syndrome

[5], and autoimmune diseases such as rheumatoid arthritis [6].

Processes involved in the regulation of neutrophil activation and potentiation are mediated by a number of cell surface receptors. Such receptors are capable of coupling extracellular signal to intracellular transduction pathways and include G protein-coupled receptors (GPCRs), Fc-receptors, adhesion receptors, various cytokines receptors and innate immune receptors. Briefly, GPCRs such as formyl peptide receptors [7], complement component 5a receptors [8] or proteinase-activated receptors (PARs) [9], are involved in host defence and inflammation. Upon recognition of microbial invasion, GPCRs strongly trigger neutrophil chemotaxis and activation by inducing ROS production and exocytosis of intracellular granules. Fc-receptors are primarily involved in the recognition of opsonised pathogens [10]. However, some members such as the low affinity Fcγ-receptors, FcγRIIA and FcγRIIIB, may also participate in immune complex-mediated activation of neutrophils [11]. Adhesion receptors such as selectins/selectin ligands and integrins are essential for neutrophil extravasation [12]. In particular, the most important integrin expressed on neutrophils is the β2 integrin Mac-1 (CD11b/CD18) [13]. Cytokine receptors recognise a broad range of small proteins including interleukins (IL), interferons and tumour necrosis factors (TNF), and therefore play a major role in intercellular communications regulating various neutrophil functions [14]. Finally, innate immune receptors such as Toll-like receptors [15] and C-type lectins [16] are pattern recognition receptors that are involved in the direct recognition of pathogens and tissue damage.

Neutrophils are therefore crucial players of innate immunity and inflammation. They express a plethora of immune effector proteins including proteases, cytokines, and receptors which contribute significantly to both the innate and adaptive immune responses. However, mechanisms underlying most of intracellular signal transduction pathways remain to be elucidated. In addition, given the major role of neutrophils in various human diseases, understanding signal transduction in neutrophils is of major biomedical importance [17, 18].

1.2 Proteinase 3

1.2.1 Neutrophil serine proteases

Neutrophils, also called neutrophilic granulocytes, are densely packed with four different type of granules: (1) primary granules, also known as azurophilic granules; (2) secondary granules, also known as specific granules; (3) tertiary granules and; (4) secretory vesicles. Granules exocytosis upon neutrophil activation is tightly regulated and starts with secretory vesicles, followed sequentially by tertiary, specific and azurophilic granules, which are the most difficult to mobilise [19]. In particular, azurophilic granules are the main storage site of the most toxic mediators and contain neutrophil serine proteases (NSPs), myeloperoxidase (MPO), cathepsins and defensins [2].

NSPs form a group of chymotrypsin-like serine proteases that consists of human neutrophil elastase (HNE), proteinase 3 (PR3), cathepsin G (CG), neutrophil serine protease 4 (NSP4) and azurocidin. Upon inflammation, release of those serine proteases can exert widespread effects, including induction of bactericidal activity, degradation of extracellular matrix proteins, promotion of neutrophil transmigration and regulation of vascular integrity [20].

1.2.2 PR3 expression in neutrophils

PR3 is the most abundant NSP with ∼3 pg of PR3 present in each neutrophil [21, 22]. Although the azurophilic granules are the major intracellular store of PR3 in neutrophils, the protein is also localised within the specific and secretory granules of neutrophils [23], and in monocytes, basophils and mast cells [24].

In contrast to its analogous NSPs, low levels of PR3 can be detected on the surface of unstimulated neutrophils [25, 26, 22], resulting in the existence of a constitutively expressed membrane-bound PR3. This membrane-bound PR3 was reported to be catalytically inactive [27], and to exist on the cell surface either via direct interaction with cell membranes through a so-called hydrophobic patch [28] or via binding to its receptor CD177 [29, 30, 31]. Moreover, this hydrophobic patch was described as a group of five positively charged residues that are located just above a group of six

hydrophobic residues, which subsequently promotes protein insertion into the negatively-charged cell membrane [32, 33, 28, 34, 35]. This hydrophobic patch was also reported to mediate CD177 binding to PR3 [36].

Strikingly, an important study described a bimodal distribution of PR3 on resting neutrophils where a low- ($PR3^{low}$) and high-expressing ($PR3^{high}$) subset is distinguished [26]. The percentage of $PR3^{high}$ subset varies in the population between 0 and 100%, is genetically determined, and stable over time [37]. It was later determined that the presence of $PR3^{high}$ is correlated with the bimodal distribution of CD177 on neutrophils, where a non-expressing ($CD177^{neg}$) and expressing ($CD177^{pos}$) subset is distinguished, therefore resulting into the two distinct subpopulations $CD177^{pos}/PR3^{high}$ and $CD177^{neg}/PR3^{low}$ [30, 31].

Upon neutrophil activation, a rapid five- to six-fold increase in cell surface expression of PR3 is observed, contrary to minimal free release of PR3 [22]. Membrane-bound PR3 therefore appears to be the predominant extracellular form of the enzyme, and has been reported to be catalytically active and resistant to inhibition by endogenous inhibitors such as α1-antitrypsin (A1AT) and elafin [22].

1.2.3 Biological roles of PR3

PR3 was initially identified in 1989 under different protein names and is therefore also known as myeloblastin [38], AGP7 [39] and p29 [40] in addition to proteinase 3 [41]. While AGP7, p29 and PR3 refer to the PR3 homolog expressed by neutrophils, myeloblastin refers to the PR3 homolog isolated from human leukaemia cell line HL-60, where the protein was shown to regulate myeloid cell differentiation [38].

While PR3 antimicrobial activity mainly arises from its ability to directly break down bacteria, the protease may also process the inactive peptide precursor hCAP-18 to create the potent antimicrobial peptide LL-37 [42].

Apart from its bactericidal activity, PR3 mediates a wide range of immune processes either through proteolytic cleavage or by protein-protein interactions. In particular, one of the major function of PR3 is in the regulation of pro-inflammatory processes. To do this, PR3 may process and activate cytokines such as tumour necrosis factor α (TNFα)

[43] and interleukin-1β [44], chemokines such as interleukin-8 (IL-8) [45], and cellular receptors such as proteinase-activated receptor 2 (PAR2) [46]. More specifically, PAR2 activation by PR3 results in the production and secretion of IL-8 [47], and IL-8 processing by PR3 results in a strong chemotactic activity responsible for neutrophil migration to inflammatory sites and degranulation [48]. By contrast, PR3 may also process and inactivate anti-inflammatory proteins such as annexin 1 [49] and C1-inhibitor [50].

Moreover, PR3 appears to play a role in endothelial activation by enhancing the production of both IL-8 and monocyte protein-1, two chemotactic stimuli for neutrophils and monocytes, respectively, by endothelial cells [51, 52].

Additionally, given that PR3 can degrade extracellular matrix components such as elastin [53], fibronectin, laminin, vitronectin and collagen type IV [54, 55], an increased presence of PR3 on the surface of neutrophils has been suggested to promote neutrophil transmigration. Similarly, proteolytic cleavage of endothelial protein C receptor (EPCR) [56] and von Willebrand factor [57] by PR3 has been reported, which suggests a role for PR3 in haemostasis.

Interestingly, several recent reports propose a new role in the clearance of apoptotic cells (*i.e.* efferocytosis). More specifically, PR3 appears to be externalised along with phosphatidylserine lipids during neutrophil apoptosis [58, 59]. Presence of PR3 on the surface of apoptotic cells then prevents them from being cleared by macrophages, which ultimately hamper resolution of inflammation and may contribute to systemic, non-resolving inflammation as well as autoimmunity [60, 61, 62].

Finally, several protein interacting partners of PR3 have been identified including CD16b/FcγRIIIB [63], Mac-1 (CD11b/CD18) [64], IL-32α [65] and a 111 kDa membrane molecule on endothelial cells [66] in addition to CD177 [30]. However, the role of such associations in mediating neutrophil functions remains to be understood.

1.2.4 Role of PR3 in Granulomatosis with polyangiitis

Shortly after its first identification, PR3 was also recognised as the main autoantigen of antineutrophil cytoplasmic autoantibodies (ANCAs) [40, 67, 68, 69, 70, 71],

which are harmful antibodies detected in patients suffering from Granulomatosis with polyangiitis (GPA).

GPA, formerly known as Wegener's granulomatosis, is a rare autoimmune disorder that causes inflammatory necrosis of small and medium blood vessels in the nose, sinuses, throat, lungs and kidneys. This condition is one of a group of blood vessel disorders called ANCA-associated vasculitis (AAV). AAV is subdivided into GPA, microscopic polyangiitis (MPA) and eosinophilic granulomatosis with polyangiitis (EGPA), and is characterised by the presence of ANCAs targeting either PR3 or MPO in patient's blood [72]. More specifically, the presence of PR3- and MPO-ANCAs is strongly associated with GPA and MPA, respectively [73, 74].

PR3-ANCAs were first recognised as disease marker of GPA in 1985, when monoclonal antibodies against extranuclear components of neutrophils were detected in GPA patients [75]. Binding of PR3-ANCAs to PR3 on the surface of neutrophils was shown to trigger ligation of the low affinity FcγRIIA and FcγRIIIB receptors [76] and engagement of the β2 integrin Mac-1 [77] resulting in respiratory burst and neutrophil degranulation [78, 79, 80, 81, 82, 77]. This was suggested to lead the generation of more ANCAs, which would in turn activate more neutrophils and ultimately contribute to self-perpetuating tissue injury caused by the release of ROS and proteases in the surrounding environment. Additionally, secreted PR3 may activate endothelial cells and promote strong chemotactic stimulus for neutrophils and monocytes [51, 52], resulting in influx and extravasation of neutrophils and mononuclear cells to the site of inflammation. There, they would gather around activated neutrophils and ultimately contribute to formation of the characteristic granulomas observed in GPA patients [83].

Abundance of PR3 is today considered as a critical risk factor for GPA given the strong association between high surface expression levels of PR3 and GPA pathogenesis [84, 85, 86]. In addition, the detection of circulating PR3-ANCAs in GPA patients and the positive correlation between disease activity and ANCA titre suggest that the immune response to PR3 plays a major role in GPA pathogenesis.

1.3 CD177

1.3.1 The Ly6/uPAR family

The lymphocyte antigen-6 (Ly6)/urokinase plasminogen activator receptor (uPAR) family consists of structurally related proteins that are characterised by the presence of Ly6/uPAR (LU) domain(s). This family of proteins was first described more than 40 years ago when the search for lymphocyte differentiation antigens in mouse led to the identification of Ly6 proteins [87]. Since then, several human homologs exhibiting varying degrees of similarities to murine Ly6 proteins were isolated and led to the recognition that they represent a well-conserved family with a wide range of expression patterns and functions.

To date, 35 genes encoding Ly6/uPAR family members have been identified in the human genome and revealed the family mainly comprises secreted or glycosylphosphatidylinositol (GPI)-anchored single-domain proteins [88]. Members typically contain one LU domain, with the exception of C4.4A [89], CD177 [90], Haldisin [91], uPAR [92] and TEX101 [93], which contain two to three direct repeats of LU domains.

1.3.1.1 LU domain architecture

The first structure of LU domain was obtained from solving the extracellular region of CD59 by nuclear magnetic resonance spectroscopy [94], and was shorty followed by several other structures solved by x-ray crystallography [95, 96, 97]. A typical LU domain consists of 60-90 amino acids and contains ten cysteines arranged in a specific pattern allowing the formation of five disulphide bridges. The overall fold consists of three β-stranded loops extending from a compact core stabilised by four disulphide bridges. As the three loops projecting from the core resemble three outstretch fingers of a hand, this was described as a three-fingered protein fold. Strikingly, this fold is highly similar to the structure adopted by the three-finger toxins (3FTs) found in snake venom [98] and first described in 1976 [99, 100] (Figure 1.2).

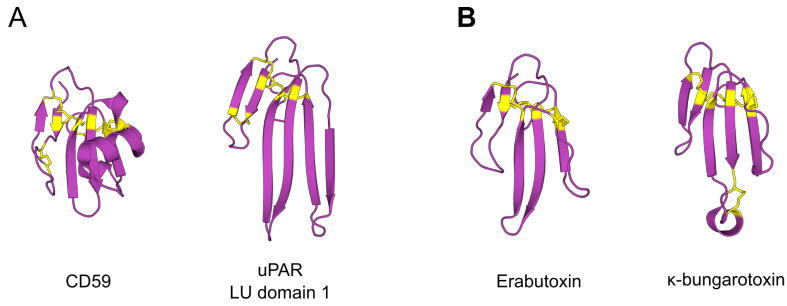


Figure 1.2: Structural comparison of LU domains with 3FTs. **(A)** Structures of two mammalian LU domains from CD59 and uPAR (PDB ID: 2J8B and 1YWH, respectively). **(B)** Structure of the snake toxins erabutoxin and κ-bungarotoxin (PDB ID: 1QKD and 1KBA, respectively).

1.3.1.2 Biological function

While the 3FTs from snake venom mainly target the central nervous system by interacting with acetylcholine receptors, acetylcholine esterases and L-type calcium channels [101], mammalian LU domain-containing proteins present a wide range of functions in nervous, immune, endocrine, vascular and reproductive systems [88].

The best characterised member is uPAR which, by binding to its ligand urokinase plasminogen activator (uPA), mediates and focuses uPA-mediated plasminogen activation to the cell surface. Other members of Ly6/uPAR family such as TEX101 and C4.4A have been suggested a role in fertility [102] and in cell differentiation [103], respectively. Lynx1 and SLURP-1 are GPI-anchored and secreted members, respectively, which modulate neuronal acetylcholine receptors [104]. GPI-anchored high-density lipoprotein-binding protein 1 (GPIHBP1) plays a major role in efficient hydrolysis of intravascular triglyceride by transporting and stabilising lipoprotein lipase [105]. In the immune system, CD59 inhibits autologous complement activation by preventing formation of the complement membrane attack complex [106], while CD177 mediates neutrophil endothelial transmigration [107, 108]. Most importantly, TGFβR1-2, BMPR1A-B, BMPR2, ACVR1A-C, ACVR2A-B, which are receptors for transforming growth factor β (TGFβ), bone morphogenic protein (BMP) and activin, respectively, play fundamental roles in the regulation of basic biological processes such as growth, development, tissue homeostasis and regulation of immune system [109].

1.3.2 Cluster of Differentiation 177

CD177, also named neutrophil-specific alloantigen 1 (NB1), is a GPI-anchored glycoprotein of ∼50 kDa and an atypical member of Ly6/uPAR family by possessing two LU domains within its ECD. The receptor is mainly expressed on neutrophils plasma membrane and secondary granules [110, 111], but can also be found on neutrophil precursors such as neutrophilic myelocytes and metamyelocytes [112, 113, 90].
Strikingly, CD177 is only expressed by a subpopulation of neutrophils, resulting in existence of $CD177^{pos}$ and $CD177^{neg}$ neutrophil subsets. Similarly to PR3, the percentage of $CD177^{pos}$ subset varies in the population between 0 and 100%, is genetically determined, and remains constant over time [112, 114]. The only exception is during pregnancy, when an increased fraction of $CD177^{pos}$ subset can be observed [115].
Notably, CD177 is also known under the alternative name PRV-1, as one of its allele has been identified and associated with the pathogenesis of the hematoproliferative disease polycythemia rubra vera (PRV) [116, 117].

1.3.2.1 Human neutrophil antigen 2

CD177 was first characterised in 1971 as the target of maternal antibodies in neonatal alloimmune neutropenia [118]. More specifically, CD177 carries the neutrophil alloantigen HNA-2, which appears to be absent from the neutrophils surface in 3-5% of the population. Exposure of those HNA-2_{null} individuals to neutrophils presenting CD177 on their surface subsequently induces immunisation against CD177 and production of alloantibodies. This is the case during pregnancy, when HNA-2_{null} women produce alloantibodies against CD177, which can then cross the placenta and react with foetal neutrophils resulting in neonatal neutropenia [118]. Such an alloimmune reaction also occurs during transfusion and transplantation, causing transfusion-related acute lung injury [119, 120, 121] and graft failure after bone marrow transplantation [118, 122, 123], respectively.

1.3.2.2 Biological roles of CD177

Although the functional roles of CD177 remain poorly understood, a few studies have suggested a main role in neutrophil transmigration. Indeed, CD177 has been identified

as a high-affinity binding partner for platelet endothelial cell adhesion molecule-1 (PECAM-1, CD31) presented by endothelial cells, and may therefore contribute to interactions between neutrophils and endothelial cells in the context of inflammatory cell recruitment [121].

CD177 has also been described as the receptor which presents PR3 on the surface of neutrophils [29, 30, 31]. However, there have been contradicting reports on the role of CD177-PR3 complex in neutrophil transmigration. While some suggested that presence of CD177-bound PR3 facilitates neutrophil transmigration by virtue of its proteolytic activity [107], others showed instead that CD177 binding to PR3 inhibits its proteolytic activity and therefore that $CD177^{pos}/PR3^{high}$ and $CD177^{neg}/PR3^{low}$ neutrophils show similar migration through endothelial cell layers [124].

Interestingly, in a PECAM-1-independent pathway, ligation of CD177 was shown to enhance surface expression of $\beta 2$ integrins and attenuate chemokine signalling, which ultimately resulted in impaired neutrophil migration and stronger adhesion [108]. Such a mechanism could be relevant in AAV, where recognition of CD177-bound PR3 by ANCAs may activate similar pathways, ultimately contributing to vascular injury. Additionally, direct association of CD177 with the $\beta 2$ integrin Mac-1 (CD11b/CD18) was suggested to form the signalling complex mediating the transduction of signals initiated by binding of ANCAs [77]. Nevertheless, CD177 does not seem to be directly involved in the pathogenesis of GPA, but rather act as a simple receptor for PR3 [77, 124, 125].

More recently, CD177 has been identified on the surface of a subpopulation of tumour-infiltrating regulatory T cells (TI T_{reg} cells) that are present in solid cancers such as breast, renal, lung and colorectal cancers [126, 127, 128]. In particular, expression of CD177 was found in 15-25% of TI T_{reg} cells and was closely associated with increased suppressive activity contributing to cancer immune evasion [128].

Remarkably, neutrophils from HNA-2_{null} individuals appear to be functioning normally, which suggest the presence of an alternative protein that duplicates the function of CD177 on neutrophils [129].

1.4 The adhesion G-protein-coupled receptor GPR97

1.4.1 Adhesion G-protein-coupled receptors

GPCRs form the largest family of membrane proteins with over 800 members encoded by the human genome [130] and divided into five major families: Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin [131]. GPCRs are activated by a diverse selection of stimuli, which leads to a large variety of cellular responses, resulting in a complex network of interactions between a ligand, its receptor and the signalling cascade.
Adhesion G protein-coupled receptors (aGPCRs) were first identified in 1995 with the description of EMR1 as epidermal growth factor (EGF) module-containing, mucin-like hormone receptor 1 [132] and CD97 as cluster of differentiation 97 [133]. Both new receptors displayed a seven transmembrane (7TM) region preceded by an extensive extracellular region containing several EGF-like domains, and were therefore grouped into a new subfamily named EGF-7TM [134]. Soon after, other homologous 7TM receptors with large extracellular regions were identified and include the calcium-independent receptor of α-latrotoxin 1 (CIRL-1) [135], cadherin EGF LAG seven-pass G-type receptors (CELSRs) [136] and brain-specific angiogenesis inhibitors (BAIs) [137, 138] among others. This confirmed the existence of a novel type of GPCR with a large extracellular domain (ECD) composed of structural motifs that are typically found in cell adhesion proteins, suggesting a role in coupling cell-cell and cell-matrix interactions to intracellular signalling. Later, an extensive phylogenetic analysis of the human GPCR superfamily finally disclosed 33 members in the new aGPCR family [139], which can be further divided into nine subfamilies I-IX [140], and which now forms the second largest family of the GPCR superfamily.

1.4.1.1 Molecular organisation of adhesion GPCRs

aGPCRs share the 7TM domain and C-terminal cytoplasmic tail of classic GPCRs. However, they are atypical by having a large N-terminal ECD that contains a variety of conserved adhesion domains such as EGF-like repeats, cadherin domains (Cad), immunoglobulin (Ig)-like domains, laminin domains (lam), hormone receptor domains (HormR), calnexin-β domains (Calxβ) and leucine-rich repeats (LRR). The presence

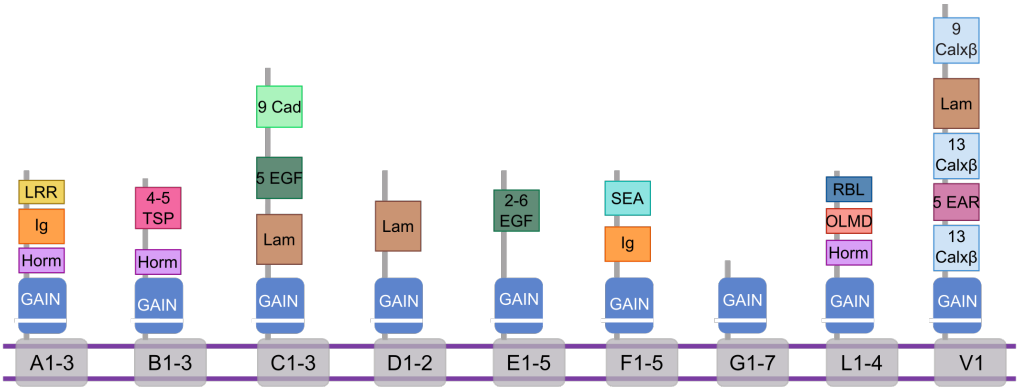


Figure 1.3: Representation of the nine distinct aGPCR subfamilies and their unique N-terminal domain architecture. Subfamily naming follows the new nomenclature, which was recently proposed by the International Union of Basic and Clinical Pharmacology to provide a more coherent and systematic naming for receptors which were historically named based on their characteristics and functions. The receptors were renamed with the ADGR prefix, followed by a letter indicating the receptor subfamily and a number for each receptor within that group [148]. They key protein domains are depicted for each group. Some subfamily members have a slightly different structure from what is indicated in the figure. ADGRA1 has no GAIN, Horm, Ig, or LRR domain; ADGRA2 has no Ig domain; ADGRB3 has an N-terminal CUB; ADGRF2-4 lack the SEA domain; ADGRG1 has a PLL domain, ADGRG6 has the CUB and LAM domains; ADGRL4 has only EGF and GAIN domains. Abbreviations: CUB, complement C1r/C1s, Uegf, Bmp1 domain; EAR, epilepsy-associated repeat; Horm, hormone receptor domain; OLMD, olfactomedin-like domain; RBL, rhamnose-binding lectin; SEA, Sperm protein, Enterokinase and Agrin domain; TSP, type-1 thrombospondin repeat. Figure adapted from [149].

and combination of those adhesion domains varies greatly between members of different subfamilies (Figure 1.3). The one ECD domain that is common to almost all aGPCRs is the GPCR autoproteolysis-inducing (GAIN) domain [141], which contains the highly conserved GPCR proteolysis site (GPS) motif [142], almost directly N-terminal to the first transmembrane helix of the 7TM domain. This unique feature makes members of the aGPCR family capable of having their large N-terminal ECD autoproteolytically cleaved from their 7TM and C-terminal tail. Both N- and C-terminal receptor protomers (named the N-terminal fragment (NTF) and C-terminal fragment (CTF)) remain non-covalently associated with each other to form a mature heterodimeric receptor complex on the cell surface [143, 135, 142, 144, 141]. Intriguingly, not all aGPCRs are cleaved at their GPS [145] and the receptor self-cleavage is not necessary for cell surface expression [146, 147].

1.4.1.2 Dual role of adhesion GPCRs

As a result of their hybrid structure, aGPCRs may play dual roles in cell adhesion by acting as cell-surface sensors to mediate and probe cell-cell or cell-matrix interactions,

Receptor (historic names/nomenclature)	**Proposed function**	**Ligand**	**Refs**
	Subfamily I (ADGRL subfamily)		
Lphn1/CIRL-1/ ADGRL1	Synapse formation	Teneurin2	[150]
	Neuronal migration	FLRT1, FLRT3	[151]
	Synapse formation	Neurexins	[152]
	Neuronal apoptosis	Contactin6	[153]
Lphn2/CIRL-2/ ADGRL2	Synapse formation	Teneurin2	[150, 154]
	Neuronal migration	FLRT3	[151]
Lphn3/CIRL-3/ ADGRL3	Synapse formation	Teneurin2	[154]
	Neuronal migration	FLRT1, FLRT2, FLRT3	[155, 151]
	Neuronal migration	Unc5D	[156]
	Subfamily II (ADGRE subfamily)		
EMR1/F4/80/ ADGRE1	Macrophage-NK cell communication	Unknown ligand on NK cells	[157]
EMR2/CD312/ ADGRE2	Cell adhesion	Dermatan sulphate	[158]
	Inflammation	Factor H-related protein 1	[159]
EMR3/ADGRE3	Myeloid-myeloid cells communication	Unknown ligand on macrophages and neutrophils	[160]
EMR4/FIRE/GPR127/ ADGRE4	Myeloid-B cells communication	Unknown B cell ligand	[161]
CD97/ADGRE5	Leukocytes-leukocytes communication	CD55 (Decay-Accelerating Factor)	[162]
	Cell adhesion	Dermatan sulphate	[158]
	Angiogenesis	Integrin $\alpha_5\beta_1$, $\alpha_v\beta_3$	[163]
	Leukocytes trafficking	CD90 (Thy-1)	[164]
	Prostate cancer cell migration	Lysophosphatidic acid receptor 1	[165]
	Colonic cancer cell migration	β-Catenin	[166]
	Subfamily III (ADGRA subfamily)		
GPR124/TEM5/ ADGRA2	Angiogenesis	Integrin $\alpha_v\beta_3$	[167]
	Cell adhesion	Glycosaminoglycans	[167]
	CNS angiogenesis and BBB regulation	Reck	[168]
	Subfamily IV (ADGRC subfamily)		
CELSR3/ADGRC3	Axon guidance	Dystroglycan	[169]
	Subfamily V (ADGRD subfamily)		
GPR133/ADGRD1	Embryo transport	Plexin Domain-Containing Protein 2	[170]
	Subfamily VI (ADGRF subfamily)		
GPR110/ADGRF1	Brain development	Synaptamide	[171]

(continued on next page)

GPR116/Ig-Hepta/ ADGRF5	Lung surfactant homeostasis	Surfactant protein D	[172]
	Glucose homeostasis	Fibronectin type III domain-containing 4	[173]
Subfamily VII (ADGRB subfamily)			
BAI1/ADGRB1	Phagocytosis of apoptotic cells	Phosphatidylserine	[174]
	Phagocytosis of Gram-negative bacteria	Lipopolysaccharide	[175]
	CNS angiogenesis	CD36	[176]
	CNS angiogenesis	Integrin $\alpha_v\beta_3$	[177]
	Synapse and dendrite formation, axon growth	RTN4/NoGo receptors	[178]
BAI2/ADGRB2	-	Glutaminase interacting protein	[179]
BAI3/ADGRB3	Synapse formation and maintenance	C1q-like proteins	[180]
Subfamily VIII (ADGRG subfamily)			
GPR56/TM7XN1/ ADGRG1	Suppression of cancer cell growth	Tissue transglutaminase 2	[181]
	Neuronal migration	Collagen III	[182]
	Colonic stem and cancer cells proliferation	Progastrin	[183]
GPR97/Pb99/ADGRG3	-	Cortisol	[184]
GPR126/DREG/ ADGRG6	Schwann cell development	Laminin-211	[185]
	Schwann cell development	CollagenIV	[186]
	Myelin homeostasis	Prion Protein	[187]

Table 1.1: Summary of known extracellular ligands. Abbreviations: BBB, blood-brain barrier; CD, cluster of differentiation; CNS, central nervous system; DREG, developmentally regulated GPCR; FIRE, F4/80-like receptor; FLRT, fibronectin leucine-rich repeat transmembrane protein; GPR, G-protein coupled receptor; Lphn, latrophilin; NK, natural killer; Pb, pre-B cell; TEM5, tumor endothelial marker; TM7XN, seven-transmembrane protein with unknown N-terminal function; Unc5D, Uncoordinated-5 Netrin Receptor D. Table adapted and updated from [188, 189].

and in signalling by acting as classical receptors, which transmit external signals into the cell interior, thereby initiating intracellular signalling cascades.

aGPCRs can bind to a broad range of ligands including membrane-bound proteins, glycosaminoglycans and lipids, which permits them to modulate diverse biological processes ranging from immunoregulation to tissue polarity, angiogenesis, tumorigenesis and brain development (Table 1.1). Furthermore, the ability of aGPCRs to bind many, often seemingly unrelated, ligands suggests that these receptors may be able to act as so-called "context recognisers" and modulate their signalling depending on

the interacting partner [188].

Although many interacting partners have been identified for the majority of aGPCRs, many of them remain orphan receptors. In addition, the molecular mechanisms by which ligand recognition leads to receptor signalling remain to be fully elucidated. The GPS motif has been shown to play an important role for receptor function [143, 161, 190, 191]. In fact, one main model for receptor activation, called tethered agonist model, relies on receptor self-cleavage at the GPS site, resulting in the release of a short and hydrophobic polypeptide designated as the stalk, *Stachel* or tethered ligand [192, 193, 165, 194, 195, 186, 196, 197, 198, 199]. This tethered ligand usually maintains its β-strand conformation while being associated with the rest of the GAIN domain. However, upon either ligand binding or mechanical forces displacing the NTF [200, 201, 145, 185], the now-exposed tethered ligand looses its secondary structure and can insert deeply into the 7TM bundle, leading to activation of the signalling cascade [202, 203, 204, 205] (Figure 1.4A).

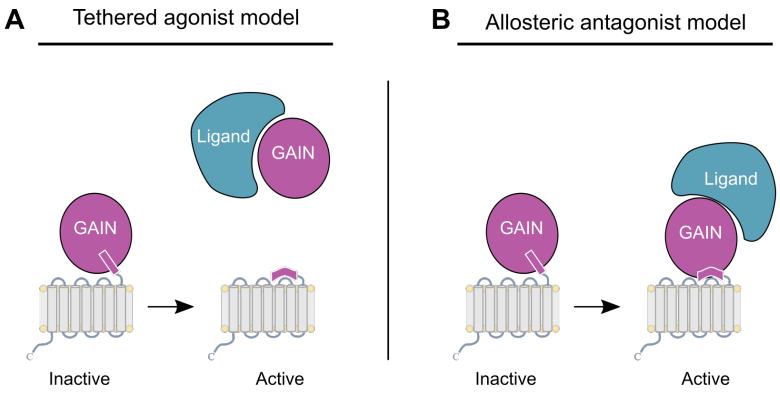


Figure 1.4: Disinhibinition models of aGPCR activation. In the resting state, the NTF acts to mask the *Stachel*. **(A)** In the tethered agonist model, ligand binding and/or mechanical forces result in dissociation of the NTD from the CTF, which exposes the *Stachel* and leads to receptor activation. **(B)** In the allosteric antagonist model, ligand binding and/or mechanical forces result in a conformational change in the NTF, allowing the *Stachel* to activate the receptor.

In another model, called the allosteric antagonist model, the release and exposure of the tethered ligand is not a requirement for receptor activation, and explains why some autoproteolysis-deficient aGPCRs are still able to transduce extracellular stimuli [180, 206, 147, 203] (Figure 1.4B). Interestingly, the tethered agonist model of activation is remarkably similar to that of another GPCR subfamily called the proteinase-activated receptors (PARs), where cleavage by exogenous proteases result in a tethered agonist,

which then physically interacts with the 7TM domain and subsequently activates the signalling cascade.

1.4.2 aGPCRs in immunology

It was the study of the immune system that led to the identification of the first aGPCRs EMR1/ADGRE1 and CD97/ADGRE5 in 1995 [132, 133]. Since then, several other aGPCRs were found to be prominently expressed by immune cells. Those include the remaining members of the ADGRE subfamily ADGRE2-4, BAI1/ADGRB1, GPR56/ADGRG1, GPR97/ADGRG3 and GPR114/ADGRG5.

1.4.2.1 ADGRE subfamily

Mouse EMR1/ADGRE1 (also known as F4/80) is most specifically expressed on the surface of macrophages, with only a limited expression pattern on monocytes, tissue macrophages and some dendritic cells [207]. It has been shown to be essential for the natural killer cell-macrophage cellular contact, which is necessary to trigger a robust immune response following exposure to bacteria [157]. The receptor also appears to play a direct role in the induction of peripheral immune tolerance by mediating the differentiation of $CD8^+$ regulatory T cells; therefore preventing the immune system from reacting to harmless food antigens and allergens and/or causing autoimmune disease [208]. Intriguingly, human EMR1 is not present on the surface of macrophages, but is instead specifically expressed by eosinophilic granulocytes, suggesting a different role for the receptor in humans [209].

Both EMR2/ADGR2 and EMR3/ADGRE3 are widely expressed by human myeloid cells (monocytes, macrophages, dendritic cells and granulocytes) [210, 211]. Interestingly, no murine homolog can be found for both of these receptors, suggesting a specific need for EMR2 and EMR3 in the human immune response [212]. EMR2 interaction with dermatan sulphate, which is found mostly in the skin and in connective tissue, suggest a function during cell adhesion and migration [158]. In fact, EMR2 has been shown to play a role in neutrophil activation by enhancing granulocyte adhesion and migration as well as production of ROS and degranulation of proteolytic enzymes in the

presence of inflammatory stimuli [213]. The receptor is also involved in adhesion and migration of myeloid cells during cell differentiation, maturation and activation [214]. In addition, its increased expression has been associated with liver cirrhosis [215] and non-infectious systemic inflammation response syndrome (SIRS), where an increased number of hyperactivated neutrophils lead to multiple organ failure [216]. Interestingly, EMR2 possesses a missense variant (C492Y) within its GAIN domain that is solely responsible for a rare type of dermal allergy caused by vibration-induced mast cell activation, called autosomal dominant vibratory urticaria [217]. This missense variant is thought to weaken the association between the NTF and CTF, therefore making it easier to dislodge the NTF upon vibrations and resulting in receptor activation.
Mouse EMR4/ADGRE4 is exclusively found on the surface of myeloid cells [218]. Human EMR4, however, does not exist on the surface of immune cells as a result of a one-nucleotide deletion truncating the 7TM domain. Interestingly, this deletion is not observed in non-human primates (such as chimpanzees), and therefore accounts for a different role for the receptor in human and non-human primates [219].
On the contrary, the structure and cellular distribution of CD97/ADGRE5 are similar between human and mouse and suggest a conserved biological role. CD97 is expressed by all leukocytes [220, 221] and has been implicated in innate and adaptive immune processes through its role in cellular migration [162, 222], proliferation [223, 224] and mobilisation from the bone marrow [225]. Notably, CD97 appears to be involved in chronic inflammatory diseases such as rheumatoid arthritis [226, 227] and multiple sclerosis [228]. In addition, overexpression of the receptor has been shown to promote tumour growth and invasion capacity [229].

1.4.2.2 BAI1/ADGRB1

Despite BAI1/ADGRB1 expression not being restricted to immune cells, the receptor plays an important role in cell clearance and immune regulation. Indeed, BAI1 is directly involved not only in the phagocytic uptake of Gram-negative bacteria by macrophages [175, 230], but also in the engulfment of apoptotic cells by macrophages [174] and microglia [231]. Upon stimulation with apoptotic cells, the receptor has also been

shown to mediate the biogenesis of high-density lipoprotein, which possesses potent anti-inflammatory functions [232]. Moreover, BAI1 can increase the uptake of apoptotic cells by intestinal epithelial cells, leading to a reduction of tissue inflammation due to uncleared apoptotic corpses [233].

1.4.2.3 ADGRG subfamily

While the GPR56/ADGRG1 expression pattern seems to be restricted to natural killer cells, T lymphocytes, as well as hematopoietic stem and progenitor cells in human, GPR97/ADGRG3 appears to be predominantly expressed on human granulocyte precursors and terminally differentiated granulocytes. By contrast, GPR114/ADGRG5 can be found in human eosinophils [234, 235].

When expressed on the surface of natural killer cells, GPR56 seems to inhibit the natural cytotoxicity of those immune cells by suppressing the production of inflammatory cytokines and cytolytic proteins [236]. The receptor is also highly expressed in quiescent hematopoietic stem cells in the adult bone marrow [237], and plays a critical role in the generation of the first hematopoietic stem cells [238] as well as their renewal and retention within the bone marrow [239]. Together with GPR97 and GPR114, GPR56 has been associated with poor survival in acute myeloid leukaemia [240], where treatment-resistant leukaemic patients were found to have particularly high GPR56 expression levels [241, 242].

1.4.3 GPR97

Human GPR97 (also named PB99 or PGR26) was identified in 2002 [243] and is a member of the subfamily VIII of aGPCRs. GPR97 is one the smallest aGPCRs with a predicted ∼30 kDa ECD comprising solely of the conserved GAIN domain, where the receptor is autoproteolysed at the GPS motif [235]. Along with GPR56 and GPR114, its closest relatives, they form one of the three gene clusters of aGPCRs predominantly expressed by immune cells [244]. GPR97 is one of the remaining aGPCRs for which an interacting partner has not yet been discovered.

Nonetheless, the receptor has been identified as the most highly and specifically expressed GPCR in mouse lymphatic endothelium, where it is thought to regulate adhesion and migration of endothelial cells via the small GTPases RhoA and Cdc4, as well as cytoskeletal reorganisation [245].

Interestingly, GPR97 and GPR56 appear to share redundant roles during hematopoietic cell development and differentiation [246].

As GPR97 gene expression was found in murine pre-B and pre-T cells in the thymus, but not in mature B and T cells, a role for the receptor in early lymphoid development was initially suggested [247, 234]. Intriguingly, following studies in independently raised knockout mice revealed no necessity for GPR97 in B or T cell development [247], but implied instead an essential role in regulating the follicular-versus-marginal zone B lymphocyte cell fate decision, which is an important event during peripheral B cell differentiation [248].

GPR97 is also involved in macrophage-related local inflammation, whereby GPR97 knockout results in an increase of infiltrating macrophages and higher TNFα levels in the liver and kidney of a high-fat diet-induced obese mouse [249]. In contrast, despite GPR97 expression in eosinophils, the receptor was found to be insignificant for inflammation in the development of allergic asthma in mice [250].

Importantly, GPR97 has been associated with acute kidney injury, where upregulation of the receptor is thought to contribute to aberrant inflammatory responses and tubular epithelial death, thus resulting in renal damage [251]. GPR97 may also have a role in the development of an autoimmune disease in the central nervous system called experimental autoimmune encephalomyelitis, by which the upregulated receptor modulates T cell differentiation and cytokine profiles of the spinal cord and spleen tissues [252].

More recently, the receptor has been shown to be highly expressed by tissue-infiltrating neutrophils in the case of inflammation and certain cancers, and to be upregulated during systemic inflammation. GPR97 is thought to play an important role in potentiating neutrophil activation by enhancing the production of ROS as well as the activity of proteolytic enzymes. The receptor also appears to increase neutrophil bacterial uptake and killing activity, therefore regulating antimicrobial activity in human neutrophils [235].

Notably, GPR97 is the only aGPCR for which activation by small molecules have been reported. These include the drug beclomethasone dipropionate [253], the inhibitor of cholesterol absorption ezetimibe, the calcium channel and histamine H1 blocker flunarizine, the non-steroidal estrogen zeranol [254], and curcumin [255]. In particular, GPR97 can be activated by glucocorticoid stress hormones such as cortisol (hydrocortisone), cortisone and 11-deoxycortisol, and has recently been the first aGPCR to reveal its 7TM domain organisation in complex with the heterotrimeric G protein G_o and two glucocorticoids, the anti-inflammatory drug beclomethasone and the endogenous steroid hormone cortisol [184].

1.4.3.1 GPR97 neutrophil surface expression in inflammation

In order to understand the role of GPR97 in human neutrophils, Prof. Hsi-Hsien Lin (Chang Gung University, Taiwan) led a new study from which some of the results are described thereafter until Section 1.4.3.4. The work is currently under revision in the *Nature Communications* journal and available as a preprint online [256].

To test and confirm whether GPR97 is upregulated in neutrophils of diverse inflammatory disorders [235], neutrophils from patients suffering from bacterial sepsis and from healthy controls were collected and compared for their level of GPR97 surface expression (Figure 1.5A). Patients with sepsis showed a higher amount of GPR97 present on the cell surface of their neutrophils, which therefore supports the observation that GPR97 is upregulated in activated neutrophils.

Additionally, neutrophils from patients suffering from either GPA or MPA, were collected. GPA and MPA are two distinct autoimmune diseases, both characterised by the presence of ANCAs directed toward PR3 and MPO, respectively. In order to find a correlation between the level of GPR97 surface expression and inflammation, the neutrophils were subsequently sorted into two groups, which were determined by the patient being in the active or remission stage of the disease. Interestingly, much higher GPR97 surface expression was detected in neutrophils from patients with signs of active disease than in those from patients in remission or from healthy controls

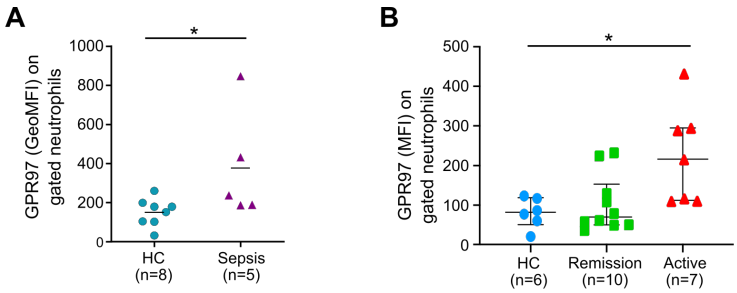


Figure 1.5: GPR97 expression during inflammatory disorders. Flow cytometry analyses comparing GPR97 surface expression levels by measuring the geometric mean fluorescence intensity (geoMFI) or mean fluorescence intensity (MFI) on the surface of gated neutrophils (mean ± SEM). **(A)** A higher GPR97 surface expression could be detected in bacterial sepsis patients compared to healthy controls (HC) (*p < 0.05). **(B)** A higher GPR97 surface expression could be detected in patients with signs of active disease compared to patients in remission stage and/or healthy controls (HC) (*p $<$ 0.05).

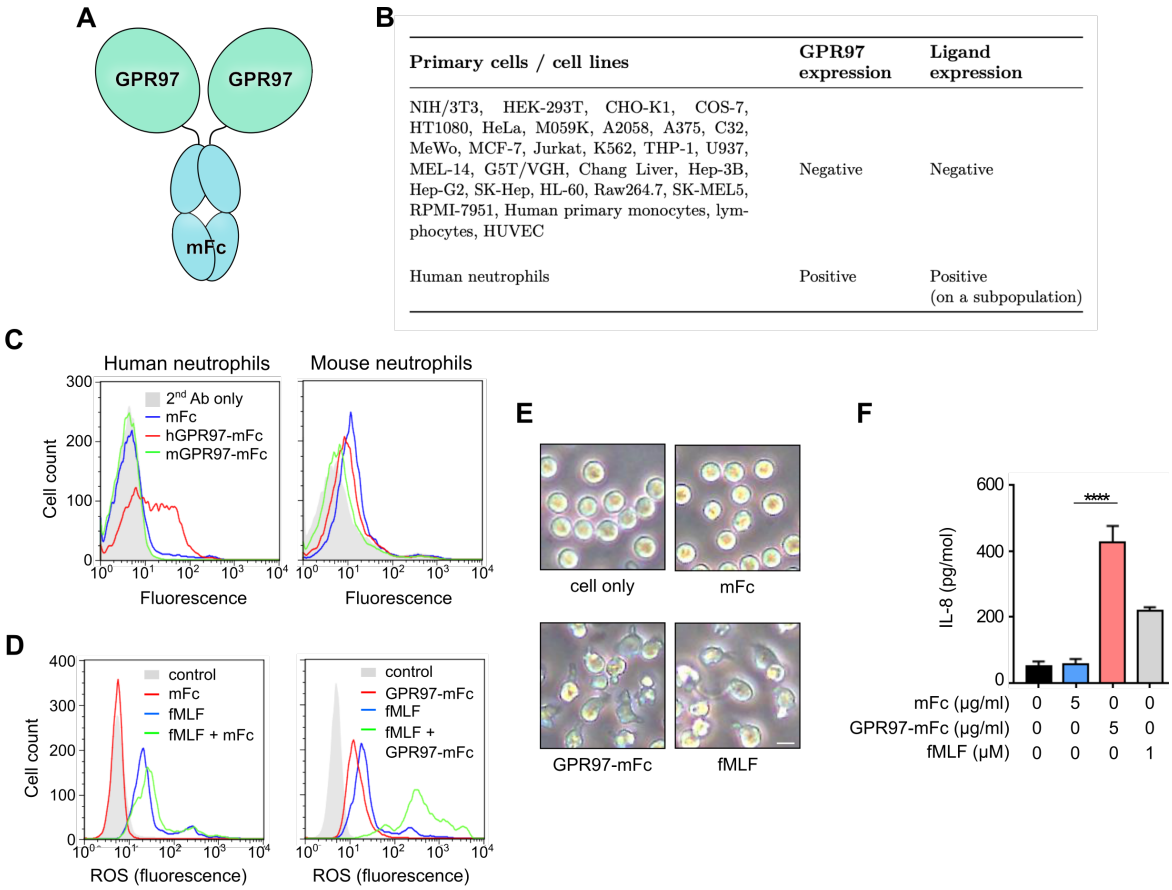
(Figure 1.5B), therefore hinting at a positive correlation between the level of GPR97 surface expression and disease activity.

1.4.3.2 GPR97 ligand(s) on the surface of neutrophils

In order to explore the role of GPR97 in neutrophil activation, the presence of a potential ligand was investigated by screening a panel of cell lines and primary cells using a fluorescence-activated cell sorting (FACS)-based ligand-binding assay [257]. To this end, a chimeric GPR97-mFc protein containing the human GPR97 extracellular region (residues 1-270) fused to a mouse fragment crystallisable (mFc) was engineered and used as a probe (Figure 1.5A).

From over 30 different cell samples screened, one positive binding signal was observed when GPR97-mFc was incubated with neutrophils (Figure 1.5B). Intriguingly, the putative ligand displayed a bimodal expression pattern, as it seemed to be solely expressed on the surface of a subpopulation of human neutrophils (Figure 1.5B,C). In addition, this interaction could not be detected when human neutrophils were probed by a mouse GPR97-mFc probe, nor when mouse neutrophils were probed by either human or mouse GPR97-mFc probes (Figure 1.5C), therefore suggesting that the interaction between GPR97 and its putative ligand is specific to human neutrophils.

Additional experiments have shown that incubation of neutrophils with GPR97-mFc triggered a robust activation, which resulted in several distinct phenotypic manifestations



Primary cells / cell lines	GPR97 expression	Ligand expression
NIH/3T3, HEK-293T, CHO-K1, COS-7, HT1080, HeLa, M059K, A2058, A375, C32, MeWo, MCF-7, Jurkat, K562, THP-1, U937, MEL-14, G5T/VGH, Chang Liver, Hep-3B, Hep-G2, SK-Hep, HL-60, Raw264.7, SK-MEL5, RPMI-7951, Human primary monocytes, lymphocytes, HUVEC	Negative	Negative
Human neutrophils	Positive	Positive (on a subpopulation)

Figure 1.6: GPR97 ligand screening. **(A)** Schematic of the chimeric GPR97-mFc probe containing the human GPR97 extracellular region (residues 1-270) fused to a mouse fragment crystallisable (mFc). **(B)** List of cell samples screened for the expression of GPR97 and the putative GPR97-ligand. **(C)** FACS-based ligand-binding assays showing the GPR97-ligand interaction is specific to human neutrophils. The experiments were performed in purified human and mouse neutrophils using human and mouse GPR97-mFc probes as indicated. **(D)** Neutrophils incubated with GPR97-mFc showed an upregulation of ROS production. mFc and fMLF were used as negative and positive controls, respectively. **(E)** Representative images of neutrophils observed by inverted light microscopy. Neutrophils were incubated with either GPR97-mFc, mFc as negative control or N-formyl-Met-Leu-Phe (fMLF), which is a potent neutrophil activator as positive control. Scale bar is 20 µm. **(F)** Neutrophils incubated with GPR97-mFc showed an upregulation of IL-8 production (mean ± SEM, ****$p < 0.0001$). mFc and fMLF were used as negative and positive controls, respectively.

including a change of cell shape as well as enhanced production of ROS and IL-8 (Figure 1.5D-F).

As a result, there seems to exist a GPR97-induced neutrophil activation which may be caused by interaction of GPR97 with a ligand present on the surface of a subpopulation of neutrophils.

1.4.3.3 Identification of GPR97 ligand(s)

So far, known ligands of other members of the aGPCR family have included lipids, glycosaminoglycans, soluble proteins as well as membrane proteins [148]. As the interaction of GPR97 with its ligand appeared to be human- and neutrophil-specific, the existence of a protein ligand was first investigated. To this end, far-western blot analyses using the GPR97-mFc probe on neutrophil extracts were performed and revealed three distinct signals in the membrane fraction, thus suggesting the ligand to be a membrane-associated protein. More specifically, there was one prominent signal at ∼30 kDa, and two weaker ones at ∼50 and ∼140 kDa (Figure 1.7A).

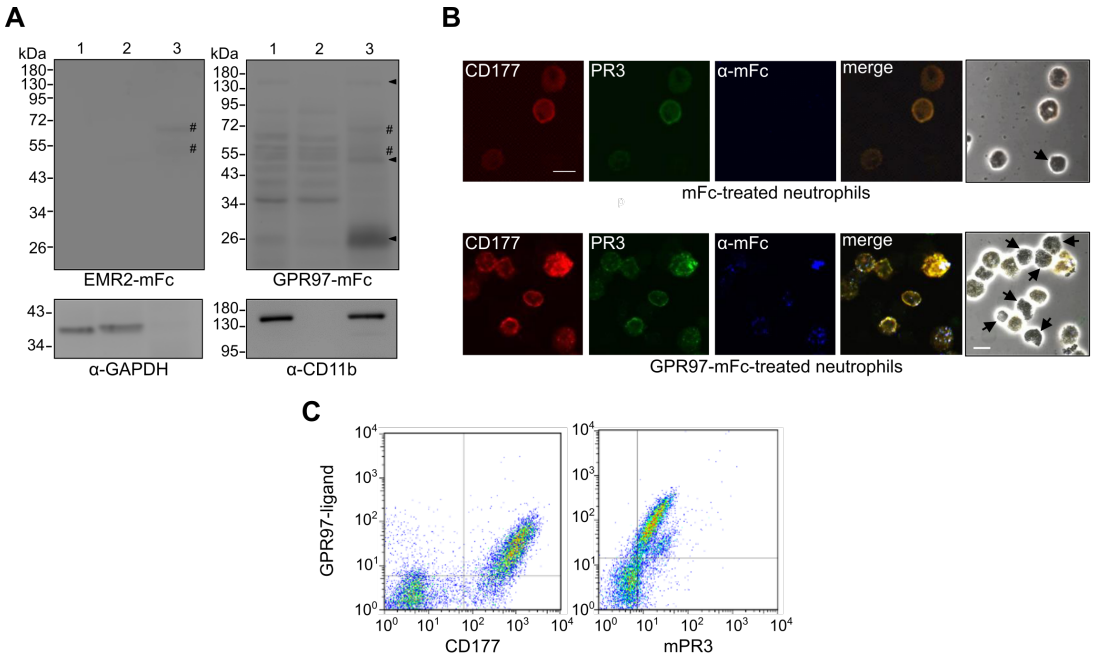


Figure 1.7: Identification of CD177 and PR3 as potential GPR97 ligands. (A) Far-western blot analyses on neutrophils using the GPR97-mFc probe. Lanes 1, 2 and 3 correspond to the total cell lysate, the cytoplasmic fraction and the membrane fraction, respectively. The EMR2-mFc probe was used as a negative control, and the α-GADPH and α-CD11b mAbs were used to detect the corresponding protein markers of cytoplasmic and membranous fractions, respectively. Putative ligands are indicated by black arrowheads, # corresponds to non-specific signals. **(B)** Representative images of neutrophils incubated with either GPR97-mFc or mFc as negative control. Cells were stained for PR3 (green) and CD177 (red). The mFc-fusion protein probes were detected using an α-mFc mAb (blue). $CD177^{neg}$ neutrophils are indicated by arrows. GPR97-mFc is shown to co-localise with $CD177^{pos}/PR3^{high}$ neutrophil subset. Scale bar is 20 µm. **(C)** Dot-plots of neutrophils double-stained with GPR97-mFc and either α-CD177 or α-PR3 mAbs show that the GPR97-mFc probe only binds to $CD177^{pos}$ and $PR3^{high}$ neutrophils. mPR3 stands for membrane-bound PR3.

Interestingly, the only proteins that have been reported to display a bimodal expression pattern on the surface of neutrophils are Olfactomedin 4 (OLM4, also known as GW112

and hGC-1) [258], CD177 (also known as NB1, PRV-1 and HNA-2a) [30] and PR3 (also known as myeloblastin, p29b, AGP7) [26], whose theoretical molecular weight are 56 kDa, 42 kDa and 28 kDa, respectively. Out of those three candidates, only CD177 and PR3 are known to be localised on the membrane: CD177 through a GPI-anchor, and PR3 mainly via direct interaction with its receptor CD177 [30, 29].

GPR97 interaction with PR3 and/or CD177

To further investigate CD177 and PR3 as potential ligands for GPR97, neutrophils were incubated with either GPR97-mFc or mFc, and were then stained for CD177, PR3 as well as mFc presence on the cell surface. Not only could a fluorescent signal corresponding to the presence of GPR97-mFc on the cell surface be observed, but GPR97-mFc also seemed to co-localise with CD177 and PR3 (Figure 1.7B).

Additionally, a repeat of the FACS-based ligand-binding assay was performed, but this time using neutrophils which were sorted into either $CD177^{pos}$ and $CD177^{neg}$ subpopulations, or into $PR3^{high}$ and $PR3^{low}$ subpopulations. As the GPR97-binding signal could only be observed in $CD177^{pos}$ and $PR3^{high}$ neutrophil subpopulations (Figure 1.7B-C), this further supported CD177 and/or PR3 to be potential GPR97 ligands.

In order to determine if both or only one of those proteins is the GPR97 binding partner, far-western blot analyses using the GPR97-mFc probe were repeated using either CD177-transfected HEK293T cell lysates or PR3 purified from neutrophils. The resulting blot revealed specific binding of GPR97-mFc to PR3 but not CD177, suggesting that PR3 is the GPR97 ligand and that the previously observed interaction of the GPR97-mFc probe with the $CD177^{pos}$ subpopulation was likely mediated through the CD177-bound PR3 present on the cell surface (Figure 1.8A).

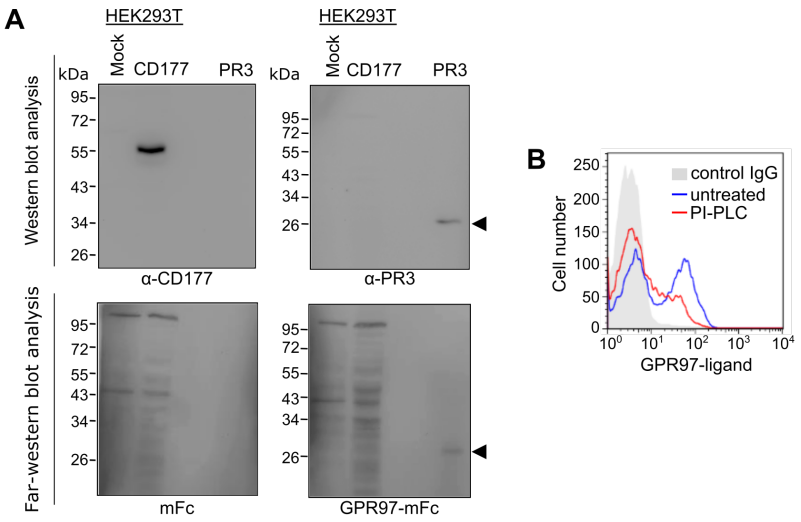


Figure 1.8: Identification of PR3 as potential GPR97 ligand. (A) Western-blot and far-western blot analyses of HEK293T cell lysates from mock- and CD177-transfected cells as well as neutrophil-purified PR3. Use of α-CD177 and α-PR3 mAb confirmed the presence of CD177 and PR3 in the analysed samples (top panel). Far-western blot analysis only shows a specific-signal detected by the GPR97-mFc probe corresponding to PR3 (black arrowheads), but not CD177. **(B)** The FACS-based ligand-binding analysis of neutrophils pre-treated with PI-PLC shows that the GPR97-mFc binding is lost after PI-PLC pre-treatment.

To confirm this, neutrophils were treated with phosphoinositide phospholipase C (PI-PLC) in order to cleave most GPI linkers before a FACS-based ligand-binding assay using the GPR97-mFc probe was repeated (Figure 1.8B). The resulting decrease in GPR97-mFc binding to the surface of neutrophils indeed confirms that GPR97 interacts with CD177-bound PR3.

1.4.3.4 GPR97 interaction with PR3 on the surface of neutrophils

GPR97 direct interaction with CD177-bound PR3

To assess whether GPR97 and PR3 are interacting directly, a FACS-based ligand-binding assay using HEK293T cells transfected with full-length GPR97 and PR3 as the probe was carried out. Surprisingly, very little to no GPR97 binding signal was detected in comparison with control cells that were transfected with full-length CD177 (Figure 1.9A). In parallel, an ELISA-like protein-protein binding assay was performed, where PR3 was added at different concentrations into wells coated with either GPR97-mFc, CD177-mFc or the negative control GPR56-mFc (Figure 1.9B). Although a much weaker PR3 binding signal was observed for GPR97 than with CD177, the GPR97-PR3 binding signal was

still significantly higher than the negative controls. This assay thus confirmed a direct, but weak GPR97-PR3 interaction. Notably, this weak GPR97-PR3 direct interaction could explain why very little signal was observed in the FACS-based ligand-binding assay.

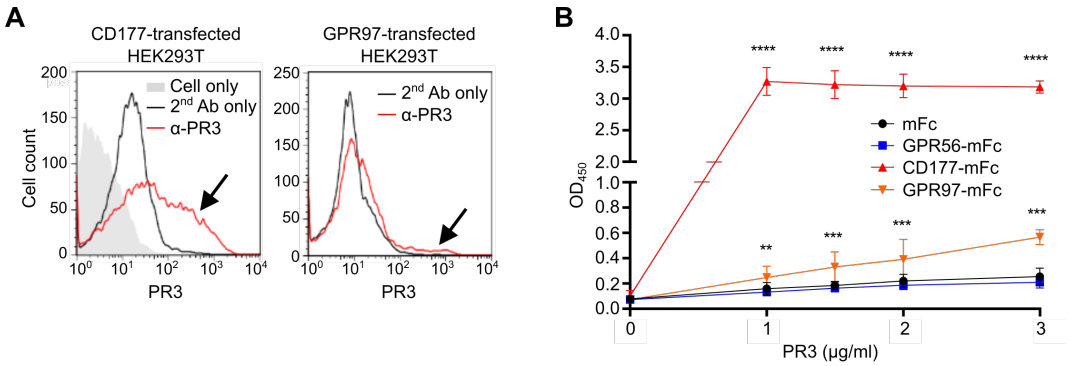


Figure 1.9: A weak GPR97-PR3 direct interaction. **(A)** CD177- or GPR97-transfected HEK293T cells were subjected to a FACS-based PR3-binding assay. The PR3-specific binding is indicated by a black arrow. **(B)** ELISA-like PR3-binding experiments were performed using either GPR97-mFc, CD177-mFc as positive control or GPR56-mFc as negative control (n = 6 repeats).

Additionally, GPR97-induced neutrophil activation (measured by IL-8 upregulation) could only be observed either in the unsorted neutrophil pool or in the CD177^{pos}-sorted subpopulation (Figure 1.10), thus further supporting the interaction of GPR97 with CD177-bound PR3 as well as the role of GPR97-PR3 interaction in neutrophil activation.

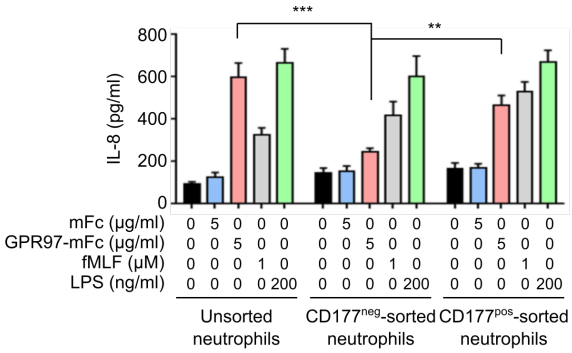


Figure 1.10: Neutrophil activation upon GPR97-PR3 interaction. Comparison of IL-8 production following neutrophil incubation with either GPR97-mFc, mFc as negative control, or fMLF and lipopolysaccharides (LPS) as positive controls (mean ± SEM). Unsorted neutrophils or CD177^{pos} neutrophils showed a significantly higher IL-8 production compared to CD177 ^{neg} neutrophils (**p < 0.01, ****p < 0.0001).

GPR97 effect on PR3 proteolytic activity

Previous studies have shown that CD177-bound PR3 possessed a decreased serine protease activity compared to free PR3, and therefore suggested that the PR3-CD177 binding interface would involve the PR3 catalytic cleft [124].

In order to determine if GPR97 binding to PR3 would result in a similar effect on PR3 proteolytic activity, the enzymatic activity of PR3 was determined *ex vivo* by incubating neutrophils with a PR3-specific FRET-based substrate [259]. Surprisingly however, incubation of neutrophils with GPR97-mFc resulted instead in an increased PR3 enzymatic activity (Figure 1.11A).

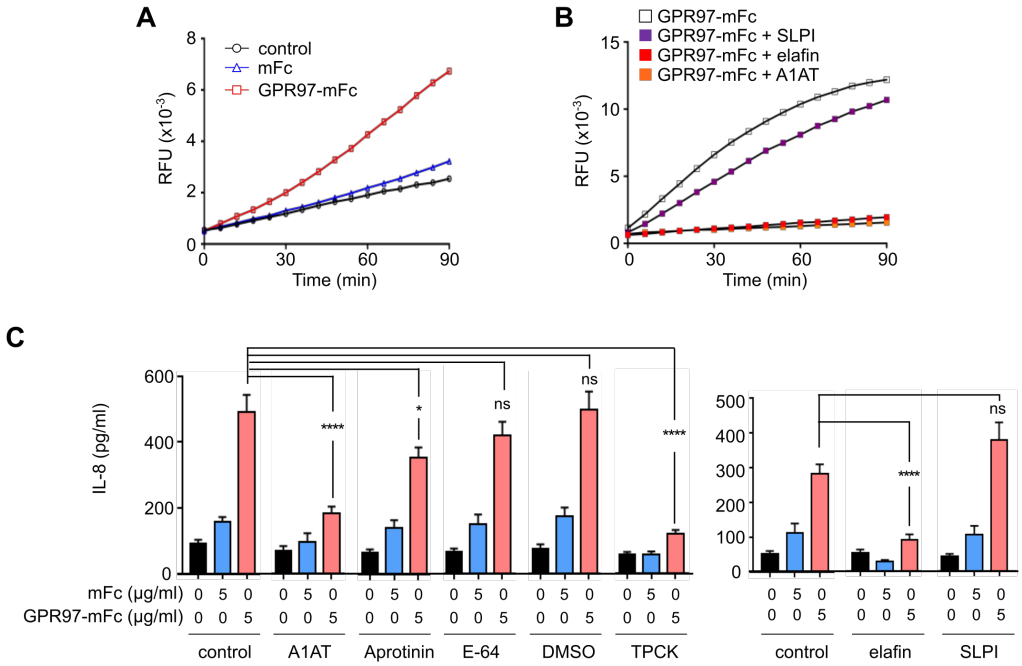


Figure 1.11: GPR97 effect on PR3 proteolytic activity. (A) Incubation of neutrophils with GPR97-mFc resulted in an increase of PR3 enzymatic activity *ex vivo*. **(B)** Incubation of neutrophils with a PR3 inhibitor effectively abolished the increased *ex vivo* PR3 enzymatic activity due to GPR97-mFc. **(C)** Comparison of IL-8 production by neutrophils following their incubation with GPR97-mFc and the indicated protease inhibitor (n = 4 repeats, mean ± SEM). Incubation of neutrophils with PR3 inhibitors such as A1AT, aprotinin, TPCK and elafin resulted in a significant decrease of IL-8 production (*$p < 0.05$; ****$p < 0.0001$; ns, non-significant).

To study whether this increased proteolytic activity was linked to GPR97-induced neutrophil activation, neutrophils were incubated with various inhibitors before addition of GPR97-mFc. The inhibitors tested included broad-spectrum serine protease inhibitors such as A1AT, aprotinin and N-tosyl-L-phenylalanyl chloromethyl ketone (TPCK), endogenous NSP-specific inhibitors such as elafin and secretory leukocyte peptidase

inhibitor (SLPI), and the cysteine protease inhibitor E-64. While elafin is known to inhibit all NSPs including PR3, SLPI only inhibits HNE and CG but not PR3 (Figure 1.11B). Interestingly, GPR97-induced neutrophil activation (measured by IL-8 production) was significantly attenuated in the presence of A1AT, aprotinin, TPCK and elafin but not in the presence of SLPI and E-64 (Figure 1.11C). In other words, GPR97-induced neutrophil activation appears to be dependent on PR3 protease activity, which may be potentiated through interaction with GPR97.

1.5 Research aims

Neutrophils are potent immune cells which act as the first line of defence against invading pathogens and whose activities are tightly regulated in order to prevent widespread collateral damage to surrounding tissue. Cell surface receptors are particularly important as they mediate various processes involved in the regulation of neutrophil activation and potentiation.

GPR97 and CD177 are two neutrophil-specific receptors, whose role in human neutrophils remain mainly unknown. Initial work led by Prof. Hsi-Hsien Lin indicate GPR97 may interact with PR3 and induce neutrophil activation. In addition, the presence of PR3 on the surface of neutrophils is unusual for an NSP and has been strongly associated with an autoimmune disease characterised by the presence of ANCAs, called GPA.

The aim of this D.Phil. project was therefore to provide molecular insights into how those proteins are arranged and interact together on the surface of neutrophils. To achieve this, several objectives were set, including:

- Producing GPR97 ECD and solving its crystal structure.
- Biophysically assessing GPR97 interaction with PR3 and functionally characterising the outcome of this interaction in neutrophils.
- Producing and obtaining high quantities of soluble PR3.
- Producing CD177 ECD and solving its crystal structure.
- Biophysically and structurally characterising CD177 interaction with PR3

- Identifying relevant surfaces that may be important for PR3 recognition in pathogenesis of GPA

Chapter 2
Materials & Methods

2.1 Molecular Cloning

2.1.1 Genes

Details of the DNAs encoding GPR97, PR3 and CD177 used in this work are detailed in Table 2.1. The genes were either gifts from collaborators or purchased from the Mammalian Gene Collection.

Protein	Species	Uniprot	Gene	
			Source	Accession
GPR97/ADGRG3	*Homo sapiens*	Q86Y34	gift	NM_170776.5
	Mus musculus	Q8R0T6	cDNA clone MGC 76424	BC064726.1
Proteinase 3	*Homo sapiens*	P24158	cDNA clone MGC 116835	BC096184.3
CD177	*Homo sapiens*	Q8N6Q3	gift	BC029167.1

Table 2.1: Summary of genes used for sub-cloning. The DNA encoding *hs*GPR97 possesses two single nucleotide polymorphisms (SNP) at nucleotides 1339 (A→G) and 1605 (C→T), resulting in the missense mutation M→V447 and the silent mutation S→S535 respectively.

