University of Oxford

Generating tools for the characterisation of cell surface proteins



MARIA KOKOLAKI

Wolfson College

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Hilary term 2023

Acknowledgements

Throughout my DPhil, I experienced rewarding and challenging moments inside and outside my academic life. Now that my time at Oxford University has come to an end, I realise that nothing would have been possible without the endless support from my colleagues, friends and family.

I would like to start by expressing my gratitude to my supervisor, Prof. Elena Seiradake, for her encouragement and support during my DPhil. She has always been present and supportive while inspiring me and guiding me through my DPhil studies. I would also like to thank my secondary supervisor and director of my DPhil programme, Prof. Matt Higgins, for his support and advice.

Furthermore, I want to thank the Wellcome Trust for funding my studies and Wolfson College for the support in various aspects of my work and life throughout the DPhil. I would also like to thank my college advisor Prof. Christina Redfield for her encouragement and advice, Dr Ed Lowe for his help and guidance on the X-ray crystallography part of my research work and Dr David Staunton for his assistance with CD, SEC-MALS and SPR.

Of course, I would like to thank my supportive and encouraging lab colleagues in the Seiradake group. Amy, Celine, Fina, Jin, Maria, Miguel, Onno and Owen have been a fantastic team, always creating an enjoyable environment in and out of the lab. Special thanks to Maria Carrasquero-Ordaz for her support and patience in teaching me many of the skills and techniques that I needed and for her incredible work at keeping the lab running smoothly. Also, many thanks to my lab partner Miguel Berbeira-Santana for his emotional support and encouragement and for our long conversations.

Most importantly, I would like to express my gratitude to my friends and family for their unconditional love and encouragement throughout these five years. First and foremost, I want to thank my beloved George for always being there for me in all the ups and downs and for supporting me and believing in me in times when it was hard to stay positive. I will always be grateful for his encouragement and love. I would also like to express my gratitude to my family, my parents Haris and Argyro, and my sister Ellie for their love and support while being abroad and away from home. Finally, I owe a big thank you to my friends back home and my friends in Oxford for standing by my side and being an outlet whenever it was needed.

Abstract

Cell migration depends on cell surface molecules that work together to define the precise direction and position of the migrating cell. To understand the cells' communication with their environment, the signaling networks that are formed between cell surface proteins are being explored. The determination of the structures of the individual proteins, as well as the complexes they form, is essential for the comprehension of these signaling networks. Recently, an interaction between the morphogen receptor Glypican 3 (GPC3) and the Uncoordinated-5 receptor D (Unc5D) was discovered. The crystal structure of the complex revealed an octamer of four GPC3 and four Unc5D molecules. Evidence in the literature about the two proteins points to a function of this newly discovered interaction in cell migration.

In my DPhil project, I developed tools to characterise the GPC3-Unc5D interaction and understand its role in cell migration. Nanobodies against GPC3 were characterised, and their effects on the GPC3-Unc5 interaction were investigated using biochemical, biophysical, and structural biology techniques. Two nanobodies were identified as potential research tools, one that enhances and one that inhibits the GPC3-Unc5 interaction. They were effectively employed in functional *in vitro* and *in vivo* experiments, which revealed the importance of the GPC3-Unc5 interaction in neuronal cell migration during cortical development and in neuroblastoma collective cell migration. In addition to the nanobodies, I also contributed to the design and characterisation of non-binding GPC3 and Unc5 mutants. These mutants were used as tools, together with the nanobodies, in functional experiments, and their contribution was critical for understanding the Unc5-GPC3 complex function *in vivo*.

In a side project during my DPhil, I worked on the adhesion G-protein coupled receptor (GPCR), VLGR1, the largest member of the GPCR family. VLGR1 has functions in cell-cell communication through its intracellular signaling and adhesion properties, and data indicate its important role in the central nervous system development. In this project, I worked towards the structural characterisation of this large adhesion GPCR. The protein construct design, expression and purification for this project are presented in this thesis.

Table of Contents

List of Figures ..... ***xi***

List of Tables ..... ***xiv***

List of Abbreviations ..... ***xv***

1. Introduction ***1***

1.1 Cell migration and cortex development 2

1.2 Essential cues for neuronal cell migration 4

1.3 Glypicans 6

1.3.1 Domain architecture and structures of the Glypican family 6

1.3.2 Glypican 3 (GPC3) 8

1.3.3 Developmental functions of GPC3 9

1.3.4 GPC3 and cancer 11

1.3.5 GPC3 as a diagnostic and therapeutic target 12

1.4 The Uncoordinated-5 (Unc5) receptors 14

1.4.1 Domain architecture and structures 14

1.4.2 The Unc5 receptors: signaling and functions 15

1.4.3 The Unc5 receptors in the dependence receptors family 17

1.5 The Unc5-GPC3 interaction 18

1.5.1 The octameric complex 18

1.5.2 The GPC3-Unc5 interaction elicits repulsion *in vitro* 21

1.5.3 The role of the GPC3-Unc5 interaction in radial neuronal migration 22

1.5.4 The role of the GPC3-Unc5 interaction in neuroblastoma cell migration 23

1.6 Nanobodies 28

1.6.1 Nanobodies: small antigen-binding fragments 28

1.6.2 Nanobodies' favourable properties and applications 29

1.7 VLGR1: an adhesion GPCR 32

1.7.1 Adhesion G protein-coupled receptors (ADGRs) 32

1.7.2 The VLGR1 domain architecture 34

1.7.3 The VLGR1 expression pattern, functions, and related disease 35

1.8 Research aims 37

1.9 Results summary 38

2. Materials and methods 39
2.1 Molecular biology and cloning 39
2.1.1 Constructs 39
2.1.2 Expression vectors and vector preparation 40
2.1.3 Construct generation 42
2.1.4 Recombinant plasmid DNA generation 42
2.2 Protein expression and purification 44
2.2.1 Mammalian cell cultures 44
2.2.2 Small-scale transient transfection 44
2.2.3 Large-scale cell preparation and transient transfection 45
2.2.4 Protein Biotinylation 46
2.2.5 Bacterial cell cultures and nanobody protein expression 46
2.2.6 Periplasmic protein extraction 47
2.2.7 Protein purification 48
2.2.8 Nanobody tetramerisation with streptavidin 49
2.2.9 SDS-PAGE gel electrophoresis and Western blot 49
2.2.10 Buffer chemicals and reagents 51
2.3 Biophysical and biochemical methods 51
2.3.1 Surface plasmon resonance (SPR) 51
2.3.2 Small scale pull-downs 53
2.3.3 Circular Dichroism Spectrophotometry (CD) 54
2.3.4 SEC-MALS 54
2.3.5 Mass spectrometry data processing after brain lysate pull-downs 55
2.4 X-ray crystallography 55
2.4.1 Protein crystallisation and crystal optimisation 55
2.4.2 Data collection and processing 56
2.4.3 AlphaFold and model preparation for Molecular Replacement (MR) 57
2.4.4 Molecular Replacement (MR) and Refinement 58
2.4.5 VMXm sample preparation and crystal screening 59
2.5 Cell Biology 59
2.6 Cryo-EM 61
2.6.1 Crosslinking 61
2.6.2 Sample and grid preparation and screening 61

3. Characterisation of nanobodies targeting GPC3 63
3.1 Introduction and Summary 63
3.2 Production and preparation of the nanobodies 64
3.3 Nanobody and GPC3 binding 66
3.4 Nb49 and Nb58 can manipulate the GPC3-Unc5 interaction 69
3.4.1 Pull-down experiments 69
3.4.2 SPR experiments 70
3.5 GPC3-Unc5-Nb49 complex stoichiometry 73
3.6 Mouse cortex lysate pull-down experiments 75
3.7 More nanobodies against GPC3 77
3.8 Discussion 81
3.8.1 The utilisation of the nanobodies in stripe assays 82
3.8.2 GPC3 immunostaining on mouse cortex using Nb49 83
3.8.3 *In vivo* expression of the Nb49 and Nb58 in the mouse cortex 84
3.8.4 The effect of the Nb49 and Nb58 in the neuroblastoma cell migration .. 86
4. Structural characterisation of the nanobody-GPC3 complexes 89
4.1 Introduction and summary 89
4.2 Crystallisation of the Nb49-GPC3 and Nb58-GPC3 complexes 89
4.2.1 Initial crystals available 90
4.2.2 New crystallisation trials of the GPC3-Nb49 complex. 92
4.2.3 Crystallisation trials of the Nb58-GPC3 complex 94
4.3 Structure determination of the GPC3-Nb49 complex 95
4.3.1 Data processing 95
4.3.2 Molecular Replacement using human GPC3 structure (7ZAW) as the search model 97
4.3.3 AlphaFold multimer prediction of the GPC3-Nb49 complex 98
4.3.4 Molecular Replacement using the AlphaFold Nb49-GPC3 complex structure as a search model 101
4.3.5 Refinement 102
4.3.6 Assessing the presence of only one nanobody in the asymmetric unit 105
4.4 Cryo-EM trials for the GPC3-Unc5-Nb49 complex structure determination 110
4.4.1 Sample grid screening 110
4.4.2 Crosslinking trials of the GPC3-Unc5 complex 112

4.5 Discussion 114
4.5.1 Validation of the Nb49-GPC3 binding interface 115
4.5.2 How does Nb49 enhance the GPC3-Unc5 interaction? 115

5. Generation and characterisation of Unc5 and GPC3 non-binding mutants 117

5.1 Introduction and summary 117
5.2 Mutation design 118
5.2.1 Mutation design on Unc5B 118
5.2.2 Mutation design on human GPC3 120
5.3 Initial characterisation of the mutants in cell binding assays 121
5.4 Further experiments for the mutants' characterisation 124
5.5 Validation of the mutants' correct folding 126
5.6 Discussion 127

6. VLGR1 131

6.1 Introduction and summary 131
6.2 Bioinformatic analysis of the VLGR1 domain architecture 131
6.3 Initial construct design, expression and purification 133
6.4 Optimisation of the purification 136
6.5 New construct design 137
6.6 Discussion 141

7. Discussion and future plans 143

7.1 Non-binding GPC3 and Unc5 mutants 143
7.2 Nanobodies against GPC3 144
7.2.1 Biochemical and biophysical characterisation 144
7.2.2 Structural characterisation of the Nanobody-GPC3 complexes 146
7.2.3 Applications of the anti-GPC3 nanobodies 148

Appendix A: Primers used in this work 153

Appendix B: Unc5 sequences alignment 159

Appendix C: GPC3 sequences alignment 163

Appendix D: Nanobody CDR3 sequences alignment and classification 165

References 167

List of Figures

Figure 1.1: A schematic of the radial cortical migration. 3
Figure 1.2. The structure of GPC3. 8
Figure 1.3: The Unc5 receptors architecture and structure. 15
Figure 1.4: The octameric structure of the GPC3-Unc5 complex. 20
Figure 1.5: Stripe assays show that the GPC3-Unc5 interaction has a contact-repulsion effect. 21
Figure 1.6: The GPC3-Unc5 interaction regulates the radial neuronal migration. 25
Figure 1.7: GPC3 siRNA affected the neuroblastoma cells migration and tumour formation. 26
Figure 1.8: The effect of the Unc5^{IgIgTSP} and non-GPC3-binding Unc5^{IgIgTSP} mutant expression in neuroblastoma cell migration. 27
Figure 1.9: Nanobodies 29
Figure 1.10: Domain architecture of the VLGR1b. 35
Figure 2.1: The mammalian expression vector, pHLsec. 40
Figure 3.1: Purification of Nb49 and Nb58 from WK6 bacterial cells. 64
Figure 3.2: Expression, biotinylation and purification of the nanobodies in HEK-293T. 65
Figure 3.3: Tetramerisation of Nb49 with Streptavidin-Alexa 594. 66
Figure 3.4: Nanobody - GPC3 binding. 67
Figure 3.5: Nb49 and Nb58 do not bind Unc5 in SPR test injections. 68
Figure 3.6: Pull-down assays show that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction. 69
Figure 3.7: SPR experiment confirms that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction. 71
Figure 3.8: Another SPR experiment confirms that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction. (Figure in previous page p. 72) 73
Figure 3.9: SEC-MALS experiment for the GPC3-Nb49 and the GPC3-Unc5-Nb49 complexes. 74
Figure 3.10: Pull-down experiment from mouse cortical lysate. 76
Figure 3.11: A pool of nanobodies against GPC3. 78
Figure 3.12: The nanobodies against GPC3 do not bind Unc5. 79

Figure 3.13: A few new nanobodies are interesting with respect to the GPC3-Unc5 interaction. 80
Figure 3.14: More nanobodies against GPC3 and how they affect the GPC3-Unc5 interaction. 80
Figure 3.15: Stripe assays show that Nb49 enhances and Nb58 inhibits the repulsive effect of the GPC3-Unc5 interaction. 83
Figure 3.16: Immunostaining of GPC3 using Nb49 on a coronal section of the E15.5 mouse cortex. 84
Figure 3.17: Nb49 and Nb58 affect the radial migration in the mouse cortex. 85
Figure 3.18: The expression of Nb49 and Nb58 in the neuroblastoma cells of this migration model affects their migration. 87
Figure 4.1: The hGPC3^{core} and Nb49 purifications. 90
Figure 4.2: Initial crystals of the GPC3-Nb49 complex. 91
Figure 4.3: Small crystals on VMXm. 93
Figure 4.4: Nb49^{short} was used in crystallisation trials. 94
Figure 4.5: GPC3-Nb58 crystals that did not diffract. 95
Figure 4.6: MR using the hGPC3^{core} crystal structure model. 98
Figure 4.7: AlphaFold model of human GPC3. 99
Figure 4.8: AlphaFold Multimer GPC3-Nb49 complex prediction models. 100
Figure 4.9: Alphafold predicted model of the GPC3-Nb49 complex. 101
Figure 4.10: The GPC3-Nb49 complex crystal structure. 104
Figure 4.11: Nanobody density in the *P*1 space group. 106
Figure 4.12: Nanobody density in the $P2_1$ space group. 107
Figure 4.13: Nanobody density in the $P2_122_1$ space group. 108
Figure 4.14: Cryo-EM screening for the GPC3-Unc5-Nb49 complex. 112
Figure 4.15: Crosslinking trials of the GPC3-Unc5 complex. 114
Figure 4.16: N-glycosylation mutations on the binding site. 115
Figure 4.17: The GPC3-Nb49 complex aligned in the octameric GPC3-Unc5 complex. 116
Figure 5.1: Design of the Unc5 mutants. 119
Figure 5.2: Design of the GPC3 mutants. 120
Figure 5.3: Cell surface expression of the full-length Unc5 WT and mutants. 121
Figure 5.4: The Unc5B mutants disrupt the GPC3 binding. 123
Figure 5.5: Additional experiments for the confirmation of the mutants' efficacy. 125

Figure 5.6: Data to support that the mutant proteins are correctly folded. 127
Figure 6.1: Predicted domain architecture of the mVLGR1b. 133
Figure 6.2: Initial mVLGR1b construct design and expression. 134
Figure 6.3: Purification of construct #7. 135
Figure 6.4: Small-scale His pull-downs to compare different histidine tags. 136
Figure 6.5: Design of new mVLGR1b constructs. 138
Figure 6.6: Large-scale purification of the newly designed mVLGR1b constructs. 139
Figure 6.7: AlphaFold database predicted models of VLGR1 fragments. 140
Figure B1: Unc5 ectodomain sequences alignment 161
Figure C1: GPC3 sequences alignment 164
Figure D1: Nanobodies CDR3 alignment 166

List of Tables

Table 2.1: Protein constructs used in this thesis's chapters 3, 4, 5. 39
Table 2.2: Buffer chemicals and reagents 51
Table 4.1: Crystallisation conditions of the initial GPC3-Nb49 crystals. 91
Table 4.2: Data collection statistics comparison of dataset #1 and dataset #2 96
Table 4.3: Data collection and refinement statistics for GPC3-Nb49 complex. .. 103
Table 4.4: Comparison of the data processing statistics in $P1$, $P2_1$ and $P2_122_1$. 109
Table 4.5: Constructs and optimisation conditions of the cryo-EM trials. 111
Table A1: Summary of the primers for the constructs cloned in this work (Chapters 3-5)..... 154
Table A2: Summary of the primers for the constructs cloned in this work (Chapter 6)..... 155

List of Abbreviations

7TM	**7 T**rans**m**embrane region
ADGR	**Ad**hesion **G**-protein-coupled **R**eceptor
AP	**A**pical **P**rogenitor
ASU	**As**ymmetric **U**nit
BSA	**B**ovine **S**erum **A**lbumin
$CC_{1/2}$	**C**orrelation **C**oefficient between random half datasets
CD	**C**ircular **D**ichroism
CDR	**C**omplementarity-**D**etermining **R**egion
CP	**C**ortical **P**late
CTF	**C**-**t**erminal **F**ragment
DCC	**D**eleted in **C**olorectal **C**arcinoma
DLP	**D**ally-**l**ike **P**rotein
DMEM	**D**ulbecco's **M**odified **E**agle's **M**edium
E	**E**mbryonic day
EAR	**E**pilepsy-**a**ssociated **R**epeat
ECD	**E**xtra**c**ellular **D**omain
ELISA	**E**nzyme-**l**inked **I**mmunosorbent **A**ssay
EPTP	**E**pitempin
FBS	**F**etal **B**ovine **S**erum
FLRT	**F**ibronectin **L**eucine-**R**ich **T**ransmembrane protein
GAG	**G**lycos**a**mino**g**lycan
GAIN	**G**PCR **A**utoproteolysis-**In**ducing
GPC	**G**ly**pi**can
GPC3	**G**ly**pi**can **3**
GPCR	**G**-**p**rotein-**c**oupled **R**eceptor
GPI	**G**lyco**p**hosphatidyl**i**nositol
GPS	**G**PCR **P**roteolysis **S**ite
HCC	**H**epato**c**ellular **C**arcinoma
HEK	**H**uman **E**mbryonic **K**idney
Hh	**H**edge**h**og
HRP	**H**orse **R**adish **P**eroxidase

HS	**H**eparan **S**ulfate
HSGAG	**H**eparan **S**ulfate **G**lycos**a**mino**g**lycan
HSPG	**H**eparan **S**ulfate **P**roteo**g**lycans
ICD	**I**ntra**c**ellular **D**omain
Ig	**I**mmuno**g**lobin-like
IZ	**I**ntermediate **Z**one
LamG	**Lam**inin **G**
LLG	**L**og **L**ikelihood **G**ain
LPHN	**L**atro**ph**ili**n**
mAU	**m**illi **A**bsorbance **U**nit
MR	**M**olecular **R**eplacement
MW	**M**olecular **W**eight
Nb	**N**ano**b**ody
NTF	**N**-**t**erminal **F**ragment
PBS	**P**hosphate **B**uffer **S**aline
PCR	**P**olymerase **C**hain **R**eaction
PDB	**P**rotein **D**ata **B**ank
pLDDT	**p**redicted **L**ocal **D**istance **D**ifference **T**est
Ptc	**P**a**tc**hed
PTX	**P**en**t**ra**x**in
RU	**R**esponse **U**nit
SDS-PAGE	**S**odium **D**odecyl **S**ulfate - **P**oly**a**crylamide **G**el **E**lectrophoresis
SEC	**S**ize **E**xclusion **C**hromatography
SEC-MALS	**SEC** - **M**ultiangle **L**ight **S**cattering
SGBS	**S**impson - **G**olabi - **B**ehmel over-growth **S**yndrome
Smo	**Smo**othened
SPR	**S**urface **P**lasmon **R**esonance
SVZ	**S**ub**v**entricular **Z**one
TFZ	**T**ranslation **F**unction **Z**-Score
TSP	**T**hrombo**sp**ondin
Unc5	**Unc**oordinated-**5**
VLGR1	**V**ery **L**arge **G**-protein-coupled **R**eceptor **1**
VZ	**V**entricular **Z**one
WT	**W**ild **T**ype

1. Introduction

After fertilisation, the development of a new organism begins. In a multicellular organism, embryonic development is characterised by rapid cell division but also extensive cell rearrangement and migration. Cell migration is a crucial step in development. Except for embryonic development and morphogenesis, cell migration is also present in an adult organism in events like wound healing, immune response, cancer metastases and angiogenesis.

Cells detect various signals in their environment, which dictate their behaviour. In cell migration, cues that provoke attraction or repulsion are needed for the precise placement of the migrating cell. Cell migration can be separated into single-cell and collective cell migration. In single-cell migration, cytoskeleton reconstruction, cell polarity and mobilisation follow the detection of the relevant cue by a cell [1]. In collective cell migration, a group of two or more cells can migrate in clusters together. This type of migration is based on the same principles as the single-cell migration. However, in this case, the cells stay connected as they move, and the polarity and cytoskeleton reconstruction apply to this body of cells. Collective cell migration is particularly important in organ morphogenesis but also in tissue regeneration, wound healing and cancer [2].

1.1 Cell migration and cortex development

The cerebral cortex is the largest part of the mammalian brain and is responsible for cognitive functions. During cortical development, the proliferation of the neuronal progenitors and their migration take place. Accurate and precise neuronal migration and cortical wiring are required for the correct development of the complex cortical structure. Dysregulation of the migration can lead to cortical malformations, which have been linked with neuropsychiatric disorders [3,4].

Cortical neuronal migration is an example of single-cell migration. Along with neuronal cell migration, axonal and dendrite growth and guidance also take place in cortical wiring. Axons and dendrites (neuronal processes) are projections from the neuron's cell body (soma) used in signal transmission. During cortical development, pyramidal neurons leave the lower layers of the cortex (ventricular zone, subventricular zone) where they were born and migrate towards the developing upper layers and the final layer, the cortical plate. Cortical migration is separated into radial migration and tangential dispersion, both very important for the right formation of the cortex [5]. Three ways of radial neuronal migration have been observed during cortical development. These are locomotion, somal translocation [6] and multipolar migration [7]. In locomotion, neurons with bipolar morphology migrate from the ventricular and subventricular zones towards the developing cortex using the scaffold that the apical radial glial cell fibres form from the ventricular zone until the cortical plate. In somal translocation, the neuron's soma moves on the already leading process, which is located further out radially, independently from the apical radial glial cell scaffold [6]. During multipolar migration, the initial phase of radial migration, a multipolar neuron that presents many processes, senses its environment and moves towards or away from specific cues. When newborn

multipolar neurons reach the intermediate zone, they switch to bipolar and continue to migrate towards the cortical plate [7]. The radial cortical migration follows an “inside out” arrangement in which the early-born neurons from the germinal zone migrate and form the first new layers above the ventricular zone, and the neurons that are born later migrate through and beyond these early-born neurons and position in the superficial cortical layers [8]. The apical radial glial cells that contribute to the neuronal radial migration are a subtype of the apical progenitors and are born at the beginning of neurogenesis [9].

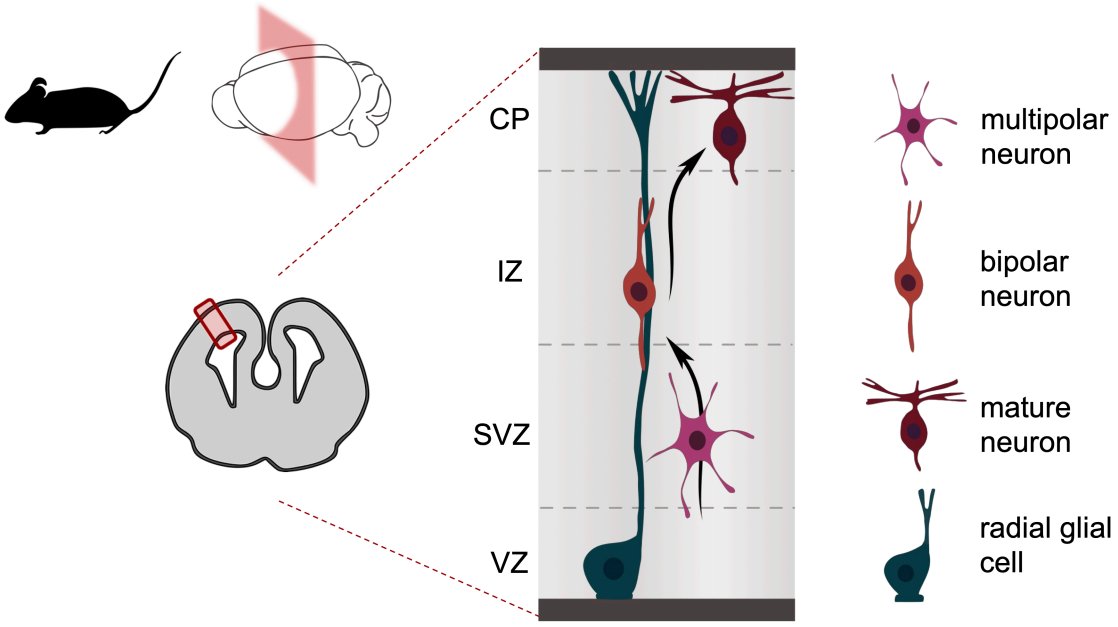


Figure 1.1: A schematic of the radial cortical migration.
A schematic representation of the cortical layers of a mouse cortex coronal section. Pyramidal neurons born from the germinal layers (ventricular zone) are initially multipolar. They migrate from the ventricular zone (VZ) radially through the subventricular zone (SVZ), and when they reach the intermediate zone (IZ), they transition to a bipolar morphology, and they use the fibres of the radial glial cells / apical progenitors as a scaffold to move towards the cortical plate (CP) and their final position. Adapted from [5].

1.2 Essential cues for neuronal cell migration

The brain's normal development and correct wiring are critical for its vital function. Like the migration of any type of cell, the migration of the neurons during brain development is precisely directed by extracellular cues. These cues are either diffusible molecules, molecules present on the cell surface or components of the extracellular matrix and can act in short or long-range manner. The migrating neurons detect these cues. The cues modulate the target signaling pathways and cause cytoskeletal rearrangements that move the neurons toward their destination. Except for cell migration, neurons' axons and dendrites use these cues for their pathfinding. In order to function and guide neurons and axons, guidance cues act on the neuronal growth cones and regulate their growth. Growth cones are dynamic structures at the tip of neuronal axons that sense the extracellular guidance cues and move and expand towards the desired direction [10].

Cell adhesion molecules and cell guidance molecules and their receptors mediate the essential attraction and repulsion which is needed during neuronal migration and axonal guidance. The Cell Adhesion Molecules (CAMs) are cell surface proteins that are either transmembrane receptors or attached to the membrane with a glycophosphatidylinositol (GPI) anchor, and provide stability to stationary cells and the essential force for migrating cells to move. The cadherins and the integrins, typical examples of CAMs, form homophilic or heterophilic interactions [11]. As an example, the homophilic interaction of the NF-protocadherin in the optic tract between the retinal ganglion cells and the neuroepithelium regulates the pathfinding of the axons of the retinal neurons [12].

Furthermore, four groups of cell guidance molecules, the Netrins, the Ephrins, the Semaphorins and the Slits, known as the "canonical cues", regulate neuronal

migration and navigation of the axonal growth cones in a repulsive or attractive way [10]. For instance, the Slit proteins act as repulsive axonal guidance cues when bound to their receptor, Robo [13,14]. The way that these guidance cues direct neuronal migration is very complex. A ligand can act differently depending on the receptor it will interact with. The Netrins, well-characterised cell guidance molecules, present chemoattractive activity when they interact with their receptors in the DCC (Deleted in Colorectal Carcinoma) family. However, they also mediate a repulsive effect when interacting with the Uncoordinated-5 (Unc5) receptors [10]. Besides the “canonical cues”, morphogens and growth factors also play a role in neuronal cell migration and axon guidance [10]. The morphogen Wnt4 presents a decreasing anterior-to-posterior gradient in the spinal cord. The gradience of Wnt4’s concentration and the signaling through its receptor, Frizzled, have been shown to guide the commissural axons in the anterior-posterior axis after crossing the midline [15].

In this thesis, tools have been developed for use in the characterisation of the GPC3-Unc5 protein interaction. The crystal structure of the GPC3-Unc5 complex revealed an octamer. Mutations on the GPC3 and the Unc5 proteins were designed, and nanobodies against GPC3 were characterised in order to be used as tools to understand the function of this interaction. Firstly, an introduction to the individual components of this complex is necessary.

1.3 Glypicans

Glypicans are cell surface proteins that belong to the heparan sulfate proteoglycans (HSPGs), whose main characteristic is the presence of a variable number of heparan sulfate (HS) glycosaminoglycan (GAG) polysaccharide chains. The glypican family in mammals has six members (GPC1-6) [16]. Based on their sequence identity, the six members of the glypican family are put into groups. Glypicans 1, 2, 4 and 6 belong to one group and glypicans 3 and 5 in the other [17]. Drosophila has two orthologs for the glypican family proteins: the Dally and the Dally-like protein (DLP) [18,19]. Based on the sequence identity, they belong to the glypican 3/5 and the glypican 1/2/4/6 subfamilies, respectively [16].

1.3.1 Domain architecture and structures of the Glypican family

The protein architecture of all the glypicans comprises an N-terminal core protein structure followed by a C-terminal disordered region which accommodates the attachment sites of the GAGs. The number of GAGs varies between the glypican family members, and GPC3 contains two of these HS-GAGs attachment sites [16]. Heparan sulfate GAGs are present in all glypicans except for GPC5, which can have both heparan and chondroitin sulfate GAGs [20]. A glycosyl-phosphatidylinositol

(GPI) anchor is located at the C-terminal end of the disordered region, and it tethers the glypican proteins to the cell membrane [16].

The core protein of all the glypican family members has 14 highly conserved cysteines, suggesting a highly similar three-dimensional structure [16]. A furin-like convertase cleavage site is also present in the glypican core protein. After cleavage, the two produced core protein subunits remain covalently attached [21]. The structure of the core protein of the human GPC1 (PDB: 4YWT, 4ACR), solved by X-ray crystallography, revealed an α-helical architecture and 7 disulphide bonds that the 14 cysteine residues form [22,23]. The furin-like cleavage site is located in a long, disordered loop between two α-helices. The multiple cysteines, the disulfide bonds and the furin-like cleavage site comprise three main characteristics of the GPC1 core protein. The core protein can be separated into three regions based on these characteristics. The Cys-rich lobe, where 12 out of the 14 cysteine residues are located. The protease-site lobe where the furin-like cleavage site is, and the central lobe in the middle of the other two. The Drosophila glypican ortholog Dally-like protein (DLP) structure (PDB: 3ODN) also revealed a helical structure similar to the GPC1 [24].

Because of the GPI anchor, glypicans were thought only to be tethered to the cell membrane. However, GPC3 can also be found in the extracellular environment [25–27]. It is believed that GPC3 can be released from the membrane by a lipase. The mammalian Notum, a phospholipase, can cleave and release the glypican from the cell membrane [28]. However, it was reported in 2015 that Notum does not cleave the GPI anchor of glypicans in order to release the soluble protein [29].

1.3.2 Glypican 3 (GPC3)

The structure of the morphogen receptor and member of the glypican family, GPC3, was recently solved, and an α-helical fold was revealed like the previously solved GPC1 and DLP structures. In 2022, Akkermans *et al.* [30] presented the crystal structures of the human and mouse GPC3 core protein whose sequences are 94% identical (Figure 1.2). The two structures were very similar. The main difference between GPC3 and the previously solved structure of GPC1 is the more bent shape of the core protein. N-linked glycosylation motifs were also predicted on the GPC3 molecule. The structures of both human and mouse GPC3 have two N-linked glycans modelled next to the predicted asparagine residues [30].

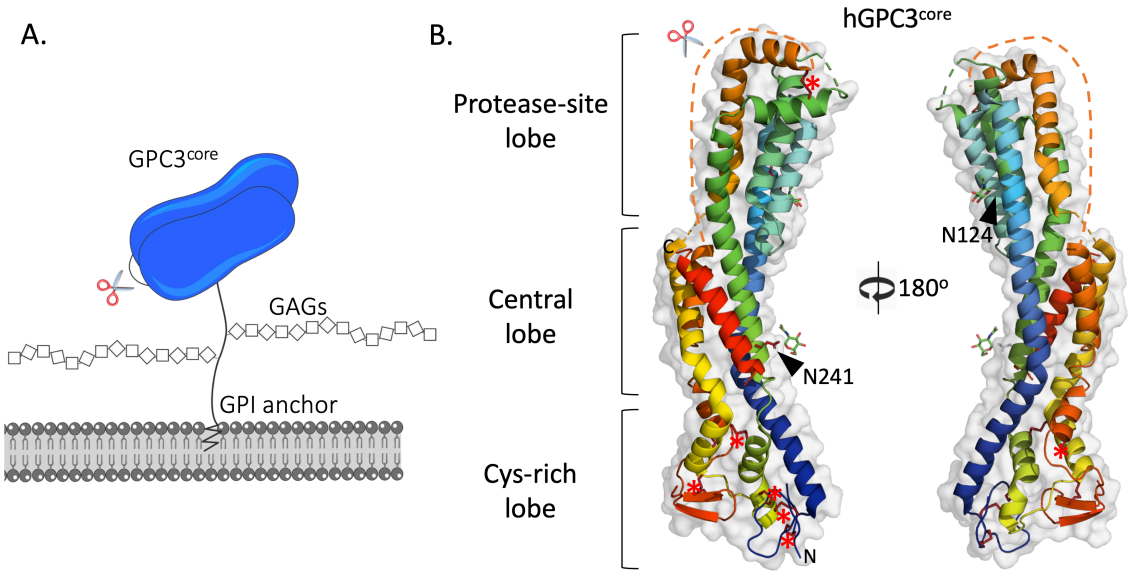


Figure 1.2. The structure of GPC3.
(A) A schematic representation of the architecture of the GPC3 protein. GPC3 is attached to the cell membrane with the GPI anchor, and two GAGs are attached to the disordered region close to the C-terminal. The core protein is coloured blue, and a scissors icon highlights the cleavage site. (B) The structure of the human GPC3 in rainbow representation from blue (N-terminal) to red (C-terminal). Red asterisks highlight the disulphide bonds. The asparagine residues with a glycan attached are also highlighted (N241, N124). Structural regions have been annotated based on the nomenclature presented in [22] since the key characteristics of the structure remain the same as in GPC1. Figure adapted from [30].

GPC3 is broadly expressed during embryonic development in a tissue and stage-specific manner. In adult tissue, GPC3 is not present or lowly expressed in specific tissues [31]. While the expression in adult tissue is low, GPC3 is highly expressed in different cancers like hepatocellular carcinoma (HCC) [32]. The expression of GPC3 in embryonic development, in many adult malignancies and in many embryonal tumours [33,34], defines GPC3 as an oncofetal protein. Recent data have also presented a more detailed expression pattern of GPC3 in the developing mouse brain. The mRNA and protein of GPC3 are mainly present in the ventricular zone and the apical progenitor cells of the developing mouse cortex [30].

1.3.3 Developmental functions of GPC3

Mutations on the GPC3 gene cause the genetic disorder, Simpson-Golabi-Behmel over-growth syndrome (SGBS) [35]. SGBS is an overgrowth syndrome, the patients of which present with visceral and skeletal abnormalities but also an increased risk of tumour development and mild/moderate intellectual deficiency. The expression pattern of GPC3 in a developing organism correlates with the SCBS syndrome phenotype. It is expressed in tissues in which its absence (SCBS syndrome) causes organ overgrowth like the mesoderm-derived tissues (cartilage, bones, head mesenchyme) [36]. At a molecular level, GPC3 has been shown to interact with different growth factors, morphogens and their receptors and modulate their signaling pathways. These molecules include the Wnts [37], the Hedgehogs (Hhs) [38] and the Bone Morphogenic Proteins (BMPs) [39].

The regulation of the Wnt pathway by GPC3 has been thoroughly studied. In the “canonical” Wnt/β-catenin pathway, the first step is the binding of the Wnt ligands to their receptor, Frizzled, followed by the activation of the subsequent signaling pathway. GPC3 facilitates this first step and the binding between Wnt and Frizzled

by interacting with Frizzled through its HS-GAG chains and with Wnt through its core protein. The HS-GAG chains of GPC3 are not required for its binding to Wnt. This binding stabilisation stimulates the activity of the Wnt signaling pathway [27,37,54]. The GPC3 cleavage step at the furine-like cleavage site is not essential for the regulation of the Wnt pathway [40]. When GPC3 is not tethered on the cell surface, but is soluble in the extracellular environment, it acts as a competitive inhibitor of the Wnt signaling [27]. The same mechanism has been observed in Drosophila as well. The Dally-like protein (Drosophila homologue of glypicans) was shown to bind to Wingless (a Drosophila Wnt) and stimulate the “canonical” Wnt signaling [41].

GPC3 also regulates the Hedgehog (Hh) signaling pathway. In this pathway, when the Hh ligand is not present, its receptor, Patched (Ptc), suppresses the GPCR Smoothened (Smo) activity, and consequently, the signal transduction. In the active state of the Hedgehog pathway, Hh binds to Ptc, and Ptc does not repress the Smoothened activity, which mediates downstream signaling. GPC3 inhibits the Hh signaling since it binds Hh. As a result of the GPC3-Hh binding, less Hh is available for Ptc and Ptc suppresses the Smo activity. Therefore, the Hh signaling is inhibited. For the binding between Hh and GPC3, the core protein of GPC3 is sufficient, and the HS-GAG chains are not required [42]. When GPC3 binds Hh, endocytosis and degradation of the complex are induced. This reduces the Hh presence and availability for the Ptc receptor [43]. Additionally, the above normal activation of the Hh pathway when GPC3 is not present has been linked with the SGBS syndrome, and it is at least partially responsible for the overgrowth phenotype [42,44]. In contrast to the Wnt pathway, the Hh pathway requires the GPC3 processing by the furin-like convertases for its inhibition [45].

1.3.4 GPC3 and cancer

GPC3, as an oncofetal protein, except for embryonic development, it is upregulated in many different malignancies, including Hepatocellular carcinomas (HCCs) [25], lung Squamous Cell Carcinoma (SqCC) [46] and many embryonal tumours [34]. However, it is downregulated in breast cancer while it is expressed normally in adult mammary tissue [47].

The presence of GPC3 in HCC and its diagnostic and therapeutic roles have been extensively explored. HCC is the most common kind of liver cancer, and liver cancer ranks second in terms of cancer-related mortality [48]. Most cases of HCC are seen in people with other liver conditions, such as hepatitis B or C viruses or liver damage from heavy alcohol consumption [49]. Its early detection and surveillance are crucial in order to prevent a poor prognosis [50]. Although GPC3 is downregulated in most adult tissues, it is specifically overexpressed in more than 70% of HCCs, whereas it is undetectable in normal liver tissue or benign liver lesions [25]. Except for the HCC tissue, GPC3 is also present and elevated in the serum of many patients with HCC [25,26,51]. Elevated GPC3 expression in HCC patients correlates with a poorer prognosis [52].

The Wnt/β-catenin pathway has been found to be dysregulated in HCC [53], and it has been established that GPC3 promotes the growth of HCC by stimulating the Wnt/β-catenin pathway. GPC3 binds the Wnts and their receptor, Frizzled, stimulates the pathway, enhances cell proliferation and promotes HCC [27,37,54]. Inhibiting the binding of GPC3 to Wnt with targeted mutants reduced the Wnt pathway's overactivation and therefore reduced the mouse liver cancer growth [54]. On the other hand, the soluble GPC3, which is present in the extracellular space, suppresses the growth of HCC [55,56]. The upregulation of GPC3 promotes the

progression of lung squamous cell carcinoma (SqCC) through the “canonical” Wnt signaling pathway as well [57].

Initial evidence that GPC3 interacts with FAT1, a cell surface protocadherin, has recently been published. FAT1 regulates cell proliferation, migration and polarity and is involved in cancer. These initial data suggest that the GPC3-FAT1 interaction is involved in the HCC cell migration [58].

1.3.5 GPC3 as a diagnostic and therapeutic target

As presented above, the overexpression of GPC3 in the liver tissue and serum of HCC patients makes it a suitable serological biomarker for the early detection of HCC. However, the elevated levels of GPC3 in the serum of patients with other types of cancer [51] suggest that it cannot be used as a single marker, but only in combination with other biomarkers. The use of different biomarkers in parallel will increase the sensitivity and specificity that are required.

Except for a potential diagnostic biomarker, the research on GPC3 and HCC has pointed out GPC3 as a potential therapeutic target. Due to the nature of cancer and the fact that the available treatments are not particularly successful, immunotherapy has been identified as a powerful strategy for cancer treatment. Several therapeutic antibodies against GPC3 have been developed. Firstly, antibodies against GPC3 have been raised for immunohistochemistry analysis of the HCC and the detection of GPC3 in the patient’s serum. One of the first developed antibodies was used to confirm the elevated expression of GPC3 in the HCC tissue and the detection in the serum of the HCC patients with an ELISA (Enzyme-Linked Immunosorbent Assay) method [25]. The soluble GPC3, present in the HCC patient serum, was also detected by an antibody against the N-terminal region of GPC3 [26]. Later, antibodies were raised against GPC3 and characterised for their therapeutic

potential. The monoclonal antibody GC33 which targets the C-terminal region of GPC3 had antitumour activity. The use of this antibody in GPC3-positive hepatoma cells induced antibody-dependent cellular toxicity (ADCC) and caused tumour growth inhibition in liver cancer xenograft models [59,60]. The humanised version of the GC33 antibody (hGC33) was equally effective [61], and it was further tested in clinical trials, but it did not present a clinical benefit [62–64]. The monoclonal antibody YP7, another antibody with the same binding epitope as the hGC33, can detect the GPC3 protein in patient serum and HCC tissue. Also, like GC33, YP7 had antitumour activity and presented growth inhibition in HCC xenografts [65].

Based on the knowledge that GPC3 interferes with developmental pathways, antibodies were raised to target specific functional regions of GPC3. For example, the monoclonal antibody HS20 which binds to the heparan sulphate chains of GPC3 disrupted the Wnt signaling pathway in HCC cells and had antitumour activity *in vivo* [66]. Another approach was the development of a “smaller” antibody which would target a small and functional epitope of GPC3 and potentially block the cell proliferation that the GPC3 promotes, directly. HN3, a human heavy-chain variable domain antibody against GPC3, was developed. In order to bind GPC3, it required the presence of both the N-terminal and C-terminal subunits of the GPC3 core protein. HN3 inhibited the proliferation of GPC3-positive cells and the tumour growth of HCC xenografts in mice [67]. Together with the antibodies, more immunotherapy strategies for HCC that target GPC3 have been explored. Some of them are CAR-T cells (Chimeric Antigen Receptor T-cells), immunotoxins and peptide/DNA vaccines [68]. Moreover, to date, four synthetic nanobodies have been generated against GPC3 and characterised as to their binding to the antigen, protein stability and effect on normal cell proliferation [69].

1.4 The Uncoordinated-5 (Unc5) receptors

The Unc5 receptors are a family of single-pass transmembrane receptors which play an essential role in several cellular processes, including axonal and cell guidance, angiogenesis and apoptosis. The vertebrate Unc5 family has four members, Unc5 A–D [70]. The function of the Unc5 receptors is initiated by the binding of their ligands: fibronectin leucine-rich transmembrane proteins (FLRT) [71,72], Netrins [73,74] and GPC3 [30].

1.4.1 Domain architecture and structures

The extracellular region of the Unc5 receptors consists of two Immunoglobin-like domains (Ig1 and Ig2) and two thrombospondin domains (TSP1 and TSP2) [70]. The intracellular part comprises a ZU5 domain, a DCC-binding motif-containing domain (UPA domain) and a death domain (DD) [75] (Figure 1.3A).

The structure of the human Unc5A isoform 1 ectodomain (PDB: 4V2A) revealed the two Ig domains and only the presence of the second of the two TSP domains that the Unc5 homologue ectodomains usually have [72]. The first two tryptophan residues in the consensus sequence $W_1xxW_2xxW_3$ in the TSP domain are C-mannosylated [30] (Figure 1.3B). The structure of the intracellular region of Unc5B (PDB: 3G5B) revealed a structural supramodule in which the ZU5 domain binds both the UPA domain and the death domain and locks this tight conformation (Figure 1.3C). This conformation represents an inactive signaling state, and its release leads to activation [75].

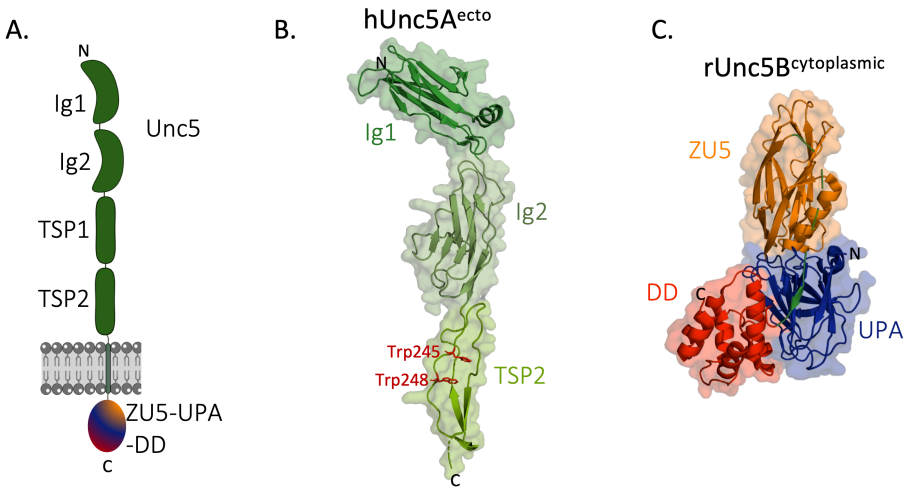


Figure 1.3: The Unc5 receptors architecture and structure.
(A) A schematic representation of the domain architecture of the Unc5 receptors. (B) Structure of the ectodomain of the human Unc5A isoform 1 (PDB: 4V2A). The Trp245 and Trp248 that are C-mannosylated are highlighted in red. (C) Structure of the intracellular domain of the rat Unc5B (PDB: 3G5B).

The complex structure of Unc5D with its ligand FLRT2 is also known (PDB: 4V2C). The Ig1 domain is only needed for the interaction with the LRR (Leucine Rich Repeat) domain of FLRT2 [72]. The Unc5 receptor also participates in a super complex with the adhesion molecules FLRT and Latrophilin (LPHN). A ternary complex of Unc5 with these two proteins was suggested to be formed [76]. However, the structure of this interaction revealed a complex of LPHN3:FLRT2:Unc5D with a stoichiometry of 2:1:1 (PDB: 5FTU), which can be further dimerised to a larger complex with a stoichiometry of 4:2:2 (PDB: 5FTT) through an Unc5-Unc5 interaction [77].

1.4.2 The Unc5 receptors: signaling and functions

The Unc5 receptors mRNA are predominantly present in the nervous system during development but in the adult tissue as well [78]. Through the interaction with their main ligand, Netrin-1, the Unc5 receptors mediate the repulsive responses of Netrin in the nervous system during embryogenesis [73,79]. Netrin-1 acts as a

chemoattractive or chemorepulsive cue depending on the receptors available on cells. The chemoattractive nature of Netrin-1 takes place through the interaction with its receptor Deleted in Colorectal Cancer (DCC) [10,80]. The chemoattractive response switches to repulsion when Netrin binds to the Unc5 receptors. After the Netrin binding, the repulsion is induced by the dimerisation of the Unc5 receptor with the DCC receptor through their cytoplasmic domains. The Netrin binding to Unc5 induces the dimerisation [73].

The FLRTs have also been identified as binding partners of the Unc5 receptors [71]. They are a family of cell surface proteins, broadly expressed in all the tissues in vertebrates, and they regulate embryonic, vascular and neural development. The FLRTs promote cell adhesion by interacting with themselves [81] or other adhesion molecules like members of the LPHN family [82]. They promote cell repulsion through their interaction with the Unc5 receptors [72,83]. The repulsive FLRT-Unc5 and the adhesive FLRT-FLRT interactions regulate radial and lateral neuronal migration, respectively, during cortical development. Metalloproteases are shedding the extracellular domains (ECD) of the FLRT proteins from the cell surface. The soluble FLRT ECDs bind the Unc5 receptors on the neurons' cell surface and regulate cortical development [83]. Moreover, the Unc5D is one of the receptors that participate in the regulation and switch of the multipolar neurons to bipolar during radial cortical migration [84].

Except for the nervous system, the Unc5B mRNA is also present in the developing vascular system, mammary glands and limbs [78]. Unc5B is expressed by endothelial tip cells of the vascular system and plays a role in its morphogenesis [85,86]. Activation of Unc5B with Netrin-1 leads to inhibition of angiogenesis [85]. However, studies have shown that Netrin-1 can have a pro-angiogenic effect [87–

89]. This discrepancy may be due to differences in Netrin-1 concentration [90]. The interaction of Unc5B with its ligand FLRT3 has also been found to regulate retinal vascularisation [72].

GPC3 is the newly discovered binding partner of Unc5, and their interaction plays a role in cell migration [30]. A detailed description and characterisation of this new interaction are presented in Chapter 1.5.

1.4.3 The Unc5 receptors in the dependence receptors family

The Unc5 receptors, as well as DCC, belong to the dependence receptors family. When their ligand is present, the dependence receptors transduce positive signals and block cell death, but in the absence, they emit negative signals and induce apoptosis [91]. In the absence of Netrin-1, Unc5 receptors are thought to promote apoptosis, but the molecular mechanism is not yet fully understood [74]. In the presence of Netrin-1, the Unc5 and DCC receptors multimerise. This step inhibits the proapoptotic effects that these receptors induce [92]. When Netrin-1 is not present, the intracellular domains of the Unc5 receptors adopt an open active conformation [75], get cleaved by caspase-3 [74], more proteins are recruited, and proapoptotic signaling is induced [93].

The Unc5 receptors, as dependence receptors, play a role in cancer. Since they can induce apoptosis in the absence of their ligand, they are potential tumour suppressors. Unc5s are downregulated in several cancers [94]. At the same time, Netrin-1 is upregulated in different tumours and may promote tumourigenesis since, in the presence of Netrin-1, Unc5 should not induce apoptosis [95–97]. Besides inhibiting apoptosis, Netrin-1 also has a role in cancer metastasis [98–100]. Therefore, the interaction between these two proteins has been aimed as a therapeutic target for cancer treatment. Grandin and colleagues developed an

antibody that inhibits the Netrin-1/Unc5 interaction, which induces apoptosis in several cancer cell lines [101].

1.5 The Unc5-GPC3 interaction

An ELISA experiment identified a new interaction between GPC3 and Unc5D, and SPR (Surface Plasmon Resonance) confirmed this interaction [30], in line with the results of earlier research [102]. SPR also showed that GPC3 could bind all the Unc5 family members (Unc5A-D) [30].

1.5.1 The octameric complex

The crystal structure of the core protein of human GPC3 ($hGPC3^{core}$) in complex with the rat Unc5D ($rUnc5D^{Ig1Ig2TSP1}$) revealed an octameric complex, which has four GPC3 chains and four Unc5D chains that are enclosed between the four GPC3s. The $rUnc5D^{Ig1Ig2TSP1}$ molecules are arranged in an alternating antiparallel configuration (Figure 1.4). Native mass spectrometry confirmed the presence of the octamer in another experiment except for the crystal lattice. Masses corresponding to a dimer (1:1 Unc5D:GPC3) and a tetramer (2:2 Unc5D:GPC3) were also present [30].

Three binding sites between GPC3 and Unc5D contribute to the formation of this octamer. Each GPC3 molecule binds three Unc5D chains. Two of them contribute to the binding interface with the region between the Ig1 and Ig2 domains, and the third with the TSP1 domain (Figure 1.4F). Molecular dynamics simulations were used to characterise these interfaces, and hydrophobic, charged, and hydrogen-bonding interactions were identified. Interestingly, SPR data showed that the Ig1 domain and the TSP1 domain alone or an Ig1-Ig2 domain Unc5 construct could not

bind GPC3, and in order for Unc5 to bind GPC3, it needs all three domains Ig1, Ig2 and TSP1 [30].

The octamer structure also revealed that a glycan interaction occurs at the centre of the complex. The members of this interaction are a GPC3 N-linked glycan chain and two α-mannoses bound to two tryptophan residues on the Unc5 molecule [30]. Previous data have shown that in the TSP1 domain of the Unc5 chains, two tryptophan residues are C-mannosylated. These residues are the first two tryptophans in the consensus sequence $W_1xxW_2xxW_3$ [103]. Mass spectrometry data confirmed that these two tryptophan residues were C-mannosylated in all the Unc5 homologues (A-D). These two α-mannoses were interacting with the N-linked glycan of the asparagine residue 241 of the hGPC3 molecule, and the interaction was present in the crystallographic data of the Unc5-GPC3 octameric complex (Figure 1.4G). The contribution of this glycan interaction to the GPC3-Unc5 complex formation was assessed by mutating the glycosylation site and therefore removing the N-linked glycan. The absence of the glycan led to the loss of affinity between the two proteins [30].

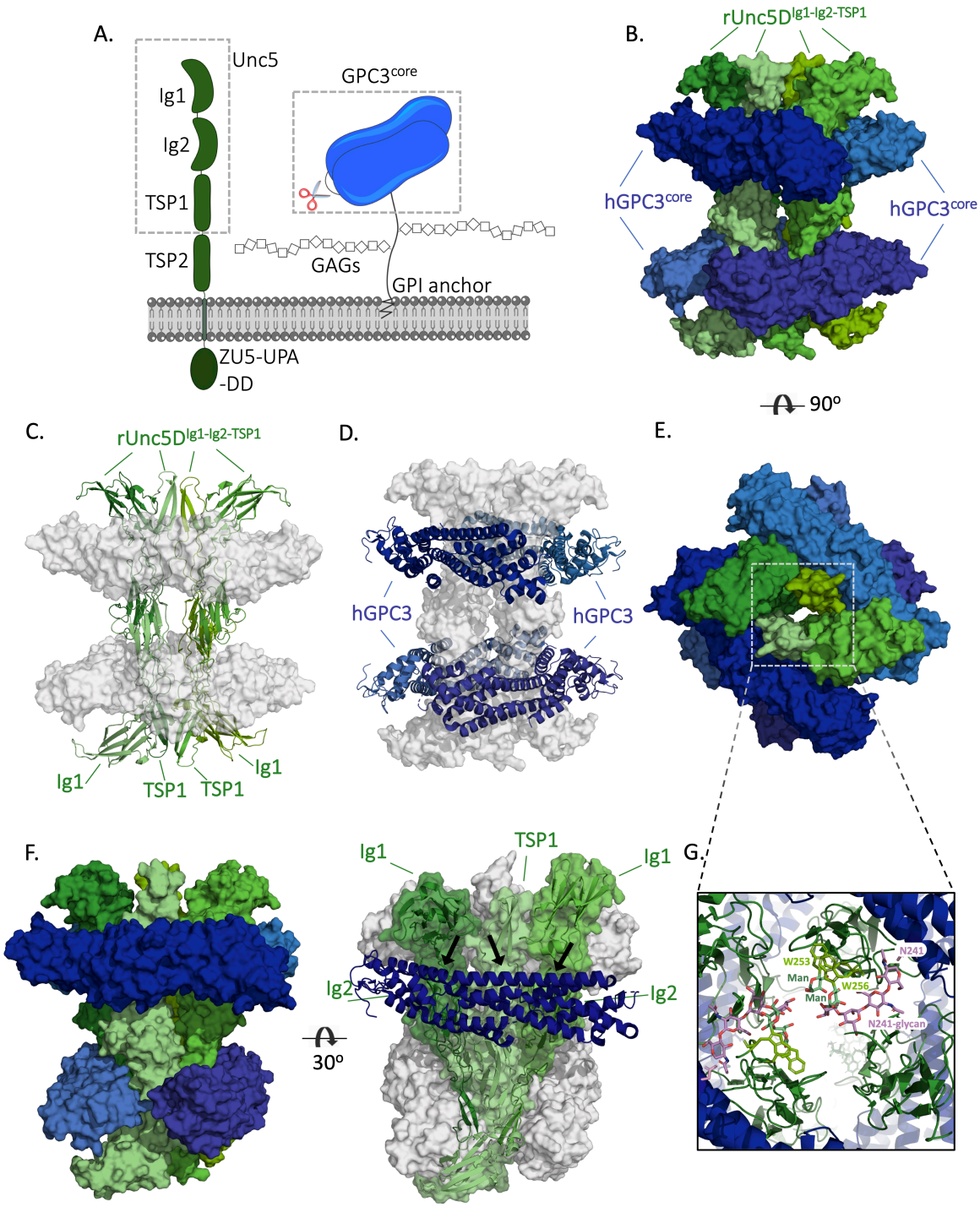


Figure 1.4: The octameric structure of the GPC3-Unc5 complex.
(A) Schematic representation of the GPC3 and Unc5 domain architecture. Dashed line boxes highlight the part of the molecules which participate in the octamer. (B) Surface representation of the GPC3-Unc5 octamer. (C) The Unc5 chains in the octamer are highlighted in green with their distinct antiparallel configuration. (D) The GPC3 molecules in the octamer are highlighted in blue. (E) Top view of the octameric structure in surface representation. (F) The three binding interfaces between one GPC3 molecule and three Unc5 chains are highlighted with three arrows. (G) The glycan-glycan interaction in the centre of the octamer in detail. Figure adapted from [30].

