



HAL
open science

Structural studies of the Roundabout protein family

Francesco Bisiak

► **To cite this version:**

Francesco Bisiak. Structural studies of the Roundabout protein family. Structural Biology [q-bio.BM]. Université Grenoble Alpes, 2018. English. NNT : 2018GREAV006 . tel-01807970

HAL Id: tel-01807970

<https://theses.hal.science/tel-01807970v1>

Submitted on 5 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESIS / THÈSE

To obtain the title of / Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ
GRENOBLE ALPES**

Discipline / Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Presented by / Présentée par

Francesco Bisiak

Thesis supervisor / Thèse dirigée par **Dr. Andrew McCarthy**

Thesis prepared at / Thèse préparée au sein du
**European Molecular Biology Laboratory (EMBL)
Grenoble Outstation**

in / dans **l'École Doctorale de Chimie et Sciences du Vivant**

Structural studies of the Roundabout protein family

*Etudes structurales de la famille des
protéines Roundabout*

Public defense on / Thèse soutenue publiquement le **06/02/2018**

Jury members / Devant le jury composé de :

Dr. Carlo Petosa President / Président
Group Leader, IBS / University of Grenoble, France

Dr. Joanna Timmins Examiner / Examineur
Team Leader, IBS, France

Dr. Valérie Castellani Reviewer / Rapporteur
Group Leader, INMG / University of Lyon, France

Prof. Elena Seiradake Reviewer / Rapporteur
Associate professor, University of Oxford, United Kingdom



Ai miei genitori

TABLE OF CONTENTS

LIST OF FIGURES.....	I
LIST OF TABLES.....	III
ABBREVIATIONS	IV
SUMMARY.....	VIII
<u>1 INTRODUCTION.....</u>	<u>1</u>
RESUME EN FRANÇAIS	1
1.1 COMMON SYSTEMS OF NEURONAL AND VASCULAR DEVELOPMENT.....	2
1.2 NEURONAL WIRING.....	3
1.3 THE ROUNDABOUT PROTEIN FAMILY	5
1.3.1 ROBO1, ROBO2 AND ROBO3 SIGNALLING	7
1.3.2 ROBO4 SIGNALLING AND ANGIOGENESIS	11
1.3.3 INVOLVEMENT OF ROBO4 IN PATHOLOGICAL ANGIOGENESIS.....	13
1.4 THE UNC5 PROTEIN FAMILY	14
1.4.1 UNC5 PROTEINS IN GUIDANCE	15
1.4.2 UNC5 PROTEINS IN APOPTOSIS.....	17
1.4.3 UNC5B IN ANGIOGENESIS	19
1.5 BASICS OF N-LINKED GLYCOSYLATION	20
1.6 SYNTHETIC ANTIBODIES	21
1.7 AIM OF THE STUDY.....	23
<u>2 MATERIALS AND METHODS.....</u>	<u>25</u>
RESUME EN FRANÇAIS	25
2.1 SUPPLIES.....	26
2.2 PREPARATION OF <i>E. COLI</i> COMPETENT CELLS	29
2.3 CLONING OF ROBO4 AND UNC5B CONSTRUCTS	29
2.3.1 MUTAGENESIS OF ROBO4 ECTO.....	32
2.3.2 VECTORS AND LIGATION	33
2.3.3 TRANSFORMATION OF BACTERIAL STRAINS AND DNA ISOLATION.....	34
2.3.4 BACMID PREPARATION FOR INSECT CELL TRANSFECTION	34
2.4 MAMMALIAN CELL CULTURE.....	35
2.4.1 ROUTINE MAINTENANCE AND STORAGE.....	35
2.4.2 MAMMALIAN CELL TRANSFECTION OPTIMIZATION PROTOCOL	37
2.4.3 MAMMALIAN CELL EXPRESSION TEST	38
2.5 INSECT CELL CULTURE	39
2.5.1 INSECT CELL TRANSFECTION	39
2.6 <i>E. COLI</i> EXPRESSION	40
2.7 PURIFICATION OF ROBO4 ECTO	41
2.8 PURIFICATION OF UNC5B CONSTRUCTS	42

2.8.1	PURIFICATION OF UNC5B ECTO	42
2.8.2	PURIFICATION OF UNC5B TSP1-2.....	42
2.9	PURIFICATION OF FABS	43
2.10	ENDO F1 EXPRESSION AND PURIFICATION.....	44
2.11	WESTERN BLOTTING.....	44
2.12	DEGLYCOSYLATION TRIALS.....	45
2.13	SEC-MALS ANALYSIS.....	45
2.14	SPR BINDING EXPERIMENTS.....	46
2.14.1	IMMOBILIZATION OF ROBO4 AND UNC5B	46
2.14.2	KINETIC CONSTANT DETERMINATION OF FAB BINDING	47
2.14.3	SPR BINDING TESTS OF UNC5B.....	48
2.15	SAXS ANALYSIS OF ROBO4 ECTO AND FAB COMPLEXES.....	49
2.16	CRYSTALLISATION EXPERIMENTS.....	51
2.16.1	UNC5B ECTO CRYSTALLISATION	51
2.16.2	CRYSTAL HARVESTING	52
2.17	STRUCTURE DETERMINATION	53
2.17.1	GENERAL DATA COLLECTION STRATEGY	53
2.17.2	UNC5B ECTO CRYSTALS DATA COLLECTION AND ANALYSIS	54
2.17.3	UNC5B ECTO SULPHUR SAD ANOMALOUS DATA COLLECTION	55
2.18	BIOINFORMATICS TOOLS.....	57
3	<u>RESULTS</u>	<u>59</u>
	RESUME EN FRANÇAIS	59
3.1	EXPRESSION AND PURIFICATION OF HUMAN ROBO4 CONSTRUCTS	60
3.2	ROBO4 ECTO DEGLYCOSYLATION	64
3.3	SEC-MALS ANALYSIS OF ROBO4 ECTO	66
3.4	PURIFICATION OF FABS	68
3.5	FAB PROBING OF ROBO4 EXTRACELLULAR DOMAIN BINDING.....	69
3.5.1	DETERMINATION OF INTERACTION CONSTANTS	71
3.6	SEC-SAXS DERIVED STRUCTURAL INFORMATION ON ROBO4 ECTO	75
3.7	CRYSTALLISATION OF ROBO4 ECTO ALONE AND IN COMPLEX WITH FABS.....	83
3.8	EXPRESSION AND PURIFICATION OF HUMAN UNC5B ECTO	84
3.9	UNC5B ECTO DEGLYCOSYLATION.....	86
3.10	EXPRESSION AND PURIFICATION OF UNC5B TSP DOMAINS	87
3.11	BINDING STUDIES OF ROBO4 AND UNC5B EXTRACELLULAR DOMAINS	88
3.12	CRYSTAL STRUCTURE OF UNC5B EXTRACELLULAR DOMAIN.....	92
3.12.1	CRYSTALLISATION AND STRUCTURE SOLUTION.....	92
3.12.2	UNC5B EXTRACELLULAR DOMAIN CRYSTAL STRUCTURE	96
3.12.3	COMPARISON OF UNC5B ECTO TO EXISTING UNC5 STRUCTURES	98
3.12.4	S-SAD ANALYSIS OF UNC5B ECTO	103
4	<u>DISCUSSION</u>	<u>105</u>

RESUME EN FRANÇAIS	105
4.1 RECOMBINANT PRODUCTION OF ROBO4 AND UNC5B EXTRACELLULAR DOMAINS.	106
4.2 THE ROBO4 AND UNC5B EXTRACELLULAR DOMAINS ARE GLYCOSYLATED AND ROBO4 GLYCOSYLATION IS NECESSARY FOR FOLDING	107
4.3 THE ROBO4 EXTRACELLULAR DOMAIN IS A FLEXIBLE MONOMER IN SOLUTION	108
4.4 FABS INTERACTION WITH THE ROBO4 EXTRACELLULAR DOMAIN	110
4.5 THE UNC5B EXTRACELLULAR DOMAIN IS A MONOMER IN SOLUTION AND DOES NOT INTERACT WITH THE ROBO4 EXTRACELLULAR DOMAIN	112
4.6 CRYSTAL STRUCTURE OF THE UNC5B EXTRACELLULAR DOMAIN	115
5 <u>CONCLUSIONS AND PERSPECTIVES</u>	118
REFERENCES.....	121
ACKNOWLEDGEMENTS.....	145

List of figures

Figure 1: Co-localisation of neurons and vessels in mouse endothelium.....	2
Figure 2: Neuron organisation and example of neuronal wiring of the optic nerves.....	3
Figure 3: Domain organisation of mammalian Roundabout proteins.....	5
Figure 4: Robo/Slit role in midline crossing.....	8
Figure 5: Putative Robo/Slit signalling mechanism.....	9
Figure 6: General structure of vertebrate UNC5 proteins.....	14
Figure 7: UNC5 and DCC influence on guidance.....	16
Figure 8: Structure of the cytoplasmic ZU5, UPA and DD domains of UNC5B.....	18
Figure 9: The three main types of N-linked glycosylation.....	20
Figure 10: General IgG and Fab structure.....	21
Figure 11: Robo4 and UNC5B constructs.....	31
Figure 12: Disorder prediction of human Robo4.....	60
Figure 13: Purification of Robo4 ectodomain.....	62
Figure 14: Ion exchange chromatography of Robo4 ecto.....	63
Figure 15: Deglycosylation of Robo4 ecto.....	64
Figure 16: SEC-MALS comparison of Robo4 ecto in HEK293 and HEK293S cell lines.....	67
Figure 17: SDS-PAGE of purified Fabs.....	68
Figure 18: SEC-MALS of Robo4 ecto and Fab5582.....	70
Figure 19: SEC-MALS of Robo4 ecto in complex with Fabs.....	70
Figure 20: pH scouting of Robo4 ecto.....	71
Figure 21: SPR analysis of Fabs: 5555, 5562, 5564 and 5570 binding to Robo4 ecto.....	72
Figure 22: Steady state analysis of Fab5585 binding to Robo4 ecto.....	73
Figure 23: SPR analysis of Fab5582 binding to Robo4 ecto.....	74
Figure 24: Robo4 ecto SAXS curves.....	75
Figure 25: SAXS curves, Guinier and Kratky plots of Robo4 ecto and all Fab complexes.....	76
Figure 26: Pair distribution function of Robo4 ecto and Fab complexes.....	78
Figure 27: SAXS bead models of Robo4 ecto, Robo4 ecto/Fab5555 and Robo4 ecto/Fab5564.....	79
Figure 28: Robo4 ecto SAXS bead model superimposition to Robo1 Ig1-2 structure.....	80
Figure 29: Homology modelling and EOM fit.....	82
Figure 30: Robo4 ecto/Fab complex crystals.....	83
Figure 31: Purification of UNC5B ecto.....	85
Figure 32: Deglycosylation of UNC5B ecto.....	86
Figure 33: Purification and size exclusion chromatography of UNC5B TSP1-2.....	87
Figure 34: Strep pull-down of UNC5B and Robo4 extracellular domains.....	88
Figure 35: SEC-MALS analysis of UNC5B ecto and after incubation with Robo4 ecto.....	89

Figure 36: pH scouting of UNC5B ecto	90
Figure 37: SPR binding test of Robo4 ecto and UNC5B ecto.....	91
Figure 38: Avenir binding test of Robo4 and UNC5B extracellular domains.....	91
Figure 39: UNC5B ecto crystals optimization	92
Figure 40: Multiple sequence alignment of known UNC5 structures and UNC5B ecto.....	95
Figure 41: UNC5B ecto crystal structure and glycosylation density.....	96
Figure 42: Estimated evolutionary conservation of UNC5B residues	97
Figure 43: Superimposition of UNC5B with UNC5D and UNC5A structures	100
Figure 44: TSP loop of UNC5 proteins	101
Figure 45: UNC5A and UNC5B TSP domains alignment	101
Figure 46: Electrostatic potential surfaces of the three known UNC5 structures	102
Figure 47: Anomalous density map at disulphide bridges of UNC5B ecto.....	103
Figure 48: New insights into Robo4 and UNC5B function.....	119

List of tables

Table 1: Supplies, kits and other materials	26
Table 2: Cell culture materials	28
Table 3: Robo4 and UNC5B cloning and mutagenesis primers.....	30
Table 4 List of Fabs CDR	43
Table 5: Fab dilution range for SPR experiment	48
Table 6: Predicted Robo4 glycosylation sites.....	61
Table 7: SEC-MALS calculated molecular weight of Robo4 ecto	66
Table 8: Map of suggested Fabs binding onto Robo4 ecto.....	68
Table 9: Observed molecular weight of Robo4 ecto/Fabs complexes from SEC-MALS.....	69
Table 10: Kinetic constants of Fab binding to Robo4 ecto.....	73
Table 11: SAXS parameters of Robo4 ecto and Fab complexes.....	77
Table 12: List of Robo4 ecto/Fab crystallisation conditions	83
Table 13: Predicted UNC5B glycosylation sites.....	84
Table 14: SEC-MALS analysis of UNC5B ecto binding to Robo4 ecto.....	89
Table 15: Initial crystallisation conditions of UNC5B ecto	93
Table 16: Best crystallisation conditions of UNC5B ecto in presence of seeds	93
Table 17: Crystallographic table of UNC5B ecto single crystal dataset	94
Table 18: DynDom RMSD, angles and hinges between UNC5B, UNC5A and UNC5D domains	98
Table 19: Crystallographic table of UNC5B ecto anomalous sulphur datasets.....	104

Abbreviations

Å	Angstrom	dNTP	Deoxynucleotide triphosphate
Abl	Abelson kinase	DPBS	Dulbecco phosphate buffer saline
ADP	Adenosine diphosphate	DynDom	Protein domain motion analysis
Arf6	ADP-ribosylation factor 6	Ecto	Extracellular domain
ATCC	American type culture collection	EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
AU	Absorbance unit	EDNA	Enhanced automated collection of data
bp	Base pair	EDTA	Ethylen diaminetetracetic acid
CAM	Cell adhesion molecule	Endo F1	Endoglycosidase F1
CC	Conserved cytoplasmic	Endo H	Endoglycosidase H
Cdc42	Cell division control protein homolog 42	EOM	Ensemble optimization method
CDR	Complementary determining region	ER	Endoplasmic reticulum
cm	Centimetre	ESRF	European synchrotron radiation facility
CNS	Central nervous system	exp	Exponent
Comm	Commissureless	Fab	Fragment antigen binding
COS-1	<i>Cercophitecus ethiops</i> kidney fibroblast-like	FBS	Fetal bovine serum
CV	Column volume	Fc	Crystallisable fragment
DAPK	Death associated protein kinase	FLRT	Fibronectin leucine rich repeat transmembrane protein
DCC	Deleted in colorectal cancer	FnIII	Fibronectin type III
DD	Death domain	FW	Forward primer
D _{max}	Maximum dimension	g	Gravitational force
DMEM	Dulbecco's modified eagle media		
DMSO	Dimethyl sulphoxide		

GFP	Green fluorescent protein	k_d	Dissociation constant
GnTI	N-acetylglucosaminyl transferase I	K_D	Equilibrium dissociation constant
GST	Glutathione S-transferase	kDa	KiloDalton
GTPase	Guanosine triphosphate hydrolase	keV	Kiloelectronvolt
HC	Heavy chain	LB	Miller's lysogeny broth
HEK	Human embryonic kidney	LC	Light chain
Hi5	<i>Trichoplusia ni</i> cell line	Lphn3	Latrophilin-3
HMVEC	Human microvascular embryonic cells	M	Molar
HRP	Horseradish peroxidase	Matn	Matrilin-1
hs	<i>Homo sapiens</i>	MALS	Multi angle light scattering
HS	Heparan sulphate	mAU	Milli absorbance unit
HSPG	Heparan sulphate proteoglycans	MBP	Maltose binding protein
HTX	High throughput crystallisation	MC	Mammalian cell
HUVEC	Human umbilical vein endothelial cells	MEM	Modified eagle media
IC	Insect cell	mg	Milligram
Ig	Immunoglobulin	min	Minute
IgG	Immunoglobulin type G	ml	Millilitre
IPTG	isopropyl β -D-1-thiogalactopyranoside	mM	Millimolar
I(s)	Scattering intensity	mm	Millimetre
K	Kelvin	MPD	2-Methyl-2,4-pentanediol
k_a	Association constant	mS	MilliSievert
kb	Kilobase	NAG	N-acetylglucosamine
		NEAA	Non-essential amino acid
		ng	Nanogram
		NHS	N-hydroxysuccinimide

nM	Nanomolar	rMMS	Random microseed matrix screening
nm	Nanometre		
NSD	Normalized spatial discrepancy	RMSD	Root-mean square deviation
OD	Optical density	rn	<i>Rattus norvegicus</i>
PAGE	Polyacrylamide gel electrophoresis	Robo	Roundabout
PBS	Phosphate buffer saline	RU	Response unit
PCR	Polymerase chain reaction	RV	Reverse primer
PDB	Protein databank	s	Seconds
PEG	Polyethylene glycol	s	Momentum transfer (only SAXS data)
PEI	Polyethylenimine	SAXS	Small angle x-ray scattering
PNGase F1	Peptide-N-glycosidase F1	SDS	Sodium dodecyl sulphate
$p(r)$	Pair distribution function	SEC	Size exclusion chromatography
PSB	Partnership for structural biology	Sf21	<i>Spodoptera frugiperda</i> cell line
PSG1	Pregnancy specific glycoprotein 1	SFM	Serum free media
PTM	Post-translational modification	SPR	Surface plasmon resonance
PVDF	Polyvinylidene difluoride	srGAP	Slit/Robo GTPase activating protein
Rac1	Ras-related C3 botulinum toxin substrate 1	S-SAD	Sulphur–single wavelength anomalous diffraction
R_g	Radius of gyration	T75	75 cm ² cell flask
RI	Refractive index	TEV	Tobacco etch virus
R_{ligand}	RU ligand response	TSP	Thrombospondin type 1
R_{max}	RU total analyte response	U	Unit
		μg	Microgram

μl	Microlitre
μM	Micromolar
μm	Micrometre
UNC5	Uncoordinated movement 5
UniProt	Universal protein resource
UPA	UNC5/PIDD/Ankyrin conserved domain
UV	Ultraviolet
V	Volt (only gel electrophoresis)
V	Volume
V_c	Correlated volume
VEGF	Vascular endothelial growth factor
V_p	Porod volume
v/v	Volume/volume
w/v	Weight/volume
w/w	Weight/weight
Z-score	Model quality score
ZU5	Zona occludens-1 like UNC5 domain

Summary

Neuronal and vascular systems require a complex network to properly perform their functions. The processes involved in creating this network rely on coordinated pathways, often activated through common protein/receptor systems, which lead to cytoskeletal remodelling. In general, neuronal and vascular cells respond to extracellular stimuli in the form of soluble secreted proteins, which interact with surface receptors to mediate attraction or repulsion towards the source of the secreted proteins. This process, called guidance, is regulated by seven families of receptors and their respective ligands, which influence each other and can act on the neuronal system, the vascular system or both.

Structural information about the extracellular region of many of these receptors, and how signal is relayed across the membrane, is lacking.

This study is focused around the extracellular domain of two single-pass transmembrane receptors of the Roundabout and UNC5 protein families that are majorly involved in angiogenesis: Robo4 and UNC5B.

Based on the findings of this study, the Robo4 and UNC5B extracellular domains are extensively glycosylated with N-linked glycans of the complex type. Site-directed mutagenesis of the predicted Robo4 glycosylation sites disrupts protein expression, indicating that they are necessary for protein stability and passage through the glycosylation pathway might be necessary for correct folding. MALS and SAXS data show that in solution the Robo4 extracellular domain is a flexible monomer with extended shape. Several Fabs binding to the extracellular domain of Robo4 were characterised, with the expectation to identify those Fabs that could inhibit the reported Robo4/UNC5B interaction for further characterisation. Complex formation was verified by SEC-MALS and SAXS, and interaction constants were determined using SPR. Crystals of some Robo4 extracellular domain/Fab complexes were produced, although the structure of the complex could not be solved at the present time.

Despite a study by another group showing otherwise, pull-down, SEC-MALS and SPR experiments show that the Robo4 and UNC5B extracellular domains do not interact with each other. It is proposed that the difference may be caused by different glycosylation patterns specific to the cell lines used for each study, or by

an undetected third party necessary for interaction. This, however, still requires further study. SEC-MALS analysis showed that the UNC5B extracellular domain is a monomer in solution and its crystal structure was solved at 3.4 Å resolution. Comparison to the existing structures of human UNC5A and rat UNC5D shows striking similarities and a high degree of evolutionary conservation of the Ig domains might be indication of the importance of this region, which is responsible for binding to the guidance cue Netrin. Although the Netrin binding region is known to be within the Ig domains, the precise binding site has not yet been determined, but it might be located in proximity, or within, the negatively charged surfaces present on the Ig domains which are observed in the UNC5B structure. It is hoped that the work presented here will give the basis for better biochemical and structural characterisation of these two receptors.

1 INTRODUCTION

Résumé en français

Les systèmes neuronaux et vasculaires nécessitent un réseau complexe pour exécuter correctement leurs fonctions. Les processus impliqués dans la création de ce réseau s'appuient sur des voies coordonnées, souvent activées par des systèmes protéine/récepteur communs, qui conduisent au remodelage du cytosquelette.

En général, les cellules neuronales et vasculaires répondent aux stimuli extracellulaires sous forme de protéines solubles sécrétées, qui interagissent avec les récepteurs de surface pour favoriser l'attraction ou la répulsion vers la source des protéines sécrétées. Ce processus est régulé par sept familles de récepteurs et leurs ligands respectifs, qui s'influencent les uns sur les autres et peuvent agir sur le système neuronal, vasculaire ou les deux ensembles. Cette étude est centrée sur deux récepteurs transmembranaires à passage unique, qui sont principalement impliqués dans l'angiogenèse : Robo4 et UNC5B.

Robo4 est caractérisé par un domaine N-terminal extracellulaire, composé de deux domaines d'immunoglobuline et deux domaines de fibronectine type III. L'extrémité C-terminale cytoplasmique est désordonnée, avec deux domaines cytoplasmiques conservés, qui sont les sites de liaison pour les facteurs de signalisation en aval.

UNC5B a deux domaines immunoglobulines et deux domaines thrombospondine dans la région N-terminale extracellulaire. Dans la région cytoplasmique, elle a trois domaines (UPA, ZU5 et DD) impliqués dans les interactions cytoplasmiques. Les mécanismes par lesquels Robo4 influence sa voie de signalisation ne sont pas clairs. Comme il a été démontré que Robo4 forme des hétéromères avec d'autres récepteurs Robo et avec UNC5B, une multimerisation du récepteur a été suggérée. En outre, Robo4 peut agir sur les effets répulsifs et attractifs en fonction de la présence d'autres partenaires de liaison. L'attraction et la répulsion sont médiées par les GTPases Rho qui induisent la réorganisation du cytosquelette.

De même, l'effet de la signalisation UNC5B est différent selon le contexte cellulaire. Les changements dans le niveau de phosphorylation médié par les interactions d'UNC5B, induisent des réarrangements du cytosquelette. En outre, UNC5B joue un rôle actif dans le contrôle de l'apoptose.

1.1 Common systems of neuronal and vascular development

The neuronal and vascular systems, while functionally different, are ordered networks that share important similarities. Both require extensive branching of their components, which often co-localise (Figure 1) and influence each other through coordinated pathways (Melani and Weinstein, 2010).

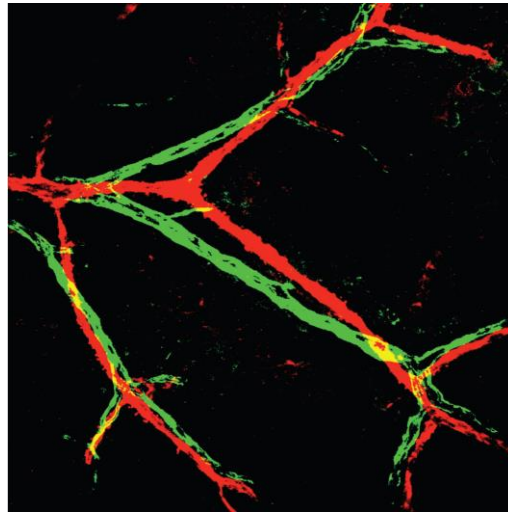


Figure 1: Co-localisation of neurons and vessels in mouse endothelium

Whole-mount immunofluorescence confocal image of endothelial mouse tissue. Arteries (red) are aligned with peripheral nerves (green) following similar branching patterns. Image from (Melani and Weinstein, 2010).

Between the two, the mechanism of assembly and patterning of neuronal cells is the more studied, while less is known about the processes that guide vessel sprouting and development. For this reason, most mechanisms described in this work will take the former as reference. There are, however, many shared pathways which lead to cytoskeletal remodelling in both cases, even using the same agents.

In general, neuronal and vascular cells respond to extracellular stimuli in the form of soluble secreted proteins (guidance cues), which interact with surface receptors to start various cytoplasmic signalling pathways. Combinations of cue/receptor binding, cross-talk between downstream pathways and regulation of protein expression, influence the response of each cell type. The work presented here will address the structural characteristics of the extracellular domains of Robo4 and UNC5B, which both play a major role in angiogenesis, to understand how their interaction can induce transmembrane signal transmission.

1.2 Neuronal wiring

Neurons are the fundamental functional units that transmit and receive nerve impulses in all nervous systems. They are a highly differentiated cell type, assuming distinct forms for electrical signal transduction. The most typical is characterized by the presence of short branched extensions (dendrites) around the main body (soma), followed by one long process (axon), which extends, even for several meters, to contact other neurons or motor cells (Figure 2A). The extension mechanism guiding axons to their destination is called axon guidance, and is fundamental for the development of the central nervous system (CNS) (Evans and Bashaw, 2010a; Garbe and Bashaw, 2004).

The CNS of bilateral organisms is divided into two mirror images around a longitudinal axis of symmetry called the midline. Contact between the two is necessary for correct function of the nervous system, and crossing of the midline is strictly regulated (Placzek and Briscoe, 2005). A precise geometry rules the relative positions of the soma and its axon.

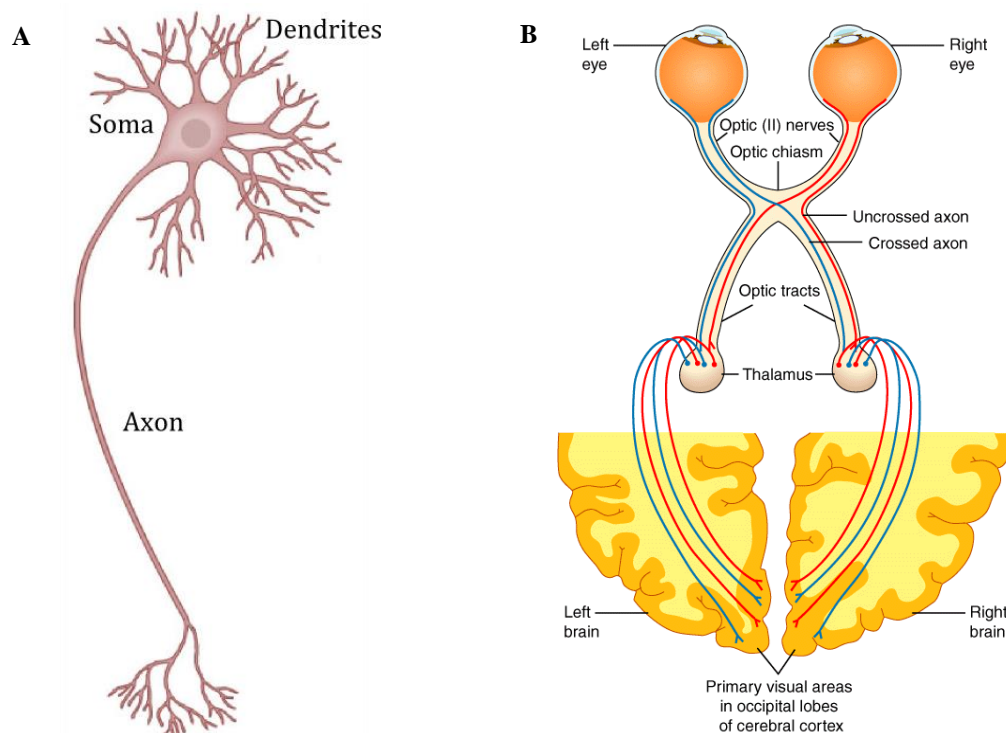


Figure 2: Neuron organisation and example of neuronal wiring of the optic nerves

A: Classical neuron structure. The main cell body (soma) sprouts several tiny processes (dendrites) for incoming signal reception. A single long process is developed for distal signal transmission (axon) which will contact dendrites of one or more different neurons.

Modified from [Human body: form and function, Eberly college of science, USA](#).

B: Simplified representation of ipsilateral (axon and cell body is on the same side of the CNS) and commissural axons (axon and cell body is on opposite sides of the CNS) in the optic nerve.

INTRODUCTION

They are either located on the same side of the CNS (ipsilateral), or the soma is on one side and the axon crosses the midline once to reach the other (commissural).

An example of this is the neuronal wiring of the optical nerves (Figure 2B). Both ipsilateral and commissural axons are necessary to completely form the optic nerve on both halves of the CNS. Guidance cue molecules, and transmembrane receptors, regulate this through chronological and context dependent activities (Raper and Mason, 2010). There are four canonical families of axon guidance cues: the Slit (Tessier-Lavigne and Goodman, 1996), the Netrin (Kennedy, 2000), the Ephrin (Kullander and Klein, 2002) and the Semaphorin (Pasterkamp and Kolodkin, 2003). Along with them, other families are also known to be involved in neuronal wiring, like the fibronectin leucine-rich repeat transmembrane proteins (FLRT), the Wnt (a portmanteau of the first two discovered members Wingless and Int-1) morphogens, as well as the cell adhesion molecules (CAM) and cadherin superfamilies.

Apart from the already large number of factors involved, cross-talk between the guidance cues and their transmembrane receptors is a fundamental factor determining the direction of axonal growth (Chen et al., 2001; Dascenco et al., 2015; Yu and Bargmann, 2001; Zelina et al., 2014).

The overall system is versatile, capable of intervening in other important processes such as cell migration, tumour development and angiogenesis (Bicknell and Harris, 2004; Blockus and Chédotal, 2016; Carmeliet and Tessier-Lavigne, 2005; Jongbloets and Pasterkamp, 2014; Navankasattusas et al., 2008; Seiradake et al., 2014; Suchting et al., 2006; Ypsilanti et al., 2010). Some of these functions will be addressed in detail later, as they were an important factor determining the basis of this study.

1.3 The Roundabout protein family

The first Roundabout gene was identified in *Drosophila* during a large scale genetic screening of mutations affecting CNS development (Seeger et al., 1993). The authors observed that mutations of this gene, caused the axons of commissural neurons to repeatedly cross the midline by circling around the point of crossing, and thus called it Roundabout (Robo). Later studies showed that the encoded protein is a transmembrane receptor involved in midline repulsion (Kidd et al., 1998; Perez and Steller, 1996).

Furthermore, it was shown the guidance cue Slit was a Robo receptor ligand in both *Drosophila* and mammals, and that their combined action is responsible for the collapse of neuronal growth cones (Brose et al., 1999; Kidd et al., 1999; Li et al., 1999; Yuan et al., 1999).

Three different Robo receptors have been identified in *Drosophila* (Robo1, Robo2 and Robo3), while four are present in mammals (Robo1/Dutt1, Robo2, Robo3/Rig1, and Robo4/Magic Roundabout) (Dickson and Gilestro, 2006). Interestingly, despite their similarity, the members of the Robo family in *Drosophila* and mammals were created by independent gene duplications. Only one common Robo was present in the last common ancestors of protostomes and deuterostomes (Evans and Bashaw, 2012).

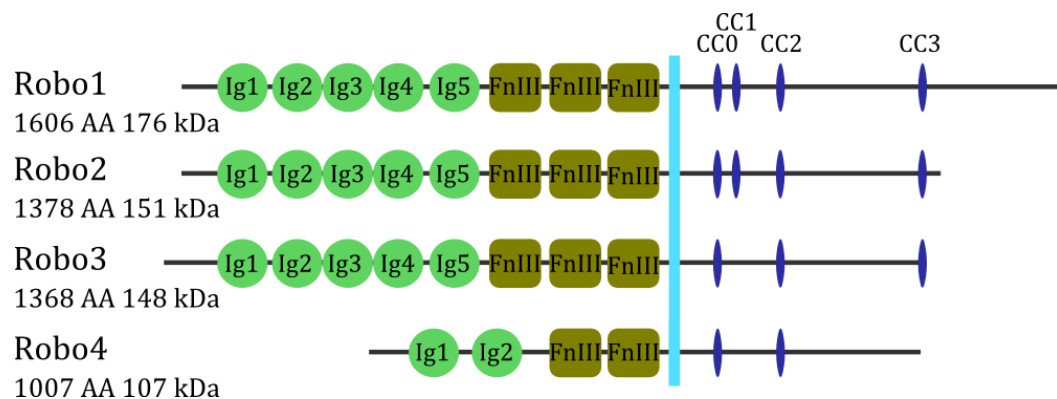


Figure 3: Domain organisation of mammalian Roundabout proteins

Robo1, Robo2 and Robo3 have five Ig domains (green) and three FnIII domains (olive) in the N-terminal extracellular region. They are followed by a single transmembrane α -helical region (sky blue) and three or four CC domains (dark blue) in the cytoplasmic region. Robo4 is the smallest of the family, with two Ig and two FnIII in the N-terminal extracellular half and two CC in the C-terminal cytoplasmic half.

INTRODUCTION

All Robos share a general organisation of immunoglobulin-like (Ig) and fibronectin type III (FnIII) domains in the extracellular N-terminal region, a single membrane spanning α -helix of roughly 20 amino acids, and finally a low order C-terminal tail containing up to four conserved cytoplasmic (CC) motifs (Figure 3).

Robo1, Robo2 and Robo3 are the most similar, containing five Ig and three FnIII domains, differing only in the composition of their cytoplasmic tail. The mammalian specific Robo4, instead, is strikingly different from the others. It is one third smaller in size, containing only the first two Ig domains (which are most similar to the first two Ig of the other Robos), and only two FnIII domains. Additionally, it has only two CC domains in its cytoplasmic region.

The Ig domains are involved in protein-protein interactions that are crucial for Robo signalling. The FnIII domains are a structurally recurring feature, found in 2% of all animal proteins and mainly on proteins which associate to the cellular membrane. All these domains are highly conserved between Robo1, Robo2 and Robo3, but are comparatively more divergent in Robo4. In *Drosophila*, the CC0 domain is a site of tyrosine phosphorylation mediated by the Abelson kinase (Abl), which is suggested to have an inhibitory function (Bashaw et al., 2000; Coleman et al., 2010). CC1 also contains a tyrosine phosphorylation site, and is the binding site of the Netrin receptor, deleted in colorectal cancer (DCC). This is the first example of the cross-talk between various guidance pathways (Bashaw et al., 2000). CC2 is the binding site for the downstream signalling molecule Enabled in *Drosophila* (Bashaw et al., 2000; Wills et al., 2002) and for its homolog, Mena, in mammals (Jones et al., 2009), which counterbalances Abl inhibition. CC3 is a polyproline stretch (Kidd et al., 1998) involved in the recruitment of Slit/Robo guanosine triphosphate hydrolase (GTPase) activating proteins (srGAP) (Li et al., 2006; Wong et al., 2001).

Apart from the different composition of their domains, the localisation of the Robo proteins differs depending on their function. While Robo1, Robo2 and Robo3 are mainly expressed in neuronal tissues and involved in neuronal development, Robo4 is more ubiquitous and assumes a major role in angiogenesis.

1.3.1 Robo1, Robo2 and Robo3 signalling

These three Robo receptors are mostly involved in the neuronal guidance of commissural and ipsilateral axons in the CNS, mostly through interaction with the Slit proteins. In *Drosophila*, there is only a single Slit protein expressed, and secreted, by midline cells (Rothberg et al., 1988, 1990) which was shown to bind all three Robos (Howitt et al., 2004). In contrast, mammals have three Slit proteins (Slit1, Slit2 and Slit3), which are well conserved. Mammalian Slit2 is the best characterized, and most commonly expressed guidance cue at the midline (Nguyen-Ba-Charvet and Chédotal, 2002; Nguyen-Ba-Charvet et al., 2001; Yuan et al., 1999; Zelina et al., 2014). However, the three mammalian Slit proteins are redundant in function, as only triple Slit1, Slit2, Slit3 knockout mutants allow unregulated axonal midline crossing (Long et al., 2004). Furthermore, the important role of the interaction between Slit2 and Robo1 is confirmed by the high degree of evolutionary conservation present, as human Slit2 can bind *Drosophila* Robo1, and *Drosophila* Slit can bind mammalian Robo1 (Ba-charvet et al., 2001; Brose et al., 1999; Li et al., 1999).

Figure 4 shows how Robo/Slit interactions affect ipsilateral and commissural axon crossing in *Drosophila* and mammals respectively. Slit mediated repulsion through Robo1 binding does not normally allow midline crossing (Pappu and Zipursky, 2010). To silence Robo1, instead, the presence of another transmembrane receptor is necessary. In *Drosophila*, this is performed by commissureless (Comm) (Georgiou and Tear, 2002), which sequesters Robo1 while in transit through the Golgi and brings it into late endosomal compartments (van den Brink et al., 2013; Keleman et al., 2002). After crossing, down regulation of comm allows Robo to be translocated to the extracellular membrane, and Robo repulsion will not allow recrossing of the midline to occur (Figure 4A). The factors involved in this switch, however, are still not clear.

In mammals, no comm homolog has been identified. However, it was discovered that Robo1 repulsion is partially regulated by Robo3 (Sabatier et al., 2004) (Figure 4B). Two alternative splicing isoforms of Robo3 with opposite effects have been identified: Robo3.1, responsible for the antagonizing effect against Robo1, and Robo3.2, which instead acts together with Robo1 and Robo2 to induce

INTRODUCTION

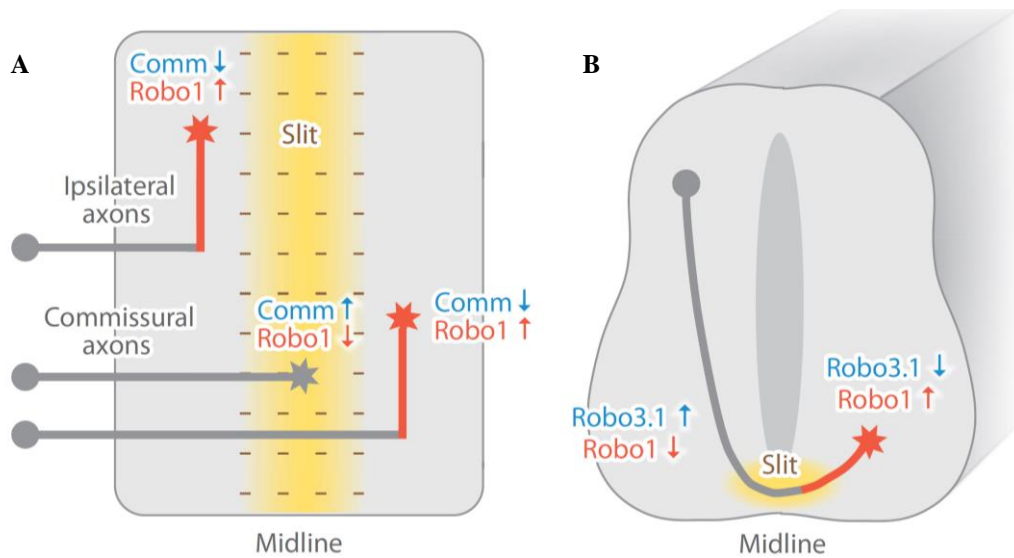


Figure 4: Robo/Slit role in midline crossing

Schematic representation of Robo/Slit mediated midline crossing.

A: In *Drosophila*, Slit expressed at the midline repels axons expressing Robo1. Commissural axons expressing comm, do not present Robo1 on the surface and allows crossing of the midline. Midline recrossing is then denied by downregulation of Comm and Robo1 delivery at the surface.

B: In mammals, the alternative splicing variant of Robo3, Robo3.1, antagonises Robo1 repulsion allowing midline crossing. After crossing, downregulation of Robo3.1 removes Robo1 inhibition and prevents recrossing.

Images from (de Wit et al., 2011).

repulsion (Camurri et al., 2005; Chen et al., 2008). Furthermore, unlike *Drosophila* Robo3, both mammalian Robo3 isoforms are unable to bind Slit2 (Zelina et al., 2014). The mechanism regulating the switch between expression of Robo3.1 and Robo3.2 remains elusive, but the specialisation of Robo3 is considered a critical element of the increased neuronal circuit complexity which distinguishes mammalian evolution (Beamish and Kennedy, 2015; Zelina et al., 2014).

Generally, binding of Slit2 to Robo1 induces recruitment of srGAP, which increases the GTPase activity of the cell division control protein homolog 42 (Cdc42), effectively inactivating it. Inactivation of Cdc42 leads to a successive reduction in the activity of factors which promote actin polymerisation (Wong et al., 2001). The repulsive effect of Slit2 through Robo1 is therefore a direct consequence of the downregulation of actin polymerisation.

Signalling through single pass transmembrane receptors is often tied to their oligomerisation state. Ligand binding (or absence of it), induces association, or dissociation, of the receptor complex, which in turn influences the downstream signalling cascade (Alberts et al., 2002; Moore et al., 2008).

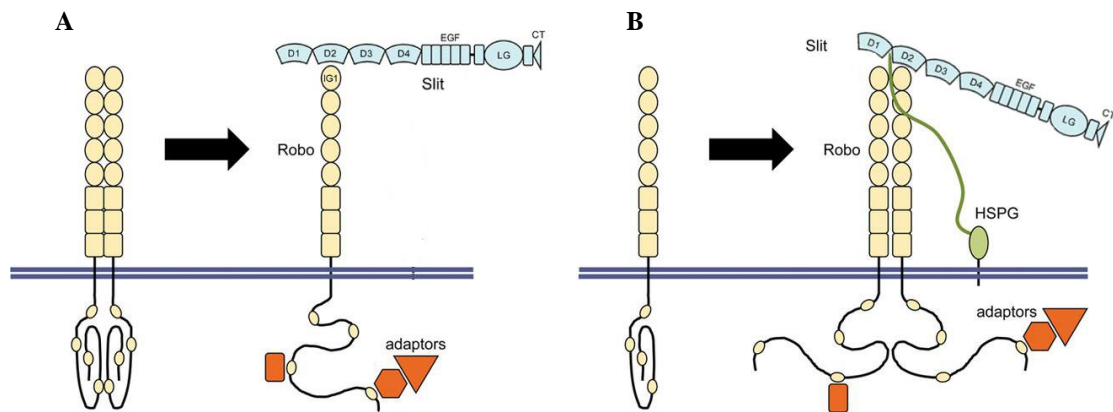


Figure 5: Putative Robo/Slit signalling mechanism

Changes in the oligomerisation state are likely responsible for the Robo signalling cascade. Monomeric and dimeric forms are shown for clarity, but higher orders of organisation are also possible, along with interactions with other partners. Interaction with Slit proteins can **(A)** induce oligomer dissociation or **(B)** oligomer association of Robo1, which in turn enables recruitment of downstream signalling molecules. Heparan sulphate proteoglycans (HSPG) also play a role in the Slit/Robo interaction.

Modified from (Hohenester, 2008).

Although Slit-Robo signalling has been intensely studied there is still a clear lack of knowledge on how exactly their interaction is relayed across the membrane. A reorganization of the Robo receptor oligomeric state upon Slit binding was suggested as the start of the signalling cascade (Hohenester, 2008) (Figure 5).

To support this, previous analysis showed that the ectodomain of Robo is responsible for Robo1-Robo1 oligomerisation (Hivert, 2002; Liu et al., 2004) and that heteromeric complexes of Robo1, Robo2 and Robo3 are also possible (Camurri et al., 2005; Hernández-Miranda et al., 2011; Hivert, 2002). For instance, it was shown that dimerisation of *Drosophila* Robo2 is mediated by its Ig3-Ig5 domains (Evans and Bashaw, 2010a). Interaction of Robo1 with Robo4 has also been suggested (Kaur et al., 2006; Sheldon et al., 2009), but surface plasmon resonance (SPR) experiments did not detect a direct interaction (Koch et al., 2011; Suchting et al., 2005).

A different study, however, illustrated that the oligomerisation state of Robo1 does not change regardless of the presence of Slit (Zakrys et al., 2014). Taken together, these studies suggest a model where Slit binding induces conformational changes necessary for receptor activation. This is further supported by negative stain electron microscopy, and *in vitro* proximity ligation assay data from our group (Aleksandrova et al., in press) that supports the idea that Robo1 exists as a dimer

INTRODUCTION

in a putative inactive conformation, which requires interaction with Slit2 to induce the changes required for downstream signalling.

To add another layer of complexity, both *Drosophila* and mammalian Robo1 can undergo cleavage by matrix metalloproteinases at the extracellular domain (Coleman et al., 2010) and by γ -secretases at the cytoplasmic domain (Seki et al., 2010). Shedding of the receptor influences the downstream signalling cascade, but, although this mechanism is supported by biochemical studies on Robo1 (Barak et al., 2014), it has not been conclusively shown to be Slit dependent, nor how important it is for signalling.

1.3.2 Robo4 signalling and angiogenesis

Robo4 is the most recently discovered protein of the Roundabout family (Huminiecki and Bicknell, 2000). It is mainly expressed in endothelial cells and at sites of active angiogenesis (Huminiecki et al., 2002; Park et al., 2003). A 3 kilobase region upstream of the *Robo4* gene contains the binding site for transcription factors of the EST family, which are lineage specific for endothelial expression (Okada et al., 2007). Robo4 was also reported to be expressed by hematopoietic stem cells and vascular smooth muscle cells (Liu et al., 2006; Shibata et al., 2009; Smith-Berdan et al., 2011).