2.1.2 Generation of constructs

Constructs of GPR97 ectodomain, PR3 and CD177 were generated using the CloneAmp^{TM} HiFi PCR Premix kit (Takara Bio, Cat.#639298). A typical reaction contained 100 ng of plasmid DNA, 7.5 pmol of both the forward and reverse primers and 12.5 µl of the 2X PCR Premix in a total volume of 25 µl. PCRs were performed over 34 thermal cycles (10 sec

of denaturation at 98℃, 15 sec of annealing at 55℃ and 10 sec of extension at 72℃). All primers used for cloning and mutagenesis were purchased from Thermo Fisher Scientific. PCR products were then separated by agarose gel electrophoresis, during which DNA fragments were electrophoresed at 150 volts for 30 min in TAE buffer (10 mM Tris base, 20 mM acetic acid, 1 mM EDTA). Isolated DNA bands were visualised using a Dark Reader® transilluminator and excised for follow-up purification using a gel clean-up kit (Thermo Fisher Scientific, Cat.#12303368).

2.1.3 Expression vectors

Vectors from the pHLsec family were used for expression in human embryonic kidney (HEK) cells [260]. The multiple cloning site (MCS) of members of the pHLsec family comprises four restriction sites of interest, namely *Eco*RI, *Age*I, *Kpn*I and *Xho*I (Figure 2.1). Two variants from the family were used depending on the experiment. For protein expression and purification, the pHLsec vector allows the fusion of a C-terminal Lys-Hexahistidine (His_6) tag to the protein insert. For surface plasmon resonance (SPR) experiments, the pHL-Avitag3 allows the fusion of a C-terminal GSS linker region followed by a biotin ligase (BirA) recognition site [261, 262] and a His_6 tag to the protein insert. Prior to cloning, 20 µg of vector plasmid were digested and linearised using a mixture of at least 10 units of two selected restriction enzymes (New England Biolabs (NEB): *Eco*RI Cat.#0101S, *Age*I Cat.#R0552S, *Kpn*I Cat.#R0142M, *Xho*I Cat.#R0146S) in the appropriate NEB buffer. The restriction reactions were incubated at 37℃ for 6 hours, and the linearised vectors were isolated by agarose gel electrophoresis and purified using a gel kit.

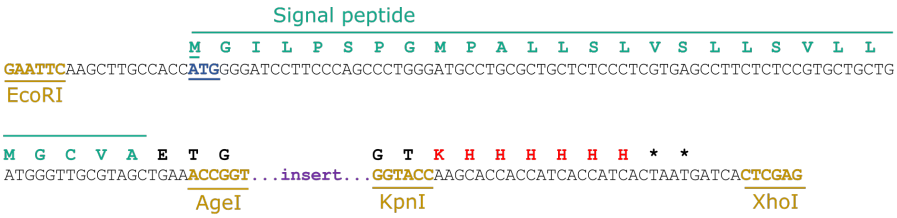


Figure 2.1: MCS of the pHLsec vector. The DNA and translated protein sequence of the signal peptide as well as the C-terminal His_6 tag of the pHLsec vector are indicated in green and red, respectively. The four restriction sites of interest *Eco*RI, *Age*I, *Kpn*I and *Xho*I are highlighted in dark yellow. The start (ATG) and stop codons (TAATGA) are highlighted in bold blue and indicated by black asterisks, respectively.

2.1.4 Generation and selection of recombinant plasmid DNA

Constructs were subcloned into either the *Eco*RI-*Kpn*I or *Age*I-*Kpn*I cloning sites of the pHLsec/pHL-Avitag3 vectors, depending on the presence of a native signal sequence. Constructs with a FLAG (DYKDDDDK) or maltose-binging protein (MBP) tag were subcloned into the *Eco*RI-*Xho*I cloning sites of the pHLsec vector.

In-Fusion® reactions were carried out using the In-Fusion® HD Cloning kit (Takara Bio, Cat.#121416). A typical reaction contained 50 ng of purified linearised vector, 50 ng of purified PCR product(s) and 1 µl of the 5X In-Fusion HD Enzyme Premix in a total volume of 5 µl.

The In-Fusion™ reactions were incubated at 50℃ for 15 min. 2.5 µl of the reaction mixture were then used to transform Stellar™ competent cells (Takara Bio, Cat.#636763). Cells were incubated on ice for 30 min, heat shocked at 42℃ for 45 sec and returned to ice for at least 2 min before spreading onto lysogeny broth (LB) agar plates (Merck, Cat.#1.10285.5000) containing the required antibiotic (100 µg/ml ampicillin, 50 µg/ml kanamycin). Plates were incubated at 37℃ overnight in a static incubator or at room temperature (r.t.) over two days. Single colonies were then selected and used to inoculate 5 ml of ampicillin-containing LB media. Cultures were incubated at 37℃ overnight in an orbital shaker. The recombinant plasmid DNA were extracted from each culture using a plasmid DNA miniprep kit (Thermo Fisher Scientific, Cat.#11932392). Each clone was analysed by Sanger sequencing (Source Biosciences). The correct clones were further transformed into Subcloning Efficiency™ DH5α™ competent cells in order to increase the DNA yield after extraction.

All constructs used throughout this thesis are summarised in Appendix tables B.1 and B.2.

2.2 Protein production

2.2.1 HEK293 cell lines and control of N-linked glycosylation

Two derivatives of the HEK293 cell line were used for production and secretion of recombinant proteins in this work.

HEK293T is a stable adherent cell line expressing the SV40 large T antigen [263, 264], and has been widely used for its fast growth rate and capacity to highly express recombinant proteins following transient transfection. Glycoproteins produced by the HEK293T cell line appear to have large and heterogeneous N-linked glycans, which consist of multi-antennary N-glycans of the complex-type (Figure 2.2) that are non sensitive to endoglycosidases (Endo) such as Endo H [265] or Endo F_1 [266]. This heterogeneity is partly due to over-expression overwhelming the cells' processing capacity.

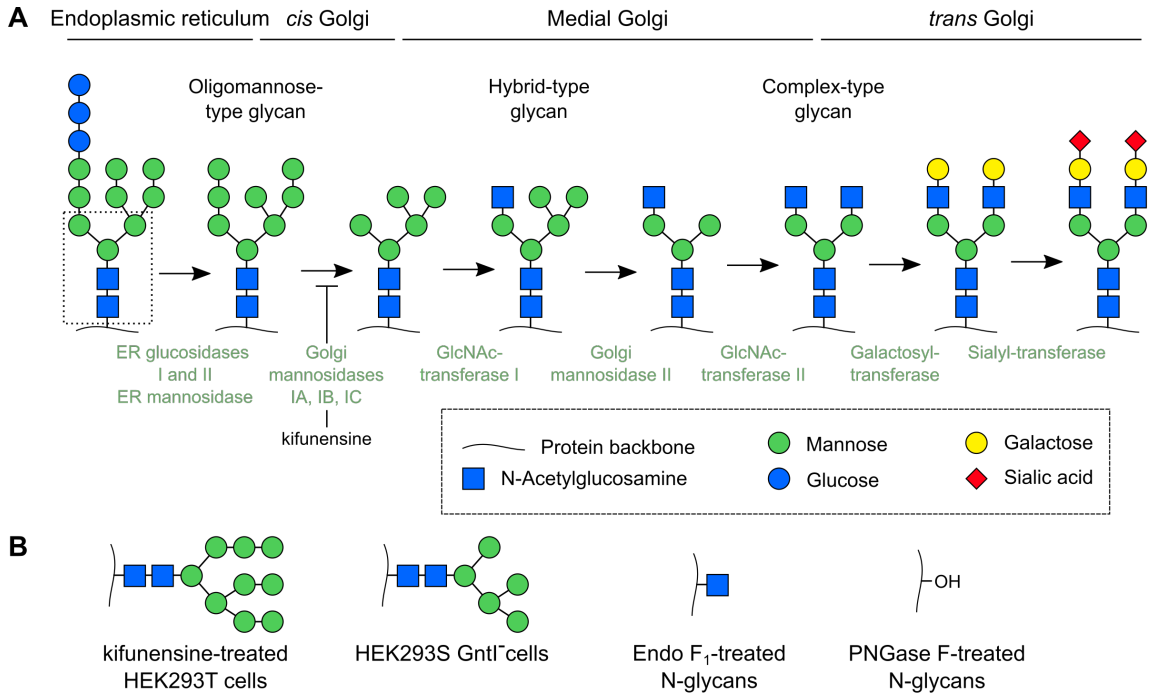


Figure 2.2: Control of N-linked glycosylation in HEK293 cells. N-linked glycosylation occurs at the amide nitrogen of asparagine side chains from the Asn-Xaa-Ser/Thr (where Xaa is any other amino acid except for Pro) consensus motif. **(A)** Schematic diagram of the primary pathway from an initial oligomannose N-linked glycan to a complex N-glycan. This pathway involves ordered steps that occur in different compartment within the cell, from the endoplasmic reticulum (ER) to the Golgi. All eukaryotic N-linked glycans share a common core structure consisting of $Man_3GlcNAc_2$ (highlighted by the black box), but then differ in their terminal elaboration extending from this core. These can be classified into three different categories: (i) oligomannose (or high-mannose), (ii) hybrid and (iii) complex. **(B)** Comparison between N-linked glycans produced by kifunensine-treated HEK293T cells, 293S GnTI$^-$ cells as well as after treatment with the endoglycosidases Endo F_1 and PNGase F. Endo F_1 cleaves between the two GlcNAc residues of the core of oligomannose- and hybrid-type N-glycans only, leaving a single GlcNAc residue at each glycosylation site. PNGase F cleaves between the innermost GlcNAc and asparagine residues of N-glycans of all types, resulting in aspartate residues. All glycans are depicted following the Symbol Nomenclature for Glycans (SNFG) [267, 268].

Using an N-glycosylation processing inhibitor such as kifunensine alters the glycosylation pathway to only produce N-glycans of the oligomannose-type. Although, this approach

still yields heterogeneous glycoproteins with a mixture of N-linked $Man_{6\rightarrow9}GlcNAc_2$ N-glycans, those are now highly sensitive to endoglycosidases, therefore allowing homogeneous deglycosylation of glycoproteins [269].

N-acetylglucosaminyltransferaseI-deficient HEK293S (HEK293S GnTI$^-$) is a suspension growth-adapted cell line which lacks the *N*-acetyl glucosamine transferase I (GnTI) enzyme [270], and therefore limits N-linked glycosylation of expressed glycoproteins to a pattern dominated by the $Man_5GlcNAc_2$ N-glycan [271, 269] (Figure 2.2). However, expression levels in HEK293S GnTI$^-$ cells are only 10%-50% of that in HEK293T cells [269]. HEK293S GnTI$^-$ cells therefore require a harvest ten days after transfection, as opposed to the four days necessary before harvest of HEK293T cells.

2.2.2 HEK293 cell culture maintenance

Both HEK293T and 293S GnTI$^-$ cells were cultured as adherent monolayers in T75 or T175 flasks (Greiner Bio-One, Cat.#658175, #660175). Cells were grown in Dulbecco's Modified Eagle Medium (DMEM, Gibco, Cat.#41966052) supplemented with 10% Fetal Bovine Serum (FBS, Gibco, Cat.#10437028), 2 mM L-Glutamine (L-Glu, Thermo Fisher Scientific, Cat.#25030024) and 10 mM non-essential amino acids (NEAA, Gibco, Cat.#11140035) at 37℃ in a 5% CO_2 atmosphere. Cells were passaged at 90-100% confluency by: (i) removal of former DMEM (ii) washing with a phosphate buffer saline (PBS) solution (Lonza, Cat.#LZBE17-516F) (iii) treatment with 0.05% Trypsin-EDTA (Gibco, Cat.#25300054) at 37℃ for 4 min or until detached (iv) resuspension into fresh DMEM to the desired density.

2.2.3 Small-scale transient protein expression

After passage from a T75 flask at 90-100% confluency, 2 ml of resuspended HEK293 cells were seeded into each well of a six-well plate (Greiner Bio-One, Cat.#657160). The next day, 3 µg of purified DNA were mixed with polyethylenimine (PEI) (CAS #9002-98-6, Sigma-Aldrich, Cat.#408727) at a DNA:PEI mass ratio of 1:3 in 100 µl of Opti-MEM^{TM} (Gibco, Cat.#31985047) to make the transfection cocktail [260]. For co-transfection of two proteins, 2 µg of each DNA were used and mixed with PEI at a DNA:PEI mass

ratio of 1:3. The transfection cocktail was then incubated at r.t. for 10 min to allow DNA-PEI complex formation before being carefully added to the side of each well of the six-well plate. After two days, the culture medium from each well was clarified by centrifugation at 15,000 rpm for 10 min using a benchtop centrifuge (Beckman Coulter). The level of protein expression was assessed by western blotting (see Section 2.3.2).

2.2.4 Large-scale transient protein expression

2.2.4.1 Roller bottle preparation and transfection

For a typical large-scale protein expression, twelve T175 flasks at 90-100% confluency were used to seed twelve roller bottles (Greiner Bio-One, Cat.#681070), each containing 250 ml of DMEM supplemented with 10% FBS, 2 mM L-Glu and 10 mM NEAA. Roller bottles were incubated at 37℃ in a 5% CO_2 atmosphere in a Wheaton roller bottle incubator for three days (HEK293T) or four days (HEK293S GnTI$^-$) prior to transfection. Plasmid DNA purified from a gigaprep kit (Invitrogen, Cat.#K210009XP) was prepared as a 1 mg/ml stock and was firstly cleared from any residual impurity by mixing 100 µl chloroform per 1 ml DNA. After centrifugation at 6,000 rpm for 30 min, the DNA supernatant was carefully recovered and added (0.5 mg / roller bottle) to serum-free DMEM supplemented with 2 mM L-Glu and 10 mM NEAA (50 ml / roller bottle) before addition of PEI (1 mg/ml) at a DNA:PEI mass ratio of 1:2. This transfection cocktail was incubated at r.t. for 10 min to allow DNA-PEI complex formation. Media from each roller bottle was then replaced with 200 ml of DMEM supplemented with 2% FBS, 2 mM L-Glu and 10 mM NEAA plus 50 ml of the transfection cocktail. If using kifunensine-treated HEK293T cells, kifunensine (CAS #109944-15-2, Cayman Chemical, Cat.#10009437) (0.25 mg / roller bottle) was added in the transfection cocktail just before addition to the roller bottles. Cells were incubated for four days (HEK293T) or ten days (HEK293S GnTI$^-$) for protein production [260, 269].

2.2.4.2 Cell culture medium harvesting

Cell culture medium, typically up to 3 L, containing the secreted proteins was harvested and cell debris were removed by centrifugation at 7,000 rpm, 4℃ for 30 min. The

medium was further clarified and sterilised by passing through a 0.22 µm filter (Merck, Cat.#S2GPT05RE) before either storage at 4℃ or processing to protein purification.

2.2.5 Protein purification

2.2.5.1 Immobilised metal-ion affinity chromatography

The culture medium was diafiltered into PBS, 20 mM Tris pH 7.5 and 150 mM NaCl using a tangential flow filtration device (Pall) and concentrated to ∼500 ml before immobilised metal-ion affinity chromatography (IMAC). After supplementing with 10 mM imidazole pH 8.0, the concentrated medium was applied to a HisTrap™ HP column (Cytiva, Cat.#GE17-5248-02) which would retain the His_6-tagged recombinant proteins. The column was washed with 100 ml of 10 mM imidazole IMAC buffer (20 mM Tris pH 7.5, 300 mM NaCl) and 100 ml of 40 mM imidazole IMAC wash buffer. Bound proteins were eluted with 500 mM imidazole IMAC elution buffer in 0.5-1.5 ml fractions.
In some cases, the eluted fractions were pooled and incubated with 100 µl of Endo F_1 at 1 mg/ml at r.t. for a few hours and then left at 4℃ overnight [272]. The deglycosylated mixture was then diluted with IMAC buffer to 10 mM imidazole concentration and purified once more via IMAC as described above to remove Endo F_1.

2.2.5.2 Size-exclusion chromatography

IMAC fractions were assessed for protein content by a quick non-quantitative Bradford assay, as well as by SDS-polyacrylamide gel electrophoresis (SDS-PAGE, see Section 2.3.1). The fractions containing the most proteins were pooled to a final volume of 4-4.5 ml before loading into a 5 ml tubing loop connected to a HiLoad® 16/600 Superdex® 200 column (Cytiva, Cat.#GE28-9893-35) previously equilibrated with size-exclusion chromatography (SEC) buffer (20 mM Tris pH 7.5, 300 mM NaCl). 1.5 ml fractions were collected and analysed by SDS-PAGE. Fractions containing pure protein were pooled and concentrated using 10,000 MWCO Amicon® Ultra centrifugal filters (Merck, Cat.#UFC800324 and #UFC900324). Pure concentrated protein was then either aliquoted and stored at -80℃ or used immediately. Protein concentration was measured in triplicates using

a NanoDrop™ spectrophotometer device and the theoretical extinction coefficient ϵ calculated by the ProtParam tool on the ExPASy Server [273].

2.2.6 Protocol adaptation for PR3 expression and purification trials

2.2.6.1 PR3 expression in CD177-conditioned medium

Purified CD177^{ecto}was sterilised using a 0.22 µm centrifugal filter (Merck, Cat.#UFC30GV0S). In small-scale experiments, 0.18 nmol (7.6 µg) was added into the well before addition of the transfection cocktail. In large-scale experiments, CD177^{ecto}(5 nmol / roller bottle) was added in the transfection cocktail just before addition to the roller bottles.

2.2.6.2 PR3 purification trials

Buffer exchange

In some cases, PR3-containing culture medium was subject to dialysis against PBS at 4°C for 48 h using 12-14,000 MWCO Visking membranes (Scientific Laboratory Supplies, Cat.#TUB2024). The dialysed culture medium supplemented with 10 mM imidazole was either applied to a HisTrap™ HP column or was incubated with 3 ml/L of a 50% slurry of His·Bind® Nickel-iminodiacetic acid (Ni-IDA) beads (Merck, Cat.#69670), previously equilibrated in 10 mM imidazole IMAC buffer, at 4°C overnight under gentle stirring. PR3 bound to a HisTrap™ HP column was washed and eluted as described previously. When using loose Ni-IDA beads, samples were collected the next day, poured into an Econo-Column® (Bio-Rad) and washed with 100 ml of 10 mM imidazole IMAC buffer and 100 ml of IMAC wash buffer. Bound PR3 was eluted in IMAC elution buffer in 150-200 µl fractions. Fractions were assessed for protein presence using a NanoDrop™ spectrophotometer device and those containing the most proteins were pooled together, concentrated to a final volume of 450 µl and loaded into a 500 µl tubing loop connected to a Superdex® 200 10/300 GL column (Cytiva, Cat.#GE17-5175-01).

Stabilisation using detergents

When purification trials were performed in presence of n-Dodecyl-β-D-Maltopyranoside (DDM, Glycon Biochemicals, Cat.#D97002), 0.03% DDM, corresponding to 3.9 times

its critical micelle concentration (cmc), was added to the harvested culture medium prior to centrifugation and filtration. After incubation at r.t. for 1 h, the culture medium was dialysed against PBS, incubated with loose Ni-IDA beads and eluted as described above. IMAC and SEC buffers were all supplemented with 0.01% DDM, corresponding to 1.3·cmc.

Co-purification with CD177^{ecto}

PR3-containing culture medium and CD177-containing culture medium, both obtained after centrifugation and filtration, were mixed in a 2:1 volume ratio. The mixture was filtered and incubated at r.t. overnight before purification as described in Section 2.2.5.

2.3 Protein analysis and detection

2.3.1 SDS-polyacrylamide gel electrophoresis

Protein samples were mixed with 5X loading dye (250 mM Tris pH 7.5, 10% sodium dodecyl sulfate (SDS), 50% glycerol, 0.05% bromophenol blue, 5% β-mercaptoethanol) and boiled at 95℃ for 2 min. The samples were then briefly spun down and 20 µl were loaded into each well of a NuPAGE™ 4-12% Bis-Tris gel (Invitrogen, Cat.#NP0336). 10 µl of a protein ladder (Invitrogen, Cat.#10747012) was loaded in the first lane of each gel to help with protein analysis and molecular weight determination. Gels were run in NuPAGE™ MES SDS running buffer (Invitrogen, Cat.#NP0002) at 150-180 volts for 45-60 min and stained with InstantBlue™ (Sigma-Aldrich, Cat.#ISB1L).

2.3.2 Western blotting

As with SDS-PAGE, the protein samples were mixed with 5X loading dye, boiled at 95℃ for 2 min and briefly spun down before being loaded into each well of a NuPAGE™ 4-12% Bis-Tris gel. 10 µl of a pre-stained protein ladder (Invitrogen, Cat.#10748010) supplemented with 0.2 µl of a His-tagged protein ladder (Invitrogen, Cat.#10747012) were loaded in the first lane of each gel to help with protein analysis and assess transfer efficiency. Following a standard SDS-PAGE, proteins from an unstained gel were transferred and blotted onto a 0.45 µm nitrocellulose membrane

(Cytiva, Cat.#GE10600013). The membrane was soaked for at least 20 min in the NuPAGE™ transfer buffer (Invitrogen, Cat.#NP0006) before assembling the transfer cassette and running the transfer at 30 volts for 70 min.

To reduce non-specific interactions, the membrane was incubated in PBS supplemented with 1% (v/v) Tween 20 (PBST) and 5% skim milk (Thermo Fisher Scientific, Cat.#16694685) at r.t. for at least 30 min. His-tagged proteins were then probed with the primary monoclonal mouse anti-His antibody (α-His mAb, Qiagen, Cat.#34660) used at a 1:1000 dilution in PBST supplemented with 3% bovine serum albumin (BSA, Sigma-Aldrich, Cat.#A7906) either at r.t. for at least 1 h or at 4℃ overnight. A monoclonal mouse anti-GPR97 antibody (α-GPR97 mAb) generated by Prof. Hsi-Hsien Lin's group was also used as a primary antibody to probe for GPR97 constructs. The membrane was washed three times with PBST before incubation with the secondary polyclonal goat anti-mouse antibodies conjugated to horseradish peroxidase (α-mouse-HRP pAb, Sigma-Aldrich, Cat.#A0168) used at a 1:10000 dilution in PBST supplemented with 5% skim milk at r.t. for 1 h. The membrane was finally washed three times with PBST and the signals were captured on autoradiography films (Cytiva, Cat.#GE28-9068-35) using ECL™ detection reagents (Cytiva, Cat.#GERPN2106).

2.3.3 PNGase F deglycosylation assay

Glycosylated proteins were incubated with PNGase F (NEB, Cat.#P0704S) in accordance with the protocol supplied with the enzyme. ∼20 µg of protein were incubated with PNGase in denaturing reaction conditions for analysis on SDS-PAGE gel, and in non-denaturing reaction conditions for SEC experiments.

2.3.4 His-tag pull-down assay

Culture media recovered from small-scale expression experiments were clarified by centrifugation at 15,000 rpm, 18℃ for 10 min using a benchtop centrifuge, filtered using a 0.22 µm syringe filter (Thermo Fisher Scientific, Cat.#10509821) and supplemented with imidazole to a final concentration of 10 mM. 100 µl of a 50% slurry of His·Bind® Ni-IDA beads, pre-equilibrated in 10 mM imidazole IMAC buffer, were added to 1 ml of

culture medium sample and the mixture was left to incubate on a rotating mixer either at 4℃ overnight or at r.t. for 1-2 h. The beads were either left to decant or were spun down at 50 x g for 5 min before carefully removing the culture medium ("flow-through") and replacing it with 1.5 ml of 20 mM imidazole IMAC wash buffer, twice. 200 µl of 500 mM imidazole IMAC elution buffer were then added to the beads, and the eluted proteins were recovered after a gentle spin. In some cases, the beads were further washed with 1.5 ml of IMAC wash buffer before incubation with 100 µl of 5X loading dye.

2.3.5 PR3 recovery assay

PR3-containing culture media was recovered and incubated with either various reagents (Table 2.2) or buffers at various pH. The mixture was allowed to incubate at r.t. for 1 h before being subject to centrifugation at 15,000 rpm, 18℃ for 10 min. The supernatants were collected for western blot analysis.

2.3.6 PR3 pull-down assay

∼114 mg of dry CNBr-activated Sepharose 4B beads (Cytiva, Cat.#17043001) were hydrated and resuspended in 1 mM HCl before addition of 2 mg of $CD177^{ecto}$ purified in 20 mM HEPES pH 7.5, 300 mM NaCl and coupling buffer (100 mM $NaHCO_3$ pH 8.3, 500 mM NaCl) to a final volume of 5 ml. The mixture was left to incubate on a rotating mixer at 4℃ overnight before washing the excess $CD177^{ecto}$ from the beads with a large amount of coupling buffer. The coupling reaction was quenched by replacing the coupling solution with 5 ml of 100 mM Tris pH 8.0 and left to incubate at r.t. for 3 h. The beads were then washed with three rounds of alternating low pH wash (100 mM acetic acid pH 4.0, 500 mM NaCl) and high pH wash (100 mM Tris pH 8.0, 500 mM NaCl) before being stored as a 50% slurry in PBS.

Culture media from small-scale expression in T175 flasks was recovered, clarified by centrifugation at 4,700 rpm, 18℃ for 30 min and filtered using a 0.22 µm syringe filter. 30 µl of a 50% slurry of CD177-coupled beads were added to 20 ml of PR3-containing culture medium in a 50 ml tube and the mixture was left to incubate on a rotating mixer at 4℃ overnight. The next day, the beads were transferred into 1.5 ml tubes

Reagent	
L-Arginine	CAS #74-79-3 Sigma-Aldrich, Cat.#A5006
3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)	CAS #75621-03-3 Sigma-Aldrich, Cat.#850500P
DDM	CAS #69227-93-6 Anatrace, Cat.#D310S
L-Glutamic acid	CAS #142-47-2 Sigma-Aldrich, Cat.#G1626
Glycerol	CAS #56-81-5 Thermo Fisher Scientific, Cat.#A16205.AP
Lauryl Maltose Neopentyl Glycol (LMNG)	CAS #1257852-96-2 Anatrace, Cat.#NG310
Non-detergent sulfobetaines 256 (NDSB-256)	Hampton Research, Cat.#HR2-705
NDSB-195	Hampton Research, Cat.#HR2-703
n-Octyl-β-D-Glucoside (OG)	CAS #29836-26-8 Anatrace, Cat.#O311
SDS	CAS #151-21-3 Sigma-Aldrich, Cat.#75746
Triton X-100	CAS #9036-19-5 Sigma-Aldrich, Cat.#T8787
Tween 20	CAS #9005-64-5 Thermo Fisher Scientific, Cat.#10246910

Table 2.2: Summary of reagents tested in the PR3 recovery assay.

and washed twice with 1 ml of PBS before addition of 1 ml of 5X loading dye, boiling at 95℃ for 2 min and recovering the supernatant for analysis.

2.4 Functional assays

2.4.1 Cell-surface staining assay

Constructs were cloned into the pHLsec vector with a N-terminal FLAG tag following on from the protein native secretion signal sequence for CD177 and a C-terminal FLAG tag for PR3. HeLa cells were cultured on glass cover slips and were transiently transfected with the plasmids. Anti-FLAG (Sigma-Aldrich, Cat.#F1804) was added to the cells at a concentration of 5 µg/ml. Cells were incubated for 1 h on ice and washed with PBS.

Cells were fixed in 4% paraformaldehyde (20 min, on ice), immunostained with Cy3-labeled anti-mouse antibody at a concentration of 50 µg/ml (Abcam, Cat.#ab97035), counter-stained with DAPI and mounted.

2.4.2 Enzyme-linked immunosorbent assay

Coating protein samples were diluted to 2 µg/ml in Dulbecco's PBS (DPBS, Sigma-Aldrich, Cat.#D8537) before incubation onto Nunc MaxiSorp^{TM} flat-bottom plates (Thermo Fisher Scientific, Cat.#735-0083) at 4°C overnight. Plates were washed four times with PBST and incubated with Blocker ^{TM} Casein in PBS buffer (Thermo Fisher Scientific, Cat.#37528) at r.t. for 1 h in order to reduce non-specific interactions. Meanwhile, patients' serum was diluted 1 to 200 with Blocker ^{TM} Casein in PBS buffer before incubation with plates at r.t. for 2 h. Plates were once again washed four times with PBST and then incubated with a secondary polyclonal goat anti-human antibody conjugated to alkaline phosphatase (α-human-AP pAb, Sigma-Aldrich, Cat.#A3187) at r.t. for 1 h. Finally, plates were washed four times with PBST and twice with PBS before colour development using 4-nitrophenyl phosphate (Signa-Aldrich, Cat.#N2765) dissolved in Pierce^{TM} diethanolamine buffer (Thermo Fisher Scientific, Cat.#34064) as substrate. The absorbance was measured at 405 nm on a Bio-tek ELx800^{TM} microplate reader with Gen5 ELISA software.

2.4.3 ANCA-depletion assay

140 µl of a 50% slurry of Pierce^{TM} streptavidin beads, pre-equilibrated in DPBS, were incubated with 210 µg purified biotinylated PR3 at r.t. for 1 h in a Mobicol spin column (MoBiTec, Cat.#M1003). Excess of PR3 was removed by three washes with 500 µl DPBS, and beads were incubated with patients' sera previously diluted 1 to 50 in DPBS at r.t. for 2 h. Flow-through was collected and analysed by ELISA for ANCA presence.

2.5 Biophysical methods

2.5.1 Size-exclusion chromatography-multiangle light scattering

110 µl of protein at 1 mg/ml was applied onto a Superdex® 200 HR 10/30 column previously equilibrated with SEC buffer. The system was connected to a Wyatt Dawn® HELEOS-II 8 angle light scattering detector as well as a Wyatt Optilab® rEX refractive index monitor.

2.5.2 SEC binding assay

4 nmol of $CD177^{ecto}$ was mixed with 8 nmol of PR3 variant in a maximum of 80 µl total volume. The two proteins were allowed to incubate at r.t. for at least 10 min before loading into a 100 µl tubing loop connected to a Superdex® 200 10/300 GL column previously equilibrated with SEC buffer. 250 µl fractions were collected and analysed by SDS-PAGE.

2.5.3 Surface Plasmon Resonance

SPR samples were prepared from small-scale expression experiments in six-well plates of constructs cloned into the pHL-Avitag3 vector. The proteins were biotinylated enzymatically *in vitro* at their C-terminal AviTag via co-transfection with 1 µg of the bacterial biotin ligase BirA in the presence of 200 nmol of biotin (Sigma-Aldrich, Cat.#B4501) supplemented in the culture medium. To improve the efficiency of this co-transfection, plasmid DNA were transfected with FuGENE® 6 (Promega, Cat.#E2691) instead of PEI at a DNA:FuGENE® mass-to-volume ratio of 1:3. The culture media was harvested after three days and clarified by centrifugation at 15,000 rpm, 18℃ for 10 min using a benchtop centrifuge. Excess biotin in the culture medium was removed by dialysis against PBS at 4℃ for 48 h. *In vivo* biotinylation efficiency was assessed by western blotting using HRP-labelled streptavidin antibody (Thermo Fisher Scientific, Cat.#21130).

SPR equilibrium binding experiments were performed at 25℃ using a Biacore T200 instrument (GE Healthcare). Series S sensor chip CM5 (Cytiva, Cat.#29149603) were

coated with streptavidin using the amine coupling kit (Cytiva, Cat.#BR100050), allowing biotinylated protein ligands to be attached to the chip. Ligands were immobilised at densities of 200-1200 response units (RUs). For dissociation constant K_D calculations, concentration series of the analytes were prepared by two-fold serial dilutions. Each concentration was injected for 200-300 s, followed by 300 s of dissociation time. The chip was regenerated with 2 M $MgCl_2$ followed by 300 s of stabilisation period between each injection. All the steps were carried out at a flow rate of 5 µl/min, in PBS supplemented with either 0.005% (v/v) Tween 20 or 0.05% (v/v) Tween 20 and 300 mM NaCl as running buffer, depending on the propensity of non-specific interactions.
When used, the catalytically active PR3 purified from neutrophils was purchased (Sigma-Aldrich, Cat.#SRP6309). In some experiments, 1 mM phenylmethanesulfonyl fluoride (PMSF, CAS #329-98-6, Sigma-Aldrich, Cat.#78830) was used in the PR3 dilution series and in the running buffer to inhibit PR3 catalytic activity.
K_D and maximum analyte binding (B_{max}) values were obtained by nonlinear curve fitting of a 1:1 Langmuir interaction model. Data were analysed using the BIAevaluation software 1.0 and Prism 7 (GraphPad).

2.6 X-ray crystallography

2.6.1 Protein crystallisation and crystal optimisation

For all crystallisation experiments, proteins were produced using HEK293S GnTI$^-$ cells in order to reduce N-glycan heterogeneity and therefore increase the chances of obtaining well diffracting crystals [274]. All crystallisation trials were performed using a mosquito® LV device (SPT Labtech) and crystallisation conditions were screened in SwissCi (MRC) 96-well 2-drop plates (SPT Labtech). Typically, crystallisation trials were set up by mixing 100 nl of concentrated protein with the crystallisation condition at a 1:1 volume ratio. In some cases, 150 nl of the protein were mixed with the crystallisation condition at a 2:1 volume ratio. When additive or protein seeds were used, the protein was mixed with the crystallisation condition and seeds/additive at a 2:1:1 or 6:2:1 volume ratio. Crystallisation trials were performed using a wide range of commercially available sparse matrix and systematic grid screens such as JCSG-plus™, MemGold™, MemGold2™,

MIDASplus™, Morpheus®, Morpheus® II, PACT premier™, PGA Screen ™, ProPlex™ (all from Molecular Dimensions), JCSG Core Suite I-IV (NeXtal), Index™ and PEG/Ion Screen™ (both from Hampton Research).

Crystal optimisation was sometimes performed to improve the yield and/or quality of the crystals. This consisted of either varying the concentration and/or pH of components from the hit condition, or by adding various additives from the Additive Screen and Silver Bullets (Hampton Research, Cat.#HR2-138 and HR2-096) to the hit condition. On some occasions, a microseed matrix screening was performed, in which initial crystals were harvested, crushed into small fragments using MicroSeed Beads (Molecular Dimensions, Cat.#MD2-14) and mixed with concentrated protein for a new broad crystallisation trial. When needed, stock solutions of crystallisation conditions were either purchased from the initial screen manufacturer or made from reagents available in the laboratory. The optimisation conditions were prepared in SwissCi (MRC) 96-well 2-drop plates and the crystallisation trials were performed using the mosquito® LV device as described above. All proteins were crystallised in sitting drops using the vapour diffusion method.

2.6.2 GPR97^{ecto} crystallisation and structure solution

2.6.2.1 Crystallisation trials

GPR97^{ecto} was concentrated to 40 mg/ml in the presence of 100 mM non-detergent sulfobetaines 256 (NDSB-256) (Hampton Research, Cat.#HR2-705) and was used to set up the following crystallisation trials: Index, JCSG I-II, Morpheus and PGA, all at a 1:1 protein:reservoir ratio and at 18°C. Crystals grew in a solution of 1.8 M tri-ammonium citrate pH 7 from the Index screen. They were harvested, cryo-protected in a solution of 3 M tri-ammonium citrate and sent to the ID30B beamline in European Synchrotron Radiation Facility (ESRF, France).

Optimisation screens included an ammonium citrate vertical gradient (1.76-1.83 M), and the use of the Additive Screen and Silver Bullets on the hit condition. Ammonium citrate stock solutions were prepared in-house using ammonium citrate tribasic (CAS #3458-72-8, Sigma-Aldrich, Cat.#A1332-100G) dissolved in water. The pH was adjusted using citric acid. Crystals grew in all three screens. They were harvested, cryo-protected

in a solution of 3 M tri-ammonium citrate and most of them were sent to the I03 and I04 beamlines in Diamond Light Source (DLS, UK). Other crystals were sent to the I23 beamline in DLS for Sulphur Single-wavelength Anomalous Dispersion (S-SAD) phasing. Endo F_1-treated $GPR97^{ecto}$ was concentrated to 40 mg/ml and was used to set up the following crystallisation trials: JCSG+, JCSG II, Morpheus, Morpheus II, Natrix, PACT, PEG/Ion and PGA, all at a 1:1 protein:reservoir ratio at 18℃.

2.6.2.2 Data collection and data processing

The best native dataset was collected at a wavelength of 0.9763 Å at the ESRF beamline ID30B. 180° of data were collected with an exposure of 0.02 s per 0.1° rotation on a PILATUS3 6M (Dectris). The data was integrated using DIALS [275] and scaled and merged using AIMLESS [276, 277]. The S-SAD dataset was collected at a wavelength of 2.7552 Å on the long-wavelength beamline I23 at DLS [278]. 360° of data were collected with an exposure of 0.1 s per 0.1° rotation on a PILATUS 12M detector (Dectris). The data showed sign of anisotropy and was therefore processed using AUTOPROC [279] and STARANISO [280].

2.6.2.3 Structure determination

The flexible loops of the GAIN domains of Lphn1, BAI3 and GPR56 (PDB ID 4DLQ, 4DLO, 5KVM respectively) were manually removed and an ensemble of those trimmed GAIN domains was used as a search model for molecular replacement using the native dataset and PHASER [281]. The solution consisted of two molecules in the asymmetric unit and displayed additional density on top of the GAIN domain, accounting for the small N-terminal domain.

To help build this N-terminal domain, molecular replacement using the S-SAD dataset and PHASER [281] was carried out with the partial model consisting of the GPR97 GAIN domain. The resulting solution was subjected to one round of refinement in autoBUSTER [282, 283]. Electron density maps were improved in PARROT [284] revealing the missing N-terminal domain. Anomalous difference maps calculated with ANODE [285] were used to locate sulphur atoms and help manual building in COOT [286]. All-atom refinement

with autoBUSTER [282, 283] and Refmac5 [287] (version 5.8.0267) with local NCS, TLS and gelly body restraints resulted in the final model containing residues 28-260.
The structure was deposited to the PDB server and was assigned the PDB ID 7QU8.

2.6.3 CD177^{ecto} crystallisation and structure solution

2.6.3.1 Crystallisation trials

CD177^{ecto} concentrated to 12 mg/ml was used at a 1:1 protein:reservoir ratio in the following crystallisation screens: Index, JCSG I-IV, Midas, Morpheus, Morpheus 2, PACT, PEG/ION, PGA at 18℃ and JCSG+, Morpheus, PACT at 4℃. Protein concentrated to 9.8 mg/ml was used at a 2:1 protein:reservoir ratio in the JCSG II, Morpheus, PACT screens at 18 ℃, and JCSG+, Morpheus, PACT at 4℃. In both cases, crystals grew in a 20% w/v PEG 3350, 0.20 M ammonium nitrate solution from the JCSG+ screen at 4℃. Crystals were harvested, cryo-protected in 25% glycerol and sent to the I24 and I03 beamlines in DLS.
Optimisation screens were performed with CD177^{ecto} concentrated to 12 mg/ml and included a combination of a PEG 3350 horizontal gradient (15-22%) with an ammonium nitrate vertical gradient (0.17-0.22 M) as well as a combination of PEG 3350 horizontal gradient (15-22%) with a vertical ten-fold dilution series of crystal seeds obtained from the initial hit condition. Additive Screen and Silver Bullets on the hit condition was also done. All optimisation screens were set up at 4℃, at both 1:1 and 2:1 protein:reservoir ratio. Crystals grew in all screens. Some of them were used for heavy atom soaking and sent to the I24 beamline in DLS while others were sent to the I23 beamline in DLS for S-SAD phasing.

2.6.3.2 Heavy-atom soaking

Crystal that were stuck to the bottom of the wells were first dislodged using an acupuncture needle before being soaked with derivatives of the heavy atoms platinum, gold, mercury and tantalum (Table 2.3). For this, the tip of a pipette tip was used to transfer a tiny amount of heavy metals to a tube containing 20 µl of the reservoir solution. In each case, about 0.2 µl of the heavy metal solution was applied next to the

drop containing the crystals until they fused. Crystals were then left to soak overnight. The next day, 2 µl of 25% glycerol was applied on top of each drop before the crystals were harvested and made sure to go through the glycerol layer for cryo-protection.

Name	Chemical formula
Potassium Tetrachloroplatine(II)	K_2PtCl_4
Potassium Hexachloroplatinate(IV)	K_2PtCl_6
Potassium Tetranitroplatinate (II)	$K_2Pt(NO_2)_4$
Potasium Tetracyanoplatinate(II) hydrate	$K_2Pt(CN)_4 \bullet xH_2O$
Potassium Tetrabromoplatinate(II)	K_2PtBr_4
Gold(I) Potassium Cyanide	$KAu(CN)_2$
Potassium Tetrachloroaurate(III) hydrate	$KAuCl_4 \bullet xH_2O$
Mersalyl acid	$C_{13}H_{17}HgNO_6$
Ethyl Mercuric Phosphate	$C_2H_5HgH_2PO_4$
Mercury(II) Potassium Iodide	K_2HgI_4
Tantalum (V) Bromide	Ta_2Br_{10}

Table 2.3: Summary of all heavy-atom derivatives used for crystal soaking.

2.6.3.3 Data collection and data processing

Native datasets were collected at a wavelength of 0.9686 Å and 1.0718 Å at the DLS beamlines I24 and I03, respectively. 180° of data were collected with an exposure of 0.05 s per 0.2° rotation on a PILATUS3 6M (Dectris) for the I24 beamline or Eiger2 XE 16M (Dectris) for I03 beamline. The data were integrated using DIALS [275] via the xia2 system [288] and scaled and merged using AIMLESS [276, 277]. Blend [276, 289, 277] was sometimes used to merge multiple datasets together. The best datasets were processed with STARANISO [280].

The K_2PtCl_6-, $K_2Pt(CN)_4 \bullet xH_2O$- and $KAu(CN)_2$-derived datasets were collected at a wavelength of 1.0713, 1.0710 and 1.0389, respectively, at the DLS beamline I24. 180° of data were collected with an exposure of 0.05 s per 0.1° rotation for $K_2Pt(CN)_4 \bullet xH_2O$ and $KAu(CN)_2$ and per 0.2° rotation for K_2PtCl_6 on a PILATUS3 6M (Dectris). The data were processed using the XDS package.

The S-SAD dataset was collected at a wavelength of 2.7552 Åat the DLS beamline I23 [278]. 360° of data were collected with an exposure of 0.1 s per 0.1° rotation on a

PILATUS 12M detector (Dectris). Twelve datasets from three different crystals were individually integrated using XDS and merged using XSCALE [290].

2.6.3.4 Structure determination

SHELXC-D within the hkl2map graphical interface [291, 292] was used for substructure determination. AutoSol was used to obtain initial density-modified electron density maps [293]. A combination of automated model building using Crank2 [294] and AutoBuild [295] with manual building in COOT [286] enabled to build the model step-by-step. The solution consisted of two molecules in the asymmetric unit and several iterations of model building and AutoSol runs resulted in a nearly full model of $CD177^{ecto}$.

2.6.4 $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystallisation and structure solution

2.6.4.1 Crystallisation trials

A 1:1 molar mixture of purified $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$ was concentrated to 7.3 mg/ml and was used at both 1:1 and 2:1 protein:reservoir ratio to set up the following crystallisation trials: Index, JCSG I-IV, Midas, Morpheus, Morpheus II, PACT, PGA, ProPlex at 18℃, and Index, JCSG+, Midas, Morpheus II, ProPlex at 4℃. Crystals grew in a solution of 0.1 M Tris pH 7, 25 % w/v SOKALAN CP 42, 10 % v/v tetrahydrofuran from the Midas screen at 4℃. They were harvested, cryo-protected in 25% glycerol and sent to the I04 beamline in DLS.

A SEC-purified $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ complex was concentrated to 7.3 mg/ml and was used to set up the following crystallisation trials: JCSG+, Midas, Morpheus, PACT, ProPlex at 18℃. Crystals grew in three different solutions from the Midas screen at 18℃, were harvested and cryo-protected in 25% glycerol before being sent to the I24 beamline in DLS.

Optimisation screens included the use of the Additive Screen and Silver Bullets on the four hits conditions from the Midas screen. A microseed matrix screening was also performed using Index, JCSG+, Midas, Morpheus and ProPlex. Crystals which grew in various conditions upon seeding were harvested, cryo-protected either in 25% glycerol or in a solution of 9 M sodium formate and sent to the I04 beamline in DLS.

2.6.4.2 Data collection and data processing

Native datasets were collected at a wavelength of 0.9795 Å and 0.9999 Å at the DLS beamlines I04 and I24, respectively. 180° of data were collected with an exposure of 0.04 s per 0.1° rotation on a Eiger2 XE 16M (Dectris) for I04 beamline or PILATUS3 6M (Dectris) for the I24 beamline. All four datasets failed to process.

2.6.5 $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$ crystallisation and structure solution

2.6.5.1 Crystallisation trials

A 1:1 molar mixture of purified $PR3^{IWG\text{-}NGT}$ and $CD177^{subA}$ was concentrated to 8.3 mg/ml and was used in the following crystallisation screens: Index, JCSG I-IV, Midas, Morpheus, Morpheus II, PACT, PGA, ProPlex 18°C. Crystals which grew in a wide range of conditions were harvested, cryo-protected in 25% glycerol and sent to the PROXIMA-1 beamline at SOLEIL (France).

2.6.5.2 Data collection and data processing

Native datasets were collected at a wavelength of 0.9786 Å at the SOLEIL beamline PROXIMA-1. 360° of data were collected with an exposure of 0.01 s per 0.1° rotation on a Eiger2 XE 16M (Dectris). The data was integrated using XDSme [296] or DIALS [275] via the xia2 system [288], and scaled and merged using AIMLESS [276, 277].

2.6.5.3 Structure determination

A model of PR3 (PDB ID 1FUJ) as well as a partial model of $CD177^{subA}$ were used as search models for molecular replacement using the best native dataset and PHASER [281]. The solution consisted of one molecule in the asymmetric unit and was rebuilt using Buccaneer [297]. All-atom refinement with Refmac5 [287] (version 5.8.0267) resulted in a full model containing residues 28-256 and 22-206 for $PR3^{IWG\text{-}NGT}$ and $CD177^{subA}$, respectively.

Chapter 3

The role of GPR97 in inflammatory activation of human neutrophils

3.1 Summary

GPR97 is an orphan adhesion GPCR (aGPCR), which has been found to be upregulated in neutrophils of diverse inflammatory disorders. More specifically, GPR97 appears to trigger neutrophil activation upon direct interaction with Proteinase 3 (PR3), a neutrophil serine protease present on the neutrophil cell surface via its GPI-anchored receptor CD177. In order to biophysically characterise the GPR97-PR3 complex, the GPR97 ectodomain was recombinantly produced using HEK293 cells and yielded crystals which diffracted to 3.37 Å resolution. The dataset could be solved by a combination of molecular replacement and S-SAD methods, and resulted in a model of the full GPR97 ectodomain consisting of a short N-terminal domain (NTD) followed by a conserved GAIN domain. Both the NTD and GAIN domains of GPR97 were found to contribute to PR3 binding, and therefore for neutrophil activation. However, as the two proteins failed to display a direct binding signal in a surface plasmon resonance experiment, other stabilising interacting partners were investigated. In fact, it appeared that the CD177-bound PR3 has to be part of the CD177/GPR97/PAR2/CD16b interactome in order to properly interact with GPR97. Overall, this work identified a novel PR3/CD177/GPR97/PAR2/CD16b interactome as a critical modulator of

neutrophil activation and propose a new aGPCR-GPCR (GPR97-PAR2) transactivation mechanism, which would direct neutrophil activation and inflammation.

Disclaimer: For the purpose of putting the biochemical and structural studies of GPR97 into context, some work performed by members of Prof. Hsi-Hsien Lin group (Chang Gung University, Taiwan) are presented in this chapter. This relates to results presented in Sections 3.4.1, 3.4.3 and 3.4.4 as well as the cell assays in Section 3.4.2. Additionally, the S-SAD experiment mentioned in Section 3.2.3.2 was performed by Dr. Kamel El-Omari from the I23 beamline in DLS.

3.2 GPR97^{ecto} production and crystallisation

3.2.1 Bioinformatic analyses

As a member of the GPCR superfamily, GPR97/ADGRG3 shares the same overall topology consisting of an N-terminal extracellular domain followed by seven transmembrane helices and a C-terminal cytoplasmic tail. As this study focuses on GPR97 interactions with extracellular partners, only constructs of the ectodomain were considered. To inform the design of suitable constructs and avoid either miss-folding of incomplete domain fragments or the presence of long flexible loops at the N- and C-termini, bioinformatic tools were used to predict disordered regions as well as structured subdomains within the protein (Figure 3.1A,B).

Initially, the first 20 residues, M1 to G20, were analysed as the protein signal peptide by the SignalP server [300]. Next, various secondary structure prediction servers predicted the ectodomain to end at residues T264 (NCBI CDD [301]), I267 (Interpro [302]), R270 (SMART [303], TMHMM [304], PSIPRED [305]), Q173 (HMMTOP [306]) or C276 (TMpred [307], MEMSAT [308]). Additionally, sequence alignment with its closest relative, and crystallised, aGPCR GPR56/ADGRG1 suggested P258 as another C-terminus of the ECD [309].

The only structural information available on GPR97 thus far is the presence of the evolutionary conserved canonical GAIN domain, almost immediately N-terminal to its 7TM region (Figure 3.1A). At the time, only four structures of the GAIN domain were available and were all solved by X-ray crystallography. The GAIN domains of

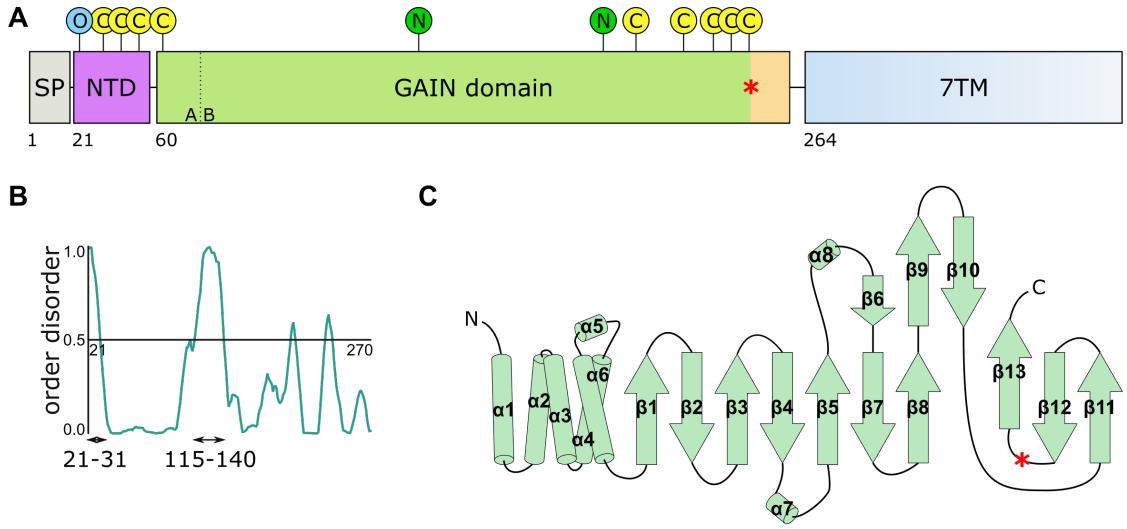


Figure 3.1: Predicted GPR97 domain architecture. **(A)** Schematic depicting the predicted extracellular domain boundaries after bioinformatic analyses. GPR97 possesses 9 cysteines (C32, C50, C54, C62, C215, C233, C244, C245, C246) that are indicated by yellow spheres, 2 predicted N-glycosylation sites [298] (N144, N210) indicated by green spheres and 1 predicted O-glycosylation site [299] (T25) indicated by a blue sphere. The GAIN domain is partitioned into subdomains A and B by the dotted line. The GPS autoproteolysis site is indicated by a red asterisk and the cleaved β13-strand is coloured in beige. **(B)** Regions predicted to be disordered by PONDR® are indicated by arrows and correspond to Q21-T31 and Q115-S140. **(C)** Typical topology of a GAIN domain. The autoproteolysis site between β12- and β13-strands is indicated by a red asterisk.

Lphn1/ADGRL1 and BAI3/ADGRB3 were the first to provide molecular details on the domain organisation and have set up the nomenclature that is used today [141]. Both of those GAIN domains contain ∼310 residues and display an N-terminal subdomain A composed of six α-helices, numbered α1 to α6, and a C-terminal subdomain B composed of a twisted β-sandwich including thirteen β-strands and two small α-helices, numbered β1 to β13 and α7 to α8, respectively (Figure 3.1C). The highly conserved GPS motif is located directly N-terminal to the 7TM region and spans the β9-β13 strands (Appendix Figure C.1). Cleavage at the GPS occurs between the β12- and β13-strands (Figure 3.1C, Appendix Figure C.1).

More recently, shorter GAIN domains of ∼260 residues from GPR126/ADGRG6 [310] and ∼220 residues from GPR56/ADGRG1 [309] were deposited in the PDB database and showed to be lacking the first α1-helix and the first α1- to α3-helices of the subdomain A for GPR126 and GPR56, respectively (Figure 3.2A).

Interestingly, GPR97 homology models generated by the Swiss-Model [311] and HHPred [312] (using the MODELLER software) servers suggested an even smaller GAIN domain

starting from residue N63, with only ∼210 residues and a subdomain A consisting of only the α6-helix. In addition, a closer look at the amino acid sequence of the receptor ectodomain showed the presence of four conserved cysteines within the N-terminal residues that are not part of the predicted GAIN domain, suggesting the existence of a small but functionally conserved N-terminal domain (NTD) (Figure 3.1A).

These cysteines are highly likely to be involved in interdomain disulfide bonds, where they would bridge and orient the NTD with respect to the GAIN domain. This has indeed been previously observed in the crystal structures of Lphn1, BAI3, GPR126 and GPR56, where disulphide bridges are stabilising the GAIN domains to its directly N-terminal HormR domain [141, 310] or PLL domain [309] (Figure 3.2B).

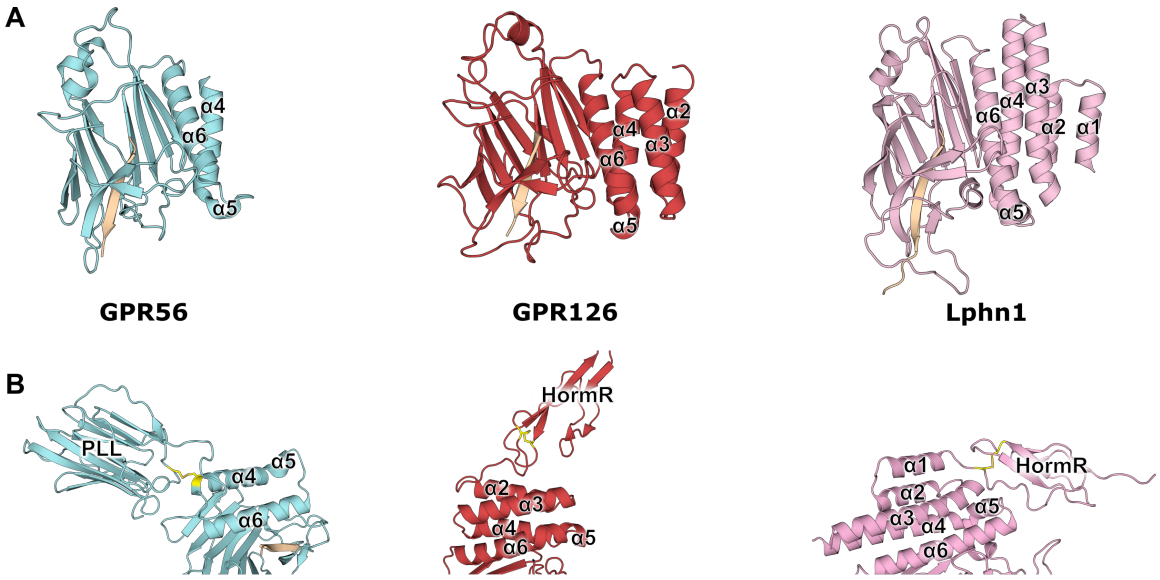


Figure 3.2: Comparison of the crystal structures of other aGPCRs. **(A)** Cartoon representation of GAIN domain models from GPR56, GPR126 and Lphn1 (PDB ID: 5KVM, 6V55 and 4DLQ, respectively). **(B)** Close-up view of the interdomain disulphide bridges linking the GAIN domain of GPR56, GPR126 and Lphn1 to their directly N-terminal PLL domain, HormR domain and HormR domain, respectively.

In addition, a part of this NTD (residues Q21-T31) was predicted to be disordered and could therefore be truncated out during construct design (Figure 3.1B). Another predicted disordered region (residues Q115-S140) maps to a long flexible loop between β3 and β4 of the GAIN domain (Figure 3.1B,C) and could potentially be removed if it appeared to hinder crystallisation trials.

3.2.2 Construct screening

Following up on the bioinformatic analyses, several constructs encompassing either the full or shorter versions of the ectodomain were designed. Constructs of the full ectodomain started either from M1 or Q21 and ended between residues T264 to G277. Constructs starting from M1 possessed the receptor native signal peptide, while those starting from Q21 were fused to the pHLsec vectors's secretion signal peptide. Shorter constructs of the ectodomain were truncated at their N-terminus and included a version starting from R29 (Figure 3.1A,B) and another starting from G58 (Figure 3.1A), containing only the predicted GAIN domain. This latter version contained a C62S mutation in order to prevent unspecific reactions from the free C62 as it would normally have formed a disulphide bond with C32, C50 or C54, which are all missing in this construct. Expression levels were assessed via small-scale protein expression trials in HEK293T cells and by western blotting.

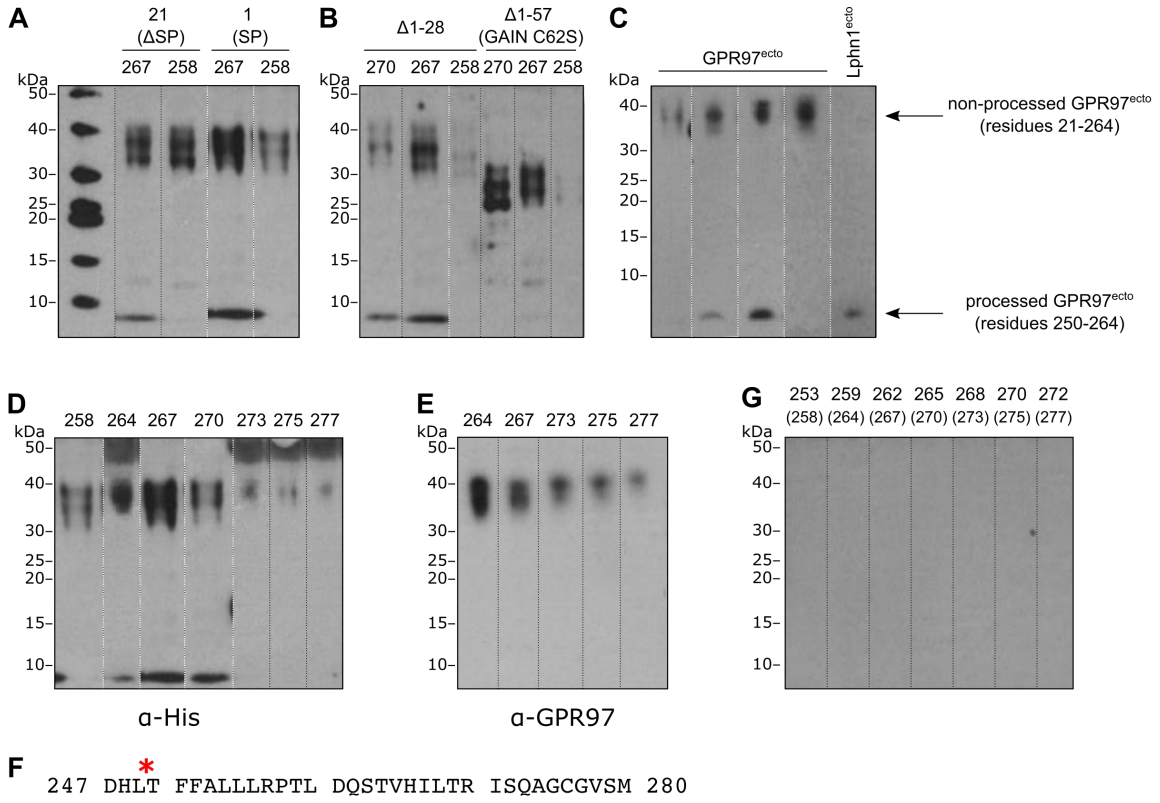


Figure 3.3: Expression trials of GPR97^{ecto} constructs. Western blot comparing the expression levels of different GPR97^{ecto} constructs. White lines separate lanes extracted from different western blots. **(A)** Constructs possessing the native signal peptide (SP; 1-258 and 1-267) do not show a significant difference with constructs possessing the vector signal peptide (ΔSP; 21-258 and 21-267). **(B)** Constructs with a truncated N-terminus (Δ1-28; 29-258, 29-267, 29-270) do not show increased expression levels while most of the constructs that only encompass the GAIN domain (Δ1-57; 58-258, 58-267, 58-270) appear to be well-expressed. **(C)** Transfection of one same constructs (1-264) resulted in varying autoproteolsyis rates between each experiment. A construct of Lphn1 ectodomain that is fully autoproteolysed is shown for comparison. **(D,E)** Expression levels of GPR97 constructs with different C-terminus boundaries were assessed using either **(D)** α-His mAb or **(E)** α-GPR97 mAb as the primary antibody. **(F)** GPR97 amino acid sequence from D247 to M280. The GPS autoproteolysis site between L249 and T250 is indicated by a red asterisk. **(G)** No expression could be observed when using mouse GPR97 constructs.

Initial tests on full ectodomain constructs compared the expression levels obtained from using the native signal peptide or the signal peptide already present in the expression vector (Figure 3.3A). No significant differences were observed and the native signal peptide was retained for all full ectodomain constructs. In addition, while trimming the flexible N-terminal boundary to R29 failed to improve the expression level, the construct only encompassing the GAIN domain appeared to be well-expressed (Figure 3.3B). Interestingly, GPR97 does not seem to undergo a complete autoproteolysis, as demonstrated by the presence of one band at ∼30 kDa corresponding to the non-processed

receptor, and one band lower than 10 kDa corresponding to the C-terminal fragment of the processed receptor. Consequently, this meant the necessity to take both bands into account in order to properly assess and compare the expression levels between various constructs (Figure 3.3A,B). However, the signal ratio between these two bands would often vary between experiments performed on different days. When using the same construct, the bottom band would sometimes disappear, or in some cases, no signal at all would be observed (Figure 3.3C). Such experimental differences could be due to different rates of autoproteolytic cleavage ocurring between different constructs, but also with the same construct when transfected and harvested in different ways (Figure 3.3C,D). In addition, small peptides have been found to easily detach from the blot by abrupt shaking or washes, which could therefore result in wrongly labelling a highly expressed and autoproteolysed construct as a low expressing one. As a result, this made it difficult to obtain reproducible blots and thus to draw conclusions regarding expression levels. While investigating the best C-terminus boundary, this issue was addressed by using a monoclonal anti-GPR97 antibody (α-GPR97 mAb) instead of the usual monoclonal anti-His antibody (α-His mAb) as the primary antibody. This enabled specific detection for GPR97 independently of its proteolytic state, and allowed 1-264 to be chosen as the best expressing construct for large-scale expression and purification (Figure 3.3D-F). In parallel, in order to maximise the chances of obtaining well-expressing construct, all the full ectodomain constructs were also transposed in mouse GPR97. Unfortunately, none of them expressed in HEK293T cells (Figure 3.3G).

3.2.3 Protein production and structure determination

3.2.3.1 Protein production and crystallisation trials

GPR97 1-264, named GPR97^{ecto}, was selected for large-scale expression in HEK293S GnTI^{-} cells to obtain homogeneously glycosylated and pure protein to set up crystallisation trials. Using a standard purification protocol consisting of IMAC and SEC, GPR97^{ecto} could be purified with a final yield ranging from 3 to 5 mg from 3 L of culture (Figure 3.4A-C). SDS-PAGE analysis confirmed a molecular weight ranging from 30 to 40 kDa. However, as the protein presented quite significant smearing on

gel, proteolytic degradation or the presence of differently glycosylated species resulting from the predicted O-linked glycosylation site at T25 and N-linked glycosylation sites at N144 and N210 were investigated. To this end, GPR97^{ecto} was incubated with the amidase PNGase F, which is known to remove the vast majority of N-linked glycans before hydrolysing the involved asparagine to an aspartate residue. The resulting mixture showed a deglycosylated GPR97^{ecto} running as a single band at 30 kDa, which corresponds to the estimated molecular weight for GPR97^{ecto}, and confirms the receptor is indeed glycosylated (Figure 3.4D).

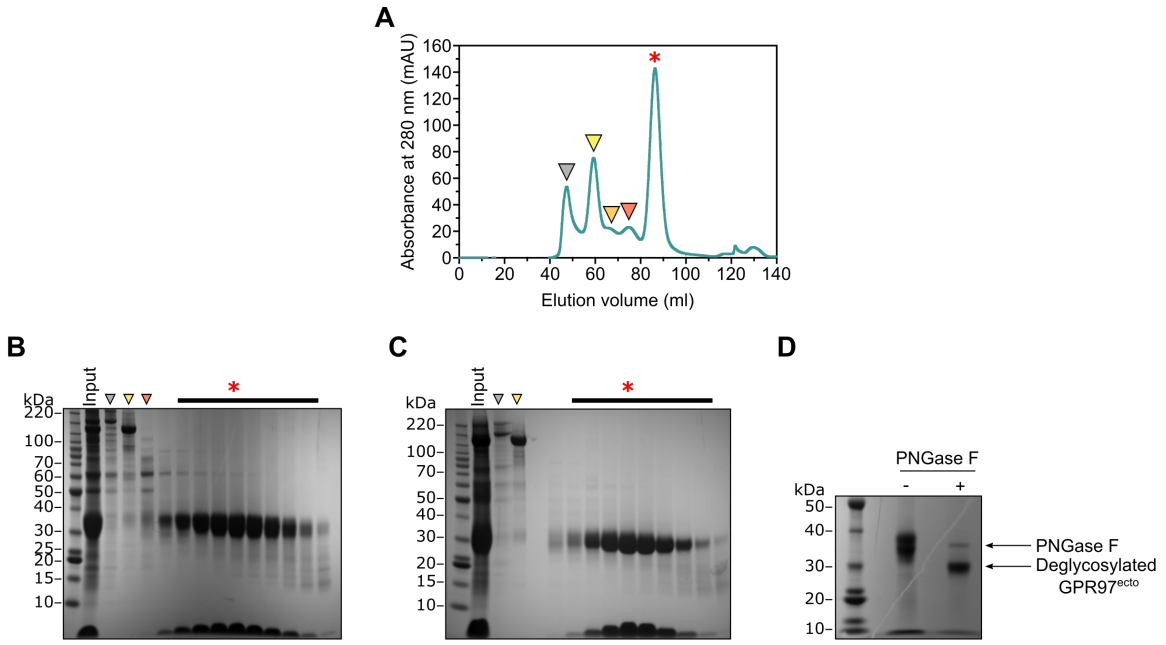


Figure 3.4: Large-scale production of GPR97^{ecto}. (A) A representative SEC purification profile shows GPR97^{ecto} elutes around 86 ml on a Superdex® 200 column. The protein peak (red asterisk) is preceded by a void peak at 48 ml (grey arrowhead), followed by impurity peaks present in all purifications from HEK293 cells at 60 ml, 67 ml and 75 ml (yellow, orange and red arrowheads, respectively). **(B)** Representative SDS-PAGE gel showing the SEC input, void and impurities peak fractions and GPR97^{ecto} pure protein fractions. GPR97^{ecto} runs as two bands due to its autoproteolysis. **(C)** SDS-PAGE gel of samples obtained after Endo F_1-treatment of GPR97^{ecto}. It shows the SEC input, void and impurities peak fractions and GPR97^{ecto} pure protein fractions. The GPR97^{ecto} fractions run at a lower molecular weight after deglycosylation. **(D)** SDS-PAGE gel of GPR97^{ecto} before and after treatment with PNGase F showing GPR97^{ecto} is heavily glycosylated.