1.5.2 The GPC3-Unc5 interaction elicits repulsion *in vitro*

Netrin-1 and FLRT, well-known Unc5 binding partners, have a repulsive effect on cells that express the Unc5 receptors [73,83]. Stripe assays showed that GPC3 has a repulsive effect in cells that express Unc5 endogenously. These cells were cortical neurons, HeLa cells, and two neuroblastoma cell lines, N2A and SY5Y. The cells presented a preference for migrating toward the non-Unc5-binding GPC3 mutant and away from the wild-type GPC3 (Figure 1.5) [30].

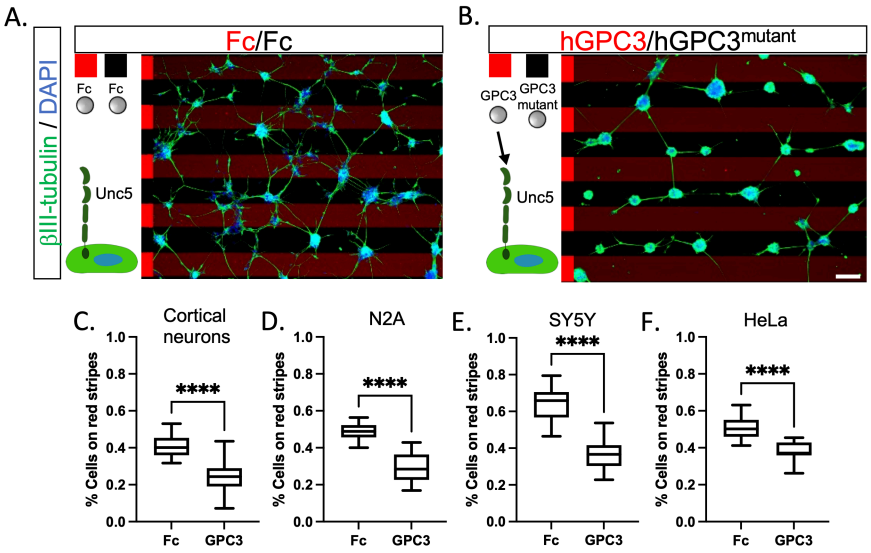


Figure 1.5: Stripe assays show that the GPC3-Unc5 interaction has a contact-repulsion effect.
(A) Control stripe assay in which both stripes (red, black) have immobilised Fc protein. Cortical neurons that endogenously express Unc5D are placed on the stripes. (B) Stripe assay where the red stripe has the wild type GPC3 and the black stripe the non-Unc5-binding GPC3 mutant immobilised. The Unc5D-expressing neurons prefer the black stripes and the non-binding mutant. (C) Quantification and statistical analysis of the cortical neurons stripe assay. Fc (= Fc/Fc control), GPC3 (=hGPC3/hGPC3^{mutant}). (D-F) Quantification and statistical analysis of the stripe assays of the N2A, SY5Y and HeLa cells. Fc (= Fc/Fc control), GPC3 (= hGPC3/hGPC3^{mutant}). ****p < 0.0001, Student T-tests. Scale bar: 100 μm. Figure adapted from [30].

However, in another stripe assay format, the cells moved away from both the wild type and the mutant GPC3 and preferred the control Fc stripe. This suggests that the repulsion that the GPC3 elicits does not depend only on the GPC3-Unc5

interaction, but another GPC3 interaction is potentially present. These stripe assay experiments showed that, at least in this *in vitro* set-up, GPC3 interacting with Unc5 has a contact-repulsion effect [30].

1.5.3 The role of the GPC3-Unc5 interaction in radial neuronal migration

The expression patterns and known functions of the GPC3 and Unc5 proteins indicate that they participate in the brain's development and function. In the developing cortex, the Unc5D-FLRT2 interaction controls the radial migration of the neurons since Unc5D-expressing neurons are being repelled by FLRT2 [72]. *In situ* hybridisation in the developing mouse cortex showed Unc5D expression in the intermediate and subventricular zones, areas of the cortex where migrating neurons are abundant [30]. Single-cell RNA seq data confirmed the enrichment of Unc5D in the migrating neurons [30,104,105].

Mutations in GPC3 lead to the Simpson-Golabi-Behmel overgrowth syndrome (SGBS). Patients with this syndrome display neuronal-related symptoms, like intellectual deficiencies and malformations in cortical development [35,106,107]. These characteristics might indicate a role of GPC3 in the nervous system formation. Expression data on the mRNA level during the mouse embryonic days E13.5 to E17.5 showed the presence of GPC3 in the germinal layers and especially at the ventricular zone, the zone where the apical progenitor cells are positioned during these embryonic days. More specifically, single-cell RNA-seq data revealed that GPC3 is present mainly in the apical progenitor cells, which agrees with *in situ* hybridisation data [30,104,105]. Based on the expression pattern, a model has been proposed in which the apical progenitors / radial glial cells express the GPC3 protein and the migrating neurons express Unc5D and use the radial glial cell fibres as a scaffold during migration (Figure 1.6A). This model was tested *in vivo*.

In vivo cortical cell migration experiments suggest that the GPC3-Unc5 interaction regulates radial neuronal migration [30]. In order to study the effect of this protein interaction, *in-utero* electroporation was used, aiming at the introduction of the desired protein DNA and its over-expression in the developing cortex. The expression of a soluble version of the Unc5D receptor was used to block the natural interaction without the signaling effect. This had as a result, the delayed migration of the neurons (Figure 1.6C, E). The over-expression of a non-GPC3-binding Unc5 mutant partially rescued this effect, suggesting that the GPC3-Unc5 interaction is at least partially responsible for the delayed neuronal migration (Figure 1.6D, E) [30].

Except for the over-expression of soluble Unc5, small hairpin RNA (shRNA) was also used to knock down the endogenous GPC3 on the radial glial cells and limit the interaction between GPC3 and Unc5D. This led to the same result of reduced neuronal migration [30].

1.5.4 The role of the GPC3-Unc5 interaction in neuroblastoma cell migration

Except for neuronal cell migration, both GPC3 and Unc5 are implicated in cancer. The Netrin-1 - Unc5 interaction affects tumour growth and spreading since the overexpression of Netrin-1 in neuroblastomas inhibits the proapoptotic activity of the Unc5 receptors. On the contrary, the overexpression of Unc5 in neuroblastoma cells suppressed the tumour cell growth, migration and invasion [97,108]. At the same time, the expression of GPC3 has been observed in aggressive neuroblastoma samples [109,110]. A neuroblastoma model was used to investigate the role of the GPC3-Unc5 interaction in this cancer migration. Neuroblastoma is a paediatric cancer with embryonal origin and derives from the developing sympathetic nervous system. The sympathoadrenal lineage of the neural crest gives birth to the neuroblastoma cells during development [111]. The expression of

the Unc5 receptors and GPC3 at the mRNA level in neuroblastoma was confirmed [110]. For the neuroblastoma cell line SY5Y, which was used in the experiments, the presence of the Unc5 receptors and GPC3 at the protein level was also confirmed [30].

An *in vivo* xenograft neuroblastoma model was used in which the neuroblastoma cells, SY5Y, were engrafted in a chick embryo's neural crest. This model was used to mimic the migration of the cells from the neural crest towards the developing sympathetic ganglia and adrenal medulla. The neuroblastoma cells will migrate and form tumours there, after following the stereotypical ventral migratory path [112].

Small interfering RNA (siRNA) was used to silence the GPC3 expression in the xenograft cells, and the results showed that these cells migrated and formed tumours outside the proximal and distal sympatho-adrenal (SA) territories which are the target areas for the neuroblastoma cells. They also appeared dispersed and not in a collective migration format (Figure 1.7) [30].

The overexpression of the rUnc5D^{Ig1Ig2TSP1} and the non-GPC3-binding rUnc5D^{Ig1Ig2TSP1} mutant in the xenograft resulted in tumour growth outside the SA territories and dispersed appearance of the cells, which shows that the Unc5 receptors affect the migration of the cells independent from the GPC3 binding (Figure 1.8). These experiments demonstrated that the GPC3-Unc5 interaction is important for the neuroblastoma cell migration *in vivo* and plays a role in cellular cohesion essential for cancer cell collective migration [30].

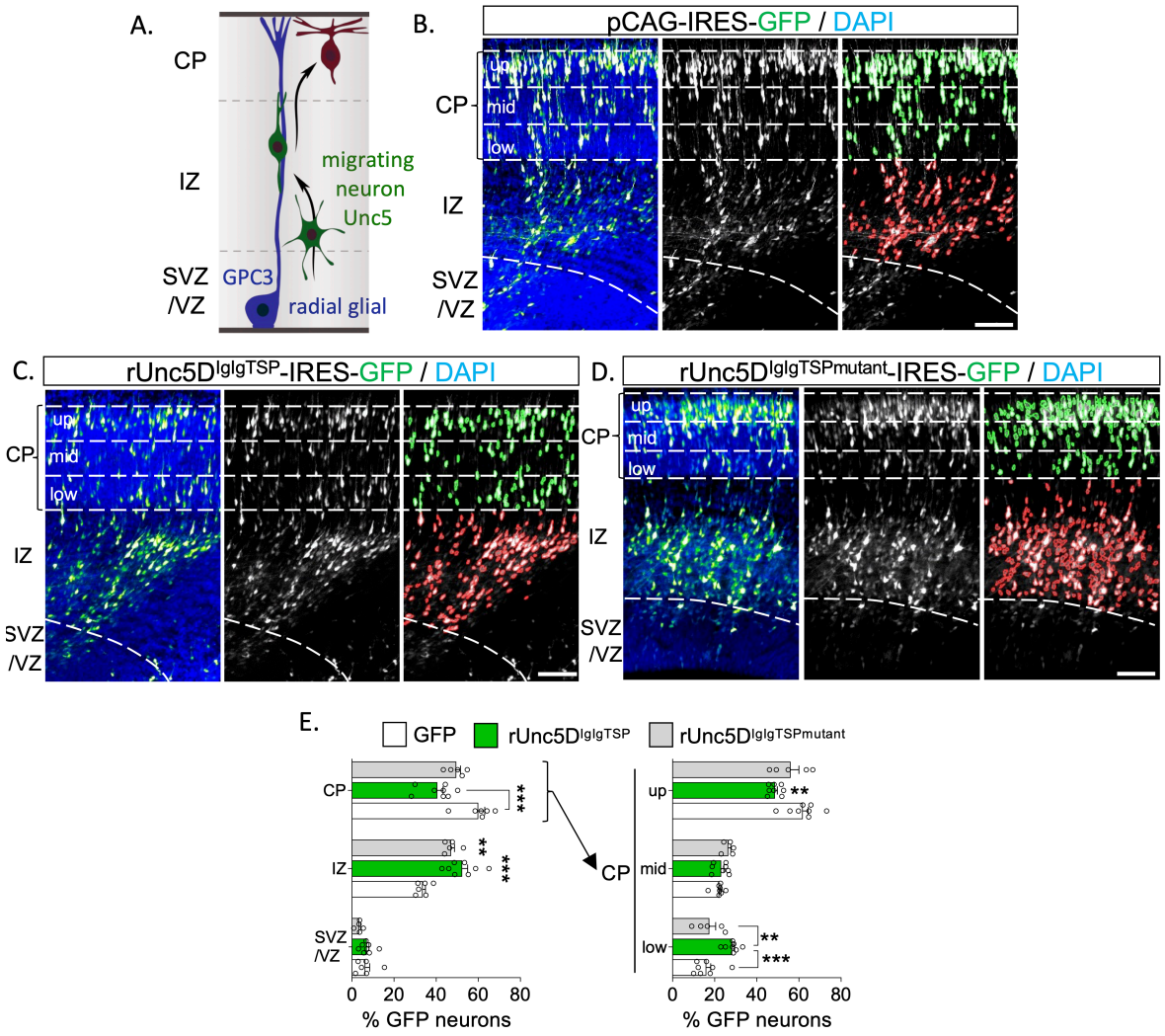


Figure 1.6: The GPC3-Unc5 interaction regulates the radial neuronal migration.

(A) Summary model of the Unc5 and GPC3 protein localisation during radial migration based on the expression data. The apical progenitors / radial glial cells express the GPC3 protein, and the migrating neurons the Unc5D protein. (B-D) Coronal sections of the mouse cortex at E16.5 after *in-utero* electroporation. B: empty vector – control C: rUnc5D^{IgIgTSP} D: rUnc5D^{IgIgTSPmutant}. DAPI staining is used to define the cortical plate. The GFP-positive cells are coloured red in the IZ and green in the CP. (E) Quantification of the experimental data shown in B, C, D. n = 8 GFP, n = 8 rUnc5D^{IgIgTSP} and n = 5 rUnc5D^{IgIgTSPmutant} electroporated brains. **p < 0.01, ***p < 0.001, one-way ANOVA test with Tukey's post hoc analysis. Scale bar: 100 µm. CP: cortical plate; IZ: intermediate zone; SVZ: subventricular zone; VZ: ventricular zone. Figure adapted from [30].

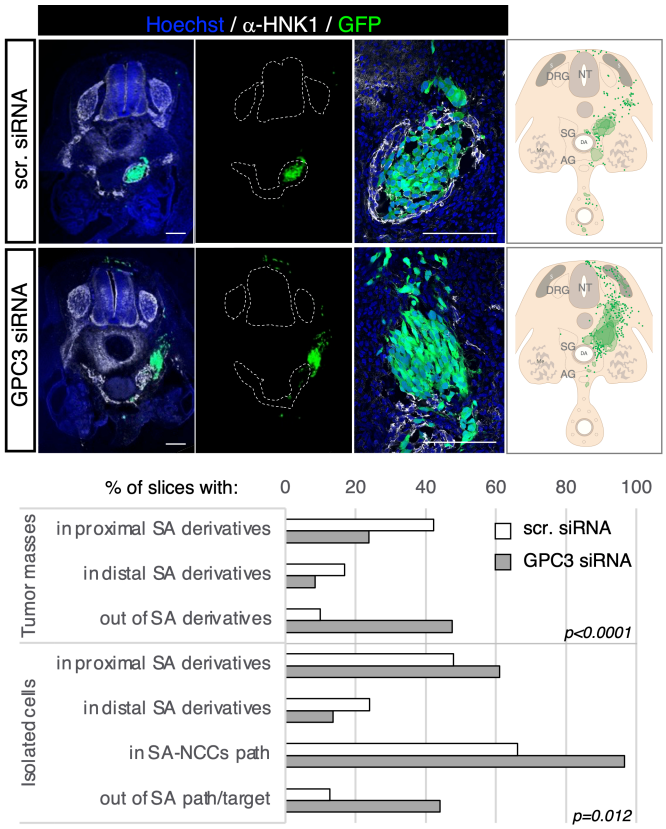


Figure 1.7: GPC3 siRNA affected the neuroblastoma cells migration and tumour formation.
The neuroblastoma SY5Y cells (GFP-green) were transfected with scramble (scr.) siRNA (control) or GPC3 siRNA and engrafted into the migratory trunk neural crest of E2 chicken embryos. Slices were analysed 2 days later. In the control experiment, the cells from the neural crest migrated and formed the structures that are labelled with the anti-HNK1 antibody (white). In the GPC3 siRNA sample, the cells migrated, and the tumours were formed outside their normal destination. The collective migration was also disturbed, and the cells appeared more dispersed. The quantification of the cell migration is present at the bottom part of the figure. (scr. siRNA: 4 embryos, 71 sections. GPC3 siRNA: 4 embryos, 59 sections.) Statistical analysis: χ^2 tests. NT: Neural Tube; S: Somite; DRG: Dorsal Root Ganglia; SG: Sympathetic Ganglia; AG: Adrenal Gland; DA: Dorsal Aorta; Me: Mesonephros; Hoechst: stains the nuclei; α-mito antibody: detects the human neuroblastoma cells (SY5Y graft). Scale bar = 200 µm. Figure adapted from [30].

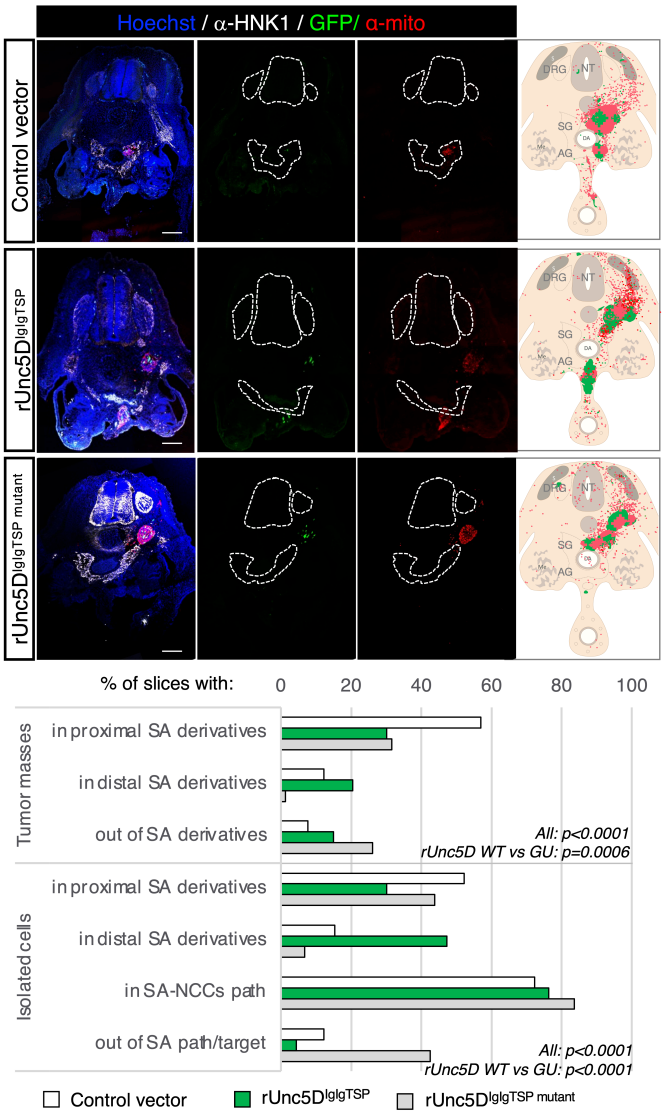


Figure 1.8: The effect of the Unc5^{IgIgTSP} and non-GPC3-binding Unc5^{IgIgTSP} mutant expression in neuroblastoma cell migration.
The neuroblastoma SY5Y cells (GFP-green) were transfected with rUnc5^{IgIgTSP} and non-GPC3-binding rUnc5^{IgIgTSP} mutant and engrafted into the migratory trunk neural crest of E2 chicken embryos. Slices were analysed 2 days later. In the control experiment, the cells from the neural crest migrated and formed the structures labelled with the anti-HNK1 antibody (white). All cells: green and red. Non-transfected cells: red. The expression of the WT and the mutant resulted in misplaced tumours and dispersed appearance. The quantification of the cell migration is present at the bottom part of the figure. (control: 7 embryos, 65 sections. rUnc5^{IgIgTSP}: 7 embryos, 93 sections. rUnc5$^{IgIgTSPmutant:}$ 6 embryos, 73 sections) Statistical analysis: χ^2 tests. NT: Neural Tube; S: Somite; DRG: Dorsal Root Ganglia; SG: Sympathetic Ganglia; AG: Adrenal Gland; DA: Dorsal Aorta; Me: Mesonephros; Hoechst: stains the nuclei; α-mito antibody: detects the human neuroblastoma cells (SY5Y graft). Scale bar = 200 μm. Figure adapted from [30].

1.6 Nanobodies

Monoclonal antibodies have been extensively studied and used in research and medicine as research, therapeutic and diagnostic tools. However, conventional antibodies have drawbacks that can be outweighed by using nanobodies instead.

1.6.1 Nanobodies: small antigen-binding fragments

Nanobodies are small (~15 kDa) antigen-binding fragments which derive from the heavy-chain-only antibodies that naturally occur in the serum of camelid and cartilage fish [113,114]. The conventional antibodies comprise two heavy and two light polypeptide chains, forming the characteristic Y-shape structure [115]. In contrast, the heavy-chain-only antibodies consist of two heavy chains with one variable domain (VHH) on each chain, which is responsible for antigen recognition. This variable domain is named nanobody and has full antigen binding ability (Figure 1.9A, B). From the structural perspective, the nanobodies maintain their core structure with their conserved framework regions formed by nine antiparallel β-strands. These nine β-strands form one β-sheet of four and one β-sheet of five strands, and the two sheets are stabilised with an internal disulphide bond. The antigen recognition mainly relies on the three variable regions of the nanobody structure. These regions are located on loops between the β-strands, and they are called complementarity-determining regions (CDRs) since they recognise the antigen and are complementary to its binding interface. CDR3 is usually the longest variable loop and plays the main role in antigen recognition and binding, while the CDR1 and CDR2 regions also contribute (Figure 1.9C, D) [116].

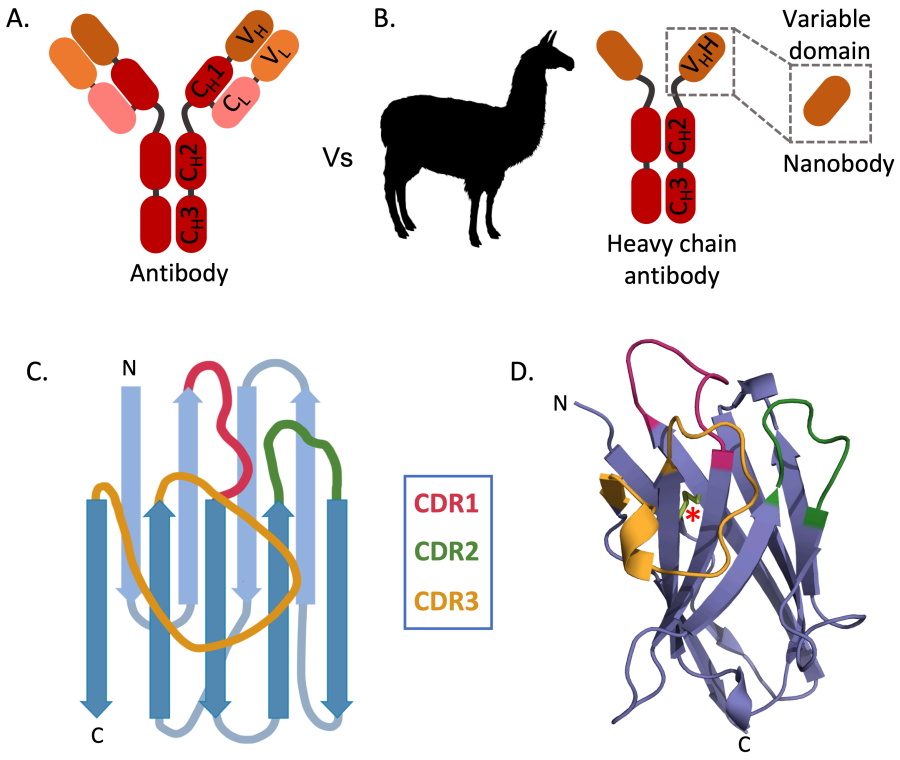


Figure 1.9: Nanobodies
(A) Schematic representation of the Y-shape structure of a conventional antibody. C_H1, 2, 3: Constant domain 1, 2, 3 of the antibody heavy chain. V_H: Variable region of heavy chain. C_L: Constant region of light chain. V_L: Variable region of light chain (B) Schematic representation of the cameloid heavy chain antibody. The nanobody fragment is highlighted. V_HH: Variable domain of the heavy chain of heavy-chain-antibody (C) Schematic representation of the folded nanobody structure, which shows the nine antiparallel β-strands. The CDR loops are highlighted in different colours. (D) 3D model structure of an example nanobody (AlphaFold model of a nanobody generated in this thesis [117,118]). The internal disulphide bond is highlighted with an asterisk. Figure inspired by [116].

1.6.2 Nanobodies' favourable properties and applications

The nanobodies present several characteristics which make them superior to other antigen-binding fragments and antibodies [119,120]. They comprise robust protein molecules that are easy to produce and manipulate, and they demonstrate high specificity and affinity to their target in the low nanomolar, even picomolar range. Nanobodies for research purposes can be obtained either by immunisation of a cameloid animal or from nanobody libraries, and protocols are already established [121,122]. One characteristic that makes them popular is the high-yield

protein production after expression and purification from conventional prokaryotic expression systems. Protocols for their expression and purification are already well-established [121]. Since they do not require any post-translational modifications, they can be expressed on large scale in the periplasm of bacterial cell systems, which is simple and cost-effective. Nanobodies' small size, short sequence, and monomeric structure make it easy to manipulate and improve their biophysical features. For example, targeted mutagenesis on their CDR loops can increase their binding affinity to their antigen [123]. Multimerisation of them can be easily achieved and lead to multispecificity and avidity increase [124,125]. Nanobodies also present high solubility, robustness and thermostability [126].

Moreover, nanobodies can be used as research tools. Because of their small size, the much greater length of the CDR3 loop and the high specificity and affinity to their target, they can bind in small and/or hidden epitopes [127], prevent specific protein-protein interactions [123], stabilise transient protein complexes [128], block enzymatic activity [129] or stabilise proteins in different conformations [130,131]. Nanobodies have also been used effectively as crystallography chaperons [132,133]. Additionally, because of their high stability, nanobodies can be expressed intracellularly, retain their binding ability, and target intracellular proteins [134]. For example, nanobodies fluorescently tagged have been used in microscopy [135]. Cryo-EM has also benefited from the nanobody technology. Nanobodies connected with larger scaffold proteins form chimeric proteins named Megabodies. These structures can be used in cryo-EM to help with the structure determination of small particles, assist in the preferential particle orientation problem and act as a marker for the alignment of small particles whose orientation is not easy to

determine [136]. A similar but larger nanobody-based tool has been developed for the structure determination of even smaller particles and protein molecules [137].

Except for research tools, the favourable properties of the nanobodies make them attractive for therapeutic and diagnostic applications. Due to their size, nanobodies can achieve deep-tissue penetration and diffuse quickly in tissues [138]. This characteristic makes them suitable for application in cancer therapeutics and diagnostics where nanobodies need to reach solid tissues in tumour formations. Their size also allows them to transport across the blood-brain barrier [139]. Additionally, because of their small size, they are quickly cleared from the circulation, which results in a low background signal, beneficial in imaging [138]. On the therapeutic side, it can be a disadvantage since a high concentration for a longer time or a higher dose frequency is needed for the nanobody to reach the target and achieve the potential therapeutic efficacy. The nanobodies also present low immunogenicity compared to the conventional monoclonal antibodies, but in case problems arise, they can be humanised [140,141]. Extensive research has been done on the use of nanobodies in diagnostic and therapeutic applications [119,120]. In cancer diagnostics, molecular imaging can utilise the nanobodies to bind to a tumour antigen in the tumour tissue or in the circulation. Nanobodies bound on their targets can be visualised with different techniques [120]. Nanobodies have been linked with numerous therapeutic applications, as well. Like the conventional monoclonal antibodies, nanobodies can block, enhance, or more generally modulate different interactions, which will lead to the targeted modification of signaling pathways important in different pathologies [142,143]. They can also target tumours and improve the efficacy of cancer therapies already established, like the CAR-T cells (Chimeric Antigen Receptor T-cells) [144] or therapeutic drug

and immunotoxin delivery [145]. Many cancer clinical trials utilising nanobodies are in progress [120].

1.7 VLGR1: an adhesion GPCR

The Very Large G-protein-coupled receptor 1 (VLGR1), also known as ADGRV1, GPR98 or MASS1, is an adhesion G-protein coupled receptor (ADGR). VLGR1 is the largest member of the GPCR family and one of the largest cell-surface proteins known. The molecular weight of the full-length VLGR1b protein isoform is around 700 kDa [146].

1.7.1 Adhesion G protein-coupled receptors (ADGRs)

The G-protein-coupled receptors (GPCRs) are the largest family of receptors in mammals, consisting of five subfamilies: Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin [147]. A seven α-helical transmembrane domain is the common structural characteristic among all the GPCRs. These receptors can detect and get activated by a wide variety of ligands and subsequently regulate many different pathways via G proteins [148].

The adhesion GPCRs subfamily are receptors with distinctive large extracellular domains with which they participate in cell adhesive functions and facilitate cell-cell and cell-matrix interactions. In addition to their adhesion function, they mediate cell signaling with their intracellular region. The main functions that these receptors regulate, include cell migration, neuronal cell function, tissue adhesion and polarity [149]. The members of the ADGRs family are grouped into nine subfamilies based on the architecture of their extracellular domain and their phylogenetic relationship. These subfamilies are: ADGRL (latrophilins; LPHLs), ADGRE (EMRs), ADGRA,

ADGRC (CELSRs), ADGRD, ADGRF, ADGRB (BAIs), ADGRG and ADGRV (GPR98) [150].

Like the GPCR receptors, the adhesion GPCRs have a canonical seven transmembrane (7TM) domain followed by an intracellular domain responsible for the signaling transduction. Their large ectodomains (ECD) comprise several different smaller repetitive subdomains. The ECDs of the ADGRs share two common features, the GPCR proteolysis site (GPS) [151] and the GAIN (GPCR autoproteolysis-inducing) domain [152] close to the cell membrane. The GPS proteolytic site is contained in the GAIN domain and is located almost directly before the first helix of the 7TM. The GAIN domain is a ~320 amino acid domain and a common and well-conserved feature among the adhesion GPCRs, except ADGRA1 (GPR123). For the autoproteolysis, the GPS motif needs to be in the GAIN domain, and it cannot function otherwise [152].

After autoproteolysis, the mature receptors are separated into two fragments, the NTF (N-terminal fragment) and CTF (C-terminal fragment). The two fragments usually remain non-covalently linked; therefore, the NTF remains attached to the cell surface. The NTF fragment contains the extracellular domain of the receptors and the biggest part of the GAIN domain up until the GPS motif. Consequently, the CTF fragment includes the small part of the GAIN domain after the GPS motif (Stachel sequence; a cryptic tethered agonist) [153], the transmembrane and the intracellular domains [148]. The activation of the ADGRs has been linked with the separation of the NTF and CTF fragments and the exposure of the Stachel fragment [154–156]. In other activity modulation models of the ADGRs, the Stachel fragment does not play a role in the activation [157,158] or it is getting exposed by a conformational change of the NTF and not its removal [159]. Except for the signal

transduction of the ADGRs through G-protein coupling, these receptors can interact with other types of proteins [160].

1.7.2 The VLGR1 domain architecture

VLGR1 has the largest extracellular domain among the adhesion GPCR family members [146]. Several protein isoforms have been reported in vertebrates. VLGR1b is the full-length isoform which is predicted to be around 700 kDa. The full-length VLGR1b isoform and the isoform VLGR1a are the only known isoforms that contain the transmembrane domain and are cell receptors. The isoform VLGR1c lacks the transmembrane domain and is predicted to be a soluble protein [161]. Two more isoforms, VLGR1d and VLGR1e, have been identified, and they both include the N-terminal end of the full-length isoform VLGR1b [162]. The last three isoforms that have been reported (MASS1.1, MASS1.2, MASS1.3) are internal pieces of the long isoform without a transmembrane domain [163].

The N-terminal region of all the VLGR1 protein isoforms comprises several Ca^{2+} binding CalX-β motif repeats [164]. These CalX-β motifs are calcium-binding domains that are present in the Na^{+}/Ca^{2+} exchangers [165]. This suggests that VLGR1 binds calcium, and experimental data support it [164,166]. Sequence-based predictions have revealed that the N-terminal extracellular region of the longest VLGR1 isoform, VLGR1b, contains 35 Ca^{2+} binding CalX-β motif repeats [161]. These repeats are being interrupted by a single LaminiG (LamG)/pentraxin (PTX) domain which is predicted to have a jellyroll fold [161,167] and seven Epitempin (EPTP) repeats [168]/epilepsy-associated repeats (EAR) [169,170]. These repeats consist of copies of a short, ~50 amino acid repeat, which is thought to fold into a seven-bladed β-propeller structure (Figure 1.10).

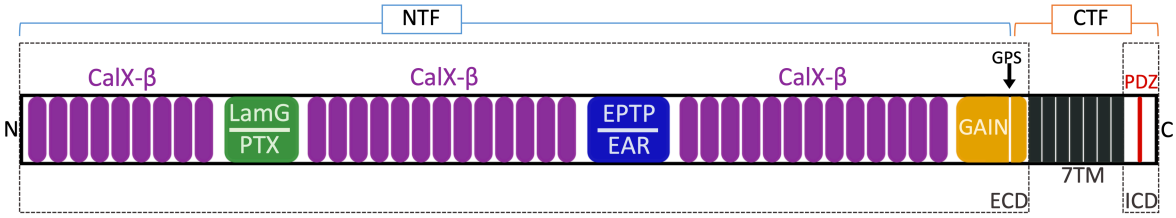


Figure 1.10: Domain architecture of the VLGR1b.
Schematic representation of the predicted domains on the largest isoform of VLGR1. CalX-β: Ca^{2+} binding motif, LamG/PTX: predicted jellyroll fold, EPTP/EAR repeats: predicted seven-bladed β-propeller structure, NTF: N-terminal fragment, CTF: C-terminal fragment, ECD: Extracellular domain, ICD: Intracellular domain, 7TM: Seven transmembrane domain, GAIN: GPCR autoproteolysis-inducing domain, GPS: autoproteolytic site.

As found in other ADGRs, VLGR1 contains a GAIN domain with the autoproteolytic site (GPS) followed by a 7-transmembrane GPCR domain and an intracellular domain with a PDZ binding motif at the C-terminal end [171] (Figure 1.10). The presence and the autoproteolysis of the GPS motif were confirmed, and the cleaved CTF can activate G-protein signaling [172]. To date, there are no known three-dimensional structures for VLGR1 or its individual domains.

1.7.3 The VLGR1 expression pattern, functions, and related disease

The VLGR1 isoforms have been detected at low levels in several tissues. However, its high expression in the developing central nervous system proposes a role of VLGR1 in brain development [161,162]. *In situ* hybridisation data have shown that it is highly expressed in the mouse central nervous system when the development begins. During mid-gestation, VLGR1 is expressed in the neuroepithelium of the ventricular zone, where neuronal precursor cells start their proliferation and differentiation. This also suggests a fundamental role for VLGR1 in the development of the central nervous system [161]. Furthermore, VLGR1 is enriched in human astrocytes [173] and mature mouse oligodendrocytes [166].

VLGR1 is also located at the base of the stereocilia of the hair cells in the cochlea. The expression of VLGR1 at the stereocilia lasts roughly from embryonic day E17 to post-embryonic day P10, which goes in line with the development time frame of the inner ear. There, together with other proteins, VLGR1 forms the ankle links of the stereocilia [174–177]. The wild-type stereocilia have a characteristic shape. In the absence of the VLGR1 protein in mutated mice, the ankle links fail to form, and the stereocilia lose their shape [174]. VLGR1 also participates in the periciliary membrane of photoreceptor cells [178].

Mutations in VLGR1 lead to the development of the Usher syndrome type 2C, which is characterised by hearing loss and progressive retinitis pigmentosa [171]. The PDZ binding motif at the C-terminal end of the VLGR1 receptor is involved in the receptor's interaction with Usher proteins like Whirlin, Harmonin and PDZD7, and its intracellular signaling may have a role in the Usher syndrome development [172,179–181].

Recent studies attempted to identify new binding partners for VLGR1 with tandem affinity purifications followed by mass spectrometry analysis using as bait different fragments of VLGR1. The results revealed new binding partners that can associate VLGR1 with several different functions like neural development, cell adhesion and photoreceptor function [182].

Mouse models with impaired VLGR1 expression are used in VLGR1 research [162,163,183]. In these models, the mice present audiogenic seizures and partially Usher syndrome symptoms. Mutations in VLGR1 have also been associated with seizures in humans [184].

1.8 Research aims

The discovery of the GPC3-Unc5 interaction and the structure of its octameric complex needed further investigation. The expression pattern and the functions of the individual proteins of the complex, GPC3 and Unc5, led the research towards neuronal and tumour cell migration. *In vivo* and *in vitro* assays were used to characterise the interaction's function in cell migration.

The aim of this DPhil was to produce and characterise tools for the investigation of the structure and function of the GPC3-Unc5 interaction. These tools were nanobodies against GPC3 and non-binding GPC3 and Unc5 mutants. Consequently, the following objectives were set.

- Biophysical characterisation of two nanobodies raised against GPC3.
- Identification of the way that these nanobodies affect the GPC3-Unc5 interaction.
- Preparation of the nanobodies in order to be used in *in vivo* and *in vitro* assays.
- Identification of more nanobodies with potential function as research tools, out of a pool of nanobodies against GPC3.
- Structural characterisation of the GPC3-Nanobody interaction.
- Design and characterisation of non-GPC3-binding Unc5 and non-Unc5-binding GPC3 mutants.

At the same time, part of this DPhil, as a side project, was to get an initial understanding of the three-dimensional structure and function of the adhesion GPCR, VLGR1. Literature suggests that VLGR1 has essential roles in the function of the nervous system. However, relatively little is known about the three-dimensional structure and function of this enigmatic GPCR.

The aim of this project was to address this lack of knowledge and reveal details of the VLGR1 structure. The objectives of this project were:

- The design, cloning and expression of protein constructs, parts of the extracellular domain of VLGR1b isoform.
- Structural characterisation of these protein constructs.

1.9 Results summary

This thesis presents the generation and characterisation of tools used to understand the function of the GPC3-Unc5 interaction.

Out of a pool of nanobodies against human GPC3, two nanobodies were thoroughly characterised biochemically, biophysically and structurally. It was revealed that one nanobody enhances and the other blocks the GPC3-Unc5 interaction. The rest of the available nanobodies were screened, and preliminary data identified more nanobodies that can be proven helpful in future research. The generation and characterisation of non-binding mutants are also presented. Cell binding assays and SPR confirmed the function of a non-GPC3-binding Unc5 and a non-Unc5-binding GPC3 mutant.

Collaborators successfully used both the non-binding mutants and the nanobodies in *in vivo* and *in vitro* assays to help understand how the Unc5-GPC3 interaction affects cell migration. These experiments, summarised in this thesis, revealed the importance of the GPC3-Unc5 interaction in cell migration, particularly in neuronal cell migration during cortical development and in neuroblastoma collective cell migration.

Finally, the construct design, expression and purification that was attempted towards the structural characterisation of the VLGR1 are presented.

2. Materials and methods

2.1 Molecular biology and cloning

2.1.1 Constructs

Different protein constructs were used for the experiments in chapters 3, 4 and 5, which include the characterisation of nanobodies against GPC3 and the design, cloning and characterisation of Unc5 and GPC3 mutants. Table 2.1 presents all the protein constructs and relevant information about them. Most of the constructs were already cloned into the Age1-Kpn1 or EcoR1-Kpn1 cloning sites of the high-level mammalian expression vector, pHLsec [185].

Table 2.1: Protein constructs used in this thesis's chapters 3, 4, 5.

Protein	Construct name	Residues	
Human Glypican 3 (cDNA clone BC035972)	hGPC3$^{full\text{-}length}$	1 – 580	[30]
	hGPC3^{core}	32 – 483	[30]
	hGPC3 K1	Q112N A114T	*
	hGPC3 K2	L157N	*
	hGPC3 K3	N241Q	*
Mouse Glypican 3 (UniProt ID: Q8CFZ4)	mGPC3^{ecto}	31 – 559	[30]
	mGPC3^{core}	31 – 482	[30]
	mGPC3^{488}	31 – 488	[30]
Mouse Unc5A (UniProt ID: Q8K1S4)	mUnc5A^{ecto}	1 – 359	[30]
Mouse Unc5B (UniProt ID: Q8K1S3)	mUnc5B$^{full\text{-}length}$	26 – 934	[30]
	mUnc5B^{ecto}	26 – 362	[30]
	mUnc5B U6	Y150N R152T	*
	mUnc5B M7	A277N L279T	*
	mUnc5B K4	I211N R213T	*
Human Unc5B (UniProt ID: Q8IZJ1)	hUnc5B^{ecto}	1 – 375	[72]
	hUnc5B^{Ecto}	1 – 355	[30]
Rat Unc5D (UniProt ID: F1LW30)	rUnc5D^{ecto}	1 – 382	[77]
	rUnc5D^{Ig1Ig2TSP1}	1 – 307	[77]

* Constructs cloned during this thesis's projects

The nanobodies were originally cloned into the phagemid vector pADL-23c. The nanobody constructs that were generated in this thesis are Nb49 in pHLsec (residues: 23-162), Nb58 in pHLsec (residues 23-163) and Nb49^{short} in pADL-23c (residues: 1-146).

All constructs produced in the VLGR1 project (Chapter 6) were subcloned from codon-optimised synthesised cDNA of the mouse VLGR1b (GenBank: AF435926.1) into the pHLsec vector. The full-length VLGR1b cDNA was synthesised into five fragments which were fused as needed for the different construct cloning. The different constructs cloned for this project are described in Chapter 6.

2.1.2 Expression vectors and vector preparation

For the mammalian cell expression, all the constructs were cloned into the Age1-Kpn1 or EcoR1-Kpn1 cloning sites of the pHLsec vector [185]. This vector has an N-terminal secretion signal peptide which was used instead of the protein's native signal peptide in most cases.

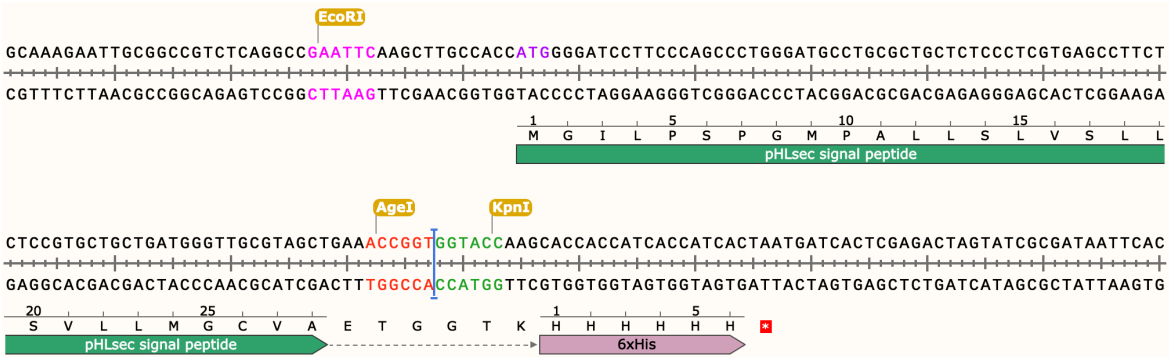


Figure 2.1: The mammalian expression vector, pHLsec.
Schematic representation of the N-terminal secretion signal peptide (green), the restriction enzyme sites (EcoRI; magenta, AgeI; red, KpnI; green) and the C-terminal 6 × histidine tag (pink) of the pHLsec vector. Figure generated with SnapGene® software. As explained by the figure, constructs cloned into the EcoRI-KpnI cloning sites were using their original signal peptide, while the constructs cloned into the AgeI-KpnI were using the signal peptide from the pHLsec vector.

Typically, a C-terminal 6 × histidine tag is present, but different N-terminal and C-terminal tags were used for different experiments. For protein biotinylation, proteins were cloned into the pHLsec vector with a C-terminal Avi tag (sequence: GLNDIFEAQKIEWHE). For the cell binding assays, the full-length mUnc5B was cloned in the pHLsec vector with an N-terminal HA tag (sequence: YPYDVPDYA) and a C-terminal fluorescent mVenus tag. The full-length hGPC3 was cloned in the pHLsec vector with an N-terminal Flag tag (sequence: DYKDDDDK), and a stop codon was cloned just before the C-terminal 6 × histidine tag. The stop codon was introduced because the full-length GPC3 is attached to the cell surface with a GPI anchor, and it cannot have a C-terminal tag attached. The protein constructs in the VLGR1 project were cloned in the pHLsec vector using the pHLsec secretion signal peptide and a C-terminal 6 × histidine tag. N-terminal and C-terminal 10 × histidine tags were also used.

For their expression in the *E. coli* WK6 strain, the nanobodies were already cloned into the phagemid vector pADL-23c. In this vector, a pelB leader sequence (sequence: MKYLLPTAAAGLLLLAAQPAMA) was included before the nanobody sequence, and it was responsible for the export of the nanobody protein in the periplasm of the bacterial cells. The expression of the nanobodies in the periplasm helps with the proper folding and the disulphide bond formation because of its oxidised environment nature and the presence of folding chaperones. The nanobody sequence was followed by a 6 × histidine tag and a myc tag (sequence: EQKLISEEDL) at the C-terminal end.

The vectors were linearised by digestion using suitable restriction enzymes for each experiment. The restriction enzymes EcoRI (R0101S), AgeI (R0552S), and KpnI (R0142S) from NEB were used. For the restriction enzyme digestion, 6 µg of

circular plasmid vector were mixed with 2 µl from each restriction enzyme, the suitable buffer for the optimal enzymatic function and 2 µl of the calf intestinal alkaline phosphatase (CIP) (NEB) in a total volume of 50 µl. The reaction mixture was then incubated at 37°C for 3 hours.

2.1.3 Construct generation

All the constructs and the mutants were generated using the CloneAmp™ HiFi PCR Premix kit (Takara Bio, Cat#639298). A typical PCR reaction consists of 100 ng plasmid DNA, 7.5 pmol of both the forward and the reverse primers and 12.5 µl of the PCR premix in a total volume of 25 µl. For a typical PCR protocol, 30 thermal cycles of 10 sec at 98°C for denaturation, 15 sec at 55°C for the primer annealing and 15, 20 or 30 sec of extension (based on the product length) at 72°C were performed.

PCR primers were designed to bind to the template with a 25 bp binding region and have an overlap region of 15 bp for the insertion into the plasmid. The mutations on mUnc5B and hGPC3 were added by inserting the change in bases between the annealing and overlap regions. All the primers are presented in Table A1 and Table A2 in the appendix.

2.1.4 Recombinant plasmid DNA generation

The PCR products and the linearised vectors were separated and cleaned by agarose gel electrophoresis at 150 V for 30 to 40 min in TAE buffer (10 mM Tris, 20 mM acetic acid, 1 mM EDTA). For the agarose gel, agarose was dissolved in TAE buffer at a concentration of 0.8% - 1.2%, depending on the size of the DNA products. The mixture was melted, SYBR™ Safe DNA Gel Stain (Invitrogen™, Cat#S33102) was added (1:10000 dilution), and the gel was let to set. Gel extraction of the correct

DNA band followed as per protocol using the Macherey-Nagel™ NucleoSpin™ Gel Clean-up kit (Thermo Scientific™, Cat#12303368) after visualisation and detection of the band under a Dark Reader® transilluminator.

The different constructs and mutants were cloned in the correct vector using the In-Fusion™ HD cloning kit (Takara Bio, Cat#639650). The purified DNA after gel extraction was used for the In-fusion reaction. For a typical reaction, in a total volume of 5 µl, 50 ng of insert DNA, 50 ng of the linearised vector and 1 µl of the 5 X In-Fusion HD Enzyme Premix were added. The reaction was incubated at 50°C for 15 min. 1.2 µl of the reaction was transformed into Stellar™ competent cells (Takara Bio, Cat#636763).

For a typical transformation, cells were mixed with 1 µl of the desired plasmid (100 ng/µl), incubated on ice for 30 min, heat shocked for 45 s at 42°C, and then returned to ice for 2 min before plating on agar plates (Thermo Scientific Chemicals, Cat#10589462) containing 100 µg/ml ampicillin (AppliChem, Cat#A0839.0025). Plates were incubated overnight at 37°C. Single colonies were selected and used to inoculate 10 ml of LB Broth (Millipore, Cat#1102850500) supplemented with 100 µg/ml ampicillin, and the cultures were incubated overnight at 37°C in a shaking incubator. The plasmid DNA was extracted from the cells using the Macherey-Nagel™ NucleoSpin™ Plasmid kit (Cat#11932392). Each clone was analysed by Sanger sequencing (Source Biosciences). The correct clones were confirmed and further transformed into DH5α competent cells (Invitrogen™, Cat#18265017) for better yield.

2.2 Protein expression and purification

2.2.1 Mammalian cell cultures

All the protein constructs (except for the non-biotinylated nanobodies) were expressed in adherent HEK-293 cells, which are highly transfectable mammalian cells derived from human embryonic kidney 293 cells. In this thesis, HEK-293T cells were used to express most of the protein constructs except for the proteins for the crystallisation trials. These cells express proteins with complex N-linked glycans present, which is not optimal for crystallisation. For this reason, all the proteins for the crystallisation trials were expressed in HEK-293S GnTI$^-$ cells. These cells lack the N-acetyl-glucosaminyltransferase I (GnTI) activity and express secreted proteins with homogeneous and non-complex N-linked glycans [186].

Cells were maintained in T75 flasks at 37°C and 5% CO_2 and were subcultured every 2-3 days when they reached 90-100% confluency. For the culture media, Dulbecco's Modified Eagle's Medium (DMEM; Gibco^{TM}, Cat#21969035) was used, supplemented with 10% Fetal Bovine Serum (FBS; Gibco^{TM}, Cat#10437028), 1% MEM non-essential amino acids (Gibco^{TM}, Cat#11140035) and 1% L-Glutamine (Gibco^{TM}, Cat#25030024). For the subculturing process, after the aspiration of the old media, the cells were gently washed with 10 ml of PBS (Lonza, Cat#BE17-516F) and were detached from the flask surface using 2 ml of 0.05% Trypsin-EDTA (Gibco^{TM}, Cat#25300054) for 4 min. Then, cells were resuspended in fresh media to the desired density and seeded into a new flask.

2.2.2 Small-scale transient transfection

For small-scale protein expression, 6-well plates were prepared from the cell stock T75 flasks. A confluent T75 flask was split as usual. 2 ml of the resulting cell

mixture were resuspended in 10 ml fresh 10% FBS-DMEM media, and 2 ml of the resuspended cells were seeded into each well of a 6-well plate. The 6-well-plated cells were incubated at 37°C, 5% CO_2 for 24 hours before transfection. For the transfection, a transfection cocktail of 100 µl Opti-MEM (Gibco™, Cat#31985047), 3 µg DNA and PEI (Polyethylenimine) (Sigma-Aldrich, Cat#408727) at a ratio of 3 to 1 (PEI:DNA) was prepared. After a 10 min incubation at room temperature, the transfection cocktail was added carefully to the side of each well to avoid cell disruption. For expression tests of secreted proteins, media were harvested 2-3 days post-transfection, followed by a 10 min centrifugation at 15,000 rpm to remove any remaining cells. Protein expression was verified by western blotting. For the cell binding assays, the cells were ready to be used 24h after transfection.

2.2.3 Large-scale cell preparation and transient transfection

For large-scale expression, HEK-293T or HEK-293S GnTI$^-$ cells were prepared, and 2125cm² expanded surface roller bottles (Grenier Bio-One, Cat#681070) were used for this purpose. Cells were seeded in T175 flasks from the T75 cell stock flask, and all the cells from one confluent T175 were used to prepare one roller bottle. 250ml of 10% FBS-DMEM media were added to each roller bottle prior to the cell addition. The roller bottles were prepared and incubated at 37°C for 4 days before transfection. Cells were transiently transfected in the roller bottles using PEI at a 2 to 1 PEI:DNA ratio. The protein construct plasmid DNA for the large-scale transfection was prepared using the PureLink™ HiPure Expi Plasmid Gigaprep Kit (Invitrogen™, Cat# K210009XP). Before transfection, the plasmid DNA was mixed (vortex) with 100 µl chloroform (Sigma-Aldrich, Cat#C2432) per 1ml DNA (1 mg/ml) and centrifuged at 6,000 rpm for 10 min. For one roller bottle, a transfection cocktail of 50 ml serum-free DMEM, 1 mg PEI and 0.5 mg DNA was prepared and incubated

for 10 min. Before adding the transfection cocktail in the roller bottle, the media was changed to 2% FBS-DMEM. After transfection, HEK-293T and HEK-293S GnTI$^-$ were allowed to grow for 4 and 10 days, respectively, in order to produce the secreted recombinant protein.

2.2.4 Protein Biotinylation

For protein biotinylation in HEK-293T cells, protein constructs were cloned in the pHLsec vector with a C-terminal Avi tag and co-transfected with a vector encoding BirA [187]. BirA, the *E. coli* biotin ligase recognises the Avi tag and covalently attaches biotin to the lysin of the 15 amino acid Avi tag peptide (sequence: GLNDIFEAQKIEWHE).

For small-scale transfection and biotinylation in a 6-well plate, 3 µg of the protein construct plasmid were co-transfected with 1µg of the BirA vector and PEI at 3 to 1 of PEI to DNA ratio. Cell culture medium was supplemented with sterile filtered 100 nM of biotin (Sigma-Aldrich, Cat#B4501). Three days post-transfection, media were harvested and dialysed against 1X PBS at 4°C to remove the excess biotin. Protein expression and biotinylation were verified by western blotting.

For large-scale transfection and biotinylation in roller bottles, the transfection cocktail for one roller bottle was prepared with 50 ml serum-free DMEM, 1 mg PEI and 0.5 mg DNA of which 90% is the plasmid of the desired protein construct and 10% the BirA vector. The fresh 2% FBS-DMEM was supplemented with 50 nM biotin before the addition of the transfection cocktail.

2.2.5 Bacterial cell cultures and nanobody protein expression

Nanobodies against the hGPC3^{core} protein were obtained by intra-muscular immunisation of a llama with the purified protein, as described in [30]. The *E. coli*

WK6 strain (ATCC, 47078™) was used for the nanobody periplasmic expression. Cells were mixed with 1 µl of the desired plasmid (100 ng/µl), incubated on ice for 30 min, heat shocked for 45 s at 42°C, and then returned to ice for 2 min before plating on agar plates containing 100 µg/ml ampicillin. Plates were incubated overnight at 37°C. Single colonies were selected and used to inoculate 10 ml of LB Broth supplemented with 100 µg/ml ampicillin, and the cultures were incubated overnight at 37°C in a shaking incubator. 5 ml of the overnight cultures were used to inoculate 500 ml of Terrific Broth (Fisher BioReagents™, Cat#12891660) large scale cultures. Cells were grown at 37°C until an OD_{600}=0.8, and then the temperature was lowered to 21°C. After one hour at 21°C, expression was induced with 150 µM IPTG (Isopropyl β-D-1-thiogalactopyranoside) (Sigma-Aldrich, Cat#I6758) and incubation at 21°C for 16 h. Cells were harvested by centrifugation (6,000 × g, 15 min).