Unlike other members of the Robo family, the extracellular domain of Robo4 only contains two Ig and two FnIII domains and, in the cytoplasmic part, only the CC0 and CC2 motifs are conserved (Figure 3). While it has no apparent role in axonal guidance, it has a fundamental role in both developmental and pathological angiogenesis (Koch et al., 2011; Yadav and Narayan, 2014).

In zebrafish, Robo4 is essential for proper vascular vessel growth during early embryonic development (Bedell et al., 2005). It was suggested that Robo4 induces cell migration through activation of Cdc42 and the Ras-related C3 botulinum toxin substrate 1 protein (Rac1) (Kaur et al., 2006). Both Cdc42 and Rac1 are Rho GTPases that induce cytoskeletal remodelling and filipodia reorganisation by intervening on actin and microtubules. In endothelial cells, this mechanism influences filipodia reorganisation, directing cell migration and adhesion (Schnittler, 2016; Wen et al., 2017). In neuronal cells the same pathways are regulated through Robo1 (albeit through Slit-dependent inhibition), and result in axonal guidance (Ghose and Van Vactor, 2002).

There are still many open questions on the signalling mechanism of Robo4, and on its downstream signalling pathway. The responses can be different depending on the cell type, but also on the nature of its extracellular ligands.

For instance, it was shown by co-immunoprecipitation that Slit2 could interact with Robo4. This interaction was further shown to inhibit cell migration of human microvascular embryonic cells (HMVEC) (Park et al., 2003). On the other hand, when studying the effect of Slit proteins on angiogenesis, other studies reported how Slit2 acts as a strong angiogenic inducer in human umbilical vein endothelial

INTRODUCTION

cells (HUVEC) in presence of Robo4, while co-presence of Robo1 and Robo4, through an unknown mechanism, inhibited this effect (Wang et al., 2003). Further studies showed that a recombinant soluble form of the Robo4 extracellular domain inhibited angiogenesis both *in vivo* and *in vitro*, and yet, no direct interaction with Slit2 was observed (Suchting et al., 2005). In a different context, another study that evaluated the effect of Robo4/Slit signalling on the vascular endothelial growth factor (VEGF), and other associated proteins (most notably, MEK, ERK and FAK), concluded that in the presence of Slit2, Robo4-induced inhibition of endothelial migration will lead to pathological angiogenesis (Seth et al., 2005). Furthermore, in mammals, Slit3 was shown to induce endothelial vascular development when interacting with Robo4 (Zhang et al., 2009).

Ultimately, Robo4 signalling, and its interaction with Slit proteins, was shown to be varied, depending on multiple factors and on the presence of other Robo receptors. In a recent study on HUVEC cells, it was demonstrated that while Slit2/Robo1 will promote HUVEC migration, the presence of Robo4 negatively regulates this effect, inhibiting migration (Enomoto et al., 2016).

While these studies link Robo4 activity with the presence of Slit, a direct interaction between the two was disproved (Koch et al., 2011), and several of the critical Slit binding residues are absent in Robo4 (Hohenester, 2008; Morlot et al., 2007). Since Robo1, Robo2 and Robo4 expression overlaps in some cell types, modern models suggest that Slit function on Robo4 is mediated through receptor heterodimers. This was indeed shown *in vitro*, where Robo1/Robo4 heterodimers promote cell migration (Sheldon et al., 2009).

A different mechanism of Robo4/Slit signalling includes recruitment of the membrane-associated protein paxillin (Turner et al., 1990), which inhibits the activity of GTPase ADP-ribosylation factor 6 (Arf6). Promoting, in this way, vascular stability and reduced angiogenesis (Jones et al., 2009).

Slit and other Robos, however, are not the only possible binding partners of Robo4 at the extracellular level. UNC5B (which is discussed in detail later) was shown to directly interact with Robo4 (Koch et al., 2011). In this case, the signalling appeared to be mediated through the UNC5B downstream signalling pathway, instead of Robo4.

1.3.3 Involvement of Robo4 in pathological angiogenesis

Many of the downstream pathways of Robo4 involve the activation, or inhibition, of VEGF (Jones et al., 2008; Marlow et al., 2010), and deregulation of Robo4 was directly observed in colorectal (Gröne et al., 2006), bladder (Li et al., 2015b) and endothelial cell cancer (Seth et al., 2005). For this reason, Robo4 is now considered a potential tumour angiogenesis marker (Legg et al., 2008; Seth et al., 2005), and there have already been studies on how to interfere with its function.

For example, enhanced Robo4 signalling can suppress breast cancer growth and metastasis (Zhao et al., 2016). *In vivo* mouse experiments, showed that vaccines against Robo4, and development of anti-Robo4 antibodies, can reduce the growth of Lewis lung carcinoma (Zhuang et al., 2015). Furthermore, anti-Robo4 antibodies can accumulate by internalisation at sites of uncontrolled angiogenesis and can be conjugated with drugs for specific targeting (Yoshikawa et al., 2013). Developing drugs against Robo4 can also help in reducing cancer development by controlling the angiogenesis of the lymphatic system (Yu et al., 2014), which is a vehicle for tumour metastasis.

The original aim of this study is to gain a deeper understanding on how Robo4 binds to its partners and induces signalling. Obtaining structural information of the extracellular domain of Robo4, however, will also allow for the development of better therapeutic strategies.

INTRODUCTION

1.4 The UNC5 protein family

The genes of the Uncoordinated (Unc) protein family were initially discovered in *C. elegans*, where it was observed that mutants in those genes showed defects in axonal growth towards specific directions within the nervous system (Hedgecock et al., 1990).

The genes identified included both the secreted guidance cue *Netrin* (formerly known as *Unc-6*, now separated in its own family) and two of its transmembrane receptors, *Unc-40* and *Unc-5*. The former, *Unc-40*, was identified as the *C. elegans* homolog of the vertebrate DCC receptor (Chan et al., 1996), which was discovered the same year (Keino-Masu et al., 1996). The latter, *Unc-5*, encodes for a single pass transmembrane protein, expressed specifically on the surface of migrating axons and growth cones (Leung-Hagesteijn et al., 1992). Homologs of *Unc-5* were later identified in *Drosophila* (Keleman and Dickson, 2001) and vertebrates (Ackerman et al., 1997; Engelkamp, 2002; Leonardo et al., 1997), where it was observed that disruption of mammalian *Unc-5* genes caused malformations during cerebellum development (Przyborski et al., 1998).

The proteins they translate are now commonly referred to as the UNC5 protein family.

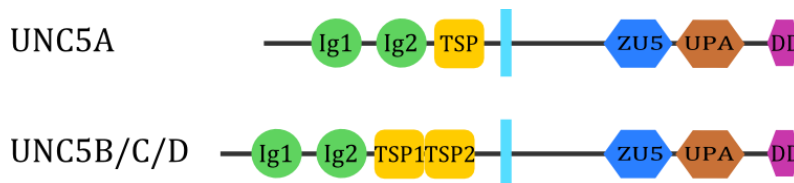


Figure 6: General structure of vertebrate UNC5 proteins

All UNC5 proteins have the same domain organisation, except UNC5A which has one less TSP domain. In the N-terminal extracellular region are the Ig (green) and TSP (yellow) domains. Then there is a single transmembrane α -helix (sky blue). In the cytoplasmic region are ZU5 (blue), UPA (light brown) and DD (purple).

In vertebrates, there are four UNC5 homologs: UNC5A, UN5CB, UNC5C and UNC5D. All four vertebrate UNC5 proteins share the same domain composition (Figure 6), with the exception of UNC5A, which has only a single TSP domain, and are between 96 and 103 kDa. The N-terminal extracellular part is composed of two Ig and two thrombospondin type 1 (TSP) domains. The Ig domains are the main interaction region of the receptor, while the TSP domains are a structurally recurring motif of

extracellular and matrix interacting proteins. The C-terminal cytoplasmic region contains three distinct domains, which are responsible for the recruitment of downstream signalling factors. The first is the ZU5 domain, which takes its name by the combination of Zona Occludens-1, which is a homologous tight junction protein, and UNC5. UPA is a conserved UNC5/PIDD/Ankyrin domain. Finally, DD is a death domain.

The functions assumed by the UNC5 protein family can be classified in three main categories: guidance (neuronal or vascular), apoptosis and angiogenesis.

1.4.1 UNC5 proteins in guidance

The four UNC5 proteins are partially redundant, but they assume different functions dependent on the cell type and ligands (extracellular and cytoplasmic) present. UNC5A is almost exclusively expressed in neuronal tissue, UNC5B and UNC5C are found in all types of tissue, while UNC5D is found mostly in neuronal and some specialised epithelium cells (Fagerberg et al., 2014). In some cases there are differences in the pathways they stimulate, but the whole array of functions they regulate has not been fully elucidated.

Neuronal guidance is the first phenomenon discovered that UNC5 proteins influence and UNC5B is the best characterised homolog, which acts on both guidance and angiogenesis. Unless a specific UNC5 homolog is mentioned from now on, it should be assumed that any characteristic applies to all of them.

UNC5 receptors exert their main functions by binding to the secreted members of the Netrin protein family (Keino-Masu et al., 1996; Leonardo et al., 1997). In mammals four members of the Netrin family are secreted: Netrin-1, Netrin2, Netrin4, and Netrin-5 (Lai Wing Sun et al., 2011). Netrin-1 is mostly involved in neuronal guidance and is the best described (Wilson et al., 2006), while Netrin-4 is more involved in angiogenesis (Lejmi et al., 2008). Two other membrane-anchored members are present (NetrinG1 and NetrinG2), but they do not interact with the UNC5 receptors (Goldman and Kennedy, 2011; Lai Wing Sun et al., 2011).

Interaction of UNC5 proteins with Netrin is mediated through the UNC5 Ig domains (Geisbrecht et al., 2003; Grandin et al., 2016). The Ig domains, however,

INTRODUCTION

are not redundant: deletion of Ig1 reduces binding, but deletion of Ig2 entirely abolishes it (Grandin et al., 2016; Kruger et al., 2004).

Attraction and repulsion are regulated by the combined actions of UNC5 receptors and the second Netrin receptor DCC (that is not discussed here), which have competing activities (Figure 7).

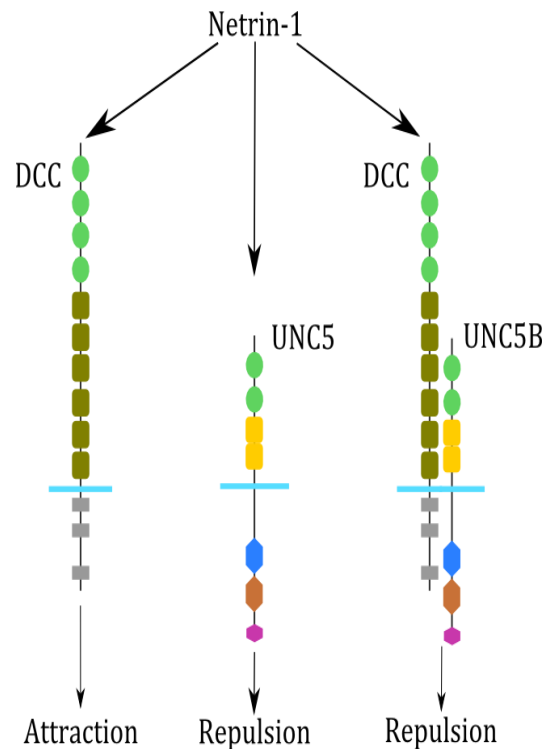


Figure 7: UNC5 and DCC influence on guidance.

Green: Ig domains. Olive: FnIII domains. Yellow: TSP domains. Grey: DCC specific cytoplasmic domains. Blue: ZU5 domain. Light brown: UPA domain. Purple: DD. Sky blue: cell membrane. Binding of Netrin to UNC5 causes repulsion, while binding to DCC causes attraction. When UNC5B and DCC form a heteromeric complex, they induce repulsion.

Netrin-1 binding to UNC5 causes repulsion (Hedgecock et al., 1990; Hong et al., 1999; Keleman and Dickson, 2001; Merz et al., 2001), while Netrin-1 binding to DCC causes attraction (De La Torre et al., 1997; Ly et al., 2008). Furthermore, apart from having an opposite effect, DCC binding to Netrin is not mediated through the Ig domains of DCC, but through its fifth FnIII domain (Geisbrecht et al., 2003).

UNC5B and DCC can also form a heteromeric complex, which has been shown to mediate repulsion upon Netrin-1 binding (Hong et al., 1999). Interestingly, the UNC5B extracellular domain is dispensable for this effect, and binding of UNC5 to DCC is mediated through the UNC5 UPA domain (Hong et al., 1999). Furthermore, the extracellular domains of UNC5B and DCC do not interact at all (Geisbrecht et

al., 2003; Hong et al., 1999). This showed that interaction of the cytoplasmic domain of UNC5B with the cytoplasmic domain of DCC is sufficient to switch DCC mediated signalling from attraction to repulsion.

The downstream signalling pathway mediated by the DCC/Netrin interaction is the best characterised (Finci et al., 2015; Gitai et al., 2003), and involves the recruitment of kinases and adaptor proteins involved in cytoskeleton remodelling. The effects of Netrin binding to UNC5 have still not been entirely addressed and remain contradictory. Some studies have shown that one of the direct effects during UNC5/Netrin interaction, is the inhibition of basal tyrosine phosphorylation levels (Kruger et al., 2004). However, others have also shown that the same interaction leads to recruitment of the SRC-1 kinase in *C. elegans* and phosphorylation of its UNC5 homolog (Lee et al., 2005). The same increase in phosphorylation has been shown in vertebrate UNC5 homologs, which also recruit tyrosine phosphatases (Tong et al., 2001). Regardless of the exact mechanisms, recruitment of kinases, and changes in phosphorylation, seem to be a common response, as already discussed in Robo signalling (§§ 1.3.1 and 1.3.2). In addition, they are also observed to occur in another class of guidance cues, the Semaphorins (Takahashi and Strittmatter, 2001).

While axon guidance mediated by UNC5 and DCC interaction with Netrin is effective at short range, another Netrin receptor was suggested to capture Netrin, and present it for recognition by other receptors, in *Drosophila*: Frazzled (Hiramoto et al., 2000; Keleman and Dickson, 2001). The mechanisms behind Frazzled interaction, and its cross-talk with the UNC and DCC pathways, have still not been completely elucidated (Akin and Lawrence Zipursky, 2016; Manhire-Heath et al., 2013).

1.4.2 UNC5 proteins in apoptosis

Due to the presence of the cytoplasmic DD (Figure 6), UNC5 are classified as dependence receptors. Furthermore, the *UNC5B* gene (also known as *p53RDL1*) is regulated by p53 (Tanikawa et al., 2003), which is well known for its involvement in cancer and apoptosis.

INTRODUCTION

Active DD domains induce a signalling cascade of proapoptotic factors (Bredesen et al., 2005; Hofmann and Tschopp, 1995; Mehlen and Bredesen, 2004; Mehlen et al., 1998; Tanikawa et al., 2003). Binding of netrin proteins is necessary to inhibit signalling through the DD (Castets et al., 2009; Thiebault et al., 2003). The crystal structure of the DD inactive conformation was solved and shows how DD inhibition depends on interaction with UPA and ZU5 (Wang et al., 2009) (Figure 8). This fold is a common feature of proteins containing these domains, and a similar structure was identified in Ankyrin-B (Wang et al., 2012). Opening of this complex exposes DD to proteolytic cleavage by caspases (Llambi et al., 2001), which frees DD for dimerisation (Wang et al., 2009), and recruitment of death associated protein kinase (DAPK) (Llambi et al., 2005). This starts a common signalling cascade which eventually leads to cell death (Park et al., 2007a, 2007b; Weber and Vincenz, 2001; Xiao et al., 1999).

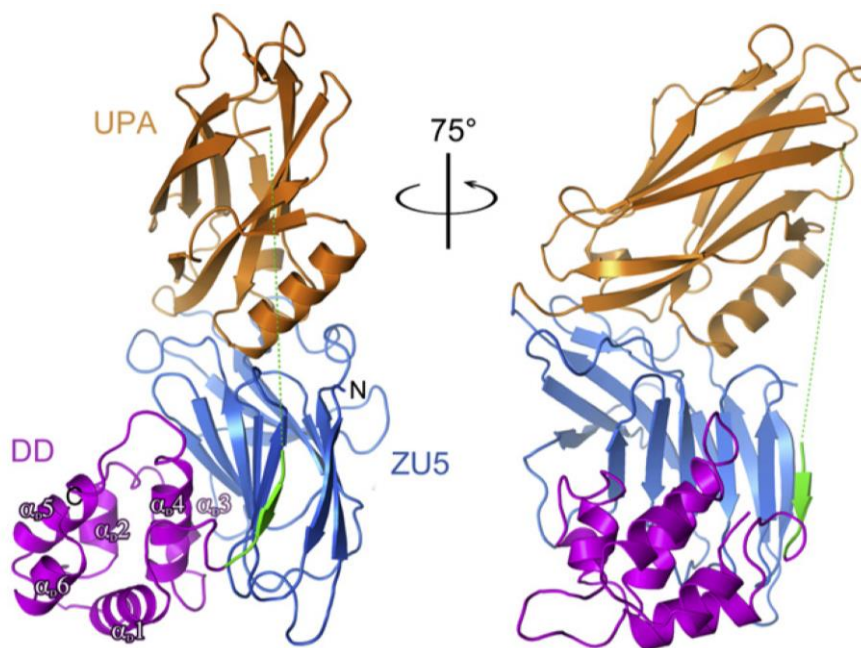


Figure 8: Structure of the cytoplasmic ZU5, UPA and DD domains of UNC5B

Cartoon representation of the closed ZU5 (blue), UPA (orange), and DD (purple) domains. The C-terminal DD domain fold over the UPA to contact the N-terminal ZU.

Image from (Wang et al., 2009).

1.4.3 UNC5B in angiogenesis

Of the UNC5 receptors, UNC5B is the one most implicated in angiogenic processes. In murine embryos, while normally downregulated in adult vasculature, UNC5B is specifically expressed in developing blood vessels (Larrivéé et al., 2007; Lu et al., 2004), and its deletion results in aberrant filipodia formation and unregulated vessel branching *in vivo*, leading to early death of mouse embryos (Lu et al., 2004). It was later shown that this lethal deletion of UNC5B *in vivo* is caused by a disruption of placental labyrinth vascularisation (Navankasattusas et al., 2008). UNC5B alone promotes endothelial cell migration (Köhler et al., 2013; Lu et al., 2004), but Netrin-1 interaction with UNC5B, instead, inhibits it (Larrivéé et al., 2007). This was confirmed upon corneal injury, where this interaction mediates anti-inflammatory effects by reducing apoptosis, neovascularisation and blocking neutrophil and macrophage infiltration (Han et al., 2012). The last, which is effectively an inhibition of cell migration, is likely related to filipodia retraction stimulated by repulsive effects (Lu et al., 2004).

On the role of Netrins in vascularisation there are still controversial opinions. For instance, it was shown that Netrin-1 positively influences vascularisation (Park et al., 2004), but not if this effect is mediated through UNC5B or DCC (Nguyen and Cai, 2006). Furthermore, both proangiogenic (Nguyen and Cai, 2006; Park et al., 2004; Wilson et al., 2006) and antiangiogenic effects were described (Larrivéé et al., 2007; Lu et al., 2004). It was suggested that the concentration of Netrin-1 is important, where low doses stimulate vascularisation and high doses inhibit it (Yang et al., 2007). Netrin-4, instead, inhibits angiogenesis by binding to Neogenin, which then sequesters UNC5B in a 1:1:1 trimeric complex (Lejmi et al., 2008).

UNC5B is also downregulated in several types of cancers, including breast, colorectal, stomach, lung and kidney cancer (Baker et al., 2006; Klagsbrun and Eichmann, 2005). Furthermore, UNC5A, UNC5B and UNC5C act as tumour suppressors, by inducing apoptosis in absence of Netrins and inhibiting malignant cell migration in its presence (Thiebault et al., 2003).

The extensive effect of UNC5B on physiological and pathological vascularisation, and cell migration, makes it an excellent target of study for therapeutic strategies aimed at influencing these processes.

1.5 Basics of N-linked glycosylation

Glycosylation is one of several post-translational modifications (PTM) that proteins undergo during or after biosynthesis, and is often necessary for their correct folding or function (Knorre et al., 2009). It's a common modification found on secreted and membrane proteins, which requires several steps of addition and removal of monosaccharides and glycans (single carbohydrates or assemblies of carbohydrates). The first step is performed within the endoplasmic reticulum (ER), where the protein is first translocated during translation (Aebi, 2013). Afterwards, consequential steps while the protein travels from the ER to the Cis-, Medial- and Trans-Golgi lead to complete maturation. Once all the correct modifications have been completed, the mature protein is exported for secretion or insertion in the plasma membrane (Caramelo and Parodi, 2015). Two main types of glycosylation are possible on glycoproteins: O-linked glycosylation that will not be discussed here, and N-linked glycosylation.

N-glycans are covalently attached to the protein backbone through an N-glycosidic bond on asparagine residues at specific sites where the signature sequence is: Asn-X-Ser/Thr (where X is any amino acid except proline).

Figure 9 shows a typical representation of the three common types of N-glycans, which share a common core of two N-acetylglucosamine (NAG) and 3 mannose glycans.

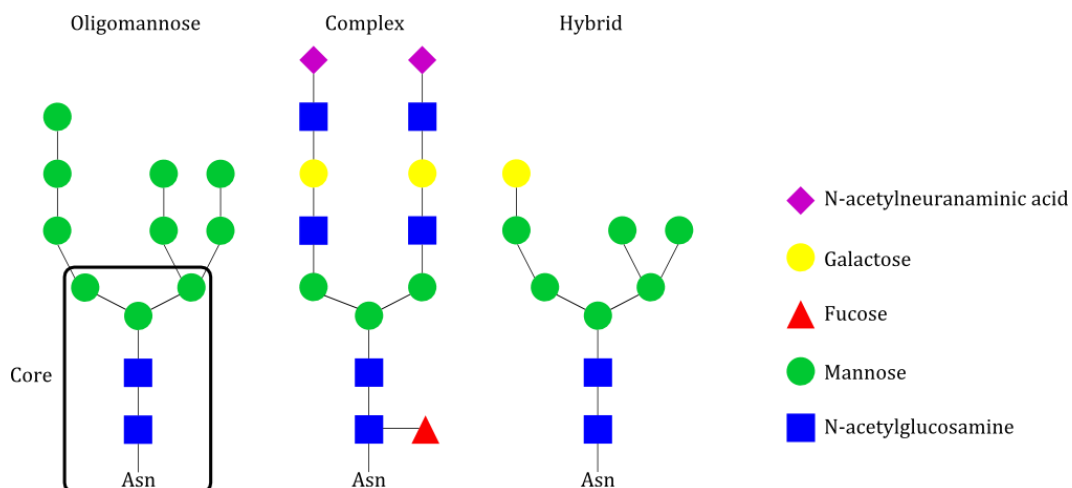


Figure 9: The three main types of N-linked glycosylation

Oligomannose, complex or hybrid glycosylation are possible on all types of glycoproteins. The core region shared by all three types of N-linked glycans is composed by two NAG and one mannose in a linear chain and two more branched mannose (black box).

Although the core is invariable, there are 18 common monosaccharides which can be combined in several variants to form the terminal glycans.

Glycosylation (and its heterogeneity) is important on several levels, influencing protein folding (Shental-Bechor and Levy, 2008), function (Parekh, 1991), binding, and receptor signalling (Arey, 2012). Furthermore, changes in protein glycosylation have an important influence in cancer development (Varki et al., 2017).

1.6 Synthetic antibodies

Thanks to their diversity, specificity and versatility, monoclonal antibodies have become an invaluable tool for use in pathological diagnosis, treatment of infectious diseases, and cancer (Keller and Stiehm, 2000; Leavy, 2010; Weiner, 2015).

An Ig type G antibody (IgG), is composed of two heavy chains connected by two disulphide bridges, and a light chain associated to each heavy chain, connected by a single disulphide bridge (Figure 10A). The C-terminal part of the antibody containing the two connected heavy chain is called the crystallisable fragment (Fc) region, which is responsible for interaction with cell surface receptors and proteins of the complement system. The remaining part, containing a fraction of

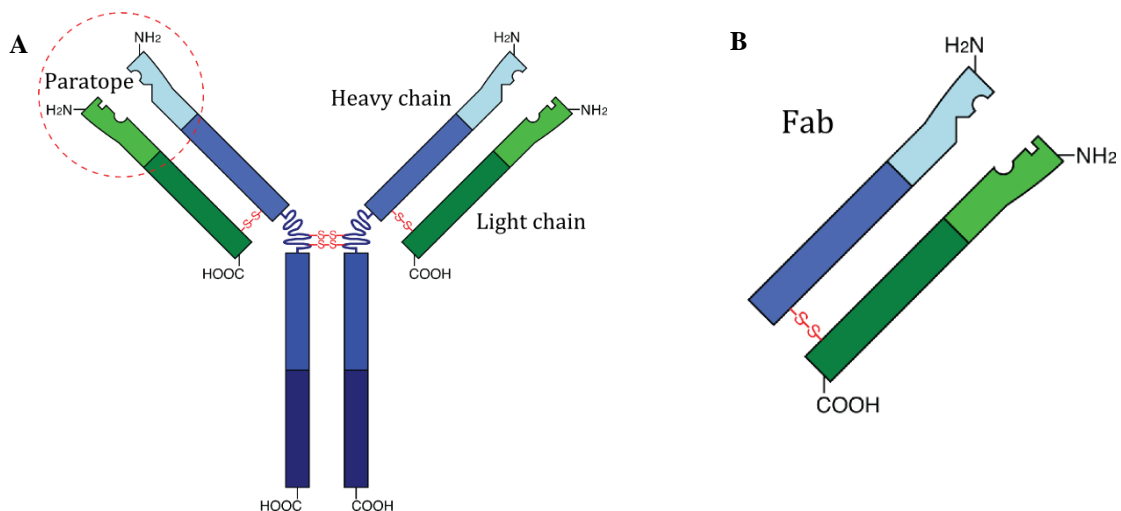


Figure 10: General IgG and Fab structure

A: Typical structure of an IgG. The two heavy chains (blue) are connected by disulphide bridges. Each light chain (green) is connected to the respective heavy chain by another disulphide bridge. The binding region (paratope) is within the variable region, highlighted by a red circle.

B: Fragment antibodies are composed by one light chain and part of the heavy chain connected by a single disulphide bridge.

Modified from ["Practical immune systems", Uppsala university, Sweden.](#)

INTRODUCTION

the heavy chain connected to the light chain, is called the fragment antigen binding (Fab) region (Figure 10B). This region contains the recognition region (paratope), and its specificity is derived by the complementary determining regions (CDR). The CDR is composed of three short amino acid stretches per chain.

Over the years, several variant antibodies have been created and the use of synthetic antibodies (produced *in vitro*) is quickly superseding the use of *in vivo* sources. The most important advantages of synthetic antibodies are the possibility of expanding the available paratopes, which can't be produced in the natural immune system (Shim, 2015). This allows for a more rational design, or selection, of particular binding regions on specific targets, and can help to avoid unwanted immune responses (Adams and Sidhu, 2014; Bradbury et al., 2011). To this end, several technologies have been developed, with a strong focus towards clinical applications (Deyev and Lebedenko, 2009). Reducing the size of synthetic antibodies to smaller functional units, furthermore, allows for easier tissue penetration and delivery of therapeutic effects (Jain, 1990; Yokota et al., 1992). Synthetic antibodies, however, are also important tools in manipulating protein function (Paduch et al., 2013), in protein screenings (Säll et al., 2016), and in crystallography (Dominik et al., 2016; Tereshko et al., 2008).

Anti-Robo4 antibodies have already been successfully used to interfere in angiogenesis (Zhuang et al., 2015), and to identify sites of abnormal vascularisation (Yoshikawa et al., 2013). Recombinant Fabs produced in *E. coli* have been used in this particular study (Figure 10B).

1.7 Aim of the study

While Robo4 binding seems important for UNC5B activation (Koch et al., 2011), there are still many open questions about how this interaction is relayed across the membrane for signalling. As both are single pass transmembrane proteins, a change in oligomerisation is likely implicated (Alberts et al., 2002; Moore et al., 2008) and several studies have shown, or inferred, the formation of homophilic and heterophilic species of both receptors (Geisbrecht et al., 2003; Hivert, 2002; Hong et al., 1999). Furthermore, endocytosis of Robo1 was suggested to be triggered by ligand binding and involved in the signalling cascade (Chance and Bashaw, 2015), and it was proposed this mechanism might be shared by all members of the Robo family (Charron, 2015). The activation of each receptor can trigger different effects, depending on the presence of other factors. For instance, DCC switches from attraction to repulsion in presence of UNC5B (Hong et al., 1999) and the proposed Slit2/Robo4 interaction is proangiogenic, but it changes to antiangiogenic in presence of Robo1 (Wang et al., 2003). Since Robo4 and UNC5B heteromerisation was described (Koch et al., 2011), a characterisation of this complex would shed light on the cross talk between different guidance pathways. The major focus of this study was to gain functional insights into how an interaction between Robo4 and UNC5B is relayed across the membrane for intracellular signalling. The work presented here is mainly focused on obtaining structural and biophysical information on the extracellular domains of the Robo4 and UNC5B receptors, individually and in complex, to ultimately determine the molecular details of their interaction. In order to achieve this, high quality recombinant Robo4 and UNC5B were produced. Synthetic Fabs produced by the Sidhu lab to bind the Robo4 extracellular domain were also employed. Initially it was hoped to identify those Fabs that inhibited the interaction of Robo4 with UNC5B for further characterisation. This was in order to prepare the groundwork for future studies on influencing the Robo4 signalling pathway, or to otherwise help the development of new classes of Fabs as potential therapeutics. Later, it was hoped that Robo4 binding Fabs could stabilise the receptor in a conformation that would facilitate its crystallisation, in the guise of crystallisation chaperones (Dominik et al., 2016; Tereshko et al., 2008).

2 MATERIALS AND METHODS

Résumé en français

Comme les modifications post-traductionnelles sont importantes pour la stabilité et les propriétés de liaison des récepteurs transmembranaires et des protéines sécrétées, des systèmes d'expression eucaryotes ont été utilisés pour la production de protéines dans cette étude. Toutes les constructions ont été exprimées sous forme de protéines solubles. Les constructions de Robo4 ont été exprimées dans un système d'expression de cellules de mammifère et les constructions d'UNC5B dans des cellules de mammifères et des cellules d'insectes.

Des techniques biochimiques telles que la chromatographie par exclusion de taille (SEC), la diffusion de lumière multi-angles (MALS) et la résonance plasmonique de surface (SPR) ont été utilisées pour étudier la liaison des domaines extracellulaires Robo4 et UNC5B l'un à l'autre. La liaison de fragments de six anticorps monoclonaux (choisis parmi notre collaborateur) au domaine extracellulaire de Robo4 a également été testée. Les constantes cinétiques de l'interaction entre le domaine extracellulaire de Robo4 et chaque fragment d'anticorps ont été déterminées en utilisant la résonance plasmonique de surface.

Des protocoles de deglycosylation ont été mis au point pour étudier l'effet de la glycosylation hétérogène et pour améliorer les chances de cristallisation des domaines extracellulaires de Robo4 et d'UNC5B.

Des informations structurales sur le domaine extracellulaire de Robo4, et le complexe formé entre ce domaine et chaque fragment d'anticorps, ont été étudiées en utilisant la diffusion des rayons X aux petits angles (SAXS).

Pour améliorer les chances de cristallisation, la technique de micro-ensemencement à matrice aléatoire a été utilisée et des techniques de cristallographie aux rayons X ont été utilisées pour obtenir des informations structurales sur le domaine extracellulaire d'UNC5B.

MATERIALS AND METHODS

2.1 Supplies

A list of materials used in this study (including kits and reagents) is provided in Table 1. Separately, a list of all materials involved in cell culture is provided in Table 2.

Table 1: Supplies, kits and other materials

Material	Supplier	Catalogue number
1 kb DNA ladder	NEB	N3232S
100 bp DNA ladder	NEB	N3231S
Acrylamide 4K solution 30%	AppliChem	A1672
Agarose D5	Euromedex	D5-D
Amino coupling kit	GE Healthcare	BR100050
Ampicillin	Euromedex	EU0400-D
Antibody anti-His primary mouse	GE Healthcare	27-4710-01
Antibody anti-His primary mouse	Sigma	H1029
Antibody anti-mouse secondary AlexaFluor532	ThermoFisher	A-11002
Antibody anti-mouse secondary HRP	ThermoFisher	31439
BioLock biotin blocking solution	IBA	2-0205-050
Bio SEC-3 gel filtration column	Agilent	5190-2511
Chloramphenicol	Euromedex	3886-A
CM5 sensor chip	GE Healthcare	BR100530
Colour prestained protein standards	NEB	P7712S
cOmplete protease inhibitor	Roche	5056489001
dNTP mix	NEB	N0447L
<i>E. coli</i> strain BL21 RIL (DE3)	Agilent	230265
<i>E. coli</i> strain DH5 α	Invitrogen	18258012
Endo H	NEB	P0702S
Glutathione Sepharose 4 Fast flow resin	GE Healthcare	71-5016-97
HiLoad Superdex 200 16/600 column	GE Healthcare	28-9893-35
HiPrep 26/10 Desalting column	GE Healthcare	17-5087-01
HiTrap MabSelect SuRe column	GE Healthcare	11-0034-94
HiTrap MBPTrap HP column	GE Healthcare	28-9136-32
Kanamycin	Euromedex	EU0420
Immobilon-P PVDF membrane	Millipore	IPVH00005
Isopropyl β -D-1-thiogalactopyranoside	Euromedex	EU0008-B
MicroMesh loop	Mitegen	M3-L18SP-10
Nickel Sepharose Excel	GE Healthcare	17-3712-01

NucleoBond Xtra plasmid Maxiprep kit	Macherey-Nagel	740414
NucleoSpin Gel and PCR clean-up kit	Macherey-Nagel	740609
NucleoSpin plasmid Miniprep kit	Macherey-Nagel	740588
PACT premier screen	Molecular Dimensions	MD1-36
pFastBac1	Invitrogen	10360014
Phusion HF Polymerase kit	NEB	M05305
Plasmid Giga kit	QIAGEN	12191
PNGase F1	NEB	P0704S
Pur-A-Lyzer mini dialysis tubes 6 kDa	Sigma	PURN60100
Purified agar agar	Euromedex	1329-D
Q5 Polymerase	NEB	M04915
SIGMAFAST diaminobenzidine tablet	Sigma	D4168
Source 15Q 4.6/100 PE column	GE Healthcare	1751801
Stericup 500 ml 0.22 µm	Millipore	SCGPU10RE
Strep-Tactin buffer E	IBA	2-1000-025
Strep-Tactin buffer W	IBA	2-1003-100
Strep-Tactin resin	IBA	2-1201-010
Superdex 200 5/150 column	GE Healthcare	28906561
Superdex 200 Increase 10/300 column	GE Healthcare	28990944
T4 DNA Ligase	NEB	M0202S
Tween20	Sigma	P9416

MATERIALS AND METHODS

Table 2: Cell culture materials

Chemicals	Supplier	Catalogue number
Di-methyl sulfoxide	Sigma	D2438
Dulbecco's modified Eagle's medium	Sigma	D5796
ExpressFIVE SFM	Gibco	10486-025
Fetal bovine serum	Gibco	10270-106
HEK293 GnTI-	ATCC	CRL-3022
HEK293T/17	ATCC	CRL-11268
Improved Neubauer counting chamber	Sigma	BR717805
L-glutamine	Gibco	25030-024
MEM Non-essential amino acids	Gibco	11140-035
Opti-MEM	Gibco	111058-021
Penicillin-Streptomycin	Gibco	15140-122
Polyethylenimine 25kDa branched	Sigma	408727
Sf900 II SFM	Gibco	10902-088
Sodium pyruvate	Gibco	11360-039
Trypan blue	Sigma	T8154
Trypsin-EDTA 0.05% (v/v)	Gibco	25300-096
X-tremeGENE HP DNA transfection reagent	Sigma	000000006366236001

2.2 Preparation of *E. coli* competent cells

Chemically competent *E. coli* cells were prepared following a modified protocol derived from (Sambrook and Russell, 2001). Briefly, a single colony was picked from a plate containing, if necessary, the suitable antibiotic(s). After overnight incubation in 5 ml of Miller's lysogeny broth (LB) media, this culture was used to inoculate 200 ml of fresh LB media and allowed to grow until reaching an optical density (OD) of 0.48 at 600 nm. The culture was then incubated on ice for 15 minutes (min) and centrifuged at 1900×g for 5 min at 4°C. The pellet was resuspended in 20 ml of solution containing 30 mM potassium acetate pH 5.8, 100 mM KCl, 10 mM CaCl₂, 50 mM MnCl₂ and 15% (v/v) glycerol before incubation on ice for 5 min. Afterwards, the solution was again centrifuged at 1900×g for 5 min at 4°C, the supernatant discarded, and the pellet resuspended in 2 ml of buffer containing 10 mM MOPS pH 6.5, 10 mM KCl, 75 mM CaCl₂, and 15% (v/v) glycerol. After 15 min incubation on ice, the bacteria containing solution was aliquoted into single use 50 µl aliquots, flash frozen and stored at -80°C.

2.3 Cloning of Robo4 and UNC5B constructs

The reference protein sequence of Robo4 used in this study was taken from the [Uniprot](#) database (entry [Q8WZ75](#)). The protein sequence was reverse translated, codon optimized, and supplied as a synthetic gene using the GeneArt service offered by Thermo Fisher Scientific. Through this process, restriction sites within the open reading frame were substituted and the codons optimized to achieve optimal protein expression in a mammalian cell expression system.

The template DNA of UNC5B was a gift of Elena Seiradake (University of Oxford) and encodes for the protein corresponding to entry [Q8IZI1](#) from [Uniprot](#).

Protein constructs (Figure 11) were amplified by polymerase chain reaction (PCR) using the primers listed in Table 3.

MATERIALS AND METHODS

Table 3: Robo4 and UNC5B cloning and mutagenesis primers

Green: restriction sites (BamHI, EcoRI, HindIII, KpnI, NcoI, NotI, and SacI). **Purple:** honey bee Melittin signal sequence. **Blue:** strep tag. **Red:** protein coding sequence. **Brown:** mutation site.

FW stand for forward primer. RV: reverse primer. Addition of “ec” indicates primers used for cloning into *E. coli* expression vectors.

	Sequence (5' to 3')
Robo4 25 FW	AAAGGTACCGGAATGGCACAGGATAGCCC
Robo4 28 FW	AAAGGTACCAGGATAGCCCC
Robo4 228 FW	AAAGGTACCCCCAGGACTACACCGAG
Robo4 245 FW	AAAGGTACCGAAAACGTGACCCTGCTGA
Robo4 130 RV	AAAGAGCTCCAGTCTGGCGCCTCTAGAC
Robo4 231 RV	AAAGAGCTCGTAGTCCTGGGGTTCCTGG
Robo4 430 RV	AAAGAGCTTTCGCCTGCGCCAG
Robo4 444 RV	AAAGAGCTTTCATGGCCTGTTCCAG
Robo4 462 RV	AAAGAGCTCCGCAGCTGTTCCAGTGT
Robo4 467 RV	AAAGAGCTCCGCTTCAGTGTGGCC
Robo4 N246D FW	CAGCTGGAAGACGTGACCCTGCTGAACCCGATC
Robo4 N246D RV	TCACGTCTTCCAGCTGAATCCGCACGGCCAGC
Robo4 N246Q FW	GCGGATTCAGCTGGAACAGGTGACCCTGCTGAAC
Robo4 N246Q RV	GTTCCAGCAGGGTCACCTGTTCCAGCTGAATCCGC
Robo4 N360D FW	GCCTGGCGACGGCACCGTGTTCGTGTCCTG
Robo4 N360D RV	TGCCGTGCCAGGCTTCAGGGTCACTTCCTG
Robo4 N360Q FW	CTGGCCAGGGCACCGTGTTCGTGTCCTGG
Robo4 N360Q RV	GGTGCCTGGCCAGGCTTCAGGGTCACTTC
Robo4 N389D FW	CTGGGCGACACTCTCTGCCCCCTGCCAATTG
Robo4 N389D RV	GAGGTGTCGCCAGGGACCACACTTGGTAGC
Robo4 N389Q FW	CTGGGCGACACTCTCTGCCCCCTGCCAATTG
Robo4 N389Q RV	AGAGGTGTCGCCAGGGACCACACTTGGTAGC
Robo4 N396D FW	CCTGCCGACTGGACCGTCGTGGGAGAGCAG
Robo4 N396D RV	GTCCAGTCGGCAGGGGGCAGAGAGGTGTTG
Robo4 N396Q FW	CCTGCCAGTGGACCGTCGTGGGAGAGCAG
Robo4 N396Q RV	GTCCACTGGCAGGGGGCAGAGAGGTGTTG
UNC5B 27 honey FW	AAAGGATCCATGAAGTTTTTGGTCAACGTCGCCTTGGTGTTCATGGTCGTGACATCA GCTACATCTATGCGGCCGCTGGCACTGATTCTGGCAGC
UNC5B 377 strep RV	TTTAAGCTTTTACTTCTCGAACTGAGGGTGGGACCAATACAGCGCCGCATCCC
UNC5B 245 FW ec	AAACCATGGTGAATGGCGGCTGGTC
UNC5B 354 RV ec	AAAGAATTCCTTACATGCACAGCCCATCTG

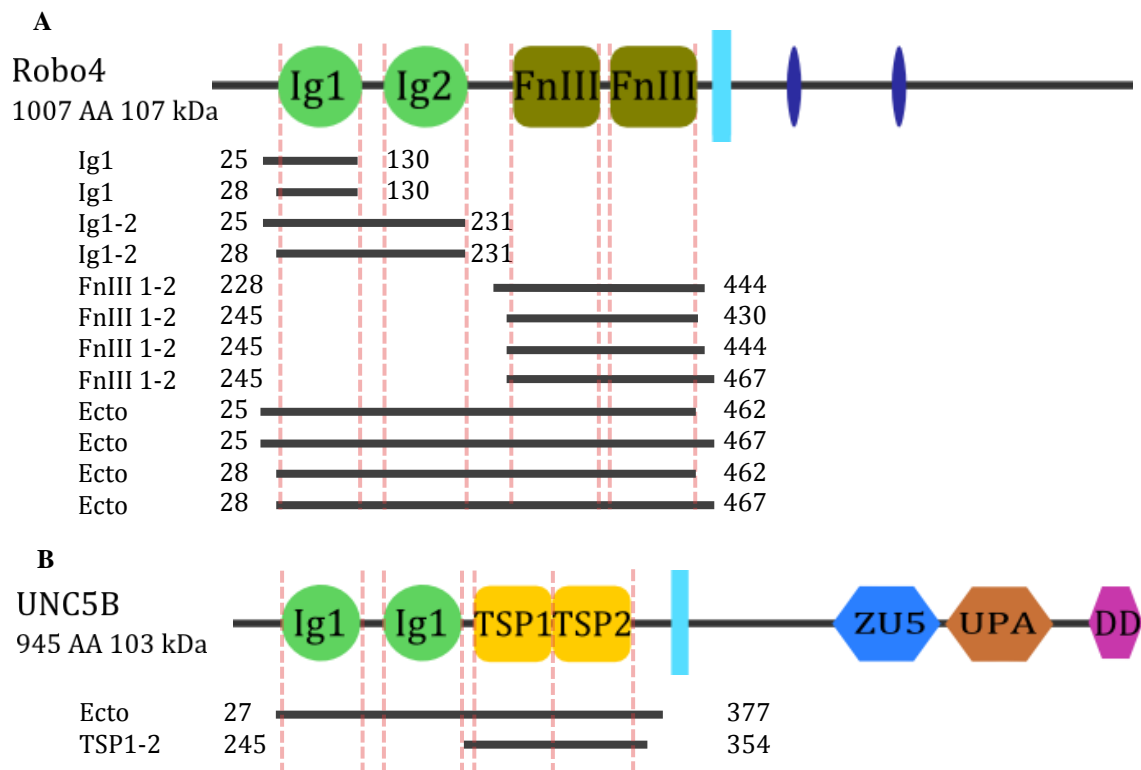


Figure 11: Robo4 and UNC5B constructs

Representation of protein constructs cloned. The red lines represent each domain boundaries.

A: Robo4.

B: UNC5B.

PCR was performed using either Phusion or Q5 polymerase in a 25 μ l reaction mix, composed of: 1x reaction buffer mix (Phusion or Q5), 200 μ M dNTPs, 0.5 μ M of forward and reverse primer and 1 ng of template DNA, occasionally supplemented with GC enhancing buffer (Phusion or Q5) or 5% (v/v) dimethyl sulphoxide (DMSO). Amplification was performed on a thermal cycler, programmed for 25 to 30 standard cycles, dissociation at 95°C for 30 seconds (s), annealing for 10 s at 50 to 58°C, and using an extension time of 20 to 30 s per kilobase (s/kb) at 72°C. Annealing temperature and extension times were adjusted depending on the polymerase and primer pair used.

The Robo4 native secretion signal sequence (amino acids 1-27) was replaced by the signal sequence of the human Pregnancy-Specific Glycoprotein 1 (PSG1), which is provided by the vector, to increase protein expression.

For the insect cell expression of the UNC5B 27-377 construct (from now on UNC5B ecto) the original signal peptide sequence (amino acids 1-26) was substituted with the honey bee Melittin signal sequence, which is more suitable for

MATERIALS AND METHODS

use in an insect cell expression system (Tessier et al., 1991), by primer addition. An additional restriction site was added to the N-terminus of the signal sequence to facilitate subsequent subcloning. At the C-terminus, a Strep tag (WSHPQFEK) was added before the restriction site.

The TSP1 and TSP2 domains of UNC5B were cloned into pETM-30 and pETM-40 plasmids for *E. coli* expression.