Fractions showing protein purity above 95% were pooled together and supplemented with 100 mM NDSB-256 to prevent early protein aggregation during concentration. When the protein reached 8 mg/ml, a first crystallisation trial was performed using the Morpheus® screen. It was considered that having about or above 50% precipitated protein in all

Morpheus® crystallisation conditions was a good estimate of the protein entering the supersaturated state. However, all conditions appeared clear of precipitation whether the protein was concentrated to 8 mg/ml, 20 mg/ml or even 40 mg/ml. Ultimately, a few crystallisation trials were prepared with GPR97^{ecto} at 40 mg/ml (Table 3.1). Although most of the conditions either remained clear or showed phase separation, one condition from the Index^{TM} screen showed small crystals growing after a week of incubation. The crystals seemed to reach full growth after three weeks and were harvested after 56 days (Figure 3.5). One full dataset was collected and could be processed to a resolution of 3.37 Å.

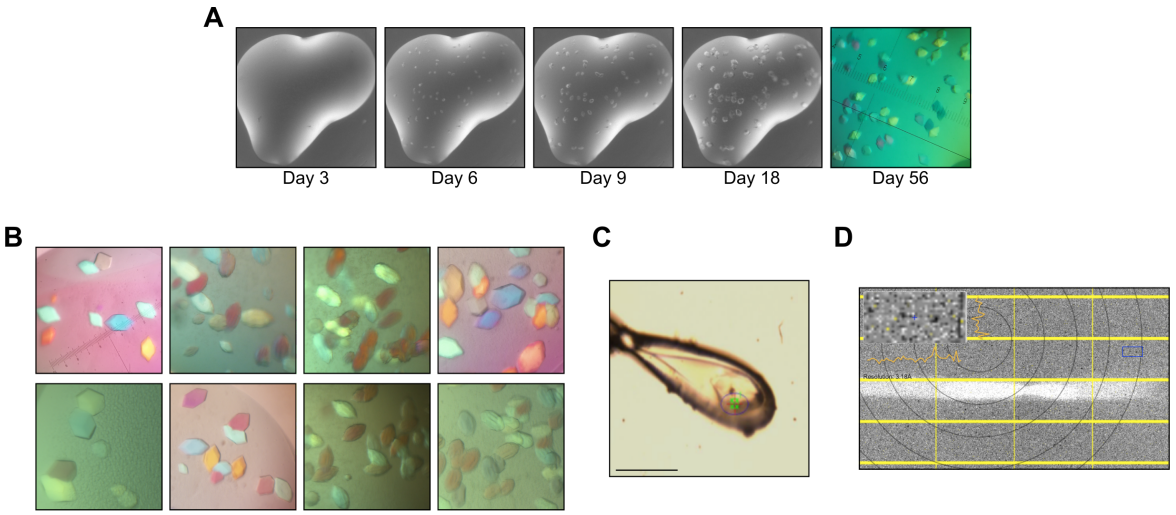


Figure 3.5: GPR97^{ecto} crystallisation trials. (A) GPR97^{ecto} crystals growing in condition B9 of the Index^{TM} crystallisation screen (1.8 M tri-ammonium citrate, pH 7) at 18°C. **(B)** GPR97^{ecto} crystals optimisation led to the obtention of crystals of various sizes and shapes. **(C)** View of one crystal cryo-protected in 3 M tri-ammonium citrate and mounted onto a SPINE pin prior to data collection. Scale bar is 100 µm. **(D)** Diffraction pattern resulting from one GPR97^{ecto} crystal, showing protein diffraction spots around 3.5-4 Å.

As the crystals were harvested belatedly, crystallisation trials were repeated in the hit condition as well as in optimised conditions derived from the initial hit with the hope of obtaining crystals diffracting at a better resolution (Table 3.1). This yielded a number of crystals of different sizes and shapes (Figure 3.5B). Unfortunately, none of them diffracted to a significantly better resolution.

In parallel, GPR97^{ecto} was purified in the presence of Endo F_1 and concentrated without NDSB-256 with the hope that this would decrease protein solubility and therefore decrease the concentration required for the protein to reach the supersaturation state

(Figure 3.4B). Numerous crystallisation screens were tested (Table 3.1) but none of them were successful in producing crystals.

Concentration	**Crystallisation screen**
GPR97^{ecto} in 20 mM Tris pH 7.5, 300 mM NaCl, 100 mM NDSB-256	
8 mg/ml (283 µM)	Morpheus
20 mg/ml (708 µM)	Morpheus
40 mg/ml (1416 µM)	Index, JCSG I, JCSG II, Morpheus, PGA
	Optimisation screens derived from the hit condition (1.8 M tri-ammonium citrate pH 7):
40 mg/ml (1416 µM)	Tri-ammonium vertical gradient (1.76 to 1.83 M) Tri-ammonium vertical gradient (1.76 to 1.83 M) at 4℃ Additive screen on hit condition Silver bullets on hit condition
GPR97^{ecto} Endo F_1-treated in 20 mM Tris pH 7.5, 300 mM NaCl	
8 mg/ml (283 µM)	Morpheus
20 mg/ml (708 µM)	Morpheus
40 mg/ml (1416 µM)	JCSG+, Morpheus, Morpheus 2, Natrix, PACT, PEG/Ion, PGA

Table 3.1: Summary of GPR97^{ecto} crystallisation trials.

3.2.3.2 Structure determination

The resulting dataset was manually processed using DIALS for integration and AIMLESS from the CCP4 suite for scaling and merging (Table 3.2). Both of those steps already proposed P 4_3 2_1 2 as the protein space group found in the crystal. An estimation of the cell content using the Matthews coefficient method indicated a probability of 12% and 87% of having two and three copies of GPR97^{ecto} in the asymmetric unit (ASU), respectively, as well as a corresponding solvent content of 64% and 46%. Knowing that a relatively low resolution for a dataset is usually indicative of a high solvent content, as the solvent could disrupt the packing of the proteins, the crystal ASU appeared to be much likely composed of two copies of GPR97^{ecto}.

Since the GAIN domain constitutes more than 80% of GPR97^{ecto}, a molecular replacement of the GAIN domain was first attempted to provide an initial electron density map. To this end, the α1- to α5-helices as well as the flexible loops of the GAIN domains of Lphn1, BAI3 and GPR56 were removed and an ensemble of those trimmed

GAIN domains was used as a search model for molecular replacement. The program was set to search for two copies of the ensemble in both the P 4_3 2_1 2 space group as well as its enantiomorph P 4_1 2_1 2. This produced one unique solution consisting of two molecules in the ASU of the P 4_3 2_1 2 protein space group. Interestingly, the resulting electron density map displayed clear additional electron density on top of the α6-helix of subdomain A, suggesting the presence of the NTD on top of the GAIN domain (Figure 3.6A). Unfortunately however, the resolution did not allow for trivial automated or even manual model building.

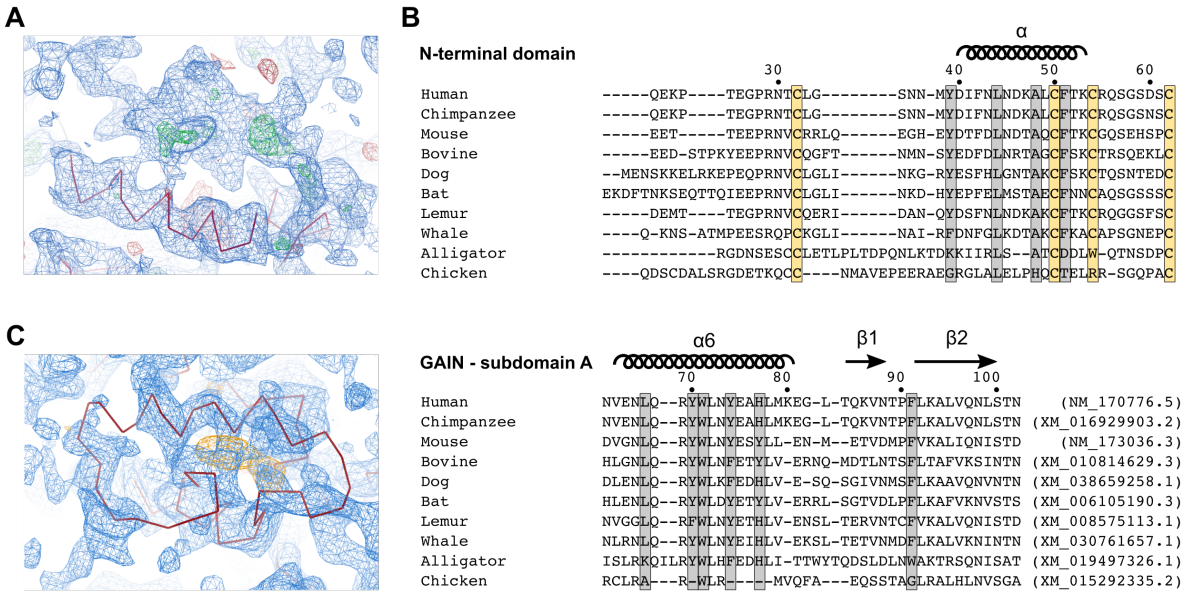


Figure 3.6: Structural building of the NTD. 2Fo-Fc electron density maps are shown at the 1 sigma level in blue. Fo-Fc electron density maps are shown at ± 3 sigma level in green and red, respectively. **(A)** The α6-helix of subdomain A is depicted as a red trace in a 2Fo-Fc electron density map. Clear additional density accounting for the NTD can be seen on top of the GAIN domain. **(B)** Sequence alignment with secondary structure of the NTD, α6-helix, β1- and β2-strands of the GAIN domain of GPR97 orthologs. The conserved cysteines involved in known disulphides bridges are highlighted in yellow. The conserved residues involved in the hydrophobic interactions stabilising the NTD to the GAIN domain are highlighted in grey. **(C)** The NTD is shown as a red trace in a 2Fo-Fc electron density map obtained after a few cycles of refinement. The sulphur anomalous difference map is shown in orange at the 3 sigma level.

To solve this issue, GPR97^{ecto} crystals were reproduced and used in S-SAD experiments. S-SAD phasing is a method which uses a longer wavelength during data collection in order to detect the anomalous signal of the sulphur atoms, which therefore results in the identification and location of the sulphur atoms in the crystal. As the anomalous

signal of sulphur is experimentally weak, this technique usually requires a high signal-to-noise ratio as well as a high multiplicity in order to obtain satisfactory measurements. Fortunately, both were the case for the GPR97^{ecto} crystals obtained, which therefore suggested the S-SAD phasing to be a suitable method to help locate the four cysteine residues (C32, C50, C54, and C62) present in the NTD.

A combination of molecular replacement using the partial model of GPR97^{ecto} along with the sulphur anomalous signal obtained by S-SAD enabled the two disulphide bridges involving C32, C50, C54, and C62 to be placed, which ultimately lead to the model solution of the NTD (Figure 3.6B,C). Several iterations of manual building and refinement resulted in a final model of GPR97^{ecto} at 3.37 Å resolution spanning P28-L260 (Table 3.2).

	Native	**Sulphur-SAD**
Data collection		
Source	ID30B (ESRF)	I23 (DLS)
Wavelength λ (Å)	0.97625	2.7552
Number of crystals	1	1
Resolution range (Å)	96.28-3.37 (3.64-3.37)	172.68-3.50 (3.83-3.50)
Space group	P 4_3 2_1 2	P 4_3 2_1 2
Cell dimensions		
a, b, c (Å)	96.3, 96.3, 172.4	95.9, 95.9, 172.7
α, β, γ (°)	90, 90, 90	90, 90, 90
Unique reflections	12030 (2398)	10793 (2519)
Multiplicity	11.6 (12.2)	22.8 (21.0)
Completeness (%)	99.7 (99.4)	100 (100)
Wilson B ($Å^2$)	103.06	140.66
R_{merge} (%)	27.3 (123.8)	22.5 (359.3)
R_{meas} (%)	28.6 (129.2)	23.0 (368.3)
R_{pim} (%)	8.3 (36.4)	4.8 (79.8)
$CC_{1/2}$ (%)	99.6 (52.1)	99.9 (61.6)
Average I/σ(I)	6.5 (1.9)	9.9 (0.9)
Refinement		
PDB accession number	7QU8	
Resolution range (Å)	84.11-3.37	
Reflections (work / free set)	10967 / 1040	
R_{work} / R_{free} (%)	26.59 / 31.30	
Number of atoms (protein)	3640	
Mean B value ($Å^2$) (protein)	134.5	
RMSD bonds (Å)	0.005	
RMSD angles (°)	1.229	
Ramachandran		
Favoured (%)	94.3	
Allowed (%)	5.7	
Outliers (%)	0.0	
Molprobity score / percentile	2.03	

Table 3.2: Data collection and refinement statistics for GPR97^{ecto}. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor, R_{meas} for the multiplicity-corrected merging R-factor and R_{pim} for the precision-indicating merging R-factor. $CC_{1/2}$ is the correlation coefficient between random half data sets and RMSD is the acronym for root-mean-square deviation from ideal geometry.

3.3 Structural analysis of GPR97^{ecto}

The resulting GPR97^{ecto} model consists of a small cysteine-stabilised helical NTD followed by an unusually small GAIN domain that consists of a short subdomain A with one α-helix and a typical subdomain B with 13 β-strands and one α-helix.

3.3.1 The N-terminal domain

The GPR97^{ecto} model shows that the single α-helix of the subdomain A (helix α6 in canonical GAIN domain) packs against subdomain B and acts as an anchor for the upstream α-helix of the NTD (Figure 3.7A). Structural comparison confirms that the NTD α-helix does not align with any of the missing α1-α5-helices of subdomain A and that the predicted NTD is indeed a new structural element unrelated to the GAIN domain (Figure 3.7B).

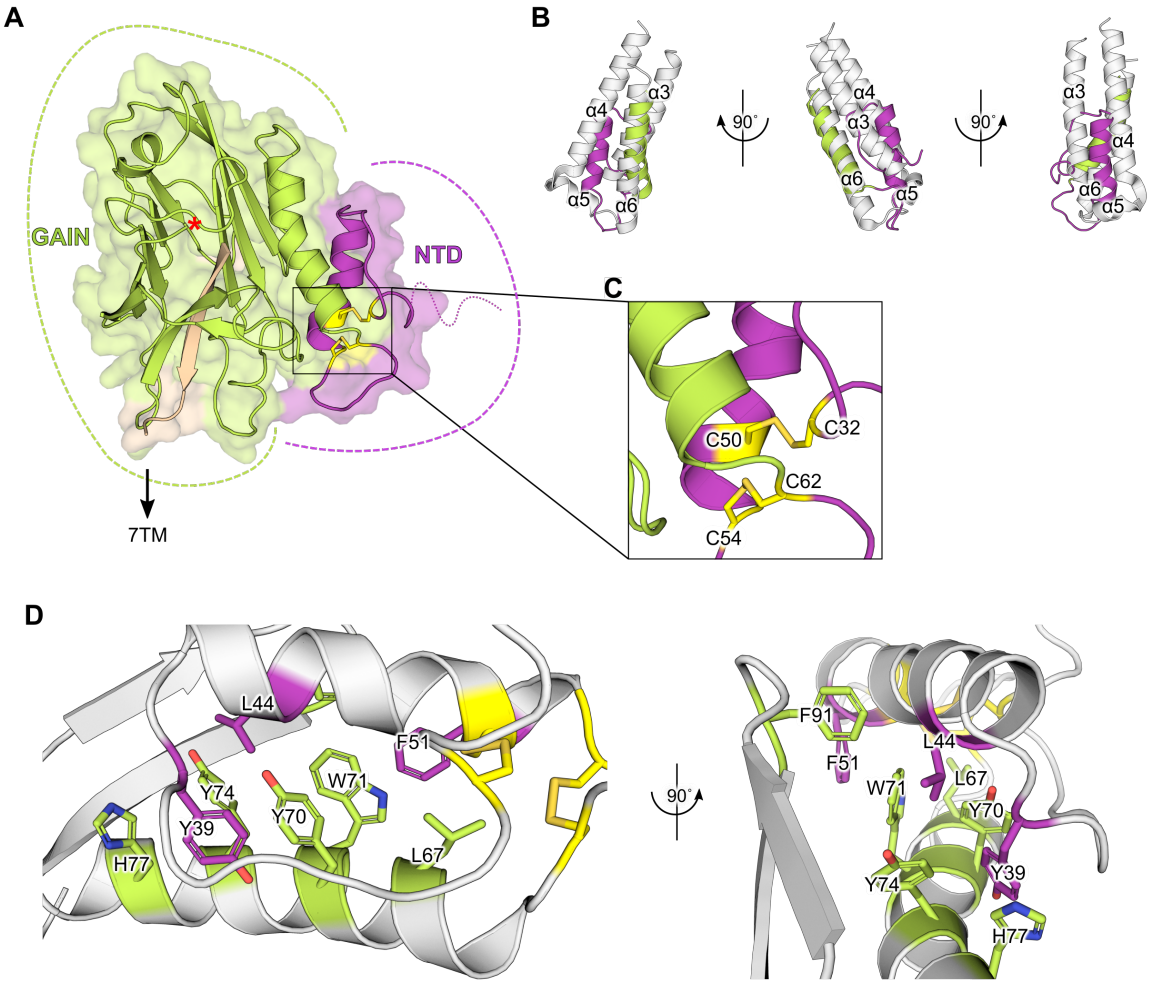


Figure 3.7: Structural analysis of the NTD. (A) Cartoon representation of GPR97^{ecto} structure obtained from X-ray crystallography experiments. The NTD is depicted in purple and the GAIN domain in green and beige (β13-strand). Cysteine residues involved in linking and stabilising the NTD are shown as yellow sticks. The red asterisk indicates the GPS cleavage site. **(B)** Structural alignment of the α6-helix (green) and NTD (purple) from GPR97 with the α3-α6-helices from Lphn1 (white, PDB ID: 4DLQ). **(C)** Close-up of the two disulphide bridges, C32-C50 and C54-C62, that are involved in the stabilisation of the NTD and GAIN domains. **(D)** Close-up on the extensive network of hydrophobic interactions stabilising the packing of the NTD against the GAIN domain. Residues from the NTD are depicted in purple and those from the GAIN domain are depicted in green.

A cysteine bridge stabilises the packing of the NTD α-helix against the subdomain A α6-helix (C54-C62). Another intradomain cysteine bridge secures the N-terminal flexible region to its α-helix (C32-C50) (Figure 3.7C). The NTD helix is further stabilised by extensive hydrophobic interactions with the GAIN domain, involving the residues Y39, L44, F51 from the NTD and L67, Y70, W71, Y74, H77 and F91 from the GAIN domain (Figure 3.7D). Strikingly, these hydrophobic residues are highly conserved specifically among GPR97 orthologs (Figure 3.6B) but not in other members of the aGPCR family (Appendix Figure C.2).

3.3.2 The GAIN domain

Overall, the GPR97 GAIN domain displays the same conserved fold observed in previously reported structures (Figures 3.7A and 3.2) and consists of a short subdomain A with only the α6-helix and a complete subdomain B. The electron density map also confirms the receptor is autoproteolysed at the conserved GPS motif (Figure 3.8A). One notable difference is the presence of an extensive loop (E112-D129) between β3- and β4-strands, which was predicted to be disordered in the initial bioinformatic analyses (Figures 3.8B and 3.1B).

Structural comparison with a model calculated by AlphaFold [313] revealed the two models to be similar, including for the long loop between β3- and β4-strands (Figure 3.8C). However, the NTD-helix is off-set compared to the experimental model. Interestingly, the N-terminal residues Q21-G27 were predicted with low confidence, which is in accordance with them being predicted as disordered during the bioinformatic analysis and corroborates the absence of electron density in the experimental data.

Intriguingly, human GPR97 possesses a non-conserved free cysteine C245 which is solvent-exposed and does not seem to be involved in a disulphide bridge (Figure 3.8D,E). Even though the two copies of GPR97 appear to have their free C245 close to each other in the ASU, the distance between the two residues measures over 7 Å, therefore making it impossible for them to form a disulphide bridge (Figure 3.8D). Instead, the presence of two positive Fo-Fc electron densities, within a continuous 2Fo-Fc electron density protruding from each of those two free cysteines suggests that C245 could have either

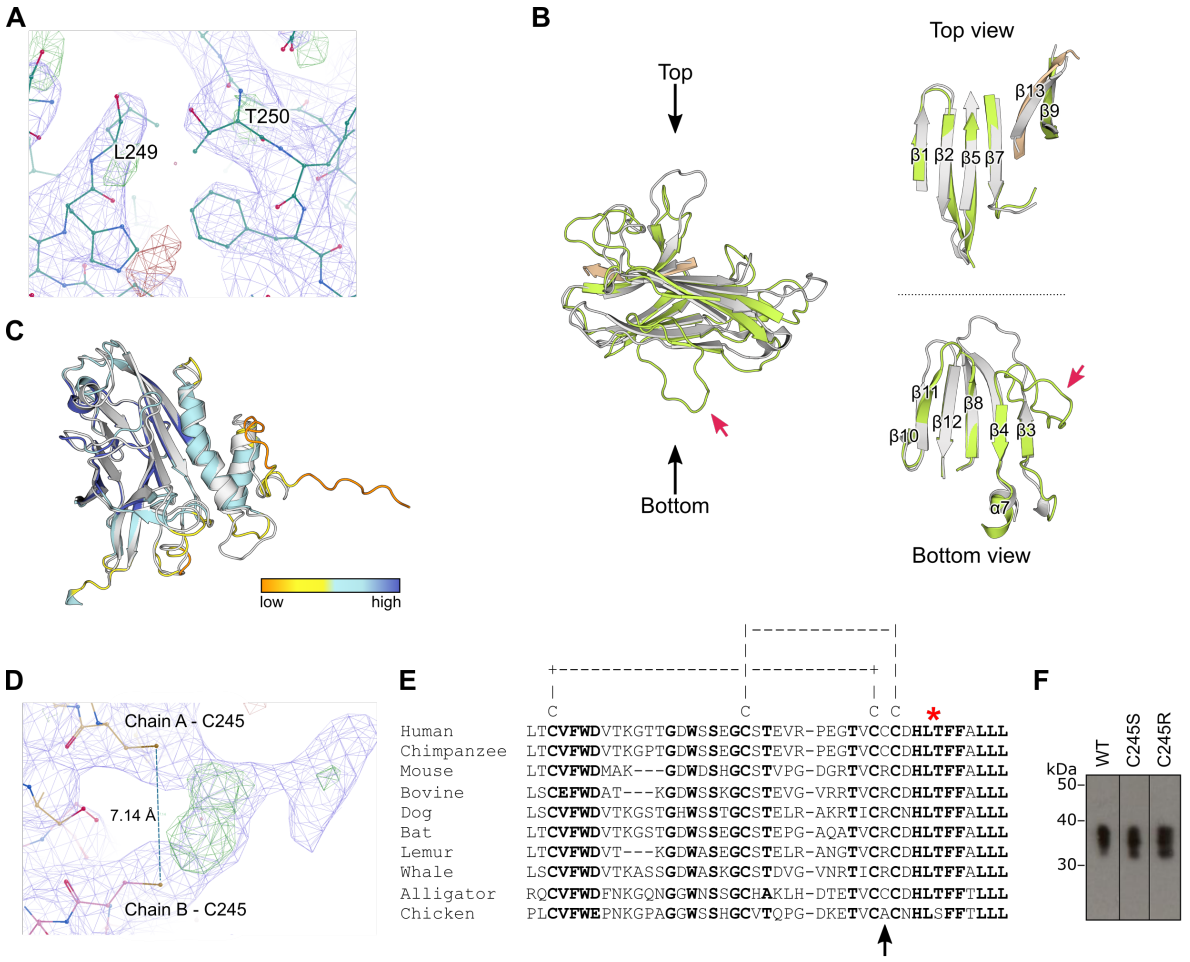


Figure 3.8: Structural analysis of GPR97^{ecto}. 2Fo-Fc electron density maps are shown at the 1 sigma level in purple. Fo-Fc electron density maps are shown at ± 3 sigma level in green and red, respectively. **(A)** Close-up on β12- and β13-strands in a 2Fo-Fc electron density map. The absence of peptide bond between L249 and T250 confirms the receptor autoproteolysis at the conserved GPS site. **(B)** Structural alignment of GPR97 GAIN domain (green and beige) with GPR56 GAIN domain (white, PDB ID: 5KVM) showing the conservation of all secondary structures. The long E112-D129 loop between β3- and β4-strands is indicated by a red arrow. **(C)** Structural comparison of the model predicted by AlphaFold with GPR97^{ecto} crystal structure (white). The AlphaFold model is coloured from very low model confidence (orange) to very high model confidence (blue). **(D)** Close-up on C245 from the two GPR97^{ecto} copies present in the crystal ASU in a 2Fo-Fc electron density map. The presence of two positive Fo-Fc electron densities at the tip of C245 side chain suggest a possible PTM. **(E)** Sequence alignment of the GPS site from various GPR97 orthologs showing the presence of five cysteines. The four cysteines forming the conserved C215-C244 and C233-C246 disulphide bridges are shown. The lone and non-conserved C245 is indicated by a black arrow. The autoproteolysis site in indicated by a red asterisk. **(F)** Western blot analysis of HEK293T supernatants comparing the expression of wild type (WT) GPR97^{ecto} with GPR97^{ecto} constructs possessing the C245S and C245R mutations.

undergone post-translational modifications (PTMs) or been neutralised by a component from our culture medium or purification buffer. Interestingly, expression trials of C245S or C245R mutants show a slight difference in protein expression with the appearance of

new species of lower molecular weight (Figure 3.8F). However, whether the presence of this free cysteine is of physiological and functional relevance remains to be elucidated.

3.4 GPR97 interaction with PR3 on the surface of neutrophils

3.4.1 Mapping the GPR97-PR3 binding interface

In order to locate the GPR97-PR3 binding interface, domain-swapped mFc-fusion probes were generated by interchanging the GAIN domains of GPR97 and GPR56, two closely related ADGRG subfamily members (Figure 3.9A). Surprisingly, both $GPR97^{NTD}$-$GPR56^{GAIN}$-mFc and $GPR56^{PLL}$-$GPR97^{GAIN}$-mFc probes showed a similar binding profile to the surface of neutrophils as for the GPR97-mFc probe (Figure 3.9B). Next, domain-truncated GPR97-mFc probes containing either the NTD (Q21-N63) or the GAIN domain (N63-R270) confirmed that both the NTD and GAIN domains of GPR97 interact with PR3 (Figure 3.9A,C). Interestingly, those domain-swapped and domain-truncated probes induced much weaker neutrophil activation phenotypes compared to GPR97-mFc, as observed by smaller morphological changes and weaker IL-8 production (Figure 3.9D,E). Whether these results suggest the presence of two independent PR3-binding sites on GPR97, or the presence of a single PR3-binding interface spanning both GPR97 NTD and GAIN domains, remains to be elucidated. It is, however, possible to conclude that the maximal GPR97-induced neutrophil activation is achieved when PR3 interacts simultaneously with both binding interfaces located in the GPR97-NTD and -GAIN domains.

Interestingly, only the truncated $GPR97^{GAIN}$-mFc probe induced a moderate increase in PR3 enzymatic activity, highlighting the importance of the GAIN binding interface (Figure 3.9F). This effect was not observed with the $GPR56^{PLL}$-$GPR97^{GAIN}$-mFc probe, possibly because of the large PLL domain on top of GPR97 GAIN domain, which could either obstruct the PR3 catalytic cleft, or sterically hinder GPR97-PR3 association. Alternatively, the NTD may be necessary for proper PR3 association and positioning on GPR97.

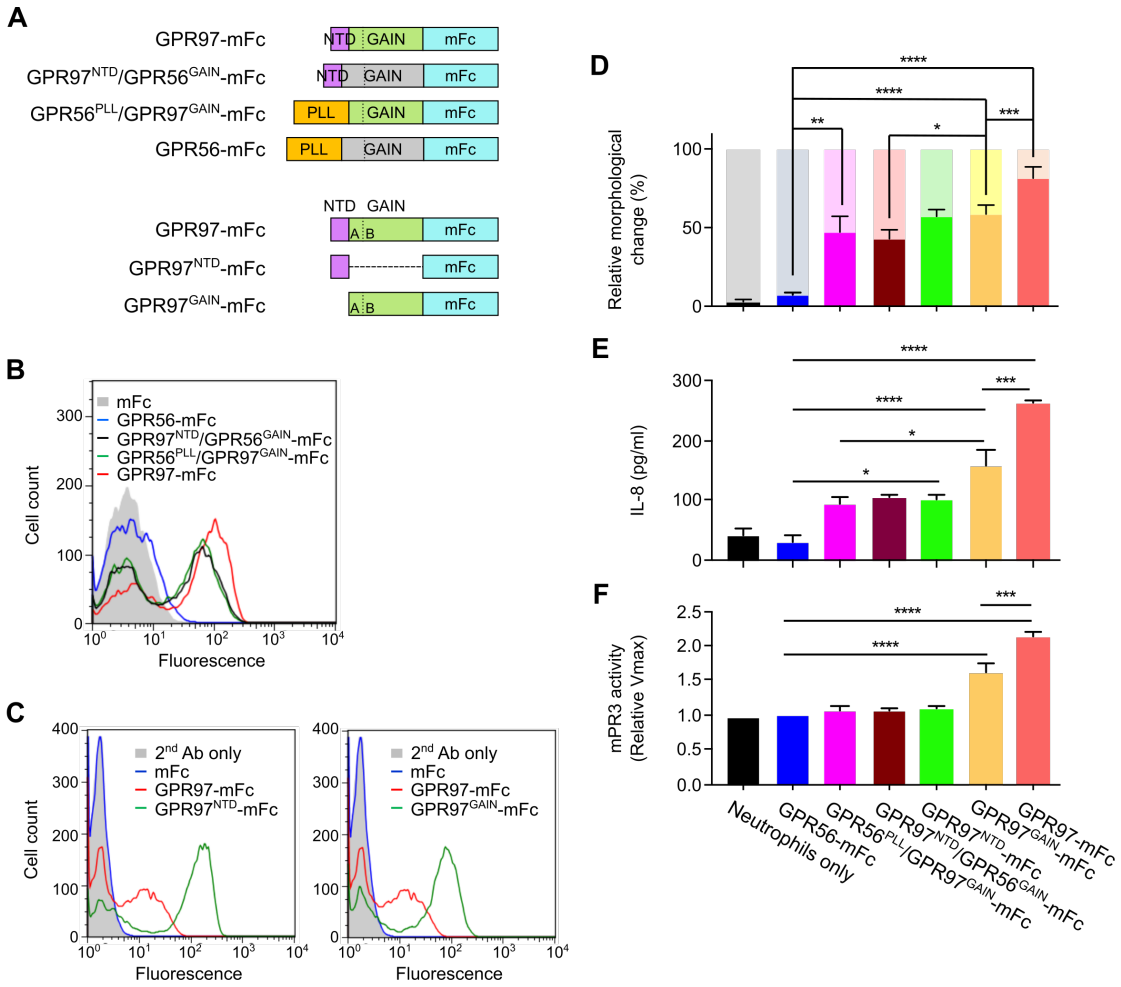


Figure 3.9: Mapping the GPR97-PR3 binding interface. **(A)** Schematic of the GPR97 and GPR56 domain-swapped mFc-fusion probes as well as of the domain-truncated GPR97-mFc probes. **(B)** FACS-based ligand binding assay showing that GPR97^{NTD}-GPR56^{GAIN}-mFc and GPR56^{PLL}-GPR97^{GAIN}-mFc probes display a similar binding profile as GPR97-mFc. **(C)** FACS-based ligand binding assay showing that GPR97^{NTD}-mFc and GPR97^{GAIN}-mFc retained their PR3 binding ability. **(D-F)** Comparison of neutrophil morphological changes (n = 3 repeats), IL-8 production (n = 4 repeats) and *ex vivo* PR3 enzymatic activity (n = 6 repeats) upon neutrophil incubation with the indicated GPR97-mFc probe (mean ± SEM, *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001). **(D)** The bright and dim colours represent the percentage of irregular-shaped and round-shaped neutrophils, respectively. **(F)** The results are shown in relative Vmax of the *ex vivo* PR3 enzymatic activity.

3.4.2 *In vitro* and *in vivo* analyses of GPR97-PR3 interaction

Direct GPR97-PR3 binding was first investigated by SPR. The binding experiment was performed with GPR97^{ecto} and CD177^{ecto} (as a positive control) immobilised on an SPR sensor chip and with neutrophil-purified PR3 (nPR3) as the solute analyte. Notably, nPR3 does not possess a mutation at its catalytic serine and is therefore catalytically active.

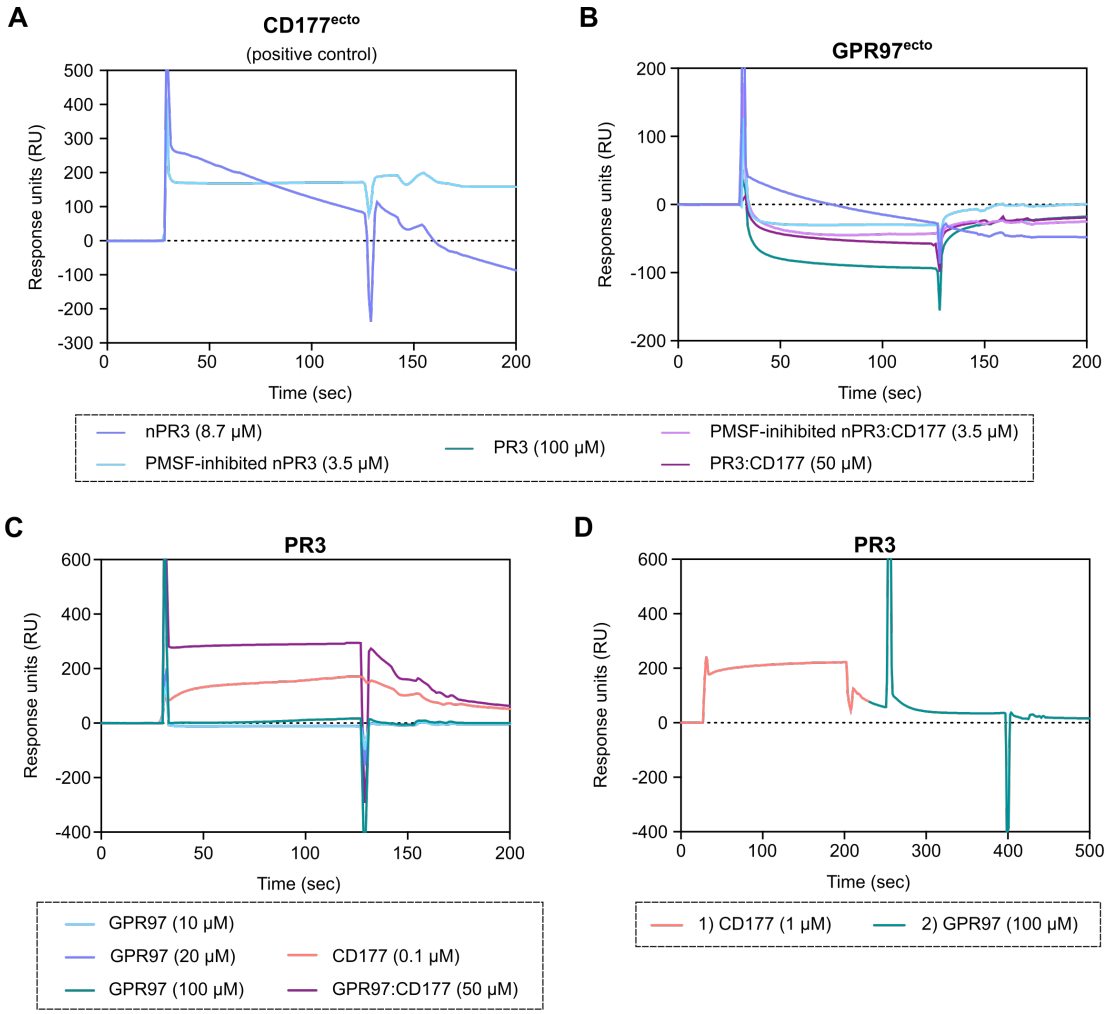


Figure 3.10: SPR analysis of GPR97-PR3 interaction. (A,B) For binding experiments using nPR3, GPR97^{ecto} and CD177^{ecto} were immobilised at a surface density of 740 and 770 RU, respectively. Addition of nPR3 resulted in an initial increase of 50 RU followed by a decrease of 80 RU for GPR97^{ecto}, and in an initial increase of 280 RU followed by a decrease of 200 RU for CD177^{ecto}, suggesting proteolysis of both receptors by nPR3. For binding experiments using PMSF-inhibited nPR3, GPR97^{ecto} and CD177^{ecto} were immobilised at a surface density of 560 and 470 RU, respectively. For all other binding experiments, GPR97^{ecto} and CD177^{ecto} were immobilised at a surface density of 390 and 560 RU, respectively. CD177^{ecto} displays strong direct binding with PR3 while GPR97^{ecto} does not show any binding signal. **(C,D)** Recombinant PR3 was immobilised at a surface density of 350 RU and showed strong binding to CD177^{ecto} but no binding to GPR97^{ecto}.

Strikingly, after an initial increase in signal suggesting a binding event, the response units (RU) drastically decreased, suggesting GPR97 proteolysis by nPR3. A similar behaviour was also observed with CD177^{ecto}, albeit with a much stronger binding, since the use of a dissociation buffer was necessary to remove bound nPR3 and therefore stop CD177 proteolysis. Unexpectedly however, when the experiment was repeated with a PMSF-inhibited nPR3, no GPR97-PR3 binding signal could be observed. This

was the case whether a 30-fold excess of PR3 or a 1:1 pre-incubated mixture of nPR3:CD177^{ecto} was injected.

A reverse binding experiment was then performed, where a recombinantly-produced and catalytically inactive PR3 (detailed in Chapter 4) was immobilised and GPR97^{ecto} and CD177^{ecto} (as positive control) were used as the solute analytes. While 0.1 µM of CD177^{ecto} was sufficient to observe a strong binding signal, no signal could be observed with 10 µM, 20 µM or even 100 µM of GPR97^{ecto}. Similarly, direct injection of a 1:1 pre-incubated mixture of GPR97^{ecto}:CD177^{ecto} or sequential injection of CD177^{ecto} directly followed by GPR97^{ecto} did not result in a GPR97-PR3 binding signal.

To further confirm these results, nPR3 was incubated with CD177-expressing HEK293T cells and the GPR97-mFc probe was used in a FACS-based ligand binding assay. While nPR3 readily bound and could be detected on the surface of the cells, no binding signal could be observed from GPR97-mFc (Figure 3.11A,B).

Together, these inconclusive binding results indicate that the current model of GPR97 binding to CD177-bound PR3 is incomplete and suggest that there exists missing components essential for GPR97-PR3 interaction (such as other protein partners or extracellular matrix components) which are likely present on the surface of neutrophils, but missing in the experimental setups. In addition, the initial binding data with GPR97 and active nPR3 is intriguing and needs to be repeated and further investigated as this could help unravel how GPR97 potentiates PR3 enzymatic activity as discussed in section 1.4.3.4.

3.4.3 Identification of a new receptor interactome

The absence of direct GPR97-PR3 interaction in binding assays suggested that other stabilising interacting partners were absent in the experiments. In addition, neutrophil activation upon GPR97 interaction strongly implicates the involvement of a signalling receptor expressed on the surface of neutrophils and capable of intracellular signal transduction. The latter would either be triggered by GPR97-PR3 interaction or by GPR97-potentiated PR3 proteolytic activity.

CD177 is unlikely to be the signalling receptor as it possesses a GPI-anchor and is therefore incapable of direct signalling. Potential protein partners and/or substrates

of the PR3:CD177 complex expressed on the surface of neutrophils include EPCR [314], CD16b (also known as FcγRIIIB) [63], Mac-1 (also known as CD11b/CD18) [77] and PAR2 [315].

In order to determine the signalling receptor/complex, all of the mentioned neutrophil receptors plus GPR97 were co-expressed in different combinations with CD177 in HEK293T cells before being subjected to a FACS-based ligand binding assay using the GPR97-mFc probe (Figure 3.11A,B). Surprisingly, a positive GPR97-PR3 binding signal was detected only in cells co-expressing CD177, GPR97, CD16b and PAR2. Moreover, the substitution of any member of this four-receptor complex to an analogous protein receptor consistently resulted in a loss of binding signal (Figure 3.11C).

A proximity ligation assay (PLA) was then performed in order to assess whether these receptors were closely associated on the surface of neutrophils, and confirmed the receptors' proximity to each other in a pairwise fashion (Figure 3.11D). Taken together, these results indicate the involvement of a minimal clustered CD177/GPR97/PAR2/CD16b receptor interactome for GPR97-PR3 binding.

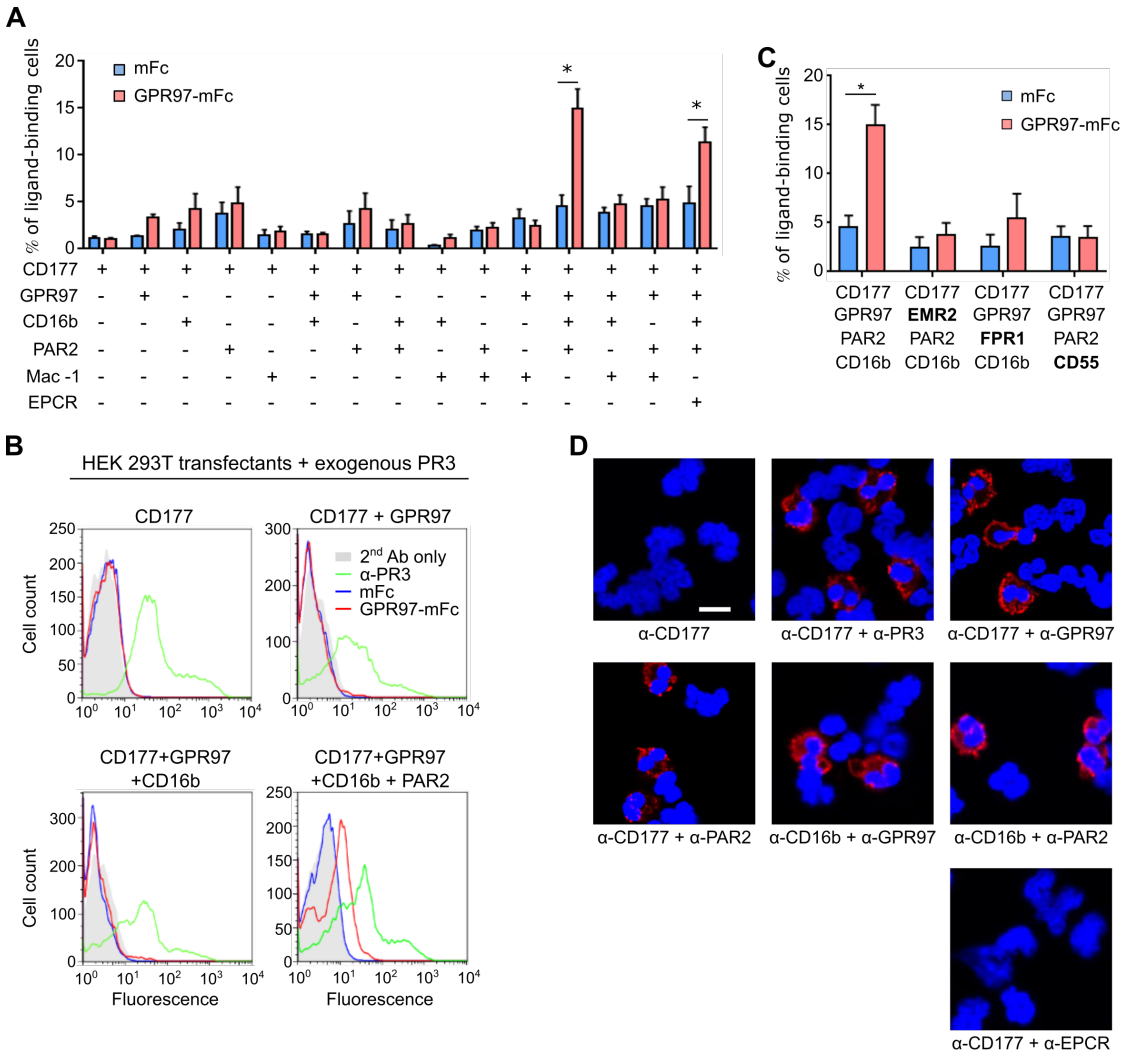


Figure 3.11: Identification of a new receptor interactome. **(A)** Summary of all FACS-based PR3-binding assays performed on HEK293T cells expressing various receptors in various combinations (n = 3 repeats, mean ± SEM). A positive GPR97-PR3 binding signal was detected only in cells co-transfected with CD177, GPR97, CD16b and PAR2 (*p < 0.05). **(B)** Representative profiles of FACS-based PR3-binding assays performed on HEK293T cells expressing various receptors as indicated. **(C)** Substitution of GPR97, CD16b and PAR2 by EMR2, the GPI-anchored CD55 and formyl peptide receptor 1 (FPR1), respectively, consistently resulted in a loss of GPR97-PR3 binding signal (n = 4 repeats, mean ± SEM, *p < 0.05). **(D)** PLA analyses showing that PR3, CD177, GPR97, PAR2 and CD16b are all closely associated with each other on the surface of resting neutrophils. By contrast, EPCR is not associated with CD177 on the surface of resting neutrophils. Scale bar is 10 µm.

3.4.4 Functional relevance of the CD177/GPR97/PAR2/ CD16b interactome

Out of the newly identified receptors, CD16b possesses a GPI-anchor and PAR2 a 7TM domain as a member of the GPCR superfamily. In particular, PAR2 is the predominant member of the PAR family that is expressed by human neutrophils. Previous studies

have shown that PAR2 activation induced similar neutrophil activation phenotypes to those triggered by GPR97-mFc, such as changes in cell morphology, enhanced IL-8 secretion and reduced trans-endothelial migration (data not shown) [9]. Interestingly, members of the PAR family are regulated by proteolytic cleavage. More specifically, the PAR2 ectodomain can be cleaved by PR3, leading to exposure of a tethered ligand, which can subsequently interact with the 7TM domain and activate the receptor signalling cascade [46]. Altogether, this highly suggests PAR2 as the signal transducer recruited and activated by GPR97-potentiated PR3.

To confirm this, neutrophils were incubated with a functionally blocking anti-PAR2 mAb (MAB3949), PAR2-specific antagonists (ENMD-1068, FSLLRY-NH2) or PAR1-specific antagonist (vorapaxar) before their level of activation was assessed by IL-8 production (Figure 3.12). Strikingly, enhanced IL-8 production by GPR97-induced neutrophils was significantly and specifically attenuated when PAR2 was inhibited, clearly demonstrating that PAR2 is the essential signalling receptor targeted by the GPR97-potentiated PR3.

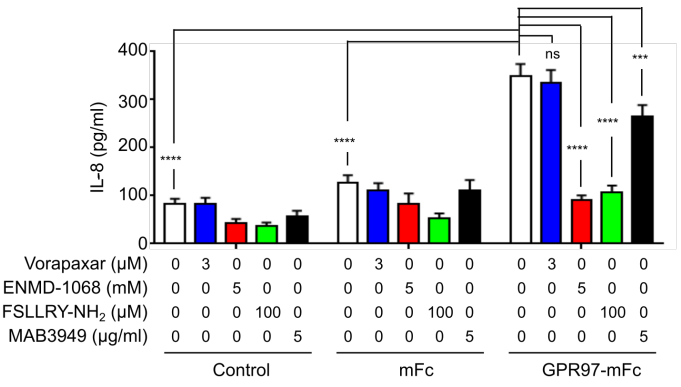


Figure 3.12: Decrease in neutrophil activation upon inhibition of PAR2. Comparison of IL-8 production following neutrophil incubation with the indicated reagents (n = 4 repeats, mean ± SEM). Inhibition of PAR2 functional activity led to a significant decrease of GPR97-induced neutrophil activation (***p < 0.001; ****p < 0.0001; ns, non-significant).

Altogether, this work suggests a novel GPR97-PAR2 transactivation mechanism, where GPR97 interacts with PR3 presented by the CD177-associated interactome and promotes its enzymatic activity, which in turn cleaves and activate PAR2.

3.5 Discussion

3.5.1 Functional importance of GPR97 NTD

GPR97/ADGRG3 is a member of the ADGRG subfamily, which contains some of the smaller-sized aGPCR members including GPR56, GPR114 and GPR97 itself. GPR97 ectodomain is comprised of a small NTD followed by the conserved GAIN domain. Interestingly, despite the small size of the NTD compared to the GAIN domain, the presence of the NTD appears to be required for PR3 binding and essential for eliciting maximal neutrophil activation.

These results fall into line with previous reports highlighting the functional importance of the N-terminal domains present in the extracellular region of aGPCRs. For instance, Lphn1 and BAI3, two well-studied aGPCRs, contain various adhesion-like motifs such as the HormR, olfactomedin and lectin domains for Lphn1 or the thrombospondin type 1 repeats and CUB domain for BAI3. In processes such as synapse development or maintenance, Lphn1 has been shown to trigger either cell adhesion or repulsion depending on the various associations with its binding partners through its olfactomedin and lectin domains [156]. On the other hand, BAI3 potentially regulates synapse formation via high affinity interactions with C1q-like proteins through its thrombospondin type 1 repeats and CUB domain [180]. Similarly, GPR56 possesses an additional N-terminal PLL domain with a suggested evolutionarily conserved function in central nervous system myelination and in particular in oligodendrocyte development [309]. A more recent study revealed that the extracellular region of GPR126 is also comprised of various and distinct domains, namely the CUB, PTX, SEA and HormR domains, which are critical for proper myelination of the peripheral nervous system, ear development and heart development [310].

Notably, in all of those cases, aGPCR functions are elicited by interactions of binding partners with those additional N-terminal domains, independently of the GAIN domain. This work, however, underlines for the first time the importance of the GAIN domain in addition to the NTD for proper receptor interaction and function.

3.5.2 A scaffolding role for subdomain A for the GAIN domain

As one of the smallest aGPCRs, GPR97 possesses an extracellular region of approximately 30 kDa. The crystal structure of GPR97^{ecto} reveals a small NTD sitting on top of a truncated GAIN domain lacking the first five α-helices of subdomain A. As truncated GAIN domains have been previously observed in GPR56 and GPR126, and given that GPR97 could not accommodate a ∼35 kDa complete GAIN domain, it was thus expected to find the GPR97 GAIN domain to be truncated. Strikingly however, GPR97 lost the majority of subdomain A and only retained the last α6-helix.

In comparison, the GAIN domains of GPR56 and GPR126 are only missing α1-α3 helices and α1-helix of subdomain A, respectively. This suggests that the full subdomain A is not necessary for proper receptor function. Additionally, a sequence alignment of the GAIN domain of all aGPCRs reveals that several other members are also likely to possess a smaller GAIN domain (Appendix Figure C.3). These include all members of the ADGRE subfamily (EMR1-4/ADGRE1-4 and CD97/ADGRE5), GPR123/ADGRA1, GPR133/ADGRD1, GPR111/ADGRF2, GPR115/ADGRF4 and all members of the ADGRG subfamily (GPR56/ADGRG1, GPR64/ADGRG2, GPR97/ADGRG3, GPR112/ ADGRG4, GPR114/ADGRG5, GPR126/ADGRG6 and GPR128/ADGRG7).

Strikingly, all GAIN domain truncations happen in subdomain A, while subdomain B remains intact. In fact, subdomain B contains highly conserved residues, cysteine bridges as well as the GPS motif, and therefore appears to be the functionally important part of the GAIN domain. This implies that subdomain A could serve mostly as a scaffolding platform, whose role is to properly orient the functionally important N-terminal domains on the surface of the cell membrane.

3.5.3 Importance of the free cysteine C245

GPR97 is the only aGPCR possessing a free and solvent-exposed cysteine residue within the highly conserved GPS motif inside of subdomain B of the GAIN domain (Appendix Figure C.1). Notably, the proximity of C245 to the 7TM region may suggest an interaction with another cysteine present in one of the 7TM extracellular loops. However, a recent cryo electron microscopy structure of GPR97 7TM region negates

such a hypothesis. Indeed, the structure encompassing residues S263 to P527 does not display the presence of another free cysteine in one of the extracellular loops and therefore confirms C245 to not be involved in an intramolecular disulphide bridge [184]. A closer look at the electron density map reveals that this cysteine seems to form a covalent bond with a still unknown small molecule. Moreover, mutating this cysteine appears to result in a different expression pattern, and thus highly suggests that the residue undergoes PTM during protein expression. Cysteine PTMs are usually not catalysed by enzymes, but instead occur via direct reaction of thiols with ROS, nitrogen, chlorine and electrophilic species, metals and other thiols. Possible PTM on cysteines include *S*-nitrosylation (SNO), *S*-glutathionylation (SSG), *S*-acylation, *S*-cysteinylation, *S*-sulfhydration, and *S*-sulphenylation (Cys-SOH), suggesting that experiments such as native mass spectrometry should enable the identification of the PTM on C245.
Although GPR97 is expressed on all granulocytes, the receptor is particularly abundant in neutrophils where it was shown to participate in the immune cell activation and degranulation, and could therefore be in contact with ROS. Oxidative PTMs on cysteines are known to play a major role in intracellular redox signalling, but little is known about their potential role in the extracellular environment. Besides, the low abundance of cysteine residues within the human proteome (about 2.2% [316]) led them to be mostly observed in functional sites of proteins. Therefore, this additional and solvent-exposed cysteine within the highly conserved GPS motif of GPR97 could account for a potential role or regulation mechanism in inflammation that needs to be uncovered.

3.5.4 A GPR97-mediated PAR2 transactivation on human neutrophils

A novel GPR97-PAR2 transactivation mechanism was identified, where interaction of GPR97 to PR3 presented by the CD177/GPR97/PAR2/CD16b interactome promotes PR3 enzymatic activity, allowing it to cleave and thus activate PAR2 (Figure 3.13). Strikingly, this combination of receptors is highly analogous to the recently reported interactome involving the GPI-anchored RECK, the aGPCR GPR124/ADGRA2, the GPCR Frizzled (FZD4) and the single transmembrane protein Low-density lipoprotein

receptor 5 (LRP5) [317, 168, 318]. More specifically, the RECK-GPR124 receptor complex was shown to mediate Wnt7a interaction with its co-receptors FZD4 and LRP5, and therefore its bioactivity in central nervous system angiogenesis and blood-brain barrier integrity.

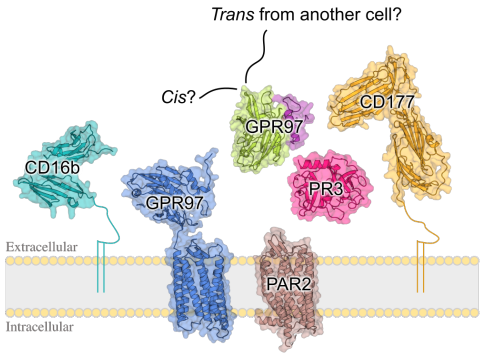


Figure 3.13: Simplified schematic of the CD177/GPR97/PAR2/CD16b interactome on the surface of neutrophils. The full-length GPR97 model was built using the experimental model of the ECD (PDB ID: 7QU8) and the 7TM domain recently reported (PDB ID: 7D76) [184]. The CD177, PR3, CD16b and PAR2 models were extracted from the AlphaFold Protein Structure and PDB Databases (PDB ID: 1FUJ [319], 1NFL [320] and 5NDD [321]), respectively.

Intriguingly, the formation of the PR3/CD177/GPR97/PAR2/CD16b interactome appears to only be possible on the surface of human neutrophils. Indeed, many interspecies differences can be observed for most members of the interactome, preventing its formation on the surface of neutrophils in other species. Human PR3 differs from gibbon and mouse PR3 by containing a distinctive hydrophobic patch which mediates its interaction with CD177 [36]. Mouse CD177 significantly differs from human CD177 with its ECD that is twice as big as the human one. Human GPR97 is only expressed on the surface of neutrophils while its mouse counterpart appears to be involved in various immune processes such as B-lymphocyte fate decision [248], obesity-associated macrophage inflammation [249], and lymphatic endothelial cell migration [245]. And finally, CD16b is a member of the Fcγ receptor family that is only expressed by human neutrophils [322]. Altogether, this explains the absence of GPR97-PR3 binding when mouse neutrophils were incubated with human GPR97-mFc, and underlines the human-specific feature of GPR97 interaction with interactome-associated PR3. The GPR97-PAR2 transactivation mechanism therefore appears to be specific to human neutrophils.

PAR2 is a GPCR that is activated through proteolysis by a number of serine proteases such as trypsin [323], mast cell tryptase [324], PR3 [325], and matriptase [326], or by transactivation through PAR1, during which the exposed tethered ligand of PAR1 interacts with PAR2 7TM and subsequently activate the receptor [327]. Identification of this novel GPR97-PAR2 transactivation mechanism presents a substantial advance in mechanistic understanding of signals leading to neutrophil activation in the inflammatory response. However, what ultimately activates this mechanism and how this is regulated remains to be elucidated. To answer those questions, investigating the role of the other components CD177, GPR97 and CD16b appears to be essential.

Chapter 4
Production and characterisation of PR3

Proteinase 3 (PR3) is a neutrophil serine protease that is presented on the neutrophil cell surface through interaction with its GPI-anchored receptor CD177. The presented work describes the recombinant production and isolation of the protein for use in further biochemical, structural and functional studies. PR3 appears to be aggregating at different stages of purification. As a result, different methods and resources including the use of additives, reducing the expression time and mutagenesis, were explored to tackle the different issues presented. In the end, a hydrophobic patch present on the surface of PR3, and the protein predisposition for dimerisation and/or oligomerisation appeared to be the main factor causing protein misbehaviour. PR3 is not only unstable upon prolonged contact with the cell culture medium, but also binds non-specifically to a plethora of protein impurities, making its purification quite challenging. Mutation of the key residues responsible for protein misbehaviour led to creation of a panoply of well-behaving PR3 variants that can be used in further studies. It also enabled to narrow down and identify important residues involved in binding with CD177.

4.1 Expression and initial purification trials of PR3

4.1.1 Bioinformatic analysis

Human PR3 is one of the four NSPs contained within the azurophilic and secretory granules of neutrophils. PR3 shares the same fold as members from the chymotrypsin

family of serine proteases [319], which consists of two β-barrels, forming the N- and C-terminal lobes respectively, between which sits the catalytic groove composed of the catalytic triad H71-D118-S203 (Figure 4.1A).

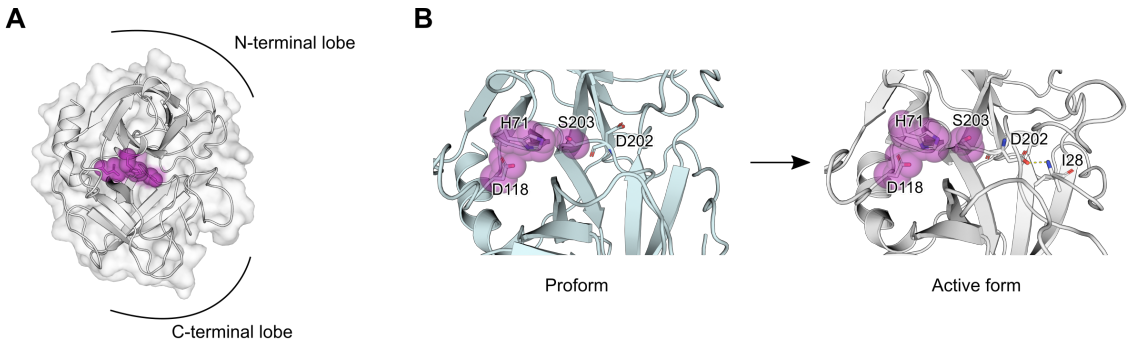


Figure 4.1: The conserved serine protease fold of PR3. (A) Cartoon representation of PR3 (PDB ID: 1FUJ) displaying the N- and C-terminal lobes. The catalytic triad H71-D118-S203 is represented as sticks and highlighted in purple spheres. **(B)** Comparison of the catalytic cleft between the active form of PR3 and a homology model of its proform obtained from chymotrypsinogen A (PDB ID: 2CGA) [328].

Similarly to other serine proteases, PR3 possesses a secretion signal peptide (residues M1-A25) followed by an N-terminal propeptide (A26-E27), whose proteolytic cleavage by cathepsin C results in a catalytically active PR3. Indeed, this newly created N-terminus inserts into the core of the enzyme where the amino group of I28 interacts with the formerly buried side chain of D202. This hydrogen bonding results in a conformational rearrangement of the catalytic S203, and therefore in the proper positioning of the H-D-S catalytic triad to perform proteolysis (Figure 4.1B). PR3 also possesses a C-terminal propeptide (residues R249-P256) which appears to be removed in the mature form of all serine proteases, although its functional significance remains unknown.
Lastly, PR3 carries two glycosylation sites at N129 in the N-terminal lobe and N174 in the C-terminal lobe [319].

4.1.2 Construct screening

To avoid cell toxicity and other practical issues resulting from strong protease activity, the catalytic S203 was mutated into alanine in all of the constructs to produce catalytically inactive variants. In addition, previous studies have reported successful recombinant

production of both proPR3 and PR3 using HEK293 cells [36, 329, 330, 331], which thus confirms the suitability of this expression system for production of PR3 variants.

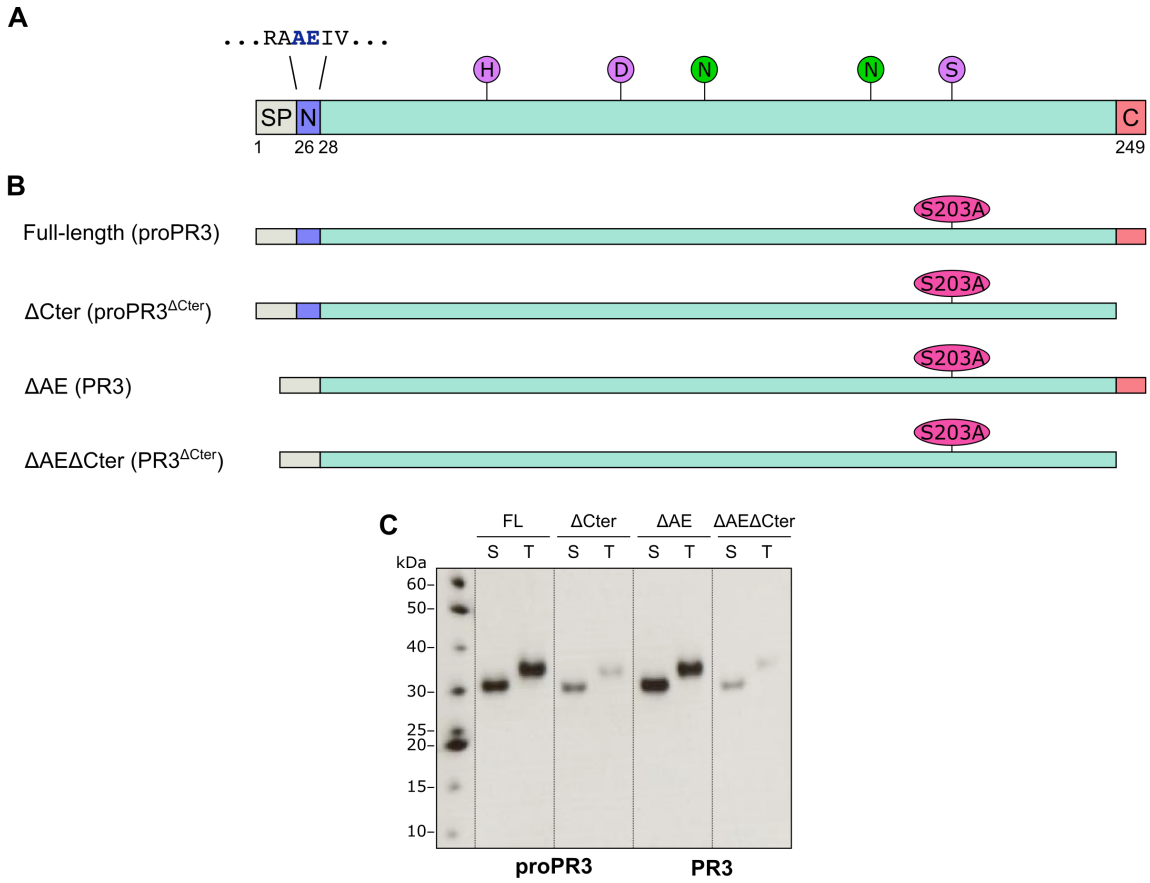


Figure 4.2: Expression trials of PR3 constructs. (A) Schematic architecture of PR3 with its amino acid numbering shown below. The H71-D118-S203 catalytic triad is indicated by purple spheres, and the two glycosylation sites N129 and N174 by green spheres. The N-terminal propeptide A26-E27 is highlighted in bold blue. Abbreviations are: SP, signal peptide; N, N-terminal propeptide; C, C-terminal propeptide. **(B)** Schematic describing the different PR3 variants used. All constructs have the catalytic S203 mutated to alanine. **(C)** Western blot comparing the expression levels of PR3 variants produced in either HEK293T (T) or HEK293S GnTI$^-$ (S) cells.

Constructs containing the N-terminal propeptide A26-E27 were selected to produce proPR3 while variants lacking those two residues were selected to produce PR3. For both proPR3 and PR3, constructs were designed with the native signal peptide, and with or without the C-terminal propeptide (Figure 4.2A-B).

Interestingly, comparison of the expression levels between the different constructs revealed a stronger expression from the constructs possessing the C-terminal propeptide. This suggests that not only the C-terminus remains uncleaved in PR3 variants expressed by HEK293 cells, but also that it seems to increase protein production (Figure 4.2C).

Another notable observation is the difference of protein size when produced by HEK293T or HEK293S GnTI$^-$ cells. HEK293T-produced PR3 appears to be up to 5 kDa bigger than HEK293S GnTI$^-$-produced PR3. As PR3 is known to only carry two glycosylation sites at N129 and N174, this suggests that HEK293T-produced PR3 is heavily glycosylated, with 2-3 kDa sugars at each glycosylation site.

4.1.3 Protein purification

As small-scale expression trials showed high level of protein expression, full-length proPR3 and PR3 constructs were selected for large-scale production by HEK293S GnTI$^-$ cells.