2.2.6 Periplasmic protein extraction

Bacterial cell pellets were resuspended in ice-cold 20% sucrose, 30 mM Tris pH 7.5, and 2 mM EDTA buffer and incubated at 4°C for 20 min on a tube rotator. The cell suspension was clarified by centrifugation (10,000 rpm, 20 min, 4°C). After collecting the supernatant, the cell pellets were resuspended in ice-cold 30 mM Tris pH 7.5, 5 mM $MgSO_4$ buffer and incubated at 4°C for 20 min on a tube rotator. The cell suspension was clarified by centrifugation (10,000 rpm, 20 min, 4°C). The supernatants after each step were mixed (periplasmic extract), filtered through 0.22 µm sterile filters (Millipore, Cat#S2GPT05RE) and used for the nanobody purification.

2.2.7 Protein purification

For the protein purification from the HEK-293 cells, the cell culture medium from the roller bottles was harvested, clarified by centrifugation at 6,000 x g for 20 min, and filtrated through 0.22 µm sterile filters. The filtrated medium was exchanged into diafiltration buffer (1X PBS, 20 mM Tris pH 7.5, 150 mM NaCl) and concentrated to ~ 600 ml. The concentrated medium was applied to a 5 ml HisTrap^{TM} HP column (Cytiva, Cat#17-5248-02) at a 5 ml/min flow rate. The column was washed with 200 ml wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 40 mM Imidazole), and the protein was eluted in elution buffer (20 mM Tris pH 7.5, 300 mM NaCl, 500 mM Imidazole) into 1 ml fractions. A Bradford assay was carried out to determine the fractions with the highest protein concentration. 90 µl of Bradford reagent (Sigma-Aldrich, Cat#B6916) was mixed with 10 µl sample. The 3ml HisTrap elution peak was loaded onto a HiLoad 16/600 Superdex 200 pg column (Cytiva) pre-equilibrated in 10 mM or 20 mM Tris pH 7.5, 150 mM or 200 mM NaCl. A lower concentration of Tris and NaCl was used for some proteins purified for crystallisation trials. 1 ml elution fractions were collected in a sterile 96-deep-well plate and analysed using SDS-PAGE. Peak fractions were pooled and concentrated. Aliquots were stored at -80°C or used directly for crystallisation trials. Biotinylated proteins were purified as previously but in ice-cold conditions.

For the bacterial cell nanobody purification, after periplasmic extraction of the proteins, the filtered supernatant was supplemented with 150 mM NaCl and 2 mM Imidazole and was applied to a 5 ml HisTrap^{TM} HP column at a flow rate of 2.5 ml/min pre-equilibrated in 30 mM Tris pH 7.5, 150 mM NaCl, 2 mM Imidazole. The column was washed with 200 ml wash buffer (30 mM Tris pH 7.5, 150 mM NaCl, 5 mM Imidazole), and the protein was eluted in elution buffer (30 mM Tris pH 7.5, 150

mM NaCl, 500 mM Imidazole) into 1 ml fractions. The 3 ml Histrap elution peak was loaded onto a HiLoad 26/600 Superdex 75 pg (Cytiva) pre-equilibrated in 10 mM or 20 mM Tris pH 7.5, 150 mM or 200 mM NaCl. 1.5 ml elution fractions were collected in a sterile 96-deep-well plate and analysed using SDS-PAGE. Peak fractions were pooled and concentrated. Aliquots were stored at -80°C or used directly for crystallisation trials. The SEC profiles were prepared for the figures with GraphPad Prism (version 9.5.0).

The protein concentration was measured using a NanoDrop™ spectrophotometer device. The theoretical molecular weight and extinction coefficient, ε, were calculated by ProtParam on the ExPaSy Server [188].

2.2.8 Nanobody tetramerisation with streptavidin

The biotinylated Nb49 was tetramerised using the streptavidin linker Alexa Fluor™ 594 streptavidin (Invitrogen™, Cat#S11227). Alexa Fluor™ 594 streptavidin is a biotin-binding protein covalently attached to an Alexa fluorescent label. The biotinylated Nb49 was mixed with the Alexa Fluor™ 594 streptavidin in a 6 to 1 molar ratio and incubated at room temperature for 30 min. The sample was then loaded onto a Superdex 200 10/300 GL column (Cytiva) in dark conditions, pre-equilibrated in 20 mM Tris pH 7.5, 200 mM NaCl. 300 µl elution fractions were collected in a sterile 96-deep-well plate and analysed using SDS-PAGE. The fractions of the complex peak were pooled, concentrated, aliquoted and stored at -80°C.

2.2.9 SDS-PAGE gel electrophoresis and Western blot

For SDS-PAGE, NuPAGE™ 4-12% Bis-Tris gels (Invitrogen™, Cat#NP0336BOX) were used. For proteins with molecular weight higher than 50

kDa, the MOPS SDS running buffer (Invitrogen™, NP0001) was used, and for proteins with molecular weight lower than 50 kDa, the MES SDS running buffer (Invitrogen™, NP0002) was used. For the separation of proteins with higher molecular weight up to 500kDa, NuPAGE™ 3-8%, Tris-Acetate gels (Invitrogen™, Cat#EA03785BOX) were used with the Tris-Acetate SDS running buffer (Invitrogen™, Cat#LA0041). The samples were prepared by adding 5 µl SDS gel loading buffer supplemented with 5% β-mercaptoethanol (Sigma-Aldrich, Cat#M6250) to 20 µl protein sample, and they were boiled for 5 minutes. The BenchMark™ Protein Ladder (Invitrogen™, Cat#10747012) and the HiMark™ Unstained Protein Standard (Invitrogen™, Cat#LC5688) were used as ladders in the SDS-PAGE. Gels were run at 180 V for 45-60 min and visualised using InstantBlue™ Coomassie stain (Sigma-Aldrich, Cat#ISB1L).

For Western blotting, from an unstained SDS-PAGE gel, the protein bands were transferred for 70 min at 30 V on a nitrocellulose membrane (Cytiva, Cat#10600013) using two blotting papers, five sponges and the 1X NuPAGE® Transfer Buffer (Invitrogen™, Cat#NP00061) in 20% ethanol in the NuPAGE® XCellII™ Blot transfer module. To confirm the transfer, the BenchMark™ Pre-stained Protein Ladder (Invitrogen™, Cat#10748010) was loaded in the gel. The membrane was blocked with 3% BSA (Sigma-Aldrich, Cat#A7906) in PBST (1X PBS, 0.1% Tween-20) for 30 min and incubated for 1 hour in the primary antibody solution, which was made in 3% BSA PBST buffer. The mouse anti-His antibody in a 1:1000 dilution (Qiagen, Inc. Cat#34660) was used as the primary antibody for the detection of His-tagged proteins. The primary antibody solution was removed, and the membrane was washed with PBST and then incubated for 1 hour in the secondary antibody solution (anti-mouse IgG - HRP, Sigma-Aldrich, Cat#A0168 dilution: 1:10000 in 3%

BSA PBST). The secondary antibody solution was removed, and the membrane was washed with PBST, developed with the ECL™ Western Blotting detection reagents (Cytiva, Cat#RPN2106) and visualised using the Amersham™ Hyperfilm™ ECL™ films (Cytiva, Cat#28-9068-35).

For the confirmation of the biotinylation of the proteins, the Pierce™ high sensitivity Streptavidin-HRP (Thermo Scientific™, Cat#21130) was used in western blotting.

2.2.10 Buffer chemicals and reagents

Table 2.2: Buffer chemicals and reagents

Buffer Reagents	Source	Cat #
Phosphate buffered saline (PBS)	Sigma-Aldrich	P4417
Sucrose	Sigma-Aldrich	S0389
Tris base	Sigma-Aldrich	T1503
EDTA (Ethylenediaminetetraacetic acid)	Sigma-Aldrich	E5134
$MgSO_4$	Sigma-Aldrich	746452
NaCl	VWR Chemicals	27810
Imidazole	Sigma-Aldrich	I3386
Tween-20 / polysorbate-20	Thermo Scientific Chemicals	10246910
Hepes	Sigma-Aldrich	H4034
$CaCl_2$	Sigma-Aldrich	C7902

2.3 Biophysical and biochemical methods

2.3.1 Surface plasmon resonance (SPR)

Equilibrium binding experiments were performed at 25°C using a Biacore™ T200 instrument (Cytiva). For all the experiments, CM5 chips (Cytiva, Cat#29104988) were used, and they were activated and coated with streptavidin via amine coupling (Amine coupling kit: Cytiva, Cat#BR100050) prior to the experiment using 1X PBS pH 7.5, 0.005% (v/v) polysorbate-20 as running buffer. To prepare the proteins immobilised on the chip, GPC3, Unc5 and nanobody constructs were expressed in

small-scale 6-well plates, biotinylated enzymatically at a C-terminal Avi tag and dialysed against 1X PBS at 4°C for 36 h. Afterwards, the biotinylated proteins were coupled to the streptavidin-coated CM5 chip at densities between 400 and 500 RU. All the SPR experiment steps were carried out at a 5 µl/min flow rate. The chip was regenerated using 2 M $MgCl_2$ (Sigma-Aldrich, Cat#M8266) for 25 sec.

To test the binding of the nanobodies to human and mouse GPC3, biotinylated Nb49 and Nb58 were immobilised on the CM5 chip. For the experiment, a 20 mM Tris pH 7.5, 200 mM NaCl, 0.005% (v/v) polysorbate-20 buffer was used as the running and dilution buffer. A 2-fold dilution series of $hGPC3^{core}$ and $mGPC3^{488}$ was injected over the chip for 500 sec, followed by 100 sec dissociation time.

For the competition SPR, biotinylated Nb49 and Nb58 were immobilised on the CM5 chip. 20 mM Tris pH 7.5, 200 mM NaCl, 0.005% (v/v) polysorbate-20 was used as the running and dilution buffer. A 2-fold dilution series of $hGPC3^{core}$ + $rUnc5D^{ecto}$ and $mGPC3^{488}$ + $rUnc5D^{ecto}$ was injected over the chip for 500 sec, followed by 100 sec dissociation time. In another competition SPR set-up, Unc5 ectodomains were immobilised on the CM5 chip at densities of 473-483 RU. 1X PBS pH 7.5, 0.005% (v/v) polysorbate-20 buffer was used as the running and dilution buffer. A 2-fold dilution series of $hGPC3^{core}$ or $mGPC3^{ecto}$ in the presence of Nb49, Nb58 or Fc was injected over the chip for 250 sec, followed by 25 sec dissociation time. The concentration of the nanobodies and Fc were kept constant in the 2-fold dilution series of GPC3.

In SPR test injections, a sample of the relevant protein was injected over the immobilised proteins on the chip for 150 sec, and the response units were measured. Regeneration of the chip using 2 M $MgCl_2$ for 25 sec was done after every injection.

Data were analysed using the BIAevaluation software (Biacore, GE Healthcare) and GraphPad Prism (version 9.5.0). The K_D and B_{max} (maximum analyte binding) values were obtained by non-linear curve fitting of a 1:1 Langmuir interaction model.

2.3.2 Small scale pull-downs

Pull-down experiments to investigate the effects of Nb49 and Nb58 on the Unc5-GPC3 interaction were performed. Biotinylated hGPC3^{core} was coupled to high-capacity streptavidin agarose resin (Thermo Scientific^{TM}, Cat#20357) after incubation of the beads with the biotinylated protein for 1 hour. hUnc5B^{Ecto} or rUnc5D^{ecto} were mixed with nanobody in 1:1, 1:5, and 1:10 molar ratios in 20 mM Tris pH 7.5, 200 mM NaCl, 1% BSA and incubated with hGPC3^{core}-coated streptavidin beads for 3 hours. Beads were washed with 20 mM Tris pH 7.5, 200 mM NaCl and the proteins were eluted with SDS gel loading buffer supplemented with 5% β-mercaptoethanol. Samples were analysed using SDS-PAGE.

His-tag pull-down experiments were used for small-scale purification of VLGR1 constructs. A His•Bind® uncharged Resin (Millipore, Cat#69670) was used, charged with $NiSO_4$ and equilibrated with the purification buffer (20 mM Tris pH 7.5, 300 mM NaCl, ± 5 mM $CaCl_2$). After the media harvest from a small-scale expression, the media with the protein were incubated with the charged resin for 3 hours in 1.7 ml tubes. The resin was settled with centrifugation at 750 rpm for 5 min, and the flow through (FT) was collected. The resin was washed with 1 ml wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 40 mM Imidazole, ± 5 mM $CaCl_2$), and the beads were settled again (750 rpm, 5 min) to collect the wash. The protein was eluted in two steps of 500 µl elution buffer (20 mM Tris pH 7.5, 300 mM NaCl, 500 mM Imidazole, ± 5 mM $CaCl_2$). Samples were analysed using western blotting.

2.3.3 Circular Dichroism Spectrophotometry (CD)

The CD experiment was performed in the Biophysics facility of the Biochemistry department with the help of Dr David Staunton. The protein samples (hGPC3^{core} WT and hGPC3^{core} K3 mutant) were expressed in the same cells (HEK-293T) and purified in 20 mM Tris pH 7.5, 200 mM NaCl. The protein samples were diluted to lower the NaCl concentration at 20 mM, as indicated by the CD's recommended protocol. The samples were diluted just before the experiment to avoid any precipitation from the low NaCl concentration. The diluted concentration of the proteins was 0.062 mg/ml, and the required sample volume was 200 µl. For the baseline data, the protein buffer was used. The CD experiment was performed by Dr David Staunton at a J-815 CD Spectropolarimeter by JASCO at 20°C. A 1 mm quartz cuvette was used, and the measurement range was 195 – 250 nm. The plots were generated with GraphPad Prism (version 9.5.0).

2.3.4 SEC-MALS

Size-exclusion chromatography – multiangle light scattering (SEC-MALS) was performed in the Biophysics facility of the Biochemistry department by Dr David Staunton. Samples of 110µl volume were loaded onto a Superdex® 200 HR 10/300 GL (Cytiva) previously equilibrated with 20 mM Tris pH 7.5, 200 mM NaCl and attached to a Shimadzu HPLC system comprising an LC-20AD pump, a SIL-20A Autosampler and an SPD20A UV/Vis detector. This system was linked to a Wyatt Dawn® HELEOS-II 8-angle light scattering detector and a Wyatt Optilab® rEX refractive index monitor. The data were processed with ASTRA 6.1 software (Wyatt Technology), and the plots were generated with GraphPad Prism (version 9.5.0).

2.3.5 Mass spectrometry data processing after brain lysate pull-downs

The del Toro lab performed the pull-down from mouse cortex lysate as described in [30]. E15.5 mouse cortices were homogenised, the cells were lysed, and the supernatant was used for the pull-downs. Biotinylated Nb49 was mixed with the cortex lysates and streptavidin agarose resin. For the control pull-down, the same resin beads were used. The next day the beads were harvested by centrifugation and mass spectrometry was used for the sample analysis.

For the mass spectrometry data processing, R-studio (2021.09.2 Build 382) and the DEP-LFQ package for R, BiocManager 1.30.19 (CRAN) were used. Volcano plots were generated with GraphPad Prism (version 9.5.0).

2.4 X-ray crystallography

2.4.1 Protein crystallisation and crystal optimisation

For the crystallisation trials, all the proteins were expressed in HEK-293S GnTI$^-$ cells except for the nanobodies, which were expressed in the *E. coli* WK6 strain. Based on the SDS-PAGE, the purest SEC fractions from both *h*GPC3^{core} and nanobodies were pooled, mixed in a 1 to 1 molar ratio and concentrated. Different concentrations (200 mM – 350 mM) of non-detergent sulfobetaine 256 (NDSB-256) (Hampton Research, Cat#HR2-705) were added during protein concentration in order to avoid protein aggregation and precipitation.

Swissci 96-well 2-sitting drop vapour diffusion crystallisation plates were prepared with either commercial X-ray crystallography screens to find initial crystallisation hits or the desired optimisation condition to improve the yield or the quality of already existing crystals. The commercially available screens used were the following. From Molecular Dimensions, the screens were Morpheus®,

Morpheus® II JCSG-plus™, ProPlex™, PACT Premier™, the PGA Screen™, MIDASplus™, MemGold™ and MemGold2™. Also, from Hampton Research, the screens Index™ and PEG/Ion Screen™ and from NeXtal, the JCSG Core suite I-IV.

For the optimisation plates, the initial crystallisation plates were screened for any crystal hits, and an optimisation strategy was decided. The promising condition's pH and concentration were slightly modified (e.g. serial dilution of the original condition) in order to find a new condition in which better crystals would grow. The required condition was either purchased from the manufacturer or prepared from its individual components and dispensed into the reservoir well of a 96-well 2-sitting drop vapour diffusion crystallisation plate. The Silver Bullets™ additive screen (Hampton Research) was also used as an optimisation step. Moreover, initial crystal hits or badly formed crystals were used for microseeding. The crystals were harvested, smashed into small fragments with MicroSeed Beads™ (Molecular Dimensions, Cat#MD2-14) and mixed with the protein sample before the new crystallisation plate preparation.

The protein sample was dispensed into the 96-well plates using a Mosquito® nanopipetter (TTP Labtech). The crystallisation plates were incubated/stored at 18°C or 4°C. Crystals were harvested, cryo-protected in 25% glycerol and cooled for screening and data collection by Dr Edward Lowe.

2.4.2 Data collection and processing

All the crystals were screened, and all the data were collected at the Diamond Light Source synchrotron science facility. Crystal screening and data collection were performed by Dr Edward Lowe. For the two hGPC3^{core}-Nb49 complex crystals (Crystal #1 and #2), native datasets were collected at the Diamond Light Source i24

beamline at a wavelength of 0.9999 Å. Data were collected with an exposure of 0.01s per 0.10° for 180° on a Pilatus3 6M detector.

After data collection, the data were processed using Dials [189] via the Xia2 system [190]. The resolution cut-off was determined using the correlation coefficient between random half datasets ($CC_{1/2}$) parameter, which is recommended to be above 30% for the highest resolution. Xia2-Dials was used through the CCP4 software suite (8.0.009) [191] and the CCP4i2 interface 1.1.0 (revision 6539) [192].

2.4.3 AlphaFold and model preparation for Molecular Replacement (MR)

AlphaFold 2.1.1. in its monomer [117,118] and multimer [193] mode was used for the generation of the search models for MR, and the AlphaFold 2.1.1 version was run remotely on the Central Oxford Structural Molecular Imaging Centre (COSMIC) server. A fasta format sequence file with the necessary protein sequences was prepared before every run.

AlphaFold 2.1.1 was loaded with the following command:

```
module load ALPHAFOLD/2.1.1
```

The job was run with the following command:

```
alphafold --fasta_paths=GPC3_Nb49.fasta --max_template_date=2022-05-14
--model_preset=multimer
```

The `--fasta_paths` loads the fasta file of the protein sequence. For the prediction, AlphaFold used PDB templates dated earlier than the `--max_template_date` date. The `--model_preset` can be set either to monomer or multimer to indicate which AlphaFold mode will be used.

The predicted molecule PDB file contains the per-residue confidence metric (pLDDT) for all the atoms representing the confidence of the prediction of each individual residue. The predicted TM-score (pTM) represents the global accuracy of

the model. Except for the pLDDT metric, AlphaFold multimer provides an additional model confidence score which is a combination of the pTM (already used in the AlphaFold monomer) and the Interface ipTM.

For the AlphaFold prediction models to be used in MR, further processing is needed. The CCP4i2 task “Process Predicted Models” is used for the AlphaFold model preparation for MR. With this task, the low confidence residues (pLDDT < 70%) are removed from the model, and the pLDDT scores are converted to B-factors.

2.4.4 Molecular Replacement (MR) and Refinement

The AlphaFold predicted model of the hGPC3^{core}-Nb49 complex was used as the search model in MR by PHASER [194]. Two copies of the hGPC3^{core} and two Nb49 were present in the asymmetric unit (ASU) based on the Matthews coefficient calculation. However, initial MR showed that probably only one of the two hGPC3^{core} molecules has a Nb49 attached. For this reason, the AlphaFold hGPC3^{core}-Nb49 complex model was edited to delete the nanobody structure. The hGPC3^{core} molecule without the Nb and the hGPC3^{core}-Nb49 complex were ultimately used for MR. After MR, 150 cycles of jelly body refinement (sigma: 0.01, max distance: 4.2) followed, using Refmac5 [195]. Several 20-cycle jelly body refinement steps followed, alternating with manual model inspection and model building in Coot 0.9.8.7 [196]. The already published crystal structure of hGPC3^{core} (PDB 7ZAW, 2.58 Å) was also used as a reference model in refinement. Density modification with Parrot was performed in order to improve the electron density map [197]. Molprobity [198] was used for the evaluation of the structure quality. PHASER, Refmac5 and Parrot were used through the CCP4 software suite (8.0.009) and the CCP4i2 interface 1.1.0 (revision 6539).

2.4.5 VMXm sample preparation and crystal screening

Small crystals of the hGPC3^{core}-Nb49 complex were screened in the micro/nanofocus macromolecular crystallography beamline VMXm (Versatile Macromolecular Crystallography Microfocus) in the Diamond Light Source. The size of the beam for this experiment was 3 x 3 μm. For the sample preparation, QUANTIFOIL™ R 2/2 200 copper mesh grids were used. The grids were first treated with glow discharge treatment. The crystal drop was diluted with the reservoir buffer and transferred on the grid. With the help of a Leica EM GP2 automatic plunge freezer, the grid was blotted with filter paper on the opposite side from where the crystal sample was applied to remove the excess liquid. After blotting, the grid was quickly plunge cooled into liquid ethane. The cooled grid with the crystals on it was transferred on the custom, for this beamline, sample holder and then the samples were inserted in the beamline with a custom mechanism. The crystal screening was performed at a wavelength of 0.5813 Å with an exposure of 0.1-0.2 s per 0.10º for 10º on an Eiger2 X 9M (CdTe) detector.

2.5 Cell Biology

Cell surface expression and cell binding assays

For the cell surface expression and cell binding assays, full-length mUnc5B WT and mutant constructs with a C-terminal intracellular mVenus tag were expressed on the cell surface in small-scale expression in 6-well plates. Before the cell seeding in the plate, two coverslips were placed in each well. Cells were transfected with the mUnc5B plasmid DNA and PEI, as described in Chapter 2.2.2. 24 hours after the transfection, the coverslips coated with cells were gently removed from the wells and moved on an ice-cold metal surface. Cells were incubated in ice-cold Hank's

Balanced Salt Solution (HBSS) (Gibco™, Cat#14170088) supplemented with 1% BSA and 20 mM Hepes pH 7.5 for 30 minutes.

For the cell surface expression test, the primary antibody solution was prepared by mixing the mouse anti-HA antibody (Sigma-Aldrich, Cat#H3663) with the HBSS/BSA/Hepes solution. For the cell binding assay, the primary antibody mixture consists of the binding proteins (mGPC3^{ecto} or mFLRT2^{ecto}) and the mouse anti-His antibody (Invitrogen™, Cat#37-2900) in HBSS/BSA/Hepes. The binding protein was previously expressed and purified as described in Chapter 2.2 and the anti-His antibody targets these proteins in order to identify their potential binding. The primary antibody and the binding protein are mixed in a 2 to 1 µgr ratio. Cells were incubated in primary antibody solution for 1 hour, washed three times with ice-cold PBS and then fixed with ice-cold 4% paraformaldehyde (PFA) (Sigma-Aldrich, Cat#158127) for 20 min. PFA was quenched with ice-cold PBS containing 20 mM ammonium chloride and subsequently washed with ice-cold PBS twice. Cyanine3-conjugated (Cy3) anti-mouse antibody (Invitrogen™, Cat#A10521) was added to the cells and incubated for 60 minutes at room temperature. The cells were washed three times with PBS and then incubated with 350 mM DAPI (4′,6-diamidino-2-phenylindole) solution in PBS for 2 minutes. Finally, cells were washed with PBS and milliQ water and the coverslips were mounted on glass slides with Immu-mount™ (Invitrogen™). The binding of mGPC3^{ecto} and mFLRT2^{ecto} was measured by the colocalisation percentage of the Cy3 and the mVenus fluorescence of the Unc5 proteins. ImageJ (Fiji) (version 1.53f51) [199] was used for the cell images inspection, colour overlay and quantification of colocalisation of the Cy3-linked antibody with the mVenus tag. The statistical analysis and the bar graph of the cell binding assays were generated with GraphPad Prism (version 9.5.0).

2.6 Cryo-EM

2.6.1 Crosslinking

The hGPC3^{core} and rUnc5D^{Ig1Ig2TSP1} protein constructs ± Nb49 crosslinking was attempted using BS3 (bis(sulfosuccinimidyl)suberate), an amine-to-amine crosslinking reagent (Thermo Scientific™, Cat#21580). The protein constructs were purified in 20 mM Hepes pH 7.5, 200 mM NaCl since a purification buffer with Tris cannot be used with BS3 because of its primary amine. The BS3 solution and its dilutions were freshly prepared before each experiment. Different BS3 concentrations and incubation times were screened, and the reaction was quenched with a final concentration of 100 mM Tris in the reaction. Crosslinking of the protein samples was analysed using SDS-PAGE.

2.6.2 Sample and grid preparation and screening

For the protein sample preparation, GPC3, Unc5 and Nb49 protein constructs were mixed in a 1:1:3 molar ratio. Before mixing the three components, the individual protein samples were centrifuged for 10 min at 21,300 × g to remove any precipitation. After centrifugation, the sample was retrieved from the tube carefully to avoid disturbing any pellet on the bottom of the tube. Different constructs and isoforms were used for both GPC3 and Unc5 in order to find the best-behaving sample. Different types of grids were also used to find the most suitable grid type for the sample. Holey carbon copper grids (Quantifoil R 1.2/1.3, 300 mesh) and graphene oxide support film copper grids (Quantifoil® R 1.2/1.3 carbon film, 300 mesh) were used, and they were freshly plasma cleaned for 2 minutes and 30 seconds, respectively, before the sample application. 3 µl of the protein sample were applied on each grid, and the sample was vitrified by plunge-freezing in liquid

ethane using a Vitrobot (FEI). The Vitrobot was set at 21°C, 100% humidity, 5 blot force, no waiting time, and a variable blot time between 3.5, 4 and 4.5 seconds in order to achieve the best ice thickness. For the optimisation trials, the detergents FOM (Fluorinated octyl maltoside) and CHAPS ((3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) were used either directly on the sample drop on the grid just before plunge freezing (1% FOM) or mixed with the protein sample (0.4% CHAPS, 0.04% FOM). The prepared grids were screened on an Arctica electron microscope (ThermoFisher – FEI) operating at 200 kV at a magnification of 150,000 x.

3. Characterisation of nanobodies targeting GPC3

3.1 Introduction and Summary

Previous work in the Seiradake lab showed that GPC3 interacts with Unc5D, and the crystal structure revealed an octameric complex. Non-Unc5-binding GPC3 mutants and non-GPC3-binding Unc5 mutants were used to help characterise this interaction in *in vitro* and *in vivo* assays (Chapters 1.5 & 5) [30]. Together with the non-binding mutants, nanobodies against GPC3 were characterised and used as tools to further understand the GPC3-Unc5 interaction.

A pool of anti-hGPC3 nanobodies, raised against the human GPC3 core protein ($hGPC3^{core}$) in a llama, was already available in the lab. The initial screening revealed a nanobody, Nb58, that disrupts the interaction between GPC3 and Unc5. At the same time, another nanobody against GPC3, Nb49, was characterised in order to be used as a control. It was revealed that these two nanobodies, Nb49 and Nb58, enhance and disrupt the GPC3-Unc5 interaction, respectively. Pull-down assays and SPR in a competition set-up were used to characterise how the nanobodies affect the GPC3-Unc5 interaction. As more nanobodies against $hGPC3^{core}$ were available from the initial pool, screening experiments were performed in which more promising for future use nanobodies were detected.

In this chapter, experiments in which these nanobodies were utilised are also presented. The nanobodies were used to confirm the presence of GPC3 in the developing mouse cortex but also to manipulate the GPC3-Unc5 interaction and understand its function *in vitro* and *in vivo*.

Disclaimer: The presented project is collaborative and therefore some aspects of the work were performed by others. Of the work presented in this chapter, I prepared all the proteins needed for the SPR and the pull-down assays. I also performed the pull-downs and the SPR experiments after being trained by Maria Carrasquero-Ordaz. The SEC-MALS experiment was performed in the Biophysics facility of the Biochemistry Department by Dr David Staunton. Experiments in mice and chicken embryos, in which the nanobodies were used, were performed by the del Toro lab and the Castellani Lab. In the brain lysate pull-down experiment, although, I processed the data after mass spectrometry myself in order to get the information and the graphs presented here.

3.2 Production and preparation of the nanobodies

The two nanobodies of interest (see section 3.1), Nb49 and Nb58, were cloned into the phagemid vector pADL-23c, expressed in the WK6 *E. coli* strain cells and purified with nickel affinity, followed by size exclusion chromatography (Figure 3.1).

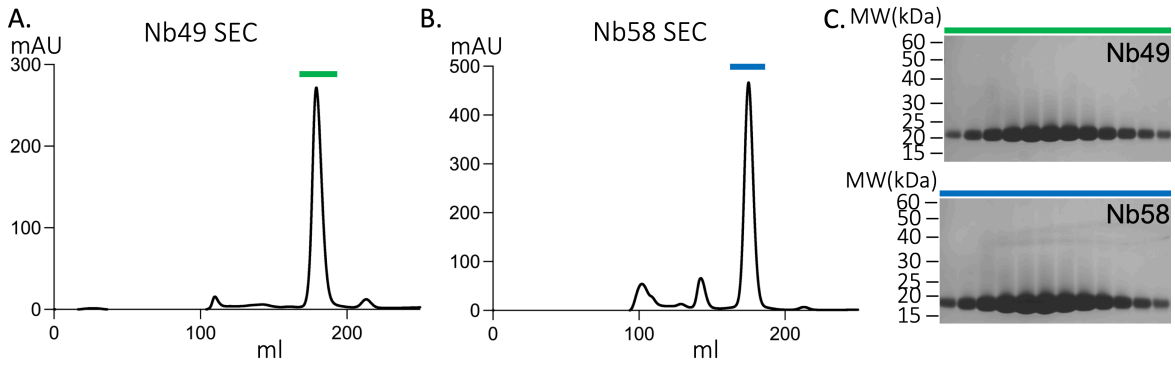


Figure 3.1: Purification of Nb49 and Nb58 from WK6 bacterial cells.
(A-B) SEC profiles of the Nb49 (17.55 kDa) and Nb58 (17.03 kDa) purification. (C) SDS-PAGE gels of the peak fractions of the Nb49 and Nb58 SEC profiles in A and B.

Besides the bacterial expression system, nanobodies were cloned into the mammalian expression vector, pHLsec, with a C-terminal Avi tag followed by a 6 × histidine tag and expressed in HEK-293T cells to get biotinylated. Biotinylation is essential for the following SPR experiments, the pull-down experiments from mouse brain cortex lysate (Chapter 3.6) and the tetramerisation with streptavidin. A small-scale expression and biotinylation test confirmed the expression of the nanobodies cloned in the pHLsec vector in the mammalian cells and their biotinylation in this expression system (Figure 3.2A). Large-scale expression, biotinylation and nanobody protein purification were achieved, and both Nb49 and Nb58 were produced in high protein yield (Figure 3.2B-D).

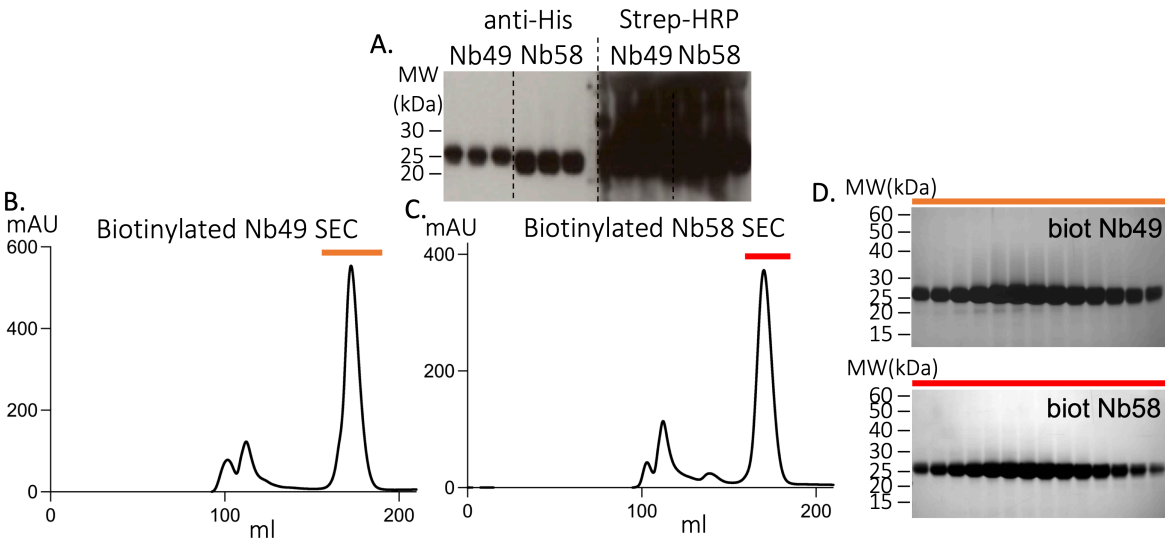


Figure 3.2: Expression, biotinylation and purification of the nanobodies in HEK-293T.
(A) Western blot of the HEK-293T cells media from a small-scale expression test of Nb49 (19.25 kDa) and Nb58 (18.73 kDa) cloned into pHLsec. The biotinylation was also tested. The expression and the biotinylation were confirmed with an anti-His antibody and streptavidin-HRP, which binds to biotin. The streptavidin-HRP has a very strong signal in a western blot which explains the very thick bands in the presented blot. (B-C) SEC profiles of the large-scale purification of the biotinylated Nb49 and Nb58 expressed in HEK-293T. (D) SDS-PAGE gels of the peak fractions of the Nb49 and Nb58 SEC profiles in B and C.

Streptavidin was used to tetramerise the nanobodies in order to improve their affinity and efficacy. The biotinylated Nb49 was mixed with Streptavidin-Alexa 594 in a 6 to 1 molar ratio, followed by size exclusion chromatography to remove the excess nanobody (Figure 3.3).

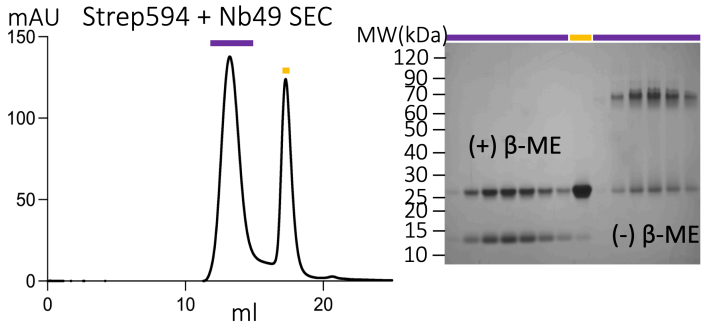


Figure 3.3: Tetramerisation of Nb49 with Streptavidin-Alexa 594.
(A) SEC profile of the Streptavidin-Alexa 594 (Strep594) plus the biotinylated Nb49. The excess nanobody peak is highlighted in yellow. (B) SDS-PAGE gel of the Strep594+Nb49 and the excess Nb49 peaks. In reducing condition (+β-mercaptoethanol, β-ME), the complex was not intact, while in the absence of β-mercaptoethanol and without the heat treatment of the samples, higher molecular weight bands were present.

3.3 Nanobody and GPC3 binding

The binding of the nanobodies to GPC3 was confirmed, and the affinity was measured by SPR. The biotinylated nanobodies were immobilised on the SPR chip, and hGPC3^{core} and mGPC3^{core} were used as the analytes (Figure 3.4A). Both Nb49 and Nb58 bind human and mouse GPC3. Nb49 presents a higher affinity (Figure 3.4B-D). The affinity of the two nanobodies to hGPC3^{core} was higher than the affinity to the mGPC3^{core}, which is explained by the fact that the nanobodies were raised against the hGPC3^{core}.

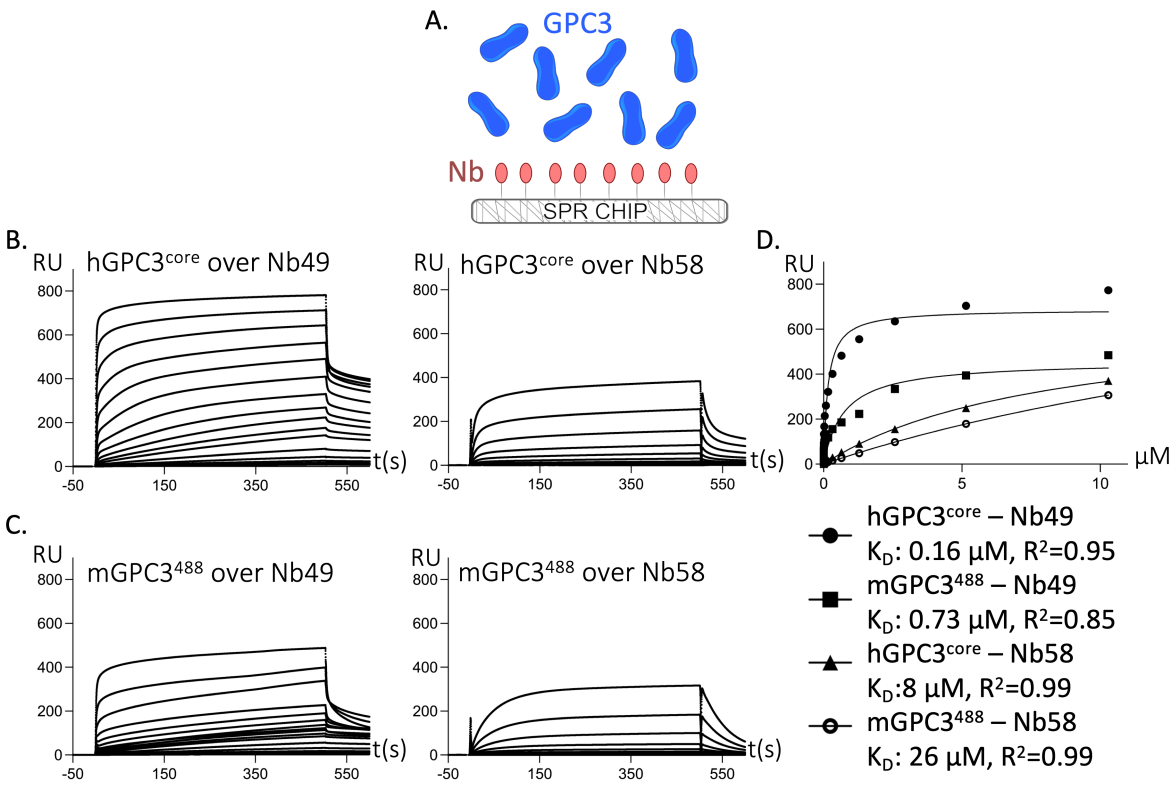


Figure 3.4: Nanobody - GPC3 binding.
(A) Schematic representation of the SPR experiment presented in this figure. Nb49 or Nb58 were immobilised on the chip at densities between 418 and 439 RU. (B) SPR data showing binding of hGPC3^{core} to Nb49 and Nb58. The highest concentration of hGPC3^{core} is 10.29 µM. (C) As in B, with mGPC3^{488}. The highest concentration of mGPC3^{488} is 10.29 µM. (D) The fitted curves of the SPR data, presented in panels B and C and the calculated K_D values. The K_D values were calculated assuming 1:1 binding.

The potential unspecific binding of the nanobodies to Unc5 was also tested since these nanobodies were planned to be used for the characterisation of the GPC3-Unc5 interaction. The rat Unc5D (rUnc5D^{ecto}) and the human Unc5B (hUnc5B^{Ecto}) extracellular domains were used as analytes for SPR test injections over Nb49 and Nb58. Both Unc5 homologues did not show any binding to either of the nanobodies (Figure 3.5 A, B). In a reverse experiment, the ectodomains of rUnc5D and mUnc5A were immobilised on the SPR chip, and Nb49 and Nb58 were flown over the chip in test injections. Likewise, in this experimental set-up, the nanobodies did not bind to rUnc5D^{ecto} or mUnc5A^{ecto} (Figure 3.5 C, D). Nb^{LPHN}, a nanobody raised against

LPHN3, was used as a negative control in several experiments as it does not bind GPC3 or Unc5. Its binding to GPC3 and Unc5 was tested in SPR, and Nb49 was used as a control, as it binds GPC3 but not Unc5. These experiments confirmed that Nb^{LPHN} does not bind either $hGPC3^{core}$ or $rUnc5D^{ecto}$ (Figure 3.5E, F).

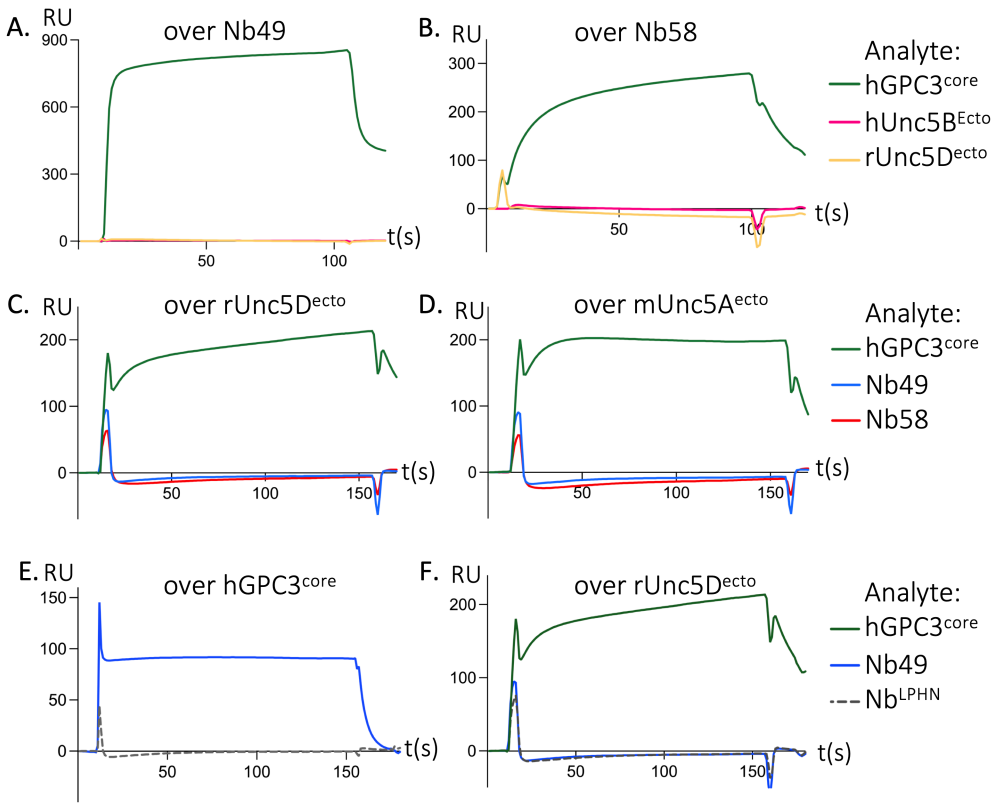


Figure 3.5: Nb49 and Nb58 do not bind Unc5 in SPR test injections.
(A-B) Nb49 and Nb58 were immobilised on the SPR chip (density: 418 - 439 RU), and $hUnc5B^{Ecto}$ (6.3 µM), $rUnc5D^{ecto}$ (6.3 µM) and $hGPC3^{core}$ (6.3 µM) were flown over. The Unc5 molecules do not bind to the nanobodies, while $hGPC3^{core}$, the positive control, binds. (C-D) Nb49, Nb58 (10 µM), and $hGPC3^{core}$ (8 µM) were used as analytes over $rUnc5D^{ecto}$ and $mUnc5A^{ecto}$ (density: 468 - 493 RU). GPC3 binds both Unc5s, but neither Nb49 nor Nb58 binds Unc5. (E-F) Nb^{LPHN} does not bind $hGPC3^{core}$ or $rUnc5D^{ecto,}$ and Nb49 (control) binds $hGPC3^{core}$ but does not bind $rUnc5D^{ecto}$. Nb^{LPHN} (10 µM), Nb49 (10 µM) and $hGPC3^{core}$ (8 µM) were flown over $hGPC3^{core}$ (density: 441 RU) and $rUnc5D^{ecto}$ (density: 468 RU).

3.4 Nb49 and Nb58 can manipulate the GPC3-Unc5 interaction

3.4.1 Pull-down experiments

The preliminary initial screening of the nanobodies performed in the lab, indicated that one of the nanobodies, Nb58, disrupts the GPC3-Unc5 interaction. Pull-down experiments were performed to confirm this initial finding. Streptavidin beads coated with biotinylated hGPC3^{core} were incubated with the human Unc5B extracellular region (hUnc5B^{Ecto}) mixed with the different nanobodies in different molar ratios (Figure 3.6 A). Nb49 was used as a control, as it binds GPC3 but should not interfere with the Unc5 interaction, according to the preliminary data.

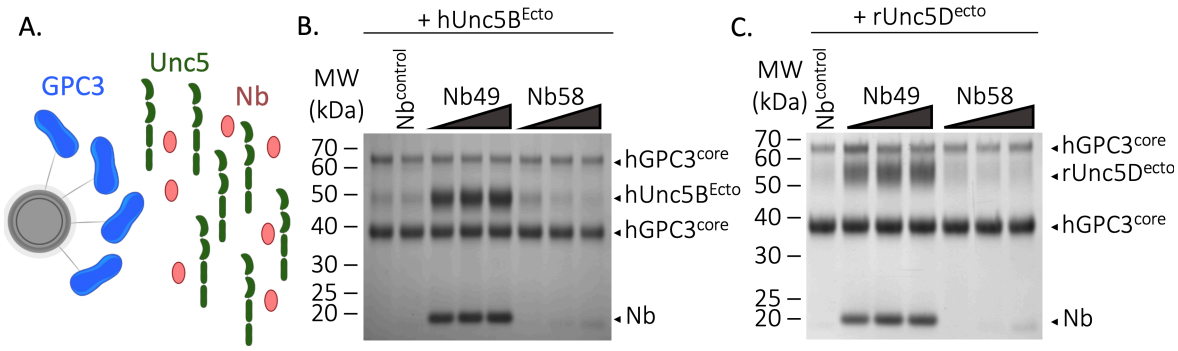


Figure 3.6: Pull-down assays show that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction.
(A) Schematic representation of the pull-down experiment. (B) SDS-PAGE of the pull-down samples with hUnc5B^{Ecto}. (C) Same as in B, but with rUnc5D^{ecto}. The increasing nanobody concentration corresponds to 1:1, 1:5 and 1:10 Unc5:Nb molar ratio. Nb^{control} = Nb^{LPHN} (Nb^{LPHN}:Unc5 10:1 molar ratio) in the control sample.

When Nb58 was mixed with hUnc5B^{Ecto} and incubated with the GPC3-coated beads, less Unc5 was bound on the GPC3 beads compared to the Nb^{LPHN} control. This effect was clearly visible when the molar ratio of hUnc5B^{Ecto}:Nb58 was 1 in 10, and the increasing Nb58 concentration showed a decrease in the Unc5B binding (Figure 3.6B). This experiment showed that Nb58 disrupts the GPC3-Unc5 interaction, but it also surprisingly showed that Nb49 enhances the interaction

(Figure 3.6B). In all the hUnc5B^{Ecto}/Nb49 molar ratio mixture samples, more hUnc5B^{Ecto} was bound on the GPC3 beads than the control sample.

Further pull-down experiments were performed with hGPC3^{core} on the beads, but this time the rUnc5D^{ecto} was used. These pull-down data agreed with the previous experiment that Nb49 enhances the interaction between the two proteins. However, they did not clearly show the disruption of the binding by Nb58, probably because the binding of the rUnc5D^{ecto} on the GPC3 beads was already rather low in this experimental set-up (Figure 3.6C).

3.4.2 SPR experiments

SPR was used to confirm the pull-down experiments finding that the two nanobodies enhance or inhibit the GPC3-Unc5 complex formation. For this purpose, SPR was performed in a competition set-up. Nb49 and Nb58 were immobilised on the SPR chip, and a mixture of hGPC3^{core} or mGPC3^{488} with rUnc5D^{ecto} was used as the analyte. This experiment confirmed that Nb49 enhances while Nb58 inhibits the GPC3-Unc5 interaction (Figure 3.7).

SPR experiments were also performed with biotinylated Unc5 extracellular domains (Unc5^{ecto}) immobilised on the chip, and a mixture of GPC3 with either Nb49, Nb58 or Fc was used as the analyte (Figure 3.8A). Different homologues of Unc5 ectodomains (mUnc5A^{ecto}, hUnc5B^{ecto}, rUnc5D^{ecto}) were used to confirm that the nanobodies had the same effect with other Unc5s as it was already known that GPC3 could bind all the Unc5 homologues (A-D) [30]. Same as in the previous SRP experiment, Nb49 enhances the binding of Unc5 to GPC3 compared to the control, and Nb58 interrupts the binding. The same pattern was observed for mouse and human GPC3 but also for mUnc5A, mUnc5B and rUnc5D (Figure 3.8B, C, D).

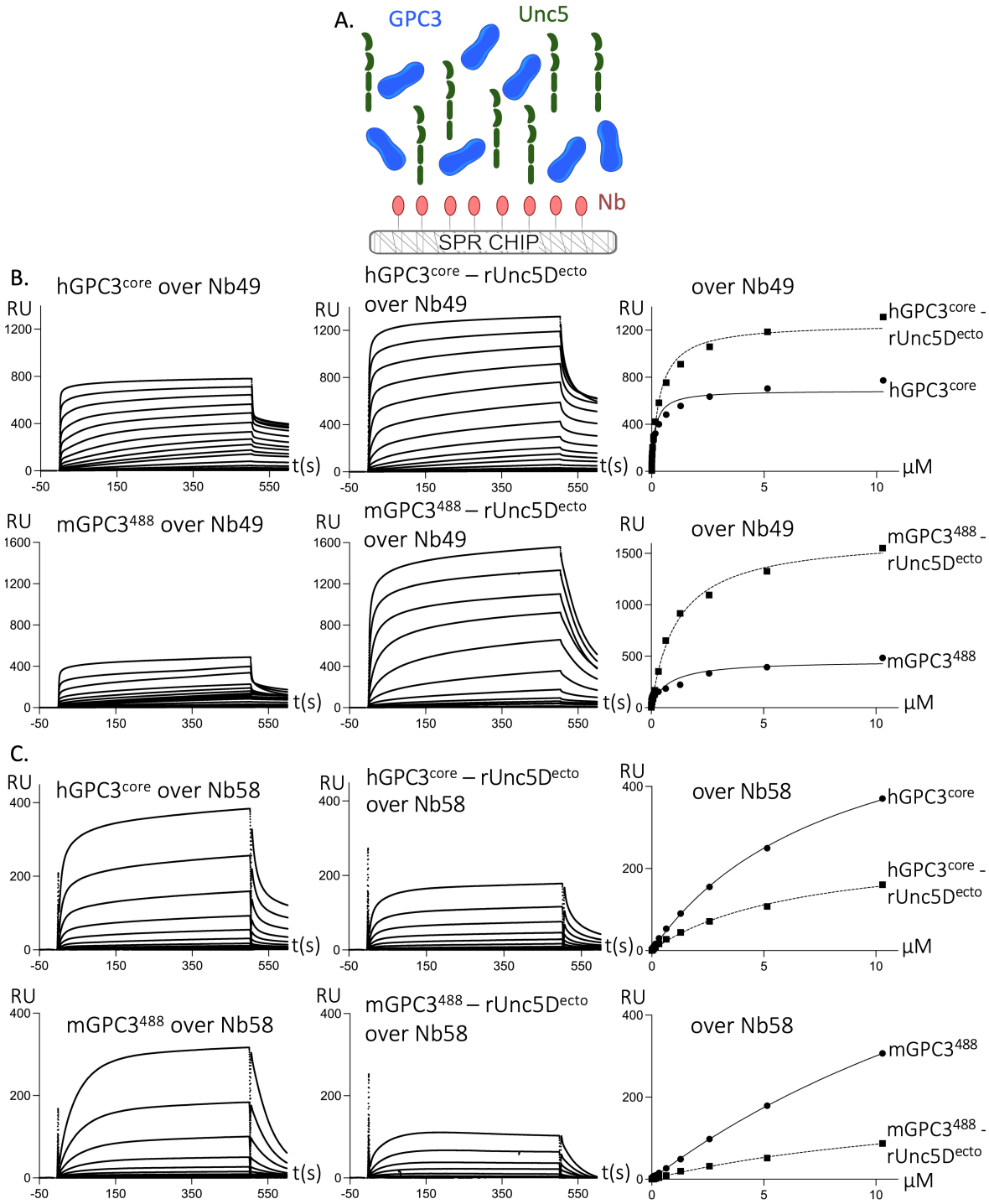


Figure 3.7: SPR experiment confirms that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction.

(A) Schematic representation of competition SPR where the nanobodies (Nb49 or Nb58) were immobilised on the SPR chip (density: 418 - 439 RU) and GPC3 alone or mixed with Unc5 were flown over. (B) Competition SPR data in which Nb49 is immobilised on the chip show that Nb49 enhances the GPC3-Unc5 interaction. (C) Competition SPR data in which Nb58 is immobilised on the chip show that Nb58 inhibits the interaction. The high concentration of $hGPC3^{core}$, $mGPC3^{488}$ and $rUnc5D^{ecto}$ is 10.29 µM.