DNA fragments were analysed by agarose gel electrophoresis on a 1.6-2% (w/v) agarose gel (depending on fragment length) made with 1x TBE (178 mM Tris pH8, 178 mM boric acid, 4 mM EDTA) containing 0.05 µl/ml of ethidium bromide. A final concentration of 1x loading dye was added to the sample before loading on agarose gel and separated at 90 V for 50 min. PCR fragments were purified using the NucleoSpin gel and PCR clean-up kit from Macherey-Nagel following the manufacturer's protocol. All DNA fragments were digested with the appropriate pair of enzymes in a 50 µl reaction of 1x CutSmart buffer. The reactions were incubated for 15 min at 37°C. To stop the digestion, the fragments were either purified a second time using the PCR purification kit, or subjected to heat inactivation at 80°C for 20 min.

2.3.1 Mutagenesis of Robo4 ecto

Mutagenesis was performed following a modified QuikChange protocol originally developed by Stratagene (Liu and Naismith, 2008; Xia et al., 2015; Zheng et al., 2004) in which amplification of the full plasmid containing the mutation occurs. Primers were designed to have a minimum of 12 overlapping bases with the mutation site located in the middle of this region, having at least 4 and 6 bases from the 5' and 3' end respectively.

Cloning was performed with 50 ng of template DNA, as described in § 2.3. Amplification was performed on a thermocycler programmed to perform 25 cycles, dissociation at 95°C for 30 s, annealing for 30 s at 55°C, and amplification for 30 s/kb at 72°C. Amplification of the backbone was confirmed by running a 0.8% (w/v) agarose gel as described previously (§ 2.3). In order to eliminate the original DNA template, DpnI was added to the reaction mix in

1x CutSmart buffer. After 1 hour incubation at 37°C, and successive inactivation for 20 min at 80°C, the mutated plasmid was directly transformed into *E. coli* (§ 2.3.3).

2.3.2 Vectors and ligation

All Robo4 constructs were cloned into the pXLGsec vector (courtesy of Rob Meijers, EMBL Hamburg), which is a modified version of the pXLG-eGFP vector, for expression in mammalian cells. This vector contains a Kozac consensus sequence followed by the PSG1 signal sequence. The gene of interest was inserted between a KpnI site, which allows *in frame* insertion after the PSG1 signal sequence, and a SacI site that is followed by a C-terminal hexa-histidine tag (6x-His) for affinity purification, and a stop codon.

The UNC5B ecto construct was cloned into pFastBac1 vector for insect cell expression, using primers that would add the honey bee Melittin signal sequence at the N-terminus, and a strep tag at the C-terminus as previously explained.

Additionally, the UNC5B construct containing the TSP1 and TSP2 domains (amino acids 245-354) was also cloned into the pETM-30 and pETM-40 vectors for bacterial expression. These vectors are part of the pETM family and encode an N-terminal His glutathione S-transferase (GST) tag and an N-terminal maltose binding protein (MBP) tag, respectively, followed by a tobacco etch virus (TEV) cleavage site for affinity purification and subsequent cleavage.

In all cases, 2 µg of vector was digested with the appropriate enzymes in 50 µl reactions in 1x CutSmart buffer, by adding 5 U of each enzyme. The vectors were then loaded on 0.7% (w/v) agarose gel in 1x TBE buffer and separated at 90 V for 30 min. The band containing the cut vector was excised from the gel and purified using the gel and PCR purification kit according to the manufacturer's protocol.

Classical ligation was performed using 150 ng of total vector and a 1:3 molar ratio of the insert in a 20 µl reaction, supplemented with 1x T4 DNA ligase buffer and 400 U of T4 DNA ligase. The reaction was either incubated at room temperature for 10 min or at 16°C overnight. 3 µl of the ligation reaction were used to transform chemically competent DH5α bacterial cells (§ 2.3.3).

2.3.3 Transformation of bacterial strains and DNA isolation

Plasmid DNA or ligation mixes were transformed in chemically competent *E. coli* using the heat shock method. A 50 µl sample of chemically competent *E. coli* cells was incubated on ice with 50 ng of plasmid DNA or 3 µl of ligation mix for 15 min. Heat shock was performed at 42°C for 45 s followed by incubation on ice for 1 min. 200 µl of LB media were added, the cells incubated at 37°C for 30 min, and evenly spread on LB agar plates supplemented with the appropriate antibiotics for positive selection.

Isolation of plasmid DNA was performed from 5 ml LB cultures of a transformed DH5α *E. coli* strain using a NucleoSpin Plasmid miniprep kit, and successful cloning confirmed by sequencing (Genewiz).

To preserve the transformed bacteria, glycerol stocks were prepared using 1 ml aliquots of an overnight culture supplemented with 40% (v/v) glycerol and stored at -80°C.

Due to the high quantity of plasmid DNA required for mammalian transfection, the DNA was prepared using the QIAGEN Plasmid Giga kit or the Macherey-Nagel NucleoBond Xtra plasmid Maxiprep kit following manufacturer's protocols.

2.3.4 Bacmid preparation for insect cell transfection

The pFastbac1 plasmid containing the construct of interest was transformed into DH10EMBaY *E. coli* strain to produce a bacmid for virus production in insect cells, as described for the MultiBac system (Bieniossek et al., 2008; Fitzgerald et al., 2006).

Briefly, this strain contains the EMBaY bacmid and a transposase expressed by a helper plasmid. The UNC5B fragment is transposed into the bacmid following Tn7 signal sequences, and is inserted, along with the gentamycin resistance gene, in the middle of a *lacZ* gene. This allows for the selection of positive colonies on LB plates containing 50 mg/ml Kanamycin, 10 mg/ml Tetracycline, 10 mg/ml Gentamycin, 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.2 mg/ml BluOGal. The bacmid isolated from positive colonies is used for transfection following established procedures as described in § 2.5.1.

2.4 Mammalian cell culture

Protein production was performed in either HEK293T/17 or HEK293S GnTI⁻ cell lines from the American Type Culture Collection (ATCC). The HEK293T/17 (from here on referred to as HEK293) is a derivative of the classic 293 cell line, which is characterized by higher transfection susceptibility. The HEK293S GnTI⁻ (from here on referred to as HEK293S), also originated from the 293 cell line, was selected for its lack of N-acetyl-glucosaminyltransferase I (GnTI) activity (Reeves et al., 2002). This enzyme, which is an important part of the glycosylation pathway, catalyses the reaction necessary to convert high mannose glycans into complex type glycans (see Figure 9). As a consequence, the proteins produced using this cell line have less heterogeneous and more accessible glycan chains for enzymatic cleavage (Chang et al., 2007).

2.4.1 Routine maintenance and storage

Both cell lines were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% (v/v) Fetal Bovine Serum (FBS), and 100 U/ml Penicillin-Streptomycin antibiotics, in an incubator maintaining a 5% (v/v) carbon dioxide (CO₂) high humidity atmosphere at 37°C. The presence of a rich CO₂ atmosphere and sodium bicarbonate in the buffer, keeps the pH of the media in the range necessary for cell viability during growth. A high humidity environment is necessary to avoid media evaporation.

Cells were routinely grown in 75 cm² clear plastic flasks (T75) with 10 ml of media, and passaged when reaching ~90% confluency. For splitting, the media was removed and the cells were washed with 5 ml of Dulbecco Phosphate Buffer Saline (DPBS) to remove residual traces of media and FBS. To detach the cells from the flask surface, 2 ml of 0.05% (v/v) Trypsin-EDTA were added and the cells incubated at 37°C for up to 3 min. At the end of the incubation, 8 ml of fresh DMEM were added to the flask to block trypsinization and all remaining clumps of cells dissociated carefully via pipetting. Finally, a 1 to 10 dilution of the original culture was prepared in a new flask containing fresh DMEM.

MATERIALS AND METHODS

To determine proliferation rate and viability of the cells, an improved Neubauer counting chamber was used in combination with 0.4% (v/v) Trypan blue. This dye only stains dead cells by traversing the damaged cell membrane. This allows determination of the percentage of live cells by exclusion. In the case where the viability was less than 90%, or the proliferation was evidently impaired, the cultures were discarded.

In order to maintain a viable stock, low passage cells with a viability >75% were prepared for freezing. The suspension of cells was centrifuged at 200×g for 3 min to avoid damaging the cells, the media was removed, and the cells resuspended in DMEM supplemented with 10% (v/v) FBS and 5% (v/v) DMSO to a final density of 3×10^6 cells/ml. The cells were then divided in 1 ml aliquots, and the temperature gradually lowered by resting 1 hour at -20°C, and overnight at -80°C before final storage in liquid nitrogen.

To prepare a new batch of cells when needed, an aliquot was taken from the nitrogen storage and quickly defrosted in a thermobath at 37°C. The aliquot was then resuspended in 10 ml of warm culture media, centrifuged at 200×g for 3 min, the media removed and then resuspended again in another 10 ml of media. This step was repeated two times in order to eliminate the DMSO of the freezing media. The final suspension was performed in 5 ml of media, and the cells transferred to a T75 flask for incubation at 37°C. Since viability is usually very low immediately after thawing, the media was changed after 24 hours without any wash to eliminate floating dead cells and avoid detaching new cells. Once ~80% confluency was reached and the media was removed, the cells were washed with 2 ml of PBS and detached using 0.5 ml of Trypsin-EDTA. All the cells were then transferred to a T175 flask with 10 ml of culture media for routine passaging. After 1 to 2 weeks, depending on batch, the cells were fully recovered and ready to be used for expression. For all experiments, cells between passage 5 and 25 were used.

For small scale expression tests, 6-well plates were used. 0.3×10^6 cells are added to each well containing 2.8 ml of culture media. After 24 to 48 hours the plate was ready for transfection.

2.4.2 Mammalian cell transfection optimization protocol

Transfection efficiency and protein production is batch-dependent, and easily influenced by DNA and transfection reagent (amount and ratio). Polyethyleneimine (PEI) has been shown to be a cheap and efficient transfection reagent, which makes it useful for large scale expression where the amount of cells required for protein production becomes considerable (Aricescu et al., 2006; Dalton and Barton, 2014). It acts by enclosing the negatively charged DNA into positively charged particles that can associate with the cell membrane for internalization via endocytosis. The endosomes containing the DNA/PEI complex then start to swell and undergo osmolysis due to pH imbalance and increased Cl⁻ intake, delivering the extraneous DNA to the cell compartment (Boussif et al., 1995; Sonawane et al., 2003). Although linear 22 kDa PEI was shown to have a slightly better efficiency when compared to branched 25 kDa PEI factor (Wiseman et al., 2003), the difference was not reported to be a determining factor and can be minimized by screening for appropriate transfection conditions. Since the latter was readily available, it was chosen as the standard DNA delivery reagent.

In order to test the best conditions and as a positive control for all transfections, two green fluorescent protein (GFP) containing plasmids were used, the pmMGFP⁺ and pmMGFP⁻. They respectively encode for a 6xHistidine-tagged and a nontagged GFP. The protocol was devised based on established guidelines for PEI transfection (Aydin et al., 2012; Longo et al., 2014). A transfection medium is prepared, composed of DMEM supplemented with only 2% (v/v) FBS, 100 U/ml of Penicillin-Streptomycin, 2 mM L-glutamine, 1 mM Sodium Pyruvate and 0.1 mM MEM Non-Essential Amino Acids (NEAA). Since it was demonstrated that formation of DNA-PEI complexes is driven by aggregation and negatively influenced by the presence of FBS (Escriou et al., 1998), formation of DNA/PEI complexes is first made in Opti-MEM, a minimal FBS-free media, before addition to the cells.

In order to screen for the necessary amount of DNA and PEI for optimal transfection, $\sim 0.3 \times 10^6$ cells were distributed to each well in a 6-well plate. After ≥ 24 hours, the cells that reached 80% confluency were ready for transfection. A range from 0.5 to 3 μ g of DNA per well, and a ratio of 2 to 4 times PEI, were tested. Only the most typical protocol that gave the best results is reported here. The

MATERIALS AND METHODS

culturing media was removed from the wells and the cells carefully washed with 1 ml of DPBS buffer while keeping the plate tilted to avoid stress and reduce detachment of the cell layer. Upon removal of the DPBS, 2 ml of fresh transfection media were added, and the plates put back in the incubator to recover while the transfection mix is prepared.

To prepare the transfection mix, one fifth of the total amount of media, in this case 0.5 ml, of Opti-MEM media were added to as many 1.5 ml tubes as the number of transfections desired. 6 μ l of PEI at 1 mg/ml pH 7 were added to each tube. After 1 min incubation and shaking, 3 μ g of DNA were added to the tube and the transfection mix incubated for 15 min at room temperature. At the end of the incubation, the transfection mix was added drop by drop on top of the wells and the cells incubated at 37°C for an additional 72 to 120 hours for expression before collection of the media and cells.

2.4.3 Mammalian cell expression test

Expression tests were performed in a 6-well plate format following the protocol detailed in § 2.4.2. On each plate, one well of untransfected cells and one of pmmGFP+ transfected cells (expressing 6xHis-tagged GFP) or pmmGFP-transfected cells were kept as negative and positive transfection controls respectively. 72 hours after transfection, the efficiency was determined visually using a microscope equipped with a mercury fluorescence lamp and suitable filters for GFP excitation and visualization of GFP expression. Transfection efficiencies of over 75% could be regularly achieved with little toxicity using this protocol. To verify expression, 1 ml of media was collected and centrifuged at $>12000\times g$ to remove dead cells. An aliquot of this was then taken for Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and western blot analysis.

The rest of the media was discarded and the cells were washed with DPBS to eliminate the excess of FBS components, resuspended in 1 ml of PBS, and sonicated to obtain a whole lysate fraction. Finally, the whole lysate was centrifuged for 10 min at $>12000\times g$ to pellet the insoluble fraction and an aliquot of supernatant taken for analysis (§ 2.11).

2.5 Insect cell culture

Virus and protein production were separated to dedicated cell lines. The *Spodoptera frugiperda* (Sf21) cell line was used for virus production and test expression, while the *Trichoplusia ni* (Hi5) cell line was used for large scale protein production. Sf21 cells were maintained in suspension in Sf900 II serum free media (SFM) media at a density between 1×10^6 cells/ml and 4×10^6 cells/ml. Hi5 cells were maintained in suspension in ExpressFIVE SFM supplemented with 4 mM L-glutamine to improve protein expression at a density between 1×10^6 cells/ml to 2.5×10^6 cells/ml.

Cells were regularly checked to ensure they were viable at all times and no contamination was present. No antibiotic was added. An improved Neubauer counting chamber was used for cell counting. Where viability or proliferation was evidently impaired, the cultures were discarded.

2.5.1 Insect cell transfection

Transfection of the bacmid into Sf21 cells was carried out using the X-tremeGENE DNA transfection reagent and approximately 3 μ g of DNA. Between 0.5 to 1×10^6 cells/ml of Sf21 cells were transfected in a 6-well plate containing 3 ml of Sf900 II SFM culture media per well to produce a V_0 virus. The first generation virus was then used to produce a V_1 higher titer virus by infecting 25 ml of Sf21 at 0.5×10^6 cells/ml in a shaker flask. After determination of the infection efficiency, the V_1 virus was used for large scale expression by infection of 1 to 6 litres of Hi5 insect cells at 0.5×10^6 cells/ml in ExpressFIVE SFM medium. Media and cells were collected 120 hours post proliferation arrest, centrifuged at $1000 \times g$ for 20 min to remove cells, and the media recovered for protein purification.

2.6 *E. coli* expression

Expression of the UNC5B TSP1-2 construct and the fragment antigen-binding (Fab) constructs were performed in a BL21 RIL *E. coli* strain.

For each plasmid, an overnight starter culture was grown from a single clone in LB media containing 50 µg/ml kanamycin and 34 µg/ml chloramphenicol. The next day, the starter culture was transferred to a large (litre scale) culture of the same media composition in a 1:100 ratio and grown to an OD of 0.6. Expression was induced by the addition of 0.05 mM IPTG for the UNC5B TSP1-2 His-GST tagged construct (pETM-30) and 0.5 mM IPTG for the MBP-tagged construct (pETM-40), and left overnight at 20°C before collecting the bacteria by centrifugation at 6000×g at 4°C for 15 min.

For each Fab construct, expression was induced by addition of 1 mM IPTG, and incubated at 37°C for 4 hours, before collecting the bacteria by centrifugation at 6000×g at 4°C for 15 min.

2.7 Purification of Robo4 ecto

The full length Robo4 ectodomain (amino acids 28-462) was directly purified from the media of growing HEK293 or HEK293S cells. After filtration using a 0.22 μm Stericup filter unit, the media was applied to a Ni Sepharose Excel, a specialized type of Nickel resin specifically manufactured to be used with raw media. This type of resin is composed of an agar matrix of cross-linked Nickel ions that can withstand mild stripping conditions, which were an issue using standard Ni resins due to undisclosed reagents in the media. The resin was washed three times with 5 column volumes (CV) of 20 mM Tris pH 7.5, 500 mM NaCl, and 25 mM Imidazole. Proteins were eluted by applying five CV of 20 mM Tris pH 7.5, 500 mM NaCl, and 500 mM Imidazole. Fractions containing the Robo4 ectodomain were analysed by SDS-PAGE analysis, pooled, concentrated, and loaded on a gel filtration Superdex 200 16/600 column equilibrated in 20 mM Tris pH 7.5 and 150 mM NaCl buffer. Fractions containing the protein were again analysed by SDS-PAGE, pooled, concentrated and flash-frozen in liquid nitrogen for storage at -80°C .

Since the expression varied depending on the cell's batch, and on the efficiency of transfection, an exact estimate of protein expression is not fully reliable, but on average it allowed the purification of ~ 0.1 mg of protein per transfected roller bottle.

An ion exchange purification step was later performed in order to separate the different glycosylated species normally present in the final product. For this the protein samples were exchanged via diafiltration in a buffer with low NaCl content (20 to 30 mM) before loading on a Source 15Q 4.6/100 column. Protein was eluted using a high, 1000 mM NaCl buffer by application of a linear gradient over a large number of CVs (5 to 40), or by a step gradient elution.

2.8 Purification of UNC5B constructs

2.8.1 Purification of UNC5B ecto

Purification of UNC5B ecto (amino acids 27-377) was performed directly from the insect cell media. BioLock biotin blocking solution was added to the media to mask free biotin normally present in insect cell media. The media was filtered using a 500 ml Stericup 0.22 μm filter unit and applied to 5 ml of Strep-Tactin resin by gravity flow. The resin was washed 2 times with 5 CV of 1x buffer W (buffers were part of the Strep purification kit from IBA), and the protein eluted by applying 3 times 1 CV of buffer E, or until no protein could be detected in the eluate.

The elution was concentrated and loaded onto a gel filtration HiLoad Superdex 200 16/600 column in 20mM Tris pH 8, 150 mM NaCl. Fractions containing the proteins were pooled, concentrated and flash-frozen in liquid nitrogen for storage at -80°C .

2.8.2 Purification of UNC5B TSP1-2

The bacterial pellets were resuspended in 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM MgCl_2 supplemented with cOmplete protease inhibitor, and sonicated on ice for 20 min (15 s pulse and 20 s rest). The crude lysates were centrifuged at $30,000\times g$ (4°C for 30 min). The soluble fraction of the His-GST tagged construct was then loaded onto a glutathione sepharose resin by gravity flow and washed with 5 CV of resuspension buffer. Elution was performed using 3 CV of resuspension buffer supplemented with 10 mM reduced glutathione and the fractions analysed by SDS-PAGE. Fractions containing the protein of interest were collected, concentrated and applied to a size exclusion chromatography column in 20 mM Tris pH 7, 300 mM NaCl, 1 mM MgCl_2 buffer.

The soluble fraction of the MBP tagged construct was loaded onto a MBP HiTrap column and washed with resuspension buffer until stabilisation of the UV 280 nm baseline. Elution was performed with 3 CV of resuspension buffer supplemented with 10 mM maltose and analysed by SDS-PAGE. Fractions containing the protein were collected, concentrated and applied to a size exclusion chromatography column in the same buffer as for GST-tagged UNC5B TSP1-2. Fractions containing

the proteins were pooled, concentrated, and flash-frozen in liquid nitrogen for storage at -80°C.

2.9 Purification of Fabs

The Fab constructs were provided by our collaborator from the laboratory of Sachdev Sidhu (University of Toronto). The Sidhu laboratory develops antibody phage display technologies (Adams and Sidhu, 2014; Sidhu et al., 2000), from which suitable Robo4 binding antibodies were selected for use in this study (Table 4).

Table 4 List of Fabs CDR

Amino acid composition of the CDR of light chains (LC, 1 to 3) and heavy chains (HC, 1 to 3) for each Fab selected for this study.

	LC 1	LC 2	LC 3	HC 1	HC 2	HC 3
Fab5555	SVSSA	SASSLYS	GYHLI	ISYYSM	SISPSSSYTY	SWGYYPPAM
Fab5562	SVSSA	SASSLYS	SWSSSSYPF	LYSYYM	SISSYSSYTY	AGYYVWYAI
Fab5564	SVSSA	SASSLYS	SYWWPI	IYSSSM	SIYPYSSYTY	YPYAASYYSYGVHYAL
Fab5570	SVSSA	SASSLYS	YSYYGSLI	LSSYYM	SISPYYSYTY	GSYPSGL
Fab5582	SVSSA	SASSLYS	YAYGYSLI	ISSYSM	SIYPSYSYTY	TVRGSKKPYFSGWAM
Fab5585	SVSSA	SASSLYS	AFSLI	ISYYYI	SISPSYGYTY	YWGYPWGYGM

Bicistronic expression vectors (Kirsch et al., 2005) based on the RH2.2 backbone were used for expression of the Fabs. Light and heavy chains, which compose each Fab, are preceded by a leader sequence that directs translocation of the polypeptides to the periplasm, where the formation of the disulphide bridge that connects the two chains will happen. This process is necessary for the folding of a functional Fab.

Bacteria containing the expressed protein were resuspended in 1x PBS lysis buffer in the presence of cOmplete protease inhibitor and sonicated on ice for 20 min (15 s pulse and 20 s rest). The total extract was centrifuged at 30,000×g (4°C for 30 min), and loaded on a HiTrap MabSelect SuRe column (containing protein A resin) using an Akta system. The column was washed with lysis buffer until stabilisation of the UV 280 nm baseline. Elution was achieved by application of a

MATERIALS AND METHODS

low pH buffer (100 mM Sodium citrate pH 3). 1.7 ml fractions were collected into tubes containing 0.3 ml of 1 M Tris pH 11 to neutralize the acidic elution. Fractions containing the Fabs were collected and loaded on a HiPrep 26/10 desalting column to facilitate the change to a Tris-salt buffer (20 mM Tris pH 7.5, 500 mM NaCl). Fractions containing the purified FABs were collected, concentrated and flash frozen in liquid nitrogen before storage at -80°C.

2.10 Endo F1 expression and purification

The plasmid containing the Endo F1 deglycosidase was a gift of Elena Seiradake (University of Oxford). Endo F1 was expressed in *E. coli* BL21 RIL strain and purified according to an established protocol (Grueninger-leitch et al., 1996). Briefly, after GST purification, the glycosidase was dialyzed in 10 mM acetic acid-sodium acetate buffer pH 5.5, and 50% glycerol (v/v). The final product was concentrated to 1 mg/ml, divided in aliquots and conserved at -80°C after flash freezing in liquid nitrogen.

2.11 Western blotting

Expression of target proteins was verified by western blot with a primary mouse anti-His antibody and a secondary anti-mouse antibody, conjugated with horseradish peroxidase (HRP) or AlexaFluor 532.

Proteins were transferred on an Immobilon-P polyvinylidene difluoride (PVDF) membrane with the wet transfer method performed at 100 V for 1 hour at 4°C. The membrane was blocked with a 5% (w/v) milk solution in PBST buffer (1x PBS buffer, 0.1% (v/v) Tween 20) for 1 hour at room temperature. The membrane was then washed three times for 5 min with PBST buffer before incubation with a 1:1000 dilution of primary antibody in PBST for 1 hour. Unbound antibody was eliminated by three washes for 5 min in PBST. The secondary antibody was then incubated 1 hour at 1:2000 or 1:10000 dilution for HRP and AlexaFluor 532 conjugated antibody respectively. Detection was performed either using the SIGMAFAST diaminobenzidine tablets for colorimetric detection, or by exposing

the membrane on a TYPHOON scanner with filters set at 532 nm wavelength for excitation and 554 nm wavelength for fluorescence detection.

2.12 Deglycosylation trials

Deglycosylation of Robo4 was performed using three different enzymes, the commercial Endo H and PNGase F1, or the *in house* produced Endo F1 (§ 2.10).

To test the effectiveness of deglycosylation, proteins were incubated at 37°C with the respective enzyme and the reaction followed by collecting samples after 1 hour, 3 hours and overnight incubations. The results were visually analysed by SDS-PAGE. The selected enzyme was added to the protein of interest for deglycosylation in a 1 to 100 (w/w) ratio for Endo F1 or 10 U/μg of protein for PNGase F1 and Endo H diluted in the protein buffer. Deglycosylated samples for crystallisation were purified by gel filtration chromatography prior to screening.

2.13 SEC-MALS analysis

Size Exclusion Chromatography Multi Angle Light Scattering (SEC-MALS) analysis was performed at the biophysical platform of the Partnership for Structural Biology (PSB) in Grenoble. This technique allows the accurate measurement of light scattering derived from particles in solution, and by careful analysis can give an absolute measurement of their molecular mass (Tarazona and Saiz, 2003).

The system and columns were equilibrated overnight with the appropriate buffer prior to each experiment in order to achieve an optimum baseline. 50 μl of protein sample at a concentration between 0.7 to 4 mg/ml (single proteins or preassembled complexes in a 1:1 molar ratio) were injected onto a Superdex 200 Increase 10/300 or a Superdex 200 5/150 column. The MALS spectrum was recorded by a laser at 690 nm wavelength using a DAWN-HELEOS detector. The refractive index (RI) was measured by an Optilab T-rEX detector. The mass was determined by analysing the differential refractive index through the elution peak of each sample using the Debye model for proteins (Edelman, 1992), integrated in the ASTRA software (version 6.5.0.3).

2.14 SPR binding experiments

Surface Plasmon Resonance (SPR) experiments for binding and determination of kinetic interaction constants were performed at the biophysical platform of the PSB in Grenoble with a Biacore T200 instrument. This technique measures of protein-protein interactions by following the change of the reflected light angle. One interaction partner is immobilized on the sensor surface, covalently or by affinity methods, while the other flows freely in solution at different concentrations. Upon binding, a change in the light reflection is measured to determine association, and dissociation, of the partner protein, and thereby infer the association (k_a), and dissociation (k_d), constants (Karlsson and Larsson, 2004).

2.14.1 Immobilization of Robo4 and UNC5B

Robo4 ecto and UNC5B ecto were immobilized with the amine coupling method on a series S CM5 sensor chip with a hydrophilic carboxymethylated dextran surface. In order to remove amine-reactive Tris, all samples for immobilization (Robo4 ecto and UNC5B ecto) were dialyzed overnight at 4°C in 10 mM HEPES pH 7.5, 150 mM NaCl and 0.05% (v/v) Tween 20 buffer, while agitating using Pur-A-Lyzer mini dialysis tubes with a 6 kDa cut-off. The same buffer was used for measurements. To obtain the best response for each experiment, the target immobilization response was decided following three assumptions, and optimized based on the result obtained:

- For binding specificity experiments, any protein ligand density that gives a proper signal from low to high is sufficient.
- For kinetic experiments, a total analyte response (R_{max}) of ~100 response units (RU) upon surface binding is desired (low ligand density). The necessary ligand density can be extrapolated in RU (R_{ligand}) from the formula:

$$R_{max} = \frac{R_{ligand} * MW_{analyte} * Valency_{ligand}}{MW_{ligand}}$$

Equation 1: Total maximum response (R_{max}) equation

- Availability of total immobilized protein for binding is 20 to 30 %, as the amino coupling method immobilizes the proteins in random orientations that may partially mask the binding sites

In all experiments, immobilization was performed with the amine coupling method at a flow rate of 10 $\mu\text{l}/\text{min}$. Chip surfaces were activated by injecting a 1 to 1 mixture of 0.1 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and 0.1 M N-hydroxysuccinimide (NHS) for 7 min. Robo4 ecto was immobilized at 3 $\mu\text{g}/\text{ml}$ in 10 mM sodium acetate buffer pH 5 with a target RU of 300, while UNC5B ecto was immobilized at 5 $\mu\text{g}/\text{ml}$ with a target RU of 400 in 10 mM sodium acetate buffer pH 4.5, by means of successive injections as calculated by the program to reach the target RU. After each injection, the chip surface was washed with a 30 s injection of 50 nM sodium hydroxide solution. The two ligands were immobilized either on flow cells 2 or 4. Flow cells 1 and 3 were subjected to blank immobilization to serve as reference cells. All surfaces were then blocked with a 7 min injection of 1 M ethanolamine pH 8.5.

2.14.2 Kinetic constant determination of Fab binding

All analytes were dialyzed as previously described (§ 2.14.1), or diluted in running buffer by at least 100-fold. To account for unspecific surface interaction, and background noise, analytes were injected on both sample and reference cell and the resulting sensorgrams subtracted to eliminate these sources of error. Only the subtracted sensorgrams are displayed, unless otherwise stated.

To collect kinetic binding data, runs for each Fab were performed at a flow rate of 30 $\mu\text{l}/\text{min}$. Each Fab was injected in a two-fold dilution series as described in Table 5, allowed to associate for 300 s (150 μl of sample per injection) and to dissociate for 600 s. Regeneration was performed after dissociation by 30 s pulses of 10 mM glycine pH 2.5. Complete regeneration was achieved for most Fabs in two pulses, except for Fab5582, which required one, and Fab5555, which required four pulses. Baseline recovery was achieved in 120 s before the next injection.

Each experiment was performed in triplicate, using two different chips, three different immobilizations and two batches of proteins. Sensorgrams for association

MATERIALS AND METHODS

and dissociation phases were recorded and the data analysed using the BiaEvaluation T200 software package.

Table 5: Fab dilution range for SPR experiment

Serial two-fold dilutions were used as described for kinetic constants determination.

	Concentration range (nM)							
Fab5555	20	10	5	2.5	1.25	0.625	0.312	0.156
Fab5562	40	20	10	5	2.5	1.25	0.625	0.312
Fab5564	20	10	5	2.5	1.25	0.625	0.312	0.156
Fab5570	-	10	5	2.5	1.25	0.625	0.312	0.156
Fab5582	100	50	25	12.5	6.25	3.125	1.562	0.781
Fab5585	100	50	25	12.5	6.25	3.125	1.562	0.781

A 1:1 Langmuir binding model, or equilibrium at steady state analysis, was used to fit the data and calculate k_a and k_d . The equilibrium dissociation constant (K_D) was determined as the ratio k_d/k_a .

2.14.3 SPR binding tests of UNC5B

As was previously reported by other authors (Koch et al., 2011), we tested the Robo4/UNC5B direct interaction by replicating the reported conditions with our samples. To this end Robo4 ecto and UNC5B ecto were immobilized, as previously stated. The same Robo4 ecto and UNC5B ecto samples were then injected as analytes at the concentrations of 50 nM and 1000 nM with a flow rate of 30 μ l/min, a contact time of 300 s (150 μ l of sample for each injection), and allowed to dissociate for 600 s. A single regeneration pulse of 30 s with 10 mM glycine pH 3 was performed after dissociation, and the system was allowed to equilibrate for 120 s before the next injection. A single injection of Fab5570 at 50 nM concentration following the same protocol, but with 2 regeneration rounds at pH 2.5, was used as a positive control of Robo4 ecto immobilization. Sensorgrams were recorded and data analysed as before (§ 2.14.2).

2.15 SAXS analysis of Robo4 ecto and Fab complexes

Small angle X-ray scattering (SAXS) data were collected using the Bio-SAXS beamline BM29 (Pernot et al., 2013) at the European Synchrotron Radiation Facility (ESRF) in Grenoble. Unlike crystallographic techniques, SAXS can provide information on the size and overall shape of biomolecules in solution by collecting the average information derived by all possible orientations in a mobile phase. Since only the intensities of the scattered X-rays can be measured, and not the phases, it's not possible to directly reconstruct the shape of the molecule by inverse Fourier transform, but it is possible to gather information of the maximum dimension, the mass, and indirectly reconstruct a model of all possible conformations of the scattering molecule (Feigin and Svergun, 1989; Glatter and Kratky, 1982).

To minimise the influence of sample aggregation, and to improve data quality, an inline gel filtration liquid chromatography system was used to perform data collection (Brennich et al., 2017). The collection was performed at room temperature, but all samples were subjected to gel filtration prior to data collection at 4°C for buffer exchange. The Robo4 ectodomain alone, and the complexes between it and Fab5555, Fab5562, Fab5564, Fab5570, Fab5582, were run on a Superdex 200 5/150 column in phosphate buffer pH 7.5, 500 mM NaCl, 3% (v/v) glycerol. The complexes with Robo4 ecto/Fab5555 and Robo4 ecto/Fab5585, were run on a Bio SEC-7.8x300 column in 20 mM Tris pH 7.5, 300 mM NaCl, 1% (v/v) sucrose. Sample concentration was between 6 mg/ml and 11 mg/ml.

Data collection was carried out with a Pilatus 1M detector (Dectris) (Broennimann et al., 2006) at a distance of 2.86 m from a 1.8 mm sample glass capillary. The wavelength of the X-rays was 0.991 Å, with an exposure time of 1 second/frame. The momentum transfer range covered was 0.008 to 0.47 Å⁻¹. The Enhanced automated collectioN of data (EDNA) framework for bioSAXS (Brennich et al., 2016) available at BM29 performs a preliminary analysis of the frames and, through the ISPyB interface (De Maria Antolinos et al., 2015), allows to visualize data quality of each frame across the elution peak of the samples. A subset of the

MATERIALS AND METHODS

frames that showed a constant radius of gyration (R_g) were collected and merged with Primus, part of the ATSAS package (Petoukhov et al., 2012).

In order to eliminate the influence of protein aggregation, and capillary spoiling in the data, two datasets of the complex Robo4 ecto/Fab5555 were merged in Primus, using the low q data range from one dataset collected at low protein concentration, and the high q data range from another dataset collected at a higher protein concentration.

The R_g values derived from Guinier analysis were calculated using Primus and verified with ScÅtter (Rambo, 2017). The dimensionless Kratky plot (Durand et al., 2010) derived from the scattering data was calculated, normalized, and scaled with ScÅtter. For all complexes, pair distance distribution functions ($p(r)$), and maximum dimension (D_{max}), were computed from the scattering curve with GNOM (Svergun, 1992), along with the R_g value that can be derived from the whole curve. Porod volumes and correlated volumes were ultimately calculated with ScÅtter. Multiple independent models of Robo4 ecto, Robo4 ecto/Fab5562 and Robo4 ecto/Fab5564 were created using the DAMMIF program (Franke and Svergun, 2009). Subsequently, the models of each complex were aligned, compared to exclude outliers, and averaged with the DAMAVER pipeline (Volkov and Svergun, 2003). The final model was refined with DAMMIN (Svergun, 1999) using the result from DAMAVER as a starting search model.

On the Robo4 ecto dataset the ensemble optimization method (EOM) was also used in an attempt to describe the experimental SAXS data by fitting the theoretical scattering intensity derived from an ensemble representation of atomic model conformations (Bernadó et al., 2007; Tria et al., 2015). The models used for this approach were created as described in § 2.18.

2.16 Crystallisation experiments

Protein crystals are grown by forming a super-saturated solution and inducing an ordered form of precipitation that results in regular 3D arrays of the protein. To supersaturate a protein solution there are two main possibilities: increasing the protein concentration, or adding a second reagent that reduces solubility. Reagents, such as salts or polyethylene glycols (PEGs), will decrease solubility while water evaporates from the crystallization drop. This vapour diffusion method is the main technique that allows formation of protein crystals.

Samples for crystallisation were sent to the high throughput crystallisation facility of EMBL Grenoble ([HTX lab](#)) to test a large number of conditions that could be favourable for crystallization. The standard protocol used by the platform has 200 nl sitting drops in a 1:1 protein:reservoir ratio stored at 20°C. Storage at 4°C was tested with a few selected conditions, but it was discarded since it showed no benefits. During optimization, and for the final crystals used in data collection, manual drops were prepared following the hanging drop method in 2 µl drops. Sample concentration was between 1 and 12 mg/ml per sample, depending on protein and optimization stage.

The Robo4 ectodomain was tested in three different forms: fully glycosylated, with only mannose-type glycosylation (HEK293S expressed), and deglycosylated. Samples of Robo4 ecto in complex with the Fabs were also tested for crystallisation. The complexes were formed by mixing in a 1 to 1 molar ratio, and further purified by gel filtration on a Superdex 200 10/300 column. Details are summarized in later sections (§ 3.7).

The UNC5B ectodomain was tested without further modification.

2.16.1 UNC5B ecto crystallisation

Because the few initial hits only produced microcrystals, the random Microseed Matrix Screening rMMS approach (Till et al., 2013), recently introduced at the HTX lab, was used to increase the screen dimensions, and achievable a better resolution. Seeds were prepared from the initial crystals obtained using the Hampton Seed Bead kit according to the manufacturer's protocol (Luft and DeTitta,

1999). The undiluted seed stock was used to perform the first random screenings on a PACT premier screen (Newman et al., 2005), and a Salt Grid screen (homemade by the HTX lab). Using this approach, several other conditions were identified, and further optimized by varying the composition of the reservoir, the protein concentration, and using 1:100 and 1:1000 dilutions of the seed stock (§ 3.12 for details). Crystallisation drops were prepared following the ratio of 1 (protein) to 0.3 (diluted seed) to 0.7 (precipitant) at a protein concentration of 1-8 mg/ml. All crystals had the same shape (parallelepiped of varying thickness), and were harvested for data collection as described in the following sections. Further details are summarized in later sections (§ 3.12)

2.16.2 Crystal harvesting

Crystals were harvested either automatically at the HTX lab (Cipriani et al., 2012; Márquez and Cipriani, 2014; Zander et al., 2016) or manually. During manual harvesting, crystals were either directly mounted on loops, or transferred to a cryo-solution drop containing the reservoir solution supplemented with either 20% (v/v) PEG 400 or ethylene glycol prior to flash freezing in liquid nitrogen.

MicroMesh loops were used to harvest multiple microcrystals of UNC5B ecto. Excess buffer was eliminated by touching the bottom of the mesh with filter paper to minimise ice formation in conditions where no cryo-protectant was present (Pellegrini et al., 2011).

2.17 Structure determination

X-ray diffraction from a crystal derives from the scattering of X-rays by the ordered arrangement of atoms in a crystal. The diffraction pattern originates from positive and destructive interference events. The intensity of the spots depends on the electrons that scatter the X-rays, while the position of the spot arises from the relative positions of the molecules within the crystal lattice. Mathematical laws dictate if diffraction in a certain vector direction will be observed. Given specific diffraction conditions, the information derived from such patterns can be used to determine the position of atoms within the crystal, and solve the 3D structure of the molecule.

2.17.1 General data collection strategy

All diffraction data were collected on ID29 (De Sanctis et al., 2012) and ID30B at the ESRF, Grenoble, using Pilatus 6M-F and Pilatus3 6M (Dectris) pixel array detectors, respectively. All macromolecular crystallography beamlines of ESRF are controlled using the latest version of the MXCuBE graphic user interface (Gabadinho et al., 2010). To mitigate radiation damage, data were collected while applying a constant stream of liquid nitrogen vapour at 100°K to the mounted crystals.

To assess the quality of the diffraction pattern before collection, the EDNA framework for X-ray crystallography (Incardona et al., 2009), as integrated in MXCuBE, was used. For each crystal, four images were collected at 90 degree rotations on the ω angle, and automatically submitted to the EDNA pipeline for characterization and determination of an optimal data collection strategy. During this process, automatic indexing of the diffraction patterns recorded is performed by LabelIt (Sauter et al., 2004) or Mosflm (Leslie, 2006; Powell et al., 2013). Afterwards, BEST (Popov and Bourenkov, 2003) is used to calculate the optimal data collection strategy. To include the contribution of the radiation damage caused by X-ray exposure, the program RADDOSE (Paithankar and Garman, 2010) is run and accounted for in the strategy calculation by BEST.

MATERIALS AND METHODS

For single crystal data collection, the strategy proposed was evaluated and corrected when necessary to achieve maximum completeness and redundancy using the graphical interface of iMosflm (Battye et al., 2011; Powell et al., 2017).

2.17.2 UNC5B ecto crystals data collection and analysis

All UNC5B Datasets were collected at 100°K at the ID30B beamline from harvested micro crystals using the MeshAndCollect method (Zander et al., 2015), or from single crystals using the appropriate strategy as described above. Since ID30B is a tuneable beamline with an energy range of 6 to 20 keV, and a variable beam sizes between 20 to 200 μm^2 , the beam size was adjusted depending on the size of the crystals.

For MeshAndCollect the best datasets were merged after integration using the program ccCluster (Santoni et al., 2017). This program uses a hierarchical cluster algorithm based on the correlation coefficients between datasets. This approach allowed an improved signal-to-noise ratio, completeness and data quality.

Programs from the CCP4 suite (Winn et al., 2011) were used for analysis. First, after image integration, the most probable Laue group was determined using Pointless (Evans, 2006, 2011). The measured intensities were then scaled and merged through Aimless (Evans and Murshudov, 2013), the structure factors were calculated with Truncate (French and Wilson, 1978) and a unique set of reflections with a corresponding 5% of the total reflections selected (R_{free}) for cross validation during refinement. The molecular replacement method was used for structure solution with the program Phaser (McCoy et al., 2007), as implemented in the ccp4i2 interface (Winn et al., 2011). The crystal structure of a homologous member of the UNC5 family, UNC5D, has already been deposited in the [PDB](#) (Jackson et al., 2016). Since it is involved in an octameric complex with several other members, the coordinates of the Ig1 and Ig2 domains of UNC5D ([PDB ID: 5FTT](#), chain A) were extracted and used independently as search models. The initial model was improved by several rounds of manual model building in Coot (Emsley et al., 2010), and refinement using REFMAC5 (Murshudov et al., 2011). During these manual building cycles the first TSP domain was built.

For the final structure, diffraction data from a single crystal was used. Data processing was performed using XDSGUI (Kabsch, 2010). The previously built model was used for molecular replacement with Phaser (McCoy et al., 2007). The final model was refined by successive rounds of manual model building in Coot (Emsley et al., 2010) and refinement with BUSTER (Bricogne et al., 2016).

2.17.3 UNC5B ecto sulphur SAD anomalous data collection

Sulphur-single wavelength anomalous diffraction (S-SAD) data was collected at the ID29 beamline (De Sanctis et al., 2012), a tuneable beamline specialized for anomalous data collection down to 5.2 keV.

During X-ray diffraction measurements the phase information required to calculate the 3D structure is lost. One method of overcoming this so called phase problem in protein crystallography is to experimentally determine these phases using anomalous scattering techniques. This scattering is dependent on the atom type and energy used. By careful selection of the experimental setup, the anomalous scattering effect of larger atoms can be measured, and used to determine the position of the anomalous scattering atoms and so derive the experimental phases.

Anomalous data, derived from the scattering of sulphur atoms at low energies, was used to validate the structure and identify the likely position of the second missing TSP domain of the UNC5B ectodomain. Although none of the sulphur absorption edges are within the range of standard macromolecular crystallography beamlines, its anomalous signal can still be measured with sufficient accuracy at around 6 keV. In this experiment the X-ray wavelength was set at 1.9 Å, corresponding to an energy of ~6.2 keV.

Both partial datasets (10° overall rotation) with the MeshAndCollect method (Zander et al., 2015), and single crystals datasets (360° overall rotation) were collected. These datasets were processed using either DIALS (Waterman et al., 2013, 2016) or XDS (Kabsch, 2010), and subjected to the ccCluster algorithm (Santoni et al., 2017). Data were collected to a resolution of 3.9 Å. The datasets were merged in Pointless (Evans, 2006, 2011), followed by scaling in Aimless (Evans and Murshudov, 2013), and converted to structure factors in Truncate

MATERIALS AND METHODS

(French and Wilson, 1978). Upon initial inspection of merging statistics, the observed anomalous signal achieved was only 8 Å. Phases were calculated by a single run of Refmac5 (Murshudov et al., 2011) using the previous structure obtained as input. The phases generated were merged with the anomalous DANO/SIGDANO columns using CAD (Winn et al., 2011). An anomalous difference map ($\Delta F_{\text{ano}}, \alpha_{\text{calc}} - 90^\circ$) at 8 Å was generated using FFT (Ten Eyck, 1973; Olthof-Hazekamp, 1978; Read and Schierbeek, 1988) through application of a Fourier transform. This map was loaded in Coot (Emsley et al., 2010) for visual inspection to verify the alignment of the sulphur containing side chains in the protein model with the anomalous density peaks observed and to help locate the missing TSP domain.

2.18 Bioinformatics tools

DNA sequences were visualized, translated and modified *in silico* with the program ApE, version 2.0.47 (Davis, 2017).

Protein disorder prediction analysis were performed using the [PrDOS](#) webserver (Ishida and Kinoshita, 2007).

Prediction of glycosylation sites was performed with the [NetNGlyc](#) webserver (Gupta et al., 2004).

Protein sequence alignments of the UNC5 proteins were performed with [MUSCLE](#) (Edgar, 2004a, 2004b; Li et al., 2015a; McWilliam et al., 2013) or [Clustal Omega](#) (Sievers et al., 2011), and graphics alignment figures created with the [ESPrInt](#) webserver (Robert and Gouet, 2014).

Protein structures and SAXS envelopes were visualized, and structure figures created using PyMOL version 1.8.6.0 (Schrödinger, LLC, 2015). Superimposition of the SAXS envelopes were performed with the aid of the SASpy plugin for PyMOL (Panjkovich and Svergun, 2016). Electrostatic potential surfaces were calculated with the aid of the APBS plugin for PyMOL (Baker et al., 2001; Dolinsky et al., 2004, 2007).