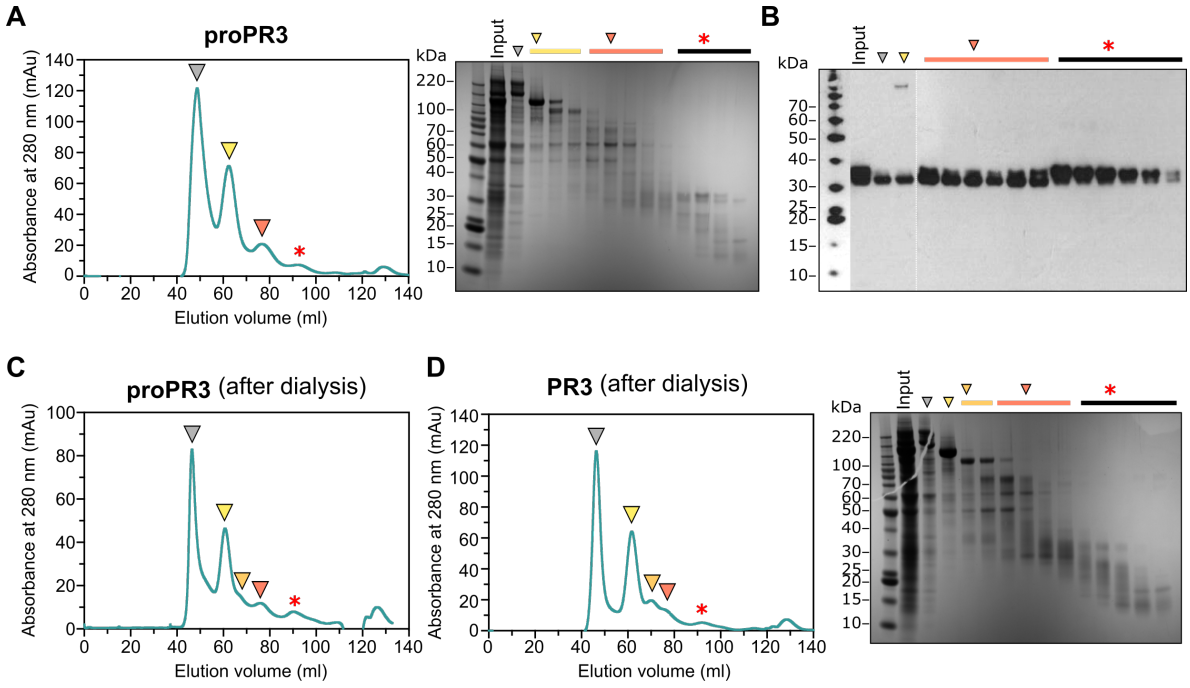


Figure 4.3: Initial large-scale purification trials of PR3. (A,C,D) Representative proPR3 and PR3 SEC purification profiles. The protein peak is indicated by a red asterisk and is preceded by a void (grey arrowhead) and impurity peaks (yellow, orange and red arrowheads), which are present in all purification from HEK293 cells. **(A)** proPR3 SEC profile obtained from a standard purification protocol. The corresponding SDS-PAGE gel shows the SEC input and fractions from the void, impurity and proPR3 protein peaks. proPR3 fractions appear to be highly contaminated with either degradation products or small protein contaminants. **(B)** Western blot analysis tracking the presence of proPR3 from the SEC input to the the different peaks obtained. A white line separates lanes extracted from different western blots. **(C)** proPR3 SEC profile obtained after dialysis replaced diafiltration as the method for buffer exchange. **(D)** PR3 SEC profile obtained after dialysis replaced diafiltration as the method for buffer exchange. The corresponding SDS-PAGE gel shows the SEC input and fractions from the void, impurity and PR3 protein peaks. Similarly to proPR3, PR3 fractions appear to be highly contaminated with either degradation products or small protein contaminants.

However, attempts at purifying proPR3 and PR3 using the same purification protocol

as previously described for GPR97^{ecto} resulted in a very limited amount of protein, whose quality and quantity were sub-optimal for structural studies (Figure 4.3). As ∼30 kDa globular proteins, both proPR3 and PR3 are expected to elute around the same volume as for GPR97, which is around 90 ml on a HiLoad® 16/600 Superdex® 200 column. Here, the observed proPR3 protein peak barely reached 10 milli absorbance units (mAu) and appeared to be highly contaminated with degradation products and/or protein contaminants according to SDS-PAGE gel analysis (Figure 4.3A). In addition, the gel revealed two distinct bands at ∼30 kDa, suggesting the existence of two glycosylated proPR3 species.

Western blot analysis of protein fractions obtained after SEC confirmed the existence of two proPR3 species and revealed that the protein peak is mainly composed of the top proPR3 species while the bottom proPR3 species is either lost in the void or elutes along with impurities (Figure 4.3B). In addition, proPR3 appeared to elute in all SEC fractions, which suggests that the protein either sticks to and elutes along with other protein contaminants or interacts with the matrix of the Superdex® column. Nevertheless, the reasons behind the different behaviour of both proPR3 species and why the bottom, hence smaller, proPR3 species elutes first remain unknown.

Furthermore, replacing the diafiltration step by dialysis to enable a more gentle buffer exchange also failed to improve the yield and purity of the protein (Figure 4.3C,D). As this was observed for both proPR3 and PR3, whether they were expressed by HEK293T or HEK293S GnTI$^-$ cells, this suggested a consistent protein misbehaviour was happening during the purification process. A further in-depth investigation of each purification step was therefore required to optimise the purification of PR3.

4.2 Optimisation of the purification protocol

Following several unsuccessful purifications of proPR3 and PR3 using a standard purification protocol, protein samples were collected at different stages of purification and analysed by western blot in order to pinpoint the step(s) that required optimisation. Interestingly, protein signal for both proPR3 and PR3 sharply decreased in the sample collected after buffer exchange, which suggested the protein was unstable either during

cell harvest, including centrifugation, filtration and storage at 4°C, or during buffer exchange (Figure 4.4.

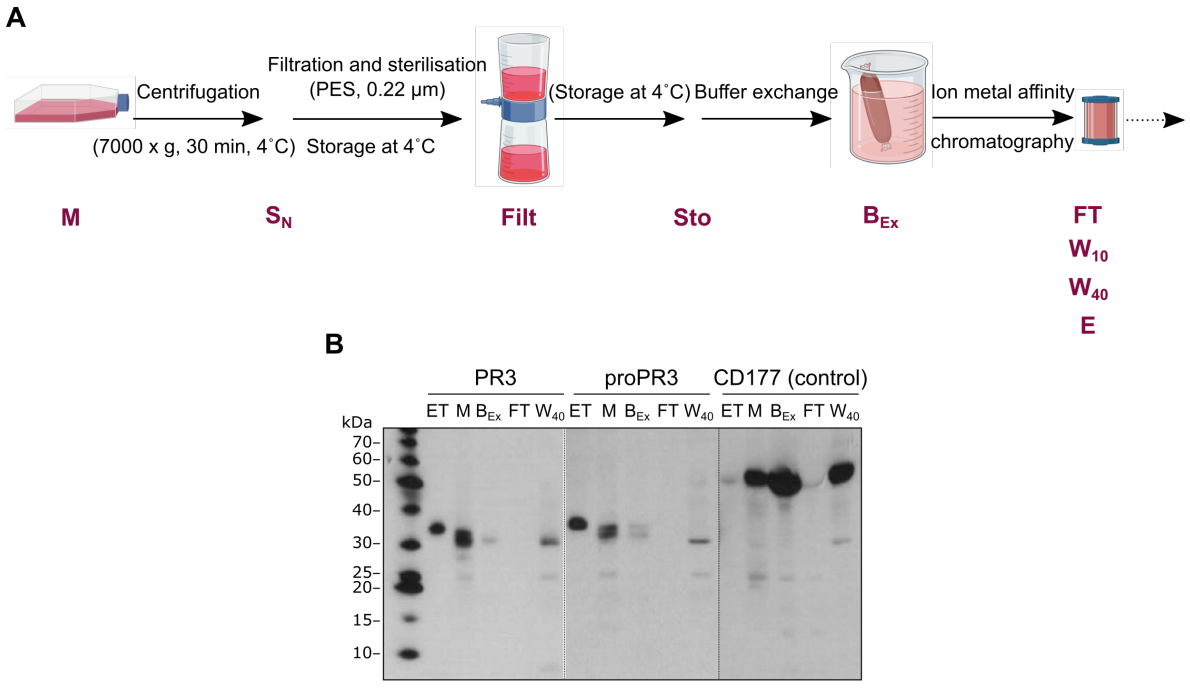


Figure 4.4: Monitoring proPR3 and PR3 during purification. **(A)** Schematic diagram depicting every step of protein harvesting and purification. The corresponding protein sample labellings (M, S_N, Filt, Sto, B_{Ex}, FT, W_{10}, W_{40} and E) are depicted below and will be used in all figures throughout the chapter. Parts of the figure were made using Biorender. **(B)** Western blot comparing the protein level between samples taken from a small-scale expression test (ET) and at different steps of the purification protocol. Both PR3 and proPR3 show a drastic reduction of protein signal in the sample after buffer exchange. The ET samples were obtained from expression trials using HEK293T cells while the large-scale expression was performed using HEK293S GnTI$^-$ cells, therefore accounting for the difference in molecular weight observed. CD177 is a well-expressed and easily purified 50 kDa protein which will be used as a positive control throughout this chapter. Abbreviations are: M, Medium; S_N, supernatant; Filt, filtrate; Sto, storage; B_{Ex}, buffer exchange; FT, flow-through; W_{10}, wash at 10 mM imidazole; W_{40}, wash at 40 mM imidazole; E, elution at 500 mM imidazole.

4.2.1 The purification buffer

As previous results indicated that the method used for buffer exchange (diafiltration or dialysis) does not impact protein recovery, the composition of the purification buffer was next investigated.

Because PR3 was highly expressed and stable in cell culture medium (DMEM supplemented with 2% FBS), but unstable after buffer exchange into PBS, 20 mM Tris pH 7.5 and 150 mM NaCl, this suggested the buffer could lack small molecules, such as ions or detergents, necessary to stabilise PR3. In order to find the missing buffer

component, the composition of cell culture medium was analysed and compared with the physiological constituents of blood plasma, where endogenously expressed PR3 is found. Interestingly, a first analysis revealed the presence of inorganic salts, namely calcium chloride ($CaCl_2$) and magnesium sulfate ($MgSO_4$), in blood plasma as well as in DMEM, but not in the minimal purification buffer used. The importance of inorganic salts was further supported by one previous study which used recombinant PR3 purified by a single step IMAC followed by dialysis in the presence of 2 mM $CaCl_2$ [36]. This prompted a small-scale dialysis trial where cell-free culture media containing secreted PR3 were dialysed into either DMEM supplemented with 2% FBS, DMEM only, DMEM supplemented with 2 mM EDTA to deplete the divalent cations, or diafiltration buffer (negative control). PR3, however, remained stable in all dialysed solutions, including the diafiltration buffer (Figure 4.5A).

This experiment was then scaled-up in roller bottles and samples of culture medium taken at collection, after storage at 4℃ (*i.e.* just before buffer exchange) and after buffer exchange were collected and analysed by western blot (Figure 4.5B). Surprisingly, the large-scale experiment did not replicate the results obtained with small-scale samples. While a protein signal was detected in the culture medium sample taken after collection, no signal could be observed in the following samples, therefore implying the protein was already misbehaving before buffer exchange.

Notably, there seems to be a difference in protein behaviour between small- and large-scale experiments. The reasons behind such a difference remain unknown but will require further investigation.

4.2.2 Protein storage stability

The pI of PR3 with a C-terminal His_6 tag is estimated to be around 8.7. Therefore, in order to first make sure that the protein was not aggregating due the pH of the solution reaching the protein pI, the pH of the culture medium from a large-scale expression was monitored throughout protein harvesting and storage at 4℃. The pH slightly increased from 6.5-7.0 for the culture medium in the incubator at 37℃, to 7.0 for the supernatant collected after centrifugation for 30 min at 4℃, to finally reach a value of 8.0-8.5 after

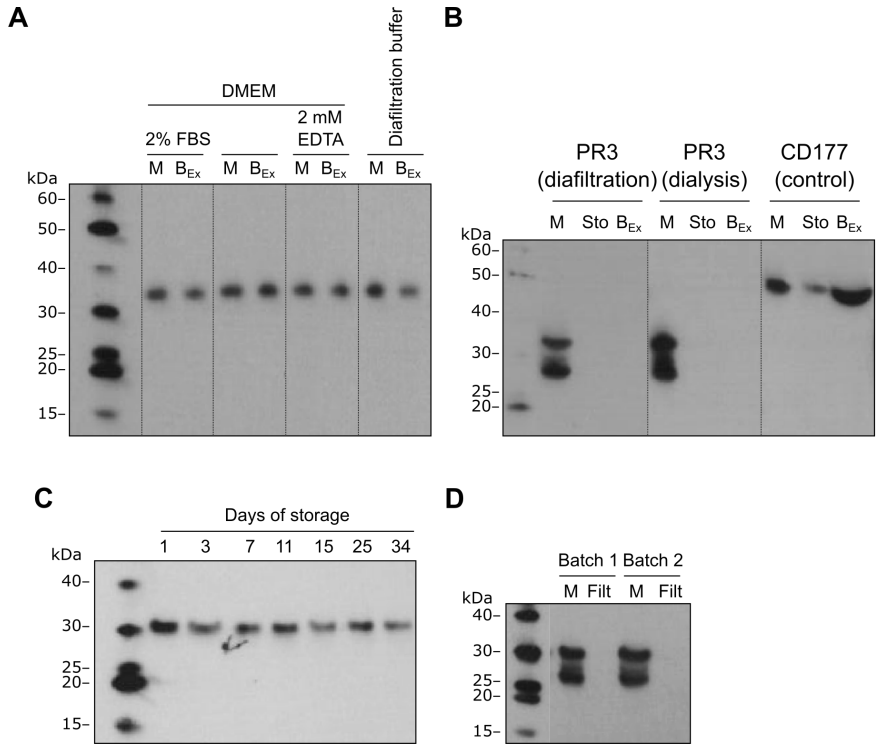


Figure 4.5: PR3 stability tests. PR3 protein signal was analysed in: **(A)** small-scale HEK293T culture medium samples before (M) and after dialysis (B_{Ex}); **(B)** large-scale HEK293T culture medium samples at collection (M), after centrifugation, filtration and storage at 4℃ (Sto) and after buffer exchange (B_{Ex}); **(C)** small-scale HEK293S GnTI$^-$ culture medium samples over 34 days of protein storage 4℃; **(D)** HEK293T culture medium samples at collection (M) and after filtration (Filt) from two different large-scale experiments which used different cells transfected with different batches of plasmid DNA.

days of storage at 4℃. Because the pH of the culture medium during storage appeared to be relatively close to the protein pI, protein stability in DMEM at 4℃ was then monitored over a month (Figure 4.5C). This was done using a sample from small-scale expression and showed that PR3 stayed stable during storage.

Once again, this experiment shows a drastic difference in protein behaviour between small- and large-scale experiments. Where a culture medium sample collected from a large-scale experiment shows no presence of PR3 after storage, a culture medium sample collected from a small-scale experiment shows that PR3 is not only present but also stays stable over a month of storage. Such a difference may indicate that PR3 becomes unstable and aggregates as a result of the procedure used for large-scale cell harvesting. By contrast, PR3 stays stable when the procedure for small-scale cell harvesting is used. Cell harvesting typically consists of three steps including culture medium collection,

centrifugation and filtration. Samples of culture medium from two large-scale productions were subsequently taken at collection and after filtration and analysed for PR3 presence (Figure 4.5D). Strikingly, both samples after filtration showed no presence of PR3 and therefore confirmed that PR3 was unstable during cell harvesting.

4.2.3 Centrifugation of culture medium

The centrifugation procedures in small- and large-scale experiments differ in centrifugal force (15,000 rpm and 7000xg, respectively), temperature (18 and 4℃, respectively) and duration (10 and 30 min, respectively). As a result, both procedures were reproduced on culture medium samples collected from a small-scale experiment (Figure 4.6). Interestingly, all supernatants displayed a similar level of protein signal, independently of the procedure used for centrifugation. In addition, their signal was quite different to the pattern observed for the sample pre-centrifugation. More particularly, this latter displayed three protein signals around ∼30 kDa, which were all stronger than the single protein signal observed for supernatant samples.

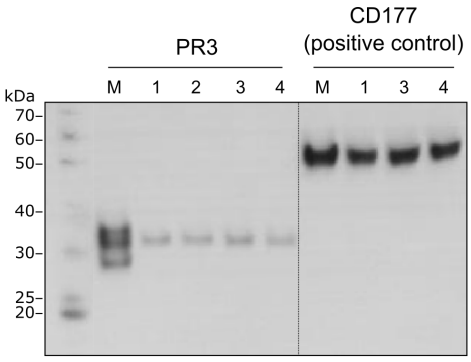


Figure 4.6: PR3 stability during centrifugation of culture medium. Comparison of PR3 protein signal between HEK293T culture medium samples at collection (M) and after centrifugation (lanes 1-4). Each lane corresponds to a different centrifugation procedure. Lane 1 is 7,000xg at 4℃ for 30 min (large-scale procedure), lane 2 is 15,000 rpm at 4℃ for 10 min, lane 3 is 7,000xg at 18℃ for 30 min and lane 4 is 15,000 rpm at 18℃ for 10 min (small-scale procedure). CD177 is shown as a positive control.

The culture medium sample consists of a mixture of culture medium and cell debris, and may therefore contain insoluble PR3 species accounting for the two additional protein signals observed. Nevertheless, the signal accounting for the soluble PR3 species has undoubtedly decreased in all samples collected after centrifugation, which

therefore suggests that PR3 either exists as aggregates in the culture medium or binds to cell membranes. Such a behaviour may also explain the drastic loss of protein observed during cell harvesting.

4.2.4 PR3 recovery from culture medium

4.2.4.1 By washing the cell membranes

PR3 has been previously suggested to behave as a peripheral membrane protein through a so-called hydrophobic patch, in which a group of five positively charged residues are located just above a group of six hydrophobic residues. This arrangement was thought to promote protein insertion into the negatively-charged cell membrane [33, 35], which may result in PR3 interaction with HEK293 cell surface and subsequently prevent the protein from being soluble in culture medium.

In order to investigate whether PR3 actually binds to the cell surface, preliminary cell-surface staining experiments were performed, during which HEK293T cells were transfected with either PR3 or full-length CD177 as positive control. The cells were then analysed for surface expression by using fluorescence microscopy (Figure 4.7A). Interestingly, cells transfected with PR3 appeared to display some fluorescent signal, suggesting possible PR3 binding to cell membranes.

To validate this observation, adherent cells obtained from a small-scale experiment were washed with either PBS, 1 M NaCl or 15% glycerol in order to recover putative membrane-bound PR3 (Figure 4.7B). Unfortunately, although washing with 15% glycerol seemed to recover a tiny fraction of membrane-bound PR3, the corresponding protein signal was insignificant compared to the one observed in the culture medium supernatant. Therefore, this indicated that PR3 interaction with cell membranes is not likely to be responsible for the loss of protein observed upon centrifugation.

4.2.4.2 By direct purification from the crude culture medium

In order to circumvent PR3 loss during centrifugation, this step was removed during purification. Culture medium obtained from a small-scale experiment was split into three fractions before each of them was subjected to either centrifugation followed by

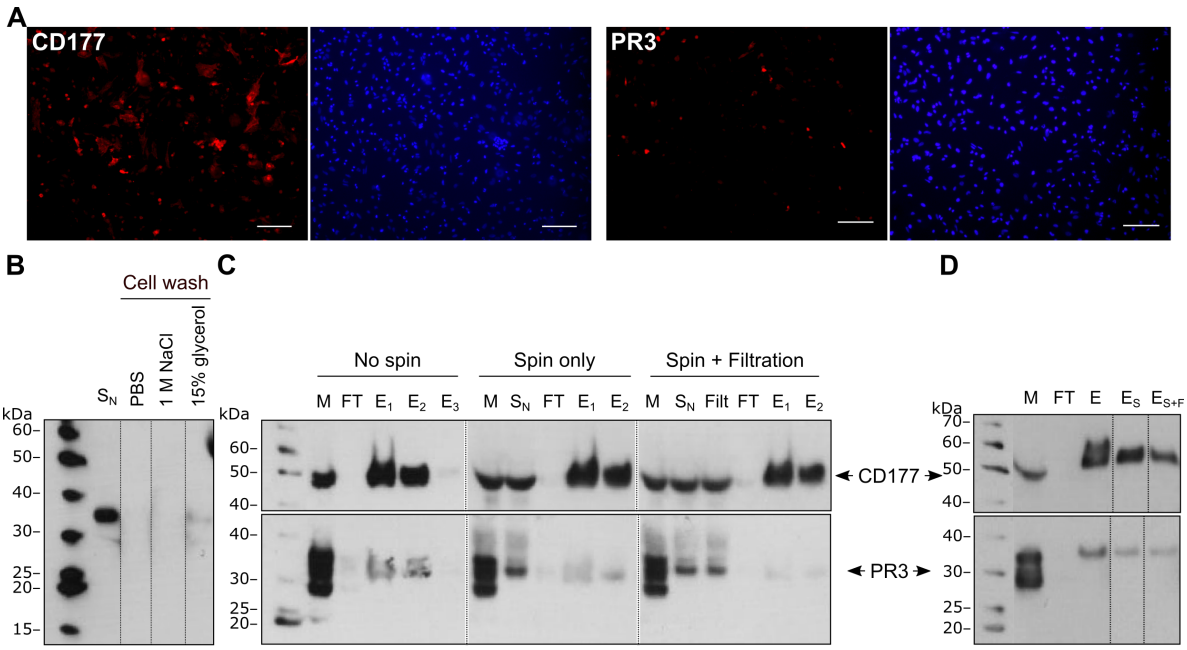


Figure 4.7: PR3 recovery from washing cell membranes. (A) Full-length CD177 (left panel) and PR3 (right panel) constructs were expressed in HeLa cells with an N-terminal FLAG tag. FLAG-tagged proteins were detected by cell-surface immunostaining (red). DAPI (blue) highlights cell nuclei. This preliminary cell-surface staining assay shows that full-length CD177 is well-expressed at the cell surface while PR3 displays some fluorescent signal. Scale bar is 100 µm. **(B)** PR3 presence was analysed in HEK293T culture medium supernatant (S_N) and in different wash buffers. **(C)** Analysis of PR3 and CD177 presence in samples collected from culture medium at collection (M), after centrifugation (S_N), after filtration (Filt) and from IMAC flow-through (FT), first elution (E_1), second elution (E_2) and third elution (E_3). White lines separate lanes extracted from different western blots. **(D)** Summarising comparison of PR3 and CD177 presence in samples collected from culture medium at collection (M), IMAC flow-through (FT) and IMAC elution (E, E_S and E_{S+F}). E, E_S and E_{S+F} correspond to the "no spin", "spin only" and "spin+filtration" samples, respectively.

filtration, centrifugation only or no centrifugation at all. All fractions then underwent IMAC using loose beads (Figure 4.7C). Interestingly, there seemed to be slightly more PR3 material eluted, and therefore recovered, from the sample which did not undergo centrifugation compared to the ones which underwent centrifugation (Figure 4.7C-D). Unfortunately, this increase remained insignificant in comparison with the amount of PR3 initially present in the culture medium at collection.

Most surprisingly however, was the very little amount recovered in all elution fractions by comparison with the starting material present in the culture medium (Figure 4.7D). This indicated that the optimised purification protocol for PR3 should not only provide a means to increase protein recovery after centrifugation, but should also take into account protein instability in subsequent purification steps.

4.2.4.3 By expression in CD177-conditioned medium

Given the previous results showing little PR3 recovery from cell membranes, the presence of PR3 aggregates was next investigated. In particular, the existence of a hydrophobic patch containing six exposed hydrophobic residues on the surface of PR3 would likely drive the formation of such aggregates.

CD177 is a GPI-anchored glycoprotein and is the receptor which presents PR3 on the surface of neutrophils. Notably, CD177 has been suggested to bind PR3 on its hydrophobic patch [36], and therefore appears as an ideal candidate to stabilise PR3 in solution. Indeed, the receptor's ability to cover the exposed hydrophobic residues on PR3 may result in the production of a well-behaving PR3.

In order to investigate the stabilising effect of CD177 on PR3, a small-scale expression trial was first performed, during which two different methods of expressing PR3 in presence of CD177 were assessed. PR3 was either co-expressed with $CD177^{ecto}$ or expressed in $CD177^{ecto}$-conditioned medium (CD177-CM) before the protein stability during centrifugation was tested (Figure 4.8A).

Interestingly, expressing in a CD177-CM resulted in a much higher protein production than co-expressing with $CD177^{ecto}$. In addition, although no improvement could be observed in the samples after centrifugation, the presence of CD177 seemed to improve PR3 behaviour during the subsequent IMAC purification. Indeed, expression in CD177-CM resulted in a dramatic decrease of PR3 aggregation on Nickel beads, indicating that PR3 does behave differently when in complex with CD177 (Figure 4.8B).

Expression of PR3 in CD177-CM was scaled-up using HEK293S $GnTI^-$ cells, and resulted in successful purification of 700 µg of PR3-$CD177^{ecto}$ complex, and of 900 µg of an excess of $CD177^{ecto}$ from 3 L of culture (Figure 4.8C,D). However, western blot analysis showed that, despite the overall better protein behaviour during purification, PR3 still misbehaves during centrifugation (Figure 4.8E-F).

Nevertheless, this hinted that protein instability during centrifugation and during purification are two separate issues. While the element responsible for the former remains unidentified, the latter seems to improve upon covering the protein hydrophobic patch in solution.

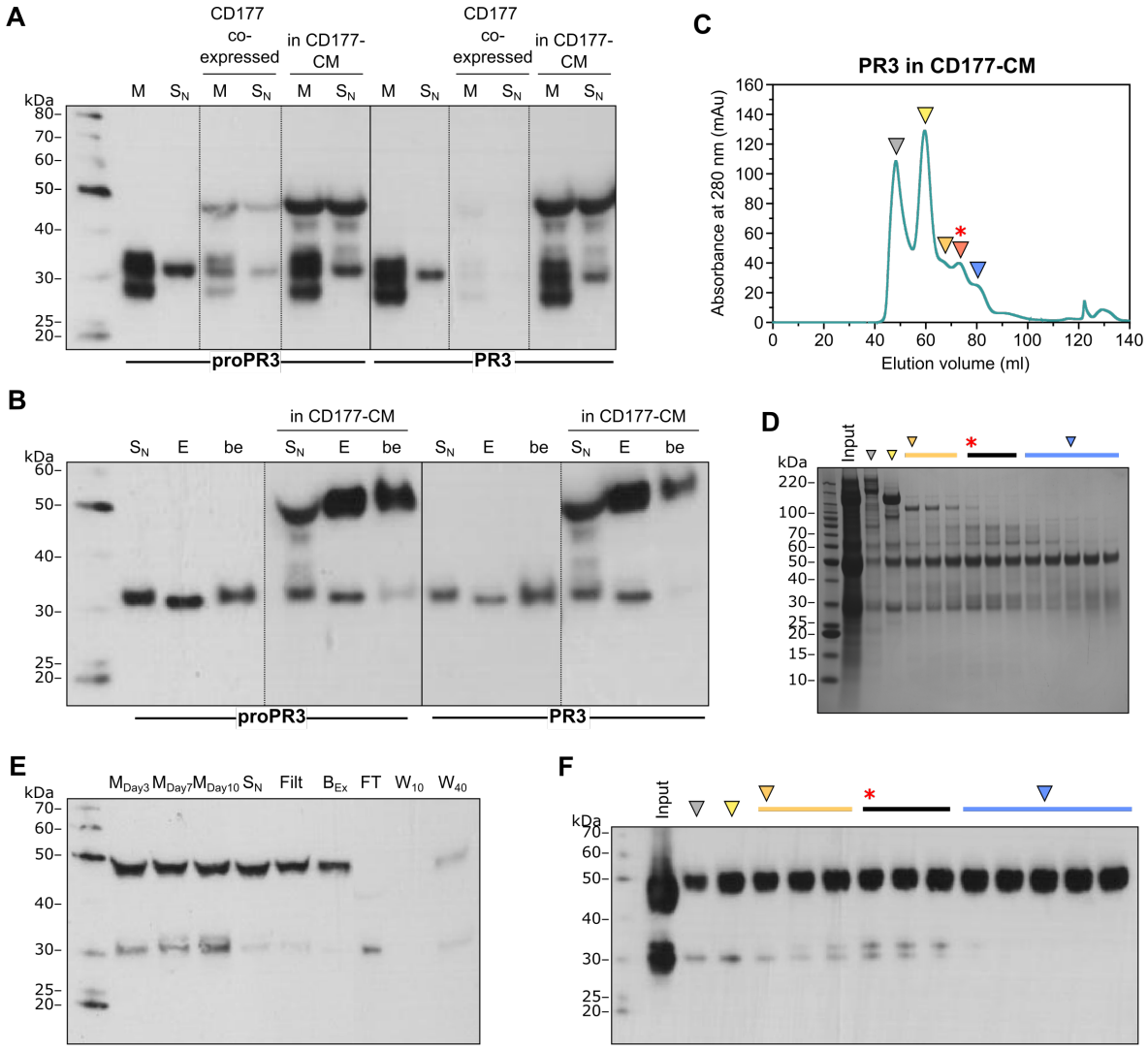


Figure 4.8: PR3 expression in CD177-conditioned medium. **(A)** Analysis of proPR3 and PR3 presence in supernatant samples recovered from PR3 expression alone, from co-expression with CD177 and in CD177-CM. **(B)** Analysis of proPR3 and PR3 presence in samples collected after centrifugation (S_N), IMAC elution (E) and on Nickel beads (be). **(C,D)** The SEC purification profile from a large-scale purification of PR3 in CD177-CM shows PR3-CD177 elutes around 75 ml on a Superdex® 200 column. The complex peak (red asterisk) is followed by an excess of CD177 (blue arrowhead), which elutes around 80 ml. The corresponding SDS-PAGE gel shows the SEC input and fractions from the void, impurity, PR3-CD177 complex and excess CD177 protein peaks. **(E,F)** Western blot analysis tracking the presence of PR3 and CD177 from protein expression in roller bottles to the different purification steps including IMAC and SEC.

4.2.4.4 By using additives before centrifugation

Previous studies have suggested a role for PR3 in the clearance of apoptotic cells, and in particular, an interaction of PR3 with phosphatidylserine lipids on the surface of apoptotic cells [58, 59, 62]. This may explain why the fraction of PR3 lost during centrifugation could not be recovered by washing the adherent cells as the protein

would instead be interacting with the cell debris. Conversely, PR3 may also exists as oligomers and/or aggregates in the culture medium.

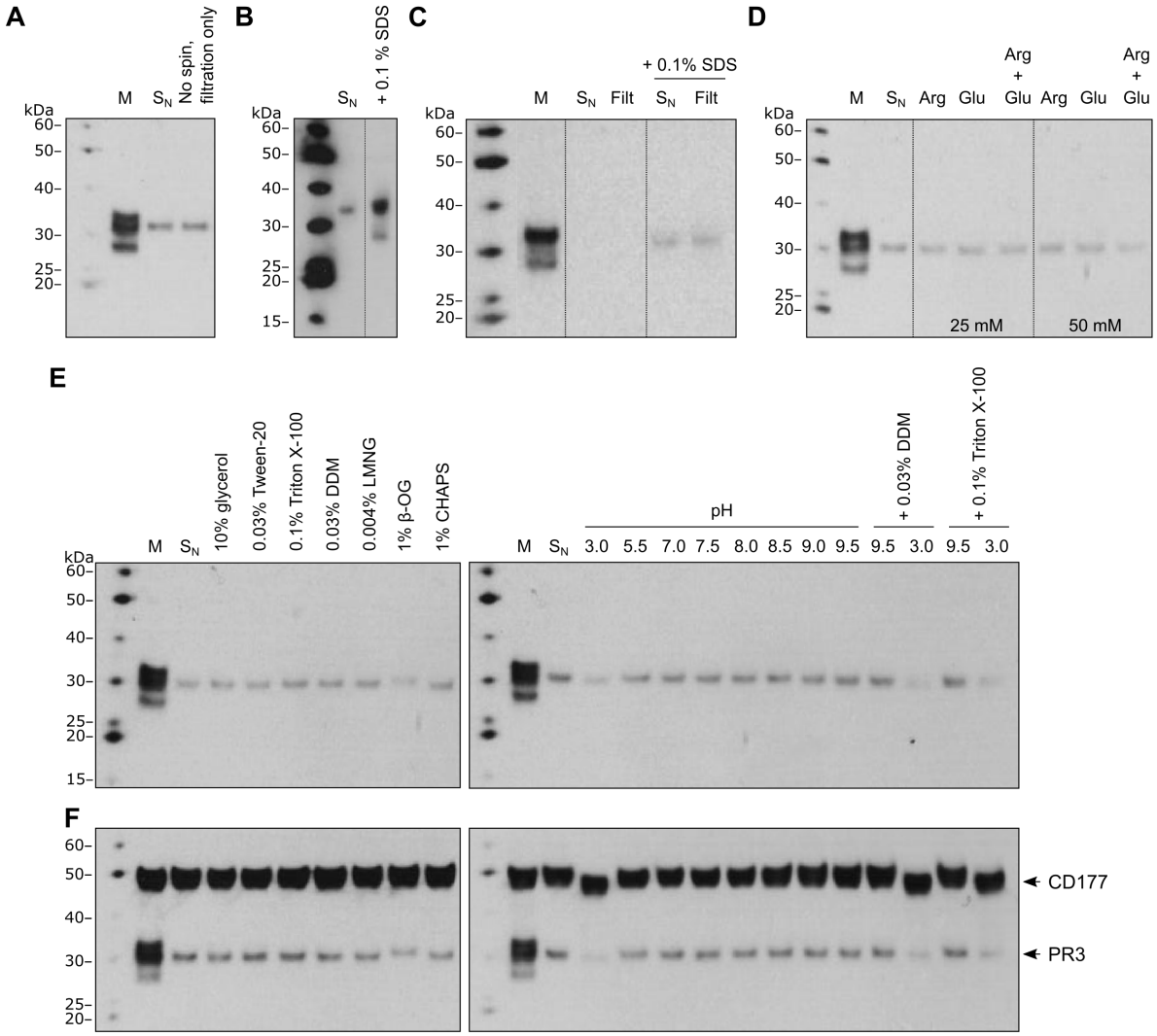


Figure 4.9: Use of additives to improve PR3 solubilisation. Western blot analyses of PR3 presence in samples collected at collection (M), after centrifugation (S_N) and filtration (Filt). **(A)** PR3 presence in a sample collected after centrifugation (S_N) is similar to that of a filtered sample. **(B,C)** Addition of 0.1% SDS in the culture medium at collection of **(B)** a small-scale and **(C)** large-scale experiment results in an increased presence of PR3 in S_N samples. **(D)** Addition of the solubilising agents L-arginine and L-glutamate does not increase the presence of PR3 in S_N samples. **(E,F)** Addition of detergents at a concentration of two to three times their critical micellar concentration, and/or changing the pH does not increase the presence of PR3 in S_N samples, whether PR3 is expressed **(E)** alone or **(F)** in CD177-CM.

Interestingly, filtering the culture medium with a 0.22 µm filter appeared to have a similar effect as centrifuging it, suggesting that PR3 might either interact with cell debris or form aggregates larger than 0.22 µm (Figure 4.9A). Moreover, preliminary experiments in which culture medium samples from small- and large-scale experiments

were first incubated with 0.1% SDS have shown some protein recovery after centrifugation (Figure 4.9B,C), therefore suggesting PR3 could be stabilised using additives. This prompted a more extensive screening where various additives known for their solubilisation properties were incubated with culture medium prior to centrifugation. The effect of pH as well as combinations of both pH and additive were also tested (Figure 4.9D,E). Unfortunately, none of the condition resulted in a better recovery. Repeating the experiments using PR3 expressed in a CD177-CM also failed to improve the results (Figure 4.9F).

Intriguingly, a recent study has reported a successful PR3 purification, during which 0.02% DDM was used to stabilise the protein [124]. In an attempt to reproduce this purification, 0.03% DDM was added to culture medium before centrifugation and kept throughout buffer exchange by dialysis and IMAC, before SEC in 0.01% DDM (Figure 4.10A,B). However, this not only failed to increase protein recovery after centrifugation, but also resulted in purification of a histidine-rich contaminant of ∼60 kDa instead of PR3 (Figure 4.10C,D).

4.2.4.5 Reproducibility between small- and large-scale experiments

Overall, challenges during optimisation of PR3 purification mostly came from variable results obtained after centrifugation, especially for samples collected in small-scale experiments. In general, using HEK293T cells over HEK293S GnTI$^-$ exacerbated the loss of PR3 during centrifugation. In addition, protein levels obtained after centrifugation varied between different experiments. Indeed, PR3 signal in supernatant would sometimes be comparable to signal of the ladder, be much lower or completely disappear (Figures 4.6, 4.5 and 4.4B, respectively). Most strikingly however, was the difference observed between samples produced in small-scale and those produced in large-scale experiments. While loss of protein after centrifugation was not always observed when using sample from small-scale experiments, it always occurred in large-scale protein productions. This suggested that differences between small- and large-scale expression procedures may be responsible for protein loss upon centrifugation. Possible factors include the density

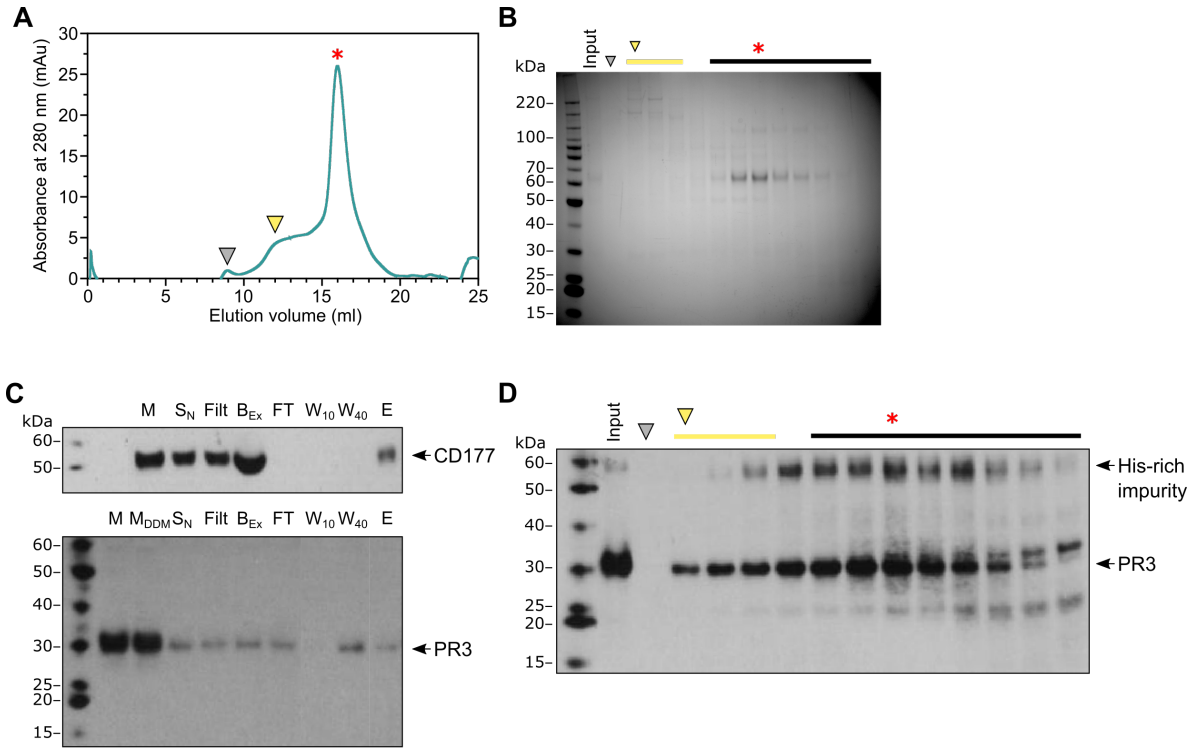


Figure 4.10: Large-scale purification of PR3 in DDM. (A,B) SEC purification profile and corresponding SDS-PAGE gel of a purification in 0.01% DDM. A 60 kDa-protein peak (red asterisk) is preceded by a void and impurity peak (grey and yellow arrowheads, respectively). **(C,D)** Western blot analysis tracking the presence of PR3 from culture medium at collection (M) to the different purification steps including IMAC and SEC. Presence of CD177 throughout those steps is used as comparison and shows the expected protein behaviour.

of cells seeded before transfection, the number of cells at the time of transfection, the transfection reagent used, the duration of expression, etc.

4.3 PR3 stability during protein expression

4.3.1 Monitoring PR3 expression in the culture medium

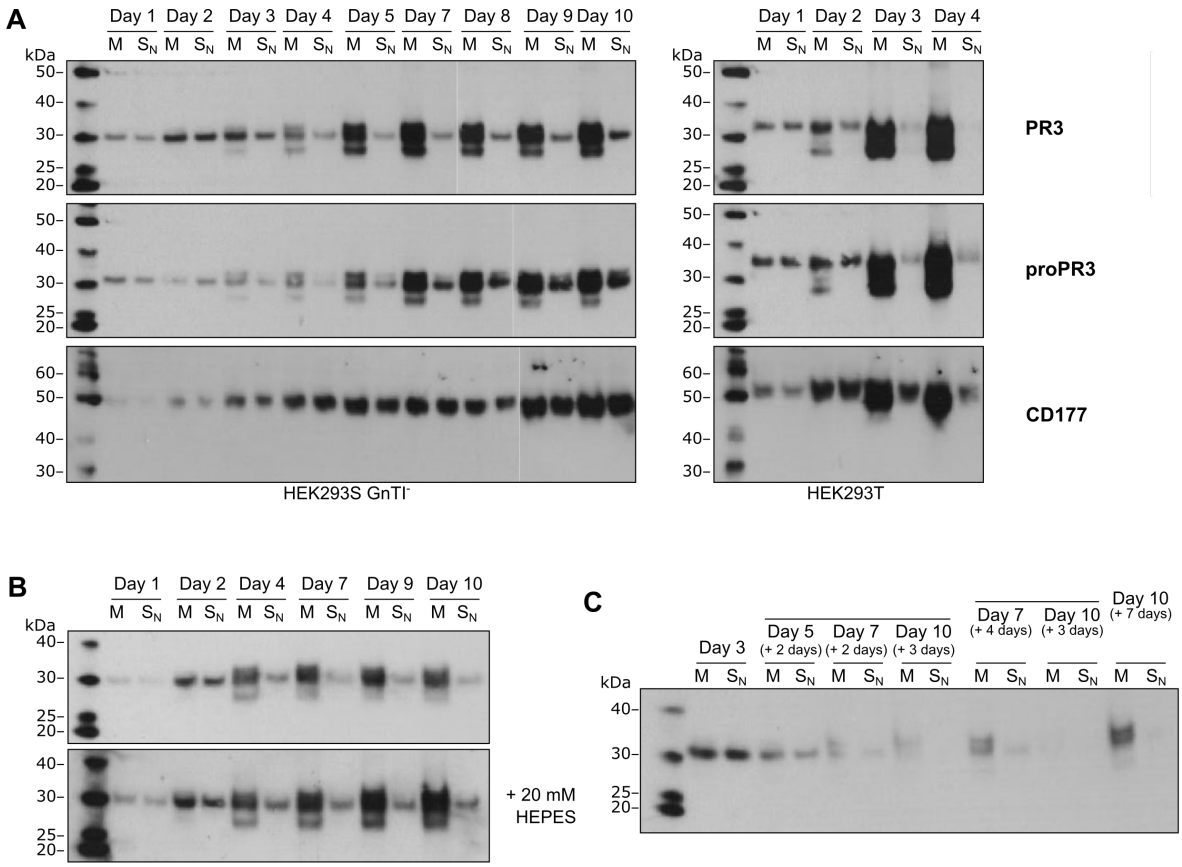


Figure 4.11: Evolution of PR3 solubility during protein expression. Analysis and comparison of PR3 presence in samples of culture medium at collection (M) and after centrifugation (S_N) taken at different days during protein expression. CD177 is shown as positive control. **(A)** Comparison of PR3 solubility over time when expressed by HEK293S GnTI$^-$ or HEK293T cells. White lines separate lanes extracted from different western blots. **(B)** Comparison of PR3 solubility in a culture medium with and without 20 mM HEPES. **(C)** Evolution of PR3 solubility upon addition of fresh DMEM supplemented with 2% FBS after an initial harvest at day three post-transfection. The fresh medium was left for various amount of time before being harvested and replaced once again in order to continue PR3 expression until day ten post-transfection. Fresh medium was either added once (harvests at day three and ten), twice (harvests at day three, seven and ten) or three times (harvests at day three, five, seven and ten).

In large-scale experiments, samples of PR3, proPR3 and CD177^{ecto} were collected throughout the protein expression phase and subjected to centrifugation in order to assess loss of PR3 over time. Strikingly, although the quantity of PR3 expressed in culture medium increased over time, the quantity of PR3 recovered after centrifugation seemed to decrease. This was particularly noticeable when proteins were expressed using

HEK293T cells (Figure 4.11A), which suggests that accumulation of secreted protein in culture medium, prolonged contact with increasing number of dead cells or changes in biochemical properties of the culture medium could be responsible for loss of soluble PR3. To investigate whether the decrease of pH happening during protein expression was responsible for this observation, the physiological pH of cell culture medium was maintained by addition of 20 mM HEPES (Figure 4.11B). As this did not improve protein recovery during the last days of expression, this not only further confirmed the protein instability upon prolonged contact with the cell culture medium, but also ruled out pH variation as the responsible factor.

Notably, western blot analysis suggested that an earlier harvest at day three post-transfection was necessary in order to avoid protein loss upon centrifugation (Figure 4.11A-B). To increase protein yield, fresh DMEM supplemented with 2% FBS was added to roller bottles after an initial harvest at day three post-transfection. This was left for two, four or seven more days of expression for a final of five, seven or ten days of expression, respectively. Unfortunately, although a second harvest after two additional days of expression showed almost no loss of protein upon centrifugation, harvesting later than that would be ineffective and would not increase protein yield (Figure 4.11C).

4.3.2 Large-scale purification trials of PR3

Purification of PR3 expressed by HEK293S GnTI$^-$, during which the culture medium was harvested after three days of expression instead of the usual ten days, yielded mixed results (Figure 4.12A,B). On one hand, PR3 was purified in sufficient quantities for unambiguous detection by SDS-PAGE, which was not the case for all previous purifications. On the other hand, PR3 appeared to be sticky and unable to elute as a single protein peak by SEC. Instead, PR3 was present in all SEC fractions (Figure 4.12B).

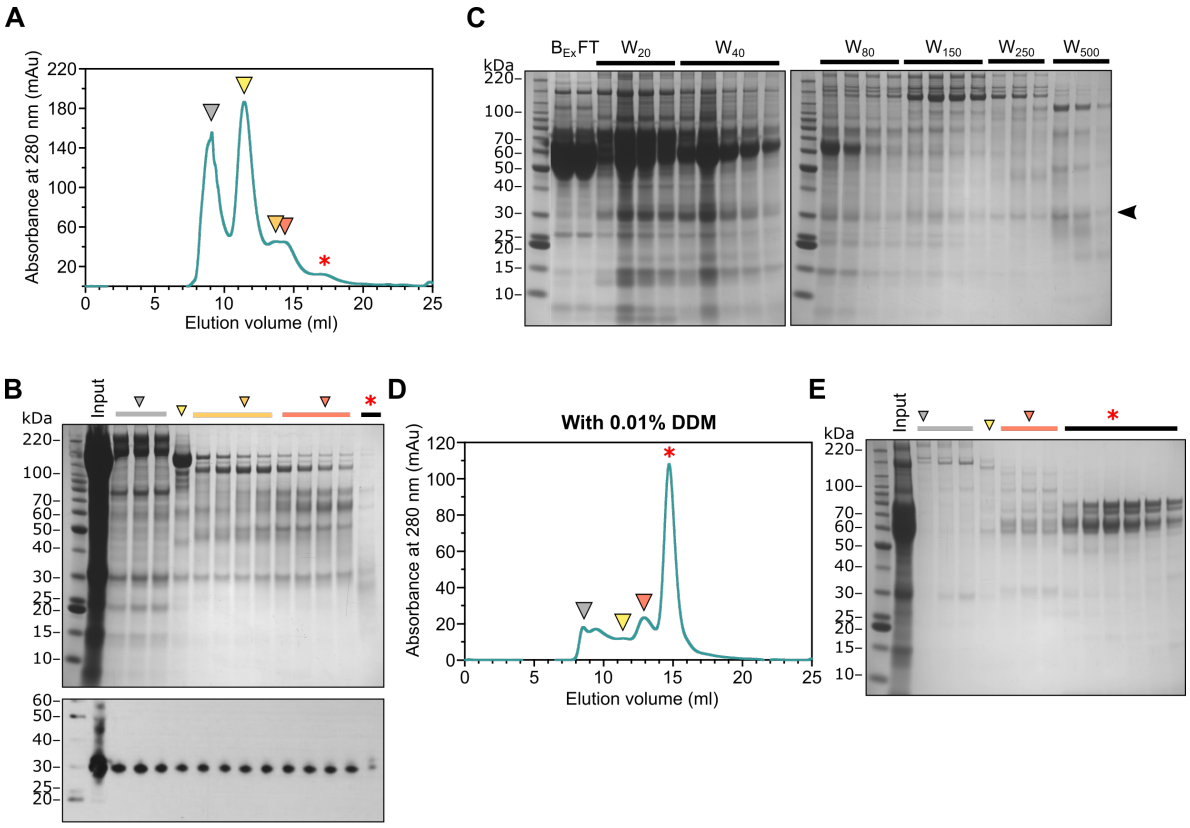


Figure 4.12: Large-scale purification of PR3 from early harvest. (A,B) The SEC purification profile shows PR3 elutes around 17 ml on a Superdex® 200 10/300 GL column. The protein peak (red asterisk) is preceded by a void and impurity peaks (grey, yellow, orange and red arrowheads, respectively). The corresponding SDS-PAGE gel and western blot shows the SEC input and fractions from the void, impurity and protein peaks. PR3 is present in all SEC fractions. **(C)** SDS-PAGE gel of samples taken during IMAC step gradient elution. **(D,E)** SEC purification profile and corresponding SDS-PAGE gel of a purification in 0.01% DDM. A 60 kDa-protein peak (red asterisk) is preceded by a void and impurity peaks (grey, yellow and red arrowheads, respectively).

A step gradient elution was then introduced during IMAC, with the hope of recovering purer PR3-containing fractions before SEC (Figure 4.12C). This, however, was unsuccessful as PR3 was already eluting along with impurities during washes at 20 mM and 40 mM imidazole.

Purification in 0.03% DDM was also repeated but yielded the same previous result where a histidine-rich impurity was purified instead of PR3 (Figures 4.12D,E and 4.10).

4.3.3 Large-scale purification trials of PR3-CD177^{ecto}

Given that co-purification with CD177 appeared to improve PR3 behaviour during purification, whether this would also decrease the number of PR3 non-specific binding

was next investigated.

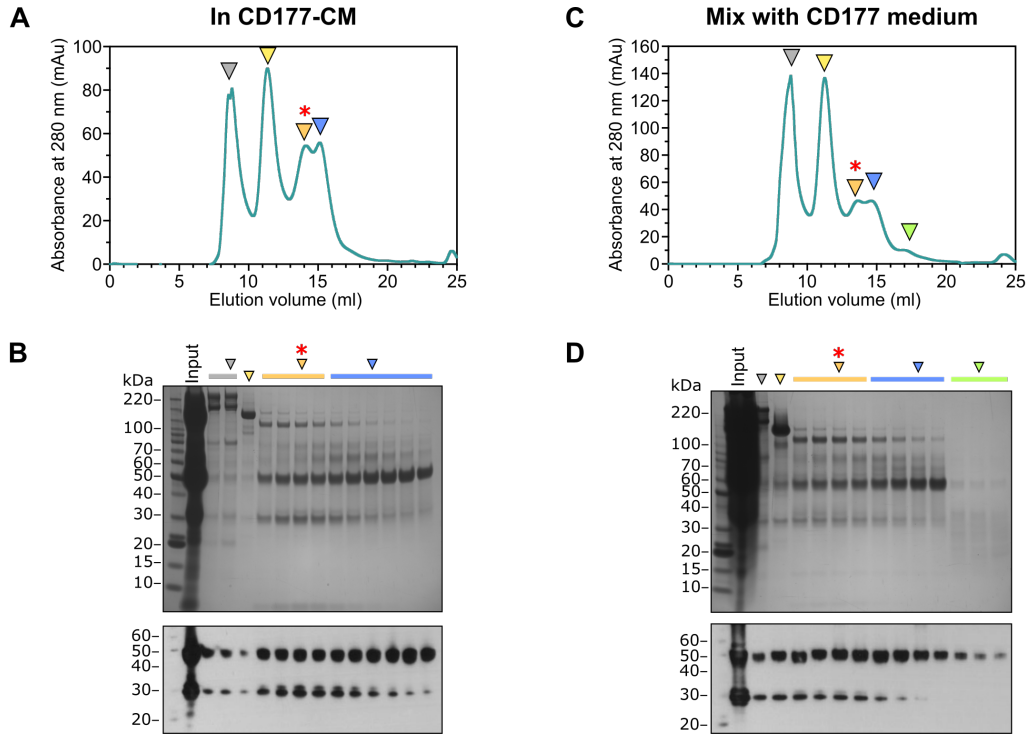


Figure 4.13: Large-scale purification of PR3 from early harvest in presence of CD177. **(A,B)** SEC purification profile and corresponding SDS-PAGE gel and western blot of a purification in CD177-CM. The complex (red asterisk) elutes along with a 100 kDa impurity (orange arrowhead) and is followed by elution of excess CD177 (blue arrowhead). **(C,D)** SEC purification profile and corresponding SDS-PAGE gel and western blot of a purification after incubation with CD177-containing culture medium. The complex (red asterisk) elutes along with a 100 kDa impurity (orange arrowhead) and is followed by elution of excess CD177 (blue arrowhead) and of a small protein (green arrowhead).

To do this, PR3 was either expressed in CD177-CM (Figure 4.13A,B), or in standard culture medium that was then mixed with half the volume of a culture medium containing CD177^{ecto} (Figure 4.13C,D). In both cases, the culture media containing PR3 were harvested after three days of expression. Overall, both methods allowed purification of PR3-CD177^{ecto} complex, albeit with a major 100 kDa impurity. Most importantly, the majority of PR3 elutes as a complex with CD177, in fractions significantly purer than those from PR3 purifications on its own (Figure 4.12A,B).

4.4 PR3 surface mutants

Although expression of PR3 in CD177-CM allowed co-purification of PR3, the aim of this work was to isolate PR3 alone. However, optimisation of the purification protocol of PR3

failed, and PR3 showed instead strong non-specific binding and potential aggregation upon prolonged contact with HEK293 culture medium. Alternatively, another way to obtain and isolate PR3 is to alter its surface properties through mutation of hydrophobic residues present on the surface of the protein to hydrophilic residues. This may improve protein solubility and decrease non-specific binding.

4.4.1 Design of PR3 surface mutants

An analysis of the only crystal structure available revealed that the crystal ASU is composed of four copies of PR3 arranged in a peculiar tetramer conformation, in which all copies have their hydrophobic patch facing inwards, and therefore hidden from the environment (Figure 4.14A). Whether this tetramer actually exists in solution or not is unknown. However, the presence of this hydrophobic patch on the surface of PR3 might be responsible for protein stickiness and most of the non-specific interactions observed.

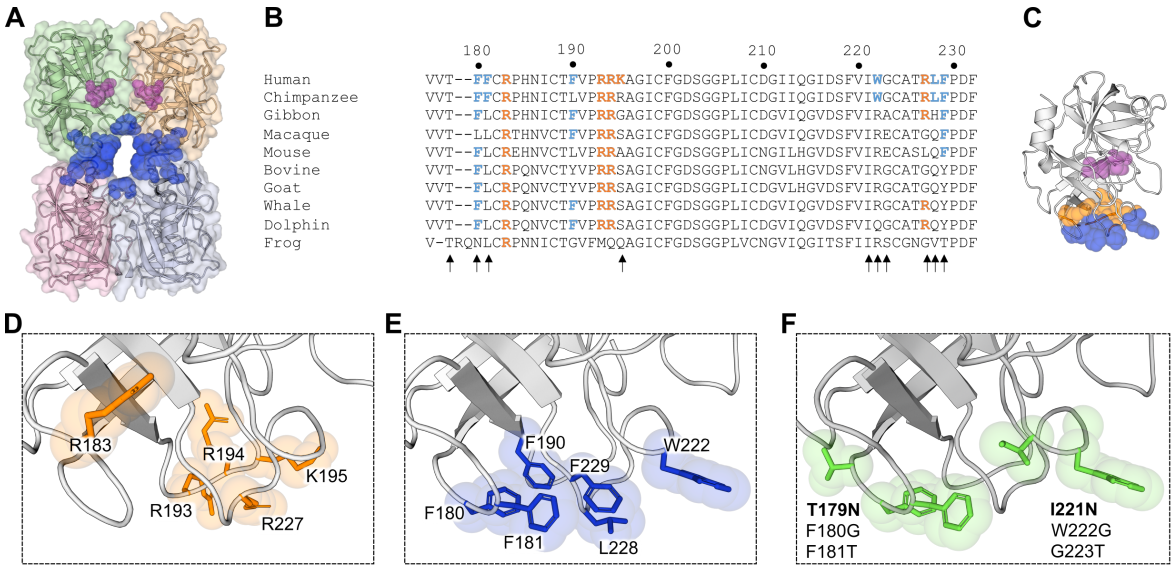


Figure 4.14: Design of PR3 surface mutants. **(A)** Cartoon representation of the tetramer arrangement of PR3 in the crystal ASU (PDB ID: 1FUJ). Each colour corresponds to one protomer. The catalytic triad and hydrophobic residues from the hydrophobic patch are indicated by purple and blue spheres, respectively. **(B)** Sequence alignment of different PR3 orthologs. Hydrophobic and positively-charged residues from the hydrophobic patch are highlighted in bold blue and orange, respectively. Residues mutated in PR3 variants are indicated by a black arrow. **(C)** Cartoon representation of PR3. The catalytic triad, hydrophobic and positively-charged residues from the hydrophobic patch are indicated by purple, blue and orange spheres, respectively. **(D,E)** Close-up on **(D)** positively-charged and **(E)** hydrophobic residues from the hydrophobic patch. **(F)** Close-up on the three consecutive residues mutated to the NGT sequon for addition of an artificial glycosylation site.

The hydrophobic patch consists of the six hydrophobic residues F180, F181, F190, L228, F229, W222, which are located below the five positively-charged residues R183, R193, R194, K195 and R227 (Figure 4.14B,C). A multiple sequence alignment of human PR3 with PR3 from other species revealed that most of those residues are not highly conserved and could therefore be mutated without the risk of altering protein folding. Despite a similar three-dimensional structure and a high sequence identity and similarity (66.8% and 78.6%, respectively), the physico-chemical properties of human and mouse PR3 have been found to be significantly different [332]. This inspired the design of PR3 variants, which consisted of mutations of the least conserved residues F181, K195, W222, R227 and L228 to their mouse counterparts (Figure 4.14D,E).

Later PR3 variants were less conservative and consisted of mutations of highly conserved hydrophobic residues such as F180 and F229 to hydrophilic and polar residues. For other PR3 variants, the three consecutive residues T179-F180-F181 and/or I221-W222-G223 were replaced with the NGT sequon in order to insert artificial glycosylation sites to improve protein solubility (Figure 4.14F). All PR3 variants had the S203A mutation and were all therefore catalytically inactive.

4.4.2 Mutant screening

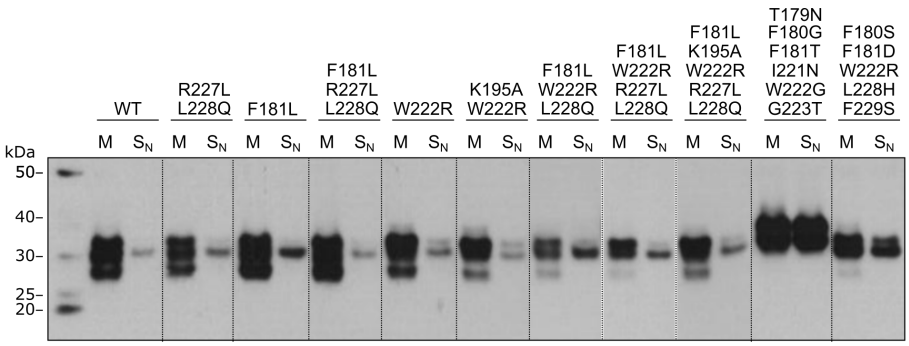


Figure 4.15: Centrifugation trial of PR3 variants. Analysis of PR3 presence in pre- (M) and post-centrifugation (S_N) samples from different PR3 variants. White lines separate lanes extracted from different western blots.

In initial centrifugation trials using samples from small-scale production, most of the initial 'mouse mutants' failed to drastically improve protein recovery after centrifugation (Figure 4.15). By contrast, more drastic mutants such as the 'double glycosylation

mutant' PR3$^{\text{TFF-NGT IWG-NGT}}$ containing the six mutations T179N, F180G, F181T, I221N, W222G, G223T or the 'synthetic mutant' PR3$^{\text{F180S F181D W222R L228H F229S}}$ containing the five mutations F180S, F181D, W222R, L228H and F229S showed high to total protein recovery in samples after centrifugation.

4.4.3 Large-scale purification of PR3 mouse mutants

In order to assess the efficacy of mutations in improving protein solubility, three mouse variants were selected and produced in large-scale. The variants had either two (K195, W222R), three (F181L, W222R, L228Q) or five (F181L, K195A, W222R, R227L, L228Q) mutations.

All three variants were first expressed and harvested after full-time expression, *i.e.* four days if using HEK293T cells or ten days if using HEK293S GnTI^{-} cells. Overall, the variants displayed a better behaviour than the wild type (WT), with more protein recovered in IMAC elution and a separate protein elution peak on the SEC profile that could be related to the protein signal observed on SDS-PAGE gels and western blots (Figures 4.3 and 4.16). However, the yields were limited to 70-130 µg/L of relatively impure protein, which remained unsuitable for crystallographic analyses.

The variants were once again expressed by HEK293S GnTI^{-}, but this time, harvested after only three days of expression. This early harvest resulted in a remarkable improvement of sample purity, with a final yield of 30-70 µg/L of pure protein (Figures 4.16).

In the end, the combination of conservative mutations with an early harvest revealed to be a promising way to obtain pure and homogeneous PR3 variants, albeit still in sub-optimal quantities for structural studies.

4.4.4 Large-scale purification of PR3 double glycosylation mutant

In addition, the double glycosylation mutant PR3$^{\text{TFF-NGT IWG-NGT}}$ was expressed and harvested after full-time expression. Remarkably, this variant appeared to not only be well-behaved, but also extremely well-expressed with a spectacular increase in yield, reaching 6 mg per L of culture (Figure 4.17A).

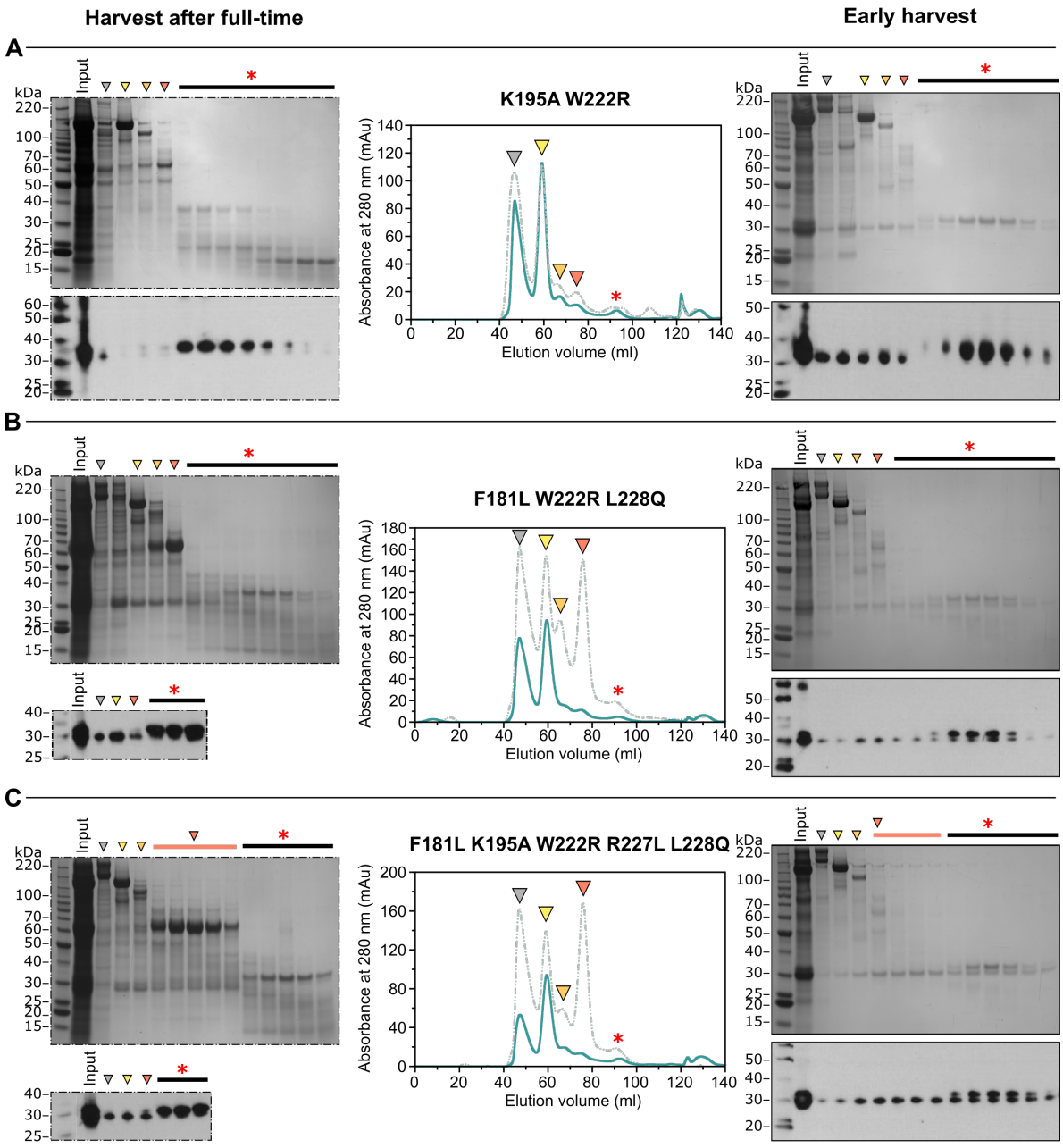


Figure 4.16: Large-scale purification of PR3 mouse mutants. The SEC purification profiles show the PR3 variants **(A)** $PR3^{K195\ W222R}$, **(B)** $PR3^{F181L\ W222R\ L228Q}$ and **(C)** $PR3^{F181L\ K195A\ W222R\ R227L\ L228Q}$ all elute around 93 ml on a Superdex® 200 column. The protein peak (red asterisk) is preceded by a void and impurity peaks (grey, yellow, orange and red, respectively). The purification profiles of proteins produced by cells harvested after full-time expression and after early harvest are depicted in dotted and full lines, respectively, along with their corresponding SDS-PAGE gels and western blots.