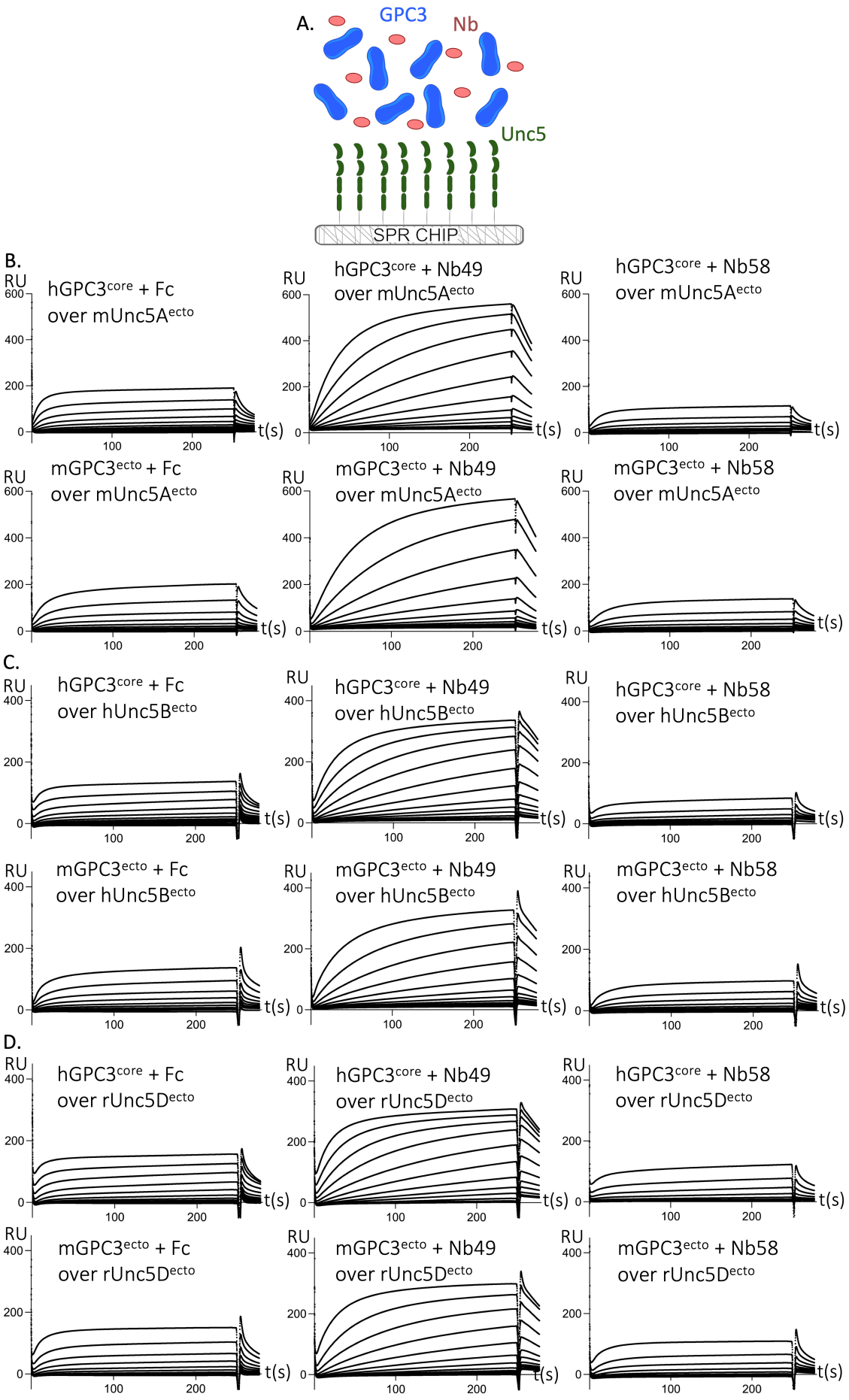
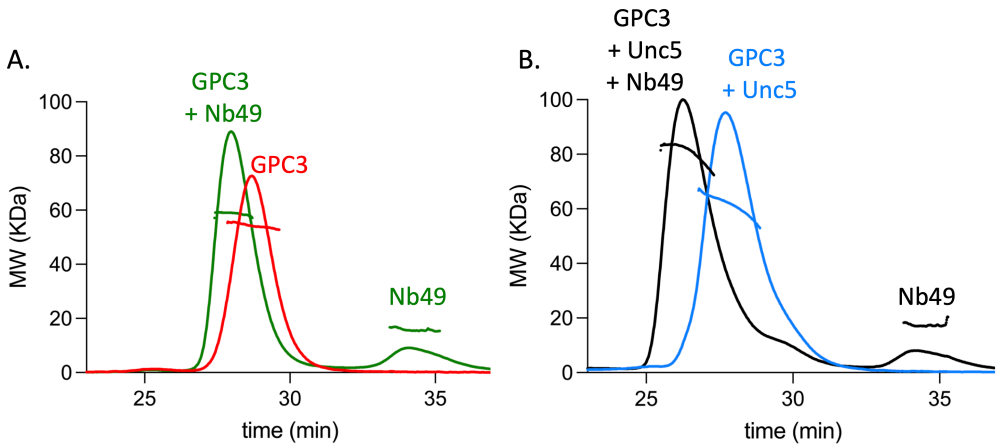

A.
GPC3
Nb
Unc5
SPR CHIP
B.
RU
hGPC3core + Fc
over mUnc5Aecto
hGPC3core + Nb49
over mUnc5Aecto
hGPC3core + Nb58
over mUnc5Aecto
mGPC3ecto + Fc
over mUnc5Aecto
mGPC3ecto + Nb49
over mUnc5Aecto
mGPC3ecto + Nb58
over mUnc5Aecto
C.
hGPC3core + Fc
over hUnc5Becto
hGPC3core + Nb49
over hUnc5Becto
hGPC3core + Nb58
over hUnc5Becto
mGPC3ecto + Fc
over hUnc5Becto
mGPC3ecto + Nb49
over hUnc5Becto
mGPC3ecto + Nb58
over hUnc5Becto
D.
hGPC3core + Fc
over rUnc5Decto
hGPC3core + Nb49
over rUnc5Decto
hGPC3core + Nb58
over rUnc5Decto
mGPC3ecto + Fc
over rUnc5Decto
mGPC3ecto + Nb49
over rUnc5Decto
mGPC3ecto + Nb58
over rUnc5Decto
t(s)

Figure 3.8: Another SPR experiment confirms that Nb49 enhances and Nb58 disrupts the GPC3-Unc5 interaction. (Figure in previous page p. 72)
(A) Schematic representation of the competition SPR in which Unc5 homologues were immobilised on the SPR chip (density: 473 – 483 RU) and GPC3 mixed with the nanobodies, or the Fc protein (control) were flown over. (B - D) Competition SPR data in which mUnc5A^{ecto} in B, hUnc5B^{ecto} in C and rUnc5D^{ecto} in D were immobilised on the chip. hGPC3^{core} or mGPC3^{ecto} (top concentration 4.5 μM) mixed with Nb49, Nb58 or Fc were flown over. The concentration of the nanobodies was kept constant in the 2-fold dilutions at 9 μM for Nb49 and 4.5 μM for Nb58.

3.5 GPC3-Unc5-Nb49 complex stoichiometry

SEC-MALS experiments were performed to further biophysically characterise the GPC3-Unc5-Nb49 complex. Nb49 appeared to enhance the interaction between Unc5 and GPC3 in pull-down and SPR experiments. SEC-MALS was used in order to determine the stoichiometry of the GPC3-Unc5-Nb49 complex. For this experiment, the four samples that were prepared were hGPC3^{core}, hGPC3^{core} mixed with Nb49 at 1:2 molar ratio (Figure 3.9A), hGPC3^{core} mixed with rUnc5D^{Ig1Ig2TSP1} at 1:1 molar ratio and the last sample was a 1:1:2 molar ratio mix of hGPC3^{core}, rUnc5D^{Ig1Ig2TSP1} and Nb49 (Figure 3.9B). An excess of the nanobody was used to increase its potency and the concentration of GPC3 was the same in all the samples.

Previous SEC-MALS experiments in the lab that aimed to determine the GPC3-Unc5 complex stoichiometry showed that the complex formation is concentration-dependant, and the interaction is not strong enough to keep the complex intact after the dilution of the sample in the size exclusion chromatography step. Therefore, SEC-MALS was not the best technique for this purpose.



	Protein complex	Apparent MW (kDa)	Expected MW (kDa)
—	$hGPC3^{core}$	54.2 ± 0.40 %	59.62
—	$hGPC3^{core}$ +Nb49	58.7 ± 0.30 %	77.17
—	$hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$	61.3 ± 0.20%	94.11
—	$hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$ + Nb49	80.7 ± 0.30%	111.66

Nb49: 17.55 kDa, $hGPC3^{core}$: 59.62 kDa, $rUnc5D^{Ig1Ig2TSP1}$: 34.49 kDa

Figure 3.9: SEC-MALS experiment for the GPC3-Nb49 and the GPC3-Unc5-Nb49 complexes.
SEC-MALS indicates that Nb49 binds to GPC3 and enhances the GPC3-Unc5 interactions since the $hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$ + Nb49 complex elutes earlier than the $hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$ and a higher molecular weight is predicted. However, the apparent MWs calculated by MALS do not match the expected MW predicted by the sequences of the constructs because the complex formation is concentration dependant. The concentration of the proteins injected was ~ 1 mg/ml, but with the SEC step, the concentrations resulted in at least a 10-fold dilution. The proteins' properties did not allow an injection at a higher concentration. The MW values include the MW of the N-linked glycans of the proteins based on the predicted N-linked glycosylation sites.

In this SEC-MALS experiment, the assumption was that Nb49, which enhances the interaction, would keep the complex intact through size exclusion chromatography prior to the MALS step. However, in agreement with the previous SEC-MALS experiment, it was not possible to measure the accurate molecular weights of these complexes. The peaks are shifted when comparing the $hGPC3^{core}$ with the $hGPC3^{core}$ + Nb49 and the $hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$ with the $hGPC3^{core}$ + $rUnc5D^{Ig1Ig2TSP1}$ + Nb49, but the calculation of the molecular weights of these

complexes does not correspond to the molecular weight estimated from the protein sequence. The same reason as in the previous SEC-MALS experiments might be responsible for this problem. The complex has a weak affinity which depends on the concentration, making it difficult for SEC-MALS to be accurate due to the very low concentration of the sample after the size exclusion chromatography step.

3.6 Mouse cortex lysate pull-down experiments

The confirmation of the expression of GPC3 in the developing mouse brain was essential before the following mouse *in vivo* experiments were performed. The del Toro lab used the biotinylated Nb49 and biotinylated hGPC3 (prepared as the biotinylated Nb49) for pull-down experiments using mouse cortex lysates on different embryonic days (Figure 3.10A). The samples after the pull-down experiment were prepared and sent for mass spectrometry in the Max Planck Institute of Biochemistry Proteomic facility in Martinsried, Germany.

After data processing, the results showed that GPC3 was pulled down by Nb49 and it was highly enriched on embryonic day E15.5, confirming its presence in the mouse cortex (Figure 3.10B). Except for the presence of GPC3 in the mouse brain cortex, this experiment also shows the ability of the Nb49 to pull down GPC3. GPC3 was not pulled down on embryonic days E13.5 and E14. Moreover, Unc5D was present in the mass spectrometry data of the Nb49 pull-down experiment on E15.5, presenting an indication that GPC3 and Unc5D interact *in vivo* (Figure 3.10B).

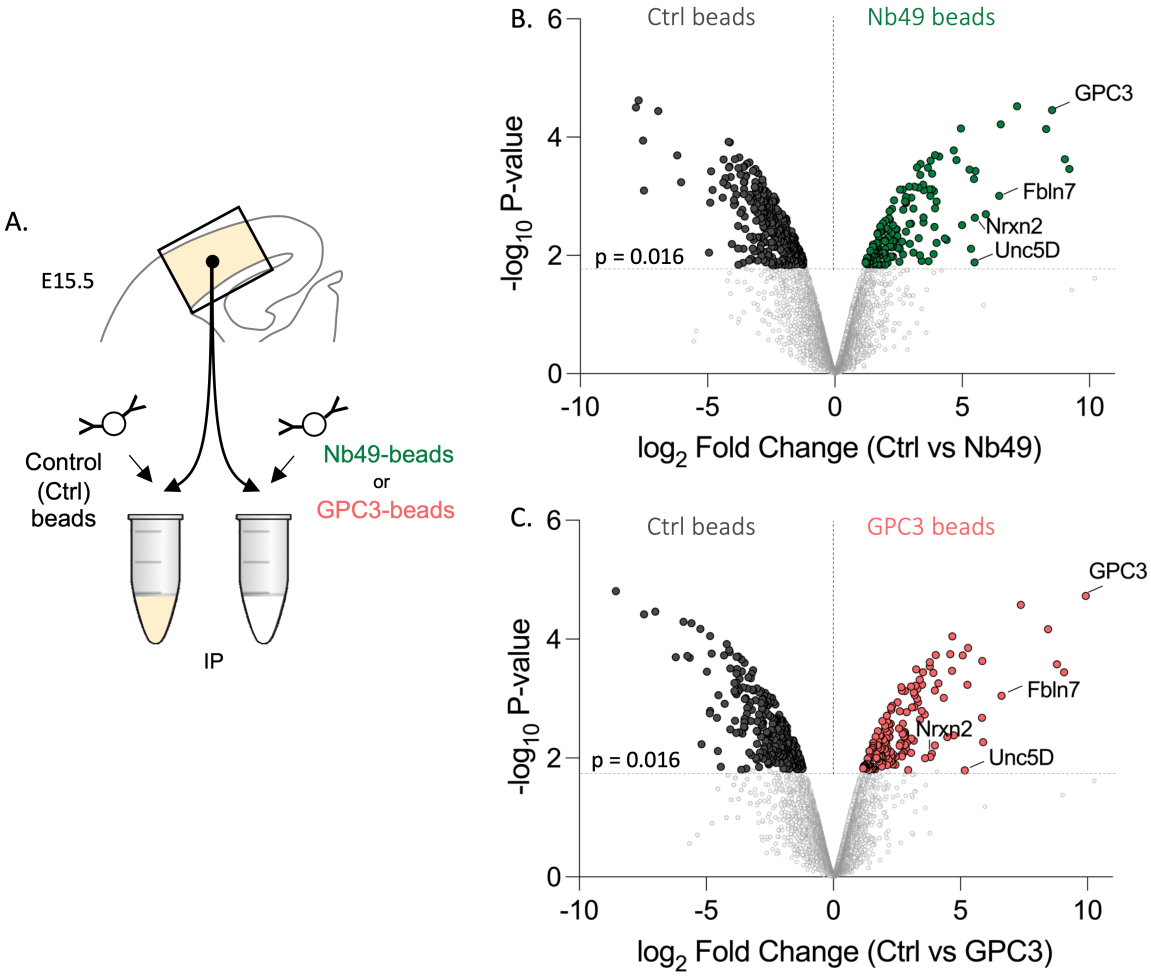


Figure 3.10: Pull-down experiment from mouse cortical lysate.
(A) Biotinylated Nb49 and GPC3 on streptavidin beads were used for pull-downs from mouse cortical lysates. The streptavidin beads were used alone as the control. (B-C) Volcano plots present the enriched proteins from the Nb49 and GPC3 pull-downs and the subsequent mass spectrometry. The enriched proteins from the control beads are presented in dark grey, and the non-significant ones in light grey. In the Nb49 pull-down, GPC3 was highly enriched, which shows that it is present in the E15.5 mouse cortex. Figure adapted from [30].

More pull-down experiments from brain lysates were performed with biotinylated GPC3. Unc5D was also present in these pull-downs (Figure 3.10C). GPC3 is highly enriched in this experiment, but it represents the protein that was used for the pull-down and not the endogenous GPC3.

Apart from GPC3 and Unc5D, many proteins were present in the mass spectrometry data after this pull-down experiment. A careful look at all the proteins

might provide indications of potential new GPC3 binding partners. A review of all these proteins showed that many of them are intracellular proteins that probably cannot interact with GPC3 physiologically. However, it also revealed extracellular proteins that can potentially be GPC3 binding partners. The two most relevant examples of these proteins that can bind GPC3 are Fibulin 7 (Fbln7) and Neurexin 2 (Nrxn2). Fibulin 7 is an extracellular protein that participates in several cellular functions like growth, adhesion, survival, and motility [200], and Neurexin 2 is a member of a membrane protein family that plays a role in synapse formation [201]. More experiments are needed to explore this.

3.7 More nanobodies against GPC3

Out of the pool of nanobodies against GPC3, already available in the lab, only two nanobodies (Nb49, Nb58) were fully characterised initially, to be used as tools in *in vitro* and *in vivo* assays. Next, other nanobodies from the original 96-well plate of nanobody sequences were investigated. The different nanobody sequences were put into groups based on their CDR3 sequence identity. This categorisation of the sequences resulted in six nanobody classes and eight individual nanobodies. For the following experiments, one representative nanobody from each group and all the individual nanobodies were expressed in the WK6 *E. coli* strain and purified (Figure 3.11A). The nanobodies were named with a number from 48 to 61 (Nb48 - Nb61), and the two nanobodies already characterised (Nb49 and Nb58) are included in this nomenclature.

A preliminary screening of all the nanobodies was conducted. Firstly, all the nanobodies were tested for binding to GPC3 with SPR. Biotinylated $hGPC3^{core}$ was immobilised on the SPR chip, and all the nanobodies were used as analytes in test

injections. Nb49 and Nb58 were the positive controls as it was already established that they bind hGPC3^{core}. Nb^{LPHN} was used as a negative control. All the nanobodies showed binding to hGPC3^{core} with different affinities (Figure 3.11B). As expected, the negative control, Nb^{LPHN}, did not bind GPC3. Several nanobodies bind GPC3 better than Nb49, a very important characteristic for the nanobody utilisation in other experiments.

In the same experiment, the potential unspecific binding of the nanobodies on Unc5 was tested. Biotinylated mUnc5A^{ecto} and rUnc5D^{ecto} were immobilised on the SPR chip, and all the nanobodies were flown over. None of the GPC3 nanobodies binds to mUnc5A or rUnc5D (Figure 3.12).

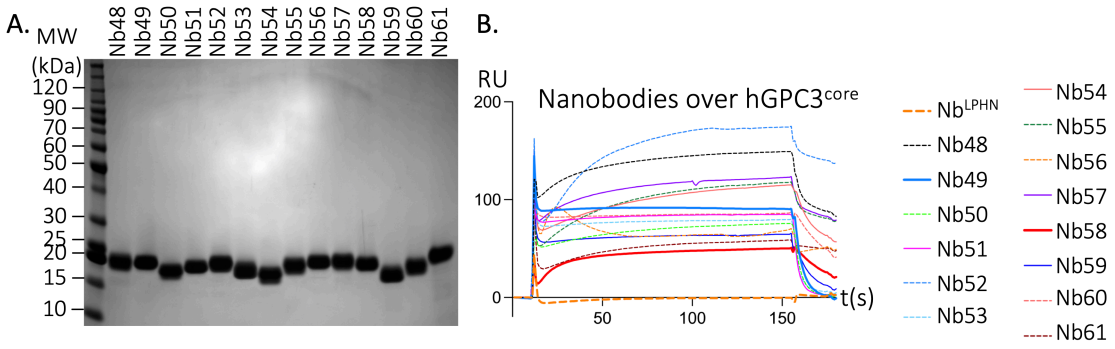


Figure 3.11: A pool of nanobodies against GPC3.
(A) SDS-PAGE gel of all the nanobodies after the final step of their purification (B) SPR test injections of all the nanobodies (10 μM) over the immobilised on the SPR chip, hGPC3^{core} (density: 441 RU). All the nanobodies that bind to hGPC3^{core} present a higher response compared to Nb58 (red). Also, several nanobodies have a response higher than the Nb49 (light blue). Nb^{LPHN} (control) does not bind hGPC3^{core} (dashed orange).

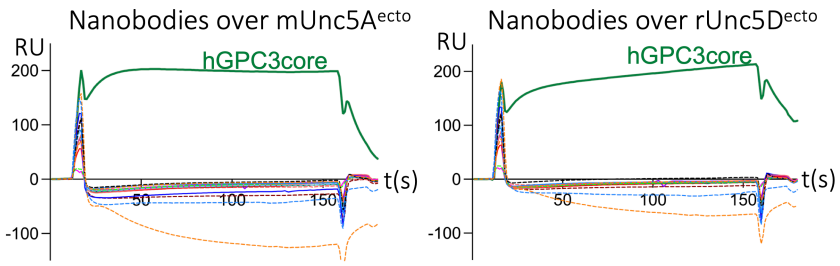


Figure 3.12: The nanobodies against GPC3 do not bind Unc5.
All the nanobodies (10 μM) and hGPC3^{core} (8 μM) were flown over immobilised mUnc5A^{ecto} and rUnc5D^{ecto} (density: 468 - 493). Only hGPC3^{core} and none of the nanobodies bind Unc5A and Unc5D.

After confirming that all the available nanobodies bind GPC3, SPR test injections in a competition format followed in order to investigate if any of these nanobodies affect the GPC3-Unc5 interaction. In this experiment, rUnc5D^{ecto} and mUnc5A^{ecto} were immobilised on the chip, and a mixture of hGPC3^{core} with each nanobody was flown over. Nb^{LPHN} was used as a control as it should not interfere with the Unc5-GPC3 interaction. Purified hGPC3^{core} was also flown over the chip without any nanobody to set the baseline.

The hGPC3^{core} mixed with the control nanobody, Nb^{LPHN}, had the same response as hGPC3^{core} alone, which shows that the control nanobody does not interfere with the GPC3-Unc5 interaction. As expected, the GPC3-Nb49 sample has the most significant response out of all the nanobodies, and it is the one already proven to enhance the GPC3-Unc5 interaction. The GPC3-Nb58 sample presents less binding of GPC3 to the immobilised Unc5s, which was also expected based on the previous data showing that Nb58 blocks the GPC3-Unc5 interaction. Nb61 and Nb53 behave like Nb58 and Nb49, respectively, but further characterisation is needed to confirm that. Interestingly, Nb60 appears to block the GPC3-Unc5 interaction completely (Figure 3.13).

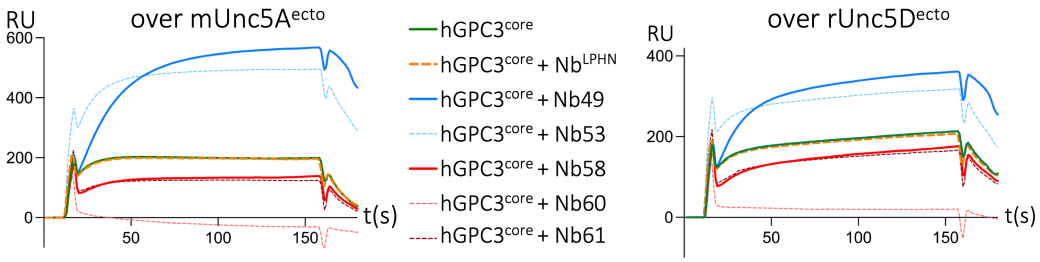


Figure 3.13: A few new nanobodies are interesting with respect to the GPC3-Unc5 interaction.
mUnc5A^{ecto} and rUnc5D^{ecto} (density: 468 - 493 RU) were immobilised on the SPR chip, and hGPC3^{core} mixed with nanobodies (1:1 molar ratio, 8 µM) were used as the analytes in SPR test injections. Nb53 and Nb49 similarly enhance the GPC3-Unc5 interaction, whereas Nb60 completely blocks the interaction.

Most of the other nanobody-GPC3 samples present a higher response compared to the baseline (hGPC3^{core}), but it is unclear how exactly they affect the GPC3-Unc5 interaction (Figure 3.14). This result might show that they help the formation of the complex like Nb49, but it can also mean that they just do not block the interaction and both the nanobody and Unc5 can bind GPC3 at the same time. Further experiments are needed to confirm their function.

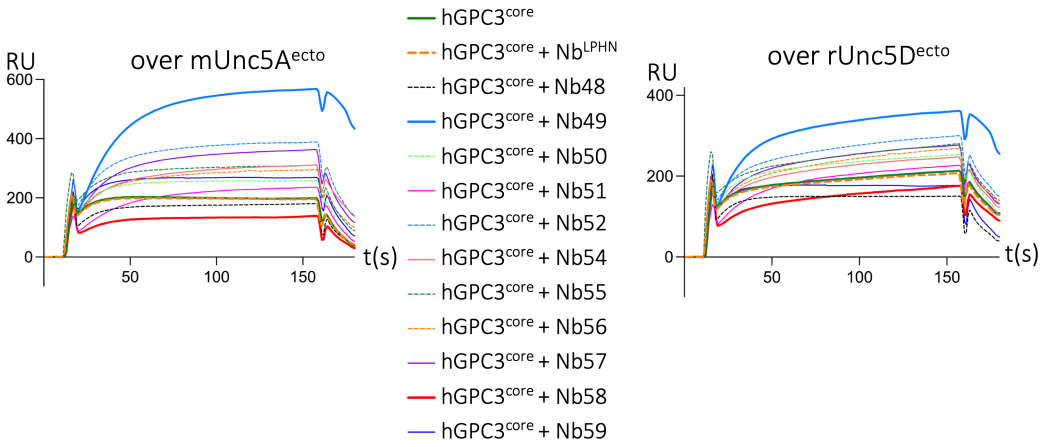


Figure 3.14: More nanobodies against GPC3 and how they affect the GPC3-Unc5 interaction.
SPR test injections with mUnc5A^{ecto} and rUnc5D^{ecto} (density: 468 - 493 RU) on the SPR chip and hGPC3^{core} mixed with nanobodies (1:1 molar ratio, 8 µM) as the analytes. The response curves of the nanobodies mixed with the hGPC3^{core} should be compared with the hGPC3^{core} (green), hGPC3-Nb49 (light blue) and hGPC3-Nb58 response curves (red).

3.8 Discussion

In this chapter, work on nanobodies against hGPC3^{core} was described. Two nanobodies, Nb49 and Nb58, were biophysically characterised. It was confirmed that both nanobodies bind human and mouse GPC3, and Nb49 binds GPC3 with higher affinity compared to Nb58. The human and mouse GPC3 sequences are 94% identical, and the structures of their core proteins are very similar [30], which explains why the nanobodies bind both homologues. The binding of the nanobodies to mouse GPC3 was essential to be confirmed because of the use of these nanobodies in *in vivo* mouse experiments.

With respect to the GPC3-Unc5 interaction, Nb49 enhances, whereas Nb58 inhibits the interaction. These characteristics were confirmed with SPR and pull-down assays. Also, Nb49 does not bind Unc5, which shows that it does not enhance the interaction because of its binding to both proteins at the same time. SEC-MALS was used to verify the stoichiometry of the Nb49-GPC3 and the GPC3-Unc5-Nb49 complexes since Nb49 enhances the GPC3-Unc5 complex formation. However, it was shown that SEC-MALS was not the right technique for this, as the complexes were not stable after the SEC dilution. SEC-MALS experiments with serial dilutions of the complexes can be used to confirm the concentration-dependent interaction since the proteins cannot get concentrated high enough.

At the time, the primary need was to characterise in detail the two nanobodies already presenting signs of being able to modulate the GPC3-Unc5 interaction. However, more anti-GPC3 nanobodies were available. SPR experiments revealed nanobodies with high affinity against GPC3 and nanobodies that affect the GPC3-Unc5 interaction that can be utilised in future research.

The aim of the characterisation of Nb49 and Nb58 was the use of these nanobodies in experiments which will help to understand better the GPC3-Unc5 interaction. These experiments are published in [30] and are summarised below.

3.8.1 The utilisation of the nanobodies in stripe assays

Stripe assays have been used widely in exploring how cells migrate, or axons grow in response to different guidance molecules [83,202,203]. Cortical neurons and Unc5-expressing cells on stripes showed a preference for migrating towards the non-Unc5-binding GPC3 mutant versus GPC3-coated stripes. The cells also preferred control Fc over GPC3. In conclusion, all these results indicate that in this *in vitro* experiment, the GPC3-Unc5 interaction has a contact-repulsion effect which does not depend only on GPC3 interacting with Unc5 (Chapter 1.5.2) [30].

The Nb49 and Nb58 were tested in the stripe assay set-up in a tetrameric format in which biotinylated nanobodies were linked with streptavidin. As the characterisation experiments indicated, Nb49 enhances the GPC3-Unc5 interaction and the cell repulsion in the stripes assays since the repulsive effect is stronger in the Nb49 presence compared to the Nb58. The cells preferred the non-Unc5-binding GPC3 mutant and migrated away from the red stripe (GPC3 WT) more when Nb49 was present compared to the control. In the stripe assays with Nb58, which inhibits the interaction, the repulsive effect is lower and the cells preferred the GPC3 WT more than in the control condition (Figure 3.15). Cortical neurons were also examined in this experimental set-up. In the presence of Nb58, they preferred the red stripe (GPC3 WT) more than in the control condition. Interestingly, Nb49 immobilised the neurons on the cell surface, and no migration was quantified.

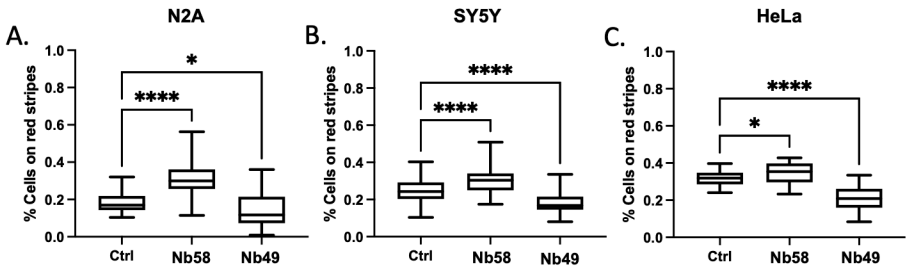


Figure 3.15: Stripe assays show that Nb49 enhances and Nb58 inhibits the repulsive effect of the GPC3-Unc5 interaction.
(A-C) The N2A, SY5Y and HeLa cells which express the Unc5 receptors endogenously, were tested in stripe assays. In this assay, the hGPC3^{core} WT is immobilised on the red stripes and the hGPC3^{core} non-Unc5-binding mutant on the black stripes. The control experiment was performed in the presence of streptavidin, and the Nb49 and Nb58 were added after tetramerisation with streptavidin. Statistical analysis: one-way ANOVA with Tukey's post hoc tests, *$p < 0.05$; ****$p < 0.0001$. Figure adapted from [30].

3.8.2 GPC3 immunostaining on mouse cortex using Nb49

As indicated by the expression data on the mRNA level, GPC3 showed expression at the ventricular zone where the apical progenitors are from E13.5 to E17.5 embryonic days. Also, single-cell RNA-seq data revealed that GPC3 is present mainly in the apical progenitor cells, which agrees with *in situ* hybridisation data [30]. At the protein level, the pull-down experiment from mouse cortex lysate using Nb49 showed that GPC3 is present in the mouse brain cortex on E15.5 and highly enriched (Chapter 3.6).

In another experiment which was used to confirm the GPC3 presence in the mouse cortex on the protein level, the Nb49-Streptavidin594 complex (Chapter 3.2) was used by the del Toro lab for GPC3 immunostaining on a coronal section of the E15.5 mouse cortex (Figure 3.16).

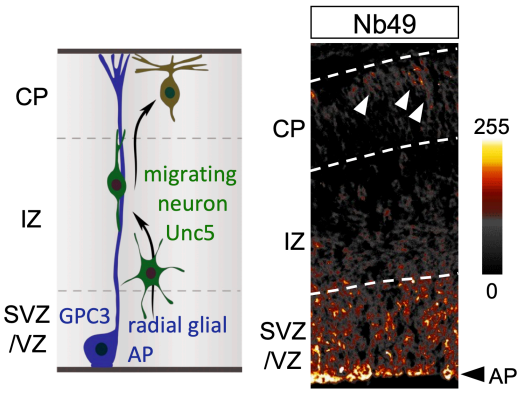


Figure 3.16: Immunostaining of GPC3 using Nb49 on a coronal section of the E15.5 mouse cortex.
GPC3 is immunostained in the mouse cortex by the fluorescent streptavidin attached to Nb49. The image is coloured by the intensity of the staining, and GPC3 is mainly present in the apical progenitor cells and the end feet of the apical progenitor fibres, as indicated by the arrow. CP: cortical plate; IZ: intermediate zone; SVZ: subventricular zone; VZ: ventricular zone; AP: apical progenitors. Figure adapted from [30].

This experiment showed that GPC3 is primarily located in the ventricular zone, where the apical progenitor cell bodies are. There is also some lower-level presence of the protein in the intermediate zone and the cortical plate where the fibres and the end feet of the apical progenitors are [30]. This result agrees with the previous expression data.

3.8.3 *In vivo* expression of the Nb49 and Nb58 in the mouse cortex

The nanobodies were used in *in vivo* experiments to help understand the function of the GPC3-Unc5 interaction. In the developing mouse cortex, GPC3 is located in the apical progenitor cells, while Unc5D is enriched in the migrating neurons. Based on the expression pattern of these proteins, a working model was developed where the scaffold apical progenitor cells express GPC3 and the migrating neurons express Unc5D (Figure 3.17A). Wanting to examine if and how the GPC3-Unc5 interaction affects neuronal migration, *in-utero* electroporation was used by the del Toro lab at E13.5 to introduce the desired constructs in the cells of the developing

cortex. The effects of the overexpression of these proteins on the development of the brain were assessed on day E16.5 when migration was already completed.

The nanobodies were overexpressed in the cortex in order to block or stabilise the GPC3-Unc5 interaction *in vivo* and evaluate how this interferes with neuronal migration and, thus, cortical development. Compared to the control, the expression of secreted Nb49 or Nb58 in the cortex delayed the migration of the neurons to the upper cortical plate (Figure 3.17B, C).

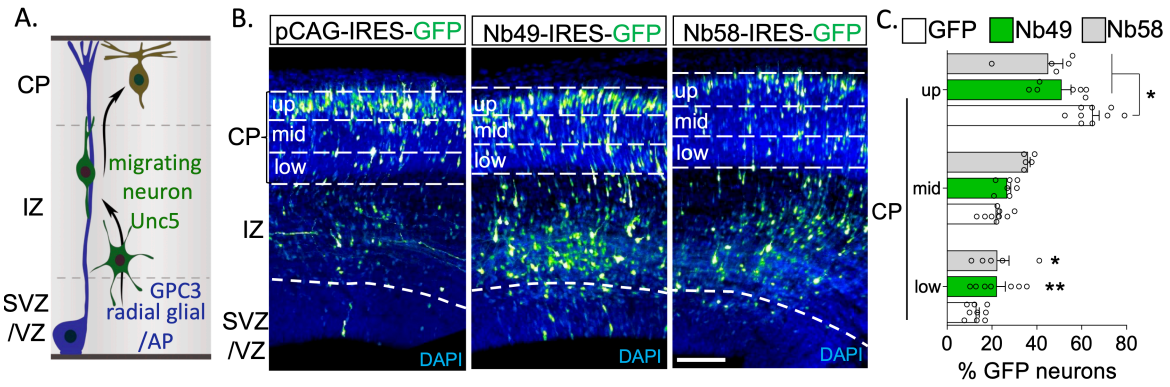


Figure 3.17: Nb49 and Nb58 affect the radial migration in the mouse cortex. (A) Schematic representation of the summary model of radial migration on a coronal section of the mouse cortex based on the Unc5D and GPC3 expression. (B) Coronal sections of the mouse cortex at E16.5 after *in-utero* electroporation of the empty pCAGIG vector (control) and the vectors for the expression of Nb49 and Nb58. The GFP-positive cells are coloured green. (C) Quantification of the experimental data shown in B. n = 9 GFP, n = 7 Nb49, and n = 5 Nb58 electroporated brains. Statistical analysis: one-way ANOVA test with Tukey's post hoc, *p < 0.05, **p < 0.01. Scale bar = 100 µm. CP: cortical plate; IZ: intermediate zone; SVZ: subventricular zone; VZ: ventricular zone; AP: apical progenitors. Figure adapted from [30].

Despite the fact that Nb49 stabilises the GPC3-Unc5 interaction, it delayed the radial cell migration like Nb58 did. The neuronal migration in cortical development is a fine-tuned process, and the cell surface molecules bind and release their partners to find the balance that is needed for the correct migration of the cells. Interactions need to be released in order for the cells to move, and it is hypothesised that Nb49 disrupted this balance, and that is what caused the delay in the radial

migration. Also, the delayed migration of the neurons in the presence of Nb49 goes in line with the stripe assay result in which the cortical neurons in the presence of Nb49 were immobilised on the stripes surface.

This experiment completes a series of *in vivo* cell migration experiments in the mouse cortex. The overexpression of a secreted construct of Unc5D and the knockdown of GPC3 using shRNA both delayed the radial neuronal migration (Chapter 1.5.3) [30]. All these results conclude that the GPC3-Unc5 interaction is an essential regulator of cortical migration *in vivo*.

3.8.4 The effect of the Nb49 and Nb58 in the neuroblastoma cell migration

A model of neuroblastoma cell migration [112] was additionally used to characterise the GPC3-Unc5 interaction effect in cell migration by the Castellani Lab. For this experiment, the neuroblastoma cells, SY5Y, were engrafted in a chick embryo's neural crest. It was confirmed that they express the Unc5 receptors and GPC3. Both the overexpression of a secreted construct of Unc5D and the silencing of GPC3 using siRNA in the SY5Y xenograft cells showed that the GPC3-Unc5 interaction is essential for the correct neuroblastoma cell migration and for the cellular cohesion of the cancer cells in the collective cell migration that characterise them [30] (Chapter 1.5.4).

Secreted Nb49 and Nb58 were overexpressed in the SY5Y xenograft cells (Figure 3.18). Nb58, which inhibits the GPC3-Unc5 interaction, affected the neuroblastoma cell migration, and only a few cells reached the target sympatho-adrenal (SA) derivatives. At the same time, it was noticed that non-transfected cells behaved the same as the transfected cells, which indicates that the secreted Nb58 have a paracrine effect as well. On the other hand, overexpression of Nb49 resulted in further migration of the cells and the formation of tumour masses in the most

distal SA territories. With Nb49, cells did not lose their cohesive phenotype and fewer cells were found outside the tumour sites. These results show that the strength of the GPC3-Unc5 interaction plays an essential role in the neuroblastoma cell migration in this model, and the alteration of it affects the migration.

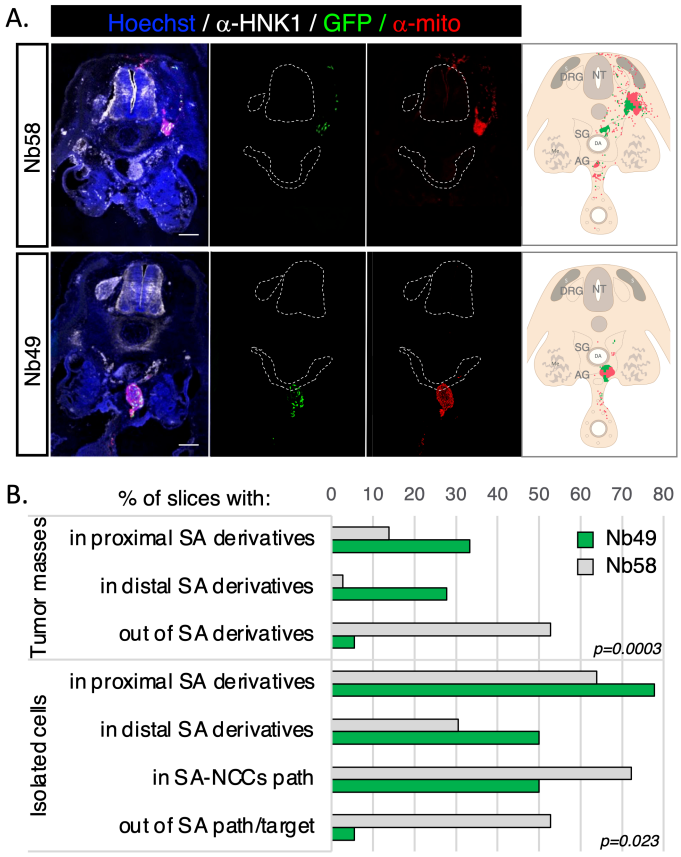


Figure 3.18: The expression of Nb49 and Nb58 in the neuroblastoma cells of this migration model affects their migration.
(A) The xenograft cells SY5Y were transfected with vectors that encode Nb49 or Nb58, and soluble nanobodies were expressed in the graft. The neural crest cells and therefore the structures that they form after migration are detected with the anti-HNK1 antibody (white). The schematic representation of the SY5Y migration in the chicken embryos shows that the cells expressing Nb58 have a delayed migration and do not reach the target territories. In contrast, cells expressing Nb49 migrate past the target. Also, cells expressing Nb49 maintain their cohesive phenotype. All cells: green and red. Non-transfected cells: red. (B) Quantification of the SY5Y cells and the tumour positions two days after grafting (Nb49: 5 embryos and 18 sections. Nb58: 8 embryos and 36 sections). Statistical analysis: χ^2 tests. Scale bar = 200 µm. NT: Neural Tube; S: Somite; DRG: Dorsal Root Ganglia; SG: Sympathetic Ganglia; AG: Adrenal Gland; DA: Dorsal Aorta; Me: Mesonephros; Hoechst: stains the nuclei; α-mito antibody: detects the human neuroblastoma cells (SY5Y graft). Figure adapted from [30].

4. Structural characterisation of the nanobody-GPC3 complexes

4.1 Introduction and summary

Two nanobodies raised against GPC3 were identified as modifiers of the GPC3-Unc5 interaction, Nb49 enhances, and Nb58 inhibits the interaction. The next step after the biophysical characterisation of these nanobodies was the structural characterisation of them in complex with GPC3 in order to understand how they enhance or block the GPC3-Unc5 interaction. Several crystallisation trials and optimisation steps were performed. Crystals of the GPC3-Nb49 complex were successfully obtained and diffracted at ~4 Å resolution. The crystal structure revealed the binding epitope of Nb49 on GPC3. Cryo-EM grids with the GPC3-Unc5-Nb49 complex were prepared and screened in order to learn more about the way that Nb49 enhances the GPC3-Unc5 interaction. Substantial progress was made in the optimisation of the Unc5-GPC3-Nb49 complex particles on cryo-EM grids, but the conditions were still not optimal for data collection.

4.2 Crystallisation of the Nb49-GPC3 and Nb58-GPC3 complexes

Following the biophysical characterisation, the next aim was the structural characterisation of the nanobodies, Nb49 and Nb58, in complex with GPC3. The human GPC3 core protein ($hGPC3^{core}$) and the nanobodies were recombinantly expressed and purified for crystallisation trials. A construct of the $hGPC3^{core}$ was already available in the lab and consisted of only the core protein of hGPC3 (residues: 32 - 483) followed by a 6 × histidine tag. The $hGPC3^{core}$ was expressed

in HEK-293S GnTI$^-$ cells, resulting in a protein homogeneously glycosylated, a characteristic important for crystallisation. The protein was purified with a standard two-step purification, starting with a nickel affinity step followed by SEC. The purest fractions were pooled and used for the crystallisation trials (Figure 4.1A). As previously described, the nanobodies were expressed in WK6 bacterial cells, were purified, and the pure fractions were used for the crystallisation trials (Figure 4.1B).

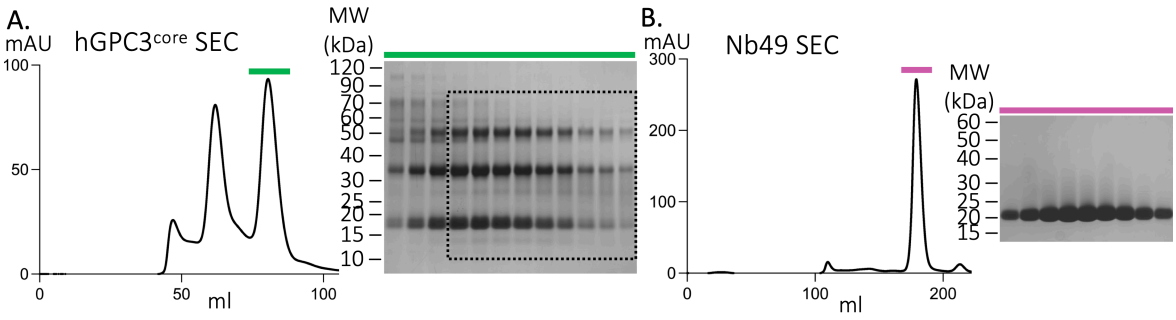


Figure 4.1: The hGPC3^{core} and Nb49 purifications.
(A) SEC profile of the hGPC3^{core} purification after expression in HEK-293S GnTI$^-$ cells and the SDS-PAGE gel of the peak fractions. The fractions in the dotted line box were pooled for the crystallisation trials. The SDS-PAGE of hGPC3^{core} presents three protein bands because of the cleavage of GPC3 on the protease site. A portion of the protein molecules are cleaved, but the two fragments remain attached with disulphide bonds. The reducing conditions in the gel separate the two fragments. (B) SEC profile of the Nb49 purification and the SDS-PAGE gel of the peak fractions.

4.2.1 Initial crystals available

Before starting this work, initial small crystals of the Nb49-hGPC3^{core} complex had already been obtained in previous crystallisation trials prepared by Maria Carrasquero-Ordaz (Figure 4.2A). Crystals from condition #1 (Table 4.1) were harvested and sent for data collection. Most of these crystals diffracted at ~6.8 Å resolution except for one which diffracted at ~4.2 Å resolution (dataset #1). When these initial crystallisation plates were rechecked on a later date, a new crystal had grown in an optimisation plate of condition #1 supplemented with the silver bullet additive screen for optimisation (Figure 4.2B, Table 4.1). The crystal from this

condition #2 diffracted at ~4 Å resolution (Figure 4.2C). The two datasets from crystal #1 and crystal #2 were assessed and compared to decide the most suitable dataset for structure determination in Chapter 4.3.1.

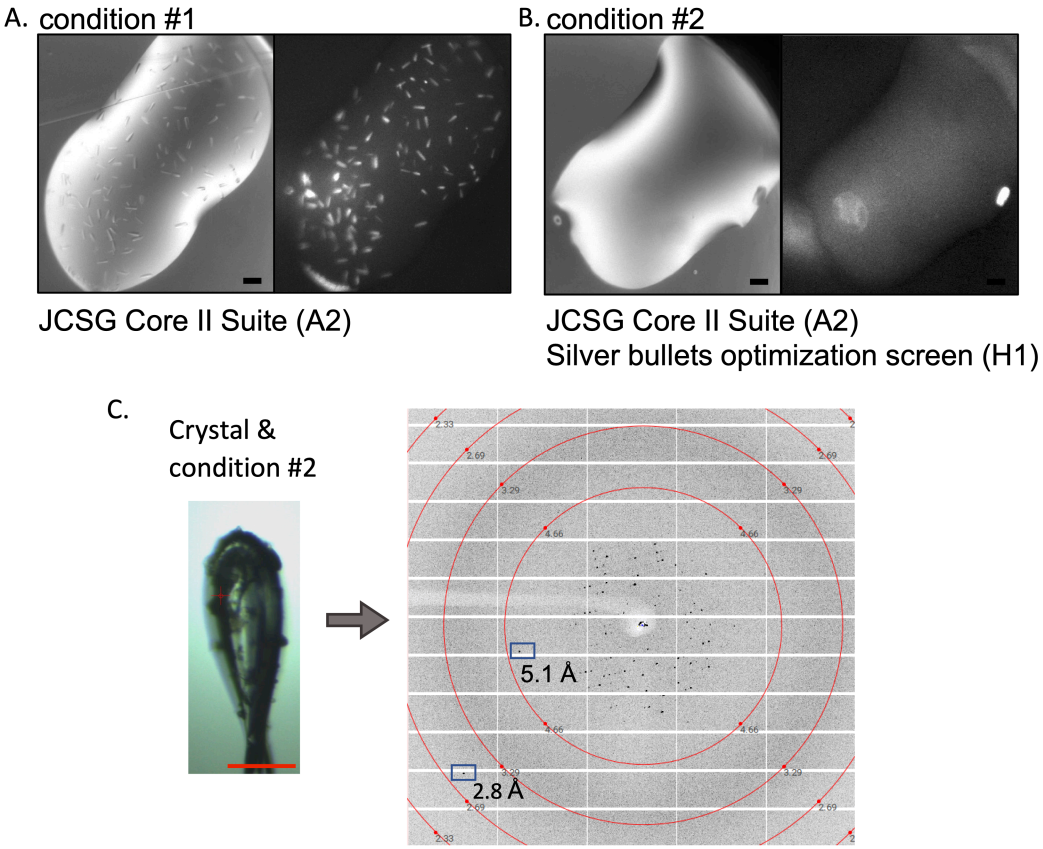


Figure 4.2: Initial crystals of the GPC3-Nb49 complex.
(A-B) Crystals of the hGPC3^{core}-Nb49 complex in their crystallisation drop (200 nl). The concentration of the GPC3-Nb49 complex (1:1 molar ratio) used to prepare these crystallisation plates was 6.42 mg/ml. In both conditions, the crystals were glowing when illuminated with UV light. (C) Crystal #2 mounted on the harvesting loop in the beamline and diffraction pattern resulting from the hGPC3^{core}-Nb49 complex crystal #2. Crystal #2 diffracted at ~4 Å resolution. Scale bar = 100 µm.

Table 4.1: Crystallisation conditions of the initial GPC3-Nb49 crystals.

Condition/ Crystal	Commercial screen	Condition
#1	JCSG Core II (A2)	0.2 M Sodium chloride, 0.1 M CHES pH 7.5, 1.26 M Ammonium sulfate
#2	JCSG Core II (A2) + Silver bullets (H1)	0.2 M Sodium chloride, 0.1 M CHES pH 7.5, 1.26 M Ammonium sulfate 0.2% w/v 1.3-Propanediol, 0.2% w/v D-(-)-3-Phosphoglyceric acid disodium salt, 0.2% w/v Gly-gly, 0.2% w/v Glycerol phosphate disodium salt hydrate, 0.2% w/v Maleic acid, 0.02 M HEPES sodium pH 6.8

4.2.2 New crystallisation trials of the GPC3-Nb49 complex.

In order to improve the existing crystals and help with solving the GPC3-Nb49 complex structure, further crystal optimisation was attempted. The hGPC3^{core} and the nanobodies were mixed and concentrated before the crystallisation plates were set up. The right concentration of the protein for the crystallisation plates was determined by the percentage of precipitation after preparing a Morpheus crystallisation screening plate. The presence of precipitation in around 50% or above of the Morpheus crystallisation conditions is a good indication that the protein concentration is suitable.

Many different optimisation strategies were tried, either by optimising conditions that were already promising or by trying different approaches from the very first screening. These attempts include the microseeding approach in optimisation plates or commercial screens, the addition of NDSB-256 during concentration to avoid precipitation, the different ratios in the drop (1:1 and 2:1 of protein:reservoir) and the mixture of the proteins in different nanobody and hGPC3^{core} molar ratio (1:1 and 1.2:1 of nanobody:GPC3). Crystallisation trials of Nb49 in complex with hGPC3^{core} and rUnc5D^{Ig1Ig2TSP1} were also attempted in order to help understand the way that Nb49 enhance the interaction. However, no crystals were obtained.

A serial dilution plate of condition #1 was prepared at a higher protein concentration. For this plate, condition #1 was diluted with purified H_2O, and new reservoir conditions from 100% to 78% (2% dilution step) of the original condition were prepared. However, only small lookalike crystals grew in the 92% dilution of condition #1 (Figure 4.3A). Because of their size (5 - 15µm), these crystals were suitable for screening at the VMXm beamline in the Diamond Light Source, where

they were placed and plunge freeze on a grid. No protein diffraction was detected from these crystals screening (Figure 4.3).

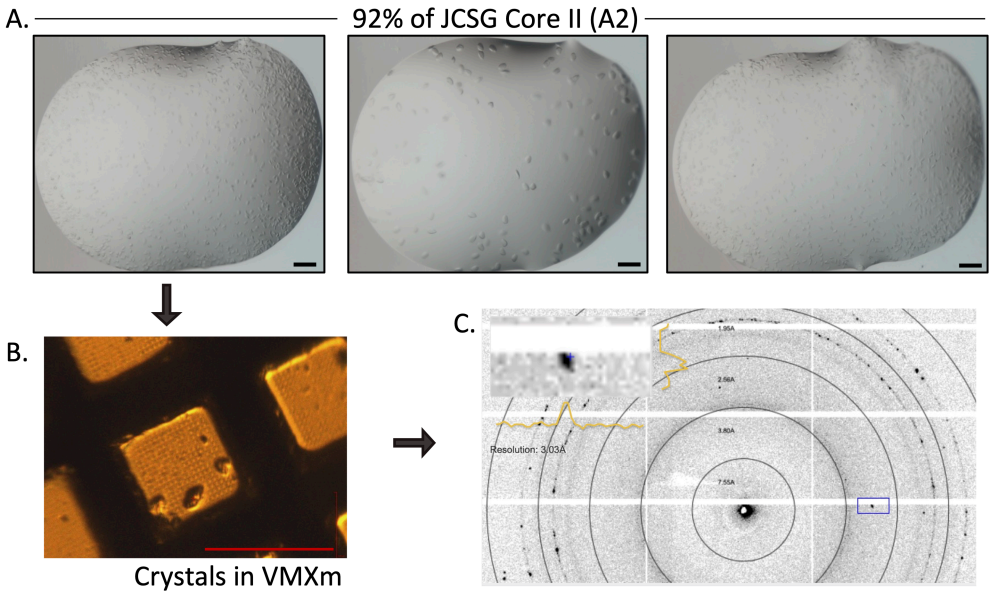


Figure 4.3: Small crystals on VMXm.
(A) Crystals < 15 µm that grew in a few wells in a condition #1 serial dilution optimisation plate. The concentration of the GPC3-Nb49 complex (1:1 molar ratio) used to prepare these crystallisation plates was 7.69 mg/ml. These crystals were glowing when illuminated with UV light. (B) The crystals in VMXm beamline on a grid after plunge freezing in liquid ethane. (C) Diffraction pattern resulting from a GPC3-Nb49 crystal in VMXm. Only two to three spots are present at around 3 Å, which does not correspond to a protein crystal. Scale bar = 100µm.

Editing the nanobody sequence was also attempted in order to make the nanobody less flexible. All the nanobodies were originally cloned into the pADL-23c vector with a C-terminal 6 × histidine tag followed by a myc tag. Additionally, a short linker peptide was present between the core nanobody sequence and the C-terminal tags originating from the heavy chain of the llama antibody. This extra linker sequence and the myc tag were not necessary either for the nanobody's function or for the purification step. Therefore, they were deleted from the sequence in order to avoid the presence of a long flexible region ($Nb49^{short}$) (Figure 4.4A). The $Nb49^{short}$ was expressed and purified, and its binding to $hGPC3^{core}$ was confirmed

(Figure 4.4B, C). Available commercial screens in the X-ray crystallography facility of the Department of Biochemistry were used for the crystallisation trials of the hGPC3^{core}-Nb49^{short} complex. However, crystallisation trials with this nanobody-GPC3 complex did not produce any crystals.

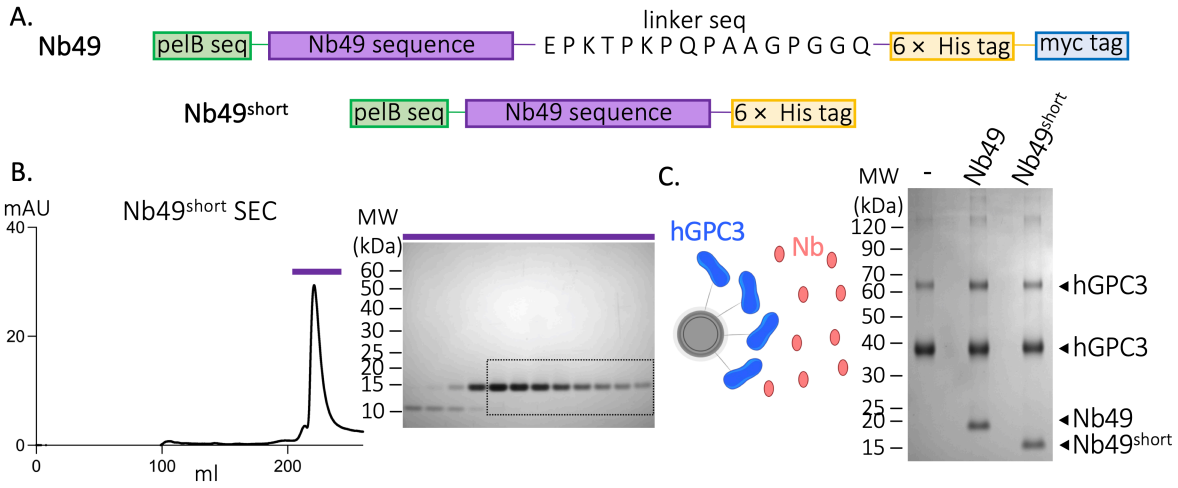


Figure 4.4: Nb49^{short} was used in crystallisation trials.
(A) Schematic representation of the sequence difference between Nb49 and Nb49^{short}. pelB seq: leader sequence for the direction of the nanobodies in the bacterial periplasm. (B) The SEC profile of the Nb49^{short} (14.89 kDa) purification after expression in WK6 *E. coli* and the SDS-PAGE gel of the peak fractions. The fractions in the dotted line box were pooled for the crystallisation trials. (C) Pull-down assay confirms the binding of Nb49^{short} to hGPC3^{core}. Nb49 and Nb49^{short} were incubated with hGPC3^{core} coated beads, and both nanobodies were bound.

4.2.3 Crystallisation trials of the Nb58-GPC3 complex

Crystallisation trials were attempted for the Nb58-GPC3 complex. For this complex, a variety of screens and optimisation steps were utilised. All the available commercial screens were used. Also, different concentrations of the protein complex, addition of NDSB-256 during concentration to avoid precipitation, different nanobody and hGPC3^{core} molar ratios (1:1 and 1.2:1 of nanobody:GPC3), different drop ratios (1:1 and 2:1 of protein:reservoir), and different protein constructs

(hGPC3^{core} and mGPC3^{core}, Nb58 and Nb58^{short}) were tried. A few potential protein crystals grew, but they did not diffract (Figure 4.5).