Homology models for use in the EOM were built with [SWISS-MODEL](#) (Arnold et al., 2006; Biasini et al., 2014; Guex et al., 2009; Kiefer et al., 2009) using the crystal structures of the closest Ig and FnIII domain homologues in the [PDB](#) as templates. In this case, [PDB](#) entries [2V9R](#) and [4HLJ](#) were chosen to model the Ig1-Ig2 domains, and the first and second FnIII domains respectively.

Evolutionary sequence conservation of human UNC5B across other species was analysed using the [ConSurf](#) webserver (Ashkenazy et al., 2016; Celniker et al., 2013; Glaser et al., 2003; Landau et al., 2005). The species included in the analysis were: *Mus musculus*, *Rattus norvegicus*, *Xenopus levi*, *Bos taurus*, *Gallus gallus*, *Canis lupus familiaris*, *Ovis aries*, *Danio rerio*, *Cavia porcellus*, *Pelodiscus sinensis*, *Ailuropoda melanoleuca*, *Papio anubis*, *Oryctolagus cuniculus*, *Felis catus*, *Salmo salar*, *Gorilla gorilla gorilla*, *Equus caballus*, *Myotis brandtii*.

3 RESULTS

Résumé en français

Des protocoles permettant d'exprimer les domaines extracellulaires complets de Robo4 et UNC5B ont été conçus. En solution, le domaine extracellulaire de Robo4 existe sous la forme d'un monomère glycosylé on azote. Des expériences de diffusion de la lumière à multi-angle (MALS) et de résonance plasmonique de surface (SPR) ont confirmé que cinq fragments d'anticorps parmi six, se lient au domaine extracellulaire de Robo4 dans un complexe de stœchiométrie 1:1. Le dernier fragment d'anticorps, bien qu'il ait montré une interaction, ne se lie pas suffisamment au domaine extracellulaire de Robo4. Des constantes cinétiques ont été calculées pour les cinq complexes et se situent dans la gamme de 1-6 nM, sauf par un qui se situe à ~30 nM. L'analyse par diffusion des rayons X aux petits angles (SAXS) confirme l'état monomérique et montre que le domaine extracellulaire de Robo4 prend une forme allongée. Cependant, en raison du haut degré de flexibilité, nous n'avons pas pu déterminer les positions relatives de chaque domaine dans les modèles calculés à partir des données de diffusion. De même, l'état oligomérique du domaine extracellulaire d'UNC5B a également été étudié, et a été trouvé monomérique en solution. De plus, son association avec le domaine extracellulaire de Robo4, qui a été décrit dans une étude précédente, a été étudiée. Nous n'avons obtenu aucune preuve de liaison directe entre les domaines extracellulaires d'UNC5B et Robo4. Un résultat confirmé par chromatographie d'exclusion de taille (SEC), MALS, SPR et un criblage d'interaction extracellulaire avec la technique AVEIXIS. Des essais de cristallisation du domaine extracellulaire de Robo4 sous sa forme native et déglycosylé, tout seul ou en complexe 1:1 avec des fragments anticorps, ont été réalisés. Les cristaux obtenus n'étaient pas de qualité suffisante pour la résolution de la structure. De plus, la structure cristallographique du domaine extracellulaire complet d'UNC5B a été résolue à 3,3 Å de résolution et validée en croisant la position de la diffusion anormale des atomes de soufre par la technique de diffusion anormale de longueur d'onde unique (SAD).

RESULTS

3.1 Expression and purification of human Robo4 constructs

One of the objectives of this thesis work was the structural characterisation of the Robo4 transmembrane receptor, in particular of its extracellular domains.

It was suggested that the cytoplasmic domain of Robo4 has the characteristics of a disordered protein (Hohenester, 2008) and disorder prediction confirms that the cytoplasmic domain (amino acids 491-1007) is substantially less ordered than the extracellular domain (amino acids 28-469) (Figure 12).

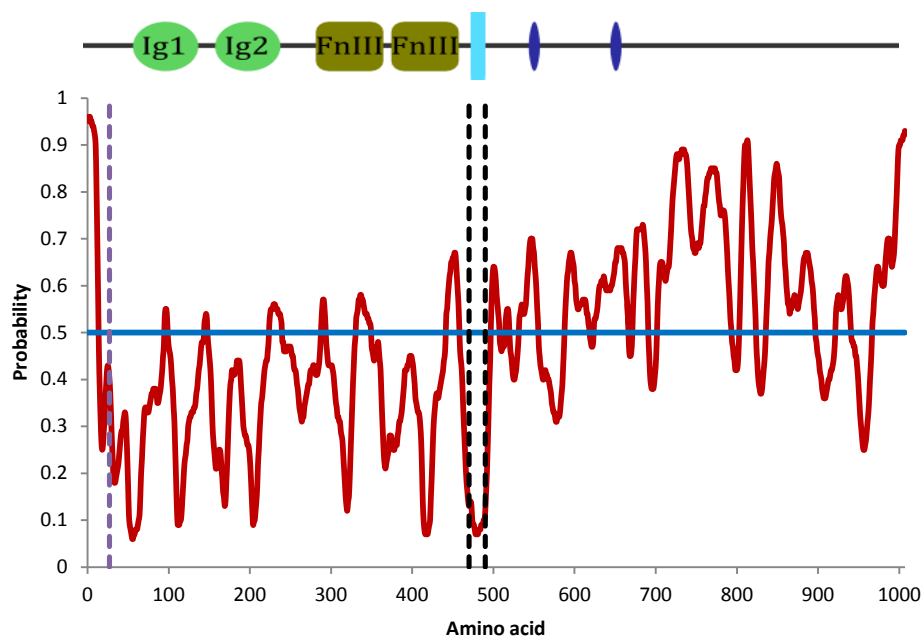


Figure 12: Disorder prediction of human Robo4

Disorder prediction of human Robo4 calculated by PrDOS. The threshold in blue is set at 0.5. At the top is the Robo4 domain organisation. The signal sequence (amino acids 1-27) and the membrane spanning region (amino acids 470-490) are delineated in purple and black dashed lines respectively. Higher disorder is predicted at the N-terminal signal sequence and the cytoplasmic region.

It was therefore decided to focus on the extracellular region of Robo4, characterized by two Ig and two FnIII domains. Since transmembrane receptors often undergo PTMs, which are necessary for their function and folding (Karve and Cheema, 2011; Knorre et al., 2009), and the extracellular domain of Robo4 has some predicted glycosylation sites (Table 6), a mammalian expression system was established to express different variants of the Robo4 extracellular domain. To avoid the difficulties of membrane protein purification and crystallisation, it was decided to design constructs excluding the transmembrane region that would be expressed as secreted proteins. This approach allows processing through the

secretion pathway, ensuring PTM addition and exportation from the endoplasmic reticulum if correctly folded, or degradation of unstable fragments (Anelli and Sitia, 2008; Hammond and Helenius, 1995). The secreted proteins can therefore be purified directly from the media. Although the protein content of FBS is high, affinity tag purification is suitable for the purification of most target proteins with significantly less contaminants than traditional purification from total cell extract. This can considerably simplify the purification protocol, depending on the tag utilised, and yield quite pure products. On the other hand, nickel affinity purification from mammalian cell media is occasionally subject to the obstacle of chelating agents in the commercial media.

Table 6: Predicted Robo4 glycosylation sites

Sites of N-linked glycosylation as predicted by NetNGlyc.

Amino acid	Type	Confidence
246	N-linked	0.72
360	N-linked	0.79
389	N-linked	0.69
396	N-linked	0.57

Several of the shorter Robo4 constructs prepared for expression (Figure 11) showed little, or no, expression. For structural characterisation, the full length Robo4 ectodomain (28-462), encompassing the membrane proximal region before the start of the membrane spanning α -helix (470-490), was the most promising target in terms of expression level and stability. Most of the work presented in this thesis is focused around this construct, referred to as Robo4 ecto.

Mutagenesis of the glycosylation sites on this construct to alanine or aspartic acid was also attempted (§ 2.3.1), but resulted in extremely poor or loss of expression of this construct, therefore it was not pursued further.

Due to its low cost and adaptability to litre scale up, the main purification method selected in this thesis was Nickel affinity purification. The standard purification protocol included a complete media exchange through a dialysis membrane to eliminate chelating agents that hinder purification by stripping of the metal ions from the affinity resin. For the experiments described in this thesis several litres of

RESULTS

media were routinely handled. This required dialysis, or diafiltration, steps that prolonged the time needed for affinity purification alone to a whole day. An optimised protocol was finally established, which instead took advantage of a special type of nickel resin, developed for resistance to mild stripping condition. This process allowed to perform the purification process in a single day.

Robo4 ecto was expressed and purified both from HEK293 and HEK293S cell lines. In both purified products, the elution volume from size exclusion chromatography, and the migration on SDS-PAGE (Figure 13), were unusual for a protein of 48 kDa. On a HiLoad Superdex 200 16/60 column, the HEK293 expressed Robo4 ecto eluted as a homogenous peak at 69.9 ml (Figure 13A), while the HEK293S expressed Robo4 ecto eluted at 71.9 ml (Figure 13D). Compared to molecular weight standards, elution peaks at these volumes would be expected from proteins between 100-150 kDa.

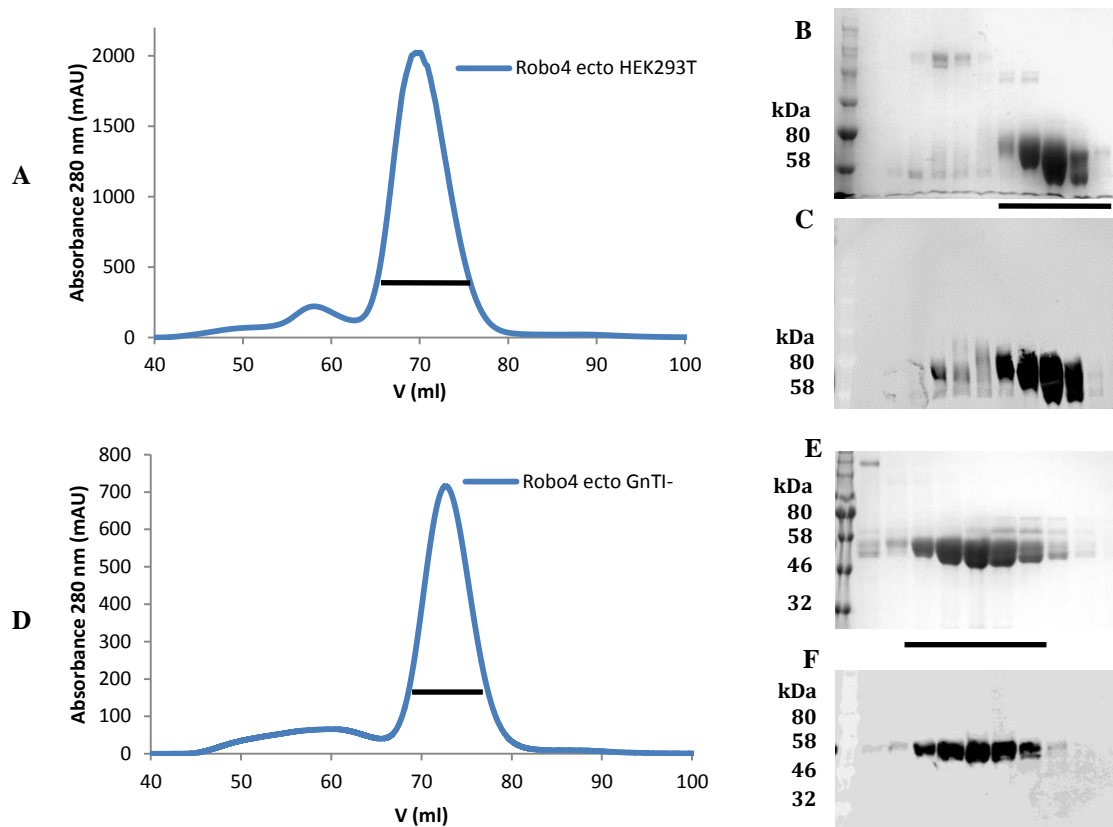


Figure 13: Purification of Robo4 ectodomain

A: Robo4 ectodomain purified after expression in HEK293 cells. Size exclusion chromatography on HiLoad Superdex 200 16/600. **B:** SDS-PAGE. **C:** Western blot anti-His.

D: Robo4 ectodomain purified after expression in HEK293S GnTI⁻ cells. Size exclusion chromatography on HiLoad Superdex 200 16/600. **E:** SDS-PAGE. **F:** Western blot anti-His.

The black line indicates the fractions of each peak from gel filtration.

On SDS-PAGE gels, both products appeared as large smears migrating between 50 and 70 kDa for the HEK293 expressed sample (Figure 13B), and between 48 and 58 kDa for the HEK293S expressed one (Figure 13E). The presence of the His-tagged protein was confirmed by western blot analysis using anti-His antibodies. Comparing the western blot membranes to the acrylamide gels confirmed that the whole smears represent the His-tagged protein of interest, Robo4 ecto (Figure 13C and Figure 13F). Although a contaminant is visible in the purified HEK293S expressed sample (Figure 13E), it is not present across the whole peak, and no other co-eluting products were identified on SDS-PAGE gel or western blot.

It is known, that heterogeneous glycosylation of transmembrane receptors and extracellular proteins can result in anomalous elution volumes and SDS-PAGE migration patterns (Chang et al., 2007; Selcuk Unal et al., 2008). Comparison of the HEK293 and HEK293S products by SDS-PAGE shows how impairing the glycosylation pathway can influence the final product (Figure 13B and E).

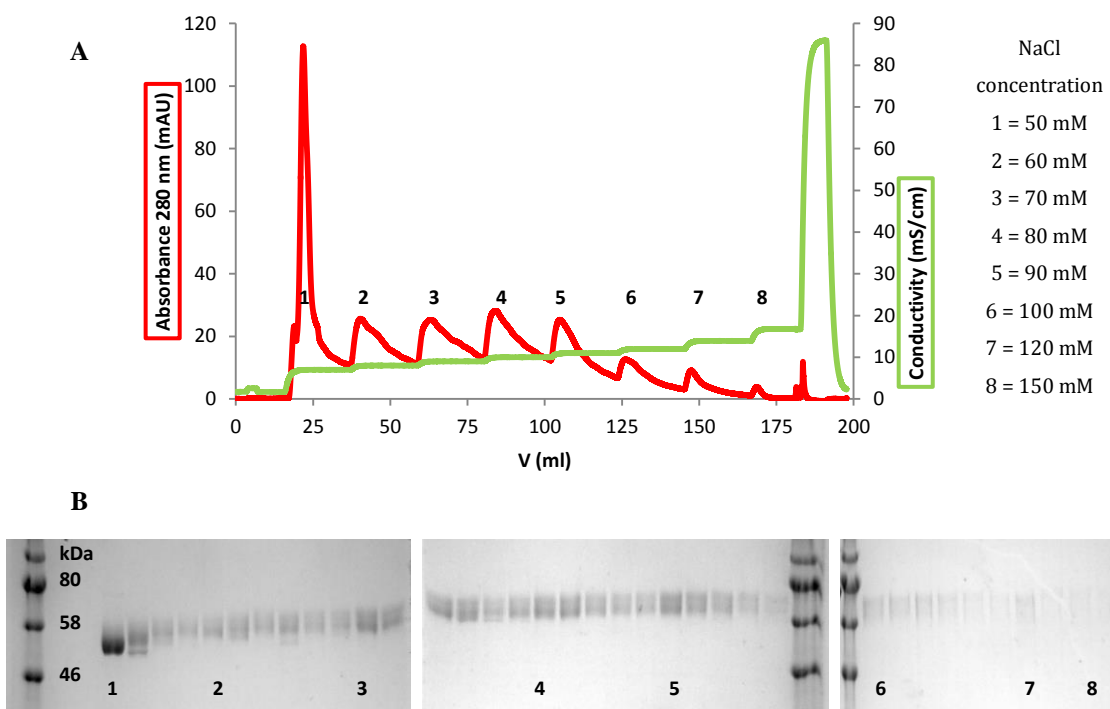


Figure 14: Ion exchange chromatography of Robo4 ecto

A: Ion exchange profile.

B: SDS-PAGE of fractions.

Each number marks one step at increasing sodium chloride concentration. At 50 mM sodium chloride the lowest band, which constitutes a large part of the fully glycosylated Robo4 ecto sample, can be separated, although part of the higher molecular weight smear is still visible. The rest of the high migrating glycosylated species elute gradually at each increasing elution step.

RESULTS

In the purified HEK293 expressed Robo4 ecto two main bands can be distinguished (see Figure 13B). Since they could not be resolved by gel filtration, ion exchange chromatography was attempted to increase the homogeneity of the sample (Figure 14A). Only the lower band could be partially separated, while the higher band was progressively eluted through increasing salt concentration (Figure 14B). This approach was eventually deemed not promising enough for further optimisation, resulting in inferior sample quality compared to enzymatic deglycosylation, which is discussed later (§ 3.2).

3.2 Robo4 ecto deglycosylation

In order to investigate the presence of heterogeneous glycosylation on Robo4 ecto, which could negatively impact on crystallisation attempts, deglycosylation trials were performed using one of three different enzymes: PNGase F1, Endo H or Endo F1. While the first two enzymes are commercially obtainable, Endo F1 was produced *in house*, as described in § 2.10.

Endo H and Endo F1 have similar activities. Both cleave high mannose and hybrid oligosaccharides (but not complex) between the first and second NAG residue of N-linked glycans, leaving one NAG attached to the asparagine. Endo F1, however, will cut sulphated high-mannose oligosaccharides, and is less sensitive to protein conformation, while Endo H can cut fucosylated oligosaccharides. Endo F1 also has a decreased (50-fold) activity in the presence of α 1-6 fucose linked to the first core NAG, which is common in vertebrates (Maley et al., 1989; Tarentino et al., 1992).

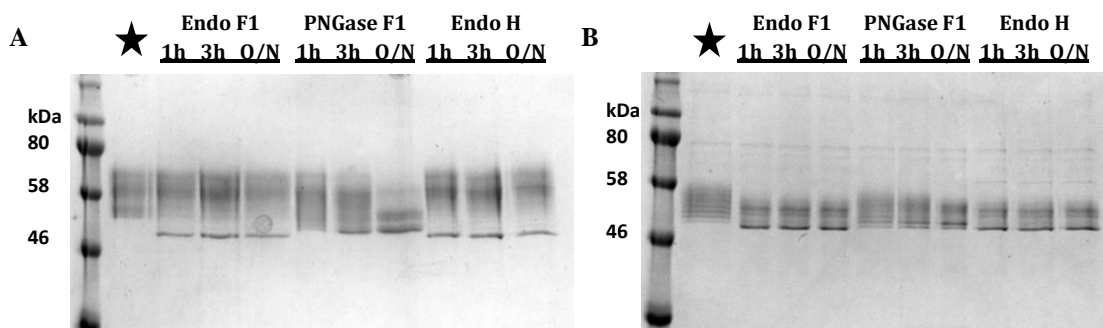


Figure 15: Deglycosylation of Robo4 ecto

Deglycosylation was performed in the protein purification buffer. For each enzyme depicted a sample was taken after 1 hour, 3 hours and overnight incubation. The star indicates time 0 for each sample.

A: Robo4 ecto overexpressed and purified from HEK293.

B: Robo4 ecto overexpressed and purified from HEK293S.

PNGase F1 is active on all three types of N-glycans (oligomannose, hybrid and complex), and in the presence of α 1-6 fucosylation, but not when there is α 1-3 fucosylation (which is commonly found in invertebrates, including insect cells, and plants). Cleavage takes place between the innermost NAG and the protein's asparagine residues, effectively converting it to an aspartate. Furthermore, PNGase F1 activity is sometimes impeded by protein folding, being most active on denatured proteins (Tarentino and Plummer, 1994).

As determined by SDS-PAGE analysis, deglycosylation of the fully glycosylated Robo4 ecto was incomplete, even after O/N incubation using either Endo F1 or Endo H, while PNGase F1 needed an overnight incubation before starting to produce a more homogenous sample (Figure 15A). In all three cases, formation of a prominent band migrating at ~48 kDa is visible, which corresponds to the expected molecular weight of deglycosylated Robo4 ecto.

Deglycosylation of the Robo4 ecto expressed in HEK293S cells showed no added benefit over the one hour incubation mark using Endo F1 or Endo H, while PNGase F1 seemed to achieve the same result only after overnight incubation (Figure 15B).

Treatment with Endo F1 and Endo H was therefore selected to use with Robo4 ecto expressed in HEK293S cells.

RESULTS

3.3 SEC-MALS analysis of Robo4 ecto

It was previously suggested that Robo4 can dimerise (Yadav and Narayan, 2014) in a process induced either by its extracellular or cytoplasmic domain (Bedell et al., 2005). SEC-MALS was performed to verify the oligomeric state of the receptor and to observe the average molecular weight difference caused by heterogeneous glycosylation between the Robo4 ecto expressed in HEK293 and HEK293S cell lines. The fully glycosylated ectodomain had a mass of 56.1 kDa, while Robo4 ecto overexpressed and purified from HEK293S had a slightly lower mass of 54.1 kDa, with a difference of at least 6 kDa compared to the expected weight of a non-modified ectodomain (see Table 7).

Table 7: SEC-MALS calculated molecular weight of Robo4 ecto

The predicted molecular weights based on the Robo4 ecto sequence, and the observed molecular weights derived from MALS measurements with the polydispersity of each sample, and errors, are shown.

	Robo4 ecto HEK293	Robo4 ecto HEK293S
Predicted MW	48.1 kDa	48.1 kDa
Observed MW	56.1 kDa \pm 1.1 %	54.6 kDa \pm 0.9 %
Polydispersity	1.000 \pm 1.5 %	1.001 \pm 1.3 %

The difference in measured weight between the two expressed samples is <2 kDa, and within instrumental error, which is set at 5% of the total mass. The HEK293S expressed sample (Figure 16, blue line), however, showed a narrower peak with a shift from 12.9 to 13.1 ml in the elution volume compared to the HEK293 expressed Robo4 ecto (Figure 16, black line). This change in the elution peak can be interpreted as a change in the overall hydrodynamic radius of the protein, which is influenced by the type of glycosylation. In the experimental conditions used, no dimerisation of the extracellular domain was observed.

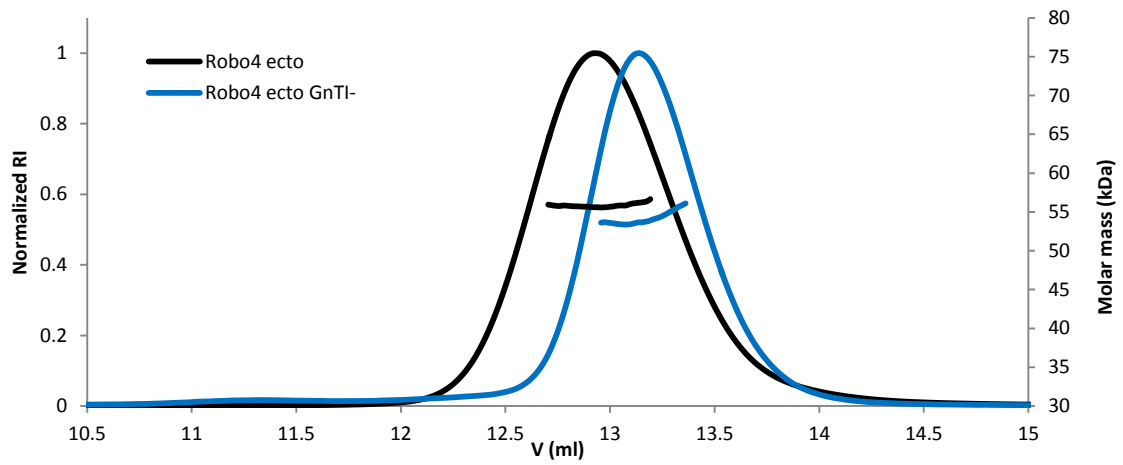


Figure 16: SEC-MALS comparison of Robo4 ecto in HEK293 and HEK293S cell lines

Black: Robo4 ectodomain overexpressed in the HEK293 cell line. Blue: the same construct overexpressed in the glycosylation deficient HEK293S cell line. The same amount of protein was injected and runs were performed consecutively.

RESULTS

3.4 Purification of Fabs

The important function of Robo4 in angiogenesis (Bedell et al., 2005; Koch et al., 2011; Park et al., 2003; Yadav and Narayan, 2014) makes it an interesting target for therapeutic strategies. A collaboration with the Sidhu group from the University of Toronto was therefore started to try and study the effect of Fab binding to the Robo4 receptor. Moreover, there was the prospect that an increase in stability induced by Fab binding would favour crystallisation of the complex. The Sidhu group selected the Fabs used in this study from their Robo4 interaction library, and Table 8 reports the binding sites of each Fab on Robo4 ecto as determined by them.

Table 8: Map of suggested Fabs binding onto Robo4 ecto

	Fab5555	Fab5562	Fab5564	Fab5570	Fab 5582	Fab5585
Robo4 ecto domains bound	Ig1	Ig1	Ig2	Ig2	Ig1 / Ig2	Ig1 / Ig2

To avoid loss of sample quality during transport, it was decided to directly produce the Fabs *in house*. All Fabs were expressed in BL21 RIL codon plus *E. coli* (§ 2.9). The yield was usually around 2 mg of protein per litre of culture.

Except for Fab5582, unspecific binding to the gel filtration resin of the free Fabs (Fabs not in a complex) was observed during purification. In order to overcome this, the final gel filtration step was omitted since it did not influence the quality of the purified product. All Fabs were purified following the same protocol as detailed in § 2.9, and purity was generally >85% (Figure 17).

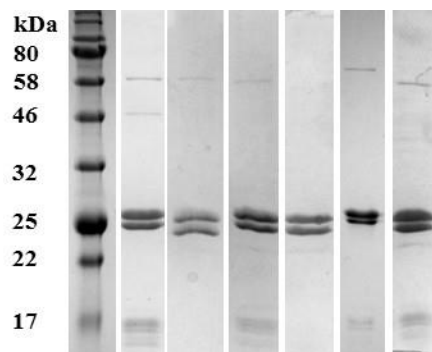


Figure 17: SDS-PAGE of purified Fabs

Composite image of final purification step of each expressed Fab. From left to right: Fab5555, Fab5562, Fab5564, Fab5570, Fab5582, Fab5585. The two visible bands correspond to light and heavy chain forming each Fab.

3.5 Fab probing of Robo4 extracellular domain binding

To verify their interaction, Robo4 ecto and Fabs were mixed in 1:1 ratio, incubated for 1 hour at 4°C, and analysed by gel filtration. SEC-MALS was performed on each Robo4/Fab to verify their complex formation (§ 2.13).

A clear shift in retention volume and mass was observed after incubating Robo4 ecto with each of the following Fabs: Fab5555, Fab5562, Fab5564, Fab5570 or Fab5585. All complexes eluted in monodisperse symmetric peaks close to an elution volume of 12 ml, except for Robo4 ecto/Fab5555 and Robo4 ecto/Fab5564 which showed a slightly skewed peak. The Robo4 ecto/Fab5585 complex eluted in a broader symmetric peak compared to the other complexes (Figure 19). The observed molecular weights are summarised in Table 9. Due to unspecific interaction with the gel filtration matrix, the molecular weight of the single Fabs could not be measured, except for Fab5582 (Figure 18). Based on their sequence, all Fabs have an expected molecular weight of ~50 kDa. Interestingly, while bound to their intended target, the Fabs did not interact with the gel filtration matrix or cause retention of the complex on the resin.

Table 9: Observed molecular weight of Robo4 ecto/Fabs complexes from SEC-MALS

Predicted molecular weights based on Robo4 ecto and Fabs sequences, the observed molecular weight derived from MALS measurements and the polydispersity of each sample, with relative errors, are reported. Fab5582 did not form a detectable complex with Robo4 ecto.

	Robo4 ecto Fab5555	Robo4 ecto Fab5562	Robo4 ecto Fab5564	Robo4 ecto Fab5570	Robo4 ecto Fab5582	Robo4 ecto Fab5585
Predicted MW	98.6 kDa	99.2 kDa	99.7 kDa	98.7 kDa	99.7 kDa	98.9 kDa
Observed MW	96.82 kDa ± 1.3 %	95.03 kDa ± 1.2 %	95.24 kDa ± 1.1 %	97.04 kDa ± 1.2 %	N/A	81.37 kDa ± 1.3 %
Polydispersity	1.000 ± 1.8 %	1.003 ± 1.7 %	1.001 ± 1.6 %	1.002 ± 1.6 %	N/A	1.010 ± 1.8 %

The Robo4 ecto/Fab5555, Robo4 ecto/Fab5562, Robo4 ecto/Fab5564 and Robo4 ecto/Fab5570 complexes all showed similar elution profiles and their mass corresponds to Robo4 ecto/Fab heterodimeric complexes. The measured polydispersity was within the expected range, confirming the validity of the measured values.

RESULTS

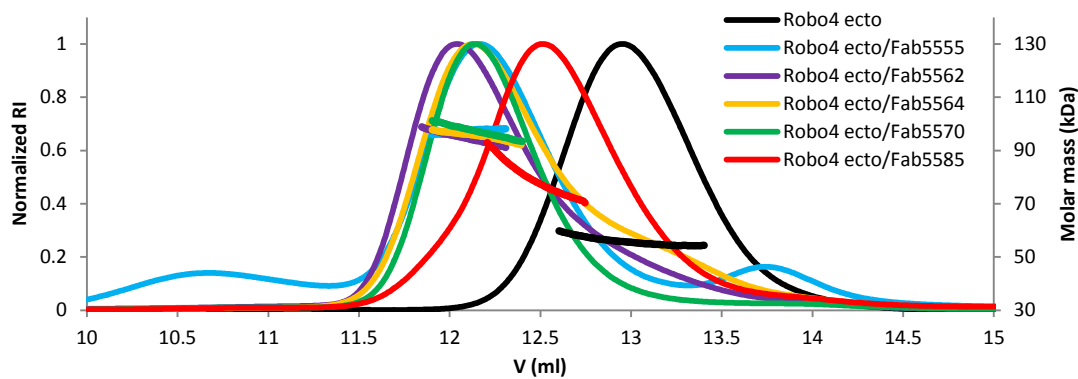


Figure 19: SEC-MALS of Robo4 ecto in complex with Fabs

SEC-MALS analysis of Robo4 ectodomain in complex with Fab5555, Fab5562, Fab5564, Fab5570 or Fab5585. Injection of Robo4 ectodomain alone (black) is shown for reference.

All complexes, except for the Robo4 ecto/Fab5585 (red) have similar profiles.

The Robo4 ecto/Fab5585 complex, instead, showed a peculiar behaviour (Figure 19, red outline). A sharp decrease in molecular weight from ~ 100 kDa to ~ 70 kDa could be observed, with an average of 81.4 kDa measured from start to end of the elution peak. The peak maximum was also shifted to 12.5 ml, which is in between the elution volume of each Robo4 ecto/Fab complex and the elution volume of Robo4 ecto alone (Figure 19, black outline). The measured polydispersity is also significantly higher than the other complexes, which all together indicate the presence of partial complex formation with a fast association/dissociation kinetic.

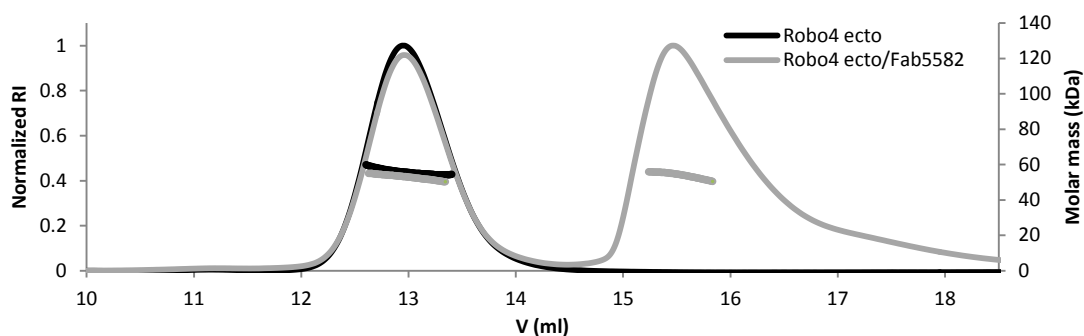


Figure 18: SEC-MALS of Robo4 ecto and Fab5582

SEC-MALS analysis of Robo ecto in complex with Fab5582 (grey). Injection of Robo4 ecto alone is shown for reference (black).

Despite being specifically selected by our collaborator, Fab5582 showed no signs of binding (Figure 18, grey). Curiously, in addition to not forming an observable complex by size exclusion chromatography, Fab5582 was the only Fab that did not bind the gel filtration matrix, and eluted close to its expected retention volume for a protein of 51.5 kDa.

3.5.1 Determination of interaction constants

To quantify the strength of interaction, the binding of Robo4 ecto to each Fab was measured by SPR. The best immobilization conditions were determined using the pH scouting method. The ligands were diluted in 10 mM sodium acetate buffer at pH 4, 4.5, 5 and 5.5. The tests were performed by serially injecting 10 $\mu\text{g}/\text{ml}$ of the protein ligand in each buffer for 120 s over a non-activated chip surface. The best immobilization strategy was determined as the condition with the highest pH that showed a significant association to the chip surface. In the case of Robo4 ecto, the best immobilisation was performed at pH 5 (Figure 20).

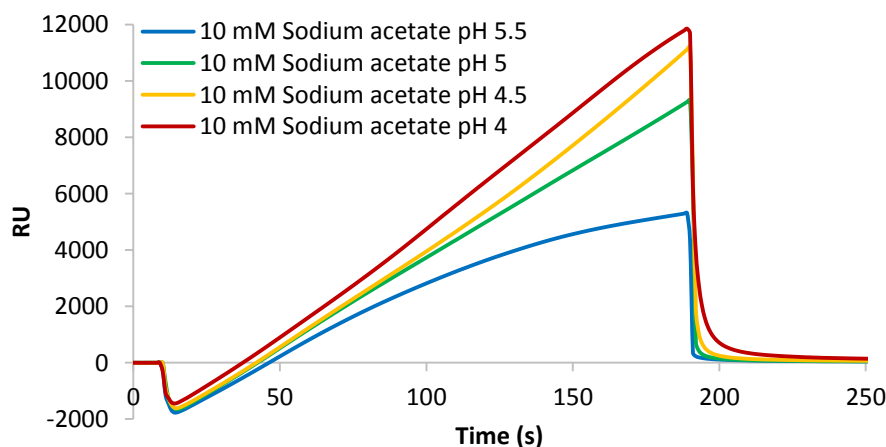


Figure 20: pH scouting of Robo4 ecto

10 mM Sodium acetate at pH 5.5 (blue), pH 5 (green), pH 4.5 (orange) and pH 4 (red) was used to scout the best immobilization strategy. pH 5 was chosen as the highest pH that allowed for sufficient immobilization.

Binding kinetics of Fab5555, Fab5562, Fab5564 and Fab5570 were determined by fitting the data against the 1:1 Langmuir binding model (Figure 21). Kinetic constants were determined by measuring the association/dissociation constants at multiple concentrations as described in § 2.14.2 and reported in Table 10.

RESULTS

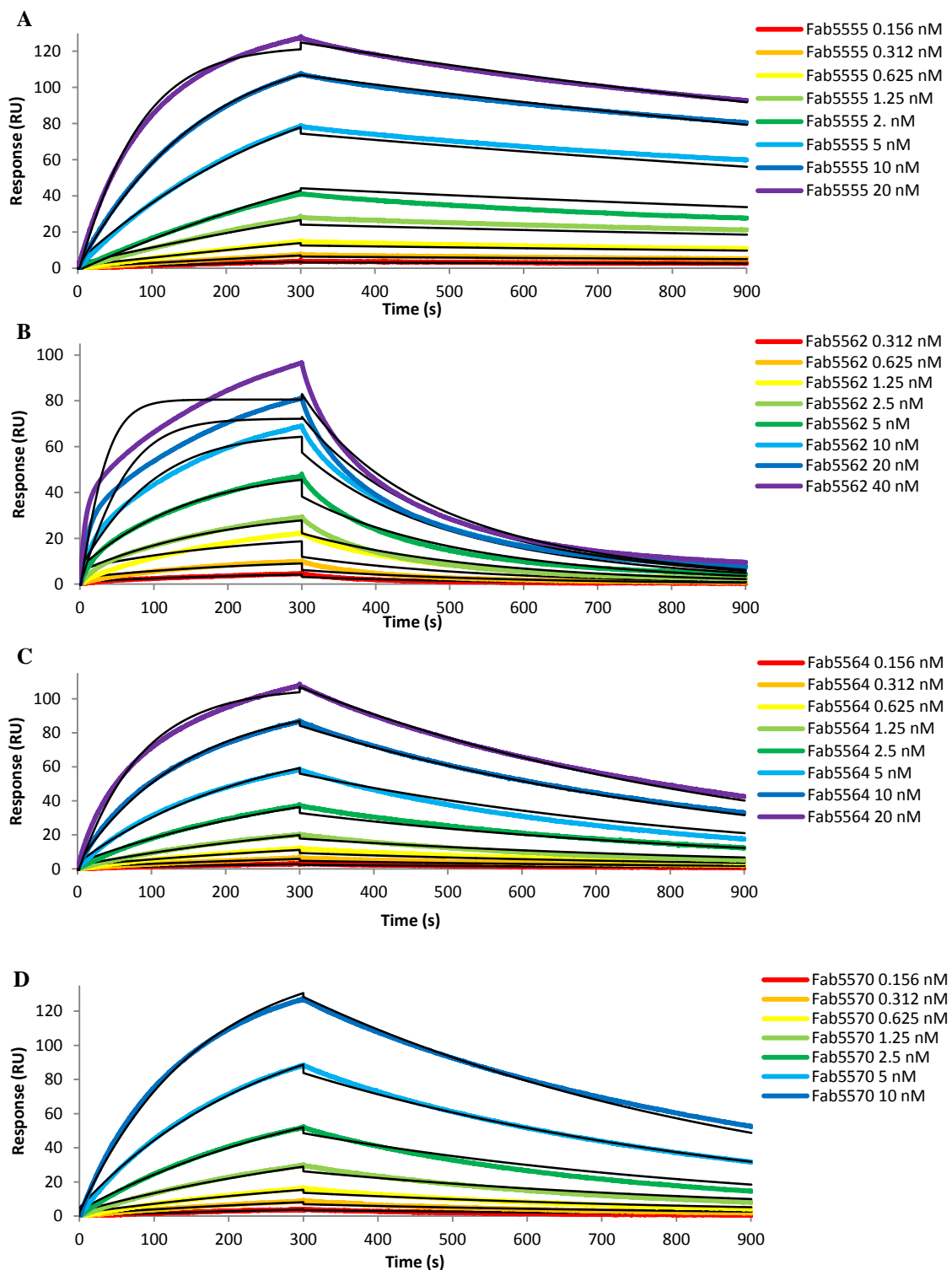


Figure 21: SPR analysis of Fabs: 5555, 5562, 5564 and 5570 binding to Robo4 ecto

Robo4 ecto was immobilized as ligand. Data was analysed with the Langmuir 1:1 binding model.

A: Fab5555.

B: Fab5562. The model did not fit the data satisfactorily. Between the concentrations of 0.625-1.25 nM, 2.5-5 nM and 5-10 nM, a sudden jump in response can be observed, instead of a linear increase.

C: Fab5564.

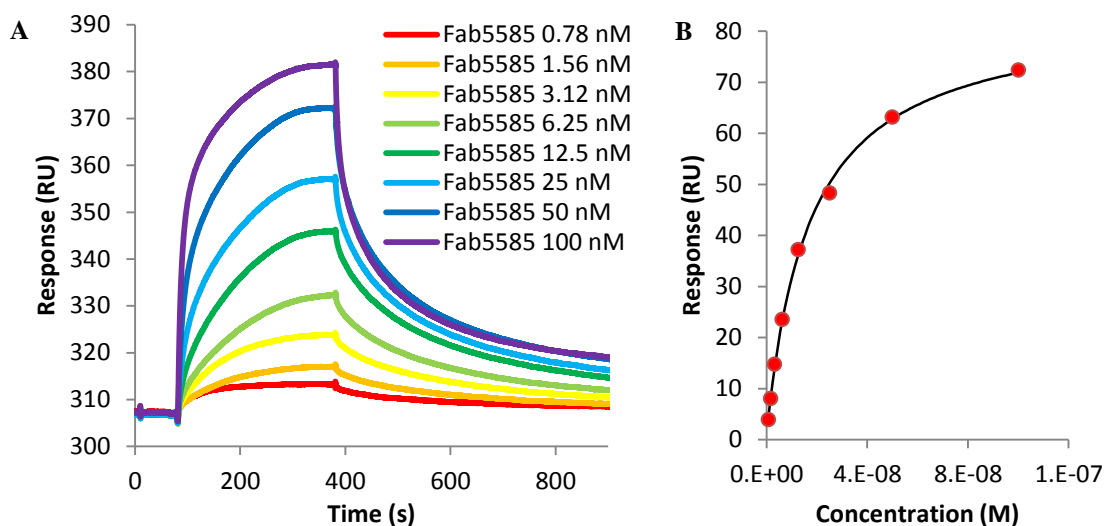
D: Fab5570.

Table 10: Kinetic constants of Fab binding to Robo4 ecto

Experimental R_{\max} , Chi^2 and U-value are reported as quality indicators of the analysis. The expected response was 100 RU. Chi^2 is a measure of the average deviation of the experimental data from the fitting curve. U-value indicates the uniqueness of the calculated kinetic constants, the lower the value the highest the confidence. The Langmuir 1:1 binding model was used to calculate the binding constant of Fab5555, Fab5562, Fab5564 and Fab5570 binding. The K_D of Fab5585 was derived from steady state analysis, so the experimental R_{\max} and U-value are not applicable in this context.

	K_D	Experimental R_{\max}	Chi^2	U-value
Fab5555	0.9 nM \pm 0.2	131	4.8	2.4
Fab5562	5.8 nM \pm 0.5	95	9.0	3.1
Fab5564	2.6 nM \pm 0.6	127	1.9	0.8
Fab5570	2.8 nM \pm 0.6	142	7.7	1.8
Fab5585	28.5 nM \pm 10	N/A	15.2	N/A

The interaction profiles with Fab5555 (Figure 21A), Fab5564 (Figure 21C) and Fab5570 (Figure 21D), confirmed that these Fabs bind to Robo4 ecto following a simple 1:1 kinetic. Binding of Fab5562 to Robo4 ecto, instead, did not fit the same model (Figure 21B). Sudden jumps can be observed in the binding profile between the injections at 0.625-1.25 nM, 2.5-5 nM and 5-10 nM concentration, which indicated that a more complex kinetic interaction was involved that cannot be described by the 1:1 Langmuir model. Although the dynamic is unclear, the affinity given can still be considered as an average value within the system.

**Figure 22: Steady state analysis of Fab5585 binding to Robo4 ecto**

A: Sensorgram of Fab5585 binding to Robo4 ecto. The maximum response obtained with Fab5585 was lower than expected and the standard model resulted in an unsatisfactory fit to the data.

B: Steady state analysis fit. Each point corresponds to one injection of Fab5585 at the concentration described in Table 5 of § 2.14.2. The fitted curve is shown in black.

Although the fitting of the curve appears satisfactory, the binding might not have been completely saturated. It is, however, sufficient for an approximation of the K_D

RESULTS

Fab5585 reached only from 50 to 80% of the expected maximum response (Figure 22A shows the best dataset) and the model could not satisfactorily fit the data. Together with the information derived by SEC-MALS (§ 3.5), this might be a case where the binding site is only partially accessible and there is a fast exchange between different molecules. Since the binding was close to saturation, the K_D was approximated by using steady state analysis (Figure 22B) and shows an affinity 10 to 40 times lower than the other Fabs.

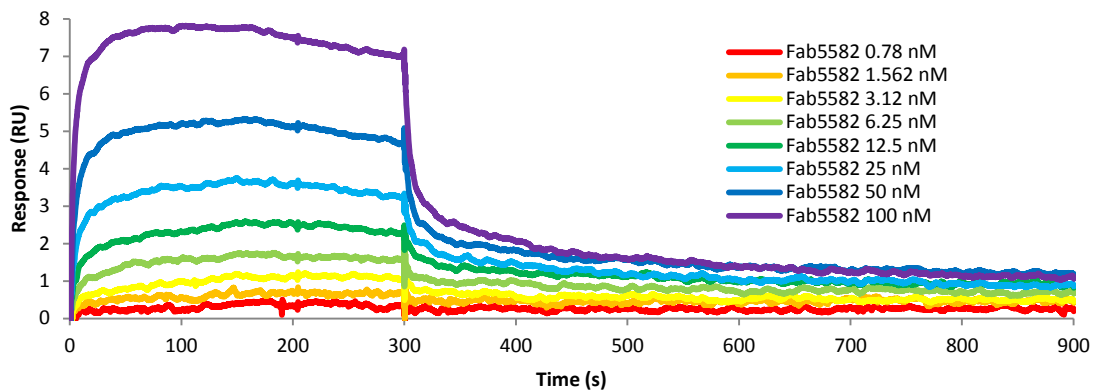


Figure 23: SPR analysis of Fab5582 binding to Robo4 ecto

Robo4 ecto was immobilized as ligand. A low response of <10 RU was observed. This data suggests partial binding of Fab5582 to Robo4 ecto. Kinetic analysis could not be performed.

Although the binding of Fab5582 was not detectable by gel filtration or SEC-MALS (§ 3.5), SPR still showed evidence of binding, as illustrated in Figure 23. The response was too low compared to the expected maximum response of 100 RU to effectively make any accurate measurement of the affinity. The data indicates that only a small proportion of Fab5582 was able to interact with Robo4 ecto.