Intriguingly, $PR3^{TFF\text{-}NGT\ IWG\text{-}NGT}$ seemed to be produced as a mixture of three main species, whose molecular weights shifted whether the protein was expressed by HEK293T or HEK293S GnTI$^-$ cells (Figure 4.17A,B). In order to confirm whether those species

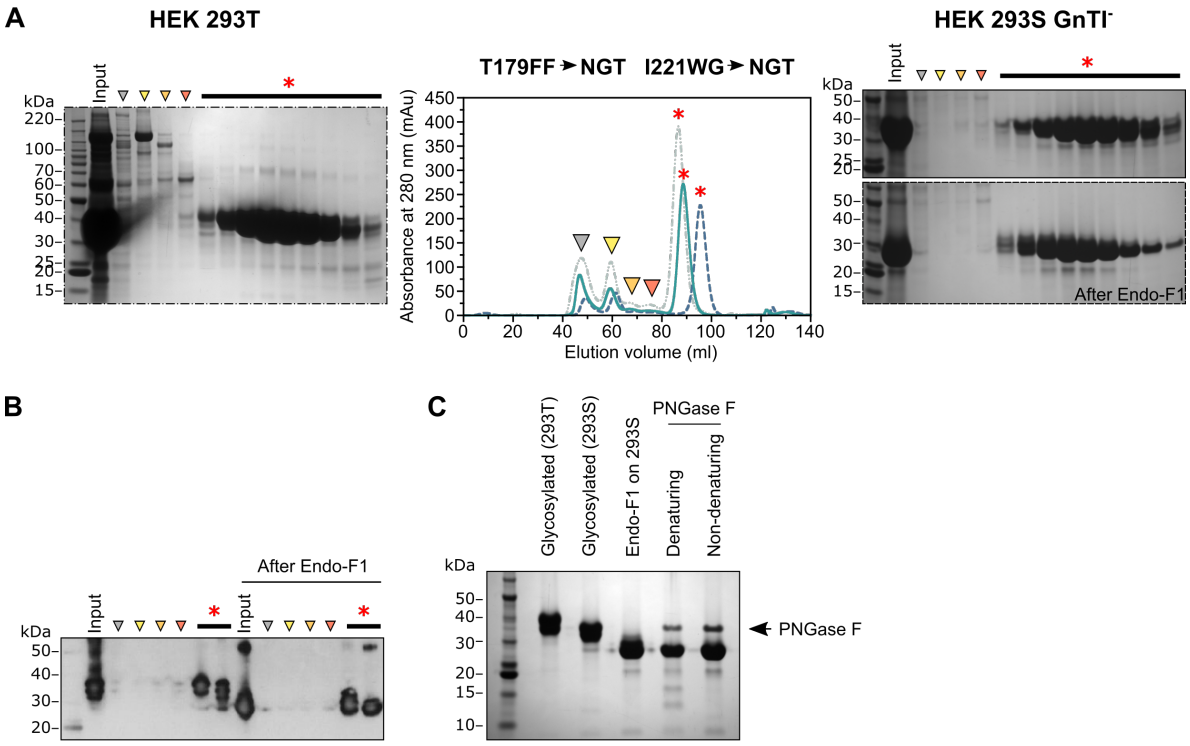


Figure 4.17: Large-scale purification of PR3 double glycosylation mutant. **(A)** SEC purification profiles with their corresponding SDS-PAGE gel and western blot of PR3$^{\text{TFF-NGT IWG-NGT}}$ expressed by HEK293T cells (dotted lines), HEK293S GnTI$^-$ cells with Endo-F_1 treatment (dashed lines) and without (full lines). The protein peak (red asterisk) is preceded by a void and impurity peaks (grey, yellow, orange and red arrowheads, respectively) and shows later elution when deglycosylated. **(B)** Western blot comparing SEC fractions of PR3$^{\text{TFF-NGT IWG-NGT}}$ expressed by HEK293S GnTI$^-$ cells with and without Endo-F_1 treatment. **(C)** SDS-PAGE gel of samples obtained after deglycosylation by Endo F_1 or PNgase F.

were related glycosylation species, HEK293S GnTI$^-$-produced PR3$^{\text{TFF-NGT IWG-NGT}}$ was incubated with either Endo F_1 or PNGase F (Figure 4.17B,C). While incubation with Endo F_1 did not result in complete protein deglycosylation, incubation with PNGase F in denaturing conditions resulted in a single PR3$^{\text{TFF-NGT IWG-NGT}}$ species, therefore confirming that the several species observed are differently glycosylated species, and not degradation products.

4.4.5 Interaction of PR3 variants with CD177

As CD177 has previously been suggested to bind to the hydrophobic patch [36], it was necessary test whether the PR3 variants retained their ability to interact with CD177 to make sure the purified variants were functionally relevant.

4.4.5.1 PR3 pull-downs

A small-scale PR3 pull-down experiment was set up, where $CD177^{ecto}$-coupled sepharose beads were incubated with culture medium containing PR3 variants. Western blot analysis revealed that all the mouse mutants retained their ability to interact with CD177 (Figure 4.18). However, the best expressing mutant $PR3^{TFF\text{-}NGT\ IWG\text{-}NGT}$ did not interact with CD177 anymore. Interestingly, the synthetic mutant $PR3^{F180S\ F181D\ W222R\ L228H\ F229S}$ appeared to still bind CD177, albeit more weakly.

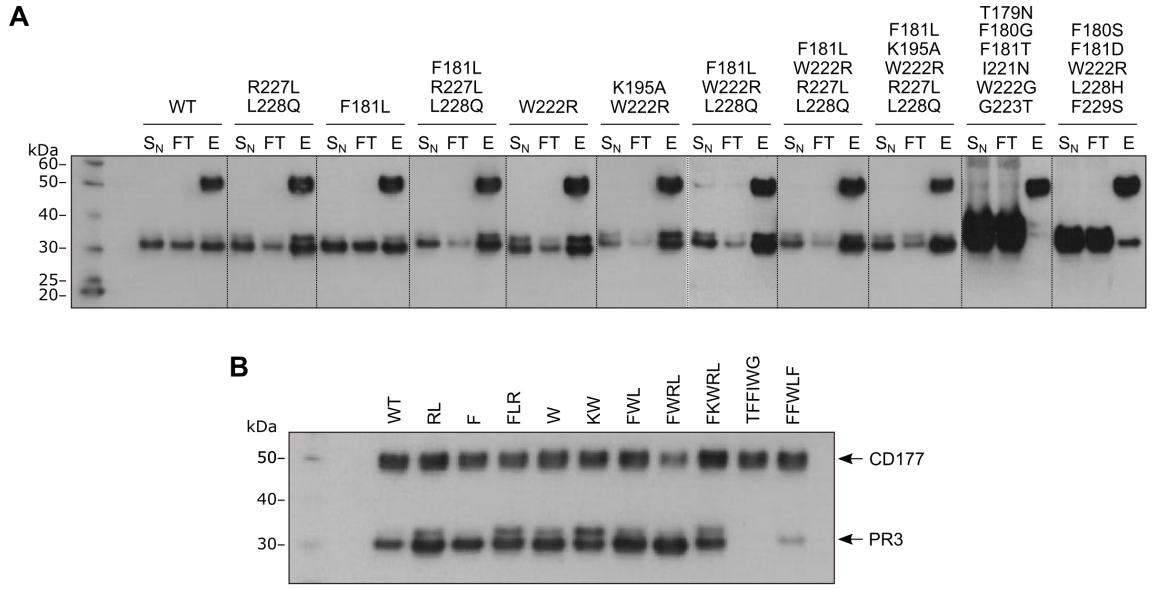


Figure 4.18: PR3 pull-down experiment. (A) Analysis of PR3 presence in supernatant (S_N), pull-down flow-through (FT) and elution (E) samples from different PR3 variants. Absence of PR3 in E sample indicates a non-CD177 binding PR3 variant. **(B)** Summary of all E samples.

4.4.5.2 SEC binding assays

To further confirm the previously observed results, purified PR3 variants were incubated with $CD177^{ecto}$ at a 2:1 molar ratio before being analysed by SEC.

Surprisingly, while the $PR3^{K195\ W222R}$ variant confirmed its ability to form a complex with CD177, the other $PR3^{F181L\ W222R\ L228Q}$ and $PR3^{F181L\ K195A\ W222R\ R227L\ L228Q}$ variants failed to form such a complex, suggesting a role of the F181, R227 and L228 residues in CD177 binding (Figure 4.19A-C). Mutating those residues may have led to a decrease in CD177-binding affinity, therefore explaining why interaction with CD177 could be observed in a PR3 pull-down but not by SEC.

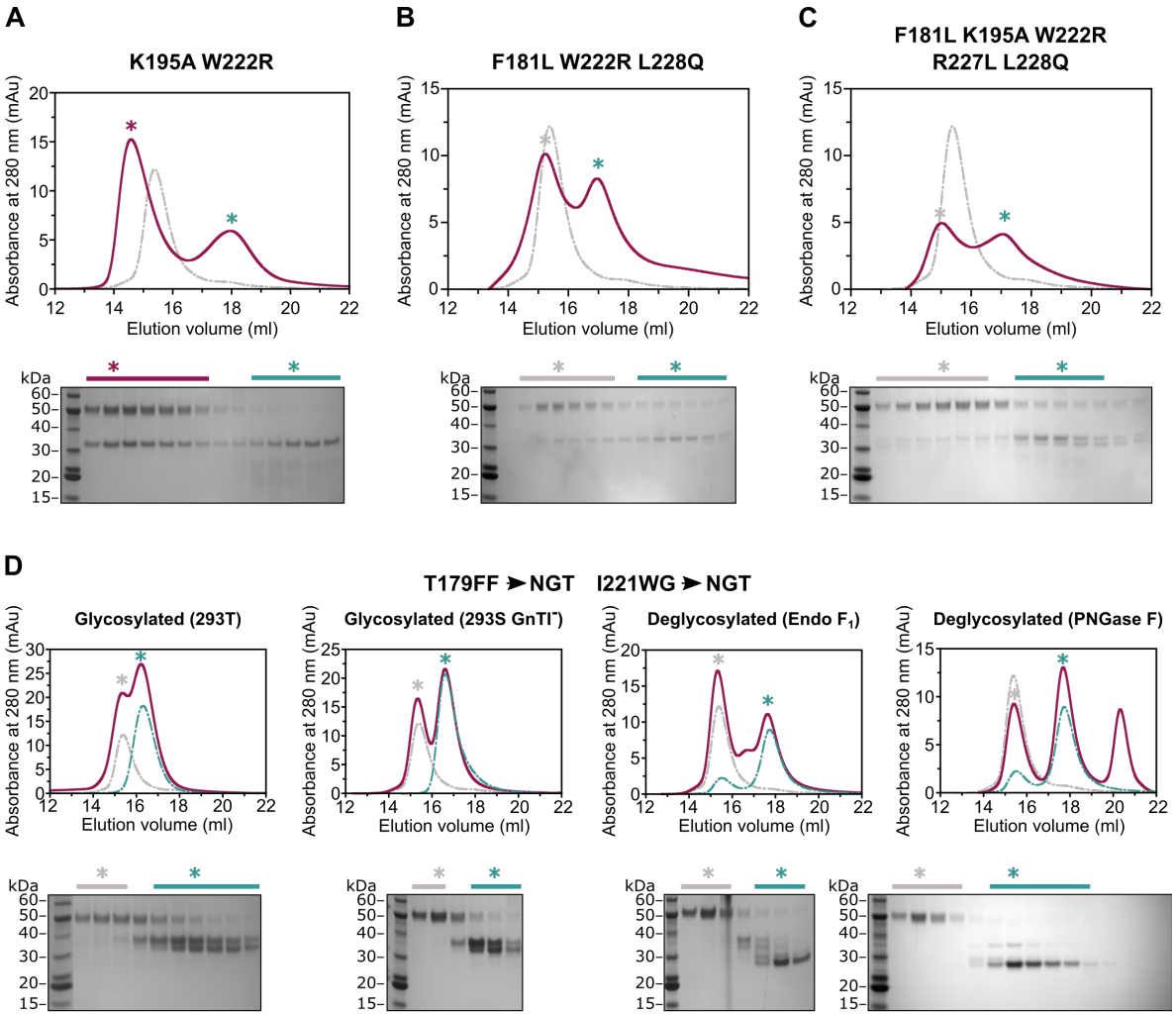


Figure 4.19: SEC binding assays of PR3 variants with CD177^{ecto}. SEC binding profile and corresponding SDS-PAGE gel of the PR3 variants **(A)** PR3$^{K195A\ W222R}$, **(B)** PR3$^{F181L\ W222R\ L228Q}$, **(C)** PR3$^{F181L\ K195A\ W222R\ R227L\ L228Q}$ and **(D)** PR3$^{TFF\text{-}NGT\ IWG\text{-}NGT}$ with CD177^{ecto}. CD177, PR3 variant and PR3-CD177 complex protein peaks are indicated by grey, green and purple asterisks, respectively. CD177 SEC profile is shown in grey dotted lines, the PR3:CD177 mixture SEC profile is shown in purple.

Moreover, no interaction with CD177 could be observed with the PR3$^{TFF\text{-}NGT\ IWG\text{-}NGT}$ variant, whether the protein was glycosylated or deglycosylated (Figure 4.19D). This highlights the importance of T179, F180, F181, I221, W222 and G223 in CD177 binding and suggests a role for some of them at the PR3-CD177 binding interface.

4.4.6 Investigation of CD177 binding site

4.4.6.1 Design and production of CD177-binding mutants

$PR3^{TFF\text{-}NGT\ IWG\text{-}NGT}$ is a double glycosylation mutant, which expressed very well but lost its ability to bind CD177. Given that both artificially added glycosylation sites are located relatively apart (Figure 4.14D), two new mutants, each bearing one single artificial glycosylation site, were designed with the hope that one of them would have the CD177 binding site restored.

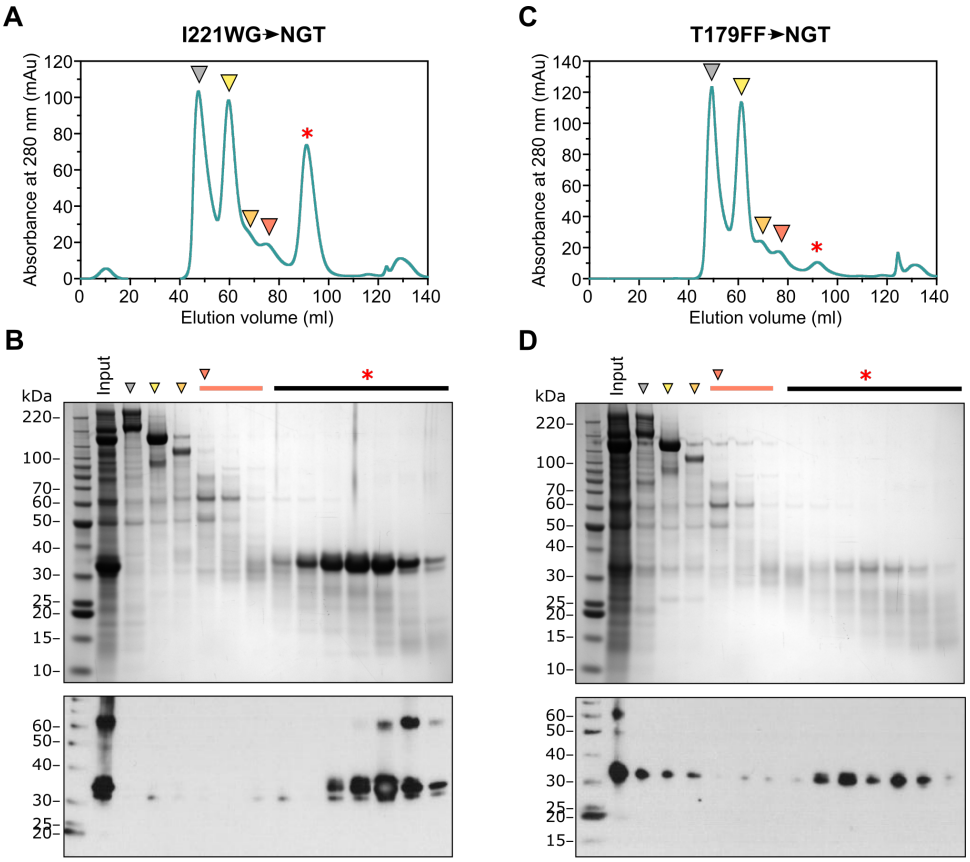


Figure 4.20: Large-scale purification of PR3 single glycosylation mutants. SEC purification profile and corresponding SDS-PAGE gel and western blot of **(A,B)** $PR3^{IWG\text{-}NGT}$ and **(C,D)** $PR3^{TFF\text{-}NGT}$. The protein peak (red asterisk) is preceded by a void and impurity peaks (grey, yellow, orange and red arrowheads, respectively).

Both $PR3^{TFF\text{-}NGT}$ and $PR3^{IWG\text{-}NGT}$ were expressed using HEK293S GnTI$^-$ and harvested after ten days of expression. $PR3^{IWG\text{-}NGT}$ appeared to be a well-expressed and soluble protein with yields reaching an average of 1.2 mg/L of culture (Figure 4.20A,B). By contrast, little amount of purified $PR3^{TFF\text{-}NGT}$ was recovered. $PR3^{TFF\text{-}NGT}$ seemed

to behave similarly to the mouse mutants as it was not only present in the void and impurity fractions, but was also produced with a yield of 100 µg/L of impure protein (Figure 4.20C,D).

Interestingly, both $PR3^{IWG\text{-}NGT}$ and $PR3^{TFF\text{-}NGT}$ appeared to be stably expressed and to stay soluble during protein expression and cell harvesting (Figure 4.21). This therefore suggests that $PR3^{TFF\text{-}NGT}$ is lost during the subsequent purification steps.

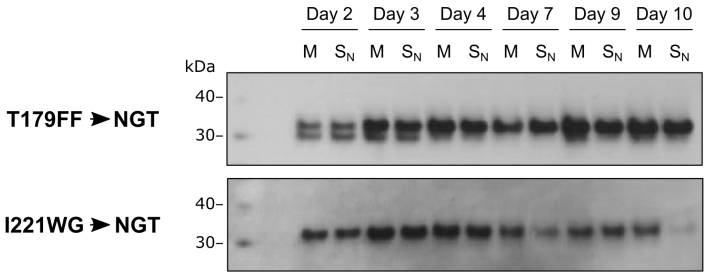


Figure 4.21: Evolution of $PR3^{IWG\text{-}NGT}$ and $PR3^{TFF\text{-}NGT}$ solubility during protein expression. Analysis and comparison of PR3 presence in samples of culture medium at collection (M) and after centrifugation (S_N) taken at different days during protein expression.

4.4.6.2 Binding assays

Both single glycosylation mutants were then subjected to a SEC binding assay with $CD177^{ecto}$. Surprisingly, both mutants showed binding to $CD177^{ecto}$, albeit with different strength. Indeed, the $PR3^{IWG\text{-}NGT}$ SEC profile displayed formation of a new complex peak along with complete disappearance of CD177 protein peak, indicating the formation of a strong complex (Figure 4.22A). By contrast, the $PR3^{TFF\text{-}NGT}$ SEC profile displayed both peaks, suggesting formation of a weaker complex, which partly dissociated after the sample was applied onto the column (Figure 4.22B).

Comparison of binding results obtained from $PR3^{IWG\text{-}NGT}$ and from the mouse mutant $PR3^{F181L\ W222R\ L228Q}$, which showed no CD177 binding, enabled identification of F181 and L228 as potential key residues in CD177 binding.

Additionally, comparison of $PR3^{TFF\text{-}NGT}$ and $PR3^{IWG\text{-}NGT}$ SEC profiles revealed a significant shift in elution volume (Figure 4.22B), suggesting that $PR3^{TFF\text{-}NGT}$ elutes as a larger or more elongated protein than $PR3^{IWG\text{-}NGT}$. The role of I221, W222 and G223 and the reasons behind such a behaviour remains to be elucidated.

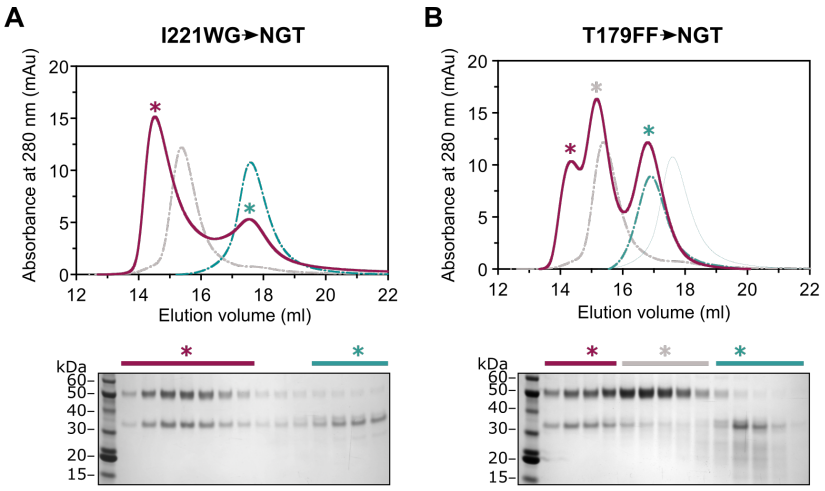


Figure 4.22: SEC binding assays of PR3 single glycosylation mutants with CD177^{ecto}. **(A,B)** SEC binding profiles with corresponding SDS-PAGE gels of PR3$^{IWG\text{-}NGT}$ and PR3$^{TFF\text{-}NGT}$ with CD177^{ecto}. CD177, PR3 variant and PR3-CD177 complex protein peaks are indicated by grey, green and purple asterisks, respectively. CD177 and PR3 variant individual SEC profiles are shown in grey and green dotted lines, respectively. The PR3:CD177 mixture SEC profile is shown in purple. **(B)** SEC binding profile of PR3$^{IWG\text{-}NGT}$ is shown in faint green line for comparison.

To further confirm the CD177- and non CD177-binding properties of PR3$^{IWG\text{-}NGT}$ and PR3$^{TFF\text{-}NGT\ IWG\text{-}NGT}$, respectively, SPR experiments were performed with both PR3 variants immobilised on the sensor chip and with CD177^{ecto} as the solute analyte. While no binding was observed for PR3$^{TFF\text{-}NGT\ IWG\text{-}NGT}$, PR3$^{IWG\text{-}NGT}$ displayed a concentration-dependent binding to CD177^{ecto}, which enabled determination of a dissociation constant (K_d) of 77.94 nM. Intriguingly, a reverse experiment where CD177^{ecto} was immobilised on the sensor chip and PR3$^{IWG\text{-}NGT}$ used as the solute analyte measured the complex dissociation constant at 10.79 nM instead. The reasons behind this seven-fold difference remain unclear, but may suggest PR3$^{IWG\text{-}NGT}$ to be more unstable on the sensor chip than CD177^{ecto}.

Nevertheless, those calculated K_d values stayed in range with previously published values of 4.1 nM for WT nPR3, and 57 nM for a variant possessing I221A and W222A mutations [124].

Overall, PR3$^{IWG\text{-}NGT}$ appeared to be the most suitable variant to use in further biophysical, structural and functional studies as it is a well-behaving variant which retained its ability to bind CD177.

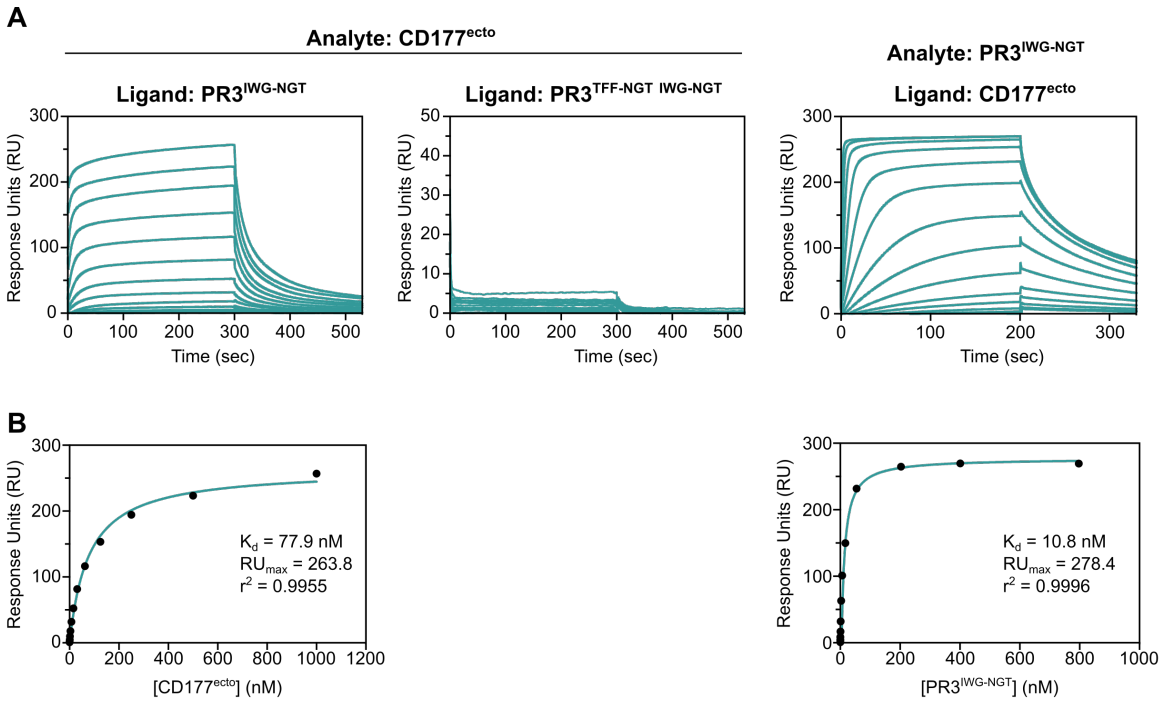


Figure 4.23: SPR analysis of PR3$^{\text{IWG-NGT}}$ and PR3$^{\text{TFF-NGT IWG-NGT}}$ interaction with CD177$^{\text{ecto}}$. (A) PR3$^{\text{IWG-NGT}}$ and PR3$^{\text{TFF-NGT IWG-NGT}}$ were both immobilised at a surface density of 330 RU. CD177$^{\text{ecto}}$ was immobilised at a surface density of 560 RU. **(B)** For K_d measurements, the binding signal was measured for 15 two-fold serial diluted samples of CD177$^{\text{ecto}}$ and PR3$^{\text{IWG-NGT}}$, with concentration ranging from 0.06 to 1000 nM and 0.05 to 800 nM, respectively.

4.5 Putative dimerisation of PR3

Following up on previous results, which suggested a role for F181 and L228 in CD177 binding, it appeared unlikely that CD177-bound PR3 would exist in the tetramer conformation observed in the crystal structure. Indeed, CD177 seemed to interact with the bottom side of the hydrophobic patch, thus preventing interaction between two 'top copies' of PR3 with two 'bottom copies' of PR3 (Figure 4.14A).

However, some evidence also points towards the existence of PR3 dimers. In particular, the W222 residue appears to play a major role in the formation of this putative dimer as each W222 inserts itself into a groove presented by the other protomer (Figure 4.24). A closer look at the crystal structure identified K195 as another potential residue involved in dimerisation as it forms a salt-bridge with E113 from the other protomer (Figure 4.24B). Interestingly, both K195 and W222 are the least conserved positively-charged and hydrophobic residue, respectively, among PR3 orthologs (Figure 4.14B).

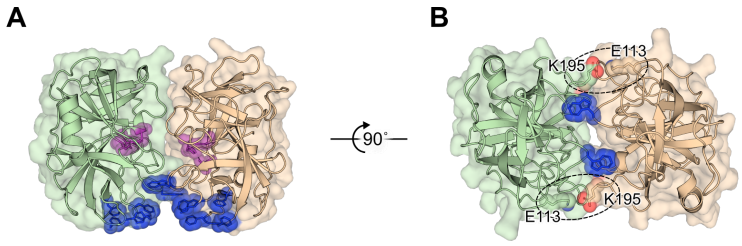


Figure 4.24: PR3 putative dimerisation interface. (A) Front view of the putative PR3 dimer. The catalytic triad and hydrophobic residues from the hydrophobic patch are depicted in purple and blue spheres, respectively. **(B)** Bottom view of the putative PR3 dimer. W222 is depicted in blue sticks and spheres. The two K195-E113 salt bridges are indicated by an oval, with residues depicted in sticks and spheres.

4.5.1 Design and production of PR3 dimer variants

The role of W222 in putative dimerisation made it necessary to use PR3 variants with unmutated W222 for further investigation. However, most of the variants possessed the W222R mutation, and the one which did not have the mutation (*e.g.* $PR3^{TFF\text{-}NGT}$) failed to produce a well-behaving protein (Figure 4.20C,D).

As a consequence, the synthetic $PR3^{F180S\ F181D\ W222R\ L228H\ F229S}$ variant, which comprises mutations to highly hydrophilic and polar residues, was used to design $PR3^{F180S\ F181D\ L228H\ F229S}$ with the hope that the four S180, D181, H228 and S229 residues would overcome the hydrophobicity generated by mutating W222 back.

Both $PR3^{F180S\ F181D\ W222R\ L228H\ F229S}$ and $PR3^{F180S\ F181D\ L228H\ F229S}$ were expressed by HEK293S GnTI$^-$ cells and harvested after ten days of expression. $PR3^{F180S\ F181D\ W222R\ L228H\ F229S}$ expressed well and behaved similarly to $PR3^{IWG\text{-}NGT}$, with a resulting yield of ∼1.2 mg/L of culture (Figure 4.25A,B). By comparison, $PR3^{F180S\ F181D\ L228H\ F229S}$'s yield was reduced by two-fold and was co-purified with impurities (Figure 4.25C,D). Interestingly however, when harvested after three days of expression, the behaviour and yield of $PR3^{F180S\ F181D\ L228H\ F229S}$ drastically improved (Figure 4.25C-E). This also resulted in the appearance of a new peak eluting 10 ml earlier than the usual monomeric PR3 peak. Interestingly, fractions collected from this new, higher molecular weight peak ran as a pure protein of 30 kDa on an SDS-PAGE gel. Taken together, these results highly suggest the existence of PR3 oligomers, and confirm the important role of W222 in PR3 oligomerisation.

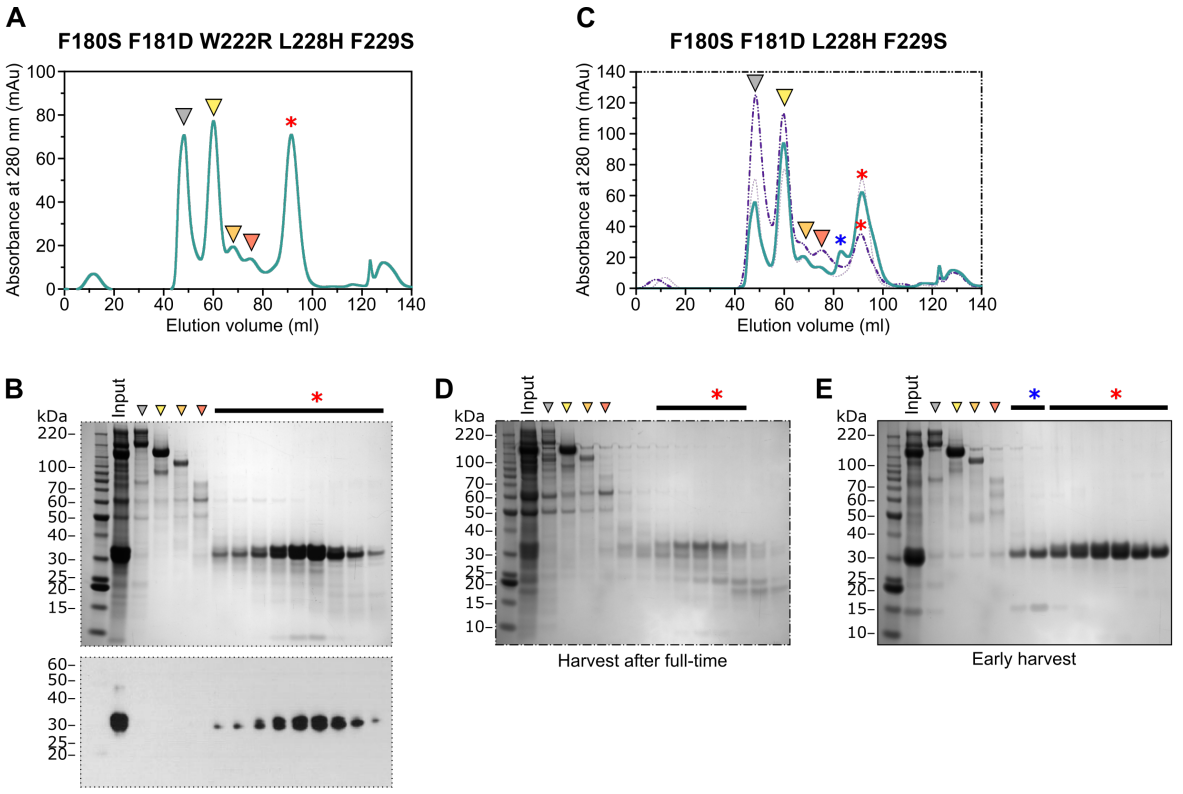


Figure 4.25: Large-scale purification of PR3 synthetic and dimer mutant. SEC purification profile with corresponding SDS-PAGE gel and western blot of **(A,B)** PR3$^{\text{F180S F181D W222R L228H F229S}}$ and **(C-E)** PR3$^{\text{F180S F181D L228H F229S}}$. The protein peak is indicated by a red asterisk, void and impurity peaks are indicated by grey, yellow, orange and red arrowheads, respectively. **(C)** Purification profiles of PR3$^{\text{F180S F181D L228H F229S}}$ harvested early and after full-time expression are depicted in green line and purple dotted line, respectively. Purification profile of PR3$^{\text{F180S F181D W222R L228H F229S}}$ is shown as thin purple dotted line for comparison.

4.5.2 Identification of PR3 variants' oligomerisation states

PR3$^{\text{F180S F181D W222R L228H F229S}}$ as well as both oligomeric and monomeric species of PR3$^{\text{F180S F181D L228H F229S}}$ purified by SEC were subjected to a SEC-multi-angle laser scattering (SEC-MALS) experiment for determination of their molecular weight.

The SEC-MALS experiment revealed that both PR3$^{\text{F180S F181D W222R L228H F229S}}$ and monomeric species of PR3$^{\text{F180S F181D L228H F229S}}$ are 30 kDa monomeric proteins. By contrast, the SEC-MALS profile showed two peaks for oligomeric species of PR3$^{\text{F180S F181D L228H F229S}}$, one of which a 30 kDa monomer and the other a 60 kDa dimer. The low resolution of the measurements would, however, require the experiment to be repeated using a higher concentration of protein. Nevertheless, the SEC-MALS data indicates the existence of PR3 dimers, which appear to be driven by W222.

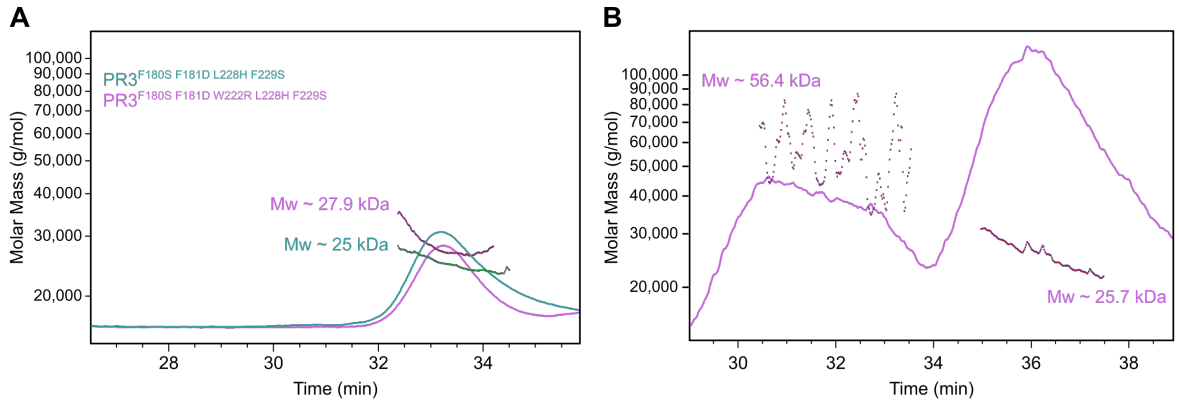


Figure 4.26: PR3 variants SEC-MALS experiment. **(A)** SEC-MALS profiles indicate that both PR3$^{\text{F180S F181D W222R L228H F229S}}$ and monomeric species of PR3$^{\text{F180S F181D L228H F229S}}$ are 30 kDa monomeric proteins. **(B)** SEC-MALS profile shows two peaks for the oligomeric PR3$^{\text{F180S F181D L228H F229S}}$ species, indicating the presence of 30 kDa monomers and 60 kDa dimers.

4.5.3 SEC binding assays

Both PR3$^{\text{F180S F181D W222R L228H F229S}}$ and PR3$^{\text{F180S F181D L228H F229S}}$ showed no complex formation with CD177$^{\text{ecto}}$ by SEC. However, as some binding signal was observed in the PR3 pull-down using PR3$^{\text{F180S F181D W222R L228H F229S}}$ (Figure 4.18), this suggests formation of a weaker complex between those PR3 variants and CD177$^{\text{ecto}}$. This also further confirms the importance of the mutated residues in CD177 binding.

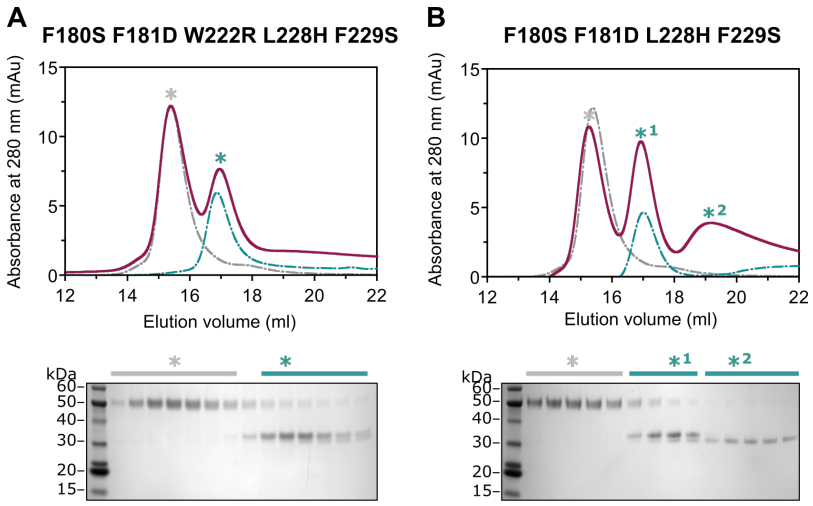


Figure 4.27: SEC binding assays of PR3 synthetic and dimer mutant. SEC binding profile with corresponding SDS-PAGE gel of **(A)** PR3$^{\text{F180S F181D W222R L228H F229S}}$ and **(B)** PR3$^{\text{F180S F181D L228H F229S}}$ with CD177$^{\text{ecto}}$. CD177, PR3 variant and PR3-CD177 complex protein peaks are indicated by grey, green and purple asterisks, respectively. CD177 and PR3 variant individual SEC profiles are shown in grey and green dotted lines, respectively. The PR3:CD177 mixture SEC profile is shown in purple.

4.6 PR3$^{\text{IWG-NGT}}$ recognition by GPA patients' sera

Because PR3$^{\text{IWG-NGT}}$ is a soluble and monomeric version of nPR3, it was necessary to test whether the variant retained its properties as the main autoantigen in the autoimmune GPA disease before using it in further functional studies.

4.6.1 ELISA experiments

Preliminary ELISA experiments were performed, in which various PR3 variants were coated and incubated with blood sera from either an ANCA-positive patient or from a healthy control (Figure 4.28A).

nPR3 as well as both PR3 variants were successfully recognised by ANCAs in this assay. Remarkably, PR3$^{\text{IWG-NGT}}$ showed a drastic improvement in signal detection compared to nPR3. PR3$^{\text{F180S F181D L228H F229S}}$ was also tested, and although its detection signal was not as strong as PR3$^{\text{IWG-NGT}}$, it was still a significant improvement from nPR3. Notably, this assay was done using patients' sera diluted by 200 instead of the usual dilution by 50 or 100 that is performed when using nPR3 [333], which may explain why the signal observed with nPR3 is so low.

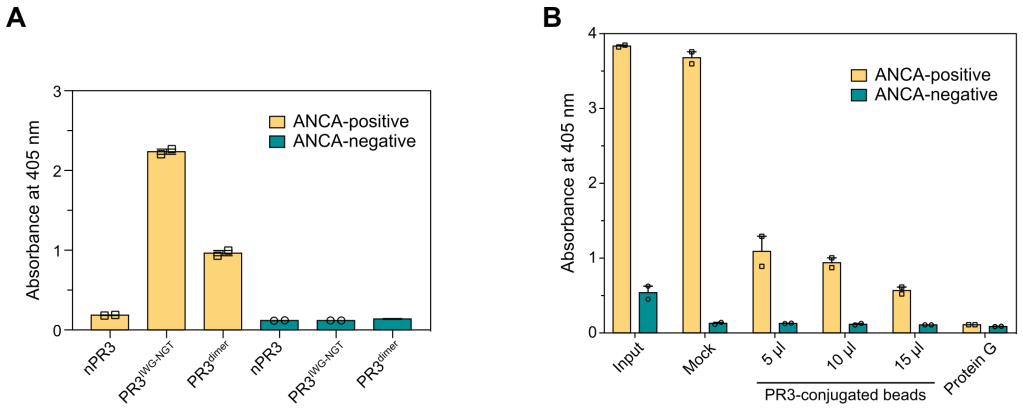


Figure 4.28: PR3$^{\text{IWG-NGT}}$ recognition by GPA patients' sera. (A) Preliminary ELISA experiments show a higher ANCA detection signal from both PR3$^{\text{IWG-NGT}}$ and PR3$^{\text{F180S F181D L228H F229S}}$ variants than nPR3 (n = 2 repeats, mean ± SEM). **(B)** Comparison of ANCA quantities in patients' sera shows that the use of PR3-conjugated beads is efficient for depleting patients' blood from their ANCAs (n = 2 repeats, mean ± SEM).

4.6.2 ANCA depletion experiment

In similar experiments, patients' sera were incubated with increasing amounts of PR3-conjugated beads before the resulting flow-through was collected and analysed by ELISA (Figure 4.28B). The data show a successful depletion of ANCAs from the serum using PR3-conjugated beads, albeit not complete when compared to the use of Protein G-conjugated beads.

Overall, $PR3^{IWG\text{-}NGT}$ retained its properties as the main autoantigen targeted by ANCAs. The results also suggest a potential usage of this variant for diagnostic, or even therapeutic, purposes.

4.7 Discussion

4.7.1 Advantages of using a recombinantly-produced PR3

In the majority of PR3 studies found in the literature, nPR3 was used. While most of the recent studies just purchase the protein from providers, older studies used to produce their own nPR3. To do this, neutrophils were either subjected to a stimulant known to cause degranulation, or burst open under high pressure. The released granules were then collected and typically incubated with 1.5% Triton X-100 in order to lyse them and recover PR3 in the supernatant. From there, different purification method have been tested and have yielded different results. In some cases, PR3 was purified by affinity chromatography using anti-PR3 mAb [40, 32], or purified ANCAs from patients [334]. In other cases, the protein was purified by dye-ligand affinity chromatography, which used a dye called orange A to act as a PR3 substrate analog [335]. This was often followed by ion exchange chromatography [53, 336] or affinity chromatography using anti-PR3 mAb [336]. The more common association of ion exchange chromatography followed by SEC has also been reported [50, 337, 338] as well the use of reverse high-performance liquid chromatography (HPLC) [41, 339] to purify PR3.

When it was indicated, yields from such purifications varied between 25 and 60 ng of PR3 from 10^6 neutrophils. Notably, a normal range of neutrophils in a healthy adult

corresponds to 2500-7000 neutrophils per µl of blood, which, in other words, translates the yield into a maximum of 0.4 ng of PR3 per µl of blood.

In comparison, the recombinantly expressed PR3 can be purified using a standard purification protocol consisting of IMAC followed by SEC. Additionally, the purification does not need to be performed at 4℃, nor in the presence of the serine protease inhibitor PMSF as the recombinant PR3 bears a mutation which renders the protease catalytically inactive. Most strikingly, the yield from a recombinant expression averages around 3.5 mg from a routine 3 L culture. By contrast, if purified from native sources, such a yield would only be possibly achieved by using 9 L of blood.

4.7.2 A challenging protein production and purification

The extremely low yield of PR3 combined with the fact that the protein costs nearly eight times as much as neutrophil-purified MPO suggests a particularly difficult and expensive purification process. Therefore, over the last 30 years, numerous attempts have been made to recombinantly produce PR3. However, while some studies managed to produce and use a recombinantly-produced protein in functional assays, very few reported successful protein purifications.

4.7.2.1 The recombinantly-produced PR3 variants are functionally relevant

Previous studies have found it necessary to use an eukaryotic expression system for expression of a properly folded PR3. Since then, PR3 has thus been recombinantly expressed using yeast [340], Sf9 insect cells [319, 341, 342, 343], human mast cell line-1 (HMC-1) [344, 345, 329, 330] and HEK293 cells [329, 330, 36, 124]. However, when the recombinantly-produced PR3 were tested for their recognition by patients' serum containing anti-PR3 ANCAs, only the proteins produced by HMC-1 and HEK293 cells were recognised by the autoantibodies [344, 345, 329, 330].

This work falls into line with those previous observations and confirms the use of HEK293 cells as a suitable expression system for the production of PR3 variants. In addition, preliminary ELISA experiments confirmed that both $PR3^{IWG\text{-}NGT}$ and

$PR3^{F180S\ F181D\ L228H\ F229D}$ have retained their ability to be recognised by ANCAs and are therefore relevant for subsequent functional studies.

Intriguingly, recombinant WT PR3 appeared to be produced as a mixture of three main species, which also appear to be the case for nPR3 [337, 338]. However, whether those are the same species and whether this holds a functional importance remains unknown.

4.7.2.2 PR3 is an aggregation-prone protein

Following several unsuccessful purification trials and failed optimisation attempts, PR3 appeared to be an exceptionally challenging protein to purify. Mounting evidence in the literature have also been pointing at PR3 as an aggregation-prone protein that is difficult to purify and handle. This is further illustrated by the unique crystal structure that has been deposited for PR3 in 1996 while all other NSPs have seen a growing number of structures being deposited until today (Table 4.1).

Protein	Number of PDBs	Recombinant expression?
Proteinase 3	1	Yes (1)
Neutrophil Elastase	23	No
Cathepsin G	4	No
Neutrophil Serine Protease 4	4	Yes (4)
Azurocidin	4	Yes (4)

Table 4.1: Total number of published structures of all NSPs. The number of structures solved from recombinantly-produced proteins is indicated in parenthesis.

Interestingly, very few reports of successful PR3 purifications exist and the few that do either omit to describe the purification procedure [319, 124, 125], or fail to present the purification SEC profiles and/or the corresponding SDS-PAGE gels [36, 124, 125].

Notably, several purification trials have been using detergents such as 0.1% Triton X-100 [346], 1% β-OG [338], 0.1% Tween-80 [343] or 0.02% DDM [124] in order to decrease the amount of aggregation. Unfortunately however, none of those detergents successfully managed to improve the protein recovery in this work. In addition, because presence of detergents in the buffer could pose an issue for subsequent functional assays (*e.g.* those involving T cells), this avenue was not further pursued.

4.7.2.3 PR3 is lost upon prolonged interaction with the cell culture medium

In addition to behave as a sticky protein, PR3 also appears to be unstable and to aggregate upon prolonged contact with cell culture medium during expression. Indeed, an early harvest managed to significantly improve the amount of protein recovered after purification. The reasons behind such a behaviour remain however unknown.

In a physiological system, the amount of PR3 in blood is usually regulated by its endogenous inhibitor A1AT. A1AT is also called a suicide inhibitor as it covalently binds to the catalytic serine of PR3 in order to deactivate and clear the blood from PR3 protease activity. Nevertheless, knowing that all PR3 variants have their catalytic serine mutated to alanine, such a mechanism cannot be responsible for the observed loss of PR3 in the medium.

There is a possibility that PR3 becomes unstable upon the increasing amount of expressed protein in the medium, or upon interaction with the increasing number of dead cells. The latter is particularly interesting as PR3 has been shown to have a role in efferocytosis and to directly interact with apoptotic cells [58, 59, 62].

4.7.3 Using surface mutants to improve the protein surface properties

Mutating the protein hydrophobic patch revealed to be successful for improving protein solubility and behaviour in solution. This led to the creation of a panel of PR3 variants, with variable level of solubility as well as variable CD177 binding ability (Figure 4.29). While the most highly expressed and soluble variant is $PR3^{\text{TFF-NGT IWG-NGT}}$, which possesses two added glycosylation sites, the best variant to use in subsequent studies is $PR3^{\text{IWG-NGT}}$ as it retained its CD177 binding ability. By contrast, the most conservative mouse mutants did not show a significant improvement and behaved just slightly better than the WT.

It has been previously observed that WT PR3 was produced as a mixture of three species, with only one remaining soluble after centrifugation. Intriguingly, mutating residues from the hydrophobic patch seemed to produce different mixtures of PR3 species (Figure 4.29D). In particular, both $PR3^{\text{TFF-NGT IWG-NGT}}$ and $PR3^{\text{TFF-NGT}}$ exist as two

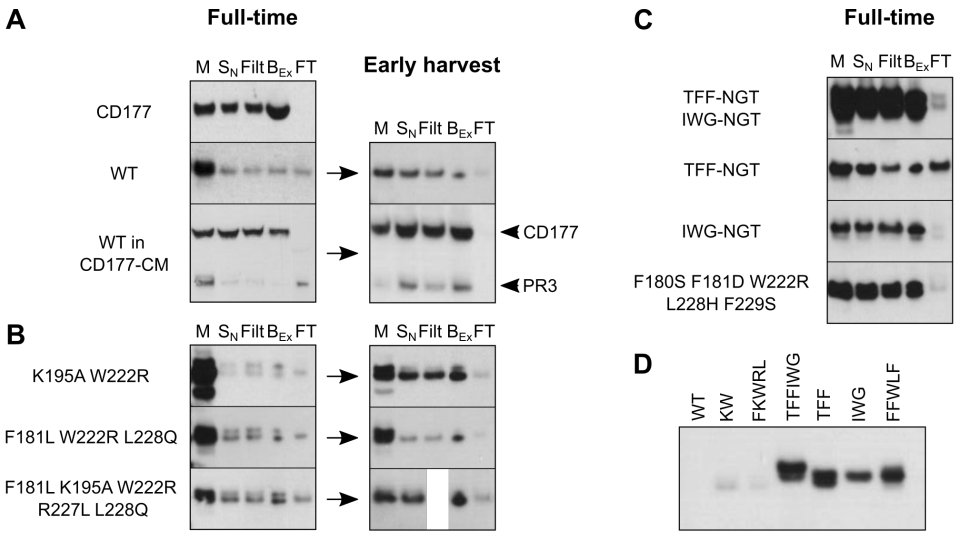


Figure 4.29: Summary and comparison of all purified PR3 variants. PR3 protein signal was analysed in culture medium samples at collection (M), after centrifugation (S_N), filtration (Filt), buffer exchange (B_{Ex}) and in IMAC flow-through (FT). of wild type PR3 and mouse mutants when harvested after full-time and **(A,B)** Early harvest of wild type and PR3 mouse mutants results in better protein behaviour throughout the purification steps. **(C)** Glycosylation and synthetic PR3 mutants show a drastic improvement of protein behaviour during purification. **(D)** Comparison of S_N samples from various PR3 variants.

soluble species while $PR3^{IWG\text{-}NGT}$ exists as one soluble species. The reason behind these differences remains unknown. However, investigating the nature of those different species could appear essential for a better understanding of PR3 biochemical properties.

4.7.4 PR3 dimerisation and/or oligomerisation

Preliminary experiments have suggested a role for W222 in the formation of PR3 dimers, confirming a previous study which observed similar results [124]. Intriguingly, the proposed model of PR3 dimers shows two molecules facing and presenting their catalytic cleft to each other, so that both catalytic sites are buried and hidden after dimer formation. Further experiments would however be required to confirm the existence of such dimers as well as the dimer interface. In particular, crystallisation trials and/or hydrogen-deuterium exchange experiments using the dimeric $PR3^{F180S\ F181D\ L228H\ F229S}$ could be performed. Nevertheless, confirmation of the existence of PR3 dimers would dramatically increase the knowledge on PR3 and would open up many possibilities for protein regulation and activation in physiological conditions.

4.7.5 The PR3-CD177 binding site

In accordance with previous studies, mutational analysis as well as binding assays have confirmed the CD177 binding interface to be located on PR3 hydrophobic patch. By comparing the CD177 binging ability of the different PR3 variants, F180, F181 and L228 are proposed as key residues involved in the binding interface with CD177. As those residues are relatively far from W222 and the putative dimerisation interface, there is a possibility that the PR3-CD177 complex actually exists as a 2:2 protein complex.
Previous studies which attempted to study the role of CD177 binding on PR3 proteolytic activities have found that despite a reduced level of proteolytic activity compared to free PR3, CD177-bound PR3 remains catalytically active [36, 124]. Apart from that, there are no other clues on the functional role of the PR3-CD177 complex on the surface of neutrophils.
There is therefore a need to understand how CD177 interacts with PR3, and whether CD177 acts as a simple receptor or possesses a more complex regulatory role for PR3.

Chapter 5

Structural studies of the PR3-CD177 complex

CD177 is a GPI-anchored member of the Ly6/uPAR family, and the known receptor for PR3 on the surface of neutrophils. In this chapter, recombinant production and purification of the ectodomain of CD177 enabled production of crystals which diffracted to 2.7 Å. As the entirety of CD177 ectodomain was predicted to be a succession of four LU domains, molecular replacement was first attempted in order to solve the structure. However, this revealed to be challenging due to the absence of high quality homology models and the relatively low resolution of the data. S-SAD phasing was instead used and resulted in a model of the full CD177 ectodomain consisting of two compact subdomains, each comprising two LU domains facing each other in a head to tail manner. Interestingly, binding assays using the soluble $PR3^{IWG\text{-}NGT}$ variant characterised in Chapter 4 showed that only the first subdomain of CD177 ($CD177^{subA}$) was necessary for PR3 binding. Co-crystallisation trials of $CD177^{subA}$ with $PR3^{IWG\text{-}NGT}$ yielded crystals diffracting to 1.7 Å whose datasets could be solved by molecular replacement and resulted in the first high-resolution model of the PR3-CD177 complex.

5.1 CD177^{ecto} production and crystallisation

5.1.1 Bioinformatic analysis

Human CD177 is a ∼50 kDa GPI-anchored glycoprotein (Figure 5.1A) and is an atypical member of the Ly6/uPAR family. Indeed, it has two predicted LU domains within its ECD [90] while most of the members only have one [88]. uPAR is the only other protein with three LU domains and whose structure has been solved [347]. LU domains are characterised by the presence of six β-strands forming a structural core stabilised by four conserved disulphide bridges, and from which extends three β-stranded loops (Figure 5.1B,C). A fifth disulphide bridge commonly exists to further stabilise β1- to β2-strand.

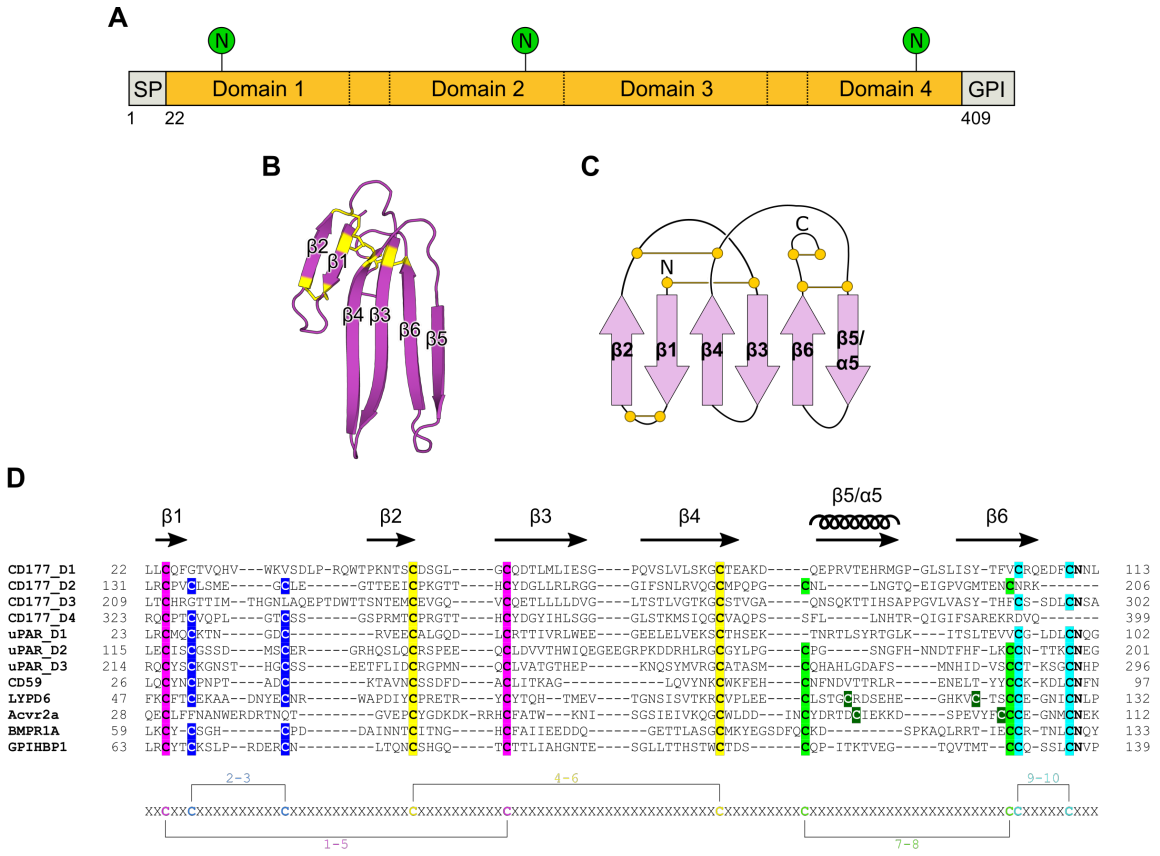


Figure 5.1: Predicted CD177 domain architecture. (A) Schematic depicting the predicted LU domain boundaries after sequence analyses. CD177 possesses 3 predicted N-glycosylation sites (N46, N189, N382) indicated by green spheres. **(B)** Cartoon representation of the first LU domain from uPAR (PDB ID: 1YWH). **(C)** Typical topology of a LU domain. Disulphide bridges are indicated by yellow spheres linked together. **(D)** Sequence alignment with secondary structure of CD177 with crystallised mammalian LU domains.

Bioinformatic tools confidently predicted the first 22 residues, M1 to A21, as the protein signal peptide, and the last 28 residues, G409 to C437, as a GPI-attachment signal sequence. However, due to the high cysteine content of CD177 (26 cysteines), delimiting the boundaries of the two LU domains was challenging. Different homology modelling servers such as Swiss-Model, HHPred, Phyre^2 and I-TASSER resulted in different models, which all appeared to be unlikely. Indeed, the resulting homology models would place some of the conserved disulphide bridges, but would also place several cysteines in the middle of β-strands or loops. Swiss-Model and I-TASSER produced homology models consisting of three LU domains using uPAR as a template, but those were misaligned with CD177 sequence as several cysteines were not properly arranged in disulphide bridges. Manual identification and annotation of cysteines enabled to define four separate LU domains, named Domain 1 to 4 (D1-D4), spanning CD177 ECD until residue 399. The delimited LU domains appeared to lack one or two of the conserved disulphide bridges (Figure 5.1D), which could explain why the bioinformatic tools failed to properly align and predict CD177 LU domains. Notably, homology modelling of individual and properly delimited LU domains was far more successful and yielded reliable models with all cysteines forming disulphide bonds.
Intriguingly, a closer look at the sequence and at the arrangement of the cysteine bridges suggested a possible sub-organisation into two subdomains, where one subdomain contains two successive LU domains linked together by a 20 residues-flexible linker.

5.1.2 Construct screening

A full-length construct of CD177 as well as constructs without the GPI-attachment signal were designed and tested for their expression levels in HEK293T cells. As expected, the full-length CD177 could not be detected in the culture medium supernatant as the protein likely remained attached to the cell membranes through its GPI-anchor (Figure 5.2A). Interestingly, keeping the native signal peptide seemed to be essential to obtain an acceptable level of protein expression (Figure 5.3A).

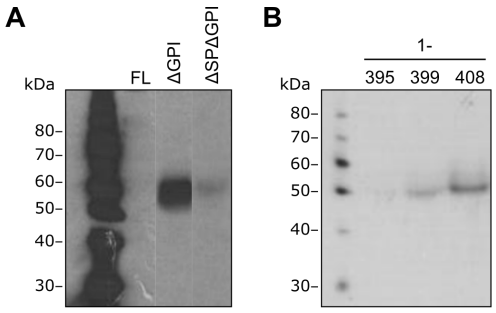


Figure 5.2: Expression trials of CD177^{ecto} constructs. Western blot comparing the expression levels of different CD177^{ecto} constructs. **(A)** Construct possessing the native signal peptide (ΔGPI) is significantly more expressed than construct possessing the vector signal peptide (ΔSPΔGPI). **(B)** Comparison of expression levels of CD177 constructs with different C-terminal boundaries.

Different C-terminal boundaries were also tested. 1-408 consists of the full CD177 ECD while the shorter, 1-395 and 1-399, constructs were designed to stop either after the last predicted β-strand or at the end of D4, respectively. The construct encompassing the full ECD (1-408) expressed the best and was therefore selected for large-scale expression and purification (Figure 5.2B).

5.1.3 Protein production and crystallisation trials

CD177 1-408, named CD177^{ecto}, was expressed in large-scale in HEK293S GnTI^{-} cells. CD177^{ecto} was successfully purified using a standard purification protocol consisting of IMAC and SEC, and yielded 3 mg of protein from 3 L of culture (Figure 5.3A,B).

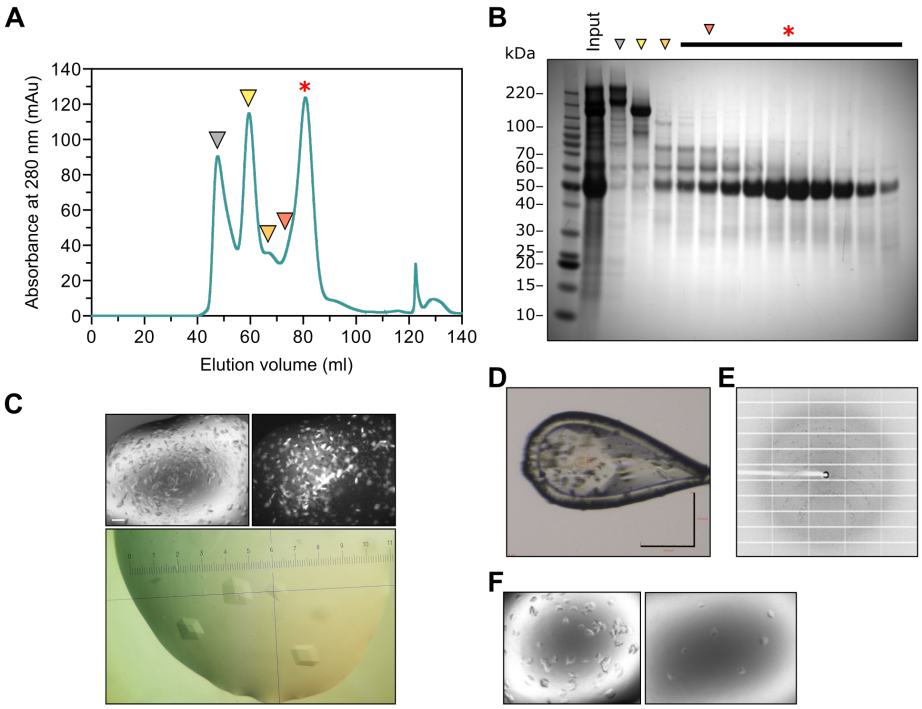


Figure 5.3: Large-scale production of CD177^{ecto}. (A,B) Representative SEC purification profile with corresponding SDS-PAGE gel showing CD177^{ecto} elutes around 80 ml on a Superdex® 200 column. The protein peak is indicated by a red asterisk. Void and impurities fractions are indicated by grey, yellow, orange and red arrowheads, respectively. **(C)** CD177^{ecto} crystals growing in condition C3 of the JCSG+ crystallisation screen (20% w/v PEG 3350, 0.20 M ammonium nitrate) at 4°C. Scale bar is 100 µm. **(D)** View of one crystal cryo-protected in 25% glycerol and mounted onto a SPINE pin prior to data collection. Scale bar is 100 µm. **(E)** Diffraction pattern resulting from one CD177^{ecto} crystal, showing protein diffraction spots around 3 Å. **(F)** New CD177^{ecto} crystals obtained from crystallisation optimisation experiments.

As CD177^{ecto} co-eluted with some impurities, the purest fractions were pooled together and concentrated for subsequent crystallisation trials (Table 5.1).

One condition from the JCSG+ screen at 4°C showed crystals growing just two days after incubation (Figure 5.3C). Notably, larger crystals were observed when a less concentrated protein (9.8 mg/ml) was used at a 2:1 protein:reservoir volume ratio instead of a more concentrated protein (12 mg/ml) used at a 1:1 ratio.

Crystals from both 2:1 and 1:1 drops were harvested at Day 8 and tested for their diffraction quality (Figure 5.3D,E).

Concentration	**Crystallisation screen**
12 mg/ml (283 µM)	Index, JCSG I-IV, Midas, Morpheus, Morpheus II, PACT, PGA
(1:1 ratio)	JCSG+, Morpheus, PACT at 4℃
9.8 mg/ml (231 µM)	JCSG II, Morpheus, PACT
(2:1 ratio)	JCSG+, Morpheus, PACT at 4℃
	Optimisation screens derived from the hit condition (20% w/v PEG 3350, 0.20 M ammonium nitrate at 4℃):
12 mg/ml (283 µM)	PEG 3350 horizontal gradient with seeding (15 to 22%) at 4℃
(1:1 ratio)	PEG 3350 horizontal gradient (15 to 22%) combined with
and	ammonium nitrate vertical gradient (0.17 to 0.22 M) at 4℃
(2:1 ratio)	Additive screen on hit condition at 4℃
	Silver bullets on hit condition at 4℃

Table 5.1: Summary of CD177^{ecto} crystallisation trials.

5.1.4 Structure determination

5.1.4.1 Data processing

Multiple crystals diffracted and enabled collection of several full datasets, which could all be autoprocessed to ∼3 Å resolution in the H 3 spacegroup. In order to compare them and choose the best dataset to use for all downstream calculations, all datasets were individually processed using DIALS for integration and AIMLESS for scaling and merging. All resolution cut-offs were chosen so that the correlation coefficient between random half datasets ($CC_{1/2}$) would not go below 30%, and were suggested by AIMLESS. The datasets were then compared by their R factors, $CC_{1/2}$ value, completeness and resolution (Table 5.2).

Overall, all datasets appeared to be relatively similar and could all be used for subsequent experiments. Nevertheless, the dataset possessing the lowest R factors, highest completeness as well as the best resolution appeared to be the merged dataset A-1+A-2.