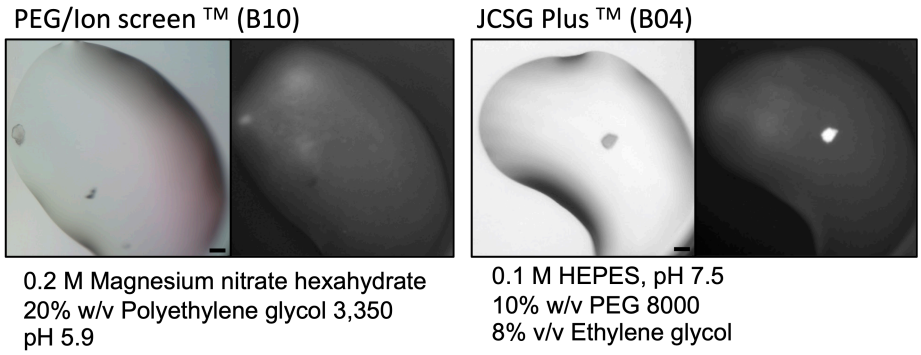


Figure 4.5: GPC3-Nb58 crystals that did not diffract.
In both crystals, the hGPC3^{core} and Nb58 were mixed in a 1 to 1.2 molar ratio and concentrated at 6.39 mg/ml. The crystals grew in a 2 to 1 protein-to-reservoir drop ratio. These crystals were glowing when illuminated with UV light. Scale bar = 100µm.

4.3 Structure determination of the GPC3-Nb49 complex

4.3.1 Data processing

All the crystals of the GPC3-nanobody complexes were screened at the Diamond Light Source synchrotron. Since the screening and data collection for this project took place during the pandemic and the 2021 lockdown, they were done by Dr Edward Lowe. The hGPC3^{core}-Nb49 complex crystals were screened at the beamline i24. Crystal #1 diffracted at ~4.2 Å and crystal #2 at ~4 Å resolution based on the statistics of the automated data processing with the pipeline Xia2-Dials right after data collection at Diamond. Both datasets were auto-processed in the space group $P2_122_1$.

The resolution cut-off can be determined by different parameters. In this case, the correlation coefficient between random half datasets ($CC_{1/2}$) was used, and it is recommended to be above 30% at the highest resolution. For crystal #2, the

automated data processing right after data collection resulted in a dataset with the highest resolution at 4.06 Å, but the $CC_{1/2}$ at this resolution was 20%. The data were processed again using Xia2-Dials [189,190], and the resolution cut-off was set manually to 4.2 Å resolution, which resulted in a dataset with the highest resolution of 4.21 Å and $CC_{1/2}$ at 37.4%. For crystal #1, the automated data processing gave a dataset at 4.2 Å resolution, and the $CC_{1/2}$ at this resolution was 33%.

A comparison of the statistics after data processing with Xia2-Dials of the two datasets is shown in Table 4.2.

Table 4.2: Data collection statistics comparison of dataset #1 and dataset #2

Data collection		
	Dataset #1	**Dataset #2**
Source	I24 (DLS)	I24 (DLS)
Wavelength λ (Å)	0.9999	0.9999
Resolution range (Å)	179.43 – 4.20 (4.27 – 4.20)	178.52 – 4.21 (4.29 – 4.21)
Space group	$P2_122_1$	$P2_122_1$
Cell dimensions		
a, b, c (Å)	82.2803, 117.195, 179.432	81.637, 117.129, 178.518
α, β, γ (°)	90, 90, 90	90, 90, 90
Unique reflections	13263 (615)	12759 (426)
Multiplicity	6.2 (6.3)	6.3 (6.4)
Completeness (%)	100 (97.77)	100 (66.56)
Wilson B ($Å^2$)	113.68	136.94
R_{merge} (%)	48.6 (360.8)	31.3 (192.2)
R_{meas} (%)	53.2 (393.8)	34.3 (209.5)
R_{pim} (%)	21.3 (155.9)	13.8 (826)
$CC_{1/2}$ (%)	96.2 (32.6)	98.4 (37.4)
Mean I/σ(I)	2.8 (0.4)	3.4 (0.3)

* The values in the parenthesis are for the high resolution shell.
DLS: Diamond Light Source, R_{merge}: merging R-factor, R_{meas}: multiplicity-corrected merging R-factor, R_{pim}: Precision-indicating merging R factor, $CC_{1/2}$: Correlation Coefficient between random half datasets

Based on the lower R-values, dataset #2 appeared to be better, and it was used for the following steps of solving the structure and refinement. R_{meas} was used instead of R_{merge} because it takes multiplicity into consideration, and it is more accurate [204]. The lower completeness in the high-resolution shell of dataset #2

was noted, but it should not affect the downstream processes since it is low only in a very narrow resolution shell (4.29 Å - 4.21 Å).

4.3.2 Molecular Replacement using human GPC3 structure (7ZAW) as the search model

The structure of the human GPC3 core protein (PDB: 7ZAW) was solved in the lab by Dr Onno Akkermans [30]. Since the structure of the $hGPC3^{core}$ was available, molecular replacement (MR) was used for phasing.

Firstly, the Matthews coefficient was used to estimate how many copies of the protein molecules there are in the asymmetric unit (ASU) [205,206]. This method gives a probability of 41% of having two copies of $hGPC3^{core}$ and two Nb49 in the ASU and a 57% probability of having three copies of $hGPC3^{core}$ and three Nb49 in the ASU. The solvent content in these two possibilities is 59% and 38%, respectively. The 38% solvent content is too low for a crystal at this relatively low resolution which suggests that it is more probable for the ASU to contain two $hGPC3^{core}$ molecules and two Nb49 molecules.

The $hGPC3^{core}$ model (PDB: 7ZAW) was used initially together with an AlphaFold predicted Nb49 model as search models for MR in the GPC3-Nb49 dataset #2, but it was not possible for PHASER [194] to find a confident and unique solution. When only the $hGPC3^{core}$ model was used as a search model, a confident, unique solution was identified with TFZ > 8 and LLG > 400 and the two GPC3 molecules in the asymmetric unit were placed clearly in the density (Figure 4.6B, C). In PHASER, the “all in point group” search setting was used, and the solution agreed on the $P2_122_1$ space group, which was assigned at the data processing with Xia2-Dials.

Even though PHASER was not able to locate the nanobody, an extra density was present next to one of the two $hGPC3^{core}$ molecules, indicating possibly the

presence of the nanobody there, but it was not clear. 150 cycles of jelly body refinement in Refmac followed, the extra density was improved and could potentially correspond to the nanobody (Figure 4.6D, E). However, it was still not possible to accurately place the nanobody.

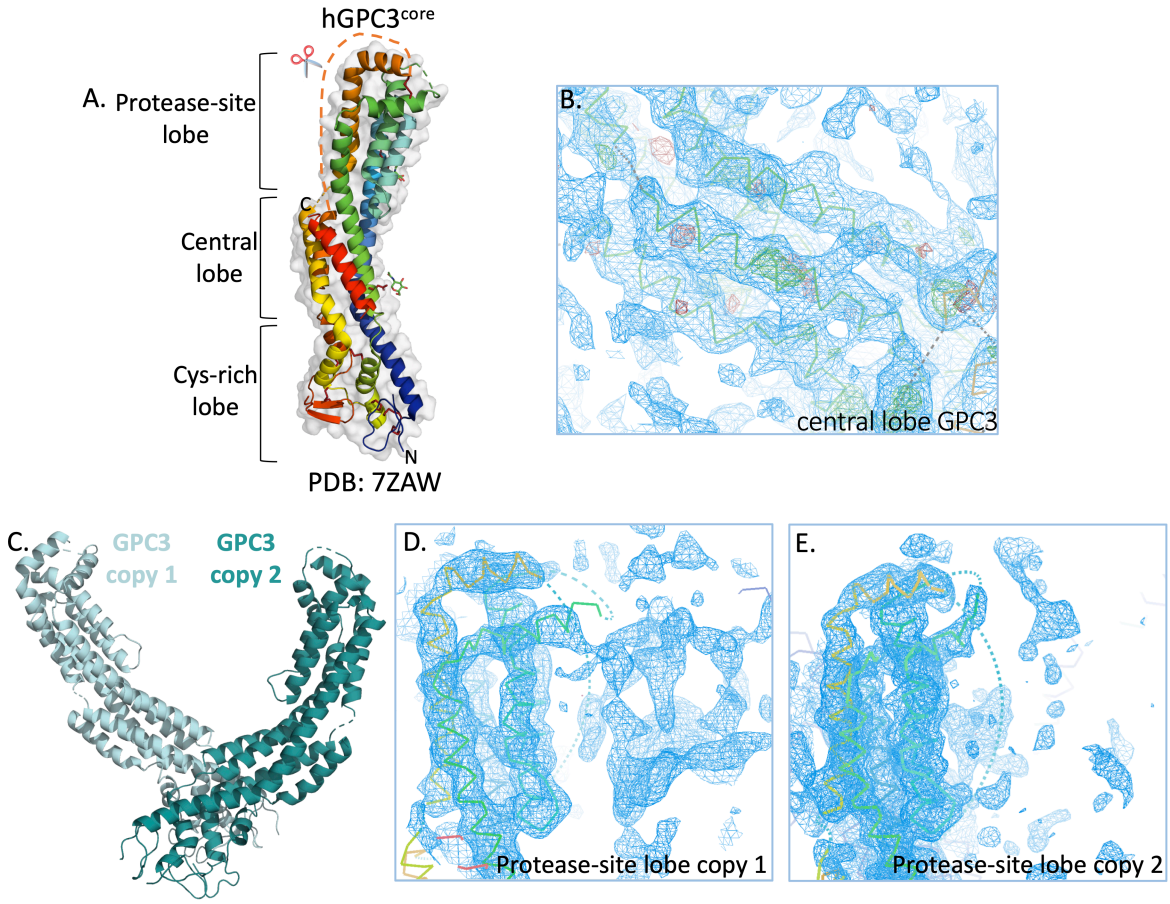


Figure 4.6: MR using the hGPC3^{core} crystal structure model.
(A) The hGPC3^{core} crystal structure model (PDB: 7ZAW) was used as a search model in MR with the dataset #2 of the hGPC3^{core}-Nb49 complex. (B) An example screenshot of the α-helices of the central lobe of GPC3 fitted into the electron density map after MR. (C) The two hGPC3^{core} molecules in the ASU after MR in dataset #2. (D-E) Extra density is present next to the protease-site lobe of one the two GPC3 copies in the ASU but this extra density is not present next to the second copy.

4.3.3 AlphaFold multimer prediction of the GPC3-Nb49 complex

In the crystallisation optimisation trials for the GPC3-Nb49 complex, no better crystals were obtained. For this reason, efforts to utilise dataset #2 and improve the model were made.

To date, the AlphaFold Protein Structure Database (AlphaFold DB) contains over 200 million protein structure predictions [117,118]. An important characteristic of the predicted structure output is the per-residue confidence metric called pLDDT (predicted Local Distance Difference Test), which indicates how confident the prediction is. A prediction of the hGPC3 is available in the AlphaFold DB. The comparison with the experimentally determined crystal structure of the hGPC3^{core} (PDB: 7ZAW) revealed two very similar models (Figure 4.7). The α-helices, the main characteristic of the GPC3 structure, were present and almost identical. The loops, that were not present in the crystal structure because they were disordered and could not be modelled, were present in the AlphaFold model but with low confidence (red).

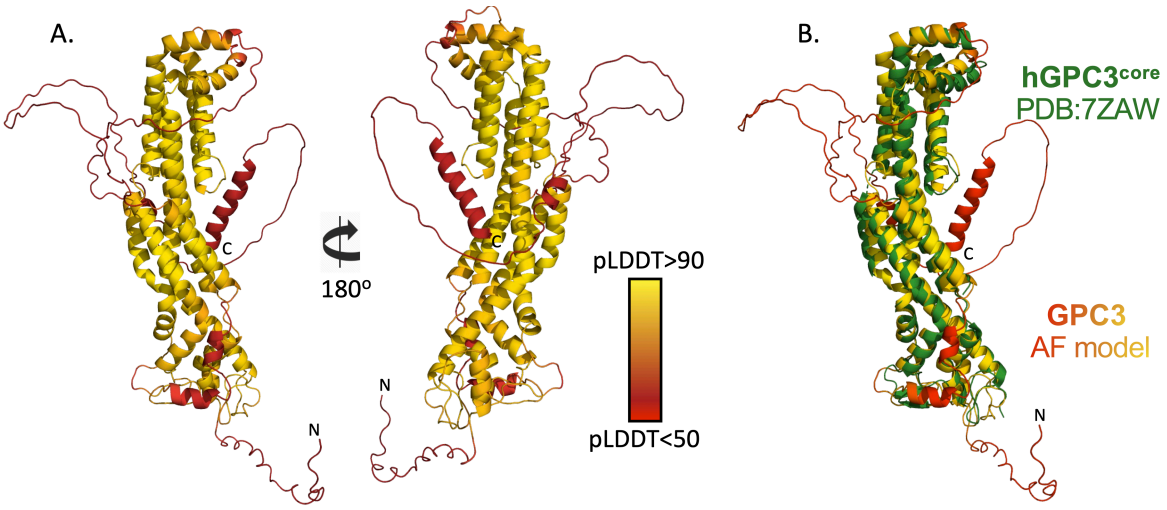


Figure 4.7: AlphaFold model of human GPC3.
(A) AlphaFold (AF) model of human GPC3 from the AlphaFold DB. The molecule is coloured based on the confidence metric pLDDT (Scale: 0 - 100). The red areas have the lowest confidence. (B) Comparison of the AlphaFold human GPC3 molecule with the human GPC3^{core} crystal structure. The loops and areas coloured in red do not exist in the crystal structure model. These areas are flexible, and they are not visible in the crystal structure. At the same time, the prediction of them by AlphaFold was not confident. The big α-helix on the right is the low-confidence model prediction of the C-terminal end of GPC3, which was not part of the crystal structure molecule.

Since AlphaFold can predict the GPC3 structure confidently and the prediction is very similar to the experimentally solved crystal structure, AlphaFold-Multimer [193] was used to predict the binding site of the Nb49 to GPC3. The five AlphaFold multimer predicted structures are ranked based on a model confidence score which makes the ranking based on how confident the binding interface prediction is. The table in Figure 4.8B presents the predicted models and their ranking. Models 2 and 5 are the two models with the highest confidence, they align almost perfectly and predict the binding site of Nb49 on the protease-site lobe of GPC3 (Figure 4.8C). The rest of the models are not very well predicted based on their score, but they all predict the Nb49 binding site very close to the one with the highest score (Figure 4.8D).

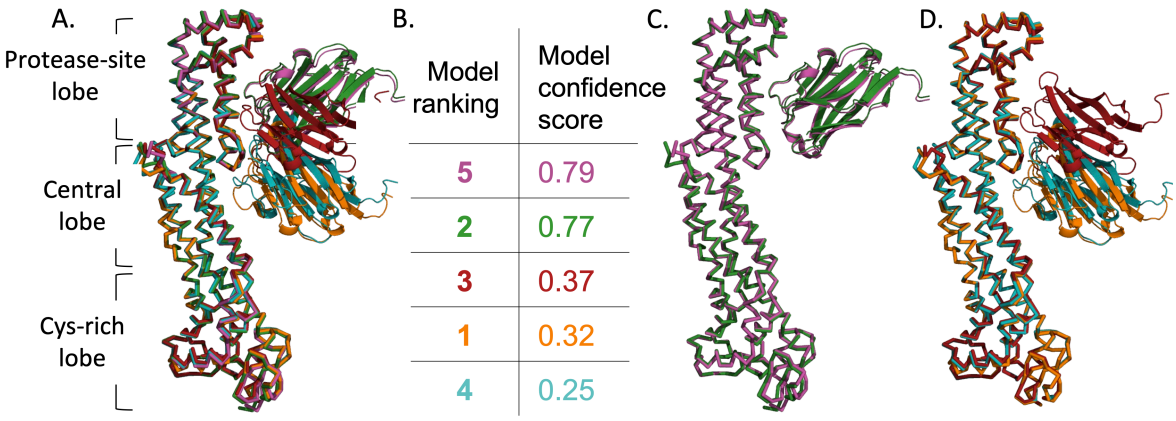


Figure 4.8: AlphaFold Multimer GPC3-Nb49 complex prediction models.
(A) All the AlphaFold multimer predictions of the GPC3-Nb49 complex aligned. (B) Table with the ranking of the predicted by AlphaFold multimer models of the GPC3-Nb49 based on their model confidence score (Scale: 0 - 1). (C) The models with the highest score (5 and 2) align very well. (D) The rest of the predicted models (3, 1, and 4) do not align with the other two with the highest score, but the binding site prediction is close.

Interestingly, the AlphaFold predicted binding location of the nanobody on the GPC3 molecule aligns with the extra density observed in the electron density map

after MR on dataset #2 (Figure 4.9). Therefore, the AlphaFold predicted complex model with the highest score (model 5) was used as a search model for MR.

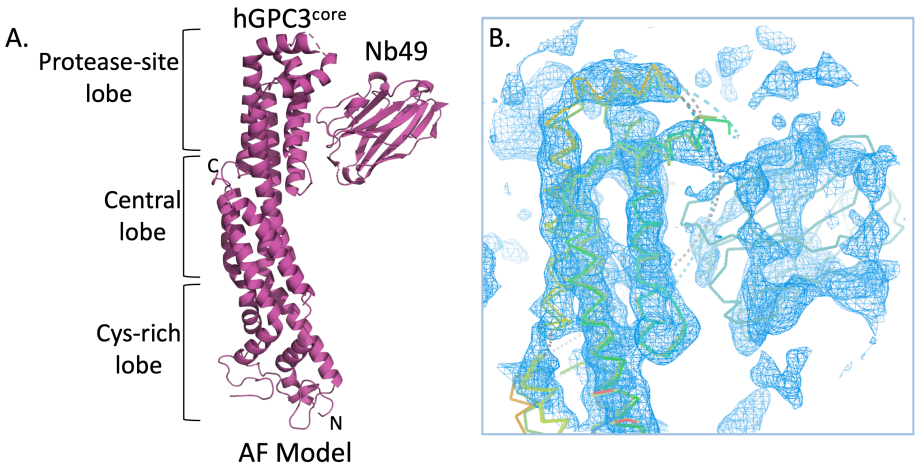


Figure 4.9: Alphafold predicted model of the GPC3-Nb49 complex.
(A) The AlphaFold (AF) multimer predicted model of the GPC3-Nb49 complex with the highest score, which will be used for MR. (B) Alignment of the AlphaFold predicted model with the hGPC3^{core} molecule of the initial MR shows that the extra density next to the protease-site lobe of GPC3 is located where the nanobody binds in the AlphaFold predicted structure.

4.3.4 Molecular Replacement using the AlphaFold Nb49-GPC3 complex structure as a search model

The AlphaFold-predicted hGPC3^{core}-Nb49 model with the highest confidence score was prepared for MR. The PDB file of an AlphaFold predicted model contains the pLDDT per-residue confidence metric instead of the B-factor in the B-factor column. Therefore, to use the model in MR, the pLDDT confidence metric has to be converted into a “B-factor”. The B-factor and the pLLDT confidence metric have an inverse relationship. When the confidence is high, the B-factor will be low. This will make it difficult for PHASER to determine the most confident part of the model, and MR will not work as great. The CCP4i2 task “Process Predicted Models” was used

to change the pLDDT values to B-factors, and the low confidence residues (pLDDT < 70%) were also removed.

The processed model was used as search model for MR in dataset #2, which once again did not result in a confident, unique solution. At the initial MR in chapter 4.3.2, an extra electron density was present next to only one of the two GPC3 molecules in the ASU and not the other one. Considering this, a model of GPC3 in complex with Nb49 and a model of GPC3 without the nanobody were used as search models for the next MR. From this MR, two solutions were identified with TFZ > 8 and LLG > 170. The two solutions were almost identical. The only difference was the slight shift of the GPC3 molecule, which does not have a nanobody bound. The solution with the best TFZ and LLG scores and the lowest clash score (clash score: 2.00 versus 6.00) was examined in Coot [196], and since it was a reasonable solution, it was used for further refinement and model building.

4.3.5 Refinement

Right after MR, a model of GPC3 in complex with Nb49 and a model of only GPC3 were fitted in the density map. However, one part of the backbone of the GPC3 molecule, the one without the nanobody, does not fit perfectly into the low-resolution density. A step of 150 cycles of jelly body refinement in Refmac [195] fixed this issue and placed the GPC3 molecule in its electron density. The refinement process continued with 20-cycle jelly body refinement steps and the use of the high-resolution GPC3 structure (PDB: 7ZAW) as a reference model. Manual model building and refinement were conducted in Coot. In Table 4.3 the statistics of data processing and refinement are presented. Given the resolution and the data quality, further refinement will not be beneficial.

Table 4.3: Data collection and refinement statistics for GPC3-Nb49 complex.

Data collection	
Source	I24 (DLS)
Wavelength λ (Å)	0.9999
Resolution range (Å)	178.52 – 4.21 (4.29 – 4.21)
Space group	$P2_122_1$
Cell dimensions	
a, b, c (Å)	81.637, 117.129, 178.518
α, β, γ (°)	90, 90, 90
Unique reflections	12759 (426)
Multiplicity	6.3 (6.4)
Completeness (%)	100 (66.56)
Wilson B ($Å^2$)	136.94
R_{merge} (%)	31.3 (192.2)
R_{meas} (%)	34.3 (209.5)
R_{pim} (%)	13.8 (826)
$CC_{1/2}$ (%)	98.4 (37.4)
Mean I/σ(I)	3.4 (0.3)
Refinement	
Resolution range (Å)	117.40 – 4.21
Reflections (work / free set)	12726 / 643
R_{work} / R_{free} (%)	29.0 / 33.0
Number of atoms (protein)	
chain A ($hGPC3^{core}$)	6061
chain B ($hGPC3^{core}$)	5977
chain C (Nb49)	1893
Mean B value ($Å^2$) (protein)	
chain A ($hGPC3^{core}$)	186.1
chain B ($hGPC3^{core}$)	180.2
chain C (Nb49)	228.3
RMSD bonds (Å)	0.0097
RMSD angles (°)	1.776
Ramachandran	
Favoured (%)	98.12
Allowed (%)	1.88
Outliers (%)	0
Molprobity score /percentile	1.49 / 100^{th}

* The values in the parenthesis are for the high resolution shell.
DLS: Diamond Light Source, R_{merge}: merging R-factor, R_{meas}: multiplicity-corrected merging R-factor, R_{pim}: Precision-indicating merging R factor, $CC_{1/2}$: Correlation Coefficient between random half datasets, RMSD: Root-Mean-Square Deviation from ideal geometry.

The electron density around the nanobody molecule is not ideal, but it is sufficient to suggest that the nanobody is present and its binding site is on the protease-site

lobe of GPC3 (Figure 4.10A, C). On the second GPC3 copy, there is no extra electron density to suggest the presence of the second nanobody. In both GPC3 chains, a small extra density is present next to N241, confirming the N-glycosylation presence on this residue (Figure 4.10D). The crystal structure of GPC3 (PDB: 7ZAW) has a NAG (N-acetyl-D-glucosamine) in two N-glycosylation sites (N241 and N124), but in this structure, extra density is only present on N241. The two GPC3 copies in the ASU appear to be different from each other. Indeed, the structural alignment of the two GPC3 presents a shift of the GPC3 protease lobe when the nanobody is present (Figure 4.10B).

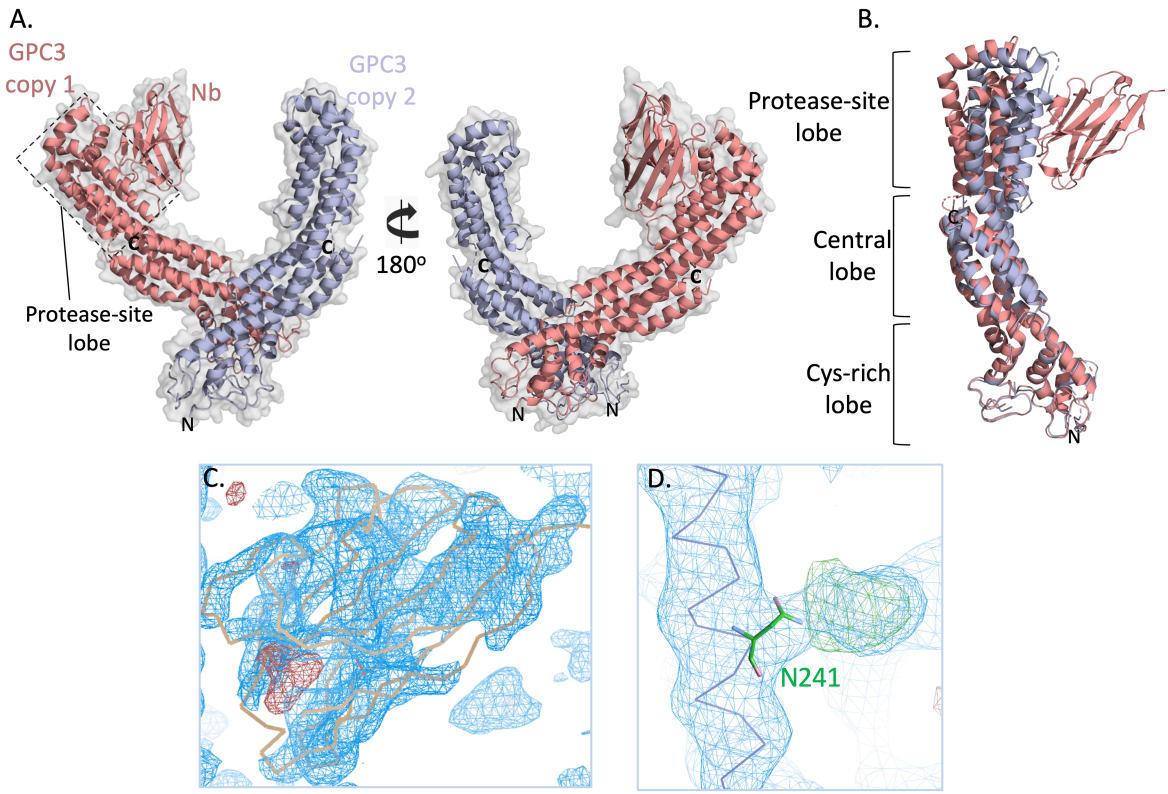


Figure 4.10: The GPC3-Nb49 complex crystal structure.
(A) Two hGPC3^{core} molecules are present in the ASU, and a nanobody is bound to one of the hGPC3^{core}. (B) The two hGPC3^{core} molecules are slightly different to each other. The protease-site lobe of the hGPC3^{core} molecule, which does not have a nanobody bound, presents a more bent shape. (C) The electron density map around the one nanobody model present in the crystal structure after refinement. (D) The extra density present next to the N241 indicates the presence of an N-linked glycan.

4.3.6 Assessing the presence of only one nanobody in the asymmetric unit

The presence of only one nanobody bound to one of the two GPC3 molecules in the ASU needed further investigation. The raw data were reindexed into lower symmetry space groups (*P*1, $P2_1$) to exclude the possibility of the second nanobody being present but vanishing by the imposed symmetry.

Preliminary refined models of the hGPC3^{core}-Nb49 and hGPC3^{core} alone were used as search models for MR in the *P*1 and $P2_1$ indexed data. In the *P*2 point group, the $P2_1$ space group was chosen because after data processing, MR in *P*2 could not find a solution, and PHASER chose $P2_1$. As it is explained by the symmetry and considering that two hGPC3^{core} molecules are present in the ASU in $P2_122_1$, eight GPC3 molecules are present in the asymmetric unit in the *P*1 space group and four molecules in the $P2_1$ space group. MR found confident, unique solutions in both space groups.

The density maps and the models after MR were inspected in Coot. In both cases, as well as in the original $P2_122_1$ space group, the protein molecules pack well in space without abnormal empty spaces (A in Figure 4.11, Figure 4.12, Figure 4.13). The density maps for both solutions present density where the nanobody is but not where the second nanobody should be. Next, two hGPC3^{core} molecules were used as search models in MR without the nanobody in order to check the extra "nanobody" density without bias from the model. In these maps, the extra density is present next to half of the hGPC3^{core} molecules, which agrees with the presence of the nanobody on one of the two hGPC3^{core} copies in the $P2_122_1$ (B, D in Figure 4.11, Figure 4.12, Figure 4.13). As an extra check, density modification of the maps was performed, and a low solvent content (10% lower than predicted) was provided to Parrot to avoid flattening density which is present but not strong. Despite the noisy

map, the distinct extra density of the nanobody was present only in half of the hGPC3^{core} copies (C, E in Figure 4.11, Figure 4.12, Figure 4.13).

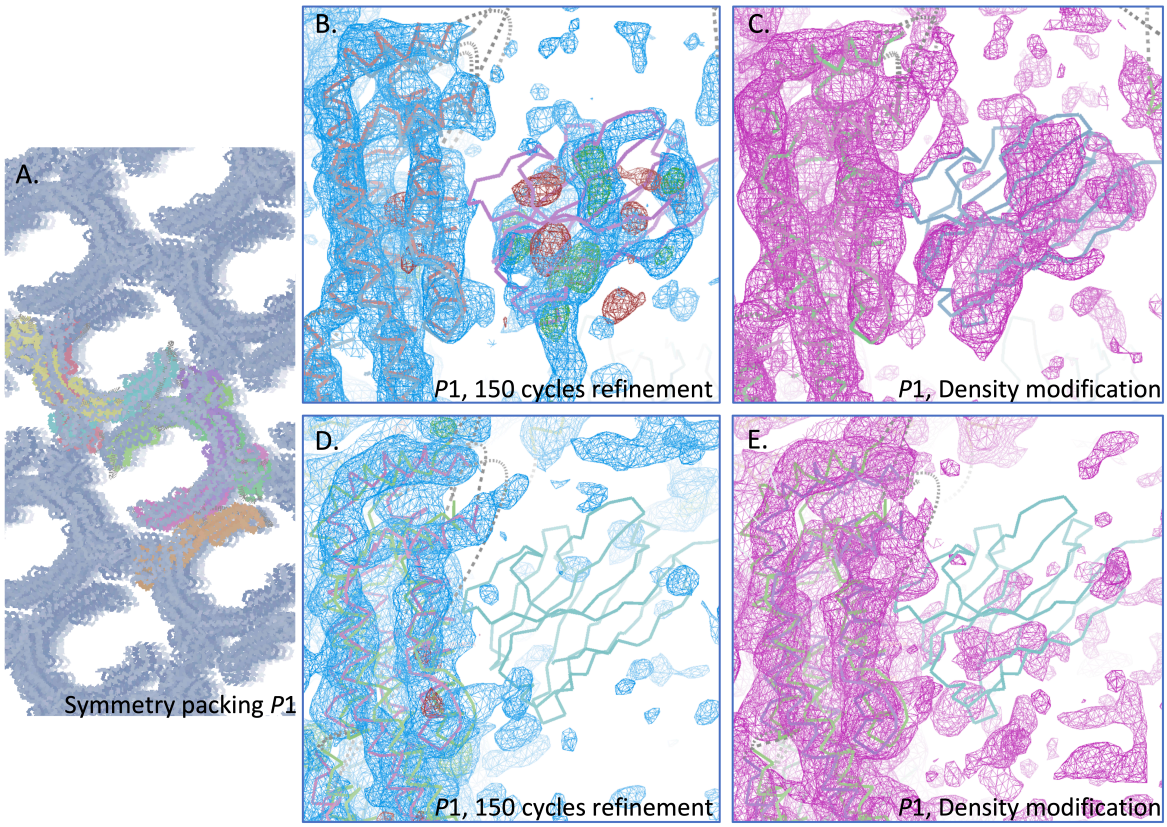


Figure 4.11: Nanobody density in the *P*1 space group.
(A) Packing of the hGPC3^{core} copies in space in *P*1. (B-C) The density of the nanobody next to the protease-site lobe of four out of eight hGPC3^{core} copies in the ASU. (D-E) No density of the nanobody is present next to the protease lobe of the other four copies of hGPC3^{core}. The blue electron density maps are a result of 150 cycles of jelly body refinement in Refmac5 right after MR and the magenta density maps are the result of density modification in Parrot right after MR. A hGPC3^{core}-Nb49 complex model was aligned in order to help visualise where the nanobody should be.

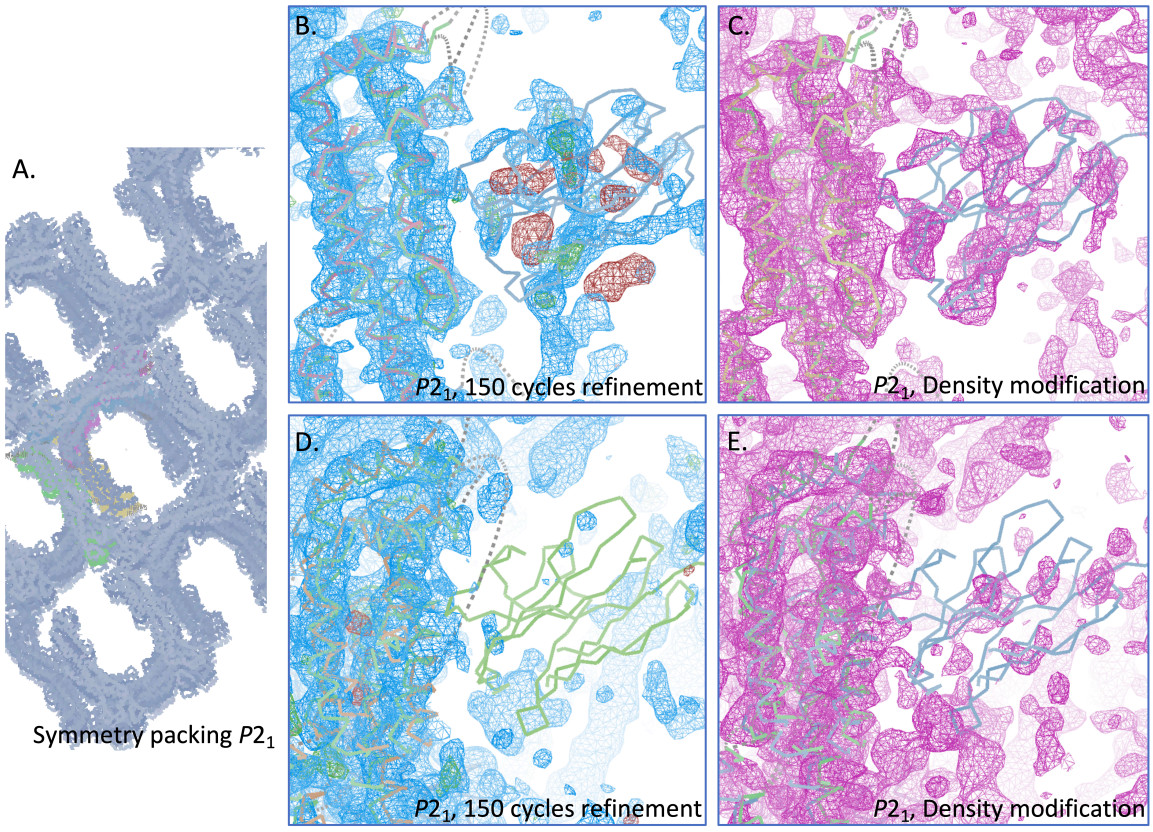


Figure 4.12: Nanobody density in the $P2_1$ space group.
(A) Packing of the hGPC3^{core} copies in space in $P2_1$. (B-C) The density of the nanobody next to the protease-site lobe of two out of four hGPC3^{core} copies in the ASU. (D-E) No density of the nanobody is present next to the protease lobe of the other two copies. of hGPC3^{core}. Blue density map: 150 cycles of jelly body refinement, magenta density map: density modification. A hGPC3^{core}-Nb49 complex model was aligned to help visualise where the nanobody should be.

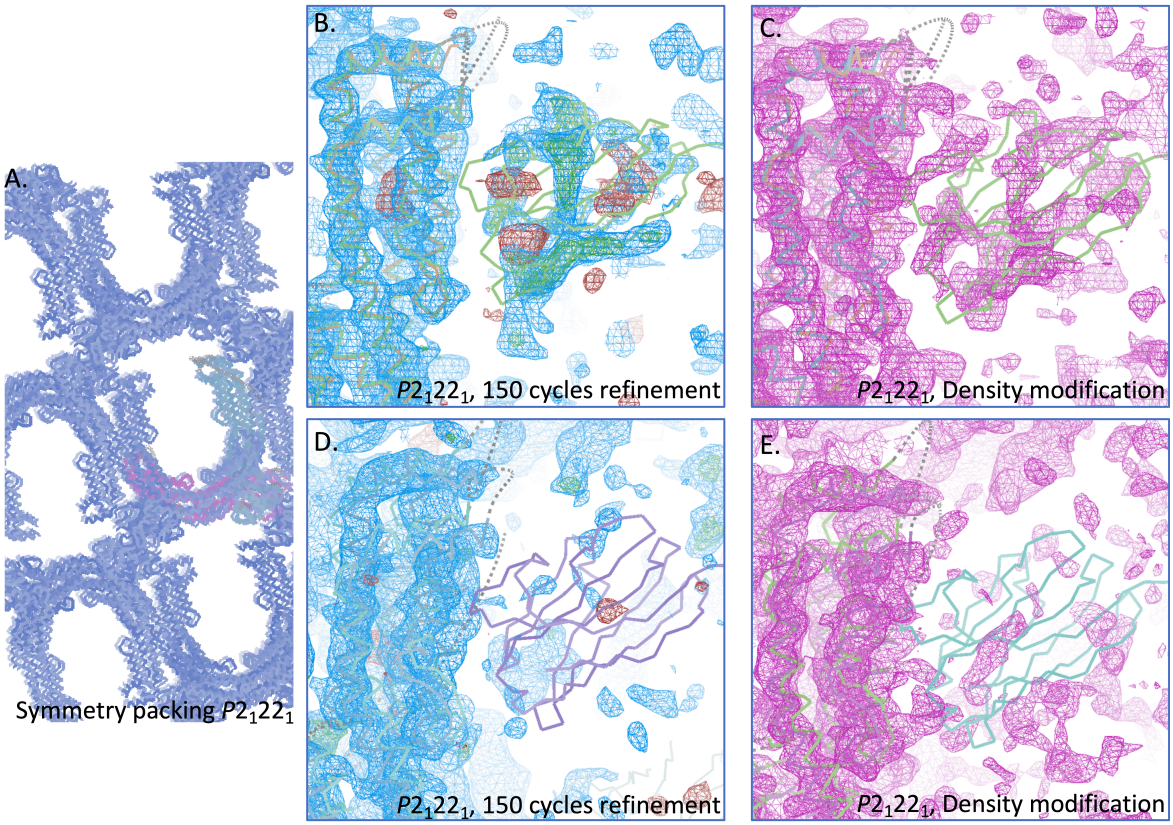


Figure 4.13: Nanobody density in the $P2_122_1$ space group.
(A) Packing of the hGPC3^{core} copies in space in $P2_122_1$. (B-C) The density of the nanobody next to the protease-site lobe of one out of the two hGPC3^{core} copies in the ASU. (D-E) No density of the nanobody is present next to the protease lobe of the other copy of hGPC3^{core}. Blue density map: 150 cycles of jelly body refinement, magenta density map: density modification. A hGPC3^{core}-Nb49 complex was aligned to help visualise where the nanobody should be.

On the other hand, the data processing statistics of the *P*1, $P2_1$ and $P2_122_1$ indexed datasets were compared (Table 4.4).

Table 4.4: Comparison of the data processing statistics in *P*1, $P2_1$ and $P2_122_1$.

	High res (Å)	High res $CC_{1/2}$	Overall R_{meas}	Low res shell (Å)	Low res shell R_{meas}
*P*1	4.34	0.204	0.308	184.45 - 11.79	0.163
$P2_1$	4.23	0.316	0.312	179.12 - 11.48	0.113
$P2_122_1$	4.21	0.343	0.374	178.68 - 11.44	0.115

* High res: High resolution, Low res: Low resolution, $CC_{1/2}$: Correlation Coefficient between random half datasets, R_{meas}: multiplicity-corrected merging R-factor.

The overall R_{meas} is being affected by the high-resolution data and therefore, the comparison of R_{meas} at the low-resolution shell followed. The $P2_1$ and $P2_122_1$ datasets look similar while the R_{meas} of *P*1 is much higher. Given that the maps of the nanobody in $P2_1$ and $P2_122_1$ are almost identical, the $P2_122_1$ space group was chosen for the structure solution and refinement as it was the higher symmetry space group. Moreover, except for the Xia2-Dials autoprocessing, the autoPROC autoprocessing integration software [207] assigned the $P2_122_1$ space group to this data in agreement with Xia2-Dials.

Additionally, dataset #1, which was not initially chosen for the downstream processes of MR and refinement, was used at this stage to examine the space group and the number of nanobodies in the ASU. Autoprocessing with Xia2-Dials determined that the space group of this dataset is $P2_122_1$, the same as dataset #2. After MR, the map and the model were examined manually in Coot. The inspection showed that in this case as well, two GPC3 molecules were present in the ASU and only one nanobody. This is one more piece of evidence that only one of the GPC3 molecules has a nanobody bound in this crystal packing.

4.4 Cryo-EM trials for the GPC3-Unc5-Nb49 complex structure determination

4.4.1 Sample grid screening

Finding that Nb49 enhances the interaction between GPC3 and Unc5 generated the idea to further understand the ternary complex of GPC3-Unc5-Nb49 and identify in detail how Nb49 enhances this interaction. The GPC3-Unc5-Nb49 complex did not form any crystals in crystallisation trials.

Cryo-EM was then employed in order to get the ternary complex structure, and grids with different samples were prepared. Once the samples were ready, the grids were prepared and screened by Dr Jin Zhou. In the first trial, grids were prepared with a sample of a mixture of the hGPC3^{core}, rUnc5D^{Ig1Ig2TSP1} and Nb49 in a 1:1:5 molar ratio. The screening of these grids showed huge aggregation patches in the grid holes. To fix this issue, several steps to improve the grids were taken. Freshly purified proteins were used for the grid preparation to improve the protein quality. Grids with freshly purified protein did not present any improvement in the aggregation level. In all the following cryo-EM trials, prior to grid preparation, the samples were centrifuged for 10 minutes at 21,300 × g to get rid of any aggregation. However, the centrifugation did not help with the aggregation problem on the grids. It was then hypothesised that this behaviour of the proteins on the grid could derive from the contact of the proteins to the liquid-air surface, which led to their denaturation. To tackle this, detergents were added either directly on the sample drop on the grip just before plunge freezing (1% FOM) or in a mixture with the protein sample (0.4% CHAPS, 0.04% FOM), but this approach did not help either. Grids with graphene oxide (GO) support film and no open holes were used as well.

This seemed to improve the appearance of the sample on the grid, but the particles were not clearly visible due to less contrast.

At the same time, many different homologues were used, aiming at a better-behaving sample. Different constructs of the human and mouse GPC3 were available, and it was already known that they could both bind Nb49 and all the Unc5 homologues (A-D). In Table 4.5, all the different construct combinations and experimental variables are presented. The combination of $mGPC3^{ecto}$-$mUnc5A^{ecto}$ behaved the best (Figure 4.14).

Table 4.5: Constructs and optimisation conditions of the cryo-EM trials.

	Protein construct combination	Extra variables
1	$rUnc5D^{Ig1Ig2TSP1}$ + $hGPC3^{core}$ + Nb49	± detergent on the grid
2	$rUnc5D^{Ig1Ig2TSP1}$ + $hGPC3^{core}$	± detergent in the protein sample
3	$rUnc5D^{ecto}$ + $hGPC3^{core}$	Open hole and GO grids
4	$rUnc5D^{ecto}$ + $mGPC3^{ecto}$	Open hole grids
5	$mUnc5B^{ecto}$ + $hGPC3^{core}$	
6	$mUnc5B^{ecto}$ + $mGPC3^{ecto}$	
7	$mUnc5A^{ecto}$ + $hGPC3^{core}$	
8	$mUnc5A^{ecto}$ + $mGPC3^{ecto}$	Open hole and GO grids
9	$mUnc5A^{ecto}$ + $mGPC3^{ecto}$ + Nb49	Open hole and GO grids

The behaviour of the $mGPC3^{ecto}$-$mUnc5A^{ecto}$ sample with open-hole grids was good. These grids are preferable because of the minimal background signal. Initially, the grids with the $mGPC3^{ecto}$-$mUnc5A^{ecto}$-Nb49 sample were slightly overcrowded. Screening several dilutions of the original sample resulted in a grid with acceptable contrast and concentration. However, the sample was not homogeneous enough for data collection (Figure 4.14C).

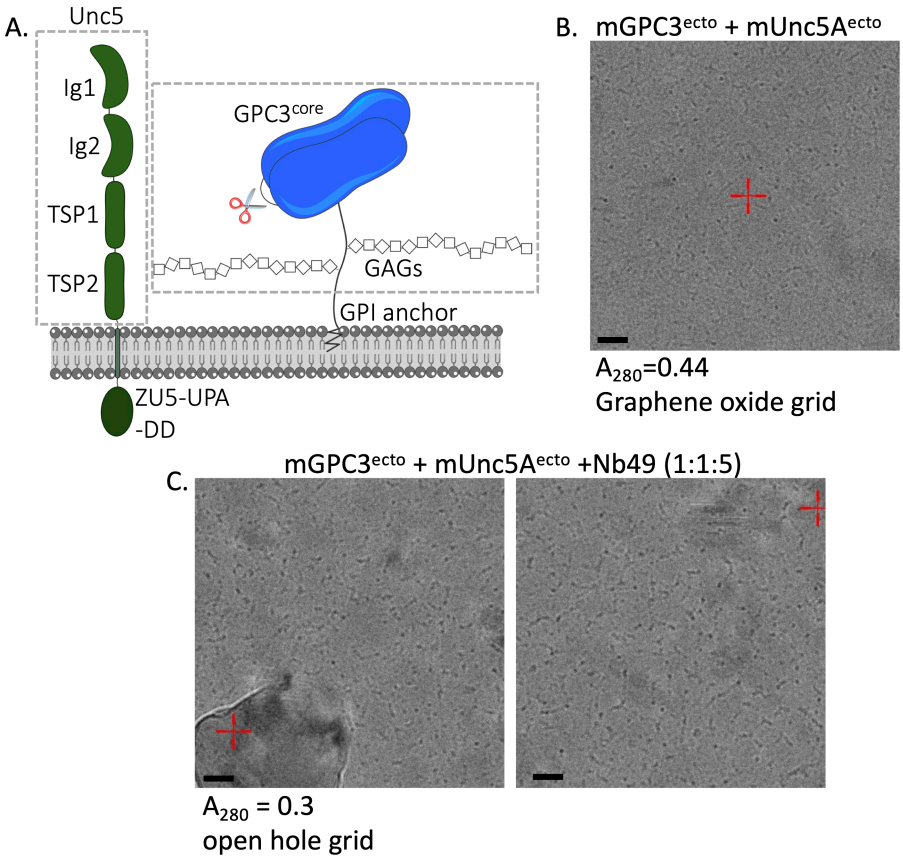


Figure 4.14: Cryo-EM screening for the GPC3-Unc5-Nb49 complex.
(A) Schematic representation of the Unc5 and GPC3 domain architecture. The dotted line boxes indicate the constructs (mGPC3^{ecto}, mUnc5A^{ecto}) that were better behaved and were used in the grids presented here. (B) The mGPC3^{ecto} + mUnc5A^{ecto} sample behaved the best among other GPC3 and Unc5 constructs and homologue combinations. Particles are visible. In this trial, graphene oxide support grids were used. (C) The mGPC3^{ecto} + mUnc5A^{ecto} + Nb49 sample on open hole grids. Particles are present and visible. Scale bar = 20 nm.

4.4.2 Crosslinking trials of the GPC3-Unc5 complex

To improve the complex stability and, therefore, the GPC3-Unc5-Nb49 complex quality for the cryo-EM experiments, crosslinking of the complex was attempted. For the initial crosslinking trials, only the hGPC3^{core} and rUnc5D^{Ig1Ig2TSP1} were used and not the nanobody in order to find the suitable conditions first (Figure 4.15A, B). Bis(sulfosuccinimidyl) suberate (BS3) is an amine-to-amine crosslinker with an 8-carbon spacer arm (11.4 Å), and it was used for these crosslinking trials.

A series of different concentrations of BS3 and different reaction times were tested. First, hGPC3^{core} and rUnc5D^{Ig1Ig2TSP1} were mixed at a one-to-one molar ratio, and BS3 concentrations of 0.5, 3 and 5 mM were tried for 1.5, 2 and 3 hours at room temperature (Figure 4.15C). The aim of these increasing concentrations of BS3 was to see a gradient of crosslinking from no crosslinked molecules to fully crosslinked molecules. In this trial, the sample with 0.5 mM BS3 is less crosslinked than the 5 mM BS3 sample, but neither the 0.5 mM sample represents a minimum crosslinking sample after the 0 mM concentration, nor the 5 mM sample represents a fully crosslinked sample. The incubation reaction time does not seem to make a difference.

In the next trial, the incubation time was not a variable, and the protein concentration was increased (Figure 4.15D). Except for the 0.5, 3 and 5 mM BS3 concentration series, a higher BS3 concentration at 15.38 mM was tried. The higher protein and BS3 concentration increased the crosslinking of the proteins, but it was not complete. The lower molecular weight bands of GPC3 disappeared, whereas the Unc5 band below 60 kDa did not seem to get crosslinked at the same rate.

While crosslinking of the Unc5-GPC3-Nb49 is possibly necessary for the cryo-EM experiment, BS3 may not be the most suitable crosslinker for this sample. Indeed, since BS3 is an amine-to-amine crosslinker, the lysine sidechains are responsible for the link formation between two molecules. Based on the GPC3-Unc5 octameric complex (PDB: 7ZA1), very few lysines are located at the binding interfaces, and the distance between them is longer than the BS3 linker, although the resolution of the structure is not enough to determine where exactly the side chains are. Crosslinking trials with a different crosslinker would be the next logical step.

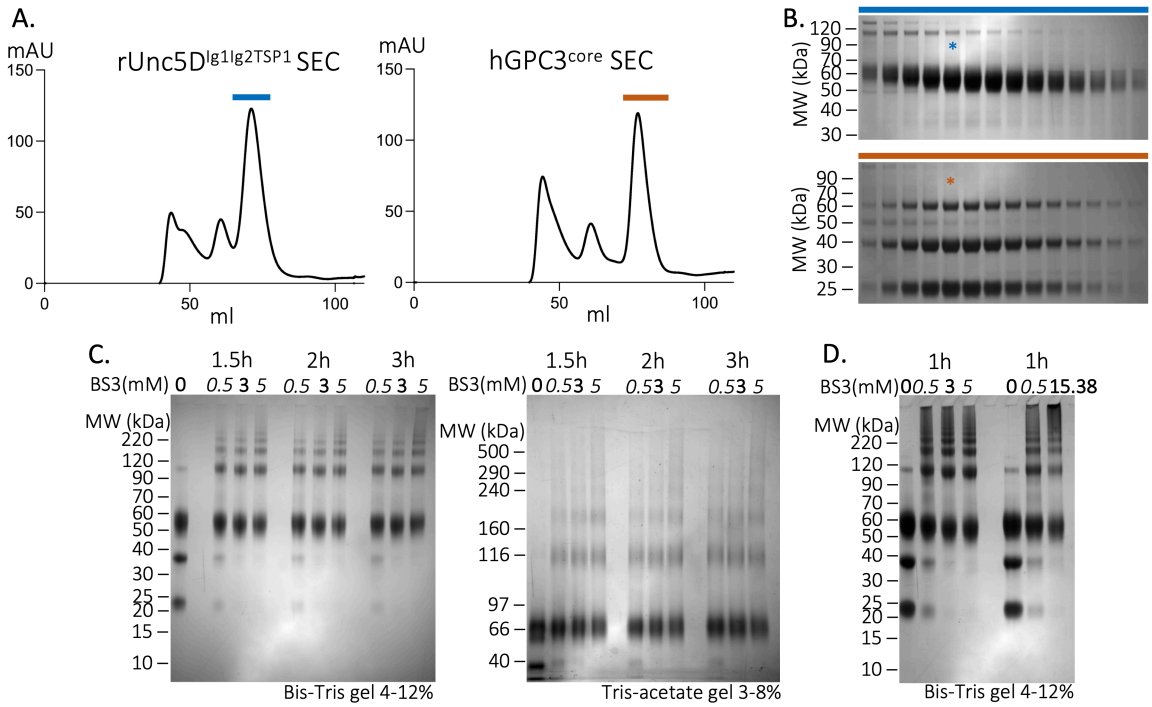


Figure 4.15: Crosslinking trials of the GPC3-Unc5 complex.
(A-B) The SEC profiles and the SDS-PAGE of the purification of hGPC3^{core} and rUnc5D^{Ig1Ig2TSP1} in 20 mM Hepes pH 7.5 and 200 mM NaCl. The fractions highlighted with an asterisk were used for the crosslinking reactions. (C) SDS-PAGE gels of the first crosslinking trial. The samples in both gels are the same. The purpose of the two gels with different polyacrylamide concentration gradients is to separate bands of higher molecular weight. The Tris-acetate 3-8% gel can separate proteins up to 500 kDa. (D) SDS-PAGE gel of the second crosslinking trial. The protein sample in this experiment has a higher concentration.

4.5 Discussion

Despite the difficulty of placing the nanobody in the electron density map initially, the AlphaFold multimer predicted GPC3-Nb49 complex helped. The low-resolution data did not impede the discovery of the nanobody binding epitope on GPC3. However, the exact binding interface and the detailed interactions could not be determined. It is important to note that the AlphaFold multimer predicted model placed the nanobody at the same position as where the nanobody is based on the experimental data and the extra density in the electron density map.

4.5.1 Validation of the Nb49-GPC3 binding interface

Despite the lack of information about the detailed interactions between hGPC3 and Nb49, the crystal structure of the hGPC3^{core}-Nb49 complex revealed the binding site of Nb49 on the GPC3 molecule. Binding studies and mutants can be used in order to confirm this binding site. Four artificial N-glycosylation triplet motifs (N-X-T/S, X ≠ P) were introduced on the binding site in order to therefore introduce an N-glycan, which will potentially block the interaction (Figure 4.16). These N-glycosylation mutants need to be characterised.

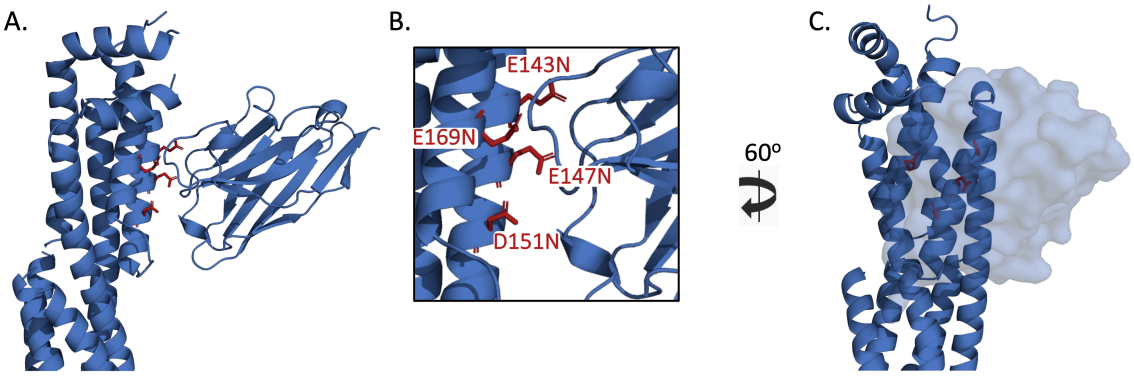


Figure 4.16: N-glycosylation mutations on the binding site.
(A) The four red in sticks representation amino acids will be mutated into asparagine to form the artificial glycosylation triplet motif (N-X-T/S, X ≠ P). (B) Detailed representation of these amino acids. (C) A side view of the complex in which the nanobody is presented in a transparent surface format at the front of the hGPC3^{core} molecule, and the red residues are behind it.