3.6 SEC-SAXS derived structural information on Robo4 ecto

To obtain low resolution structural information on the Robo4 extracellular domain and to identify the relative positions of its domains, SAXS analysis of Robo4 ecto alone, and in complex with each different Fab was performed (Figure 24 and Figure 25A). Data were collected using an inline chromatography system to reduce the eventual effects of particle aggregation. As judged by the Guinier plot calculated from the scattering data, all samples were aggregate free and did not show interparticle interactions (Figure 25B to G).

Analysis of the dimensionless Kratky plot shows that Robo4 ecto alone does not assume the shape of a globular particle (which would be a bell shape), but exhibits a very high flexibility, as observed by a slow decay on the Kratky plot (Figure 24).

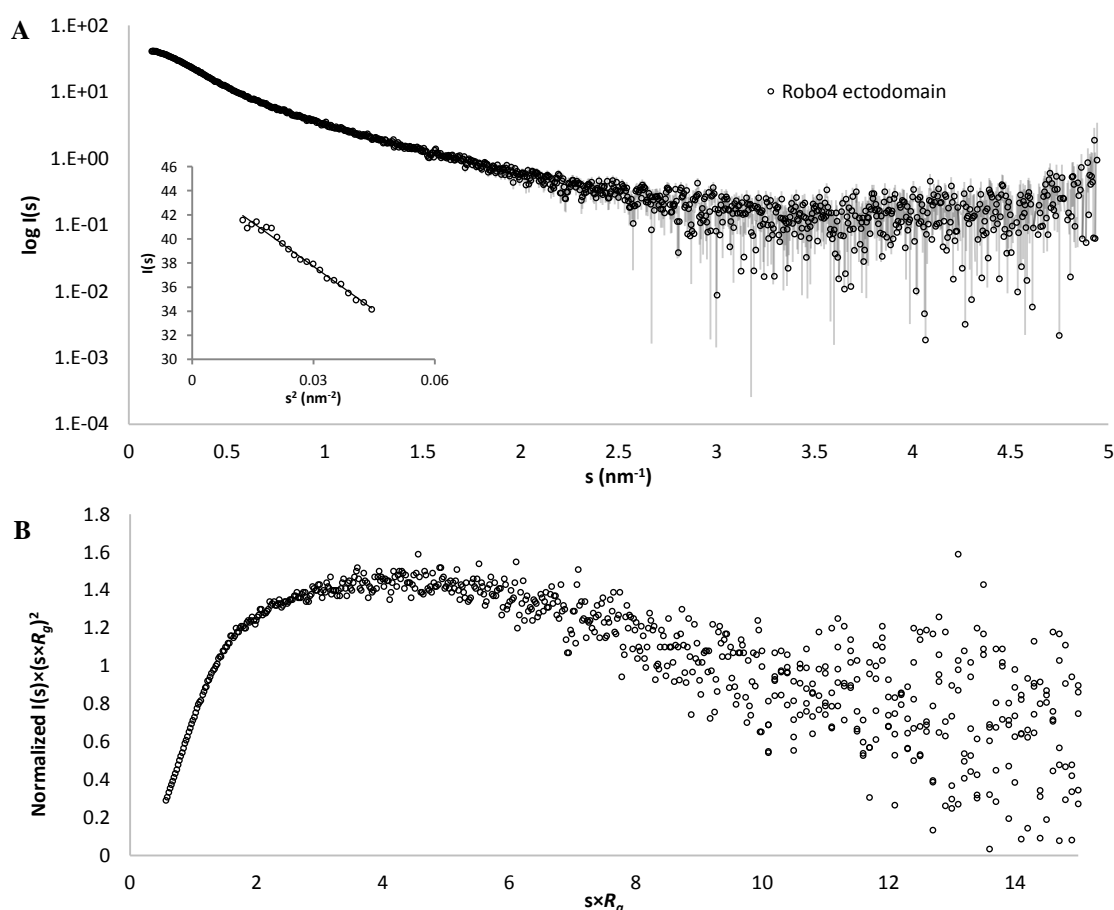


Figure 24: Robo4 ecto SAXS curves

A: Experimental SAXS curve of the Robo4 extracellular domain. In grey: experimental error bars. The small graph contains the calculated Guinier plot of the data within the $s \leq 1.3$ limit. No signs of aggregation or interparticle interaction were observed.

B: Dimensionless Kratky plot derived from the scattering curve, showing the characteristic profile of a very flexible non globular particle.

RESULTS

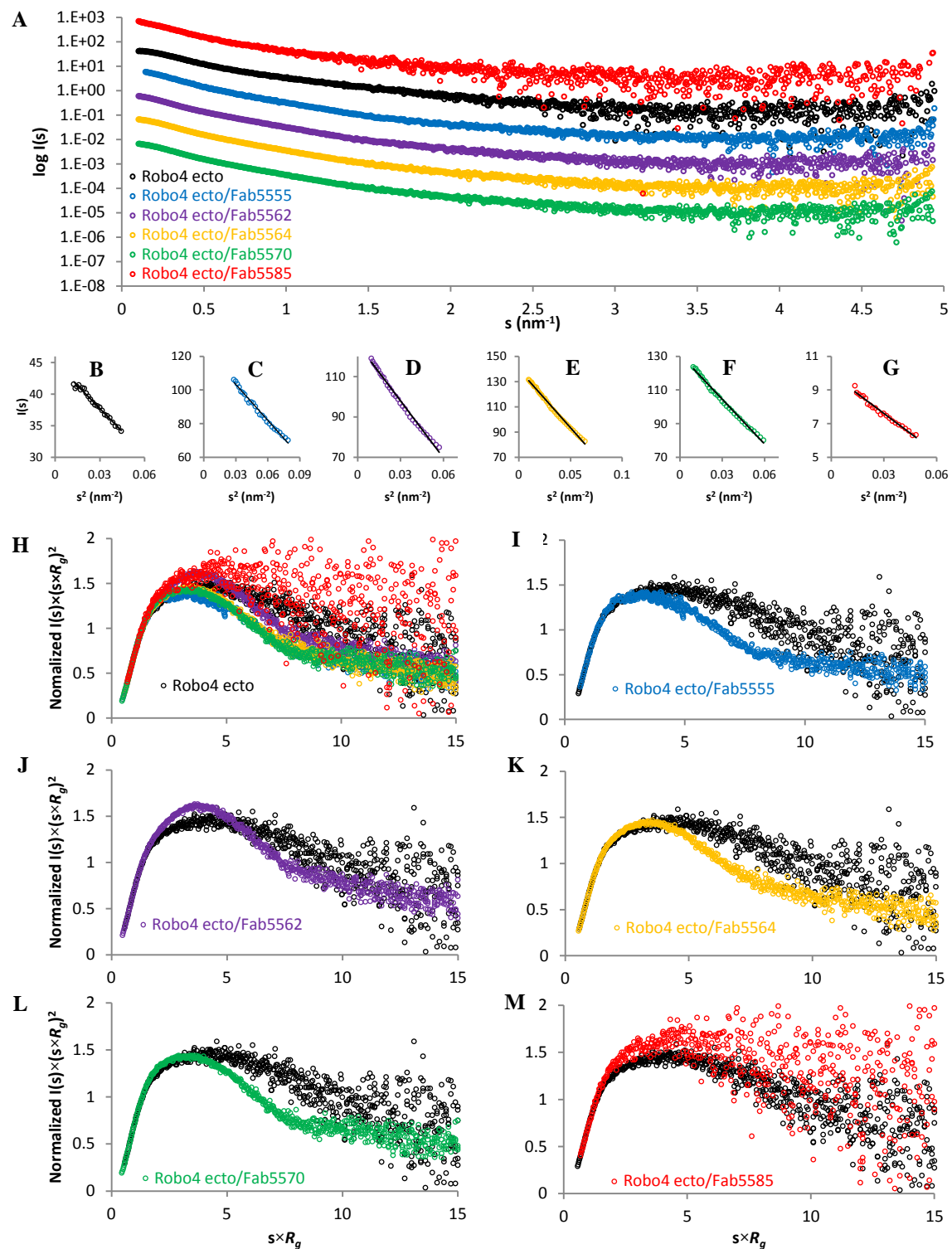


Figure 25: SAXS curves, Guinier and Kratky plots of Robo4 ecto and all Fab complexes

Black: Robo4 ectodomain. Blue: Robo4 ecto/Fab5555. Purple: Robo4 ecto/Fab5562. Yellow: Robo4 ecto/Fab5564. Green: Robo4 ecto/Fab5570. Red: Robo4 ecto/Fab5585.

A: SAXS curves of Robo4 ectodomain and each Robo4 ecto/Fab complex. Curves were scaled to better show the differences in curvature.

B to G: Guinier plot analysis of each curve. The Guinier range was respected within the $s \leq 1.3$ limit.

H: dimensionless Kratky plot of Robo4 ecto and each Fab complex superimposed.

I to M: dimensionless Kratky plot of each Robo4 ecto/Fab complex superimposed with the Robo4 ecto Kratky plot (in black) to highlight the differences in shape.

Robo4 ecto in complex with Fab5555, Fab5562, Fab5564 or Fab5570 shows a more accentuated bell-shaped curve, accompanied by a decrease in flexibility, as shown by a faster approach to 0 of the dimensionless Kratky plot (Figure 25H to M). This is interpreted to indicate an increased globularity caused by complex formation. Instead, formation of a complex between Robo4 ecto and Fab5585 results in an increase of disorder of the protein complex, as can be seen in Figure 25M. This could be due to Fab5585 having a detrimental effect on the folding of Robo4 ecto, or because Fab5585 itself was partially unfolded, summing up to an overall instability of the system.

Table 11: SAXS parameters of Robo4 ecto and Fab complexes

R_g : radius of gyration. GNOM R_g : GNOM calculated radius of gyration from $p(r)$. D_{max} : maximum dimension of the particle. V_p : Porod volume and in brackets the Porod exponent (ideally <3). V_c : correlated volume.

	Robo4 ecto	Robo4 ecto Fab5555	Robo4 ecto Fab5562	Robo4 ecto Fab5564	Robo4 ecto Fab5570	Robo4 ecto Fab5585
R_g	4.53 nm ± 0.02	5.14 nm ± 0.06	5.43 nm ± 0.02	5.11 nm ± 0.01	5.11 nm ± 0.02	5.5 nm ± 0.09
GNOM R_g	4.76 nm	5.46 nm	5.7 nm	5.33 nm	5.32 nm	5.43 nm
D_{max}	16.5 nm	19 nm	18.7 nm	17.4 nm	18.3 nm	19.11 nm
V_p (Porod exp)	107.84 nm ³ (3.1)	211.91 nm ³ (3.2)	207.74 nm ³ (2.8)	201.74 nm ³ (2.8)	209.46 nm ³ (2.5)	172.76 nm ³ (2.5)
V_c	571.6 nm ²	815.9 nm ²	810.9 nm ²	784.9 nm ²	792.3 nm ²	631.9 nm ²

The SAXS statistics of each dataset are reported in Table 11. The radius of gyration (R_g) was calculated from the Guinier approximation. Formation of the complex can be clearly followed by the change in R_g , which is consistently higher for all complexes compared to Robo4 ecto alone (4.53 nm), and is indicative of a change in the distribution of the particle mass. Except for the Robo4 ecto/Fab5562 and the Robo4 ecto/Fab5585 complexes, which were measured to have an R_g of 5.43 nm and 5.5 nm respectively, the other three Fab complexes displayed similar values of ~5.12 nm. The GNOM R_g (the radius of gyration calculated from the pair distribution function $p(r)$) is also reported in Table 11 and is consistently larger in the Robo4 ecto/Fab complexes datasets. The Porod volumes (V_p) are also reported, however, they can only be considered reliable when the Porod exponent is <3. Furthermore, the Porod volume depends on a parameter that matches the area

RESULTS

under the peak in the Kratky plot, which, in this case, is not always a clearly finite surface (see Figure 24B and Figure 25I to M). In these cases, the correlated volume (V_c) is a more reliable estimate and is reported. Despite some discrepancy, most notably the Robo4 ecto/Fab5585 complex, the increase in volume is in agreement with the formation of a 1:1 dimeric complex.

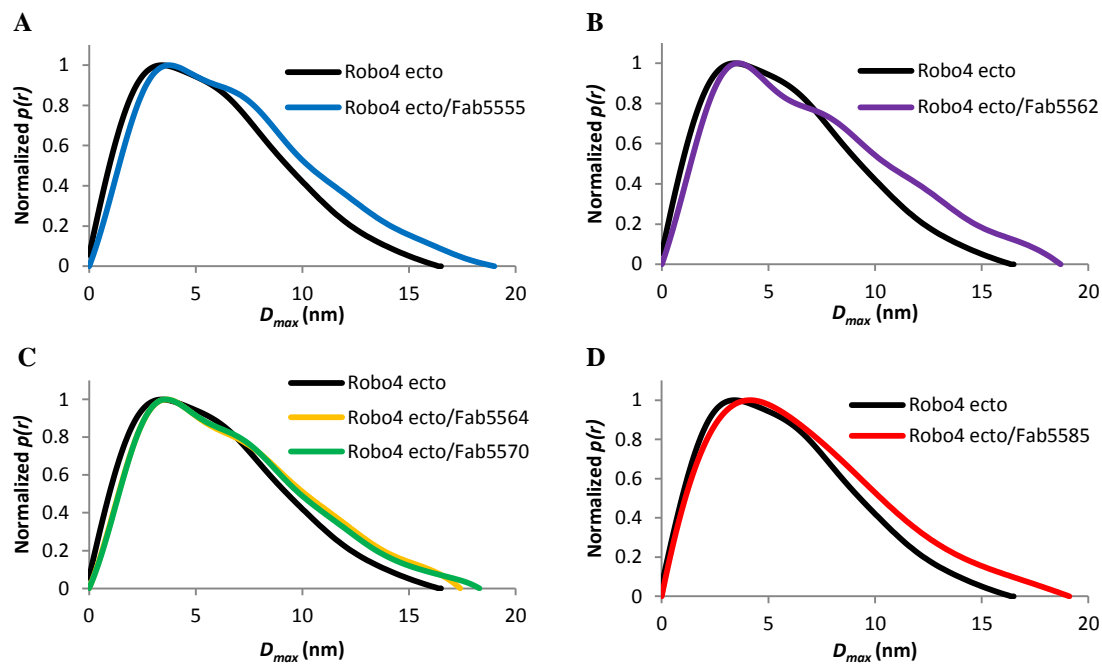


Figure 26: Pair distribution function of Robo4 ecto and Fab complexes

The $p(r)$ of Robo4 ecto alone in black is shown in all figures for comparison.

A: Robo4 ecto/Fab5555 (blue).

B: Robo4 ecto/Fab5562 (purple).

C: Robo4 ecto/Fab5564 (yellow) and Robo4 ecto/Fab5570 (green).

D: Robo4 ecto/Fab5585 (red).

Figure 26 shows the $p(r)$ of each Robo4 ecto/Fab complex compared to Robo4 ecto alone. These functions reveal an elongation in one direction that is characteristic of a prolate particle, with a maximum dimension (D_{max}) of 16.5 nm for Robo4 ecto, and a slight increase for all other complexes to ~18-19 nm (Table 11). With the exception of Robo4 ecto/Fab5585 (Figure 26D), all other complexes' curves show an additional shoulder, indicating a transition to a dumbbell shaped particle. The Robo4 ecto/Fab5564 and Robo4 ecto/Fab5570 complexes have similar shapes (Figure 26C), which might be expected since their interacting regions are both located within the Ig2 domain of Robo4 ecto (see Table 8, page 68). However, the Fab5555 and Fab5562 complexes, both interacting with the Ig1 domain (Table 8,

page 68), are not that similar (Figure 26A and B). In this case, the $p(r)$ of the second peak in Robo4 ecto/Fab5562 is lower, and more pronounced, at 8 nm.

Considering the quality of the data, *ab initio* models were built for Robo4 ecto alone, Robo4 ecto/Fab5562 complex, and Robo4ecto/Fab5564 complex. The other complexes were excluded either because the data was not suitable (as in the case of Robo4 ecto/Fab5585), or because it was deemed that no additional information could be extracted from them considering the high flexibility and similarity observed.

The normalized spatial discrepancy (NSD) calculated by SUPCOMB (part of the DAMAVER pipeline) between all models was 0.96 to 1.2, where good NSD should be lower than 1 and ideally lower than 0.7. The obtained NSD values reported here indicate that the models should be considered with care.

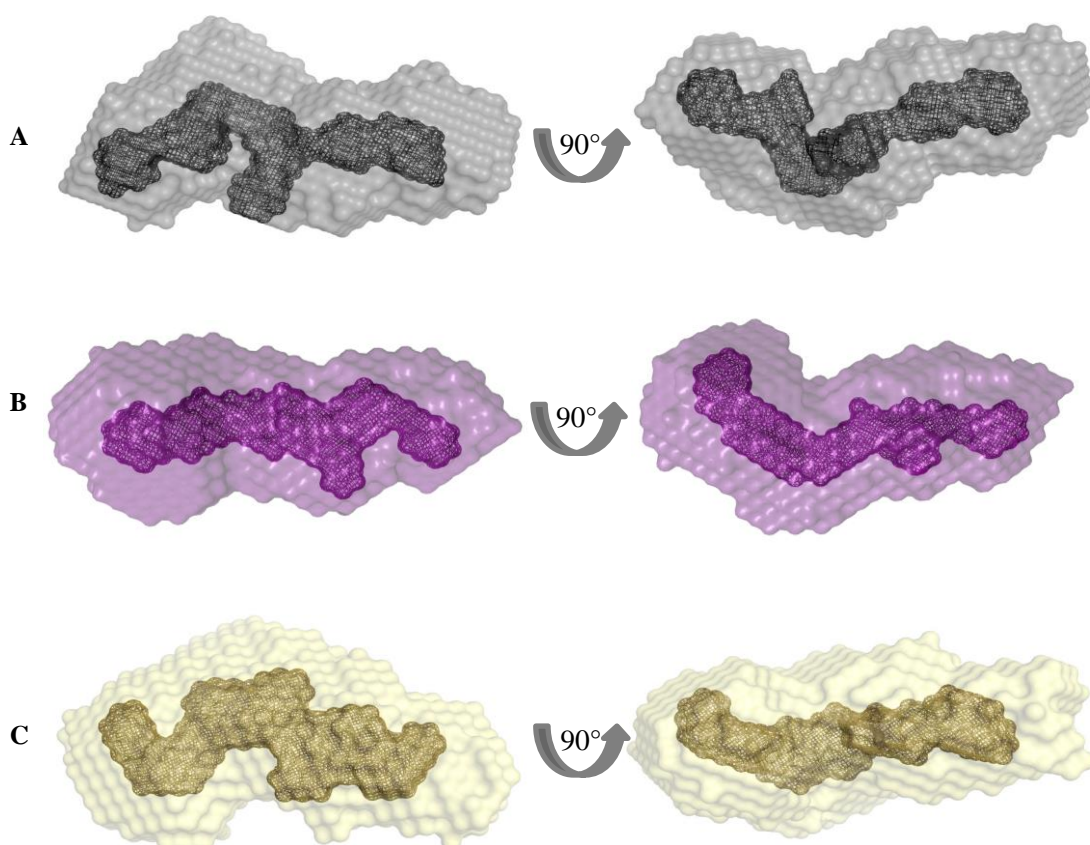


Figure 27: SAXS bead models of Robo4 ecto, Robo4 ecto/Fab5555 and Robo4 ecto/Fab5564

A: Robo4 ecto. **B:** Robo4 ecto/Fab5562. **C:** Robo4 ecto/Fab5564.

The models are shown as meshes, the transparent envelopes surrounding them shows the total volume occupied by all possible models. Even between the Robo4 ecto model alone and the models of the Robo4 ecto in complex with one of the Fabs, there is no significant difference to be observed, despite the data confirming complex formation.

RESULTS

As expected from the $p(r)$ functions (Figure 26A to D), and the calculated D_{max} (Table 11), all models presented an elongated shape with some protrusions (Figure 27). Half of the density can be manually filled by rigid body modelling the existing crystal structure of the Ig1-2 domains of Robo1 (Figure 28). Considering that FnIII domains are roughly the same size of Ig domains, and would occupy a similar volume at this resolution, it can be argued that the models, although only approximate, are a correct estimation of the state of Robo4 ecto in solution. However, it was not possible to assign a direction, or identify the position of any single domains so no additional conclusions can be determined with the overall data quality achieved.

The Robo4 ecto/Fab5562 and Robo4 ecto/Fab5564 complexes did not present any significant protrusion, or other distinguishable features, when compared to Robo4 ecto alone. The flexibility of the protein (or Fab complexes) does not allow any significant feature to be unambiguously distinguished. Even when superimposing the model of each complex with the model of Robo4 ecto, no additional density corresponding to a Fab is visible.

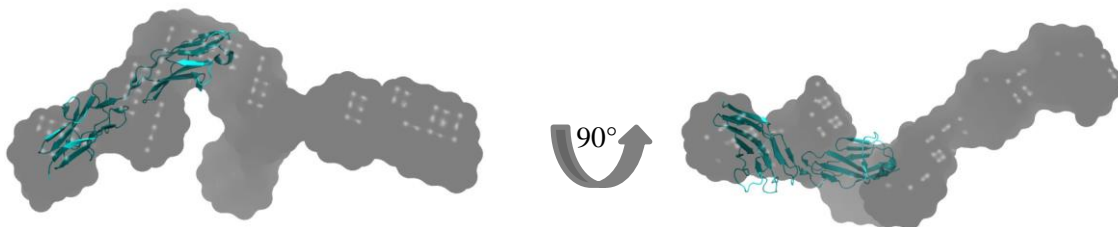


Figure 28: Robo4 ecto SAXS bead model superimposition to Robo1 Ig1-2 structure

The structure was manually superimposed to the Robo4 ecto bead model. The dimensions are comparable with the presence of domains of similar size.

In order to build a better Robo4 ectodomain model, the EOM approach was also employed to generate, and select, an ensemble of possible conformations that could fit the data. Since this approach requires prior knowledge of the structure of the domains involved, or of an approximation in the form of homology models, SWISS-MODEL was used to create homology models to provide as input. Unfortunately, the quality of the models was not very high, as can be observed from the statistics provided from the SWISS-MODEL calculations (Figure 29A to D). While the local quality of the Ig1-Ig2 model (Figure 29A), and the overall quality compared to a set of real crystal structures (Figure 29B) are acceptable (although

poor), in the case of the first and second FnIII model, the local quality (Figure 29C), and the model comparison (Figure 29D) are considerably worse. As a result of either the poor homology model quality, the complexity of the system, or a combination of these factors, this approach did not generate a curve with a satisfactory fit to the data (Figure 29E).

RESULTS

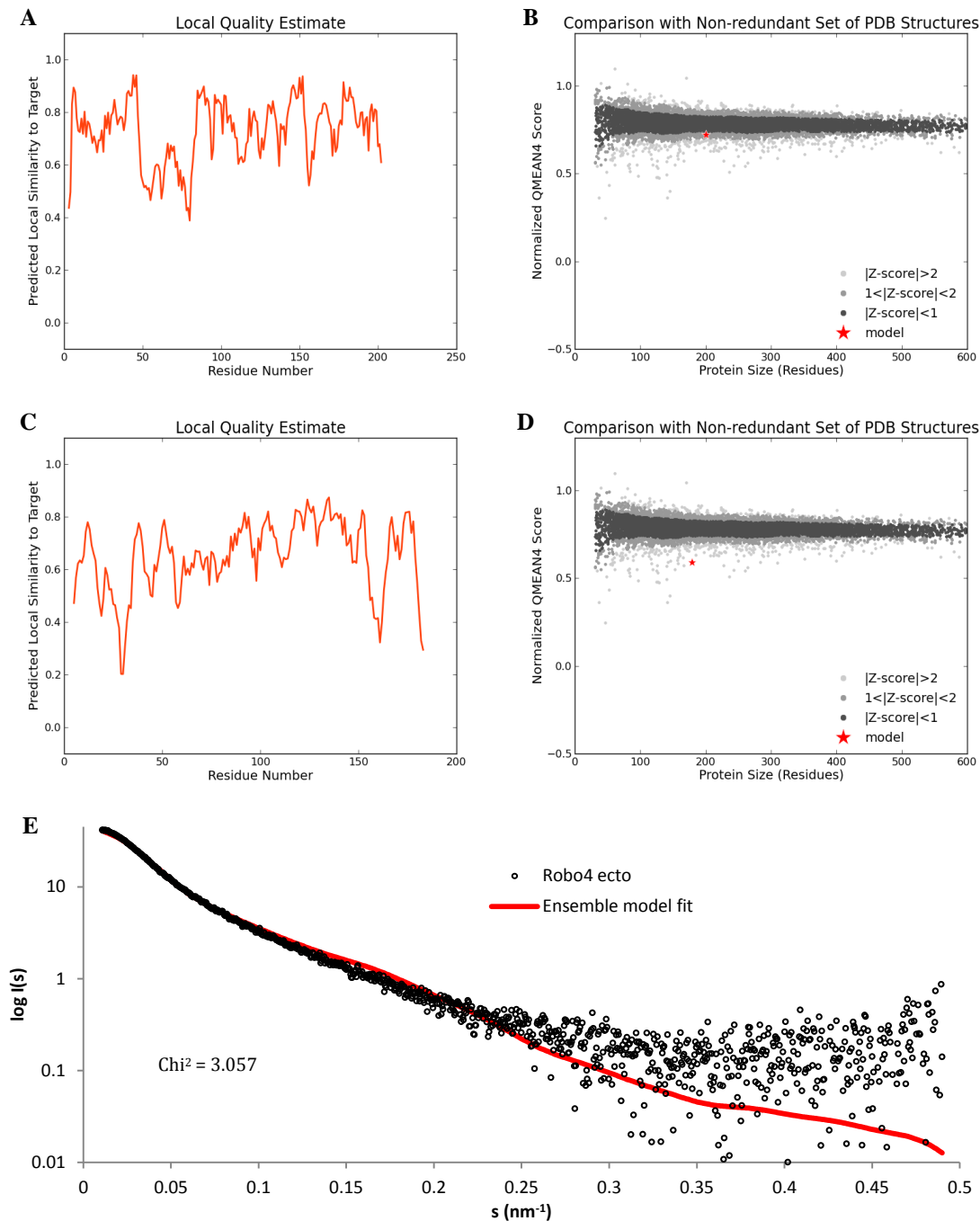


Figure 29: Homology modelling and EOM fit

The “local quality estimate plots” show the predicted similarity of each residue of the model to the native structure. A score below 0.6 is considered low quality.

The “comparison with non-redundant set of PDB structures” plots shows model quality scores as Z-scores in comparison to scores obtained for high-resolution structures. Every dot is one protein structure. The red star represents the model.

A and B: local quality and comparison plot of the Ig1-Ig2 model respectively.

C and D: local quality and comparison plot of the two FnIII model respectively.

E: EOM calculation plot. Black: Robo4 ecto experimental scattering curve. Red: Calculated scattering curve of ensemble model fit.

3.7 Crystallisation of Robo4 ecto alone and in complex with Fabs

Several crystallisation trials of the Robo4 ectodomain alone (fully glycosylated, reduced glycosylation, or deglycosylated), and in complex with the interacting Fabs were performed. Details on sample preparation are in § 2.16. Conditions that produced crystal growth over a period of 2 to 4 weeks are listed in Table 12. The crystals either showed very weak, or no diffraction. Attempts to reproduce and optimize the crystallisation conditions did not improve the crystal size or the quality of diffraction and are, therefore, not reported (Figure 30).

Table 12: List of Robo4 ecto/Fab crystallisation conditions

	Buffer	Salt	Precipitant	Additional component
Robo4 ecto Fab5562	-	0.17 M ammonium sulphate	25.5% (w/v) PEG4000	15% (v/v) Glycerol
Robo4 ecto Fab5562	0.1 M Bis-Tris pH 5.5	0.2 M calcium chloride	45% (v/v) MPD	-
Robo4 ecto Fab5562	0.1 M sodium acetate pH 5	0.2 M calcium chloride	20% (w/v) PEG6000	-
Robo4 ecto Fab5564	-	0.17 M ammonium sulphate	25.5% (w/v) PEG4000	15% (v/v) glycerol
Robo4 ecto Fab5570	0.1 M Tris pH 8.5	0.2 M lithium sulphate	1.26 M ammonium sulphate	-

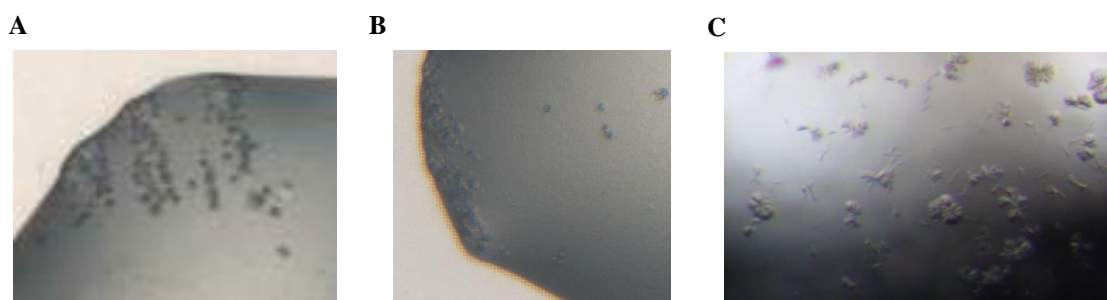


Figure 30: Robo4 ecto/Fab complex crystals

A: Robo4 ecto/Fab5562.

B: Robo4 ecto/Fab5564.

C: Robo4 ecto/Fab5570.

RESULTS

3.8 Expression and purification of human UNC5B ecto

Another objective of this work became the structural characterisation of UNC5B, a proposed binding partner of Robo4 (Koch et al., 2011). The interaction between Robo4 and UNC5B is still poorly understood at the molecular level, although their binding regions were mapped to the Ig domains of Robo4, and the TSP domains of UNC5B. A stable complex with a K_D of 12 nM was reported (Koch et al., 2011), suggesting a tight interaction that could improve the chance of crystallisation. UNC5B contains two predicted glycosylation sites (Table 13), therefore, as described before (§ 3.1), an expression system that could provide the necessary PTM modifications was necessary.

Table 13: Predicted UNC5B glycosylation sites

Sites of N-linked glycosylation as predicted by NetNGlyc.

Amino acid	Type	Confidence
222	N-linked	0.63
347	N-linked	0.42

In this case, the mammalian cell and the insect cell expression systems were used and the complete UNC5B ectodomain (amino acids 27-377, referred to as UNC5B ecto from now on) encompassing the membrane proximal region was produced in both. Similar to mammalian expression, the UNC5B ecto was secreted into the insect cell media by the use of the honey bee Melittin signal sequence added to the construct. Compared to mammalian cell media, the insect cell media has a lower overall protein content, and strep-tag affinity purification was sufficient to obtain a pure sample suitable for crystallisation studies after a single purification step.

Gel filtration was required, but not necessary, to remove the small amount of aggregated protein present in the purified sample and UNC5B ecto eluted as a peak at 13 ml of retention volume in a Superdex 200 Increase 10/300 column (Figure 31A). Similar to Robo4 ecto (§ 3.1), this retention volume would be expected for a much larger protein of ~150 kDa, while UNC5B ecto has an expected molecular weight of 40.1 kDa based on its sequence. The final product was more than 95% pure as assessed by SDS-PAGE (Figure 31B) and it migrated at an apparent molecular weight of ~50 kDa, without any visible co-eluting protein.

The same construct was also produced in mammalian cells with a histidine tag for purification, following the same general protocol as for Robo4 ecto (§ 2.7), but gave a significantly lower yield. During gel filtration on a HiLoad Superdex 200 Increase column, UNC5B ecto eluted homogeneously at an elution volume of 72.4 ml (Figure 31C), which is, again, characteristic of a protein with a molecular weight in the 120-150 kDa range. SDS-PAGE analysis confirmed the absence of co-eluting proteins (Figure 31D).

Due to the large difference in yield between the two systems, all following experiments were performed with the insect cell expressed UNC5B extracellular domain, unless otherwise stated.

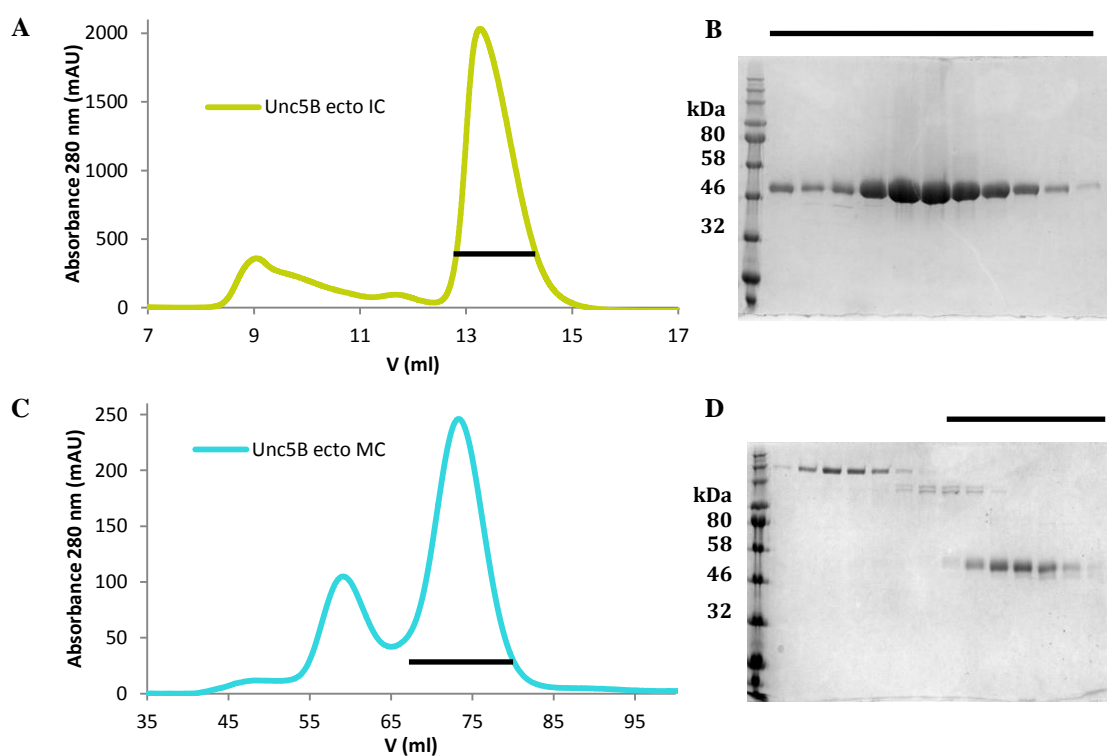


Figure 31: Purification of UNC5B ecto

A: Size exclusion chromatography on Superdex 200 Increase 10/300 of UNC5B ecto overexpressed and purified from insect cells. **B:** SDS-PAGE gel.

C: Size-exclusion chromatography on HiLoad Superdex 200 16/600 of UNC5B ecto overexpressed and purified from mammalian HEK293 cells. **D:** SDS-PAGE gel.

The black line indicates the fractions corresponding to the relevant peak loaded on SDS-PAGE gel.

RESULTS

3.9 UNC5B ecto deglycosylation

To improve the chance of crystallisation and to investigate the extent of heterogeneous glycosylation on UNC5B, deglycosylation of UNC5B ecto was tested with a similar protocol to that used for Robo4 ecto (§ 3.2). In this case, deglycosylation of UNC5B ecto overexpressed and purified from insect cells did not seem to be affected, apart from a very minor shift in the apparent migration on a SDS-PAGE gel after incubation with PNGase (Figure 32A).

UNC5B overexpressed and purified from HEK293 cells showed the presence of a faint lower molecular weight band with all three enzymes. PNGase F1 was the most active, and seemed to complete its action after 1 hour of incubation time. After overnight incubation, the same low molecular band appears slightly more abundant (Figure 32B).

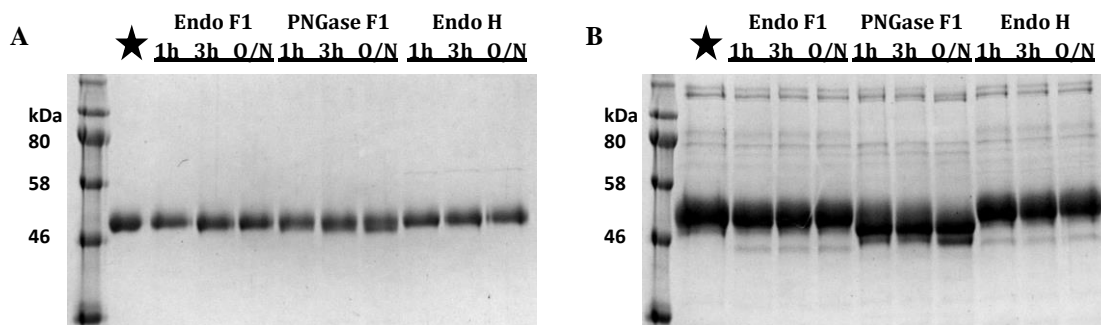


Figure 32: Deglycosylation of UNC5B ecto

Deglycosylation was performed in the protein purification buffer. For each enzyme depicted a sample was taken after 1 hour, 3 hours and overnight incubation. The star indicates time 0 of each sample.

A: UNC5B ecto overexpressed and purified from insect cells.

B: UNC5B ecto overexpressed and purified from HEK293 cells.

3.10 Expression and purification of UNC5B TSP domains

The TSP domains of the TSP-1 protein have previously been produced by *E. coli* expression (Klenotic et al., 2011). Since the UNC5B interaction site with Robo4 is located in its TSP domains (Klenotic et al., 2011), it was attempted to individually produce a construct incorporating this region in *E. coli*. Two different solubility tags were tested: a His-GST (Figure 33A) and MBP fusion tag (Figure 33C). The highest level of expression was observed in presence of the MBP tag, but size exclusion chromatography showed that both tag expressed products were eluting in the void volume at ~8 ml (Figure 33B and D). Dynamic light scattering analysis confirmed the presence of soluble aggregates. Several attempts to overcome this, such as expression temperature, different *E. coli* strains and lysis buffer screening were unsuccessful, so this approach was eventually discarded.

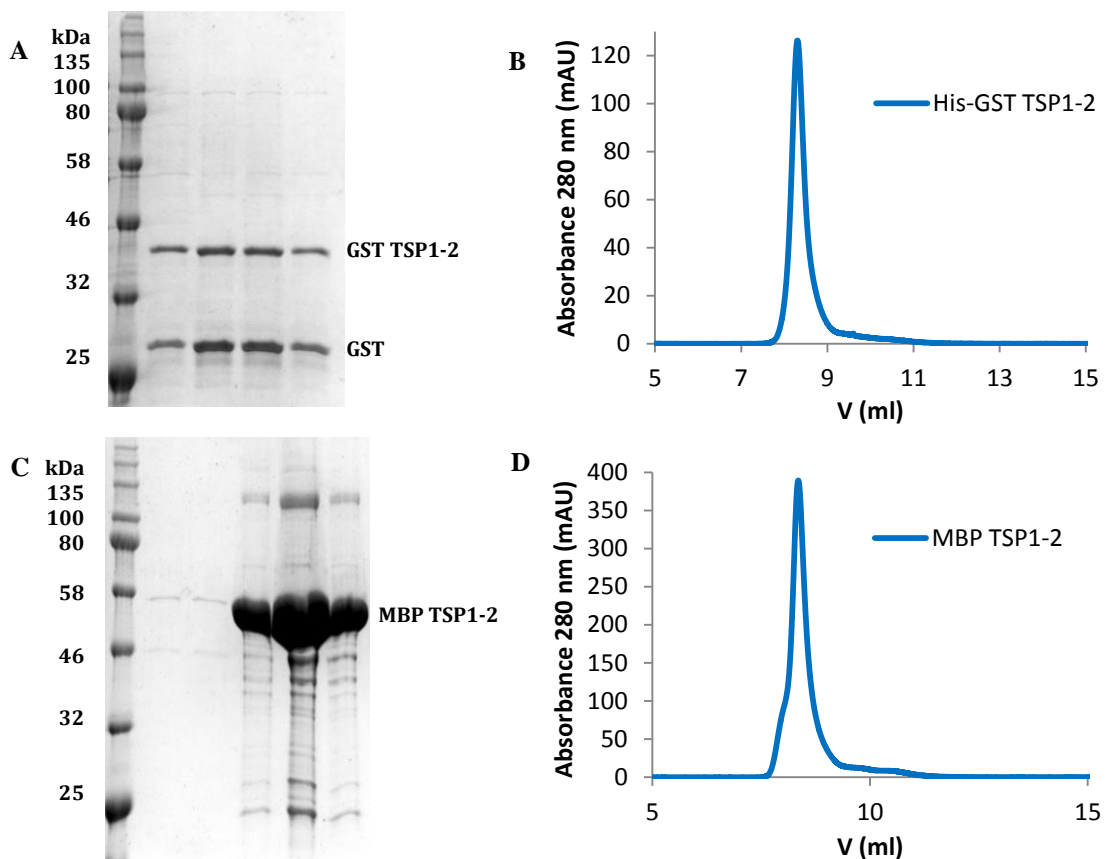


Figure 33: Purification and size exclusion chromatography of UNC5B TSP1-2

A: GST affinity purification of TSP1-2, cleaved GST is visible at the bottom.

B: Size exclusion chromatography on Superdex 75 column. All of the expressed protein eluted in the void volume at 8.3 ml.

C: MBP affinity purification of TSP1-2.

D: Size exclusion chromatography on Superdex 75 column. All of the expressed protein eluted in the void volume at 8.3 ml.

RESULTS

3.11 Binding studies of Robo4 and UNC5B extracellular domains

As mentioned earlier, an interaction between UNC5B and Robo4 was previously reported (Koch et al., 2011). To verify this, several experiments were carried out to probe the association of full-length glycosylated Robo4 and UNC5B extracellular domains, as these are the most biologically relevant forms.

Pull-down essays were first performed using the Strep tag present on UNC5B. A pull-down using the histidine tag of Robo4 ecto was also performed, but since the nickel resin presents unspecific binding, these data are not shown. No association using either pull down tag was observed (Figure 34).

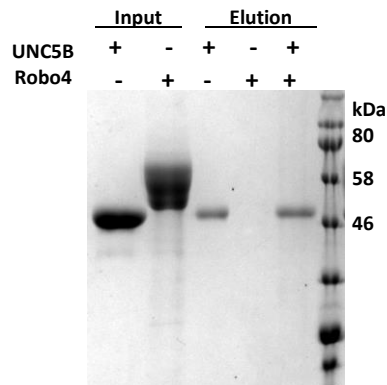


Figure 34: Strep pull-down of UNC5B and Robo4 extracellular domains

Lane 1 and 2: input proteins, UNC5B and Robo4 respectively. Lane 3: elution of Strep beads incubated with UNC5B ecto alone. Lane 4 elution of Strep beads incubated with Robo4 ecto alone. Lane 5: elution of Strep beads after incubation with preincubated UNC5B ecto/Robo4 ecto.

SEC-MALS analysis was also used to verify their interaction. The experiment depicted in Figure 35A was performed on a Superdex 200 10/300 column using UNC5B ecto overexpressed and purified from insect cells and Robo4 ecto. The experiment in Figure 35B was instead performed using a Superdex 200 5/150 column and UNC5B ecto overexpressed and purified from mammalian cells (UNC5B ecto MC). Observed molecular weights are reported in Table 14.

Based on its sequence, the predicted molecular weight of UNC5B ecto is 40.2 kDa. The measured molecular weight of UNC5B ecto overexpressed in insect cells (UNC5B ecto IC) was 47.4 kDa (Figure 35A), while that in mammalian cells was 49.7 kDa (Figure 35B). Considering the previously measured molecular weight of Robo4 ecto of 56.1 kDa (§ 3.3, Table 7), a shift of the elution volume peak, and an almost doubling of the measured weight was expected to be observed upon

complex formation. However, this was not observed using UNC5B ecto sample from either expression system (Figure 35, Table 14). In both cases, the observed mass was only roughly equivalent to the higher molecular mass species of the larger Robo4 ecto coupled to a concomitant broadening of the elution peak. This indicates a lack of interaction, and is consistent with the pull down assays above.

Table 14: SEC-MALS analysis of UNC5B ecto binding to Robo4 ecto

Listed in the table are the predicted molecular weights based on UNC5B ecto and the UNC5B ecto/Robo4 ecto complex sequences, the observed molecular weight derived from MALS measurements and the polydispersity of each sample, with errors.