5.1.4.2 Molecular replacement

Estimation of cell content using the Matthews coefficient method indicated a probability of 52% and 46% of having two and three copies of CD177^{ecto} in the ASU, respectively, as well as a corresponding solvent content of 58% and 37%. Similarly to GPR97^{ecto},

	Statistics when $CC_{1/2}$ is cut to 0.3				
	R_{merge} (%)	**R_{pim} (%)**	**$CC_{1/2}$ (%)**	**Completeness (%)**	**Resolution (Å)**
A-1	20.3 (237.7)	11.9 (134.6)	97.8 (42.0)	92.4 (92.5)	2.85
A-2	18.0 (202.2)	9.7 (110.9)	99.0 (33.5)	99.9 (99.7)	2.82
B	19.5 (172.9)	13.8 (121.2)	94.4 (37.9)	99.8 (99.7)	2.85
C	21.0 (204.1)	10.4 (100.7)	98.2 (62.2)	100 (100)	2.76
D	28.8 (192.4)	16.6 (112.2)	98.4 (45.9)	99.8 (99.6)	3.34
E	26.9 (194.8)	15.7 (115.4)	97.4 (38.1)	99.9 (100)	3.07
F	32.2 (137.8)	21.5 (92.6)	91.1 (36.3)	94.3 (95.9)	3.21
G	37.5 (117.6)	25.7 (81.9)	89.4 (43.1)	94.6 (94.8)	3.50
H	16.9 (160.7)	9.6 (92.0)	99.1 (47.7)	100 (99.9)	2.89
I	18.6 (114.4)	15.2 (94.1)	96.9 (46.4)	98.6 (97.7)	3.13
A-1+A-2	20.9 (292.5)	8.3 (120)	99.4 (36.3)	100 (100)	2.70
B+C	25.7 (488.4)	9.8 (192.4)	99.1 (35.4)	99.8 (98.7)	2.59

Table 5.2: Comparison between several native datasets. Crystals A to D were grown at 9.8 mg/ml in a 2:1 protein:reservoir ratio. Crystals E to I were grown at 12 mg/ml in a 1:1 protein:reservoir ratio. Addition of a number to a crystal name indicates that several datasets were collected from the same crystal. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor and R_{pim} for the precision-indicating merging R factor. R values were determined with all I+ and I-.

the relatively low resolution suggests the ASU is more likely to be composed of two copies of $CD177^{ecto}$.

Given that the entirety of $CD177^{ecto}$ consists of four LU domains, molecular replacement was first attempted by searching for eight copies of a LU domain-derived search model. To do this, models of previously published LU domains such as ACVR2A [95], BMPR1A [96], CD59 [348], LYPD6 [349] and uPAR D1-D3 [347] were used. These models were all first tested individually, then with their flexible loops truncated and also as an ensemble. Unfortunately, none of those models yielded good solutions.

In order to create better search models, sequence alignments of CD177 D1-D4 with all other crystallised LU domains were performed. The most similar LU domain template was chosen for each of CD177 LU domain, and was subjected to either Chainsaw [350] or Sculptor [351] in order to prune non-conserved sides chains and/or loops. Additionally, homology models were produced by HHPred, Swiss model, $Phyre^2$ and I-TASSER. However, none of those models managed to improve previous results.

A search model only containing the β-structural core with the four disulphide bridges was tested but also failed. Using another approach, the program was set to search two copies

of the three consecutive LU domains D1-D3 from uPAR, but failed to yield good solutions. Overall, molecular replacement revealed to be challenging. The relatively low resolution of the dataset combined with the fact that the search model is not only sub-optimal, but also only accounts for one eighth of the total protein in the ASU are probably the main reasons hindering successful molecular replacement.

5.1.4.3 Heavy-atom soaking

In order to obtain more CD177^{ecto} crystals, crystallisation trials were repeated in the hit condition as well as in optimised conditions derived from the initial hit (Figure 5.3F, Table 5.1). Newly grown crystals were then soaked with a variety of heavy-atoms including platinum, gold, mercury and tantalum (Table 2.3) in order to obtain heavy-atom derivatised crystals for SAD and/or single or multiple isomorphous replacement with anomalous dispersion (SIRAS or MIRAS, respectively) experiments. All of those techniques involve determination of a substructure created by the heavy-atoms, which can then be used to help phase, and therefore, solve the structure of the protein. Where anomalous dispersion uses the different interaction properties of the heavy-atoms with the X-rays, isomorphous replacement uses the comparison between a native dataset with one or more isomorphous heavy-atom derivatised datasets to help determine the heavy-atom structure within the crystal.

Datasets from heavy-atom derivatised crystals were collected at one of the X-ray absorption edge of the corresponding heavy-atom and were processed using the XDS package. The best diffracting crystals were crystals which were soaked with Potassium tetracyanoplatinate(II) hydrate ($K_2Pt(CN)_4 \bullet xH_2O$), Potassium hexachloroplatinate(IV) (K_2PtCl_6) and Gold(I) Potassium Cyanide ($KAu(CN)_2$) (Table 5.3).

As the intensity of the anomalous signal needs to be greater than 1 for it to be reliably used in substructure determination, each dataset was analysed in order to determine the resolution at which such a signal could be observed.

The anomalous signal of the two platinum-derived metals, potassium hexachloroplatinate(IV) and potassium tetracyanoplatinate(II) hydrate, appeared to be very low and required to work at the resolution of 7 and 10 Å, respectively. However, the one of the

	Statistics when $CC_{1/2}$ is cut to 0.3				
	R_{merge} (%)	R_{pim} (%)	$CC_{1/2}$ (%)	Completeness (%)	Resolution (Å)
$K_2Pt(CN)_4 \bullet xH_2O$					
J	23.7 (133.5)	16.7 (100.2)	87.2 (35.7)	97.5 (91.5)	2.82
K	16.2 (137.7)	9.7 (93.5)	94.6 (38.9)	97.4 (92.7)	2.80
L	11.2 (149.4)	8.1 (111.3)	98.1 (41.8)	98.0 (88.5)	2.80
M	11.2 (149.4)	8.1 (111.3)	98.1 (41.8)	98.0 (88.5)	2.80
K_2PtCl_6					
N	40.8 (309.3)	19.5 (161.8)	98.3 (31.6)	99.9 (99.6)	3.70
$KAu(CN)_2$					
O	13.4 (197.4)	8.9 (152.6)	99.3 (31.3)	96.8 (81.7)	2.70
P-1	16.3 (142.1)	9.7 (102.7)	98.7 (41.3)	97.3 (82.9)	2.70
P-2	17.3 (173.9)	10.5 (120.1)	98.5 (27.5)	98.9 (93.8)	2.80
M+P-1	16.8 (184.8)	7.5 (108.5)	99.1 (45.0)	99.8 (98.7)	2.70
O+P-2	19.2 (249.5)	8.3 (133.4)	99.3 (34.6)	99.1 (93.9)	2.70
M+O+ P-1+P-2	21.2 (324.3)	6.4 (139.8)	99.5 (30.3)	99.7 (97.8)	2.60

Table 5.3: Comparison between several heavy-atom derivatised datasets. Addition of a number to a crystal name indicates that several datasets were collected from the same crystal. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor and R_{pim} for the precision-indicating merging R factor. R values were determined with all I+ and I-.

gold-derived metal, gold(I) potassium cyanide, appeared slightly better and enabled to work at the resolution of 4 Å(Figure 5.4A).

Numerous attempts at determining the number of gold atom sites were performed using the automated structure solution pipeline in the CCP4 suite and at different resolution cut-offs (from 5 to 9 Å). All calculations suggested the presence of 6 sites. Unfortunately, the ambiguity of the resulting experimental phases always failed to be resolved.

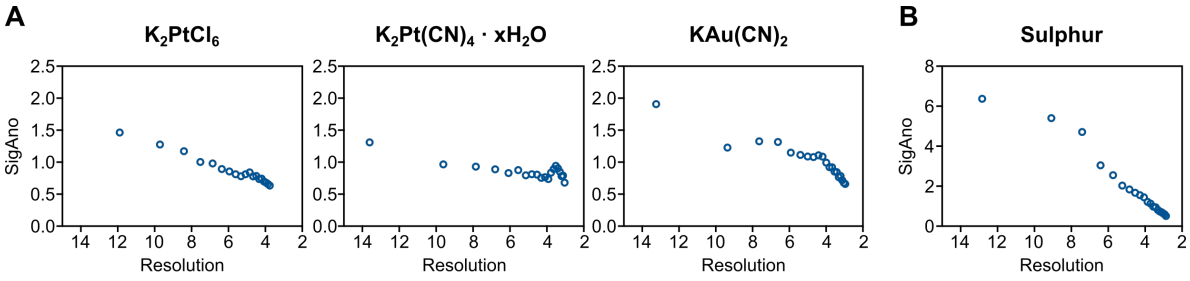


Figure 5.4: Anomalous signal in function of resolution. Anomalous signals from **(A)** various heavy-atoms and from **(B)** sulphur atoms were plotted against resolution.

5.1.4.4 S-SAD phasing

In parallel to heavy-atom derivatisation, native crystals were subjected to S-SAD phasing. As the signal-to-noise ratio of the crystals was generally quite low, a total of 16 full datasets was collected from 4 crystals with the hope that merging multiple datasets from isomorphous crystals would result in an improved diffraction signal.

All datasets were manually integrated using XDS and 12 datasets were scaled and merged together using XSCALE. SHELXC was then used to analyse the data and determined a reliable sulphur anomalous signal could be detected from ~4 Å resolution for downstream calculations (Figure 5.5A). In general, the sulphur anomalous signal from this merged dataset appeared to be much stronger and could be detected at a much higher resolution than the anomalous signal from heavy-atom derivatised datasets (Figure 5.4).

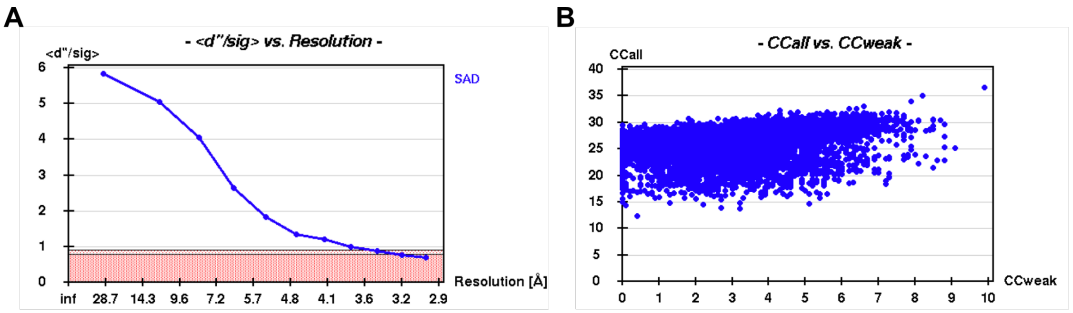


Figure 5.5: Substructure determination of CD177^{ecto} sulphur atoms. **(A)** Analysis by SHELXC indicates a reliable sulphur anomalous signal can be detected from 4 Å resolution. **(B)** Calculations performed by SHELXD shows one solution with high correlation coefficient between the observed reflections (CC_{all}) and those calculated from the substructure found by a given try (CC_{weak}).

Given that CD177^{ecto} possesses 26 cysteines and 9 methionines, 70 sulphur atoms were expected in the ASU. However, as the resolution of the dataset is lower than the length of a disulphide bridge, which is typically around 2.05 Å, it is likely that both sulphur atoms within a disulphide will not be resolved and should instead be considered as one site for a sulphur atom. Thus in theory, 44 sulphur atoms are to be expected in the ASU. Substructure determination was performed by SHELXD, which calculates and plots the correlation coefficient between observed reflections (both all and the weak ones) and those calculated from the substructure found by a given try (CC_{all} and CC_{weak}, respectively). A successful solution is one with a relatively high CC_{all} and CC_{weak} (Figure 5.5B).

AutoSol was then used to refine the coordinates of the sulphur atoms before using those in order to calculate experimental phases. It also carried density modification in order to resolve the ambiguity and provide an initial electron density map (Figure 5.6). The quality of the map was good enough to confidently place six group of β-strands, which seemed to correspond to six LU domains.

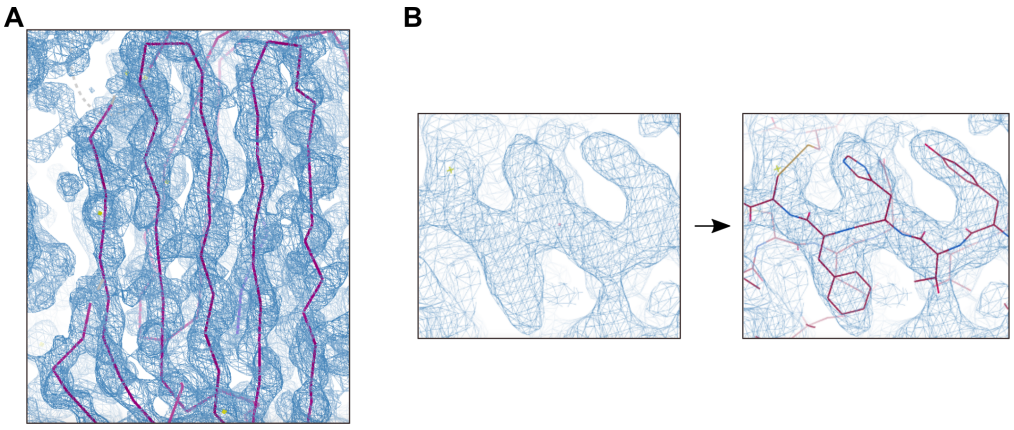


Figure 5.6: Initial electron density map obtained from S-SAD phasing. The 2Fo-Fc electron density map is shown at the 1 sigma level in blue. Yellow stars indicate the position of sulphur atoms. **(A)** A preliminary model of a LU domain was built in the map and is shown in dark red. **(B)** Close-up on one part of the map where the density was good enough to identify and build side chains (shown in dark red).

This starting model was used in another AutoSol run in order to further refine the sulphur sites and to obtain better experimental electron density maps. Each LU domain could then be identified by taking into account the position of disulphide bridges, methionine residues and surrounding side chains. In the end, several iterations of AutoSol and model building enabled to build a partial model of $CD177^{ecto}$ (Figure 5.7A). The model is, however, still incomplete as 353 out of 387 residues have been modelled so far.

5.1.5 $CD177^{ecto}$ structure analysis

The current model of $CD177^{ecto}$ comprises residues 24 to 110 and 132 to 397, which corresponds to most of the protein except for a long loop between D1 and D2, and the last 11 residues (Figure 5.7A, Table 5.4).

Nonetheless, the model confirms the presence of four LU domains spanning the whole receptor ectodomain and reveals these four LU domains are arranged into two compact subdomains, namely subdomain A and subdomain B. Each subdomain comprises two

successive LU domains, namely D1-D2 and D3-D4, linked together by a 20 residue-flexible linker. Each LU domain possesses the characteristic three-finger fold formed by a core of six β-strands stabilised by four disulphide bridges and from which extends three β-stranded loops (Figure 5.7).

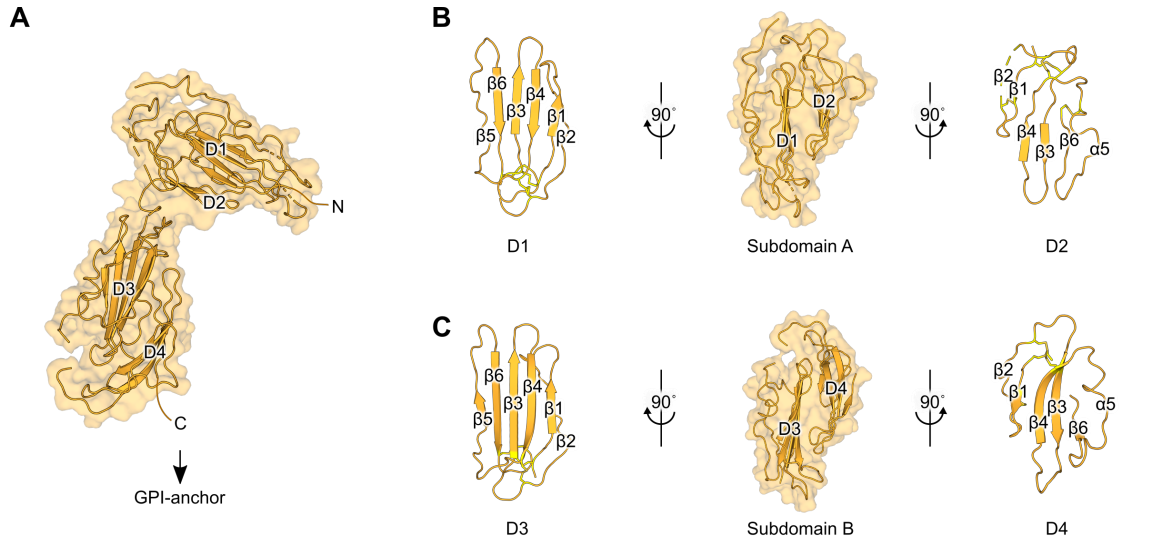


Figure 5.7: CD177^{ecto} crystal structure. (A) Cartoon representation of CD177^{ecto} structure obtained from X-ray crystallography experiments. Each LU domain (D1 to D4) is indicated. **(B,C)** Within each subdomain, two LU domains are facing each other in a head to tail manner. All four LU domains possess the canonical LU domain fold and disulphide bridges (shown in yellow sticks).

Within one subdomain, the last β-strand, β6-strand, of the first LU domain extends into a long loop, which seems to wrap and circle around the LU domain before placing the second LU domain so that both domains face each other in a head to tail manner (Figure 5.7B,C). Both LU domains appear to be tightly packed against each other, forming a globular unit stabilised by an extensive network of hydrophobic interactions (Figure 5.8).

Notably, both CD177^{ecto} protomers in the ASU adopt the same configuration of subdomain A relative to subdomain B, which may suggest a lack of flexibility between the two subdomains despite the presence of a short loop between those two.

On the other hand, structural comparison with a high confidence model calculated by AlphaFold [313] reveals a significant movement of subdomain B with respect to subdomain A, and suggests instead some degree of flexibility between subdomains (Figure 5.9). Nevertheless, structural comparison of individual subdomains showed high similarity between the two models.

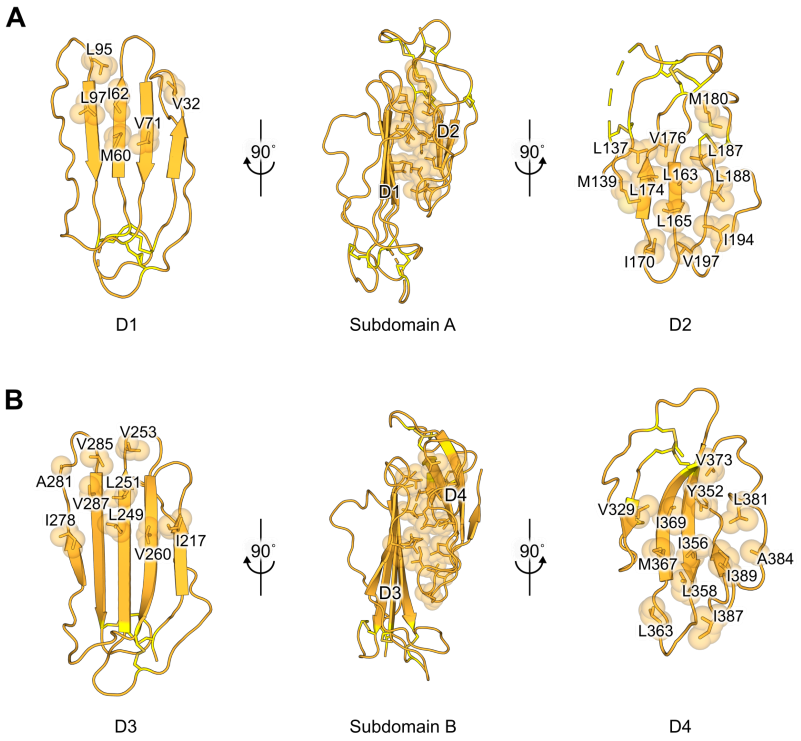


Figure 5.8: Network of hydrophobic interactions stabilising each CD177 subdomain. Packing of two LU domains within a subdomain is stabilised by an extensive network of hydrophobic interactions involving a number of residues from both domains (shown in sticks and spheres).

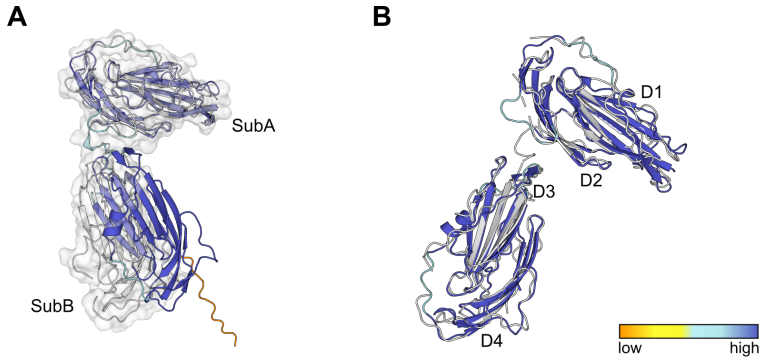


Figure 5.9: Comparison with AlphaFold model. Structural comparison of the model predicted by AlphaFold with CD177^{ecto} crystal structure (white). The AlphaFold model is coloured from very low model confidence (orange) to very high model confidence (blue). **(A)** Alignment of full-length CD177 AlphaFold model (shown in cartoon representation) with full-length CD177^{ecto} crystal structure (shown in cartoon and surface representation). A significant deviation of the position of subdomain B can be observed. **(B)** Subdomains A and B were aligned separately and showed high similarity between the two models.

	Native	**S-SAD**
Data collection		
Source	I03 (DLS)	I23 (DLS)
Wavelength λ (Å)	1.0718	2.7552
Number of crystals	1 (A-1 + A-2)	3
Resolution range (Å)	96.92-2.70 (2.83-2.70)	96.46-2.87 (3.01-2.87)
Space group	H 3	H 3
Cell dimensions		
a, b, c (Å)	193.846, 193.846, 71.851	192.930, 192.930, 71.980
α, β, γ (°)	90, 90, 120	90, 90, 120
Unique reflections	27651 (3693)	22589 (3071)
Multiplicity	7.0 (6.8)	100.8 (56.5)
Completeness (%)	100 (100)	98.9 (92.4)
Wilson B (Å^2)	76.25	85.22
R_{merge} (%)	20.9 (292.5)	21.0 (668.7)
R_{meas} (%)	22.5 (316.7)	21.2 (679.9)
R_{pim} (%)	8.3 (120)	2.9 (118.0)
$CC_{1/2}$ (%)	99.4 (36.3)	100 (56.7)
Average I/σ(I)	4.2 (0.5)	22.2 (1.1)
Refinement		
Resolution range (Å)	38.95-2.7	
Reflections (work / free set)	27358 / 1296	
R_{work} / R_{free} (%)	40.62 / 46.53	
Number of atoms (protein)	5246	
Mean B value (Å^2) (protein)	87.93	
RMSD bonds (Å)	0.007	
RMSD angles (°)	1.04	

Table 5.4: Data collection and initial refinement statistics for CD177^{ecto}. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor, R_{meas} for the multiplicity-corrected merging R factor and R_{pim} for the precision-indicating merging R factor. RMSD is the acronym for root-mean-square deviation from ideal geometry.

5.2 PR3 interaction with CD177

5.2.1 $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystallisation trials

5.2.1.1 Co-purification of $PR3^{IWG\text{-}NGT}$ with $CD177^{ecto}$

Previous attempts at co-purifying PR3 with CD177 always failed to produce a sufficiently pure protein complex to use in crystallisation trials (Figures 4.8C,D and 4.13). Here, the variants $PR3^{K195A\ W222R}$, $PR3^{F181L\ K195A\ W222R\ R227L\ L228Q}$ and $PR3^{IWG\text{-}NGT}$ were expressed in HEK293S GnTI$^-$ cells and harvested either after three days ($PR3^{K195A\ W222R}$ and $PR3^{F181L\ K195A\ W222R\ R227L\ L228Q}$) or ten days ($PR3^{IWG\text{-}NGT}$) before being mixed with half the volume of a culture medium containing secreted $CD177^{ecto}$. After overnight incubation, the resulting mixture was subjected to a standard purification protocol consisting of buffer exchange followed by IMAC and SEC.

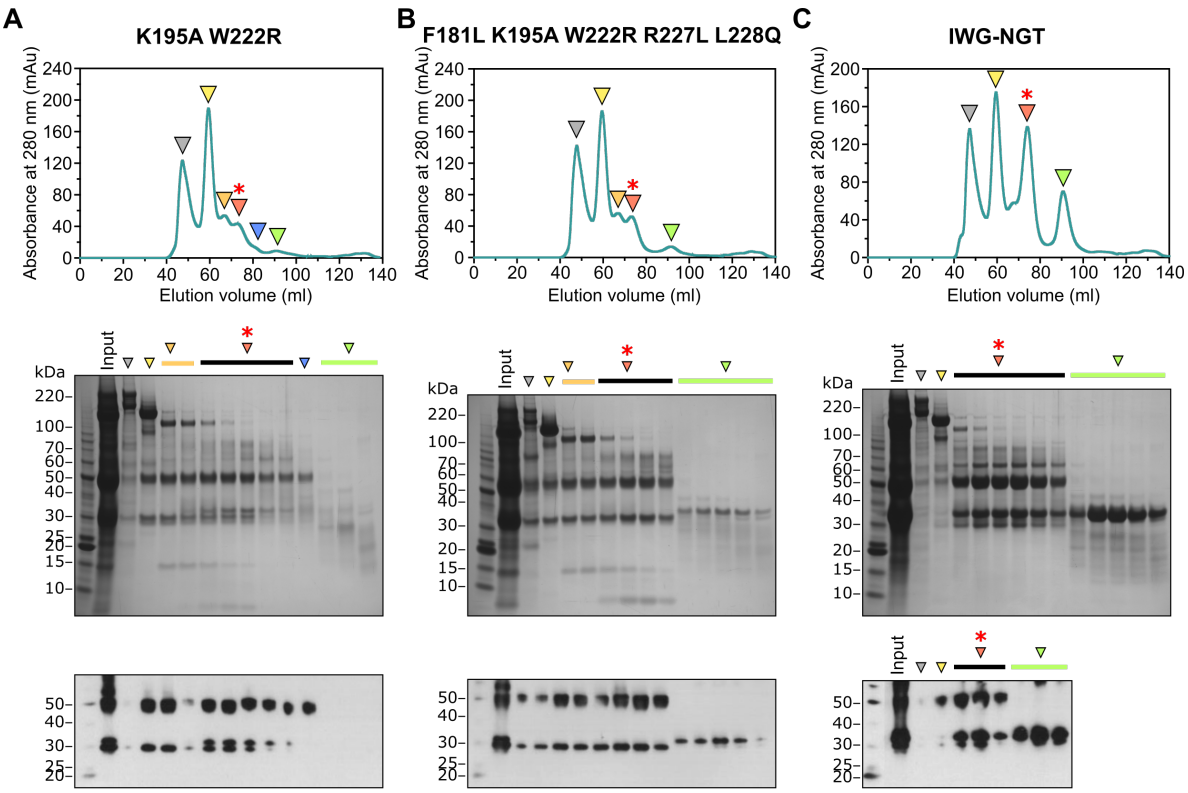


Figure 5.10: Large-scale purifications of PR3-CD177 complex. SEC purification profile with corresponding SDS-PAGE gel and western blot of **(A)** $PR3^{K195A\ W222R}$, **(B)** $PR3^{F181L\ K195A\ W222RR227L\ L228Q}$ and **(C)** $PR3^{IWG\ NGT}$. The complex peak is indicated by a red asterisk and co-elutes along with a 100 kDa impurity. Void and impurity peaks are indicated by grey, yellow, orange and red arrowheads, respectively. Excess of CD177 and PR3 variant are indicated by blue and green arrowheads, respectively.

Although all mutants were able to form stable complexes with CD177, $PR3^{IWG\text{-}NGT}$ yielded the most abundant and pure PR3-CD177 complex (Figure 5.10).
Nevertheless, the PR3-CD177 complex appears to elute along with one of the high molecular weight impurities, which is sub-optimal if the complex is to be used in crystallisation trials.

5.2.1.2 1:1 mixture of $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$

In order to obtain large amounts of pure PR3-CD177 complex, both $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$ were first purified separately before mixing them together at a 1:1 molar ratio for use in crystallisation trials.
This method however relies on both theoretical values of extinction coefficient and molecular weight to calculate the molar concentration of the proteins. In particular, the theoretical molecular weight is unreliable due to the presence of three glycosylation sites on both $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$. Such lack of precision can result in a mixture with a slight excess of one of the two protein partner.
Nevertheless, growing crystals were found in one condition from the Midas screen at 4℃ after 15 days (Figure 5.11). They were harvested at Day 40 and a full dataset was collected and autoprocessed to ∼8 Å resolution.

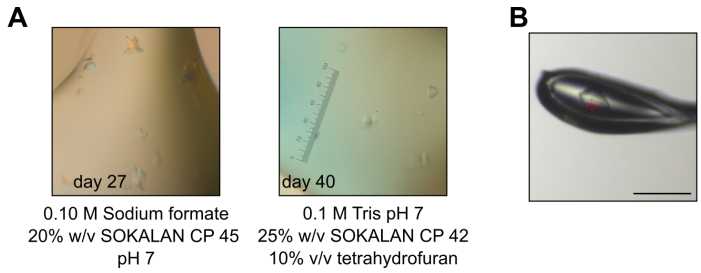


Figure 5.11: Crystal hits obtained from 1:1 mixture of $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$. (A) $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystals growing in conditions from the Midas crystallisation screen. **(B)** View of one crystal cryo-protected in 25% glycerol and mounted onto a SPINE pin prior to data collection. Scale bar is 100 µm.

Despite the low resolution of the dataset, molecular replacement was performed to search for copies of PR3 and CD177, and determine the structure of the binding interface. Unfortunately, this was challenging due to the uncertainty of the space group to which the dataset belongs to. Indeed, it was impossible to determine and choose into which

lattice the dataset had to be indexed as they all possessed similar root mean square deviation (RMSD) and correlation coefficient values (Table 5.5).

Sol	Metric fit	RMSD	Min/Max CC	#spots	Lattice	Unit cell
22	0.145	0.065	0.631/0.707	1651	cP	197.84 197.84 197.84 90.00 90.00 90.00
21	0.145	0.063	0.664/0.670	1654	hR	279.75 279.75 342.15 90.00 90.00 120.00
20	0.145	0.063	0.631/0.647	1636	hR	279.52 279.52 342.69 90.00 90.00 120.00
19	0.145	0.062	0.649/0.707	1640	tP	197.67 197.67 197.48 90.00 90.00 90.00
18	0.145	0.062	0.692/0.707	1639	oC	279.57 279.56 197.50 90.00 90.00 90.00
17	0.145	0.06	0.702/0.702	1628	mC	279.37 279.35 197.38 90.00 90.09 90.00
16	0.1189	0.065	0.659/0.665	1651	hR	279.60 279.60 341.99 90.00 90.00 120.00
15	0.1189	0.066	0.655/0.668	1668	hR	279.36 279.36 342.67 90.00 90.00 120.00
14	0.1249	0.065	0.597/0.703	1651	tP	197.87 197.87 197.83 90.00 90.00 90.00
13	0.1249	0.06	0.615/0.690	1624	tP	197.35 197.35 197.66 90.00 90.00 90.00
12	0.122	0.062	0.660/0.714	1631	oC	278.45 278.84 197.42 90.00 90.00 90.00
11	0.1249	0.06	0.690/0.707	1625	oP	197.20 197.16 197.49 90.00 90.00 90.00
10	0.1189	0.062	0.692/0.692	1638	mC	279.12 279.17 197.20 90.00 90.09 90.00
9	0.1249	0.06	0.707/0.707	1626	mP	197.21 197.19 197.51 90.00 90.00 90.00
8	0.122	0.062	0.690/0.690	1631	mP	197.18 197.58 197.19 90.00 90.08 90.00
7	0.1024	0.063	0.647/0.703	1655	oC	279.03 279.71 197.47 90.00 90.00 90.00
6	0.0927	0.06	0.714/0.714	1644	mC	278.99 278.76 197.51 90.00 89.91 90.00
5	0.1024	0.063	0.693/0.693	1660	mC	279.72 279.05 197.47 90.00 90.01 90.00
4	0.0997	0.063	0.647/0.647	1657	mC	279.02 279.73 197.47 90.00 89.99 90.00
3	0.0793	0.059	0.660/0.660	1612	mC	278.50 278.73 197.40 90.00 90.10 90.00
2	0.0299	0.058	0.703/0.703	1616	mP	197.10 197.10 197.39 90.00 90.13 90.00
1	0	0.058	-/-	1625	aP	197.02 197.02 197.36 90.12 90.01 90.03

Table 5.5: Challenging indexing during data processing.

As a result, the data was manually indexed into all of the suggested lattices (aP, mP, mC, oP, oC, tP, hR and cP), integrated using DIALS and scaled and merged using AIMLESS. Molecular replacements were then performed with each of the differently-processed dataset, but none of the tests yielded promising results.

5.2.1.3 SEC-purified PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ complex

An additional SEC purification step was added in order to make sure the PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ was as pure as possible, with no other protein contaminants (Figure 5.12A). Additionally, a SEC-MALS experiment confirmed the PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ complex exists as a 1:1 complex (Figure 5.12B).

Growing crystals could be observed in the same crystallisation conditions where the 1:1 mixture of PR3$^{\text{IWG-NGT}}$ and CD177$^{\text{ecto}}$ yielded small crystals (Figure 5.12C). Notably,

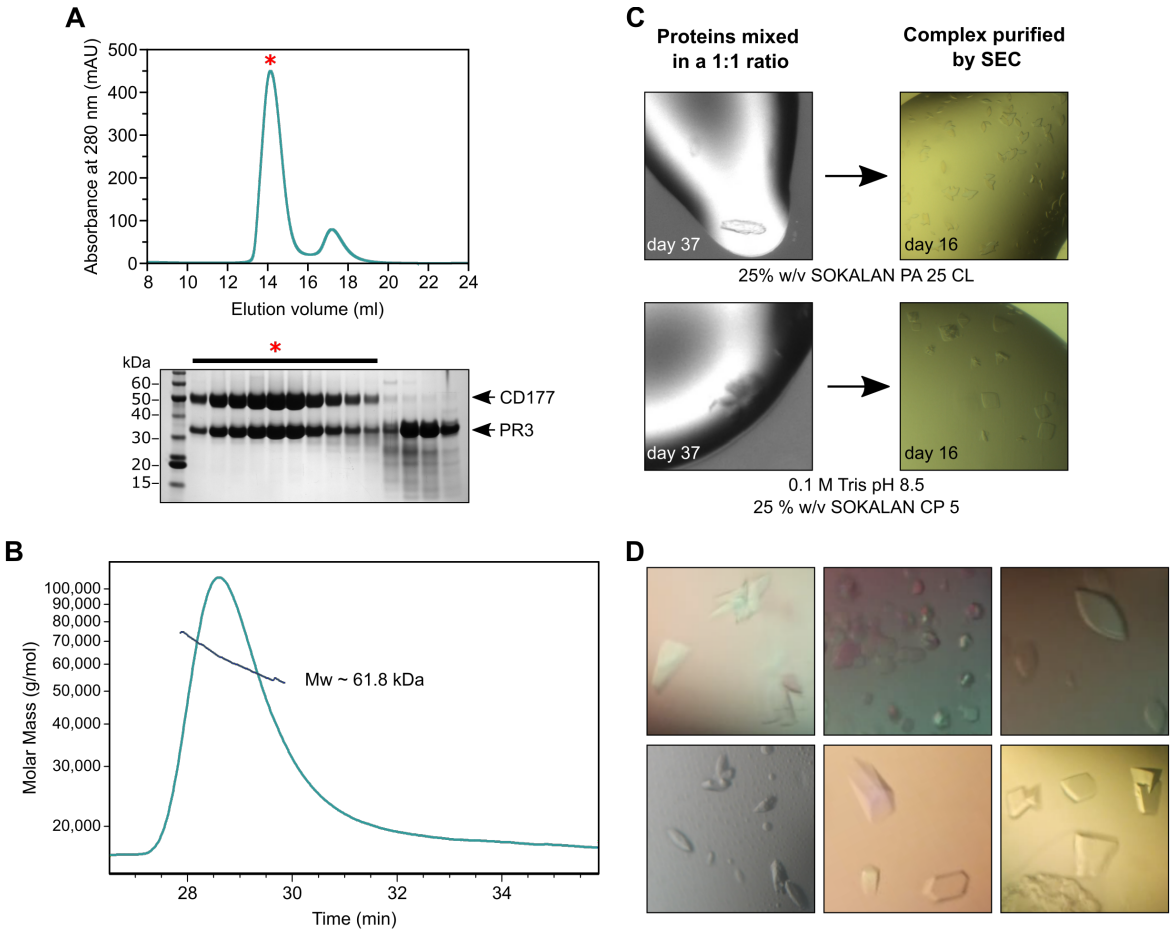


Figure 5.12: Production of pure PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ complex. **(A)** SEC profile of PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ complex with corresponding SDS-PAGE gel. **(B)** SEC-MALS experiment profile showing PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ is a 1:1 complex. **(C)** Optimised crystals obtained with SEC-purified PR3$^{\text{IWG-NGT}}$-CD177$^{\text{ecto}}$ complex. **(D)** Crystal optimisation led to the obtention of crystals of various sizes and shapes. Images were taken at day 13.

the newer crystals look sharper and grew within two to four days while the older ones only appeared after three weeks.

Because none of those crystals managed to diffract at a higher resolution than 20-25 Å, several optimisation screens were set-up (Table 5.6). In particular, the microseed matrix screening yielded numerous crystals of different size and shape (Figure 5.12D). Unfortunately, none of those crystals managed to diffract at a good resolution.

Concentration	**Crystallisation screen**
Using purified $PR3^{IWG\text{-}NGT}$ and $CD177^{ecto}$ mixed at a 1:1 ratio	
7.3 mg/ml (106 µM) (1:1 and 2:1 protein:reservoir ratio)	Index, JCSG I-IV, Midas, Morpheus, Morpheus II, PACT, PGA, ProPlex Index, JCSG+, Midas, Morpheus II, ProPlex at 4℃
Using SEC-purified $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ complex	
7.3 mg/ml (106 µM)	JCSG+, Midas, Morpheus, PACT, ProPlex
	Optimisation screens derived from hits condition:
7.3 mg/ml (106 µM)	Additive screen and Silver bullets on 0.1 M sodium formate, 20 % w/v SOKALAN CP 45 pH 7.0 Additive screen and Silver bullets on 25 % w/v SOKALAN PA 25 CL Additive screen and Silver bullets on 0.1 M Tris pH 8.5, 25 % w/v SOKALAN CP 5
7.3 mg/ml (106 µM)	Additive screen and Silver bullets on 0.1 M Tris pH 7, 25 % w/v SOKALAN CP 42, 10 % v/v tetrahydrofuran at 4℃
	Microseed matrix screening into:
7.3 mg/ml (106 µM)	Index, JCSG+, Midas, Morpheus, ProPlex

Table 5.6: Summary of $PR3^{IWG\text{-}NGT}$-$CD177^{ecto}$ crystallisation trials.

5.2.2 Design of shorter PR3-binding CD177 constructs

Two shorter CD177 constructs consisting of subdomain A (1-206) and subdomain B (209-408) were designed and tested for their binding to PR3. Narrowing the PR3 binding site to one subdomain would enable to perform new crystallisation trials and enhance the chances of obtaining more compact crystal configurations leading to higher resolution datasets.

5.2.2.1 SPR binding assays

SPR experiments were performed with both $CD177^{subA}$ and $CD177^{subB}$ immobilised on the sensor chip, and with $PR3^{IWG\text{-}NGT}$ as the solute analyte. Strikingly, $CD177^{subA}$ showed a binding affinity similar to that of $CD177^{ecto}$, with K_d values of 18 nM and 10.8 nM, respectively (see Section 4.4.6.2 of Chapter 4). By contrast, $CD177^{subB}$ displayed no binding at all, suggesting that the PR3 binding site is exclusively located on subdomain A of CD177.

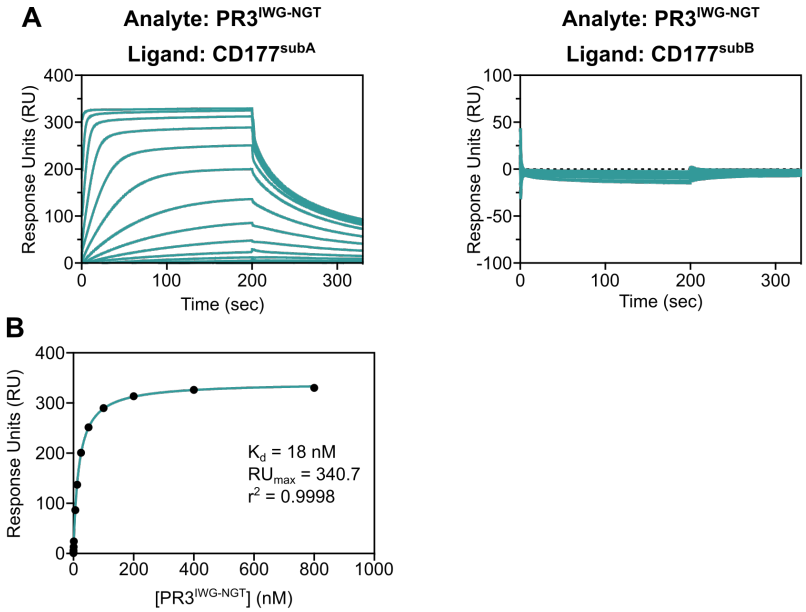


Figure 5.13: SPR analysis of PR3$^{\text{IWG-NGT}}$ interaction with CD177 constructs. CD177$^{\text{subA}}$ and CD177$^{\text{subB}}$ were immobilised at a surface density of 340 and 320 RU, respectively. For K_d measurements, the binding signal was measured for 15 two-fold serial diluted samples of PR3$^{\text{IWG-NGT}}$, with concentration ranging from 0.05 to 800 nM.

5.2.2.2 CD177$^{\text{subA}}$ production

CD177$^{\text{subA}}$ was produced in HEK293S GnTI^{-} cells and purified following a standard purification protocol. The protein could be produced with a yield of 1.3 mg/L of culture and purified to >95% purity (Figure 5.14).

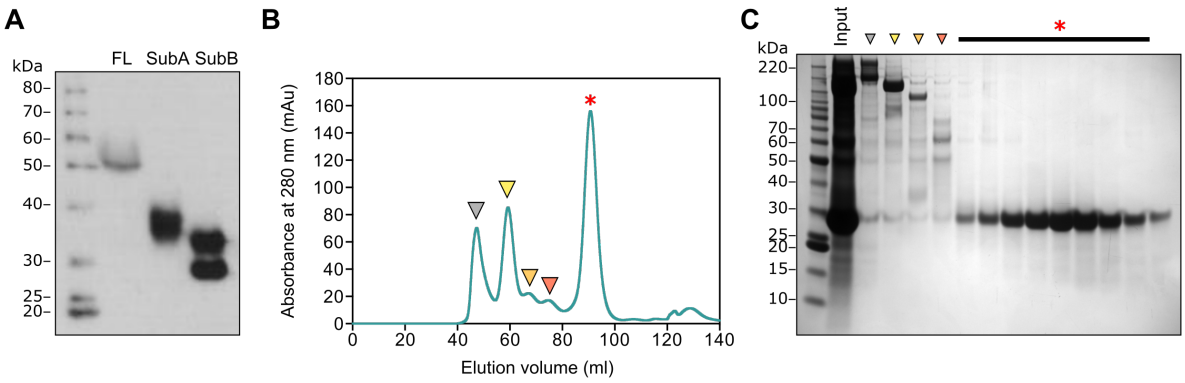


Figure 5.14: Large-scale purification of CD177$^{\text{subA}}$. (A) Expression trials of CD177$^{\text{subA}}$ and CD177$^{\text{subB}}$ constructs. **(B, C)** SEC purification profile with corresponding SDS-PAGE gel of CD177$^{\text{subA}}$. The protein elutes at around 90 ml on a Superdex® 200 column (red asterisk) and is preceded by void (grey arrowhead) and impurity peaks (yellow, orange and red arrowheads).

5.2.2.3 PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ crystallisation trials

Purified PR3$^{\text{IWG-NGT}}$ was mixed with purified CD177$^{\text{subA}}$ at a 1:1 ratio, and was used to set up a wide range of crystallisation trials (Table 5.7).

Concentration	Crystallisation screen
8.3 mg/ml (175 µM)	Index, JCSG I-IV, Midas, Morpheus, Morpheus II, PACT, PGA, ProPlex

Table 5.7: Summary of PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ crystallisation trials.

Strikingly, many crystals appeared in a broad range of conditions from multiple crystallisation screens (Figure 5.15) and enabled the collection of five full datasets which could be processed to 1.7-2.5 Å resolution.

Autoprocessing by XDSme produced three orthorhombic (P 2_1 2_1 2_1) datasets at 2.3-2.5 Å resolution and two monoclinic (P 1 2_1 1) datasets at ∼1.7 Å resolution (Table 5.8). The monoclinic dataset U was chosen as the best one, reprocessed using DIALS to 1.50 Å resolution and was used to solve the structure of the complex.

	Statistics when $CC_{1/2}$ is cut to 0.3				
	R_{merge} (%)	R_{pim} (%)	$CC_{1/2}$ (%)	Completeness (%)	Resolution (Å)
Space group P 2_1 2_1 2_1					
Q	26.1 (477.4)	7.3 (134.4)	99.5 (29.6)	100 (100)	2.30
R	19.0(135.2)	5.3 (37.8)	99.7 (83.1)	99.8 (98)	2.35
S	35.7 (421.6)	11.0 (134.2)	98.7 (42.3)	100 (100)	2.50
Space group P 1 2_1 1					
T	13.0 (336.1)	5.4 (139.9)	99.7 (31.1)	100 (100)	1.85
U	9.2 (86.3)	3.7 (35.1)	99.8 (65.9)	99.6 (93.4)	1.74
U-reprocessed	9.3 (189)	3.9 (78.7)	99.8 (65.2)	99.9 (99.9)	1.50

Table 5.8: Comparison between several PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ datasets. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor and R_{pim} for the precision-indicating merging R factor. R values were determined with all I+ and I-.

5.2.3 PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ structure determination

Estimation of cell content indicated a probability of 99% of having one copy of PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ in the ASU, as well as a corresponding solvent content of 60%.

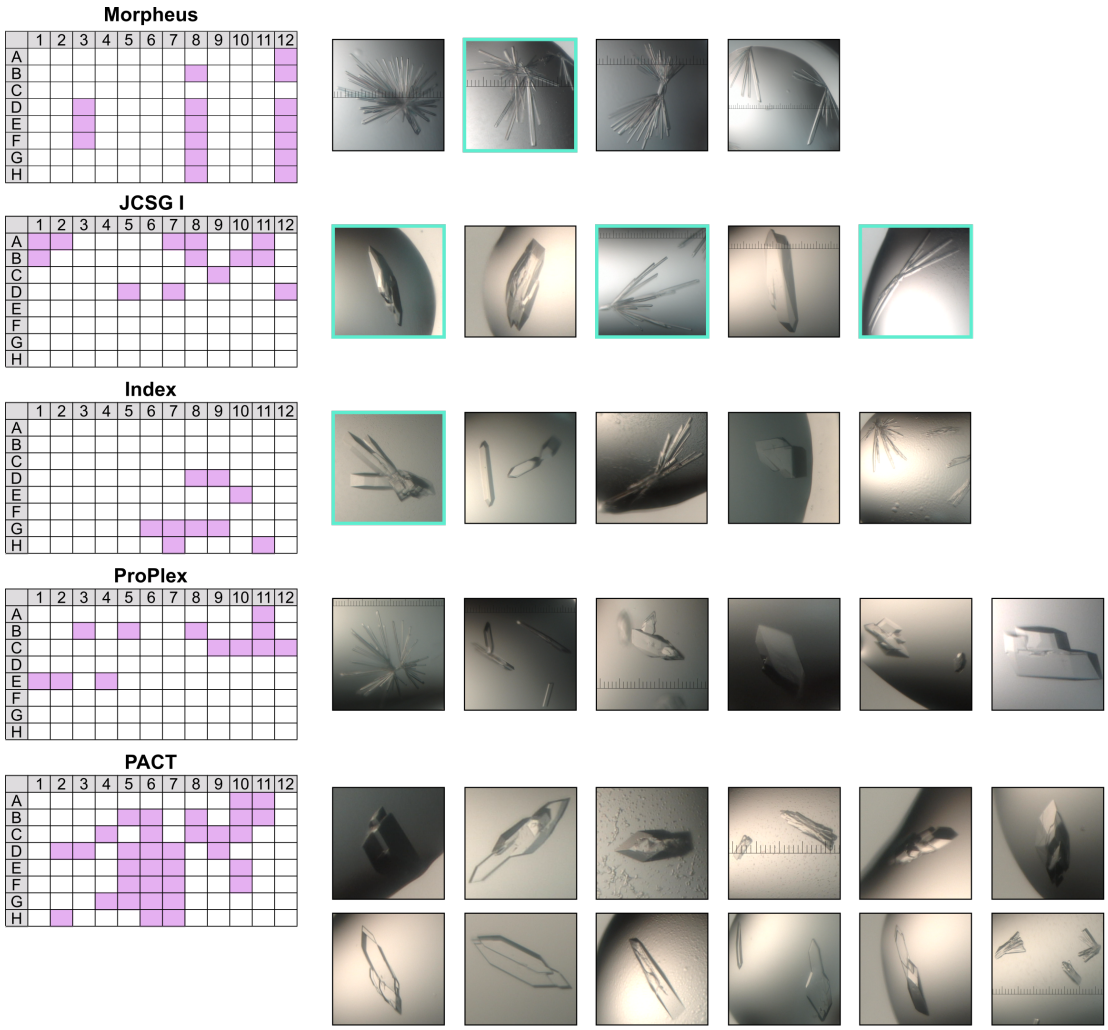


Figure 5.15: PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ crystals. High-resolution datasets were collected from crystals highlighted in blue.

Molecular replacement was performed using the available crystal structure of PR3 [319] and the partial model of CD177$^{\text{subA}}$ built in section 5.1.4.4. This resulted in one successful solution consisting of one copy of the complex in the ASU. Given the high quality of the electron density map, Bucaneer was used to build an improved model of the complex. The resulting model comprises residues I28 to P256 for PR3$^{\text{IWG-NGT}}$ and L22 to C203 for CD177$^{\text{subA}}$ (Figure 5.16A, Table 5.9). Notably, clear electron density accounting for *N*-glycans at CD177 N189, PR3 N129, N174, and more importantly N221 could be observed.

Data collection	
Source	PX-1 (SOLEIL)
Wavelength λ (Å)	0.9786
Number of crystals	1
Resolution range (Å)	60.37-1.50 (1.53-1.50)
Space group	$P12_11$
Cell dimensions	
a, b, c (Å)	39.760, 120.735, 61.983
α, β, γ (°)	90, 96.992, 90
Unique reflections	92506 (4607)
Multiplicity	6.6 (6.7)
Completeness (%)	99.9 (99.9)
Wilson B ($Å^2$)	19.23
R_{merge} (%)	9.3 (189.0)
R_{meas} (%)	10.1 (205.2)
R_{pim} (%)	3.9 (78.7)
$CC_{1/2}$ (%)	99.8 (65.2)
Average I/σ(I)	8.6 (0.7)
Refinement	
Resolution range (Å)	43.13-1.50
Reflections (work / free set)	87775 / 4684
R_{work} / R_{free} (%)	17.0 / 21.1
Number of atoms (protein)	5976
Number of atoms (water)	441
Mean B value ($Å^2$) (protein)	27.65
RMSD bonds (Å)	0.008
RMSD angles (°)	1.54

Table 5.9: Data collection and initial refinement statistics for $PR3^{IWG\text{-}NGT}$-$CD177^{subA}$. Values in parenthesis are for the highest resolution shell. R_{merge} stands for the merging R-factor, R_{meas} for the multiplicity-corrected merging R factor and R_{pim} for the precision-indicating merging R factor. RMSD is the acronym for root-mean-square deviation from ideal geometry.

5.2.4 Improvement of $CD177^{ecto}$ model

In parallel, $CD177^{subA}$ was used to produce a homology model of $CD177^{subB}$ by using MODELLER. A full model $CD177^{ecto}$ could then be assembled and was manually improved in Coot using the experimental electron density map obtained previously. One round of refinement using this model produced R_{work} and R_{free} values of 37.75 and 42.31, respectively. Those are significantly improved compared to the values obtained with manual building in section 5.1.5, and therefore suggest to use this model for further refinement cycles (Table 5.10).

Refinement	
Resolution range (Å)	38.95-2.7
Reflections (work / free set)	27358 / 1296
R_{work} R_{free} (%)	37.75 / 42.31
Number of atoms (protein)	5570
Mean B value (Å^2) (protein)	96.69
RMSD bonds (Å)	0.008
RMSD angles (°)	1.10

Table 5.10: Improved refinement statistics for CD177^{ecto}. Values in parenthesis are for the highest resolution shell. RMSD is the acronym for root-mean-square deviation from ideal geometry.

5.2.5 Analysis of the PR3-CD177 binding interface

The crystal structure revealed a 655.1 Å^2 PR3-CD177 interface which mainly involves residues from the PR3 hydrophobic patch and from the long CD177 linker connecting D1 to D2 in subdomain A (Figure 5.16A).

The binding interface appears to be governed by an extensive network of hydrophobic interactions driven by F180, F181, F190 and L228 from PR3 and P39, L117, P118 and W120 from CD177 (Figure 5.16B). Additional hydrophobic interactions involve H148, V178, P192, R193, R227, F229, P230 from PR3 and V86, I98, Y100, L113, V114, P122 from CD177 (Figure 5.16C). Additionally, three hydrogen bonds formed by residues located at the rim of the interface (R193 and L228 from PR3 with L119 and N115 of CD177, respectively) further stabilise the complex.

Although the residues mutated in PR3$^{IWG\text{-}NGT}$ are located near the interface, they do not appear to be directly involved in the formation of the complex, therefore explaining why PR3$^{IWG\text{-}NGT}$ can still bind CD177 with high affinity.

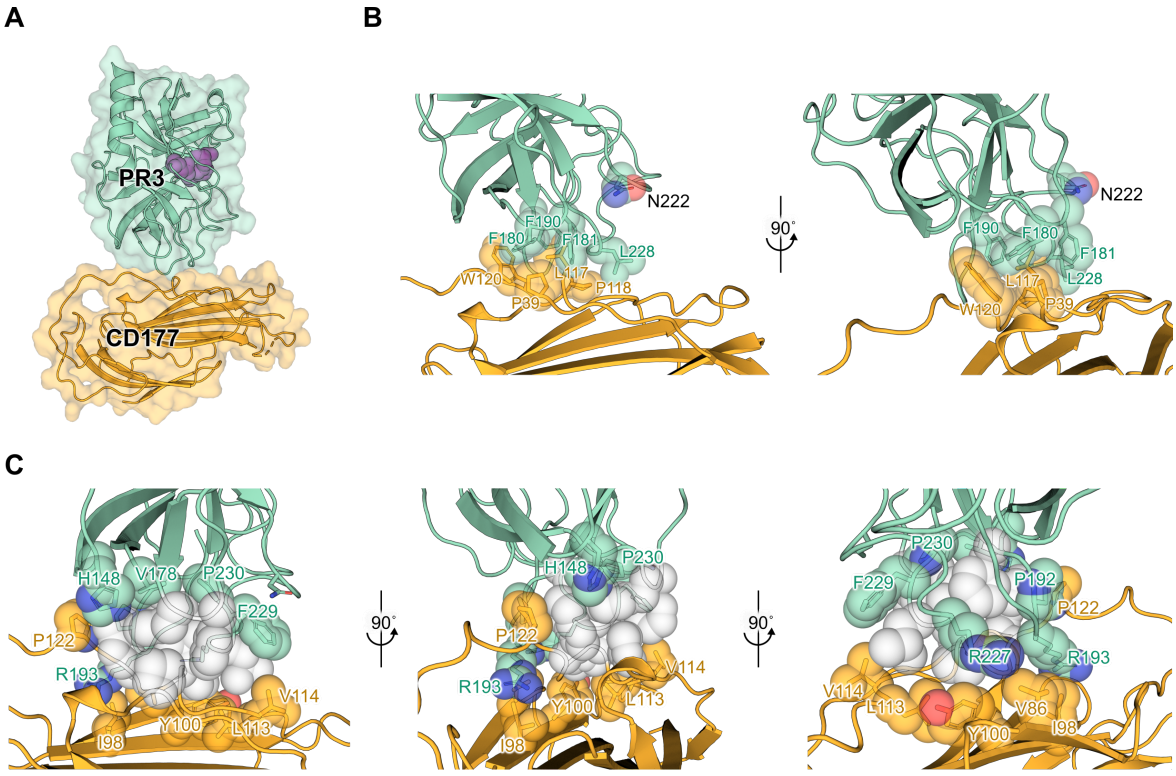


Figure 5.16: Crystal structure of PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ complex. (A) Cartoon representation of PR3$^{\text{IWG-NGT}}$-CD177$^{\text{subA}}$ structure obtained from X-ray crystallography experiments. PR3$^{\text{IWG-NGT}}$ and CD177$^{\text{subA}}$ are depicted in green and orange, respectively. PR3 catalytic triad is depicted in purple spheres. **(B)** Core residues forming the PR3-CD177 binding interface are depicted in sticks and spheres. The I221N mutation is also shown. **(C)** The binding interface is further stabilised by an extensive network of hydrophobic interactions surrounding the core residues. Stabilising residues are depicted in sticks and spheres, core residues are depicted as white spheres.

5.3 Discussion

5.3.1 CD177$^{\text{ecto}}$ reveals a tandem arrangement of LU domains

As a member of the Ly6/uPar family, CD177 was initially predicted to have two LU domains. In fact however, sequence analysis suggested four consecutive LU domains spanning the whole receptor ectodomain. The crystal structure confirms the presence of these four LU domains, which are arranged into two compact subdomains, each comprising two successive LU domains linked together by a 20 residue-flexible linker. A typical LU domain contains five conserved disulphide bridges, namely the 1-5, 2-3, 4-6, 7-8 and 9-10 disulphide bonds. However, CD177 D1 and D3 only possess three conserved disulphide bridge and lack both the 2-3 and 7-8 disulphide bonds. CD177 D4 also only possesses three conserved disulphide bridge but lacks the last two 7-8

and 9-10 disulphide bonds instead. While the loss of the 7-8 disulphide bond has been previously observed in the most N-terminal LU domain of multi-LU domain-containing proteins such as uPAR, TEX101 and C4.4, it is much less common to lose the 2-3 and 9-10 disulphide bonds (Figure 5.1D).

Additionally, CD177 is the first and only member of the Ly6/uPAR family that contains four LU domains and that has been crystallised. Each CD177 subdomain is composed of two successive LU domains that are linked together by a long flexible linker. Strikingly, the linker wraps and circles around the first LU domain before placing the second LU domain so that the domains are facing each other in a head to tail manner. The closeness and extensive network of hydrophobic interaction between these two consecutive domains probably explain why molecular replacement was challenging and failed to properly place homology models.

At the time, such a configuration of LU domains was completely novel and unprecedented. However, recently deposited crystal structures of TEX101 and C4.41, which are two other members of the Ly6/uPar family possessing two LU domains, display a strikingly similar arrangement of their two LU domains [352, 353]. Notably, both structures required either the use of a Fab fragment [352] or heavy-atom derivatised crystals [353] in order to obtain phase information and solve the structure. The DALI web server calculated TEX101 as being the model most similar to $CD177^{subA}$ with 35% sequence identity and an RMSD of 1.7 Å [354]. In comparison, C4.4A shows 19% sequence identity and an RMSD of 3.1 Å. Finally, mouse CD177 is surprisingly different to its human counterpart with an ECD that is twice as big and comprising no less than eight predicted LU domains. A sequence analysis suggests those eight LU domains are arranged into four subdomains, with the last two subdomains being most similar to those of human CD177. Whether mouse CD177 retains the same biological function as human CD177, or possesses new roles via its additional two subdomains has not been studied so far and remains to be elucidated.

5.3.2 Flexibility between subdomains A and B

Both protomers of CD177 in the ASU adopt the same configuration of subdomain A with respect to subdomain B, suggesting a locked configuration of the two subdomains.

However, this may also be a structural artefact as the presence of a seven-residue linker (NRKDFLT) bridging both subdomains could allow for inter-subdomain flexibility in solution. Further small-angle X-ray scattering experiments or molecular dynamics simulations could be powerful techniques to investigate this.

Additionally, flexibility between both subdomains may explain why molecular replacement using this CD177^{ecto} model failed to solve the low resolution dataset obtained from PR3$^{IWG\text{-}NGT}$-CD177^{ecto} crystals.

5.3.3 PR3-CD177 binding interface

The functional diversity of the three-finger toxins have been mostly connected with the variability of the loop regions. This was the case for the snake toxins α-neurotoxins which inhibit the nicotinic acetylcholine receptors (nAChRs) [355], or with micrurotoxin1 (MmTX1) and MmTX2 which increase the sensitivity of γ-aminobutyric acid receptors ($GABA_A$ receptors) to agonists [356].

By contrast, mammalian LU domain-containing proteins display a much broader range, and therefore variability, of interactions with their partners. Indeed, the binding interface of TGF-βR1, CD59 and GPIHBP1 with their interacting partner TGF-β3 [357], vaginolysin [358] and lipoprotein lipase [359], respectively, are all located at different sites of their LU domain. In the case of uPAR, its ligand uPA inserts into the deep cavity formed by the three consecutive LU domains of the receptor [360].

Here, a new mode of interaction is presented, whereby PR3 mainly interacts with a long interdomain loop rather than directly with a LU domain of CD177. The PR3-CD177 binding interface primarily involves the hydrophobic residues F180, F181, F190, L228 from PR3, and P39, L117, P118, W120 from CD177. This explains why the PR3$^{IWG\text{-}NGT}$ variant as well as all the PR3 mouse mutants retained their CD177 binding ability as the F181L, K195A, W222R, R227L and L228Q mutations would hardly disrupt the binding interface. This also supports the decrease of binding observed when using PR3$^{F180S\ F181D\ W222R\ L228H\ F229S}$ and PR3$^{TFF\text{-}NGT}$ variants where the key F180, F181, L228 residues were mutated to polar and charged residues.

Interestingly, the crystal structure appears to contradict one previous study which suggested a role for W222 in complex formation [124]. Indeed, the study reported a seven-fold decrease in PR3-CD177 binding affinity upon W222A mutation. This is, however, not supported by the current PR3-CD177 model as W222 does not appear to be directly involved in the binding interface. One possible explanation for the difference of binding affinity observed between WT PR3 and $PR3^{W222A}$ could lie in the oligomerisation ability of WT PR3. Indeed, both CD177 molecules brought together by a dimeric PR3 could potentially interact together and result in a higher affinity 2:2 PR3-CD177 complex. However, further investigation requires the design of new PR3 variants displaying both dimerisation and CD177 binding abilities.

Chapter 6
Discussion

6.1 Identification of new signalling interactome in human neutrophils

A novel PR3/CD177/GPR97/PAR2/CD16b interactome was identified to be a critical modulator of neutrophil activation upon GPR97 interaction. More particularly, a new aGPCR-GPCR transactivation mechanism was proposed to direct neutrophil activation, in which GPR97 interaction with PR3 presented by the interactome promotes PR3 proteolytic activity, allowing it cleave and ultimately activate PAR2 [256]. Remarkably, the presence of PAR2, CD16b and another molecule of GPR97 was required for efficient binding of GPR97 to CD177-bound PR3 leading to signal transduction. While PAR2 appears to be the signal transducer, the roles of CD16b and GPR97 in this signalling complex remain unknown. CD177 may also serve other roles in addition to being a simple receptor presenting PR3 on the cell surface.

There have been contradicting studies on the proteolytic activity of CD177-bound PR3 [36, 124]. It is, therefore, unknown whether CD177-bound PR3 is catalytically active or not on neutrophil surface. Nevertheless, interaction with GPR97 appears to potentiate its enzymatic activity. The receptor may either allosterically modulate or reorient PR3 on the cell surface, resulting in a catalytic cleft more exposed to its environment and to protein substrates such as PAR2. Alternatively, if PR3 exists as dimers on the cell surface, binding of GPR97 may disrupt the dimer interface, which would all the same

result in exposure of the hidden catalytic cleft. However, the regulatory mechanisms preventing GPR97 from activating all neutrophils remain to be elucidated. It is also unknown whether GPR97 interacts in *cis* or *trans* with the interactome.

Interestingly, the main role of CD177-bound PR3 in this signalling complex indicates this interactome only exists on $CD177^{pos}/PR3^{high}$ neutrophils. Therefore, there must exist another signalling complex performing a redundant role on $CD177^{neg}/PR3^{low}$ neutrophils. As GPR97 failed to activate $CD177^{neg}$ neutrophils, such a signalling complex would not only involve other mediators than CD177 and PR3, but would also be activated upon binding of another protein than GPR97.

Previous studies suggested PR3/CD177/Mac-1 to be another signalling complex leading to neutrophil activation [77]. Mac-1 is, however, not involved in this interactome, which therefore suggests those two signalling complexes to be independent. More specifically, Mac-1 was presumed to be the signal transducer leading neutrophil activation upon PR3 recognition by ANCAs. It would thus be interesting to see whether binding of ANCAs to PR3 would also result in neutrophil activation through CD177/GPR97/PAR2/CD16b interactome.

Finally, the presence of so many protein partners in this interactome implies this pathway is tightly regulated. This also presents many potential therapeutic targets for neutrophil-mediated inflammatory diseases. In particular, GPR97 is not only the main mediator of this signalling complex, but is also overexpressed in neutrophils of diverse inflammatory disorders. Therefore, preventing GPR97 from interacting with PR3 may be relevant for the treatment of neutrophil-mediated inflammatory diseases.

6.2 The role of CD177-PR3 in pathogenesis of GPA

As PR3 recognition by ANCAs is the main leading cause of pathogenesis of GPA, knowing how PR3 is presented by CD177 on the cell surface brings crucial molecular insights into the disease. So far, a number of studies have attempted to map PR3 epitopes targeted by ANCAs by using small amounts of recombinantly-produced proteins in immunoassays such as western blotting and ELISA. However, proper biophysical and structural studies have been hampered by the aggregation-prone nature of native PR3

and by the large variety of ANCAs. Indeed, ANCAs are polyclonal antibodies, which means they all bind to different epitopes on PR3 [361, 329, 362, 363]. Therefore, in addition to the challenge of obtaining large amounts of PR3, there is also the technical challenge of isolating all ANCAs before studying them individually.

In this context, both the crystal structure of PR3-CD177 and the production of a highly soluble and monomeric PR3 variant bring dramatic advances to understanding the molecular mechanisms underlying PR3 surface presentation. CD177 binds PR3 on its hydrophobic patch, and although this leaves PR3 catalytic cleft free and exposed to the extracellular environment, the possibility of a 2:2 PR3-CD177 complex burying the catalytic cleft may explain the absence of activity observed in some studies. Nevertheless, using both the 1:1 and 2:2 PR3-CD177 complex models may help better understand and identify the functionally relevant PR3 epitopes which are targeted by ANCAs. Additionally, it will now be possible to collect peripheral blood mononuclear cells (PBMCs) from GPA patients and use $PR3^{IWG\text{-}NGT}$ to select those expressing ANCAs for subsequent single-cell sequencing. This will allow recombinant production of individual ANCAs and structural characterisation of their binding site on PR3.

Given that 57-95% of GPA patients with active disease show presence of PR3-ANCAs [364, 334] and that ANCA titre positively correlates with disease activity [365, 75], patients displaying clinicopathological symptoms belonging to GPA have been routinely tested for presence of PR3-ANCAs. However, the observation that some patients in acute disease display low ANCA titre while some patients in remission display high ANCA titre led to the supposition that some ANCAs are more pathogenic than others. This was confirmed by several studies which showed that the nature and specificity of ANCAs evolve along with the disease [362, 363], but also that some ANCAs are non-pathogenic as they do not trigger neutrophil activation [366]. Therefore, characterisation and comparison of ANCAs between patients in acute disease and those in remission might identify pathogenic epitopes on PR3.