4.5.2 How does Nb49 enhance the GPC3-Unc5 interaction?

In order to explain how Nb49 enhances the GPC3-Unc5 interaction, the model of the hGPC3^{core}-Nb49 complex structure was aligned with each GPC3 molecule in the model of the GPC3-Unc5 octamer. The octameric complex revealed three binding interfaces between GPC3 and Unc5. When Nb49 is bound to GPC3, interface 1 is blocked, but the other two Unc5 molecules can still bind to GPC3 (Figure 4.17).

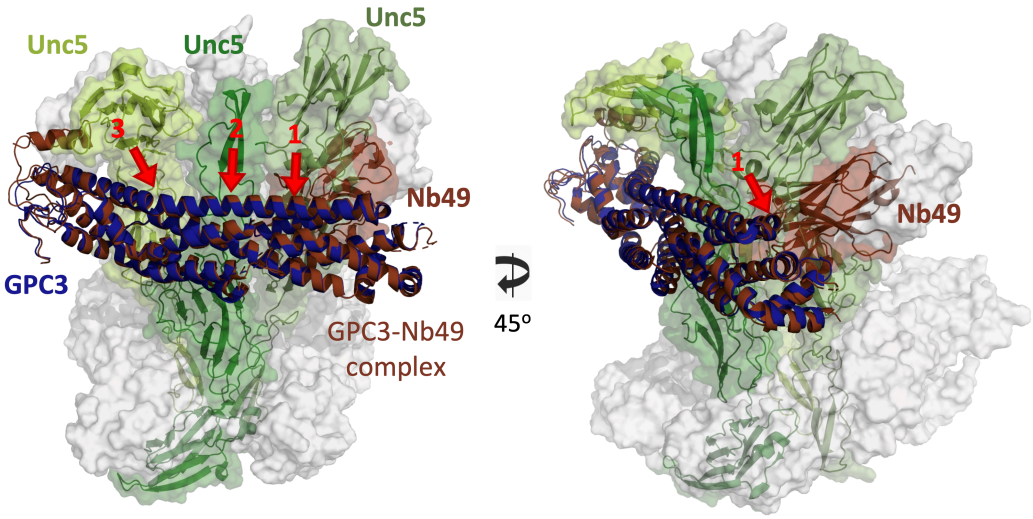


Figure 4.17: The GPC3-Nb49 complex aligned in the octameric GPC3-Unc5 complex.
The hGPC3^{core} molecule of the hGPC3^{core}-Nb49 complex (brown) is aligned to the hGPC3^{core} molecule of the GPC3-Unc5D octameric complex. The three binding interfaces are highlighted with red arrows. Based on this alignment, the binding of Nb49 seems to block the binding interface 1 between GPC3 and Unc5D.

However, the GPC3-Unc5 complex can exist in lower stoichiometry as a 2:2 or 1:1 complex [30]. Therefore, Nb49 can stabilise any of these subcomplexes. Based on the *in vitro* stripe assays presented in Chapter 3, Nb49 enhances the cell repulsion from the GPC3 WT-coated red stripes in line with the pull-downs and the SPR data that shows that Nb49 enhances the Unc5-GPC3 interaction. Since the octameric complex of Unc5 and GPC3 cannot form in the presence of Nb49, these stripe assays data show that the octamer is not needed for the complex to function and lower stoichiometry complexes can have biological relevance.

The interaction might be stabilised by the slight change of the GPC3 structural conformation after the Nb49 binding. More data are needed to better understand how Nb49 enhances the GPC3-Unc5 interaction. Cryo-EM and crosslinking can be further optimised and used to solve the structure of the GPC3-Unc5-Nb49 complex.

5. Generation and characterisation of Unc5 and GPC3 non-binding mutants

5.1 Introduction and summary

The crystal structure of the GPC3-rUnc5D complex revealed an octamer which consists of four rUnc5D^{Ig1Ig2TSP1} chains enclosed between four GPC3^{core} molecules. To confirm and study the interaction between Unc5 and GPC3, non-binding GPC3 and Unc5 mutants were produced and employed in *in vivo* and *in vitro* assays, in addition to the nanobodies, to characterise the interaction. In this chapter, the design and characterisation of the mutants are described. Artificial N-glycosylation sites were designed for each protein at the binding interfaces. Interestingly, a glycan-glycan interaction was discovered in the centre of the octamer. Deleting the GPC3 glycosylation that contributes to this interaction, effectively disrupted the binding. For the initial characterisation of the mutants, cell binding assays were performed. SPR confirmed the mutants' efficacy and more experiments were performed to help confirm the mutants' correct folding.

Disclaimer: The presented project is collaborative, and therefore some aspects of the work were performed by others. Of the work presented in this chapter, I co-designed the mUnc5B and hGPC3 mutants with Dr Onno Akkermans, and I cloned them. I also performed the initial characterisation with cell binding assays. Further cell binding assays and SPR experiments (Chapter 5.4, 5.5) were performed by Miguel Berbeira-Santana and Maria Carrasquero-Ordaz, respectively. I also

performed the CD experiment in the Biophysics facility of the Biochemistry department with the help of Dr David Staunton.

5.2 Mutation design

The mutations in this project introduce an artificial glycosylation triplet motif (N-X-T/S), where X is any amino acid except proline. They aim to disrupt protein-protein interactions by adding an N-glycan to the binding site. This is an established approach of complex disruption through protein mutagenesis when the resolution of the complex structure is not satisfactory to design point mutations [72,77,202,203,208].

5.2.1 Mutation design on Unc5B

Three mutations were designed and cloned for Unc5. The protein constructs that crystallised and revealed the octameric structure were the complex of the human or mouse $GPC3^{core}$ and the $rUnc5D^{Ig1Ig2TSP1}$; therefore, the mutations were designed on the rat Unc5D. After their design, the mutations were cloned in the mouse Unc5B (mUnc5B). The rUnc5D and mUnc5B sequences are highly conserved on the binding interfaces, and it is also known that they both bind GPC3 [30] which supports the fact that it is possible to clone the mutants in the mUnc5B after their design on rUnc5D.

The Unc5 mutants cloned in the $mUnc5B^{full\text{-}length}$ are the following: $mUnc5B^{Y150N\text{-}R152T}$(U6), $mUnc5B^{A277N\text{-}L279T}$(M7), and $mUnc5B^{I211N\text{-}R213T}$(K4) (Figure 5.1) and they all aim at the introduction of the artificial glycosylation site.

The mutants U6 and M7 were previously designed and cloned in the lab on Unc5D. The U6 mutation is located between the domains Ig1 and Ig2, and the M7 mutation is located between the domains Ig2 and TSP1. Both mutations aim to

disrupt the binding interface with GPC3. It is important to note that the artificial triplet motif of the M7 mutant contains a proline residue between the asparagine and threonine residues of the motif. This proline could hinder the glycosylation on the site. The K4 mutation is located on the Ig2 domain of Unc5. The rationale behind the design of this mutation is to prevent the two neighbouring Unc5 chains in the octamer from approaching one another closely enough to form the interaction with GPC3.

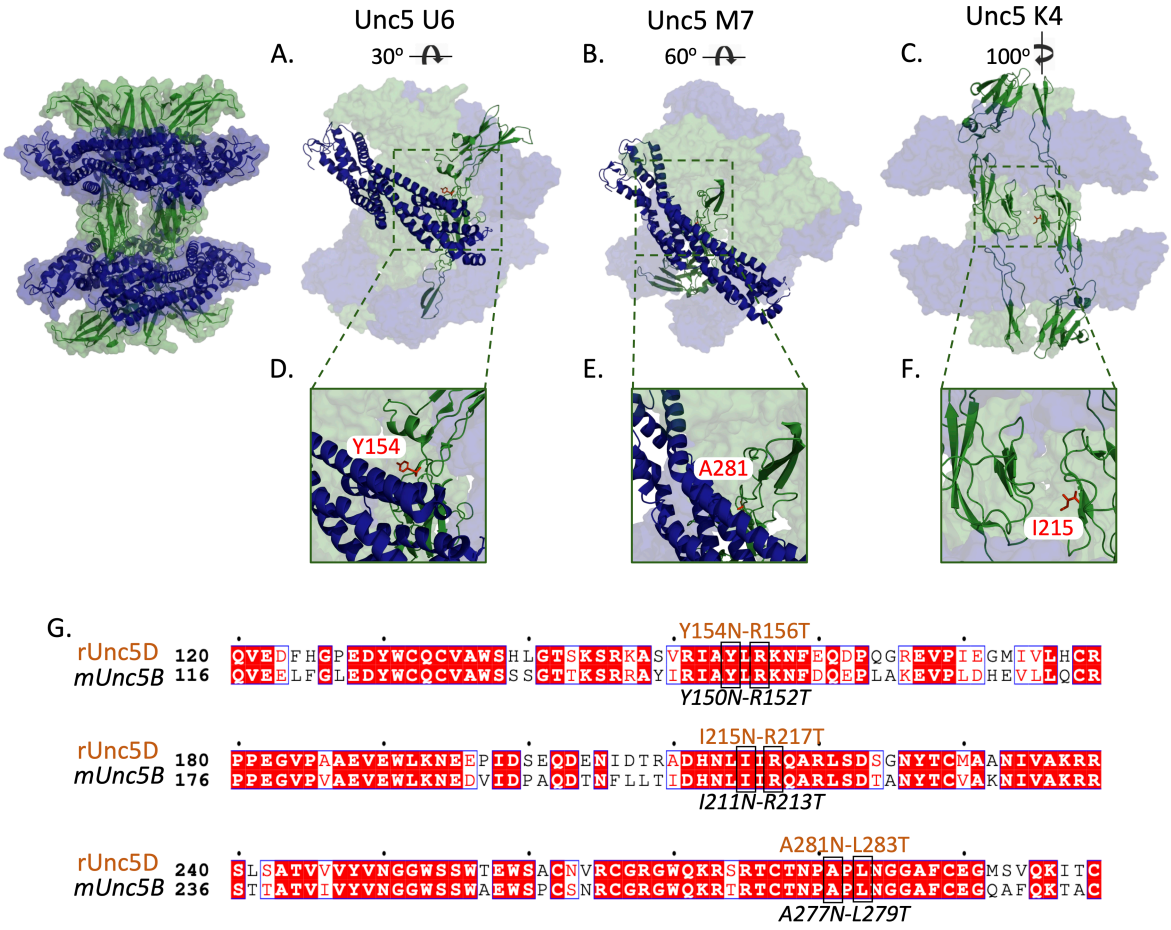


Figure 5.1: Design of the Unc5 mutants.
In these protein models of the octameric complex of hGPC3^{core} - rUnc5D^{Ig1Ig2TSP1}, the residues that are going to be mutated into an asparagine (N) are highlighted in red and in stick representation. (A) The location of the Unc5 U6 mutation in the octamer. (B) The location of the Unc5 M7 mutation in the octamer. (C) The location of the Unc5 K4 mutation in the octamer. (D-F) Zoom-in view of each mutation site. The residue which will get mutated in N is indicated. (G) rUnc5D and mUnc5B alignment and the corresponding mutations on rUnc5D and mUnc5B.

5.2.2 Mutation design on human GPC3

Mutations on the binding interfaces of the GPC3 molecule were also cloned. The GPC3 mutants cloned in the full-length hGPC3 are the following: hGPC3$^{Q112N\text{-}A114T}$ (K1), hGPC3^{L157N} (K2), and hGPC3^{N241Q} (K3) (Figure 5.2).

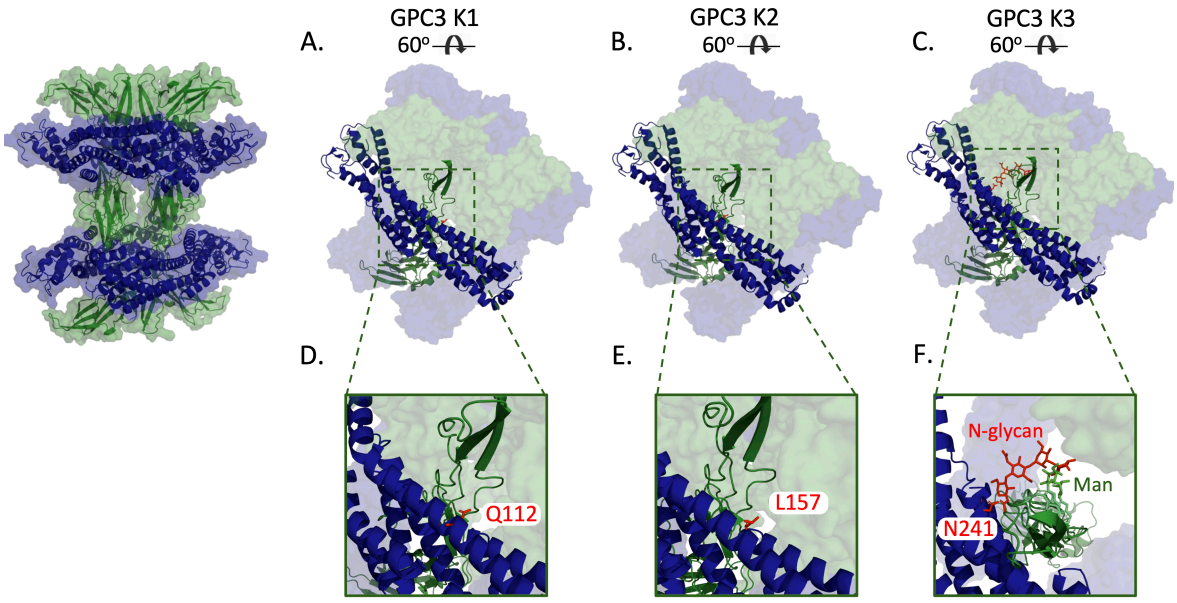


Figure 5.2: Design of the GPC3 mutants.
In these protein models of the octameric complex of hGPC3^{core} - rUnc5D^{Ig1Ig2TSP1}, the residues that are going to be mutated into an asparagine (N) are highlighted in red and in stick representation. (A) Location of the GPC3 K1 mutation in the octamer. (B) Location of the GPC3 K2 mutation in the octamer. (C) Location of the GPC3 K3 mutation in the octamer, in which the asparagine (N241) will be mutated into Glutamine (Q) for the deletion of this existing N-glycosylation site. (D - F) Zoom-in view of each mutation site. The residue which will get mutated in N is indicated. In panel F, the N-glycan attached to N241 is present, and its interaction with the Unc5 mannoses is also visible.

Mutation K1 and mutation K2 are additions of an N-glycosylation site and, by extension, a glycan on a binding interface with Unc5. Mutation K3, different from the others, is a deletion of a native glycosylation site since, based on the crystal structure, this N-linked glycan may stabilise the binding with rUnc5D by interacting with C-mannosylated tryptophans 253 and 256 of the TSP1 domain (Chapter 1.5.1).

5.3 Initial characterisation of the mutants in cell binding assays

As the mutants were initially characterised by cell binding assays, the expression of the mutants on the cell surface with the suitable tags is essential. For this purpose, the Unc5 mutants were cloned initially in the mUnc5B$^{full\text{-}length}$, containing an N-terminal HA tag and a C-terminal mVenus tag. HEK-293T cells were transfected with the mUnc5B$^{full\text{-}length}$ WT and mutants. All the protein mutants are expressed and present on the cell surface, as confirmed by cell surface staining (Figure 5.3).

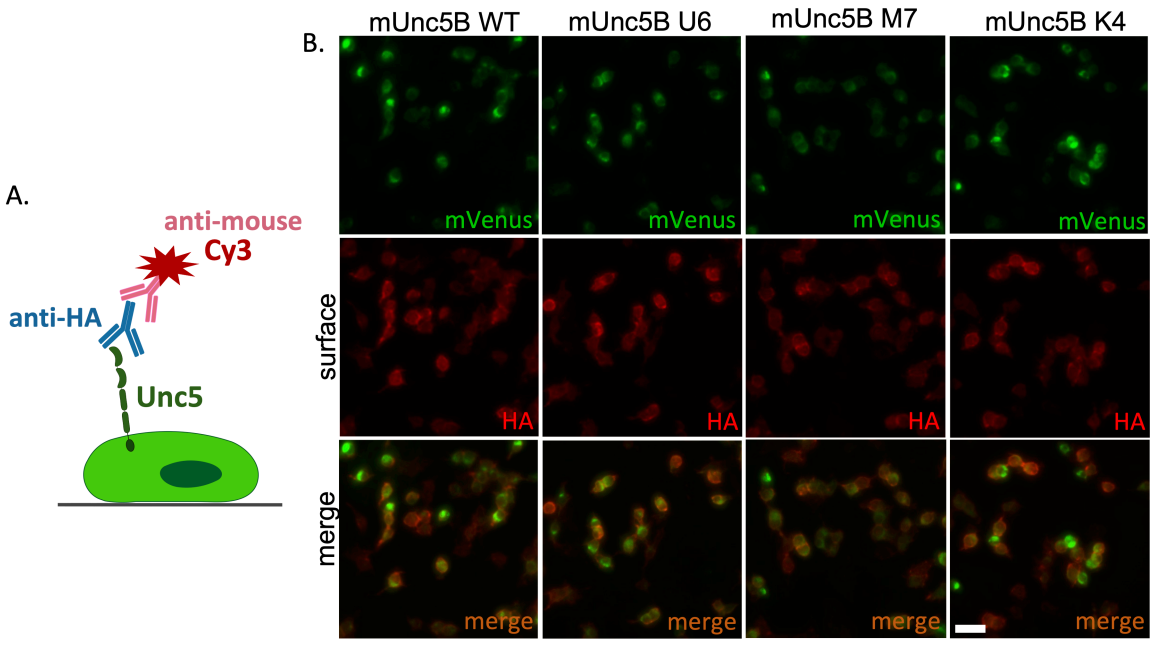


Figure 5.3: Cell surface expression of the full-length Unc5 WT and mutants. (A) Schematic representation of the cell surface staining experiment. A primary mouse anti-HA antibody detects the N-terminal HA tag of the Unc5 molecule, and a secondary anti-mouse Cy3-conjugated antibody detects the primary. (B) Confirmed cell surface expression of mUnc5B$^{full\text{-}length}$ WT and mutants. The Cy3 anti-HA signal confirms the presence on the cell surface, and the green of the cells (mVenus) colocalises with the red (anti-HA) in the merged image. Scale bar = 30 µm.

For this experiment to work, it is essential to confirm the presence of the proteins on the cell surface. It is particularly important for the mutant constructs since the behaviour of the protein might have changed after the introduction of an additional

N-linked glycan. The green colour of the cells confirmed the expression of the proteins in the cells because of the C-terminal mVenus tags of the Unc5s. To confirm the cell surface expression, an antibody against the N-terminal HA tag and a Cy3-linked secondary antibody was used (Figure 5.3A). The wild-type mUnc5B and all the mUnc5B mutants showed expression on the cell surface (Figure 5.3B).

For the binding studies, purified mGPC3^{ecto} (lacks the GPI anchor) was added to the cells that express the mUnc5B WT and the mutants on the surface. A primary anti-His antibody against the histidine tag of the mGPC3^{ecto} and a Cy3-linked secondary antibody were used. Since the mUnc5B expresses the mVenus tag, the colocalisation of the mVenus (green-coloured cells) and the Cy3 signal will demonstrate the binding of the two proteins on the cell. In the first trials, the amount of mGPC3^{ecto} added was not enough, and the signal was not sufficiently strong. Adjustments to the soluble mGPC3^{ecto} concentration fixed that problem. After the optimisation of the experiment, the results showed that all the mutations disrupt the binding between mUnc5B and mGPC3. The disruption of the interaction by the mUnc5B U6 mutant was less effective. The binding of mFLRT2^{ecto} to the Unc5, an already established binding partner of Unc5, was used as a positive control (Figure 5.4). As a negative control, purified mGPC3^{ecto} , the primary anti-His antibody and the Cy3-linked secondary antibody were added to untrasfected cells and no binding was detected.

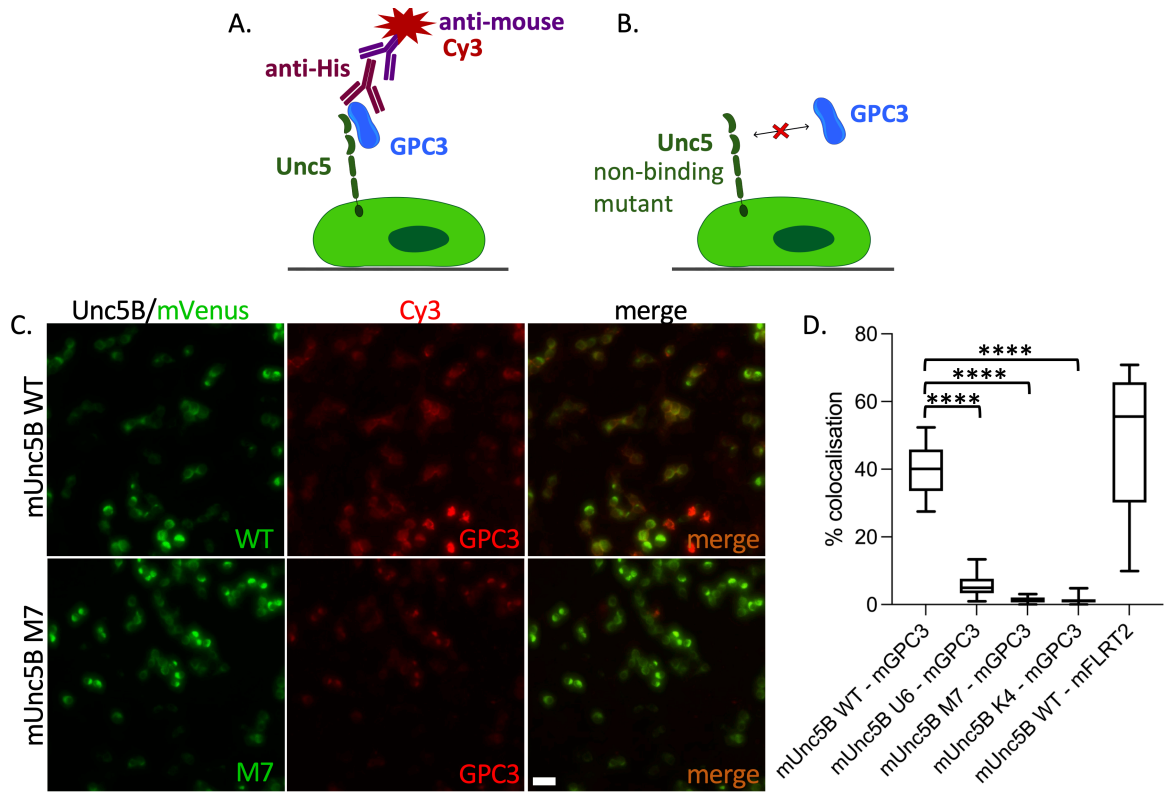


Figure 5.4: The Unc5B mutants disrupt the GPC3 binding.
(A-B) Schematic representation of the cell binding assay. Because of the mVenus tag, the cells expressing the mUnc5B$^{full\text{-}length}$ WT and mutants are green. The soluble mGPC3^{ecto} was tagged by a primary anti-His and a secondary Cy3-conjugated antibody. When GPC3 binds Unc5, the Cy3 signal is detected on the cell surface and colocalises with the mVenus. When GPC3 does not bind, the Cy3 signal is not present. (C) An example of cells from this cell binding assay. Cells expressing the mUnc5B WT have the Cy3 signal colour, whereas the cells expressing the mUnc5B M7 mutant do not. (D) Quantification of the cell binding assay and statistical analysis. **** p <0.0001, one-way ANOVA with Tukey's post hoc tests. Scale bar = 30 µm.

The GPC3 mutants were initially cloned in the full-length hGPC3, containing an N-terminal Flag tag. These mutants were also cloned in this template in order to express on the cell surface for the preliminary cell binding assays. An intracellular tag, like the mVenus for the Unc5 constructs, could not be added because GPC3 does not have transmembrane and intracellular domains, but it is attached to the cellular membrane with a GPI anchor [16]. The hGPC3$^{full\text{-}length}$ WT and its mutants did not show enough expression on the cell surface to proceed with the binding studies.

5.4 Further experiments for the mutants' characterisation

After the cell binding assays, SPR experiments were used to further confirm the Unc5 mutants' efficacy. SPR experiments showed that the binding of the hGPC3^{core} to the mUnc5B U6 mutant is not much different from the wild type, which partially agrees with the cell binding assay result in Chapter 5.3. The mUnc5B M7 mutant abolishes the binding completely. This non-binding characteristic of the M7 mutation on the mUnc5B appeared to be the same in all the Unc5 homologues (Figure 5.5 E-H). Therefore, the Unc5 M7 mutant was chosen as a non-GPC3-binding Unc5 mutant for further experiments.

The GPC3 mutants were tested with SPR since it was not possible to use the cell binding assays for their characterisation (as described above). The hGPC3^{N241Q} K3 was chosen for further experiments as a non-Unc5-binding GPC3 mutant as it abolishes the interaction with rUnc5D and hUnc5B (Figure 5.5A-D). The efficacy of this mutant on the cell surface was tested with cell aggregation assays. The results showed that the GPC3 K3 mutant disrupts the binding between GPC3 and Unc5 at a cellular level (Figure 5.5I-J). The different cell type in this assay (K562 vs HEK-293T in the cell binding assays in Chapter 5.3) was able to express the GPC3 constructs on the cell surface.

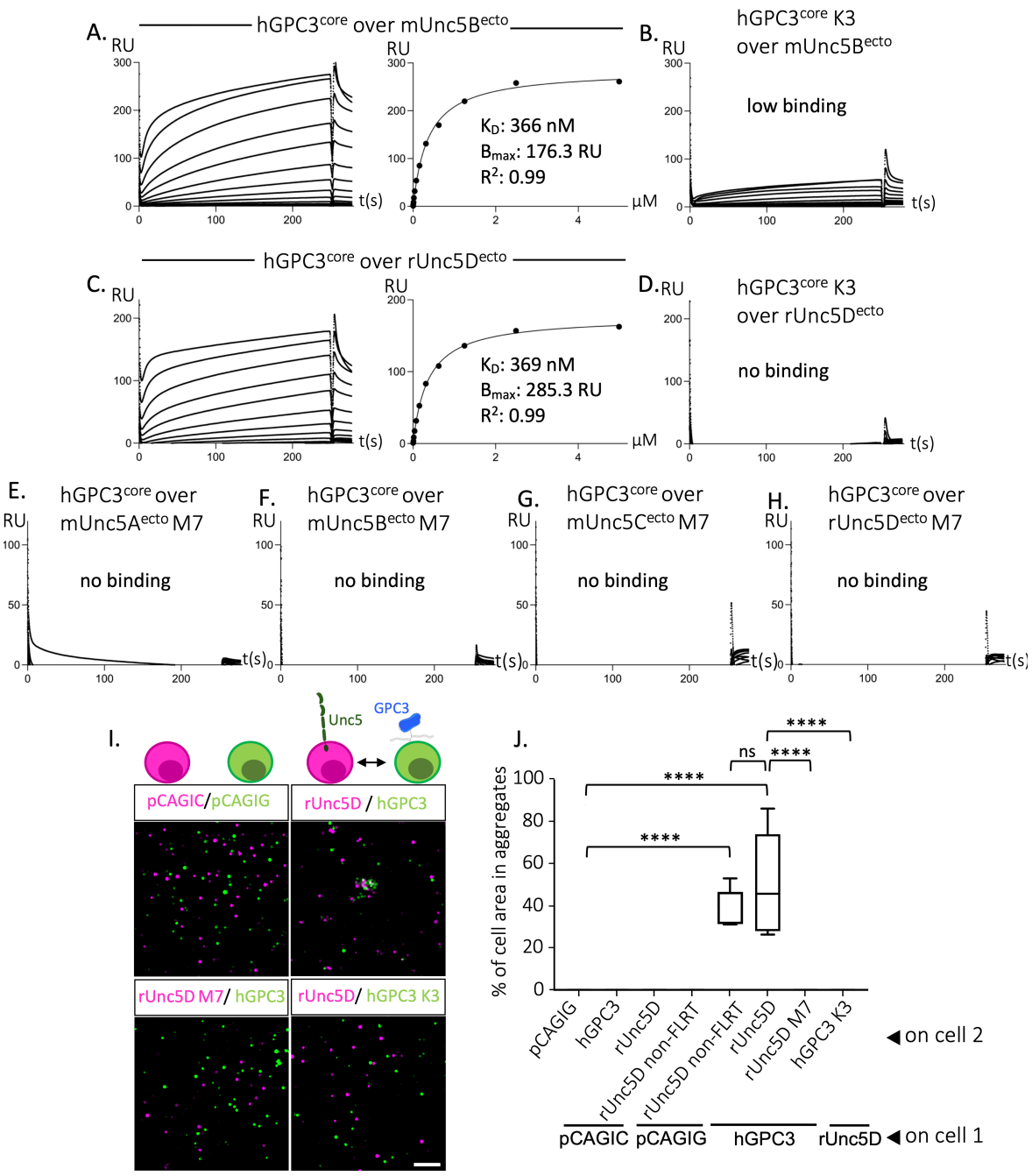


Figure 5.5: Additional experiments for the confirmation of the mutants' efficacy.

(A-D) SPR experiments confirm the non-binding of the hGPC3^{core} K3 to mUnc5B^{ecto} and rUnc5D^{ecto}. (E-H) SPR experiments confirm that the Unc5 M7 mutants disrupt the binding with hGPC3^{core} cloned in any Unc5 homologue. (I-J) Cell aggregation assays confirm the efficacy of the GPC3 K3 mutant at the cellular level. The cells do not interact in trans and do not form aggregates when a group of cells express the hGPC3 K3 mutant and another group the rUnc5D WT. Statistical analysis: one-way ANOVA with Tukey's post hoc tests. ****$p < 0.0001$, Scale bar = 100 μm. Figure adapted from [30].

5.5 Validation of the mutants' correct folding

After the characterisation of the mutants by SPR experiments, the rUnc5D$^{A277N-L279T}$ M7 and the hGPC3^{N241Q} K3 mutants were selected to be used in other *in vivo* and *in vitro* experiments. Adding an N-glycosylation site to the sequence and, therefore, an N-linked glycan to a specific site on a protein which was not present before, can affect the protein's folding. The same is true for the deletion of an N-glycosylation site in the case of hGPC3 K3 mutation. All the steps of protein expression, post-translational modification and secretion can secure the protein's proper folding, but it is not always the case.

More experiments were used to assess if the proteins were folded correctly after adding or deleting a glycan. The proper folding of the rUnc5D M7 mutant was assessed by checking the binding of FLRT2 since FLRT2 is an established Unc5 binder. A cell binding assay showed that the mUnc5B M7 mutant can still bind FLRT2 as much as the mUnc5B WT (Figure 5.6A). This provides evidence to support the concept that the protein is folded correctly, but it also shows that the M7 mutation specifically disrupts the GPC3 binding.

Circular Dichroism (CD) was used to examine the secondary structure of the GPC3 K3 mutant. This experiment showed that the hGPC3^{core} K3 mutant still has its characteristic helical folding since its CD spectrum exhibits a positive peak at 195 nm and two negative peaks at 208 and 222 nm, which is typical for proteins containing α-helices [209]. The hGPC3^{core} WT was also tested to compare with the hGPC3^{core} K3 spectrum at the same concentration. The two CD spectra do not present any differences in this experiment, indicating that at least in the secondary structure level, the hGPC3^{core} WT and K3 mutant remain the same (Figure 5.6B).

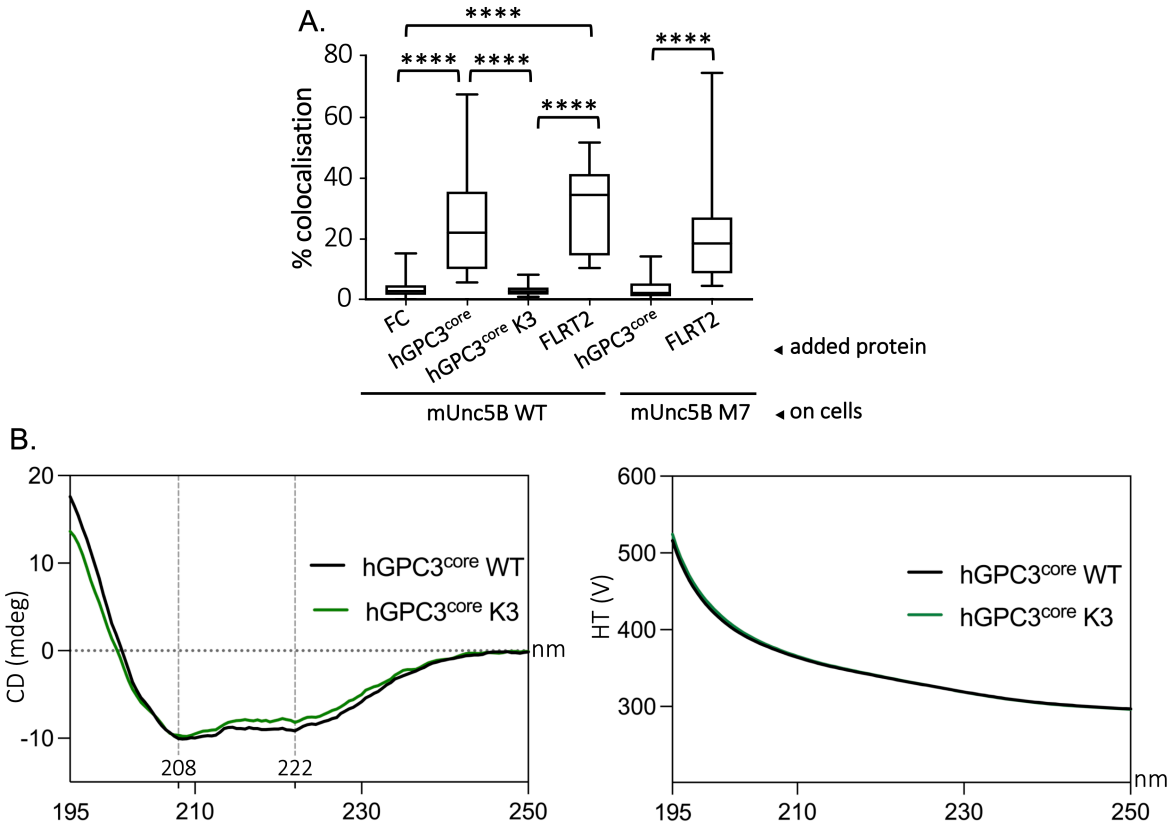


Figure 5.6: Data to support that the mutant proteins are correctly folded. (A) Cell binding assays show that the mUnc5B M7 mutant cannot bind GPC3 but can still bind FLRT2, while the mUnc5B WT can bind both GPC3 and FLRT2. Adapted from [30]. (B) CD experiment shows that the hGPC3^{core} K3 mutant spectrum displays the same peaks as the α-helical CD spectral signature but also the same as the hGPC3^{core} WT.

5.6 Discussion

The GPC3-Unc5 crystal structure revealed an octameric complex of four GPC3 and four Unc5 protein chains. In order to probe and validate the binding of these two proteins, N-linked glycosylation sites were introduced on the binding interfaces in both the GPC3 and Unc5 molecules. The addition of an N-glycan to the binding site between two proteins is an established approach to disrupt a protein-protein interaction [72,77,202,203,208].

The Unc5$^{A277N-L279T}$ M7 and the GPC3^{N241Q} K3 mutants were chosen among other GPC3 and Unc5 mutants that were designed. Cell assays and SPR experiments

were used to characterise the mutants at the cellular level (cell binding and cell aggregation assays) and in solution (SPR). It is important to note that the artificial triplet motif of the M7 mutant contains a proline residue between the asparagine and threonine residues of the motif. This proline could hinder the glycosylation on the site. Both the M7 and the K3 mutants disrupted the GPC3-Unc5 interaction in all experiments.

The GPC3^{N241Q} K3 mutants did not comprise the addition of an artificial glycosylation site but the deletion of one. Since the crystal structure revealed a close contact between an N-linked glycan from the GPC3 molecule and two tryptophan α-mannoses from the Unc5 molecule, the deletion of the GPC3 glycosylation site was the next step to evaluate this glycan-glycan interaction. The cell assays and SPR confirmed this glycan interaction. The addition or deletion of an N-glycosylation site and, consequently, the presence or absence of an N-glycan on a protein molecule could result in structural changes. The known Unc5 binding partner, FLRT2, could still bind the Unc5B M7 mutant, which supports the correct folding of the Unc5 M7 mutant. The GPC3 K3 mutant's secondary structure was tested with Circular Dichroism, and the spectrum showed no changes compared to the GPC3 WT.

The chosen mutants were used in other experiments to investigate the GPC3-Unc5 interaction. Firstly, the non-binding mutants confirmed the Unc5-GPC3 interaction in the SPR experiments and the cell assays. Afterwards, the mutants were used in stipe assays (Chapter 1.5.2), in *in vivo* cell migration assays in the mouse cortex (Chapter 1.5.3) and in a neuroblastoma *in vivo* migration model (Chapter 1.5.4) presented in detail in the introduction. The mutants precisely interrupted the GPC3-Unc5 interaction, and the effects were observed. Together

with other tools, like the nanobodies, the mutants contributed to the understanding of the importance of the GPC3-Unc5 interaction in neuronal and cell migration [30].

6. VLGR1

6.1 Introduction and summary

The adhesion GPCR, VLGR1 (also referred to as ADGRV1, GPR98), is the largest member of its receptor family with the longest isoform to exceed the molecular weight of 700 kDa. The literature suggests that this adhesion GPCR has important roles in the nervous system, but there is little information about the three-dimensional structure and function of this enigmatic GPCR. Several protein constructs of small parts of the extracellular domain of the longest mouse VLGR1b isoform were designed, cloned, and expressed in mammalian cells. However, their purification proved to be difficult. Steps were taken to solve this problem by optimising the purification and the construct design without improvement. Today, in the AlphaFold database, a few predicted models of protein fragments of the VLGR1 receptor are available, but there is no model of the full-length protein. In this chapter, the construct design and their expression and purification are presented.

6.2 Bioinformatic analysis of the VLGR1 domain architecture

The aim of this project was initially the structural characterisation of the longest mouse VLGR1 isoform, mVLGR1b (Uniprot: Q8VHN7-1). This isoform is a very large protein that comprises 6298 amino acids. Due to its large size, the structural characterisation of the full-length protein would be challenging. Therefore, a small fragment approach was attempted, and the structural characterisation of parts of the full-length VLGR1 ectodomain was the project's first step. At the same time, the

cloning, expression, and purification of the full-length mVLGR1b and the entire extracellular domain were attempted.

Before any construct design or further biophysical and structural studies on this protein, a better understanding of its domain architecture was necessary. In the literature, structural features of the VLGR1 protein were described based on predictions from its amino acid sequence, and they are summarised in Chapter 1.7.2. Several repetitions of CalX-β motifs constitute the receptors' long extracellular domain, and they are interrupted by a LaminiG/pentraxin domain and EPTP/EAR repeats. The typical in the adhesion GPCRs family, GAIN domain and GPS motif are also present.

For the construct design of this project, domain prediction and secondary structure prediction tools were used to give a more detailed understanding and determination of the domain boundaries on the sequence level. The InterPro database was used to identify the different domains on the mVLGR1b extracellular domain and get the predicted boundaries. Through InterPro, search is being done into 13 different protein signature databases [210]. Except for InterPro, the SMART (Simple Modular Architecture Research Tool) and Pfam were also used to help identify and annotate domains on the mVLGR1b protein [211,212]. At the time, SMART and Pfam were not part of Interpro databases. The PSIPRED server and its PSIPRED secondary structure prediction method were also used to refine the exact boundaries of the constructs [213,214]. The different predictions did not agree on the domain boundaries or the number of repetitions for most of the extracellular region subdomains. Only determining the structure would give an accurate location of the domain boundaries. A summary of the predictions is presented in Figure 6.1.

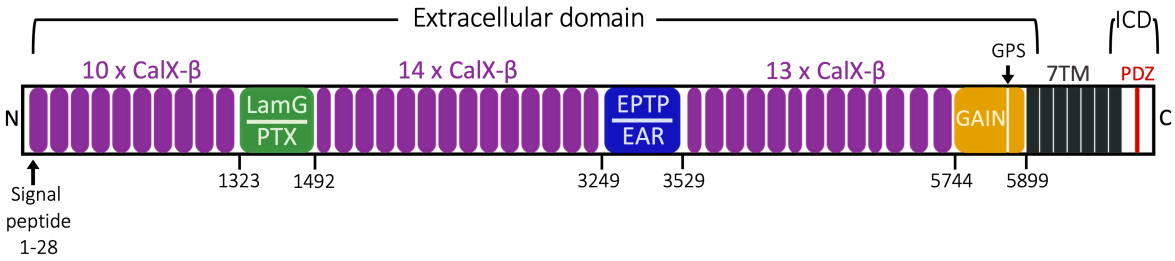


Figure 6.1: Predicted domain architecture of the mVLGR1b.
Schematic representation of the mVLGR1b domains, which comprise a combination of the predictions by the different domain prediction tools. The number of the CalX-β motifs and the predicted boundaries of the LaminiG, EPTP and GAIN domains are noted. CalX-β: Ca^{2+} binding motif, LamG/PTX: predicted jellyroll fold, EPTP/EAR repeats: predicted seven-bladed β-propeller structure, ICD: Intracellular domain, 7TM: Seven transmembrane domain, GAIN: GPCR autoproteolysis-inducing domain, GPS: autoproteolytic site.

The domain prediction tools predicted the autoproteolytic site, GPS, as it is conserved and easy to identify (Thr5846 - Ser5896). The same applies to the seven-transmembrane domain (7TM). The GAIN domain was predicted by Interpro and annotated on the mVLGR1b sequence from Leu5744 to Asn5899. However, with the help of homology modelling in SWISSMODEL [215] and the secondary structure prediction with PSIPRED, the boundaries of the GAIN domain were predicted to be Glu5634 and Ser5896, including the GPS motif. The extracellular region of mVLGR1b is predicted to be heavily glycosylated based on the literature but also an N-glycosylation prediction tool [216].

6.3 Initial construct design, expression and purification

Nine initial constructs were designed based on the information from Interpro, Pfam and SMART but also from PSIPRED in order to avoid the cloning of only small parts of structural features like an α-helix or a β-strand (Figure 6.2A). Except for the help from domain prediction tools, constructs #2 and #3 were designed based on the fact that they correspond to the isoforms VLGR1 isoforms, d and e, respectively.

This increased the chases of cloning constructs that express in mammalian cells and are well-folded.

A small-scale expression test showed that 6 of these 9 initial protein constructs were expressed (Figure 6.2B). The long construct #1 and the two constructs with similar boundaries for the EPTP/EAR repeats (#8, #9) showed no expression.

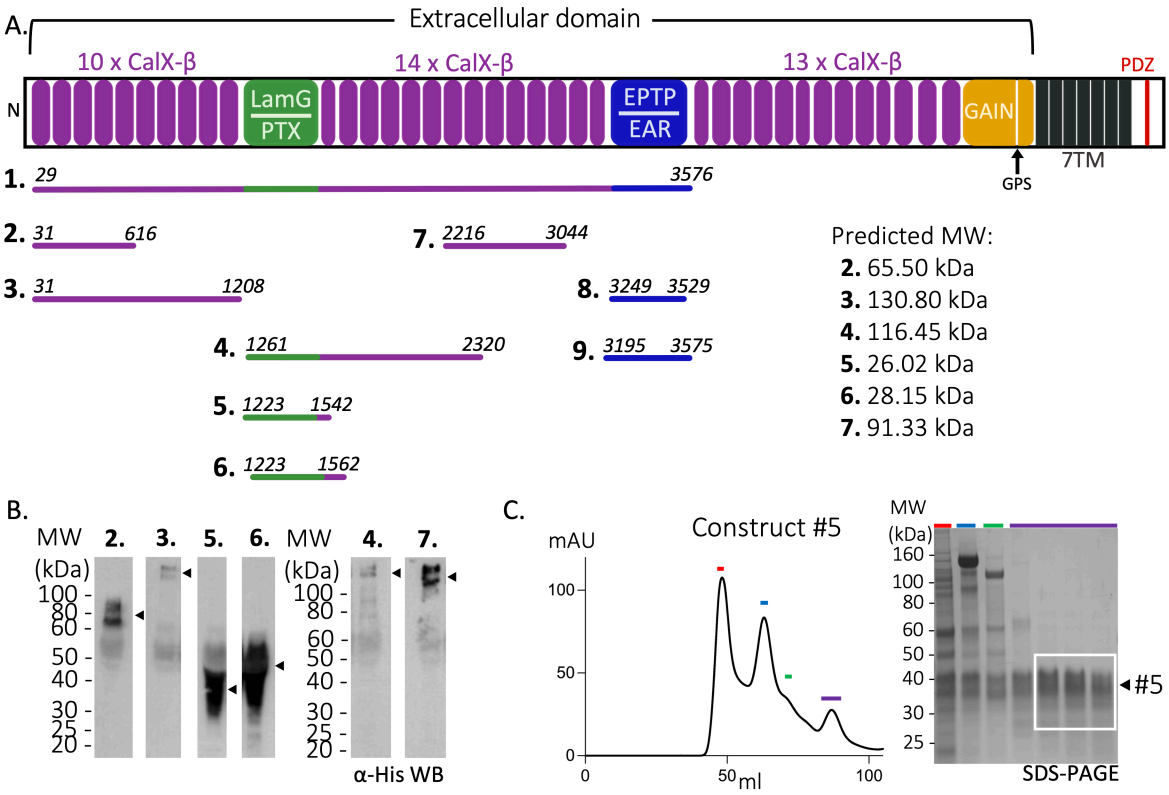


Figure 6.2: Initial mVLGR1b construct design and expression.
(A) Schematic representation of the constructs designed (#1 - #9) on the extracellular region of the mVLGRb. They were all cloned into pHLsec with a C-terminal 6 × histidine tag. Their boundaries are marked at the edges of each construct with the residue number. The molecular weight of the constructs that express was predicted based on their sequence without considering the predicted glycosylation. (B) Anti-His western blots of the small-scale expression tests (HEK-293T cells) of the constructs that showed expression. (C) SEC profile and SDS-PAGE of the SEC fractions of construct #5. The white box indicates the fractions that were used for crystallisation trials.

The next step was the large-scale expression and purification of the constructs that showed expression on the small-scale test. The aim was the crystallisation and structure determination of each protein construct. Construct #5, which comprises the predicted LaminG domain, was expressed and purified successfully, and crystallisation plates were prepared, but crystals did not grow (Figure 6.2C).

The large-scale expression and purification for the rest of the constructs (constructs #2, #3, #4, #7) were not successful. In more detail, the example SEC profile of construct #7 and the SDS-PAGE of the SEC fractions show that it was not possible to identify the protein of interest. An anti-His western blot showed that the protein was not well expressed and was present in all the SEC profile peaks. Also, a significant amount of the protein was not binding to the column and was present in the flow through (FT) but also present in the 40 mM Imidazole wash step (W) of the His-trap purification (Figure 6.3). This pattern was the same for the rest of the constructs.

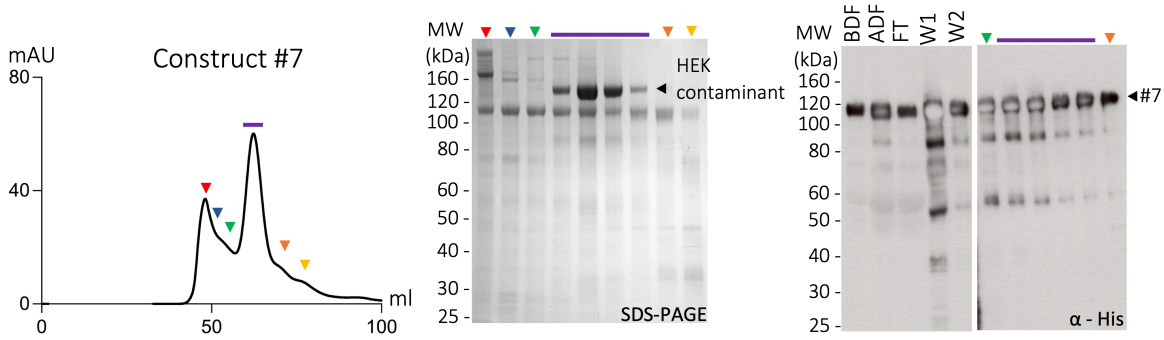


Figure 6.3: Purification of construct #7.
SEC profile of the construct #7 purification. SDS-PAGE and anti-His western blot of the purification fractions. This purification is an example of the purification of the rest of the constructs (#2, #3, #4), as they all behaved similarly. BDF: Before Diafiltration (buffer exchange before His-trap), ADF: After Diafiltration, FT: Flow Through, W1, W2: First and second wash with 40mM imidazole buffer. HEK contaminant: a usual contaminant from the expression and purification of protein constructs in HEK-293T and HEK-293S cells.

6.4 Optimisation of the purification

Steps were taken to minimise the protein loss in the flow through and wash steps but also to improve the buffer conditions for the protein. A common structural feature of all the constructs which failed to be purified on large-scale is the CalX-β motifs. CalX-β motifs have been shown to bind calcium in other proteins but in the VLGR1 as well [164,166]. Based on this, $CaCl_2$ was added to the purification buffers and the effect of the $CaCl_2$ was tested. The addition of $CaCl_2$ did not improve the protein loss in the FT and Wash steps in small-scale His-trap purifications (Figure 6.4).

Another approach was the addition of a 10 × histidine tag either on the N-terminal or the C-terminal end of the protein constructs instead of the already existing C-terminal 6 × histidine tag to increase the binding strength to the column and decrease the protein loss. The 10 × histidine tag construct showed an improvement as less protein was being lost in the FT and Wash, but it did not solve the problem completely (Figure 6.4).

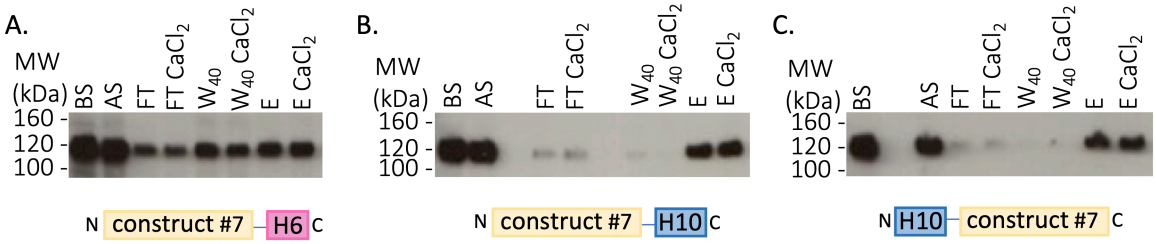


Figure 6.4: Small-scale His pull-downs to compare different histidine tags. Anti-His western blots of the samples of small-scale His pull-downs (A) Pull-down of construct #7 with a C-terminal 6 × histidine tag (H6) (B) Pull-down of construct #7 with a C-terminal 10 × histidine tag (H10) (C) Pull-down of construct #7 with an N-terminal 10 × histidine tag. BS: Before Spin, AS: After Spin (input), FT: Flow Through, W_{40}: Wash with 40mM Imidazole, E: Elution, $CaCl_2$: the buffers for this pull-down contain $CaCl_2$.

6.5 New construct design

Since these optimisation steps did not improve the purification of the mVLGR1b protein constructs, designing new constructs was the step that followed. The purification optimisation steps were utilised in the new construct design, and the new constructs were all cloned into pHLsec with a C-terminal 10 × histidine tag. Initially, by observing the protein sequence of the mouse VLGR1b extracellular domain, a repetitive amino acid pattern was noticed. Homology modelling using SWISSMODEL predicted this pattern to be the linker between the CalX-β motifs, which indicated different domain boundaries for the CalX-β motifs than was predicted before by the domain prediction tools. At the time, AlphaFold was not available.

These different domain boundaries were used to design new constructs (Figure 6.5A). All the new constructs were cloned into a pHLsec vector with a 10 × histidine tag. A small-scale expression test showed expression for 6 of these 13 new constructs (Figure 6.5B).

The purification of the new constructs was tested first on a small-scale His pull-downs. Construct #10, one of the newly designed constructs, behaved well, and there was no protein loss at the FT and W steps of this affinity purification (Figure 6.5C).

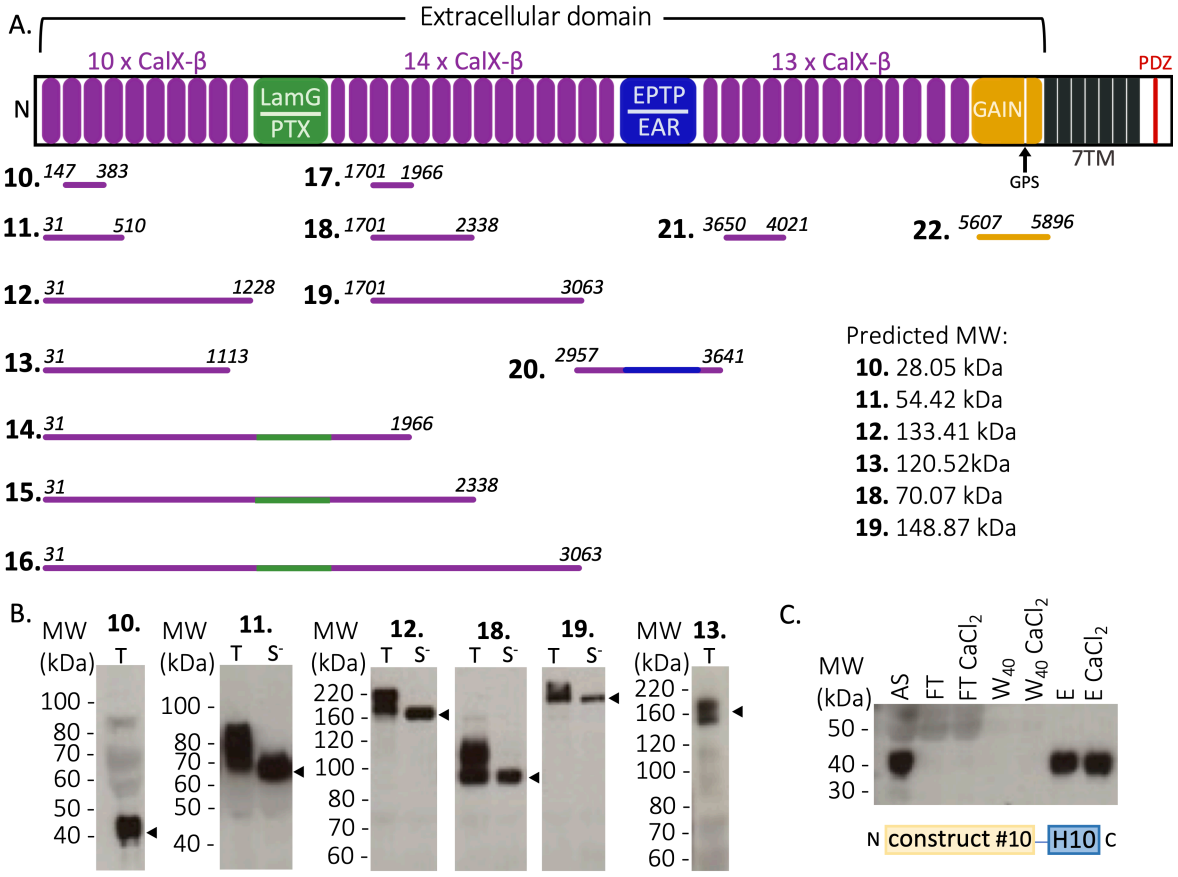


Figure 6.5: Design of new mVLGR1b constructs.
(A) Schematic representation of the constructs designed (#10 - #22) on the extracellular region of the mVLGRb. They were all cloned into pHLsec with a C-terminal 10 × histidine tag. Their boundaries are marked at the edges of each construct with the residue number. The molecular weight of the constructs that express was predicted based on their sequence without considering the predicted glycosylation. (B) Anti-His western blots of the small-scale expression tests of the constructs that showed expression. The expression was tested in HEK-293T (T) and HEK-293S GnTI$^-$ (S^-) cells. The expression in HEK-293S GnTI$^-$ cells present a solid, less smeary band because of the no complex, homogenous glycosylation. (C) Anti-His western blot of the samples of small-scale His pull-downs of construct #10. H10: 10 × histidine tag, AS: After Spin (input), FT: Flow Through, W_{40}: Wash with 40mM Imidazole, E: Elution, $CaCl_2$: the buffers for this pull-down contain $CaCl_2$.