	UNC5B ecto IC	UNC5B ecto IC Robo4 ecto	UNC5B ecto MC	UNC5B ecto MC Robo4 ecto
Predicted MW	40.2 kDa	88.2 kDa	40.2 kDa	88.2 kDa
Observed MW	47.42 kDa ± 1.5 %	57.69 kDa ± 1.2 %	49.72 kDa ± 1.8 %	57.15 kDa ± 1.3 %
Polydispersity	1.000 ± 2.2 %	1.001 ± 1.7 %	1.002 ± 2.5 %	1.003 ± 1.9 %

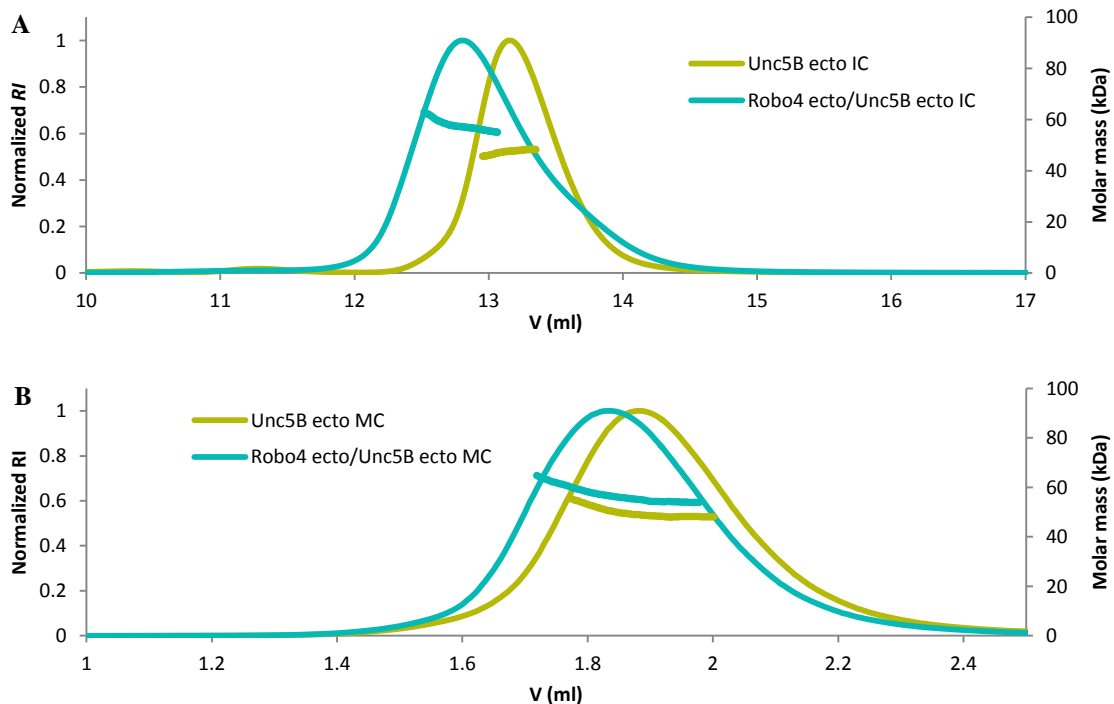


Figure 35: SEC-MALS analysis of UNC5B ecto and after incubation with Robo4 ecto

A: UNC5B ecto overexpressed and purified from insect cells was injected alone (dark yellow) and after incubation with Robo4 ecto (teal) on a Superdex 200 Increase 10/300 column. Addition of the Robo4 showed a skewed peak in size exclusion chromatography, but no change in overall mass.

B: UNC5B ecto overexpressed and purified from HEK293 was injected alone (dark yellow) and after incubation with Robo4 ecto (teal) on a Superdex 200 5/150 column. No significant change was observed in the peak or the overall mass.

RESULTS

To further probe their interaction, SPR experiments were carried out in order to duplicate the previously published result ((Koch et al., 2011) and § 2.14.3). Robo4 ecto and UNC5B ecto were immobilized on two different flow cells, and injected serially with 50 and 1000 nM of each protein. UNC5B ecto was immobilised in 10 mM sodium acetate buffer at pH 4.5 (Figure 36). For Robo4 ecto immobilisation see § 3.5.1 or § 2.14.1.

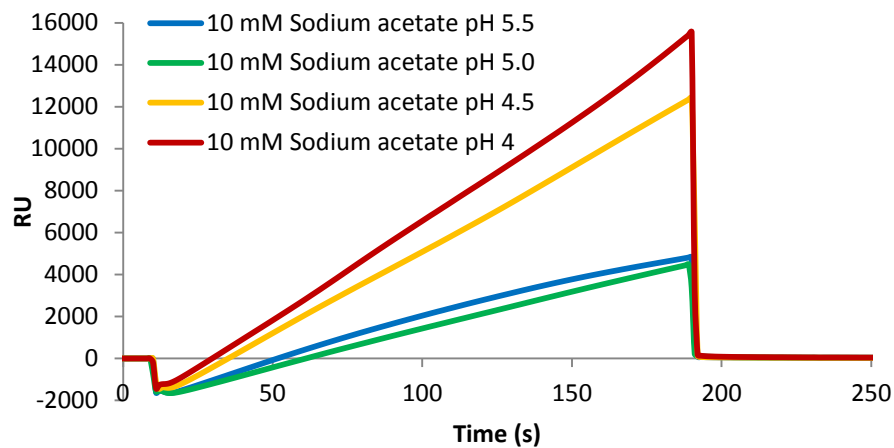


Figure 36: pH scouting of UNC5B ecto

10 mM Sodium acetate at pH 5.5 (blue), pH 5 (green), pH 4.5 (orange) and pH 4 (red) was used to scout the best immobilization strategy. pH 4.5 was chosen for the immobilisation strategy as the highest pH allowing for sufficient immobilisation.

No association was observed when injecting UNC5B ecto on immobilized Robo4 ecto (Figure 37A), or when injecting Robo4 ecto on immobilized UNC5B ecto (Figure 37B), contradictory to previous reports. It was interesting to observe an apparent weak homophilic binding of Robo4 ecto (Figure 37A), which was not observed in SEC-MALS or SAXS (§§ 3.3 and 3.6), but this was not investigated further. A higher concentration of Tween20 detergent was used compared to the published studies (0.05% v/v instead of 0.005% v/v), but this is unlikely to have negatively influenced the experiment. This again confirms a lack of interaction between the UNC5B and Robo4 extracellular domains.

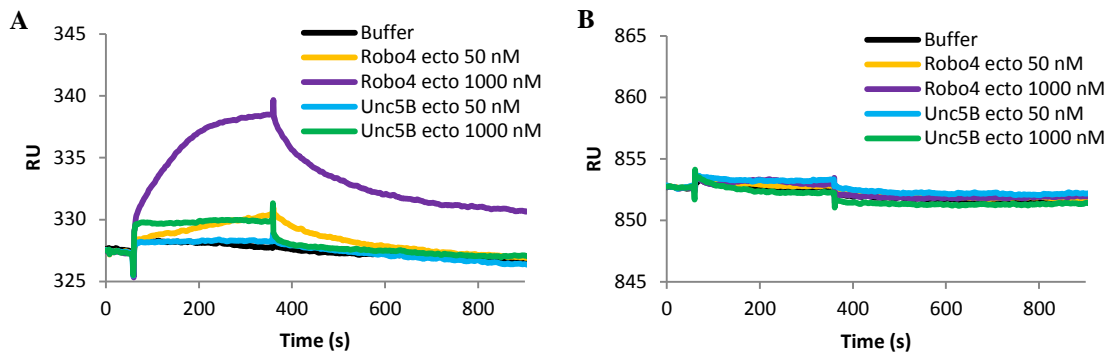


Figure 37: SPR binding test of Robo4 ecto and UNC5B ecto

Soluble Robo4 ecto and UNC5B were injected on the chip at 50 and 1000 nM concentration. The two sensorgrams show the response obtained from immobilized Robo4 ecto and UNC5B ecto as described. The baseline was manually subtracted to compensate for baseline drift between regeneration cycles.

A: Chip with immobilized Robo4 ecto. Homophilic binding and dissociation can be observed with an increased amount of injected Robo4 ecto.

B: Chip immobilized with UNC5B ecto. No binding event observed.

As an additional test, the AVEXIS approach was used by the group of Rob Meijers at EMBL Hamburg (Bushell et al., 2008). Here, the Robo4 ecto construct used in this study was cloned into the bait and prey vectors, and sent to Hamburg, where the experiment was performed by Xuefan Gao following established protocols (Kerr and Wright, 2012).

These results are presented in Figure 38 (courtesy of Xuefan Gao, EMBL Hamburg). Again no evidence of binding with the UNC5B ectodomain was observed. Netrin-1 binding to bait UNC5B and DCC, and binding of Robo4 ecto to bait Matrilin-1 (Matn) serve as positive controls. Only when using a 20 times excess of Robo4 ecto, was a very weak signal observed (Figure 38B).

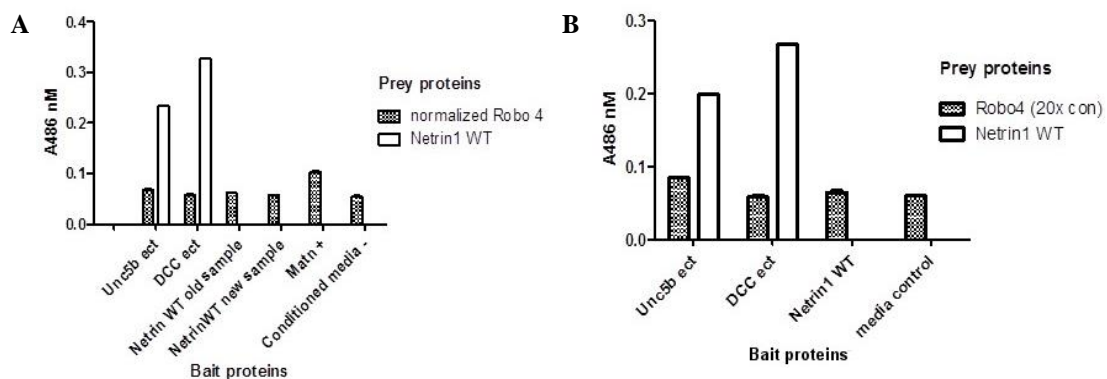


Figure 38: Avexis binding test of Robo4 and UNC5B extracellular domains

Avexis binding tests, images courtesy of Xuefan Gao (EMBL Hamburg).

A: Test under standard conditions.

B: 20x concentrated Robo4 ecto prey.

The UNC5B/Netrin and DCC/Netrin interactions are used as positive controls. Matn is a general positive control for prey proteins, having a characteristic slow interaction with low response. Conditioned cell media served as negative control and is indicative of the background signal level.

RESULTS

3.12 Crystal structure of UNC5B extracellular domain

3.12.1 Crystallisation and structure solution

Crystals of UNC5B ecto were obtained in several crystallisation conditions as micro crystals. The best diffracting crystals grew after application of the seeding method, and further optimization. Some examples are shown in Figure 39, and a full description can be found in § 2.16.

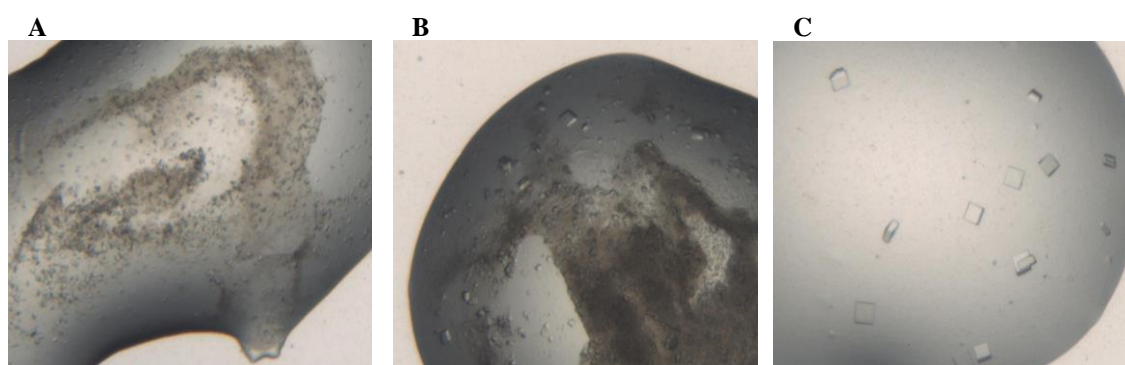


Figure 39: UNC5B ecto crystals optimization

A: Micro crystals obtained in initial screening after 17 days of growth in 0.1 M sodium acetate pH 5, 0.2 M ammonium chloride, 20% (w/v) PEG6000.

B: Same condition as before with the addition of undiluted seed stock, crystals of larger size could grow in 4 days.

C: After optimization and seeding, large single crystals after 7 days of growth in 0.1 M SPG buffer pH 7, 19% (w/v) PEG1500.

Initially, parallelepiped microcrystals of UNC5B ecto could be obtained from the conditions reported in Table 15 after a growth period of >3-4 weeks. After optimization by rMMS several more conditions were identified, resulting in crystals of the same shape that grew larger, faster (from 4 days to 2 weeks), and had better diffraction qualities. These conditions are reported in Table 16.

The crystals were very fragile and manual manipulation of the crystals generally lowered the resolution achieved, or resulted in no visible diffraction. The best datasets were obtained from crystals automatically harvested using the CrystalDirect™ robot of the HTX facility, without addition of cryoprotecting agents (Pellegrini et al., 2011).

Table 15: Initial crystallisation conditions of UNC5B ecto

Buffer	Salt	Precipitant
-	0.2 M ammonium citrate	20% (w/v) PEG3350
0.1 M Tris pH 8.5	0.2 M ammonium phosphate	50% (v/v) MPD
0.1 M sodium acetate pH 4.6	-	25% (w/v) PEG3000
0.1 M sodium acetate pH 5	0.2 M magnesium chloride	20% (w/v) PEG6000
0.1 M sodium acetate pH 5	-	20% (w/v) PEG6000
0.1 M sodium acetate pH 5	0.2 M calcium chloride	20% (w/v) PEG6000
-	0.2 M ammonium sulphate	20% (w/v) PEG3350

Table 16: Best crystallisation conditions of UNC5B ecto in presence of seeds

Buffer	Salt	Precipitant
0.1 M Sodium acetate pH 4.5-5.5	-	10-20% (w/v) PEG3000
-	0-0.2 M ammonium citrate	10-20% (w/v) PEG3350
SPG pH 7	-	19-29% (w/v) PEG1500
MIB pH 7	-	19-29% (w/v) PEG1500
MMT pH 7-8	-	19-29% (w/v) PEG1500
HEPES pH 7	0-0.2 M lithium chloride	14-24% (w/v) PEG6000
0.1 M Tris pH 8	0-0.1 M calcium chloride	14-24% (w/v) PEG6000
-	-	14-24% (w/v) PEG3350
-	0.2 M sodium bromide	14-24% (w/v) PEG3350
-	0.2 M sodium iodide	14-24% (w/v) PEG3350
-	0.2 M potassium thiocyanate	14-24% (w/v) PEG3350
-	0.2 M sodium nitrate	14-24% (w/v) PEG3350
-	0.2 M potassium sodium tartrate	14-24% (w/v) PEG3350
0-01 M BIS-Tris propane pH 7.5	0.02 M sodium potassium phosphate	14-24% (w/v) PEG3350
0.1 M sodium acetate pH 4.6-5	-	19-27% (w/v) PEG3000

Homologous structures of human UNC5A and rat UNC5D are present in the [PDB](#) ([PDB ID 4V2A](#) and [5FTT](#) chain A respectively, see alignment in Figure 40). UNC5B was solved with the molecular replacement method using an initial multi crystal dataset collected at 3.8 Å resolution (§ 2.17.2). For this the Ig1 and Ig2 domains of UNC5D were used as independent search models. The final structure was built and refined using a dataset collected from a single crystal that grew in 25% (v/v)

RESULTS

PEG1500 and 0.1 M SPG buffer pH7 over a period of 7 days at 293°K. For this the initial structure (at 3.8 Å) was positioned in the new high resolution (3.3 Å) dataset and refined to a final R_{work}/R_{free} of 0.27/0.33 respectively with good geometry. All crystallographic data is summarised in Table 17.

Table 17: Crystallographic table of UNC5B ecto single crystal dataset

Values in parenthesis are for the highest resolution shell.

Space group	I 4 ₁ 2 2
Cell dimensions	
a, b, c (Å)	70.45, 70.45, 396.17
α , β , γ (°)	90, 90, 90
Statistics	
Resolution (Å)	48.34–3.30 (3.69–3.30)
Unique reflections	7780 (2162)
Completeness (%)	96.4 (97.5)
I/ σ	6.81 (0.64)
CC (1/2)	99.9 (26.3)
<i>R</i> _{meas} (%)	17.7 (255.6)
<i>R</i> _{work} (%) / <i>R</i> _{free} (%)	0.27/0.33 (0.28/0.30)
Average B-factors	148.37
RMS deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.13
Ramachandran	
Favoured/Outliers (%)	84/4
Number of atoms	
Protein/Heterogen	1723/28

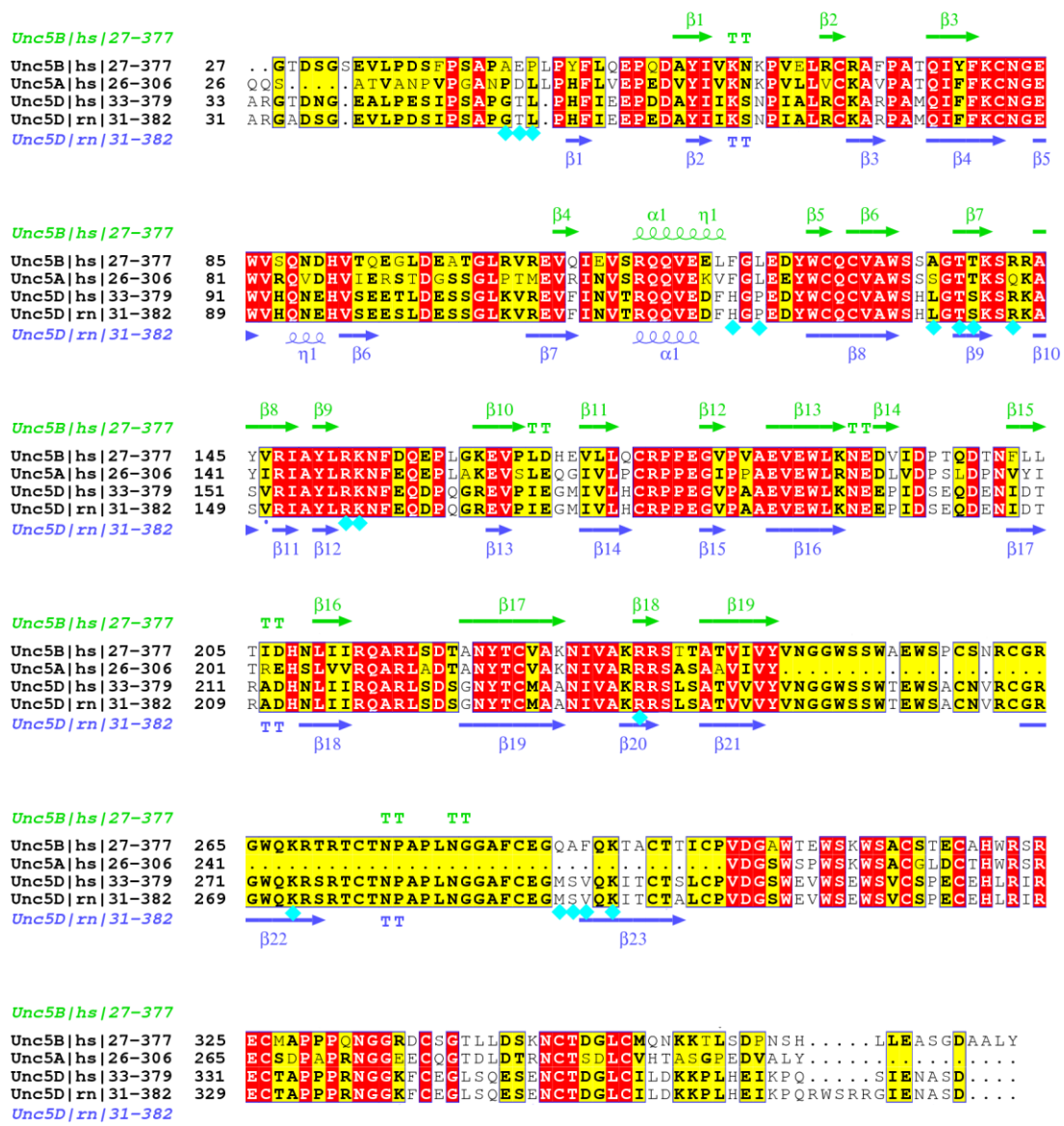


Figure 40: Multiple sequence alignment of known UNC5 structures and UNC5B ecto
 The extracellular domains of UNC5A and UNC5D were aligned to UNC5B. “hs” stands for *Homo sapiens*, “rn” for *Rattus norvegicus*. Secondary structural elements of UNC5B derived from the model solved in this thesis are at the top of the sequence (green), while secondary structure elements derived from the *Rattus norvegicus* UNC5D structure (PDB ID: 5FTT, chain A) are at the bottom (blue). Red boxes highlight identical residues. Yellow boxes contain similar residues. Bold highlighting indicates residues with similar chemical properties. The aqua squares indicate the UNC5D residues which are involved in interactions with Latrophilin-3, FLRT2 and the other UNC5D, in the octameric complex described in (Jackson et al., 2016).

RESULTS

3.12.2 UNC5B extracellular domain crystal structure

As expected of an extra cellular multidomain receptor, UNC5B ecto assumes an extended conformation (Figure 41). The two Ig domains, and first TSP domain, are fully resolved, except for two short loops between amino acids 89-92 and 98-103, for which no density could be observed. The second TSP domain is located in a large solvent region, where no discernible electron density could be used to position the domain. At the moment, there is also no structure known of the second TSP domain of UNC5D, while UNC5A only possesses a single TSP.

Electron density for the carbohydrate chain linked to Asn 222 is clearly visible (Figure 41). At this resolution, each carbohydrate residue, and their relative orientation, cannot be unambiguously determined. But the core carbohydrates are always composed of two NAG and one mannose in a linear chain, followed by two branched mannose residues (§ 1.5, Figure 9). The presence of discontinuous density further into the cavity, also at hydrogen bond distance from symmetry

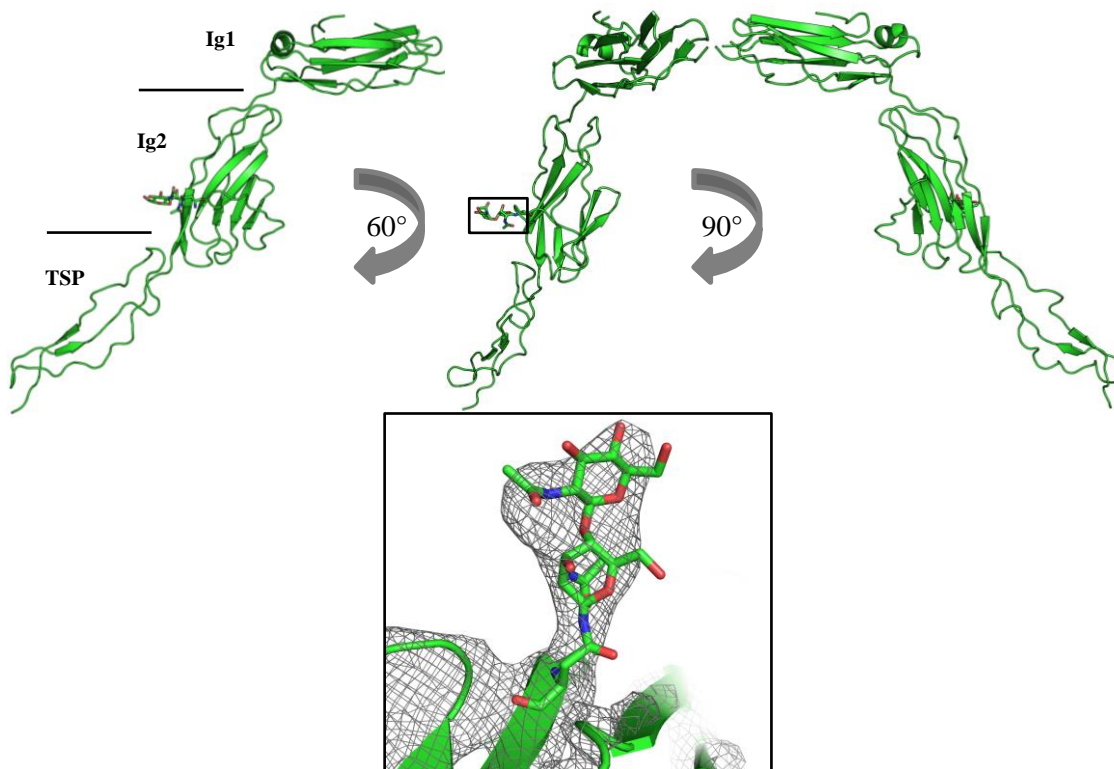


Figure 41: UNC5B ecto crystal structure and glycosylation density

The disconnected not modelled loop is visible in the Ig1 domain near the top of the left image. The modelled part of the carbohydrate chain is shown as a stick model along with the electron density maps in the black box. The first two NAG easily fill the density. The side chain of Asn222 is also shown to highlight the connection to the protein main chain. 2Fo-Fc density map contoured at 2σ .

molecules, is an indication that the carbohydrate chain probably continues further, but it was not possible to model any more at the current resolution.

Conservation analysis of UNC5B from multiple other species (§ 2.18), shows that UNC5B is quite evolutionary conserved. The Ig1 and Ig2 domains appear more conserved than the TSP1 domain, which is probably a result of the Ig domains being the main extracellular interaction site of UNC5B, and other Ig containing receptors, eg. Robo1. Still, there seems to be a more variable region on one specific side of the Ig domains, in particular for Ig1. The “internal” face (the part of the Ig domains facing the closed angle between Ig1 and Ig2) of the Ig domains is less conserved within respect to the “external” side.

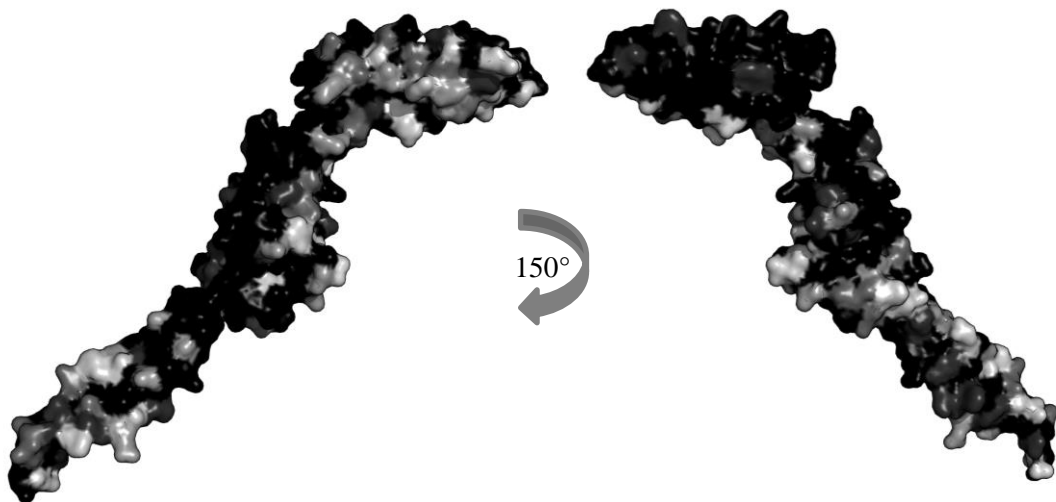


Figure 42: Estimated evolutionary conservation of UNC5B residues

From least conserved (light grey) to most conserved (black). Based on multiple sequence alignment between UNC5B homologues (§ 2.18).

RESULTS

3.12.3 Comparison of UNC5B ecto to existing UNC5 structures

A clear difference in the angles between each domain is visible when comparing the crystal structure of the extracellular domain of human UNC5B to those of human UNC5A (PDB ID: [4V2A](#)) and rat UNC5D (PDB ID: [5FTT](#), chain A) (Figure 43A to F). The online tool DynDom ([Protein Domain Motion Analysis](#)) was used to measure the angles between each domain (Hayward and Berendsen, 1998). Table 18 reports the root-mean square deviation (RMSD) values of the fixed domain atom positions against the position of the equivalent domain in UNC5A and UNC5B, together with the angles and suggested amino acids hinge regions.

Between UNC5B and UNC5D, the Ig1 and Ig2 domain angle differed by 19.4° (Figure 43C). While the Ig2 and TSP1 domain had a 30.5° angular difference (Figure 43D). Comparison of UNC5B and UNC5A revealed that the angular difference between Ig1 and Ig2 was 10.9° (Figure 43E), while the angle between Ig2 and TSP1 differed by 22.6° (Figure 43F).

Table 18: DynDom RMSD, angles and hinges between UNC5B, UNC5A and UNC5D domains

	UNC5B / UNC5A		UNC5B / UNC5D	
	Ig1/Ig2	Ig2/TSP1	Ig1/Ig2	Ig2/TSP1
RMSD	2.07 Å	0.92 Å	1.01 Å	0.99 Å
Rotation angle	10.9°	22.6°	19.4°	30.5°
Hinge region	V146-R147-I148	Y244-V245-N246	V146-R147-I148	V245-N246-G247

By comparing their structures, UNC5D and UNC5A also have an angular difference between their Ig1 and Ig2 domains (Figure 43A and B). It should be noted that the UNC5D structure used for these comparisons is part of a large octameric complex. Therefore the tight conformation between the Ig domains of UNC5D is likely induced, at least in part, by its participation in this complex (Jackson et al., 2016). Most notably, Latrophilin-3 (Lphn3) interacts with the loop between the Ig1 and Ig2 domains of UNC5D, and the α -helix of Ig1 (Figure 43A), which might contribute in a closing of the Ig1 domain upon Ig2.

Several UNC5D residues were identified to be important for complex formation (Jackson et al., 2016). Some of these are different in both UNC5B and UNC5A. For example, UNC5D his 125 and pro 127, which are involved in the UNC5D/Lphn3

interaction are located at the end of the Ig1 α -helix, and are respectively substituted by a Phe and a Leu in both UNC5B and UNC5A. Furthermore, Gly 49, Thr 50, Leu 140, Ser 143, Met 292, Ser 293 and Val 294 of UNC5D, which are all involved in homophilic interactions important for the overall stability of the complex, are different in UNC5B and UNC5A (see Figure 40, aqua squares).

RESULTS

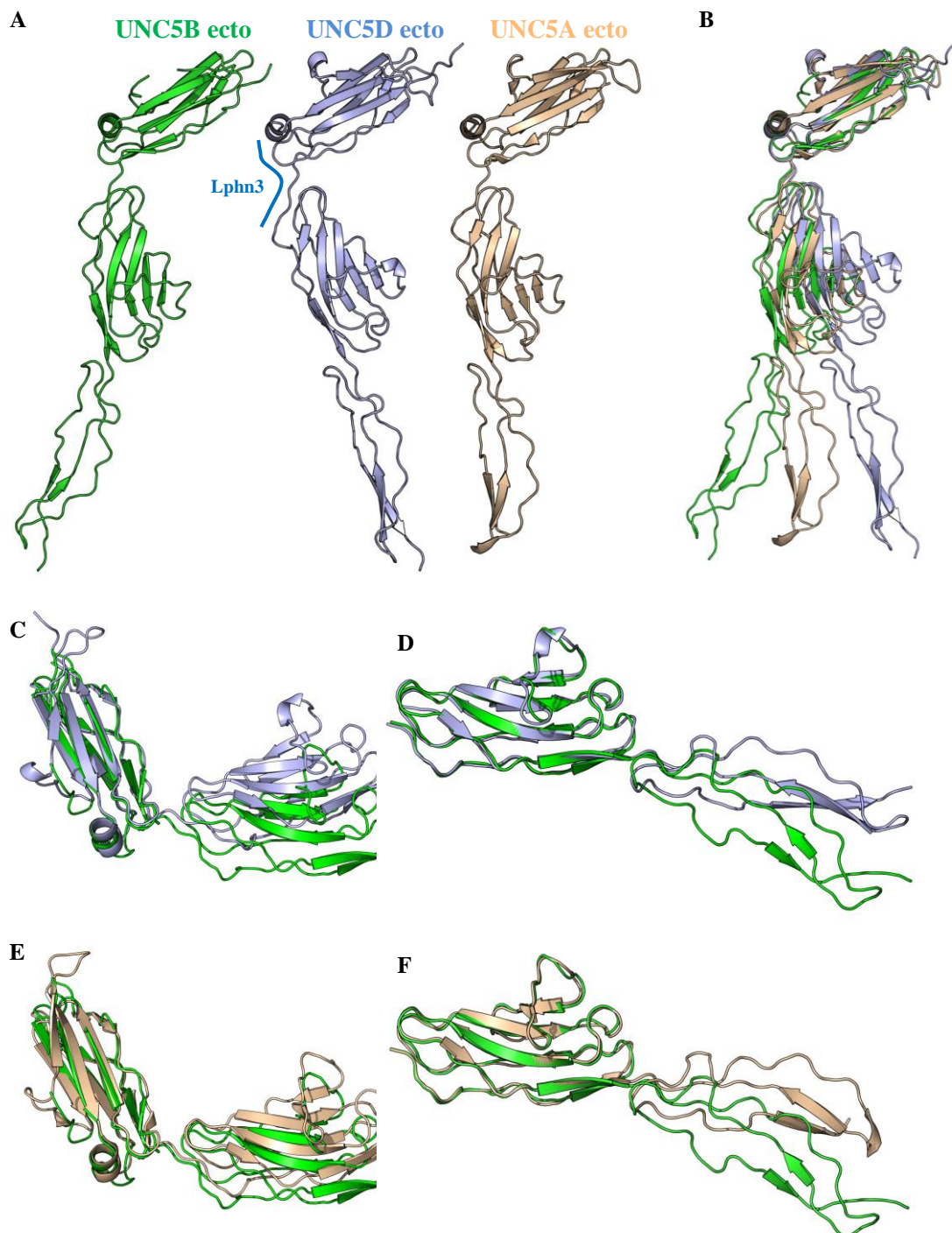


Figure 43: Superimposition of UNC5B with UNC5D and UNC5A structures

A: Human UNC5B (green), rat UNC5D (light blue) and human UNC5A (light brown) extracellular domains structures, aligned relatively to their Ig1 domain. The bright blue curve close to the UNC5D structure highlights the interacting interface between UNC5D and Lphn3.

B: superimposition of the three structures relative to their Ig1 domains.

C: superimposition of Ig1-Ig2 of UNC5B and UNC5D relative to Ig1.

D: superimposition of Ig2-TSP1 of UNC5B and UNC5D relative to Ig2.

E: superimposition of Ig1-Ig2 of UNC5B and UNC5A relative to Ig1.

F: superimposition of Ig2-TSP1 of UNC5B and Ig2-TSP of UNC5A relative to Ig2.

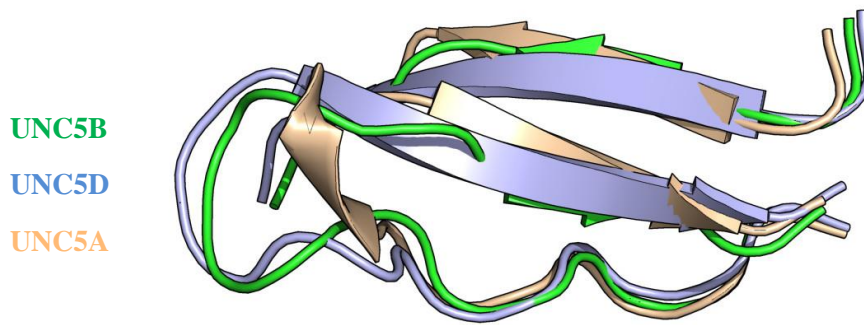


Figure 44: TSP loop of UNC5 proteins

Close up of superimposed C-terminal loop of the TSP domain of UNC5B (green), UNC5A (light brown) and UNC5D (light blue) were superimposed. The terminal loop of UNC5A appears to be formed by a short β -strand, while UNC5B and UNC5D are more extended.

Comparison between the TSP domain structures of UNC5B, UNC5D and UNC5A shows that the last loop of UNC5A (amino acids 253-260) is shorter (Figure 44). Local sequence alignments of the three UNC5 receptors, aligned the TSP1 domain of UNC5A with the TSP2 domains of UNC5B and UNC5D, disregarding their TSP1 domains (Figure 40). Forcing the alignment of the TSP domain of UNC5A with TSP1 of UNC5B (Figure 45A) reveals that the TSP domain of UNC5A is three amino acids shorter (black rectangle, Figure 45A), accounting for the difference in structure observed. Furthermore, the TSP2 domain of UNC5B is clearly more similar to the TSP of UNC5A (Figure 45B).

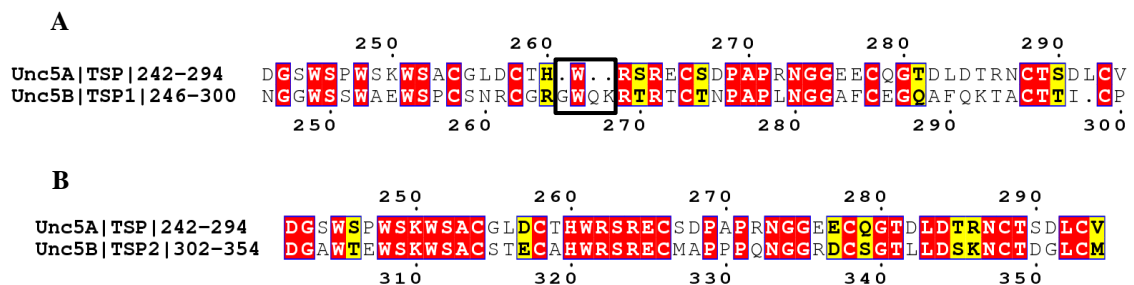


Figure 45: UNC5A and UNC5B TSP domains alignment

A: Alignment of UNC5A TSP domain with UNC5B TSP1 domain.

B: Alignment of UNC5A TSP domain with UNC5B TSP2 domain.

Red boxes highlight identical residues, yellow boxes contain 4 similar or less than 4 identical residues and bold highlighting indicates residues with similar chemical properties. The black rectangle highlights the position of the three missing amino acids close to the loop forming region.

RESULTS

The electrostatic potential surfaces between UNC5B, UNC5D and UNC5A are mostly similar (Figure 46). UNC5A shows a more pronounced positive patch on the external side of the Ig1-Ig2 domains, while UNC5D has a larger negative interface located on the internal side under the Ig1 and Ig2 domains. However, on UNC5D, these surfaces are not involved in any interaction within the octameric complex previously mentioned (Jackson et al., 2016).

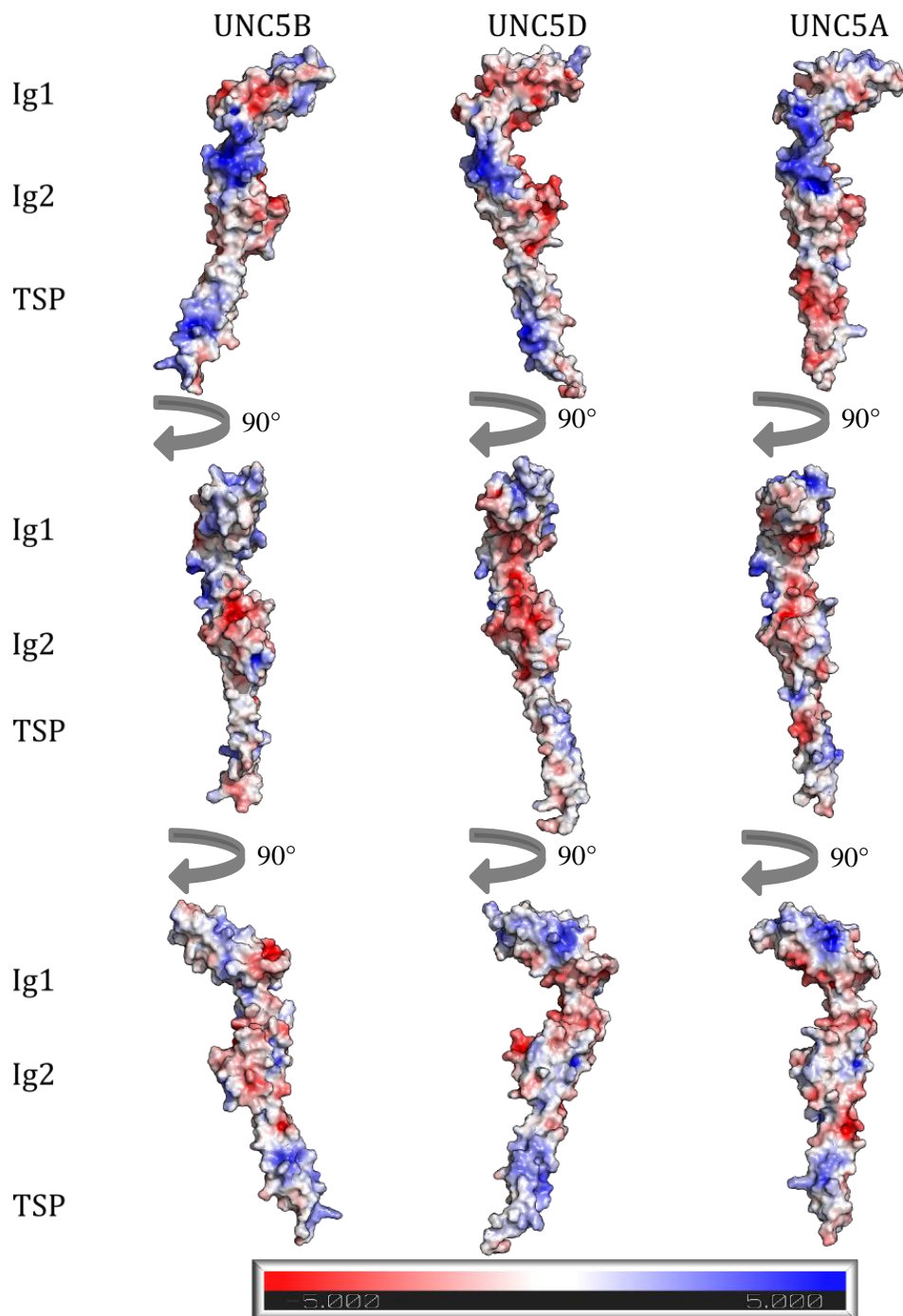


Figure 46: Electrostatic potential surfaces of the three known UNC5 structures
From negative (red) to positive (blue). Each structure is aligned in respect to their first Ig domain.

3.12.4 S-SAD analysis of UNC5B ecto

A S-SAD experiment was performed to help locate the missing TSP2 domain. Using anomalous difference maps derived from a multi crystal S-SAD data set, we resolved and validated the positions of disulphide bridges in the UNC5B ecto structure (Figure 47). Crystallographic data is summarised in Table 19. By comparing several datasets obtained using the clustering algorithm, we identified the positions of all disulphide bridges within the Ig1 (C69-C130, C81-C128, Figure 47A) and Ig2 (C174-C225, Figure 47B) domains, and the first and second disulphide bridge of the TSP1 domain (C258-265, C273-C285, Figure 47C).

Towards the C-term of the TSP1 domain, there are four other cysteine residues able to form disulphide bridges (C258, C262, C295 and C299). In the presented model, they were paired as C258-C295 and C262-C299. In this case, the anomalous density of C258-C295 was present in only one of the datasets (Figure 47C), but the C262-C299 density was not visible in any dataset. Unfortunately no anomalous density corresponding to disulphide bridges in the TSP2 domain was visible.

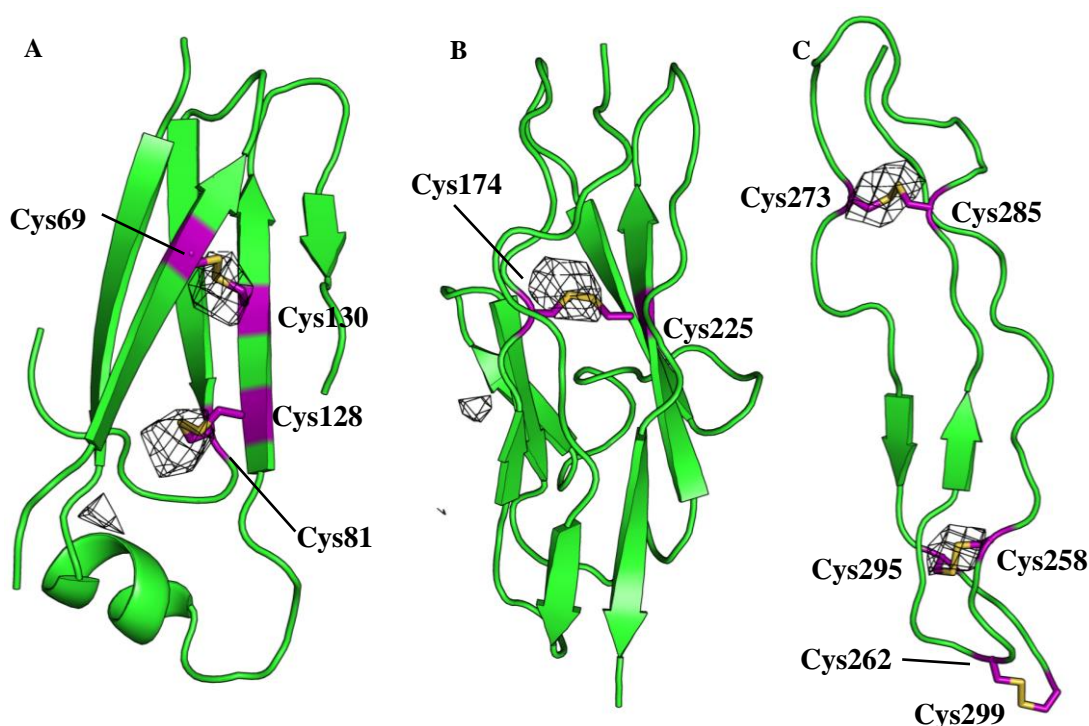


Figure 47: Anomalous density map at disulphide bridges of UNC5B ecto

Sulphur forming bridges are shown in purple and yellow as stick representations.

A: cross-section of Ig1 domain. Disulphide bridges are formed by Cys 66-130 and 81-128. Anomalous Fo-Fc map contoured at 3σ .

B: Ig2 domain with bridge forming Cys 174-225. Anomalous Fo-Fc map contoured at 3σ .

C: TSP1 domain with bridge forming Cys 273-285, 258-295 and 262-299. Anomalous Fo-Fc map contoured at 2σ . The sulphur bridge C262-C299 did not have visible anomalous density.

Table 19: Crystallographic table of UNC5B ecto anomalous sulphur datasets

The data were processed either with DIALS or XDS as indicated. Values for each clustered dataset are reported. Although data were collected up to the resolution indicated. Maps were created after applying an 8 Å cut-off, as that was the limit of the anomalous signal.