Additionally, using PR3-CD177 instead of just PR3 in ANCA-detection assays might result in the identification of a new population of ANCAs which would have remained

undetected. This could be particularly important for patients who are tested ANCA-negative but who present all clinicopathological symptoms belonging to GPA. Such a case has been previously reported, when the presence of a serum component was masking a crucial and particularly pathogenic epitope targeted by MPO-ANCAs on MPO [367]. In the case of GPA and PR3-ANCAs, hidden epitopes could include the binding interface of PR3-CD177, the dimerisation and/or oligomerisation interfaces of PR3 or, on the contrary, the faces on PR3 which are buried upon oligomerisation.

Despite its relatively low incidence and prevalence (0.4-11.9 cases per one million people per year and 2.3-146 cases per one million people, respectively [72]), GPA represents a major challenge as it is difficult to recognise, and early full-blown cases have a poor prognosis, especially if treated inadequately [368]. There is therefore a need for proper diagnosis tools and the use of an array of PR3 variants with different properties, such as $PR3^{\text{IWG-NGT}}$ and $PR3^{\text{F180S F181D L228H F229S}}$, may be useful for improving patients' diagnosis.

APPENDICES

Appendix A

Supplementary Materials & Methods

Reagents and antibodies

All chemicals and reagents were purchased from Sigma (St. Louis, MO, USA) and Invitrogen (CA, USA) unless specified otherwise. E-64 (Cat.#78434) was from Thermo Fisher Scientific. Human PR3 (Cat.#16-14-161820) was purchased from Athens Research & Technology, Inc. (Athens, Georgia, USA). Human Fc receptor blocker (Human BD Fc Block™, Cat.#564220) was from BD Biosciences (New Jersey, USA).

Construction of expression vectors and cell transfection

Mammalian expression constructs encoding the mFc only as well as various GPR97- and GPR56-mFc fusion proteins were generated in the pcDNA3.1-mFc vector. For transient cell transfection, HEK293T cells were transfected with purified plasmid DNAs using Lipofectamine™ or Lipofectamine™ 2000 (Invitrogen). All cell culture media and supplements, including fetal calf serum (FCS), L-glutamine, penicillin and streptomycin were purchased from Invitrogen.

Generation and purification of the mFc-fusion proteins

All recombinant mFc-fusion proteins were purified from conditioned media of transiently transfected HEK293T cells using Protein A-Sepharose affinity chromatography. Briefly, HEK293T cells were transfected with the expression constructs of interest using the calcium phosphate-based transfection method. Following transfection, cells were cultured in serum-free Opti-MEM™ medium for four-five days and conditioned media was collected, centrifuged, and passed through Protein A-Sepharose affinity column (nProtein A Sepharose™ 4 Fast Flow, GE Healthcare), followed by extensive washes with washing buffer (50 mM Tris-HCl pH 7.4, 10 mM $CaCl_2$, 150 mM NaCl). The mFc-fusion protein was eluted and subjected to dialysis in the Slide-A-Lyzer® Dialysis Cassettes (Thermo Scientific) in washing buffer for 24 h at 4℃, filtered, and stored at -80℃ until use.

Purification of blood neutrophils

Human neutrophils were isolated from fresh venous blood donated by healthy volunteers and diseased patients as indicated. All experimental procedures were approved by the Chang Gung Memorial Hospital Ethics Committee (CGMH IRB No: 201701852B0,

201901358B0, 201901293B0, and 202002255B0) and performed according to the guidelines set by the Committee. In brief, blood samples were drawn into the collection tubes coated with sodium heparin and neutrophil isolation was performed using Polymorphprep™ density gradient centrifugation (Axis-Shield, Oslo, Norway). Following the lysis of the remnant erythrocytes, isolated neutrophils were checked for purity by flow cytometry with specific surface markers and resuspended in Roswell Park Memorial Institute (RPMI) medium containing 10% FBS for all following experiments unless otherwise denoted. The isolation procedure routinely produced a >95% pure and viable neutrophil population.

Flow cytometry analysis and the FACS-based ligand-binding assay

All procedures were performed at 4℃. Cells were blocked in ice-cold blocking buffer (1% BSA/5% normal goat serum (NGS) in PBS) for 1 h. For the standard flow cytometry analysis, cells were incubated sequentially with the primary antibody of interest and appropriate fluorescence-labeled secondary antibody diluted in blocking buffer at pre-determined optimal concentrations for 1 h. For the cellular ligand-binding assay, cells were incubated with the mFc-fusion protein probe (10 μg/mL) in blocking buffer for 2 h. Where indicated, cells were pretreated or treated simultaneously with various reagents in the presence of the probes. Fluorescence-conjugated goat anti-mouse IgG (2 μg/mL) was used as the secondary antibody. Following extensive washes between each incubation step, cells were washed lastly in cold PBS and subjected to analysis by FACScan flow cytometer (BD Biosciences). Positive ligand-binding was determined as the shift of fluorescence signals. The mFc protein was always included as a negative control.

Phenotypic analyses of neutrophils

Freshly isolated human neutrophils ($1X10^6$ cells/mL) were suspended in RPMI complete medium and incubated with soluble GPR97-mFc, control mFc protein (10 μg/mL), or fMLF (10 μM) for 3 h at 37℃. For the expressional analysis of CD markers, cells were fixed by 4% paraformaldehyde/PBS for 20 min at 4℃ and stained with fluorescence-conjugated CD markers as indicated using the standard flow cytometry protocol. For the analysis of neutrophil morphological changes, live cells were either collected at 1 h after incubation and photographed under the inverted microscopy, or analyzed directly

by FACScan flow cytometer and shown as FSC/SSC. Alternatively, cells were fixed and subjected to the Wright-Giemsa stain, or permeabilized and blocked in the blocking buffer (1% BSA, 5% NGS, 0.1% saponin in PBS) for 30 min, then incubated with Phalloidin-TRITC (10 µg/mL) and analyzed by flow cytometer as well as Fluoview FV10i (Olympus). For IL-8 production, culture supernatant was collected at 3 h incubation and subjected to the ELISA analysis.

Measurement of IL-8 by sandwich ELISA

Freshly isolated human neutrophils ($1X10^6$ cells/mL) were cultured in RPMI/10% FBS in the presence of fMLF (10 µM), mFc or GPR97-mFc (10 µg/mL) for 3 h at 37℃. When indicated, culture plates were coated with the mFc fusion proteins at 4℃ overnight before incubation with neutrophils. When necessary, neutrophils were pretreated with various reagents/inhibitors such as A1AT (5 mg/mL), aprotinin (50 µg/mL), E-64 (50 µg/mL) and N-p-Tosyl-L-phenylalanine chloromethyl ketone (TPCK, 10 µM) for 1 h at 37℃ or incubated with elafin (4 µM) and SLPI (4 µM) simultaneously. Cell-free supernatants were collected to determine the concentration of IL-8 by using the ELISA kit (DuoSet® ELISA, R&D Systems, Cat.#DY208) according to the protocols suggested by the manufacturer.

FACS-based PR3-binding assay

HEK293T cells were transfected with various expression constructs as indicated. HEK293T transfectants were collected, washed, and incubated in blocking buffer (1% BSA, 5% NGS/PBS) for 1 h at 4℃ before incubating with commercially available purified PR3 (5 µg/mL) diluted in blocking buffer for 1 h at 4℃. Cells were subjected to extensive washes in cold PBS, and then incubated with the anti-PR3 mAb (5 µg/ml, clone MCPR3-2, Thermo Fisher Scientific Cat.#MA5-11945) or PeliCluster ANCA (5 µg/ml)(clone CLB-12.8, Sanquin, Cat.#M1574) for 1h at 4℃. Cells were extensively washed and then incubated with fluorescence-labeled secondary antibody for 1 h at 4℃. Finally, cells were washed three times in cold PBS and subjected to analysis by FACScan flow cytometer (BD Biosciences). For the analysis of GPR97-mPR3 binding in transfected HEK293T cells, cells were transfected as described. Half of the transfected cells were checked

by the flow cytometry analysis using receptor-specific antibodies for the expression of the specific receptors. The other half of the transfected cells were incubated first with purified PR3 (5 µg/mL) diluted in blocking buffer for 1 h at 4℃, washed, and followed by incubation with mFc (10 µg/mL) or GPR97-mFc (10 µg/mL) diluted in blocking buffer for 2 h at 4℃. Cells were washed and reacted with fluorescence-labeled goat anti-mouse IgG for 1 h at 4℃. Finally, cells were washed extensively in cold PBS and subjected to analysis by FACScan flow cytometer (BD Biosciences).

The *ex vivo* membrane PR3 activity assay

The *ex vivo* neutrophil membrane PR3 activity was measured using the PR3-specific FRET substrate, Abz-VADnV-RDRQ-EDDNP (Cayman, MI, USA). Briefly, freshly isolated neutrophils ($1X10^6$ cells/well) were washed twice with PBS. Cells were then suspended in 150 µL PBS and incubated with recombinant mFc-fusion proteins of interest in appropriate concentration for 10 min at 37℃ before adding the PR3-specific FRET substrate (20 µM). Fluorescence intensity was detected by the SpectraMax M2e ELISA reader (Molecular Devices, CA, USA) for at least 90 min during the reaction. When necessary, cells were treated with various protease inhibitors as indicated. The corresponding Vmax was analyzed by the Softmax Pro 5.3 software.

The proximity-ligation assay

To detect the close interaction of CD177, PR3, GPR97, PAR2, and CD16b in human neutrophils, the Duolink in situ PLA (Sigma-Aldrich, Cat.#DUO92013) was employed according to the manufacturer's protocol. Purified human neutrophils were cytospun on Poly-D-Lysine-coated coverslips and blocked in blocking buffer (1% BSA/5% NGS in PBS) for 60 min at 4℃. Cells were incubated with the receptor-specific antibodies (15 µg/mL) pre-conjugated with the PLA-PLUS or -MINUS oligonucleotide probe (Sigma-Aldrich, Cat.#DUO92009 and DUO92010) for 60 min at r.t.. The enzymatic ligation of the PLA-PLUS and -MINUS probes were performed for 30 min at 37℃, followed by circle amplification for 90 min at 37℃ and finally incubated with Hoechst 33342 (Invitrogen) for nuclear staining. Cells were observed by the LSM780 confocal microscope system (ZEISS) at 40X magnification.

Statistics

All results were analyzed using GraphPad Prism version 6.0 (San Diego, CA, USA) and expressed as means ± standard error of the mean (SEM) with the number of experimental replicates (n) provided. Differences between groups were determined by Student's t-test, one-way and two-way ANOVA as indicated. In all cases, a probability (p) value of <0.05 was accepted to reject the null hypothesis and considered as significant.

Appendix B

Constructs and primers used in this work

Boundaries	Mutation		Primer
		human GPR97	
1-258		Fwd	TCTCAGGCCGAATTCAAGCTTGCCACCATGGCGACGCCCAGGGGCCTG
		Rev	GTGGTGCTTGGTACCGGGTCTCAGGAGCAGGGCGAAAAAGG
1-264		Rev	GTGGTGCTTGGTACCCGTGGACTGGTCCAAGGTGGGTCTCAGG
1-264	C245S	Rev_{mut}	GTGGTGCTTGGTACCCGTGGACTGGTCCAAGGTGGGTCTCAGGAGCAG GGCGAAAAAGGTCAGGTGGTCACAGGAGCACACGGTCCCCTCAGG
1-264	C245R	Rev_{mut}	GTGGTGCTTGGTACCCGTGGACTGGTCCAAGGTGGGTCTCAGGAGCAG GGCGAAAAAGGTCAGGTGGTCACAGCGGCACACGGTCCCCTCAGG
1-267		Rev	GTGGTGCTTGGTACCGATATGCACCGTGGACTGGTCCAAGGTGG
1-270		Rev	GTGGTGCTTGGTACCGCGTGTGAGGATATGCACCGTGGACTGG
1-273		Rev	GTGGTGCTTGGTACCCTGGGAGATGCGTGTGAGGATATGCAC
1-275		Rev	GTGGTGCTTGGTACCGCCCGCCTGGGAGATGCGTGTGAGG
1-277		Rev	GTGGTGCTTGGTACCCCCACAGCCCGCCTGGGAGATGCG
21-258		Fwd	GTAGCTGAAACCGGTCAGGAAAAGCCCACCGAAGGGCCAAG
21-267			
29-258		Fwd	GTAGCTGAAACCGGTAGAAACACCTGCCTGGGGAGCAACAAC
29-267			
29-270			
58-258	C62S	Fwd	GTAGCTGAAACCGGTGGCAGCGACTCCTCCAATGTGGAAAACTTGC
58-267	C62S		
58-270	C62S		
		mouse GPR97	
1-253		Fwd	TCTCAGGCCGAATTCAAGCTTGCCACCATGGCGACAGCCAGGAGTCTGG
		Rev	GTGGTGCTTGGTACCCGGCCTCAGCAGCAAGGCGAAGAAGGTC
1-259		Rev	GTGGTGCTTGGTACCCGTGGCCAGGTCCAAGATCGGCCTCAGC
1-262		Rev	GTGGTGCTTGGTACCAGTCTGTGCCGTGGCCAGGTCCAAG
1-265		Rev	GTGGTGCTTGGTACCGCGAGTGAGAGTCTGTGCCGTGGCCAG
1-268		Rev	GTGGTGCTTGGTACCCTGGGAGATGCGAGTGAGAGTCTGTGC
1-270		Rev	GTGGTGCTTGGTACCGCCTGCCTGGGAGATGCGAGTGAGAG
1-272		Rev	GTGGTGCTTGGTACCGGCACTGCCTGCCTGGGAGATGCGAG
		CD177	
1-437		Fwd	TCTCAGGCCGAATTCAAGCTTGCCACCATGAGCCCGGTATTACTGCTGG
		Rev	GTGGTGCTTGGTACCGCAGGAAGGGCAAACCACTCCCCACCAC
1-408		Rev	GTGGTGCTTGGTACCACCTCCCTCATGCTGAGAGGCAGG
1-399		Rev	GTGGTGCTTGGTACCCTGCACATCACGCTTCTCACGCGCAGAG
1-395		Rev	GTGGTGCTTGGTACCCTTCTCACGCGCAGAGAAGATCC
1-206		Rev	GTGGTGCTTGGTACCTTTCCTATTGCAGTTCTCAGTCATACC
22-408		Fwd	GTAGCTGAAACCGGTCTGCTCTGCCAGTTTGGGACAGTTCAGC
209-408		Fwd	ATGAGCCCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCACTGCC AGGAGTGCAGGCGCTGACCTGTCATCGGGGGACCACC

Table B.1: Summary of GPR97^{ecto} and CD177^{ecto} constructs used in this work.

Boundaries	Mutation	Primer	
		PR3	
1-256	S203A	Fwd	TCTCAGGCCGAATTCAAGCTTGCCACCATGGCTCACCGGCCCCCCAG
		Rev	GTGGTGCTTGGTACCGGGGCGGCCCTTGGCCTCCACACGG
		Fwd_{mut}	CGTCCCTCGCCGCAAGGCCGGCATCTGCTTCGGAGACGCAGGTGGC
		Rev_{mut}	CCTTGGATGATGCCATCACAGATCAGGGGGCCACCTGCGTCTCCG
1-248	S203A	Rev	GTGGTGCTTGGTACCGCGCAGCGTGGAACGGATCCAGTCC
28-256	S203A	Fwd	GTAGCTGAAACCGGTATCGTGGGCGGGCACGAGGCGCAGCC
28-248	S203A		
1-256	ΔA26E27 S203A	Fwd	TCTCAGGCCGAATTCAAGCTTGCCACCATGGCTCACCGGCCCCCCAG CCCTGCCCTGGCGTCCGTGCTGCTGGCCTTGCTGCTGAGCGGTGCT GCCCGAGCTATCGTGGGCGGGCACGAG
1-256	ΔA26E27 S203A R227L L228Q	Fwd_{mut}	GCATCATCCAAGGAATAGACTCCTTCGTGATCTGGGGATGTGCCACC CTCCAATTCCCTGAC
		Rev_{mut}	CGGATCCAGTCCACGTAGAGGGCTACCCGCGTGAAGAAGTCAGG GAATTGGAGGGTGGCAC
1-256	ΔA26E27 F181L S203A	Fwd_{mut}	CCAGCCCAGGTCCTGCAGGAGCTCAATGTCACCGTGGTCACCTTCTT GTGCCGGCCAC
		Rev_{mut}	CCTTGCGGCGAGGGACGAAAGTGCAAATGTTATGTGGCCGGCACAAG AAGGTGACC
1-256	ΔA26E27 F181L S203A R227L L228Q		
1-256	ΔA26E27 S203A W222R	Fwd_{mut}	CCTGATCTGTGATGGCATCATCCAAGGAATAGACTCCTTCGTGATCCG GGGATGTG
		Rev_{mut}	GCTACCCGCGTGAAGAAGTCAGGGAAAAGGCGGGTGGCACATCCCCG GATCACG
1-256	ΔA26E27 K195A S203A W222R	Fwd_{mut}	CCTTCTTCTGCCGGCCACATAACATTTGCACTTTCGTCCCTCGCCGCG CGGCCGGCATC
		Rev_{mut}	CCATCACAGATCAGGGGGCCACCTGCGTCTCCGAAGCAGATGCCGGC CGCGCGGCGAG
1-256	ΔA26E27 F181L S203A W222R L228Q		Synthetic gene
1-256	ΔA26E27 F181L S203A W222R R227L L228Q	Rev_{mut}	GCTACCCGCGTGAAGAAGTCAGGGAATTGGAGGGTGGCACATCCCCG GATCACG
1-256	ΔA26E27 F181L K195A S203A W222R R227L L228Q	Fwd_{mut}	CCTTCTTGTGCCGGCCACATAACATTTGCACTTTCGTCCCTCGCCGCG CGGCCGGCATC
1-256	ΔA26E27 T179N F180G F181T S203A I221N W222G G223T		Synthetic gene
1-256	ΔA26E27 T179N F180G F181T S203A	Fwd_{mut}	CCTGATCTGTGATGGCATCATCCAAGGAATAGACTCCTTCGTGATCTG GGGATGTGCCACC
		Rev_{mut}	GCTACCCGCGTGAAGAAGTCAGGGAAAAGGCGGGTGGCACATCCCCA GATCACGAAGGAG

(continued on next page)

1-256	ΔA26E27 S203A I221N W222G G223T	Fwd_{mut}	CCAGCCCAGGTCCTGCAGGAGCTCAATGTCACCGTGGTCACCTT CTTCTGCCGGCCAC
		Rev_{mut}	CCTTGCGGCGAGGGACGAAAGTGCAAATGTTATGTGGCCGGCAGAAG AAGGTGACCACGG
1-256	ΔA26E27 F180S F181D W222R L228H F229S		Synthetic gene
1-256	ΔA26E27 F180S F181D W222R L228H F229S	Fwd_{mut}	GCATCATCCAAGGAATAGACTCCTTCGTGATCTGGGGATGTGC
		Rev_{mut}	CGTGAAGAAGTCAGGGGAGTGGCGGGTGGCACATCCCCAGATCACG

Table B.2: Summary of PR3 constructs used in this work.

Appendix C

GPR97 sequence alignments

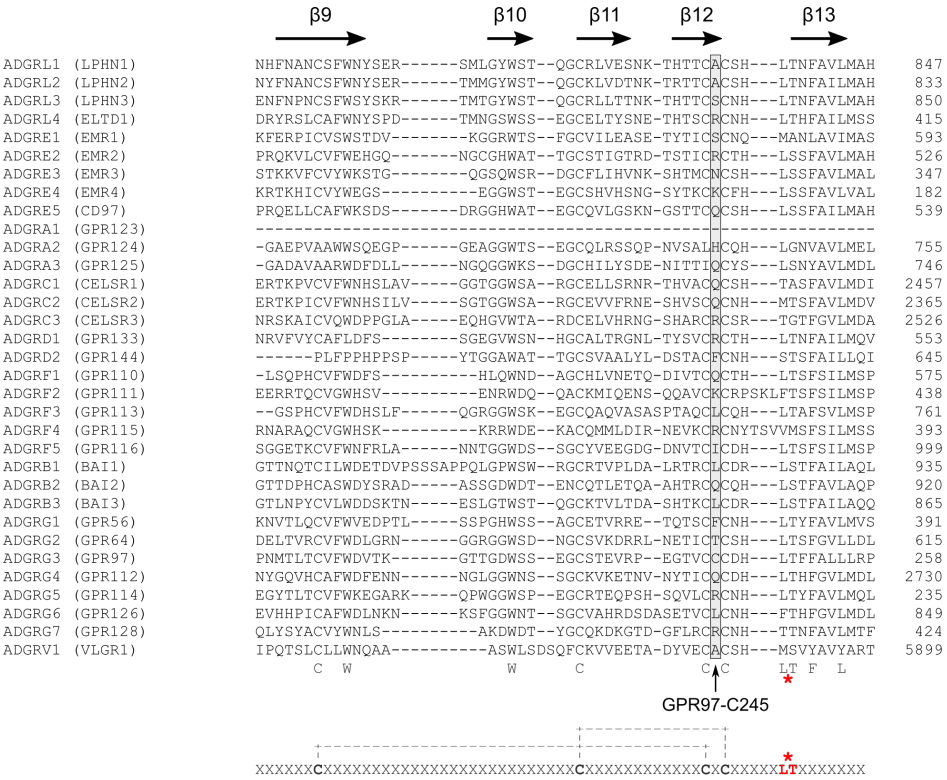


Figure C.1: Predicted GPS for all aGPCRs. Sequence alignment with secondary structure showing no other aGPCRs possess an additional free cysteine within the GPS. Sequence alignment was created using Promals3D server [369].

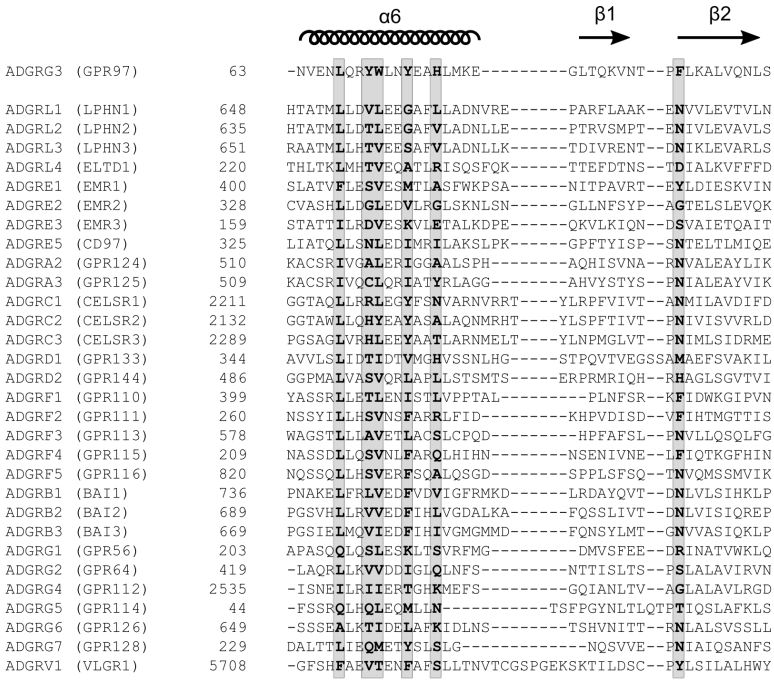


Figure C.2: GPR97 hydrophobic residues involved in NTD-GAIN stabilisation. Sequence alignment with secondary structure showing hydrophobic residues involved in NTD-GAIN stabilisation are not conserved among other members of aGPCR family



Figure C.3: Predicted subdomain A of GAIN domains for all aGPCRs. Sequence alignment with secondary structure showing truncated subdomain A for several aGPCRs. The N-terminal boundaries were selected based on the predicted secondary structure and presence of inter-domain cysteines. Sequence alignment was created using Promals3D server [369] and edited by manual inspection.

Appendix D

PR3 sequence alignment

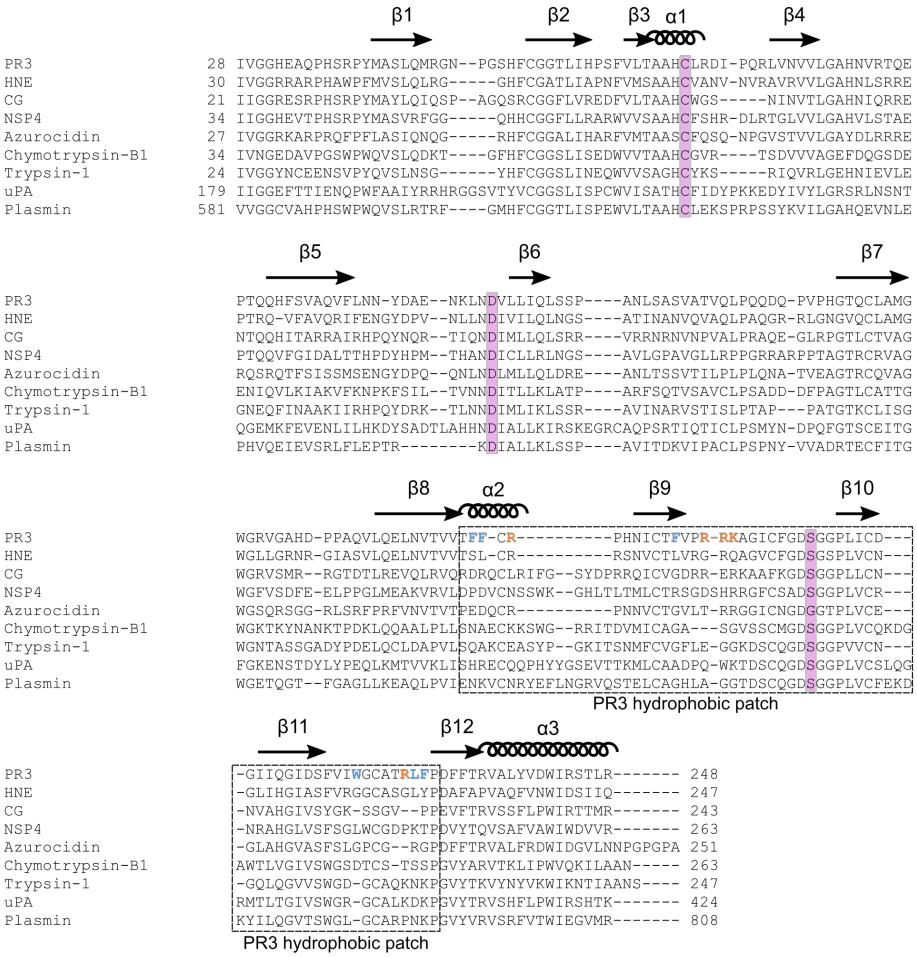


Figure D.1: Sequence alignment of serine proteases. Sequence alignment with secondary structure of all human NSP (PR3, NE, CG, NSP4 and azurocidin) with common serine proteases (chymotrypsin, trypsin, uPA and plasmin). Catalytic residues are shown in bold and highlighted in purple. Hydrophobic and positively-charged residues from PR3 hydrophobic patch are shown in bold blue and orange, respectively. Sequence alignment was created using Promals3D server [369].

References

[1] M. Schultze (1865), Ein heizbarer Objecttisch und seine Verwendung bei Untersuchungen des Blutes. *Archiv für mikroskopische Anatomie*, 1(1):1–42.

[2] N. Borregaard & J. B. Cowland (1997), Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood*, 89(10):3503–3521.

[3] T. Németh, M. Sperandio & A. Mócsai (2020), Neutrophils as emerging therapeutic targets. *Nature Reviews Drug Discovery*, 19(4):253–275.

[4] K. Brown, S. Brain, J. Pearson, J. Edgeworth, S. Lewis & D. Treacher (2006), Neutrophils in development of multiple organ failure in sepsis. *Lancet*, 368(9530):157–169.

[5] M. F. Jimenez, R. W. G. Watson, J. Parodo, D. Evans, D. Foster, M. Steinberg, O. D. Rotstein & J. C. Marshall (1997), Dysregulated expression of neutrophil apoptosis in the systemic inflammatory response syndrome. *Archives of Surgery*, 132(12):1263–1270.

[6] H. L. Wright, R. J. Moots & S. W. Edwards (2014), The multifactorial role of neutrophils in rheumatoid arthritis. *Nature Reviews Rheumatology*, 10(10):593–601.

[7] R. D. Ye, F. Boulay, M. W. Ji, C. Dahlgren, C. Gerard, M. Parmentier, C. N. Serhan & P. M. Murphy (2009), International union of basic and clinical pharmacology. LXXIII. Nomenclature for the formyl peptide receptor (FPR) family. *Pharmacological Reviews*, 61(2):119–161.

[8] H. Lee, P. L. Whitfeld & C. R. Mackay (2008), Receptors for complement C5a. The importance of C5aR and the enigmatic role of C5L2. *Immunology and Cell Biology*, 86(2):153–160.

[9] V. Shpacovitch, M. Feld, M. D. Hollenberg, T. A. Luger & M. Steinhoff (2008), Role of protease-activated receptors in inflammatory responses, innate and adaptive immunity. *Journal of Leukocyte Biology*, 83(6):1309–1322.

[10] P. Bruhns (2012), Properties of mouse and human IgG receptors and their contribution to disease models. *Blood*, 119(24):5640–5649.

[11] M. J. Zhou, D. M. Lublin, D. C. Link & E. J. Brown (1995), Distinct tyrosine kinase activation and triton X-100 insolubility upon FcγRII or FcγRIIIB ligation in human polymorphonuclear leukocytes: Implications for immune complex activation of the respiratory burst. *Journal of Biological Chemistry*, 270(22):13553–13560.

[12] M. Sperandio, C. A. Gleissner & K. Ley (2009), Glycosylation in immune cell trafficking. *Immunological Reviews*, 230(1):97–113.

[13] J. Schymeinsky, A. Mócsai & B. Walzog (2007), Neutrophil activation via β2 integrins (CD11/CD18): Molecular mechanisms and clinical implications. *Thrombosis and Haemostasis*, 98(2):262–273.

[14] M. Gadina, D. Hilton, J. A. Johnston, A. Morinobu, A. Lighvani, Y. J. Zhou, R. Visconti & J. J. O'Shea (2001), Signaling by Type I and II cytokine receptors: Ten years after. *Current Opinion in Immunology*, 13(3):363–373.

[15] R. Medzhitov (2001), Toll-like receptors and innate immunity. *Nature Reviews Immunology*, 1(2):135–145.

[16] A. M. Kerrigan & G. D. Brown (2009), C-type lectins and phagocytosis. *Immunobiology*, 214(7):562–575.

[17] H. L. Malech & J. I. Gallin (1987), Neutrophils in Human Diseases. *New England Journal of Medicine*, 317(11):687–694.

[18] J. W. Leiding (2017), Neutrophil evolution and their diseases in humans. *Frontiers in Immunology*, 8(AUG).

[19] H. Sengeløv, L. Kjeldsen & N. Borregaard (1993), Control of exocytosis in early neutrophil activation. *The Journal of Immunology*, 150(4):1535–1543.

[20] R. Kettritz (2016), Neutral serine proteases of neutrophils. *Immunological Reviews*, 273(1):232–248.

[21] V. Witko-Sarsat, L. Halbwachs-Mecarelli, A. Schuster, P. Nusbaum, I. Ueki, S. Canteloup, G. Lenoir, B. Descamps-Latscha & J. A. Nadel (1999), Proteinase 3, a potent secretagogue in airways, is present in cystic fibrosis sputum. *American Journal of Respiratory Cell and Molecular Biology*, 20(4):729–736.

[22] E. J. Campbell, M. A. Campbell & C. A. Owen (2000), Bioactive Proteinase 3 on the Cell Surface of Human Neutrophils: Quantification, Catalytic Activity, and Susceptibility to Inhibition. *The Journal of Immunology*, 165(6):3366–3374.

[23] V. Witko-Sarsat, E. M. Cramer, C. Hieblot, J. Guichard, P. Nusbaum, S. Lopez, P. Lesavre & L. Halbwachs-Mecarelli (1999), Presence of proteinase 3 in secretory vesicles: Evidence of a novel, highly mobilizable intracellular pool distinct from azurophil granules. *Blood*, 94(7):2487–2496.

[24] M. G. Braun, E. Csernok, W. L. Gross & H. K. Müller-Hermelink (1991), Proteinase 3, the target antigen of anticytoplasmic antibodies circulating in Wegener's granulomatosis: Immunolocalization in normal and pathologic tissues. *American Journal of Pathology*, 139(4):831–838.

[25] E. Csernok, M. Ernst, W. Schmitt, D. F. Bainton & W. L. Gross (1994), Activated neutrophils express proteinase 3 on their plasma membrane in vitro and in vivo. *Clinical and experimental immunology.*

[26] L. Halbwachs-Mecarelli, G. Bessou, P. Lesavre, S. Lopez & V. Witko-Sarsat (1995), Bimodal distribution of proteinase 3 (PR3) surface expression reflects a constitutive heterogeneity in the polymorphonuclear neutrophil pool. *FEBS Letters*, 374(1):29–33.

[27] B. Korkmaz, J. Jaillet, M. L. Jourdan, A. Gauthier, F. Gauthier & S. Attucci (2009), Catalytic activity and inhibition of Wegener antigen proteinase 3 on the cell surface of human polymorphonuclear neutrophils. *Journal of Biological Chemistry*, 284(30):19896–19902.

[28] N. Hu, J. Westra, M. G. Huitema, M. Bijl, E. Brouwer, C. A. Stegeman, P. Heeringa, P. C. Limburg & C. G. Kallenberg (2009), Coexpression of CD177 and membrane proteinase 3 on neutrophils in antineutrophil cytoplasmic autoantibody-associated systemic vasculitis: Anti-proteinase 3-mediated neutrophil activation is independent of the role of CD177-expressing neutrophils. *Arthritis and Rheumatism*, 60(5):1548–1556.

[29] S. Bauer, M. Abdgawad, L. Gunnarsson, M. Segelmark, H. Tapper & T. Hellmark (2007), Proteinase 3 and CD177 are expressed on the plasma membrane of the same subset of neutrophils. *Journal of leukocyte biology*, 81(2):458–464.

[30] S. Von Vietinghoff, G. Tunnemann, C. Eulenberg, M. Wellner, M. C. Cardoso, F. C. Luft & R. Kettritz (2007), NB1 mediates surface expression of the ANCA antigen proteinase 3 on human neutrophils. *Blood*, 109(10):4487–4493.

[31] M. Abdgawad, L. Gunnarsson, A. A. Bengtsson, P. Geborek, L. Nilsson, M. Segelmark & T. Hellmark (2010), Elevated neutrophil membrane expression of proteinase 3 is dependent upon CD177 expression. *Clinical and Experimental Immunology*, 161(1):89–97.

[32] W. H. Goldmann, J. L. Niles & M. A. Arnaout (1999), Interaction of purified human proteinase 3 (PR3) with reconstituted lipid bilayers. *European Journal of Biochemistry*, 261(1):155–162.

[33] E. Hajjar, M. Mihajlovic, V. Witko-Sarsat, T. Lazaridis & N. Reuter (2007), Computational prediction of the binding site of proteinase 3 to the plasma membrane. *Proteins: Structure, Function, and Bioinformatics*, 71(4):1655–1669.

[34] T. Broemstrup & N. Reuter (2010), How does Proteinase 3 interact with lipid bilayers? *Physical Chemistry Chemical Physics*, 12(27):7487–7496.

[35] C. Kantari, A. Millet, J. Gabillet, E. Hajjar, T. Broemstrup, P. Pluta, N. Reuter & V. Witko-Sarsat (2011), Molecular analysis of the membrane insertion domain of proteinase 3, the Wegener's autoantigen, in RBL cells: implication for its pathogenic activity. *Journal of Leukocyte Biology*, 90(5):941–950.

[36] B. Korkmaz, A. Kuhl, B. Bayat, S. Santoso & D. E. Jenne (2008), A hydrophobic patch on proteinase 3, the target of autoantibodies in Wegener granulomatosis, mediates membrane binding via NB1 receptors. *Journal of Biological Chemistry*, 283(51):35976–35982.

[37] A. Schreiber, A. Busjahn, F. C. Luft & R. Kettritz (2003), Membrane expression of proteinase 3 is genetically determined. *Journal of the American Society of Nephrology*, 14(1):68–75.

[38] D. Bories, M. C. Raynal, D. H. Solomon, Z. Darzynkiewicz & Y. E. Cayre (1989), Down-regulation of a serine protease, myeloblastin, causes growth arrest and differentiation of promyelocytic leukemia cells. *Cell*, 59(6):959–968.

[39] J. E. Gabay, R. W. Scott, D. Campanelli, J. Griffith, C. Wilde, M. N. Marra, M. Seeger & C. F. Nathan (1989), Antibiotic proteins of human polymorphonuclear leukocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 86(14):5610–5614.

[40] J. L. Niles, R. T. McCluskey, M. F. Ahmad & M. A. Arnaout (1989), Wegener's granulomatosis autoantigen is a novel neutrophil serine proteinase. *Blood*, 74(6):1888–1893.

[41] D. Campanelli, P. A. Detmers, C. F. Nathan & J. E. Cabay (1990), Azurocidin and a homologous serine protease from neutrophils. Differential antimicrobial and proteolytic properties. *Journal of Clinical Investigation*, 85(3):904–915.

[42] O. E. Sørensen, P. Follin, A. H. Johnsen, J. Calafat, G. S. Tjabringa, P. S. Hiemstra & N. Borregaard (2001), Human cathelicidin, hCAP-18, is processed to the antimicrobial peptide LL-37 by extracellular cleavage with proteinase 3. *Blood*, 97(12):3951–3959.

[43] S. Robache-Gallea, V. Morand, J. M. Bruneau, B. Schoot, E. Tagat, E. Realo, S. Chouaib & S. Roman-Roman (1995), In vitro processing of human tumor necrosis factor-α. *Journal of Biological Chemistry*, 270(40):23688–23692.

[44] C. Coeshott, C. Ohnemus, A. Pilyavskaya, S. Ross, M. Wieczorek, H. Kroona, A. H. Leimer & J. Cheronis (1999), Converting enzyme-independent release of tumor necrosis factor α and IL-1β from a stimulated human monocytic cell line in the presence of activated neutrophils or purified proteinase 3. *Proceedings of the National Academy of Sciences of the United States of America*, 96(11):6261–6266.

[45] M. Padrines, M. Wolf, A. Walz & M. Baggiolini (1994), Interleukin-8 processing by neutrophil elastase, cathepsin G and proteinase-3. *FEBS Letters*, 352(2):231–235.

[46] E. Csernok, M. X. Ai, W. L. Gross, D. Wicklein, A. Petersen, B. Lindner, P. Lamprecht, J. U. Holle & B. Hellmich (2006), Wegener autoantigen induces maturation of dendritic cells and licenses them for Th1 priming via the protease-activated receptor-2 pathway. *Blood*, 107(11):4440–4448.

[47] A. Uehara, K. Muramoto, H. Takada & S. Sugawara (2003), Neutrophil serine proteinases activate human nonepithelial cells to produce inflammatory cytokines through protease-activated receptor 2. *The Journal of Immunology*, 170(11):5690–5696.

[48] M. Baggiolini & I. Clark-Lewis (1992), Interleukin-8, a chemotactic and inflammatory cytokine. *FEBS letters*, 307(1):97–101.

[49] L. Vong, F. D'Acquisto, M. Pederzoli-Ribeil, L. Lavagno, R. J. Flower, V. Witko-Sarsat & M. Perretti (2007), Annexin 1 cleavage in activated neutrophils: A pivotal role for proteinase 3. *Journal of Biological Chemistry*, 282(41):29998–30004.

[50] R. W. Leid, B. E. Ballieux, I. van der Heijden, C. Kleyburg-van der Keur, E. C. Hagen, L. A. van Es, F. J. van der Woude & M. R. Daha (1993), Cleavage and inactivation of human C1 inhibitor by the human leukocyte proteinase, proteinase 3. *European Journal of Immunology*, 23(11):2939–2944.

[51] S. P. Berger, M. A. Seelen, P. S. Hiemstra, J. S. Gerritsma, E. Heemskerk, F. J. Van Der Woude & M. R. Daha (1996), Proteinase 3, the major autoantigen of Wegener's granulomatosis, enhances IL-8 production by endothelial cells in vitro. *Journal of the American Society of Nephrology*, 7(5):694–701.

[52] M. E. Taekema-Roelvink, C. Van Kooten, S. Van Der Kooij, E. Heemskerk & M. R. Daha (2001), Proteinase 3 enhances endothelial monocyte chemoattractant protein-1 production and induces increased adhesion of neutrophils to endothelial cells by upregulating intercellular cell adhesion molecule-1. *Journal of the American Society of Nephrology*, 12(5):932–940.

[53] R. C. Kao, N. G. Wehner, K. M. Skubitz, B. H. Gray & J. R. Hoidal (1988), Proteinase 3. A distinct human polymorphonuclear leukocyte proteinase that produces emphysema in hamsters. *Journal of Clinical Investigation*, 82(6):1963–1973.

[54] N. V. Rao, N. G. Wehner, B. C. Marshall, A. B. Sturrock, T. P. Huecksteadt, G. V. Rao, B. H. Gray & J. R. Hoidal (1991), Proteinase-3 (PR-3): A Polymorphonuclear Leukocyte Serine Proteinase. *Annals of the New York Academy of Sciences*, 624(1):60–68.

[55] N. V. Rao, N. G. Wehner, B. C. Marshall, W. R. Gray, B. H. Gray & J. R. Hoidal (1991), Characterization of proteinase-3 (PR-3), a neutrophil serine proteinase: Structural and functional properties. *Journal of Biological Chemistry*, 266(15):9540–9548.

[56] a. Villegas-Mendez, R. Montes, L. R. Ambrose, a. N. Warrens, M. Laffan & D. a. Lane (2007), Proteolysis of the endothelial cell protein C receptor by neutrophil proteinase 3. *Journal of thrombosis and haemostasis : JTH*, 5(5):980–988.

[57] T. J. Raife, W. Cao, B. S. Atkinson, B. Bedell, R. R. Montgomery, S. R. Lentz, G. F. Johnson & X. L. Zheng (2009), Leukocyte proteases cleave von Willebrand factor at or near the ADAMTS13 cleavage site. *Blood*, 114(8):1666–1674.

[58] S. Durant, M. Pederzoli, Y. Lepelletier, S. Canteloup, P. Nusbaum, P. Lesavre & V. Witko-Sarsat (2004), Apoptosis-induced proteinase 3 membrane expression is independent from degranulation. *Journal of Leukocyte Biology*, 75(1):87–98.

[59] C. Kantari, M. Pederzoli-Ribeil, O. Amir-Moazami, V. Gausson-Dorey, I. C. Moura, M. C. Lecomte, M. Benhamou & V. Witko-Sarsat (2007), Proteinase 3, the Wegener autoantigen, is externalized during neutrophil apoptosis: Evidence for a functional association with phospholipid scramblase 1 and interference with macrophage phagocytosis. *Blood*, 110(12):4086–4095.

[60] J. Gabillet, A. Millet, M. Pederzoli-Ribeil, P. Tacnet-Delorme, L. Guillevin, L. Mouthon, P. Frachet & V. Witko-Sarsat (2012), Proteinase 3, the Autoantigen in Granulomatosis with Polyangiitis, Associates with Calreticulin on Apoptotic Neutrophils, Impairs Macrophage Phagocytosis, and Promotes Inflammation. *The Journal of Immunology*, 189(5):2574–2583.

[61] A. Millet, K. R. Martin, F. Bonnefoy, P. Saas, J. Mocek, M. Alkan, B. Terrier, A. Kerstein, N. Tamassia, S. K. Satyanarayanan, A. Ariel, J. A. Ribeil, L. Guillevin, M. A. Cassatella, A. Mueller, N. Thieblemont, P. Lamprecht, L. Mouthon, S. Perruche & V. Witko-Sarsat (2015), Proteinase 3 on apoptotic cells disrupts immune silencing in autoimmune vasculitis. *Journal of Clinical Investigation*, 125(11):4107–4121.

[62] P. Tacnet-Delorme, J. Gabillet, S. Chatfield, N. Thieblemont, P. Frachet & V. Witko-Sarsat (2018), Proteinase 3 interferes with C1q-mediated clearance of apoptotic cells. *Frontiers in Immunology*, 9(APR).

[63] A. David, R. Fridlich & I. Aviram (2005), The presence of membrane Proteinase 3 in neutrophil lipid rafts and its colocalization with FcγRIIIb and cytochrome b558. *Experimental Cell Research*, 308(1):156–165.

[64] A. David, Y. Kacher, U. Specks & I. Aviram (2003), Interaction of proteinase 3 with CD11b/CD18 (β 2 integrin) on the cell membrane of human neutrophils . *Journal of Leukocyte Biology*, 74(4):551–557.

[65] D. Novick, M. Rubinstein, T. Azam, A. Rabinkov, C. A. Dinarello & S. H. Kim (2006), Proteinase 3 is an IL-32 binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, 103(9):3316–3321.

[66] M. E. Taekema-Roelvink, C. Van Kooten, E. Heemskerk, W. Schroeijers & M. R. Daha (2000), Proteinase 3 interacts with a 111-kD membrane molecule of human umbilical vein endothelial cells. *Journal of the American Society of Nephrology*, 11(4):640–648.

[67] R. Goldschmeding, C. E. Van der Schoot, D. Ten Bokkel Huinink, C. E. Hack, M. E. Van den Ende, C. G. Kallenberg & A. E. Von dem Borne Kr. (1989), Wegener's granulomatosis autoantibodies identify a novel diisopropylfluorophosphate-binding protein in the lysosomes of normal human neutrophils. *Journal of Clinical Investigation*, 84(5):1577–1587.

[68] N. Rasmussen (1989), Alpha-elisa determination of anca and characterization of the anca-related antigen. *Apmis*, 97(6):40.

[69] J. C. Jennette, J. R. Hoidal & R. J. Falk (1990), Specificity of anti-neutrophil cytoplasmic autoantibodies for proteinase 3 (I). *Blood*, 75(11):2263–2264.

[70] J. Ludemann, B. Utecht & W. L. Gross (1990), Anti-neutrophil cytoplasm antibodies in Wegener's granulomatosis recognize an elastinolytic enzyme. *Journal of Experimental Medicine*, 171(1):357–362.

[71] J. Ludemann, B. Utecht & W. L. Gross (1991), Anti-cytoplasmic antibodies in wegener's granulomatosis are directed against proteinase 3. *Advances in Experimental Medicine and Biology*, 297:141–150.

[72] A. R. Kitching, H. J. Anders, N. Basu, E. Brouwer, J. Gordon, D. R. Jayne, J. Kullman, P. A. Lyons, P. A. Merkel, C. O. Savage, U. Specks & R. Kain (2020), ANCA-associated vasculitis. *Nature Reviews Disease Primers*, 6(1):37–72.

[73] P. A. Lyons, T. F. Rayner, S. Trivedi, J. U. Holle, R. A. Watts, D. R. Jayne, B. Baslund, P. Brenchley, A. Bruchfeld, A. N. Chaudhry, J. W. Cohen Tervaert, P. Deloukas, C. Feighery, W. L. Gross, L. Guillevin, I. Gunnarsson, L. Harper, Z. Hrušková, M. A. Little, D. Martorana, T. Neumann, S. Ohlsson, S. Padmanabhan, C. D. Pusey, A. D. Salama, J.-S. F. Sanders, C. O. Savage, M. Segelmark, C. A. Stegeman, V. Tesař, A. Vaglio, S. Wieczorek, B. Wilde, J. Zwerina, A. J. Rees, D. G. Clayton & K. G. Smith (2012), Genetically Distinct Subsets within ANCA-Associated Vasculitis. *New England Journal of Medicine*, 367(3):214–223.

[74] S. Lionaki, E. R. Blyth, S. L. Hogan, Y. Hu, B. A. Senior, C. E. Jennette, P. H. Nachman, J. C. Jennette & R. J. Falk (2012), Classification of antineutrophil cytoplasmic autoantibody vasculitides: The role of antineutrophil cytoplasmic

autoantibody specificity for myeloperoxidase or proteinase 3 in disease recognition and prognosis. *Arthritis and Rheumatism*, 64(10):3452–3462.

[75] F. J. Van Der Woude, S. Lobatto, H. Permin, M. Van Der Giessen, N. Rasmussen, A. Wiik, L. A. Van Es, G. K. Van Der Hem & T. H. The (1985), Autoantibodies Against Neutrophils and Monocytes: Tool for Diagnosis and Marker of Disease Activity in Wegener'S Granulomatosis. *The Lancet*, 325(8426):425–429.

[76] R. J. Falk, R. S. Terrell, L. A. Charles & J. C. Jennette (1990), Anti-neutrophil cytoplasmic autoantibodies induce neutrophils to degranulate and produce oxygen radicals in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 87(11):4115–4119.

[77] U. Jerke, S. Rolle, G. Dittmar, B. Bayat, S. Santoso, A. Sporbert, F. Luft & R. Kettritz (2011), Complement receptor Mac-1 is an adaptor for NB1 (CD177)-mediated PR3-ANCA neutrophil activation. *Journal of Biological Chemistry*, 286(9):7070–7081.

[78] R. J. Falk & J. C. Jennette (1988), Anti-Neutrophil Cytoplasmic Autoantibodies with Specificity for Myeloperoxidase in Patients with Systemic Vasculitis and Idiopathic Necrotizing and Crescentic Glomerulonephritis. *New England Journal of Medicine*, 318(25):1651–1657.

[79] M. T. Keogan, V. L. Esnault, A. J. Green, C. M. Lockwood & D. L. Brown (1992), Activation of normal neutrophils by anti-neutrophil cytoplasm antibodies. *Clinical and Experimental Immunology*, 90(2):228–234.

[80] W. L. Gross, E. Csernok & B. K. Flesch (1993), 'Classic' anti-neutrophil cytoplasmic autoantibodies (cANCA), 'Wegener's autoantigen' and their immunopathogenic role in Wegener's granulomatosis. *Journal of Autoimmunity*, 6(2):171–184.

[81] R. Kettritz, J. C. Jennette & R. J. Falk (1997), Crosslinking of ANCA-Antigens Stimulates Superoxide Release by Human Neutrophils. *Journal of the American Society of Nephrology*, 8(3):386–394.

[82] A. P. Van Rossum, Y. M. Van Der Geld, P. C. Limburg & C. G. Kallenberg (2005), Human anti-neutrophil cytoplasm autoantibodies to proteinase 3 (PR3-ANCA) bind to neutrophils. *Kidney International*, 68(2):537–541.

[83] J. C. Jennette (1991), Antineutrophil Cytoplasmic Autoantibody-Associated Diseases: A Pathologist's Perspective. *American Journal of Kidney Diseases*, 18(2):164–170.

[84] V. Witko-Sarsat, P. Lesavre, S. Lopez, G. Bessou, C. Hieblot, B. Prum, L. H. Noël, L. Guillevin, P. Ravaud, I. Sermet-Gaudelus, J. Timsit, J. P. Grünfeld & L. Halbwachs-Mecarelli (1999), A large subset of neutrophils expressing membrane proteinase 3 is a risk factor for vasculitis and rheumatoid arthritis. *Journal of the American Society of Nephrology*, 10(6):1224–1233.

[85] A. A. Rarok, C. A. Stegeman, P. C. Limburg & C. G. Kallenberg (2002), Neutrophil membrane expression of proteinase 3 (PR3) is related to relapse in PR3-ANCA-associated vasculitis. *Journal of the American Society of Nephrology*, 13(9):2232–2238.

[86] J. J. Yang, W. F. Pendergraft, D. A. Alcorta, P. H. Nachman, S. L. Hogan, R. P. Thomas, P. Sullivan, J. C. Jennette, R. J. Falk & G. A. Preston (2004), Circumvention of normal constraints on granule protein gene expression in peripheral blood neutrophils and monocytes of patients with antineutrophil cytoplasmic autoantibody-associated glomerulonephritis. *Journal of the American Society of Nephrology*, 15(8):2103–2114.

[87] I. F. McKenzie, J. Gardiner, M. Cherry & G. D. Snell (1977), Lymphocyte antigens: Ly-4, Ly-6, and Ly-7. *Transplantation Proceedings*, 9(1):667–669.

[88] C. L. Loughner, E. A. Bruford, M. S. McAndrews, E. E. Delp, S. Swamynathan & S. K. Swamynathan (2016), Organization, evolution and functions of the human and mouse Ly6/uPAR family genes. *Human genomics*, 10:10.

[89] J. Würfel, S. Seiter, M. Stassar, A. Claas, R. Kläs, M. Rösel, R. Marhaba, L. Savelyeva, M. Schwab, S. Matzku & M. Zöller (2001), Cloning of the human homologue of the metastasis-associated rat C4.4A. *Gene*, 262(1-2):35–41.

[90] D. F. Stroncek (2007), Neutrophil-specific antigen HNA-2a, NB1 glycoprotein, and CD177. *Current Opinion in Hematology*, 14(6):688–693.

[91] H. Gårdsvoll, M. C. Kriegbaum, E. P. Hertz, W. Alpízar-Alpízar & M. Ploug (2013), The Urokinase Receptor Homolog Haldisin Is a Novel Differentiation Marker of Stratum Granulosum in Squamous Epithelia. *Journal of Histochemistry and Cytochemistry*, 61(11):802–813.

[92] H. W. Smith & C. J. Marshall (2010), Regulation of cell signalling by uPAR. *Nature Reviews Molecular Cell Biology*, 11(1):23–36.

[93] A. Kurita, T. Takizawa, T. Takayama, K. Totsukawa, S. Matsubara, H. Shibahara, M.-C. Orgebin-Crist, F. Sendo, Y. Shinkai & Y. Araki (2001), Identification, cloning, and initial characterization of a novel mouse testicular germ cell-specific antigen. *Biology of reproduction*, 64(3):935–945.

[94] C. M. Fletcher, R. A. Harrison, P. J. Lachmann & D. Neuhaus (1994), Structure of a soluble, glycosylated form of the human complement regulatory protein CD59. *Structure*, 2(3):185–199.

[95] J. Greenwald, W. H. Fischer, W. W. Vale & S. Choe (1999), Three-finger toxin fold for the extracellular ligand-binding domain of the type II activin receptor serine kinase. *Nature Structural Biology*, 6(1):18–22.

[96] T. Kirsch, W. Sebald & M. K. Dreyer (2000), Crystal structure of the BMP-2-BRIA ectodomain complex. *Nature Structural Biology*, 7(6):492–496.

[97] C. C. Boesen, S. Radaev, S. A. Motyka, A. Patamawenu & P. D. Sun (2002), The 1.1 Å crystal structure of human TGF-β type II receptor ligand binding domain. *Structure*, 10(7):913–919.

[98] R. M. Kini & R. Doley (2010), Structure, function and evolution of three-finger toxins: Mini proteins with multiple targets. *Toxicon*, 56(6):855–867.

[99] D. Tsernoglou & G. A. Petsko (1976), The crystal structure of a post-synaptic neurotoxin from sea snake at 2.2 Å resolution. *FEBS Letters*, 68(1):1–4.

[100] B. W. Low, H. S. Preston, A. Sato, L. S. Rosen, J. E. Searl, A. D. Rudko & J. S. Richardson (1976), Three dimensional structure of erabutoxin b neurotoxic protein: inhibitor of acetylcholine receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 73(9):2991–2994.

[101] V. I. Tsetlin (2015), Three-finger snake neurotoxins and Ly6 proteins targeting nicotinic acetylcholine receptors: Pharmacological tools and endogenous modulators. *Trends in Pharmacological Sciences*, 36(2):109–123.

[102] Y. Fujihara, K. Tokuhiro, Y. Muro, G. Kondoh, Y. Araki, M. Ikawa & M. Okabe (2013), Expression of TEX101, regulated by ACE, is essential for the production of fertile mouse spermatozoa. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20):8111–8116.

[103] M. C. Kriegbaum, B. Jacobsen, A. Hald & M. Ploug (2011), Expression of C4.4A, a structural uPAR homolog, reflects squamous epithelial differentiation in the adult mouse and during embryogenesis. *Journal of Histochemistry and Cytochemistry*, 59(2):188–201.

[104] H. Morishita, J. M. Miwa, N. Heintz & T. K. Hensch (2010), Lynx1, a cholinergic brake, limits plasticity in adult visual cortex. *Science*, 330(6008):1238–1240.

[105] S. Mysling, K. K. Kristensen, M. Larsson, A. P. Beigneux, H. Gårdsvoll, L. G. Fong, A. Bensadouen, T. J. Jørgensen, S. G. Young & M. Ploug (2016), The acidic domain of the endothelial membrane protein GPIHBP1 stabilizes lipoprotein lipase activity by preventing unfolding of its catalytic domain. *eLife*, 5.

[106] Y. Nevo, B. Ben-Zeev, A. Tabib, R. Straussberg, Y. Anikster, Z. Shorer, A. Fattal-Valevski, A. Ta-Shma, S. Aharoni, M. Rabie, S. Zenvirt, H. Goldshmidt, Y. Fellig, A. Shaag, D. Mevorach & O. Elpeleg (2013), CD59 deficiency is associated with chronic hemolysis and childhood relapsing immune-mediated polyneuropathy. *Blood*, 121(1):129–135.

[107] C. J. Kuckleburg, S. B. Tilkens, S. Santoso & P. J. Newman (2012), Proteinase 3 Contributes to Transendothelial Migration of NB1-Positive Neutrophils. *The Journal of Immunology*, 188(5):2419–2426.

[108] M. Bai, R. Grieshaber-Bouyer, J. Wang, A. B. Schmider, Z. S. Wilson, L. Zeng, O. Halyabar, M. D. Godin, H. N. Nguyen, A. Levescot, P. Cunin, C. T. Lefort, R. J. Soberman & P. A. Nigrovic (2017), CD177 modulates human neutrophil migration through activation-mediated integrin and chemoreceptor regulation. *Blood*, 130(19):2092–2100.

[109] A. Herpin, C. Lelong & P. Favrel (2004), Transforming growth factor-β-related proteins: An ancestral and widespread superfamily of cytokines in metazoans. *Developmental and Comparative Immunology*, 28(5):461–485.

[110] D. F. Stroncek, K. M. Skubitz & J. J. McCullough (1990), Biochemical characterization of the neutrophil-specific antigen NB1. *Blood*, 75(3):744–755.

[111] R. Goldschmeding, C. M. van Dalen, N. Faber, J. Calafat, T. W. Huizinga, C. E. van der Schoot, L. T. Clement & A. E. von dem Borne (1992), Further characterization of the NB 1 antigen as a variably expressed 56–62 kD GPI-linked glycoprotein of plasma

membranes and specific granules of neutrophils. *British Journal of Haematology*, 81(3):336–345.

[112] D. F. Stroncek, R. A. Shankar, P. A. Norën, G. P. Herr & L. T. Clement (1996), Analysis of the expression of NB1 antigen using two monoclonal antibodies. *Transfusion*, 36(2):168–174.

[113] D. F. Stroncek, R. Shankar, C. Litz & L. Clement (1998), The expression of the NB1 antigen on myeloid precursors and neutrophils from children and umbilical cords. *Transfusion Medicine*, 8(2):119–123.

[114] K. Matsuo, A. Lin, J. L. Procter, L. Clement & D. Stroncek (2000), Variations in the expression of granulocyte antigen NB1. *Transfusion*, 40(6):654–662.

[115] L. Caruccio, M. Bettinotti, K. Matsuo, V. Sharon & D. Stroncek (2003), Expression of human neutrophil antigen-2a (NB1) is increased in pregnancy. *Transfusion*, 43(3):357–363.

[116] S. Temerinac, S. Klippel, E. Strunck, S. Röder, M. Lübbert, W. Lange, M. Azemar, G. Meinhardt, H.-E. Schaefer & H. L. Pahl (2000), Cloning of PRV-1, a novel member of the uPAR receptor superfamily, which is overexpressed in polycythemia rubra vera. *Blood*, 95(8):2569–2576.

[117] L. Caruccio, M. Bettinotti, A. E. Director-Myska, D. C. Arthur & D. Stroncek (2006), The gene overexpressed in polycythemia rubra vera, PRV-1, and the gene encoding a neutrophil alloantigen, NB1, are alleles of a single gene, CD177, in chromosome band 19q13.31. *Transfusion*, 46(3):441–447.

[118] P. Lalezari, G. B. Murphy & F. H. Allen (1971), NB1, a new neutrophil-specific antigen involved in the pathogenesis of neonatal neutropenia. *The Journal of clinical investigation*, 50(5):1108–1115.

[119] J. Bux, F. Becker, W. Seeger, D. Kilpatrick, J. Chapman & A. Waters (1996), Transfusion-related acute lung injury due to HLA-A2-specific antibodies in recipient and NB1-specific antibodies in donor blood. *British Journal of Haematology*, 93(3):707–713.

[120] R. Leger, S. Palm, H. Wulf, A. Vosberg & J. Neppert (1999), Transfusion-related lung injury with leukopenic reaction caused by fresh frozen plasma containing anti-NB1. *Anesthesiology*, 91(5):1529–1532.

[121] U. J. Sachs, C. L. Andrei-Selmer, A. Maniar, T. Weiss, C. Paddock, V. V. Orlova, Y. C. Eun, P. J. Newman, K. T. Preissner, T. Chavakis & S. Santoso (2007), The neutrophil-specific antigen CD177 is a counter-receptor for platelet endothelial cell adhesion molecule-1 (CD31). *Journal of Biological Chemistry*, 282(32):23603–23612.

[122] D. F. Stroncek, G. P. Herr & L. B. Plachta (1994), Neutrophil-specific antigen nb1 inhibits neutrophil-endothelial cell interactions. *The Journal of laboratory and clinical medicine*, 123(2):247–255.

[123] D. Stroncek, G. Herr, R. Maguire, G. Eiber & L. Clement (1994), Characterization of the neutrophil molecules identified by quinine-dependent antibodies from two patients. *Transfusion*, 34(11):980–985.

[124] U. Jerke, S. F. Marino, O. Daumke & R. Kettritz (2017), Characterization of the CD177 interaction with the ANCA antigen proteinase 3. *Scientific Reports*, 7(February):1–11.

[125] S. F. Marino, U. Jerke, S. Rolle, O. Daumke & R. Kettritz (2022), Competitively disrupting the neutrophil-specific receptor-autoantigen CD177:proteinase 3 membrane complex reduces anti-PR3 antibody-induced neutrophil activation. *Journal of Biological Chemistry*, 298(3).

[126] G. Plitas, C. Konopacki, K. Wu, P. D. Bos, M. Morrow, E. V. Putintseva, D. M. Chudakov & A. Y. Rudensky (2016), Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. *Immunity*, 45(5):1122–1134.

[127] D. Zemmour, R. Zilionis, E. Kiner, A. M. Klein, D. Mathis & C. Benoist (2018), Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR article. *Nature Immunology*, 19(3):291–301.

[128] M. C. Kim, N. Borcherding, K. K. Ahmed, A. P. Voigt, A. Vishwakarma, R. Kolb, P. N. Kluz, G. Pandey, U. De, T. Drashansky, E. Y. Helm, X. Zhang, K. N. Gibson-Corley, J. Klesney-Tait, Y. Zhu, J. Lu, J. Lu, X. Huang, H. Xiang, J. Cheng, D. Wang, Z. Wang, J. Tang, J. Hu, Z. Wang, H. Liu, M. Li, H. Zhuang, D. Avram, D. Zhou, R. Bacher, S. G. Zheng, X. Wu, Y. Zakharia & W. Zhang (2021), CD177 modulates the function and homeostasis of tumor-infiltrating regulatory T cells. *Nature Communications*, 12(1).

[129] D. F. Stroncek, L. B. Plachta, M. E. Clay & A. P. Dalmasso (1993), Polyclonal antibodies against the NB1-bearing 58- to 64-kDa glycoprotein of human neutrophils do not identify an NB2-bearing molecule. *Transfusion*, 33(5):399–404.

[130] J. Craig Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural et al. (2001), The sequence of the human genome. *Science*, 291(5507):1304–1351.

[131] T. K. Bjarnadóttir, D. E. Gloriam, S. H. Hellstrand, H. Kristiansson, R. Fredriksson & H. B. Schiöth (2006), Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. *Genomics*, 88(3):263–273.

[132] V. Baud, S. L. Chissoe, E. Viegas-Péquignot, S. Diriong, V. C. N'guyen, B. A. Roe & M. Lipinski (1995), EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments. *Genomics*, 26(2):334–344.

[133] J. Hamann, W. Eichler, D. Hamann, H. Kerstens, P. Poddighe, J. Hoovers, E. Hartmann, M. Strauss & R. Van Lier (1995), Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretin receptor superfamily with an unusual extracellular domain. *Journal of Immunology*, 155(4).

[134] A. J. McKnight & S. Gordon (1996), EGF-TM7: A novel subfamily of seven-transmembrane-region leukocyte cell-surface molecules. *Immunology Today*, 17(6):283–287.

[135] V. G. Krasnoperov, M. A. Bittner, R. Beavis, Y. Kuang, K. V. Salnikow, O. G. Chepurny, A. R. Little, A. N. Plotnikov, D. Wu, R. W. Holz & A. G. Petrenko (1997), α-Latrotoxin stimulates exocytosis by the interaction with a neuronal G-protein-coupled receptor. *Neuron*, 18(6):925–937.

[136] T. Usui, Y. Shima, Y. Shimada, S. Hirano, R. W. Burgess, T. L. Schwarz, M. Takeichi & T. Uemura (1999), Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell*, 98(5):585–595.

[137] H. Nishimori, T. Shiratsuchi, T. Urano, Y. Kimura, K. Kiyono, K. Tatsumi, S. Yoshida, M. Ono, M. Kuwano, Y. Nakamura & T. Tokino (1997), A novel brain-specific p53-target gene, BAI1, containing thrombospondin type 1 repeats inhibits experimental angiogenesis. *Oncogene*, 15(18):2145–2150.

[138] T. Shiratsuchi, H. Nishimori, H. Ichise, Y. Nakamura & T. Tokino (1997), Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI1). *Cytogenetic and Genome Research*, 79(1-2):103–108.

[139] R. Fredriksson, M. C. Lagerström, L.-G. Lundin & H. B. Schiöth (2003), The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Molecular Pharmacology*, 63(6):1256–1272.

[140] T. K. Bjarnadóttir, R. Fredriksson, P. J. Höglund, D. E. Gloriam, M. C. Lagerström & H. B. Schiöth (2004), The human and mouse repertoire of the adhesion family of G-protein-coupled receptors. *Genomics*, 84(1):23–33.

[141] D. Araç, A. A. Boucard, M. F. Bolliger, J. Nguyen, S. M. Soltis, T. C. Südhof & A. T. Brunger (2012), A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *EMBO Journal*, 31(6):1364–1378.

[142] V. Krasnoperov, Y. Lu, L. Buryanovsky, T. A. Neubert, K. Ichtehenko & A. G. Petrenko (2002), Post-translational proteolytic processing of the calcium-independent receptor of α-latrotoxin (CIRL), a natural chimera of the cell adhesion protein and the G protein-coupled receptor: Role of the G protein-coupled receptor proteolysis site (GPS) motif. *Journal of Biological Chemistry*, 277(48):46518–46526.

[143] J. X. Gray, M. Haino, M. J. Roth, J. E. Maguire, P. N. Jensen, A. Yarme, M. A. Stetler-Stevenson, U. Siebenlist & K. Kelly (1996), CD97 is a processed, seven-transmembrane, heterodimeric receptor associated with inflammation. *Journal of immunology (Baltimore, Md. : 1950)*, 157(12):5438–47.

[144] H. H. Lin, G. W. Chang, J. Q. Davies, M. Stacey, J. Harris & S. Gordon (2004), Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-coupled receptor proteolytic site motif. *Journal of Biological Chemistry*, 279(30):31823–31832.

[145] C. Wilde, L. Fischer, V. Lede, J. Kirchberger, S. Rothemund, T. Schöneberg & I. Liebscher (2016), The constitutive activity of the adhesion GPCR GPR114/ADGRG5 is mediated by its tethered agonist. *FASEB Journal*, 30(2):666–673.