However, large-scale expression and purification of the new constructs were not improved (Examples in Figure 6.6). At the SEC step, all the new protein constructs were eluted in every single collected fraction, and they were not concentrated at the fractions that should be based on their shape and size. Neither the 10 × histidine tag nor the presence of $CaCl_2$ in the buffers improved the purification.

In the His trap purification of construct #10, protein is present in the FT and Wash steps (Figure 6.6B), and the presence of 5 mM $CaCl_2$ in the purification buffers of construct #11 did not improve the SEC outcome (Figure 6.6C).

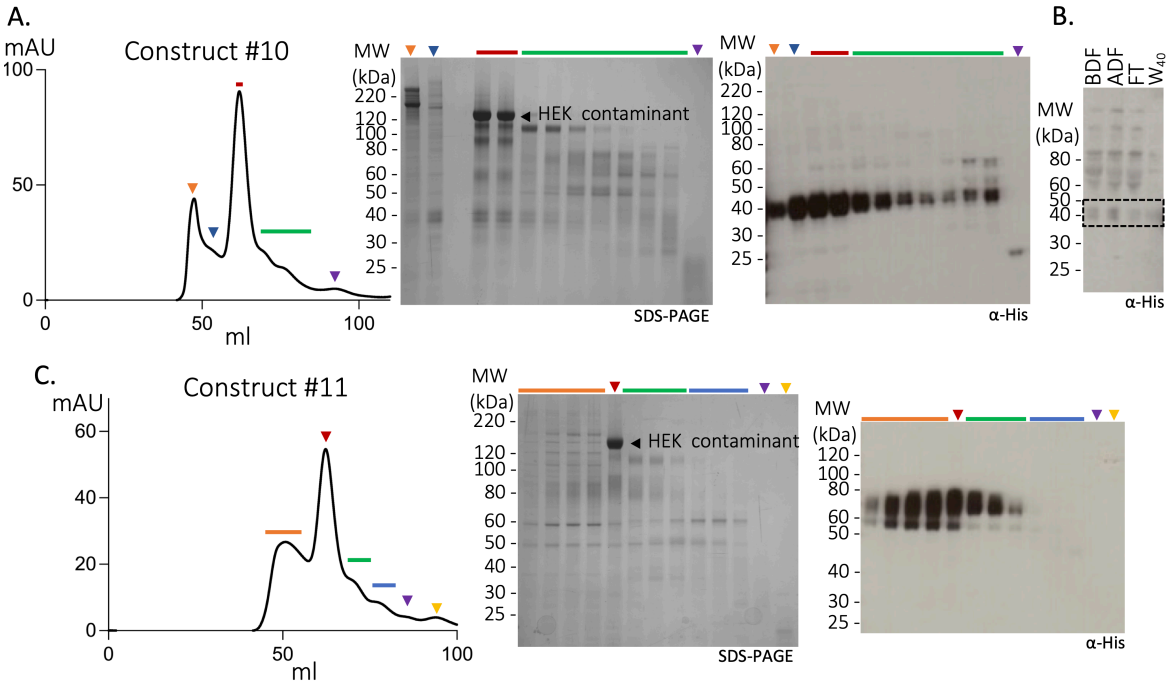


Figure 6.6: Large-scale purification of the newly designed mVLGR1b constructs.
(A) SEC profile of the construct #10 purification. SDS-PAGE and anti-His western blot of the SEC fractions. (B) Anti-His western blot of the His-trap purification samples of construct #10. BDF: Before Diafiltration (buffer exchange before His-trap), ADF: After Diafiltration, FT: Flow Through, W_{40}: Wash with 40 mM Imidazole buffer. (C) SEC profile of the construct #11 purification in buffers supplemented with 5 mM $CaCl_2$. SDS-PAGE and anti-His western blot of the SEC fractions. HEK contaminant: a usual contaminant from the expression and purification of protein constructs in HEK-293T and HEK-293S cells.

To date, the AlphaFold database does not contain a predicted model for the entire length of the long VLGR1b isoform. Few shorter fragments of the extracellular domains of VLGR1b are available. The new constructs designed for the extracellular domain of the mVLGR1b, and the new boundaries, agree with the AlphaFold-predicted models. Also, based on the AlphaFold prediction of the first 10

CalX-β motifs (Figure 6.7B), the constructs #10, #11, #12 and #13 contain 2, 4, 10 and 9 CalX-β repeats, respectively, as predicted. The 10th CalX-β motif in construct #12 based on the AlphaFold prediction is not complete. A potential new construct should be designed longer in order to include the full 10th CalX-β motif.

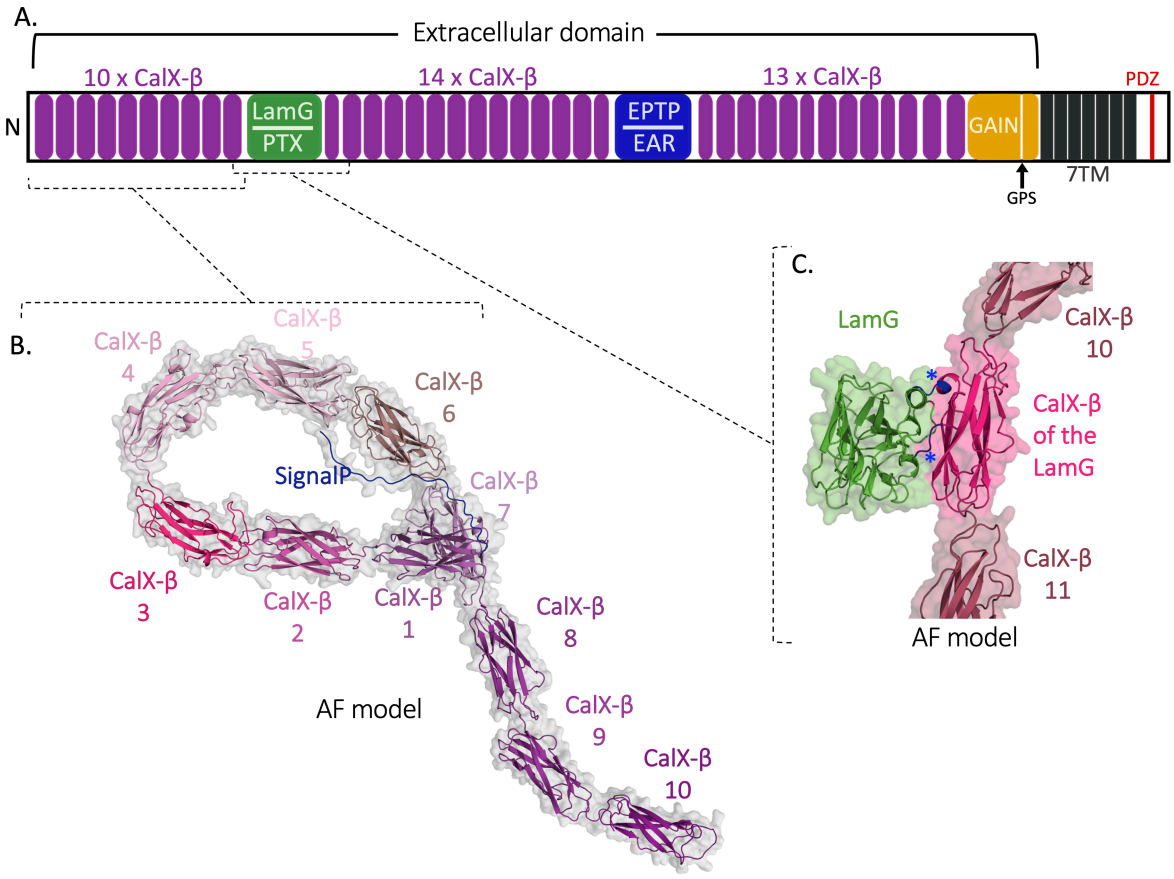


Figure 6.7: AlphaFold database predicted models of VLGR1 fragments.
(A) Schematic representation of the mouseVLGR1b domain architecture. (B) AlphaFold-predicted model of the first ten CalX-β motifs of the mVLGR1b, UniProt: D3YYF1 (C) Part of an AlphaFold-predicted model that includes the LamininG predicted domain. This model is a prediction from a VLGR1 fragment of the Pongo abelii (Sumatran orangutan), UniProt: A0A2J8UX75. The predicted LamininG domain is included in a CalX-β motif sequence and structure.

For the mouse VLGR1, few predicted models are available in the AlphaFold database. Predicted models of VLGR1 fragments of other species are available as well. An interesting finding was the AlphaFold predicted model of the LamininG domain of the Pongo abelii (Sumatran orangutan) VLGR1. The protein sequence of

the fragment of this predicted model that includes the LamininG domain is very similar to the mouse equivalent sequence. The predicted model of the sequence after the 10th CalX-β motif revealed a domain that combines a CalX-β motif and a LamininG (Figure 6.7C). This architecture would be impossible to predict without the contribution of AlphaFold.

6.6 Discussion

Several protein fragments of the extracellular domain of the mVLGR1b were designed, cloned and expressed. However, the purification of these constructs was not successful. The expressed protein constructs did not behave well, resulting in proteins that could not get purified. Today, AlphaFold can be used to predict the individual subdomains of the extracellular region of the VLGR1b more accurately and help in construct design. The already predicted models in the AlphaFold database showed some promising results. Therefore, AlphaFold can help with a better prediction of the domain boundaries and potentially the design of well-behaved small domain constructs.

The extracellular domain of the VLGR1b adhesion GPCR is a very large protein of almost 640 kDa and 5900 amino acids long. The folding of a protein that is that long and has several repetitive motifs is very difficult to predict. The small size fragments designed in this project might need a binding partner in order to behave well, be soluble and easy to handle. It is also possible that this protein's extracellular region forms a highly organised three-dimensional structure, and it does not present as an elongated protein physiologically without a binding partner. The next promising step in this project would be the cloning, expression, purification, and

structural characterisation of the entire length of the extracellular region. Its size makes it a suitable candidate for cryo-EM.

7. Discussion and future plans

The structural characterisation of the newly identified GPC3-Unc5D interaction revealed an octameric complex structure comprising four hGPC3^{core} and four rUnc5D^{Ig1Ig2TSP1} molecules. Following the discovery of the interaction, its structure and function were questioned. In this thesis, work was done on the production and characterisation of non-binding mutants and nanobodies against hGPC3 to be used as tools for the investigation of the function of the Unc5-GPC3 interaction.

7.1 Non-binding GPC3 and Unc5 mutants

Non-GPC3-binding Unc5 mutants and non-Unc5-binding GPC3 mutants were generated to help probe and validate the binding of these two proteins, GPC3 and Unc5. Mutagenesis is a common way of protein binding disruption to help study the functions of an interaction. The addition of artificial N-glycosylation sites on the protein binding interfaces was used as it is an established approach of mutagenesis aiming at complex disruption [77,202,203,208]. Except for the addition of an artificial N-glycosylation site, in this case, a deletion of a glycosylation site also abolished the interaction between the two proteins.

One non-GPC3-binding Unc5 mutant (M7) and one non-Unc5-binding GPC3 mutant (K3) were chosen among the GPC3 and Unc5 mutants produced and characterised. Both mutations abolish the Unc5-GPC3 interaction. For the Unc5 M7 non-GPC3-binding mutant, an artificial N-glycosylation site was introduced on a GPC3-Unc5 binding interface between the Ig2 and TSP1 Unc5 domains. On the other hand, the GPC3 K3 non-Unc5-binding mutant was generated by deleting an N-glycosylation site located at the N241 of the hGPC3. The crystal structure

revealed this interaction between the N241 glycosylation on hGPC3 and the α-mannoses on two tryptophan residues of the rUnc5D molecule. Usually, an N-linked glycan chain is not visible in a crystal structure because of its flexibility. In this case, the interaction with Unc5 stabilised the glycan and made it visible in the density map close to the C-mannosylated tryptophans. This led to the design of the deletion mutant GPC3 K3 which resulted in disruption of the GPC3-Unc5 interaction and confirmed the presence of the glycan-glycan interaction. At the same time, this result highlighted the significance of glycosylation for the cell membrane proteins' function, which has been documented before in the literature in other cell membrane protein systems [217,218].

The selected mutants had a significant role and were used successfully in all the *in vitro* and *in vivo* assays conducted by Akkermans and colleagues to understand the GPC3-Unc5 interaction [30]. The modulation of the GPC3-Unc5 interaction with the mutants revealed the importance of this interaction in neuronal and cell migration.

7.2 Nanobodies against GPC3

Apart from the use of the mutants in functional experiments, nanobodies against human GPC3 that modify the Unc5-GPC3 interaction were used successfully as tools to characterise the interaction. In recent years, the nanobodies have been considered highly versatile and powerful tools widely used in research [119,120].

7.2.1 Biochemical and biophysical characterisation

In this thesis, nanobodies against hGPC3^{core} were produced and characterised. Among several anti-hGPC3^{core} nanobodies, two nanobodies (Nb49, Nb58) were identified that affect the Unc5-GPC3 interaction. Nb58 disrupts, and Nb49,

interestingly, enhances the interaction. The affinity of these nanobodies to hGPC3^{core} was measured at K_D = 160 nM for Nb49 and K_D = 8 μM for Nb58. For the purposes of this research, the affinity of these nanobodies to their target was sufficient. Affinity maturation would be needed in order to achieve affinities in the lower nM or pM affinities. Pull-downs and SPR-based competition assays verified the effects of these nanobodies in the GPC3-Unc5 interaction. The potential unspecific binding of the nanobodies to the Unc5 proteins was also tested, and neither Nb49 nor Nb58 binds to Unc5. SEC-MALS experiments were used to understand the stoichiometry of the hGPC3^{core}-Nb49 and the hGPC3^{core}-Unc5D-Nb49 complexes. However, because of the low affinity of the interactions, SEC-MALS did not help to get a definite answer to this question.

Additional SPR test injections were used as a screening experiment to get a first estimation of the characteristics of the remaining available nanobodies against hGPC3^{core}. Based on these SPR test injections, the rest of the nanobodies range at the same levels of affinity as the Nb49 and Nb58. Nanobodies that bind GPC3 as strong as Nb49 or with even higher affinity were recognised. Test injections in a competition set-up also revealed promising nanobodies for future use. A nanobody that abolishes the GPC3-Unc5 interaction completely and, therefore, more effective than Nb58, was identified. Nb58 inhibits the interaction, but it does not completely abolish it. More experiments are needed to fully characterise these additional nanobodies, as they can be very useful for experiments in the future. For example, it is known that GPC3 interacts with the morphogens Wnt and Hedgehog and modulates their signaling pathways [27,38]. The interaction of the nanobodies with GPC3 can be characterised under the prism of the GPC3 interaction with Wnt and

Hh, and nanobody modulators of these interactions could be identified, like Nb49 and Nb58 modulate the GPC3-Unc5 interaction.

7.2.2 Structural characterisation of the Nanobody-GPC3 complexes

The biochemical and biophysical assays successfully characterised how the nanobodies affect the Unc5-GPC3 interaction, and the structural characterisation was needed to fully understand and confirm their function. In this thesis, the X-ray crystallography approach was selected. The fact that the $hGPC3^{core}$ crystal structure was available was the key factor in this decision since it would make the structure determination with X-ray crystallography easier. Additionally, the initial objective was to obtain the detailed crystal structure of the GPC3-nanobody complexes in order to use protein engineering, design mutations on the binding interface and manipulate the interaction.

The crystallisation of GPC3 in complex with its nanobodies proved to be challenging. With the help of AlphaFold multimer predictions in molecular replacement, the structure of the GPC3-Nb49 complex was obtained. Despite the relatively low resolution and the moderate electron density of the nanobody molecule, the binding epitope of the nanobody on GPC3 was revealed. To confirm the binding interface, mutagenesis is needed and N-glycosylation site mutations on the binding interface have already been designed.

The crystal structure of the GPC3-Nb49 structure revealed the binding epitope of Nb49 to GPC3, but it did not help explain how the nanobody enhances the GPC3-Unc5 interaction. Since Nb49 does not bind Unc5 and the nanobodies were generated against the $hGPC3^{core}$ and not a GPC3-Unc5 complex, this enhancement does not depend on a dual nanobody interaction with GPC3 and Unc5. This suggests that the binding enhancement results from conformational changes of the

hGPC3^{core} structure after Nb49 binding. The crystal structure supports this idea since the Nb49-bound hGPC3^{core} presents a slightly different conformation to the non-Nb49-bound GPC3. The determination of the structure of the GPC3-Unc5-Nb49 complex can give a more definite answer, and progress was made in the optimisation of Unc5-GPC3-Nb49 complex particles on cryo-EM grids.

When a detailed structure is not necessary, different approaches can be used to help get the needed information about the binding epitope of a nanobody in a more rapid and efficient way. Hydrogen-deuterium exchange (HDX) mass spectrometry (MS) can be used to identify the binding epitopes of nanobodies to their protein antigen. HDX-MS has already been used for antibodies and nanobodies [219]. Using this approach, many nanobodies can be screened, and conformational changes of the antigen protein can also be detected in addition to the binding epitope.

From the biophysical characterisation, the Nb49 appeared to enhance the Unc5-GPC3 interaction. However, as the data show, it does not stabilise the Unc5-GPC3 complex enough for it to stay intact in solution. The Nb49-GPC3 complex crystal structure also shows that the nanobody blocks the formation of the octameric complex. For the detection and stabilisation of the octameric complex, new nanobodies could be generated. In the structure determination of the active β2 adrenergic receptor (β2AR) in complex with the G protein, Gs, a nanobody was used to stabilise the complex. For the generation of this nanobody, the β2AR-Gs interaction was stabilised with cross-linking before the llama immunisation [128]. In the case of the GPC3-Unc5 complex, cross-linking can be used to stabilise the octameric structure or other multimeric complexes that might be biologically relevant, as native mass spectrometry data showed that the complex could exist in

a dimeric, tetrameric or octameric stoichiometry [30]. The presence of the Nb49, which enhances the GPC3-Unc5 complex formation, is not compatible with the octameric structure of the complex (Chapter 4). However, when Nb49 is present, the GPC3-Unc5 complex affects the cell migration in stripe assays in agreement with the GPC3-Unc5 complex enhancement, which supports the idea that different stoichiometry complexes might have a biological role. New nanobodies can be raised against the different cross-linked complexes either by llama immunisation or nanobody libraries.

These new nanobodies can be useful tools for the stabilisation of the octameric complex in solution but also the detection of the octameric complex on the cell surface. Cell aggregation assays showed that the GPC3-Unc5 complex could be formed either *"in cis"* on the same cell or *"in trans"* between two different cells, but when the *"in cis"* interaction occurs, the *"in trans"* interaction is blocked [30]. Detection and structure determination of the GPC3-Unc5 complex on the cell surface using cryo-ET could reveal how this large protein complex is formed naturally. Nanobodies linked to the suitable label can be used to identify the protein complex in the crowded extracellular environment within tomograms [220].

7.2.3 Applications of the anti-GPC3 nanobodies

In the literature, nanobodies have been used in a wide variety of experiments and with numerous different purposes [119,120]. The two characterised nanobodies against hGPC3^{core} were used by Akkermans et al. for the characterisation of the GPC3-Unc5 interaction and the revelation of its role in cell migration [30]. The Unc5 guidance receptors have been extensively studied for their repulsive effects on cell migration in response to their other extracellular ligands, Netrin-1 and FLRT. It is well-established that Unc5D plays an essential role in radial cell migration during

cortical development [72,83,221]. Based on this, the effects of the Unc5-GPC3 interactions in radial migration were studied.

Pull-down experiments and immunostaining using Nb49 confirmed the presence of the GPC3 protein in the developing mouse cortex. GPC3 was highly enriched in mouse cortex lysate pull-down experiments using Nb49 followed by mass spectrometry. The nanobody detected GPC3 in the brain lysate and confirmed its presence at the protein level. Nb49 was also successfully used in immunostaining GPC3 in the mouse cortex. For this experiment, Nb49 was tetramerised with fluorescent streptavidin to increase its affinity and potency.

With *in-utero* electroporation, the cells in the germinal zones of the mouse brain were transfected with either Nb49 or Nb58, and the two nanobodies were expressed in the developing mouse cortex. In both cases, the radial migration was delayed. All the *in vivo* experiments in the developing mouse cortex using tools like the non-GPC3-binding Unc5 mutant, shRNA against GPC3 and the nanobodies showed that the GPC3-Unc5 interaction regulates the cortical migration [30].

Another model of cell migration, which was used for the characterisation of the GPC3-Unc5 interaction, was the neuroblastoma collective migration. In this model, apart from the correct migration to the target, the cellular cohesion of the migrating cells is also observed. Both nanobodies were expressed in the xenograft tumour cells. These neuroblastoma migration experiments showed that the strength of the GPC3-Unc5 interaction, affected by the nanobodies, is essential for the collective migration and cohesion of these tumour cells [30].

Together, these experiments confirm that the nanobodies could be useful tools for the understanding and interpretation of protein interactions, as indicated by the literature.

The nanobodies were employed in these studies to investigate the role of the GPC3-Unc5 interaction in cell migration [30]. However, the Unc5 receptors, except for their involvement in cell migration, also function as dependence receptors. In the absence of their ligand, Netrin-1, they can induce apoptosis [91]. This function links them with cancer. Their expression is also downregulated in several cancers [94]. Netrin-1, at the same time, is upregulated, which suggests that it might promote tumourigenesis [95–97]. GPC3, as well, is upregulated in many different malignancies, and it has been characterised as an oncofetal protein [25,34,46]. Experiments could be performed in order to study the effect of the GPC3-Unc5 interaction on cell survival. Does GPC3 act like Netrin-1 when it binds Unc5? Theoretically, Nb49 can stabilise the interaction and inhibit the Unc5 apoptotic signaling, while Nb58 can block the interaction and, in the absence of GPC3, Unc5 can induce apoptosis. If the interruption of the GPC3-Unc5 interaction leads to cell apoptosis, nanobodies that block the interaction can potentially be used in cancer treatment. An antibody that inhibits the Netrin-1/Unc5 interaction has already been developed, and it induced apoptosis in several cancer cell lines [101]. Similarly, Nb58 can inhibit the GPC3/Unc5 interaction and possibly induce apoptosis. The non-binding GPC3 and Unc5 mutants can also be utilised in these experiments.

Since GPC3 is upregulated in many different malignancies, it has been broadly studied for its role in cancer and, more extensively, in hepatocellular carcinoma [25,27,37,54]. Motivated by this, researchers have selected GPC3 as a diagnostic and therapeutic target in hepatocellular carcinoma, and antibodies against GPC3 have already been developed. The implication of GPC3 in the Wnt signaling pathway has also inspired the generation of antibodies that block the interaction of GPC3 with Wnt or its receptor Frizzled [68]. GPC3 is a popular target, and the use

of nanobodies instead of antibodies could be beneficial as they present advantages that antibodies lack. The nanobodies' favourable properties make them suitable tools for therapeutic and diagnostic applications. Their small size allows them to penetrate tissues that are difficult to access, like the microenvironment of solid tumours and the brain through the blood brain barrier [138].

The anti-GPC3 nanobodies characterised in this thesis can potentially be used in cancer diagnostics and therapeutics. For the nanobodies to be used for these purposes, their affinity to their antigen has to be very high (pM). Antibody affinity maturation is a critical step of the antibody development process. Similarly, affinity maturation of the nanobodies is often essential. Many times, nanobodies present high affinity to their target without further alterations. However, affinity improvement might be necessary. If a high-resolution structure of the nanobody-antigen complex is available, site-directed mutagenesis based on the structure can be used to improve the affinity [222]. If the structure is not available, targeted random mutagenesis on the CDR3 loop can be used. The CDR3 loop contributes the most to the binding of nanobodies to their epitope, and its mutagenesis can bring the desired outcome [223]. The output of all the affinity maturation approaches needs screening and selection to identify the affinity-improved nanobodies. After the selection, each nanobody's function must be re-evaluated. Before the affinity maturation, nanobodies might have been characterised for a specific function and, even with site-directed mutagenesis, this function could have changed. For example, if the Nb49's affinity to GPC3 is improved after affinity maturation, it is not implied that it still enhances the GPC3-Unc5 interaction. Alternatively, the multimerisation of nanobodies can lead to increased affinity to their target due to avidity effects [124,224]. In this case, a multivalent nanobody multimer of Nb49 and

Nb58 can provide a high-affinity bivalent nanobody. Work in this project already utilised tetrameric nanobodies using streptavidin as a linker.

Tools generated in this thesis for the characterisation and investigation of the functions of the GPC3-Unc5 interaction were successfully used in various *in vitro* and *in vivo* experiments. Since they were proven useful, they can also be used in future experiments to further understand the GPC3-Unc5 interaction. The nanobodies can also be used in work related to the exploration of other GPC3 functions, or its related applications in cancer.

Appendix A

Primers used in this work

Table A1: Summary of the primers for the constructs cloned in this work (Chapters 3-5).

Construct	Residues	Vector		Primer
hGPC3$^{full-length}$	1 – 580	pHLsec EcoRI-KpnI	Fwd	TCTCAGGCCGAATTCATGGCCGGGACCGTGCGCACCGCGTG
			Rev	GTGGTGCTTGGTACCTCA GTGCACCAGGAAGAAGAAGCACACC
Mutation K1	Q112N A114T		Fwd_{mut}	GAATGCTGCGGTTTTCAATGAGACCTTTGAAATTGTTGTTCGCCATGCCAAG
			Rev_{mut}	GAAAACCGCAGCATTCTGAATAATTAAG
Mutation K2	L157N		Fwd_{mut}	GTGTCTCTCTACATCAACGGTTCTGACATCAATGTAGATGACATG
			Rev_{mut}	GATGTAGAGAGACACATCTGTGAAAAATTC
Mutation K3	N241Q		Fwd_{mut}	GGAATTGAAGTGATCCAGACAACTGATCACCTGAAGTTCAGTAAG
			Rev_{mut}	GATCACTTCAATTCCAAGATTCAGAG
mUnc5B$^{full-length}$	26 – 934	pHLsec AgeI-KpnI	Fwd	GTAGCTGAAACCGGTTACCCATACGATGTTCCAGATTACG
			Rev	CACCTCCAGGGTACCGCAATCGCCATCTGTGGCCATGGCTACC
Mutation U6	Y150N R152T		Fwd_{mut}	TACATCCGCATTGCCAACTTGACGAAGAACTTTGACCAGGAGCCTCTGG
			Rev_{mut}	GGCAATGCGGATGTAGGCTCGGCGAC
Mutation M7	A277N L279T		Fwd_{mut}	GACCTGCACCAATCCAAACCCAACCAATGGAGGCGCCTTCTGTGAGGGAC
			Rev_{mut}	TGGATTGGTGCAGGTCCGAGTACGC
Mutation K4	I211N R213T		Fwd_{mut}	CATTGACCACAACCTCAACATCACCCAGGCGCGCCTCTCAGACACGGCCAAC
			Rev_{mut}	GAGGTTGTGGTCAATGGTGAGCAGG
Nb49	23 - 162	pHLsec	Fwd	GTAGCTGAAACCGGTCAGGTGCAGCTGGTCGAGTCTGGGG
Nb58	23 - 163	AgeI-KpnI	Rev	GGAACCTCCGGTACCTTGGCCTCCCGGGCCGGCCGCTGGTTG
Nb49^{short}	1-146	pADL-23c	Fwd	CCTATTGCCTACGGCGGCCGCTGGATTGTTATTACTCGCGGCCCAG
			Rev	GATGGTGGTGATGGTGCTTGGTACCCGAGGAGACGGTGACCTGGGTCCCC

Table A2: Summary of the primers for the constructs cloned in this work (Chapter 6)

Constructs Mouse VLGR1b		Residues	Vector		Primer
#1	Fragment 1	29-3576	pHLsec AgeI-KpnI His_6	Fwd	GTAGCTGAAACCGGTGAGGCCGAGATCAGATTCACCGGCC
				Rev	GCTGATATTCACCAGTCTCAGG
	Fragment 2			Fwd	CTGGTGAATATCAGCGGACCAGGAGGACAGCTGGCAGAGAC
				Rev	CTCCTCCACGGCCCTGGCCACCAC
	Fragment 3			Fwd	CAGGGCCGTGGAGGAGGGCGAGGATGTGCTGAGCTACTATG
				Rev	GTGGTGCTTGGTACCTGTCAGCTCATACAGGTGGGAGTCG
#2		31-616		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GTGGTGCTTGGTACCCTCCAGCTGCACCAGAAAGTGGGCG
#3		31-1208		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GTGGTGCTTGGTACCCACCAGTCTCAGGATATAAAACTCG
#4		1261-2320		Fwd	GTAGCTGAAACCGGTTACATCACCAGCTTCACAATCCTGAG
				Rev	GTGGTGCTTGGTACCAGCCAGCTGCACCTGCACCACCTCC
#5		1223-1542		Fwd	GTAGCTGAAACCGGTGTGCTGTATTTCTCTGGACTGGAGG
				Rev	GTGGTGCTTGGTACCCACCAGCTTCAGCAGAAACAGCTCC
#6		1223-1562		Fwd	GTAGCTGAAACCGGTGTGCTGTATTTCTCTGGACTGGAGG
				Rev	GTGGTGCTTGGTACCCTGGATTCTCAGCCGGGCTGTGGTG
#7		2216-3044		Fwd	GTAGCTGAAACCGGTGAGGCCGTGATCACAATCGAGGCCAG
				Rev	GTGGTGCTTGGTACC GGTCAGCAGCAGCTGGAACCGCTCG
#8		3249-3529		Fwd	GTAGCTGAAACCGGTGATGTGGTGTTCAGCATCTTTCAGTC
				Rev	GTGGTGCTTGGTACCGTTCCAACAGTACAGGGCGCTGCCG
#9		3195-3575		Fwd	GTAGCTGAAACCGGTCTGTTTTCTATCAGCGCCGTGGAG
				Rev	GTGGTGCTTGGTACCCAGCTCATACAGGTGGGAGTCGGTTG
#7		2216-3044	pHLsec AgeI-KpnI C-term His_{10}	Fwd	GTAGCTGAAACCGGTGAGGCCGTGATCACAATCGAGGCCAG
				Rev	GATACTAGTCTCGAGTGATCATTAATGGTGGTGATGGTGATGGTGATGGTGGTG

(continued on the next page)

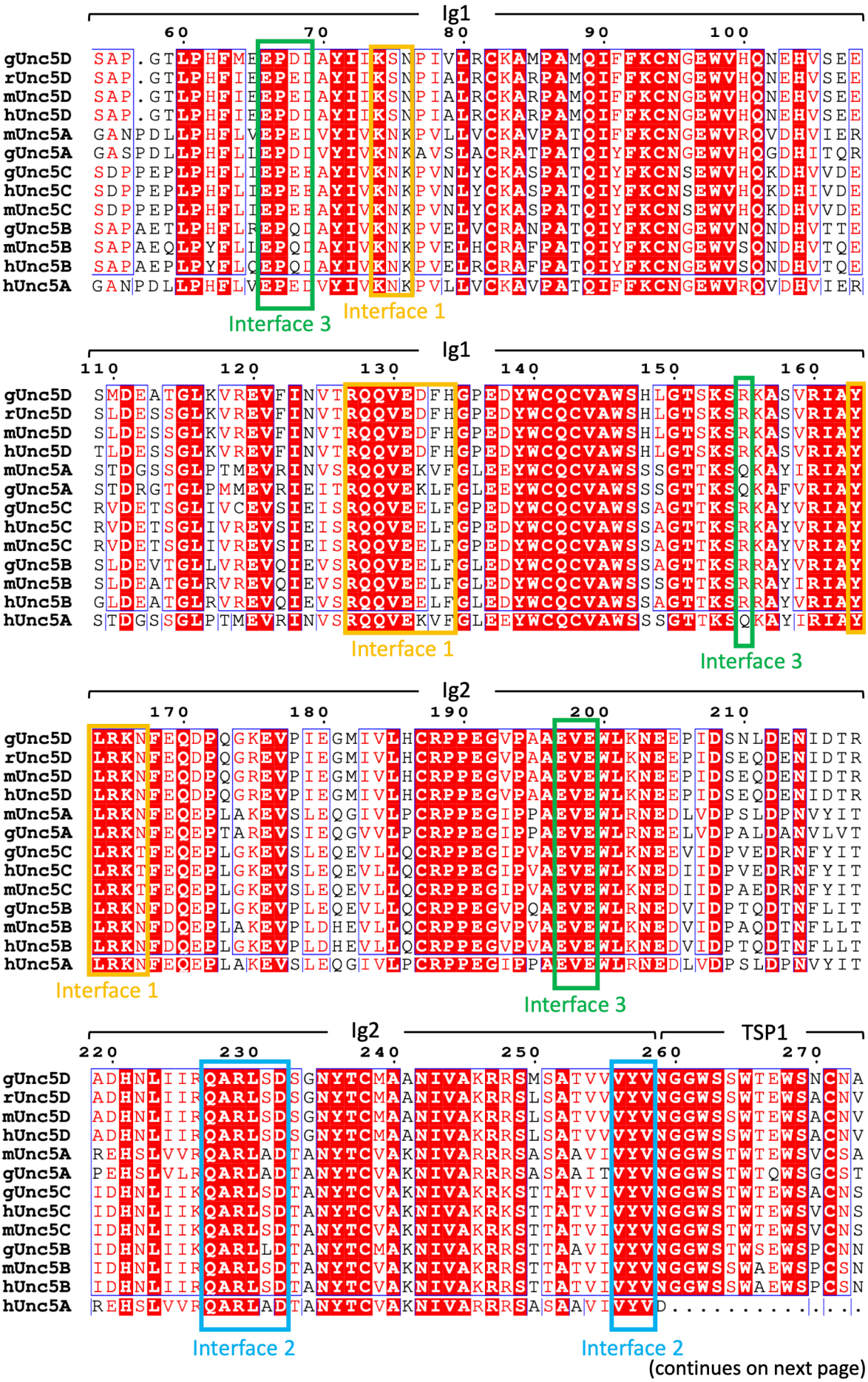
Constructs Mouse VLGR1b		Residues	Vector		Primer
#7		2216-3044	pHLsec AgeI-KpnI N-term His_{10}	Fwd	GATGGGTTGCGTAGCTCACCACCATCACCATCACCATCACCACCATGAAACCGGT GAGGCCGTGATCACAATC
				Rev	GATGGTGGTGCTTGGTACCTCATTAGGTCAGCAGCAGCTGGAACCGCTCGTCG
#10		147-383	pHLsec AgeI-KpnI C-term His_{10}	Fwd	GTAGCTGAAACCGGTTCCTTTTCTACACCAAGCTCCATCTC
				Rev	GTGGTGCTTGGTACCCTTGATTGTCACCTGCACGGTGGATG
#11		31-510		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GTGGTGCTTGGTACCCTGGATGTAGAACAGCAGCTGGGCAG
#12		31-1228		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GTGGTGCTTGGTACCGATCATCACGGTCACCTGGAAATTTG
#13		31-1113		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GTGGTGCTTGGTACCCTCGATGATCACGGTGGCCACGCCATAC
#14	Fragment 1	31-1966		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GCTGATATTCACCAGTCTCAGG
	Fragment 2			Fwd	CTGGTGAATATCAGCGGACCAGGAGGACAGCTGGCAGAGAC
				Rev	GTGGTGCTTGGTACCGCTGGCCAGGATTGTCAGGGTGGCGTTG
#15	Fragment 1	31-2338		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GCTGATATTCACCAGTCTCAGG
	Fragment 2			Fwd	CTGGTGAATATCAGCGGACCAGGAGGACAGCTGGCAGAGAC
				Rev	GTGGTGCTTGGTACCGATCACGATATTGGCCACCCTATCC
#16	Fragment 1	31-3063		Fwd	GTAGCTGAAACCGGTGAGATCAGATTCACCGGCCAGACAG
				Rev	GCTGATATTCACCAGTCTCAGG
	Fragment 2			Fwd	CTGGTGAATATCAGCGGACCAGGAGGACAGCTGGCAGAGAC
				Rev	GTGGTGCTTGGTACCCAGCACGGTGATCAGGGCGATTGTATTC
#17		1701-1966		Fwd	GTAGCTGAAACCGGTCTGATGCAGTTTTCTACCGGCCTGC
				Rev	GTGGTGCTTGGTACCGCTGGCCAGGATTGTCAGGGTGGCGTTG
#18		1701-2338		Fwd	GTAGCTGAAACCGGTCTGATGCAGTTTTCTACCGGCCTGC
				Rev	GTGGTGCTTGGTACCGATCACGATATTGGCCACCCTATCC

(continued on the next page)

Constructs Mouse VLGR1b	Residues	Vector	Primer	
#19	1701-3063	pHLsec AgeI-KpnI C-term His_{10}	Fwd	GTAGCTGAAACCGGTCTGATGCAGTTTTCTACCGGCCTGC
			Rev	GTGGTGCTTGGTACCCAGCACGGTGATCAGGGCGATTGTATTC
#20	2957-3641		Fwd	GTAGCTGAAACCGGTATGATCGAGTGGCAGAGGAACCGCTTC
			Rev	GTGGTGCTTGGTACCCACTGTCACCCGCACGCTAGAATTG
#21	3650-4021		Fwd	GTAGCTGAAACCGGTGTGGTGGCCTTTGCCCAGAATAGCC
			Rev	GTGGTGCTTGGTACCGATGACCACGTTCACGCTCACGTCATC
#22	5607-5896		Fwd	CACCATGAAACCGGTACAGCCAACATCACCCTGATCTCCG
			Rev	GATCATTACTTGGTACCTCATTAGGAGCTGTTGTCGGTCTGGGCATAC

Appendix B

Unc5 sequences alignment



(continues on next page)

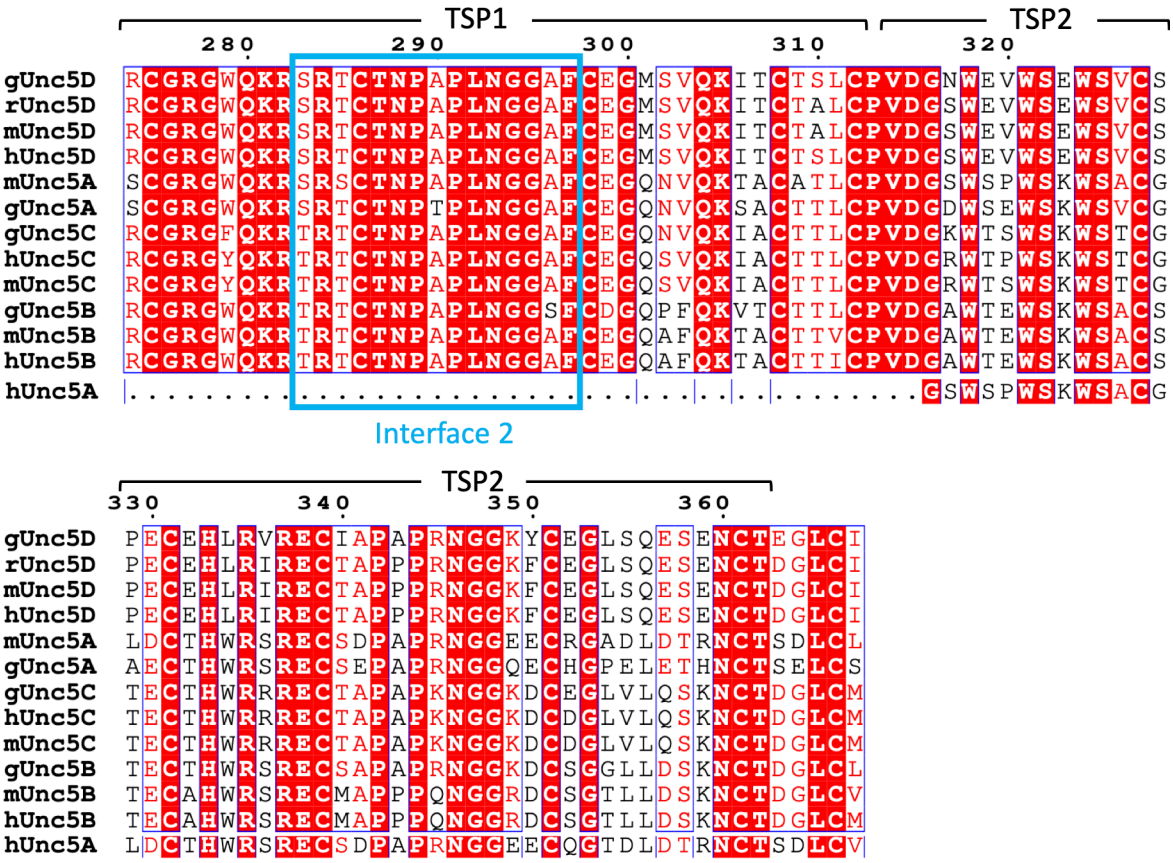


Figure B1: Unc5 ectodomain sequences alignment
Sequence alignment of the extracellular domains of the Unc5 receptors homologues A, B, C and D from different species. h: homo sapiens, m: mus musculus, g: gallus gallus, r: rattus norvegicus. The different extracellular domains (Ig1, Ig2, TSP1, TSP2) are indicated. The binding interfaces with GPC3 are also indicated. The sequences were aligned with Clustal Omega [225] and Espript [226] was used to visualise the alignment.

Appendix C

GPC3 sequences alignment

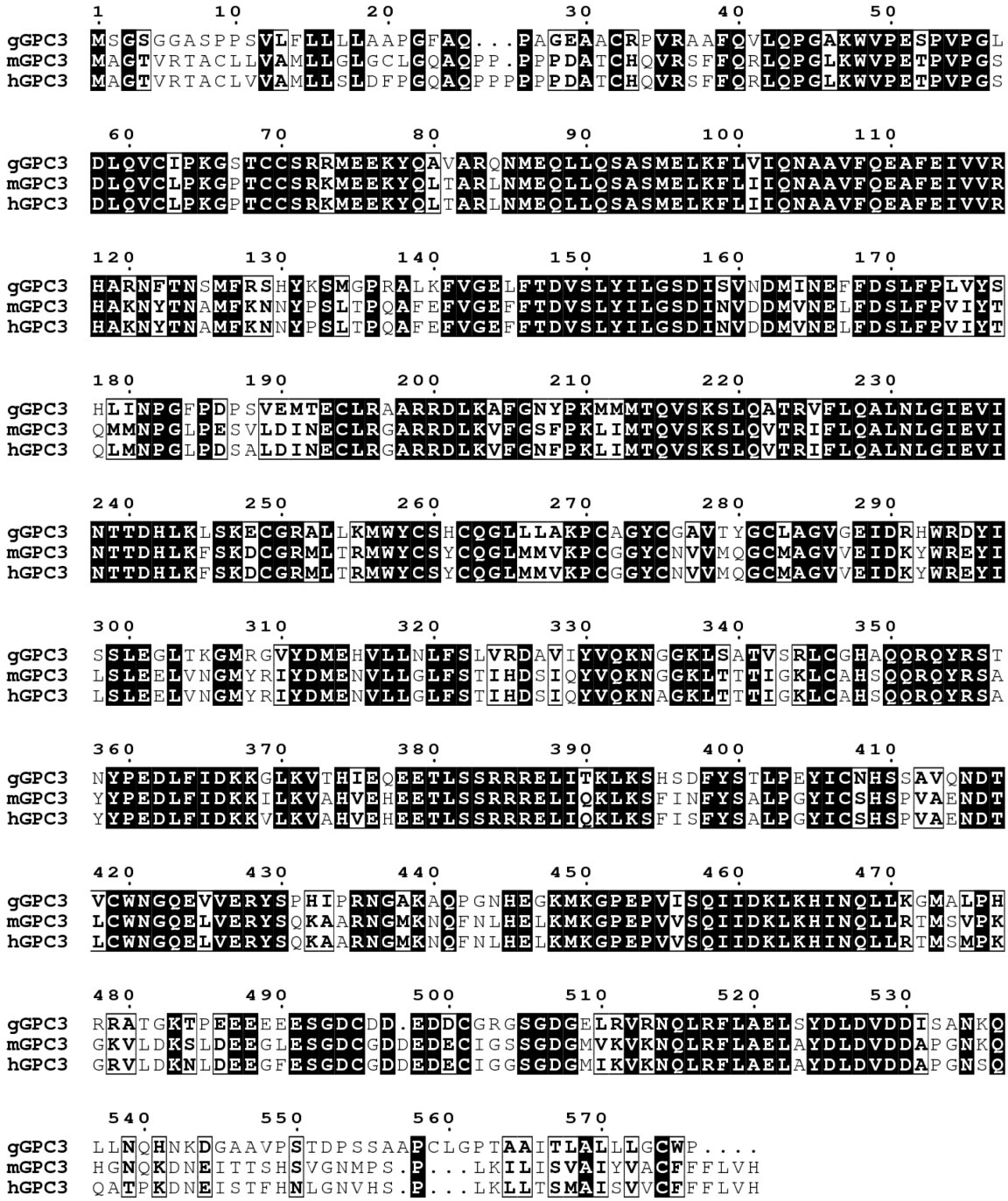


Figure C1: GPC3 sequences alignment
Sequence alignment of GPC3 sequences from different species. h: homo sapiens, m: mus musculus, g: gallus gallus. The sequences were aligned with Clustal Omega [225] and Espript [226] was used to visualise the alignment.

Appendix D

Nanobody CDR3 sequence alignment and classification

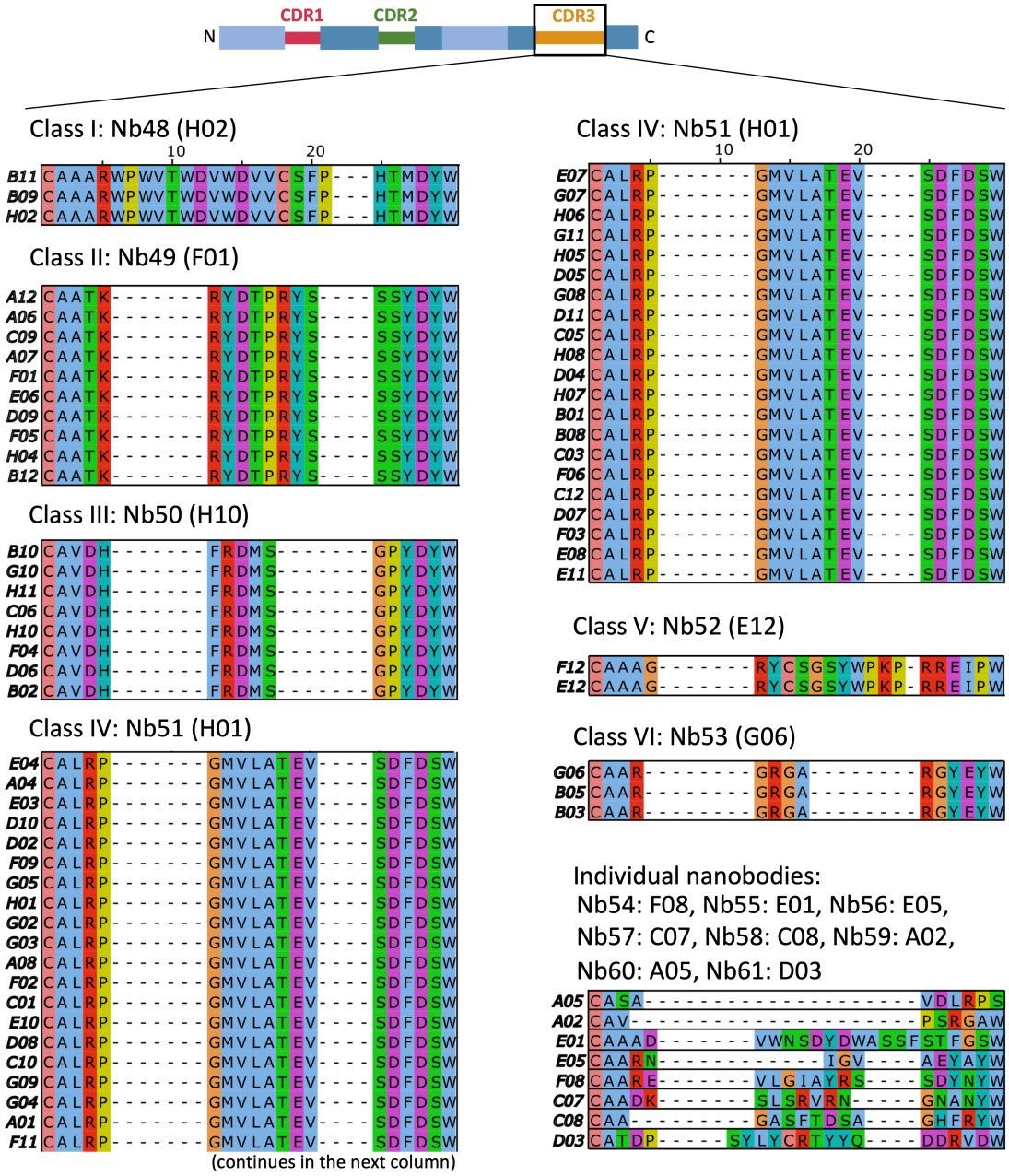


Figure D1: Nanobodies CDR3 alignment
The available nanobody sequences (96-well plate) were grouped into classes based on their CDR3 loop sequence identity. Since they were generated from a llama after the immunization of the animal with the hGPC3^{core} protein, the framework sequence and the CDR1 and CDR2 loops are not identical, but they are similar enough in order to use only the CDR3 loop sequence identity for their classification. The CDR3 loops play the main role in the nanobody binding to their binding epitope. The nanobodies that did not show enough binding at the initial ELISA were excluded from the classification. Jalview was used for the generation of these alignments [227].

References

[1] Ridley, A.J., Schwartz, M.A., Burridge, K., Firtel, R.A., Ginsberg, M.H., Borisy, G., Parsons, J.T., & Horwitz, A.R. (2003), Cell Migration: Integrating Signals from Front to Back. *Science*. 302 (5651) : 1704–1709.

[2] Friedl, P. & Gilmour, D. (2009), Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology*. 10 (7) : 445–457.

[3] Subramanian, L., Calcagnotto, M.E., & Paredes, M.F. (2020), Cortical Malformations: Lessons in Human Brain Development. *Frontiers in Cellular Neuroscience*. 13 : 576.

[4] Fukuda, T. & Yanagi, S. (2017), Psychiatric behaviors associated with cytoskeletal defects in radial neuronal migration. *Cellular and Molecular Life Sciences*. 74 (19) : 3533–3552.

[5] Peregrina, C. & del Toro, D. (2020), FLRTing Neurons in Cortical Migration During Cerebral Cortex Development. *Frontiers in Cell and Developmental Biology*. 8 (578506).

[6] Nadarajah, B., Brunstrom, J.E., Grutzendler, J., Wong, R.O.L., & Pearlman, A.L. (2001), Two modes of radial migration in early development of the cerebral cortex. *Nature Neuroscience*. 4 (2) : 143–150.

[7] Tabata, H. & Nakajima, K. (2003), Multipolar Migration: The Third Mode of Radial Neuronal Migration in the Developing Cerebral Cortex. *Journal of Neuroscience*. 23 (31) : 9996–10001.

[8] Hirota, Y. & Nakajima, K. (2017), Control of Neuronal Migration and Aggregation by Reelin Signaling in the Developing Cerebral Cortex. *Frontiers in Cell and Developmental Biology*. 5.

[9] Arai, Y. & Taverna, E. (2017), Neural progenitor cell polarity and cortical development. *Frontiers in Cellular Neuroscience*. 11 (384).

[10] Kolodkin, A.L. & Tessier-Lavigne, M. (2011), Mechanisms and molecules of neuronal wiring: A primer. *Cold Spring Harbor Perspectives in Biology*. 3 (a001727).

[11] Moreland, T. & Poulain, F.E. (2022), To Stick or Not to Stick: The Multiple Roles of Cell Adhesion Molecules in Neural Circuit Assembly. *Frontiers in Neuroscience*. 16 (889155).

[12] Leung, L.C., Urbančič, V., Baudet, M.L., Dwivedy, A., Bayley, T.G., Lee, A.C., Harris, W.A., & Holt, C.E. (2013), Coupling of NF-protocadherin signaling to

axon guidance by cue-induced translation. *Nature Neuroscience*. 16 (2) : 166–173.

[13] Brose, K., Bland, K.S., Kuan, H.W., Arnott, D., Henzel, W., Goodman, C.S., Tessier-Lavigne, M., & Kidd, T. (1999), Slit proteins bind robo receptors and have an evolutionarily conserved role in repulsive axon guidance. *Cell*. 96 (6) : 795–806.

[14] Kidd, T., Bland, K.S., & Goodman, C.S. (1999), Slit is the midline repellent for the Robo receptor in Drosophila. *Cell*. 96 (6) : 785–794.

[15] Lyuksyutova, A.I., Lu, C.C., Milanesio, N., King, L.A., Guo, N., Wang, Y., Nathans, J., Tessier-Lavigne, M., & Zou, Y. (2003), Anterior-Posterior Guidance of Commissural Axons by Wnt-Frizzled Signaling. *Science*. 302 (5652) : 1984–1988.

[16] Filmus, J., Capurro, M., & Rast, J. (2008), Glypicans. *Genome Biology*. 9 (5) : 224.

[17] Veugelers, M., De Cat, B., Ceulemans, H., Bruystens, A.M., Coomans, C., Dürr, J., Vermeesch, J., Marynen, P., & David, G. (1999), Glypican-6, a new member of the glypican family of cell surface heparan sulfate proteoglycans. *Journal of Biological Chemistry*. 274 (38) : 26968–26977.

[18] Nakato, H., Futch, T.A., & Selleck, S.B. (1995), The division abnormally delayed (daily) gene: A putative integral membrane proteoglycan required far cell division patterning during postembryonic development of the nervous system in Drosophila. *Development*. 121 (11) : 3687–3702.

[19] Khare, N. & Baumgartner, S. (2000), Dally-like protein, a new Drosophila glypican with expression overlapping with wingless. *Mechanisms of Development*. 99 : 199–202.

[20] Saunders, S., Paine-Saunders, S., & Lander, A.D. (1997), Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb, and brain. *Developmental Biology*. 190 (1) : 78–93.

[21] De Cat, B., Muyldermans, S.Y., Coomans, C., Degeest, G., Vanderschueren, B., Creemers, J., Biemar, F., Peers, B., & David, G. (2003), Processing by proprotein convertases is required for glypican-3 modulation of cell survival, Wnt signaling, and gastrulation movements. *Journal of Cell Biology*. 163 (3) : 625–635.

[22] Svensson, G., Awad, W., Håkansson, M., Mani, K., & Logan, D.T. (2012), Crystal structure of N-glycosylated human glypican-1 core protein: Structure of two loops evolutionarily conserved in vertebrate glypican-1. *Journal of Biological Chemistry*. 287 (17) : 14040–14051.

[23] Awad, W., Adamczyk, B., Örnros, J., Karlsson, N.G., Mani, K., & Logan, D.T. (2015), Structural aspects of N-glycosylations and the C-terminal region in human glypican-1. *Journal of Biological Chemistry*. 290 (38) : 22991–23008.