CLUSTER	1	2	3	4	5	6
Space group	I 4 ₁ 2 2	I 4 ₁ 2 2	I 4 ₁ 2 2	I 4 ₁ 2 2	I 4 ₁ 2 2	I 4 ₁ 2 2
Processing program	DIALS	DIALS	DIALS	DIALS	XDS	XDS
Cell dimensions						
a (Å)	70.25	70.14	70.12	70.11	70.23	70.23
b (Å)	70.25	70.14	70.12	70.11	70.23	70.23
c (Å)	394.88	394.28	394.10	394.08	396.42	396.46
α (°)	90	90	90	90	90	90
β (°)	90	90	90	90	90	90
γ (°)	90	90	90	90	90	90
Statistics						
Resolution (Å)	20.07-3.90	20.15-3.89	20.22-3.89	20.22-3.89	20.00-3.9	29.20-3.88
Unique reflections	4906	4907	4908	4908	4733	4954
Ano completeness (%)	99.6	99.7	99.7	99.7	92.7	95.2
Ano multiplicity	13.6	15.4	15.7	15.9	13.6	14.1
DelAnom correlation	0.012	0.019	0.015	0.018	-0.012	-0.198
Mid-slope Anom prob	0.983	0.949	0.954	0.963	1.015	0.990

4 DISCUSSION

Résumé en français

La deglycosylation enzymatique a confirmé que les domaines extracellulaires de Robo4 et UNC5B (Robo4 ecto et UNC5B ecto) sont largement glycosylés avec des glycans liés en azote du complexe type, qui ne peuvent pas être coupés par Endo H et Endo F1. La mutagenèse dirigée des sites de glycosylation prédits de Robo4 perturbe son expression, indiquant que ces résidus sont nécessaires pour la stabilité de la protéine et que leur glycosylation, ou leur passage dans la voie de glycosylation, pourrait être nécessaire pour un repliement correct. Les données MALS et SAXS montrent qu'en solution, Robo4 ecto est un monomère flexible de forme allongée. En utilisant des Fab, des cristaux du complexe Robo4 ecto / Fab ont été produits, mais la qualité n'était pas suffisante pour la détermination de la structure. Alors qu'un Fab de notre collaborateur était capable de se lier à des fragments plus courts du domaine extracellulaire, mais il était incapable de se lier au Robo4 complet, ce qui réduit son utilité pour une utilisation clinique. De même, l'analyse SAXS a montré qu'un autre Fab influence négativement le repliement de Robo4 ecto, ce qui le rend inapproprié aux études structurales. Les expériences de pull-down, SEC-MALS et SPR montrent que Robo4 ecto et UNC5B ecto n'interagissent pas entre elles, malgré une étude par un autre groupe montrant le contraire. Étant donné que différentes lignées cellulaires ont été utilisées, des modèles de glycosylation spécifiques, ou une tierce partie non détectée, pourraient être nécessaires pour l'interaction. En raison de leur implication avec les récepteurs extracellulaires, les héparanes sulfates sont un candidat probable, mais d'autres partenaires devraient être envisagés. La structure cristallographique de l'UNC5B ecto est similaire aux structures existantes d'UNC5A et UNC5D. Le haut degré de conservation d'un côté spécifique des domaines de l'Ig pourrait être une indication de l'importance de cette région, qui est responsable de la liaison à Latrophilin-3 dans UNC5D. Cependant, certains des résidus impliqués dans cette liaison sont mutés dans UNC5B. C'est ne toujours pas clair si UNC5B conserve cette interaction ou si c'est une caractéristique commune de toutes les protéines UNC5. Le travail présenté ici devrait servir de base à une meilleure caractérisation biochimique et structurale des récepteurs extracellulaires Robo4 et UNC5B.

4.1 Recombinant production of Robo4 and UNC5B extracellular domains

While it remains unclear if Slit proteins can bind Robo4 (Hohenester, 2008; Morlot et al., 2007; Park et al., 2003; Turner et al., 1990; Zhang et al., 2009), it has been shown that Robo4 can act as a UNC5B ligand (Koch et al., 2011). How Robo4 binding to UNC5B is relayed across the membrane to initiate signalling still remains elusive. As a single pass transmembrane receptor, a change in UNC5B oligomerisation could be implicated in signalling, as suggested for other receptors (Alberts et al., 2002; Moore et al., 2008). However, even when discussing other Robo proteins, there are still controversial opinions if this is mediated through homomerisation, heteromerisation, or a combination of these states upon Slit binding (Camurri et al., 2005; Evans and Bashaw, 2010b; Hivert, 2002; Hohenester, 2008; Kaur et al., 2006; Liu et al., 2004; Sheldon et al., 2009). Although recent studies are suggesting that a conformational change is required for signalling (Aleksandrova et al., in press).

To elucidate this mechanism, one of the main objectives of this study was to obtain structural information by X-ray crystallography on the extracellular domains of Robo4 and UNC5B individually, and in complex. Protocols to produce milligram amounts of recombinant Robo4 and UNC5B extracellular domains in mammalian (Robo4 ecto and UNC5B ecto) and insect cell (UNC5B ecto) expression system were devised, and successfully optimised for crystallographic and biophysical characterisation (Figure 13 and Figure 31). Both Robo4 ecto and UNC5B ecto presented anomalous elution SEC profiles, with elution volumes expected for much larger proteins (Figure 13A and D, Figure 31A and C). SDS-PAGE migration was also higher than that for proteins of their size (Figure 13B and E, Figure 31B and D). Both are common characteristics displayed by glycosylated proteins (Selcuk Unal et al., 2008).

4.2 The Robo4 and UNC5B extracellular domains are glycosylated and Robo4 glycosylation is necessary for folding

Robo4 ecto was extensively glycosylated (Figure 13B and C), but attempts to obtain a more homogenous sample by separating the various glycosidic species by ion exchange chromatography were unsuccessful (Figure 14). Furthermore, site-directed mutagenesis of the predicted glycosylation sites of Robo4 ecto entirely abolished expression (§§ 2.3.1 and 3.1, Table 6). To overcome this issue, deglycosylation of Robo4 ecto was performed using PNGase F1, which was identified as the best enzyme tested (Figure 15A), although it did not result in an entirely homogenous sample. This shows that the Robo4 extracellular domain contains complex N-linked glycosidic chains (Figure 9), as the other enzymes tested (Endo H and Endo F1), are not able to cleave this type of N-linked glycans (Maley et al., 1989; Tarentino and Plummer, 1994; Tarentino et al., 1992).

Deglycosylation does not affect the stability of the protein itself, but a mutation of these asparagine residues to aspartate or glutamine abolished their expression, as no protein could be detected in whole cell lysates. This suggests that the presence of glycosylation, or of these residues, is necessary for the stability and correct folding of Robo4 that is most likely mediated by passage through the endoplasmic reticulum (Aebi, 2013; Shental-Bechor and Levy, 2008). However, the presence of a glycosylated chain on each of those residues (Table 6) was not verified during this study. Simple mass spec analysis was not possible, as PNGase F1 digestion was not complete (Figure 15A and B) (Khoshnoodi et al., 2007) and, since more complex protocols are otherwise required (An et al., 2009; Nettleship et al., 2007), this matter should be addressed in further studies.

Robo4 ecto was also produced in the HEK293S GnTI⁻ cell line (Figure 13E and F), resulting in simpler glycosylation patterns (Reeves et al., 2002). In this case, Endo F1 and Endo H could more efficiently deglycosylate the protein when compared to PNGase F1 (Figure 15B), indicating that no complex N-linked chains are present, as expected from expression in this cell line.

The complete UNC5B extracellular domain was produced in large quantities using insect cells (Figure 31A and B) and in much smaller quantities using mammalian cells (Figure 31C and D). UNC5B ecto is also glycosylated and has two predicted

DISCUSSION

N-linked glycosylation sites in its extracellular region (Table 13). Here, only deglycosylation by PNGase F1 appeared to have any effect on SDS-PAGE migration on the insect cell expressed UNC5B (Figure 32A). HEK293 produced UNC5B ecto was more heterogeneously glycosylated, but less than that observed for Robo4 ecto. Similar to Robo4 ecto, PNGase F1 was more effective compared to Endo H and Endo F1 (Figure 32B).

Due to its central role in protein function, the glycosylation pathway is well conserved between different species, so insect cells can still offer similar, albeit simpler, N-linked glycosylation and induce functional folding (Rendic et al., 2008; Shi and Jarvis, 2007).

4.3 The Robo4 extracellular domain is a flexible monomer in solution

SEC-MALS measurement showed that the Robo4 extracellular domain exists as a monomer in solution and has a mass of ~56 kDa (Figure 16, Table 7). As extensive crystallisation attempts of Robo4 ecto alone in several different forms (distinguished by extent of glycosylation) were unsuccessful (§ 2.16), a SAXS analysis was undertaken to investigate its low resolution structural features (§ 3.6). The SAXS data shows that Robo4 ecto adopts an extended conformation as expected (Figure 26), and that it's highly flexible, as shown by the Kratky plot (Figure 24B). This is consistent with the bead model built from the SAXS data, which indicates that the Robo4 extracellular domain mostly exists in an elongated conformation, rather than a globular protein (Figure 27A). The shape and volume on the bead model produced can accommodate two Ig and two FnIII domains, as expected from its features (Figure 28). Averaging of all the possible conformations, which are inferred from the SAXS data, also shows a high degree of freedom for the extracellular domain (Figure 27A). Considering the flexibility of the Robo4 extracellular domain as measured by SAXS, it is also possible that multiple conformations are present, without a clear preference for one (or few). In this case, the use of the EOM approach, which allows for the selection of multiple conformers based on the experimental SAXS data, was also attempted. The low quality of the

models used, and the poor fit of the calculated ensemble to the scattering data, did not allow further analysis (Figure 29).

Other Robo receptors are reported to form monomeric, as well as homo and heterophilic assemblies (Camurri et al., 2005; Evans and Bashaw, 2010b; Hernández-Miranda et al., 2011; Hivert, 2002; Hohenester, 2008; Kaur et al., 2006; Koch et al., 2011; Liu et al., 2004; Sheldon et al., 2009). One of the current questions surrounding Robo4 is its oligomeric state at the membrane, and how this may correlate with signalling. The dimerisation of Robo4 has been suggested (Yadav and Narayan, 2014), based on similarity with Robo1 and Robo2 function (Hivert, 2002), but not directly proved. Other authors have suggested that Robo4 dimers are possible in absence of ligand and are mediated by the cytoplasmic domain, however, the data to support this hypothesis is not publicly available (Bedell et al., 2005).

Interestingly, SPR analysis shows that a small portion of Robo4 may form a weak homophilic interaction (Figure 37A), which was visible only at the highest concentration used (1 μM). No evidence of Robo4 oligomerisation has been reported until now, while SAXS and SEC-MALS data indicate that Robo4 is in monomeric state when at concentrations of $\sim 15\text{-}90 \mu\text{M}$. It should be noted that the observed effect might be caused by a transient interaction of Robo4 with the dextran matrix at the chip's surface, rather than by the formation of a homophilic interaction of Robo4. Due to its characteristics, dextran is a common immobilisation substrate for many types of proteins in SPR analysis, but it was shown to influence cell adhesion, which is mediated by its interaction with surface proteins (Massia et al., 2000; Neu et al., 2008). Possibly, further study through analytical ultracentrifugation might be helpful in providing further proof of the oligomeric state of Robo4 ecto.

Based on the current data from this study, the extracellular domain of Robo4 alone is not capable of dimerisation. At present it cannot be excluded that other factors, or the cytoplasmic domain, might be necessary to induce the supposed oligomerisation of Robo4.

4.4 Fabs interaction with the Robo4 extracellular domain

Synthetic antibodies are a useful tool not only for the development of therapeutic strategies (Adams and Sidhu, 2014; Deyev and Lebedenko, 2009), but also for structural applications (Dominik et al., 2016; Tereshko et al., 2008).

In order to improve the chance of crystallising Robo4 ecto, several Fabs that were selected by our collaborator (§ 2.9) were produced in *E. coli* (Figure 17). Complex formation between Fabs and Robo4 ecto were confirmed by SEC-MALS (Figure 18, Table 9) and SPR (Figure 21, Figure 22, Table 10). SEC-MALS showed that each Fab, except Fab5582, forms a 1:1 dimeric complex with the extracellular domain of Robo4 (Table 9, Figure 19). Fab5582 was the only one to show no interaction (Figure 18), despite being selected by our collaborator (Table 8) based on binding to their own Robo4 extracellular domain and library fragments screenings of the Robo4 extracellular domain. It's also possible this Fab binds a region of Robo4 ecto that is masked in the complete extracellular domain, either by folding or glycosylation.

The elution volume, and the measured mass, of the Robo4 ecto/Fab5585 complex were also unusual. The elution volume is an average between the elution volume of Robo4 ecto alone, and any of the other Robo4 ecto/Fab complexes (Figure 19), while the measured molecular weight is 81.37 kDa, instead of the expected 98.9 kDa. (Table 9). This might have been caused by a very fast association/dissociation kinetic, that can be seen from the steep slopes of the sensorgram during association and dissociation (Figure 22A), coupled with partial complex formation, as determined by the higher polydispersity in SEC-MALS compared to the other complexes (Table 9).

SPR was used to verify Fab/Robo4 ecto complex formation, and to determine their kinetic interaction and derived binding constants (Table 10). Fab5555, Fab5564 and Fab5570 follow the Langmuir 1:1 binding model (Figure 21A, C and D). Fab5562 did not correctly fit this model, but the K_D was determined to be similar to those of the other Fabs (Figure 21B), being in the low nanomolar range. The affinity of Fab5585 was lower than the others at 28 nM (Figure 22), again showing some peculiar characteristics compared to the other samples, as it did not entirely reach the targeted RU. In this case the K_D was calculated using the steady state

model (Figure 22B). Despite no interaction being observed by SEC-MALS, Fab5582 did show a very weak binding to Robo4 ecto by SPR (Figure 23), further reinforcing the idea that its binding site on the full length Robo4 extracellular domain is inaccessible. Thus Fab5585 was not used for further analyses.

Because Robo4 was observed to be flexible in solution (Figure 24A), it was hoped that an interaction with Fabs might reduce this intrinsic effect, resulting in a more stable complex amenable for crystallisation. The co-crystallisation of each Fab (except Fab5585) with Robo4 ecto was attempted and crystals of Robo4 ecto/Fab5562, Robo4 ecto/Fab5564 and Robo4 ecto/Fab5570 complexes were obtained (§ 3.7, Figure 30). Unfortunately thus far, the crystals were of insufficient quality to provide a complete dataset for structure determination. In order to obtain low resolution structural information, these complexes were also analysed by SAXS (Figure 25). Such a strategy has been successfully reported for other protein-antibody complexes (Chen et al., 2016; Walter et al., 2013).

All complexes showed a slight change in D_{max} (Figure 26), and a concomitant increase in volume consistent with complex formation (Table 11). In addition, the Robo4 ecto/Fab5555, Robo4 ecto/Fab5562, Robo4 ecto/Fab5564 and Robo4 ecto/Fab5570 complexes show an increased compactness, and a reduced flexibility compared to Robo4 ecto alone (Figure 25H to L). These results are also consistent with the appearance of crystals for Robo4 ecto/Fab5562, Robo4 ecto/Fab5564 and Robo4 ecto/Fab5570 that may one day yield diffraction quality crystals.

The complex Robo4 ecto/Fab5585, instead, showed a marked increase in disorder (Figure 25M), indicating that binding of this Fab was either disrupting Robo4 ecto folding, or the Fab itself is partially unfolded.

Frustratingly, the *ab initio* bead models derived from the SAXS data did not provide any additional insight into the Robo4 ecto domain composition (Figure 27B and C). Despite all statistics clearly indicating complex formation, no additional density could be unambiguously assigned for the Fabs that could be used to identify their Robo4 interaction region.

Although the use of the Fabs in this instance did not provide structurally relevant information, SAXS and SPR analysis provided useful data which are relevant for

DISCUSSION

their future uses. Fab5555, Fab5562, Fab5564 and Fab5570 were shown to be promising binding reagents with a very high affinity, which could potentially be used to specifically target the Robo4 extracellular domain in cellular assays. Since trafficking of Robo4 from the membrane to vesicles seems to be involved in its signalling (Sheldon et al., 2009; Zhang et al., 2016), these Fabs could be used to follow Robo4 movement within the cell (Kriebel et al., 2008; Prada et al., 2006).

While our collaborator implied that Fab5582 is able to bind isolated fragments of the Robo4 extracellular domain, the data presented here shows it is unable to bind the complete extracellular domain, greatly reducing its utility for biologically relevant applications. Similarly, Fab5585 had lower affinity in binding Robo4 ecto as compared to the other Fabs, and SAXS analysis suggested that the Robo4 ecto/Fab5585 complex is less ordered than Robo4 ecto alone. This deleterious effect is either caused by an instability induced by Fab5585, or by the intrinsic instability of Fab5585 itself. Either case makes this Fab unsuitable for use in crystallisation and other applications.

4.5 The UNC5B extracellular domain is a monomer in solution and does not interact with the Robo4 extracellular domain

Koch et al., showed that the interaction of Robo4 with the UNC5B extracellular domain is mediated through their Ig and TSP domains respectively, with an affinity of 12 nM (Koch et al., 2011). Successive studies have indirectly linked the activities of Robo4 and UNC5B together, but there has been no reported confirmation of this direct interaction. For instance, in their review, Yadav and Narayan (Yadav and Narayan, 2014) connect the findings of Suchting (Suchting et al., 2005) that the inhibition of VEGF by Robo4, is an UNC5B mediated effect, based on the retrospective assumption that Robo4 and UNC5B interact. While Zhang et al. (Zhang et al., 2016) show that UNC5B is unable to modulate VEGFR2 activity in absence of Robo4 by knockdown experiments, but do not otherwise provide evidence of a direct interaction between the two.

The main objective of this work was to obtain a high resolution structure of the complex of Robo4 and UNC5B to better elucidate their interaction. This was to be complemented by the identification of synthetic antibodies from the Sidhu lab that

Chapter 4.5 The UNC5B extracellular domain is a monomer in solution and does not interact with the Robo4 extracellular domain

inhibited this interaction. As the expression of the TSP domains, which should mediate UNC5B binding to Robo4, did not produce a sample viable for study (Figure 33), the full length UNCB extracellular domain was produced and used. The expression of UNC5B in insect cells was eventually pursued because of the extremely low yield obtained from mammalian cell expression.

SEC-MALS analysis confirmed that the extracellular domain of UNC5B is monomeric in solution and, similar to Robo4, its molecular weight is higher than expected (47.4 kDa) due to the presence of glycosylation (Figure 35, Table 14). Although it was not analysed by SAXS, the crystal structure of UNC5B ecto is consistent with a monomeric species and will be discussed in the next section (§ 4.6).

Despite the high K_D previously reported (12 nM) (Koch et al., 2011), no interaction between the UNC5B and Robo4 extracellular domains was detected in the experiments performed in this study. Several techniques were used to observe if this reported interaction was detectable under different conditions (eg. pH, salt concentration, and buffer). None of the assays tried, varying from pull-down (Figure 34) and SEC-MALS (Figure 35), to SPR (Figure 37), which was performed in the same experimental conditions as Koch's work (Koch et al., 2011), did not show any evidence of their interaction. To independently confirm this result, the Robo4 ecto construct used in this study was cloned into a prey vector for the AVEXIS screening (Bushell et al., 2008; Kerr and Wright, 2012) and sent to Rob Meijers' group in EMBL Hamburg. There, the Meijers group performed the AVEXIS interaction screen using their own bait UNC5B. Consistent with the results presented here, their assay also did not show any interaction in standard conditions (Figure 38A). Only when using a 20 times excess of Robo4, was a very weak signal observed (Figure 38B).

Based on the results obtained here, and in Hamburg, the Robo4 and UNC5B extracellular domains are unlikely to interact as described previously. Nevertheless, Koch's experiments show convincing evidence of the opposite, including association at the cell surface by fluorescent confocal microscopy on live cells *in vitro* (Koch et al., 2011).

DISCUSSION

While the constructs used to perform binding assays were similar (Koch's Robo4 and UNC5B were 7 and 3 amino acids shorter respectively), the only major difference were the cell lines used, where Koch used the kidney fibroblast-like *Cercopithecus ethiops* (COS-1) cell line for expression, and this study used the HEK293 cell line (Koch et al., 2011). Differences in the folding seem unlikely, but their glycosylation patterns might differ, which may influence binding. Another possibility is the involvement of an undetected third party, which either mediates the interaction between the extracellular domains, or induces the conformational change necessary to trigger such an interaction. One possible candidate could be heparin polysaccharides, which have been shown to participate in several extracellular complexes (Dreyfuss et al., 2009; Kreuger et al., 2006; Shimada et al., 1981; Trindade et al., 2008; Yayon et al., 1991). The role of heparin is already known in the Netrin/DCC interaction (Finci et al., 2015; Geisbrecht et al., 2003), and has been recently shown to play a role in the Robo1/Slit interaction (Li et al., 2015c). The binding of heparin, or heparan sulphate, to UNC5B has never been demonstrated. Furthermore, the residues of Robo1 that bind heparin (Gao et al., 2016; Zhang et al., 2013) are missing in Robo4, so it is also possible that another factor is responsible for the interaction.

While this, and the study from Alexander Koch (Koch et al., 2011), only take into consideration the extracellular domains, other possible interactions at the cytoplasmic level should not be excluded, and have not yet been investigated. For instance the cytoplasmic domains have been shown to be sufficient for UNC5/DCC interaction (Geisbrecht et al., 2003; Hong et al., 1999).

Considering how the Robo4/Slit2 interaction has been debated for several years, and is still under scrutiny (§ 1.3.2), further studies directly aimed at the characterisation of the putative Robo4/UNC5B interaction are necessary, in order to better understand the mechanisms they mediate.

4.6 Crystal structure of the UNC5B extracellular domain

The crystal structure of the UNC5B extracellular domain, including the first two Ig domains and up to the first TSP domain (45-299), was solved at 3.4 Å resolution (§ 3.12). The extracellular receptor assumes an extended conformation and is glycosylated at Asn 222, which can be modelled in the electron density (Figure 41). By using anomalous diffraction data of sulphur atoms the position of disulphide bridges internal to the Ig1-Ig2 domains and the TSP1 domain were validated, which led to increased confidence in the proposed model (Figure 47). The last TSP domain could not be modelled, as it is located in a large solvent pocket where it could move freely, making it impossible to trace at the current resolution.

The overall structures of human UNC5A, UNC5B and rat UNC5D are remarkably similar, differing mostly in the angles between each domain (Figure 43 and Table 18). If these characteristics represent the effective state *in vivo*, or artefacts due to crystal packing, is however up for debate. UNC5A, in respect to UNC5B and UNC5D, presents a shorter TSP domain with a slightly different folding loop (Figure 44), whose sequence is more similar to the TSP2 domain of UNC5B (Figure 45B). A shorter loop may also be due to UNC5A close proximity to the plasma membrane, as UNC5A contains only a single TSP domain. However, no structural information about the receptors associated to the membrane is currently available. An analysis of the evolutionary conservation between UNC5B homologues of different species shows a preferred conservation for the external side of the first Ig domain, while the internal face seems more variable. Ig2 has relatively good conservation, while the TSP1 domain is less conserved (Figure 42). Previous work from the Seiradake group (Jackson et al., 2016) showed that the rat UNC5D region, which interacts with Lphn3, is well conserved between all UNC5 homologues and between UNC5D homologues. It was therefore suggested that Latrophilin interactions might play an integral role in UNC5 signalling. Similarly, the corresponding region in UNC5B also has very high conservation between UNC5B homologues (Figure 42 and Figure 43A). However, two of the five proposed UNC5D residues involved in the interaction (His 125 and Pro 127) are mutated in UNC5B, in favour of Phe and Leu (Figure 40). While the His-Phe substitution might not have major effects on the overall structure, apart from the eventual loss of

DISCUSSION

polar interactions provided by His, the Pro-Leu substitution might introduce steric hindrance, lowering the affinity of the binding region. A more detailed analysis of the Lphn3/UNC5D binding interface, UNC5B, and a study of its eventual interaction with Lphn3, however, are necessary before further discussion. The current resolution of either structure does not allow for a better definition of the characteristics of these regions.

Interestingly, the less evolutionarily conserved internal face of the Ig domains (Figure 42) shows a more pronounced electrostatic negative surface charge in rat UNC5D compared to human UNC5B (Figure 46). This region, however, is not involved in any contacts in the octameric complex previously mentioned (Jackson et al., 2016) and it is currently unknown if it is involved in other interactions. However, some of the Netrin-1 residues involved in UNC5B binding have been identified, and are located in the Laminin V-2 domain of Netrin. Furthermore, specific arginines forming a positive charged surface in the Laminin V-2 domain have been shown to be necessary for binding (Grandin et al., 2016). Although the precise binding interface of Netrin on UNC5B (or the other UNC5 homologues) is still unknown, the binding region has been restricted to the Ig domains of UNC5B and, more specifically, to Ig2 (Geisbrecht et al., 2003; Grandin et al., 2016; Kruger et al., 2004). The negative charged surfaces that are present on Ig1 and Ig2 of UNC5A, UNC5B and UNC5D (Figure 46), are likely to have an important role in the Netrin/UNC5 interaction and should be further studied.

5 Conclusions and perspectives

While *in vivo* and *in vitro* cell studies, knockdowns and mRNA quantification give a global understanding of the relationship between cell processes, they are often not sufficient to determine if specific protein/protein interactions are present in a biological context. Similarly to the Slit2/Robo4 interaction debate, which is still on going, the Robo4/UNC5B interaction also needs to be more thoroughly investigated. In order to improve our general understanding of Robo4 and UNC5B signalling, a better characterisation of these receptors, and their interaction partners, are necessary in order to identify the characteristics that distinguish them from the other members of their families.

Biochemical and structural information can help to resolve these issues, by providing direct information on the interaction of these important receptors, or by predicting possible interaction surfaces on the single proteins. Recurring structural features and electrostatically charged surfaces can be used to identify regions of interest potentially involved in protein-protein, or protein-extracellular matrix, interaction. By correlating this information with the existing wealth of *in vitro* and *in vivo* studies, more focused experiments can be devised to test the possible interactions inferred by knock-down and functional experiments. Furthermore, by specifically developing antibodies or drugs aimed at these regions, we could directly intervene on their signalling pathways to influence cancer development or, to better study the cross-talk between different pathways. While further studies are necessary, the work presented in this thesis enriches the basic knowledge about the Robo4 and UNC5B extracellular domains. To the best of my knowledge it reports the first crystal structure of the UNC5B extracellular domain and biophysical characterisation of several high affinity Robo4 binding Fabs (Figure 48). I hope these results will provide a solid basis for further biochemical and molecular biology studies on the role of these two receptors, in order to draw a fuller picture on their mechanisms of action.

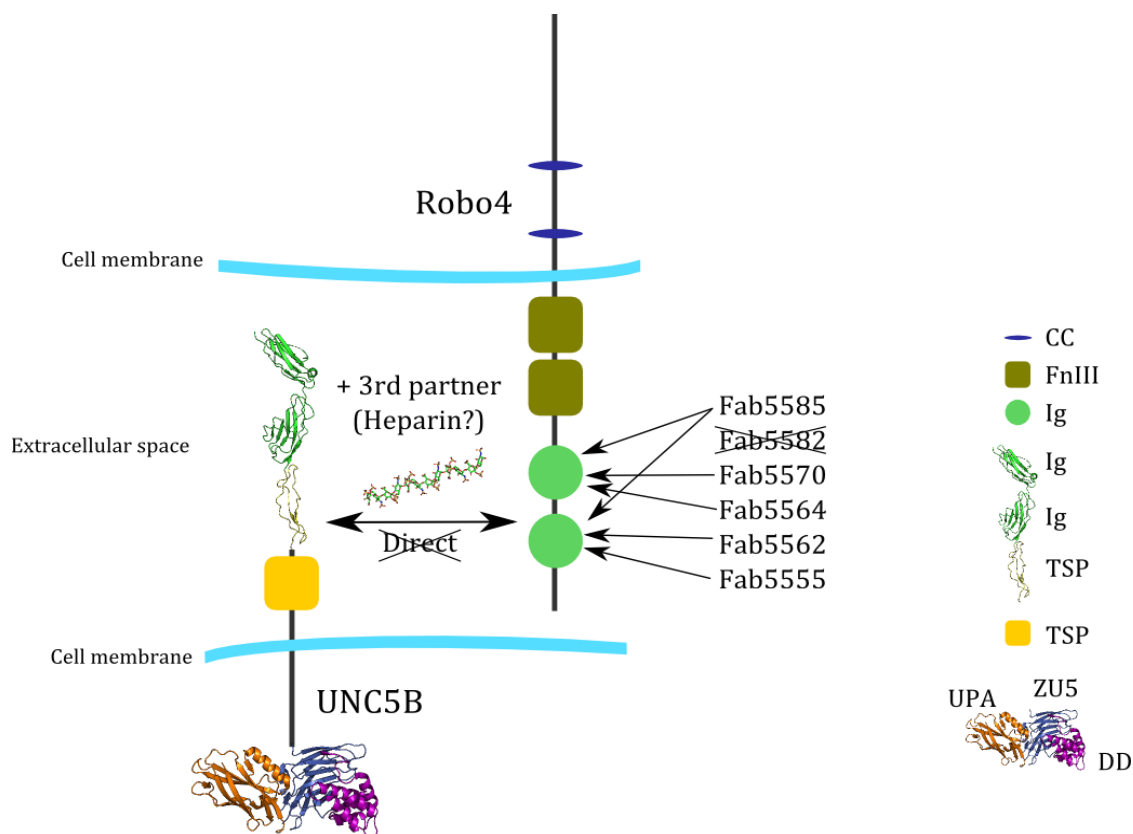


Figure 48: New insights into Robo4 and UNC5B function

The interaction of Robo4 and UNC5B has been described and is mediated by the Robo4 Ig domains and the UNC5B TSP domains (Koch et al., 2011). However, direct interaction is in contrast with the data presented here. Presence of a third partner is proposed, but is yet to be determined. In the figure is shown an example heparin structure ([PDB ID: 1HPN](#)) between Robo4 and UNC5B, but this is only meant for reflection and not as an actually present partner.

In light blue is represented the cytoplasmic membrane.

The crystal structure of the Ig1, Ig2 and TSP1 domain of UNC5B obtained in this study is shown.

The structure of the rat UNC5B cytoplasmic domains is from [PDB entry 3G5B](#).

The Fabs used in the course of this study and their proposed binding sites (Table 8) are indicated.

Fab5582 is crossed out to indicate absence of binding.

REFERENCES

- Ackerman, S.L., Kozak, L.P., Przyborski, S.A., Rund, L.A., Boyer, B.B., and Knowles, B.B. (1997). The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein. *Nature* 386, 838–842.
- Adams, J.J., and Sidhu, S.S. (2014). Synthetic antibody technologies. *Curr. Opin. Struct. Biol.* 24, 1–9.
- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta - Mol. Cell Res.* 1833, 2430–2437.
- Akin, O., and Lawrence Zipursky, S. (2016). Frazzled promotes growth cone attachment at the source of a Netrin gradient in the *Drosophila* visual system. *Elife* 5, 1–28.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Membrane proteins. In *Molecular Biology of the Cell*, (Garland Science), pp. 593–613.
- An, H.J., Froehlich, J.W., and Lebrilla, C.B. (2009). Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* 13, 421–426.
- Anelli, T., and Sitia, R. (2008). Protein quality control in the early secretory pathway. *EMBO J.* 27, 315–327.
- Arey, B.J. (2012). The Role of Glycosylation in Receptor Signaling. In *Glycosylation*, (InTech), p. 450.
- Aricescu, A.R., Lu, W., and Jones, E.Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62, 1243–1250.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.
- Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* 44, W344–W350.
- Aydin, H., Azimi, F.C., Cook, J.D., and Lee, J.E. (2012). A Convenient and General Expression Platform for the Production of Secreted Proteins from Human Cells. *J. Vis. Exp.* 1–7.
- Ba-charvet, K.T.N., Brose, K., Ma, L., Wang, K.H., Sotelo, C., Tessier-lavigne, M., and Che, A. (2001). Diversity and Specificity of Actions of Slit2 Proteolytic Fragments in Axon Guidance. *Development* 128, 4281–4289.
- Baker, K.A., Moore, S.W., Jarjour, A.A., and Kennedy, T.E. (2006). When a diffusible

REFERENCES

- axon guidance cue stops diffusing: roles for netrins in adhesion and morphogenesis. *Curr. Opin. Neurobiol.* *16*, 529–534.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 10037–10041.
- Barak, R., Lahmi, R., Gevorkyan-Airapetov, L., Levy, E., Tzur, A., and Opatowsky, Y. (2014). Crystal structure of the extracellular juxtamembrane region of Robo1. *J. Struct. Biol.* *186*, 283–291.
- Bashaw, G.J., Kidd, T., Murray, D., Pawson, T., and Goodman, C.S. (2000). Repulsive axon guidance: Abelson and Enabled play opposing roles downstream of the roundabout receptor. *Cell* *101*, 703–715.
- Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *67*, 271–281.
- Beamish, I. V., and Kennedy, T.E. (2015). Robo3: The road taken. *Dev. Cell* *32*, 3–4.
- Bedell, V.M., Yeo, S.-Y., Park, K.W., Chung, J., Seth, P., Shivalingappa, V., Zhao, J., Obara, T., Sukhatme, V.P., Drummond, I.A., et al. (2005). Roundabout4 Is Essential for Angiogenesis in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 6373–6378.
- Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D.I. (2007). Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* *129*, 5656–5664.
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* *42*, 252–258.
- Bicknell, R., and Harris, A.L. (2004). Novel angiogenic signaling pathways and vascular targets. *Annu. Rev. Pharmacol. Toxicol.* *44*, 219–238.
- Bieniossek, C., Richmond, T.J., and Berger, I. (2008). MultiBac: Multigene baculovirus-based eukaryotic protein complex production. *Curr. Protoc. Protein Sci.* 1–26.
- Blockus, H., and Chédotal, A. (2016). Slit-Robo signaling. *Development* *143*, 3037–3044.
- Boussif, O., Lezoualc’h, F., Zanta, M.A., Mergny, M.D., Scherman, D., Demeneix, B., and Behr, J.P. (1995). A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci.* *92*, 7297–7301.
- Bradbury, A.R.M., Sidhu, S.S., Dübel, S., and McCafferty, J. (2011). Beyond natural

- antibodies: the power of in vitro display technologies. *Nat. Biotechnol.* *29*, 245–254.
- Bredesen, D.E., Mehlen, P., and El-Deiry, W.S. (2005). Receptors that mediate cellular dependence. *Cell Death Differ.* *12*, 1031–1043.
- Brennich, M.E., Kieffer, J., Bonamis, G., De Maria Antolinos, A., Hutin, S., Pernot, P., and Round, A. (2016). Online data analysis at the ESRF bioSAXS beamline, BM29. *J. Appl. Crystallogr.* *49*, 203–212.
- Brennich, M.E., Round, A.R., and Hutin, S. (2017). Online Size-exclusion and Ion-exchange Chromatography on a SAXS Beamline. *J. Vis. Exp.* 1–9.
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., et al. (2016). BUSTER version 2.10.3. Cambridge, United Kingdom Glob. Phasing Ltd.
- van den Brink, D.M., Banerji, O., and Tear, G. (2013). Commissureless Regulation of Axon Outgrowth across the Midline Is Independent of Rab Function. *PLoS One* *8*.
- Broennimann, C., Eikenberry, E.F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., et al. (2006). The PILATUS 1M detector. *J. Synchrotron Radiat.* *13*, 120–130.
- Brose, K., Bland, K.S., Wang, K.H., Arnott, D., Henzel, W., Goodman, C.S., Tessier-Lavigne, M., and Kidd, T. (1999). Slit proteins bind robo receptors and have an evolutionarily conserved role in repulsive axon guidance. *Cell* *96*, 795–806.
- Bushell, K.M., Söllner, C., Schuster-Boeckler, B., Bateman, A., and Wright, G.J. (2008). Large-scale screening for novel low-affinity extracellular protein interactions. *Genome Res.* *18*, 622–630.
- Camurri, L., Mambetisaeva, E.T., Davies, D., Parnavelas, J., Sundaresan, V., and Andrews, W. (2005). Evidence for the existence of two Robo3 isoforms with divergent biochemical properties. *Mol. Cell. Neurosci.* *30*, 485–493.
- Caramelo, J.J., and Parodi, A.J. (2015). A sweet code for glycoprotein folding. *FEBS Lett.* *589*, 3379–3387.
- Carmeliet, P., and Tessier-Lavigne, M. (2005). Common mechanisms of nerve and blood vessel wiring. *Nature* *436*, 193–200.
- Castets, M., Coissieux, M.-M., Delloye-Bourgeois, C., Bernard, L., Delcros, J.-G., Bernet, A., Laudet, V., and Mehlen, P. (2009). Inhibition of Endothelial Cell Apoptosis by Netrin-1 during Angiogenesis. *Dev. Cell* *16*, 614–620.
- Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2013). ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Isr. J. Chem.* *53*, 199–206.

REFERENCES

- Chan, S.S.-Y., Zheng, H., Su, M.-W., Wilk, R., Killeen, M.T., Hedgecock, E.M., and Culotti, J.G. (1996). UNC-40, a *C. elegans* Homolog of DCC (Deleted in Colorectal Cancer), Is Required in Motile Cells Responding to UNC-6 Netrin Cues. *Cell* *87*, 187–195.
- Chance, R.K., and Bashaw, G.J. (2015). Slit-Dependent Endocytic Trafficking of the Robo Receptor Is Required for Son of Sevenless Recruitment and Midline Axon Repulsion. *PLoS Genet.* *11*.
- Chang, V.T., Crispin, M., Aricescu, A.R., Harvey, D.J., Nettleship, J.E., Fennelly, J.A., Yu, C., Boles, K.S., Evans, E.J., Stuart, D.I., et al. (2007). Glycoprotein Structural Genomics: Solving the Glycosylation Problem. *Structure* *15*, 267–273.
- Charron, F. (2015). Signaling from Within: Endocytic Trafficking of the Robo Receptor Is Required for Midline Axon Repulsion. *PLoS Genet.* *11*, 1–4.
- Chen, E., Salinas, N.D., Huang, Y., Ntumngia, F., Plasencia, M.D., Gross, M.L., Adams, J.H., and Tolia, N.H. (2016). Broadly neutralizing epitopes in the Plasmodium vivax vaccine candidate Duffy Binding Protein. *Proc. Natl. Acad. Sci.* *113*, 6277–6282.
- Chen, J., Wen, L., Dupuis, S., Wu, J.Y., and Rao, Y. (2001). The N-terminal leucine-rich regions in Slit are sufficient to repel olfactory bulb axons and subventricular zone neurons. *J. Neurosci.* *21*, 1548–1556.
- Chen, Z., Gore, B.B., Long, H., Ma, L., and Tessier-Lavigne, M. (2008). Alternative Splicing of the Robo3 Axon Guidance Receptor Governs the Midline Switch from Attraction to Repulsion. *Neuron* *58*, 325–332.
- Cipriani, F., Röwer, M., Landret, C., Zander, U., Felisaz, F., and Márquez, J.A. (2012). CrystalDirect: A new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *68*, 1393–1399.
- Coleman, H.A., Labrador, J.-P., Chance, R.K., and Bashaw, G.J. (2010). The Adam family metalloprotease Kuzbanian regulates the cleavage of the roundabout receptor to control axon repulsion at the midline. *Development* *137*, 2417–2426.
- Dalton, A.C., and Barton, W.A. (2014). Over-expression of secreted proteins from mammalian cell lines. *Protein Sci.* *23*, 517–525.
- Dascenco, D., Erfurth, M.L., Izadifar, A., Song, M., Sachse, S., Bortnick, R., Urwyler, O., Petrovic, M., Ayaz, D., He, H., et al. (2015). Slit and Receptor Tyrosine Phosphatase 69D Confer Spatial Specificity to Axon Branching via Dscam1. *Cell* *162*, 1140–1154.
- Davis, M.W. (2017). ApE, A plasmid Editor.
- Deyev, S.M., and Lebedenko, E.N. (2009). Modern Technologies for Creating Synthetic Antibodies for Clinical application. *Acta Naturae* *1*, 32–50.

- Dickson, B.J., and Gilestro, G.F. (2006). Regulation of Commissural Axon Pathfinding by Slit and its Robo Receptors. *Annu. Rev. Cell Dev. Biol.* 22, 651–675.
- Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32, 665–667.
- Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007). PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 35, 522–525.
- Dominik, P.K., Borowska, M.T., Dalmas, O., Kim, S.S., Perozo, E., Keenan, R.J., and Kossiakoff, A.A. (2016). Conformational Chaperones for Structural Studies of Membrane Proteins Using Antibody Phage Display with Nanodiscs. *Structure* 24, 300–309.
- Dreyfuss, J.L., Regatieri, C. V, Jarrouge, T.R., Cavalheiro, R.P., Sampaio, L.O., and Nader, H.B. (2009). Heparan sulfate proteoglycans: structure, protein interactions and cell signaling. *An. Acad. Bras. Cienc.* 81, 409–429.
- Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P., and Fieschi, F. (2010). NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* 169, 45–53.
- Edelman, J. (1992). The low temperature heat capacity of solid proteins. *Biopolymers* 32, 209–218.
- Edgar, R.C. (2004a). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Edgar, R.C. (2004b). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K.D. (2010). Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66, 486–501.
- Engelkamp, D. (2002). Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation. *Mech. Dev.* 118, 191–197.
- Enomoto, S., Mitsui, K., Kawamura, T., Iwanari, H., Daigo, K., Horiuchi, K., Minami, T., Kodama, T., and Hamakubo, T. (2016). Suppression of Slit2/Robo1 mediated HUVEC migration by Robo4. *Biochem. Biophys. Res. Commun.* 469, 797–802.
- Escriou, V., Ciolina, C., Lacroix, F., Byk, G., Scherman, D., and Wils, P. (1998). Cationic lipid-mediated gene transfer: Effect of serum on cellular uptake and intracellular fate of lipopolyamine/DNA complexes. *Biochim. Biophys. Acta - Biomembr.* 1368, 276–288.

REFERENCES

- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *62*, 72–82.
- Evans, P.R. (2011). An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *67*, 282–292.
- Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* *69*, 1204–1214.
- Evans, T.A., and Bashaw, G.J. (2010a). Axon guidance at the midline: of mice and flies. *Curr. Opin. Neurobiol.* *20*, 79–85.
- Evans, T.A., and Bashaw, G.J. (2010b). Functional Diversity of Robo Receptor Immunoglobulin Domains Promotes Distinct Axon Guidance Decisions. *Curr. Biol.* *20*, 567–572.
- Evans, T.A., and Bashaw, G.J. (2012). Slit/Robo-mediated axon guidance in *Tribolium* and *Drosophila*: Divergent genetic programs build insect nervous systems. *Dev. Biol.* *363*, 266–278.
- Ten Eyck, L.F. (1973). Crystallographic fast Fourier transforms. *Acta Crystallogr. Sect. A* *29*, 183–191.
- Fagerberg, L., Hallström, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., et al. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteomics* *13*, 397–406.
- Feigin, and Svergun, D.I. (1989). Structure Analysis by Small-Angle X-Ray and Neutron Scattering. *Acta Polym.* *40*, 224–224.
- Finci, L., Zhang, Y., Meijers, R., and Wang, J.H. (2015). Signaling mechanism of the netrin-1 receptor DCC in axon guidance. *Prog. Biophys. Mol. Biol.* *118*, 153–160.
- Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., and Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* *3*, 1021–1032.
- Franke, D., and Svergun, D.I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* *42*, 342–346.
- French, S., and Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallogr. Sect. A* *34*, 517–525.
- Gabardinho, J., Beteva, A., Guijarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M.W., Brockhauser, S., Flot, D., Gordon, E.J., Hall, D.R., et al. (2010). MxCuBE: A synchrotron beamline control environment customized for macromolecular crystallography experiments. *J. Synchrotron Radiat.* *17*, 700–707.

- Gao, Q., Chen, C.-Y.Y., Zong, C., Wang, S., Ramiah, A., Prabhakar, P., Morris, L.C., Boons, G.-J.J., Moremen, K.W., and Prestegard, J.H. (2016). Structural aspects of heparan sulfate binding to Robo1-Ig1-2. *ACS Chem. Biol.* *11*, 3106–3113.
- Garbe, D.S., and Bashaw, G.J. (2004). Axon guidance at the midline: from mutants to mechanisms. *Crit. Rev. Biochem. Mol. Biol.* *39*, 319–341.
- Geisbrecht, B. V., Dowd, K.A., Barfield, R.W., Longo, P.A., and Leahy, D.J. (2003). Netrin binds discrete subdomains of DCC and UNC5 and mediates interactions between DCC and heparin. *J. Biol. Chem.* *278*, 32561–32568.
- Georgiou, M., and Tear, G. (2002). Commissureless is required both in commissural neurones and midline cells for axon guidance across the midline. *Development* *129*, 2947–2956.
- Ghose, A., and Van Vactor, D. (2002). Gaps in Slit-Robo signaling. *BioEssays* *24*, 401–404.
- Gitai, Z., Yu, T.W., Lundquist, E.A., Tessier-Lavigne, M., and Bargmann, C.I. (2003). The netrin receptor UNC-40/DCC stimulates axon attraction and outgrowth through enabled and, in parallel, Rac and UNC-115/abLIM. *Neuron* *37*, 53–65.
- Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor, D., and Martz, E. (2003). ConSurf: identification of functional regions in proteins by surface mapping of phylogenetic information. *Bioinformatics* *19*, 163–164.
- Glatter, V.O., and Kratky, O. (1982). Small Angle X-ray Scattering. *Small Angle X-Ray Scatt.* *36*, 1985.
- Goldman, J.S., and Kennedy, T.E. (2011). The clasp between NetrinG and NGL becomes crystal clear. *EMBO J.* *30*, 4342–4344.
- Grandin, M., Meier, M., Delcros, J.G., Nikodemus, D., Reuten, R., Patel, T.R., Goldschneider, D., Orriss, G., Krahn, N., Boussouar, A., et al. (2016). Structural decoding of the Netrin-1/UNC5 interaction and its therapeutical implications in cancers. *Cancer Cell* *29*, 173–185.
- Gröne, J., Doeblner, O., Loddenkemper, C., Hotz, B., Buhr, H.J., and Bhargava, S. (2006). Robo1/Robo4: Differential expression of angiogenic markers in colorectal cancer. *Oncol. Rep.* *15*, 1437–1443.
- Grueninger-leitch, F., D’Arcy, A., D’Arcy, B., and Chene, C. (1996). Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases. *Protein Sci.* 2617–2622.
- Guex, N., Peitsch, M.C., and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* *30*, 162–173.
- Gupta, R., Jung, E., and Brunak, S. (2004). Prediction of N-glycosylation sites in human proteins. *46*, 203–206.