[146] M. C. Peeters, M. Fokkelman, B. Boogaard, K. L. Egerod, B. van de Water, A. P. IJzerman & T. W. Schwartz (2015), The adhesion G protein-coupled receptor G2 (ADGRG2/GPR64) constitutively activates SRE and NFκB and is involved in cell adhesion and migration. *Cellular Signalling*, 27(12):2579–2588.

[147] A. Kishore, R. H. Purcell, Z. Nassiri-Toosi & R. A. Hall (2016), Stalk-dependent and stalk-independent signaling by the adhesion G protein-coupled receptors GPR56 (ADGRG1) and BAI1 (ADGRB1). *Journal of Biological Chemistry*, 291(7):3385–3394.

[148] J. Hamann, G. Aust, D. Arac, F. B. Engel, C. Formstone, R. Fredriksson, R. A. Hall, B. L. Harty, C. Kirchhoff, B. Knapp, A. Krishnan, I. Liebscher, H.-H. Lin, D. C. Martinelli, K. R. Monk, M. C. Peeters, X. Piao, S. Promel, T. Schoneberg, T. W. Schwartz, K. Singer, M. Stacey, Y. A. Ushkaryov, M. Vallon, U. Wolfrum, M. W. Wright, L. Xu, T. Langenhan & H. B. Schioth (2015), International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G Protein-Coupled Receptors. *Pharmacological Reviews*, 67(2):338–367.

[149] R. H. Purcell & R. A. Hall (2018), Adhesion G Protein-Coupled Receptors as Drug Targets. *Annual Review of Pharmacology and Toxicology*, 58:429–449.

[150] J. P. Silva, V. G. Lelianova, Y. S. Ermolyuk, N. Vysokov, P. G. Hitchen, O. Berninghausen, M. A. Rahman, A. Zangrandi, S. Fidalgo, A. G. Tonevitsky, A. Dell, K. E. Volynski & Y. A. Ushkaryov (2011), Latrophilin 1 and its endogenous ligand Lasso/teneurin-2 form a high-affinity transsynaptic receptor pair with signaling capabilities. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29):12113–12118.

[151] M. L. O'Sullivan, J. de Wit, J. N. Savas, D. Comoletti, S. Otto-Hitt, J. R. Yates & A. Ghosh (2012), FLRT Proteins Are Endogenous Latrophilin Ligands and Regulate Excitatory Synapse Development. *Neuron*, 73(5):903–910.

[152] A. A. Boucard, J. Ko & T. C. Südhof (2012), High affinity neurexin binding to cell adhesion G-protein-coupled receptor CIRL1/latrophilin-1 produces an intercellular adhesion complex. *Journal of Biological Chemistry*, 287(12):9399–9413.

[153] A. Zuko, A. Oguro-Ando, H. Post, R. L. Taggenbrock, R. E. van Dijk, A. F. Altelaar, A. J. Heck, A. G. Petrenko, B. van der Zwaag, Y. Shimoda, R. J. Pasterkamp & J. P. Burbach (2016), Association of cell adhesion molecules contactin-6 and latrophilin-1 regulates neuronal apoptosis. *Frontiers in Molecular Neuroscience*, 9(DEC2016).

[154] D. del Toro, M. A. Carrasquero-Ordaz, A. Chu, T. Ruff, M. Shahin, V. A. Jackson, M. Chavent, M. Berbeira-Santana, G. Seyit-Bremer, S. Brignani, R. Kaufmann, E. Lowe, R. Klein & E. Seiradake (2020), Structural Basis of Teneurin-Latrophilin Interaction in Repulsive Guidance of Migrating Neurons. *Cell*, 180(2):323–339.e19.

[155] V. A. Jackson, D. Del Toro, M. Carrasquero, P. Roversi, K. Harlos, R. Klein & E. Seiradake (2015), Structural basis of latrophilin-FLRT interaction. *Structure*, 23(4):774–781.

[156] V. A. Jackson, S. Mehmood, M. Chavent, P. Roversi, M. Carrasquero, D. Del Toro, G. Seyit-Bremer, F. M. Ranaivoson, D. Comoletti, M. S. Sansom, C. V. Robinson, R. Klein & E. Seiradake (2016), Super-complexes of adhesion GPCRs and neural guidance receptors. *Nature Communications*, 7:11184.

[157] H. Warschkau & a. F. Kiderlen (1999), A monoclonal antibody directed against the murine macrophage surface molecule F4/80 modulates natural immune response to Listeria monocytogenes. *Journal of immunology (Baltimore, Md. : 1950)*, 163(6):3409–3416.

[158] M. Stacey, G. W. Chang, J. Q. Davies, M. J. Kwakkenbos, R. D. Sanderson, J. Hamann, S. Gordon & H. H. Lin (2003), The epidermal growth factor-like domains of the human

EMR2 receptor mediate cell attachment through chondroitin sulfate glycosaminoglycans. *Blood*, 102(8):2916–2924.

[159] S. Irmscher, S. R. Brix, S. L. Zipfel, L. D. Halder, S. Mutlutürk, S. Wulf, E. Girdauskas, H. Reichenspurner, R. A. Stahl, B. Jungnickel, T. Wiech, P. F. Zipfel & C. Skerka (2019), Serum FHR1 binding to necrotic-type cells activates monocytic inflammasome and marks necrotic sites in vasculopathies. *Nature Communications*, 10(1).

[160] M. Stacey, H. H. Lin, K. L. Hilyard, S. Gordon & A. J. McKnight (2001), Human Epidermal Growth Factor (EGF) Module-containing Mucin-like Hormone Receptor 3 is a New Member of the EGF-TM7 Family that Recognizes a Ligand on Human Macrophages and Activated Neutrophils. *Journal of Biological Chemistry*, 276(22):18863–18870.

[161] M. Stacey, G. W. Chang, S. L. Sanos, L. R. Chittenden, L. Stubbs, S. Gordon & H. H. Lin (2002), EMR4, a novel epidermal growth factor (EGF)-TM7 molecule up-regulated in activated mouse macrophages, binds to a putative cellular ligand on B lymphoma cell line A20. *The Journal of biological chemistry*, 277(32):29283–29293.

[162] J. Hamann, B. Vogel, G. M. Van Schijndel & R. A. Van Lier (1996), The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). *Journal of Experimental Medicine*, 184(3):1185–1189.

[163] T. Wang, Y. Ward, L. Tian, R. Lake, L. Guedez, W. G. Stetler-Stevenson & K. Kelly (2005), CD97, an adhesion receptor on inflammatory cells, stimulates angiogenesis through binding integrin counterreceptors on endothelial cells. *Blood*, 105(7):2836–2844.

[164] E. Wandel, A. Saalbach, D. Sittig, C. Gebhardt & G. Aust (2012), Thy-1 (CD90) Is an Interacting Partner for CD97 on Activated Endothelial Cells. *The Journal of Immunology*, 188(3):1442–1450.

[165] Y. Ward, R. Lake, J. J. Yin, C. D. Heger, M. Raffeld, P. K. Goldsmith, M. Merino & K. Kelly (2011), LPA receptor heterodimerizes with CD97 to amplify LPA-initiated RHO-dependent signaling and invasion in prostate cancer cells. *Cancer Research*, 71(23):7301–7311.

[166] D. Hilbig, N. Dietrich, E. Wandel, S. Gonsior, D. Sittig, J. Hamann & G. Aust (2018), The interaction of CD97/ADGRE5 with β-catenin in adherens junctions is lost during colorectal carcinogenesis. *Frontiers in Oncology*, 8(MAY).

[167] M. Vallon & M. Essler (2006), Proteolytically processed soluble tumor endothelial marker (TEM) 5 mediates endothelial cell survival during angiogenesis by linking integrin αvβ3 to glycosaminoglycans. *Journal of Biological Chemistry*, 281(45):34179–34188.

[168] C. Cho, P. M. Smallwood & J. Nathans (2017), Reck and Gpr124 Are Essential Receptor Cofactors for Wnt7a/Wnt7b-Specific Signaling in Mammalian CNS Angiogenesis and Blood-Brain Barrier Regulation. *Neuron*, 95(5):1056–1073.e5.

[169] L. B. Lindenmaier, N. Parmentier, C. Guo, F. Tissir & K. M. Wright (2019), Dystroglycan is a scaffold for extracellular axon guidance decisions. *eLife*, 8.

[170] E. Bianchi, Y. Sun, A. Almansa-Ordonez, M. Woods, D. Goulding, N. Martinez-Martin & G. J. Wright (2021), Control of oviductal fluid flow by the G-protein coupled receptor Adgrd1 is essential for murine embryo transit. *Nature Communications*, 12(1).

[171] J. W. Lee, B. X. Huang, H. S. Kwon, M. A. Rashid, G. Kharebava, A. Desai, S. Patnaik, J. Marugan & H. Y. Kim (2016), Orphan GPR110 (ADGRF1) targeted by N-docosahexaenoylethanolamine in development of neurons and cognitive function. *Nature Communications*, 7:1–16.

[172] T. Fukuzawa, J. Ishida, A. Kato, T. Ichinose, D. M. Ariestanti, T. Takahashi, K. Ito, J. Abe, T. Suzuki, S. Wakana, A. Fukamizu, N. Nakamura & S. Hirose (2013), Lung Surfactant Levels are Regulated by Ig-Hepta/GPR116 by Monitoring Surfactant Protein D. *PLoS ONE*, 8(7).

[173] A. Georgiadi, V. Lopez-Salazar, R. E. Merahbi, R. A. Karikari, X. Ma, A. Mourão, K. Klepac, L. Bühler, A. J. Alfaro, I. Kaczmarek, A. Linford, M. Bosma, O. Shilkova, O. Ritvos, N. Nakamura, S. Hirose, M. Lassi, R. Teperino, J. Machado, M. Scheideler, A. Dietrich, A. Geerlof, A. Feuchtinger, A. Blutke, K. Fischer, T. D. Müller, K. Kessler, T. Schöneberg, D. Thor, S. Hornemann, M. Kruse, P. Nawroth, O. Pivovarova-Ramich, A. F. H. Pfeiffer, M. Sattler, M. Blüher & S. Herzig (2021), Orphan GPR116 mediates the insulin sensitizing effects of the hepatokine FNDC4 in adipose tissue. *Nature Communications*, 12(1):1–21.

[174] D. Park, A. C. Tosello-Trampont, M. R. Elliott, M. Lu, L. B. Haney, Z. Ma, A. L. Klibanov, J. W. Mandell & K. S. Ravichandran (2007), BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. *Nature*, 450(7168):430–434.

[175] S. Das, K. A. Owen, K. T. Ly, D. Park, S. G. Black, J. M. Wilson, C. D. Sifri, K. S. Ravichandran, P. B. Ernst & J. E. Casanova (2011), Brain angiogenesis inhibitor 1 (BAI1) is a pattern recognition receptor that mediates macrophage binding and engulfment of Gram-negative bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 108(5):2136–2141.

[176] B. Kaur, S. M. Cork, E. M. Sandberg, N. S. Devi, Z. Zhang, P. A. Klenotic, M. Febbraio, H. Shim, H. Mao, C. Tucker-Burden, R. L. Silverstein, D. J. Brat, J. J. Olson & E. G. Van Meir (2009), Vasculostatin inhibits intracranial glioma growth and negatively regulates in vivo angiogenesis through a CD36-dependent mechanism. *Cancer Research*, 69(3):1212–1220.

[177] J. T. Koh, H. Kook, H. J. Kee, Y. W. Seo, B. C. Jeong, J. H. Lee, M. Y. Kim, K. C. Yoon, S. Jung & K. K. Kim (2004), Extracellular fragment of brain-specific angiogenesis inhibitor 1 suppresses endothelial cell proliferation by blocking α vβ5 integrin. *Experimental Cell Research*, 294(1):172–184.

[178] J. Wang, Y. Miao, R. Wicklein, Z. Sun, J. Wang, K. M. Jude, R. A. Fernandes, S. A. Merrill, M. Wernig, K. C. Garcia & T. C. Südhof (2021), RTN4/NoGo-receptor binding to BAI adhesion-GPCRs regulates neuronal development. *Cell*, 184(24):5869–5885.e25.

[179] S. Zencir, M. Ovee, M. J. Dobson, M. Banerjee, Z. Topcu & S. Mohanty (2011), Identification of brain-specific angiogenesis inhibitor 2 as an interaction partner of

glutaminase interacting protein. *Biochemical and Biophysical Research Communications*, 411(4):792–797.

[180] M. F. Bolliger, D. C. Martinelli & T. C. Sudhof (2011), The cell-adhesion G protein-coupled receptor BAI3 is a high-affinity receptor for C1q-like proteins. *Proceedings of the National Academy of Sciences*.

[181] L. Xu, S. Begum, J. D. Hearn & R. O. Hynes (2006), GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, TG2, and inhibits melanoma tumor growth and metastasis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24):9023–9028.

[182] R. Luo, Z. Jin, Y. Deng, N. Strokes & X. Piao (2012), Disease-associated mutations prevent GPR56-collagen III interaction. *PLoS ONE*, 7(1).

[183] G. Jin, K. Sakitani, H. Wang, Y. Jin, A. Dubeykovskiy, D. L. Worthley, Y. Tailor & T. C. Wang (2017), The G-protein coupled receptor 56, expressed in colonic stem and cancer cells, binds progastrin to promote proliferation and carcinogenesis. *Oncotarget*, 8(25):40606–40619.

[184] Y. Q. Ping, C. Mao, P. Xiao, R. J. Zhao, Y. Jiang, Z. Yang, W. T. An, D. D. Shen, F. Yang, H. Zhang, C. Qu, Q. Shen, C. Tian, Z. jian Li, S. Li, G. Y. Wang, X. Tao, X. Wen, Y. N. Zhong, J. Yang, F. Yi, X. Yu, H. E. Xu, Y. Zhang & J. P. Sun (2021), Structures of the glucocorticoid-bound adhesion receptor GPR97–Go complex. *Nature*, 589(7843):620–626.

[185] S. C. Petersen, R. Luo, I. Liebscher, S. Giera, S. J. Jeong, A. Mogha, M. Ghidinelli, M. L. Feltri, T. Schöneberg, X. Piao & K. R. Monk (2015), The Adhesion GPCR GPR126 Has Distinct, Domain-Dependent Functions in Schwann Cell Development Mediated by Interaction with Laminin-211. *Neuron*, 85(4):755–769.

[186] K. J. Paavola, H. Sidik, J. B. Zuchero, M. Eckart & W. S. Talbot (2014), Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. *Science Signaling*, 7(338).

[187] A. Küffer, A. K. Lakkaraju, A. Mogha, S. C. Petersen, K. Airich, C. Doucerain, R. Marpakwar, P. Bakirci, A. Senatore, A. Monnard, C. Schiavi, M. Nuvolone, B. Grosshans, S. Hornemann, F. Bassilana, K. R. Monk & A. Aguzzi (2016), The prion protein is an agonistic ligand of the G protein-coupled receptor Adgrg6. *Nature*, 536(7617):464–468.

[188] T. Langenhan, G. Aust & J. Hamann (2013), Sticky Signaling–Adhesion Class G Protein-Coupled Receptors Take the Stage. *Science Signaling*, 6(276):re3–re3.

[189] A. Vizurraga, R. Adhikari, J. Yeung, M. Yu & G. G. Tall (2020), Mechanisms of adhesion G protein–coupled receptor activation. *Journal of Biological Chemistry*, 295(41):14065–14083.

[190] S. Shashidhar, G. Lorente, U. Nagavarapu, A. Nelson, J. Kuo, J. Cummins, K. Nikolich, R. Urfer & E. D. Foehr (2005), GPR56 is a GPCR that is overexpressed in gliomas and functions in tumor cell adhesion. *Oncogene*, 24(10):1673–1682.

[191] Y.-S. Huang, N.-Y. Chiang, C.-H. Hu, C.-C. Hsiao, K.-F. Cheng, W.-P. Tsai, S. Yona, M. Stacey, S. Gordon, G.-W. Chang & H.-H. Lin (2012), Activation of Myeloid Cell-Specific Adhesion Class G Protein-Coupled Receptor EMR2 via Ligation-Induced Translocation and Interaction of Receptor Subunits in Lipid Raft Microdomains. *Molecular and Cellular Biology*, 32(8):1408–1420.

[192] D. Okajima, G. Kudo & H. Yokota (2010), Brain-specific angiogenesis inhibitor 2 (BAI2) may be activated by proteolytic processing. *Journal of Receptors and Signal Transduction*.

[193] K. J. Paavola, J. R. Stephenson, S. L. Ritter, S. P. Alter & R. A. Hall (2011), The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. *Journal of Biological Chemistry*, 286(33):28914–28921.

[194] J. R. Stephenson, K. J. Paavola, S. A. Schaefer, B. Kaur, E. G. Van Meir & R. A. Hall (2013), Brain-specific angiogenesis inhibitor-1 signaling, regulation, and enrichment in the postsynaptic density. *Journal of Biological Chemistry*, 288(31):22248–22256.

[195] I. Liebscher, J. Schön, S. C. Petersen, L. Fischer, N. Auerbach, L. M. Demberg, A. Mogha, M. Cöster, K. U. Simon, S. Rothemund, K. R. Monk & T. Schöneberg (2014), A Tethered Agonist within the Ectodomain Activates the Adhesion G Protein-Coupled Receptors GPR126 and GPR133. *Cell Reports*, 9(6):2018–2026.

[196] Q. X. Hu, J. H. Dong, H. B. Du, D. L. Zhang, H. Z. Ren, M. L. Ma, Y. Cai, T. C. Zhao, X. L. Yin, X. Yu, T. Xue, Z. G. Xu & J. P. Sun (2014), Constitutive Gαi coupling activity of very large G protein-coupled receptor 1 (VLGR1) and its regulation by PDZD7 protein. *Journal of Biological Chemistry*, 289(35):24215–24225.

[197] H. M. Stoveken, A. G. Hajduczok, L. Xu & G. G. Tall (2015), Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *Proceedings of the National Academy of Sciences*, 112(19):6194–6199.

[198] L. M. Demberg, S. Rothemund, T. Schöneberg & I. Liebscher (2015), Identification of the tethered peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. *Biochemical and Biophysical Research Communications*, 464(3):743–747.

[199] L. M. Demberg, J. Winkler, C. Wilde, K. U. Simon, J. Schön, S. Rothemund, T. Schöneberg, S. Prömel & I. Liebscher (2017), Activation of adhesion G protein-coupled receptors: Agonist specificity of Stachel sequence-derived peptides. *Journal of Biological Chemistry*, 292(11):4383–4394.

[200] J. P. White, C. D. Wrann, R. R. Rao, S. K. Nair, M. P. Jedrychowski, J. S. You, V. Martínez-Redondo, S. P. Gygi, J. L. Ruas, T. A. Hornberger, Z. Wu, D. J. Glass, X. Piao & B. M. Spiegelman (2014), G protein-coupled receptor 56 regulates mechanical overload-induced muscle hypertrophy. *Proceedings of the National Academy of Sciences of the United States of America*, 111(44):15756–15761.

[201] N. Scholz, J. Gehring, C. Guan, D. Ljaschenko, R. Fischer, V. Lakshmanan, R. J. Kittel & T. Langenhan (2015), The Adhesion GPCR Latrophilin/CIRL Shapes Mechanosensation. *Cell Reports*, 11(6):866–874.

[202] Y.-Q. Ping, P. Xiao, F. Yang, R.-J. Zhao, S.-C. Guo, X. Yan, X. Wu, C. Zhang, Y. Lu, F. Zhao, F. Zhou, Y.-T. Xi, W. Yin, F.-Z. Liu, D.-F. He, D.-L. Zhang, Z.-L. Zhu, Y. Jiang, L. Du, S.-Q. Feng, T. Schöneberg, I. Liebscher, H. E. Xu & J.-P. Sun (2022), Structural basis for the tethered peptide activation of adhesion GPCRs. *Nature*, 604:763–770.

[203] X. Qu, N. Qiu, M. Wang, B. Zhang, J. Du, Z. Zhong, W. Xu, X. Chu, L. Ma, C. Yi, S. Han, W. Shui, Q. Zhao & B. Wu (2022), Structural basis of tethered agonism of the adhesion GPCRs ADGRD1 and ADGRF1. *Nature*, 604:779–785.

[204] P. Xiao, S. Guo, X. Wen, Q.-T. He, H. Lin, S.-M. Huang, L. Gou, C. Zhang, Z. Yang, Y.-N. Zhong, C.-C. Yang, Y. Li, Z. Gong, X.-N. Tao, Z.-S. Yang, Y. Lu, S.-L. Li, J.-Y. He, C. Wang, L. Zhang, L. Kong, J.-P. Sun & X. Yu (2022), Tethered peptide activation mechanism of the adhesion GPCRs ADGRG2 and ADGRG4. *Nature*, 604:771–778.

[205] X. Barros-Álvarez, R. M. Nwokonko, A. Vizurraga, D. Matzov, F. He, M. M. Papasergi-Scott, M. J. Robertson, O. Panova, E. H. Yardeni, A. B. Seven, F. E. Kwarcinski, H. Su, M. C. Peroto, J. G. Meyerowitz, M. Shalev-Benami, G. G. Tall & G. Skiniotis (2022), The tethered peptide activation mechanism of adhesion GPCRs. *Nature*, 604:757–762.

[206] S. Prömel, M. Frickenhaus, S. Hughes, L. Mestek, D. Staunton, A. Woollard, I. Vakonakis, T. Schöneberg, R. Schnabel, A. P. Russ & T. Langenhan (2012), The GPS Motif Is a Molecular Switch for Bimodal Activities of Adhesion Class G Protein-Coupled Receptors. *Cell Reports*, 2(2):321–331.

[207] D. A. Hume, A. P. Robinson, G. G. Macpherson & S. Gordon (1983), The mononuclear phagocyte system of the mouse defined by immunohistochemical localization of antigen F4/80: Relationship between macrophages, langerhans cells, reticular cells, and dendritic cells in lymphoid and hematopoietic organs. *Journal of Experimental Medicine*, 158(5):1522–1536.

[208] H.-H. Lin, D. E. Faunce, M. Stacey, A. Terajewicz, T. Nakamura, J. Zhang-Hoover, M. Kerley, M. L. Mucenski, S. Gordon & J. Stein-Streilein (2005), The macrophage F4/80 receptor is required for the induction of antigen-specific efferent regulatory T cells in peripheral tolerance. *The Journal of Experimental Medicine*, 201(10):1615–1625.

[209] J. Hamann, N. Koning, W. Pouwels, L. H. Ulfman, M. van Eijk, M. Stacey, H. H. Lin, S. Gordon & M. J. Kwakkenbos (2007), EMR1, the human homolog of F4/80, is an eosinophil-specific receptor. *European Journal of Immunology*, 37(10):2797–2802.

[210] M. J. Kwakkenbos, G.-W. Chang, H.-H. Lin, W. Pouwels, E. C. de Jong, R. A. W. van Lier, S. Gordon & J. Hamann (2002), The human EGF-TM7 family member EMR2 is a heterodimeric receptor expressed on myeloid cells. *Journal of leukocyte biology*, 71(5):854–62.

[211] M. Matmati, W. Pouwels, R. van Bruggen, M. Jansen, R. M. Hoek, A. J. Verhoeven & J. Hamann (2007), The human EGF-TM7 receptor EMR3 is a marker for mature granulocytes. *Journal of Leukocyte Biology*, 81(2):440–448.

[212] M. J. Kwakkenbos, M. Matmati, O. Madsen, W. Pouwels, Y. Wang, R. E. Bontrop, P. J. Heidt, R. M. Hoek, J. Hamann, M. J. Kwakkenbos, M. Matmati, O. Madsen,

W. Pouwels, Y. Wang, R. E. Bontrop, P. J. Heidt, R. M. Hoek & J. Hamann (2006), An unusual mode of concerted evolution of the EGF-TM7 receptor chimera EMR2. *The FASEB Journal*, 20(14):2582–2584.

[213] S. Yona, H. Lin, P. Dri, J. Q. Davies, R. P. G. Hayhoe, S. M. Lewis, S. E. M. Heinsbroek, K. A. Brown, M. Perretti, J. Hamann, D. F. Treacher, S. Gordon & M. Stacey (2008), Ligation of the adhesion-GPCR EMR2 regulates human neutrophil function. *The FASEB Journal*, 22(3):741–751.

[214] G. W. Chang, J. Q. Davies, M. Stacey, S. Yona, D. M. Bowdish, J. Hamann, T. C. Chen, C. Y. Lin, S. Gordon & H. H. Lin (2007), CD312, the human adhesion-GPCR EMR2, is differentially expressed during differentiation, maturation, and activation of myeloid cells. *Biochemical and Biophysical Research Communications*, 353(1):133–138.

[215] C. H. Huang, W. J. Jeng, Y. P. Ho, W. Teng, Y. C. Hsieh, W. T. Chen, Y. C. Chen, H. H. Lin, I. S. Sheen & C. Y. Lin (2016), Increased EMR2 expression on neutrophils correlates with disease severity and predicts overall mortality in cirrhotic patients. *Scientific Reports*, 6.

[216] S. M. Lewis, D. F. Treacher, J. Edgeworth, G. Mahalingam, C. S. Brown, T. A. Mare, M. Stacey, R. Beale & K. A. Brown (2015), Expression of CD11c and EMR2 on neutrophils: Potential diagnostic biomarkers for sepsis and systemic inflammation. *Clinical and Experimental Immunology*, 182(2):184–194.

[217] S. E. Boyden, A. Desai, G. Cruse, M. L. Young, H. C. Bolan, L. M. Scott, A. R. Eisch, R. D. Long, C.-C. R. Lee, C. L. Satorius, A. J. Pakstis, A. Olivera, J. C. Mullikin, E. Chouery, A. Mégarbané, M. Medlej-Hashim, K. K. Kidd, D. L. Kastner, D. D. Metcalfe & H. D. Komarow (2016), Vibratory Urticaria Associated with a Missense Variant in ADGRE2 . *New England Journal of Medicine*, 374(7):656–663.

[218] I. Caminschi, K. M. Lucas, M. A. O'Keeffe, H. Hochrein, Y. Laâbi, F. Köntgen, A. M. Lew, K. Shortman & M. D. Wright (2001), Molecular Cloning of F4/80-Like-Receptor, a Seven-Span Membrane Protein Expressed Differentially by Dendritic Cell and Monocyte-Macrophage Subpopulations. *The Journal of Immunology*, 167(7):3570–3576.

[219] J. Hamann, M. J. Kwakkenbos, E. C. de Jong, H. Heus, A. S. Olsen & R. A. van Lier (2003), Inactivation of the EGF-TM7 receptor EMR4 after the Pan-Homo divergence. *European Journal of Immunology*, 33(5):1365–1371.

[220] M. Van Pel, H. Hagoort, J. Hamann & W. E. Fibbe (2008), CD97 is differentially expressed on murine hematopoietic stem- and progenitor-cells. *Haematologica*, 93(8):1137–1144.

[221] E. N. Kop, M. Matmati, W. Pouwels, G. Leclercq, P. P. Tak & J. Hamann (2009), Differential expression of CD97 on human lymphocyte subsets and limited effect of CD97 antibodies on allogeneic T-cell stimulation. *Immunology Letters*, 123(2):160–168.

[222] J. C. Leemans, A. a. te Velde, S. Florquin, R. J. Bennink, K. de Bruin, R. a. W. van Lier, T. van der Poll & J. Hamann (2004), The epidermal growth factor-seven transmembrane (EGF-TM7) receptor CD97 is required for neutrophil migration and host defense. *Journal of immunology*, 172(2):1125–1131.

[223] M. Capasso, L. G. Durrant, M. Stacey, S. Gordon, J. Ramage & I. Spendlove (2006), Costimulation via CD55 on Human CD4 + T Cells Mediated by CD97 . *The Journal of Immunology*, 177(2):1070–1077.

[224] H. Veninga, R. M. Hoek, A. F. de Vos, A. M. de Bruin, F. Q. An, T. van der Poll, R. A. van Lier, M. E. Medof & J. Hamann (2011), A novel role for CD55 in granulocyte homeostasis and anti-bacterial host defense. *PLoS ONE*, 6(10).

[225] M. Van Pel, H. Hagoort, M. J. Kwakkenbos, J. Hamann & W. E. Fibbe (2008), Differential role of CD97 in interleukin-8-induced and granulocyte-colony stimulating factor-induced hematopoietic stem and progenitor cell mobilization. *Haematologica*, 93(4):601–604.

[226] J. Hamann, J. O. Wishaupt, R. A. Van Lier, T. J. Smeets, F. C. Breedveld & P. P. Tak (1999), Expression of the activation antigen CD97 and its ligand CD55 in rheumatoid synovial tissue. *Arthritis and Rheumatism*, 42(4):650–658.

[227] R. M. Hoek, D. De Launay, E. N. Kop, A. S. Yilmaz-Elis, F. Lin, K. A. Reedquist, J. S. Verbeek, M. E. Medof, P. P. Tak & J. Hamann (2010), Deletion of either CD55 or CD97 ameliorates arthritis in mouse models. *Arthritis and Rheumatism*, 62(4):1036–1042.

[228] L. Visser, A. F. De Vos, J. Hamann, M. J. Melief, M. Van Meurs, R. A. Van Lier, J. D. Laman & R. Q. Hintzen (2002), Expression of the EGF-TM7 receptor CD97 and its ligand CD55 (DAF) in multiple sclerosis. *Journal of Neuroimmunology*, 132(1-2):156–163.

[229] J. Galle, D. Sittig, I. Hanisch, M. Wobus, E. Wandel, M. Loeffler & G. Aust (2006), Individual cell-based models of tumor-environment interactions: Multiple effects of CD97 on tumor invasion. *American Journal of Pathology*, 169(5):1802–1811.

[230] E. A. Billings, C. S. Lee, K. A. Owen, R. S. D'Souza, K. S. Ravichandran & J. E. Casanova (2016), The adhesion GPCR BAI1 mediates macrophage ROS production and microbicidal activity against Gram-negative bacteria. *Science Signaling*, 9(413).

[231] F. Mazaheri, O. Breus, S. Durdu, P. Haas, J. Wittbrodt, D. Gilmour & F. Peri (2014), Distinct roles for BAI1 and TIM-4 in the engulfment of dying neurons by microglia. *Nature Communications*, 5.

[232] A. M. Fond, C. S. Lee, I. G. Schulman, R. S. Kiss & K. S. Ravichandran (2015), Apoptotic cells trigger a membrane-initiated pathway to increase ABCA1. *Journal of Clinical Investigation*, 125(7):2748–2758.

[233] C. S. Lee, K. K. Penberthy, K. M. Wheeler, I. J. Juncadella, P. Vandenabeele, J. J. Lysiak & K. S. Ravichandran (2016), Boosting Apoptotic Cell Clearance by Colonic Epithelial Cells Attenuates Inflammation In Vivo. *Immunity*, 44(4):807–820.

[234] Y.-M. Peng, M. D. B. van de Garde, K.-F. Cheng, P. A. Baars, E. B. M. Remmerswaal, R. A. W. van Lier, C. R. Mackay, H.-H. Lin & J. Hamann (2011), Specific expression of GPR56 by human cytotoxic lymphocytes. *Journal of Leukocyte Biology*, 90(4):735–740.

[235] C. C. Hsiao, T. Y. Chu, C. J. Wu, M. van den Biggelaar, C. Pabst, J. Hébert, T. W. Kuijpers, B. P. Scicluna, K. Y. I, T. C. Chen, I. Liebscher, J. Hamann & H. H. Lin (2018), The Adhesion G Protein-Coupled Receptor GPR97/ADGRG3 Is Expressed in

Human Granulocytes and Triggers Antimicrobial Effector Functions. *Frontiers in immunology*, 9(2830).

[236] G. W. Chang, C. C. Hsiao, Y. M. Peng, F. A. Vieira Braga, N. A. Kragten, E. B. Remmerswaal, M. D. van de Garde, R. Straussberg, G. M. König, E. Kostenis, V. Knäuper, L. Meyaard, R. A. van Lier, K. P. van Gisbergen, H. H. Lin & J. Hamann (2016), The Adhesion G Protein-Coupled Receptor GPR56/ADGRG1 Is an Inhibitory Receptor on Human NK Cells. *Cell Reports*, 15(8):1757–1770.

[237] T. N. Rao, J. Marks-Bluth, J. Sullivan, M. K. Gupta, V. Chandrakanthan, S. R. Fitch, K. Ottersbach, Y. C. Jang, X. Piao, R. N. Kulkarni, T. Serwold, J. E. Pimanda & A. J. Wagers (2015), High-level Gpr56 expression is dispensable for the maintenance and function of hematopoietic stem and progenitor cells in mice. *Stem Cell Research*, 14(3):307–322.

[238] P. S. Kartalaei, T. Yamada-Inagawa, C. S. Vink, E. de Pater, R. van der Linden, J. Marks-Bluth, A. van der Sloot, M. van den Hout, T. Yokomizo, M. Lucila van Schaick-Solernó, R. Delwel, J. E. Pimanda, W. F. van IJcken & E. Dzierzak (2015), Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *Journal of Experimental Medicine*, 212(1):93–106.

[239] Y. Saito, K. Kaneda, A. Suekane, E. Ichihara, S. Nakahata, N. Yamakawa, K. Nagai, N. Mizuno, K. Kogawa, I. Miura, H. Itoh & K. Morishita (2013), Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. *Leukemia*, 27(8):1637–1649.

[240] F. O. Bagger, D. Sasivarevic, S. H. Sohi, L. G. Laursen, S. Pundhir, C. K. Sønderby, O. Winther, N. Rapin & B. T. Porse (2016), BloodSpot: A database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Research*, 44(D1):D917–D924.

[241] C. Pabst, A. Bergeron, V. P. Lavallée, J. Yeh, P. Gendron, G. L. Norddahl, J. Krosl, I. Boivin, E. Deneault, J. Simard, S. Imren, G. Boucher, K. Eppert, T. Herold, S. K. Bohlander, K. Humphries, S. Lemieux, J. Hébert, G. Sauvageau & F. Barabé (2016), GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. *Blood*, 127(16):2018–2027.

[242] H. R. Saha, K. Kaneda-Nakashima, S. Shimosaki, A. Suekane, B. Sarkar, Y. Saito, H. Ogoh, S. Nakahata, K. Inoue, T. Watanabe, H. Nagase & K. Morishita (2018), Suppression of GPR56 expression by pyrrole-imidazole polyamide represents a novel therapeutic drug for AML with high EVI1 expression. *Scientific Reports*, 8(1):13741.

[243] R. Fredriksson, M. C. Lagerström, P. J. Höglund & H. B. Schiöth (2002), Novel human G protein-coupled receptors with long N-terminals containing GPS domains and Ser/Thr-rich regions. *FEBS Letters*, 531(3):407–414.

[244] J. Hamann, C. C. Hsiao, C. S. Lee, K. S. Ravichandran & H. H. Lin (2016), Adhesion GPCRs as modulators of immune cell function. *Handbook of Experimental Pharmacology*, 234:329–350.

[245] N. Valtcheva, A. Primorac, G. Jurisic, M. Hollmén & M. Detmar (2013), The orphan adhesion G protein-coupled receptor GPR97 regulates migration of lymphatic endothelial cells via the small GTPases RhoA and Cdc42. *Journal of Biological Chemistry*, 288(50):35736–37548.

[246] A. Maglitto, S. A. Mariani, E. de Pater, C. Rodriguez-Seoane, C. S. Vink, X. Piao, M. L. Lukke & E. Dzierzak (2021), Unexpected redundancy of Gpr56 and Gpr97 during hematopoietic cell development and differentiation. *Blood Advances*, 5(3):829–842.

[247] B. P. Sleckman, W. N. Khan, W. Xu, C. H. Bassing, B. A. Malynn, N. G. Copeland, C. G. Bardon, T. M. Breit, L. Davidson, E. M. Oltz, N. A. Jenkins, J. E. Berman & F. W. Alt (2000), Cloning and Functional Characterization of the Early-Lymphocyte-Specific Pb99 Gene . *Molecular and Cellular Biology*, 20(12):4405–4410.

[248] J. J. Wang, L. L. Zhang, H. X. Zhang, C. L. Shen, S. Y. Lu, Y. Kuang, Y. H. Wan, W. G. Wang, H. M. Yan, S. Y. Dang, J. Fei, X. L. Jin & Z. G. Wang (2013), Gpr97 is essential for the follicular versus marginal zone B-lymphocyte fate decision. *Cell Death and Disease*, 4(10).

[249] J. Shi, X. Zhang, S. Wang, J. Wang, B. Du, Z. Wang, M. Liu, W. Jiang, M. Qian & H. Ren (2016), Gpr97 is dispensable for metabolic syndrome but is involved in macrophage inflammation in high-fat diet-induced obesity in mice. *Scientific Reports*, 6(April):1–11.

[250] J. P. Shi, X. N. Li, X. Y. Zhang, B. Du, W. Z. Jiang, M. Liu, J. J. Wang, Z. G. Wang, H. Ren & M. Qian (2015), Gpr97 is dispensable for inflammation in OVA-induced asthmatic mice. *PLoS ONE*, 10(7):1–13.

[251] W. Fang, Z. Wang, Q. Li, X. Wang, Y. Zhang, Y. Sun, W. Tang, C. Ma, J. Sun, N. Li & F. Yi (2018), Gpr97 exacerbates AKI by mediating Sema3A signaling. *Journal of the American Society of Nephrology*, 29(5):1475–1489.

[252] J. Wang, X. Wang, X. Chen, S. Lu, Y. Kuang, J. Fei & Z. Wang (2018), Gpr97/Adgrg3 ameliorates experimental autoimmune encephalomyelitis by regulating cytokine expression. *Acta Biochimica et Biophysica Sinica*, 50(7):666–675.

[253] J. Gupte, G. Swaminath, J. Danao, H. Tian, Y. Li & X. Wu (2012), Signaling property study of adhesion G-protein-coupled receptors. *FEBS Letters*, 586(8):1214–1219.

[254] C. Southern, J. M. Cook, Z. Neetoo-Isseljee, D. L. Taylor, C. A. Kettleborough, A. Merritt, D. L. Bassoni, W. J. Raab, E. Quinn, T. S. Wehrman, A. P. Davenport, A. J. Brown, A. Green, M. J. Wigglesworth & S. Rees (2013), Screening β-arrestin recruitment for the identification of natural ligands for orphan G-protein-coupled receptors. *Journal of Biomolecular Screening*, 18(5):599–609.

[255] N. Harada, Y. Arahori, M. Okuyama, P. B. Luis, A. I. Joseph, T. Kitakaze, N. Goshima, C. Schneider, H. Inui & R. Yamaji (2022), Curcumin activates G protein-coupled receptor 97 (GPR97) in a manner different from glucocorticoid. *Biochemical and Biophysical Research Communications*, 595:41–46.

[256] T.-Y. Chu, C. Zheng-Gérard, K.-Y. Huang, Y.-C. Chang, Y.-W. Chen, K.-Y. I, Y.-L. Lo, N.-Y. Chiang, H.-Y. Chen, M. Stacey, S. Gordon, W.-Y. Tseng, C.-Y. Sun, Y.-M. Wu, Y.-S. Pan, C.-H. Huang, C.-Y. Lin, T.-C. Chen, K. El Omari, M. Antonelou, S. R. Henderson, A. Salama, E. Seiradake & H.-H. Lin (2022), Gpr97 triggers inflammatory processes in human neutrophils via a macromolecular complex upstream of par2 activation. *Nature Communications*, 13(1):6385.

[257] N. Y. Chiang, C. C. Hsiao, Y. S. Huang, H. Y. Chen, I. J. Hsieh, G. W. Chang & H. H. Lin (2011), Disease-associated GPR56 mutations cause bilateral frontoparietal polymicrogyria via multiple mechanisms. *Journal of Biological Chemistry*, 286(16):14215–14225.

[258] S. N. Clemmensen, C. T. Bohr, S. Rørvig, A. Glenthøj, H. Mora-Jensen, E. P. Cramer, L. C. Jacobsen, M. T. Larsen, J. B. Cowland, J. T. Tanassi, N. H. H. Heegaard, J. D. Wren, A. N. Silahtaroglu & N. Borregaard (2012), Olfactomedin 4 defines a subset of human neutrophils. *Journal of Leukocyte Biology*, 91(3):495–500.

[259] B. Korkmaz, S. Attucci, M. A. Juliano, T. Kalupov, M. L. Jourdan, L. Juliano & F. Gauthier (2008), Measuring elastase, proteinase 3 and cathepsin G activities at the surface of human neutrophils with fluorescence resonance energy transfer substrates. *Nature Protocols*, 3(6):991–1000.

[260] A. R. Aricescu, W. Lu & E. Y. Jones (2006), A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica Section D: Biological Crystallography*, 62(10):1243–1250.

[261] M. Fairhead & M. Howarth (2015), Site-specific biotinylation of purified proteins using BirA. *Methods in Molecular Biology*, 1266:171–184.

[262] A. R. Aricescu, W. C. Hon, C. Siebold, W. Lu, P. A. Van Der Merwe & E. Y. Jones (2006), Molecular analysis of receptor protein tyrosine phosphatase μ-mediated cell adhesion. *EMBO Journal*, 25(4):701–712.

[263] D. C. Rio, S. G. Clark & R. Tjian (1985), A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction. *Science*, 227(4682):23–28.

[264] Y. C. Lin, M. Boone, L. Meuris, I. Lemmens, N. Van Roy, A. Soete, J. Reumers, M. Moisse, S. Plaisance, R. Drmanac, J. Chen, F. Speleman, D. Lambrechts, Y. Van De Peer, J. Tavernier & N. Callewaert (2014), Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications*, 5.

[265] F. Maley, R. B. Trimble, A. L. Tarentino & T. H. Plummer (1989), Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Analytical Biochemistry*, 180(2):195–204.

[266] R. B. Trimble & A. L. Tarentino (1991), Identification of distinct endoglycosidase (endo) activities in Flavobacterium meningosepticum: Endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans. *Journal of Biological Chemistry*, 266(3):1646–1651.

[267] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J. J. Prestegard, R. L. Schnaar, H. H. Freeze, J. D. Marth, C. R. Bertozzi, M. E. Etzler, M. Frank, J. F. Vliegenthart, T. Lütteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach, K. F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi & S. Kornfeld (2015), Symbol nomenclature for graphical representations of glycans. *Glycobiology*, 25(12):1323–1324.

[268] S. Neelamegham, K. Aoki-Kinoshita, E. Bolton, M. Frank, F. Lisacek, T. Lütteke, N. O'Boyle, N. H. Packer, P. Stanley, P. Toukach, A. Varki & R. J. Woods (2019), Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology*, 29(9):620–624.

[269] V. T. Chang, M. Crispin, A. R. Aricescu, D. J. Harvey, J. E. Nettleship, J. A. Fennelly, C. Yu, K. S. Boles, E. J. Evans, D. I. Stuart, R. A. Dwek, E. Y. Jones, R. J. Owens & S. J. Davis (2007), Glycoprotein Structural Genomics: Solving the Glycosylation Problem. *Structure*, 15(3):267–273.

[270] P. J. Reeves, N. Callewaert, R. Contreras & H. G. Khorana (2002), Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21):13419–13424.

[271] A. R. Aricescu, R. Assenberg, R. M. Bill, D. Busso, V. T. Chang, S. J. Davis, A. Dubrovsky, L. Gustafsson, K. Hedfalk, U. Heinemann, I. M. Jones, D. Ksiazek, C. Lang, K. Maskos, A. Messerschmidt, S. Macieira, Y. Peleg, A. Perrakis, A. Poterszman, G. Schneider, T. K. Sixma, J. L. Sussman, G. Sutton, N. Tarboureich, T. Zeev-Ben-Mordehai & E. Y. Jones (2006), Eukaryotic expression: Developments for structural proteomics. *Acta Crystallographica Section D: Biological Crystallography*, 62(10):1114–1124.

[272] F. Grueninger-Leitch, A. D'Arcy, B. D'Arcy & C. Chène (1996), Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases. *Protein Science*, 5(12):2617–2622.

[273] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel & A. Bairoch (2005), Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, 571–607.

[274] S. J. Davis, M. J. Puklavec, D. A. Ashford, K. Harlos, E. Y. Jones, D. I. Stuart & A. F. Williams (1993), Expression of soluble recombinant glycoproteins with predefined glycosylation: Application to the crystallization of the t-cell glycoprotein cd2. *Protein Engineering, Design and Selection*, 6(2):229–232.

[275] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter & G. Evans (2018), DIALS: Implementation and evaluation of a new integration package. *Acta Crystallographica Section D: Structural Biology*, 74:85–97.

[276] P. R. Evans (2011), An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):282–292.

[277] P. R. Evans & G. N. Murshudov (2013), How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography*, 69(7):1204–1214.

[278] A. Wagner, R. Duman, K. Henderson & V. Mykhaylyk (2016), In-vacuum long-wavelength macromolecular crystallography. *Acta Crystallographica Section D: Structural Biology*, 72(3):430–439.

[279] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack & G. Bricogne (2011), Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):293–302.

[280] I. Tickle, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, O. Smart, C. Vonrhein & G. Bricogne (2018), STARANISO. Cambridge, UK: Global Phasing Ltd.

[281] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni & R. J. Read (2007), Phaser crystallographic software. *Journal of Applied Crystallography*, 40(4):658–674.

[282] P. W. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P. & W. T. O. Roversi P, Sharff A., Smart O.S., Vonrhein C. (2011), BUSTER version 2.10.3 Cambridge, United Kingdom: Global Phasing Ltd.

[283] O. S. Smart, T. O. Womack, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C. Vonrhein & G. Bricogne (2012), Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallographica Section D: Biological Crystallography*, 68(4):368–380.

[284] K. Cowtan (2010), Recent developments in classical density modification. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):470–478.

[285] A. Thorn & G. M. Sheldrick (2011), ANODE: Anomalous and heavy-atom density calculation. *Journal of Applied Crystallography*, 44(6):1285–1287.

[286] P. Emsley & K. Cowtan (2004), Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12 I):2126–2132.

[287] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long & A. A. Vagin (2011), REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):355–367.

[288] G. Winter (2010), Xia2: An expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, 43(1):186–190.

[289] J. Foadi, P. Aller, D. Axford, R. Owen, Y. Alguel, A. Cameron, W. Armour, D. Waterman & G. Evans (2012), Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallographica Section A Foundations of Crystallography*, 68(a1):s17–s17.

[290] W. Kabsch (2010), *XDS. Acta Crystallographica Section D*, 66(2):125–132.

[291] T. Pape & T. R. Schneider (2004), HKL2MAP : a graphical user interface for macromolecular phasing with SHELX programs . *Journal of Applied Crystallography*, 37(5):843–844.

[292] G. M. Sheldrick (2010), Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):479–485.

[293] T. C. Terwilliger, P. D. Adams, R. J. Read, A. J. McCoy, N. W. Moriarty, R. W. Grosse-Kunstleve, P. V. Afonine, P. H. Zwart & L. W. Hung (2009), Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallographica Section D: Biological Crystallography*, 65(6):582–601.

[294] N. S. Pannu, W. J. Waterreus, P. Skubák, I. Sikharulidze, J. P. Abrahams & R. A. De Graaff (2011), Recent advances in the CRANK software suite for experimental phasing. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):331–337.

[295] T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L. W. Hung, R. J. Read & P. D. Adams (2007), Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D: Biological Crystallography*, 64(1):61–69.

[296] P. Legrand (2017), Xdsme: Xds made easier. *GitHub repository.*

[297] K. Cowtan (2006), The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D*, 62(9):1002–1011.

[298] R. Gupta & S. Brunak (2002), Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 310–322.

[299] C. Steentoft, S. Y. Vakhrushev, H. J. Joshi, Y. Kong, M. B. Vester-Christensen, K. T. Schjoldager, K. Lavrsen, S. Dabelsteen, N. B. Pedersen, L. Marcos-Silva, R. Gupta, E. Paul Bennett, U. Mandel, S. Brunak, H. H. Wandall, S. B. Levery & H. Clausen (2013), Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO Journal*, 32(10):1478–1488.

[300] H. Nielsen (2017), Predicting secretory proteins with signaIP. *Methods in Molecular Biology*, 1611:59–73.

[301] S. Lu, J. Wang, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, G. H. Marchler, J. S. Song, N. Thanki, R. A. Yamashita, M. Yang, D. Zhang, C. Zheng, C. J. Lanczycki & A. Marchler-Bauer (2020), CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research*, 48(D1):D265–D268.

[302] M. Blum, H. Y. Chang, S. Chuguransky, T. Grego, S. Kandasaamy, A. Mitchell, G. Nuka, T. Paysan-Lafosse, M. Qureshi, S. Raj, L. Richardson, G. A. Salazar, L. Williams, P. Bork, A. Bridge, J. Gough, D. H. Haft, I. Letunic, A. Marchler-Bauer, H. Mi, D. A. Natale, M. Necci, C. A. Orengo, A. P. Pandurangan, C. Rivoire, C. J. Sigrist, I. Sillitoe, N. Thanki, P. D. Thomas, S. C. Tosatto, C. H. Wu, A. Bateman & R. D. Finn (2021), The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, 49(D1):D344–D354.

[303] I. Letunic, S. Khedkar & P. Bork (2021), SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Research*, 49(D1):D458–D460.

[304] A. Krogh, B. Larsson, G. Von Heijne & E. L. Sonnhammer (2001), Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, 305(3):567–580.

[305] D. W. Buchan & D. T. Jones (2019), The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*, 47(W1):W402–W407.

[306] G. E. Tusnády & I. Simon (2001), The HMMTOP transmembrane topology prediction server. *Bioinformatics*, 17(9):849–850.

[307] K. Hofmann & W. Stoffel (1993), TMbase: A Database of Membrane Spanning Protein Segments. *Biological Chemistry*, 374:166.

[308] D. T. Jones, W. R. Taylor & J. M. Thornton (1994), A Model Recognition Approach to the Prediction of All-Helical Membrane Protein Structure and Topology. *Biochemistry*, 33(10):3038–3049.

[309] G. S. Salzman, S. D. Ackerman, C. Ding, A. Koide, K. Leon, R. Luo, H. M. Stoveken, C. G. Fernandez, G. G. Tall, X. Piao, K. R. Monk, S. Koide & D. Araç (2016), Structural Basis for Regulation of GPR56/ADGRG1 by Its Alternatively Spliced Extracellular Domains. *Neuron*, 91(6):1292–1304.

[310] K. Leon, R. L. Cunningham, J. A. Riback, E. Feldman, J. Li, T. R. Sosnick, M. Zhao, K. R. Monk & D. Araç (2020), Structural basis for adhesion G protein-coupled receptor Gpr126 function. *Nature Communications*, 11(1).

[311] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. De Beer, C. Rempfer, L. Bordoli, R. Lepore & T. Schwede (2018), SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1):W296–W303.

[312] F. Gabler, S. Z. Nam, S. Till, M. Mirdita, M. Steinegger, J. Söding, A. N. Lupas & V. Alva (2020), Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Current Protocols in Bioinformatics*, 72(1).

[313] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli & D. Hassabis (2021), Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589.

[314] S. Kurosawa, C. T. Esmon & D. J. Stearns-Kurosawa (2000), The soluble endothelial protein C receptor binds to activated neutrophils: involvement of proteinase-3 and CD11b/CD18. *Journal of Immunology*, 165:4697–4703.

[315] Y. C. Kim, J. E. Shin, S. H. Lee, W. J. Chung, Y. S. Lee, B. K. Choi & Y. Choi (2011), Membrane-bound proteinase 3 and PAR2 mediate phagocytosis of non-opsonized bacteria in human neutrophils. *Molecular Immunology*, 48(15-16):1966–1974.

[316] A. Miseta & P. Csutora (2000), Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Molecular Biology and Evolution*, 17(8):1232–1239.

[317] E. Posokhova, A. Shukla, S. Seaman, S. Volate, M. B. Hilton, B. Wu, H. Morris, D. A. Swing, M. Zhou, E. Zudaire, J. S. Rubin & B. St.Croix (2015), GPR124 functions as a WNT7-specific coactivator of canonical β-catenin signaling. *Cell Reports*, 10(2):123–130.

[318] M. Vallon, K. Yuki, T. D. Nguyen, J. Chang, J. Yuan, D. Siepe, Y. Miao, M. Essler, M. Noda, K. C. Garcia & C. J. Kuo (2018), A RECK-WNT7 Receptor-Ligand Interaction Enables Isoform-Specific Regulation of Wnt Bioavailability. *Cell Reports*, 25(2):339–349.e9.

[319] M. Fujinaga, M. M. Chernaia, R. Halenbeck, K. Koths & M. N. James (1996), The crystal structure of PR3, a neutrophil serine proteinase antigen of Wegener's granulomatosis antibodies. *Journal of Molecular Biology*, 261(2):267–278.

[320] Y. Zhang, C. C. Boesen, S. Radaev, A. G. Brooks, W. H. Fridman, C. Sautes-Fridman & P. D. Sun (2000), Crystal structure of the extracellular domain of a human FcγRIII. *Immunity*, 13(3):387–395.

[321] R. K. Cheng, C. Fiez-Vandal, O. Schlenker, K. Edman, B. Aggeler, D. G. Brown, G. A. Brown, R. M. Cooke, C. E. Dumelin, A. S. Doré, S. Geschwindner, C. Grebner, N. O. Hermansson, A. Jazayeri, P. Johansson, L. Leong, R. Prihandoko, M. Rappas, H. Soutter, A. Snijder, L. Sundström, B. Tehan, P. Thornton, D. Troast, G. Wiggin, A. Zhukov, F. H. Marshall & N. Dekker (2017), Structural insight into allosteric modulation of protease-activated receptor 2. *Nature*, 545(7652):112–115.

[322] J. C. Unkeless, Z. Shen, C. W. Lin & E. DeBeus (1995), Function of human FcγRIIA and FcγRIIIB. *Seminars in Immunology*, 7(1):37–44.

[323] C. M. Belham, R. J. Tate, P. H. Scott, A. D. Pemberton, H. R. Miller, R. M. Wadsworth, G. W. Gould & R. Plevin (1996), Trypsin stimulates proteinase-activated receptor-2-dependent and -independent activation of mitogen-activated protein kinases. *Biochemical Journal*, 320(3):939–946.

[324] M. Molino, E. S. Barnathan, R. Numerof, J. Clark, M. Dreyer, A. Cumashi, J. A. Hoxie, N. Schechter, M. Woolkalis & L. F. Brass (1997), Interactions of mast cell tryptase with thrombin receptors and PAR-2. *Journal of Biological Chemistry*, 272(7):4043–4049.

[325] C. J. Kuckleburg & P. J. Newman (2013), Neutrophil proteinase 3 acts on protease-activated receptor-2 to enhance vascular endothelial cell barrier function. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 33(2):275–284.

[326] T. Takeuchi, J. L. Harris, W. Huang, K. W. Yan, S. R. Coughlin & C. S. Craik (2000), Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. *Journal of Biological Chemistry*, 275(34):26333–26342.

[327] H. Lin & J. Trejo (2013), Transactivation of the PAR1-PAR2 heterodimer by thrombin elicits β-arrestin-mediated endosomal signaling. *Journal of Biological Chemistry*, 288(16):11203–11215.

[328] D. Wang, W. Bode & R. Huber (1985), Bovine chymotrypsinogen A. X-ray crystal structure analysis and refinement of a new crystal form at 1.8 Å resolution. *Journal of Molecular Biology*, 185(3):595–624.

[329] J. Sun, D. N. Fass, M. A. Viss, A. M. Hummel, H. Tang, H. A. Homburger & U. Specks (1998), A proportion of proteinase 3 (PR3)-specific anti-neutrophil cytoplasmic antibodies (ANCA) only react with PR3 after cleavage of its N-terminal activation dipeptide. *Clinical and Experimental Immunology*.

[330] Y. M. Van Der Geld, W. Oost-Kort, P. C. Limburg, U. Specks & C. G. Kallenberg (2000), Recombinant proteinase 3 produced in different expression systems: Recognition by anti-PR3 antibodies. *Journal of Immunological Methods*, 244(1-2):117–131.

[331] S. A. Capizzi, M. A. Viss, A. M. Hummel, D. N. Fass & U. Specks (2003), Effects of carboxy-terminal modifications of proteinase 3 (PR3) on the recognition by PR3-ANCA. *Kidney International*, 63(2):756–760.

[332] O. Wiesner, R. D. Litwiller, A. M. Hummel, M. A. Viss, C. J. McDonald, D. E. Jenne, D. N. Fass & U. Specks (2005), Differences between human proteinase 3 and neutrophil elastase and their murine homologues are relevant for murine model experiments. *FEBS Letters*, 579(24):5305–5312.

[333] J. Vanderlocht, J. J. van Beers, P. C. Limburg, J. Damoiseaux & C. Roozendaal (2019), Antigen-specific detection of autoantibodies against myeloperoxidase (MPO) and proteinase 3 (PR3). *Methods in Molecular Biology*, 1901:153–176.

[334] J. Lüdemann, B. Utecht & W. L. Gross (1988), Detection and quantitation of anti-neutrophil cytoplasm antibodies in wegener's granulomatosis by elisa using affinity-purified antigen. *Journal of immunological methods*, 114(1-2):167–174.

[335] L. Halbwachs-Mecarelli, P. Nusbaum, L. Noel, D. Reumaux, S. Erlinger, J. Grünfeld & P. Lesavre (1992), Antineutrophil cytoplasmic antibodies (anca) directed against cathepsin g in ulcerative colitis, crohn's disease and primary sclerosing cholangitis. *Clinical & Experimental Immunology*, 90(1):79–84.

[336] E. C. Hagen, K. Andrassy, E. Csernok, M. R. Daha, G. Gaskin, W. L. Gross, B. Hansen, Z. Heigl, J. Hermans, D. Jayne, C. G. Kallenberg, P. Lesavre, C. M. Lockwood, J. Lüdemann, F. Mascart-Lemone, E. Mirapeix, C. D. Pusey, N. Rasmussen, R. A. Sinico, A. Tzioufas, J. Wieslander, A. Wiik & F. J. Van Der Woude (1996), Development and standardization of solid phase assays for the detection of anti-neutrophil cytoplasmic antibodies (ANCA). A report on the second phase of an international cooperative study on the standardization of ANCA assays. *Journal of Immunological Methods*, 196(1):1–15.

[337] B. Baslund, J. Petersen, H. Permin, A. Wiik & J. Wieslander (1994), Measurements of proteinase 3 and its complexes with α1-proteinase inhibitor and Anti-Neutrophil Cytoplasm Antibodies (ANCA) in plasma. *Journal of Immunological Methods*, 175(2):215–225.

[338] L. Stummann & A. Wiik (1997), A simple high yield procedure for purification of human proteinase 3, the main molecular target of cANCA. *Journal of Immunological Methods*, 206(1-2):35–42.

[339] G. Gaskin, H. Kendal, A. Coulthart, N. Turner & C. D. Pusey (1995), Use of proteinase 3 purified by reverse phase hplc to detect autoantibodies in systemic vasculitis. *Journal of immunological methods*, 180(1):25–33.

[340] M. C. Harmsen, P. Heeringa, Y. M. Van Der Geld, M. G. Huitema, A. Klimp, A. Tiran & C. G. Kallenberg (1997), Recombinant proteinase 3 (Wegener's antigen) expressed in Pichia pastoris is functionally active and is recognized by patient sera. *Clinical and Experimental Immunology*, 110(2):257–264.

[341] C. H. Szymkowiak, T. W. Johnston, E. Csernok & W. L. Gross (1996), Expression of the human autoantigen of Wegener's granulomatosis (PR3) in a baculovirus expression system. *Biochemical and Biophysical Research Communications*, 219(2):283–289.

[342] V. Witko-Sarsat, L. Halbwachs-Mecarelli, R. P. Almeida, P. Nusbaum, M. Melchior, G. Jamaleddine, P. Lesavre, B. Descamps-Latscha & J. E. Gabay (1996), Characterization of a recombinant proteinase 3, the autoantigen in Wegener's granulomatosis and its reactivity with anti-neutrophil cytoplasmic autoantibodies. *FEBS Letters*, 382(1-2):130–136.

[343] Y. M. Van Der Geld, M. L. Smook, M. G. Huitema, M. C. Harmsen, P. C. Limburg & C. G. Kallenberg (2002), Expression of recombinant proteinase 3, the autoantigen in Wegener's granulomatosis, in insect cells. *Journal of Immunological Methods*, 264(1-2):195–205.

[344] U. Specks, D. N. Fass, M. P. Fautsch, A. M. Hummel & M. A. Viss (1996), Recombinant human proteinase 3, the Wegener's autoantigen, expressed in HMC-1 cells is enzymatically active and recognized by c-ANCA. *FEBS Letters*, 390(3):265–270.

[345] D. Garwicz, A. Lindmark, T. Hellmark, M. Gladh, J. Jögi & U. Gullberg (1997), Characterization of the processing and granular targeting of human proteinase 3 after transfection to the rat RBL or the murine 32D leukemic cell lines. *Journal of Leukocyte Biology*, 61(1):113–123.

[346] H. R. F, K. MICHAEL, P. CARL, J. D. A & K. K. E, Recombinant pr-3 and compositions thereof.

[347] Q. Huai, A. P. Mazar, A. Kuo, G. C. Parry, D. E. Shaw, J. Callahan, Y. Li, C. Yuan, C. Bian, L. Chen, B. Furie, B. C. Furie, D. B. Cines & M. Huang (2006), Structure of human urokinase plasminogen activator in complex with its receptor. *Science*, 311(5761):656–659.

[348] K. J. Leath, S. Johnson, P. Roversi, T. R. Hughes, R. A. Smith, L. Mackenzie, B. P. Morgan & S. M. Lea (2007), High-resolution structures of bacterially expressed soluble human CD59. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 63(8):648–652.

[349] Y. Zhao, J. Ren, W. Lu, K. Harlos & E. Y. Jones (2018), Structure of the Wnt signaling enhancer LYPD6 and its interactions with the Wnt coreceptor LRP6. *FEBS Letters*, 592(18):3152–3162.

[350] N. Stein (2008), CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography*, 41(3):641–643.

[351] G. Bunkóczi & R. J. Read (2011), Improvement of molecular-replacement models with Sculptor. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):303–312.

[352] Y. Jiang, L. Lin, S. Chen, L. Jiang, M. C. Kriegbaum, H. Gårdsvoll, L. V. Hansen, J. Li, M. Ploug, C. Yuan & M. Huang (2020), Crystal structures of human C4.4A reveal the unique association of Ly6/uPAR/α-neurotoxin domain. *International Journal of Biological Sciences*, 16(6):981–993.

[353] M. Masutani, S. Sakurai, T. Shimizu & U. Ohto (2020), Crystal structure of TEX101, a glycoprotein essential for male fertility, reveals the presence of tandemly arranged Ly6/uPAR domains. *FEBS Letters*, 594(18):3020–3031.

[354] L. Holm (2020), DALI and the persistence of protein shape. *Protein Science*, 29(1):128–140.

[355] S. Huang, S. X. Li, N. Bren, K. Cheng, R. Gomoto, L. Chen & S. M. Sine (2013), Complex between α-bungarotoxin and an α7 nicotinic receptor ligand-binding domain chimaera. *Biochemical Journal*, 454(2):303–310.

[356] D. S. Kudryavtsev, I. V. Shelukhina, L. V. Son, L. O. Ojomoko, E. V. Kryukova, E. N. Lyukmanova, M. N. Zhmak, D. A. Dolgikh, I. A. Ivanov, I. E. Kasheverov, V. G. Starkov, J. Ramerstorfer, W. Sieghart, V. I. Tsetlin & Y. N. Utkin (2015), Neurotoxins from snake venoms and α-Conotoxin ImI inhibit functionally active Ionotropic γ-aminobutyric acid (GABA) receptors. *Journal of Biological Chemistry*, 290(37):22747–22758.

[357] J. Groppe, C. S. Hinck, P. Samavarchi-Tehrani, C. Zubieta, J. P. Schuermann, A. B. Taylor, P. M. Schwarz, J. L. Wrana & A. P. Hinck (2008), Cooperative Assembly of TGF-β Superfamily Signaling Complexes Is Mediated by Two Disparate Mechanisms and Distinct Modes of Receptor Binding. *Molecular Cell*, 29(2):157–168.

[358] S. L. Lawrence, M. A. Gorman, S. C. Feil, T. D. Mulhern, M. J. Kuiper, A. J. Ratner, R. K. Tweten, C. J. Morton & M. W. Parker (2016), Structural Basis for Receptor Recognition by the Human CD59-Responsive Cholesterol-Dependent Cytolysins. *Structure*, 24(9):1488–1498.

[359] G. Birrane, A. P. Beigneux, B. Dwyer, B. Strack-Logue, K. K. Kristensen, O. L. Francone, L. G. Fong, H. D. Mertens, C. Q. Pan, M. Ploug, S. G. Young & M. Meiyappan (2019), Structure of the lipoprotein lipase–GPIHBP1 complex that mediates plasma triglyceride hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 116(5):1723–1732.

[360] Q. Huai, A. Zhou, L. Lin, A. P. Mazar, G. C. Parry, J. Callahan, D. E. Shaw, B. Furie, B. C. Furie & M. Huang (2008), Crystal structures of two human vitronectin, urokinase and urokinase receptor complexes. *Nature Structural and Molecular Biology*, 15(4):422–423.

[361] W. L. Gross, E. Csernok & W. H. Schmitt (1991), Antineutrophil cytoplasmic autoantibodies: Immunobiological aspects. *Klinische Wochenschrift*, 69(13):558–566.

[362] A. Kuhl, B. Korkmaz, B. Utecht, A. Kniepert, U. Schönermarck, U. Specks & D. E. Jenne (2010), Mapping of Conformational Epitopes on Human Proteinase 3, the Autoantigen of Wegener's Granulomatosis. *The Journal of Immunology*, 185(1):387–399.

[363] F. Silva, A. M. Hummel, D. E. Jenne & U. Specks (2010), Discrimination and variable impact of ANCA binding to different surface epitopes on proteinase 3, the Wegener's autoantigen. *Journal of Autoimmunity*, 35(4):299–308.

[364] G. Lüdemann & W. Gross (1987), Autoantibodies against cytoplasmic structures of neutrophil granulocytes in wegener's granulomatosis. *Clinical and experimental immunology*, 69(2):350.

[365] A. Wiik (1995), Antineutrophil cytoplasmic antibodies in wegener's granulomatosis: some clinical and pathogenetic aspects. *Sem Clin Immunol*, 9:5–16.

[366] J. Granel, R. Lemoine, E. Morello, Y. Gallais, J. Mariot, M. Drapeau, A. Musnier, A. Poupon, M. Pugnière, S. Seren, D. Nouar, V. Gouilleux-Gruart, H. Watier, B. Korkmaz & C. Hoarau (2020), 4C3 Human Monoclonal Antibody: A Proof of Concept for Non-pathogenic Proteinase 3 Anti-neutrophil Cytoplasmic Antibodies in Granulomatosis With Polyangiitis. *Frontiers in Immunology*, 11.

[367] A. J. Roth, J. D. Ooi, J. J. Hess, M. M. Van Timmeren, E. A. Berg, C. E. Poulton, J. A. McGregor, M. Burkart, S. L. Hogan, Y. Hu, W. Winnik, P. H. Nachman, C. A. Stegeman, J. Niles, P. Heeringa, A. R. Kitching, S. Holdsworth, J. C. Jennette, G. A. Preston & R. J. Falk (2013), Epitope specificity determines pathogenicity and detectability in anca-associated vasculitis. *Journal of Clinical Investigation*, 123(4):1773–1783.

[368] A. S. Fauci, B. F. Haynes, P. Katz & S. M. Wolff (1983), Wegener's granulomatosis: Prospective clinical and therapeutic experience with 85 patients for 21 years. *Annals of Internal Medicine*, 98(1):76–85.

[369] J. Pei, B. H. Kim & N. V. Grishin (2008), PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*, 36(7):2295–2300.