[24] Kim, M.S., Saunders, A.M., Hamaoka, B.Y., Beachy, P.A., & Leahya, D.J. (2011), Structure of the protein core of the glypican Dally-like and localization of a region important for hedgehog signaling. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (32) : 13112–13117.

[25] Capurro, M., Wanless, I.R., Sherman, M., Deboer, G., Shi, W., Miyoshi, E., & Filmus, J. (2003), Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. *Gastroenterology*. 125 (1) : 89–97.

[26] Hippo, Y., Watanabe, K., Watanabe, A., Midorikawa, Y., Yamamoto, S., Ihara, S., Tokita, S., Iwanari, H., Ito, Y., Nakano, K., Nezu, J.I., Tsunoda, H., Yoshino, T., Ohizumi, I., Tsuchiya, M., Ohnishi, S., Makuuchi, M., Hamakubo, T., Kodama, T., & Aburatani, H. (2004), Identification of Soluble NH2-Terminal Fragment of Glypican-3 as a Serological Marker for Early-Stage Hepatocellular Carcinoma. *Cancer Research*. 64 (7) : 2418–2423.

[27] Capurro, M.I., Xiang, Y.Y., Lobe, C., & Filmus, J. (2005), Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical Wnt signaling. *Cancer Research*. 65 (14) : 6245–6254.

[28] Traister, A., Shi, W., & Filmus, J. (2008), Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface. *Biochemical Journal*. 410 (3) : 503–511.

[29] Kakugawa, S., Langton, P.F., Zebisch, M., Howell, S.A., Chang, T.H., Liu, Y., Feizi, T., Bineva, G., O'Reilly, N., Snijders, A.P., Jones, E.Y., & Vincent, J.P. (2015), Notum deacylates Wnts to suppress signalling activity. *Nature*. 519 (7542) : 187.

[30] Akkermans, O., Delloye-Bourgeois, C., Peregrina, C., Carrasquero-Ordaz, M., Kokolaki, M., Berbeira-Santana, M., Chavent, M., Reynaud, F., Raj, R., Agirre, J., Aksu, M., White, E.S., Lowe, E., Ben Amar, D., Zaballa, S., Huo, J., Pakos, I., McCubbin, P.T.N., Comoletti, D., Owens, R.J., Robinson, C.D., Castellani, V., del Toro, D., & Seiradake E. (2022), GPC3-Unc5 receptor complex structure and role in cell migration. *Cell*. 185 : 3931–3949.

[31] Iglesias, B. V., Centeno, G., Pascuccelli, H., Ward, F., Peters, M.G., Filmus, J., Puricelli, L., & de Kier Joffé, E.B. (2008), Expression pattern of glypican-3 (GPC3) during human embryonic and fetal development. *Histology and Histopathology*. 23 (11) : 1333–1340.

[32] Hsu, H.C., Cheng, W., & Lai, P.L. (1997), Cloning and expression of a developmentally regulated transcript MXR7 in hepatocellular carcinoma:

Biological significance and temporospatial distribution. *Cancer Research*. 57 (22) : 5179–5184.

[33] Saikali, Z. & Sinnett, D. (2000), Expression of glypican 3 (GPC3) in embryonal tumors. *International Journal of Cancer*. 89 (5) : 418–422.

[34] Ortiz, M. V., Roberts, S.S., Bender, J.G., Shukla, N., & Wexler, L.H. (2019), Immunotherapeutic targeting of GPC3 in pediatric solid embryonal tumors. *Frontiers in Oncology*. 9 (108).

[35] Pilia, G., Hughes-Benzie, R.M., MacKenzie, A., Baybayan, P., Chen, E.Y., Huber, R., Neri, G., Cao, A., Forabosco, A., & Schlessinger, D. (1996), Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nature Genetics*. 12 (3) : 241–247.

[36] Pellegrini, M., Pilia, G., Pantano, S., Lucchini, F., Uda, M., Fumi, M., Cao, A., Schlessinger, D., & Forabosco, A. (1998), Gpc3 expression correlates with the phenotype of the Simpson-Golabi- Behmel syndrome. *Developmental Dynamics*. 213 (4) : 431–439.

[37] Capurro, M., Martin, T., Shi, W., & Filmus, J. (2014), Glypican-3 binds to Frizzled and plays a direct role in the stimulation of canonical Wnt signaling. *Journal of Cell Science*. 127 (7) : 1565–1575.

[38] Filmus, J. & Capurro, M. (2014), The role of glypicans in Hedgehog signaling. *Matrix Biology*. 35 : 248–252.

[39] Dwivedi, P.P., Grose, R.H., Filmus, J., Hii, C.S.T., Xian, C.J., Anderson, P.J., & Powell, B.C. (2013), Regulation of bone morphogenetic protein signalling and cranial osteogenesis by Gpc1 and Gpc3. *Bone*. 55 (2) : 367–376.

[40] Capurro, M.I., Shi, W., Sandal, S., & Filmus, J. (2005), Processing by convertases is not required for glypican-3-induced stimulation of hepatocellular carcinoma growth. *The Journal of Biological Chemistry*. 280 (50) : 41201–41206.

[41] Yan, D., Wu, Y., Feng, Y., Lin, S.C., & Lin, X. (2009), The Core Protein of Glypican Dally-Like Determines Its Biphasic Activity in Wingless Morphogen Signaling. *Developmental Cell*. 17 (4) : 470–481.

[42] Capurro, M.I., Xu, P., Shi, W., Li, F., Jia, A., & Filmus, J. (2008), Glypican-3 Inhibits Hedgehog Signaling during Development by Competing with Patched for Hedgehog Binding. *Developmental Cell*. 14 (5) : 700–711.

[43] Capurro, M.I., Shi, W., & Filmus, J. (2012), LRP1 mediates Hedgehog-induced endocytosis of the GPC3-Hedgehog complex. *Journal of Cell Science*. 125 (14) : 3380–3389.

[44] Capurro, M.I., Li, F., & Filmus, J. (2009), Overgrowth of a mouse model of Simpson- Golabi-Behmel syndrome is partly mediated by Indian Hedgehog. *EMBO Reports*. 10 (8) : 901–907.

[45] Capurro, M., Shi, W., Izumikawa, T., Kitagawa, H., & Filmus, J. (2015), Processing by convertases is required for glypican-3-induced inhibition of Hedgehog signaling. *Journal of Biological Chemistry*. 290 (12) : 7576–7585.

[46] Aviel-Ronen, S., Lau, S.K., Pintilie, M., Lau, D., Liu, N., Tsao, M.S., & Jothy, S. (2008), Glypican-3 is overexpressed in lung squamous cell carcinoma, but not in adenocarcinoma. *Modern Pathology*. 21 (7) : 817–825.

[47] Xiang, Y.Y., Ladeda, V., & Filmus, J. (2001), Glypican-3 expression is silenced in human breast cancer. *Oncogene*. 20 (50) : 7408–7412.

[48] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., & Jemal, A. (2018), Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 68 (6) : 394–424.

[49] Allen, R. & Li, D. (2019), Hepatocellular carcinoma. *Oncology in the Precision Medicine Era: Value-Based Medicine*. 380 (15) : 259–271.

[50] Singal, A.G., Pillai, A., & Tiro, J. (2014), Early Detection, Curative Treatment, and Survival Rates for Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis: A Meta-analysis. *PLoS Medicine*. 11 (4) : e1001624.

[51] Chen, M., Li, G., Yan, J., Lu, X., Cui, J., Ni, Z., Cheng, W., Qian, G., Zhang, J., & Tu, H. (2013), Reevaluation of glypican-3 as a serological marker for hepatocellular carcinoma. *Clinica Chimica Acta*. 423 : 105–111.

[52] Shirakawa, H., Suzuki, H., Shimomura, M., Kojima, M., Gotohda, N., Takahashi, S., Nakagohri, T., Konishi, M., Kobayashi, N., Kinoshita, T., & Nakatsura, T. (2009), Glypican-3 expression is correlated with poor prognosis in hepatocellular carcinoma. *Cancer Science*. 100 (8) : 1403–1407.

[53] Bengochea, A., De Souza, M.M., Lefrançois, L., Le Roux, E., Galy, O., Chemin, I., Kim, M., Wands, J.R., Trepo, C., Hainaut, P., Scoazec, J.Y., Vitvitski, L., & Merle, P. (2008), Common dysregulation of Wnt/Frizzled receptor elements in human hepatocellular carcinoma. *British Journal of Cancer*. 99 (1) : 143–150.

[54] Li, N., Wei, L., Liu, X., Bai, H., Ye, Y., Li, D., Li, N., Baxa, U., Wang, Q., Lv, L., Chen, Y., Feng, M., Lee, B., Gao, W., & Ho, M. (2019), A Frizzled-Like Cysteine-Rich Domain in Glypican-3 Mediates Wnt Binding and Regulates Hepatocellular Carcinoma Tumor Growth in Mice. *Hepatology*. 70 (4) : 1231–1245.

[55] Saad, A., Liet, B., Joucla, G., Santarelli, X., Charpentier, J., Claverol, S., Grosset, C.F., & Trézéguet, V. (2018), Role of Glycanation and Convertase Maturation of Soluble Glypican-3 in Inhibiting Proliferation of Hepatocellular Carcinoma Cells. *Biochemistry*. 57 (7) : 1201–1211.

[56] Zittermann, S.I., Capurro, M.I., Shi, W., & Filmus, J. (2010), Soluble glypican 3 inhibits the growth of hepatocellular carcinoma in vitro and in vivo. *International Journal of Cancer*. 126 (6) : 1291–1301.

[57] Wang, D., Gao, Y., Zhang, Y., Wang, L., & Chen, G. (2019), Glypican-3 promotes cell proliferation and tumorigenesis through up-regulation of β-catenin expression in lung squamous cell carcinoma. *Bioscience Reports*. 39 (6).

[58] Meng, P., Zhang, Y.F., Zhang, W., Chen, X., Xu, T., Hu, S., Liang, X., Feng, M., Yang, X., & Ho, M. (2021), Identification of the atypical cadherin FAT1 as a novel glypican-3 interacting protein in liver cancer cells. *Scientific Reports*. 11 (1) : 1–12.

[59] Ishiguro, T., Sugimoto, M., Kinoshita, Y., Miyazaki, Y., Nakano, K., Tsunoda, H., Sugo, I., Ohizumi, I., Aburatani, H., Hamakubo, T., Kodama, T., Tsuchiya, M., & Yamada-Okabe, H. (2008), Anti-glypican 3 antibody as a potential antitumor agent for human liver cancer. *Cancer Research*. 68 (23) : 9832–9838.

[60] Nakano, K., Orita, T., Nezu, J., Yoshino, T., Ohizumi, I., Sugimoto, M., Furugaki, K., Kinoshita, Y., Ishiguro, T., Hamakubo, T., Kodama, T., Aburatani, H., Yamada-Okabe, H., & Tsuchiya, M. (2009), Anti-glypican 3 antibodies cause ADCC against human hepatocellular carcinoma cells. *Biochemical and Biophysical Research Communications*. 378 (2) : 279–284.

[61] Nakano, K., Ishiguro, T., Konishi, H., Tanaka, M., Sugimoto, M., Sugo, I., Igawa, T., Tsunoda, H., Kinoshita, Y., Habu, K., Orita, T., Tsuchiya, M., Hattori, K., & Yamada-Okabe, H. (2010), Generation of a humanized anti-glypican 3 antibody by CDR grafting and stability optimization. *Anti-Cancer Drugs*. 21 (10) : 907–916.

[62] Zhu, A.X., Gold, P.J., El-Khoueiry, A.B., Abrams, T.A., Morikawa, H., Ohishi, N., Ohtomo, T., & Philip, P.A. (2013), First-in-man phase i study of GC33, a novel recombinant humanized antibody against glypican-3, in patients with advanced hepatocellular carcinoma. *Clinical Cancer Research*. 19 (4) : 920–928.

[63] Ikeda, M., Ohkawa, S., Okusaka, T., Mitsunaga, S., Kobayashi, S., Morizane, C., Suzuki, I., Yamamoto, S., & Furuse, J. (2014), Japanese phase I study of GC33, a humanized antibody against glypican-3 for advanced hepatocellular carcinoma. *Cancer Science*. 105 (4) : 455–462.

[64] Abou-Alfa, G.K., Puig, O., Daniele, B., Kudo, M., Merle, P., Park, J.W., Ross, P., Peron, J.M., Ebert, O., Chan, S., Poon, T.P., Colombo, M., Okusaka, T., Ryoo, B.Y., Minguez, B., Tanaka, T., Ohtomo, T., Ukrainskyj, S., Boisserie, F., Rutman, O., Chen, Y.C., Xu, C., Shochat, E., Jukofsky, L., Reis, B., Chen, G., Di Laurenzio, L., Lee, R., & Yen, C.J. (2016), Randomized phase II placebo controlled study of codrituzumab in previously treated patients with advanced hepatocellular carcinoma. *Journal of Hepatology*. 65 (2) : 289–295.

[65] Phung, Y., Gao, W., Man, Y.G., Nagata, S., & Ho, M. (2012), High-affinity monoclonal antibodies to cell surface tumor antigen glypican-3 generated through a combination of peptide immunization and flow cytometry screening. *MAbs*. 4 (5) : 592–599.

[66] Gao, W., Kim, H., Feng, M., Phung, Y., Xavier, C.P., Rubin, J.S., & Ho, M. (2014), Inactivation of Wnt signaling by a human antibody that recognizes the heparan sulfate chains of glypican-3 for liver cancer therapy. *Hepatology*. 60 (2) : 576–587.

[67] Feng, M., Gao, W., Wang, R., Chen, W., Man, Y.G., Figg, W.D., Wang, X.W., Dimitrov, D.S., & Ho, M. (2013), Therapeutically targeting glypican-3 via a conformation-specific single-domain antibody in hepatocellular carcinoma. *PNAS*. 110 (12) : E1083–E1091.

[68] Zheng, X., Liu, X., Lei, Y., Wang, G., & Liu, M. (2022), Glypican-3: A Novel and Promising Target for the Treatment of Hepatocellular Carcinoma. *Frontiers in Oncology*. 12 (824208).

[69] Wang, W., Xu, C., Wang, H., & Jiang, C. (2021), Identification of nanobodies against hepatocellular carcinoma marker glypican-3. *Molecular Immunology*. 131 : 13–22.

[70] Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L., & Tessier-Lavigne, M. (1997), Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors. *Nature*. 386 : 833–838.

[71] Karaulanov, E., Böttcher, R.T., Stannek, P., Wu, W., Rau, M., Ogata, S., Cho, K.W.Y., & Niehrs, C. (2009), Unc5B Interacts with FLRT3 and Rnd1 to Modulate Cell Adhesion in Xenopus Embryos. *PLoS ONE*. 4 (5) : e5742.

[72] Seiradake, E., delToro, D., Nagel, D., Cop, F., Härtl, R., Ruff, T., Seyit-Bremer, G., Harlos, K., Border, E.C., Acker-Palmer, A., Jones, E.Y., & Klein, R. (2014), FLRT Structure: Balancing Repulsion and Cell Adhesion in Cortical and Vascular Development. *Neuron*. 84 (2) : 370–385.

[73] Hong, K., Hinck, L., Nishiyama, M., Poo, M.M., Tessier-Lavigne, M., & Stein, E. (1999), A ligand-gated association between cytoplasmic domains of UNC5 and DCC family receptors converts netrin-induced growth cone attraction to repulsion. *Cell*. 97 (7) : 927–941.

[74] Llambi, F., Causeret, F., Bloch-Gallego, E., & Mehlen, P. (2001), Netrin-1 acts as a survival factor via its receptors UNC5H and DCC. *The EMBO Journal*. 20 (11) : 2715–2722.

[75] Wang, R., Wei, Z., Jin, H., Wu, H., Yu, C., Wen, W., Chan, L.N., Wen, Z., & Zhang, M. (2009), Autoinhibition of UNC5b Revealed by the Cytoplasmic Domain Structure of the Receptor. *Molecular Cell*. 33 (6) : 692–703.

[76] Lu, Y.C., Nazarko, O. V., Sando, R., Salzman, G.S., Südhof, T.C., & Araç, D. (2015), Structural Basis of Latrophilin-FLRT-UNC5 Interaction in Cell Adhesion. *Structure*. 23 (9) : 1678–1691.

[77] Jackson, V.A., Mehmood, S., Chavent, M., Roversi, P., Carrasquero, M., Del Toro, D., Seyit-Bremer, G., Ranaivoson, F.M., Comoletti, D., Sansom, M.S.P., Robinson, C. V., Klein, R., & Seiradake, E. (2016), Super-complexes of adhesion GPCRs and neural guidance receptors. *Nature Communications*. 7 (1) : 11184.

[78] Engelkamp, D. (2002), Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation. *Mechanisms of Development*. 118 : 191–197.

[79] Leung-Hagesteijn, C., Spence, A.M., Stern, B.D., Zhou, Y., Su, M.-W., Hedgecock, E.M., & Culotti, J.G. (1992), UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in C. elegans. *Cell*. 71 (2) : 289–299.

[80] Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S.Y., Culotti, J.G., & Tessier-Lavigne, M. (1996), Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell*. 87 (2) : 175–185.

[81] Karaulanov, E.E., Böttcher, R.T., & Niehrs, C. (2006), A role for fibronectin-leucine-rich transmembrane cell-surface proteins in homotypic cell adhesion. *EMBO Reports*. 7 (3) : 283.

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

[83] Yamagishi, S., Hampel, F., Hata, K., Del Toro, D., Schwark, M., Kvachnina, E., Bastmeyer, M., Yamashita, T., Tarabykin, V., Klein, R., & Egea, J. (2011), FLRT2 and FLRT3 act as repulsive guidance cues for Unc5-positive neurons. *EMBO Journal*. 30 (14) : 2920–2933.

[84] Miyoshi, G. & Fishell, G. (2012), Dynamic FoxG1 Expression Coordinates the Integration of Multipolar Pyramidal Neuron Precursors into the Cortical Plate. *Neuron*. 74 (6) : 1045–1058.

[85] Larrivée, B., Freitas, C., Trombe, M., Lv, X., DeLafarge, B., Yuan, L., Bouvrée, K., Bréant, C., Del Toro, R., Bréchot, N., Germain, S., Bono, F., Dol, F., Claes, F., Fischer, C., Autiero, M., Thomas, J.L., Carmeliet, P., Tessier-Lavigne, M., & Eichmann, A. (2007), Activation of the UNC5B receptor by Netrin-1 inhibits sprouting angiogenesis. *Genes and Development*. 21 (19) : 2433–2447.

[86] Lu, X., Le Noble, F., Yuan, L., Jiang, Q., De Lafarge, B., Sugiyama, D., Bréant, C., Claes, F., De Smet, F., Thomas, J.L., Autiero, M., Carmeliet, P., Tessier-Lavigne, M., & Eichmann, A. (2004), The netrin receptor UNC5B mediates guidance events controlling morphogenesis of the vascular system. *Nature*. 432 (7014) : 179–186.

[87] Kye, W.P., Grouse, D., Lee, M., Karnik, S.K., Sorensen, L.K., Murphy, K.J., Kuo, C.J., & Li, D.Y. (2004), The axonal attractant Netrin-1 is an angiogenic factor. *PNAS*. 101 (46) : 16210–16215.

[88] Nguyen, A. & Cai, H. (2006), Netrin-1 induces angiogenesis via a DCC-dependent ERK1/2-eNOS feed-forward mechanism. *PNAS*. 103 (17) : 6530–6535.

[89] Wilson, B.D., Ii, M., Park, K.W., Suli, A., Sorensen, L.K., Larrieu-Lahargue, F., Urness, L.D., Suh, W., Asai, J., Kock, G.A.H., Thorne, T., Silver, M., Thomas, K.R., Chien, C. Bin, Losordo, D.W., & Li, D.Y. (2006), Netrins promote developmental and therapeutic angiogenesis. *Science*. 313 (5787) : 640–644.

[90] Yang, Y., Zou, L., Wang, Y., Xu, K.S., Zhang, J.X., & Zhang, J.H. (2007), Axon guidance cue netrin-1 has dual function in angiogenesis. *Cancer Biology and Therapy*. 6 (5) : 743–748.

[91] Goldschneider, D. & Mehlen, P. (2010), Dependence receptors: a new paradigm in cell signaling and cancer therapy. *Oncogene*. 29 (13) : 1865–1882.

[92] Mille, F., Llambi, F., Guix, C., Delloye-Bourgeois, C., Guenebeaud, C., Castro-Obregon, S., Bredesen, D.E., Thibert, C., & Mehlen, P. (2009), Interfering with multimerization of netrin-1 receptors triggers tumor cell death. *Cell Death & Differentiation*. 16 (10) : 1344–1351.

[93] Guenebeaud, C., Goldschneider, D., Castets, M., Guix, C., Chazot, G., Delloye-Bourgeois, C., Eisenberg-Lerner, A., Shohat, G., Zhang, M., Laudet, V., Kimchi, A., Bernet, A., & Mehlen, P. (2010), The Dependence Receptor UNC5H2/B Triggers Apoptosis via PP2A-Mediated Dephosphorylation of DAP Kinase. *Molecular Cell*. 40 (6) : 863–876.

[94] Thiébault, K., Mazelin, L., Pays, L., Llambi, F., Joly, M.O., Scoazec, J.Y., Saurin, J.C., Romeo, G., & Mehlen, P. (2003), The netrin-1 receptors UNC5H

are putative tumor suppressors controlling cell death commitment. *Proceedings of the National Academy of Sciences of the United States of America*. 100 (7) : 4173–4178.

[95] Mehlen, P., Delloye-Bourgeois, C., & Chédotal, A. (2011), Novel roles for Slits and netrins: axon guidance cues as anticancer targets? *Nature Reviews Cancer*. 11 (3) : 188–197.

[96] Delloye-Bourgeois, C., Brambilla, E., Coissieux, M.-M., Guenebeaud, C., Pedeux, R., Firlej, V., Cabon, F., Brambilla, C., Mehlen, P., & Bernet, A. (2009), Interference With Netrin-1 and Tumor Cell Death in Non–Small Cell Lung Cancer. *JNCI: Journal of the National Cancer Institute*. 101 (4) : 237–247.

[97] Delloye-Bourgeois, C., Fitamant, J., Paradisi, A., Cappellen, D., Douc-Rasy, S., Raquin, M.-A., Stupack, D., Nakagawara, A., Rousseau, R., Combaret, V., Puisieux, A., Valteau-Couanet, D., Bénard, J., Bernet, A., & Mehlen, P. (2009), Netrin-1 acts as a survival factor for aggressive neuroblastoma. *Journal of Experimental Medicine*. 206 (4) : 833–847.

[98] Han, P., Fu, Y., Liu, J., Wang, Y., He, J., Gong, J., Li, M., Tan, Q., Li, D., Luo, Y., Han, J., Liu, J., Tu, W., Wang, Y., Tian, D., & Yan, W. (2015), Netrin-1 promotes cell migration and invasion by down-regulation of BVES expression in human hepatocellular carcinoma. *American Journal of Cancer Research*. 5 (4) : 1396–1409.

[99] Yin, K., Wang, L., Zhang, X., He, Z., Xia, Y., Xu, J., Wei, S., Li, B., Li, Z., Sun, G., Li, Q., Xu, H., & Xu, Z. (2017), Netrin-1 promotes gastric cancer cell proliferation and invasion via the receptor neogenin through PI3K/AKT signaling pathway. *Oncotarget*. 8 (31) : 51177–51189.

[100] Yin, K., Dang, S., Cui, L., Fan, X., Xie, R., Qu, J., Shang, M., & Chen, J. (2018), Netrin-1 promotes metastasis of gastric cancer by regulating YAP activity. *Biochemical and Biophysical Research Communications*. 496 (1) : 76–82.

[101] Grandin, M., Meier, M., Delcros, J.G., Nikodemus, D., Reuten, R., Patel, T.R., Goldschneider, D., Orriss, G., Krahn, N., Boussouar, A., Abes, R., Dean, Y., Neves, D., Bernet, A., Depil, S., Schneiders, F., Poole, K., Dante, R., Koch, M., Mehlen, P., & Stetefeld, J. (2016), Structural Decoding of the Netrin-1/UNC5 Interaction and its Therapeutical Implications in Cancers. *Cancer Cell*. 29 (2) : 173–185.

[102] Verschueren, E., Husain, B., Yuen, K., Sun, Y., Paduchuri, S., Senbabaoglu, Y., Lehoux, I., Arena, T.A., Wilson, B., Lianoglou, S., Bakalarski, C., Franke, Y., Chan, P., Wong, A.W., Gonzalez, L.C., Mariathasan, S., Turley, S.J., Lill, J.R., & Martinez-Martin, N. (2020), The Immunoglobulin Superfamily

Receptome Defines Cancer-Relevant Networks Associated with Clinical Outcome. *Cell*. 182 (2) : 329- 344.e19.

[103] Shcherbakova, A., Preller, M., Taft, M.H., Pujols, J., Ventura, S., Tiemann, B., Buettner, F.F., & Bakker, H. (2019), C-mannosylation supports folding and enhances stability of thrombospondin repeats. *ELife*. 8 (e52978).

[104] Di Bella, D.J., Habibi, E., Stickels, R.R., Scalia, G., Brown, J., Yadollahpour, P., Yang, S.M., Abbate, C., Biancalani, T., Macosko, E.Z., Chen, F., Regev, A., & Arlotta, P. (2021), Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature*. 595 (7868) : 554–559.

[105] Florio, M., Albert, M., Taverna, E., Namba, T., Brandl, H., Lewitus, E., Haffner, C., Sykes, A., Wong, F.K., Peters, J., Guhr, E., Klemroth, S., Prüfer, K., Kelso, J., Naumann, R., Nüsslein, I., Dahl, A., Lachmann, R., Pääbo, S., & Huttner, W.B. (2015), Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science*. 347 (6229) : 1465–1470.

[106] Barkovich, A.J., Guerrini, R., Kuzniecky, R.I., Jackson, G.D., & Dobyns, W.B. (2012), A developmental and genetic classification for malformations of cortical development: Update 2012. *Brain*. 135 (5) : 1348–1369.

[107] Cottereau, E., Mortemousque, I., Moizard, M.-P., Bürglen, L., Lacombe, D., Gilbert-Dussardier, B., Sigaudy, S., Boute, O., David, A., Faivre, L., Amiel, J., Robertson, R., Viana Ramos, F., Bieth, E., Odent, S., Demeer, B., Mathieu, M., Gaillard, D., Van Maldergem, L., Baujat, G., Maystadt, I., Héron, D., Verloes, A., Philip, N., Cormier-Daire, V., Frouté, MF., Pinson, L., Blanchet, P., Sarda, P., Willems, M., Jacquinet, A., Ratbi, I., Van Den Ende, J,. Lackmy-Port Lis, M., Goldenberg, A., Bonneau, D., Rossignol, S., & Toutain, A. (2013), Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*. 163 (2) : 92–105.

[108] Wang, H., Zhang, B., Gu, M., Li, S., Chi, Z., & Hao, L. (2014), Overexpression of the dependence receptor UNC5H4 inhibits cell migration and invasion, and triggers apoptosis in neuroblastoma cell. *Tumor Biology*. 35 (6) : 5417–5425.

[109] Saikali, Z. & Sinnett, D. (2000), Expression of glypican 3 (GPC3) in embryonal tumors. *International Journal of Cancer*. 89 (5) : 418–422.

[110] Dong, R., Yang, R., Zhan, Y., Lai, H.D., Ye, C.J., Yao, X.Y., Luo, W.Q., Cheng, X.M., Miao, J.J., Wang, J.F., Liu, B.H., Liu, X.Q., Xie, L.L., Li, Y., Zhang, M., Chen, L., Song, W.C., Qian, W., Gao, W.Q., Tang, Y.H., Shen, C.Y., Jiang, W., Chen, G., Yao, W., Dong, K.R., Xiao, X.M., Zheng, S., Li, K., & Wang, J. (2020), Single-Cell Characterization of Malignant Phenotypes and

Developmental Trajectories of Adrenal Neuroblastoma. *Cancer Cell*. 38 (5) : 716- 733.

[111] Cheung, N.K. V. & Dyer, M.A. (2013), Neuroblastoma: Developmental biology, cancer genomics and immunotherapy. *Nature Reviews Cancer*. 13 (6) : 397–411.

[112] Delloye-Bourgeois, C., Bertin, L., Thoinet, K., Jarrosson, L., Kindbeiter, K., Buffet, T., Tauszig-Delamasure, S., Bozon, M., Marabelle, A., Combaret, V., Bergeron, C., Derrington, E., & Castellani, V. (2017), Microenvironment-Driven Shift of Cohesion/Detachment Balance within Tumors Induces a Switch toward Metastasis in Neuroblastoma. *Cancer Cell*. 32 (4) : 427-443.e8.

[113] Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N., & Hammers, R. (1993), Naturally occurring antibodies devoid of light chains. *Nature*. 363 (6428) : 446–448.

[114] Dooley, H. & Flajnik, M.F. (2006), Antibody repertoire development in cartilaginous fish. *Developmental & Comparative Immunology*. 30 (1–2) : 43–56.

[115] Padlan, E.A. (1994), Anatomy of the antibody molecule. *Molecular Immunology*. 31 (3) : 169–217.

[116] Muyldermans, S. (2013), Nanobodies: Natural single-domain antibodies. *Annual Review of Biochemistry*. 82 (1) : 775–797.

[117] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, T., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., & Hassabis D. (2021), Highly accurate protein structure prediction with AlphaFold. *Nature*. 596 (7873) : 583–589.

[118] Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., & Velankar, S. (2022), AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*. 50 (D1) : D439–D444.

[119] Muyldermans, S. (2021), Applications of Nanobodies. *Annual Review of Animal Biosciences*. 9 : 401–421.

[120] Yang, E.Y. & Shah, K. (2020), Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Frontiers in Oncology*. 10 : 1182.

[121] Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G.F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W.G.J., Kobilka, B.K., & Steyaert, J. (2014), A general protocol for the generation of Nanobodies for structural biology. *Nature Protocols*. 9 (3) : 674–693.

[122] Zimmermann, I., Egloff, P., Hutter, C.A.J., Kuhn, B.T., Bräuer, P., Newstead, S., Dawson, R.J.P., Geertsma, E.R., & Seeger, M.A. (2020), Generation of synthetic nanobodies against delicate proteins. *Nature Protocols*. 15 (5) : 1707–1741.

[123] Huo, J., Le Bas, A., Ruza, R.R., Duyvesteyn, H.M.E., Mikolajek, H., Malinauskas, T., Tan, T.K., Rijal, P., Dumoux, M., Ward, P.N., Ren, J., Zhou, D., Harrison, P.J., Weckener, M., Clare, D.K., Vogirala, V.K., Radecke, J., Moynié, L., Zhao, Y., Gilbert-Jaramillo, J., Knight, M.L., Tree, J.A., Buttigieg, K.R., Coombes, N., Elmore, M.J., Carroll, M.W., Carrique, L., Shah, P.N.M., James, W., Townsend, A.R., Stuart, D.I., Owens R.J., & Naismith, J.H. (2020), Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural & Molecular Biology*. 27 (9) : 846–854.

[124] Ma, H., Zeng, W., Meng, X., Huang, X., Yang, Y., Zhao, D., Zhou, P., Wang, X., Zhao, C., Sun, Y., Wang, P., Ou, H., Hu, X., Xiang, Y., & Jin, T. (2021), Potent Neutralization of SARS-CoV-2 by Hetero-Bivalent Alpaca Nanobodies Targeting the Spike Receptor-Binding Domain. *Journal of Virology*. 95 (10).

[125] Zhang, J., Tanha, J., Hirama, T., Khieu, N.H., To, R., Tong-Sevinc, H., Stone, E., Brisson, J.R., & MacKenzie, C.R. (2004), Pentamerization of Single-domain Antibodies from Phage Libraries: A Novel Strategy for the Rapid Generation of High-avidity Antibody Reagents. *Journal of Molecular Biology*. 335 (1) : 49–56.

[126] Dumoulin, M., Conrath, K., Van Meirhaeghe, A., Meersman, F., Heremans, K., Frenken, L.G.J., Muyldermans, S., Wyns, L., & Matagne, A. (2001), Single-domain antibody fragments with high conformational stability. *Protein Science*. 11 (3) : 500–515.

[127] De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., & Wyns, L. (2006), Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *PNAS*. 103 (12) : 4586–4591.

[128] Rasmussen, S.G.F., Devree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., Mathiesen, J.M., Shah, S.T.A., Lyons, J.A., Caffrey, M., Gellman, S.H., Steyaert, J.,

Skiniotis, G., Weis, W.I., Sunahara, R.K., & Kobilka, B.K.. (2011), Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature*. 477 (7366) : 549–557.

[129] Kromann-Hansen, T., Oldenburg, E., Yung, K.W.Y., Ghassabeh, G.H., Muyldermans, S., Declerck, P.J., Huang, M., Andreasen, P.A., & Ngo, J.C.K. (2016), A Camelid-derived Antibody Fragment Targeting the Active Site of a Serine Protease Balances between Inhibitor and Substrate Behavior. *Journal of Biological Chemistry*. 291 (29) : 15156–15168.

[130] Kruse, A.C., Ring, A.M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hübner, H., Pardon, E., Valant, C., Sexton, P.M., Christopoulos, A., Felder, C.C., Gmeiner, P., Steyaert, J., Weis, W.I., Garcia, K.C., Wess, J., & Kobilka, B.K. (2013), Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. 504 (7478) : 101–106.

[131] Uchański, T., Pardon, E., & Steyaert, J. (2020), Nanobodies to study protein conformational states. *Current Opinion in Structural Biology*. 60 : 117–123.

[132] Bukowska, M.A. & Grütter, M.G. (2013), New concepts and aids to facilitate crystallization. *Current Opinion in Structural Biology*. 23 (3) : 409–416.

[133] Park, Y.J., Budiarto, T., Wu, M., Pardon, E., Steyaert, J., & Hol, W.G.J. (2012), The structure of the C-terminal domain of the largest editosome interaction protein and its role in promoting RNA binding by RNA-editing ligase L2. *Nucleic Acids Research*. 40 (14) : 6966–6977.

[134] Wagner, T.R. & Rothbauer, U. (2020), Nanobodies right in the MIDDLE: INTRABODIES as toolbox to visualize and modulate antigens in the living cell. *Biomolecules*. 10 (12) : 1–19.

[135] Traenkle, B. & Rothbauer, U. (2017), Under the microscope: Single-domain antibodies for live-cell imaging and super-resolution microscopy. *Frontiers in Immunology*. 8 : 1030.

[136] Uchański, T., Masiulis, S., Fischer, B., Kalichuk, V., López-Sánchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkönig, A., Zögg, T., Remaut, H., Naismith, J.H., Nury, H., Vranken, W., Aricescu, A.R., Pardon, E., & Steyaert, J. (2021), Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nature Methods*. 18 (1) : 60–68.

[137] Wu, X. & Rapoport, T.A. (2021), Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *PNAS*. 118 (41) : e2115001118.

[138] Debie, P., Lafont, C., Defrise, M., Hansen, I., van Willigen, D.M., van Leeuwen, F.W.B., Gijsbers, R., D'Huyvetter, M., Devoogdt, N., Lahoutte, T.,

Mollard, P., & Hernot, S. (2020), Size and affinity kinetics of nanobodies influence targeting and penetration of solid tumours. *Journal of Controlled Release*. 317 : 34–42.

[139] Muruganandam, A., Tanha, J., Narang, S., & Stanimirovic, D. (2002), Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium. *The FASEB Journal*. 16 (2) : 1–22.

[140] Klarenbeek, A., El Mazouari, K., Desmyter, A., Blanchetot, C., Hultberg, A., de Jonge, N., Roovers, R.C., Cambillau, C., Spinelli, S., Del-Favero, J., Verrips, T., de Haard, H.J., & Achour, I. (2015), Camelid Ig V genes reveal significant human homology not seen in therapeutic target genes, providing for a powerful therapeutic antibody platform. *MAbs*. 7 (4) : 693–706.

[141] Vincke, C., Loris, R., Saerens, D., Martinez-Rodriguez, S., Muyldermans, S., & Conrath, K. (2009), General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *Journal of Biological Chemistry*. 284 (5) : 3273–3284.

[142] Menzel, S., Schwarz, N., Haag, F., & Koch-Nolte, F. (2018), Nanobody-based biologics for modulating purinergic signaling in inflammation and immunity. *Frontiers in Pharmacology*. 9 : 266.

[143] Arezumand, R., Alibakhshi, A., Ranjbari, J., Ramazani, A., & Muyldermans, S. (2017), Nanobodies as novel agents for targeting angiogenesis in solid cancers. *Frontiers in Immunology*. 8 : 1746.

[144] Safarzadeh Kozani, P., Naseri, A., Mirarefin, S.M.J., Salem, F., Nikbakht, M., Evazi Bakhshi, S., & Safarzadeh Kozani, P. (2022), Nanobody-based CAR-T cells for cancer immunotherapy. *Biomarker Research*. 10 (1) : 1–18.

[145] Wang, C., Gao, W., Feng, M., Pastan, I., & Ho, M. (2017), Construction of an immunotoxin, HN3-mPE24, targeting glypican-3 for liver cancer therapy. *Oncotarget*. 8 (20) : 32450–32460.

[146] White, P.C. & McMillan, D.R. (2014), Discovery and Function of the Very Large G Protein-Coupled Receptor. in: C. Stevens (Ed.), G Protein-Coupled Receptor Genetics. Methods in Pharmacology and Toxicology, Humana Press, Totowa, NJ, pp. 67–83.

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

[148] Hamann, J., Aust, G., Araç, D., Engel, F.B., Formstone, C., Fredriksson, R., Hall, R.A., Harty, B.L., Kirchhoff, C., Knapp, B., Krishnan, A., Liebscher, I.,

Lin, H.H., Martinelli, D.C., Monk, K.R., Peeters, M.C., Piao, X., Prömel, S., Schöneberg, T., Schwartz, T.W., Singer, K., Stacey, M., Ushkaryov, Y.A., Vallon, M., Wolfrum, U., Wright, M.W., Xu, L., Langenhan, T., & Schiöth, H.B. (2015), International union of basic and clinical pharmacology. XCIV. adhesion G protein-coupled receptors. *Pharmacological Reviews*. 67 : 338–367.

[149] Knapp, B. & Wolfrum, U. (2016), Adhesion GPCR-Related Protein Networks. in: T. Langenhan, T. Schöneberg (Eds.), Handbook of Experimental Pharmacology, Springer New York LLC, pp. 147–178.

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

[151] Krasnoperov, V., Lu, Y., Buryanovsky, L., Neubert, T.A., Ichtehenko, K., & Petrenko, A.G. (2002), Post-translational proteolytic processing of the calcium-independent receptor of alpha-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. *The Journal of Biological Chemistry*. 277 (48) : 46518–46526.

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

[153] Liebscher, I., Schön, J., Petersen, S.C., Fischer, L., Auerbach, N., Demberg, L.M., Mogha, A., Cöster, M., Simon, K.U., Rothemund, S., Monk, K.R., & Schöneberg, T. (2014), A Tethered Agonist within the Ectodomain Activates the Adhesion G Protein-Coupled Receptors GPR126 and GPR133. *Cell Reports*. 9 (6) : 2018–2026.

[154] Demberg, L.M., Winkler, J., Wilde, C., Simon, K.U., Schön, J., Rothemund, S., Schöneberg, T., Prömel, S., & Liebscher, I. (2017), Activation of adhesion G protein-coupled receptors: Agonist specificity of Stachel sequence-derived peptides. *Journal of Biological Chemistry*. 292 (11) : 4383–4394.

[155] Stoveken, H.M., Hajduczok, A.G., Xu, L., & Tall, G.G. (2015), Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *PNAS*. 112 (19) : 6194–6199.

[156] Liebscher, I., Schön, J., Petersen, S.C., Fischer, L., Auerbach, N., Demberg, L.M., Mogha, A., Cöster, M., Simon, K.U., Rothemund, S., Monk, K.R., & Schöneberg, T. (2014), A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. *Cell Reports*. 9 (6) : 2018–2026.

[157] Kishore, A., Purcell, R.H., Nassiri-Toosi, Z., & Hall, R.A. (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.

[158] Salzman, G.S., Zhang, S., Gupta, A., Koide, A., Koide, S., & Araç, D. (2017), Stachel-independent modulation of GPR56/ADGRG1 signaling by synthetic ligands directed to its extracellular region. *PNAS*. 114 (38) : 10095–10100.

[159] Scholz, N., Guan, C., Nieberler, M., Grotemeyer, A., Maiellaro, I., Gao, S., Beck, S., Pawlak, M., Sauer, M., Asan, E., Rothemund, S., Winkler, J., Prömel, S., Nagel, G., Langenhan, T., & Kittel, R.J. (2017), Mechano-dependent signaling by latrophilin/CIRL quenches cAMP in proprioceptive neurons. *ELife*. 6 (e28360).

[160] Stephenson, J.R., Paavola, K.J., Schaefer, S.A., Kaur, B., Van Meir, E.G., & Hall, R.A. (2013), Brain-specific Angiogenesis Inhibitor-1 Signaling, Regulation, and Enrichment in the Postsynaptic Density. *The Journal of Biological Chemistry*. 288 (31) : 22248–22256.

[161] Randy McMillan, D., Kayes-Wandover, K.M., Richardson, J.A., & White, P.C. (2002), Very large G protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. *Journal of Biological Chemistry*. 277 (1) : 785–792.

[162] Yagi, H., Takamura, Y., Yoneda, T., Konno, D., Akagi, Y., Yoshida, K., & Sato, M. (2005), Vlgr1 knockout mice show audiogenic seizure susceptibility. *Journal of Neurochemistry*. 92 (1) : 191–202.

[163] Skradski, S.L., Clark, A.M., Jiang, H., White, H.S., Fu, Y.H., & Ptáček, L.J. (2001), A novel gene causing a mendelian audiogenic mouse epilepsy. *Neuron*. 31 (4) : 537–544.

[164] Nikkila, H., McMillan, D.R., Nunez, B.S., Pascoe, L., Curnow, K.M., & White, P.C. (2000), Sequence similarities between a novel putative G protein-coupled receptor and Na+/Ca2+ exchangers define a cation binding domain. *Molecular Endocrinology*. 14 (9) : 1351–1364.

[165] Schwarz, E.M. & Benzer, S. (1997), Calx, a Na-Ca exchanger gene of Drosophila melanogaster. *PNAS*. 94 : 10249–10254.

[166] Shin, D., Lin, S.T., Fu, Y.H., & Ptácek, L.J. (2013), Very large G protein-coupled receptor 1 regulates myelin-associated glycoprotein via Gαs/Gαq-mediated protein kinases A/C. *PNAS*. 110 (47) : 19101–19106.

[167] Beckmann, G., Hanke, J., Bork, P., & Reich, J.G. (1998), Merging extracellular domains: fold prediction for laminin G-like and amino-terminal

thrombospondin-like modules based on homology to pentraxins. *Journal of Molecular Biology*. 275 (5) : 725–730.

[168] Staub, E., Pérez-Tur, J., Siebert, R., Nobile, C., Moschonas, N.K., Deloukas, P., & Hinzmann, B. (2002), The novel EPTP repeat defines a superfamily of proteins implicated in epileptic disorders. *Trends in Biochemical Sciences*. 27 (9) : 441–444.

[169] Scheel, H., Tomiuk, S., & Hofmann, K. (2002), A common protein interaction domain links two recently identified epilepsy genes. *Human Molecular Genetics*. 11 (15) : 1757–1762.

[170] Gibert, Y., McMillan, D.R., Kayes-Wandover, K., Meyer, A., Begemann, G., & White, P.C. (2005), Analysis of the very large G-protein coupled receptor gene (Vlgr1/Mass1/USH2C) in zebrafish. *Gene*. 353 (2) : 200–206.

[171] Weston, M.D., Luijendijk, M.W.J., Humphrey, K.D., Möller, C., & Kimberling, W.J. (2004), Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome type II. *American Journal of Human Genetics*. 74 (2) : 357–366.

[172] Hu, Q.X., Dong, J.H., Du, H.B., Zhang, D.L., Ren, H.Z., Ma, M.L., Cai, Y., Zhao, T.C., Yin, X.L., Yu, X., Xue, T., Xu, Z.G., & Sun, J.P. (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.

[173] Zhang, Y., Sloan, S.A., Clarke, L.E., Caneda, C., Plaza, C.A., Blumenthal, P.D., Vogel, H., Steinberg, G.K., Edwards, M.S.B., Li, G., Duncan, J.A., Cheshier, S.H., Shuer, L.M., Chang, E.F., Grant, G.A., Gephart, M.G.H., & Barres, B.A. (2016), Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron*. 89 (1) : 37–53.

[174] McGee, J.A., Goodyear, R.J., McMillan, D.R., Stauffer, E.A., Holt, J.R., Locke, K.G., Birch, D.G., Legan, P.K., White, P.C., Walsh, E.J., & Richardson, G.P. (2006), The very large G-protein-coupled receptor VLGR1: A component of the ankle link complex required for the normal development of auditory hair bundles. *Journal of Neuroscience*. 26 (24) : 6543–6553.

[175] Yang, J., Liu, X., Zhao, Y., Adamian, M., Pawlyk, B., Sun, X., McMillan, R.D., Liberman, M.C., & Li, T. (2010), Ablation of whirlin long isoform disrupts the USH2 protein complex and causes vision and hearing loss. *PLoS Genetics*. 6 (5) : 9.

[176] Yagi, H., Tokano, H., Maeda, M., Takabayashi, T., Nagano, T., Kiyama, H., Fujieda, S., Kitamura, K., & Sato, M. (2007), Vlgr1 is required for proper

stereocilia maturation of cochlear hair cells. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*. 12 (2) : 235–250.

[177] Michalski, N., Michel, V., Bahloul, A., Lefevre, G., Barral, J., Yagi, H., Chardenoux, S., Weil, D., Martin, P., Hardelin, J.-P., Sato, M., & Petit, C. (2007), Molecular Characterization of the Ankle-Link Complex in Cochlear Hair Cells and Its Role in the Hair Bundle Functioning. *Journal of Neuroscience*. 27 (24) : 6478–6488.

[178] Maerker, T., van Wijk, E., Overlack, N., Kersten, F.F.J., Mcgee, J., Goldmann, T., Sehn, E., Roepman, R., Walsh, E.J., Kremer, H., & Wolfrum, U. (2008), A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells. 17 (1) : 71–86.

[179] Ebermann, I., Phillips, J.B., Liebau, M.C., Koenekoop, R.K., Schermer, B., Lopez, I., Schäfer, E., Roux, A.F., Dafinger, C., Bernd, A., Zrenner, E., Claustres, M., Blanco, B., Nürnberg, G., Nürnberg, P., Ruland, R., Westerfield, M., Benzing, T., & Bolz, H.J. (2010), PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. *The Journal of Clinical Investigation*. 120 (6) : 1812–1823.

[180] Reiners, J., van Wijk, E., Märker, T., Zimmermann, U., Jürgens, K., te Brinke, H., Overlack, N., Roepman, R., Knipper, M., Kremer, H., & Wolfrum, U. (2005), Scaffold protein harmonin (USH1C) provides molecular links between Usher syndrome type 1 and type 2. *Human Molecular Genetics*. 14 (24) : 3933–3943.

[181] van Wijk, E., van der Zwaag, B., Peters, T., Zimmermann, U., te Brinke, H., Kersten, F.F.J., Märker, T., Aller, E., Hoefsloot, L.H., Cremers, C.W.R.J., Cremers, F.P.M., Wolfrum, U., Knipper, M., Roepman, R., & Kremer, H. (2006), The DFNB31 gene product whirlin connects to the Usher protein network in the cochlea and retina by direct association with USH2A and VLGR1. *Human Molecular Genetics*. 15 (5) : 751–765.

[182] Knapp, B., Roedig, J., Roedig, H., Krzysko, J., Horn, N., Güler, B.E., Kusuluri, D.K., Yildirim, A., Boldt, K., Ueffing, M., Liebscher, I., & Wolfrum, U. (2022), Affinity Proteomics Identifies Interaction Partners and Defines Novel Insights into the Function of the Adhesion GPCR VLGR1/ADGRV1. *Molecules*. 27 (10) : 3108.

[183] McMillan, D.R. & White, P.C. (2004), Loss of the transmembrane and cytoplasmic domains of the very large G-protein-coupled receptor-1 (VLGR1 or Mass1) causes audiogenic seizures in mice. *Molecular and Cellular Neuroscience*. 26 (2) : 322–329.

[184] Nakayama, J., Fu, Y.-H.H., Clark, A.M., Nakahara, S., Hamano, K., Iwasaki, N., Matsui, A., Arinami, T., Ptáček, L.J., & Ptác ˇek, L.J. (2002), A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. *Annals of Neurology*. 52 (5) : 654–657.

[185] Aricescu, A.R., Lu, W., & Jones, E.Y. (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.

[186] Reeves, P.J., Callewaert, N., Contreras, R., & Khorana, H.G. (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. *PNAS*. 99 (21) : 13419–13424.

[187] Fairhead, M. & Howarth, M. (2015), Site-specific biotinylation of purified proteins using BirA. *Methods in Molecular Biology*. 1266 : 171–184.

[188] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., & Bairoch, A. (2005), Protein Identification and Analysis Tools on the ExPASy Server. in: The Proteomics Protocols Handbook, John M. Walker (Ed), Humana Press, pp. 571–607.

[189] Winter, G., Waterman, D.G., Parkhurst, J.M., Brewster, A.S., Gildea, R.J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I.D., Sauter, N.K., & Evans, G. (2018), DIALS: Implementation and evaluation of a new integration package. *Acta Crystallographica Section D: Structural Biology*. 74 : 85–97.

[190] Winter, G. (2010), Xia2: An expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*. 43 (1) : 186–190.

[191] Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., & Wilson, K.S. (2011), Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*. 67 (4) : 235–242.

[192] Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S.J., Nicholls, R.A., Noble, M., Pannu, N.S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., et al. (2018), CCP4i2: The new graphical user interface to the CCP4 program suite. *Acta Crystallographica Section D: Structural Biology*. 74 : 68–84.

[193] Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., et al. (2022), Protein complex prediction with AlphaFold-Multimer. *BioRxiv* .

[194] McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., & Read, R.J. (2007), Phaser crystallographic software. *Journal of Applied Crystallography*. 40 : 658–674.

[195] Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., & Vagin, A.A. (2011), REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography*. D67 : 355–367.

[196] Emsley, P. & Cowtan, K. (2004), Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. D60 : 2126–2132.

[197] Cowtan, K. (2010), Recent developments in classical density modification. *Acta Crystallographica Section D: Biological Crystallography*. 66 (4) : 470–478.

[198] Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall, W.B., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S., & Richardson, D.C. (2018), MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science*. 27 (1) : 293–315.

[199] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012), Fiji: An open-source platform for biological-image analysis. *Nature Methods*. 9 (7) : 676–682.

[200] Chakraborty, P., Dash, S.P., & Sarangi, P.P. (2020), The role of adhesion protein Fibulin7 in development and diseases. *Molecular Medicine*. 26 (47).

[201] Lin, P.Y., Chen, L.Y., Jiang, M., Trotter, J.H., Seigneur, E., & Südhof, T.C. (2023), Neurexin-2: An inhibitory neurexin that restricts excitatory synapse formation in the hippocampus. *Science Advances*. 9 (1).

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

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

[204] Diederichs, K. & Karplus, P.A. (1997), Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nature Structural Biology*. 4 (4) : 269–275.

[205] Matthews, B.W. (1968), Solvent content of protein crystals. *Journal of Molecular Biology*. 33 (2) : 491–497.

[206] Kantardjieff, K.A. & Rupp, B. (2003), Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Science : A Publication of the Protein Society*. 12 (9) : 1865.

[207] Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., & Bricogne, G. (2011), Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica. Section D, Biological Crystallography*. D67 : 293–302.

[208] Jackson, V.A., Meijer, D.H., Carrasquero, M., Van Bezouwen, L.S., Lowe, E.D., Kleanthous, C., Janssen, B.J.C., & Seiradake, E. (2018), Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. *Nature Communications*. 9 (1079).

[209] Greenfield, N. & Fasman, G.D. (1969), Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation. *Biochemistry*. 8 (10) : 4108–4116.

[210] Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D.H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Orengo, C.A., Pandurangan, A.P., Rivoire, C., et al. (2023), InterPro in 2022. *Nucleic Acids Research*. 51 (D1) : D418–D427.

[211] Letunic, I., Khedkar, S., & Bork, P. (2021), SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research*. 49 (D1) : D458–D460.

[212] Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., & Bateman, A. (2021), Pfam: The protein families database in 2021. *Nucleic Acids Research*. 49 (D1) : D412–D419.

[213] Buchan, D.W.A. & Jones, D.T. (2019), The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*. 47 (W1) : W402–W407.

[214] Jones, D.T. (1999), Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*. 292 (2) : 195–202.

[215] Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., De Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018), SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*. 46 (W1) : W296–W303.

[216] Gupta, R. & Brunak, S. (2002), Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing*. 7 : 310–322.

[217] Hamming, O.J., Kang, L., Svensson, A., Karlsen, J.L., Rahbek-Nielsen, H., Paludan, S.R., Hjorth, S.A., Bondensgaard, K., & Hartmann, R. (2012), Crystal structure of interleukin-21 receptor (IL-21R) bound to IL-21 reveals that sugar chain interacting with WSXWS motif is integral part of IL-21R. *Journal of Biological Chemistry*. 287 (12) : 9454–9460.

[218] Wang, J., Miao, Y., Wicklein, R., Sun, Z., Wang, J., Jude, K.M., Fernandes, R.A., Merrill, S.A., Wernig, M., Garcia, K.C., & Südhof, T.C. (2021), RTN4/NoGo-receptor binding to BAI adhesion-GPCRs regulates neuronal development. *Cell*. 184 : 5869–5885.

[219] Rathinaswamy, M.K., Fleming, K.D., Dalwadi, U., Pardon, E., Harris, N.J., Yip, C.K., Steyaert, J., & Burke, J.E. (2021), HDX-MS-optimized approach to characterize nanobodies as tools for biochemical and structural studies of class IB phosphoinositide 3-kinases. *Structure*. 29 : 1371–1381.

[220] Beales, C.T. & Medalia, O. (2022), Gold nanomaterials and their potential use as cryo-electron tomography labels. *Journal of Structural Biology*. 214 (3) : 107880.

[221] Yamagishi, S., Bando, Y., & Sato, K. (2021), Involvement of Netrins and Their Receptors in Neuronal Migration in the Cerebral Cortex. *Frontiers in Cell and Developmental Biology*. 8.

[222] Wang, X., Chen, Q., Sun, Z., Wang, Y., Su, B., Zhang, C., Cao, H., & Liu, X. (2020), Nanobody affinity improvement: Directed evolution of the anti-ochratoxin A single domain antibody. *International Journal of Biological Macromolecules*. 151 : 312–321.

[223] Miao, X., Li, A., Chen, W., Qi, H., Qiu, Z., Zhang, Y., Zhang, J., & Wang, M. (2013), Optimization and modification of anti-rhTNF-α single chain variable fragment antibody: Effective in vitro affinity maturation and functional expression of chimeric Fab. *Biomedicine and Pharmacotherapy*. 67 (5) : 437–444.

[224] Zhang, J., Tanha, J., Hirama, T., Khieu, N.H., To, R., Tong-Sevinc, H., Stone, E., Brisson, J.R., & MacKenzie, C.R. (2004), Pentamerization of single-domain antibodies from phage libraries: A novel strategy for the rapid

generation of high-avidity antibody reagents. *Journal of Molecular Biology*. 335 (1) : 49–56.

[225] Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022), Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Research. 50 (W1) : W276–W279.

[226] Gouet, P., Robert, X., & Courcelle, E. (2003), ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research*. 31 (13) : 3320–3323.

[227] Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., & Barton, G.J. (2009), Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 25 (9) : 1189–1191.