REFERENCES

- Hammond, C., and Helenius, A. (1995). Quality control in the secretory pathway. *Curr. Opin. Cell Biol.* 7, 523–529.
- Han, Y., Shao, Y., Lin, Z., Qu, Y.L., Wang, H., Zhou, Y., Chen, W., Chen, Y., Chen, W.L., Hu, F.R., et al. (2012). Netrin-1 simultaneously suppresses corneal inflammation and neovascularization. *Investig. Ophthalmol. Vis. Sci.* 53, 1285–1295.
- Hayward, S., and Berendsen, H.J.C. (1998). Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins Struct. Funct. Genet.* 30, 144–154.
- Hedgecock, E.M., Culotti, J.G., and Hall, D.H. (1990). The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. *Neuron* 4, 61–85.
- Hernández-Miranda, L.R., Cariboni, A., Faux, C., Ruhrberg, C., Cho, J.H., Cloutier, J.-F., Eickholt, B.J., Parnavelas, J.G., and Andrews, W.D. (2011). Robo1 regulates semaphorin signaling to guide the migration of cortical interneurons through the ventral forebrain. *J. Neurosci.* 31, 6174–6187.
- Hiramoto, M., Hiromi, Y., Giniger, E., and Hotta, Y. (2000). The *Drosophila* Netrin receptor Frazzled guides axons by controlling Netrin distribution. *Nature* 406, 886–889.
- Hivert, B. (2002). Robo1 and Robo2 Are Homophilic Binding Molecules That Promote Axonal Growth. *Mol. Cell. Neurosci.* 21, 534–545.
- Hofmann, K., and Tschopp, J. (1995). The death domain motif found in Fas (Apo-1) and TNF receptor is present in proteins involved in apoptosis and axonal guidance. *FEBS Lett.* 371, 321–323.
- Hohenester, E. (2008). Structural insight into Slit-Robo signalling. *Biochem. Soc. Trans.* 36, 251–256.
- Hong, K., Hinck, L., Nishiyama, M., Poo, M.M., Tessier-Lavigne, M., and 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, 927–941.
- Howitt, J.A., Clout, N.J., and Hohenester, E. (2004). Binding site for Robo receptors revealed by dissection of the leucine-rich repeat region of Slit. *EMBO J.* 23, 4406–4412.
- Humniecki, L., and Bicknell, R. (2000). In silico cloning of novel endothelial-specific genes. *Genome Res.* 10, 1796–1806.
- Humniecki, L., Gorn, M., Suchting, S., Poulsom, R., and Bicknell, R. (2002). Magic roundabout is a new member of the roundabout receptor family that is endothelial specific and expressed at sites of active angiogenesis. *Genomics* 79, 547–552.

- Incardona, M.F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., and Svensson, O. (2009). EDNA: A framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron Radiat.* *16*, 872–879.
- Ishida, T., and Kinoshita, K. (2007). PrDOS: Prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* *35*, 460–464.
- 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., et al. (2016). Super-complexes of adhesion GPCRs and neural guidance receptors. *Nat. Commun.* *7*, 11184.
- Jain, R.K. (1990). Physiological barriers to delivery of monoclonal antibodies and other macromolecules in tumors. *Cancer Res.* *50*, 814s – 819s.
- Jones, C.A., London, N.R., Chen, H., Park, K.W., Sauvaget, D., Stockton, R.A., Wythe, J.D., Suh, W., Larrieu-Lahargue, F., Mukoyama, Y., et al. (2008). Robo4 stabilizes the vascular network by inhibiting pathologic angiogenesis and endothelial hyperpermeability. *Nat. Med.* *14*, 448–453.
- Jones, C.A., Nishiya, N., London, N.R., Zhu, W., Sorensen, L.K., Chan, A.C., Lim, C.J., Chen, H., Zhang, Q., Schultz, P.G., et al. (2009). Slit2-Robo4 signalling promotes vascular stability by blocking Arf6 activity. *Nat. Cell Biol.* *11*, 1325–1331.
- Jongbloets, B.C., and Pasterkamp, R.J. (2014). Semaphorin signalling during development. *Development* *141*, 3292–3297.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *66*, 125–132.
- Karlsson, R., and Larsson, A. (2004). Affinity Measurement Using Surface Plasmon Resonance. *Antib. Eng.* *248*, 389–415.
- Karve, T.M., and Cheema, A.K. (2011). Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *J. Amino Acids* *2011*, 1–13.
- Kaur, S., Castellone, M.D., Bedell, V.M., Konar, M., Gutkind, J.S., and Ramchandran, R. (2006). Robo4 signaling in endothelial cells implies attraction guidance mechanisms. *J. Biol. Chem.* *281*, 11347–11356.
- Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S.-Y., Culotti, J.G., and Tessier-Lavigne, M. (1996). Deleted in Colorectal Cancer (DCC) Encodes a Netrin Receptor. *Cell* *87*, 175–185.
- Keleman, K., and Dickson, B.J. (2001). Short- and long-range repulsion by the *Drosophila* Unc5 Netrin receptor. *Neuron* *32*, 605–617.
- Keleman, K., Rajagopalan, S., Cleppien, D., Teis, D., Paiha, K., Huber, L.A., Technau, G.M., and Dickson, B.J. (2002). Comm sorts Robo to control axon guidance at the *Drosophila* midline. *Cell* *110*, 415–427.

REFERENCES

- Keller, M.A., and Stiehm, E.R. (2000). Passive immunity in prevention and treatment of infectious diseases. *Clin. Microbiol. Rev.* *13*, 602–614.
- Kennedy, T.E. (2000). Cellular mechanisms of netrin function: long-range and short-range actions. *Biochem. Cell Biol.* *78*, 569–575.
- Kerr, J.S., and Wright, G.J. (2012). Avidity-based Extracellular Interaction Screening (AVEXIS) for the Scalable Detection of Low-affinity Extracellular Receptor-Ligand Interactions. *J. Vis. Exp.* 1–8.
- Khoshnoodi, J., Hill, S., Tryggvason, K., Hudson, B., and Friedman, D.B. (2007). Identification of N-linked glycosylation sites in human nephrin using mass spectrometry. *J. Mass Spectrom.* *42*, 370–379.
- Kidd, T., Brose, K., Mitchell, K.J., Fetter, R.D., Tessier-Lavigne, M., Goodman, C.S., and Tear, G. (1998). Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell* *92*, 205–215.
- Kidd, T., Bland, K.S., and Goodman, C.S. (1999). Slit Is the Midline Repellent for the Robo Receptor in *Drosophila*. *Cell* *96*, 785–794.
- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* *37*, 387–392.
- Kirsch, M., Zaman, M., Meier, D., Dübel, S., and Hust, M. (2005). Parameters affecting the display of antibodies on phage. *J. Immunol. Methods* *301*, 173–185.
- Klagsbrun, M., and Eichmann, A. (2005). A role for axon guidance receptors and ligands in blood vessel development and tumor angiogenesis. *Cytokine Growth Factor Rev.* *16*, 535–548.
- Klenotic, P.A., Page, R.C., Misra, S., and Silverstein, R.L. (2011). Expression, purification and structural characterization of functionally replete thrombospondin-1 type 1 repeats in a bacterial expression system. *Protein Expr. Purif.* *80*, 253–259.
- Knorre, D.G., Kudryashova, N. V., and Godovikova, T.S. (2009). Chemical and functional aspects of posttranslational modification of proteins. *Acta Naturae* *1*, 29–51.
- Koch, A.W., Mathivet, T., Larrivé, B., Tong, R.K., Kowalski, J., Pibouin-Fragner, L., Bouvrée, K., Stawicki, S., Nicholes, K., Rathore, N., et al. (2011). Robo4 Maintains Vessel Integrity and Inhibits Angiogenesis by Interacting with UNC5B. *Dev. Cell* *20*, 33–46.
- Köhler, D., Streißenberger, A., König, K., Granja, T., Roth, J.M., Lehmann, R., de Oliveira Franz, C.B., and Rosenberger, P. (2013). The Uncoordinated-5 Homolog B (UNC5B) Receptor Increases Myocardial Ischemia-Reperfusion Injury. *PLoS One* *8*, 1–9.

- Kreuger, J., Spillmann, D., Li, J.P., and Lindahl, U. (2006). Interactions between heparan sulfate and proteins: The concept of specificity. *J. Cell Biol.* *174*, 323–327.
- Kriebel, P.W., Barr, V.A., Rericha, E.C., Zhang, G., and Parent, C.A. (2008). Collective cell migration requires vesicular trafficking for chemoattractant delivery at the trailing edge. *J. Cell Biol.* *183*, 949–961.
- Kruger, R.P., Lee, J., Li, W., and Guan, K.-L. (2004). Mapping netrin receptor binding reveals domains of Unc5 regulating its tyrosine phosphorylation. *J. Neurosci.* *24*, 10826–10834.
- Kullander, K., and Klein, R. (2002). Mechanisms and functions of eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* *3*, 475–486.
- De La Torre, J.R., Höpker, V.H., Ming, G.L., Poo, M.M., Tessier-Lavigne, M., Hemmati-Brivanlou, A., and Holt, C.E. (1997). Turning of retinal growth cones in a netrin-1 gradient mediated by the netrin receptor DCC. *Neuron* *19*, 1211–1224.
- Lai Wing Sun, K., Correia, J.P., and Kennedy, T.E. (2011). Netrins: versatile extracellular cues with diverse functions. *Development* *138*, 2153–2169.
- Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* *33*, 299–302.
- Larrivé, B., Freitas, C., Trombe, M., Lv, X., Delafarge, B., Yuan, L., Bouvrée, K., Bréant, C., Toro, R. Del, Bréchet, N., et al. (2007). Activation of the UNC5B receptor by Netrin-1 inhibits sprouting angiogenesis. 2433–2447.
- Leavy, O. (2010). Therapeutic antibodies: past, present and future. *Nat Rev Immunol* *10*, 297.
- Lee, J., Li, W., and Guan, K.-L. (2005). SRC-1 mediates UNC-5 signaling in *Caenorhabditis elegans*. *Mol. Cell Biol.* *25*, 6485–6495.
- Legg, J.A., Herbert, J.M., Clissold, P., and Bicknell, R. (2008). Slits and Roundabouts in cancer, tumour angiogenesis and endothelial cell migration. *Angiogenesis* *11*, 13–21.
- Lejmi, E., Leconte, L., Pédrón-Mazoyer, S., Ropert, S., Raoul, W., Lavalette, S., Bouras, I., Feron, J.-G., Maitre-Boube, M., Assayag, F., et al. (2008). Netrin-4 inhibits angiogenesis via binding to neogenin and recruitment of Unc5B. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 12491–12496.
- Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L., and Tessier-Lavigne, M. (1997). Vertebrate homologues of *C. elegans* UNC-5 are candidate netrin receptors. *Nature* *386*, 833–838.
- Leslie, A.G.W. (2006). The integration of macromolecular diffraction data. *Acta*

REFERENCES

- Crystallogr. Sect. D Biol. Crystallogr. *62*, 48–57.
- Leung-Hagesteijn, C., Spence, A.M., Stern, B.D., Zhou, Y., Su, M.W., Hedgecock, E.M., and 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*, 289–299.
- Li, H., Chen, J., Wu, W., Fagaly, T., Zhou, L., Yuan, W., Dupuis, S., Jiang, Z., Nash, W., Gick, C., et al. (1999). Vertebrate Slit, a Secreted Ligand for the Transmembrane Protein Roundabout, Is a Repellent for Olfactory Bulb Axons. *Cell* *96*, 807–818.
- Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., and Lopez, R. (2015a). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* *43*, W580–W584.
- Li, X., Chen, Y., Liu, Y., Gao, J., Gao, F., Bartlam, M., Wu, J.Y., and Rao, Z. (2006). Structural basis of Robo proline-rich motif recognition by the srGAP1 Src homology 3 domain in the Slit-Robo signaling pathway. *J. Biol. Chem.* *281*, 28430–28437.
- Li, Y., Cheng, H., Xu, W., Tian, X., Li, X., and Zhu, C. (2015b). Expression of Robo protein in bladder cancer tissues and its effect on the growth of cancer cells by blocking Robo protein. *Int. J. Clin. Exp. Pathol.* *8*, 9932–9940.
- Li, Z., Moniz, H., Wang, S., Ramiah, A., Zhang, F., Moremen, K.W., Linhardt, R.J., and Sharp, J.S. (2015c). High structural resolution hydroxyl radical protein footprinting reveals an extended Robo1-heparin binding interface. *J. Biol. Chem.* *290*, 10729–10740.
- Liu, H., and Naismith, J.H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* *8*, 91.
- Liu, D., Hou, J., Hu, X., Wang, X., Xiao, Y., Mou, Y., and De Leon, H. (2006). Neuronal Chemorepellent Slit2 Inhibits Vascular Smooth Muscle Cell Migration by Suppressing Small GTPase Rac1 Activation. *Circ. Res.* *98*, 480–489.
- Liu, Z., Patel, K., Schmidt, H., Andrews, W.D., Pini, A., and Sundaresan, V. (2004). Extracellular Ig domains 1 and 2 of Robo are important for ligand (Slit) binding. *Mol. Cell. Neurosci.* *26*, 232–240.
- Llambi, F., Causeret, F., Bloch-Gallego, E., and Mehlen, P. (2001). Netrin-1 acts as a survival factor via its receptors UNC5H and DCC. *EMBO J.* *20*, 2715–2722.
- Llambi, F., Lourenço, F.C., Gozuacik, D., Guix, C., Pays, L., Del Rio, G., Kimchi, A., and Mehlen, P. (2005). The dependence receptor UNC5H2 mediates apoptosis through DAP-kinase. *EMBO J.* *24*, 1192–1201.
- Long, H., Sabatier, C., Ma, L., Plump, A., Yuan, W., Ornitz, D.M., Tamada, A., Murakami, F., Goodman, C.S., and Tessier-Lavigne, M. (2004). Conserved roles

- for Slit and Robo proteins in midline commissural axon guidance. *Neuron* *42*, 213–223.
- Longo, P. a, Kavran, J.M., Kim, M., and Leahy, D.J. (2014). Transient Mammalian Cell Transfection with Polyethylenimine. *Methods Enzymol.* *529*, 227–240.
- Lu, X., Le Noble, F., Yuan, L., Jiang, Q., De Lafarge, B., Sugiyama, D., Bréant, C., Claes, F., De Smet, F., Thomas, J.-L., et al. (2004). The netrin receptor UNC5B mediates guidance events controlling morphogenesis of the vascular system. *Nature* *432*, 179–186.
- Luft, J.R., and DeTitta, G.T. (1999). A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *55*, 988–993.
- Ly, A., Nikolaev, A., Suresh, G., Zheng, Y., Tessier-Lavigne, M., and Stein, E. (2008). DSCAM Is a Netrin Receptor that Collaborates with DCC in Mediating Turning Responses to Netrin-1. *Cell* *133*, 1241–1254.
- Maley, F., Trimble, R.B., Tarentino, A.L., and Plummer, T.H. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal. Biochem.* *180*, 195–204.
- Manhire-Heath, R., Golenkina, S., Saint, R., and Murray, M.J. (2013). Netrin-dependent downregulation of Frazzled/DCC is required for the dissociation of the peripodial epithelium in *Drosophila*. *Nat. Commun.* *4*, 1–10.
- De Maria Antolinos, A., Pernot, P., Brennich, M.E., Kieffer, J., Bowler, M.W., Delageniere, S., Ohlsson, S., Malbet Monaco, S., Ashton, A., Franke, D., et al. (2015). ISPyB for BioSAXS, the gateway to user autonomy in solution scattering experiments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *71*, 76–85.
- Marlow, R., Binnewies, M., Sorensen, L.K., Monica, S.D., Strickland, P., Forsberg, E.C., Li, D.Y., and Hinck, L. (2010). Vascular Robo4 restricts proangiogenic VEGF signaling in breast. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 10520–10525.
- Márquez, J.A., and Cipriani, F. (2014). CrysDirect: A novel approach for Automated Crystal harvesting based on photoablation of thin films. In *Structural Genomics: General Applications*, Y.W. Chen, ed. (Totowa, NJ: Humana Press), pp. 197–203.
- Massia, S.P., Stark, J., and Letbetter, D.S. (2000). Surface-immobilized dextran limits cell adhesion and spreading. *Biomaterials* *21*, 2253–2261.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* *41*, 597–600.

REFERENCES

- Mehlen, P., and Bredesen, D.E. (2004). The dependence receptor hypothesis. *Apoptosis* 9, 37–49.
- Mehlen, P., Rabizadeh, S., Snipas, S.J., Assa-Munt, N., Salvesen, G.S., and Bredesen, D.E. (1998). The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis. *Nature* 395, 801–804.
- Melani, M., and Weinstein, B.M. (2010). Common factors regulating patterning of the nervous and vascular systems. *Annu. Rev. Cell Dev. Biol.* 26, 639–665.
- Merz, D.C., Zheng, H., Killeen, M.T., Krizus, A., and Culotti, J.G. (2001). Multiple signaling mechanisms of the UNC-6/netrin receptors UNC-5 and UNC-40/DCC in vivo. *Genetics* 158, 1071–1080.
- Moore, D.T., Berger, B.W., and DeGrado, W.F. (2008). Protein-Protein Interactions in the Membrane: Sequence, Structural, and Biological Motifs. *Structure* 16, 991–1001.
- Morlot, C., Thielens, N.M., Ravelli, R.B.G., Hemrika, W., Romijn, R.A., Gros, P., Cusack, S., and McCarthy, A.A. (2007). Structural insights into the Slit-Robo complex. *Proc Natl Acad Sci USA* 104, 14923–14928.
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 355–367.
- Navankasattusas, S., Whitehead, K.J., Suli, A., Sorensen, L.K., Lim, A.H., Zhao, J., Park, K.W., Wythe, J.D., Thomas, K.R., Chien, C.-B., et al. (2008). The netrin receptor UNC5B promotes angiogenesis in specific vascular beds. *Development* 135, 659–667.
- Nettleship, J.E., Aplin, R., Radu Aricescu, A., Evans, E.J., Davis, S.J., Crispin, M., and Owens, R.J. (2007). Analysis of variable N-glycosylation site occupancy in glycoproteins by liquid chromatography electrospray ionization mass spectrometry. *Anal. Biochem.* 361, 149–151.
- Neu, B., Wenby, R., and Meiselman, H.J. (2008). Effects of Dextran Molecular Weight on Red Blood Cell Aggregation. *Biophys. J.* 95, 3059–3065.
- Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Jelloul, M. Ben, Sussman, J.L., Stuart, D.I., and Perrakis, A. (2005). Towards rationalization of crystallization screening for small- To medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 61, 1426–1431.
- Nguyen, A., and Cai, H. (2006). Netrin-1 induces angiogenesis via a DCC-dependent ERK1/2-eNOS feed-forward mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6530–6535.
- Nguyen-Ba-Charvet, K.T., and Chédotal, A. (2002). Role of slit proteins in the

- vertebrate brain. *J. Physiol. Paris* 96, 91–98.
- Nguyen-Ba-Charvet, K.T., Brose, K., Marillat, V., Sotelo, C., Tessier-Lavigne, M., and Chédotal, A. (2001). Sensory Axon Response to Substrate-Bound Slit2 Is Modulated by Laminin and Cyclic GMP. *Mol. Cell. Neurosci.* 17, 1048–1058.
- Okada, Y., Yano, K., Jin, E., Funahashi, N., Kitayama, M., Doi, T., Spokes, K., Beeler, D.L., Shih, S.C., Okada, H., et al. (2007). A three-kilobase fragment of the human Robo4 promoter directs cell type-specific expression in endothelium. *Circ. Res.* 100, 1712–1722.
- Olthof-Hazekamp, R. (1978). Crystallographic computing techniques edited by F. R. Ahmed. *Acta Crystallogr. Sect. A* 34, 350–351.
- Paduch, M., Koide, A., Uysal, S., Rizk, S.S., Koide, S., and Kossiakoff, A.A. (2013). Generating conformation-specific synthetic antibodies to trap proteins in selected functional states. *Methods* 60, 3–14.
- Paithankar, K.S., and Garman, E.F. (2010). Know your dose: RADDPOSE. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66, 381–388.
- Panjikovich, A., and Svergun, D.I. (2016). SASpy: A PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data. *Bioinformatics* 32, 2062–2064.
- Pappu, K.S., and Zipursky, S.L. (2010). Axon Guidance: Repulsion and Attraction in Roundabout Ways. *Curr. Biol.* 20, R400–R402.
- Parekh, R.B. (1991). Effects of glycosylation on protein function. *Curr. Opin. Struct. Biol.* 1, 750–754.
- Park, H.H., Lo, Y.-C., Lin, S.-C., Wang, L., Yang, J.K., and Wu, H. (2007a). The Death Domain Superfamily in Intracellular Signaling of Apoptosis and Inflammation. *Annu. Rev. Immunol.* 25, 561–586.
- Park, H.H., Logette, E., Raunser, S., Cuenin, S., Walz, T., Tschopp, J., and Wu, H. (2007b). Death domain assembly mechanism revealed by crystal structure of the oligomeric PIDDosome core complex. *Cell* 128, 533–546.
- Park, K.W., Morrison, C.M., Sorensen, L.K., Jones, C.A., Rao, Y., Chien, C. Bin, Wu, J.Y., Urness, L.D., and Li, D.Y. (2003). Robo4 is a vascular-specific receptor that inhibits endothelial migration. *Dev. Biol.* 261, 251–267.
- Park, K.W., Crouse, D., Lee, M., Karnik, S.K., Sorensen, L.K., Murphy, K.J., Kuo, C.J., and Li, D.Y. (2004). The axonal attractant Netrin-1 is an angiogenic factor. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16210–16215.
- Pasterkamp, R.J., and Kolodkin, A.L. (2003). Semaphorin junction: Making tracks toward neural connectivity. *Curr. Opin. Neurobiol.* 13, 79–89.
- Pellegrini, E., Piano, D., and Bowler, M.W. (2011). Direct cryocooling of naked

REFERENCES

- crystals: Are cryoprotection agents always necessary? *Acta Crystallogr. Sect. D Biol. Crystallogr.* *67*, 902–906.
- Perez, S.E., and Steller, H. (1996). Migration of glial cells into retinal axon target field in *Drosophila melanogaster*. *J. Neurobiol.* *30*, 359–373.
- Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., et al. (2013). Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* *20*, 660–664.
- Petoukhov, M. V, Franke, D., Shkumatov, A. V, Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P. V, and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* *45*, 342–350.
- Placzek, M., and Briscoe, J. (2005). The floor plate: multiple cells, multiple signals. *Nat. Rev. Neurosci.* *6*, 230–240.
- Popov, A.N., and Bourenkov, G.P. (2003). Choice of data-collection parameters based on statistic modelling. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* *59*, 1145–1153.
- Powell, H.R., Johnson, O., and Leslie, A.G.W. (2013). Autoindexing diffraction images with iMosflm. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *69*, 1195–1203.
- Powell, H.R., Battye, T.G.G., Kontogiannis, L., Johnson, O., and Leslie, A.G.W. (2017). Integrating macromolecular X-ray diffraction data with the graphical user interface iMosflm. *Nat. Protoc.* *12*, 1310–1325.
- Prada, I., Ongania, G.N., Buonsanti, C., Panina-Bordignon, P., and Meldolesi, J. (2006). Triggering receptor expressed in myeloid cells 2 (TREM2) trafficking in microglial cells: Continuous shuttling to and from the plasma membrane regulated by cell stimulation. *Neuroscience* *140*, 1139–1148.
- Przyborski, S.A., Knowles, B.B., and Ackerman, S.L. (1998). Embryonic phenotype of *Unc5h3* mutant mice suggests chemorepulsion during the formation of the rostral cerebellar boundary. *Development* *125*, 41–50.
- Rambo, R. (2017). Bioisis: SCATTER.
- Raper, J., and Mason, C. (2010). Cellular strategies of axonal pathfinding. *Cold Spring Harb. Perspect. Biol.* *2*, 1–21.
- Read, R.J., and Schierbeek, A.J. (1988). A phased translation function. *J. Appl. Crystallogr.* *21*, 490–495.
- Reeves, P.J., Callewaert, N., Contreras, R., and 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. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 13419–13424.
- Rendic, D., Wilson, I.B.H., and Paschinger, K. (2008). The Glycosylation Capacity of Insect Cells. *Croat. Chem. Acta* *81*, 7–21.
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* *42*, 320–324.
- Rothberg, J.M., Hartley, D.A., Walther, Z., and Artavanis-Tsakonas, S. (1988). slit: An EGF-homologous locus of *D. melanogaster* involved in the development of the embryonic central nervous system. *Cell* *55*, 1047–1059.
- Rothberg, J.M., Roger Jacobs, J., Goodman, C.S., and Artavanis-Tsakonas, S. (1990). Slit: An extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains. *Genes Dev.* *4*, 2169–2187.
- Sabatier, C., Plump, A.S., Ma, L., Brose, K., Tamada, A., Murakami, F., Lee, E.Y.H.P., and Tessier-Lavigne, M. (2004). The divergent robo family protein Rig-1/Robo3 is a negative regulator of slit responsiveness required for midline crossing by commissural axons. *Cell* *117*, 157–169.
- Säll, A., Walle, M., Wingren, C., Müller, S., Nyman, T., Vala, A., Ohlin, M., Borrebaeck, C.A., and Persson, H. (2016). Generation and analyses of human synthetic antibody libraries and their application for protein microarrays. *Protein Eng. Des. Sel.* *29*, 427–437.
- Sambrook, J., and Russell, D.W. (2001). *Molecular cloning: A laboratory manual*. Mol. Cloning A Lab. Man.
- De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., et al. (2012). ID29: A high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J. Synchrotron Radiat.* *19*, 455–461.
- Santoni, G., Zander, U., Mueller-Dieckmann, C., Leonard, G., and Popov, A. (2017). Multiple crystals hierarchical clustering for macromolecular crystallography: the ccCluster program.
- Sauter, N.K., Grosse-Kunstleve, R.W., and Adams, P.D. (2004). Robust indexing for automatic data collection. *J. Appl. Crystallogr.* *37*, 399–409.
- Schnittler, H. (2016). Contraction of endothelial cells: 40 years of research, but the debate still lives. *Histochem. Cell Biol.* *146*, 651–656.
- Schrödinger, LLC (2015). The {PyMOL} Molecular Graphics System, Version~1.8.6.0.
- Seeger, M., Tear, G., Ferres-Marco, D., and Goodman, C.S. (1993). Mutations

REFERENCES

- affecting growth cone guidance in drosophila: Genes necessary for guidance toward or away from the midline. *Neuron* *10*, 409–426.
- Seiradake, E., delToro, D., Nagel, D., Cop, F., Härtl, R., Ruff, T., Seyit-Bremer, G., Harlos, K., Border, E.C., Acker-Palmer, A., et al. (2014). FLRT Structure: Balancing Repulsion and Cell Adhesion in Cortical and Vascular Development. *Neuron* *84*, 370–385.
- Seki, M., Watanabe, A., Enomoto, S., Kawamura, T., Ito, H., Kodama, T., Hamakubo, T., and Aburatani, H. (2010). Human ROBO1 is cleaved by metalloproteinases and γ -secretase and migrates to the nucleus in cancer cells. *FEBS Lett.* *584*, 2909–2915.
- Selcuk Unal, E., Zhao, R., Qiu, A., and Goldman, I.D. (2008). N-linked glycosylation and its impact on the electrophoretic mobility and function of the human proton-coupled folate transporter (HsPCFT). *Biochim. Biophys. Acta - Biomembr.* *1778*, 1407–1414.
- Seth, P., Lin, Y., Hanai, J.I., Shivalingappa, V., Duyao, M.P., and Sukhatme, V.P. (2005). Magic roundabout, a tumor endothelial marker: Expression and signaling. *Biochem. Biophys. Res. Commun.* *332*, 533–541.
- Sheldon, H., Andre, M., Legg, J. a, Heal, P., Herbert, J.M., Sainson, R., Sharma, A.S., Kitajewski, J.K., Heath, V.L., and Bicknell, R. (2009). Active involvement of Robo1 and Robo4 in filopodia formation and endothelial cell motility mediated via WASP and other actin nucleation-promoting factors. *FASEB J.* *23*, 513–522.
- Shental-Bechor, D., and Levy, Y. (2008). Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 8256–8261.
- Shi, X., and Jarvis, D.L. (2007). Protein N-glycosylation in the baculovirus-insect cell system. *Curr. Drug Targets* *8*, 1116–1125.
- Shibata, F., Goto-Koshino, Y., Morikawa, Y., Komori, T., Ito, M., Fukuchi, Y., Houchins, J.P., Tsang, M., Li, D.Y., Kitamura, T., et al. (2009). Roundabout 4 is expressed on hematopoietic stem cells and potentially involved in the niche-mediated regulation of the side population phenotype. *Stem Cells* *27*, 183–190.
- Shim, H. (2015). Synthetic approach to the generation of antibody diversity. *BMB Rep.* *48*, 489–494.
- Shimada, K., Gill, P.J., Silbert, J.E., Douglas, W.H., and Fanburg, B.L. (1981). Involvement of cell surface heparin sulfate in the binding of lipoprotein lipase to cultured bovine endothelial cells. *J. Clin. Invest.* *68*, 995–1002.
- Sidhu, S.S., Lowman, H.B., Cunningham, B.C., and Wells, J.A. (2000). [21] Phage display for selection of novel binding peptides. pp. 333–363.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam,

- H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539.
- Smith-Berdan, S., Nguyen, A., Hassanein, D., Zimmer, M., Ugarte, F., Ciriza, J., Li, D., García-Ojeda, M.E., Hinck, L., and Forsberg, E.C. (2011). Robo4 cooperates with Cxcr4 to specify hematopoietic stem cell localization to bone marrow niches. *Cell Stem Cell* 8, 72–83.
- Sonawane, N.D., Szoka, F.C., and Verkman, A.S. (2003). Chloride Accumulation and Swelling in Endosomes Enhances DNA Transfer by Polyamine-DNA Polyplexes. *J. Biol. Chem.* 278, 44826–44831.
- Suchting, S., Heal, P., Tahtis, K., Stewart, L.M., and Bicknell, R. (2005). Soluble Robo4 receptor inhibits in vivo angiogenesis and endothelial cell migration. *FASEB J.* 19, 121–123.
- Suchting, S., Bicknell, R., and Eichmann, A. (2006). Neuronal clues to vascular guidance. *Exp. Cell Res.* 312, 668–675.
- Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25, 495–503.
- Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* 76, 2879–2886.
- Takahashi, T., and Strittmatter, S.M. (2001). PlexinA1 autoinhibition by the Plexin sema domain. *Neuron* 29, 429–439.
- Tanikawa, C., Matsuda, K., Fukuda, S., Nakamura, Y., and Arakawa, H. (2003). p53RDL1 regulates p53-dependent apoptosis. *Nat. Cell Biol.* 5, 216–223.
- Tarazona, M.P., and Saiz, E. (2003). Combination of SEC/MALS experimental procedures and theoretical analysis for studying the solution properties of macromolecules. *J. Biochem. Biophys. Methods* 56, 95–116.
- Tarentino, A.L., and Plummer, T.H. (1994). Enzymatic deglycosylation of asparagine-linked glycans: purification, properties and specificity of oligosaccharide-cleaving enzymes from flavobacterium meningosepticum. *Methods Enzymol.* 230, 44–57.
- Tarentino, A.L., Quinones, G., Schrader, W.P., Changchien, L.M., and Plummer, T.H. (1992). Multiple endoglycosidase (Endo) F activities expressed by *Flavobacterium meningosepticum*. Endo F1: molecular cloning, primary sequence, and structural relationship to Endo H. *J. Biol. Chem.* 267, 3868–3872.
- Tereshko, V., Uysal, S., Koide, A., Margalef, K., Koide, S., and Kossiakoff, A.A. (2008). Toward chaperone-assisted crystallography: protein engineering

REFERENCES

- enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. *Protein Sci.* *17*, 1175–1187.
- Tessier, D.C., Thomas, D.Y., Khouri, H.E., Laliberié, F., and Vernet, T. (1991). Enhanced secretion from insect cells of a foreign protein fused to the honeybee melittin signal peptide. *Gene* *98*, 177–183.
- Tessier-Lavigne, M., and Goodman, C.S. (1996). The molecular biology of axon guidance. *Science* *274*, 1123–1133.
- Thiebault, K., Mazelin, L., Pays, L., Llambi, F., Joly, M.-O., Scoazec, J.-Y., Saurin, J.-C., Romeo, G., and Mehlen, P. (2003). The netrin-1 receptors UNC5H are putative tumor suppressors controlling cell death commitment. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 4173–4178.
- Till, M., Robson, A., Byrne, M.J., Nair, A. V, Kolek, S.A., Shaw Stewart, P.D., and Race, P.R. (2013). Improving the success rate of protein crystallization by random microseed matrix screening. *J. Vis. Exp.* e50548.
- Tong, J., Killeen, M., Steven, R., Binns, K.L., Culotti, J.G., and Pawson, T. (2001). Netrin stimulates tyrosine phosphorylation of the UNC-5 family of netrin receptors and induces Shp2 binding to the RCM cytodomain. *J. Biol. Chem.* *276*, 40917–40925.
- Tria, G., Mertens, H.D.T., Kachala, M., and Svergun, D.I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* *2*, 207–217.
- Trindade, E.S., Oliver, C., Jamur, M.C., Rocha, H.A.O., Franco, C.R.C., Bouças, R.I., Jarrouge, T.R., Pinhal, M.A.S., Tersariol, I.L.S., Gouvêa, T.C., et al. (2008). The binding of heparin to the extracellular matrix of endothelial cells up-regulates the synthesis of an antithrombotic heparan sulfate proteoglycan. *J. Cell. Physiol.* *217*, 328–337.
- Turner, C.E., Glenney, J.R., and Burridge, K. (1990). Paxillin: A new vinculin-binding protein present in focal adhesions. *J. Cell Biol.* *111*, 1059–1068.
- Varki, A., Kannagi, R., Toole, B., and Stanley, P. (2017). Glycosylation Changes in Cancer. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, et al., eds. (Cold Spring Harbor Laboratory Press).
- Volkov, V. V, and Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* *36*, 860–864.
- Walter, J.D., Werther, R.A., Polozova, M.S., Pohlman, J., Healey, J.F., Meeks, S.L., Lollar, P., and Spiegel, P.C. (2013). Characterization and solution structure of the factor VIII C2 domain in a ternary complex with classical and non-classical inhibitor antibodies. *J. Biol. Chem.* *288*, 9905–9914.
- Wang, B., Xiao, Y., Ding, B.B., Zhang, N., Yuan, X. Bin, Gui, L., Qian, K.X., Duan, S.,

- Chen, Z., Rao, Y., et al. (2003). Induction of tumor angiogenesis by Slit-Robo signaling and inhibition of cancer growth by blocking Robo activity. *Cancer Cell* 4, 19–29.
- Wang, C., Yu, C., Ye, F., Wei, Z., and Zhang, M. (2012). Structure of the ZU5-ZU5-UPA-DD tandem of ankyrin-B reveals interaction surfaces necessary for ankyrin function. *Proc. Natl. Acad. Sci.* 109, 4822–4827.
- Wang, R., Wei, Z., Jin, H., Wu, H., Yu, C., Wen, W., Chan, L.N., Wen, Z., and Zhang, M. (2009). Autoinhibition of UNC5b Revealed by the Cytoplasmic Domain Structure of the Receptor. *Mol. Cell* 33, 692–703.
- Waterman, D.G., Winter, G., Parkhurst, J.M., Fuentes-Montero, L., Hattne, J., Brewster, A.S., Sauter, N.K., and Evans, G. (2013). The DIALS framework for integration software. *CCP4 Newsl. Protein Crystallography* 16–19.
- Waterman, D.G., Winter, G., Gildea, R.J., Parkhurst, J.M., Brewster, A.S., Sauter, N.K., and Evans, G. (2016). Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. Sect. D Struct. Biol.* 72, 558–575.
- Weber, C.H., and Vincenz, C. (2001). The death domain superfamily: A tale of two interfaces? *Trends Biochem. Sci.* 26, 475–481.
- Weiner, G.J. (2015). Building better monoclonal antibody-based therapeutics. *Nat. Rev. Cancer* 15, 361–370.
- Wen, S., Zhang, W., Ni, N., Wu, Q., Wang, X., Lin, Y., and Sun, J.-F. (2017). Expression of Rho GTPases family in melanoma cells and its influence on cytoskeleton and migration. *Oncotarget* 8, 30112–30122.
- Wills, Z., Emerson, M., Rusch, J., Bikoff, J., Baum, B., Perrimon, N., and Van Vactor, D. (2002). A *Drosophila* homolog of cyclase-associated proteins collaborates with the Abl tyrosine kinase to control midline axon pathfinding. *Neuron* 36, 611–622.
- 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., et al. (2006). Netrins promote developmental and therapeutic angiogenesis. *Science* 313, 640–644.
- 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., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 235–242.
- Wiseman, J.W., Goddard, C. a, McLelland, D., and Colledge, W.H. (2003). A comparison of linear and branched polyethylenimine (PEI) with DCChol/DOPE liposomes for gene delivery to epithelial cells in vitro and in vivo. *Gene Ther.* 10, 1654–1662.
- de Wit, J., Hong, W., Luo, L., and Ghosh, A. (2011). Role of Leucine-Rich Repeat Proteins in the Development and Function of Neural Circuits. *Annu. Rev. Cell*

REFERENCES

- Dev. Biol. 27, 697–729.
- Wong, K., Ren, X.R., Huang, Y.Z., Xie, Y., Liu, G., Saito, H., Tang, H., Wen, L., Brady-Kalnay, S.M., Mei, L., et al. (2001). Signal transduction in neuronal migration: Roles of GTPase activating proteins and the small GTPase Cdc42 in the Slit-Robo pathway. *Cell* 107, 209–221.
- Xia, Y., Chu, W., Qi, Q., and Xun, L. (2015). New insights into the QuikChange™ process guide the use of Phusion DNA polymerase for site-directed mutagenesis. *Nucleic Acids Res.* 43, e12.
- Xiao, T., Towb, P., Wasserman, S. a, and Sprang, S.R. (1999). Three-dimensional structure of a complex between the death domains of Pelle and Tube. *Cell* 99, 545–555.
- Yadav, S.S., and Narayan, G. (2014). Role of ROBO4 signalling in developmental and pathological angiogenesis. *Biomed Res. Int.* 2014.
- Yang, Y., Zou, L., Wang, Y., Xu, K.-S., Zhang, J.-X., and Zhang, J.-H. (2007). Axon guidance cue Netrin-1 has dual function in angiogenesis. *Cancer Biol. Ther.* 6, 743–748.
- Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P., and Ornitz, D.M. (1991). Cell-surface, heparin-like molecules are required for binding of basic fibroblast growth-factor to its high-affinity receptor. *Cell* 64, 841–848.
- Yokota, T., Milenic, D.E., Whitlow, M., and Schlom, J. (1992). Rapid Tumor Penetration of a Single-Chain Fv and Comparison with Other Immunoglobulin Forms. *Cancer Res.* 52, 3402–3408.
- Yoshikawa, M., Mukai, Y., Okada, Y., Tsumori, Y., Tsunoda, S.I., Tsutsumi, Y., Aird, W.C., Yoshioka, Y., Okada, N., Doi, T., et al. (2013). Robo4 is an effective tumor endothelial marker for antibody-drug conjugates based on the rapid isolation of the anti-Robo4 cell-internalizing antibody. *Blood* 121, 2804–2813.
- Ypsilanti, A.R., Zagar, Y., and Chédotal, A. (2010). Moving away from the midline: new developments for Slit and Robo. *Development* 137, 1939–1952.
- Yu, T.W., and Bargmann, C.I. (2001). Dynamic regulation of axon guidance. *Nat. Neurosci.* 4 *Suppl*, 1169–1176.
- Yu, J., Zhang, X., Kuzontkoski, P.M., Jiang, S., Zhu, W., Li, D.Y., and Groopman, J.E. (2014). Slit2N and Robo4 regulate lymphangiogenesis through the VEGF-C/VEGFR-3 pathway. *Cell Commun. Signal.* 12, 25.
- Yuan, W., Zhou, L., Chen, J.H., Wu, J.Y., Rao, Y., and Ornitz, D.M. (1999). The mouse SLIT family: secreted ligands for ROBO expressed in patterns that suggest a role in morphogenesis and axon guidance. *Dev. Biol.* 212, 290–306.
- Zakrys, L., Ward, R.J., Padiani, J.D., Godin, A.G., Graham, G.J., and Milligan, G. (2014). Roundabout 1 exists predominantly as a basal dimeric complex and this is

- unaffected by binding of the ligand Slit2. *Biochem. J.* *461*, 61–73.
- Zander, U., Bourenkov, G., Popov, A.N., De Sanctis, D., Svensson, O., McCarthy, A.A., Round, E., Gordeliy, V., Mueller-Dieckmann, C., and Leonard, G.A. (2015). MeshAndCollect: An automated multi-crystal data-collection workflow for synchrotron macromolecular crystallography beamlines. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *71*, 2328–2343.
- Zander, U., Hoffmann, G., Cornaciu, I., Marquette, J.-P., Papp, G., Landret, C., Seroul, G., Sinoir, J., Röwer, M., Felisaz, F., et al. (2016). Automated harvesting and processing of protein crystals through laser photoablation. *Acta Cryst* *72*, 454–466.
- Zelina, P., Blockus, H., Zagar, Y., Péres, A., Friocourt, F., Wu, Z., Rama, N., Fouquet, C., Hohenester, E., Tessier-Lavigne, M., et al. (2014). Signaling switch of the axon guidance receptor Robo3 during vertebrate evolution. *Neuron* *84*, 1258–1272.
- Zhang, B., Dietrich, U.M., Geng, J.-G., Bicknell, R., Esko, J.D., and Wang, L. (2009). Repulsive axon guidance molecule Slit3 is a novel angiogenic factor. *Blood* *114*, 4300–4309.
- Zhang, F., Moniz, H.A., Walcott, B., Moremen, K.W., Linhardt, R.J., and Wang, L. (2013). Characterization of the interaction between Robo1 and heparin and other glycosaminoglycans. *Biochimie* *95*, 2345–2353.
- Zhang, F., Prahst, C., Mathivet, T., Pibouin-Fragner, L., Zhang, J., Genet, G., Tong, R., Dubrac, A., and Eichmann, A. (2016). The Robo4 cytoplasmic domain is dispensable for vascular permeability and neovascularization. *Nat. Commun.* *7*, 13517.
- Zhao, H., Ahirwar, D.K., Oghumu, S., Wilkie, T., Powell, C.A., Nasser, M.W., Satoskar, A.R., Li, D.Y., and Ganju, R.K. (2016). Endothelial Robo4 suppresses breast cancer growth and metastasis through regulation of tumor angiogenesis. *Mol. Oncol.* *10*, 272–281.
- Zheng, L., Baumann, U., and Reymond, J.-L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* *32*, e115.
- Zhuang, X., Ahmed, F., Zhang, Y., Ferguson, H.J., Steele, J.C., Steven, N.M., Nagy, Z., Heath, V.L., Toellner, K.-M., and Bicknell, R. (2015). Robo4 vaccines induce antibodies that retard tumor growth. *Angiogenesis* *18*, 83–95.

Acknowledgements

Working on this project has surely been an exhausting and rewarding job and I have met a lot of people who helped me through it all.

The first thanks should go to my supervisor, Andrew, that allowed me to work on this project and has supported me with his knowledge and help over the years. It has been challenging work, but also a great opportunity to learn a great deal of things.

I am grateful to EMBL and its PhD program as a whole, and to the members of my TAC committee (Rob Meijers, Daniel Panne and Carlo Petosa) for overseeing my progress during my PhD. Many thanks to my thesis committee (Carlo Petosa, Elena Seiradake, Joanna Timmins and Valérie Castellani) for taking their time to evaluate my work presented here. Thanks to the HTX lab, the biophysical platform and the SPR platform of IBS for the use of their facilities and support. I am thankful to Stephen for sharing the lab space of his group with me when needed, and for allowing me to participate in their group meetings where I was able to receive additional feedback. Similarly, thanks to the Cusack group as a whole for listening to my presentations.

Finally, thanks to my group (Martha, Matthew, Jens, Pauline and especially Uli) for all their help during the last stage of my PhD.

Now that the customary, but honest, first round of thanks is done for, the next ones are for the people that have more directly impacted not only my professional life, but also my personal one.

Of the people that started the PhD with me in Grenoble, many thanks to Wiebke (for her scientific input, friendship and the numerous red marks covering the drafts of this thesis) and Paul (for his comments and fries, but especially for being the best climbing partner).

In no particular order, thank you to: Erika and Paolo (la miglior polenta di sempre), Lahari (because Bolliwood dancing), Aldo (ice ice, forefront), Estelle (for your company and sharp humour), Ziad (for helping with my French translations), Audrey (for the funny corridor moments), Petra (for bonus craziness), Valentina (for the Montenegro supply), Emiko, Etienne, Alex, Mathilde, Peter and Elise.

Of the people not in Grenoble, thanks to Vlada (best dinner organizer), Thomas (furniture hoarding man), Jo Ran (we are having a moment here), Olga (because firefighters) and Chris (for hosting me in times of need).

Thanks to my parents, my brother and Anna, that supported me from afar. This work is dedicated to them all.

Special thanks go to Frank Bradke, without whom I might not have been able to continue until the end.



The real result

